Modularity: Genes, Development, and Evolution

Diogo Melo,[1,*] Arthur Porto,[2,*] James M. Cheverud,[3] and Gabriel Marroig[1]

[1]Laboratório de Evolução de Mamíferos, Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, 05508-090, Brazil; email: gmarroig@usp.br

[2]Department of Biology, Washington University in St Louis, St Louis, Missouri 63130

[3]Department of Biology, Loyola University Chicago, Chicago, Illinois 60660

Annu. Rev. Ecol. Evol. Syst. 2016. 47:463–86

First published online as a Review in Advance on September 7, 2016

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi: 10.1146/annurev-ecolsys-121415-032409

*These authors have contributed equally to this work.

Keywords

macroevolution, genotype–phenotype map, G-matrix, adaptive landscape, morphological integration

Abstract

Modularity has emerged as a central concept for evolutionary biology, thereby providing the field with a theory of organismal structure and variation. This theory has reframed long-standing questions and serves as a unified conceptual framework for genetics, developmental biology, and multivariate evolution. Research programs in systems biology and quantitative genetics are bridging the gap between these fields. Although this synthesis is ongoing, some major themes have emerged, and empirical evidence for modularity has become abundant. In this review, we look at modularity from a historical perspective, highlighting its meaning at different levels of biological organization and the different methods that can be used to detect it. We then explore the relationship between quantitative genetic approaches to modularity and developmental genetic studies. We conclude by investigating the dynamic relationship between modularity and the adaptive landscape and how this relationship potentially shapes evolution and can help bridge the gap between micro- and macroevolution.

INTRODUCTION

Modularity has become a central concept in evolutionary biology (Wagner et al. 2007). A system is modular if it can be divided into multiple sets of strongly interacting parts that are relatively autonomous with respect to each other. This concept has been applied in developmental biology, in which modules either are different parts of the embryo that interact with each other, as with induction and morphogenesis, or are sets of interacting molecules that act independently in the patterning of multiple tissues. This concept can be extended to adult functional relationships, in which modules consist of parts that act together in the performance of some physiological function. In this review, we focus on the role of variational modules in evolutionary processes. Variational modules are sets of traits that vary together and somewhat independently from other modules.

Modular concepts emerged early in evolutionary thinking with Darwin's consideration of the "correlations of growth" in which he noted that slight evolutionary variations in one part of an organism would result in other parts also being modified. Later, Weldon (1893, p. 329) noted that "before we can properly estimate the changes at present going on in a race or species we must know . . . the degree of abnormality of other organs which accompanies a given abnormality of one." (For abnormality, read variations.) Pearson (1896) then derived the parameter for describing the degree of relationship between two characters that we use today, the Pearson Product Moment correlation.

Despite this very early interest, a multivariate understanding and consideration of evolving characters was not common at the time (Simpson 1958). Certainly before the development of digital computers, the amount of computational work involved in even a small multivariate study on a small sample required the calculation of enormous numbers of variances, covariances, and correlations. These herculean efforts were often not deemed worth the value derived from the research (Simpson 1958). Olson & Miller (1958) stood out by considering the variational relationships between traits as a central feature of evolution and incorporating a more holistic, systems view of the phenotype and evolution. Olson & Miller (1958) hypothesized that the degree of interdependence in development and function among morphological characters is directly related to their degree of morphological integration as measured by the statistical correlation between trait distributions. Hence, they predicted that developmentally and functionally related traits will be relatively highly intercorrelated. An early example of such phenomena was found in the flower morphology of angiosperms, which can be divided into two sets of highly intercorrelated traits (or correlation pleiades): a vegetative set and a reproductive one (Berg 1960). Further theoretical and empirical work on this concept (Cheverud 1982, 1984; Lande 1979) showed that developmental and functional integration results in correlational selection that leads to genetic integration (genetic correlations). In turn, this genetic correlation leads to evolutionary integration, the correlated evolution of traits. The concept of morphological integration maintained some currency in evolution and systematics from the 1960s through the 1990s. However, interest greatly increased in the new millennium. Much of this increased attention occurred after the publication of several papers on the role of modularity in evolution, especially that of Wagner & Altenberg (1996) and the 1999 University of Chicago Press reissue of Olson & Miller's book, *Morphological Integration*. Wagner & Altenberg (1996) argued that modularity was important in facilitating the evolution of morphological diversity. If all features of an organism are completely integrated, the parts will be prevented from evolving independent adaptations. A modular variational structure permits the evolution of complexity and diversity as observed in the natural world.

Concurrently, important developments were taking place in evolutionary quantitative genetics. In the late 1970s and early 1980s, Lande and colleagues (Lande 1979, Lande & Arnold 1983) reintroduced models of multivariate evolution that had been ignored in evolutionary biology and

systematics since Pearson's time (Pearson 1896), although they were better known in agricultural genetics (e.g., Hazel 1943). Lande (1979) also showed how quantitative genetic evolutionary models could be used in systematics to investigate the evolutionary causes of diversification on a macroevolutionary scale by providing expectations for the diversification of species under genetic drift and under directional selection. Here we review the genetics of variational modularity, its relationship with development, how it can evolve, and its consequences for evolution and systematics.

METHODOLOGICAL CONSIDERATIONS

Representations of Morphology

Although most of what is said in this review can be applied to any continuous traits, much of the work related to modularity is concerned with morphological traits. The traditional way of representing morphological structures is to use a suite of linear distances (Olson & Miller 1958), preferably taken within a single homologous structure such as a bone, to represent a given morphological structure in a specimen. This representation captures local developmental and functional factors in a single homologous trait. The last 20 years, however, saw a shift in methodology in favor of using landmark-based methods, especially generalized Procrustes analysis (GPA) (Bookstein 1997, Kendall 1984). GPA takes a set of 2-D or 3-D landmarks measured in a group of specimens, scales all specimens to a common size, and uses an interactive procedure to superimpose the scaled configurations by minimizing the squared distance between the landmarks in all specimens and a mean shape. From this superimposed set we can calculate the distance between each specimen and the mean shape, and this calculated distance is used to represent them. Although this procedure has many desirable mathematical properties (Bookstein 1997) and is a powerful way of describing a morphological structure, its appropriateness for the study of variation and covariation has never been fully established. In particular, because changes in a single landmark will cause changes in the whole configuration, locality of variation is not necessarily preserved in the GPA covariance matrix. This problem has long been recognized in the morphometrics community (Adams et al. 2013), but its consequences for the study of modularity and evolution have only recently become apparent (Márquez et al. 2012, van der Linde & Houle 2009). Because local variation is not preserved, it is hard to detect local associations and covariation in populations using GPA, thereby limiting its use for the study of modularity. Not preserving local variation is also a problem when relating genetic variation to morphological variation, because again variation will be spread out over the whole morphological structure, and local genetic factors will appear to have widespread effect (Berner et al. 2011). Promising efforts have been made to reconcile landmark-based methods and local variation, such as the local shape variables described by Márquez et al. (2012), finite element scaling analysis (Cheverud & Richtsmeier 1986), and Euclidean distance matrix analysis (Lele & Richtsmeier 1991), but these methods have not yet been widely adopted. With this problem in mind, we do not discuss approaches that make use of GPA to study covariation and modularity. Instead, we focus on representations of morphology that preserve local variation, like linear distances and local shape variables.

Detecting Variational Modularity

Variational modularity has been used in several different contexts; therefore, a wide range of methods for detecting and quantifying variational modularity in multivariate data are available. At their core, most methods are based on some measure of association among traits (e.g., covariances or correlations), and modularity has often been inferred through the analysis of patterns and

magnitudes of association (e.g., Armbruster et al. 2004, Porto et al. 2009). Given a set of traits in a population and a correlation (or covariance) matrix between them, we might ask which sets of traits are grouped in modules, or if a particular partition of traits is supported by the observed statistical correlation between trait associations. We discuss methods for the detection of modules in these two situations: (*a*) extracting putative groupings of traits without a prior hypothesis and (*b*) testing if a particular partition established on different grounds is supported by the observed correlation matrix.

Detecting putative modules is very common in systems biology (Ayroles et al. 2009, Ihmels et al. 2002), in which traits are frequently expression data for thousands of genes and a priori hypotheses are impractical or impossible. Methods for partitioning traits into modules usually derive from network and graph theory. These methods treat the correlation matrix as a fully connected weighted graph and use either algorithms designed for community detection in graphs (Langfelder & Horvath 2008, Reichardt & Bornholdt 2006) or clustering algorithms coming from other contexts (like Potts model clustering or neighbor joining). Network-based models search for partitions in which members in the same partition share more connections than expected in a random network. Currently, these methods work well in high-dimensional problems, in which misclassification of some individual traits is not a serious problem. Although these methods have been used in much lower-dimensional problems (Magwene 2001), results are not always easy to interpret and partitions can group seemingly unrelated traits together. This can be partially explained by the origin of the methods: Because these methods are borrowed from graph theory, most of the methods and definitions relate to properties of random graphs, and how to translate these assumptions to correlation matrices is not obvious. For example, few methods consider the possibility of a trait belonging to two modules, or that modules might have a nested or hierarchical organization. Recently some effort has been made to produce module detection algorithms tailored for correlation matrices (MacMahon & Garlaschelli 2015), but these algorithms have not been applied to biological systems, and more work is needed to develop tools that can deal with complex modularity structures.

In morphological systems with lower dimensionality, and for which information on development or function of the measured traits is available, we may use this information to formulate putative partitions of the traits into modules. These a priori partitions can then be ranked by the support given to them by the observed associations between traits or tested for compatibility with the observed associations using a significance test. One approach to this problem is to compare the proposed partition with random partitions, using some statistic dependent on the partitions and the correlation matrix. The correlation test proposed by Cheverud (1989) compares the within-module correlations with between-module correlations. If the observed difference in within- and between-module correlations is higher than the difference for random partitions, the modular structure is considered valid. The RV coefficients (Klingenberg 2009) are a generalization of the squared Pearson correlation coefficient to multiple dimensions and can be used to quantify the degree of independence between two groups of traits. The RV statistic is calculated for a proposed partition and compared with random partitions via permutations. Although this statistic was proposed in the context of landmark data, it can be used with linear distances or local shape variables. Márquez (2008) presents a framework that allows the simultaneous testing of many competing modularity hypotheses, including overlapping and hierarchical modules. The main idea is to use a modularity hypothesis to generate a modeled covariance matrix, in which within-module covariances are set to the observed values and the between-module covariances are set to zero. This modeled matrix is compared with the original covariance matrix with a multivariate measure of similarity. For a recent and straightforward approach based directly on the correlation matrices and model comparison, see Goswami & Finarelli (2016).

Both module detection and module validation using correlations are made difficult by the presence of global integrating factors, like size variation or growth, that increase between-module correlations (Marroig et al. 2004, Mitteroecker & Bookstein 2007, Porto et al. 2013). These are discussed below.

GENETICS OF MODULARITY

Genetic associations among traits can be explained by two different phenomena: pleiotropy and linkage disequilibrium (LD). LD refers to the nonrandom association of alleles at different loci. In large populations, and in the absence of selection, LD will be eliminated by recombination after several generations of random mating. For that reason, LD is considered a transient source of genetic association (Cheverud 1996), except in species with only a few segregating chromosomes. Pleiotropy, in our context, refers to the manifold phenotypic effects of a single unit of inheritance (Stearns 2010). The word context is used here to emphasize the difficulties in finding a universal definition for the term (see Paaby & Rockman 2013). Pleiotropy, when defined in this way, is considered an important source of genetic association, because it causes traits to be inherited together and, depending on the structure of pleiotropic effects of other contributing loci, to vary together within populations.

Given the importance of pleiotropy as a source of association among traits, one might be interested in the structure of pleiotropic effects of loci underlying modular trait variation (**Figure 1**) (Wagner & Altenberg 1996). Two prominent questions are whether pleiotropic effects are also modular and whether modular pleiotropy facilitates evolvability—the ability of a population to respond in the direction of selection (sensu Hansen 2003). From a theoretical standpoint, several models for the structure of the genotype–phenotype (GP) map have been put forward (Hansen 2003, Mitteroecker 2009, Pavlicev & Hansen 2011). The general consensus is that, given certain assumptions, multiple different models of the GP map are equally capable of explaining observed genetic associations among traits (Mitteroecker 2009). Similarly, although modular GP maps can maximize evolvability in stochastic environments, they do not necessarily maximize it under more stable environments (e.g., Hansen 2003).

A clearer picture of whether GP maps are modular and whether they promote the evolvability of organisms came with the collection of large empirical data sets in mice, yeast, and nematodes (see Wang et al. 2010). These large data sets allowed for a systematic investigation of the pleiotropic effects of genes on the phenotype across a variety of approaches, including quantitative trait loci (QTL) mapping and gene knockout studies. The picture emerging from these large data sets is that most mutational effects are modular, with different sets of genes affecting different sets of functionally and developmentally related traits (Wang et al. 2010). In other words, the variational modularity observed in the phenotype can be explained by modularity in the GP map (sensu Wagner & Altenberg 1996). A minority of mutations affect large groups of traits as they are associated with global genetic factors (**Figure 1**). More importantly, in these same studies, modular pleiotropy was shown to maximize the rate of adaptation and promote the evolution of complexity, owing to the scaling of mutational effects with the degree of pleiotropy (Wang et al. 2010).

EVOLUTION OF MODULARITY

Genetic Variation in Pleiotropy

Modularity can evolve through changes in the pleiotropic effects of alleles on traits themselves. Genetic variation in pleiotropy has long been recognized as playing an important role in evolutionary processes. Mayr (1963) noted the importance of epistatic interactions in ameliorating

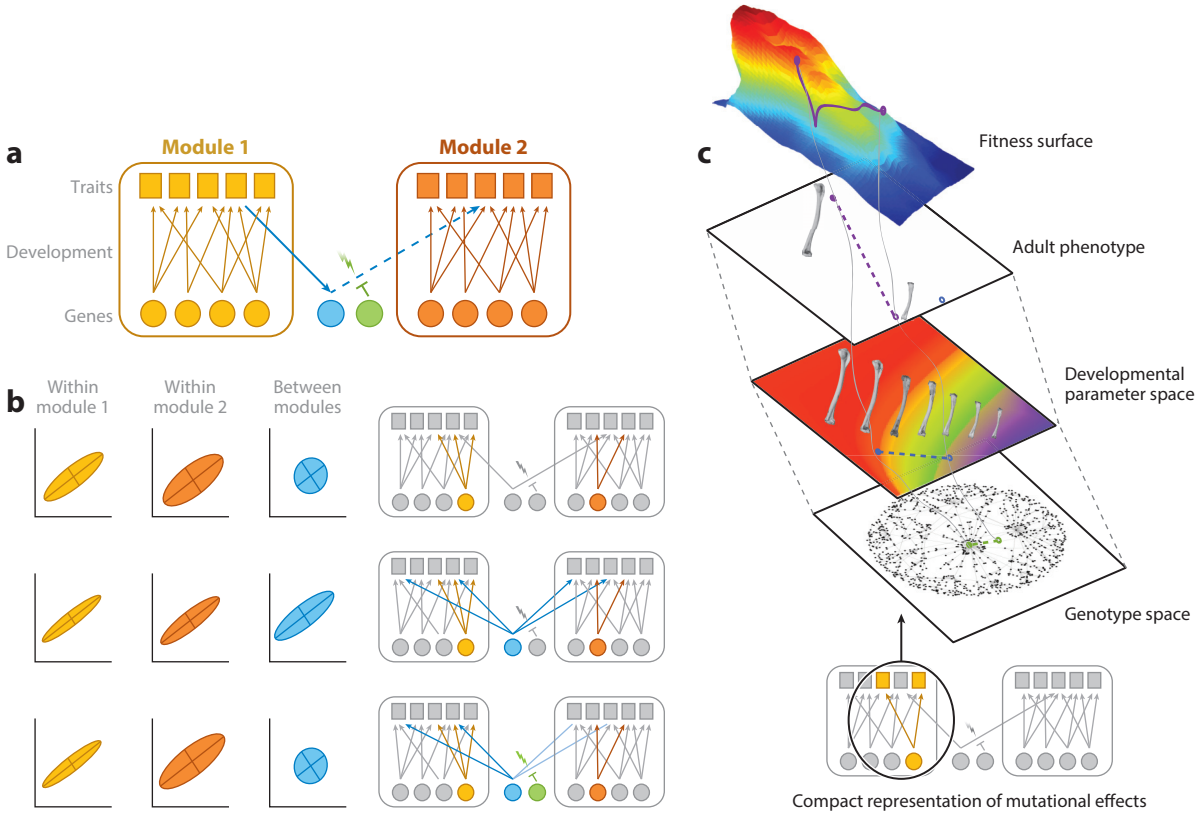


Figure 1

(*a*) Typical representation of modularity in the genotype–phenotype map. Yellow and orange circles represent modular genetic factors, and blue circles represent global ones. The green circle represents a genetic locus capable of preventing the global factor from affecting module 2 (rQTL). Squares represent phenotypic traits, and arrows represent the relationship between genotype and phenotype (polygeny and pleiotropy). (*b*) Trait correlation as a function of the underlying genetic variation. In case 1 (*first row*), genetic variation is only present for local factors. Consequently, modular patterns of covariation emerge in the phenotype. In case 2 (*second row*), genetic variation in both global and local genetic factors is present. Consequently, covariation patterns are less modular. Finally, in case 3 (*bottom row*), modular covariation patterns emerge again as a consequence of the rQTL preventing the global genetic factor from affecting module 2. (*c*) The nature of gene effects across the different levels of the biological hierarchy. In this panel, we present a case in which a certain mutation, represented at the level of the genotype, causes changes in the developmental parameter space (e.g., rate of cell division), which in turn leads to changes in the selected phenotype and movement along a fitness surface. Abbreviation: rQTL, relationship quantitative trait loci.

the deleterious pleiotropic effects of alleles on fitness and enhancing their positive fitness effects. Variation in allelic effects at a target locus is produced by differential epistasis, a phenomenon in which epistatic interactions between the target and modifier loci on multiple traits differ in their effects from one trait to the next (Cheverud 1996; Pavlicev et al. 2008, 2011b). Several examples of differential epistasis have been recognized during the last 30 years. The abnormal abdomen (*aa*) locus in *Drosophila mercatorum* is a classic example of this phenomenon. In laboratory experiments, the *aa* locus was found to have a wide variety of pleiotropic effects on morphological and life history traits (Templeton et al. 1985), but these effects were not manifested in wild populations due to modifier loci. Differential epistasis has also been described in several other systems, including coronary artery disease (Maxwell et al. 2013) and viral reproductive success (Pepin et al. 2006).

Although the importance of epistasis has long been appreciated in evolution, only recently has the major part that epistatic pleiotropy plays in shaping covariation become apparent. Wolf et al. (2005) used QTL mapping in experimental crosses of inbred Large (LG/J) and Small (SM/J) mice strains to investigate the genetic architecture in several late and early skull traits. Covariation between traits was strongly affected by epistatic variation in pleiotropy, and the genetic architecture determining the pattern of association between traits can be attributed to a complex pattern of genetic interactions. In these mice, most epistatic effects on pleiotropy reduced covariation and led to a more modular genetic organization. Pavlicev et al. (2008) investigated the allometric relation between body weight and long bone length in mice, and the authors identified several relationship QTLs (rQTLs), which are QTLs that do not necessarily affect the mean value of traits but affect the relationships between traits (Wagner et al. 2007). This widespread evidence of genetic variation in the covariation between traits due to epistatic interactions provides ample scope for natural selection to change associations between traits and, hence, modularity patterns.

Modeling Changes in Covariation

The availability of variation in the associations between traits led Pavlicev et al. (2011a) to develop a deterministic model for the evolution of pleiotropic gene effects under directional selection. In this model, there is genetic variation in the strength of the correlation between two continuous traits in the form of a polymorphic rQTL that has no effect on the trait mean. Traits that are selected in the same direction tend to become more strongly correlated, even if selection is fluctuating. Conversely, if the two traits are under corridor selection, in which one of the traits is selected to either increase or decrease and the other is kept constant, the rQTL allele representing low correlation is positively selected and the traits become independent. The main conclusion of their model is that the nature of pleiotropic allelic effects is expected to evolve to match adaptive patterns of selection.

Using the existence of variation in pleiotropic relations, Melo & Marroig (2015) developed an explicit individual-based stochastic model for the evolution of continuous traits in finite populations, in which pleiotropic associations between genetic loci and phenotypic traits are free to change under mutation. This model has the advantage of being able to include a number of complications, such as a large number of traits, drift, different patterns of selection, recombination, and mutation. The possibility of simulating several traits is especially interesting because it permits the investigation of complex modular patterns. The authors evaluated the evolution of modularity under a series of evolutionary scenarios and reached a number of conclusions. Drift and stabilizing selection were not capable of creating lasting modular patterns of covariation, whereas divergent directional selection (during which one group of traits is selected in one direction and another group is selected in the opposite direction) created modularity—traits selected in the same direction became more strongly correlated and formed clear variational modules. Under corridor selection, the group of traits under directional selection becomes more correlated, whereas the group of traits under stabilizing selection maintains intermediate levels of correlation, and the correlations between these two groups become very low. These findings suggest corridor selection is a powerful mechanism for creating complex modular patterns.

Epistasis has also been implicated in the evolution of the mutation matrix for continuous traits. Jones et al. (2014) developed a model for the evolution of two continuous traits under genetic control of several pleiotropic loci. The model also included a stable pattern of epistatic interaction between the loci affecting the quantitative traits. The traits were subjected to correlated and independent stabilizing selection. Under uncorrelated selection the traits presented

correlations near zero, and the mutation matrix also had zero correlations between traits. Under correlated stabilizing selection, however, the traits' genetic correlations changed and mirrored the pattern of stabilizing selection. The mutation matrix also aligned with the selection surface and led to a situation in which the effects of new mutations are biased by previous selective history.

These different models allow us to draw general conclusions regarding the expected evolution of covariation between quantitative traits, regardless of the specific model used. First, variation in pleiotropic relations is essential to the evolution of modularity, and this variation can be attained through epistatic interactions. Second, under directional selection, traits that are jointly selected in the same direction tend to become more strongly correlated. Third, because selection can change the pleiotropic and developmental relations between traits, future evolutionary changes can be biased by previous selective history.

Empirical Evidence for Changes in Covariation

Several instances in which evolution, and presumably selection, has broken down patterns of association among traits to produce major adaptive shifts have been reported. Hallgrímsson et al. (2012, p. 502) classified these changes in patterns of modularity as a form of evolutionary novelty, as there is "a breakdown of ancestral developmental constraints such that variation is generated in a new direction or dimension." Indeed, this type of novel variation has been documented in several systems. Young & Hallgrímsson (2005) found a common pattern of strong covariation between the forelimb and hindlimb elements in quadrupedal mammals that constrains the independent evolution of the limbs. However, two mammals with highly derived limb morphologies, the brachiating gibbons and the flying bats, show a reduction in the cross-limb correlation that accompanies their extreme limb individuation.

Examples of artificial selection overcoming the initial pattern of genetic associations have also been found. Beldade et al. (2002) used the eyespots in butterfly wings as a target of selection. Initially, the anterior and posterior eyespots were correlated, and selection for coordinated change of both eyespots produced a rapid and linear response, whereas selection in the uncoupling direction, for increase in one eyespot and decrease of the other, led to a response that was much less linear and more irregular. This study illustrates that there are preferential directions for evolution produced by the pattern of modularity, and that evolution is faster in these directions, but these are not absolute restrictions. Similar results were obtained for artificial selection experiments aimed at reducing between-sex genetic correlations for flower size in *Silene latifolia* (Delph et al. 2011). Recently, artificial selection experiments using *Drosophila melanogaster* revealed yet another important role for selection in influencing patterns of association among traits. By selecting on allometric relationships in drosophilid wing shape, Bolstad et al. (2015) produced laboratory lineages presenting larger differences in allometric slopes than the ones observed across a large clade of drosophilids. This evolutionary response to selection in the allometric slopes was, however, quickly lost after selection pressures were suspended. This finding indicates that internal selection might be responsible for maintaining conserved allometric slopes on a macroevolutionary timescale.

These results illustrate the complex interactions between modularity and selection. Evolutionary restrictions imposed by genetic associations are rarely absolute, and selection that privileges uncoupling of associated traits can lead to a reorganization of variational patterns. At the same time, covariation patterns can also be largely maintained due to either internal selective pressures (as in the allometric relations in *Drosophila*) or differences in the availability of rQTL variation to change pleiotropy.

DEVELOPMENT AS THE LINK BETWEEN GENES AND PHENOTYPE

Understanding the mechanics and regulation of development is becoming increasingly essential to elucidating the relationship between modularity and trait evolution (e.g., Salazar-Ciudad & Jernvall 2010). Development occurs not only through the molecular interaction among many gene products in a particular environmental context but also through the mechanical interactions between the developing cells and tissues, all of which can create significant nonlinearities in the GP map (Alberch 1991, Polly 2008, Watson et al. 2014). In this section, we review recent literature that explicitly addresses the connection between quantitative approaches to modularity and the underlying developmental genetics. We are particularly interested in studies that explicitly incorporate the mechanics of development into an evolutionary framework or that identify genes that contribute to change in canalization. We also highlight the emergence of the new field of system genetics.

Causes of a Phenotype Versus Causes of Phenotypic Variation

In a developmental context, it is particularly important to distinguish between the causes of a phenotype and causes of phenotypic variation. A developmental process might be essential for a trait to emerge (cause of a phenotype), but as long as this process is conserved across individuals, it will not be a cause of phenotypic variation in a population. Empirical evidence overwhelmingly suggests that causes of a phenotype are modular in nature. The whole concept of character relies on it. According to Wagner (2007), character identity is specified by gene regulatory units that control the developmental program. These regulatory units, termed character identity networks, imply that we can truly recognize something as a distinct character only if it shows some degree of modularity at the developmental genetic level.

Modularity in the causes of variation of a phenotype is another matter altogether. The notion that causes of phenotypic variation might also be modular stems from the concept of morphological integration, as seen above (in the section, Genetics of Modularity), and from the imitatory epigenotype hypothesis (Riedl 1978), which predicts that the pattern of developmental constraints imitates the pattern of functional constraints, thereby leading to covariation among functionally related traits within populations. Empirical evidence suggests that there can be a correspondence between variational and developmental modularity, but this correspondence is by no means guaranteed, as seen in previous sections.

Incorporating Development into Evolutionary Studies

Although variational modularity is often assumed to be a consequence of variation in the underlying developmental mechanisms, explicitly modeling developmental systems or even inferring developmental processes from variational modules are not simple tasks (Hallgrímsson et al. 2009, Pavličev & Cheverud 2015). Interactions among tissues, as well as local and global genetic factors acting at different time points, are all superimposed during development and contribute to the final phenotype (Hallgrímsson et al. 2009, Mitteroecker & Bookstein 2007). Similarly, a single modular pattern can emerge through multiple independent developmental pathways (Mitteroecker 2009) and make the prediction across levels of the hierarchy difficult. To our knowledge, the most successful empirical case of incorporating developmental parameters in evolutionary models is Salazar-Ciudad & Jernvall's (2002, 2010) model of tooth development. Tooth development is a relatively well understood process, as several of the genetic interactions and cellular processes that lead to tooth formation are known. Consequently, Salazar-Ciudad & Jernvall (2010)

created mathematical models describing tooth morphology as a consequence of perturbations in the underlying genetic and developmental parameters, such as the rate of cell proliferation or cell adhesion. This model was successful at producing accurate predictions of tooth morphology for several mammalian groups (Salazar-Ciudad & Jernvall 2010). Other successful cases of the use of developmental mechanisms to explain phenotypic variation come from *Drosophila* wing venation patterns (Matamoro-Vidal et al. 2015) and butterfly wing eyespots (Beldade & Brakefield 2002), both of which involve changes in several key developmental processes, such as the distribution of morphogens in the wing disc or the establishment of planar cell polarity.

A theoretical approach that has also undertaken a more explicit incorporation of developmental information into the evolution of the GP map was developed by Watson et al. (2014). In their model, multivariate traits are produced by a GP map with multiple independent developmental steps connecting the phenotype to the genotype. This conceptualization produces a nonlinear ontogeny and allows the model to capture interesting behaviors of these GP maps, such as the ability to recall multiple phenotypes that were selected in the past or the ability to produce new combinations of features from modular developmental processes. Traits in this model tend to become more associated throughout development when they are selected in the same direction and become independent when they are selected in different directions, thereby reinforcing the role of directional selection in shaping modularity.

In conclusion, although approaches relating variational modularity to the mechanics of development are relatively rare, these different empirical studies and theoretical models clearly show that incorporating the more complex developmental interactions into studies of morphological variation and evolution greatly increases our ability to understand and even predict the evolutionary dynamics of complex systems. The main challenge going forward will be to create models capable of describing more complex structures, such as the skull; allowing specific connections between DNA sequences, developmental networks, and variational modularity; and incorporating the possibility of changes in the topology of genetic and developmental networks.

System Genetics—A Systematic Approach

A promising way to integrate modularity with the underlying developmental genetics in a systematic way is currently gaining traction with the system genetics approach (Ayroles et al. 2009, Mackay et al. 2009). The idea behind system genetics is simple. It interrogates the relationship between genome and phenome under different contexts (e.g., environments or conditions). Consequently, it attempts to hit at the core of the context dependency of gene effects, which not only is fundamental for the evolution of modularity (Pavličev & Cheverud 2015), as seen in previous sections, but also emphasizes its developmental basis by potentially uncovering important modular signaling cascades.

System genetic approaches have been applied to several different model organisms (Ihmels et al. 2002, Juenger et al. 2005, Wang et al. 2010). In *Drosophila* (Ayroles et al. 2009), it led to the identification of several transcriptional modules that not only are connected to genomic variation but also underlie variation in ecologically relevant traits, such as fecundity and metabolism. Those transcriptional modules are strongly influenced by environmental, developmental, and genetic background effects, thereby highlighting the fact that context-dependent effects are the norm and therefore are responsible for most phenotypic variation. System genetic approaches have also recently been used to map changes in the amount of variation for a given phenotype (Ayroles et al. 2015). Genetic variation in phenotypic variance represents genetic variation in developmental canalization, a topic that is especially relevant to studies of threshold characters or threshold selection (Ayroles et al. 2015). Among the challenges faced by system genetics, two

should be highlighted. Due to the ambitious nature of scoring multivariate traits and entire transcriptomes/genomes, system genetics studies are inherently expensive studies. Also, multivariate statistics are often dependent on large samples and are often estimated with considerable error, an aspect that needs to be taken into account.

MODULARITY AND THE ADAPTIVE LANDSCAPE

How are modularity and integration relevant for phenotypic evolution? Having discussed and characterized modularity and the possibility of its evolution, we now address its evolutionary consequences. There are short-term and potentially long-term consequences of modularity for evolutionary change. We start by introducing the quantitative theory dealing with the short-term consequences and identifying under which circumstances this theory can be extended to macroevolutionary time. Because modularity in patterns of genetic associations between continuous traits is captured by the additive genetic variance/covariance matrix, the G-matrix (Lande 1979), this section focuses on the relationship between the G-matrix and the adaptive landscape, which relates possible phenotypes in the morphospace with fitness (sensu Arnold et al. 2001, Simpson 1944).

Why So Much Interest in the G-Matrix?

Evolution, regardless of which evolutionary process is involved, depends on genetic variation. The G-matrix summarizes the amount and pattern of additive genetic variation and covariation among traits and is, therefore, essential to our understanding of the connection between genetics and evolution (Lande 1979). Genetic covariation among traits is particularly important because of its potential to affect the course of phenotypic evolution (**Figure 2**). Unlike the univariate view of evolution, in which a single trait's value can be optimized without constraint from selection on other traits, genetic covariation among traits causes correlated responses to selection. In this situation traits will change and evolve together, often in a direction that is different from the one favored by selection (**Figure 2**) (Grant & Grant 1995, Lande 1979). Thus, the pattern and magnitude of the G-matrix elements can deflect the path of evolution from its optimal trajectory. Whether or not this short-term effect on the evolutionary responses has enduring consequences depends on the degree of stability of the G-matrix and on its relationship with the adaptive landscape (Steppan et al. 2002; see also section on Evolutionary Change in Rugged Landscapes).

Long-term stability of the G-matrix is one of the most fundamental assumptions of the research program described here, and questions related to the long-term stability and estimation of G-matrices are still open (Houle & Meyer 2015, Jones et al. 2012). Many feel uncomfortable with this assumption of stability (Björklund et al. 2013). This discomfort stems in part from empirical evidence that suggests that no two populations have identical G-matrices and that G-matrices can fluctuate over short time periods (Björklund et al. 2013, Eroukhmanoff & Svensson 2011). Biological populations are finite and almost surely differ in their gene frequencies, especially considering the potentially large number of genes affecting complex traits and the correlations between them (Phillips et al. 2001, Whitlock et al. 2002). We suggest that in many instances the assumption that population covariance matrices are identical be rejected out of hand. But the mere presence of a statistically significant difference is not the critical issue. Instead, the more interesting and relevant questions are: How similar are two covariance patterns with respect to their predicted evolutionary responses? Do some quantitative traits have more stable G-matrices than others? Fortunately, these are questions that can be examined empirically (Calsbeek & Goodnight 2009, Cheverud & Marroig 2007) and are a critical first step for the use of the G-matrix in multivariate evolution and systematics.

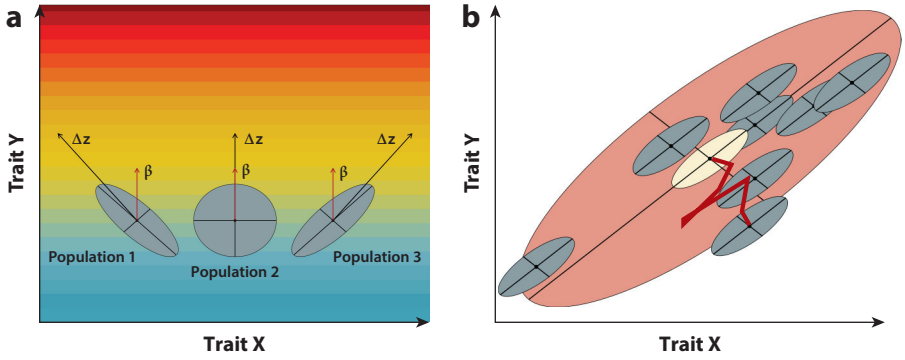


Figure 2

To illustrate the interaction of modularity (captured in the G-matrix) and evolutionary processes (selection and drift), we display two panels illustrating population averages, G-matrices, adaptive peak(s), and selection gradients (β). G-matrices are represented by ellipses with the axes of major genetic variation embedded. The major axis corresponds to the line of least resistance (Schluter 1996), that is, the direction that holds most of the genetic variation in the trait space. Selection gradients are represented by straight arrows and measure the relationship between fitness and individual traits while holding the other traits constant. Responses to selection (Δz) are also shown as arrows and indicate changes in the trait averages across time. The multivariate response to selection equation ($\Delta z = G\beta$) captures the relationship between the response to selection (Δz), inheritance (G-matrix), and selection (β). The direction of increase in average fitness is indicated by a color gradient from blue (lower fitness) to red (higher fitness). Panel *a* shows three populations under the same adaptive landscape, in which an increase in trait Y is favored and trait X does not affect fitness. Notice that the response of each population (*gray arrows*) to the same selection gradient (*red arrows*) differs. Population 1 increases Y but decreases X values, population 3 increases Y and X values, and population 2 only increases Y, although with a smaller displacement of Y than the other two populations. These different responses under the same selection gradient are due to differences in the G-matrices. Populations 1 and 3 have their responses deflected from the optimal path, due to the covariation between Y and X (negative in population 1, and positive in population 3). Thus, traits that are not under direct selection will contribute to the response owing to their shared inheritance. If traits are independent (population 2), each can be optimized separately. But note that the displacement in the Y average is smaller than in the other populations, because the genetic variation available in that direction is smaller. Panel *b* shows the consequence of a flat adaptive landscape (random genetic drift) on the averages of descendent populations (*gray ellipses*) of an ancestral population (*yellow ellipse*). At the end of the drift process, 95% of all evolution (divergence among means) is captured by the larger ellipse (*red*). Notice that substantially more divergence appears along the axis holding most of the within-population genetic variation. One population trajectory under genetic drift is shown in red to illustrate the random process.

So, what do empirical studies tell us about the relative stability of the G-matrix over macroevolutionary timescales? Empirical evidence varies greatly depending on the study system in question. One of the most thoroughly explored cases is the mammalian skull, from which empirical evidence strongly suggests that G-matrices are stable across mammalian taxa (Garcia et al. 2014, Marroig & Cheverud 2001, Porto et al. 2009). By stable we do not mean that heritable variation patterns are identical across species, but that they will deflect the phenotypic response to selection in a similar way.

Although it is possible that extant species variation patterns are fairly similar, it is also possible that stochastic fluctuations in the G-matrix over generations are large enough to render the inferences we might make from extant patterns useless. These fluctuations are possible for many reasons, like segregating alleles with large effects or linkage disequilibrium caused by periods of strong fluctuating selection (Bulmer 1971, Turelli 1988). Although this is certainly a theoretical possibility, the critical question is whether or not changes in G-matrix structure are of sufficient

magnitude to affect evolutionary inferences (Arnold et al. 2008). Fortunately, simulations quantifying these problems suggest that their effect may often be small (Jones et al. 2004, 2012).

The Missing Link: Adaptive Landscapes and the G-Matrix

We have discussed genetic associations, the genetic and developmental origins of modularity, and their influence in evolutionary response. One missing, and arguably the most important, link is the relation between covariation and the adaptive landscape. By examining the relationship between modularity and adaptive landscapes, we can put micro- and macroevolution into a common unifying theory (Arnold 2014, Arnold et al. 2001). This theory not only explains the relationship between development, function, and inheritance in shaping modularity patterns but also allows us to explore the evolution of multivariate phenotypes in deep-time and suggest future research directions.

Although many studies have been conducted at a local scale (within-population), we have precious little information about how adaptive landscapes vary between species (Pfaender et al. 2016). One of the most important developments in the past 30 years of evolutionary theory was the development of multivariate regression methods for inferring individual selection surfaces from multivariate data (Lande & Arnold 1983). Although not without critics (Shaw & Geyer 2010), these methods have allowed researchers to directly investigate adaptive landscapes and to empirically measure selection on multivariate trait sets. In this approach, the adaptive landscape is described by two main terms: a linear term related to directional selection and a quadratic term related to stabilizing/disruptive selection (Lande & Arnold 1983). The quadratic component affects the variances and covariances among traits and, along with directional selection, is thought to be the major evolutionary process shaping modularity (Melo & Marroig 2015). The selection gradient (linear term) is the direction of maximum increase in fitness and is the vector of partial regression coefficients of fitness on traits. This framework can also be used to study directional selection retrospectively: By measuring extant species means and covariance matrices we can estimate ancestral states, and by solving the Lande equation, the selection gradient that would have resulted in the observed diversification can be estimated (Lande 1979, equation 9) (**Figure 2**).

Even though this toolkit for explicitly characterizing selection has now been available for decades, empirical characterizations of multivariate selection are still rare, despite their acknowledged importance. Yet the past 20 years of reconstructing selection gradients from extant diversity have given us two fundamental insights into the nature of multivariate evolution. First, estimates of the strength of both stabilizing and directional selection are usually weaker than we previously assumed (Kingsolver et al. 2001, 2012). But, more importantly, the empirical evidence suggests that the direction of evolutionary divergence and the direction of selection are rarely the same and, often times, present little resemblance (see the section, Does Alignment with Lines of Least Resistance Imply Constraint?). An important work illustrating this notion comes from studies of *Drosophila serrata* (Chenoweth et al. 2010). In a study of sexual selection, the authors note that even though local processes of sexual selection varied considerably across their nine populations, evolutionary divergence occurred primarily along a single trait combination. Variation in sexual selection had little influence on evolutionary divergence. Instead, genetic covariation among traits caused the evolutionary response to be significantly deflected from its optimal path. Studies on *Dalechampia* blossom morphology have also emphasized that only a portion of evolutionary divergence patterns can be accounted for by estimates of external selective factors, such as community composition and availability of resources. Rather, constraints imposed by covariation patterns seem to be essential for our understanding of the evolution of blossom traits (Bolstad et al. 2014, Hansen et al. 2003).

We should point out that selection gradients reported in these retrospective works are net gradients, that is, estimates of the cumulative sum of all selection gradients acting over the generations of divergence. If the G-matrix is stable, this net selection should be a reasonably accurate estimate of the sum of individual gradients (Jones et al. 2004). Another issue with reconstructing selection is that G- and P-matrices are often estimated with substantial error, frequently resulting in poorly conditioned or negative semidefinite matrices. Thus, the inversion step in the reconstruction analysis can lead to very large errors (Marroig et al. 2012). Fortunately, matrix estimation or regularization methods can vastly improve the selection gradient estimates, and these methods should be used whenever selection estimates are made (Marroig et al. 2012, Schäfer & Strimmer 2005).

Evolutionary Change in Simple Landscapes

Most of our knowledge of the relationship between modularity and the adaptive landscape comes from simulation studies. In simulations carried out on simple landscapes, patterns and magnitudes of association among traits affect the direction, magnitude, and rate of evolutionary change under selection (e.g., Marroig & Cheverud 2010). The effect of the G-matrix on evolutionary change depends critically on its structure in relation to the adaptive landscape (Conner 2012, Laughlin & Messier 2015) and can either augment or slow the evolutionary response relative to a situation with fully independent traits. If selection is along dimensions unaligned with modularity/integration patterns, the response is deflected toward the lines of least resistance (Schluter 1996). If selection is aligned with modularity, however, the evolutionary response is greatly facilitated (Beldade et al. 2002, Bolstad et al. 2014). The closer the alignment with the major line of least resistance, the quicker and more direct the evolutionary response (**Figure 3**). However, simulations are highly concordant in showing that these effects are restricted to the microevolutionary scale, and, given sufficient time and a simple adaptive landscape, the population will eventually reach the selective peak, unless there is no genetic variation at all in that direction (an absolute constraint) (Blows & Hoffmann 2005). But theoretical work suggests that even if there is an apparent lack of genetic variation along some dimension, genetic variation is frequently hidden in the form of epistasis that can fuel evolutionary change in subsequent generations (Hansen 2013, Hansen et al. 2006). Therefore, given the possibility of adaptive changes in the G-matrix through time and the understanding that constraints imposed by G-matrices are usually microevolutionary, the emerging picture would be one in which G-matrices should not have any enduring macroevolutionary consequences (termed the transient constraints model, from now on). But what happens when we consider complex adaptive landscapes?

Evolutionary Change in Rugged Landscapes

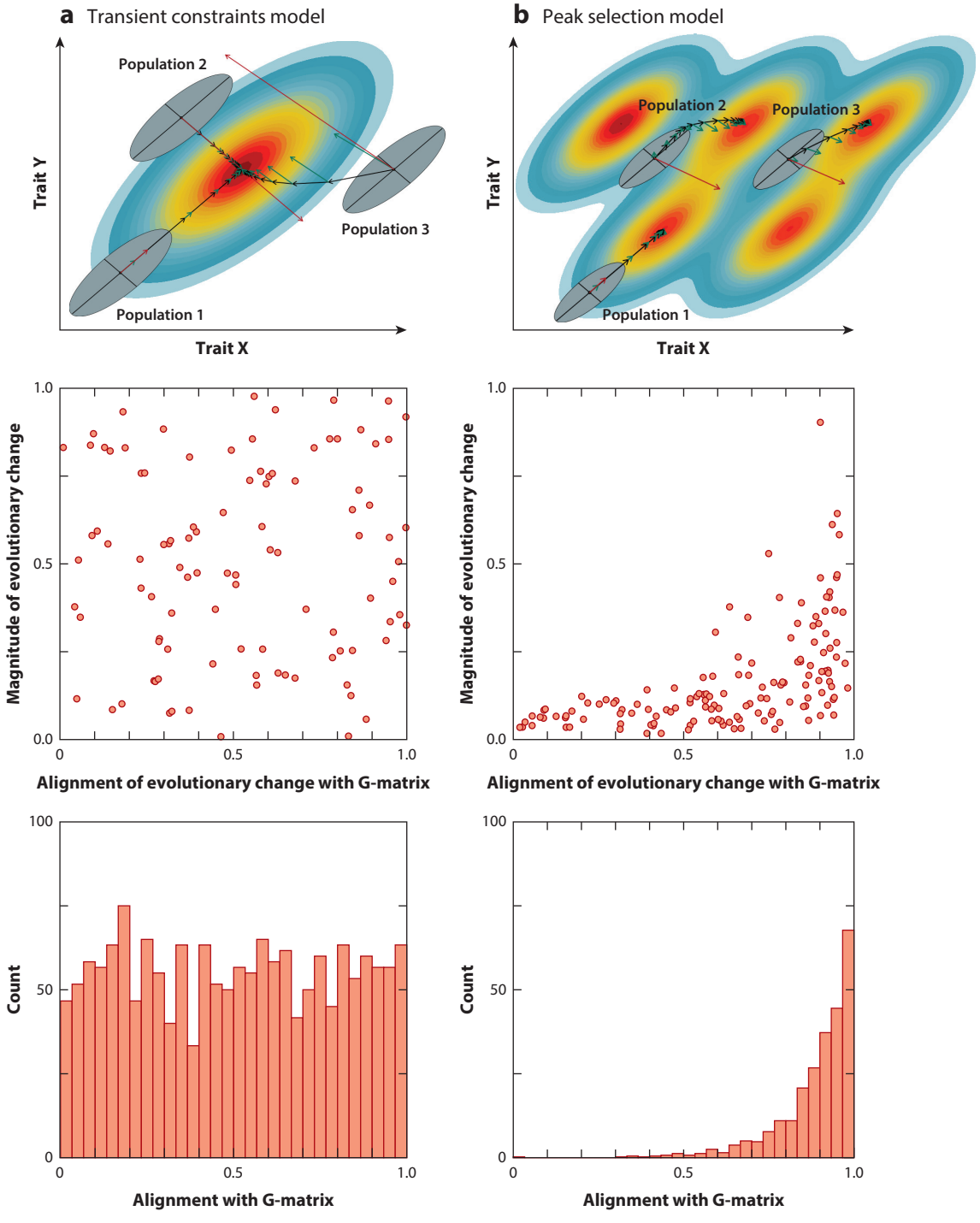
Although single-peaked adaptive landscapes are convenient for model building purposes, adaptive landscapes are thought to be very rugged, that is, they have many adaptive peaks and valleys (Kauffman & Levin 1987, Martin & Wainwright 2013, Wright 1932). When the adaptive landscape is rugged and genetic associations are stable through time, macroevolutionary dynamics are shaped by the interaction between the G-matrix and the adaptive landscape (**Figure 3**). This interaction implies that, in rugged and multiple-peaked adaptive landscapes, the G-matrix can have a major influence in determining which peak will be reached by a given population, even if in theory the effect of the G-matrix is microevolutionary (Steppan et al. 2002). This argument was already present in Lande (1979, p. 407) but in a somewhat opaque formulation: "However, the adaptive topography for each population or species generally has multiple peaks . . . . Genetic correlations

can alter the long-term result of selection by influencing the direction of evolution at critical periods when a population approaches a threshold (or saddlepoint) between adaptive zones, as by random genetic drift or by environmental fluctuations which directly affect the phenotype or alter the adaptive topography." This argument can be easily understood noting that, in evolutionary terms, the distance between the population average position and the peak is not a simple linear (Euclidean) distance between the start position and end position of the species averages but is a weighted distance, with the weight being given by the patterns of genetic association. Given the influence of genetic correlations, the distance of a population from a peak is measured in units of genetic variation. Thus, the closest peak, the peak the population eventually reaches, is not necessarily the highest or even the closest in Euclidian distance but is the closest in genetic-scaled distance. We refer to this idea as the peak selection model.

What would we expect in terms of empirical patterns under each of these scenarios? If G-matrices impose only microevolutionary constraints and population/species eventually reach their single-adaptive peak, we would expect no particular relationship between the magnitude and direction of evolutionary change and its alignment with the G-matrix. We would also not expect any significant alignment of the species response to selection with the major axis of variation of the G-matrix. Evolution in this scenario would depend only on the position of the adaptive peak in relation to the population average. Alternatively, if G-matrices have enduring consequences at the macroevolutionary level by influencing the choice of peak, we would expect an association between the magnitude and direction of evolutionary change and its alignment with the major axis of the G-matrix (see Porto et al. 2015). Furthermore, species diversification should be biased in the directions of highest variation in the G-matrix. Evolutionary change would depend not only on the position of the adaptive peaks in regard to the current position of the population averages but also on the G-matrix structure, which would affect the probability of reaching the various peaks (**Figure 3**). A possible complication of the single-peak model is one in which we have only one peak, but this single adaptive peak is not fixed, instead fluctuating randomly in the morphospace over time (Jones et al. 2012). What would be the expectations under this model? The answer would depend on the frequency and magnitude of the peak fluctuation over time, assuming that the G-matrix is stable (as discussed in the section, Why So Much Interest in the G-Matrix?). If peak displacements are small and rare, the expectations would be more in line with the transient constraints model. Conversely, if peak fluctuations are common or large in magnitude, populations possibly never quite reach equilibrium and traverse the morphospace walking on the line of least resistance, thus approaching the expectations from the rugged peak model.

Does Alignment with Lines of Least Resistance Imply Constraint?

Comparisons of G-matrix orientation with the observed direction of evolutionary change, as described in the previous section, can be a fruitful way to test these ideas. Several studies have compared morphological diversification with available genetic variation: In several instances diversification was aligned with the lines of least resistance, whereas other instances showed diversification in alternate directions (Berner et al. 2010, Marroig & Cheverud 2010, Renaud et al. 2006, Schluter 1996). But this alignment is not necessarily due to constraints, because selection and constraint can act in the same direction (Conner 2012, Marroig and Cheverud 2010). This would imply either that species lie near the axis of major evolvability due not to constraint but to a ridge in the fitness surface (Conner 2012) or that at least some of the available peaks happened to be aligned with that direction and thus the pattern is adaptive (Arnold et al. 2001, Marroig & Cheverud 2010). Likewise, macroevolutionary diversification that is not aligned with variation does not negate the possibility that the G-matrix imposed microevolutionary restrictions—it could



be that the position of adaptive peaks had some other pattern. Perhaps a more complete picture of what we are observing is one close to the peak selection model. Species don't tend to follow the line of least resistance because they are constrained in that direction, in the sense of lacking variation in other directions of the morphospace (Marroig & Cheverud 2005, 2010). Instead, G-matrix and peak distribution interact, thereby making the realized morphospace coverage much smaller than the full range of possibilities. We now turn our attention to whether or not we can gain any information on past peak distribution from comparative quantitative genetic studies.

Differentiating Between Constraints, Co-Selection, and Drift

If the covariation between species is mirrored by the G-matrix, can we attribute this to constraints or to a common pattern of selection and covariation? Notice that examining the potential relationship between the evolutionary change and the divergence time between populations is not enough to separate constraints, selection, and drift. If changes among populations are due to directional selection separating them into different adaptive peaks and if after the initial displacement their averages are kept constant by stabilizing selection, then this sequence of directional and stabilizing selection may result in observed evolutionary rates that are consistent with drift (Lemos et al. 2001). Thus, information on the adaptive landscape and explicit tests for drift are necessary. So, can we examine the alignment of the orientation of the G-matrix with the distribution of peaks in the adaptive landscape? In theory, it should be possible to estimate covariation between selection in different clades (**Figure 4**) on the basis of observed selection gradients, given some assumptions (Felsenstein 1988, Zeng 1988). Although this method gives us access only to the peaks that were eventually reached and are currently occupied by living species, it provides valuable information that can help us to explain whether macroevolution is dominated by constraints or by an interaction between constraints and selection, as in the peak selection model (Marroig & Cheverud 2010).

Figure 3

Left panels (*a*) show the transient constraints model, and right panels (*b*) show the peak selection model. Top left panel shows selection gradients (β) per generation (*green arrows*) and average responses (Δz) per generation (*black arrows*) of populations sharing similar G-matrix structure but at different starting points of a single-peaked adaptive landscape. Red arrows represent the net selection gradient (the sum of all selection gradients). Thus, the alignment between the direction of selection and the orientation of the G-matrix differs for each population. Responses to selection will thus vary between populations, in terms of direction and magnitude. Some populations will evolve rapidly and directly to the peak (population 1), whereas others will evolve slowly (population 2) owing to differences in the amount of variation aligned with selection (evolvability sensu Hansen & Houle 2008). Population 3 will approach the peak in a nonlinear way, and its trajectory will be strongly deflected by the G-matrix in the direction of the line of least resistance. The line of least resistance equals the first principal component of a G-matrix and acts as an attractor of short-term evolutionary responses. We can expand this notion to multivariate systems and think of linear combinations of the first principal components as representing hyperplanes of least resistance. This notion is related to modularity, because principal components are related to modules but do not carry a one-to-one relation with each module (Berner 2011). Usually, principal components are contrasts between modules (positive loadings for one module and negative loadings for the other), and linear combinations between these contrasts define directions of independent change for each module. Note that the net selection gradients (*red arrows*) are much larger when selection is not aligned with the G-matrix's main axis. The top right panel shows the same three populations but in a rugged adaptive landscape. In this scenario, populations won't always evolve to the closest peak (Euclidean distance) but instead will evolve to those that are closest given the covariation among traits. The central panels illustrate the predictions of each model (transient constraints, *left*; peak selection, *right*). Each point represents one species, with Y being the total magnitude of evolution, and X being the alignment of the evolutionary response (z) with the G-matrix. In the transient constraints model (*a*), you would not expect any particular relationship between the magnitude of evolutionary change and its direction, because every species would eventually reach the peak. Conversely, under the peak selection model (*b*), species' evolutionary trajectories may or may not be aligned with the G-matrix, but the magnitude of evolutionary change will be small when not aligned. Bottom panels show the predictions for both models (transient constraints, *left*; peak selection, *right*) of the number of species observed in terms of their evolutionary response (z) and their alignment with the G-matrix.

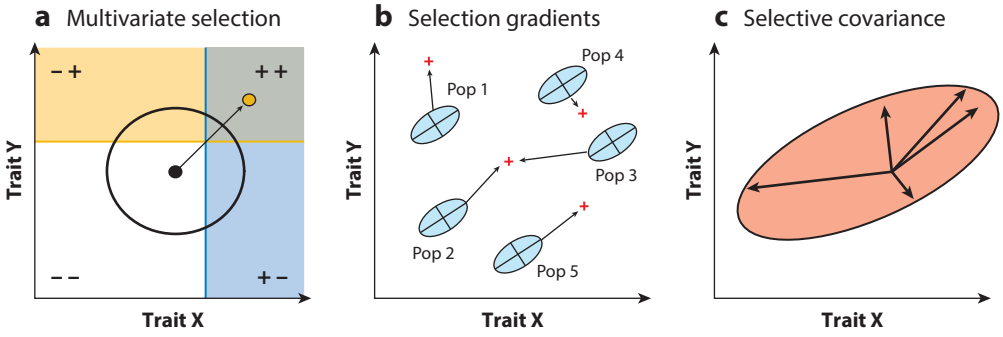


Figure 4

Traits will evolve together either because they are inherited together (G-matrix) or because they are selected together (selective covariance). Panel *a* illustrates the idea of selective covariance. Traits X and Y are genetically independent. The black dot indicates the average before selection, and the plus (+) and minus (–) signs indicate the direction of increase in fitness for each trait. Thus, selection is favoring the joint increase of X and Y, and the population will evolve a new average phenotype (*yellow dot*). The term selective covariance was coined by Felsenstein (1988) (see also Zeng 1988). If we have evidence that the G-matrix is relatively stable during macroevolution, the equation $V = GCG$ captures the covariance of changes in the averages of the species (V-matrix) in terms of its two potential (nonexcluding) sources: inheritance (G-matrix) and selective covariance (C-matrix). Theoretically, if we have a reasonable estimate of the G-matrix and of the phylogenetic relationships, we can compute the V-matrix and thus solve the Zeng–Felsenstein equation to compute $C = G^{-1}VG^{-1}$, where C is the covariance of slopes of log W (e.g., the covariance among the selection gradients operating upon each species). Panels *b* and *c* illustrate how to capture the C-matrix, with selection gradients (β) shown in panel *b* and the selective covariance matrix represented in panel *c* (*red ellipse*). Different populations are shown in panel *b*.

Under transient constraints, we should not expect any alignment between the G-matrix and the selective covariance matrix (covariance between selection gradients) because, given enough time, the populations should eventually reach their respective adaptive peaks. Conversely, under peak selection, we would expect an alignment between the G-matrix and selective covariance matrix. We are aware of only two such tests reported to date (Hohenlohe & Arnold 2008, Marroig & Cheverud 2010).

Comparative approaches establishing a relationship between genetic lines of least resistance and divergence patterns have other important limitations. Most significantly, random genetic drift can create an association between the orientation of the G-matrix and the patterns of between-species divergence, given stable patterns of genetic covariation. This association occurs because evolutionary divergence under drift is expected to be proportional to the ancestral pattern of variation and covariation among traits (**Figure 2**); therefore, an observed association between the orientation of the G-matrix and divergence can be a direct product of neutral evolution. Although most biologists would agree that morphology is usually under selection, it is useful to examine the potential consequences of drift and how it relates to modularity. Simulation work suggests that if genetic drift is the only operating evolutionary process, modularity patterns would not be stable and patterns of association would vary widely across closely related populations or taxa (Jones et al. 2003, Melo & Marroig 2015). This is clearly not observed in nature. But what if modularity is maintained by stabilizing selection and trait means are free to change by genetic drift? In this situation, divergence among populations would be largest along directions in which ancestral genetic variation is abundant and smaller in directions of low ancestral variation (Arnold et al. 2001, Lande 1976). There are methods for distinguishing drift from selection in quantitative traits (Ackermann & Cheverud 2004, Bartoszek et al. 2012, Hohenlohe & Arnold 2008, Karhunen et al. 2013), but most of them are not well suited to high dimensional systems, do not take the influence of genetic covariation into account, or require a large number of individuals

distributed in a large number of species. For example, Hohenlohe & Arnold's (2008) approach explicitly models evolution under drift to predict a probability distribution for the divergence of population averages, given a phylogeny, the G-matrix, and an estimate of effective population size. This elegant solution can test whether divergence among groups is compatible with drift and the current G-matrix, but it can be applied in full force only with two characters at a time. With few exceptions (Bartoszek et al. 2012, Hohenlohe & Arnold 2008), most phylogenetic methods fail to take genetic covariation into account, thereby limiting our understanding of macroevolution. By modeling evolution under a univariate Brownian motion model, for example, we assume that no selection is operating (but see Butler & King 2004, Hansen 1997) and that traits are evolving independently. Some advances have been made in the past decade (Bartoszek et al. 2012, Cressler et al. 2015, Hohenlohe & Arnold 2008), but we still lack comparative methods that balance the external aspect of selection (niche shifts on Ornstein-Uhlenbeck models) (Butler & King 2004, Hansen 1997) with the populational consequences of modularity.

CONCLUSION

Placing micro- and macroevolution into a common framework is essential for our understanding of the influence of genetic and developmental constraints on multivariate evolution. Quantitative genetic theory has long been interested in the variational properties of organisms, and recent studies using the conceptual umbrella of modularity have extrapolated its breadth to include long-term evolutionary change. Although empirical results led us to discard the notion that variational patterns are set in stone and act as absolute constraints, they have also made us abandon the ideas that adaptive landscapes can be characterized by simple and stable selective peaks or that variational properties are largely unimportant considerations for evolutionary change. Embracing the dynamic nature of variational patterns, their context dependency, and their relationships with genetics, development, and evolution will allow us to bridge these two levels of the hierarchy in a systematic way. One of the challenges going forward is the incorporation of mechanistic models of development into models of how variation emerges and how it influences the shape of population variation and adaptive landscapes. Another major challenge will be discriminating the relative contributions of constraints, selection, and neutral processes in determining the path of multivariate evolution. We propose that this challenge will be met only when we know more about the true shape of adaptive landscapes, including the number, height, and distribution of peaks (see Laughlin & Messier 2015, Pfaender et al. 2016), and when we incorporate modularity into our thinking.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Thanks to Dr. Douglas Futuyma, Dr. Jason Wolf, Dr. Stevan J. Arnold, Madeline Keleher, and Devin Dobias for discussions and helpful comments and to Ana Paula Assis and Guilherme Garcia for help with figures. We would like to thank Barbara Costa, Glauco Machado, and Paulo Guimarães for discussions that help the macroevolutionary part of the paper to become clearer and to all people in the mammalian evolution laboratory. D.M. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant 2014/26262-4). A.P. was supported

by the National Institute of Dental and Craniofacial Research of the National Institutes of Health (award number F31DE024944). G.M. was supported by FAPESP (grant 2011/14295-7).

LITERATURE CITED

Ackermann RR, Cheverud JM. 2004. Detecting genetic drift versus selection in human evolution. *PNAS* 101(52):17946–51

Adams D, Rohlf F, Slice D. 2013. A field comes of age: geometric morphometrics in the 21st century. *Hystrix Ital. J. Mammal.* 24:7–14

Alberch P. 1991. From genes to phenotype: dynamical systems and evolvability. *Genetica* 84(1):5–11

Armbruster S, Pelabon C, Hansen T, Mulder C. 2004. Floral integration, modularity, and accuracy: distinguishing complex adaptations from genetic constraints. In *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*, ed. M Pigliucci, K Preston, pp. 23–49. Oxford, UK: Oxford Univ. Press

Arnold SJ. 2014. Phenotypic evolution: the ongoing synthesis (American Society of Naturalists address). *Am. Nat.* 183(6):729–46

Arnold SJ, Bürger R, Hohenlohe PA, Ajie BC, Jones AG. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* 62(10):2451–61

Arnold SJ, Pfrender ME, Jones AG. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112–113:9–32

Ayroles JF, Buchanan SM, O'Leary C, Skutt-Kakaria K, Grenier JK, et al. 2015. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *PNAS* 112(21):6706–11

Ayroles JF, Carbone MA, Stone EA, Jordan KW, Lyman RF, et al. 2009. Systems genetics of complex traits in *Drosophila melanogaster*. *Nat. Genet.* 41(3):299–307

Bartoszek K, Pienaar J, Mostad P, Andersson S, Hansen TF. 2012. A phylogenetic comparative method for studying multivariate adaptation. *J. Theor. Biol.* 314:204–15

Beldade P, Brakefield PM. 2002. The genetics and evo-devo of butterfly wing patterns. *Nat. Rev. Genet.* 3(6):442–52

Beldade P, Koops K, Brakefield PM. 2002. Modularity, individuality, and evo-devo in butterfly wings. *PNAS* 99(22):14262–67

Berg RL. 1960. The ecological significance of correlation pleiades. *Evolution* 14(2):171–80

Berner D. 2011. Size correction in biology: How reliable are approaches based on (common) principal component analysis? *Oecologia* 166(4):961–71

Berner D, Kaeuffer R, Grandchamp A-C, Raeymaekers JAM, Räsänen K, Hendry AP. 2011. Quantitative genetic inheritance of morphological divergence in a lake-stream stickleback ecotype pair: implications for reproductive isolation. *J. Evol. Biol.* 24(9):1975–83

Berner D, Stutz WE, Bolnick DI. 2010. Foraging trait (co)variances in stickleback evolve deterministically and do not predict trajectories of adaptive diversification. *Evolution* 64(8):2265–77

Björklund M, Husby A, Gustafsson L. 2013. Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. *J. Evol. Biol.* 26(1):1–13

Blows MW, Hoffmann AA. 2005. A reassessment of genetic limits to evolutionary change. *Ecology* 86(6):1371–84

Bolstad GH, Cassara JA, Márquez E, Hansen TF, van der Linde K, et al. 2015. Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *PNAS* 112(43):13284–89

Bolstad GH, Hansen TF, Pélabon C, Falahati-Anbaran M, Pérez-Barrales R, Armbruster WS. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. *Philos. Trans. R. Soc. B* 369(1649):20130255

Bookstein FL. 1997. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge, UK: Cambridge Univ. Press

Bulmer MG. 1971. The effect of selection on genetic variability. *Am. Nat.* 105(943):201–11

Butler MA, King AA. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164(6):683–95

Calsbeek B, Goodnight CJ. 2009. Empirical comparison of G matrix test statistics: finding biologically relevant change. *Evolution* 63(10):2627–35

Chenoweth SF, Rundle HD, Blows MW. 2010. The contribution of selection and genetic constraints to phenotypic divergence. *Am. Nat.* 175(2):186–96

Cheverud JM. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution* 36(3):499

Cheverud JM. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* 110(2):155–71

Cheverud JM. 1989. A comparative analysis of morphological variation patterns in the papionins. *Evolution* 43(8):1737

Cheverud JM. 1996. Developmental integration and the evolution of pleiotropy. *Am. Zool.* 36(1):44–50

Cheverud JM, Marroig G. 2007. Comparing covariance matrices: random skewers method compared to the common principal components model. *Genet. Mol. Biol.* 30(2):461–69

Cheverud JM, Richtsmeier JT. 1986. Finite-element scaling applied to sexual dimorphism in rhesus macaque (*Macaca mulatta*) facial growth. *Syst. Biol.* 35(3):381–99

Conner JK. 2012. Quantitative genetic approaches to evolutionary constraint: how useful? *Evolution* 66(11):3313–20

Cressler CE, Butler MA, King AA. 2015. Detecting adaptive evolution in phylogenetic comparative analysis using the Ornstein-Uhlenbeck model. *Syst. Biol.* 64(6):953–68

Delph LF, Steven JC, Anderson IA, Herlihy CR, Brodie ED III. 2011. Elimination of a genetic correlation between the sexes via artificial correlational selection. *Evolution* 65(10):2872–80

Eroukhmanoff F, Svensson EI. 2011. Evolution and stability of the G-matrix during the colonization of a novel environment. *J. Evol. Biol.* 24(6):1363–73

Felsenstein J. 1988. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* 19:445–71

Garcia G, Hingst-Zaher E, Cerqueira R, Marroig G. 2014. Quantitative genetics and modularity in cranial and mandibular morphology of *Calomys expulsus*. *Evol. Biol.* 41(4):619–36

Goswami A, Finarelli JA. 2016. EMMLi: a maximum likelihood approach to the analysis of modularity. *Evolution* 70(7):1622–37

Grant PR, Grant BR. 1995. Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49(2):241

Hallgrímsson B, Jamniczky HA, Young NM, Rolian C, Schmidt-Ott U, Marcucio RS. 2012. The generation of variation and the developmental basis for evolutionary novelty. *J. Exp. Zool. B Mol. Dev. Evol.* 318(6):501–17

Hallgrímsson B, Jamniczky H, Young NM, Rolian C, Parsons TE, et al. 2009. Deciphering the palimpsest: studying the relationship between morphological integration and phenotypic covariation. *Evol. Biol.* 36(4):355–76

Hansen TF. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51(5):1341–51

Hansen TF. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *BioSystems* 69(2–3):83–94

Hansen TF. 2013. Why epistasis is important for selection and adaptation. *Evolution* 67(12):3501–11

Hansen TF, Alvarez-Castro JM, Carter AJR, Hermisson J, Wagner GP. 2006. Evolution of genetic architecture under directional selection. *Evolution* 60(8):1523–36

Hansen TF, Houle D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Biol.* 21(5):1201–19

Hansen TF, Pélabon C, Armbruster WS, Carlson ML. 2003. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. *J. Evol. Biol.* 16(4):754–66

Hazel LN. 1943. The genetic basis for constructing selection indexes. *Genetics* 28(6):476–90

Hohenlohe PA, Arnold SJ. 2008. MIPoD: a hypothesis-testing framework for microevolutionary inference from patterns of divergence. *Am. Nat.* 171(3):366–85

Houle D, Meyer K. 2015. Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. *J. Evol. Biol.* 28(8):1542–49

Ihmels J, Friedlander G, Bergmann S, Sarig O, Ziv Y, Barkai N. 2002. Revealing modular organization in the yeast transcriptional network. *Nat. Genet.* 31(4):370–77

Jones AG, Arnold SJ, Bürger R. 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* 57(8):1747–60
Jones AG, Arnold SJ, Bürger R. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* 58(8):1639–54
Jones AG, Bürger R, Arnold SJ. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* 5:3709
Jones AG, Bürger R, Arnold SJ, Hohenlohe PA, Uyeda JC. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. *J. Evol. Biol.* 25(11):2210–31
Juenger T, Pérez-Pérez JM, Bernal S, Micol JL. 2005. Quantitative trait loci mapping of floral and leaf morphology traits in *Arabidopsis thaliana*: evidence for modular genetic architecture. *Evol. Dev.* 7(3):259–71
Karhunen M, Merilä J, Leinonen T, Cano JM, Ovaskainen O. 2013. DRIFTSEL: an R package for detecting signals of natural selection in quantitative traits. *Mol. Ecol. Resour.* 13(4):746–54
Kauffman S, Levin S. 1987. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* 128(1):11–45
Kendall DG. 1984. Shape manifolds, procrustean metrics, and complex projective spaces. *Bull. Lond. Math. Soc.* 16:81–121
Kingsolver JG, Diamond SE, Siepielski AM, Carlson SM. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* 26(5):1101–18
Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, et al. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157(3):245–61
Klingenberg CP. 2009. Morphometric integration and modularity in configurations of landmarks: tools for evaluating a priori hypotheses. *Evol. Dev.* 11(4):405–21
Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314
Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33(1):402
Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* 37(6):1210
Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559
Laughlin DC, Messier J. 2015. Fitness of multidimensional phenotypes in dynamic adaptive landscapes. *Trends Ecol. Evol.* 30(8):487–96
Lele S, Richtsmeier JT. 1991. Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. *Am. J. Phys. Anthropol.* 86:415–27
Lemos B, Marroig G, Cerqueira R. 2001. Evolutionary rates and stabilizing selection in large-bodied opossum skulls (Didelphimorphia: Didelphidae). *J. Zool.* 255:181–89
Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10(8):565–77
MacMahon M, Garlaschelli D. 2015. Community detection for correlation matrices. *Phys. Rev. X* 5(2):021006
Magwene PM. 2001. New tools for studying integration and modularity. *Evolution* 55(9):1734–45
Márquez EJ. 2008. A statistical framework for testing modularity in multidimensional data. *Evolution* 62(10):2688–708
Márquez EJ, Cabeen R, Woods RP, Houle D. 2012. The measurement of local variation in shape. *Evol. Biol.* 39(3):419–39
Marroig G, Cheverud JM. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. *Evolution* 55(12):2576–600
Marroig G, Cheverud JM. 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. *Evolution* 59(5):1128–42
Marroig G, Cheverud J. 2010. Size as a line of least resistance II: direct selection on size or correlated response due to constraints? *Evolution* 64(5):1470–88
Marroig G, de Vivo M, Cheverud JM. 2004. Cranial evolution in sakis (Pithecia, Platyrrhini) II: evolutionary processes and morphological integration. *J. Evol. Biol.* 17:144–55

Marroig G, Melo DAR, Garcia G. 2012. Modularity, noise, and natural selection. *Evolution* 66(5):1506–24
Martin CH, Wainwright PC. 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* 339(6116):208–11
Matamoro-Vidal A, Salazar-Ciudad I, Houle D. 2015. Making quantitative morphological variation from basic developmental processes: Where are we? The case of the *Drosophila* wing. *Dev. Dyn.* 244(9):1058–73
Maxwell TJ, Ballantyne CM, Cheverud JM, Guild CS, Ndumele CE, Boerwinkle E. 2013. APOE modulates the correlation between triglycerides, cholesterol, and CHD through pleiotropy, and gene-by-gene interactions. *Genetics* 195(4):1397–405
Mayr E. 1963. *Animal Species and Evolution.* Cambridge, MA: Harvard Univ. Press
Melo D, Marroig G. 2015. Directional selection can drive the evolution of modularity in complex traits. *PNAS* 112(2):470–75
Mitteroecker P. 2009. The developmental basis of variational modularity: insights from quantitative genetics, morphometrics, and developmental biology. *Evol. Biol.* 36(4):377–85
Mitteroecker P, Bookstein F. 2007. The conceptual and statistical relationship between modularity and morphological integration. *Syst. Biol.* 56(5):818–36
Olson R, Miller E. 1958. *Morphological Integration*. Chicago: Univ. Chicago Press
Paaby AB, Rockman MV. 2013. The many faces of pleiotropy. *Trends Genet.* 29(2):66–73
Pavlicev M, Cheverud JM, Wagner GP. 2011a. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. *Proc. R. Soc. B* 278(1713):1903–12
Pavlicev M, Hansen TF. 2011. Genotype-phenotype maps maximizing evolvability: modularity revisited. *Evol. Biol.* 38(4):371–89
Pavlicev M, Kenney-Hunt JP, Norgard EA, Roseman CC, Wolf JB, Cheverud JM. 2008. Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution* 62(1):199–213
Pavlicev M, Norgard EA, Fawcett GL, Cheverud JM. 2011b. Evolution of pleiotropy: Epistatic interaction pattern supports a mechanistic model underlying variation in genotype–phenotype map. *J. Exp. Zool. B Mol. Dev. Evol.* 316(5):371–85
Pavličev M, Cheverud JM. 2015. Constraints evolve: Context dependency of gene effects allows evolution of pleiotropy. *Annu. Rev. Ecol. Evol. Syst.* 46:413–34
Pearson K. 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. A* 187:253–318
Pepin KM, Samuel MA, Wichman HA. 2006. Variable pleiotropic effects from mutations at the same locus hamper prediction of fitness from a fitness component. *Genetics* 172(4):2047–56
Pfaender J, Hadiaty RK, Schliewen UK, Herder F. 2016. Rugged adaptive landscapes shape a complex, sympatric radiation. *Proc. R. Soc. B* 283(1822):20152342
Phillips PC, Whitlock MC, Fowler K. 2001. Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* 158(3):1137–45
Polly PD. 2008. Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution? *Evol. Biol.* 35(2):83–96
Porto A, de Oliveira FB, Shirai LT, De Conto V, Marroig G. 2009. The evolution of modularity in the mammalian skull I: morphological integration patterns and magnitudes. *Evol. Biol.* 36(1):118–35
Porto A, Sebastião H, Pavan SE, VandeBerg JL, Marroig G, Cheverud JM. 2015. Rate of evolutionary change in cranial morphology of the marsupial genus *Monodelphis* is constrained by the availability of additive genetic variation. *J. Evol. Biol.* 28(4):973–85
Porto A, Shirai LT, de Oliveira FB, Marroig G. 2013. Size variation, growth strategies, and the evolution of modularity in the mammalian skull. *Evolution* 67(11):3305–22
Reichardt J, Bornholdt S. 2006. Statistical mechanics of community detection. *Phys. Rev. E* 74(1):016110
Renaud S, Auffray J-C, Michaux J. 2006. Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to selection in fossil rodents. *Evolution* 60(8):1701–17
Riedl R. 1978. *Order in Living Systems: A Systems Analysis of Evolution*. New York: Wiley
Salazar-Ciudad I, Jernvall J. 2002. A gene network model accounting for development and evolution of mammalian teeth. *PNAS* 99(12):8116–20

Salazar-Ciudad I, Jernvall J. 2010. A computational model of teeth and the developmental origins of morphological variation. *Nature* 464(7288):583–86
Schäfer J, Strimmer K. 2005. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* 4:32
Schluter D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50(5):1766
Shaw RG, Geyer CJ. 2010. Inferring fitness landscapes. *Evolution* 64(9):2510–20
Simpson GG. 1944. *Tempo and Mode in Evolution*. New York: Columbia Univ. Press
Simpson G. 1958. Book reviews: *Morphological Integration* by Everett C. Olson, Robert L. Miller. *Science* 128(3316):138
Stearns FW. 2010. One hundred years of pleiotropy: a retrospective. *Genetics* 186(3):767–73
Steppan SJ, Phillips PC, Houle D. 2002. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* 17(7):320–27
Templeton AR, Crease TJ, Shah F. 1985. The molecular through ecological genetics of abnormal abdomen in *Drosophila mercatorum*. I. Basic genetics. *Genetics* 111(4):805–18
Turelli M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution* 42(6):1342
van der Linde K, Houle D. 2009. Inferring the nature of allometry from geometric data. *Evol. Biol.* 36(3):311–22
Wagner GP. 2007. The developmental genetics of homology. *Nat. Rev. Genet.* 8(6):473–79
Wagner GP, Altenberg L. 1996. Perspective: complex adaptations and the evolution of evolvability. *Evolution* 50(3):967
Wagner GP, Pavlicev M, Cheverud JM. 2007. The road to modularity. *Nat. Rev. Genet.* 8(12):921–31
Wang Z, Liao B-Y, Zhang J. 2010. Genomic patterns of pleiotropy and the evolution of complexity. *PNAS* 107(42):18034–39
Watson RA, Wagner GP, Pavlicev M, Weinreich DM, Mills R. 2014. The evolution of phenotypic correlations and "developmental memory." *Evolution* 68(4):1124–38
Weldon WFR. 1893. On certain correlated variations in *Carcinus maenas*. *Proc. R. Soc. B* 54(326–330):318–29
Whitlock MC, Phillips PC, Fowler K. 2002. Persistence of changes in the genetic covariance matrix after a bottleneck. *Evolution* 56(10):1968–75
Wolf JB, Leamy LJ, Routman EJ, Cheverud JM. 2005. Epistatic pleiotropy and the genetic architecture of covariation within early and late-developing skull trait complexes in mice. *Genetics* 171(2):683–94
Wright S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. VI Int. Congr. Genet.* 1:356–66
Young NM, Hallgrímsson B. 2005. Serial homology and the evolution of mammalian limb covariation structure. *Evolution* 59(12):2691–704
Zeng Z-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. *Evolution* 42(2):363–74

Annual Review of Ecology, Evolution, and Systematics

Volume 47, 2016

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Ecology, Evolution, and Systematics* articles may be found at http://www.annualreviews.org/errata/ecolsys