Comparative quantitative genetics: evolution of the G matrix

Scott J. Steppan, Patrick C. Phillips and David Houle

Quantitative genetics provides one of the most promising frameworks with which to unify the fields of macroevolution and microevolution. The genetic variance–covariance matrix (G) is crucial to quantitative genetic predictions about macroevolution. In spite of years of study, we still know little about how G evolves. Recent studies have been applying an increasingly phylogenetic perspective and more sophisticated statistical techniques to address G matrix evolution. We propose that a new field, comparative quantitative genetics, has emerged. Here we summarize what is known about several key questions in the field and compare the strengths and weaknesses of the many statistical and conceptual approaches now being employed. Past studies have made it clear that the key question is no longer whether G evolves but rather how fast and in what manner. We highlight the most promising future directions for this emerging field.

10.1016/S0169-5347(02)02505-3

Organisms are complex systems comprising interacting characters underlain by shared functional, developmental and genetic processes. Within quantitative genetics (the study of inheritance at the phenotypic level), these relationships are summarized in the additive genetic variance–covariance matrix G (see Glossary). The usefulness of the quantitative genetic approach to long-term evolution depends, to a large extent, on whether G remains constant or evolves in a predictable manner. For this reason, quantitative geneticists have increasingly turned their attention to the evolution of G. Together with natural selection (the adaptive landscape) it determines the direction and rate of evolution.

The most productive approach to the study of evolutionary change is dictated by the importance of genetic details in determining the nature of that change. In some cases, genetics might be irrelevant, and evolution might be best approached as an optimization problem [1]. In other cases, only genetic mechanisms might be worth studying [2]. Quantitative genetics is useful for intermediate cases where genetics matters, but where genetic details do not. The basic quantitative genetic model (Box 1) captures the influence of genetics through G and indirectly through the selection gradient, which depends on the phenotypic matrix P [3]. If G is stable, it can be used to predict the evolutionary potential of a population or to reconstruct the form of selection that has led to divergence among populations [4] (Box 1). Quantitative genetic parameters can also be integrated with phylogenetic information within a likelihood framework to test more precisely for adaptation [5]. If stochastic events, such as genetic drift, fluctuating adaptive landscapes and rare mutations, are more important, then quantitative genetics might not be informative and macroevolution might be decoupled from microevolution. Resolution of this issue is crucial to evolutionary biology as a whole.

Until recently, the usefulness of a quantitative genetic approach to evolution has been asserted or rejected mostly on faith. Neither the high-quality data nor the analytical tools to evaluate possible changes in G have been available. Here, we highlight recent advances that are beginning to allow informative comparisons of G matrices and discuss the questions of if, how, how fast and why G might evolve. We suggest that a new field of study has emerged, comparative quantitative genetics, which has built upon traditional comparisons of genetic variances and covariances but which is distinguished by incorporating phylogenetic information using the comparative method and an emphasis on covariation among traits.

Does G evolve?
Yes. With some important statistical caveats in mind (Box 2), there are clearly some cases where G matrices, or some of their elements, are unequal [3,6–8]. The significant changes in G sometimes detected by rather small studies imply that real differences are frequently large. Laboratory studies have demonstrated significant divergence at the population level given strong selection [3,9] and/or drift [10,11]. Although matrix correlations do not test the hypothesis of inequality, nonsignificant matrix correlations can be interpreted as evidence for departures from equality, if one has confidence in the precision of the estimates. Although comparable studies at multiple systematic levels are few, comparisons among rodent genera [12,13] have shown nonsignificant correlations, whereas comparisons among and within species were significantly correlated [14] (see reviews in [7,8,15]). Comparisons within species usually show significant correlation or insignificant differences [7,8,15]. Comparisons of P matrices find significant differences even more frequently [16–19]. In summary, one cannot assume that G is constant [6,20].

How do G matrices differ?
Although G can change, understanding those aspects of G that change could allow many informative predictions about evolution to be made. For example,
Box 1. Introduction to the $G$ matrix and quantitative genetics

Quantitative genetics provides a means for predicting the evolution of suites of traits given information about directional selection and the degree of resemblance among relatives. When only one character is selected, say $z_1$, the response to selection is predicted by the familiar breeder’s equation (Eqn I),

$$\Delta z_1 = (G_z/P_z) S_1$$  \hspace{1cm} (Eqn I)

where $\overline{z}_1$ is the population mean; $G_z$ is the additive genetic variance in trait 1 and sums up the degree of resemblance between relatives; $P_z$ is the phenotypic variance, and $S_1$ is the covariance between $z_1$ and fitness $[a]$. Alternatively, this equation can be represented as (Eqn II),

$$\Delta z_1 = G_z(S_z/P_z) = G_z b_1$$  \hspace{1cm} (Eqn II)

where $b_1$ is the slope of the regression of fitness on trait 1. If $z_1$ is genetically correlated with any other traits, then the change in frequencies of genotypes affecting $z_1$ will also affect these other traits. This indirect response of another trait, say $z_2$, to selection on $z_1$ is $\Delta z_2 = G_{z_1z_2}$, where $G_{z_1z_2}$ is the additive genetic covariance between $z_1$ and $z_2$.

In general, directional selection can affect more than one trait, so our focal trait is affected both directly by selection on that trait and by the selection on all other traits correlated with it. The result is a complicated bookkeeping problem solved by means of matrix algebra. The vector of responses to selection is (Eqn III)

$$\Delta \overline{z} = GP \beta$$  \hspace{1cm} (Eqn III)

where $G$ is the additive genetic variance–covariance matrix, $P$ is the phenotypic variance–covariance matrix, and $S$ is the vector of covariances between traits and fitness. Variance–covariance matrices are square symmetric matrices with as many rows and columns as there are traits under study. The diagonal entries are the variances, and the off-diagonal elements give the covariances between traits. Equivalently (Eqn IV),

$$P^{-1}S = \beta$$  \hspace{1cm} (Eqn IV)

where $\beta$ is the vector of partial regression coefficients of fitness on the traits. The elements of $G$ reflect the relationship of each trait to fitness, holding the values of other traits constant. Lande [b,c] extended this multivariate approach and was the first to apply it to evolutionary problems. $G$ is useful for predicting which kinds of evolutionary changes are most readily accomplished. $G$ deflects the response to selection toward those trait combinations that have more genetic variance. $G$ will therefore affect the amount of time required to reach a novel state and could determine which state the population will ultimately achieve [b,d] (Fig. I). Persistent absence of additive variation for particular combinations of phenotypes would suggest that evolution in certain directions in phenotype space is not possible [b,e]. If $G$ and the adaptive landscape are indeed constant over long periods, $G$ might be used to predict the evolutionary potential of a population or to reconstruct the form of selection that has led to divergence among populations.

References
A major issue for any comparative quantitative genetic study is statistical power. The genetic variance components that comprise $G$, the genetic variance–covariance matrix, have large sampling errors [a], and so measurements of hundreds of families are usually necessary to provide reasonable power for comparisons. By contrast, sample size for $P$, the phenotypic variance–covariance matrix, is the number of individuals, where sample size for $G$ depends on the number of families, usually far less than the number of individuals. Estimating $G$ normally requires controlled breeding programs; $P$ does not. Because most $G$ matrix studies use fewer families, studies are biased towards confirming the null hypothesis. Interpretation is, however, complicated by the use of diametrically opposed null hypotheses, common among older methods of comparing $G$ matrices. Matrix correlations test the null hypothesis of no similarity between matrices [b], whereas maximum-likelihood [c] or element-by-element comparisons test the null hypothesis that matrices or a subset of their elements are equal. Few studies have adequately addressed limitations of power when trying to compare covariances [d,e], and the power of more versatile methods, such as common principal components analysis (CPCA), is currently unknown [f].

Findings of matrix similarity are also highly dependent on the model being tested. Principal components analysis (PCA), the parent technique on which CPCA depends, transforms the data from the space of the original variables, which are correlated, to a set of vectors that are uncorrelated. It captures all of the variation in the original data, whilst concentrating the variation explained in a few vectors. The forte of PCA is therefore summarizing high-dimensional data with fewer, uncorrelated variables. Flury developed CPCA to summarize multigroup data in as few vectors as possible [g–j], but evolutionary biologists often have the loftier goal of diagnosing and understanding the differences between matrices, and the method has significant shortcomings for this purpose [f].

The default implementation of CPCA orders vectors to be compared by the amount of variance explained. If these first vectors differ, matrices are declared unrelated. It is biologically plausible that populations might differ in the first vector, often size in morphological data sets, but have similarities in other aspects of variation. CPC vectors can be considered in any order, and the Phillips software [k] allows such reordering. PCA also constrains all of the vectors to be orthogonal (uncorrelated), so the vectors with large amounts of variation constrain the directions of all other vectors. Flury [h,l] proposed a more general approach, called common space analysis, which, in principle, allows any set of vectors to be compared. We know of no other implementations of common space analysis. Alternative methods of finding hidden similarities in matrix structure are needed.

References

Table 1. Methods for matrix comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>Approach</th>
<th>Strengths</th>
<th>Weaknesses</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element x element</td>
<td>t-test Correlation, permutation</td>
<td>Detailed, isolates specific elements</td>
<td>No synthesis, ignores nonindependence of elements</td>
<td>[6,55]</td>
</tr>
<tr>
<td>Matrix correlation</td>
<td>Correlation, permutation tests</td>
<td>Overall measure of similarity</td>
<td>Does not distinguish among many types of difference; ignores proportional changes; pairwise comparisons only; easily influenced by a few shared values; improper hypothesis-testing framework</td>
<td>[12–14,54,56]</td>
</tr>
<tr>
<td>Matrix regression</td>
<td>Regression</td>
<td>Estimates of proportionality</td>
<td>Can be strongly influenced by outliers (especially in covariance matrices); improper hypothesis-testing framework</td>
<td>[4,15]</td>
</tr>
<tr>
<td>Disparity</td>
<td></td>
<td>Overall measure of difference, most easily applied to phylogenetic data</td>
<td>No clear metric; no integral statistical model</td>
<td>[33]</td>
</tr>
<tr>
<td>Maximum likelihood</td>
<td>Likelihood</td>
<td>Statistical power, applicable at several levels</td>
<td>Pairwise comparisons only, does not compare matrix structure</td>
<td>[3,6,23,57]</td>
</tr>
<tr>
<td>CPC for dependent vectors</td>
<td>Principal components</td>
<td>Statistical power, hierarchy of models; multiple comparisons</td>
<td>Orthogonality of components might not reflect biology; does not incorporate nonindependence owing to phylogeny</td>
<td>[8,11,16–19,24,25,29–31,40]</td>
</tr>
<tr>
<td>Matrix pattern</td>
<td>Correlation</td>
<td>Based on CPC method; takes extra covariance patterns (e.g. growth, environment) into account</td>
<td>Same limitations as CPC approach, with even more restrictions on the pattern of shared relationships across covariance sets; has yet to be extended for genetic covariance components</td>
<td>[58,59]</td>
</tr>
<tr>
<td>Confirmatory factor analysis</td>
<td>Factor analysis; linear models</td>
<td>Nonparametric model of matrix structure derived from functional and/or developmental models</td>
<td>Shares all of the problems of matrix correlation; unclear how to compare different models</td>
<td>[60]</td>
</tr>
</tbody>
</table>

*Please note that the references are examples and are not necessarily comprehensive.

unrelated) with more inclusive clades, rather than a sequential loss of the smaller eigenvectors. That is, PC1 often differs among very similar matrices, leading CPCA to declare them ‘unrelated’. The development of more general models of similarity is needed (Table 1). An alternative approach is to explore the dimensionality of G (Box 4), which might be able to detect conservation of underlying structure that the CPCA model might miss.

How fast do G matrices change?
This now appears to be the crucial question given the observation that G can evolve. The diversity of methods used in published studies makes comparison of their results difficult. Comparisons among closely related populations most frequently show no significant differences [7]. Considering P as well as G in the increasing number of CPCA studies (only three sets of studies [8,25,28,29] have applied CPCA to G), findings have ranged from proportionality [19,29] to no shared structure [18,30], to intermediate conditions with several CPCs [11,16,18,31]. In some cases, the degree of shared CPC structure depends strongly on the method of data standardization [32]. Above the species level, most studies [7] find significant differences in G. In addition, no published study (of G or P) among subspecies or at more inclusive taxonomic levels has accepted shared structure at more than the first two to three eigenvectors [16,17,19,28]. Thus, G or P are usually not significantly different among phenotypically similar populations, but statistically significant differences are the norm among phenotypically divergent populations (e.g. subspecies and species [33]). Comprehensive multitaxon studies are needed to confirm these tentative conclusions about the rates of G matrix evolution.

The rate of G matrix evolution is best determined through the comparative approach. When characters evolve slowly relative to cladogenesis, the COMPARATIVE METHOD is needed to account for phylogeny (to avoid correlations among observations [34]) – in this case matrices – and to estimate more accurately the direction and rate of evolution. Just as several methods can be used to compare two matrices, there are several approaches to structuring comparative analyses: single-pair comparisons, hierarchical multigroup comparison and ancestral reconstruction.

All but one [14] study of G have compared just two taxa. Such comparisons contain no phylogenetic information and therefore cannot determine direction of change.

A second approach is the hierarchical application of matrix-comparison methods, particularly those that can analyse multiple matrices, such as CPCA. In CPCA, for example, all members of a clade are analysed together for shared structure, and the analysis is repeated for all clades [16]. Another application is to conduct all pairwise comparisons among members of a clade, taxon, or taxonomic
category, [16] and partition the comparisons among categories or ranks. However, this incorporates minimal phylogenetic information and the degrees of freedom must be reduced to reflect the multiple comparisons. Interpretation of hierarchical analyses becomes more difficult as the number of lineages increases (Box 3).

The third and potentially most powerful method is ancestral reconstruction, which allows change to be partitioned among branches of a phylogeny [33] (Box 3). Once ancestral matrices are estimated, any of the matrix-comparison methods (Table 1) can be employed. Unfortunately, significant error can also arise in estimating ancestors [35], and that problem is likely to be exacerbated with correlated multivariate data. Uncertainty can be accommodated in a likelihood or bayesian framework [36]. In addition, although comparative studies of individual characters can sometimes verify ancestral conditions from fossils, such verification will be difficult for variances, which are properties of populations. G will almost always be impossible to measure for ancestors (but see [37]), but P can sometimes be estimated. The most commonly used application of this general approach, independent contrasts [34], is unlikely to be appropriate to the questions asked by quantitative geneticists, because it tests evolutionary correlations over time rather than decomposing the nature of changes. Maximum likelihood methods, already applied to univariate data [35], can be modified for correlated multivariate data, although the errors involved in those estimates can be very high for biologically interesting features that vary significantly.

Why do G matrices change?
Given the many evolutionary forces that are expected to buffet G, the observed differences in G matrix structure are not surprising. Mutation, selection, genetic drift and migration are all expected to affect G [20,38]. A more productive focus might therefore be on cases in which G might be expected to retain shared structure over time. Genetic drift provides an obvious starting point, because drift in a population of reduced effective size is expected to cause a proportional shrinking of all elements in G. This expectation has been proposed by Roff as a way to distinguish the effects of drift and selection [15]: proportional changes in G matrix structure are ascribed to drift and nonproportional changes to selection. Although appealing, this dichotomy has a flaw. Although proportionality is the theoretical expectation for drift, a great deal of variation around this expectation is likely. A large study on the effects of drift on wing morphology in Drosophila has demonstrated the extent of this variation [11]. Given enough time, any pattern of divergence among G matrices would probably be compatible with the hypothesis of drift. More generally, matrices may diverge by drift even when the effective sizes of the populations are equal. We note that CPCA and

### Box 3. Evolution of matrices

Figure 1 illustrates the hierarchy of models in common principal components analysis (CPCA). Compared to the root ancestor A, ancestor C and descendents 1, 2 and 3 are unequal but proportional. The eigenvectors (orientation of axes) are the same, whereas the eigenvalues (variances along each axis) all differ by a scalar amount. Descendant 5 shares the eigenvectors with ancestor A, but the eigenvalues for the two axes do not differ by the same amount. Thus, they share a common principal component (CPC) structure (proportionality is a special case of CPC). Descendant 4 differs by both orientation and relative variances and therefore for these two dimensions, shares no common structure with A (unrelated). If, however, they did share other axes for dimensions not plotted here, then descendant 4 and ancestor A would share partial CPC (PCPC). Thus, several levels in the hierarchy are portrayed by taxa in reference to the root ancestor A: equality (B), proportionality (C, 1–3), CPC (5) and unrelated (4). Further discussion of the CPC hierarchy can be found in [a,b].

**References**

Comparative quantitative genetics should be strongly influenced by, and potentially influence, the emerging synthesis between functional developmental genomics and studies of quantitative variation (e.g. [41]). First, finer-scale genetic information, especially regarding relationships among traits, is needed for modeling long-term evolution of $G$. Second, developmental models can be turned into statistical models of covariance structure [42], which will be needed for more meaningful comparisons among matrices. Observations about the evolution of $G$ might provide insights into the forces affecting $G$ and more importantly into the underlying processes that generate the covariance structure in the first place [43]. A comparative approach to $G$ matrix evolution should provide insights into macroevolutionary changes in developmental structure.

**Are $G$ and $P$ matrices similar?**

The elements of $P$ can be estimated much more accurately than can those of $G$ (Box 2), because sampling errors scale with the inverse of the sample size. This has led some researchers to suggest that the $P$ matrix will provide a more precise estimate of the form of $G$ [8,44,45] should they be proportional. The $P$ matrix is the sum of the $G$ matrix and all other sources of covariation, including genetic covariance not contained in $G$, and environmental covariation ($E$). Both genetic and nongenetic causes of covariance can be structured by the functional architecture that underlies the traits. If each hormone or regulatory gene that helps to build a trait provides an opportunity for both genetic and nongenetic effects to occur, then genetic and nongenetic variances will be correlated. This hypothesis can be tested by a direct comparison of $G$ and $E$ [8]. Consistent with this notion, genetic and phenotypic variances are very highly correlated [46,47]. However, there are many reasons why $G$ and $P$ might depart from proportionality [48]. Comparisons of $P$ have typically found more divergence more frequently than have comparisons of $G$, particularly within species.

Roff [44] tested the correlation of $P$ and $G$ by a survey of the literature. He found that phenotypic and genetic correlations were as correlated with each other as could be expected if they only differed because of sampling errors. The correspondence was particularly good for morphological traits, which tend to have high heritabilities (i.e. $G$ is a large proportion of $P$). Evolutionary forces, such as genetic drift, might be expected to have different effects on $G$ and $E$, and thereby lead to divergence in $P$ even if these underlying matrices share similarities [11]. Particular phenotypic and genetic correlations certainly differ significantly in some cases, but, overall, there is surprisingly little empirical evidence to reject the hypotheses that $P$ and $G$ are proportional. Further testing of this conjecture is still needed. Nearly all of the multitaxon comparisons to date have involved $P$ matrices [16,17,19,30,33,49].
Glossary

Adaptive landscape: a representation of the forces of natural selection where phenotypic trait values are the X and Y coordinates and mean fitness is the elevation.

Comparative quantitative genetics: the comparative study of quantitative genetic parameters, especially covariance matrices, across populations or species.

Comparative method: the application of phylogenetic information to cross-taxon comparisons.

CPA: common principal components analysis; a generalization of principal components analysis extended to multiple matrices.

Eigenvector: latent or characteristic roots of the variance–covariance matrix; they define the orientation in multidimensional space of the orthogonal axes of maximum variation.

Genetic variance–covariance matrix: a symmetrical matrix that summarizes the additive genetic contribution to the variances of and covariances between phenotypic traits. G matrix or G covariance matrix are shorthand references.

Phenotypic matrix: phenotypic variance–covariance matrix, measured directly for a population without partitioning out genetic and environmental contributions.

Conclusions and future directions

Clearly, G can evolve. The important questions now are what parts of G evolve, what is the rate at which G evolves, and how does that rate compare to the rate of speciation, population differentiation and changes in the adaptive landscape? Empirical and theoretical studies are needed, as are new or improved analytical methods. Empirical studies are needed to test assumptions about the relationship between G and P, for example. Perhaps the greatest need is for studies that robustly estimate G for multiple taxa. The most efficient approach might be to study taxa related to those that have already been studied and to build on earlier studies rather than duplicating them. With thoughtful species selection, this approach can also be used to expand morphological diversity. The clade containing the well-studied mouse Mus and rat Rattus, for example, includes many ecologically and morphologically divergent species, including grazing and earthworm specialists. Greater diversity can also be achieved by including groups outside the model organisms that have been the primary focus of past studies. Developmental [50] and integrative [51] approaches have great potential to provide explicit hypotheses, which would provide stronger theoretical and mechanistic frameworks for the study of changes in G.

The greatest need on the analytical side is for improved methods of matrix comparison that are statistically powerful, biologically meaningful, robust and that allow decomposition of the data. Although CPA has been widely adopted, we see it as an interim method that will remain useful only until more appropriate methods are developed. Modifications of factor analytic methods, such as confirmatory factor analysis [42] or common space analysis [24], which relax the assumption of orthogonality in PCA methods, are potential next steps. Provided that G does not evolve quickly with respect to the species and clades of interest to evolutionary biologists, improved methods of ancestral reconstruction for multivariate data should be a focus of comparative studies. The field is also hampered even in formulating scientific questions by the difficulties of visualizing such complex data. New visualization techniques are being developed (e.g. [52]) and could be adapted to, or new ones developed for, evolutionary studies.

Finally, we have only begun to ask some of the most interesting questions. For example, is G evolution decoupled from phenotypic evolution? That appears to be the case with P matrices in Phyllotis [16] and G within Clarkia dudleyana [53]. The G matrix, treated as a character in its own right, can be used to explore the evolution of developmental systems and their role in phenotypic evolution. A comparative quantitative genetic approach should provide a natural linkage between studies concentrating primarily on genetic details and those focusing on long-term phenotypic outcomes.

References


http://tree.trends.com