Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2016, **7**, 565–572 doi: 10.1111/2041-210X.12511

Evaluating modularity in morphometric data: challenges with the RV coefficient and a new test measure

Dean C. Adams[1,2]*

[1]*Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA; and* [2]*Department of Statistics, Iowa State University, Ames, IA 50011, USA*

Summary

1. Modularity describes the case where patterns of trait covariation are unevenly dispersed across traits. Specifically, trait correlations are high and concentrated within subsets of variables (modules), but the correlations between traits across modules are relatively weaker. For morphometric data sets, hypotheses of modularity are commonly evaluated using the RV coefficient, an association statistic used in a wide variety of fields.

2. In this article, I explore the properties of the RV coefficient using simulated data sets. Using data drawn from a normal distribution where the data were neither modular nor integrated in structure, I show that the RV coefficient is adversely affected by attributes of the data (sample size and the number of variables) that do not characterize the covariance structure between sets of variables. Thus, with the RV coefficient, patterns of modularity or integration in data are confounded with trends generated by sample size and the number of variables, which limits biological interpretations and renders comparisons of RV coefficients across data sets uninformative.

3. As an alternative, I propose the covariance ratio (CR) for quantifying modular structure and show that it is unaffected by sample size or the number of variables. Further, statistical tests based on the CR exhibit appropriate type I error rates and display higher statistical power relative to the RV coefficient when evaluating modular data.

4. Overall, these findings demonstrate that the RV coefficient does not display statistical characteristics suitable for reliable assessment of hypotheses of modular or integrated structure and therefore should not be used to evaluate these patterns in morphological data sets. By contrast, the covariance ratio meets these criteria and provides a useful alternative method for assessing the degree of modular structure in morphological data.

Key-words: geometric morphometrics, modularity, morphological integration, trait covariation

Introduction

A perennial topic in evolutionary biology is determining the degree to which traits covary and deciphering what developmental, genetic, and functional mechanisms are responsible for the correlations among traits. It has long been recognized that levels of covariation differ between the parts of organisms (Olson & Miller 1958), with some traits exhibiting high correlations while other traits are more independent of one another. *Morphological integration* describes the correlation among traits and occurs when changes in one trait are accompanied by changes in other traits that are affected by common mechanisms such as functional activities, genetic linkages, pleiotropy or common developmental pathways (Wagner 1984a; Cheverud 1996; Bookstein *et al.* 2003; Mitteroecker & Bookstein 2007; Klingenberg 2008; Goswami & Polly 2010). However, integration is not always observed uniformly and, in many instances, may be concentrated within subsets of traits that are less correlated with other subsets of traits. In these cases, *modularity* is displayed, where integration is found within subsets of traits (modules), but where the covariation between those modules is relatively weaker (Cheverud 1982; Wagner 1996; Wagner & Altenberg 1996; Wagner, Pavlicev & Cheverud 2007; Klingenberg 2014).

Over the past several decades, the study of morphological integration and modularity has enjoyed a renaissance, in part because of the recognition that patterns of modularity and integration can have profound effects on the direction of phenotypic evolution and on how novel morphologies originate (Atchley & Hall 1991; Cheverud 1996; Wagner & Altenberg 1996). Not surprisingly, recent emphasis has been placed on characterizing patterns of modularity and integration in multiple phenotypic traits and across a wide variety of taxa (e.g. Mitteroecker *et al.* 2004, 2012; Young & Badyaev 2006; Zelditch *et al.* 2008; Nogueira, Peracchi & Monteiro 2009; Monteiro & Nogueira 2010; Meloro *et al.* 2011; Gómez-Robles & Polly 2012; Parsons, Márquez & Albertson 2012; Clune, Mouret & Lipson 2013). In addition, researchers have compared patterns across different empirical systems to determine the extent to which levels of modularity or integration are consistent, and to decipher how these trends may differ across

*Correspondence author. E-mail: dcadams@iastate.edu

© 2015 The Author. Methods in Ecology and Evolution © 2015 British Ecological Society

levels of biological organization (e.g. Bookstein *et al.* 2003; Young & Badyaev 2006; Jamniczky & Hallgrimsson 2009; Drake & Klingenberg 2010; Ivanovic & Kalezic 2010; Kolbe *et al.* 2011; Renaud, Alibert & Auffray 2012; Sanger *et al.* 2012; Goswami *et al.* 2014). Much of this empirical work has been facilitated by the development of analytical tools for evaluating patterns of modularity and integration in high-dimensional data sets (e.g. Magwene 2001; Bookstein *et al.* 2003; Monteiro, Bonato & Reis 2005; Mitteroecker & Bookstein 2007; Márquez 2008; Klingenberg 2009; Pavlicev, Cheverud & Wagner 2009; Klingenberg & Marugán-Lobón 2013; Adams & Felice 2014). Generally, these approaches fall into two categories. First, exploratory approaches attempt to identify patterns consistent with integrated or modular structure without reference to pre-defined modules (e.g. Cheverud 1982; Wagner 1984b; Magwene 2001; see also Bookstein 2015). By contrast, other methods evaluate covariance patterns across subsets of traits defined *a priori* to determine whether the observed covariance patterns correspond to what is expected under a hypothesis of modular structure, or of integration among modules (Bookstein *et al.* 2003; Klingenberg 2008; for discussion see: Mitteroecker & Bookstein 2007; Goswami & Polly 2010).

When testing hypotheses of modularity in morphological data, Escoufier's (1973) RV coefficient is frequently utilized (Klingenberg 2009). The RV coefficient is a ratio that describes the degree of covariation between sets of variables relative to the variation and covariation within sets of variables. The RV coefficient is a multivariate generalization of the squared correlation coefficient, and attains values that range between zero and one. Low RV values are found from data that express relatively less covariation between sets of variables, which may be expected under a hypothesis of modular structure. By contrast, larger RV values describe relatively greater covariation between sets of variables, which may imply that there is higher integration among them. Permutation tests may then be used to evaluate these patterns (Klingenberg 2009). For morphometric data, use of the RV coefficient is appealing, as the same analytical tool may be utilized to evaluate either hypotheses of modular structure or hypotheses of integration among modules (for examples see: Jojic, Blagojevic & Vujosevic 2012; Santanta & Lofgren 2013; Goswami *et al.* 2014; Sorenson *et al.* 2014; Tsuboi, Gonzalez-Voyer & Kolm 2014; Urbanova *et al.* 2014). Unfortunately, while RV-based procedures are intuitive and conceptually straightforward, they suffer several mathematical deficiencies that limit their ability to accurately assess the degree of modularity or integrated structure in phenotypic data sets. As shown below, values of the RV coefficient are sensitive to both the sample size of the data set and the number of variables examined. Thus, when using the RV coefficient, any patterns of covariation present in data are confounded with trends generated by sample size and the number of variables. This limits biological interpretations based on this statistic and renders comparisons of RV coefficients across data sets uninformative. As an alternative, I propose a new measure (the covariance ratio) for quantifying modular structure and show that it is insensitive to such effects. Further, tests based on the covariance ratio display appropriate type I error rates, and exhibit higher statistical power relative to the RV coefficient when used to identify modular structure. Computer code for implementing the new procedure is found in the R package *geomorph* (Adams & Otárola-Castillo 2013; Adams, Collyer & Sherratt 2015).

Methods and results

PROBLEMS WITH THE RV COEFFICIENT

To understand the properties of the RV coefficient, I used a series of computer simulations. Each simulated data set was obtained by generating random variables drawn from a normal distribution $\sim N(0,1)$. Variables were then randomly assigned to modules. Thus, these data represented what was expected under the null hypothesis of a random association of variables, where neither modular structure nor integration among modules was present. For the first series of simulations, I generated data sets containing 32 random variables divided equally between two modules. One hundred data sets were simulated at each level of sample size, which ranged from 5 to 500 specimens, and from each, the RV coefficient was obtained. Then, at each level of sample size, the expected value of the RV coefficient was calculated as the mean of the RV coefficients obtained from the 100 simulations. Next, I performed a second series of simulations, where data sets of 100 specimens each were generated that contained differing numbers of variables (20–500). As before, all variables were drawn from a normal distribution $\sim N(0,1)$ and were divided equally between two modules. One hundred data sets were simulated at each level of variable number. All simulations were performed in R 3.2.0 (R Core Team 2015).

As is evident from Fig 1(a), the RV coefficient depends highly on sample size, with its values spanning nearly the entire range between zero and one. Note that this wide range of RV values is obtained for data with no input covariation (i.e. a random associations of variables), and confirms earlier results of Smilde *et al.* (2009; see also Fruciano, Franchini & Meyer 2013). In these simulations, data sets with low sample sizes yielded large RV coefficients, while data sets with large sample sizes yielded low RV coefficients. The reason for this pattern is that the denominator of the RV coefficient contains both the covariation within modules, and additionally, the variation in all traits, the latter of which scales with sample size (see explanation in Smilde *et al.* 2009). Next, I examined the effect of the number of variables, and found that values of the RV coefficient depend highly on this as well (Fig 1b). Specifically, when the number of variables examined increased, the RV coefficient estimated from the data also increased. Thus, for a given sample size, higher RV coefficients will be found simply when more variables are examined.

Together, these results clearly demonstrate that the RV coefficient is adversely affected by attributes of the data (sample size and the number of variables) that do not characterize the covariance structure between sets of variables. As a consequence, it is difficult to determine whether small RV values are

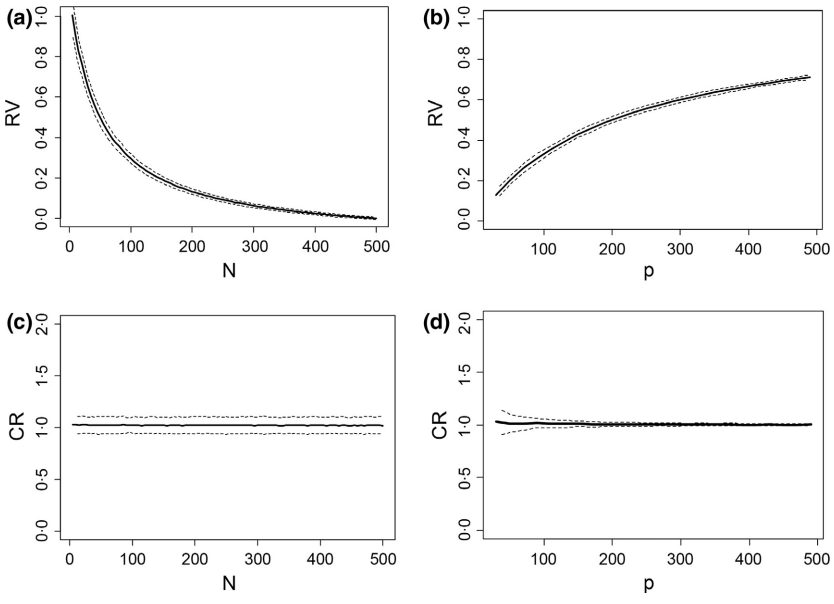


Fig. 1. Evaluation of the RV and CR coefficients under the hypothesis of random associations of variables (i.e. neither modular nor integrated structure). Mean and 95% confidence intervals of RV values obtained from (a) 100 data sets simulated across a range of sample sizes and from (b) 100 data sets simulated across a range of variable number. Mean and 95% confidence intervals of CR values obtained from (c) 100 data sets simulated across a range of sample sizes and from (d) 100 data sets simulated across a range of variable number.

the result of greater independence among modules, a large sample size or a small number of variables and whether high RV values describe data with higher covariation among modules, a small sample size or a large number of variables. This renders biological interpretations of the RV coefficient with regard to hypotheses of modularity and integration challenging. Further, while it is possible to use permutation procedures to statistically evaluate the RV coefficient obtained from a particular data set (cf. Klingenberg 2009), interpreting the results from such procedures is limited in scope. That is, one can determine whether the observed RV coefficient differs from values obtained via permutation, but one cannot discern whether this value differs from what is expected under the null hypothesis where neither integration nor modularity is present, because the baseline value of the RV coefficient for random data changes with both sample size and the number of variables. This necessarily leads to weaker biological inferences concerning modularity when using the RV coefficient.

Finally, these observations reveal that comparisons of RV coefficients across data sets are not informative. Increasingly, researchers are characterizing patterns of modularity and integration in multiple data sets, and comparing the resulting RV coefficients to determine the relative degree of integration or modularity in different species, or in different phenotypic traits (e.g. Drake & Klingenberg 2010; Ivanovic & Kalezic 2010; Renaud, Alibert & Auffray 2012; Sanger *et al.* 2012; Goswami *et al.* 2014). However, because the RV coefficient is sensitive to variation in sample size and the number of variables, it is difficult to discern whether a lower RV coefficient in one data set corresponds to relatively lower levels of association between modules (i.e. greater modularity), or whether the difference is simply due to differences in samples sizes or the number of variables. Thus, this practice should be avoided, as utilizing the RV coefficient to compare trends in modularity or integration across data sets may lead to improper biological inferences.

THE COVARIANCE RATIO AND A TEST OF MODULARITY

The RV coefficient is sensitive to changes in sample size and to changes in the number of variables. Both of these undesirable properties are observed because the RV coefficient does not isolate the covariation among traits from the overall variation within each trait, since the denominator of the RV coefficient contains the covariation within modules as well as the variation in each variable. However, the degree of variation in each variable is not strictly necessary for characterizing either modularity or the integration among modules, as both of these biological concepts describe the relative pattern of *covariation* among traits within and among modules. Thus, the RV coefficient quantifies aspects of the data that are not required for describing modular structure, and it is these very components that obfuscate biological interpretations when using this measure.

As an alternative to the RV coefficient, I propose a new measure for characterizing the degree of modularity in morphological data sets: the covariance ratio (CR). Unlike the RV coefficient, the covariance ratio uses only the pairwise covariances between variables to quantify modular structure. To estimate the covariance ratio, variables in each of two modules ($\mathbf{Y_1}$ and $\mathbf{Y_2}$) are obtained for N specimens. The data matrices may then be concatenated into a single matrix ($\mathbf{Y}$), from which an overall covariance matrix is constructed. This covariance matrix can be expressed as the partitioned matrix:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \qquad \text{eqn 1}$$

where $\mathbf{S_{11}}$ and $\mathbf{S_{22}}$ are the covariance matrices within modules $\mathbf{Y_1}$ and $\mathbf{Y_2}$ respectively, and $\mathbf{S_{12}}$ and $\mathbf{S_{21}}$ describe the covariation between modules.The covariance ratio is then found as follows:

$$CR = \sqrt{\frac{trace(\mathbf{S}_{12}\mathbf{S}_{21})}{\sqrt{trace(\mathbf{S}^*_{11}\mathbf{S}^*_{11})\,trace(\mathbf{S}^*_{22}\mathbf{S}^*_{22})}}} \qquad \text{eqn 2}$$

where $\mathbf{S}^*_{11}$ and $\mathbf{S}^*_{22}$ are the covariance matrices within modules with zeroes replacing the diagonal elements.

Examining equation 2, it can be seen that the CR coefficient is a ratio of the overall covariation between modules relative to the overall covariation within modules. Specifically, the numerator is found from the between-module covariance matrix ($\mathbf{S_{12}}$), and describes the total sum of squared covariance between $\mathbf{Y_1}$ and $\mathbf{Y_2}$. By contrast, the denominator is found from the within-module covariance matrices ($\mathbf{S_{11}}$ and $\mathbf{S_{22}}$), with the diagonal elements set to zero. As such, the denominator describes the total sum of squared covariation within each module minus the variation in each trait dimension. Therefore, when the square root is taken, the CR coefficient is a ratio of the covariation between modules relative to the covariation within them (for a full conceptual and mathematical derivation of CR, see the Supplemental Information).

The astute reader will also recognize the mathematical connection between the numerator of the CR coefficient and several other widely used association measures. Specifically, both partial least squares, which is commonly used to evaluate integration (*sensu* Bookstein *et al.* 2003) and the inertia statistic (Abdi & Williams 2013), utilize the between-module covariance matrix ($\mathbf{S_{12}}$) or its standardized version ($\mathbf{R_{12}}$) to characterize the degree of covariation between sets of variables. However, because the CR describes covariation between modules *relative* to covariation within modules, it is appropriate for evaluating tests of modular structure, rather than simply describing covariation between modules. Thus, the CR coefficient may be thought of as a complementary analytical tool, useful for different purposes (see Supporting information for further discussion).

Empirically, the CR coefficient ranges from zero to positive values, with low values expressing relatively less covariation between modules, and higher values describing data sets with relatively higher covariation between modules (the CR coefficient is undefined for the identity matrix, but for both real and simulated data, the CR coefficient attains positive values). Further, for random sets of variables, the CR coefficient has an expected value of one, as levels of covariation between modules should be, on average, the same as the covariation within modules. In addition, CR values between zero and one describe data sets where the degree of covariation between modules is less than that found within modules, which characterizes relatively more modular structure. By contrast, CR values larger than one describe greater covariation between modules relative to within modules. Thus, values of the CR coefficient are easily interpretable. Finally, extending the CR coefficient to the case of more than two modules may be accomplished by obtaining the mean from the CR coefficients obtained from all pairs of modules (*sensu* Klingenberg 2009).

It is important to recognize that because the variance of each trait is not included, the CR coefficient is insensitive to changes in sample size or the number of variables. This is easily verified through simulation. Using the simulated data sets described above, I reran analyses using the CR coefficient and found that the expected CR coefficient remains at 1·0 as sample size increases (Fig. 1c). Further, as the number of variables increases, the CR coefficient quickly asymptotes to one (Fig. 1d). Therefore, unlike the RV coefficient, the covariance ratio displays a constant expected value (CR = 1·0) under the null hypothesis of random associations of variables, regardless of sample size or the number of variables.

For empirical data, statistically evaluating the CR coefficient is accomplished via permutation, where variables are randomly assigned to modules, and each time the CR coefficient is recalculated (for the case of geometric morphometric data, landmarks are permuted relative to module designation). The proportion of permuted values lower than the original is then treated as an estimate of the significance of the test. This columnwise permutation procedure correctly generates a distribution of values expected under random associations of variables because with each iteration, the covariance patterns within and between modules are randomized, thereby dissociating the relationship between them. As such, levels of covariation within modules are not expected to differ from levels of covariation between modules. Thus, on average, the CR coefficients obtained through permutation should be centred on a value of one, which, as demonstrated previously, represents the expected value under the null hypothesis of a random association of variables (neither modular structure nor integration among modules).

Figure 2(a) demonstrates this property empirically. Here, I simulated a data set with a known level of modularity and examined the distribution of CR values obtained via the proposed permutation procedure. First, I generated initial covariance matrix where the covariation between modules was less than that within modules (s_b = 0·45 and s_w = 0·6). The values populating the covariance matrix were drawn from a normal distribution as (μ = s, σ = 0·3), and the variances along the diagonal were drawn from a normal distribution ~N(0,1). From this initial covariance matrix, the 500 specimens were then simulated, with 100 variables split evenly between two modules. The permutation method above was then implemented. As can be seen in Fig. 2(a), the distribution of CR values obtained via permutation was centred near one, confirming that the permutation procedure generates a distribution of possible outcomes that correspond to what is expected under the null hypothesis of random associations of variables. Because of this, tests based on the CR coefficient may be used to evaluate whether the pattern obtained under a particular hypothesis of modularity differs from what is expected under the null hypothesis of random associations of variables (i.e. no modular structure in the data set). This extends the level of biological inference beyond that which is possible when using the RV coefficient for similar tasks.

Finally, it should be noted that for landmark data, estimates of the CR coefficient change slightly with different orientations of the data set. However, this does not present a difficulty, because across rotation angles the CR coefficient changes in a predictable manner; generating a sine wave with every 90° of

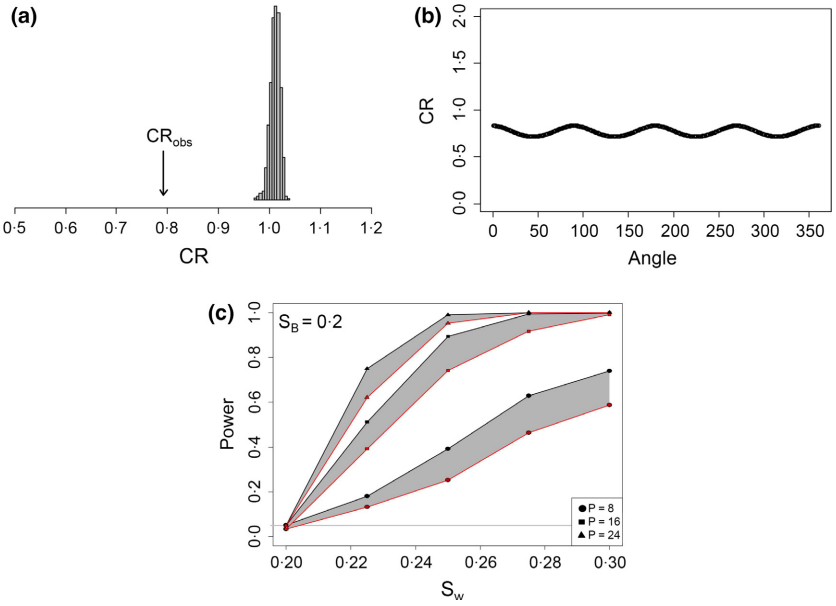


Fig. 2. Statistical performance of tests using the CR coefficient. (a) Demonstration that the distribution of CR values obtained from a permutation test corresponds to that expected under the null hypothesis of random associations of variables. (b) Demonstration that for geometric morphometric data, the CR coefficient changes predictably with orientation angle, enabling this to be accounted for during the analysis (data from Fig. 3a). (c) Simulation results evaluating the type I error and statistical power of hypothesis testing procedures for evaluating patterns of modularity using the CR (black) and RV (red) coefficients for data simulated with differing numbers of variables (S_b = covariation between modules, S_w = covariation within modules). The light grey line denotes 0·05, and the dark grey region denotes the difference in power between methods for each simulation.

rotation (see Fig. 2b). Thus, one may use the average CR coefficient across 90° of rotation angles as the observed test value, rotate the data to the angle that corresponds to the average CR, and evaluate the observed CR value while accounting for orientation of the data set using permutation procedures.

STATISTICAL PERFORMANCE OF TESTS OF THE CR AND RV COEFFICIENTS

To compare the performance of the two test statistics, I conducted a series of computer simulations. To evaluate type I error, I generated data sets containing random variables obtained from a normal distribution ~$N(0,1)$, which were then divided equally between two modules. The levels of variable number used were within the typical range found in empirical morphometric studies (P = 8, 12, 16, 20, 24, 32), and one thousand data sets of 100 specimens each were simulated for each level. To evaluate statistical power, I simulated data sets that contained a known difference between the covariation within modules (s_w) and the covariation between modules (s_b). For each level of variable number, an initial covariance matrix was constructed, where the covariation between pairs of variables among modules was drawn from a normal distribution with an average value of 0·2 (μ = 0·2, σ = 0·03), and where the covariance between pairs of variables within modules was drawn from a normal distribution ($\mu = s_w$, σ = 0·03) with an average value that exceeded the among module covariance by a set amount, depending upon simulation conditions (s_w = 0·225, 0·25, 0·275, 0·30). The elements were then adjusted as needed so that the resulting matrix conformed to the properties of a valid covariance matrix (i.e. symmetric, positive definite). Using this procedure, a set of initial covariance matrices was constructed (**S**) that displayed increasing amounts of modularity. From each, 1000 data sets of 100 specimens were then simulated. For each simulated data set, the observed CR and RV coefficients were obtained, which were evaluated using the permutation procedure above. The proportion of significant results (of 1000) was then treated as an estimate of the Type I error (random data) or power of the test (when $s_b < s_w$).

Results

For the simulation conditions examined here, both the CR and RV coefficients displayed appropriate type I error rates, near the nominal value of α = 0·05 (Fig. 2c). Additionally, in all cases, the statistical power of tests based on the CR coefficient exceeded that of the RV coefficient (a subset of results is found in: Fig. 2c). The conclusion from these simulations is that tests based on the CR coefficient display acceptable statistical properties for evaluating hypotheses of modularity and also exhibit higher statistical power for identifying such trends in data, as compared to the RV coefficient.

BIOLOGICAL EXAMPLES

Here, I provide two empirical examples that illustrate the utility of the CR coefficient for evaluating modular structure in morphological data sets. The first example is from mouse mandibles. Prior research has revealed several regions of this structure with distinct embryonic origins (e.g. Atchley & Hall 1991; Cheverud *et al.* 1991; Leamy 1993; Klingenberg, Mebus & Auffray 2003), which have been hypothesized to function as two modules: (anterior) the aveolar region and (posterior) the ascending ramus (Fig. 3a). For the current example, I examined the left and right sides of 226 adult individuals of the yellow-necked mouse, *Apodemus flavicollis* (data from Jojic, Blagojevic & Vujosevic 2012). The second example examines head shape in *Plethodon* salamanders, where landmarks from the left-lateral side of the head were digitized from 289 specimens of *P. jordani* from the southern Appalachian mountains (data from Adams 2004, 2010). These landmarks represented two distinct anatomical structures; the cranium and the mandible (Fig. 3b).

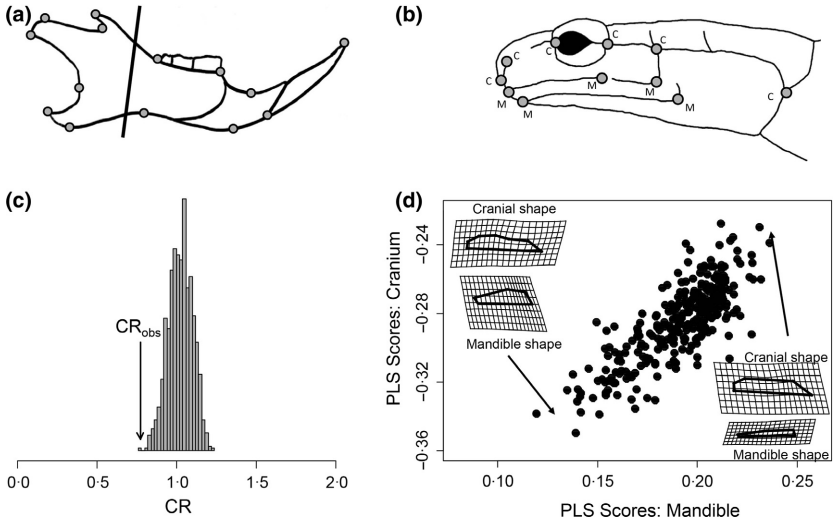


Fig. 3. Graphical summary of results from the empirical examples. (a) Locations of 14 landmarks on a mouse mandible. The solid line demarcates the separation of the two hypothesized modules. (b) Locations of 11 anatomical landmarks used to characterize head shape in *Plethodon* salamanders (image from Adams, West & Collyer 2007). (c) Histogram of CR coefficients obtained from a permutation test of alternative partitions of the mouse mandible data example, with the observed CR coefficient designated. (d) Plot of PLS scores for the first axis of mandible shape versus cranial shape found from a partial least squares analysis. Thin-plate spline deformation grids along each PLS axis for each structure are displayed.

For both examples, landmark-based geometric morphometric methods were used for the analyses (Bookstein 1991; Mitteroecker & Gunz 2009; Adams, Rohlf & Slice 2013). First, a generalized Procrustes analysis was performed to align the specimens and remove the effects of non-shape variation. For the mouse mandible data set, replicate configurations of each side of each mandible were then averaged to obtain a single estimate of the shape of each side for subsequent analyses. Covariance matrices were then obtained for each data set, and the landmarks were assigned to modules based on the hypotheses described in Fig. 3(a,b). Next, the hypothesis of modularity was evaluated using the CR coefficient, using 999 iterations of the permutation procedure to evaluate the statistical significance. All analyses were performed in R 3.2.0 (R Core Team 2015) using the package *geomorph* (Adams & Otárola-Castillo 2013; Adams, Collyer & Sherratt 2015).

For the mouse mandible data set, the observed CR coefficient was significantly lower than one (CR = 0·772 ± 0·0009; $P = 0{\cdot}001$; Fig 3c), suggesting that there was a strong degree of independence between the two modules. Thus, there was support for the hypothesis that the mouse mandible displays significant modularity when compared to the null hypothesis of no modular structure. By contrast, there was little evidence that the cranium and mandible vary independently in *Plethodon*, as CR value was greater than one and was not significant. (CR = 1·056 ± 0·00063, $P = 0{\cdot}473$). As such, the patterns in this data set implied that the two structures may be integrated. Indeed, a PLS analysis revealed strong integration between the cranium and mandible ($PLS_{corr} = 0{\cdot}802$, $P = 0{\cdot}001$: Fig. 3d), implying that anatomical changes in the mandible are accompanied by corresponding changes in the cranium.

Discussion

A ubiquitous characteristic of organisms is that their phenotypic traits covary, which may manifest as patterns of modularity or integration. Paramount to the study of integration and modularity is the availability of analytical tools that facilitate the quantification of patterns in morphological data. In this article, I evaluated the statistical properties of one such tool (the RV coefficient) and found that it was sensitive to variation in sample size and the number of variables examined, thereby confounding patterns of covariation in empirical data sets with their sample size and number of variables and complicating biological interpretations of any trends that are identified. As an alternative, I proposed the covariance ratio, a measure that is insensitive to sample size and the number of variables, and when used on random data, it has an expected value of one. Further, tests based on the covariance ratio have appropriate type I error rates and display higher statistical power as compared to the RV coefficient. As such the covariance ratio provides a more useful measure for characterizing and evaluating the degree of modularity in biological data sets than is the RV coefficient, thereby filling a critical analytical gap in our morphometric toolkit.

At the other extreme, evaluating integration among modules may already be accomplished using partial least squares (*sensu* Mitteroecker & Bookstein 2007), which summarizes the covariation between modules using $\mathbf{S_{12}}$, and is thus a direct estimate of such patterns. While the RV coefficient has also been used in this context, it provides no advantage over PLS. The reason for this is that the statistical assessment of integration for both measures is performed with permutation, where the rows (individuals) are shuffled in one module while leaving the rows in the other module constant. This correctly disassociates the covariation between modules, but leaves the variation within each module unaffected. Thus, the denominator of RV will remain constant across all permutations, adding no information to the test. Additionally, since the numerator of RV is based entirely on $\mathbf{S_{12}}$, nothing is gained by using the RV coefficient in place of PLS for tests of integration, as PLS already evaluates covariation in $\mathbf{S_{12}}$. Thus, for future studies, I recommend that evaluating the degree of morphological integration and modularity in morphometric data sets should be accomplished using a pair of analytical tools: the covariance ratio for evaluating patterns of modular structure and partial least squares for evaluating the degree of integration between modules.

For the case of geometric morphometric data, one possible extension to the method described here would be to more explicitly account for the spatial proximity of the landmark in the observed covariance matrix. Prior work has shown that landmarks that are physically close to one another are more highly correlated than are those found more distantly on the structure (Goswami 2006), suggesting that even if the landmarks are otherwise uncorrelated from the perspective of underlying biological mechanisms, they will not be independent. At present, virtually no analytical methods account for the spatial proximity of the variables when initially characterizing patterns of integration or modular structure (Mitteroecker & Bookstein 2007; but see Bookstein 2015). Thus, an important future advance will be to extend modularity methods so that they can account for this information during the analysis.

Another possible extension of the method described here is to enable the simultaneous comparison of multiple modular hypotheses (e.g. Márquez 2008). Presently, the covariance ratio may be used to statistically evaluate patterns of covariation within and among modules as described by a pre-defined set of modules. However, while this approach facilitates inferences on whether the observed pattern differs from what is expected relative to a null hypothesis of random associations of variables, it does not allow one to determine whether some other modular hypothesis provides a better fit to the data (e.g. Is the greater support for a two-module or a three-module hypothesis?). Extensions of the approach described here to allow for explicit comparisons between alternative models (*sensu* Burnham & Anderson 2002) would broaden the morphometrician's toolkit to enable comparisons of a broader set of modular hypotheses.

Finally, the procedure of examining the performance of analytical methods with both empirical and simulated data sets (including data simulated randomly) provides valuable information regarding the ability of analytical tools to address biological hypotheses under a variety of conditions can be revealed. Such a multifaceted procedure has proven useful in objectively comparing alternative methods and for distinguishing when one method should be preferred over another (in the context of morphometrics, see: Rohlf 2000a, b, 2003; also Adams & Collyer 2015). Additionally, keeping this procedure in mind when developing methods ensures that the new analytical methods meet a minimum threshold of utility for a wide variety of data sets and underlying conditions (e.g. Adams 2013, 2014a,b,c). For the case of methods for studying modularity, the results presented here demonstrate that the RV coefficient does not meet these requirements, while the covariance ratio satisfies them. Thus, the covariance ratio should provide a powerful and useful alternative approach for quantifying modular patterns in data and evaluating them relative to the null hypothesis of no modular structure.

Acknowledgements

M. Collyer and A. Kaliontzopoulou provided extensive discussion and comment during the development of this work, and I thank them for their patience in engaging in those discussions. M. Collyer, A. Kaliontzopoulou, F.J. Rohlf, M. Zelditch and several anonymous reviewers provided critical and insightful comments on versions of this manuscript. I also thank C.P. Klingenberg for sharing his perspectives. This work was sponsored in part by the National Science Foundation DEB-1257287.

Data accessibility

The r script for implementing the covariance ratio procedure is found in the Supplemental Information. Data for the empirical examples found in dryad (http://dx.doi.org/10.5061/dryad.51696 for the mouse data; http://dx.doi.org/10.5061/dryad.2kt43 for the *Plethodon* data).

References

Abdi, H. & Williams, L.J. (2013) Partial least squares methods: partial least squares correlation and partial least square regression. *Methods in Molecular Biology: Computational Toxicology* (eds B. Reisfeld & A. Mayeno), pp. 549–579. Springer Verlag, New York.

Adams, D.C. (2004) Character displacement via aggressive interference in Appalachian salamanders. *Ecology*, **85**, 2664–2670.

Adams, D.C. (2010) Parallel evolution of character displacement driven by competitive selection in terrestrial salamanders. *BMC Evolutionary Biology*, **10**, 1–10.

Adams, D.C. (2013) Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. *Systematic Biology*, **62**, 181–192.

Adams, D.C. (2014a) A generalized Kappa statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Systematic Biology*, **63**, 685–697.

Adams, D.C. (2014b) A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution*, **68**, 2675–2688.

Adams, D.C. (2014c) Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. *Systematic Biology*, **63**, 166–177.

Adams, D.C. & Collyer, M.L. (2015) Permutation tests for phylogenetic comparative analyses of high-dimensional shape data: what you shuffle matters. *Evolution*, **69**, 823–829.

Adams, D.C., Collyer, M. & Sherratt, E. (2015) geomorph: Software for geometric morphometric analyses. R package version 2.1.7-1. http://CRAN.R-project.org/package = geomorph.

Adams, D.C. & Felice, R.N. (2014) Assessing trait covariation and morphological integration on phylogenies using evolutionary covariance matrices. *PLoS One*, **9**, e94335.

Adams, D.C. & Otárola-Castillo, E. (2013) geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, **4**, 393–399.

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

Adams, D.C., West, M.E. & Collyer, M.L. (2007) Location-specific sympatric morphological divergence as a possible response to species interactions in West Virginia *Plethodon* salamander communities. *Journal of Animal Ecology*, **76**, 289–295.

Atchley, W.R. & Hall, B.K. (1991) A model for development and evolution of complex morphological structures. *Biological Reviews*, **66**, 101–157.

Bookstein, F.L. (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge.

Bookstein, F.L. (2015) Integration, disintegration, and self-similarity: characterizing the scales of shape variation in landmark data. *Evolutionary Biology*, **42**, 395–426.

Bookstein, F.L., Gunz, P., Mitteroecker, P., Prossinger, H., Schaefer, K. & Seidler, H. (2003) Cranial integration in Homo: singular warps analysis of the midsagittal plane in ontogeny and evolution. *Journal of Human Evolution*, **44**, 167–187.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*, 2nd edn. Springer-Verlag, New York.

Cheverud, J.M. (1982) Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution*, **36**, 499–516.

Cheverud, J.M. (1996) Developmental integration and the evolution of pleiotropy. *American Zoologist*, **36**, 44–50.

Cheverud, J.M., Hartman, S.E., Richtsmeier, J.T. & Atchley, W.R. (1991) A quantitative genetic analysis of localized morphology in mandibles of inbred mice using finite element scaling. *Journal of Craniofacial Genetics and Developmental Biology*, **11**, 122–137.

Clune, J., Mouret, J.B. & Lipson, H. (2013) The evolutionary origins of modularity. *Proceedings of the Royal Society B Biological Sciences*, **280**, 20122863.
Drake, A.G. & Klingenberg, C.P. (2010) Large scale diversification of skull shape in domestic dogs: disparity and modularity. *American Naturalist*, **175**, 289–301.
Escoufier, Y. (1973) Le traitement des variables vectorielles. *Biometrics*, **29**, 751–760.
Fruciano, C., Franchini, P. & Meyer, A. (2013) Resampling-based approaches to study variation in morphological modularity. *PLoS ONE*, **8**, e69376.
Gómez-Robles, A. & Polly, P.D. (2012) Morphological integration in the hominin dentition: evolutionary, developmental, and functional factors. *Evolution*, **66**, 1024–1043.
Goswami, A. (2006) Cranial modularity shifts during mammalian evolution. *American Naturalist*, **168**, 270–280.
Goswami, A. & Polly, P.D. (2010) Methods for studying morphological integration, modularity and covariance evolution. *Quantitative Methods in Paleobiology* (eds J. Alroy & G. Hunt), pp. 213–243. The Paleontological Society, Ithaca, New York, USA.
Goswami, A., Smaers, J.B., Soligo, C. & Polly, P.D. (2014) The macroevolutionary consequences of phenotypic integration: from development to deep time. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, **369**, 20130254.
Ivanovic, A. & Kalezic, M.L. (2010) Testing the hypothesis of morphological integration on a skull of a vertebrate with a biphasic life cycle: a case study of the alpine newt. *Journal of Experimental Zoology*, **314**, 527–538.
Jamniczky, H.A. & Halligrimsson, B. (2009) A comparison of covariance structure in wild and laboratory muriod crania. *Evolution*, **63**, 1540–1556.
Jojic, V., Blagojevic, J. & Vujosevic, M. (2012) Two-module organization of the mandible in the yellow-necked mouse: a comparison between two different morphometric approaches. *Journal of Evolutionary Biology*, **25**, 2489–2500.
Klingenberg, C.P. (2008) Morphological integration and developmental modularity. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 115–132.
Klingenberg, C.P. (2009) Morphometric integration and modularity in configurations of landmarks: tools for evaluating *a priori* hypotheses. *Evolution and Development*, **11**, 405–421.
Klingenberg, C.P. (2014) Studying morphological integration and modularity at multiple levels: concepts and analysis. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, **369**, 20130249.
Klingenberg, C.P. & Marugán-Lobón, J. (2013) Evolutionary covaration in geometric morphometric data: analyzing integration, modularity, and allometry in a phylogenetic context. *Systematic Biology*, **62**, 591–610.
Klingenberg, C.P., Mebus, K. & Auffray, J.C. (2003) Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evolution and Development*, **5**, 522–531.
Kolbe, J.J., Revell, L.J., Szekely, B., Brodie, E.D.I.I.I. & Losos, J.B. (2011) Convergent evolution of phenotypic integration and its alignment with morphological diversification in Caribbean *Anolis* ecomorphs. *Evolution*, **65**, 3608–3624.
Leamy, L. (1993) Morphological integration of fluctuating asymmetry in the mouse mandible. *Genetica*, **89**, 139–153.
Magwene, P.M. (2001) New tools for studying integration and modularity. *Evolution*, **55**, 1734–1745.
Márquez, E.J. (2008) A statistical framework for testing modularity in multidimensional data. *Evolution*, **62**, 2688–2708.
Meloro, C., Raia, P., Carotenuto, F. & Cobb, S.N. (2011) Phylogenetic signal, function and integration in the subunits of the carnivoran mandible. *Evolutionary Biology*, **38**, 465–475.
Mitteroecker, P. & Bookstein, F.L. (2007) The conceptual and statistical relationship between modularity and morphological integration. *Systematic Biology*, **56**, 818–836.
Mitteroecker, P. & Gunz, P. (2009) Advances in geometric morphometrics. *Evolutionary Biology*, **36**, 235–247.
Mitteroecker, P., Gunz, P., Bernhard, M., Schaefer, K. & Bookstein, F.L. (2004) Comparison of cranial ontogenetic trajectories among great apes and humans. *Journal of Human Evolution*, **46**, 679–698.
Mitteroecker, P., Gunz, P., Meubauer, S. & Muller, G. (2012) How to explore morphological integration in human evoluion and development? *Evolutionary Biology*, **39**, 536–553.
Monteiro, L.R., Bonato, V. & Reis, S.F.D. (2005) Evolutionary integration and morphological diversification in complex morphological structures: mandible shape divergence in spiny rats (Rodentia, Echimyidae). *Evolution and Development*, **7**, 429–439.
Monteiro, L.R. & Nogueira, M.R. (2010) Adaptive radiations, ecological specialization, and the evolutionary integration of complex morphological structures. *Evolution*, **64**, 724–744.
Nogueira, M.R., Peracchi, A.L. & Monteiro, L.R. (2009) Morphological correlates of bite force and diet in the skull and mandible of phyllostomid bats. *Functional Ecology*, **23**, 715–723.
Olson, E.C. & Miller, R.L. (1958) *Morphological Integration*. University of Chicago Press, Chicago.
Parsons, K.J., Márquez, E. & Albertson, R.C. (2012) Constraint and opportunity: the genetic basis and evolution of modularity in the cichlid mandible. *American Naturalist*, **179**, 64–78.
Pavlicev, M., Cheverud, J.M. & Wagner, G.P. (2009) Measuring morphological integration using eigenvalue variance. *Evolutionary Biology*, **36**, 157–170.
R Core Team (2015) *R: A Language and Environment for Statistical Computing*. *Version 3.2.0*. R Foundation for Statistical Computing, Vienna, http://cran.R-project.org.
Renaud, S., Alibert, P. & Auffray, J.C. (2012) Modularity as a source of new morphological variation in the mandible of hybrid mice. *BMC Evolutionary Biology*, **12**, 1–16.
Rohlf, F.J. (2000a) On the use of shape spaces to compare morphometric methods. *Hystrix*, **11**, 9–25.
Rohlf, F.J. (2000b) Statistical power comparisons among alternative morphometric methods. *American Journal of Physical Anthropology*, **111**, 463–478.
Rohlf, F.J. (2003) Bias and error in estimates of mean shape in geometric morphometrics. *Journal of Human Evolution*, **44**, 665–683.
Sanger, T.J., Mahler, D.L., Abzhanov, A. & Losos, J.B. (2012) Roles for modularity and constraint in the evolution of cranial diversity among *Anolis* lizards. *Evolution*, **66**, 1525–1542.
Santanta, S.E. & Lofgren, S.E. (2013) Does nasal echolocation influence the modularity of the mammal skull? *Journal of Evolutionary Biology*, **26**, 2520–2526.
Smilde, A.K., Kiers, H.A.L., Biklsma, S., Rubingh, C.M. & vanErk, M.J. (2009) Matrix correlations for high-dimensional data: the modified RV-coefficient. *Bioinformatics*, **25**, 401–405.
Sorenson, D.W., Butkus, C., Cooper, L.N., Cretekos, C.J., Rasweiler, J.J. IV & Sears, K.E. (2014) Palate variation and evolution in new world leaf-nosed and old world fruit bats (order Chiroptera). *Evolutionary Biology*, **41**, 595–605.
Tsuboi, M., Gonzalez-Voyer, A. & Kolm, N. (2014) Phenotypic integration of brain size and head morphology in Lake Tanganyika cichlids. *BMC Evolutionary Biology*, **14**, 39.
Urbanova, P., Hejna, P., Zatopkova, L. & Safr, M. (2014) The asymmetry and modularity of the hyoid bone. *International Journal of Morphology*, **32**, 251–260.
Wagner, G.P. (1984a) Coevolution of functionally constrained characters: prerequisites of adaptive versatility. *BioSystems*, **17**, 51–55.
Wagner, G.P. (1984b) On the eigenvalue distribution of genetic and phenotypic dispersion matrices: evidence for a nonrandom organization of quantitative character variation. *Journal of Mathematical Biology*, **21**, 77–95.
Wagner, G.P. (1996) Homologues, natural kinds and the evolution of modularity. *American Zoologist*, **36**, 36–43.
Wagner, G.P. & Altenberg, L. (1996) Complex adaptations and the evolution of evolvability. *Evolution*, **50**, 967–976.
Wagner, G.P., Pavlicev, M. & Cheverud, J.M. (2007) The road to modularity. *Nature Reviews Genetics*, **8**, 921–931.
Young, R.L. & Badyaev, A.V. (2006) Evolutionary persistence of phenotypic integration: influence of developmental and functional relationships on complex trait evolution. *Evolution*, **60**, 1291–1299.
Zelditch, M.L., Wood, A.R., Bonett, R.M. & Swiderski, D.L. (2008) Modularity of the rodent mandible: integrating bones, muscles, and teeth. *Evolution and Development*, **10**, 756–768.

Received 9 October 2015; accepted 13 November 2015
Handling Editor: Pedro Peres-Neto

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Data S1. Supplemental Information.