Oecologia (2006) 148: 547–554
DOI 10.1007/s00442-006-0403-6

METHODS

**Michael W. McCoy · Benjamin M. Bolker
Craig W. Osenberg · Benjamin G. Miner
James R. Vonesh**

Size correction: comparing morphological traits among populations and environments

Received: 26 May 2005 / Accepted: 15 February 2006 / Published online: 8 April 2006

Abstract Morphological relationships change with overall body size and body size often varies among populations. Therefore, quantitative analyses of individual traits from organisms in different populations or environments (e.g., in studies of phenotypic plasticity) often adjust for differences in body size to isolate changes in allometry. Most studies of among population variation in morphology either (1) use analysis of covariance (ANCOVA) with a univariate measure of body size as the covariate, or (2) compare residuals from ordinary least squares regression of each trait against body size or the first principal component of the pooled data (shearing). However, both approaches are problematic. ANCOVA depends on assumptions (small variance in the covariate) that are frequently violated in this context. Residuals analysis assumes that scaling relationships within groups are equal, but this assumption is rarely tested. Furthermore, scaling relationships obtained from pooled data typically mischaracterize within-group scaling relationships. We discuss potential biases imposed by the application of ANCOVA and residuals analysis for quantifying morphological differences, and elaborate and demonstrate a more effective alternative: common principal components analysis combined with Burnaby's back-projection method.

Keywords Analysis of covariance · Common principal components · Residuals · Size correction · Shearing

Electronic Supplementary Material Supplementary material is available for this article at http://www.dx.doi.org/10.1007/s00442-006-0403-6 and is accessible for authorized users.

Communicated by Diethart Matthies

M. W. McCoy (✉) · B. M. Bolker · C. W. Osenberg
B. G. Miner · J. R. Vonesh
Department of Zoology, University of Florida,
Gainesville, FL 32611-8525, USA
E-mail: mmccoy@zoo.ufl.edu
Tel.: +1-352-3921685
Fax: +1-352-3923704

Present address: B. G. Miner
Biology Department, Western Washington University,
Bellingham, WA 98255-9160, USA

Present address: J. R. Vonesh
Department of Biology, Washington University,
St. Luis, MO 63130, USA

Introduction

Individuals can differ markedly in morphology within and between populations, due (for example) to genetic variation, environmental effects on development, or sexual dimorphism. Whether we are interested in developmental mechanisms, their ecological implications, or the evolution of morphological variation, we must be able to reliably estimate the magnitude of phenotypic differences among populations. Most developmental processes and morphological traits increase with body size, but researchers often want to separate differences in shape from differences in body size, and must therefore perform some kind of size-corrected analysis. Unfortunately, many of the statistical techniques currently used for size correction in ecological studies have serious flaws that appear to be unappreciated by many investigators. In this paper, we discuss traditional techniques for size correction and introduce an alternative approach.

The issues we discuss apply to many topics in ecology where size correction (or correction for any set of multivariate covariates) is required. Differences in phenotypic traits within or among populations can change the rate and direction of evolution, influence population dynamics, and determine the outcome of ecological interactions (Tollrian and Harvell 1999; Bolker et al. 2003; Werner and Peacor 2003), but we must measure phenotypic differences appropriately in order to determine their causes and effects. For example, sea urchin larvae increase the size of their feeding structures when reared at low food concentrations compared to larvae reared with higher food concentrations

(Boidron-Mètarion 1988; Miner 2005). These differences may result either from overall differences in body size or from disproportionately longer arms relative to the expected scaling with body size. Because these two hypotheses lead to different conclusions about sea urchin ecology and evolution, it is critical to separate differences in body size per se from differences in size-corrected morphology.

Three approaches are routinely used to control for the effects of size in ecological research: residuals analysis, shearing (a multivariate analogue of residuals analysis), and analysis of covariance (ANCOVA). Because most developmental and physiological processes increase as a power of body size, these analyses are typically performed on log transformed data. In residuals analysis, data for traits of interest are first regressed against a univariate description of body size (e.g., mass). The residuals from the regression are then compared across treatment groups using, for example, analysis of variance (ANOVA; e.g., Relyea and Hoverman 2003; Relyea 2004). Shearing is a specific form of residuals analysis that uses a multivariate description of body size. Measurements of multiple traits (typically all increasing with body size) are pooled across treatment groups and analyzed with a principal components analysis (PCA) to yield the first principal component, PC1, which is assumed to represent "body size" (i.e., a common allometry shared among groups: Somers 1986; Jolicoeur 1963; Humphries et al. 1981; Bookstein 1991). As in residuals analysis, traits are then regressed against this (multivariate) measure and the residuals compared using ANOVA (e.g., Van Buskirk and Relyea 1998; Relyea 2001). ANCOVA uses a univariate descriptor of body size as a covariate (e.g., Dahl and Peckarsky 2002), but differs from residuals analysis in using a pooled within-group regression coefficient (i.e., estimating a common slope for multiple groups) instead of using a regression coefficient obtained from pooled data (i.e., ignoring group structure by aggregating data: Winer et al. 1991; Sokal and Rohlf 1995; Garcia-Berthou 2001). Unfortunately, all three techniques have flaws that limit their use in analyses of size-corrected morphology. Despite these limitations, all three approaches are frequently applied in analyses of morphometric data. For example, the majority of papers published in *Ecology*, *Oecologia*, and *Oikos* between 1993 and 2004 that used body size correction to analyze morphological responses to environmental conditions used either residuals analysis (10/82, 12.2%), shearing (20/82, 24.4%), or ANCOVA (42/82, 51.2%: Table 1).

In the first part of this paper, we review the flaws implicit in standard approaches to size correction in ecology. We then develop an approach described by Klingenberg (1996) that combines common principal components analysis (CPCA) with Burnaby's back-projection method (BBPM). In addition, we develop and evaluate appropriate error propagation techniques that overcome the problems of both pooled regression approaches (such as residuals analysis and shearing) and Model I regression approaches (including ANCOVA).

Shortcomings of standard methods

By characterizing the (assumed) shared allometry of multiple groups using pooled data, residuals analyses confound within- and between-group patterns and underestimate effect sizes when the dependent and confounding variables are correlated, which is typically true in studies where size correction is needed (Darlington and Smulders 2001). Although it has not previously been noted, shearing (a multivariate residuals approach) similarly underestimates effect sizes (see below). PCA provides a multivariate description of allometry for a single group, not multiple groups; when data from multiple groups are pooled, PC1 does not necessarily reflect the underlying scaling relationships within any of the original groups (Fig. 1b, c). In studies that use size-corrected data, researchers are typically looking for changes in allometry (characterized by between-group differences) over and above those related to size (often characterized by within-group variation), and thus the working hypothesis is that within- and between-group principal components are different. Shearing (one of the most commonly used multivariate size-correction techniques) fails under these circumstances, and the only situation in which the within-group relationship ($PC1_I$) will be the same as the pooled relationship ($PC1_p$) is when groups are only displaced along a single axis of allometric variation (Fig. 1a). In such cases, traits differ between groups only because organisms differ in overall size. Therefore, shearing is only appropriate in the case where size-correction cannot reveal any other differences among groups.

Analysis of covariance resolves some of the problems associated with residuals analysis (and shearing). Although ANCOVA also requires an estimate of a shared allometric axis, it tests for differences in allometry by testing for heterogeneity of slopes and it estimates this slope by pooling within-group regression coefficients rather than by obtaining a single regression coefficient from pooled data. ANCOVA also reduces the associated degrees of freedom by one because the slopes of the within-group regression lines are estimated from the data (Garcia-Berthou 2001). Size correction is a particularly interesting application of ANCOVA because size (the covariate) is an inherent part of the morphological measurements (which comprise the response variables). However, in all applications, ANCOVA assumes the covariate is measured without error and that its distribution is similar among groups (Huitema 1980; Sokal and Rohlf 1995). When the distribution varies among groups and the covariate is measured with error, which is presumably true in most studies of size correction, ANCOVA may produce biased estimates of effect sizes (Huitema 1980; Sokal and Rohlf 1995). This bias occurs because ANCOVA is based on Model I regression,

Table 1 Summary of the quantitative techniques used to analyze morphological response variables and perform size correction in studies of phenotypic plasticity

Method	Pooled data[a]	Size[b]	Error model[c]	No. of studies	Percentage of univariate studies	Percentage of multivariate studies
Residuals analysis	Yes	Univariate	Type I	10	18	–
ANCOVA	No	Univariate	Type I	41	73	–
Other[d]	Yes	Univariate	Type I	5	9	–
Shearing	Yes	Multivariate	Type I	19	–	73
ANCOVA	No	Multivariate	Type I	1	–	4
CPC/BBPM	No	Multivariate	Type II	0	–	0
Other[e]	Yes/No	Multivariate	–	6	–	23

We focused on phenotypic plasticity because it is a common ecological context in which morphology is compared among different groups. Data were obtained from a review of 82 papers published in *Ecology*, *Oecologia*, and *Oikos* between 1993 and 2004

[a]Indicates whether analyses were based on pooled data without respect to group identity

[b]Size was either a univariate measure (e.g., mass) or multivariate (e.g., from PCA)

[c]Type I error (no error in the size measure); Type II (assumes error in X and Y)

[d]Includes ratio and regression analyses

[e]Includes use of PCA and thin spline multiple warp analyses

which attributes all error to the dependent variable (for a detailed discussion of the effects of violating ANCOVA assumptions, see Appendix I). Sokal and Rohlf (1995, p. 519) have warned that ANCOVA's assumptions are frequently violated in biological applications and they did not recommend its use in situations where the covariate cannot be measured exactly until Model II regression techniques (which allow for error in both the dependent and independent variables) are developed. We present such methods here.

We advocate a size-correction approach that overcomes these three shortcomings of prior methods: it (1) uses a multivariate measure of body size; (2) avoids pooling (and residual analysis); and (3) uses a Model II approach that incorporates error in the measurement of body size. In particular, we describe CPCA as a method to test assumptions of shared allometry and estimate allometric axes common to multiple groups (Flury 1988; Klingenberg 1996; Phillips and Arnold 1999) and BBPM as a tool for size correction (sensu Burnaby 1966; Klingenberg 1996). We also provide improved estimation procedures using back-projection that account for error in the estimation of the allometric patterns described using CPCA. We compare results obtained from simulated datasets using shearing (the most common multivariate approach to date), ANCOVA, and CPCA/BBPM.

Although some of the techniques presented here have been recognized and widely used in morphometric and evolutionary genetic studies to compare variance–covariance matrices, they have rarely been used to estimate magnitudes of change and are apparently unknown to ecologists quantifying differences in morphological traits (e.g., Klingenberg and Spence 1993; Klingenberg 1996). We hope this discussion and subsequent applications of CPCA/BBPM will improve ecologists' approaches to size correction and estimation of morphological differences among groups: e.g., in studies of phenotypic plasticity where the use of potentially problematic approaches is the norm (Table 1).

Estimating an allometric axis

Common principal components analysis

Common principal components analysis is a generalization of PCA to applications involving more than one group. It can be used to test multivariate relationships among groups (i.e., to determine if covariance matrices are similar and to what degree) and can thus be used to infer if groups share common patterns of allometry. CPCA provides a test analogous to the test for homogeneity of slopes in ANCOVA—checking whether a common scaling relationship exists. CPCA has become popular in evolutionary biology for comparing phenotypic and genotypic covariance matrices (Steppan 1997a, b; Arnold and Phillips 1999; Phillips and Arnold 1999) and has also been applied to problems in multivariate allometry, although less frequently (Klingenberg and Spence 1993; Klingenberg and Zimmerman 1992).

Common principal components analysis defines levels of similarity among covariance matrices (Flury 1988; Phillips and Arnold 1999). For most studies of morphological plasticity and size correction, it is only the first CPC (CPC1) that is of interest because it describes the general scaling of traits with body size. For our purposes, Flury's (1988) hierarchical test of similarities can be reduced to three cases: (1) equality or proportionality of all principal components (i.e., all principal components are identical among groups, although the variances of the data in any particular dimension may differ among groups); (2) equality or proportionality of eigenvectors for CPC1, but not necessarily other components; and (3) dissimilarity among covariance matrices (covariance matrices do not share CPC1).

If the first principal components are not shared (Fig. 1c), then the patterns of morphological variation are so fundamentally different that there can be no globally applied size correction because "size" does not

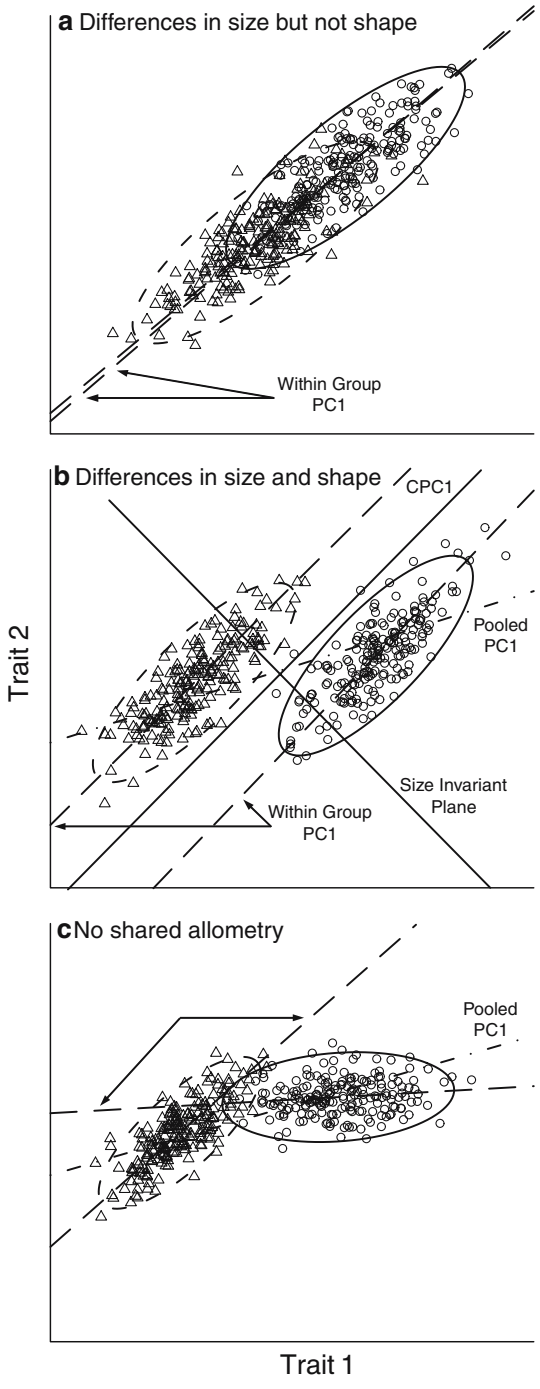

Fig. 1 Simulated two-variable data sets for two groups (*open triangles* and *open circle*) depicting three possible patterns. **a** The two groups share all of their principal components in common but are offset along the allometric or size axis. Thus, the within-group PC1s, pooled PC1 and CPC1 are the same (i.e., groups differ in size but not in shape). **b** Groups share CPC1 but have different allometry, so the within-group PC1s and CPC1 are parallel but pooled PC1 is not. The data are size adjusted by projecting them onto the size invariant plane (i.e., setting CPC1 equal to zero) and then projecting back into their original units. In other words, the size adjusted data are the perpendicular residuals expressed in terms of the original axes. **c** The groups have arbitrary covariance structure. Therefore, only within-group PC1s and pooled PC1 are shown because the groups do not have a first common principal component. *Solid lines* CPC1. *Dashed-dotted lines* PC1 from pooled data. *Dotted lines* within-group PC1. *Ellipses* demarcate 95% confidence limits of the two groups

have a common meaning across groups. On the other hand, if all principal components are common to all groups (Fig. 1a), then the groups show identical patterns of within-group covariation, indicating the same allometries of traits with size in all groups. As long as the groups to be compared share CPC1 (Fig. 1a, b) and all the traits have strong loadings on CPC1, then it can be interpreted as a common body size dimension: i.e., a "size axis" that can be used to evaluate between-group differences in morphology that go beyond simple (isometric) changes in size (Klingenberg and Spence 1993; Klingenberg and Zimmerman 1992; Klingenberg 1996). Thus, CPCA provides an initial test of a common body size (growth) dimension and a quantitative means to describe the growth axis and thus achieve body size corrections.

As with tests of heterogeneity of slopes in ANCOVA, CPCA may fail to detect between-group differences in CPC1 even when they exist. In reality, CPC1 will always differ to some degree among groups (as do the slopes of traits with respect to covariates in ANCOVA), but the hope is that any real differences are so small that the assumption of a common growth axis can still be used to draw robust inferences about size-corrected traits. Ultimately, this depends on the power of the analysis to detect problematic levels of heterogeneity. Houle et al. (2002) have raised concerns about the statistical power of CPCA. Our analyses (Appendix II) suggest that in size-correction contexts, where the main goal is to correct for the effects of a growth axis that explains a large fraction of the total variation (e.g., >90%), plausible sample sizes as small as 20 individuals per group provide sufficient power to detect small differences in CPC1.

Size correction

Burnaby's back-projection method

Burnaby (1966) proposed a general procedure for removing the effect of growth from multivariate morphometric data sets for multiple populations. BBPM involves projecting data points onto a plane (if the data set has more than two dimensions (variables), or a line for two dimensions) orthogonal to the size axis common to all groups (Fig. 1b; Klingenberg 1996). Because CPC1 can be interpreted as an allometric pattern shared by all groups, the size invariant plane can be calculated by setting CPC1 equal to zero, which is equivalent to calculating the perpendicular residuals around CPC1 (Fig. 1b; Burnaby 1966; Klingenberg 1996). The perpendicular residuals are then projected back into the original units and treated as growth-adjusted or growth-

invariant data (Klingenberg 1996) and can be analyzed via standard MANOVA or ANOVA techniques.

These growth-adjusted data can be calculated using the method developed by Burnaby (1966):

$$\mathbf{X}\left(\mathbf{I}-\boldsymbol{\mu}(\boldsymbol{\mu}'\boldsymbol{\mu})^{-1}\boldsymbol{\mu}'\right), \quad (1)$$

where $\mathbf{X}$ is an $n\times p$ data matrix (n is the number of observations and p the number of traits/variables measured), $\mathbf{I}$ is a $p\times p$ identity matrix and $\boldsymbol{\mu}$ is a p-element vector representing the growth axis (i.e., CPC1). Because CPCA components are normalized vectors with $\boldsymbol{\mu}'\boldsymbol{\mu}=1$, this equation can be simplified to $\mathbf{X}(\mathbf{I}-\boldsymbol{\mu}\boldsymbol{\mu}')$ (Klingenberg 1996). We can break up this equation to see that the first multiplication $\mathbf{X}\boldsymbol{\mu}$ projects $\mathbf{X}$ onto the first principal direction (calculating a scalar that is the score for CPC1) and the second multiplication (multiplying by $\boldsymbol{\mu}'$) translates principal component scores back into the original coordinate system (Burnaby 1966).

Materials and methods

To investigate how CPCA/BBPM may change analyses of morphological differences among groups, we simulated 300 three-dimensional data sets and analyzed them using (1) PCA and shearing, (2) ANCOVA, and (3) CPCA/BBPM. Because ANCOVA is a univariate analysis we illustrate the comparison for the second and third traits using trait one as the covariate (i.e., the measure of body size) because trait one had the highest correlation with the true size axis (Appendix I provides a detailed discussion of the conditions under which the use of ANCOVA may be inappropriate).

Simulated data sets illustrate scenarios that might be expected in empirical studies: (1) two groups that share all of their principal components in common (identical covariance matrices) but are offset (10×) along the allometric axis (i.e., the groups differ in overall size but not in size-corrected shape: Fig. 1a); (2) data from two groups with common covariance structure but with one group offset along the CPC2 axis (20×; *note*-the data depicted in Fig. 1b are only offset by 10 along the CPC2 axis) (i.e., the groups differ in size-corrected shape: Fig. 1b); and (3) data from two groups with arbitrary covariance structure (i.e., the groups share no common allometry: Fig. 1c).

Data were randomly drawn from a multivariate normal distribution using specified covariance matrices (Appendix III). Therefore, true differences in our simulated data were known, permitting us to compare the estimates of effect size in our analyses with true differences. For each of 300 simulations per scenario, we drew 200 values per group and applied shearing and CPC/BBPM. We repeated simulations for a wide set of conditions [e.g., up to eight traits, and varying offsets along CPC1 (overall size differences) and perpendicular to CPC1 (size-corrected differences in morphology)]. For simulations where there was no difference in shape (scenario 1 above), we determined Type I error rates by performing a one-way ANOVA on the back-projected data values for a particular trait (variable), but replacing the uncorrected sum of squares in the denominator of the F-ratio with the sum of the uncorrected and back-projection sums of squares (Appendix IV). Under all cases CPC/BBPM yielded Type I error rates close to the nominal value of 5%.

All analyses were performed using the R statistical programming environment (R Development Core Team 2004; http://www.R-project.org). Our program code (as an R library), which includes BBPM, the CPCA program developed by Patrick Phillips (Phillips and Arnold 1999) and a variety of other functions is available at http://www.zoo.ufl.edu/bolker/R/windows/ (see Appendix III).

Error propagation

Error in estimates of the magnitude of morphological divergence using CPCA incorporates both among-individual variation within groups (standard errors of group means) and additional variation caused by error in the estimated value of CPC1 used to size-correct the data. Both sources of error should be included in confidence intervals of back-projected trait values (Appendix IV); ignoring error in CPC1 will lead to inappropriately narrow confidence intervals (the depictions of CPCA in Fig. 1 ignore this second source of variation, but the confidence intervals presented in Fig. 2 incorporate it).

Results

Differences in body size only

When groups shared the same primary axis of variation (Fig. 1a), pooled PCA and CPCA both provided the same estimate of the allometric axis, as expected. Furthermore, analysis of size-corrected data using both PCA/shearing and CPCA/BBPM found no difference in size-corrected traits for the two groups (Fig. 2a). In this scenario, both PCA/shearing and CPCA/BBPM led to the correct conclusion that the groups differed only in overall size and not in shape (Fig. 2a). In contrast, ANCOVA incorrectly identified differences in traits 2 and 3. These differences stemmed from the error in the covariate that resulted from the lack of perfect correlation between the covariate and the true size axis (the observed bias would decrease if the groups were more broadly overlapping in size or if we used a covariate more correlated with true size: Appendix I).

Differences in size-corrected shape

When groups shared a common size axis, but were offset from this primary axis of variation (Fig. 1b), all three

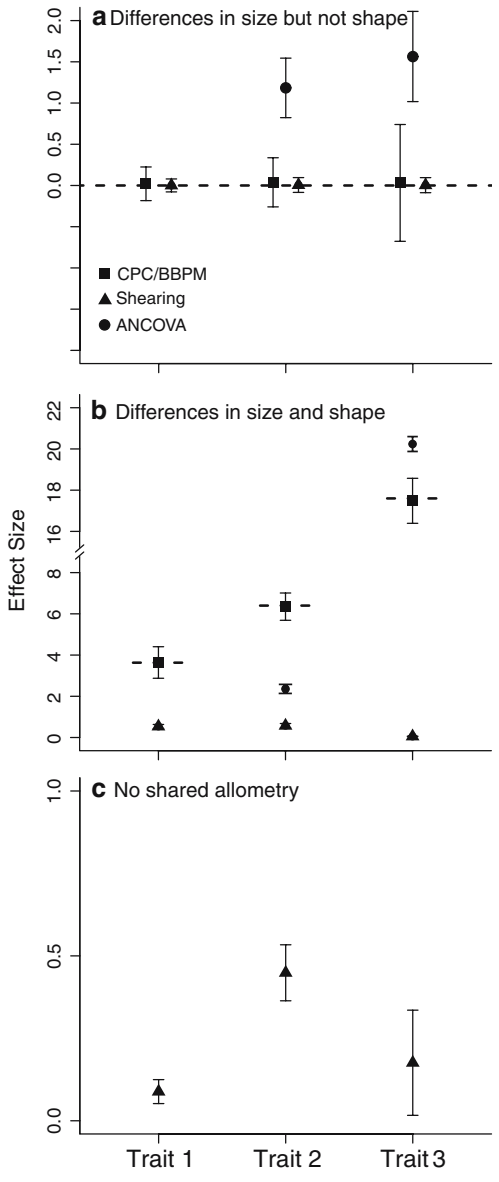

Fig. 2 Results from analyses of simulated data sets representing three scenarios (Fig. 1) using CPC/BBPM, ANCOVA, and shearing. Results illustrate the average difference from 300 simulations (±1 standard deviation), for CPC/BBMP (*filled square*), shearing (*filled circle*), and ANCOVA (*filled triangle*). *Dashed lines* give the true differences between the two groups. **a** Results for groups that differed in size but not shape (Fig. 1a). The true difference in size corrected traits between the two groups equaled zero for all traits. **b** Results for groups that differed in size and shape (Fig. 1b). For illustration we have presented the negative of the effect size for Trait 2 in panel **b**. True differences in size corrected traits between the two groups were: Trait 1 = 3.6, Trait 2 = −6.4, and Trait 3 = 18.6. **c** Results for groups that did not share CPC1. Because groups did not share CPC1 we have only shown the results from shearing. In this case the correct conclusion is that the scaling of traits was different between the two groups. No size corrected comparisons were possible (i.e., there is no size-corrected difference to be estimated) and thus BBPM was not applied

techniques detected significant differences in size-corrected morphology between groups. However, the approaches gave grossly different estimates of these differences. Only CPCA/BBPM provided estimates of between-group differences that were accurate and unbiased. In contrast, shearing severely underestimated the true effect sizes (Fig. 2b: although very small (< 1), the differences were significantly different from zero). ANCOVA performed better than shearing, but underestimated the magnitude of differences in trait 2 while over-estimating the differences in trait 3. These biases stemmed from the imperfect correlation between the covariate and the true size (Appendix I, Fig. A1).

No common body size axis

When the groups did not share a common growth axis, the inferences based on shearing were qualitatively inconsistent with those of CPCA/BBPM (and ANCOVA) and quantitatively biased with respect to the true differences between the groups. Although the differences in Fig. 2c are significant, shearing generated the spurious conclusion that the groups were not very different morphologically. Conversely, analyzing these data using CPCA/BBPM or ANCOVA resulted in the correct conclusion that the two groups did not share a common size axis and therefore no further analysis of size-corrected morphological traits was possible—there was no single way to adjust for size that is applicable to both groups, so such an attempt would be invalid (Fig. 1c). In this scenario, one would conclude that the groups are divergent in the allometric relationships among traits, and that the scaling of traits (e.g., the exponents in power functions that relate two traits) and not their offset (i.e., coefficients in the power functions) is heterogeneous among populations. CPCA/BBPM (and ANCOVA to a lesser degree) allows investigators to separately evaluate these two different forms of morphological divergence. For example, if CPC1 differs among groups (i.e., there is no common, single measure of "body size"), then further analysis should focus on the heterogeneity of the scaling relationships themselves (e.g., how the exponents in the power functions vary among groups). For these cases, one could apply an approach described by Krzanowski (1979, 1988) and Blows et al. (2004) to quantify the degree to which groups differ and to what extent individual traits contribute to differences in allometry.

Discussion

Pooled PCA and shearing (or other forms of residuals analysis) are frequently used to estimate and correct for size in groups of organisms. For example, in our review of recent papers on phenotypic plasticity, 76% that used a multivariate estimate of body size used pooled data,

while only one used CPCA and none used CPCA/BBPM. Because the application of PCA to multiple groups (and residuals analysis in general) tacitly assumes that groups have identical allometries, results derived from PCA/shearing or residuals analyses will often be biased (e.g., Fig. 1b, 2b). ANCOVA was used in 73% of the studies that used a univariate estimate of body size. This technique is preferable to shearing/residuals analysis; however, it can also lead to considerable bias and elevated Type I and Type II error rates (Fig. 2a, b, Appendix I).

The problems with PCA and shearing that we have identified result from the basic mathematical assumptions of these techniques. PC1 describes the axis that incorporates the greatest amount of variance in the data. Therefore, when calculated for pooled data, it is influenced by variation both within and among groups and thus may confound interpretation of allometry with among-group variation (Klingenberg 1996). The problem of confounding within and among group variance is most severe when groups have arbitrary covariance structure (Fig. 1c). Similar problems arise in residuals analyses in which univariate measures of body size are used to represent the allometric axis (Smith 1999; Darlington and Smulders 2001; Garcia-Berthou 2001).

Common principal components analysis is clearly more powerful, and appropriate, than PCA for summarizing common allometry for multiple groups and as a mechanism for testing assumptions of similarity of covariance matrices. In addition, the use of CPCA in conjunction with BBPM provides an effective tool for comparing the size of specific traits among groups and for quantifying differences in shape. As more studies are added to the literature, it also becomes possible to use meta-analysis (Osenberg et al. 1999) to compare the magnitude of variation in morphology among different taxa, cues, or environments (e.g., Van Buskirk 2002). However, our results raise cautions about using existing studies that have used residuals or shearing analyses for quantitative syntheses of the literature. Additionally, the use of different analytic tools (PCA/shearing vs CPCA/BBPM) is likely to introduce significant variation to the literature that is unrelated to the biology being investigated (see Osenberg et al. 1997, 1999 for related cautions). Furthermore, restricting meta-analyses to the most common approaches (e.g., shearing) will give severely biased results that may obscure real patterns. Indeed, some recent studies have demonstrated substantial survival and growth consequences resulting from (apparently) very small differences in morphology (e.g., Relyea and Hoverman 2003; Van Buskirk and McCollum 2000). However, these studies typically used shearing or residuals analysis and therefore it is likely that the true morphological differences among groups were far greater than was originally estimated (e.g., Fig. 2b). More accurate estimates of effect size in studies of morphology (through the application of CPC/BBPM) will allow us to better compare variation among studies and link functional responses with fitness consequences.

Acknowledgments We thank members of the St. Mary-Osenberg-Bolker lab group and two anonymous reviewers for helpful discussions and comments on previous drafts of this manuscript. We also thank Brian Langerhans and Jonathan Losos for their comments on an earlier draft. We thank Rich Kiltie for pointing us to the literature on CPCA. This work was partially funded by NSF (OCE 0325028 to S. Morgan and B.G.M. and OCE 0242312 to C.W.O., B.M.B. and C. St. Mary) and EPA (STAR Fellowship to J.R.V.). Contribution number 2307 Bodega Marine Laboratory, University of California, Davis, USA.

References

Arnold SJ, Phillips PC (1999) Hierarchical comparison of genetic variance covariance matrices. II. Coastal inland divergence in the garter snake, *Thamnophis elegans*. Evolution 53:1516–1527

Blows MW, Chenoweth SF, Hine E (2004) Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. Am Nat 163:329–340

Boidron-Mètairon IF (1988) Morphological plasticity in laboratory-reared chinoplutei of *Dendraster excentricus* (Eschscholtz) and *Lytechinus variegatus* (Lamarck) in response to food conditions. J Exp Mar Biol Ecol 119:31–41

Bolker BM, Holyoak M, Křivan V, Rowe L, Schmitz O (2003) Connecting theoretical and empirical studies of trait-mediated interactions. Ecology 84:1101–1114

Bookstein FL (1991) Morphometric tools for landmark data: geometry and biology. Cambridge University Press, New York, 435 pp

Burnaby TP (1966) Growth-invariant discriminant functions and generalized distances. Biometrics 22:96–107

Dahl J, Peckarsky BL (2002) Induced morphological defenses in the wild: predator effects on a mayfly, *Drunella coloradensis*. Ecology 83:1620–1634

Darlington RB, Smulders TV (2001) Problems with residuals analysis. Anim Behav 62:599–602

Flury B (1988) Common principal components and related multivariate models. Wiley, New York

Garcia-Berthou E (2001) On the misuse of residuals in ecology: testing regression residuals vs. the analysis of covariance. J Anim Ecol 70:708–711

Houle, D, Mezey, J, Galpern, P (2002) Interpretation of the results of common principal components analyses. Evolution 56:433–440

Huitema BE (1980) The analysis of covariance and alternatives. Wiley, New York

Humphries JM, Bookstein FL, Chernoff B, Smith GR, Elder RL, Elder SG (1981) Multivariate discrimination by shape in relation to size. Syst Zool 30:291–308

Jolicoeur P (1963) The multivariate generalization of the allometry equation. Biometrics 19:497–499

Klingenberg CP (1996) Multivariate allometry. In: Marcus LF, Corti M, Loy A, Naylor GJP, Slice DE (eds) Advances in morphometrics. Plenum Press, New York, pp 23–49

Klingenberg CP, Spence JR (1993) Heterochrony and allometry—lessons from the water strider genus *Limnoporus*. Evolution 47:1834–1853

Klingenberg CP, Zimmerman M (1992) Static, ontogenetic and evolutionary allometry: a multivariate comparison in 9 species of water striders. Am Nat 140:601–620

Krzanowski WJ (1979) Between-group comparisons of principal components. J Am Stat Assoc 74:703–707

Krzanowski WJ (1988) Principles of multivariate analysis: a user's perspective. Clarendon, Oxford

Miner BG (2005) Evolution of feeding structure plasticity in marine invertebrate larvae: a possible trade-off between arm length and stomach size. J Exp Mar Biol Ecol 315:117–125

Osenberg CW, Sarnelle O, Cooper SD (1997) Effect size in ecological experiments: the application of biological models in meta-analysis. Am Nat 150:798–812

Osenberg CW, Sarnelle O, Cooper SD, Holt RD (1999) Resolving ecological questions through meta-analysis: goals, metrics and models. Ecology 80:1105–1117

Phillips PC, Arnold SJ (1999) Hierarchical comparison of genetic variance covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515

Relyea RA (2001) Morphological and behavioral plasticity of larval anurans in response to different predators. Ecology 82:523–540

Relyea R (2004) Fine-tuned phenotypes: tadpole plasticity under 16 combinations of predators and competitors. Ecology 85:172–179

Relyea RA, Hoverman JT (2003) The impact of larval predators and competitors on the morphology and fitness of juvenile treefrogs. Oecologia 134:596–604

Smith RJ (1999) Statistics of sexual size dimorphism. J Hum Evol 36:423–459

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. Freeman, New York

Somers KM (1986) Multivariate allometry and removal of size with principal components analysis. Syst Zool 35:359–368

Steppan SJ (1997a) Phylogenetic analysis of phenotypic covariance structure. 1. Contrasting results from matrix correlation and common principal component analyses. Evolution 51:571–586

Steppan SJ (1997b) Phylogenetic analysis of phenotypic covariance structure. 2. Reconstructing matrix evolution. Evolution 51:587–594

Tollrian RR, Harvell CD (1999) The ecology and evolution of inducible defenses. Princeton University Press, New Jersey

Van Buskirk J (2002) A comparative test of the adaptive plasticity hypothesis: relationships between habitat and phenotype in anuran larvae. Am Nat 160:87–102

Van Buskirk J, McCollum SA (2000) Functional mechanisms of an inducible defense in tadpoles: morphology and behavior influence mortality risk from predation. J Evol Biol 13:336–347

Van Buskirk J, Relyea RA (1998) Natural selection for phenotypic plasticity: predator-induced morphological responses in tadpoles. Biol J Linn Soc 65:301–328

Werner EE, Peacor SD (2003) A review of trait-mediated indirect interactions in ecological communities. Ecology 84:1083–1100

Winer BJ, Brown DR, Michels KM (1991) Statistical principles in experimental design. McGraw Hill, New York

ELECTRONIC SUPPLEMENTARY MATERIAL

FOR

Size correction: comparing morphological traits among populations and environments

By

Michael W. McCoy[1], Benjamin M. Bolker[1], Craig W. Osenberg[1], Benjamin G. Miner[1,2], and James R. Vonesh[1,3]

TABLE OF CONTENTS

APPENDIX I
CONSEQUENCES OF VIOLATING ASSUMPTIONS OF ANALYSIS OF COVARIANCE

Analysis of Covariance (ANCOVA), when applied as a method of size correction, uses a univariate descriptor of body size as a covariate to compare the sizes of morphological traits of different groups. ANCOVA uses a pooled within-group regression coefficient (i.e., estimating a common slope for multiple groups), and corrects the degrees of freedom used for comparisons because the slopes of the within-group regression lines are estimated from the same data. ANCOVA tests for differences in allometry by testing for heterogeneity of slopes. In principle, this approach should overcome the most obvious criticism of residuals analysis and shearing (which are predicated on pooled regressions).

The standard ANCOVA model is:

$$Y_{ij} = \mu + \alpha_i + \beta\left(X_{ij} - \overline{X}\right) + \varepsilon_{ij},$$

where the j^{th} observation in group i of the independent variable Y, is a function of four primary components: 1) the grand mean of all observations of the dependent variable μ; 2) the effect of treatments α_i; 3) the effect of the covariate $\beta\left(X_{ij} - \overline{X}\right)$; and 4) the error ε_{ij} (Huitema 1980). Applications of this model require a number of assumptions (Huitema 1980 and Sokal and Rohlf 1995), but here we deal explicitly with two assumptions unique to ANCOVA that are important for its application in size correction: 1) the covariate and the effect of the treatment are independent, and 2) the covariate is fixed and measured without error. Although we discuss these two assumptions separately they are intimately linked and the potential for introducing bias in ANCOVA is greatest when they are both violated.

When the covariate and the effect of treatment are correlated (e.g., if treatments affect body size as well as size-corrected morphology) the interpretation of ANCOVA results can be ambiguous. In such cases, the within-group covariate means will be different from the grand covariate mean and when averaged may generate a grand covariate mean that is not (and potentially cannot be modified to be) shared by the treatment groups (Huitema 1980, Sokal and Rohlf 1995). Therefore, comparing adjusted means may be problematic because the covariate may either add or mask differences between treatment groups that may be misinterpreted as real treatment effects. Fortunately, the assumption of independence of covariate and treatments in ANCOVA can be tested by performing an analysis of variance (ANOVA) on the covariate; however it is difficult to know the degree of bias (if any) that is introduced by a significant result (Huitema 1980)

The second important assumption is that the covariate is fixed and measured without error (Huitema 1980, Sokal and Rohlf 1995). If the covariate is in fact measured with error, which is presumably true in morphological studies, then ANCOVA may lead either to positive or negative bias or to an inflated Type I error rate. The problem is that ANCOVA is based on Model I regression, which attributes all error to the independent variable. However there are three potential sources of error (the third is specific to applications for size correction):

1. error in the dependent variable;
2. measurement error in the covariate;
3. imperfect correlation between the variable being used as the covariate and the true size growth axis;

We investigated these potential sources of bias in ANCOVA for two different scenarios. First, we tested for bias when two groups were offset along the covariate but did not differ once the covariate was incorporated. Second, we tested for bias

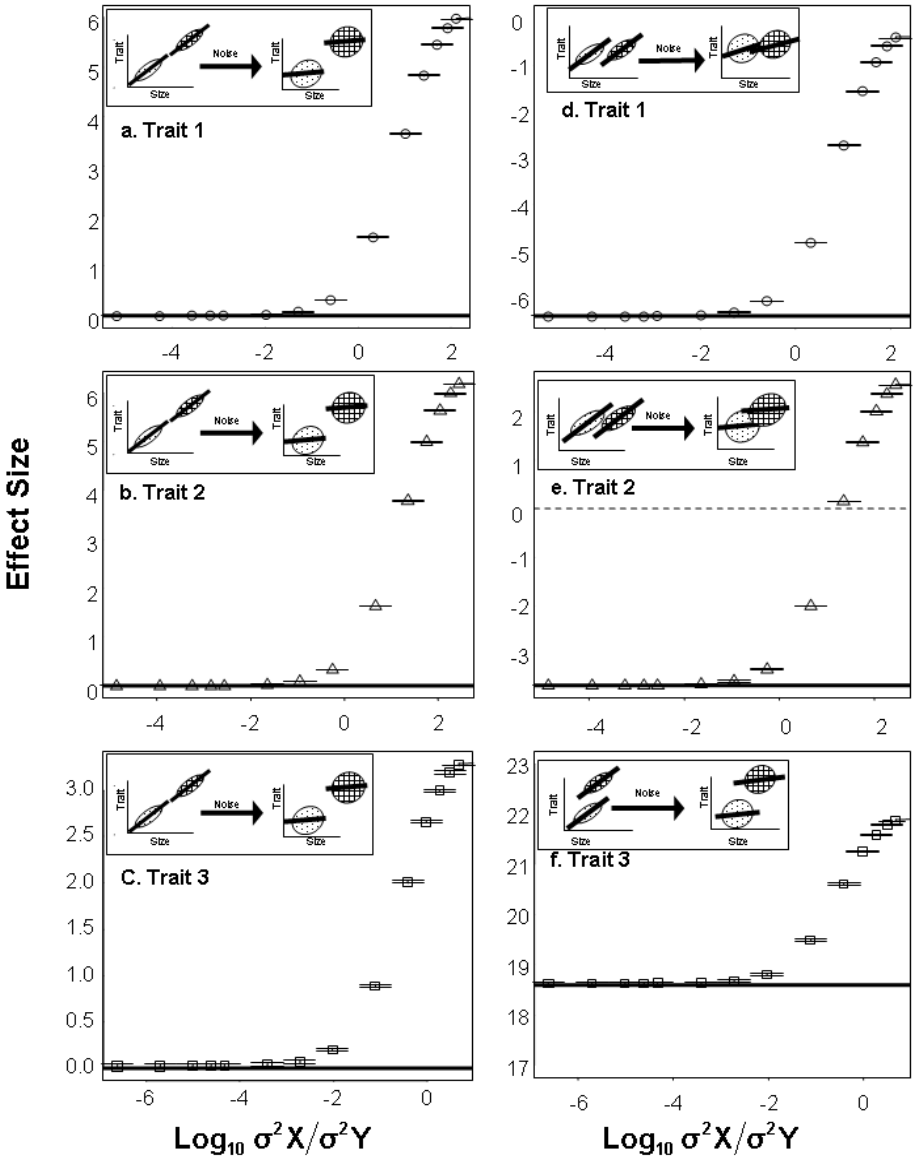

Figure A1: Results from simulations of ANCOVA analyses. Data represent means and 95% confidence intervals for estimates of effect size from 500 runs. Inset figures illustrate schematically how adding error ("Noise") to the covariate re-sults in a particular direction of bias. In panels **a**, **b**, and **c** there is no true difference in the two groups. In panels **d**, **e**, and **f** the true diff-erences are indicated by the solid horizontal line. Panels **a** and **d** illustrate the estimates for trait 1, **b** and **e** illustrate trait 2 and **c** and **f** illustrate the results for trait 3. For each trait we see that non-negligible bias occurs when the error in the covariate is 1 to 10% of the error in the response variable.

when two groups were offset along the covariate dimension and in trait space. For each scenario we generated 500 3 dimensional data sets for each of two groups. Each data set consisted of 200 points randomly drawn from a multivariate normal distribution (see appendix III) (these data were generated using the same parameters as the data used for figures 1 and 2). For each set of simulated data we ran an ANCOVA and saved estimates of effect size. We repeated this procedure 10 times for each scenario but added error to the covariate for each successive run (the error added to the covariate incorporates sources 2 and 3 from above).

We found that adding measurement error to the covariate increased the probability of finding a difference between the two groups when there was no difference (i.e. Type I error rate was increased) (Figure A1a-c). When there was a true difference between two groups, we found that adding error to the covariate could result in underestimation, overestimation, and even a shift in the direction of effect (Figure A1d-f). These results stem from the problems

associated with using a Model I regression (error only in the dependent variable) to analyze data with error in both the dependent and independent variables. Adding error to the covariate spreads the data horizontally, which reduces the slope of the within-group regression lines. For the null case this results in an inability to accept the null and in all other cases it makes interpretation of ANCOVA results ambiguous. ANCOVA does not always give misleading results, even when the covariate is measured with error. For example, when the range in the covariate was the same for both groups, we found no effects of error in the covariate (results not illustrated). For the scenarios used here (where the groups are offset along the covariate axis) it appears that ANCOVA performs well until the error in the covariate is more than 60% of the error in the response variable. Furthermore, in some systems, body size measured in mass may involve less error than linear traits. However, the problem of accurately representing the true size growth axis is not solved by precise measurement. Lack of correlation between a chosen covariate and the true body size/growth axis can also lead to biased results (source 3 above). We urge researchers who use ANCOVA as a method of size-correction to be aware of the potential biases imposed and to interpret their results cautiously.

Literature Cited

Huitema B (1980) *The Analysis of Covariance and Alternatives*. John Wiley and Sons, New York.

Sokal, RR, Rohlf FJ (1995). *Biometry: The Principles and Practice of Statistics in Biological Research.* 3rd edition. Freeman, NewYork.

APPENDIX II

ESTIMATING POWER OF CPCA FOR IDENTIFYING SHARED ALLOMETRY

Houle et al. (2002) expressed concern that CPCA (as implemented by Phillips and Arnold 1999) might lack power and therefore lead to the conclusion that there are common PCs when none actually exist. For example, Houle et al. (2002) found that relatively large sample sizes were required to detect differences in the orientation of first principal components. However, the correlation structure of the data affects the validity and power of CPCA. Houle et al.'s data sets had approximately 75% of the variance explained by within-group PC1s. In contrast, in studies of multivariate allometry, within-group PC1s should explain >85% of the variance in the data to be considered a good estimate of size (Jolicoeur 1963). Indeed, in most morphological datasets, within-group PC1 often explains over 95% of the variance. If high loadings on CPC1 increase the power of CPCA to detect differences in CPC1, then Houle et al.'s concerns may not be relevant to most studies using size-correction.

Therefore, we determined how the proportion of the variance explained by PC1 affected the power of CPCA to detect a fixed divergence of two groups in PC1 with a given sample size. We constructed 3-variable variance-covariance matrices for two groups where the variance associated with each principal component decreased geometrically (e.g. 1, 0.2, 0.04) and where the first and second principal components of the second group were rotated by a specified angle relative to those of the first group. We varied the fraction of variance explained by PC1 on a log series between 0.72 and 0.98, and varied sample sizes per group in a log series between 10 and 200 (both ranges are typical for studies of morphological plasticity). For each fraction of variance/sample size combination we tested 10 angles (evenly spaced between 0 and 45 degrees), and each angle was run 500 times to determine power: we interpolated the power curve to calculate the angle between true PC1 for which we had 80% power to detect a difference in the PC1s between groups (Figure A2).

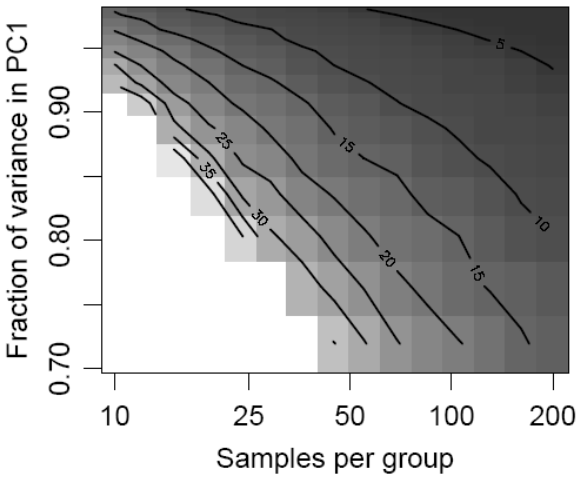

Figure A2. Contour of the critical angle (angle for which there is 80% power to detect a difference in PC1 between groups) for α=0.05 and the specified combination of sample size and fraction of variance in PC1.

Power increased with sample size and angle (i.e., the difference in the body size dimension). If 75% of the variance was associated with PC1, a 15^o angle required a sample size of ~150 per group, and a sample of 20/group could not even detect a 45^o displacement 80% of the time. These results confirm the results of Houle et al. (2002), who used a dataset with 75% of the variance associated with PC1. However, most morphometric studies have >90% (often >95%) of the variance loading on PC1. For 95%, a 15^o angle can be reliably detected with a sample size of only 20/group. Samples of 50/group can detect displacements of <10^o. Thus, we conclude that CPC is a powerful tool in morphometric analyses and that failure to detect substantial violation of common allometry is likely to be rare in morphometric studies using CPCA/BBMP for size-correction.

Literature Cited

Jolicoeur P, (1963) The multivariate generalization of the allometry equation. Biometrics 19:497-9.

Houle D, Mezey J, Galpern P, (2002) Interpretation of the results of common principal components analyses. Evolution 56:433-440.

APPENDIX III

PARAMETER VALUES AND THE "R" PACKAGE

We simulated data that represented a multivariate set of morphological measurements for two groups of individuals (e.g. prey exposed to predators and prey not exposed to predators). We assumed that there was some underlying allometry that determined how individuals that change in size also changed in shape as a result. We also assumed that the data have been appropriately transformed (e.g. by logging all trait values) to make them multivariate normally distributed with constant variance-covariance matrices, independent of size. To generate these data we wrote a function that created identically shaped but offset multivariate normal groups. We used the variance matrix

$$\begin{pmatrix} 10 & 8 & 2 \\ 8 & 10 & 3 \\ 2 & 3 & 10 \end{pmatrix},$$

with principal directions (eigenvectors) (-0.65, -0.67, -0.35) [eigenvalue=19.3]; (0.32, 0.18, -0.93) [eigenvalue=8.7]; and (0.68, -0.72, 0.09) [eigenvalue=1.93].

We used standard algorithms coded in the MASS package of R to draw 200 points for each group and shifted the mean of group 2 (group 1's mean was located at the origin, (0,0,0)) by specifying distances (offsets) along the first and second principal directions. For our null cases (difference only in size but not size-corrected shape), the offset was 10 along PC1 and 0 along PC2, leading to a group 2 mean of 10 e_1 = (-6.5,-6.7,-3.5). For cases where there were differences in size but not shape, the offset was 10 along PC1 and 20 along PC2, leading to a group 2 mean of 10 e_1 + 20 e_2 = (-0.11,-3.08,-22.14).

The R Environment and Package cpcbp

Source or binary versions of the latest version of R, which is a free and open source program, can be downloaded from the R website along with documentation (http://www.r-project.org). The binary or source code for R Package "cpcbp" is available for download via the Internet at

http://www.zoo.ufl.edu/bolker/R/windows/cbpbp_0.1.2.zip

or

http://www.zoo.ufl.edu/bolker/R/src/cpcbp_0.1.2.tgz

R includes instructions for installing add-on packages across an Internet connection or from a local ZIP file. Below is a list of the functions that are available in our CPC/BBMP package (this list can be obtained during an R session by typing "package (help=cpcbp)").

Description of package:

Package: cpcbp

Title: Common principal components and back-projection analysis

Author: Ben Bolker

Maintainer: Ben Bolker

Depends: R (>= 2.0.0)

Description: Auxiliary functions for
CPC and Flury back-projection analysis

Index (function names and descriptions):

APPENDIX IV
ESTIMATING ERROR IN CPC1

To quantify and account for the error arising from estimation of the size axis (CPC1) we generated a multivariate dataset that represented morphological measurements on individuals from two groups of individuals (e.g. prey exposed to predators and prey unexposed to predators). We assume that there is some underlying allometry by which individuals that change in size will also change in shape (on an appropriate scale, e.g. log-transformed trait values). We aim to separate changes in shape caused by phenotypic plasticity from changes that are simply due to changes in size.

To do this, we calculate common principal components (CPCA) for within-group variation, back-project to eliminate the effects of the first CPC (CPC1), and perform univariate analyses of the resulting size-standardized traits separated by group. We make two assumptions here (1) within-group allometric variation in size-related traits is a good proxy for between-group variation in size, and (2) CPC1 characterizes effects of size (e.g. CPC1 has positive loadings for all traits).

The back-projection equation is:

$$\mathbf{X}(\mathbf{I} - \boldsymbol{\beta}_1\boldsymbol{\beta}_1') \quad (1)$$

where $\mathbf{X}$ is an $n \times p$ data matrix, $\mathbf{I}$ is a $p \times p$ identity matrix (n is the total number of observations and p is the number of traits/variables measured) and $\boldsymbol{\beta}_1$ (following Flury's (1988) notation, $\boldsymbol{\beta}_j$ is the jth eigenvector, treated as a column vector and $\boldsymbol{\beta}_{ij}$ is the ith element of the jth eigenvector) is the estimated first principal direction (eigenvector), scaled so that $\boldsymbol{\beta}_1'\boldsymbol{\beta}_1 = 1$. To understand this formula, break up equation 6 to see that the first multiplication $\mathbf{X}\boldsymbol{\beta}_1$ projects $\mathbf{X}$ onto the first principal direction (calculating a scalar that is the score for CPC1). The second multiplication (multiplying by $\boldsymbol{\beta}_1'$) translates this score back into the original coordinate system (Klingenberg 1996, Burnaby 1966). Given this back projection process, any error present in CPC1 is propagated into the back projected data (i.e. size-corrected trait values) (Fig. 2; main text).

To account for this error we compute the errors on the elements of the eigenvector $\boldsymbol{\beta}_1$ (see Flury 1988: the following discussion up to equation 5 recapitulates pp. 74-85 and equation numbers 2.12 – 4.8). We start by computing

$$\hat{\theta}^{i}_{jh} = r_i^{-1} \frac{\hat{\lambda}_{ij}\hat{\lambda}_{ih}}{\left(\hat{\lambda}_{ij} - \hat{\lambda}_{ih}\right)^2} \quad (2)$$

where $r_i = n_i/n$ (fraction of total data points in group i) and $\hat{\lambda}_{ij}$ and $\hat{\lambda}_{ih}$ are estimates of the j^{th} and h^{th} eigenvalues of group i's variance-covariance matrix. Given $\hat{\theta}^{i}_{jh}$ we can calculate a harmonic mean across k groups

$$\hat{\theta}_{jh} = \left(\sum_{i=1}^{k} \left(\hat{\theta}^{i}_{jh}\right)^{-1} \right)^{-1} \quad (3)$$

and find the large-sample estimate of the standard error of β_{mh} to be

$$s(\beta_{mh}) = \left(\frac{1}{n} \sum_{j=1, j\neq h}^{p} \hat{\theta}_{jh}\beta_{mj}^2 \right)^{1/2} \quad (4)$$

where β_{mh} is the m^{th} element of the h^{th} principal component. More generally we know that the variance-covariance matrix of the elements in $\boldsymbol{\beta}_1$ is:

$$\frac{1}{n}\sum_{h=2}^{p} \hat{\theta}_{1h}\boldsymbol{\beta}_h\boldsymbol{\beta}_h' \quad (5)$$

Suppose we have calculated the error variances $\sigma^2_{\beta_{1j}}$ for each component of the first eigenvector. Then the ij^{th} element of the outer-product matrix $\boldsymbol{\beta}_1\boldsymbol{\beta}_1'$ is $b_{ij} = \beta_{1i}\beta_{1j}$. In general, the errors for two quantities can be combined by (Lyons 1991)

$$V(f(a,b)) \approx V(a)\left(\frac{\partial f}{\partial a}\right)^2 + V(b)\left(\frac{\partial f}{\partial b}\right)^2 + 2C(a,b)\left(\frac{\partial f}{\partial a}\frac{\partial f}{\partial b}\right) \quad (6)$$

which for $f(a,b) = a \cdot b$ reduces to

$$V(a)b^2 + V(b)a^2 + 2C(a,b)ab = a^2b^2\left(\frac{V(a)}{a^2} + \frac{V(b)}{b^2}\right) + 2C(a,b)ab \quad (7)$$

The approximation in (6) is based on a Taylor series expansion; when $f(a,b) = a \cdot b$ the only missing term is $E\left[(a-\bar{a})^2 \cdot (b-\bar{b})^2\right]$ which in turn is equal to $V(a) \cdot V(b) + C(a^2, b^2)$; we found this term to be generally negligible, as shown by our good type I error results based on (7).

Therefore the error variance of $\beta_{1i}\beta_{1j}$ is approximately

$$\sigma^2_{b_{ij}} = \sigma^2_{\beta_{1i}\beta_{1j}} = (\beta_{1i}\beta_{1j})^2\left(\frac{\sigma^2_{\beta_{1i}}}{\beta^2_{1i}} + \frac{\sigma^2_{\beta_{1j}}}{\beta^2_{1j}}\right) + 2\sigma_{\beta_{1i},\beta_{1j}}\beta_{1i}\beta_{1j} \quad (8)$$

where $\sigma_{a,b}$ denotes the covariance between *a* and *b*. The covariances of the elements of the back-projection matrix ($b_{ij} = \beta_{1i}\beta_{1j}$) with each other must also be calculated. We derive these covariances by expanding $E\left[b_{ij}b_{ik}\right] = E\left[\beta_{1i}\beta_{1j} \cdot \beta_{1i}\beta_{1k}\right]$ and assuming all third and fourth central moments are zero:

$$\sigma_{b_{ij},b_{ik}} = 2\left(\beta_{1i}\beta_{1j}\sigma_{\beta_{1i},\beta_{1k}} + \beta_{1i}\beta_{1k}\sigma_{\beta_{1i},\beta_{1j}}\right) + \beta_{1i}^2\sigma_{\beta_{1j},\beta_{1k}} + \beta_{1j}\beta_{1k}\sigma^2_{\beta_{1i}} \quad (9)$$

Now, we calculate the error variance introduced into the mean of the i^{th} variable by

$$\sigma^2_{i,BP} = \sigma^2_{1i}\bar{x}_i^2 + \sum_{k \neq j} \bar{x}_j\bar{x}_k\sigma_{b_{ij}b_{jk}} \quad (10)$$

and combine the back projection error and the within group variance.

$$\sigma^2_{x_i} = \sigma^2_{i,BP} + x_i^2 \cdot \sigma^2_{\bar{x}} \quad (12)$$

This now provides us with an estimate of the combined variances of each group. These combined variances can then be used as the variance terms in a t-test. Alternatively, back-projection error can be incorporated into an analysis of variance by adding the back projection sums of squares to the error sum of squares. In order to calculate the overall back-projection sum of squares for use in a corrected ANOVA we substitute the sum of the absolute deviations of the trait means of each group ($\bar{x}_{i,j}$ for trait *i* in group *j*) from the overall mean of each trait ($\bar{\bar{x}}_i$), or $\sum_j \left|\bar{x}_{i,j} - \bar{\bar{x}}_i\right|$, for the mean trait values in the calculations above. This procedure is necessary because the back-projection error affects the back-projected data in every group in the same direction, not independently. (An R package that implements this algorithm is available via http://www.zoo.ufl.edu/bolker/R/windows/).

Literature Cited

Burnaby T P (1966) Growth-invariant discriminant functions and generalized distances. Biometrics 22: 96 – 107.

Flury B, (1988) *Common Principal Components and Related Multivariate Models*. Wiley, New York.

Klingenberg CP (1996) Multivariate allometry. In LF Marcus, Corti M, Loy A, Naylor GJP, Slice DE (eds.). *Advances in Morphometrics*. Plenum Press, New York, pp. 23-49.

Lyons L (1991) *A practical guide to data analysis for physical science students*, Cambridge University Press, Cambridge, England.