Evolutionary Ecology Research, 2007, **9**: 261–282

A phylogenetic approach to determining the importance of constraint on phenotypic evolution in the neotropical lizard *Anolis cristatellus*

Liam J. Revell,[1]* Luke J. Harmon,[2] R. Brian Langerhans[1] and Jason J. Kolbe[3]

[1]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA,* [2]*Biodiversity Centre, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada and* [3]*Department of Biology, Indiana University, Bloomington, IN 47405-3700, USA*

ABSTRACT

Question: Is the pattern of phenotypic divergence among populations influenced by constraint in the form of the genetic covariances among characters?

Background: Quantitative genetic theory predicts that when evolutionary lineages diverge simultaneously by genetic drift, the pattern of among-population divergence will parallel the pattern of within-population genetic variation and covariation. Among-population divergence is measured by the variance–covariance matrix of population means (the **D** matrix), or by the variance–covariance matrix of independent contrasts ($\mathbf{D}_{\mathbf{IC}}$). The latter avoids the assumption of simultaneous divergence by incorporating phylogenetic non-independence among lineages and was developed expressly for this study. Within-population genetic variation and covariation are measured by the additive genetic variance–covariance matrix (the **G** matrix).

Organism: The Puerto Rican crested anole (*Anolis cristatellus*).

Methods: We sampled *A. cristatellus* from seven divergent populations widely dispersed across the species' range. These populations are sufficiently highly diverged to be considered evolutionarily independent lineages. We substituted the phenotypic variance–covariance matrix (**P** matrix) for **G** in evolutionary analysis. (Empirical studies have shown that **P** and **G** are frequently highly correlated for morphological traits.) In two populations, we estimated phenotypic variance–covariance matrices (**P** matrices) for 13 skeletal morphological traits, while in the remaining five we estimated mean phenotypes for the same traits. To test the hypothesis of constraint, we first calculated a pooled phenotypic variance–covariance matrix ($\bar{\mathbf{P}}$) from all populations. We compared $\bar{\mathbf{P}}$ to the variance–covariance matrix of population means (**D**) and of independent contrasts ($\mathbf{D}_{\mathbf{IC}}$). Independent contrasts were calculated using a molecular phylogeny of the included lineages.

Results: Comparison of **P** matrices between populations showed evidence that covariance structure is highly conserved in conspecific populations of *A. cristatellus*. Comparison of $\bar{\mathbf{P}}$ with **D** and of $\bar{\mathbf{P}}$ with $\mathbf{D}_{\mathbf{IC}}$ indicated significant similarity in both cases, suggesting that constraint has influenced phenotypic evolution and thus probably genotypic evolution in this species.

Keywords: evolutionary constraint, genetic variance–covariance matrix, independent contrasts, matrix comparison, phenotypic variance–covariance matrix, quantitative genetics.

* Author to whom all correspondence should be addressed. e-mail: lrevell@fas.harvard.edu

INTRODUCTION

A perennial challenge in evolutionary biology is to quantify the effects of constraint on the course and outcome of phenotypic evolution (Maynard Smith *et al.*, 1985; Barton and Partridge, 2000). Genetic constraint resulting from additive genetic correlations among characters is often postulated to play an important role in evolution (Cheverud, 1984; Arnold, 1992; Arnold *et al.*, 2001). Such constraint tends to cause genetically correlated traits to evolve together, biasing the path of multivariate phenotypic evolution away from that which maximizes the increase in mean fitness (Lande, 1979; Schluter, 1996, 2000; Arnold *et al.*, 2001; Fornori *et al.*, 2003). In particular, the single generation response to selection in a given trait is a function of the additive genetic variance for the trait, the selection gradient, β, or partial regression coefficient of fitness on phenotype, the linear selection gradients on other traits, and the genetic covariances between them (Lande, 1979; Lande and Arnold, 1983). Expressed another way, $\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$, where $\Delta\bar{\mathbf{z}}$ is a vector of the change in population means for traits, $\mathbf{G}$ is a square matrix with additive genetic variances on the diagonal and covariances elsewhere, and $\boldsymbol{\beta}$ is a vector of selection gradients (Lande, 1979; Lande and Arnold, 1983). Consequently, the short-term phenotypic response to selection is rarely in the direction of greatest increase in mean fitness, $\boldsymbol{\beta}$. As such, the genetic variances for and covariances between traits can play a central role in constraining the course of phenotypic evolution.

By extension, when many closely related populations or species simultaneously radiate from a common ancestor while experiencing similar genetic constraints, the pattern of multivariate phenotypic differentiation among populations is also influenced by genetic constraint. In particular, under genetic drift and selection, the pattern of divergence is a function of $\mathbf{G}$, the elapsed time (t), the effective population size (N_e), and the variance–covariance matrix of multivariate selection gradients (cov[$\boldsymbol{\beta}$]) (Lande, 1979; Felsenstein, 1985; Hansen and Martins, 1996; Arnold *et al.*, 2001). With no net selection, this function simplifies to $\mathbf{D} = (t/N_e)\mathbf{G}$ under genetic drift alone, where $\mathbf{D}$ is the variance–covariance matrix of population means for traits summarizing the multivariate divergence among populations (Lande, 1979; Blows and Higgie, 2003). In general, then, if evolution is primarily by genetic drift, the pattern of multivariate phenotypic differentiation among populations will be correlated with the pattern of genetic variation within them (Kluge and Kerfoot, 1973; Sokal, 1978; Lande, 1979; Björklund and Merilä, 1993; Schluter, 1996; Arnold *et al.*, 2001; Bégin and Roff, 2004). Genetic constraint thus makes some patterns of phenotypic diversification among populations or species more probable than others.

Lande's (1979) equations rely implicitly on a simultaneous radiation of the populations under study from a single common ancestor. However, they can be readily extended to the more realistic case of bifurcating phylogenetic relationships among taxa. Using the assumption of a Brownian motion model of character evolution (which is the appropriate model for genetic drift under many circumstances), the phenotypic variance among taxa descended from a common ancestor is proportional to the product of the time elapsed and the mean rate of evolution for that trait in the involved lineages (Lande, 1979). When taxa radiate according to a bifurcating phylogeny, the mean rate of phenotypic evolution for a character can be estimated by the mean square of the independent contrasts (Felsenstein, 1985; Garland, 1992; Garland *et al.*, 1992; Martins, 1994). Similarly, the mean cross-product of the independent contrasts for two characters has an expected value proportional to their mean rate of co-evolution. Thus, phylogenetic non-independence among included species or evolutionarily independent populations can be incorporated by calculating the mean squares and

mean cross-products (MS-MCP) matrix of independent contrasts, which we designate herein as $\mathbf{D}_{IC}$, and substituting this matrix for **D** in tests of the constraint hypothesis. This is similar to the test of Baker and Wilkinson (2003) in which the correlation matrix of independent contrasts was used. However, using the MS-MCP matrix allows the direct test of the hypothesis that the rate and direction of correlated character evolution among populations corresponds to the amount and direction of phenotypic variation within populations, whereas the correlation matrix of independent contrasts tests only the hypothesis that genetically correlated characters evolve together.

Some previous tests of the constraint hypothesis revealed evidence that genetic constraint is important (Sokal and Riska, 1981; Schluter, 1996; Baker and Wilkinson, 2003; Bégin and Roff, 2003, 2004; Marroig *et al.*, 2004), while others found no or only weak correspondence between **D** and **G** (Lofsvold, 1988; Venable and Burquez, 1990; Merilä and Björklund, 1999; Badyaev and Hill, 2000). Although contributing greatly to our understanding of the general importance of constraint, a shortcoming of these studies (Baker and Wilkinson, 2003, excluded) is their failure to account for the phylogenetic non-independence of the taxa included in the study.

Ideally, studies of genetic constraint should compare the structure of $\mathbf{D}_{IC}$ with that of **G**, the additive genetic variance–covariance matrix. Unfortunately, in many species **G** is not particularly amenable to estimation. In these cases, the pattern of phenotypic variation and covariation – the **P** matrix or phenotypic variance–covariance matrix – is sometimes used in lieu of **G** for evolutionary analyses (Cheverud, 1988; Reusch and Blanckenhorn, 1998; Waitt and Levin, 1998). Since $\mathbf{P} = \mathbf{G} + \mathbf{E}$, where **E** is a matrix of environmental variances and covariances, there is considerable debate about under what circumstances **P** will resemble **G** (Cheverud, 1988; Willis *et al.*, 1991). Some empirical evidence suggests that they are often highly correlated, particularly for morphological traits (Cheverud, 1988; Roff, 1995; Reusch and Blanckenhorn, 1998). Furthermore, assuming that **E** and **D** are uncorrelated, discordance between **P** and **G** will probably increase only type II, not type I, error rates in the test for constraint employed in this study.

In this study, we utilize the pattern of phenotypic variances for and covariances among quantitative traits (the **P** matrix) to study multivariate evolution in a neotropical lizard, *Anolis cristatellus*, the Puerto Rican crested anole. Lizards in the genus *Anolis*, or anoles, are a model adaptive radiation, with over 150 species of diverse morphologies known from the Caribbean isles alone (Williams, 1972). Repeated evolution of similar forms on different islands, with inter-island convergents termed 'ecomorphs', has often been cited as evidence for the potency of natural selection in an adaptive radiation (Losos *et al.*, 1998). However, previous studies of phenotypic diversification in *Anolis* have largely neglected the possible importance of genetic constraint. This study focuses on characterizing the importance of constraint at the level of phenotypically differentiating populations of a single species. As such, this paper constitutes the first direct test of the importance of constraint in phenotypic evolution in anoles.

This study uses data collected from seven highly divergent populations of *A. cristatellus* to (1) test the hypothesis that covariance structure is conserved at the intraspecific level in a widespread species, and (2) test the hypothesis that the pattern of genetic variances and covariances within populations is aligned with the pattern of phenotypic differentiation among them. The latter is a test of the constraint hypothesis (Bégin and Roff, 2004), and thus positive evidence in this test would suggest that genetic constraint has influenced the phenotypic differentiation of populations of *A. cristatellus*. In addition to characterizing the role of constraint on morphological evolution in a species of anole, we also (3) compare the results from tests of the constraint hypothesis conducted in both non-phylogenetic and

phylogenetic contexts, and (4) evaluate the extent to which correspondence between **P** and the divergence matrices derived from the population means for traits (**D** and $\mathbf{D}_{IC}$) could arise due to sampling error in the estimation of population means.

METHODS

Overview

To investigate the stability of the pattern of phenotypic covariation at an intraspecific level, we estimated **P**, the phenotypic variance–covariance matrix, for each of two highly divergent populations of *A. cristatellus*, chosen for the substantial intraspecific genetic differentiation known to exist between them (R.E. Glor *et al.*, in preparation; see also Results). Using matrix comparison (Mantel test, random skewers, CPCA, the *T* method), we then evaluated the evolutionary stability of covariance structure at an intraspecific level.

To investigate the role of constraint in morphological differentiation, we estimated multivariate mean phenotypes for five additional populations widely spaced across the species' range. In all, we estimated within-group phenotypic covariance matrices (**P** matrices) for two populations, and multivariate mean phenotypes for seven populations.

To test the hypothesis that intraspecific differentiation among populations is influenced by constraint imposed by the pattern of variation and covariation within populations, we compared the alignment of the variance–covariance matrix of population means for traits, **D**, with the phenotypic variance–covariance matrix, **P**. When **D** is estimated from the multivariate phenotypes of individuals drawn from populations or species that share common history, calculating **D** ignoring the phylogeny or phylogeographic tree relating species or populations in the sample may not be appropriate. Thus, we estimate a bifurcating tree from molecular sequence data for the seven populations of interest in this study. We also calculate $\mathbf{D}_{IC}$, the mean squares and mean cross-products (MS-MCP) matrix of multivariate independent contrasts (Felsenstein, 1985; Baker and Wilkinson, 2003), and compare **P** to $\mathbf{D}_{IC}$.

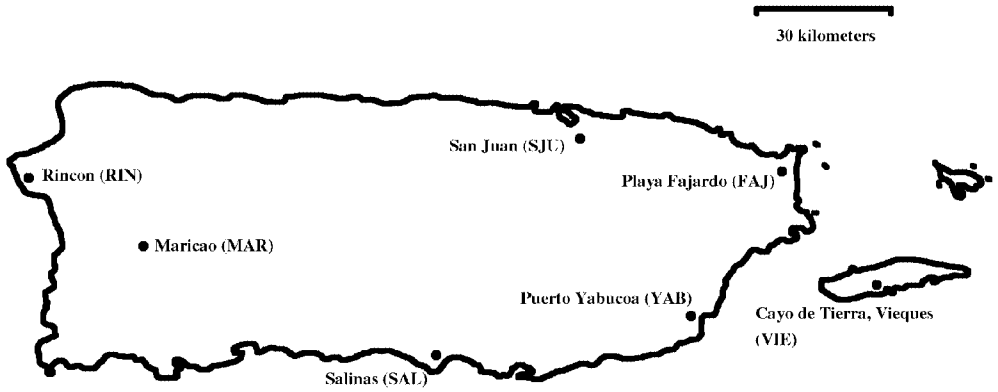
Finally, **P** and **D** or $\mathbf{D}_{IC}$ may also be aligned because phenotypic traits with high variance within populations have means subject to high error in their estimation. Thus, we use a bootstrap randomization procedure to test the hypothesis that the correlation between **P** and **D** and that between **P** and $\mathbf{D}_{IC}$ exceed the correlations expected as a consequence of sampling error in the estimation of **D** and $\mathbf{D}_{IC}$.

Morphological measurements

We collected a total of 357 adult males from seven populations of *A. cristatellus*. Two of these populations were focal populations for **P** matrix estimation. The first ($n = 156$) was located on the southeast coast of the main island of Puerto Rico, near Puerto Yabucoa (YAB; Fig. 1A). The second ($n = 141$) was on a peninsular cay on the south coast of the island of Vieques, a satellite island located approximately 20 km east of the main island (VIE; Fig. 1A). The remaining five populations, widely dispersed across the main island, were San Juan (SJU; $n = 12$), Fajardo (FAJ; $n = 12$), Salinas (SAL; $n = 12$), Maricao (MAR; $n = 11$), and Rincon (RIN; $n = 13$) (Fig. 1).

We fixed all specimens in 85% ethyl-alcohol, and collected X-ray radiographs using a Hewlett-Packard Faxitron cabinet X-ray System. We digitized all X-rays using a flatbed film

A.



B.

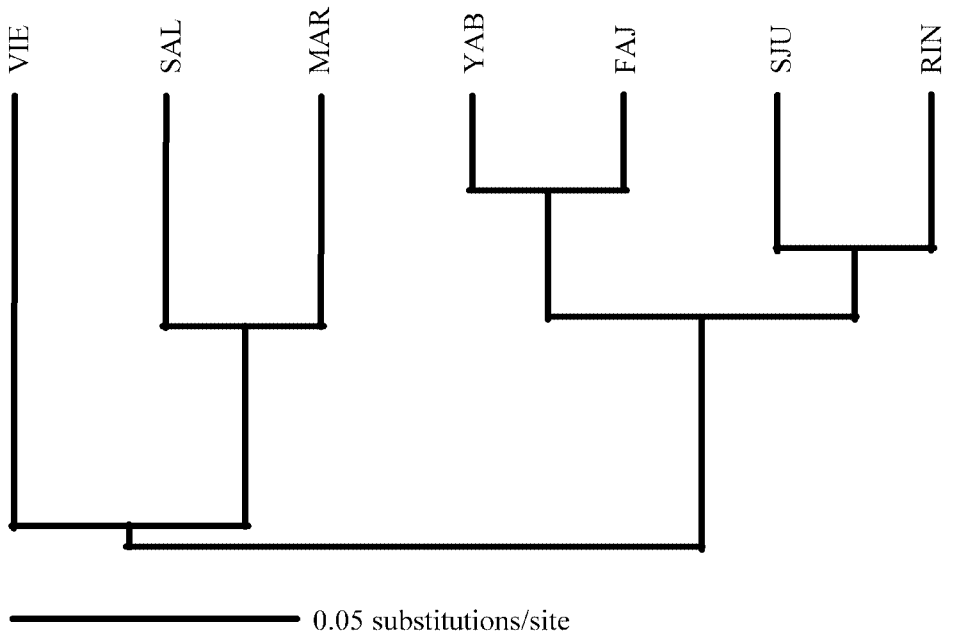


Fig. 1. (A) A map of Puerto Rico depicting the localities from which specimens were collected. Focal populations for **P** matrix estimation were Puerto Yabucoa (YAB) and Cayo de Tierra, Vieques (VIE). (B) A phylogeny of the sampled populations, estimated from ND2 (mtDNA NADH 2 gene), $tRNA^{Trp}$, and $tRNA^{Ala}$ sequence using ML in PAUP*.

scanner. We collected measurements for 13 morphological traits from each X-rayed lizard using the computer program tpsDig (Rohlf, 2005). The 13 traits were skull length, skull width, first metacarpal length, radius length, ulna length, humerus length, pectoral girdle width, pelvis width, femur length, tibia length, fibula length, and first and second metatarsal lengths. We measured body length skeletally by digitizing the posterior processes of each vertebra between the base of the skull and the base of the pelvis and summing the distances between them. We also measured snout-to-vent length (SVL) externally. Snout-to-vent length and body length measured skeletally were highly correlated ($r = 0.96$), but skeletal body length had slightly lower standardized variation [CV(skeletal body length) = 0.071; CV(SVL) = 0.081], which implies that it was measured with less error, and as such is used in all subsequent analyses.

For all bilaterally symmetrical traits, we took measurements from both right and left sides of each individual and averaged them. This procedure minimized the measurement error variance as well as other non-genetic causes of variation such as fluctuating asymmetry (Lynch and Walsh, 1998).

P matrix estimation and comparison

We natural logarithm transformed all morphological measurements and removed the effect of size from all traits by calculating the residuals from standard least squares linear regressions of each trait on skeletally measured body length. We calculated **P** using standard methods for the estimation of variances and covariances. We estimated the standard errors of the elements of **P** by a delete-one jackknife approach (Manly, 1997).

Several different matrix comparison methods exist (Steppan *et al.*, 2002), and there is considerable contention about their relative merits (e.g. Bégin *et al.*, 2004). It is not our intention to compare existing methods; however, due to their potential to yield different information about the manner in which matrices vary (Steppan *et al.*, 2002; L.J. Revell, submitted), we compared **P** matrices using several methods. Methods used included the Mantel test (Mantel, 1967), random skewers (Cheverud *et al.*, 1983), common principal components analysis [CPCA (Flury, 1988; Steppan, 1997; Phillips and Arnold, 1999)], and the T method (Roff *et al.*, 1999).

We used a Mantel test to test for similarity of the phenotypic correlation matrices. This test involves calculating the matrix correlation ($r_{\mathbf{M}}$) in the typical way, and then randomly permuting the rows and columns together in the dependent matrix, each time recalculating $r'_{\mathbf{M}}$ from the permuted matrices. One-tailed significance is evaluated as the fraction of times in which $r'_{\mathbf{M}} \geq r_{\mathbf{M}}$.

A shortcoming of the Mantel test is that it cannot be used for covariance matrices (Dietz, 1983). Random skewers (Cheverud *et al.*, 1983; Cheverud, 1996; Marroig and Cheverud, 2001) is a method that can be applied to covariance matrices. With random skewers, matrices are multiplied by random selection vectors and the correlation coefficient between the resultant response vectors (r_s) is calculated and averaged across a large number of random vectors. The test statistic is thus the mean vector correlation coefficient, and its significance is determined by comparison to the expected distribution of correlations between random vectors (Cheverud, 1996). We used random skewers to test for similarity of the phenotypic covariance matrices.

Common principal components analysis provides a very useful framework in which to test a series of alternative hypotheses of matrix similarity beginning with the test of one common principal component versus unrelated structure, and ending with matrix equality versus matrix proportionality (Flury, 1988; Steppan, 1997; Phillips and Arnold, 1999). At each step in the

hierarchy, a likelihood ratio test is used to determine whether adding additional parameters (e.g. proportionality has one more parameter than equality: the proportionality constant) significantly improves the model fit to the data (Phillips and Arnold, 1999). The likelihood ratio is evaluated against a χ^2 distribution with degrees of freedom determined by the difference in the number of parameters in the CPCA models. Models are evaluated hierarchically, and the appropriate CPCA model is selected when the likelihood ratio test is not significant. An alternative approach to CPCA model selection ['model-building approach' (Phillips and Arnold, 1999)] utilizes the Akaike information criterion [AIC (Akaike, 1973; Phillips and Arnold, 1999)], a statistic that weighs the goodness of fit of a model against the number of parameters used to fit that model. Under this approach, the best fitting CPCA model is selected because it has the lowest value for the AIC (Phillips and Arnold, 1999). Common principal components analysis provides considerable information about the manner in which matrices share common structure. However, it is affected by limitations identified in previous studies (Steppan, 1997; Phillips and Arnold, 1999; Ackermann and Cheverud, 2000; Marroig and Cheverud, 2001; Steppan *et al.*, 2002). We used CPCA to compare phenotypic covariance matrices.

With the exception of CPCA (in which several hierarchical hypotheses are tested, including one of matrix equality) the matrix comparison methods discussed in the above paragraphs use a null hypothesis of no shared structure. However, this null may be inappropriate for evolutionary data in which covariance matrices are estimated for populations or species connected by gene flow or united by common history (Turelli, 1988). As such, we also use the T method of Roff *et al.* (1999), which explicitly utilizes a null hypothesis of matrix equality. In this test, the observed sum of absolute values of the differences between the elements of two matrices, $\mathbf{P_1}$ and $\mathbf{P_2}$ (T_{obs}), is compared to the sum of the absolute values of the differences between the elements of phenotypic covariance matrices obtained when individuals (as multivariate deviations from their population means) are permuted randomly between populations (T_r). The significance of the test is calculated as the proportion of times that the value of T_r from the randomization test exceeds T_{obs} (Roff *et al.*, 1999). The sum of squared differences between the elements of $\mathbf{P_1}$ and $\mathbf{P_2}$ (or T^2_{obs}) can alternatively be used (Roff *et al.*, 1999). We compared phenotypic covariance matrices using both the T and T^2 tests.

Phylogeny estimation

We sampled for mitochondrial DNA (mtDNA) sequence one specimen per population from each of the seven populations represented in this study (Fig. 1A). Other genetic results from this species have generally found little mtDNA genetic variability within populations (R.E. Glor *et al.*, in preparation; J.J. Kolbe, unpublished), thus we deemed one individual per population to be adequate phylogeographic sampling. We extracted genomic DNA from liver or muscle tissue stored in 95% ethanol using Viogene Tissue Extraction Kits (Viogene, Inc.). Using the polymerase chain reaction, we amplified approximately 1200 bases of the mitochondrial genes ND2 (NADH dehydrogenase subunit 2), $tRNA^{Trp}$, and $tRNA^{Ala}$ using an initial denaturation of 95°C for 35 s, annealing at 53°C for 35 s, and extension at 70°C for 150 s for 30 cycles. We purified all PCR products on 1.0% low-melt agarose gel and extracted template using Viogene Gel Extraction Kits (Viogene, Inc.). We sequenced template in both directions using ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kits with AmpliTaq DNA polymerase (Perkin Elmer) using the primers H5730, L4882c, and L4437. We analysed sequencing products on an ABI Prism 3700 DNA Analyser (Applied

Biosystems) according to the manufacturer's protocols. DNA sequence alignment was mostly unambiguous as there were no insertions or deletions through much of the sequenced region among our samples, with the exception of within the tRNA genes which we aligned manually according to secondary structure models.

We estimated the tree and branch lengths using Maximum Likelihood (ML) in PAUP* 4.0b10 (Swofford, 2002) under the GTR + I + Γ model of nucleotide substitution and assuming a molecular clock. The phylogenetic analysis included the seven sequences generated for this study, as well as ND2, $tRNA^{Trp}$, and $tRNA^{Ala}$ sequence from closely related *Anolis scriptus* (GenBank accession: AY296200). We used *A. scriptus* sequence only to root the phylogeny and excluded it from all subsequent analyses.

Comparison of P and D

To assess the extent to which populations were differentiated in phenotype, we performed a multivariate analysis of variance (MANOVA) with the morphological data as dependent variables and population of origin as the independent variable.

Due to the high matrix similarity between the focal populations for **P** matrix estimation (see Results), data from all seven populations were pooled to estimate a phenotypic variance–covariance matrix for use in all comparisons with **D**. The pooled matrix was calculated in a standard way by estimating **P** for all populations and calculating the element by element weighted mean, $\bar{\mathbf{P}}$, of all separate **P** matrices (Manly, 2005). Weights were proportional to the sample sizes of the populations. This procedure generates a pooled within-population variance–covariance matrix correcting for the mean differences among populations. To test the hypothesis that constraint has influenced the pattern of phenotypic differentiation among populations, we compared this matrix, $\bar{\mathbf{P}}$, with **D**, the variance–covariance matrix of population means for traits. In addition to calculating **D**, we also computed $\mathbf{D}_{IC}$, a phylogenetically independent variance–covariance matrix of population means for traits. $\mathbf{D}_{IC}$ was calculated as the mean squares and mean cross-products matrix of standardized multivariate independent contrasts for traits.

The conceptual justification for the estimation of the among-population variance–covariance matrix in this way arises from the prediction, derived from Lande (1979), that if many species or populations simultaneously radiate from a common ancestor under drift and constraint, they have an expected variance–covariance matrix of population means (**D**) proportional to **G**. Assuming that the ancestral state is equivalent to the mean of the descendants, **D** in this analysis is exactly equivalent to the mean squares and mean cross-products matrix of the set of $\Delta\bar{\mathbf{z}}$ vectors, in which $\Delta\bar{\mathbf{z}}_i$ is a vector of the net change in multivariate mean phenotype between the common ancestor and the taxon, *i*, under study (Fig. 2). In the contrasts matrix, calculating the mean squares and cross-products matrix of all $\Delta\bar{\mathbf{z}}$ vectors along internodes and terminal branches in the phylogeny would inflate the degrees of freedom of the variances and covariances, since ancestral states for nodes are not known independently of the states at descendant nodes. Instead, we utilize independent contrasts, which do not inflate the degrees of freedom of the test (Fig. 2). We standardized the contrasts by dividing them by the square-root of the corrected branch lengths, following Felsenstein (1985), and then mulitplying by the square-root of two times the total tree depth. This is equivalent to standardizing all contrasts so that they have an expected variance equivalent to a contrast taken through the root of the tree, but contrasts can be standardized in any way so long as they have the same expected variances after correction. This

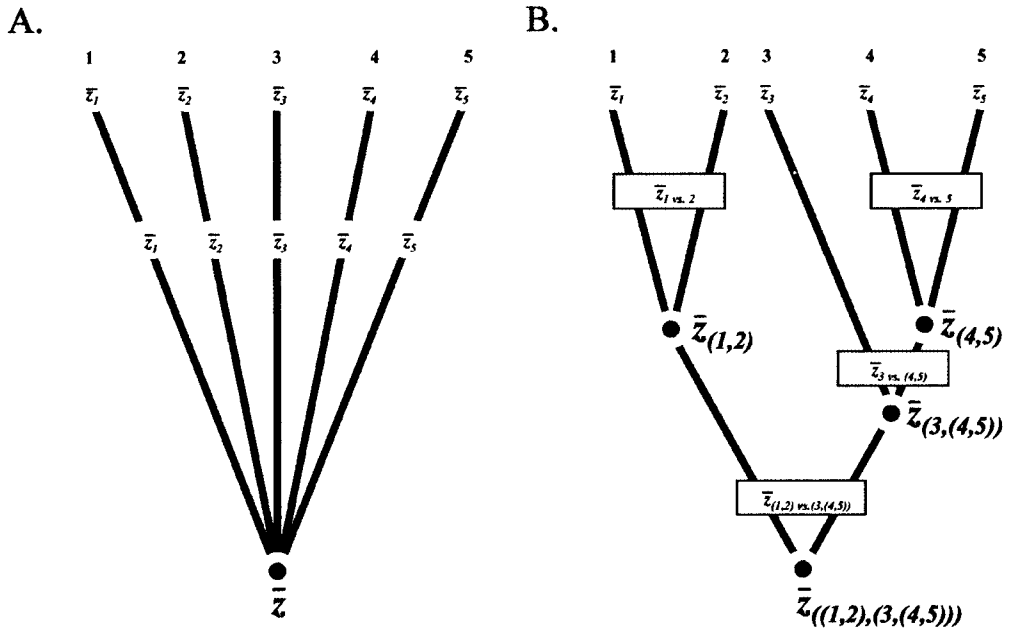


Fig. 2. An illustration of the two methods used to measure the effects of constraint on phenotypic differentiation. In (A), **D**, the variance–covariance matrix of population means for traits, is calculated ignoring the non-independence of population means due to common ancestry. If the true phylogeny is a star as pictured, and the ancestral state of all descendants is assumed to be their multivariate arithmetic means, then **D** is exactly equivalent to the mean squares and cross-products matrix of the vectors of change in mean phenotype, $\Delta\bar{\mathbf{z}}$. In (B), the mean squares and cross-products matrix of the vectors of change in mean phenotype are estimated directly from the independent contrasts of the mean phenotype vectors. This provides $\mathbf{D}_{IC}$, the variance–covariance matrix (or, more accurately, the mean squares and cross-products matrix) of independent contrasts. Here, $\bar{\mathbf{z}}$ at a node followed by a subscripted nested set of all the descendants of that node indicates the vector of reconstructed character states at that node. Furthermore, the notation $\Delta\bar{\mathbf{z}}_{i\ vs.\ j}$ indicates the independent contrast between nodes *i* and *j*. Although not illustrated here, in our analyses all contrasts for a given trait were corrected following Felsenstein (1985) to account for error in the estimation of ancestral character states, as well as to have equivalent expected variances.

standardization procedure is exactly as described by Felsenstein (1985), but contrasts are not standardized to have unit variance.

This procedure is similar to that used by Baker and Wilkinson (2003), but we substitute the MS-MCP matrix of independent contrasts ($\mathbf{D}_{IC}$) for the correlation matrix of independent contrasts. This substitution is preferred because Lande's (1979) equations predict alignment of within- and among-population variance–covariance matrices under genetic drift – a more specific prediction than the alignment of correlation matrices. Here and elsewhere, alignment is measured by several methods including matrix correlation and eigenanalysis.

We conducted matrix comparisons using the Mantel test on correlation matrices and random skewers on covariance and MS-MCP matrices (described above). The first three eigenvalues account for the vast majority of variance in both the $\bar{\mathbf{P}}$ and **D** matrices. To quantitatively compare the eigenstructure of $\bar{\mathbf{P}}$ with that of **D** and $\mathbf{D}_{IC}$, we calculated the

first three eigenvectors of all three matrices and calculated all pairwise vector correlation coefficients between the eigenvectors of $\bar{\mathbf{P}}$ and $\mathbf{D}$ and between the eigenvectors of $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$.

Alignment of $\mathbf{D}$ or $\mathbf{D}_{IC}$ with $\bar{\mathbf{P}}$ may be the result of biological processes, such as genetic constraint on phenotypic divergence (Lande, 1979; Arnold *et al.*, 2001). However, it also might result as an artifact of the sampling procedure if traits with high variance have means subject to large sampling error. To assess the extent to which sampling error has contributed to correspondence between $\bar{\mathbf{P}}$ and $\mathbf{D}$, we conducted a bootstrap randomization procedure in which population samples were recreated by re-sampling observations from all populations with replacement. Since sample sizes of the seven populations were extremely uneven, we performed bootstrap randomization by first picking a random population and then a random individual from within that population. We used this randomization procedure because re-sampling generates a set of pseudo-values for population means in which each individual observation has equivalent probability of having been drawn from any sampled population. Thus it assumes no real differences among populations, which is the desired null distribution for $\mathbf{D}$ and $\mathbf{D}_{IC}$. We evaluated the significance of the correspondence between $\bar{\mathbf{P}}$ and $\mathbf{D}$ or $\mathbf{D}_{IC}$ by comparing the observed matrix covariance to the distribution of matrix covariances generated by the randomization procedure. The covariance is preferred as a test statistic over the correlation coefficient because the correlation might be high between $\bar{\mathbf{P}}$ and $\mathbf{D}$ or $\mathbf{D}_{IC}$ even if the variances and covariances for traits among populations are negligible, so long as they are structured similarly to the pattern of variances and covariances within populations.

With the exception of CPCA, for which CPC (Phillips, 1998) was used, all tests were implemented using computer programs available from the authors.

RESULTS

P matrix estimation and comparison

Phenotypic variance–covariance matrices from the two focal populations were highly similar. Standard matrix comparison methods (Mantel test, random skewers, CPCA) showed little evidence for matrix dissimilarity, although CPCA found that common principal components (in which all eigenvectors were shared, but eigenvalues were not related by a common proportionality constant) was the preferred hypothesis over matrix proportionality and equality (Table 1). The same CPCA model was found whether the likelihood ratio test or the AIC were used for model selection. The *T* methods yielded similar results, revealing no significant differences in $\mathbf{P}$ between the two populations (Table 1).

With the effect of overall size removed by linear regression, phenotypic correlations in the pooled phenotypic variance–covariance matrix were highest among long bone limb elements (Table 2). They were also moderate to high between head dimensions, and between digital bones and long bones. Although some correlations were very low, for example between girdle widths and some long bone lengths, none were negative (Table 2).

Phylogeny estimation

The estimated phylogeny of the sampled populations is shown in Fig. 1B. The relationships among populations are largely concordant with geographic proximity, although the island population of VIE is more closely related to the south coast population (SAL) and interior

Table 1. Tests of matrix similarity between phenotypic variance–covariance matrices estimated for two populations, Puerto Yabucoa (YAB) and Cayo de Tierra, Vieques (VIE), of *Anolis cristatellus*

Test	Null hypothesis	Test statistic	Significance	Reference
Mantel test on correlation matrices	Matrices uncorrelated	Matrix correlation: $r_M = 0.93$	$P < 0.001$	Mantel (1967)
Random skewers on covariance matrices	No shared structure	Mean vector correlation of response vectors: $r_s = 0.93$	$P < 0.001$	Cheverud *et al.* (1983), Cheverud (1996)
Common principal components analysis	Matrices equal [actual test implemented: CPC (null) vs. CPC(11)]	$\chi^2 = 0.001$, AIC(CPC) = 142.37	$P = 0.98$	Flury (1988), Steppan (1997), Phillips and Arnold (1999)
T method, T^2 method	Matrices equal	$T_{obs} = 0.019$, $T^2_{obs} = 1.9 \times 10^{-5}$	$P_T = 0.18$, $P_{T^2} = 0.51$	Roff *et al.* (1999)

Note: All randomization tests used 10,000 randomizations. Random skewers used 10,000 random selection vectors. In CPCA, the selected model is CPC due to the insignificant likelihood ratio test in the test of CPC vs. CPC(11) and due to the minimum AIC for that model.

montane population (MAR), than it is to its closer, east coast neighbours. Nonetheless, the divergence of VIE is considerable (>15% corrected sequence divergence from its closest relative). Substantial sequence divergence among even the most closely related populations (YAB and FAJ; ~3.3% corrected divergence) makes it likely that the populations under study have been evolving more or less independently from one another for a substantial period of time.

Comparison of P and D

Populations differed significantly in body shape (MANOVA: $F_{78,1869.7} = 4.27$, Wilks' $\lambda = 0.404$, $P < 0.0001$). All univariate ANOVAs were also significant with the exception of jaw width and pectoral girdle width.

The variance–covariance matrix of population means, uncorrected for phylogeny (**D**), and the pooled variance–covariance matrix ($\bar{\mathbf{P}}$), were significantly correlated (mean random skewers $r_s = 0.71$, $P = 0.0026$; Mantel test on standardized matrices $r_M = 0.59$, $P = 0.0008$). **D** and the mean squares and mean cross-products matrix of independent contrasts, $\mathbf{D}_{IC}$, were quite highly correlated ($r = 0.92$). As such, when $\mathbf{D}_{IC}$ is substituted for **D** in the analysis the results are fairly similar. Both the mean vector correlation from random skewers and the Mantel matrix correlation are smaller than those calculated in the non-phylogenetic analysis, although in the case of random skewers the difference is very slight (random skewers $r_s = 0.70$, $P = 0.0031$; Mantel test $r_M = 0.34$, $P = 0.0279$).

Table 2. Pooled phenotypic variances, correlations (above the diagonal), and covariances (below the diagonal), with standard errors in parentheses

Trait	variance	jaw length	jaw width	3rd digit	radius	ulna	humerus	pectoral	pelvic	femur	tibia	fibula	4th toe-2	4th toe-1
jaw length	1.10 (0.08)	—	0.60 (0.04)	0.21 (0.06)	0.48 (0.04)	0.43 (0.05)	0.35 (0.05)	0.23 (0.05)	0.36 (0.04)	0.39 (0.05)	0.37 (0.05)	0.39 (0.05)	0.22 (0.08)	0.22 (0.06)
jaw width	1.71 (0.13)	0.82 (0.08)	—	0.24 (0.04)	0.31 (0.05)	0.33 (0.05)	0.34 (0.05)	0.32 (0.05)	0.38 (0.05)	0.35 (0.05)	0.29 (0.05)	0.29 (0.05)	0.19 (0.10)	0.24 (0.05)
3rd digit	5.18 (1.66)	0.49 (0.12)	0.72 (0.18)	—	0.49 (0.06)	0.53 (0.09)	0.46 (0.10)	0.10 (0.05)	0.23 (0.07)	0.38 (0.11)	0.43 (0.12)	0.42 (0.12)	0.40 (0.12)	0.53 (0.09)
radius	1.83 (0.16)	0.68 (0.09)	0.54 (0.10)	1.51 (0.24)	—	0.91 (0.01)	0.67 (0.04)	0.07 (0.05)	0.30 (0.05)	0.63 (0.04)	0.68 (0.03)	0.72 (0.03)	0.39 (0.10)	0.54 (0.05)
ulna	1.50 (0.15)	0.56 (0.08)	0.53 (0.09)	1.49 (0.20)	1.51 (0.15)	—	0.76 (0.03)	0.12 (0.05)	0.34 (0.05)	0.69 (0.04)	0.75 (0.03)	0.78 (0.03)	0.44 (0.12)	0.63 (0.05)
humerus	1.33 (0.16)	0.43 (0.08)	0.51 (0.09)	1.22 (0.19)	1.05 (0.15)	1.07 (0.15)	—	0.21 (0.05)	0.38 (0.05)	0.83 (0.02)	0.77 (0.03)	0.76 (0.03)	0.47 (0.11)	0.66 (0.05)
pectoral	3.36 (0.24)	0.44 (0.10)	0.77 (0.14)	0.42 (0.20)	0.18 (0.13)	0.27 (0.12)	0.44 (0.11)	—	0.35 (0.04)	0.21 (0.05)	0.19 (0.05)	0.20 (0.05)	0.08 (0.08)	0.16 (0.05)
pelvic	2.60 (0.19)	0.61 (0.09)	0.80 (0.12)	0.86 (0.18)	0.66 (0.13)	0.67 (0.12)	0.71 (0.13)	1.04 (0.15)	—	0.32 (0.06)	0.30 (0.06)	0.31 (0.06)	0.27 (0.05)	0.30 (0.05)
femur	1.31 (0.13)	0.47 (0.07)	0.53 (0.09)	0.99 (0.17)	0.98 (0.14)	0.96 (0.13)	1.10 (0.14)	0.45 (0.10)	0.59 (0.13)	—	0.83 (0.02)	0.82 (0.02)	0.50 (0.12)	0.68 (0.04)
tibia	1.45 (0.15)	0.46 (0.08)	0.46 (0.09)	1.17 (0.17)	1.11 (0.14)	1.11 (0.13)	1.07 (0.14)	0.42 (0.12)	0.59 (0.13)	1.15 (0.13)	—	0.97 (0.00)	0.56 (0.12)	0.77 (0.03)
fibula	1.50 (0.15)	0.50 (0.08)	0.46 (0.10)	1.18 (0.18)	1.19 (0.14)	1.16 (0.14)	1.08 (0.14)	0.45 (0.12)	0.60 (0.13)	1.15 (0.13)	1.43 (0.14)	—	0.53 (0.12)	0.75 (0.03)
4th toe-2	4.11 (1.37)	0.47 (0.11)	0.49 (0.17)	1.85 (0.26)	1.06 (0.20)	1.10 (0.19)	1.10 (0.20)	0.30 (0.22)	0.87 (0.20)	1.15 (0.18)	1.37 (0.18)	1.32 (0.18)	—	0.60 (0.14)
4th toe-1	2.06 (0.24)	0.33 (0.10)	0.45 (0.11)	1.74 (0.25)	1.05 (0.18)	1.11 (0.18)	1.10 (0.19)	0.42 (0.13)	0.69 (0.16)	1.12 (0.16)	1.33 (0.17)	1.32 (0.17)	1.75 (0.25)	—

Note: Traits not self-explanatory in their names include: the first metacarpal length of the third digit on the forefoot (3rd digit); the pectoral girdle width (pectoral); the pelvic girdle width (pelvic); the second metacarpal length of the fourth digit on the hindfoot (4th toe-2); and the first metacarpal length of the fourth digit on the hindfoot (4th toe-1). The data were ln-transformed and size-corrected before analysis. Pooling was performed across seven populations with sample sizes $n_{YAB} = 156$, $n_{VIE} = 141$, $n_{FAJ} = 12$, $n_{SAL} = 12$, $n_{MAR} = 11$, $n_{SJU} = 12$, $n_{RIN} = 13$ (see Fig. 1 and text for translation of population subscripts). Standard errors were calculated using a delete-one jackknife approach. Measurements were made in millimetres and variances and covariances (with their standard errors) are shown $\times 10^3$.

Eigenstructure comparison showed that the primary eigenvectors of $\bar{\mathbf{P}}$ and $\mathbf{D}$ are correlated (vector-correlation $r = 0.75$), as are the primary eigenvectors of $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$ ($r = 0.82$). The second and third eigenvectors of $\bar{\mathbf{P}}$ and $\mathbf{D}$ are also correlated with each other ($r = 0.44$ and $r = 0.43$, respectively), whereas between $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$ the order of the most highly aligned eigenvectors reverses between eigenvectors two and three. Here, the third eigenvector of $\bar{\mathbf{P}}$ is most highly correlated with the second eigenvector of $\mathbf{D}_{IC}$ ($r = 0.59$) and vice versa ($r = 0.41$).

Although a general correspondence between $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$ is revealed by these analyses (Fig. 3B), the strength of the pattern varies considerably among pairs of traits (Fig. 3A). For example, between pelvic width and jaw length, the major axis of among-population phenotypic divergence is nearly orthogonal to the major axis of the within-population covariance for the same traits, whereas between ulna length and tibia length, covariances within and among populations are tightly aligned (Fig. 3A).

In the test for whether correlation of within- and among-population variance–covariance matrices resulted from sampling error, both between $\bar{\mathbf{P}}$ and $\mathbf{D}$ and between $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$ the observed covariance was significantly greater than expected by sampling error in the estimation of $\mathbf{D}$ or $\mathbf{D}_{IC}$ alone [$\text{cov}(\bar{\mathbf{P}},\mathbf{D}) = 1.21 \times 10^{-5}$, $P = 0.034$; $\text{cov}(\bar{\mathbf{P}},\mathbf{D}_{IC}) = 1.53 \times 10^{-4}$, $P = 0.014$; Fig. 4].

DISCUSSION

Matrix comparisons

In this study, we found that phenotypic covariance structure was conserved between distantly related populations of *A. cristatellus*. Various matrix comparison methods yielded largely concordant results (see Table 1). A notable exception involved the tests for matrix equality and proportionality: common principal components analysis suggested that matrices were neither identical nor proportional (although they shared all principal components in common), whereas permutation tests [the *T* methods of Roff *et al.* (1999)] could not reject the hypothesis of matrix equality (see Table 1). Certainly, any real differences between the matrices were slight, as evidenced by their high correlation (Table 1).

Conservation of covariance structure over substantial evolutionary time scales is quite intriguing. Theory makes no general prediction about the stability of genetic variances and covariances (Turelli, 1988; Barton and Turelli, 1989; Shaw *et al.*, 1995). Genetic variances and covariances are expected to be stable under some circumstances and unstable under others (Turelli, 1988). As a consequence, stability of covariance structure becomes fundamentally an empirical question (Turelli, 1988; Camara *et al.*, 2000; Jones *et al.*, 2003).

Simulation studies have shown that genetic covariance structure stability is enhanced by a consistent pattern of correlated stabilizing natural selection (Jones *et al.*, 2003; L.J. Revell, submitted). Thus, assuming quadratic selection is important in the production and maintenance of genetic covariances, stable covariance structure could suggest that the pattern of quadratic natural selection is fairly constant between evolutionarily divergent and phenotypically differentiated populations. Alternative explanations for conserved covariance structure include consistent pleiotropic mutation (Lande, 1980; Jones *et al.*, 2003), strong developmental constraints (Arnold, 1992), and large effective population size (Jones *et al.*, 2003). In *A. cristatellus*

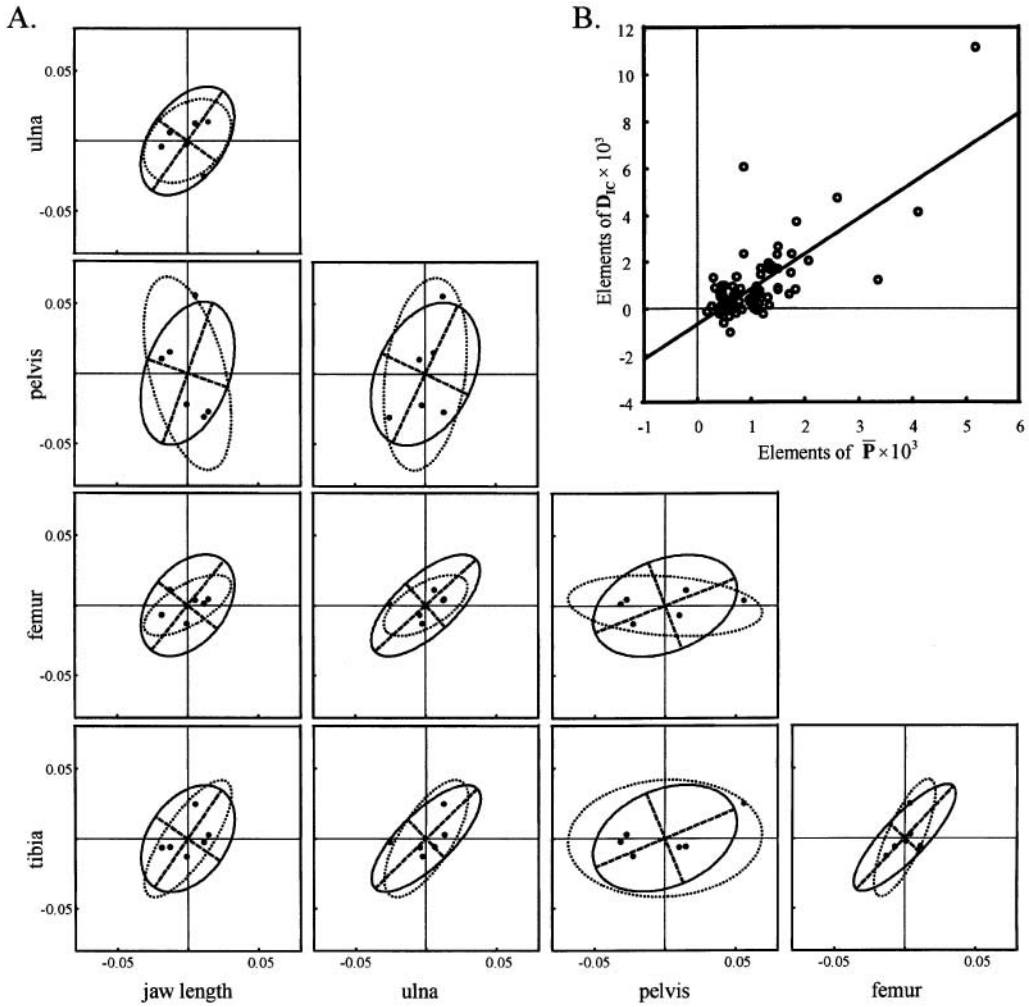


Fig. 3. (A) In a subsample of 4 of 13 traits, the pattern of trait covariation within populations (solid ellipses) sometimes aligns with the pattern of independent contrast covariation (dashed ellipses), as between tibia length and jaw length or between tibia length and ulna length. However, the within- and among-population patterns of covariation are sometimes quite different, as between pelvic girdle width and jaw length or between femur length and pelvic girdle width. In this figure, the ellipses have a length and width equivalent to the standard deviation in independent directions of the co-distribution of trait values and independent contrasts for solid and dashed ellipses respectively. Points are absolute values of independent contrasts. Ellipses and points are re-centred on the origin. (B) When the entire 13 trait phenotypic variance–covariance ($\bar{\mathbf{P}}$) and independent contrast mean squares and mean cross-products ($\mathbf{D}_{IC}$) matrices are evaluated, their alignment is highly significant (random skewers mean response vector correlation, $r_s = 0.70$, $P = 0.0031$; Mantel test on standardized matrices, $r_M = 0.34$, $P = 0.0279$).

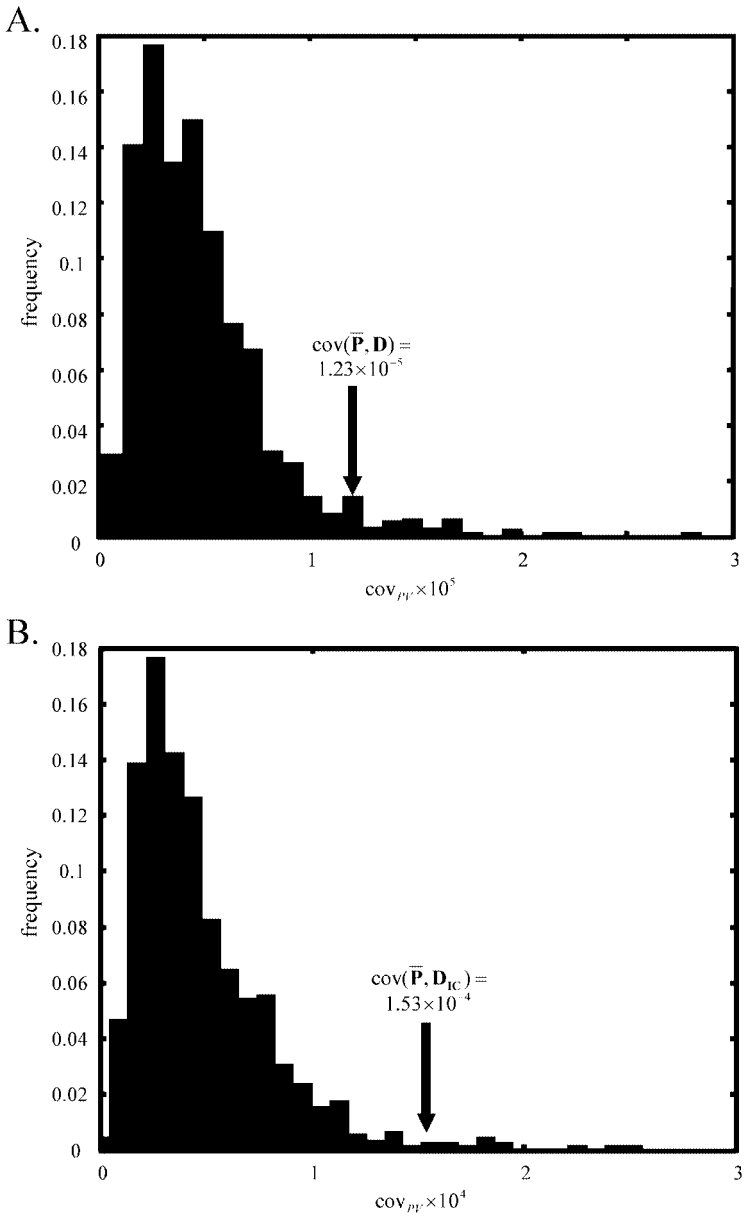


Fig. 4. Randomization distribution and test-statistic value in the bootstrap randomization test for significance of the alignment between (A) $\bar{\mathbf{P}}$ and $\mathbf{D}$ and (B) $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$. Bootstrap phenotypic means data were generated by randomly re-sampling populations and individuals to produce 1000 data sets containing population means for traits. By comparing the distribution of pseudo-values of the covariance between $\bar{\mathbf{P}}$ and $\mathbf{D}$ or $\mathbf{D}_{IC}$ (cov_{PV}) to the observed value of that covariance, this test evaluates the significance of the correspondence between $\bar{\mathbf{P}}$ and $\mathbf{D}$ or $\mathbf{D}_{IC}$ over that expected on the basis of higher sampling error for traits with high variance alone. The result [$\text{cov}(\bar{\mathbf{P}}, \mathbf{D}) = 1.23 \times 10^{-5}$, $P = 0.034$; $\text{cov}(\bar{\mathbf{P}}, \mathbf{D}_{IC}) = 1.53 \times 10^{-4}$, $P = 0.014$] indicates that the correspondence both between $\bar{\mathbf{P}}$ and $\mathbf{D}$ and between $\bar{\mathbf{P}}$ and $\mathbf{D}_{IC}$ significantly exceeds that expected due to sampling error in the estimation of $\mathbf{D}$ and $\mathbf{D}_{IC}$ alone.

census population densities in excess of 500 animals per hectare have been measured (Genet, 2002), however population structure is high (R.E. Glor *et al.*, in preparation) and genetic effective population sizes for local populations are unknown.

Due to the small number of **P** matrices estimated in this study (2), we are uncertain how general **P** matrix stability is in this species.

Testing the constraint hypothesis

In this study, we found that the pattern of variation and covariation within populations resembles the pattern of phenotypic divergence between them. This result was found whether phenotypic differentiation among populations was measured by calculating the variance–covariance matrix, **D**, of population means for traits in the standard way, or if divergence was measured by calculating the mean squares and mean cross-products matrix of independent contrasts, $\mathbf{D}_{\mathrm{IC}}$. This suggests that genetic constraint has influenced the pattern of intraspecific differentiation in *A. cristatellus*.

To our knowledge, this study is the first to compare the MS-MCP matrix of independent contrasts with **P** or **G** as a test for genetic constraint, although Baker and Wilkinson (2003) performed a similar analysis but used the correlation matrix of independent contrasts. We find that the result of our comparison of $\bar{\mathbf{P}}$ and $\mathbf{D}_{\mathrm{IC}}$ is relatively less correspondence than between $\bar{\mathbf{P}}$ and **D**.

In this study, **D** and $\mathbf{D}_{\mathrm{IC}}$ are quite highly correlated. Circumstances where **D** and $\mathbf{D}_{\mathrm{IC}}$ are poorly correlated may be rare in nature. One scenario in which **D** and $\mathbf{D}_{\mathrm{IC}}$ may be poorly correlated is when the deepest split in a balanced phylogeny also involved chance differentiation on a minor axis of variance in **G**. Under these circumstances, **G** or **P** may be poorly correlated with **D** but well correlated with $\mathbf{D}_{\mathrm{IC}}$. In this study, the opposite pattern is seen: $\bar{\mathbf{P}}$ is more strongly correlated with **D** than with $\mathbf{D}_{\mathrm{IC}}$. This pattern highlights another difficulty with using a non-phylogenetic differentiation matrix, which is that so doing artificially inflates the number of observations in the study, and thus statistical power, by utilizing sets of non-independent observations.

Bégin and Roff (2004) describe, but do not apply, a similar test to that used in this study. However, they find no phylogenetic signal in their data and thus use only the simpler **P** and **D** comparison. We agree that in practice if phylogenetic signal is low, **D** is very likely highly correlated with $\mathbf{D}_{\mathrm{IC}}$. However, under the assumption of Brownian motion evolution, the mean squares and cross-products of independent contrasts provide improved measures of the rate of character evolution and co-evolution than the rates estimated if phylogeny is ignored (Martins, 1994) and so we tentatively prefer the phylogenetic approach under circumstances in which a star phylogeny can be rejected as the history of the included taxa. Nonetheless, the general usefulness of $\mathbf{D}_{\mathrm{IC}}$ in tests of the constraint hypothesis should be the subject of further, more detailed study.

In our analyses, we use an ultrametric phylogenetic tree of included populations. An ultrametric phylogeny is one in which the evolutionary distance is the same from the root of the phylogeny to any tip. Nothing in the phylogenetic test of the constraint hypothesis applied in this study requires that only an ultrametric phylogeny be used. Non-ultrametric phylogenetic trees might be more appropriate when the sample includes an extinct taxon or if some prior hypothesis about variation in the rate of character evolution exists, for example correlated rates of molecular and morphological evolution (Omland, 1997). Another scenario appropriate to the latter case would be an instance in which there existed

independent information about the effective population sizes of the evolving lineages. In this situation, branch lengths should be scaled by the inverse of the harmonic mean of the effective population size for that branch. A cautionary consideration should be that factors causing bias in molecular phylogenetic branch-length estimates, such as model misspecification (Revell *et al.*, 2005), will create error in $\mathbf{D}_{IC}$ that would likely be further compounded by any additional branch-length manipulation.

In the test for constraint employed in this study, two types of error were ignored. Both the phenotypic variance–covariance matrix and the phylogenetic tree of the included populations are estimated with error, and excluding consideration of this error from subsequent analyses artificially inflates our confidence in the conclusions reached by those analyses. In the former case, ignoring error in the estimation of $\bar{\mathbf{P}}$ was justified on the basis of the fact that the standard errors of the elements of $\bar{\mathbf{P}}$ were small (Table 2). As **G** is usually estimated with larger error than **P** (Cheverud, 1988), future studies using **G** for comparable analyses might consider estimation error. In the latter case, ignoring error in tree estimation probably did not substantively influence the results of our analyses, as results were qualitatively similar even between phylogenetic and non-phylogenetic tests. Nonetheless, we cannot completely dismiss the possibility that phylogenetic error has biased the results from our test for constraint. Future studies might consider assessing the importance of phylogenetic error by randomly sampling pseudo-values of $\mathbf{D}_{IC}$ from the posterior distribution of phylogenies in a Bayesian analysis.

The methods used in this study make at least two important assumptions. The first assumption pertains to the substitution of **P** for **G** in evolutionary analyses. $\mathbf{P} = \mathbf{G} + \mathbf{E}$, where **E** is the environmental variance–covariance matrix. In testing the constraint hypothesis, we looked for a correlation between $\bar{\mathbf{P}}$ and **D** or $\mathbf{D}_{IC}$ where $\bar{\mathbf{P}}$ is used as a substitute for the **G** matrix. The correlation between **P** and **D** in terms of **G** and **E** is equivalent to: $\text{corr}(\mathbf{G} + \mathbf{E}, \mathbf{D}) = [\text{cov}(\mathbf{G}, \mathbf{D}) + \text{cov}(\mathbf{E}, \mathbf{D})]/\sqrt{[\text{var}(\mathbf{G}) + \text{var}(\mathbf{E}) + 2\text{cov}(\mathbf{G}, \mathbf{E})]\ [\text{var}(\mathbf{D})]}$. Since **E** only appears once in the numerator, specifically in the term $\text{cov}(\mathbf{E}, \mathbf{D})$, so long as the covariance between the environmental variance–covariance matrix and **D** is zero or negative, substitution of **P** for **G** will tend to decrease the correlation of **P** and **D** relative to the correlation of **G** and **D** and thus increase type I, not type II, error in the test for constraint. Conversely, if **E** and **D** are positively correlated, this will normally tend to increase the correlation of **P** and **D** relative to the correlation of **G** and **D** and thus increase type II error. The authors can think of no reason to presuppose that **E** and **D** are correlated; however, we cannot dismiss this possibility with our data.

The second assumption applies to the use of independent contrasts for the calculation of $\mathbf{D}_{IC}$ in the test for constraint. The independent contrasts method assumes correlated Brownian motion as a model of the evolutionary process (Felsenstein, 1985). This model is appropriate for random genetic drift in correlated characters (Lande, 1979). Under genetic drift, the **G** matrix (for which **P** is used as a proxy in this study) and **D** or $\mathbf{D}_{IC}$ are expected to be correlated (Lande, 1979; Hansen and Martins, 1996; Arnold *et al.*, 2001). Since testing for a correlation between $\bar{\mathbf{P}}$ and **D** or $\mathbf{D}_{IC}$ (which we purport was produced by drift) is the primary focus of this study, we felt that Brownian motion was a suitable assumption. That said, two additional issues should be kept in mind. In particular, (1) evolution might be by Brownian motion not only if it occurs by drift, but also if the position of the optimum moves by Brownian motion. Under this scenario, for $\bar{\mathbf{P}}$ and **D** or $\mathbf{D}_{IC}$ to be correlated would require that the position of the optimum moves according to a Brownian motion process with a variance–covariance matrix proportional to $\bar{\mathbf{P}}$. The data collected in this study can

neither support nor refute this possibility, however the authors can think of no reason to presuppose that if the position of the optimum moves by Brownian motion that it would move according to a Brownian motion process with a variance–covariance matrix proportional to $\bar{\mathbf{P}}$. Furthermore, (2) if evolution does not occur by Brownian motion, an important assumption of our method is violated. Other models have been developed to describe the evolutionary process aside from Brownian motion (e.g. Butler and King, 2004). It is possible that these could be applied to testing the constraint hypothesis, although we feel that this is beyond the scope of our study.

Other authors have proposed to explain a lack of strong correlation between **P** or **G** and **D** as indicative of selection overriding constraint (e.g. Ackermann and Cheverud, 2002). Although we find only moderate, if highly significant, correspondence between $\bar{\mathbf{P}}$ and **D** or $\mathbf{D}_{\mathrm{IC}}$, given the small number of populations ($n = 7$) used in this study, we are hesitant to ascribe this pattern, in particular the absence of a high correlation between $\bar{\mathbf{P}}$ and $\mathbf{D}_{\mathrm{IC}}$, to natural selection. Rather, we suggest only that the significant correlation between $\bar{\mathbf{P}}$ and **D** and between $\bar{\mathbf{P}}$ and $\mathbf{D}_{\mathrm{IC}}$ indicates that constraint has played some role in phenotypic differentiation in this group. This is a shortcoming of our data both due to sample design and limitations imposed by the number of evolutionarily distinct lineages in our study species. Future studies might consider sampling more populations from a group with more evolutionarily distinct lineages.

Finally, an alternative explanation for the correspondence between within- and among-population variance–covariance matrices is that they are both shaped by the same underlying pattern of natural selection (Arnold *et al.*, 2001; Bégin and Roff, 2004). Only one empirical study to date has compared the **G** matrix to the matrix of quadratic selection, and in that case no correspondence was uncovered (Blows *et al.*, 2004). In the absence of such information, we tentatively prefer as more parsimonious a hypothesis of constraint to explain the correspondence of $\bar{\mathbf{P}}$ with **D** and $\mathbf{D}_{\mathrm{IC}}$; however, this conclusion is subject to modification when information about natural selection becomes available.

Conclusions

Overall, the results of this study indicate that the evolution of covariance structure is conservative, at least at the intraspecific level, in a neotropical anole. The significant alignment of $\bar{\mathbf{P}}$ and **D** or $\mathbf{D}_{\mathrm{IC}}$ suggests that constraint has played some role in the phenotypic differentiation of lizard populations in this species. This is the first test for the importance of genetic constraint in the adaptive radiation of anoles. Our finding of significant evidence for genetic constraints on phenotypic evolution is the first for a group in which adaptive explanations for phenotypic divergence are generally proffered (e.g. Losos, 1990a, 1990b; Larson and Losos, 1996; Glor *et al.*, 2003). Future studies might consider using a similar analytic context to investigate the role of constraint in the differentiation of *Anolis* ecomorphs.

Although the non-phylogenetic variance–covariance matrix of population means, **D**, and the MS-MCP matrix of phylogenetically independent contrasts, $\mathbf{D}_{\mathrm{IC}}$, were highly correlated, their respective correlations with $\bar{\mathbf{P}}$ were substantively different, with **D** being more tightly aligned with $\bar{\mathbf{P}}$. This indicates the possibility that ignoring phylogenetic history in a study of constraint can artificially strengthen the constraint hypothesis.

ACKNOWLEDGEMENTS

This work was supported in part by a grant from the National Science Foundation (DEB-0519777). We thank R. Glor for preliminary data, T. Hrbek for field support, J. Losos, J. Jaenike, and anonymous reviewers for helpful comments, and members of the Losos lab for much useful discussion.

REFERENCES

Ackermann, R.R. and Cheverud, J.M. 2000. Phenotypic covariance structure in tamarins (genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal components analysis. *Am. J. Phys. Anthropol.*, **111**: 489–501.

Ackermann, R.R. and Cheverud, J.M. 2002. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. *Am. J. Phys. Anthropol.*, **117**: 260–271.

Akaike, H. 1973. Information theory and an extension of the maximum-likelihood principle. In *Second International Symposium on Information Theory* (B.N. Petrov and F. Csaki, eds.), pp. 267–281. Budapest: Akademiai Kiado.

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

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

Badyaev, A.V. and Hill, G.E. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. *Evolution*, **54**: 1784–1794.

Baker, R.H. and Wilkinson, G.S. 2003. Phylogenetic analysis of correlation structure in stalk-eyed flies (*Diasemopsis*, Diopsidae). *Evolution*, **57**: 87–103.

Barton, N. and Partridge, L. 2000. Limits to natural selection. *Bioessays*, **22**: 1075–1084.

Barton, N.H. and Turelli, M. 1989. Evolutionary quantitative genetics: how little do we know? *Annu. Rev. Genet.*, **23**: 337–370.

Bégin, M. and Roff, D.A. 2003. The constancy of the **G** matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: a case study in crickets. *Evolution*, **57**: 1107–1120.

Bégin, M. and Roff, D.A. 2004. From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution*, **58**: 2287–2304.

Bégin, M., Debat, V. and Roff, D.A. 2004. The effect of temperature and wing morphology on quantitative genetic variation in the cricket *Gryllus firmus*, with an appendix examining the statistical properties of the Jackknife–MANOVA method of matrix comparison. *J. Evol. Biol.*, **17**: 1255–1267.

Björklund, M. and Merilä, J. 1993. Morphological differentiation in *Carduelis* finches: adaptive vs. constraint models. *J. Evol. Biol.*, **6**: 359–373.

Blows, M.W. and Higgie, M. 2003. Genetic constraints on the evolution of mate recognition under natural selection. *Am. Nat.*, **161**: 240–253.

Blows, M.W., Chenoweth, S.F. and 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.

Butler, M.A. and King, A.A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.*, **164**: 683–695.

Camara, M.D., Ancell, C.A. and Pigliucci, M. 2000. Induced mutations: a novel tool to study phenotypic integration and evolutionary constraints in *Arabidopsis thaliana. Evol. Ecol. Res.*, **2**: 1009–1029.

Cheverud, J.M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.*, **110**: 155–171.

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

Cheverud, J.M. 1996. Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *J. Evol. Biol.*, **9**: 5–42.

Cheverud, J.M., Rutledge, J.J. and Atchley, W.R. 1983. Quantitative genetics of development: genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution*, **37**: 895–905.

Dietz, E.J. 1983. Permutation tests for association between two distance matrices. *Syst. Zool.*, **32**: 21–26.

Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.*, **125**: 1–15.

Fornori, J., Valverde, P.L. and Núñez-Farfán, J. 2003. Quantitative genetics of plant tolerance and resistance against natural enemies of two natural populations of *Datura stramonium*. *Evol. Ecol. Res.*, **5**: 1049–1065.

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

Garland, T., Jr. 1992. Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Am. Nat.*, **140**: 509–519.

Garland, T., Jr., Harvey, P.H. and Ives, A.R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.*, **41**: 18–32.

Genet, K.S. 2002. Structural habitat and ecological overlap of the Puerto Rican lizards *Anolis cristatellus* and *A. cooki*, with comments on the long-term survival and conservation of *A. cooki*. *Carib. J. Sci.*, **38**: 272–278.

Glor, R.E., Kolbe, J.J., Powell, R., Larson, A. and Losos, J.B. 2003. Phylogenetic analysis of ecological and morphological diversification in Hispaniolan trunk-ground anoles (*Anolis cybotes* group). *Evolution*, **57**: 2383–2397.

Hansen, T.F. and Martins, E.P. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution*, **50**: 1404–1417.

Jones, A.G., Arnold, S.J. and Bürger, R. 2003. Stability of the **G**-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution*, **57**: 1747–1760.

Kluge, A.G. and Kerfoot, W.C. 1973. The predictability and regularity of character divergence. *Am. Nat.*, **107**: 426–442.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. *Evolution*, **33**: 402–416.

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*, **94**: 203–215.

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

Larson, A. and Losos, J.B. 1996. Phylogenetic systematics of adaptation. In *Adaptation* (M.R. Rose and G.V. Lauder, eds.), pp. 187–220. San Diego, CA: Academic Press.

Lofsvold, D. 1988. Quantitative genetics of morphological differentiation in *Peromyscus*. II. Analysis of selection and drift. *Evolution*, **42**: 54–67.

Losos, J.B. 1990a. Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards: an evolutionary analysis. *Ecol. Monogr.*, **60**: 369–388.

Losos, J.B. 1990b. The evolution of form and function: morphology and locomotor performance in West Indian *Anolis* lizards. *Evolution*, **44**: 1189–1203.

Losos, J.B., Jackman, T.R., Larson, A., de Queiroz, K. and Rodríguez-Schettino, L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, **279**: 2115–2118.

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

Manly, B.F.J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. New York: Chapman & Hall.

Manly, B.F.J. 2005. *Multivariate Statistical Methods: A Primer*. New York: Chapman & Hall.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.*, **27**: 209–220.

Marroig, G. and Cheverud, J.M. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. *Evolution*, **55**: 2576–2600.

Marroig, G., de Vivo, M. and Cheverud, J.M. 2004. Cranial evolution in sakis (*Pithecia*, Platyrrhini) II: Evolutionary processes and morphological integration. *J. Evol. Biol.*, **17**: 144–155.

Martins, E.P. 1994. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.*, **144**: 193–209.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B. *et al.* 1985. Developmental constraints and evolution. *Q. Rev. Biol.*, **60**: 265–287.

Merilä, J. and Björklund, M. 1999. Population divergence and morphometric integration in the greenfinch (*Carduelis chloris*): evolution against the trajectory of least resistance? *J. Evol. Biol.*, **12**: 103–112.

Omland, K.E. 1997. Correlated rates of molecular and morphological evolution. *Evolution*, **51**: 1381–1393.

Phillips, P.C. 1998. *CPC: Common Principal Ccomponents Analysis*. Eugene, OR: University of Oregon.

Phillips, P.C. and Arnold, S.J. 1999. Hierarchical comparison of genetic variance–covariance matrices. I. Using the Flury hierarchy. *Evolution*, **53**: 1506–1515.

Revell, L.J., Harmon, L.J. and Glor, R.E. 2005. Underparameterized model of sequence evolution leads to bias in the estimation of diversification rates from molecular phylogenies. *Syst. Biol.*, **54**: 973–983.

Reusch, T. and Blanckenhorn, W.U. 1998. Quantitative genetics of the dung fly *Sepsis cynipsea*: Cheverud's conjecture revisited. *Heredity*, **81**: 111–119.

Roff, D.A. 1995. The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity*, **74**: 481–490.

Roff, D.A., Mousseau, T.A. and Howard, D.J. 1999. Variation in the genetic architecture of calling song among populations of *Allonemobius socius, A. fasciatus*, and a hybrid population: drift or selection? *Evolution*, **53**: 216–224.

Rohlf, F.J. 2005. *tpsDig, Version 2.04*. Stony Brook, NY: Department of Ecology and Evolution, State University of New York at Stony Brook.

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

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

Shaw, F.H., Shaw, R.G., Wilkinson, G.S. and Turelli, M. 1995. Changes in genetic variances and covariances: **G** whiz! *Evolution*, **49**: 1260–1267.

Sokal, R.R. 1978. Population differentiation: something new or more of the same? In *Ecological Genetics: The Interface* (P.F. Brussard, ed.), pp. 215–239. New York: Springer.

Sokal, R.R. and Riska, B. 1981. Geographic variation in *Pemphigus populitransversus* (Insecta: Aphididae). *Biol. J. Linn. Soc.*, **15**: 201–233.

Steppan, S.J. 1997. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution*, **51**: 571–586.

Steppan, S.J., Phillips, P.C. and Houle, D. 2002. Comparative quantitative genetics: evolution of the **G** matrix. *Trends Ecol. Evol.*, **17**: 320–327.

Swofford, D.L. 2002. *PAUP* 4.0b10. Phylogenetic Analysis Using Parsimony* (** and Other Methods*), Version 4.0b10. Sunderland, MA: Sinauer Associates.

Turelli, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution*, **42**: 1342–1347.

Venable, D.L. and Burquez, M.A. 1990. Quantitative genetics of size, shape, life-history, and fruit characteristics of the seed heteromorphic composite *Heterosperma pinnatum*. II. Correlation structure. *Evolution*, **44**: 1748–1763.

Waitt, D.E. and Levin, D.A. 1998. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. *Heredity*, **80**: 310–319.

Willis, J.H., Coyne, J.A. and Kirkpatrick, M. 1991. Can one predict the evolution of quantitative characters without genetics? *Evolution*, **45**: 441–444.

Williams, E.E. 1972. The origin of faunas. Evolution of lizard congeners in a complex island fauna: a trial analysis. *Evol. Biol.*, **6**: 47–89.