



Comparative Quantitative Genetics of the Pelvis in Four-Species of Rodents and the Conservation of Genetic Covariance and Correlation Structure

Carl J. Saltzberg¹ · Laura I. Walker² · Lee E. Chipps-Walton¹ · Bárbara M. A. Costa^{1,3} · Ángel E. Spotorno² · Scott J. Steppan¹

Accepted: 19 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Quantitative genetics is a powerful tool for predicting phenotypic evolution on a microevolutionary scale. This predictive power primarily comes from the Lande equation ($\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$), a multivariate expansion of the breeder's equation, where phenotypic change ($\Delta\bar{\mathbf{z}}$) is predicted from the genetic covariances ($\mathbf{G}$) and selection ($\boldsymbol{\beta}$). Typically restricted to generational change, evolutionary biologists have proposed that quantitative genetics could bridge micro- and macroevolutionary patterns if predictions were expanded to longer timescales. While mathematically possible, making quantitative genetic predictions across generations or species is contentiously debated, principally in assuming long-term stability of the $\mathbf{G}$ -matrix. Here we tested stability at a macroevolutionary timescale by conducting full- and half-sib breeding programs in two species of sigmodontine rodents from South America, the leaf-eared mice *Phyllotis vaccarum* and *P. darwini* and estimated the $\mathbf{G}$ -matrices for eight pelvic traits. To expand our phylogenetic breadth, we incorporated two additional $\mathbf{G}$ -matrices measured for the same traits from Kohn & Atchley's 1988 study of the murine rodents *Mus musculus* and *Rattus norvegicus*. Using a phylogenetic comparative framework and four separate metrics of matrix divergence or similarity, we found no significant association between evolutionary divergence among species $\mathbf{G}$ -matrices and time, supporting the assumption of stability for at least some structures. However, the phylogenetic sample size is necessarily small. We suggest that small fluctuations in covariance structure can occur rapidly, but underlying developmental regulation prevents significant divergence at macroevolutionary scales, analogous to an Ornstein–Uhlenbeck pattern. Expanded taxonomic sampling will be needed to test this suggestion.

Keywords Comparative biology · Covariance · Mammals · Phylogenetic constraint · Quantitative genetics · Rodents

Introduction

Understanding the factors that bias the phenotypic differences among species is a long-standing question in evolutionary biology. One promising area of research has come from quantitative genetic predictions of how multiple traits

will evolve at the microevolutionary level between generations (de Oliveira et al., 2009; Hansen & Houle, 2008; Steppan et al., 2002). This predictive power largely comes from the Lande equation ($\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$) which is a multivariate expansion of the univariate breeder's equation (Arnold et al., 2008; Lande, 1979). In this, $\Delta\bar{\mathbf{z}}$ is the predicted evolutionary response, a vector of change in each trait from the population mean; $\boldsymbol{\beta}$ is the selection gradient, a vector of the selection on each trait; and $\mathbf{G}$ is the additive genetic matrix which describes the variances and covariances, or correlations among traits, hereafter referred to as the $\mathbf{G}$ -matrix. While the primary use of the $\mathbf{G}$ -matrix in the Lande equation is to estimate selection or the change in a population across generations within a population, these components could also be used to explore the causes of macroevolutionary differences among species (Arnold et al., 2008; Lande, 1979; Rossoni et al., 2017). Although mathematically straightforward,

✉ Carl J. Saltzberg
csaltzberg@bio.fsu.edu

¹ Department of Biological Science, Florida State University, Tallahassee, FL 32306-4295, USA

² Laboratorio de Citogenética Evolutiva, Programa de Genética Humana, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, 70061 Santiago, Chile

³ Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, São Paulo, Brazil

the application of the Lande equation at this scale has been debated (Aguirre et al., 2014; Arnold et al., 2008; Hansen & Houle, 2008; Mezey & Houle, 2003; Pigliucci, 2006; Revell, 2007; Stepan et al., 2002; Turelli, 1988). A critical component to this debate rests on the assumption that the **G**-matrix is stable, as predicting the selection or change across multigenerational time scales such as those separating species would require a near-constant **G**-matrix (Porto et al., 2016). Evidence for the stability of the **G**-matrix is limited, involving either empirical observations across generations, or comparative analysis across species (Matta & Bitner-Mathé, 2004; Roff & Fairbairn, 2012; Stepan et al., 2002). It is worth noting that the few empirical studies of the stability of the **G**-matrix have often been unable to sufficiently address the pattern at a comparative scale, either lacking sufficient taxonomic sampling using two or three closely related species (Arnold et al., 2008); restricting the taxonomic scale to populations within a single species (Aguirre et al., 2014); using an indirect approach to evaluate if the phenotypic disparity across species match the macroevolutionary predictions of a single **G**-matrix (Baker & Wilkinson, 2003; Porto et al., 2009), or through substitution of the **G**-matrix entirely by the more accessible phenotypic covariances of the **P**-matrix (Agrawal & Stinchcombe, 2009). These compromises arise from the difficulty in estimating even a single **G**-matrix, often involving the maintenance of large breeding populations in captivity, or intensive field sampling over multiple generations (Aguirre et al., 2014; Matta & Bitner-Mathé, 2004; Roff, 1995, 2012). This demand has resulted in very few estimates of **G**-matrices among vertebrates and especially mammals.

The long-term stability of the **G**-matrix remains a fundamentally unresolved question. In this paper we expand on a classic study by including **G**-matrices for additional species for comparison. Kohn and Atchley (1988) compared two species of rodents within the subfamily Murinae for eight pelvic traits recorded from lab-reared colonies of the house mouse, *Mus musculus* (Linnaeus 1758), and the brown rat, *Rattus norvegicus* (Berkenhout 1769). While Kohn and Atchley (1988) concluded that the genetic covariances influencing the pelvic morphology in both murines had remained reasonably constant, they emphasized caution in assuming stability, noting both conflicting results across metrics, and whether the covariance or correlation **G**-matrix was compared. While not explicitly phylogenetic in their approach, they did comment on the evolutionary nature of their comparison and noted it could be expanded with additional species. We saw in Kohn and Atchley (1988) a foundation for incorporating additional species into a larger comparative analysis, with the two-species comparison representing but a single phylogenetic contrast between murine taxa that diverged approximately 10 mya (Stepan & Schenk, 2017). By combining two new species with those from Kohn and Atchley (1988) we were able to

triple the phylogenetic information provided by either data set alone. The pelvis is an informative structure for this question because it is functionally important and morphologically complex enough to allow multiple divergent transformations while also being developmentally simpler than the more commonly studied skull (Cheverud, 1982, 1995). In this study however, we do not predict strong divergence due to functional differences because all the species considered are ambulatory generalists (Jones et al., 2003; Turelli, 1988).

Here we test how evolutionarily stable the **G**-matrix is for the pelvis across a broader region of Rodentia, increasing the total number of phylogenetic contrasts from one to three. To accomplish this, we conducted full- and half-sib breeding programs for two species of South American rodents within the subfamily Sigmodontinae: the Vacas leaf-eared mouse *Phyllotis vaccarum* (Thomas 1912; recently elevated from *P. xanthopygus vaccarum*, see Jayat et al., 2021), and Darwin's leaf-eared mouse *Phyllotis darwini* (Waterhouse 1837). The combination of both sigmodontines with the previous murines creates two additional contrasts, one across the split between subfamilies at 18.61 mya and a second between the *Phyllotis* species at 2.91 mya (Stepan & Schenk, 2017). Using both sigmodontine breeding colonies, we estimated the **G**-matrix for the same eight pelvic traits that Kohn and Atchley (1988) had. When combined with reconstructions of the ancestral **G**-matrices in both subfamilies we performed phylogenetically structured pairwise **G**-matrix comparisons among all four-species. This sample design represents one of the largest phylogenetic comparative analyses of the **G**-matrix in mammals.

Quantitative genetic research has seen a proliferation of metrics proposed for comparing **G**-matrices, but no clear consensus has emerged (Roff et al., 2012). Some metrics describe a specific biologically meaningful aspect of the **G**-matrix, others are more nuanced, and they also vary in their null expectations (Lofsvold, 1986; Porto et al., 2016; Stepan et al., 2002). No single metric can capture all the ways and degrees by which **G**-matrices can differ from each other (Cheverud & Marroig, 2007; Houle et al., 2002; Stepan, 2004), but in combination a more complete understanding can be realized. Therefore, we elected to use five metrics derived from the literature to evaluate the degree of evolutionary divergence among **G**-matrices: (1) Correlation, (2) Orientation, (3) Random Skewers, (4) Disparity, and (5) Conditional Evolvability.

Materials and Methods

Study Species and Breeding Design

The murine animals measured by Kohn and Atchley (1988) had been obtained from breeding colonies generated by

previous investigators. The specimens of *R. norvegicus* were reared by Park et al. (1966) using a cross-fostering design from lab lines that had been random bred for > 40 generations, measured on 189-day-old individuals who had been released from selection for weight gain as described in Atchley and Rutledge (1980), Cheverud et al. (1983b), and Rutledge and Chapman (1975). Specimens of *M. musculus* were reared by Riska et al. (1984) from established random bred ICR lines, also using a cross-fostering design, measured on 70-day-old individuals who had been examined for variation in growth rates, and further described by Cheverud et al. (1983a).

The sigmodontine species used in this study were obtained from wild-caught specimens, or their direct laboratory descendants. All captured animals were collected within Chile; *P. darwini* from the Coquimbo Region approximately 350 km north of Santiago, between 20 and 200 m in elevation; *P. vaccarum* from the Metropolitana Region approximately 50 km east of Santiago, between 1700 and 2400 m in elevation. The mice were transferred to the Laboratorio de Citogenética Evolutiva in Santiago, Chile, where they were paired and bred using a mixed full- and half sib design. All laboratory pairs were raised in a standard housing consisting of a plastic cage and a bed of shavings. Food pellets and water were provided ad libitum, and a natural daylight regime was followed. Cages were checked twice a week to record births, litter sizes, and deaths. Following parturition, the male of the pair was kept in the cage for another three to five days then removed from the family. All offspring were raised for 180-days and then sacrificed along with their wild-caught parents, after which they were fully skeletonized. Complete voucher specimens of both species will be housed in the Field Museum of Natural History in Chicago, USA. The sigmodontine breeding colonies produced 305 *P. vaccarum* individuals in 54 families and 348 *P. darwini* in 62 families. Detailed pairings for these two species, the number and sex of the resulting offspring are

provided in Online Resource 1. An overview of all four-species used in this study, sample sizes, and breeding design are shown in Table 1.

Measurement

For both *M. musculus* and *R. norvegicus* we used the species data reported in Kohn and Atchley (1988) instead of the individual specimens, because neither the individuals nor the raw data were available (L. Kohn, pers. comm., June 5th 2014). For all *Phyllotis*, eight pelvic traits were measured following the anatomical descriptions and procedures described by Kohn and Atchley (1988) shown in (Fig. 1), and described in Table 2. All trait measurements on both *Phyllotis* species were made on the right innominate bone of the pelvis, substituting with the left side only when damaged or absent. If a trait was damaged or absent on both sides, its measurement was skipped, and its value left blank. The total number of specimens within species recorded for each trait are shown in Table 3. To reduce error, all measurements were performed by a single researcher using digital calipers with an accuracy of ± 0.01 mm (Mitutoyo Caliper: 573-221-20, NTD12-6" CX) and transcribed directly to a computer spreadsheet using a USB interface cable (Mitutoyo Input Tool: 264-005, IT-005D).

Phylogeny

When performing any comparative analysis among species, it is important to correct for non-independence in the data that may have resulted from shared evolutionary history (Felsenstein, 1985). Comparative studies of **G**-matrices are no exception, because relationships among traits can be influenced by phylogeny and inherited through common ancestry (Steppan, 1997a). We pruned all but the branches connecting the four-species of rodents used in this study from a large dated molecular phylogeny of the superfamily

Table 1 Sample sizes and breeding design used to estimate the **G**-matrix in each of the four-species compared

Species	Ind.	Fam. ^a	Reared age	Design	Notes	References
<i>M. musculus</i>	707	108	70-day	Cross-foster	Used in a growth rate variation study; ICR starting strains	Cheverud et al. (1983a), Riska et al. (1984)
<i>R. norvegicus</i>	522	53	189-day	Cross-foster	Used in a selection on weight gain study; four inbred starting strains	Atchley and Rutledge (1980), Baker et al. (1975), Park et al. (1966)
<i>P. darwini</i>	348	62	180-day ^b	Full-and half-sib	Designed for this study; Random-bred offspring reared from wild-caught parents	This manuscript
<i>P. vaccarum</i>	305	54	180-day ^b	Full- and half-sib	Designed for this study; Random-bred offspring reared from wild-caught parents	This manuscript

^aTotal number of family pairings generated

^bSacrificed offspring age only, wild caught parents of unknown age

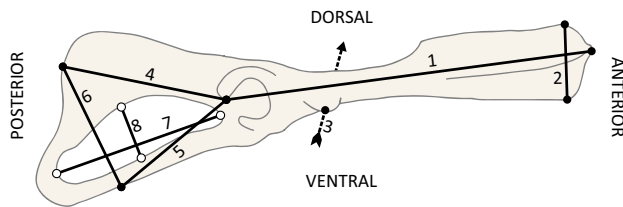


Fig. 1 Lateral view of the right innominate bone of the pelvis and visualization of the eight pelvic trait measurements (1–8) as described in Kohn and Atchley (1988). A detailed description of each linear measurement is given in Table 2. Points represent the approximate location to place the caliper jaws when taking measurements for each linear distance. Point fill indicates the method of measurement used: open=calipers expanded to the greatest interior distance; filled=calipers retracted to the greatest exterior distance. Line style indicates visibility of the measurement plane viewed in this projection: solid=no obstruction; dashed=obscured

Muroidea (Steppan & Schenk, 2017). The resulting time calibrated phylogeny after pruning (hereafter chronogram) had a total of four operational taxonomic units and three dated

ancestral nodes (Fig. 2). We regressed matrix comparisons on to these dates to correct for phylogenetic history. The age of the basal contrast was calculated as the summed branch lengths divided by two, to be comparable to the values of the other two contrasts.

G-matrix Estimation and Reconstruction

For both murine species the **G**-matrices were previously estimated using an ANOVA model in which the genetic and environmental variance components are contained in multiple levels of the nested ANOVA based on the degree of relationship shared between full-sibling individuals within, and half-siblings among families (Kohn & Atchley, 1988). In both sigmodontine species, the genetic covariances and correlations were estimated using an animal model and Restricted Maximum Likelihood (ReML) methods implemented in the software WOMBAT (Meyer, 2007). All eight pelvic traits were first natural log transformed and then multiplied by 1000 prior to the estimating to aid in model

Table 2 Trait names, abbreviations, and descriptions for eight pelvic traits according to those recorded by Kohn and Atchley (1988)

#	Trait	Name	Description
1	ILLE	Ilium length	Maximum ilium length, from the most superior aspect of the iliac crest to the inferior interior margin of the acetabulum
2	ILWD	Ilium width	Maximum width of the iliac crest
3	ILDI	Ilium diameter	Maximum diameter of the ilium at the inferior ventral spine
4	ISLE	Ischium length	Maximum length of the ischium, from the inferior interior margin of the acetabulum to the inferior aspect of the ischial tuberosity
5	PUBL	Pubis length	Maximum length of the superior pubic ramus, measured from the inferior interior margin of the acetabulum to the superior point of the pubic symphysis
6	ISPU	Ischium-pubis length	Maximum length between inferior aspect of the ischial tuberosity and the superior point of the pubic symphysis
7	OBLE	Obturator-foramen length	Maximum length of the Obturator foramen
8	OBWD	Obturator-foramen width	Maximum width of the Obturator foramen

See (Fig. 1) for a visual representation of each trait measurement

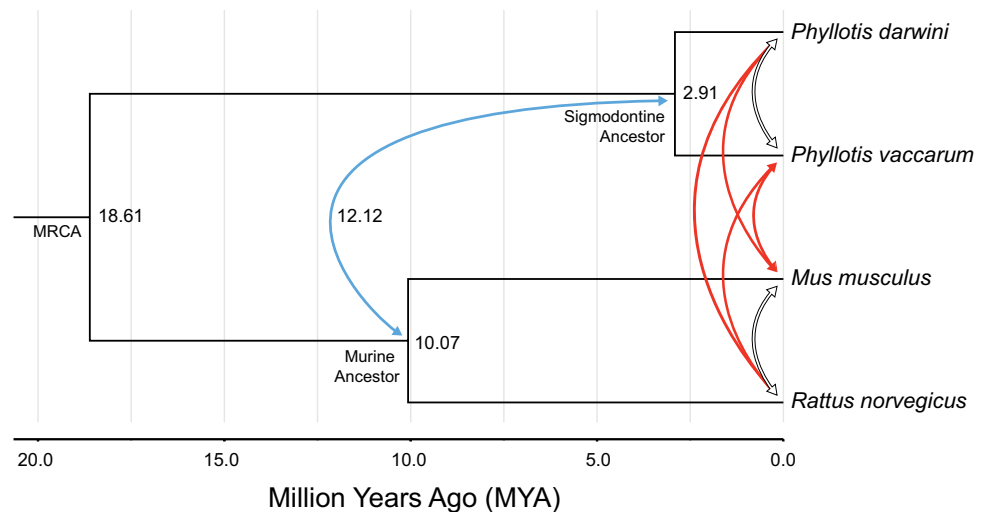
Table 3 The means (in millimeters) of all eight pelvic traits in the four-species of rodents, including standard deviation and sample size

#	Trait	<i>P. darwini</i>			<i>P. vaccarum</i>			<i>M. musculus</i> ^a			<i>R. norvegicus</i> ^a		
		$\bar{x}$	$\pm \sigma$	n	$\bar{x}$	$\pm \sigma$	n	$\bar{x}$	$\pm \sigma$	n	$\bar{x}$	$\pm \sigma$	n
1	ILLE	15.64	0.06	345	15.77	0.06	304	12.18	0.80	497	27.39	0.85	522
2	ILWD	4.19	0.13	351	3.46	0.13	305	3.19	1.37	497	6.82	1.87	522
3	ILDI	2.86	0.09	351	2.73	0.09	305	1.97	1.22	685	4.62	1.54	522
4	ISLE	9.1	0.08	351	8.6	0.09	300	5.75	1.06	685	14.15	1.39	522
5	PUBL	9.53	0.11	329	9.85	0.09	283	6.55	1.07	685	13.74	1.07	522
6	ISPU	8.82	0.09	329	9.21	0.07	282	7.24	0.81	685	16.44	1.06	522
7	OBLE	8.29	0.07	342	8.33	0.08	285	5.58	0.83	685	11.82	1.12	522
8	OBWD	3.6	0.09	350	3.54	0.08	296	2.25	1.60	685	5.53	1.32	522

Trait abbreviations and numbers are defined in Table 2

^aData from Kohn and Atchley (1988), with values shown in millimeters obtained by taking the antilog of the original log units published

Fig. 2 The four-species chronogram used in this study, pruned from the dated molecular phylogeny by Steppan and Schenk (2017). Branch lengths are scaled to millions of years ago (MYA), and the age of each ancestral node and basal contrast is displayed. Comparison pairs between **G**-matrices are shown as double ended arrow lines, and its inclusion within a set of comparisons is indicated by the color: All-nodes = blue, red, and white; Contrasts = blue and white; Tip-species = red and white (Color figure online)



convergence. The animal model is a mixed linear model that incorporates all the potential genetic variance contributions of similarity between individuals across a pedigree (Kruuk et al., 2000; Lynch & Walsh, 1998). As in Kohn and Atchley (1988), sex was treated as a fixed effect. Matrices were estimated for both full (eight dimensions) and reduced-rank models (seven to two dimensions), and the best-fitting rank model was selected on the basis of Akaike's Information Criterion (AICc).

When comparing **G**-matrices between species, Kohn and Atchley (1988) noted that rejecting the null of matrix equality depended on whether the covariance or correlation matrices were compared. To produce a comparable evaluation, we used both the covariance and correlation **G**-matrices for all comparisons among species. While we were able to directly estimate the covariance and correlation **G**-matrices in the sigmodontines, this was not possible for the murines because the original measurements were not available. Therefore, in each of the murines we "back calculated" the covariance **G**-matrix from the genetic variances and correlations presented in Kohn and Atchley (1988) using Eq. 2 found in Online Resource 2.

We estimated the ancestral covariances and correlations for each node of our chronogram (Fig. 2; branch lengths proportional to time) using a Brownian motion-based likelihood estimator (Schluter et al., 1997), treating each cell within the **G**-matrix as independent elements. This was accomplished by rearranging the upper triangle cells of each of the four **G**-matrices into a single 1×36 vector. Each element of the **G**-matrices was estimated independently throughout the chronogram, using the function "fastAnc()" from the "phytools" R package (Revell, 2012). After ancestral reconstruction, these vectors were then reorganized into square symmetric matrices, resulting in reconstructed ancestral covariance and correlation **G**-matrices for each node of the pruned tree. The eight pelvic traits means were reconstructed

with squared change parsimony in Mesquite ver. 3.1 (Maddison & Maddison, 2017).

Matrix Similarity and Divergence

We identified seven pairwise comparisons among "all-nodes" in the four-species chronogram that had independent branching histories, indicated by the arrows in (Fig. 2). From the seven "all-nodes" comparisons two subsets were made; among "tip-species" with six comparisons among all four-species, and among the three phylogenetic independent "contrasts" (Felsenstein, 1985). Using these sets we evaluated two hypotheses: (1) that the **G**-matrix is stable, by testing for similarity among "all-nodes" using matrix correlation; and (2) that the **G**-matrix divergence increases with time, by regressing metrics of matrix divergence among "tip-species" or phylogenetic "contrasts" (calculated as the absolute value of the difference between matrices) on the time since divergence (i.e. evolutionary distance). Time was calculated as the average of the branch lengths separating the tip-species or of the subtending branch lengths, respectively. All metrics used in these comparisons are detailed in the following methods and summarized in Table 4.

Metric 1 Correlation

Several methods of matrix correlation were used to test hypotheses both of a stable **G**-matrix using pairwise comparisons, and how the **G**-matrix evolved by regressing the magnitude of the matrix correlations against the time since divergence. While matrix correlations have been commonly used when testing the homogeneity between matrices, genetic or otherwise (Kohn & Atchley, 1988; Lofsvold, 1986) the suitability of comparing matrices this way has been criticized (Legendre & Fortin, 2010). However, for comparability, we elected to use the same three matrix correlation tests as those

Table 4 The four metrics used to compare evolutionary divergence between **G**-matrices

Metric	Description	References
Correlation	Matrix correlation between corresponding cell values using; Mantel ($\mathbf{r}_M$), Spearman's rank ($\mathbf{r}_s$), and Dietz statistic of association ($\mathbf{K}_c$)	Dietz (1983), Kohn and Atchley (1988)
Orientation	Deviation between evolutionary Lines of Least Resistance. Evaluated here using the angle of corresponding ranked orthogonal vectors of greatest genetic variation (i.e. the eigenvectors, of a singular value decomposition)	Schluter (1996)
Random skewers	Vector correlation between the response vectors produced from the same set of random selection vectors applied to each matrix, using the Lande equation	Cheverud (1996), (Cheverud & Marroig, 2007), Lande (1979)
Disparity	The mean of the absolute differences between the off-diagonal cells in each matrix	Steppan (1997b)
Conditional evolvability	Euclidean distance between matrices, after conversion to conditional evolvability matrices that describe the ability to respond to directional selection with stabilizing selection on all other trait combinations	Hansen and Houle (2008)

For each metric, a description and relevant methodological references are provided

used in the original comparison between *M. musculus* and *R. norvegicus* (Kohn & Atchley, 1988). These tests were the Pearson product-moment correlation coefficient ($\mathbf{r}_M$), Spearman's rank ($\mathbf{r}_s$), and Dietz statistic of association ($\mathbf{K}_c$). Each of the tests evaluate the degree to which the corresponding elements of two matrices share an overall similar pattern of magnitude. All matrix correlations between species were performed with R (R Core Team, 2017), using either the function “mantel()” in the package “vegan” (Oksanen et al., 2013) by setting the method of comparison to “pearson” and “spearman” respectively, or the function “rowwiseKrtest()” in the package “DyaDA” (Leiva et al., 2010). Statistical significance in each test was evaluated using matrix permutation set to the maximum number possible ($\mathbf{r}_M$ and $\mathbf{r}_s = 5040$, $\mathbf{K}_c = 40,319$). In each test, the null hypothesis is no similarity in structure.

Metric 2 Orientation

Second, we compared divergence in the magnitude and orientation of genetic associations between **G**-matrix lines of least resistance (LLR), capturing information regarding the shape of matrices. The LLR are an important evolutionary consideration as they describe the primary direction of phenotypic change possible given the genetic variation present in a population (Schluter, 1996). The LLR were obtained for each **G**-matrix by determining the greatest axis of genetic variation using an eigen decomposition performed in the program R using the base function “eigen()”. Following the eigen decomposition the 1st and 2nd eigenvectors were retained, which respectively represent LLR of the two largest uncorrelated orthogonal directions of genetic variation, hereafter referred to as **g-max** and **g-2nd**. Then the difference in orientation between **G**-matrices was quantified as the deviation in the angle (θ) for **g-max** and **g-2nd** calculated using the arc

cosine between corresponding vectors. Finally, we tested evolutionary divergence of **G**-matrices in the directions of **g-max** and **g-2nd** by regressing θ on the time since divergence.

Metric 3 Random Skewers

Third, we tested for divergence in shape among **G**-matrices by regressing the correlation of repeated $\Delta\bar{\mathbf{z}}$ vectors against the time since divergence using random skewers (Cheverud, 1996). This method tests the correlation between $\Delta\bar{\mathbf{z}}$ when solved for the same set of randomly generated β selection vectors (i.e. “skewers”) for a pair of matrices. Lower correlations result from different responses, which are only possible if the structure of the matrices is divergent as well (Cheverud & Marroig, 2007). For each matrix comparison, we calculated the response vectors to each **G**-matrix for the same set of 10,000 randomly generated selection vectors, performed in the program R using the function “RandomSkewers()” in the package “evolqg” (Melo et al., 2015). The metric was calculated as the average vector correlation among all matrix response vectors (Porto et al., 2009).

Metric 4 Disparity

Fourth, we tested divergence among **G**-matrices by regressing matrix (dis)similarity against time using the matrix disparity method presented in Steppan (1997b). We calculated disparity as the sum of the absolute value differences between the off-diagonal elements in both upper triangles. The disparity calculation is shown in Eq. 3 found in Online Resource 2, with minor modifications to the notation originally used by Steppan (1997b). Disparity captures both shape and size differences.

Metric 5 Conditional Evolvability

We also consider here one evolutionary statistic that could assess the evolutionary consequences of the selection on the **G**-matrices: the mean conditional evolvability. This metric captures the ability of a population to respond to directional selection in the presence of stabilizing selection on other trait combinations (Hansen & Houle, 2008) and was calculated using the function “*evolvabilityMeans()*” implement in the R package “*evolvability*” (Bolstad et al., 2014). We applied a shrink method to the **G**-matrices to keep the matrix positive-definite using the function “*shrink()*” available in Haber (2016). The mean conditional evolvability accounts with both total variance and also traits covariance and it is proportional to matrix shape and size. We estimated the Euclidean distance matrix between species mean conditional evolvability values and use those values to evaluate the matrix shape and size changes (divergence) by regressing with time since divergence (Online Resource 5).

Results

G-matrix Estimation and Reconstruction

Convergence of the REML estimates was achieved in both *Phyllotis* species. For *P. vaccarum*, a model with genetic variance in two dimensions fits up to 17.76 AICc units better than the full- and the other reduced-rank models. For *P. darwini*, the full-rank model was superior by up to 21.4 units to the lower-rank models. The estimated matrices for both *P. darwini* and *P. vaccarum* from WOMBAT (Meyer, 2007), with each variance component sampling error, and the reconstructed ancestral matrices for both sigmodontine

and murine ancestors are provided in Online Resource 3. The ratio between the estimate of variance components and sampling errors of are is mostly greater than two for *P. vaccarum* and less than two for *P. darwini*.

G-matrix Similarity

The seven pairwise matrix comparisons showed mixed results depending on which of the three methods were used, and whether the comparison involved the covariance or correlation **G**-matrix (Table 5, and detailed in Online Resource 4). The null hypothesis is of no similarity. Significant similarity between covariance **G**-matrices was supported in two of the seven comparisons; *R. norvegicus*–*P. darwini* ($P \leq 0.05$ [r_M , K_c]), and Murine–Sigmodontine ($P \leq 0.05$ [r_M , r_s , K_c]). In contrast, significant similarity between correlation **G**-matrices was supported in five of the seven all-nodes comparisons for at least one test with a modal average of three. The two non-significant exceptions were *M. musculus*–*P. vaccarum* and *P. darwini*–*P. vaccarum*. Notably, Random Skewers exhibited significant similarity among covariance **G**-matrices in five pairwise comparisons. The two non-significant exceptions were between *M. musculus*–*P. vaccarum* and *R. norvegicus*–*P. vaccarum* (Online Resource 4).

G-matrix Divergence

Only five out of a total of 30 regressions using different divergence metrics were found to have significant support. A summary of the regression results for the five matrix comparison metrics against time since divergence are presented in Table 6 and detailed in Online Resource 5.

Table 5 Results for the three **G**-matrix similarity tests

G -matrix comparison pairwise set	r_M (Pearson)		r_s (Spearman)		K_c (Dietz)	
	G_{cor}	G_{cov}	G_{cor}	G_{cov}	G_{cor}	G_{cov}
<i>M. musculus</i> – <i>R. norvegicus</i>	0.52*	0.08	0.60*	−0.04	55*	8
<i>M. musculus</i> – <i>P. darwini</i>	0.48*	0.10	0.49*	0.13	39	6
<i>M. musculus</i> – <i>P. vaccarum</i>	0.63	0.10	0.50	0.30	35	12
<i>R. norvegicus</i> – <i>P. darwini</i>	0.51*	0.37*	0.51*	0.27	56*	46*
<i>R. norvegicus</i> – <i>P. vaccarum</i>	0.10	−0.14	0.55*	−0.17	−10	−4
<i>P. darwini</i> – <i>P. vaccarum</i>	0.30	0.28	0.25	0.25	20	26
ANC _{MusRattus} –ANC _{Phyllotis}	0.87*	0.75*	0.86*	0.66*	100*	94*

The test statistic and significance are reported here for each set of **G**-matrices compared using either the correlation (G_{cor}) or covariance (G_{cov}) values. All three tests here evaluate against a null that corresponding matrix cells are dissimilar in magnitude, therefore significance implies greater similarity observed between the structure of the compared **G**-matrices then due to chance

r_M = Pearson product-moment correlation coefficient, r_s = Spearman's rank, K_c = Dietz statistic of association, G_{cor} = correlation **G**-matrix, G_{cov} = covariance **G**-matrix

* $P \leq 0.05$

None of the matrix correlation regressions among tip-species or contrasts on the time since divergence using the covariance or correlation **G**-matrix were significant. Two *g-max* regressions were significant: among contrasts using the correlation **G**-matrix (Fig. 3A) ($F_{1,1}=750$, $P=0.0232$, $r^2=0.999$) suggesting increasing similarity (*g-max* θ) with time since divergence, and among tips (Fig. 3B) ($F_{1,1}=9.09$, $P=0.0394$, $r^2=0.694$) using the covariance **G**-matrix, suggesting decreasing similarity (*g-max* θ) with time. None of the *g-2nd* regressions among tip-species or contrasts were significant. Also, both random skewers regressions among contrasts were significant, again suggesting increasing similarity between the response to selection and time in both the correlation **G**-matrix (Fig. 4A) ($F_{1,1}=9949$, $P=0.0063$, $r^2=0.999$), and the covariance **G**-matrix (Fig. 4B) ($F_{1,1}=1692$, $P=0.01547$, $r^2=0.999$). A single disparity regression among contrasts was significant, suggesting decreasing disparity in the correlation **G**-matrices with time since divergence (Fig. 5A) ($F_{1,1}=20,427$, $P=0.0044$, $r^2=0.999$). Our analysis also reveals a non-significant association between the mean conditional evolvability and time since divergence, although interestingly, trending in the opposite direction to most of the other metrics (increasing divergence with time; Fig. 5B).

Discussion

Understanding the nature of the **G**-matrix may have far reaching implications beyond just quantitative genetics. For example, an unstable or evolving **G**-matrix may allow

Table 6 Results for **G**-matrix divergence

Regression metric	Contrasts		Tip-species	
	G_{cor}	G_{cov}	G_{cor}	G_{cov}
Correlation				
r_M	0.0528	0.0309	0.0053	-0.0092
r_s	0.0618	0.0232	0.0134	-0.0030
K_c	7.6957	4.6416	0.0740	-0.4939
Orientation				
<i>g-max</i>	-1.9752*	-0.5639	-0.1326	1.8278*
<i>g-2nd</i>	-3.1416	0.3772	-1.2376	-0.4906
Random skewers	0.0408*	0.0250*	0.0050	-0.0041
Disparity	-0.0387*	-0.0297	-0.0048	-0.0027
Conditional evolvability ^a	na	0.5582	na	0.0075

Elements are the slopes of linear regressions for the four metrics regressed on divergence time (in millions of years), either for the sets of independent contrasts or tip-species matrix comparisons

* $P \leq 0.05$

^aConditional evolvabilities can only be calculated on covariance matrices

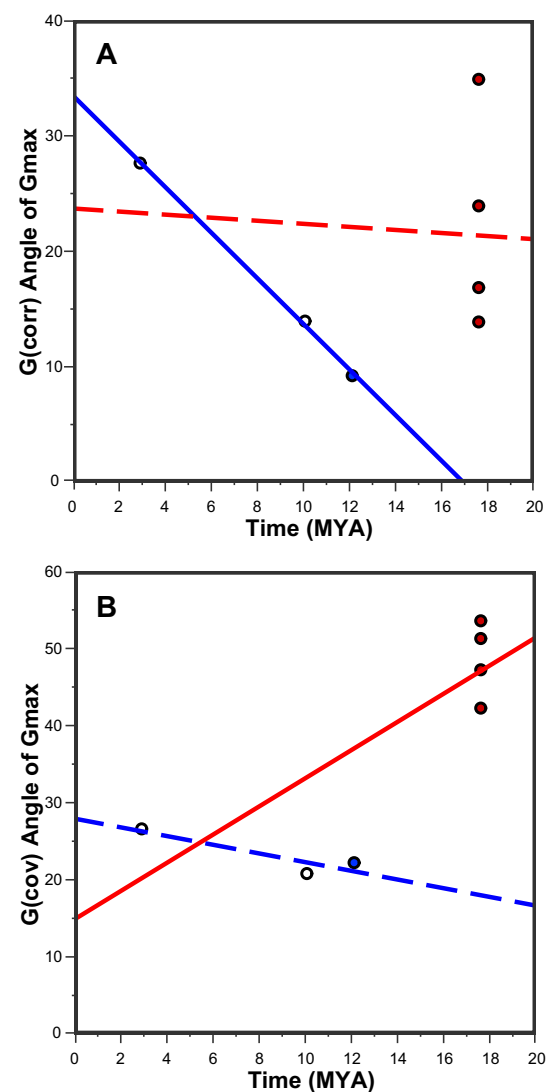


Fig. 3 The angle of *g-max* divergence among **A** correlation, or **B** covariance **G**-matrices against the time since divergence in millions of years. Dashed regression lines indicate non-significant slopes. Each regression set and the associated points (a comparison pair) are indicated by line color and point fill: Contrasts=blue and white; Tip-species=red and white (Color figure online)

greater independent phenotypic evolution, which could, in turn, promote adaptive radiations (Schluter, 1996; Selz et al., 2014). Alternatively, a stable **G**-matrix could be used to improve neutral evolutionary models (Arnold et al., 2008). Here we tested several hypotheses on the stability of the **G**-matrix by expanding on the number of species in Kohn and Atchley's (1988) comparison between the murine rodents *M. musculus* and *R. norvegicus*. To triple the number of phylogenetic contrasts we generated new **G**-matrices for a set of pelvic traits in two additional rodents *P. vaccarum* and *P. darwini* in Sigmodontinae, spanning 36 million years of divergence when analyzing all four-species together.

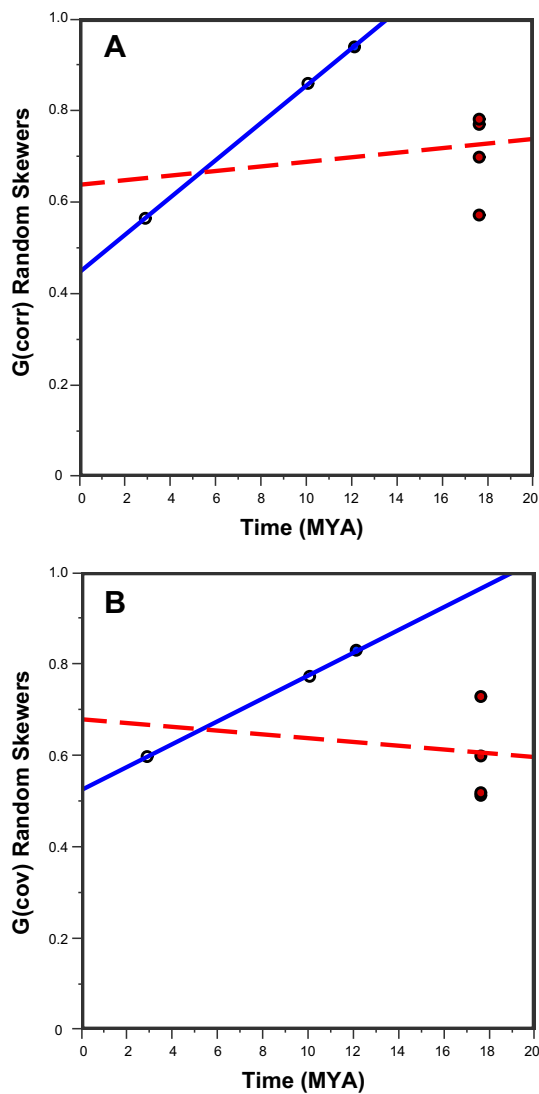


Fig. 4 Random skewers deviation between the **A** correlation, or **B** covariance **G**-matrices against the time since divergence in millions of years. Dashed regression lines indicate non-significant slopes. Each regression set and the associated points (a comparison pair) are indicated by line color and point fill: Contrasts=blue and white; Tip-species=red and white (Color figure online)

Pairwise Matrix Comparisons

We tested the hypothesis that the **G**-matrix was stable across species, in part by replicating the approach taken by Kohn and Atchley (1988). The results of our comparisons were inconsistent in rejection or support of the hypothesis; while we were generally unable to reject the null of different structure in the covariance **G**-matrices, we overwhelmingly supported similarity in the correlation **G**-matrices. Our reanalysis confirmed the results for all comparisons of the **G**-matrix between *M. musculus* and *R. norvegicus* as those in Kohn and Atchley (1988), finding no significant similarity among

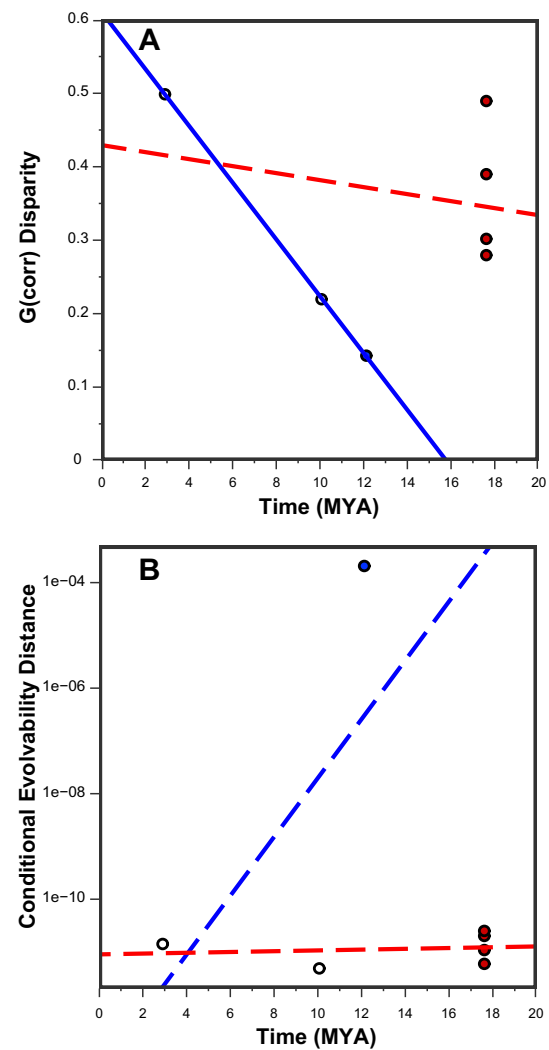


Fig. 5 **A** Disparity between correlation **G**-matrices against the time since divergence in millions of years. **B** Euclidean distances among the conditional evolvabilities of each species regressed against time since divergence. Dashed regression lines indicate non-significant slopes. Each regression set and the associated points (a comparison pair) are indicated by line color and point fill: Contrasts=blue and white; Tip-species=red and white (Color figure online)

covariances while supporting similarity when using the correlations. Such conflicting outcomes could result from the greater range of magnitudes among covariances when compared to the variance-scaled correlations, a common factor to consider when working with any set of phenotypic traits that may be measured in different units or scales. This artifact is unlikely here because the raw pelvic trait measurements were recorded in millimeters and differed at most by a single order of magnitude before log transformation. One notable contrast was between the reconstructed ancestors which were similar in all tests for both the covariance and correlation **G**-matrix, this result however could be a consequence of the reconstruction method used to estimate

ancestral values, where, in the absence of an outgroup, any optimization will tend to minimize change along the long basal branch, making the two reconstructed ancestors more similar.

Matrix Divergence Against Time

Whether selection or drift are dominant processes, a general expectation is that among species, divergence in shared traits results from accumulating independent changes, a function of the time since their common ancestor (Felsenstein, 1985; Revell et al., 2008). In fact, this pattern is a key argument to test for phylogenetic signal, as it may explain a large degree of the phenotypic disparity across species (Blomberg et al., 2003; Münkemüller et al., 2012). Because the **G**-matrix is the result of inherited phenotypic traits, an expectation that the covariances and correlations themselves would exhibit phylogenetic signal in a similar pattern is a logical conclusion (Roff & Mousseau, 2005).

We tested the hypothesis that divergence in **G**-matrices increased with time, in particular, following a Brownian motion pattern. This null model would apply to the situation where either most evolution was due to genetic drift or where the lineages have tracked adaptive peaks but where those peaks have themselves shifted optima in inconsistent directions. In contrast to the previous pairwise evaluations of similarity, our focus here was on the degree to which change in the **G**-matrix reflected phylogenetic history. With only five of the 30 total regressions being significant at $P=0.05$, we conclude that no general relationship exists between **G**-matrix divergence and time for this data set. We will however consider the results of the five metrics separately, as each was chosen to capture different aspects of the **G**-matrix. Caution should be taken in the interpretation of any significant regressions here as the probability of recovering false positives increases with multiple comparisons.

As most of the previous pairwise comparisons had shown considerable support for similarity across correlation **G**-matrices, it was not surprising that all regressions of these values on divergence time were not significant. Nonetheless, the same result was not expected considering the covariance **G**-matrices, as the majority of the same pairwise comparisons had also failed to support equality. We can conclude that while the pattern within the correlation **G**-matrix has remained stable across these four-species, there still exists considerable divergence in the covariance **G**-matrix, independent of time, and low in phylogenetic signal. One possible explanation for this pattern is that the developmental processes guiding the formation of the rodent pelvis is conserved and under stabilizing selection, but that the magnitude of heritable covariation in pelvic traits can fluctuate along lineages on much shorter time scales than even the most recent of these contrasts. Orientation of the

major axes of the **G**-matrix can be highly unstable because they are defined by only the current genetic covariation available in the population (Berner, 2012). Additionally, the divergence in orientation of major axes may not sufficiently capture the actual structural changes (Phillips & Arnold, 1999). Together these factors help erase phylogenetic signal.

The expectation for any of these phylogenetic comparisons is that at the initial divergence time of 0 years between two species the **G**-matrices should be identical (assuming a well estimated **G**-matrix). Thus, as independent evolution proceeds, the lineages can only become more dissimilar, resulting in a positive correlation between time and divergence. Although studies of **G**-matrices among diverging populations within a single species have often found similar covariance structures (Arnold et al., 2008; Lofsvold, 1986), this is clearly an overly simplified scenario, because the factors that drove divergence may have also biased or resulted from a partitioning of the genetic variance (Walter et al., 2018). However, the random skewers regression among contrasts suggested an increase in similarity with greater divergence time, the opposite expectation of change in the **G**-matrix under drift or divergent selection (McGlothlin et al., 2018). Similarly, we found that disparity between the correlation **G**-matrices decreased with time since divergence increases, the opposite of the expectation of diverging **G**-matrices. These results indicate that the intercept is not at zero disparity, and suggests that variation among populations in a species, or variation in our estimates of those matrices, is comparable to among species. We also suggest that although three phylogenetically independent contrasts is better than one, the reconstruction of ancestral matrices in such a small sample might artifactually underestimate their divergence, contributing to the trends we observed. These results emphasize the need for even greater phylogenetic sampling.

Notably, some of the phylogenetic patterns among the analyses of independent contrasts are driven in part by the greater variation contained in the *P. vaccarum* covariance matrix, which with a trace of 10.4 is almost twice the size of the next largest species matrix (*P. darwini* at 5.4). All metrics that incorporate size will show a larger difference between the two *Phyllotis* species than shape differences (variation among covariances) alone would suggest. Therefore, the *Phyllotis* contrast, which occurs over the shortest branch, will tend to be large for most metrics, contributing to the pattern of decreasing difference with time. The fact that both *Phyllotis* had wild-caught parents while the murines were the products of lab-reared animals over many generations could be a factor, but why *P. vaccarum* should have a larger trace than *P. darwini* is unclear.

We expanded upon a previous analysis by Kohn and Atchley (1988) to triple the number of phylogenetic comparisons of **G**-matrices. The primary result is that among

the four-species separated by up to 18 million years of evolutionary history, the pelvic **G**-matrix has remained essentially stable, or if there was any divergence, that divergence did not accumulate significantly over macroevolutionary time scales, as would be expected if **G**-matrices evolved like typical phenotypic traits under a Brownian motion model (Arnold et al., 2008). Similar **G**-matrices were estimated despite different breeding systems among species and statistical models of estimation. Correlation structure in particular was stable, while covariance structure showed some divergence even among close relatives. These results are suggestive of an Ornstein–Uhlenbeck pattern (Hansen, 1997; Uhlenbeck & Ornstein, 1930), where a broad adaptive peak permits small but rapid deviation in structure (occurring over thousands of years or among populations within species), but that a constraint ultimately prevents any large divergence from accumulating on scales of millions of years. Such constraints could arise from conservation of the developmental and locomotory function of the innominate and pelvis, and may not equate across other skeletal structures (Coutinho et al., 2013; Pomikal & Streicher, 2009). For example, covariance structure in the skull appears to be more variable (Atchley et al., 1981; Lofsvold, 1986). These results support the assumption that the **G**-matrix remains stable over macroevolutionary scales for at least some structures, and that multiple structures need to be examined before drawing generalizations about the entire organism or genome.

While it may be interesting to utilize more advanced, or biologically informative metrics on **G**-matrices (Aguirre et al., 2014; Garcia, 2012), we argue that it is more valuable to increase the phylogenetic sample size in comparative quantitative genetics (McGlothlin et al., 2018; Steppan, 1997b; Steppan et al., 2002). Particularly, we encourage future work with taxa that have experienced stronger divergent selection that may have been a greater factor in reshaping the **G**-matrix than among the relatively non-specialized rodents we have compared here (Jones et al., 2003; Kingsolver et al., 2001; Roff & Fairbairn, 2012). All four-species are generalist terrestrial quadrupeds, although *R. norvegicus* spends more time both climbing and swimming than do the other three species, especially relative to *Phyllotis*. Inclusion of more specialized species would be illuminating.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11692-022-09559-z>.

Acknowledgements This research was supported by US National Science Foundation Grants DEB-0108422 and DEB-1754748 to SJS. We are deeply grateful to the staff of the Mammal Division at the Field Museum of Natural History, especially Bill Stanley and Bruce Patterson for the preparation of the voucher specimens used in this study, and to Juan Oyarce, for his technical assistance in capturing and caring for the animals. To David Houle for the time spent providing invaluable advice throughout this study, particularly in **G**-matrix estimation and

statistical comparisons. Chris Klingenberg, Mihaela Pavlicev, and two anonymous reviewers provided helpful feedback on the manuscript.

Author Contributions LIW, AES and SJS conceived the study. SJS acquired funding. LIW and AES managed the breeding colonies. LEC-W and CJS collected data. CJS and BMC performed analyses. CJS wrote the manuscript. LIW, BMC, and SJS edited and contributed revisions.

Funding US National Science Foundation Grants DEB 0108422 and DEB-1754748 to Scott Steppan.

Data Availability All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Code Availability Statistical code available by request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Ethical Approval Not applicable.

References

- Agrawal, A. F., & Stinchcombe, J. R. (2009). How much do genetic covariances alter the rate of adaptation? *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1659), 1183–1191.
- Aguirre, J., Hine, E., McGuigan, K., & Blows, M. (2014). Comparing **G**: Multivariate analysis of genetic variation in multiple populations. *Heredity*, 112(1), 21.
- Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C., & Jones, A. G. (2008). Understanding the evolution and stability of the **G**-matrix. *Evolution*, 62(10), 2451–2461.
- Atchley, W. R., & Rutledge, J. J. (1980). Genetic components of size and shape. I. Dynamics of components of phenotypic variability and covariability during ontogeny in the laboratory rat. *Evolution*, 34, 1161–1173.
- Atchley, W. R., Rutledge, J. J., & Cowley, D. E. (1981). Genetic components of size and shape. II. Multivariate covariance patterns in the rat and mouse skull. *Evolution*, 35(6), 1037–1055.
- Baker, R. H., & Wilkinson, G. S. (2003). Phylogenetic analysis of correlation structure in stalk-eyed flies (*Diasemopsis diopsidae*). *Evolution*, 57(1), 87–103.
- Baker, R. L., Chapman, A. B., & Wardell, R. T. (1975). Direct response to selection for postweaning gain in the rat. *Genetics*, 80(1), 171–189. <https://doi.org/10.1093/genetics/80.1.171>
- Berner, D. (2012). How much can the orientation of **G**'s eigenvectors tell us about genetic constraints? *Ecology and Evolution*, 2(8), 1834–1842.
- Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, 57(4), 717–745.
- Bolstad, G. H., Hansen, T. F., Pélabon, C., Falahati-Anbaran, M., Pérez-Barrales, R., & Armbruster, W. S. (2014). Genetic

- constraints predict evolutionary divergence in *Dalechampia* blossoms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1649), 20130255.
- Cheverud, J. M. (1982). Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution*, 36(3), 499–516.
- Cheverud, J. M. (1995). Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *The American Naturalist*, 145(1), 63–89.
- Cheverud, J. M. (1996). Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *Journal of Evolutionary Biology*, 9(1), 5–42.
- Cheverud, J. M., & Marroig, G. (2007). Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genetics and Molecular Biology*, 30(2), 461–469.
- Cheverud, J. M., Leamy, L. J., Atchley, W. R., & Rutledge, J. (1983a). Quantitative genetics and the evolution of ontogeny: I. Ontogenetic changes in quantitative genetic variance components in random bred mice. *Genetics Research*, 42(1), 65–75.
- Cheverud, J. M., Rutledge, J., & Atchley, W. R. (1983b). Quantitative genetics of development: Genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution*, 37(5), 895–905.
- Coutinho, L. C., de Oliveira, J. A., & Pessôa, L. M. (2013). Morphological variation in the appendicular skeleton of Atlantic forest sigmodontine rodents. *Journal of Morphology*, 274(7), 779–792. <https://doi.org/10.1002/jmor.20134>
- De Oliveira, F. B., Porto, A., & Marroig, G. (2009). Covariance structure in the skull of Catarrhini: A case of pattern stasis and magnitude evolution. *Journal of Human Evolution*, 56(4), 417–430.
- Dietz, E. J. (1983). Permutation tests for association between two distance matrices. *Systematic Zoology*, 32, 21–26.
- Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, 125(1), 1–15.
- Garcia, C. (2012). A simple procedure for the comparison of covariance matrices. *BM Evolutionary Biology*, 12(1), 222.
- Haber, A. (2016). Phenotypic covariation and morphological diversification in the ruminant skull. *The American Naturalist*, 187(5), 576–591.
- Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, 51(5), 1341–1351.
- Hansen, T. F., & Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology*, 21(5), 1201–1219.
- Houle, D., Mezey, J., & Galpern, P. (2002). Interpretation of the results of common principal components analyses. *Evolution; International Journal of Organic Evolution*, 56(3), 433–440.
- Jayat, P. J., Teta, P., Ojeda, A. A., Ortiz, P. E., Novillo, A., Lanzone, C., Osland, J. M., Stepan, S. J., & Ojeda, R. A. (2021). Systematics of the *Phyllotis xanthopygus* species complex (Rodentia, Cricetidae) in the central Andes, with the description of a new species from central-western Argentina. *Zoologica Scripta*, 50(6), 689–706.
- Jones, A. G., Arnold, S. J., & Bürger, R. (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution*, 57(8), 1747–1760.
- Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C., Hoang, A., Gibert, P., & Beerli, P. (2001). The strength of phenotypic selection in natural populations. *The American Naturalist*, 157(3), 245–261.
- Kohn, L. A. P., & Atchley, W. R. (1988). How similar are genetic correlation structures? Data from mice and rats. *Evolution*, 42(3), 467–481.
- Kruuk, L. E., Clutton-Brock, T. H., Slate, J., Pemberton, J. M., Brotherton, S., & Guinness, F. E. (2000). Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), 698–703.
- Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution*, 33(1Part2), 402–416.
- Legendre, P., & Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, 10(5), 831–844.
- Lofsvold, D. (1986). Quantitative genetics of morphological differentiation in *Peromyscus*. I. Tests of the homogeneity of genetic covariance structure among species and subspecies. *Evolution*, 40, 559–573.
- Matta, B., & Bitner-Mathé, B. (2004). Genetic architecture of wing morphology in *Drosophila simulans* and an analysis of temperature effects on genetic parameter estimates. *Heredity*, 93(4), 330.
- McGlothlin, J. W., Kobiela, M. E., Wright, H. V., Mahler, D. L., Kolbe, J. J., Losos, J. B., & Brodie, E. D., III. (2018). Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. *Evolution Letters*, 2(4), 310–322.
- Melo, D., Garcia, G., Hubbe, A., Assis, A. P., & Marroig, G. (2015). Evolq: An R package for evolutionary quantitative genetics [version 3]. *F1000Research*, 4, 925.
- Meyer, K. (2007). WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University Science B*, 8(11), 815–821. <https://doi.org/10.1631/jzus.2007.B0815>
- Mezey, J. G., & Houle, D. (2003). Comparing G matrices: Are common principal components informative? *Genetics*, 165(1), 411–425.
- Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffrers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, 3(4), 743–756.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H., Wagner, H., & Oksanen, M. J. (2013). Package 'vegan.' *Community Ecology Package, Version*, 2(9), 1–295.
- Park, Y., Hansen, C., Chung, C., & Chapman, A. B. (1966). Influence of feeding regime on the effects of selection for postweaning gain in the rat. *Genetics*, 54(6), 1315–1327.
- Phillips, P. C., & Arnold, S. J. (1999). Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution*, 53, 1506–1515.
- Pigliucci, M. (2006). Genetic variance-covariance matrices: A critique of the evolutionary quantitative genetics research program. *Biology and Philosophy*, 21(1), 1–23.
- Pomikal, C., & Streicher, J. (2009). 4D-analysis of early pelvic girdle development in the mouse (*Mus musculus*). *Journal of Morphology*, 271(1), 116–126. <https://doi.org/10.1002/jmor.10785>
- Porto, A., de Oliveira, F. B., Shirai, L. T., De Conto, V., & Marroig, G. (2009). The evolution of modularity in the mammalian skull I: Morphological integration patterns and magnitudes. *Evolutionary Biology*, 36(1), 118–135.
- Porto, A., Schmelter, R., VandeBerg, J. L., Marroig, G., & Cheverud, J. M. (2016). Evolution of the genotype-to-phenotype map and the cost of pleiotropy in mammals. *Genetics*, 204(4), 1601–1612.
- Revell, L. J. (2007). The G matrix under fluctuating correlational mutation and selection. *Evolution*, 61(8), 1857–1872.
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecology and Evolution*, 3, 217–223.
- Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, 57(4), 591–601.
- Riska, B., Atchley, W. R., & Rutledge, J. J. (1984). A genetic analysis of targeted growth in mice. *Genetics*, 107(1), 79–101.

- Roff, D. A. (1995). The estimation of genetic correlations from phenotypic correlations: A test of Cheverud's conjecture. *Heredity*, 74(5), 481.
- Roff, D. A., & Fairbairn, D. J. (2012). A test of the hypothesis that correlational selection generates genetic correlations. *Evolution*, 66(9), 2953–2960.
- Roff, D. A., & Mousseau, T. (2005). The evolution of the phenotypic covariance matrix: Evidence for selection and drift in *Melanoplus*. *Journal of Evolutionary Biology*, 18(4), 1104–1114.
- Roff, D. A., Prokkola, J., Krams, I., & Rantala, M. (2012). There is more than one way to skin a G matrix. *Journal of Evolutionary Biology*, 25(6), 1113–1126.
- Rossoni, D. M., Assis, A. P. A., Giannini, N. P., & Marroig, G. (2017). Intense natural selection preceded the invasion of new adaptive zones during the radiation of New World leaf-nosed bats. *Scientific Reports*, 7(1), 11076.
- Rutledge, J. J., & Chapman, A. B. (1975). Systematic cross-fostering within control lines. *Journal of Animal Science*, 40, 70–74.
- Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 50(5), 1766–1774.
- Schluter, D., Price, T., Mooers, A. Ø., & Ludwig, D. (1997). Likelihood of ancestor states in adaptive radiation. *Evolution*, 51(6), 1699–1711. <https://doi.org/10.1111/j.1558-5646.1997.tb05095.x>
- Selz, O. M., Lucek, K., Young, K. A., & Seehausen, O. (2014). Relaxed trait covariance in interspecific cichlid hybrids predicts morphological diversity in adaptive radiations. *Journal of Evolutionary Biology*, 27(1), 11–24. <https://doi.org/10.1111/jeb.12283>
- Steppan, S. J. (1997a). Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution*, 51(2), 571–586.
- Steppan, S. J. (1997b). Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. *Evolution*, 51(2), 587–594.
- Steppan, S. J. (2004). Phylogenetic comparative analysis of multivariate data. In M. Pigliucci & K. A. Preston (Eds.), *Phenotypic integration* (pp. 325–344). Oxford University Press.
- Steppan, S. J., & Schenk, J. J. (2017). Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. *PLoS ONE*, 12(8), e0183070.
- Steppan, S. J., Phillips, P. C., & Houle, D. (2002). Comparative quantitative genetics: Evolution of the G matrix. *Trends in Ecology & Evolution*, 17(7), 320–327.
- Turelli, M. (1988). Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution*, 42(6), 1342–1347.
- Uhlenbeck, G., & Ornstein, L. (1930). On the theory of the Brownian motion. *Physical Review*, 36, 823–841.
- Walter, G. M., Aguirre, J. D., Blows, M. W., & Ortiz-Barrientos, D. (2018). Evolution of genetic variance during adaptive radiation. *The American Naturalist*, 191(4), 108–128.
- Leiva, D., Solanas, A., de Vries, H., & Kenny, D. A. (2010). DyaDA: An R package for dyadic data analysis. In *Proceedings of measuring behavior 2010* (pp. 162–165).
- Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Vol. 1). Sinauer Sunderland.
- Maddison, W. P., & Maddison, D. R. (2017). Mesquite: A modular system for evolutionary analysis. Version 3.31. <http://mesquiteproject.org>
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria. <https://www.R-project.org/>
- Roff, D. A. (2012). *Evolutionary quantitative genetics*. Springer Science & Business Media.