Macroevolutionary Divergence along Allometric Lines of Least Resistance in Frog Hindlimb Traits and Its Effect on Locomotor Evolution

Monique N. Simon,[1,*] Elodie A. Courtois,[2] Anthony Herrel,[3] and Daniel S. Moen[1,4]

1. Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, California 92521; and Department of Integrative Biology, Oklahoma State University, Stillwater, Oklahoma 74078; 2. CNRS-Guyane USR 3456, Résidence Le Relais, Cayenne, French Guiana, France; and Station d'Ecologie Expérimentale du CNRS à Moulis, USR 2936, Saint Girons, France; 3. UMR 7179 CNRS and Muséum National d'Histoire Naturelle, Département Adaptations du Vivant, Paris, France; 4. Institut de Biologie de l'École Normale Supérieure (IBENS), École Normale Supérieure, Paris, France

Submitted June 3, 2024; Accepted February 14, 2025; Electronically published May 2, 2025

Online enhancements: supplemental PDF.

ABSTRACT: Understanding whether and why microevolutionary patterns of trait covariation match macroevolutionary divergence is essential for linking evolution at different timescales. However, recent work has focused on developmental constraints for alignment between intraspecific variation and divergence, neglecting a potential role of natural selection on function to connect these scales. Here, we compare the support for the selection and constraint hypotheses to explain both phenotypic trait covariation and species divergence. To test these hypotheses, we collected data on hindlimb and jumping performance traits within and across species of two frog genera. We compared patterns of within-species phenotypic variation (the **P** matrix) with divergence and selective covariance matrices, from which we could extract the major axes of the realized adaptive landscape (AL), the directions in which adaptive peaks shifted the most over evolutionary time. We also tested whether the major axes of the AL were related to selection on jumping performance. We found high alignment between patterns of variation across scales. Most divergence occurred in allometric size, defined as the first eigenvector of the **P** matrix. However, jumping performance gradients were unaligned with the major axes of the AL and the **P** matrix. Across species, however, evolution of maximum acceleration showed a strong negative relationship with changes in allometric size. We infer that the jumping peak evolved under fluctuating selection, and species have tracked the peak along the direction of most within-species variation, allometric size. We conclude that long-term hindlimb divergence was constrained by developmental interactions among traits associated with growth and not net directional selection. Nonetheless, divergence on size indirectly influenced jumping evolution.

* Corresponding author; email: monique.nouailhetas@gmail.com.

ORCIDs: Simon, https://orcid.org/0000-0003-0106-2660; Courtois, https://orcid.org/0000-0002-9801-4771; Herrel, https://orcid.org/0000-0003-0991-4434; Moen, https://orcid.org/0000-0003-1120-0043.

Keywords: Anura, functional performance, genetic constraints, macroevolution, microevolution, morphological evolution.

Introduction

Most phenotypes affecting fitness are complex, composed of several traits and influenced by both genetic and environmental factors (Lande 1979; Lande and Arnold 1983; Wagner and Altenberg 1996; Lynch and Walsh 1998). Although the short-term evolution of complex phenotypes is well characterized (but sometimes hard to predict), understanding the evolution of complex phenotypes at larger evolutionary timescales is challenging. Connecting microevolution to macroevolutionary divergence is complicated by the evolution of the genetic architecture of traits, in which trait covariances are widespread in complex phenotypes. Genetic architecture can remain stable or change over long periods of time (Arnold et al. 2008), with each outcome affecting phenotypic divergence differently. Unraveling which evolutionary processes shape genetic covariances and correlations over long timescales is essential to better understand the evolution of complex phenotypes and, ultimately, to better connect micro- and macroevolution.

At microevolutionary scales, the divergence of complex phenotypes depends on the genetic architecture of multiple traits, represented by the additive genetic variance-covariance matrix (the **G** matrix; Lande 1979; Arnold et al. 2001, 2008). The **G** matrix is composed of genetic variation of traits on its diagonal and genetic covariation between traits in the off-diagonal, reflecting the coinheritance of multiple traits (Lande 1979; Arnold 1992; Lynch and Walsh 1998). Higher

 https://doi.org/10.1086/735569

genetic covariation between traits may bias evolutionary trajectories of phenotypic means away from the direction of selection, toward directions that have the most genetic variation (Lande 1979; Lande and Arnold 1983; Schluter 1996; Hansen and Houle 2008; Walsh and Blows 2009). Yet given enough time, phenotypic means may reach adaptive optima and the **G** matrix would no longer affect the evolution of those means (Arnold et al. 2001).

However, genetic architecture can evolve over long timescales, making its role in shaping long-term phenotypic evolution less straightforward than at microevolutionary scales (Turelli 1988; Steppan et al. 2002; Arnold et al. 2008). Both theoretical (Jones et al. 2004, 2012, 2014; Revell 2007; Arnold et al. 2008) and empirical (Cano et al. 2004; Eroukhmanoff and Svensson 2011; Björklund et al. 2013; Careau et al. 2015; Renaud et al. 2021; Chantepie et al. 2024) studies have shown that the **G** matrix can change, although the changes may be subtle even over millions of years (McGlothlin et al. 2022). On the other hand, several processes can promote **G** matrix stability, such as strong correlated mutational effects (Lande 1979; Hohenlohe and Arnold 2008), persistent correlational selection molding genetic correlations (Cheverud 1984; Jones et al. 2003; Revell 2007), and persistent directional selection that aligns the **G** matrix and the adaptive landscape (AL; Jones et al. 2004, 2007, 2014; Pavlicev et al. 2011; Melo and Marroig 2015; Houle et al. 2017). Therefore, connecting the evolution of the **G** matrix with phenotypic divergence at the macroevolutionary scale is difficult and depends on the relative strengths of distinct evolutionary forces shaping the **G** matrix and species' phenotypic means.

Empirical evidence of species divergence aligned with the axis of most genetic variation (also known as $\mathbf{g}_{max}$) was first shown by Schluter (1996) across different vertebrates. Yet divergence was less aligned with $\mathbf{g}_{max}$ in distantly related species (around 4 million years of divergence) compared with species that diverged more recently (Schluter 1996). Subsequent work also showed a relatively short-term effect of genetic constraints on divergence in different systems (Bégin and Roff 2004; McGuigan et al. 2005; Hansen and Houle 2008; Chenoweth et al. 2010; Bolstad et al. 2014; Opedal et al. 2022, 2023; Holstad et al. 2024). However, recent articles have shown that such effects can persist much longer than previously thought, spanning tens of millions of years (Houle et al. 2017; McGlothlin et al. 2018). More extensive evidence shows long-term divergence aligned with the axis of most phenotypic variation ($\mathbf{p}_{max}$), which is frequently aligned with $\mathbf{g}_{max}$ (Marroig and Cheverud 2005, 2010; Renaud et al. 2006; Hunt 2007; Simon et al. 2016; Machado 2020; Mongle et al. 2022). Such evidence suggests that the **G** matrix preserves its structure and its influence on phenotypic divergence over long periods of time (Arnold et al. 2001; McGlothlin et al. 2018).

What evolutionary scenarios might explain such a persistent relation between the **G** matrix and divergence? First, the constraint hypothesis states that a persistent correlation of mutational effects (pleiotropic mutations expressed on the mutational matrix, the **M** matrix) promotes stability of genetic correlations (Lande 1980; Jones et al. 2003; Houle et al. 2017), biasing most divergence along axes of most genetic variation, whether divergence was mainly caused by random drift or selection. If selection has been the main driver of divergence, the persistent effect of constraints may indicate the existence of an AL with multiple adaptive peaks or with a fluctuating peak, in which peaks are easier to reach in directions of most genetic variance or lineages chase the peak along directions of most variance, respectively (Steppan et al. 2002; Melo et al. 2016*b*; Holstad et al. 2024). Second, the selection hypothesis implicates evolution along selective lines of least resistance (SLLR), which represent ridges in the AL, that is, directions under weakest stabilizing selection, along which changes in phenotype will produce little or no change in average fitness (Arnold et al. 2001; Hohenlohe and Arnold 2008; Arnold 2023). SLLR can be estimated as the last eigenvector of a matrix of multivariate stabilizing selection coefficients (the $\boldsymbol{\gamma}$ matrix), which represents the curvature and orientation of the AL (Lande and Arnold 1983; Phillips and Arnold 1989; Arnold et al. 2001). The SLLR would then favor peak shifts in their own direction to shape phenotypic means while also aligning the **G** matrix with the AL (Arnold et al. 2001; Jones et al. 2003, 2012; Arnold 2023).

All of these scenarios are plausible, yet those that include selection have been most supported in the literature. Such supporting studies have either shown phenotypic divergence to be driven by selection (Marroig and Cheverud 2005, 2010; Simon et al. 2016; Machado 2020) or have been done with morphological traits that had clear functional relevance, implying diversifying selection (e.g., Renaud et al. 2006; Kolbe et al. 2011; McGlothlin et al. 2018; Mongle et al. 2022; Opedal et al. 2022). Moreover, most of these studies considered feeding or locomotor traits, suggesting a role of selection on function mediating the effects of constraints on divergence. Long-term correlational selection is expected to shape patterns of trait correlations to maintain functional integration, that is, to maintain higher correlations among traits that act together to perform a specific function (Olson and Miller 1958; Cheverud 1984; Estes and Arnold 2007). Thus, selection on functional performance may align long-term phenotypic divergence with intraspecific variation by shaping both trait correlations and phenotypic means.

However, functional performance is rarely incorporated in studies aiming to link micro- to macroevolution. Most studies that compare both scales have focused on scaling relationships (effects of changes in size on performance; e.g., Toro et al. 2003) or considered few species (e.g., Taverne

et al. 2021). Moreover, recent work has focused on developmental reasons for alignment between intraspecific variation and interspecific divergence (e.g., Mongle et al. 2022; Machado et al. 2023). Finally, no empirical work has addressed how selection on functional performance may shape both trait correlations and phenotypic divergence, precluding a test of selection on function as a mechanism to explain the lasting effect of genetic constraints on divergence.

In this article, we pursue a deeper understanding of the link between genetic architecture within populations and phenotypic divergence across species by directly considering selection on trait function. We studied 60 species of frogs belonging to two genera, one arboreal (*Boana*) and the other terrestrial (*Leptodactylus*). We collected morphological data (on hindlimb traits) for all species and jumping performance data (maximum acceleration and velocity) for 18 of those species (fig. 1). We also collected intraspecific data to estimate phenotypic within-species patterns (**P** matrices), which allowed us to test whether within-species trait covariation patterns matched between-species divergence and whether divergence in both morphology and performance were driven by selection. Although we could not estimate a jumping performance $\boldsymbol{\gamma}$ matrix and extract SLLR, we estimated the major axes of the AL shaping hindlimb morphology by estimating a selective covariance matrix, the covariance matrix of peak movements over evolutionary time (Felsenstein 1988; Marroig and Cheverud 2010; Machado 2020). Then we estimated functional performance gradients—the effects of hindlimb morphology on jumping performance—and tested how well they aligned with phenotypic lines of least resistance and with major axes of the AL. If functional performance gradients aligned with both within-species patterns of variation and with the AL, we would support the selection hypothesis and infer that selection on function mediates the alignment of micro- and macroevolution. If functional performance gradients aligned with neither within-species patterns nor the AL, we would support the constraint hypothesis and infer that constraints unrelated to function drove the alignment between micro- and macroevolution.

Methods

Specimens, Sample Sizes, and Taxonomy

We caught individuals in the wild from seven species of *Boana* and eight of *Leptodactylus* in French Guiana in 2013 and 2015. We generally targeted males to reduce potential effects of sexual dimorphism (see below) and because males were calling. In *Boana*, mature males could be identified by the presence of prepollical spines (Pinheiro et al. 2022). However, some species of *Leptodactylus* did not show clear sexual dimorphism, so we verified sex via dissection after performance data collection (see below). We sampled a total of 131 individuals across both genera, collecting 1–17 individuals per species (mean $n = 7.6$) for those we sampled only for interspecific analyses (table S1; tables S1–S10 are available online). To improve estimates of within-species trait covariation patterns, we more extensively sampled the most abundant species at our field sites for each genus: $n = 48$ for *Leptodactylus mystaceus* and $n = 29$ for *Boana boans*. These samples represent data for both morphology and jumping performance.

To better estimate morphological trait covariation within species, we measured museum specimens collected from localities close to French Guiana, so that final sample sizes were $n = 53$ for *L. mystaceus* and $n = 45$ for *B. boans* for morphology-only analysis. Likewise, we included data from the literature (Moen et al. 2013, 2016) and collected new morphological data from additional museum specimens (19 species of *Boana* and 16 species of *Leptodactylus*). In sum, for morphological analysis we analyzed 479 individuals across 33 species of *Boana* and 27 species of *Leptodactylus*. For morphological and performance analysis, we had 167 individuals representing nine *Boana* and nine *Leptodactylus* species (fig. 1). See table S1 for the sample size for each species and associated numbers of males and females. Finally, we confirmed or updated species names based on collection localities and recent taxonomic literature (for details, see the supplemental data in our Dryad repository; Simon et al. 2025).

Jumping Performance

We measured jumping forces using a Kistler piezoelectric force plate (Squirrel, model Z17097) and charge amplifier (model 9865), collecting data at 500 Hz. Body mass was measured before jumping trials on an Ohaus balance precise to 0.1 g. Each frog was then stimulated to jump from the force plate repeatedly until at least three strong jumps were recorded per individual, with one day of rest between trials. To extract maximum jumping forces, we first isolated the highest force peaks in the vertical direction in each force trace. We then extracted the corresponding horizontal (X, Y) and vertical (Z) forces as text files using BioWare software (ver. 5.4.8.0; Kistler).

We describe our analysis of force profiles in detail in the supplemental PDF (figs. S1, S2; figs. S1–S5 are available online). In brief, we used a custom Matlab script (ver. R2020a, MathWorks; script developed by Craig McGowan) to smooth, center, and integrate the X, Y, and Z force traces independently, then calculated single-dimensional jumping velocity and acceleration profiles (Schwaner et al. 2021). We then extracted the peak acceleration and velocity for each trace for each individual as our two performance variables. We used the peak values from the single best jump, defined as

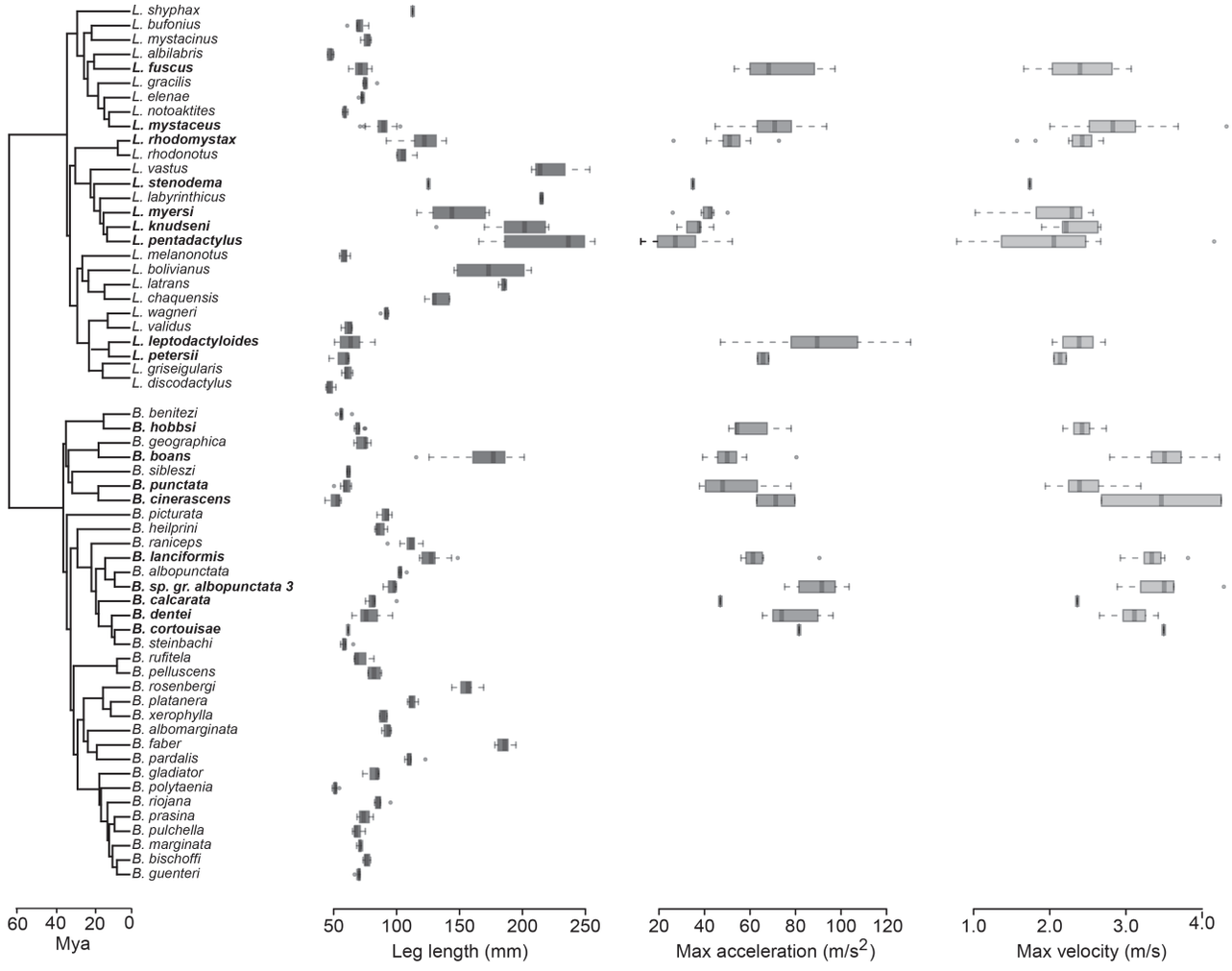


Figure 1: Frog species in the genera *Boana* and *Leptodactylus* for which we collected only morphological data (60 species) or both morphological and jumping performance data (18 species, in boldface). Morphological data were bone lengths (shown here as their sum, leg length) and muscle cross-sectional areas of hindlimbs. Performance data were maximum jumping acceleration and velocity. The phylogeny was from Portik et al. (2023). Boxplots show the median (lines), interquartile range (boxes), 95% distribution (whiskers), and points outside the 95% distribution. Mya = million years ago.

the jump with maximum acceleration, as data for each individual (Wilson et al. 2000; James and Wilson 2008). The force trace data for each jumping peak and the Matlab script are provided in Dryad and Zenodo (Simon et al. 2025). All subsequent analyses were performed in the R programming environment (ver. 4.2.2; R Core Team 2022) and can be reproduced using the data file and supplemental R scripts associated with this article (Simon et al. 2025).

Morphology

After performance trials, we euthanized individuals, preserved them in formalin, and then stored them in 70% ethanol. We primarily collected data on hindlimb traits because hindlimb length and muscle mass strongly predict frog jumping performance within species (e.g., Emerson 1978; Wilson et al. 2000; James et al. 2005, 2007) and especially across species (Gomes et al. 2009; Moen et al. 2013, 2021; Citadini et al. 2018; Moen 2019; Juarez et al. 2020). We used a digital caliper (to 0.1 mm) to externally measure snout-to-vent length, femur length, tibiofibula length, tarsus length, foot length, the largest width of the upper and lower leg, and the width perpendicular to the last two measurements (Juarez et al. 2023; Morinaga et al. 2023). Using the leg widths as axes of an ellipse to calculate area, we estimated the upper- and lower-leg cross-sectional areas (CSA), which should be proportional to muscle CSA, a key component of jumping performance in anurans (Hellam and Podolsky 1969; Peplowski and Marsh 1997; Olson and Marsh 1998; Astley 2016). We took the square root of the

muscle CSA to put it on the same linear scale as our other measurements. Moreover, to be able to interpret morphological patterns in an allometric context (given that size is an important axis of variation in all species; see "Results"), we natural log transformed all trait values.

Within-Species Phenotypic Matrix and Matrix Similarity

To characterize within-species trait covariation, we estimated three within-species variance-covariance matrices (**P** matrices) for the species with highest sample sizes: *Boana boans*, *Boana raniceps*, and *Leptodactylus mystaceus* for morphology, with sample sizes of 45, 44, and 53 individuals, respectively. We first tested for effects of sex within each species on the six hindlimb traits. We used a multivariate linear regression model with all traits as the dependent variables and only sex as the fixed factor for *B. boans* and *L. mystaceus* (our sample of *B. raniceps* had no females). We then used the residuals of the linear models to estimate the **P** matrices in the function CalculateMatrix (EvolQG package, ver. 3.1; Melo et al. 2016*a*).

To determine whether patterns of within-species trait covariation were shared by the three species, we compared **P** matrices using two matrix similarity indices. First, random skewers (Cheverud and Marroig 2007) give the similarity of a pair of matrices as the average vector correlation of response vectors obtained from applying random selection vectors. Second, the Krzanowski projection (Krzanowski 1979) gives similarity in the shared eigenvector subspace (three pairs of eigenvectors) between the matrices being compared. We provide full details of these methods in the supplemental PDF. Both indices showed high similarity among all three **P** matrices (above 0.89 for all pairwise comparisons; table S2). Because of this high similarity, we estimated pooled **P** matrices for each genus as our best estimate of within-species patterns. By pooling all species from the same genus, we achieved higher sample sizes to estimate six-by-six intraspecific **P** matrices: $n = 242$ for *Boana* and $n = 215$ for *Leptodactylus*. We used multivariate linear regression models with both species and sex as fixed factors, as well as their interaction, and then used the residuals of the models to estimate the pooled **P** matrices for *Boana* and *Leptodactylus* using the function CalculateMatrix from EvolQG. However, we also used the **P** matrices from our three species with high intraspecific samples in some analyses to test the sensitivity of our results to using pooled matrices (see below).

After estimating the pooled **P** matrices, we performed an eigenanalysis to interpret which combinations of traits were associated with the most phenotypic variation. We focused on only the first two eigenvectors because they represented more than 85% of all trait variation in the three species. We used those first two eigenvectors of the pooled **P** matrices as surrogates for genetic lines of least resistance because **G** matrices have never been estimated for our study taxa and traits. For morphological traits, **P** matrices are good approximations of **G** matrices in a diversity of biological organisms, for five reasons. First, morphological traits show moderate to high heritability (Phillips and Arnold 1989; Roff 1996; Porto et al. 2009; Styga et al. 2019). Second, phenotypic and genetic variation of log-transformed traits tend to scale isometrically for several organisms (Holstad et al. 2024). Third, a high similarity between **P** matrices of distantly related species, as we found between *B. boans* and *L. mystaceus*, suggests high stability of trait covariation patterns over evolutionary time, compatible with stable **G** matrices (Phillips and Arnold 1999; Marroig and Cheverud 2001). Fourth, a single point estimate of the **G** matrix may be less predictive of divergence than a pooled **P** matrix that represents a long-term average of covariation patterns (Mongle et al. 2022). Finally, traits closely associated with functional performance are likely under selection to favor developmental architectures that do not disrupt functional integration, channeling environmental variation to match genetic patterns (Cheverud 1984).

Divergence Variance-Covariance Matrix and Alignment with Within-Species Covariation Patterns

We estimated the divergence morphological matrix (**D** matrix) for each genus accounting for phylogeny. We used species' means as the data and the anuran phylogeny of Portik et al. (2023). This phylogeny was estimated with phylogenomic analysis on more than 300 markers, is the most recent comprehensive phylogeny available, and has branch lengths in units of time. Most of our sampled species were in the phylogeny (see the supplemental PDF for taxon-substitution details). We then pruned this phylogeny to contain only our sampled species of *Boana* ($n = 33$ species) or *Leptodactylus* ($n = 27$ species).

We calculated the **D** matrix under a Brownian motion (BM) model of evolution (Lande 1979; Revell 2007; Simon et al. 2016; Machado 2020). The mean rate of evolution of each trait was calculated as the mean square of phylogenetic independent contrasts (PICs; Felsenstein 1985), whereas the mean cross product of the PICs for two traits was proportional to their mean rate of coevolution (Revell et al. 2007). We used the function ratematrix in the R package GEIGER (ver. 2.0.11; Harmon et al. 2008) to estimate the **D** matrix for each genus. We also calculated a **D** matrix assuming a single-peak Ornstein-Uhlenbeck (OU) model of evolution. **D** matrices assuming BM and OU were highly similar to each other within and across both genera (for details, see the supplemental PDF). Thus, we chose to use **D** matrices under BM because it is the simplest model.

We conducted an eigenanalysis to examine patterns of most divergence in trait combinations and also to test whether directions of most divergence were aligned with directions of most intraspecific variation. We compared the **D** matrices with pooled **P** matrices, but we also compared them with the three within-species **P** matrices. We used the KrzCor method to test for similarity between **P** and **D** matrices using four pairs of eigenvectors; we excluded the last two eigenvectors because KrzCor will give a similarity of 1 if all eigenvectors are included (see the supplemental PDF). Additionally, to check specifically for correlations between the first two eigenvectors of the **D** matrix (which corresponded to 98%–99% of the variation across species) with the first two eigenvectors of within-species **P** matrices and the pooled **P** matrix, we also quantified their alignment using vector correlations (the absolute value of the dot product between two normalized vectors). We calculated such correlations separately for each genus. To account for estimation uncertainty for both matrices in a given genus-specific comparison, we constructed 95% confidence intervals (CIs) for the vector correlations using bootstrapping resampling (for details, see the supplemental PDF). Given that vector correlations may be high by chance, we compared these 95% CIs with a null distribution of vector correlations constructed with 1,000 unit-length random six-element vectors drawn from a normal distribution of mean zero and unit SD. We considered observed vector correlations as significant when the 95% CIs were above the upper 5% tail of the null distribution (above 0.76). We performed all analyses with ln-transformed traits to guarantee that covariances were scale invariant and mean independent (Bookstein et al. 1985; McGuigan et al. 2005).

Random Drift Test

To assess which evolutionary process most likely caused divergence in hindlimb traits, we first tested whether species divergence was consistent with genetic drift (Lande 1976, 1979; Lovsfold 1988). More variable traits within an ancestral population evolving by genetic drift will lead to more divergence in species' means in those same traits compared with less variable traits. Thus, modeling the divergence of population means under genetic drift is a function of the ancestral **G** matrix, the effective population size (N_e), and the time since divergence (t; Lande 1979; Lovsfold 1988):

$$\mathbf{D} = \mathbf{G}\left(\frac{t}{N_e}\right). \quad (1)$$

Given that t and N_e are constants, we expect the **D** matrix to be proportional to the **G** matrix under drift. For morphological data, it is reasonable to assume that the **G** matrix can be represented by the pooled **P** matrix (Ackermann and Cheverud 2002), so we tested this proportionality expectation using the variance in the direction of eigenvectors of the pooled **P** matrix and the variation in species' scores projected onto the same eigenvectors (representing the **D** matrix). We also performed drift tests using within-species **P** matrices instead of the pooled **P** matrix. On a log scale, we may then test for proportionality of within- and between-species variation by calculating the regression slope (β) that relates their variances along the eigenvectors of the pooled **P** matrix or within-species **P** matrix (Ackermann and Cheverud 2002). Under drift, we expect $\beta = 1$. In contrast, if the slope significantly deviates from 1, we reject drift and infer that some form of selection would have to have either increased (e.g., diversifying) or reduced (e.g., stabilizing) divergence relative to the expectation under drift (Arnold et al. 2001). We took phylogeny into account by using PICs of the species means, which we projected onto the eigenvectors of the pooled **P** matrix or of the species-specific **P** matrix. For further details on this method, see the supplemental PDF.

Major Axes of the Macroevolutionary Adaptive Landscape

Even if within- and between-species patterns were aligned and random drift were rejected, the divergence across species could still reflect two distinct scenarios: genetic constraints biasing the response to selection (constraint hypothesis) or selection shaping both trait covariance patterns and differences in species' means, aligning the **P** matrix with the AL (selection hypothesis). To infer genetic constraints, the first two eigenvectors of the pooled **P** matrix ($\mathbf{p}_{max}$ and $\mathbf{p}_2$), representing evolutionary lines of least resistance, must not align with the major axes of the AL, representing the directions of most movement of adaptive peaks. Alternatively, if both the pooled **P** matrix and the **D** matrix were aligned with the AL, it would indicate that trait divergence patterns were produced by the covariance of movement of morphological adaptive peaks. If we found support for the latter scenario, the major axes of the AL could in principle coincide with SLLR. However, we could not test this hypothesis because we would need many more individuals within species (at least $n = 500$) than our highest sample size ($n = 53$) to properly estimate a populational $\boldsymbol{\gamma}$ matrix and extract its last eigenvectors to represent SLLR (Simon et al. 2022).

To estimate the major axes of the AL, we first calculated the selection covariance matrix (the **W** matrix), the covariance matrix of directional net selection gradients, reflecting the covariance of selective pressures acting on multiple traits and producing covariances in trait changes across species (Felsenstein 1988; Marroig and Cheverud 2010; Machado 2020):

$$\mathbf{W} = \mathbf{G}^{-1}\mathbf{D}\,\mathbf{G}^{-1} \quad (2)$$

in which $\mathbf{G}^{-1}$ is the inverse of the **G** matrix and **D** is the divergence matrix, the variance-covariance matrix of evolutionary changes in species' means. Notice that we use the same matrix parameters as the drift test described above but now assume a multivariate directional selection model interacting with genetic constraints and no stabilizing selection (model II in Zeng 1988). This model, $\mathbf{D} = \mathbf{G}\mathbf{W}\mathbf{G}$, describes how independent populations can show correlated evolutionary changes in the **D** matrix due to shared constraints (represented by the **G** matrix), similar directional selection pressures (represented by the **W** matrix), or both (Felsenstein 1988). It is important to realize that the **D** matrix calculated with PICs is equivalent to the average cross products of $\Delta\mathbf{z}$ (response to selection vectors; Lande 1979) from ancestor-descendent lineages, and thus the **W** matrix is a covariance matrix of net selection gradients (i.e., minimum amount of directional selection required to produce differences in species' means; Felsenstein 1988; Arnold et al. 2001; Machado 2020). Therefore, the eigenvectors of the **W** matrix represent the major axes of the realized AL, the main directions of shifts in adaptive peaks. To estimate the **W** matrix, we used the pooled **P** matrix as a surrogate of the **G** matrix, and the **D** matrix representing between-species patterns, for each frog genus. We scaled both **G** and **D** with their respective traces (the sum of their diagonals; Bookstein and Mitteroecker 2014) to obtain reasonable magnitudes of variation and covariation of net selection gradients. This procedure does not change the eigenvectors of the **W** matrix.

We then performed an eigendecomposition of the **W** matrix to obtain the major axes of the AL ($\mathbf{w}_{max}$ and $\mathbf{w}_2$), then correlated them with the eigenvectors of the pooled **P** matrix and of the **D** matrix. Once again, to account for uncertainty in estimating these matrices, we used the 1,000 bootstrapped pooled **P** matrices and **D** matrices to calculate resampled **W** matrices and compute the 95% CI for the correlations. A vector correlation was considered significant if the 95% CI was above the 5% upper tail of the null distribution of six-element vector correlations (0.76). However, we interpreted these vector correlations as a simple quantification of the match among directions of the AL, divergence, and constraints, rather than strict statistical tests.

Selection on Jumping Performance

To infer whether selection on jumping performance relates to the hindlimb divergence across species, we first tested whether jumping performance showed evidence of adaptive evolution. We aimed to test whether performance showed an optimal state predicted by species variation in the axis of most divergence across both *Leptodactylus* and *Boana* (i.e., the first axis of the **D** matrices, which we inferred to reflect size variation; see "Results"). In this case, performance would be evolving under an OU process, in which adaptation is not necessarily instantaneous. In contrast, the predictor—here, species means projected on $\mathbf{d}_{max}$—would be evolving under a BM model (the OUBM model; Hansen et al. 2008). We can interpret the OU process as describing movement of the adaptive optima of jumping performance over evolutionary time in response to changes in morphology, while accounting for constraints and lag in adaptation (Hansen et al. 2008; Hansen 2012).

We compared evolutionary models separately for each performance variable using the package slouch (ver. 2.1.4; Hansen et al. 2008), which allows OU models with continuous predictors. We tested models with $\mathbf{d}_{max}$ as the predictor, assuming either a BM model with a trend or an OUBM model. We compared these models to an intercept-only BM, an OU model with a single regime for all species, and an OU model with two regimes (arboreal or terrestrial microhabitat). For the latter, we set internal ancestral states for model fitting by using maximum likelihood and an equal-rates model, which showed the best fit to the data, using the function ace (ape R package, ver. 5.7-1; Paradis and Schliep 2019). We also performed a multivariate OU analysis with both performance variables as response variables. However, our results seemed unreliable (see the supplemental PDF), so for our data we favor the single-variable models described above.

Prior to analysis, we logged both performance variables and the predictor because we tested for an adaptive allometric relationship among them. To account for sampling error when estimating the regression slopes, we used squared standard errors of both jumping performance variables and species scores on $\mathbf{d}_{max}$, following Ives et al. (2007, app. 3). We then used the Akaike information criterion corrected for small sample sizes (AICc) for model comparison, using AICc weights to check the relative support for each model. To be conservative, we interpreted parameters for all models that showed AICc weights above 0.2, indicating modest to high support (Burnham and Anderson 2004).

Functional Performance Gradients

To further test the relationship between selection on jumping performance and both the pattern of trait covariation and the realized AL, we estimated jumping performance gradients and their correlation with $\mathbf{p}_{max}$ and $\mathbf{p}_2$ and with $\mathbf{w}_{max}$ and $\mathbf{w}_2$, respectively. Linear performance gradients are partial linear regression coefficients that reflect the effects of morphology on performance and can be interpreted as part of the total selection on morphology (Arnold 1983). To estimate performance gradients within

species, we fit multiple linear regressions of maximum acceleration and velocity on the hindlimb traits, for both *Boana boans* and *Leptodactylus mystaceus*. After calculating the jumping performance gradients, we normalized them to a length of 1. We then tested their vector correlation with $\mathbf{w}_{max}$ and $\mathbf{w}_2$ and with $\mathbf{p}_{max}$ and $\mathbf{p}_2$. We performed the same analysis for between-species performance gradients, which quantified the hindlimb traits that most contributed to differences in jumping across species (see the supplemental PDF). Comparing between-species performance gradients with $\mathbf{w}_{max}$ and $\mathbf{w}_2$ may be more relevant (than within-species analysis) to test whether shifts in morphological adaptive peaks relate to species' differences in jumping performance. We calculated the 95% CI for these vector correlations using the resampled distributions of the pooled **P** matrix, divergence matrix, and **W** matrix, from which we extracted eigenvectors.

We first conducted these analyses using all six hindlimb traits for morphology. However, such gradients may have been poorly estimated with the sample sizes we had within species, and they were hard to interpret functionally. Thus, we also estimated performance gradients using just two morphological traits: leg length (the sum of the four bone lengths) and the square root of the total muscle CSA (the sum of the two muscle measurements). This simpler analysis made performance gradients more reliable and easier to interpret while still retaining the bone length by muscle CSA contrast found as relevant in the species morphological divergence (see "Results"). We also estimated pooled **P**, divergence, and **W** matrices with just leg length and muscle CSA to compare with the simplified performance gradients.

Results

Patterns of Trait Covariation Are Shared within and between Species

The patterns of trait covariation were similar across all within-species **P** matrices and, accordingly, also between pooled **P** matrices for each genus (table S2). More importantly, trait covariation patterns represented by either pooled **P** matrices or within-species **P** matrices were moderately to highly similar to the patterns of divergence matrices (KrzCor ranged from 0.74 to 0.90; table S4). The eigenvector correlations between the two first eigenvectors of pooled **P** and **D** matrices were very high, ranging from 0.92 to 0.99 (table S4; fig. S3), whereas correlations were particularly high only for the first eigenvector when using within-species **P** matrices (table S4). All empirical vector correlations were significant when accounting for uncertainty in estimating the pooled **P** and **D** matrices, while the same was true for only the first eigenvector when accounting for uncertainty in estimating species **P** matrices (table S4).

The first eigenvector (Egv1), which corresponds to $\mathbf{p}_{max}$, can be interpreted as a size vector: all coefficients had the same sign, indicating that all traits increase or decrease together in this dimension (Jolicoeur 1963). A size vector can be either isometric, in which increases in body size do not change shape, or allometric, in which shape changes with size. To test whether Egv1 was isometric, we divided each of its coefficients by the isometric expectation of 0.408 (i.e., $1/\sqrt{6}$, the expected coefficients for a six-element vector; Jolicoeur 1963). The scaled coefficients for bone lengths were significantly lower than 1 (except two bones in divergence matrices), whereas those for the square root of muscle CSA, especially from the upper leg, were significantly higher than 1 (table S5). This result indicated positive allometry for muscle CSA (stronger within species) and negative allometry for leg length (fig. 2). Therefore, we conclude that Egv1 corresponds to allometric size, the relative contribution of each variable to absolute size (Marroig and Cheverud 2005).

Variation along the second eigenvector (Egv2) represented a contrast between bone lengths and muscle CSA (table S5). This axis can be interpreted as variation in hindlimb shape unassociated with size changes, in which individuals and species with higher muscle CSA at a given body size tended to have relatively shorter legs compared with individuals and species with lower CSA that tended to have longer legs (for that same body size; fig. 2).

Random Drift Is Insufficient for Explaining Divergence in Hindlimb Traits

Regression slopes of the observed morphological divergence on within-species variation were above 1 for both genera, whether using pooled **P** matrices or within-species **P** matrices as representing ancestral patterns (except when using *Boana boans*; see table 1). However, only for *Leptodactylus* did the 95% CI of the empirical slopes exclude the 95% CI under BM simulations, which assumed within-species variance-covariance patterns for the rate parameters (table 1; fig. S4). For *Boana*, the empirical slopes were generally above the 95% CI under BM, but the 95% CI of empirical slopes were broader than in *Leptodactylus*. Thus, we could confidently reject random drift as driving species divergence for *Leptodactylus*, but we have less support for rejecting drift in *Boana*. Rejection of drift was mainly driven by higher divergence along Egv1 than expected based on within-species variation in this dimension (table 1).

*Pooled **P** and Divergence Matrices Are Not Aligned with the Major Axes of the Adaptive Landscape*

The major axes of the AL, represented by the first two eigenvectors of the **W** matrix, were not shared between

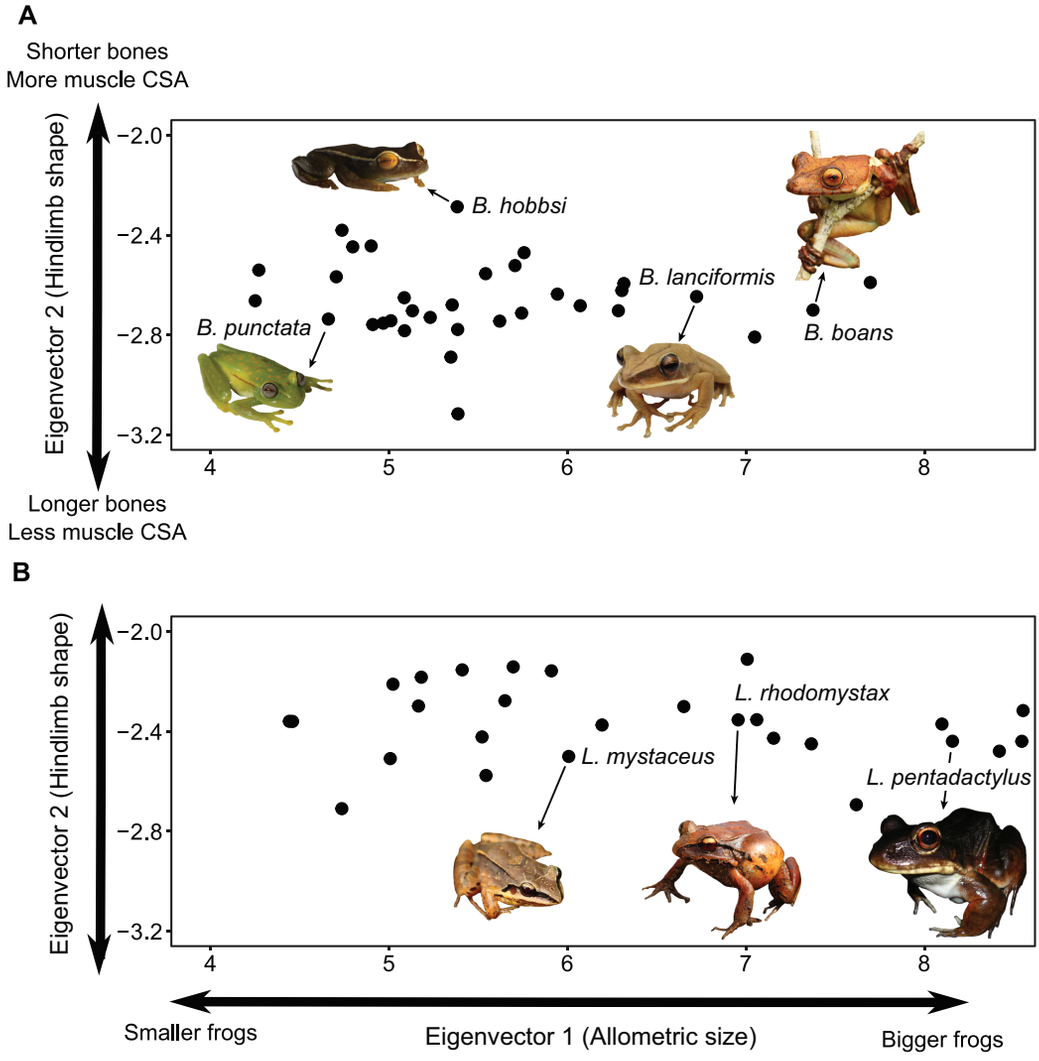


Figure 2: Species distributions along axes of most phenotypic divergence for *Boana* (*A*) and *Leptodactylus* (*B*). The first axis (eigenvector 1) corresponds to $\mathbf{d}_{max}$ and represents variation in allometric size, which explains 95% and 97% of divergence across species in *Boana* and *Leptodactylus*, respectively. The second axis (eigenvector 2) corresponds to $\mathbf{d}_2$ and represents hindlimb shape, as a contrast between bone length and muscle cross-sectional areas (CSA). $\mathbf{d}_2$ explains only 3% and 2% of divergence across species in *Boana* and *Leptodactylus*, respectively. These divergence axes are similar to axes of most variation within species. Arrows point from circles of species in photos, which are not scaled to body size. All photos by D. S. Moen.

Boana and *Leptodactylus* (table S6). In general, these axes favored opposite changes in lengths of specific bones in both genera, and directional net selection has favored opposite changes in bone and muscle CSA only in *Leptodactylus* (table S6).

Neither the axes of most variation within species ($\mathbf{p}_{max}$ and $\mathbf{p}_2$) nor the axes of most divergence across species ($\mathbf{d}_{max}$ and $\mathbf{d}_2$) showed significant vector correlations with the major axes of the AL (table S7). This result also held for our simplified analysis using just leg length and muscle CSA (table 2). Whereas both pooled **P** and divergence matrices showed a significantly high positive correlation between leg length and muscle CSA, the **W** matrices showed a very low or negative correlation between these two traits. However, this correlation was significant only in *Leptodactylus* (table 2).

Evolution of Maximum Jumping Acceleration, but Not Velocity, Tracks Allometric Size

The model with highest support for maximum acceleration was a BM regression model (table 3) showing a significant negative relationship between the evolution of maximum acceleration and evolutionary changes in allometric size

Table 1: Random drift tests for hindlimb traits in the two frog genera

	Empirical slope	Empirical 95% CI	BM simulations 95% CI
Drift test:			
Pooled **P**:			
Boana	*1.35*	.77 to 1.92	.87 to 1.13
Leptodactylus	**1.44**	1.14 to 1.73	.85 to 1.14
Within-species **P**:			
Boana (*B. boans*)	1.00	.69 to 1.32	.90 to 1.10
Boana (*B. raniceps*)	*1.30*	.75 to 1.85	.85 to 1.15
Leptodactylus	*1.41*	.89 to 1.93	.86 to 1.15
Drift test excluding $\mathbf{p}_{max}$:			
Pooled **P**:			
Boana	.94	.21 to 1.67	.78 to 1.21
Leptodactylus	*1.25*	.64 to 1.86	.71 to 1.24
Within-species **P**:			
Boana (*B. boans*)	.72	.48 to .96	.83 to 1.18
Boana (*B. raniceps*)	1.03	.05 to 2.01	.76 to 1.23
Leptodactylus	*1.21*	.00 to 2.43	.72 to 1.15

Note: We performed a regression test of between-species divergence on within-species variation. We strongly rejected drift (values in boldface) when the empirical slope 95% confidence interval (CI) was outside the 95% CI of Brownian motion (BM) simulations that used the corresponding **P** matrices as the rate matrix. Values in italics indicate empirical slopes that were higher than the 95% CI of BM simulations. We also performed the test when excluding variation in the first eigenvector ($\mathbf{p}_{max}$) to show that drift was rejected in the overall (two-axis) drift test mostly because of higher-than-expected variation along $\mathbf{p}_{max}$, rather than both axes.

($\mathbf{d}_{max}$; fig. 3*A*). This negative relationship indicated that smaller species, with less proportional muscle CSA, produced higher maximum accelerations than bigger species. An OU regression model with no differences between genera/microhabitat (OU1BM) had modest support for the evolution of maximum acceleration (table 3), also showing a significant negative evolutionary relationship between maximum acceleration and allometric size.

In contrast, the evolution of jumping peak velocity was unrelated to $\mathbf{d}_{max}$ (table 3; fig. 3*B*). An intercept-only BM model showed the most support. A two-regime OU model also showed modest support, suggesting a different optimum

Table 2: Pooled **P**, divergence, and **W** matrices when using total leg length and muscle cross-sectional areas (CSA)

	Leg length	Muscle CSA	Leg length	Muscle CSA
	Pooled *Boana*		Pooled *Leptodactylus*	
Pooled **P**:				
Leg length	**.004** (.0002 to .006)	**.660** (.602 to .788)	**.008** (.005 to .0086)	**.758** (.626 to .830)
Muscle CSA	**.004** (.0026 to .0065)	**.010** (.006 to .012)	**.010** (.0055 to .010)	**.021** (.013 to .021)
	Boana		*Leptodactylus*	
Divergence:				
Leg length	**.004** (.0023 to .0059)	**.939** (.884 to .973)	**.005** (.0028 to .0086)	**.957** (.911 to .980)
Muscle CSA	**.004** (.0024 to .0069)	**.006** (.0033 to .0089)	**.006** (.0033 to .010)	**.008** (.0044 to .014)
W matrix:				
Leg length	**3.565** (2.06 to 4.40)	.025 (−.18 to .50)	**5.685** (3.83 to 9.13)	**−.383** (−.72 to −.13)
Muscle CSA	.036 (−.32 to .52)	**.557** (.33 to .92)	**−.677** (−2.40 to −.17)	**.550** (.32 to 1.45)

Note: Pooled **P** matrices show within-species trait patterns, whereas divergence matrices show between-species patterns calculated using phylogenetic independent contrasts (i.e., standardized to time). **W** matrices show patterns of coselection between the traits. Diagonal elements are variances, below-diagonal elements are covariances, and above-diagonal elements are correlations. Values in parentheses are 95% confidence intervals (CIs) from bootstrapped matrices. Boldface values are significantly higher or lower than zero according to the 95% CI.

Table 3: Comparison of Brownian motion (BM) and Ornstein-Uhlenbeck (OU) models for maximum jumping acceleration and velocity

Univariate models	*k*	loglik	AICc	AICc weight
Max acceleration:				
BM	2	−3.79	12.40	.02
OU1	3	−3.35	14.40	.01
OU2	4	−3.17	17.40	.00
BM regression[a]	**3**	**1.33**	**5.05**	**.69**
OU1BM[b]	**4**	**1.99**	**7.10**	**.25**
OU2BM	5	1.99	11.00	.03
Max velocity:				
BM[c]	**2**	**6.08**	**−7.35**	**.51**
OU1	3	6.10	−4.48	.12
OU2[d]	**4**	**8.29**	**−5.50**	**.20**
BM regression	3	6.08	−4.44	.12
OU1BM	4	6.10	−1.12	.02
OU2BM	5	8.29	−1.57	.03

Note: Intercept-only models, BM and OU1, have no predictors. Regression models have species' scores on the axis of most divergence—allometric size—as the predictor. The OU regression models can be single- (OU1BM) or two-optima (OU2BM: arboreal or terrestrial) models, which correspond to one or two intercepts, respectively. Models in boldface show AICc weights more than 0.2. *k* = number of parameters; loglik = log likelihood of the model; AICc = Akaike information criterion corrected for small sample sizes; AICc weight = relative support for each model. Lettered footnotes show parameter estimates of the most supported models (maximum likelihood estimates with 95% confidence interval).

[a] $\sigma^2 = 0.046$ (0.011 to 0.15), slope = −0.21 (−0.32 to −0.10).

[b] Half-life = 21% of the tree length (0% to 199%), slope = −0.22 (−0.32 to −0.10), stationary variance = 0.032 (0.02 to 0.99).

[c] $\sigma^2 = 0.037$ (0.016 to 0.097).

[d] Half-life = 14% of the tree length (0.001% to 500%); optima: arboreal = 1.10 m/s (1.00 to 1.21); terrestrial = 0.83 m/s (0.71 to 0.93); stationary variance = 0.016 (0.01 to 0.21).

jumping velocity for arboreal (*Boana*) and terrestrial species (*Leptodactylus*; table 3).

Jumping Performance Gradients Are Aligned with Neither Constraints nor the AL

Very few performance gradients showed significant effects of hindlimb traits on jumping performance when considering all six traits (table S8). Hindlimb traits explained 22% of the variation in maximum acceleration in *B. boans* and 16% of the variation in maximum velocity in *Leptodactylus mystaceus*. In contrast, when analyzing just leg length and muscle CSA, we found more significant relationships (table S9). Reducing leg length increased maximum acceleration in *B. boans* and across species of *Leptodactylus* (fig. S5). Increasing muscle CSA increased maximum acceleration and velocity in *L. mystaceus*, but we found the opposite pattern across species of *Leptodactylus* (fig. S5). Although not all performance gradients were significant, several showed a pattern of opposite signs between leg length and muscle CSA. This result agrees with the low or negative correlation found between these two traits in the **W** matrix (table 2).

The full multivariate performance gradients had very low vector correlations with both $\mathbf{p}_{max}$ and $\mathbf{p}_2$ (table S10). Some performance gradients showed empirical correlations above the null expectation with the first major axis of the AL (*B. boans* for maximum acceleration) or the second major axis of the AL (*L. mystaceus* for maximum velocity). However, the 95% CIs were large, making these correlations nonsignificant (table S10).

Discussion

We found an overall pattern of hindlimb trait covariation that was highly similar within and between species in both *Boana* and *Leptodactylus* (table S4), despite a deep divergence around 60 million years ago between the genera (fig. 1). Our study adds to the mounting evidence that variational patterns within species, reflected by the **P** or **G** matrix, influence species divergence over long evolutionary timescales (Marroig and Cheverud 2005, 2010; Renaud et al. 2006; Simon et al. 2016; Houle et al. 2017; McGlothlin et al. 2018; Machado 2020, Mongle et al. 2022; Opedal et al. 2023; Machado et al. 2023). However, understanding which factors have shaped both the major axes of the AL and those of the **G** matrix is challenging. By including functional performance in our study, we were able to infer that the realized AL was not aligned with hindlimb trait correlations, and directional net selection seemed unaligned with selection associated with jumping performance. Yet our inference that the adaptive peak of jumping performance randomly drifted across species, coupled with stability of the **G** matrix, may explain the persistent effect of patterns of within-species hindlimb trait covariation on divergence.

Stability of Genetic Architecture of Hindlimb Traits in Both Frog Genera

We found high similarity among all **P** matrices across both genera (tables S2, S3). This consistency suggests that the **G** matrices of all these frog species have remained generally stable over long evolutionary timescales, at least in terms of their major axes of variation. This overall stability of trait covariation seems to be common in nature and has been found in the morphology of many species that have diverged for over tens of millions of years (e.g., Marroig and Cheverud 2001; Porto et al. 2009; Kolbe et al. 2011; Simon et al. 2016; Machado et al. 2018; McGlothlin et al. 2018, 2022).

Such stability of trait correlations may be maintained by selective, functional, or developmental constraints (Arnold

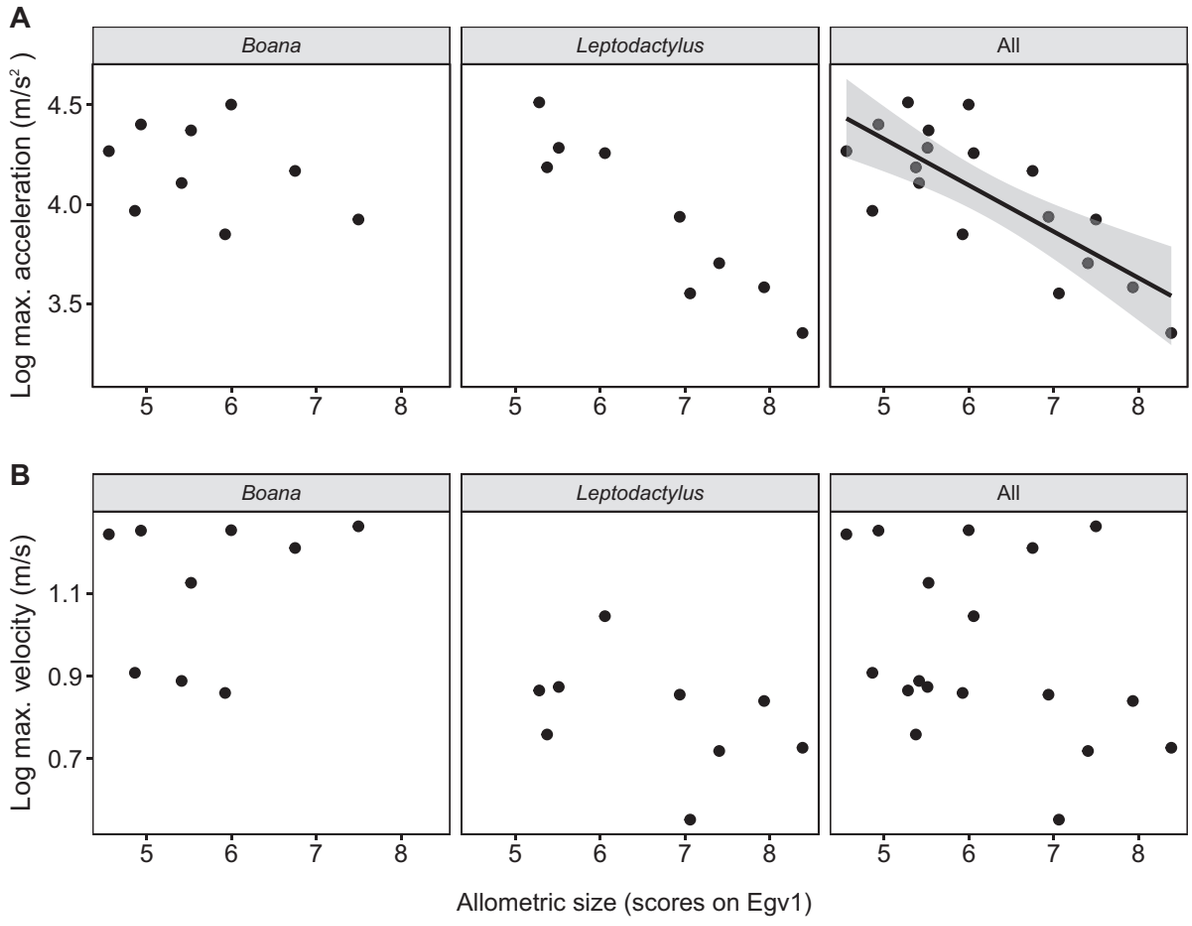


Figure 3: Scatterplots of maximum jumping acceleration (*A*) and velocity (*B*) versus allometric size. We ran different evolutionary models in the R package slouch, assuming a Brownian motion (BM) or an Ornstein-Uhlenbeck model for jumping performance. While the BM regression was the best-supported model for maximum acceleration (table 3; $\sigma^2 = 0.046 \pm 0.010$ [maximum likelihood estimate $\pm$ SE], slope $= -0.21 \pm 0.05$, $R^2 = 0.48$), an intercept-only BM model showed the highest support for maximum velocity (table 3; $\sigma^2 = 0.037 \pm 0.010$). Egv1 is the first eigenvector of the divergence matrix, corresponding to $\mathbf{d}_{max}$, the direction of most divergence across species.

1992). Because we did not find alignment between within-species lines of least resistance and the AL (table S8), stability of hindlimb trait correlations could have been maintained by a stable pattern of correlational selection (Riedl 1977; Cheverud 1984; Brodie 1992; Arnold 2005; McGlothlin et al. 2005; Roff and Fairbairn 2012)—promoting functional integration of trait complexes (Olson and Miller 1958; Hohenlohe and Arnold 2008; Simon et al. 2019)—instead of persistent directional selection acting on mutational effects (Jones et al. 2007, 2014; Pavlicev et al. 2011). We may thus expect correlational selection on hindlimb traits in frogs because unbalanced trait values (e.g., greatly differing leg bone lengths) may hinder locomotion in general, including jumping. Therefore, the positive correlation between leg length and leg muscle CSA could reflect their coordinated effect on locomotor performance favored by correlational selection. We do emphasize, however, that our estimates of the AL only take net directional selection into account and do not estimate the curvature of the AL (i.e., the **W** matrix does not take stabilizing and correlational selection into account). Therefore, the misalignment between the AL and trait correlations does not rule out a role of correlational selection promoting stability of hindlimb trait correlations over evolutionary time.

However, because we found allometric size to be the major axis of variation, the stability of trait correlations likely also reflects, at least in part, developmental constraints. Variation in limb growth rates is regulated by several growth factors related to chondrogenesis and bone formation (Hentschel et al. 2004), and also to chondrocyte proliferation and differentiation (Wang et al. 2004; Cooper et al. 2013). The expression of growth factors is further regulated by several genes that can be expressed with different patterns of timing and spatial positioning (Towers and Tickle 2009).

Hence, genetic and developmental variation regulating overall limb growth seems abundant and highly polygenic, and this variation may explain why individuals vary most in allometric size. Allometric size often acts as a line of least evolutionary resistance in several groups, along which species predominantly diverge (Schluter 1996; Marroig and Cheverud 2005, 2010; Simon et al. 2016; Machado 2020; Feiner et al. 2021; Tejero-Cicuéndez et al. 2023). This may be due to the ecological relevance of size—in diet, for example (Moen and Wiens 2009; Marroig and Cheverud 2010; Machado 2020)—but also to conserved developmental regulation of growth patterns, creating developmental lines of least resistance (Salazar-Ciudad 2006). Although developmental lines of least resistance can change over time (Feiner et al. 2021), they likely evolve slowly (Voje et al. 2014; Houle et al. 2019), which fits our findings.

Divergence in Allometric Size and Selection on Jumping Performance

Our finding that both the **G** matrix and divergence were unaligned with the AL (tables 2, S7) indicates that the high interspecific divergence in allometric size likely reflects genetic and developmental constraints biasing the response to net selection instead of diversifying selection on allometric size. This interpretation favors the hypothesis that constraint has shaped hindlimb divergence in both genera. In contrast, direct selection on allometric size may not necessarily have shaped hindlimb divergence, although indirect selection on allometric size likely happened because of positive correlations within the hindlimbs.

Our interpretation that selection has influenced hindlimb divergence in *Boana* may be surprising given the lower support for rejecting drift in hindlimb divergence in *Boana* (table 1). Our weakest support resulted from using the *Boana boans* **P** matrix to represent within-species patterns. Yet part of this result may be explained by having fewer individuals of *B. boans*, which can result in overestimation of the first eigenvalue (Hayden and Twede 2002; Meyer and Kirkpatrick 2008). Consequently, higher intraspecific variance on the first eigenvector of the *B. boans* **P** matrix likely made it harder to reject drift. Consistent with this explanation, the empirical slope for the drift test was higher than the 95% CI for BM simulations when using both the pooled **P** matrix and the *Boana raniceps* **P** matrix, for which we had higher sampling than in *B. boans*. Moreover, the large 95% CI for the empirical regression slopes likely reflect a low number of traits (and, therefore, eigenvectors) to estimate the regression (six in total), which makes this test conservative. Hence, despite the lower support for rejecting drift in *Boana*, we also infer selection as the main driver of divergence in this genus.

Moreover, our analysis of performance gradients showed further evidence for selection on hindlimb traits (tables S8, S9; fig. S5). Nonetheless, the performance gradients were generally aligned with neither directional net selection nor hindlimb divergence (tables 2, S10). These results may indicate that variation in jumping performance was not a major contributor to differences in fitness across the species over evolutionary time. While seemingly counterintuitive for organisms that primarily move by jumping, this interpretation is consistent with previous work. For example, frogs showed only a single optimum for jumping performance despite diversity in limb shape at a broader evolutionary scale, suggesting many-to-one mapping of morphology onto performance (Moen 2019). However, we cannot discard poor estimation of the AL (**W** matrix) and of within-species performance gradients as contributing factors to this misalignment. For the estimation of the **W** matrix, a potential issue is the necessity of inverting the pooled **P** matrix. Such a procedure makes the trailing eigenvalues, which are poorly estimated, dominate the inverted matrix (Hayden and Twede 2002). We tried to account for this issue by using an eigenvalue extension procedure (replacing the sixth eigenvalue of the pooled **P** matrix by the fifth, which is better estimated) on the pooled **P** matrix, following Marroig et al. (2012), to better estimate the **W** matrices (although this procedure is recommended for matrices with 10 or more traits). Yet the major axes of the AL were still unaligned with within-species variation and with performance gradients (results shown in "R script 1.R" in our Dryad repository; Simon et al. 2025). Moreover, other unmeasured traits may be important for jumping, potentially obscuring the relationships between performance gradients and the AL.

Although our evolutionary analyses of maximum jumping velocity and acceleration across species in *Boana* and *Leptodactylus* suggest that jumping performance evolved randomly (table 3), this pattern can alternatively be interpreted as jumping performance evolving by directional selection to track an adaptive peak that itself has moved randomly (Felsenstein 1988; Hansen and Martins 1996; Arnold et al. 2001; Hansen 2012; Holstad et al. 2024; Uyeda and McGlothlin 2024). We favor this adaptive explanation for the evolution of jumping performance because of its relevance in frog ecology (Gans and Parsons 1966; Emerson 1978; Duellman and Trueb 1994; Marsh 1994; Moen 2019). Adaptive peaks may move in a random fashion when the environment fluctuates, and these fluctuations in the direction of selection may explain a persistent effect of genetic variance on divergence if species mostly track the peaks along genetic lines of least resistance (Holstad et al. 2024). Such fluctuations in the jumping performance peak may relate to changes in the balance of functional trade-offs (Garland et al. 2022) on the hindlimbs, which also influence other locomotor behaviors (e.g., swimming, climbing, burrowing, walking; Moen et al. 2013; Reynaga et al. 2018; Moen 2019; Mendoza et al. 2020).

Evolution of Maximum Jumping Acceleration and Allometric Size

The negative relationship between the evolution of maximum acceleration and allometric size suggests a functional constraint associated with size scaling. Generally, larger animals have larger muscles, which produce higher forces. These forces depend on muscle CSA, which scales to the second power of length. However, such larger animals have proportionally larger masses to move with those muscles because mass scales to the third power of length. Overall, this scaling could result in lower whole-body acceleration in larger animals. One potential consequence is the positive allometry for muscle CSA that we found within and across species (table S5). If larger animals have proportionally more muscle CSA for their size, they may be able to partially offset the reduced jumping acceleration due to muscle CSA–body mass scaling.

However, the pattern of smaller species showing higher peak jumping acceleration than larger species may also be interpreted as an evolutionary innovation in response to size evolution in frogs. Smaller species may more frequently use power amplification by elastic energy storage to achieve higher accelerations, despite having lower absolute muscle mass, than larger species (Astley 2016; Ilton et al. 2018; Sutton et al. 2019; Mendoza et al. 2020). Those smaller animals, with shorter legs and thus contact times, need higher acceleration than larger animals to achieve a comparable takeoff velocity. Yet such higher acceleration is limited by both their reduced muscle CSA and the reduced force produced by fast muscular contractions (which are necessary if the contact time is very short in small animals; Bennet-Clark 1977; Vogel 2009). On the other hand, such small species may be more likely to release elastic energy stored in hindlimb connective tissue (Sutton et al. 2019) and do so more effectively (Mendoza and Azizi 2021). Hence, species divergence being restricted mostly to allometric size may have strongly affected the evolution of jumping acceleration, as an evolutionary solution in smaller species to a functional constraint of muscle function and leg length. This evolutionary innovation may also explain why the performance gradients in *Leptodactylus mystaceus* showed an opposite pattern than within the genus *Leptodactylus* (fig. S5). While bigger muscles produce more force to increase acceleration and velocity within species (Vogel 2009; Biewener and Patek 2018), this is not the pattern seen across species when smaller species may be more likely and more effective in releasing elastic energy to achieve higher locomotor accelerations and velocity (Sutton et al. 2019; Mendoza and Azizi 2021).

Conclusions

We found support for the constraint hypothesis in shaping long-term divergence in hindlimb traits on frogs, with a potentially strong effect of conserved developmental constraints, channeling divergence mostly to allometric size. This pattern is consistent with development having a primary role in aligning micro- and macroevolution, especially in a view of a dynamic AL. In this view, fluctuating selection on functional performance can reinforce a link between these scales over larger timescales by making species phenotypic means chase these adaptive peak fluctuations mostly along genetic lines of least resistance. However, developmental constraints can be broken by episodes of strong selection, which may occur when lineages invade a new environment that demands great changes in specific behaviors, such as locomotion (i.e., a new adaptive zone). Therefore, studying anuran species that differ more strongly in ecology and performance, such as aquatic and burrowing species, may favor a pure-selection hypothesis and a role of selection on performance in shaping both the genetic architecture of traits and species' phenotypic means. Overall, it is likely that both development and function play a role in long-term phenotypic divergence. However, knowing how often each of these factors dominate macroevolutionary patterns demands directly studying the macroevolutionary AL and its connection with both developmental and functional processes acting within populations and species.

Acknowledgments

We thank Anne-Claire Fabre for help in collecting frogs in 2015, Menelia Vasilopoulou-Kampitsi for help in early analysis of force traces, and Craig McGowan for assistance in smoothing force traces. Antoine Fouquet contributed a key specimen for performance testing and helped with identification of specimens and systematics of anurans from French Guiana. Finally, for loans of museum specimens, we thank Jeremy Jacobs, Kevin de Queiroz, and Rob Wilson of the Division of Amphibians and Reptiles at the Smithsonian Institution; Allison Whiting from Brigham Young University; and Alan Resetar and Joshua Mata from the Field Museum. We also thank Editor-in-Chief Volker Rudolf and Associate Editor Joel McGlothlin for great suggestions on the manuscript, as well as two anonymous reviewers who helped improve our work.

Statement of Authorship

Conceptualization: D.S.M., A.H., M.N.S.; funding acquisition: D.S.M, A.H., E.A.C.; experimental design: D.S.M., A.H.; data collection: E.A.C., D.S.M., A.H., M.N.S.; data analysis: M.N.S., D.S.M.; coding: M.N.S.; data validation and visualization: M.N.S.; writing—original draft: M.N.S., D.S.M.; writing—review and editing: M.N.S., D.S.M., A.H., E.A.C.

Data and Code Availability

The data associated with this work have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.rn8pk0pnp; Simon et al. 2025), and the code has been deposited in Zenodo (https://doi.org/10.5281/zenodo.14757462; also linked with the Dryad submission).

Literature Cited

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

Arnold, S. J. 1983. Morphology, performance and fitness. American Zoologist 23:347–361.

———. 1992. Constraints on phenotypic evolution. American Naturalist 140:S85–S107.

———. 2005. The ultimate causes of phenotypic integration: lost in translation. Evolution 59:2059–2061.

———. 2023. Evolutionary quantitative genetics. Oxford University Press, Oxford.

Arnold, S. J., R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones. 2008. Understanding the evolution and stability of the **G**-matrix. Evolution 62:2451–2461.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Pages 9–32 *in* A. P. Hendry and M. T. Kinnison, eds. Microevolution rate, pattern, process, contemporary issues in genetics and evolution. Springer Netherlands, Dordrecht.

Astley, H. C. 2016. The diversity and evolution of locomotor muscle properties in anurans. Journal of Experimental Biology 219:3163–3173.

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

Bennet-Clark, H. C. 1977. Scale effects in jumping animals. Pages 185–201 *in* T. J. Pedley, ed. Scale effects in animal locomotion. Academic Press, London.

Biewener, A., and S. Patek. 2018. Animal locomotion. Oxford University Press, Oxford.

Björklund, M., A. Husby, and L. Gustafsson. 2013. Rapid and unpredictable changes of the **G**-matrix in a natural bird population over 25 years. Journal of Evolutionary Biology 26:1–13.

Bolstad, G. H., T. F. Hansen, C. Pélabon, M. Falahati-Anbaran, R. Perez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. Philosophical Transactions of the Royal Society B 369:20130255.

Bookstein, F. L., B. C. Chernoff, R. L. Elder, J. M. Humphries, G. R. Smith, and R. E. Strauss. 1985. Morphometrics in evolutionary biology. Academy of Natural Sciences Philadelphia Special Publication 15:1–277.

Bookstein, F. L., and P. Mitteroecker. 2014. Comparing covariance matrices by relative eigenanalysis, with applications to organismal biology. Evolutionary Biology 41:336–350.

Brodie, E. D. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. Evolution 46:1284–1298.

Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods and Research 33:261–304.

Cano, J. M., A. Laurila, J. Palo, and J. Merilä. 2004. Population differentiation in **G** matrix structure due to natural selection in *Rana temporaria*. Evolution 58:2013–2020.

Careau, V., M. E. Wolak, P. A. Carter, and T. Garland Jr. 2015. Evolution of the additive genetic variance–covariance matrix under continuous directional selection on a complex behavioural phenotype. Proceedings of the Royal Society B 282:20151119.

Chantepie, S., A. Charmantier, B. Delahaie, F. Adriaensen, E. Matthysen, M. E. Visser, E. Álvarez, et al. 2024. Divergence in evolutionary potential of life history traits among wild populations is predicted by differences in climatic conditions. Evolution Letters 8:29–42.

Chenoweth, S. F., H. D. Rundle, and M. W. Blows. 2010. The contribution of selection and genetic constraints to phenotypic divergence. American Naturalist 175:186–196.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. Journal of Theoretical Biology 110:155–171.

Cheverud, J. M., and G. Marroig. 2007. Comparing covariance matrices: random skewers method compared to the common principal components model. Genetics and Molecular Biology 30:461–469.

Citadini, J. M., R. Brandt, C. R. Williams, and F. R. Gomes. 2018. Evolution of morphology and locomotor performance in anurans: relationships with microhabitat diversification. Journal of Evolutionary Biology 31:371–381.

Cooper, K. L., S. Oh, Y. Sung, R. R. Dasari, M. W. Kirschner, and C. J. Tabin. 2013. Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. Nature 495:375–378.

Duellman, W. E., and L. Trueb. 1994. Biology of amphibians. John Hopkins University Press, Baltimore.

Emerson, S. B. 1978. Allometry and jumping in frogs: helping the twain to meet. Evolution 32:551–564.

Eroukhmanoff, F., and E. I. Svensson. 2011. Evolution and stability of the **G**-matrix during the colonization of a novel environment. Journal of Evolutionary Biology 24:1363–1373.

Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. American Naturalist 169:227–244.

Feiner, N., I. S. Jackson, E. Van der Cruyssen, and T. Uller. 2021. A highly conserved ontogenetic limb allometry and its evolutionary significance in the adaptive radiation of *Anolis* lizards. Proceedings of the Royal Society B 288:20210226.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

———. 1988. Phylogenies and quantitative characters. Annual Review of Ecology and Systematics 19:445–471.

Gans, C., and T. S. Parsons. 1966. On the origin of the jumping mechanism in frogs. Evolution 20:92–99.

Garland, T., C. J. Downs, and A. R. Ives. 2022. Trade-offs (and constraints) in organismal biology. Physiological and Biochemical Zoology 95:82–112.

Gomes, F. R., E. L. Rezende, M. B. Grizante, and C. A. Navas. 2009. The evolution of jumping performance in anurans: morphological correlates and ecological implications. Journal of Evolutionary Biology 22:1088–1097.

Hansen, T. F. 2012. Adaptive landscapes and macroevolutionary dynamics. Pages 205–221 *in* E. Svensson and R. Calsbeek, eds. The adaptive landscape in evolutionary biology. Oxford University Press, Oxford.

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology 21:1201–1219.

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

Hansen, T. F., J. Pienaar, and S. H. Orzack. 2008. A comparative method for studying adaptation to a randomly evolving environment. Evolution 62:1965–1977.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hayden, A. F., and D. R. Twede. 2002. Observations on the relationship between eigenvalues, instrument noise, and detection performance. Proceedings of the SPIE 4816:355–362.

Hellam, D. C., and R. J. Podolsky. 1969. Force measurements in skinned muscle fibres. Journal of Physiology 200:807–819.

Hentschel, H. G. E., T. Glimm, J. A. Glazier, and S. A. Newman. 2004. Dynamical mechanisms for skeletal pattern formation in the vertebrate limb. Proceedings of the Royal Society B 271:1713–1722.

Hohenlohe, P. A., and S. J. Arnold. 2008. MIPoD: a hypothesis-testing framework for microevolutionary inference from patterns of divergence. American Naturalist 171:366–385.

Holstad, A., K. L. Voje, Ø. H. Opedal, G. H. Bolstad, S. Bourg, T. F. Hansen, and C. Pélabon. 2024. Evolvability predicts macroevolution under fluctuating selection. Science 384:688–693.

Houle, D., G. H. Bolstad, K. van der Linde, and T. F. Hansen. 2017. Mutation predicts 40 million years of fly wing evolution. Nature 548:447–450.

Houle, D., L. T. Jones, R. Fortune, and J. L. Sztepanacz. 2019. Why does allometry evolve so slowly? Integrative and Comparative Biology 59:1429–1440.

Hunt, G. 2007. Evolutionary divergence in directions of high phenotypic variance in the ostracode genus *Poseidonamicus*. Evolution 61:1560–1576.

Ilton, M., M. S. Bhamla, X. Ma, S. M. Cox, L. L. Fitchett, Y. Kim, J. Koh, et al. 2018. The principles of cascading power limits in small, fast biological and engineered systems. Science 360:eaao1082.

Ives, A. R., P. E. Midford, and T. Garland Jr. 2007. Within-species variation and measurement error in phylogenetic comparative methods. Systematic Biology 56:252–270.

James, R. S., C. A. Navas, and A. Herrel. 2007. How important are skeletal muscle mechanics in setting limits on jumping performance? Journal of Experimental Biology 210:923–933.

James, R. S., and R. S. Wilson. 2008. Explosive jumping: extreme morphological and physiological specializations of Australian rocket frogs (*Litoria nasuta*). Physiological and Biochemical Zoology 81:176–185.

James, R. S., R. S. Wilson, J. E. de Carvalho, T. Kohlsdorf, F. R. Gomes, and C. A. Navas. 2005. Interindividual differences in leg muscle mass and pyruvate kinase activity correlate with interindividual differences in jumping performance of *Hyla multilineata*. Physiological and Biochemical Zoology 78:857–867.

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

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

———. 2004. Evolution and stability of the **G**-matrix on a landscape with moving optimum. Evolution 58:1639–1654.

———. 2007. The mutation matrix and the evolution of evolvability. Evolution 61:727–745.

Jones, A. G., R. Bürger, and S. J. Arnold. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. Nature Communications 5:3709.

Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the **G**-matrix and the response of the trait mean to selection. Journal of Evolutionary Biology 25:2210–2231.

Juarez, B. H., D. S. Moen, and D. C. Adams. 2020. A morphological method to approximate jumping performance in anurans for macroevolutionary studies. Evolutionary Biology 47:260–271.

———. 2023. Ecology, sexual dimorphism, and jumping evolution in anurans. Journal of Evolutionary Biology 36:829–841.

Kolbe, J. J., L. J. Revell, B. Szekely, E. D. Brodie III, and J. B. Losos. 2011. Convergent evolution of phenotypic integration and its alignment with morphological diversification in Caribbean *Anolis* ecomorphs. Evolution 65:3608–3624.

Krzanowski, W. J. 1979. Between-group comparison of principal components. Journal of the American Statistical Association 74:703–707.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

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

———. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

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

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

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Machado, F. A. 2020. Selection and constraints in the ecomorphological adaptive evolution of the skull of living Canidae (Carnivora, Mammalia). American Naturalist 196:197–215.

Machado, F. A., C. S. Mongle, G. Slater, A. Penna, A. Wisniewski, A. Soffin, V. Dutra, and J. C. Uyeda. 2023. Rules of teeth development align microevolution with macroevolution in extant and extinct primates. Nature Ecology and Evolution 7:1729–1739.

Machado, F. A., T. M. G. Zahn, and G. Marroig. 2018. Evolution of morphological integration in the skull of Carnivora (Mammalia): changes in Canidae lead to increased evolutionary potential of facial traits. Evolution 72:1399–1419.

Marroig, G., and J. M. Cheverud. 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.

———. 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. Evolution 59:1128–1142.

———. 2010. Size as a line of least resistance II: direct selection on size or correlated response due to constraints? Evolution 64:1470–1488.

Marroig, G., D. Melo, and G. Garcia. 2012. Modularity, noise and natural selection. Evolution 66:1506–1524.

Marsh, R. L. 1994. Jumping ability of anuran amphibians. Pages 51–111 *in* J. H. Jones, ed. Advances in veterinary science and comparative medicine. Academic Press, New York.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, J. J. Kolbe, J. B. Losos, and E. D. Brodie III. 2022. Conservation and convergence of genetic architecture in the adaptive radiation of *Anolis* lizards. American Naturalist 200:E207–E220.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, D. L. Mahler, J. J. Kolbe, J. B. Losos, and E. D. Brodie. 2018. Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. Evolution Letters 2:310–322.

McGlothlin, J. W., P. G. Parker, V. Nolan, and E. D. Ketterson. 2005. Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. Evolution 59:658–671.

McGuigan, K., S. F. Chenoweth, and M. W. Blows. 2005. Phenotypic divergence along lines of genetic variance. American Naturalist 165:32–43.

Melo, D., G. Garcia, A. Hubbe, A. P. Assis, and G. Marroig. 2016*a*. EvolQG—an R package for evolutionary quantitative genetics. F1000Research 4:925.

Melo, D., and G. Marroig. 2015. Directional selection can drive the evolution of modularity in complex traits. Proceedings of the National Academy of Sciences of the USA 112:470–475.

Melo, D., A. Porto, J. M. Cheverud, and G. Marroig. 2016*b*. Modularity: genes, development, and evolution. Annual Review of Ecology, Evolution, and Systematics 47:463–486.

Mendoza, E., and E. Azizi. 2021. Tuned muscle and spring properties increase elastic energy storage. Journal of Experimental Biology 224:jeb243180.

Mendoza, E., E. Azizi, and D. S. Moen. 2020. What explains vast differences in jumping power within a clade? diversity, ecology and evolution of anuran jumping power. Functional Ecology 34: 1053–1063.

Meyer, K., and M. Kirkpatrick. 2008. Perils of parsimony: properties of reduced-rank estimates of genetic covariance matrices. Genetics 180:1153–1166.

Moen, D. S. 2019. What determines the distinct morphology of species with a particular ecology? the roles of many-to-one mapping and trade-offs in the evolution of frog ecomorphology and performance. American Naturalist 194:E81–E95.

Moen, D. S., D. J. Irschick, and J. J. Wiens. 2013. Evolutionary conservatism and convergence both lead to striking similarity in ecology, morphology and performance across continents in frogs. Proceedings of the Royal Society B 280:20132156.

Moen, D. S., H. Morlon, and J. J. Wiens. 2016. Testing convergence versus history: convergence dominates phenotypic evolution for over 150 million years in frogs. Systematic Biology 65:146–160.

Moen, D. S., R. N. Ravelojaona, C. R. Hutter, and J. J. Wiens. 2021. Testing for adaptive radiation: a new approach applied to Madagascar frogs. Evolution 75:3008–3025.

Moen, D. S., and J. J. Wiens. 2009. Phylogenetic evidence for competitively driven divergence: body-size evolution in Caribbean treefrogs (Hylidae: *Osteopilus*). Evolution 63:195–214.

Mongle, C. S., A. Nesbitt, F. A. Machado, J. B. Smaers, A. H. Turner, F. E. Grine, and J. C. Uyeda. 2022. A common mechanism drives the alignment between the micro- and macroevolution of primate molars. Evolution 76:2975–2985.

Morinaga, G., J. J. Wiens, and D. S. Moen. 2023. The radiation continuum and the evolution of frog diversity. Nature Communications 14:7100.

Olson, E. C., and R. L. Miller. 1958. Morphological integration. University of Chicago Press, Chicago.

Olson, J. M., and R. L. Marsh. 1998. Activation patterns and length changes in hindlimb muscles of the bullfrog *Rana catesbeiana* during jumping. Journal of Experimental Biology 201:2763–2777.

Opedal, Ø. H., W. S. Armbruster, T. F. Hansen, A. Holstad, C. Pélabon, S. Andersson, D. R. Campbell, et al. 2023. Evolvability and trait function predict phenotypic divergence of plant populations. Proceedings of the National Academy of Sciences of the USA 120:e2203228120.

Opedal, Ø. H., L. S. Hildesheim, and W. S. Armbruster. 2022. Evolvability and constraint in the evolution of three-dimensional flower morphology. American Journal of Botany 109:1906–1917.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Pavlicev, M., J. M. Cheverud, and G. P. Wagner. 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proceedings of the Royal Society B 278:1903–1912.

Peplowski, M. M., and R. L. Marsh. 1997. Work and power output in the hindlimb muscles of Cuban tree frogs *Osteopilus septentrionalis* during jumping. Journal of Experimental Biology 200:2861–2870.

Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. Evolution 43:1209–1222.

———. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.

Pinheiro, P. D., B. L. Blotto, S. R. Ron, E. L. Stanley, P. C. Garcia, C. F. Haddad, T. Grant, et al. 2022. Prepollex diversity and evolution in Cophomantini (Anura: Hylidae: Hylinae). Zoological Journal of the Linnean Society 195:995–1021.

Portik, D. M., J. W. Streicher, and J. J. Wiens. 2023. Frog phylogeny: a time-calibrated, species-level tree based on hundreds of loci and 5,242 species. Molecular Phylogenetics and Evolution 188:107907.

Porto, A., F. B. de Oliveira, L. T. Shirai, V. De Conto, and G. Marroig. 2009. The evolution of modularity in the mammalian skull I: morphological integration patterns and magnitudes. Evolutionary Biology 36:118–135.

R Core Team. 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Renaud, S., J.-C. Auffray, and J. Michaux. 2006. Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to selection in fossil rodents. Evolution 60:1701–1717.

Renaud, S., C. Girard, and A.-B. Dufour. 2021. Morphometric variance, evolutionary constraints and their change through time in Late Devonian *Palmatolepis* conodonts. Evolution 75:2911–2929.

Revell, L. J. 2007. The **G** matrix under fluctuating correlational mutation and selection. Evolution 61:1857–1872.

Reynaga, C. M., H. C. Astley, and E. Azizi. 2018. Morphological and kinematic specializations of walking frogs. Journal of Experimental Zoology A 329:87–98.

Riedl, R. 1977. A systems-analytical approach to macro-evolutionary phenomena. Quarterly Review of Biology 52:351–370.

Roff, D. A. 1996. The evolution of genetic correlations: an analysis of patterns. Evolution 50:1392–1403.

Roff, D. A., and D. J. Fairbairn. 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution 66:2953–2960.

Salazar-Ciudad, I. 2006. Developmental constraints vs. variational properties: how pattern formation can help to understand evolution and development. Journal of Experimental Zoology B 306:107–125.

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

Schwaner, M. J., G. A. Freymiller, R. W. Clark, and C. P. McGowan. 2021. How to stick the landing: kangaroo rats use their tails to reorient during evasive jumps away from predators. Integrative and Comparative Biology 61:442–454.

Simon, M. N., R. Brandt, T. Kohlsdorf, and S. J. Arnold. 2019. Bite performance surfaces of three ecologically divergent Iguanidae lizards: relationships with lower jaw bones. Biological Journal of the Linnean Society 127:810–825.

Simon, M. N., E. Courtois, A. Herrel, and D. S. Moen. 2025. Data from: Macroevolutionary divergence along allometric lines of least resistance in frog hindlimb traits and its effect on locomotor evolution. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.rn8pk0pnp.

Simon, M. N., F. A. Machado, and G. Marroig. 2016. High evolutionary constraints limited adaptive responses to past climate changes in toad skulls. Proceedings of the Royal Society B 283:20161783.

Simon, M. N., G. Marroig, and S. J. Arnold. 2022. Detecting patterns of correlational selection with sampling error: a simulation study. Evolution 76:207–224.

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

Styga, J. M., T. M. Houslay, A. J. Wilson, and R. L. Earley. 2019. Development of **G**: a test in an amphibious fish. Heredity 122:696–708.

Sutton, G. P., E. Mendoza, E. Azizi, S. J. Longo, J. P. Olberding, M. Ilton, and S. N. Patek. 2019. Why do large animals never actuate their jumps with latch-mediated springs? because they can jump higher without them. Integrative and Comparative Biology 59: 1609–1618.

Taverne, M., H. Dutel, M. Fagan, A. Štambuk, D. Lisičić, Z. Tadić, A.-C. Fabre, et al. 2021. From micro to macroevolution: drivers of shape variation in an island radiation of *Podarcis* lizards. Evolution 75:2685–2707.

Tejero-Cicuéndez, H., I. Menéndez, A. Talavera, G. Mochales-Riaño, B. Burriel-Carranza, M. Simó-Riudalbas, S. Carranza, et al. 2023. Evolution along allometric lines of least resistance: morphological differentiation in *Pristurus* geckos. Evolution 77:2547–2560.

Toro, E., A. Herrel, B. Vanhooydonck, and D. J. Irschick. 2003. A biomechanical analysis of intra- and interspecific scaling of jumping and morphology in Caribbean *Anolis* lizards. Journal of Experimental Biology 206:2641–2652.

Towers, M., and C. Tickle. 2009. Growing models of vertebrate limb development. Development 136:179–190.

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

Uyeda, J. C., and J. W. McGlothlin. 2024. The predictive power of genetic variation. Science 384:622–623.

Vogel, S. 2009. Glimpses of creatures in their physical worlds. Princeton University Press, Princeton, NJ.

Voje, K. L., T. F. Hansen, C. K. Egset, G. H. Bolstad, and C. Pélabon. 2014. Allometric constraints and the evolution of allometry. Evolution 68:866–885.

Wagner, G. P., and L. Altenberg. 1996. Complex adaptations and the evolution of evolvability. Evolution 50:967–976.

Walsh, B., and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annual Review of Ecology, Evolution, and Systematics 40:41–59.

Wang, C.-K.L., M. Omi, D. Ferrari, H.-C. Cheng, G. Lizarraga, H.-J. Chin, W. B. Upholt, et al. 2004. Function of BMPs in the apical ectoderm of the developing mouse limb. Developmental Biology 269:109–122.

Wilson, R. S., C. E. Franklin, and R. S. James. 2000. Allometric scaling relationships of jumping performance in the striped marsh frog *Limnodynastes peronii*. Journal of Experimental Biology 203:1937–1946.

Zeng, Z.-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. Evolution 42:363–374.

References Cited Only in the Online Enhancements

Bartoszek, K., J. Fuentes-González, V. Mitov, J. Pienaar, M. Piwczyński, R. Puchałka, K. Spalik, et al. 2023. Model selection performance in phylogenetic comparative methods under multivariate Ornstein–Uhlenbeck models of trait evolution. Systematic Biology 72:275–293.

Bartoszek, K., J. Pienaar, P. Mostad, S. Andersson, and T. F. Hansen. 2012. A phylogenetic comparative method for studying multivariate adaptation. Journal of Theoretical Biology 314:204–215.

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. American Naturalist 163:329–340.

Caminer, M. A., and S. R. Ron. 2020. Systematics of the *Boana semilineata* species group (Anura: Hylidae), with a description of two new species from Amazonian Ecuador. Zoological Journal of the Linnean Society 190:149–180.

Carvalho, T. R., A. Fouquet, M. L. Lyra, A. A. Giaretta, C. E. Costa-Campos, M. T. Rodrigues, C. F. Haddad, et al. 2022. Species diversity and systematics of the *Leptodactylus melanonotus* group (Anura, Leptodactylidae): review of diagnostic traits and a new species from the Eastern Guiana Shield. Systematics and Biodiversity 20:1–31.

Da Silva, L. A., F. M. Magalhaes, H. Thomassen, F. S. Leite, A. A. Garda, R. A. Brandao, C. F. Haddad, et al. 2020. Unraveling the species diversity and relationships in the *Leptodactylus mystaceus* complex (Anura: Leptodactylidae), with the description of three new Brazilian species. Zootaxa 4779:151–189.

de Almeida, A. P., L. J. Moraes, R. R. Rojas, I. J. Roberto, V. T. de Carvalho, R. W. Avila, L. Frazao, et al. 2021. Phylogenetic relationships of the poorly known treefrog *Boana hobbsi* (Cochran & Goin, 1970) (Anura: Hylidae), systematic implications and remarks on morphological variations and geographic distribution. Zootaxa 4933:301–323.

de Sá, R.O., T. Grant, A. Camargo, W. R. Heyer, M. L. Ponssa, and E. Stanley. 2014. Systematics of the Neotropical genus *Leptodactylus* Fitzinger, 1826 (Anura: Leptodactylidae): phylogeny, the relevance of non-molecular evidence, and species accounts. South American Journal of Herpetology 9:S1–S128.

Escalona, M., E. La Marca, M. Castellanos, A. Fouquet, A. Crawford, F. J. Rojas-Runjaic, A. Giaretta, et al. 2021. Integrative taxonomy reveals a new but common Neotropical treefrog, hidden under the name *Boana xerophylla*. Zootaxa 4981:401–448.

Faivovich, J., P. D. Pinheiro, M. L. Lyra, M. O. Pereyra, D. Baldo, A. Munoz, S. Reichle, et al. 2021. Phylogenetic relationships of the *Boana pulchella* group (Anura: Hylidae). Molecular Phylogenetics and Evolution 155:106981.

Ho, L. S. T., and C. Ané. 2014. Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. Methods in Ecology and Evolution 5:1133–1146.

Koehler, J., D. Koscinski, J. M. Padial, J. C. Chaparro, P. Handford, S. C. Lougheed, and I. De la Riva. 2010. Systematics of Andean gladiator frogs of the *Hypsiboas pulchellus* species group (Anura, Hylidae). Zoologica Scripta 39:572–590.

Magalhães, F. de M., M. L. Lyra, T. R. De Carvalho, D. Baldo, F. Brusquetti, P. Burella, G. R. Colli, et al. 2020. Taxonomic review of South American butter frogs: phylogeny, geographic patterns, and species delimitation in the *Leptodactylus latrans* species group (Anura: Leptodactylidae). Herpetological Monographs 34:131–177.

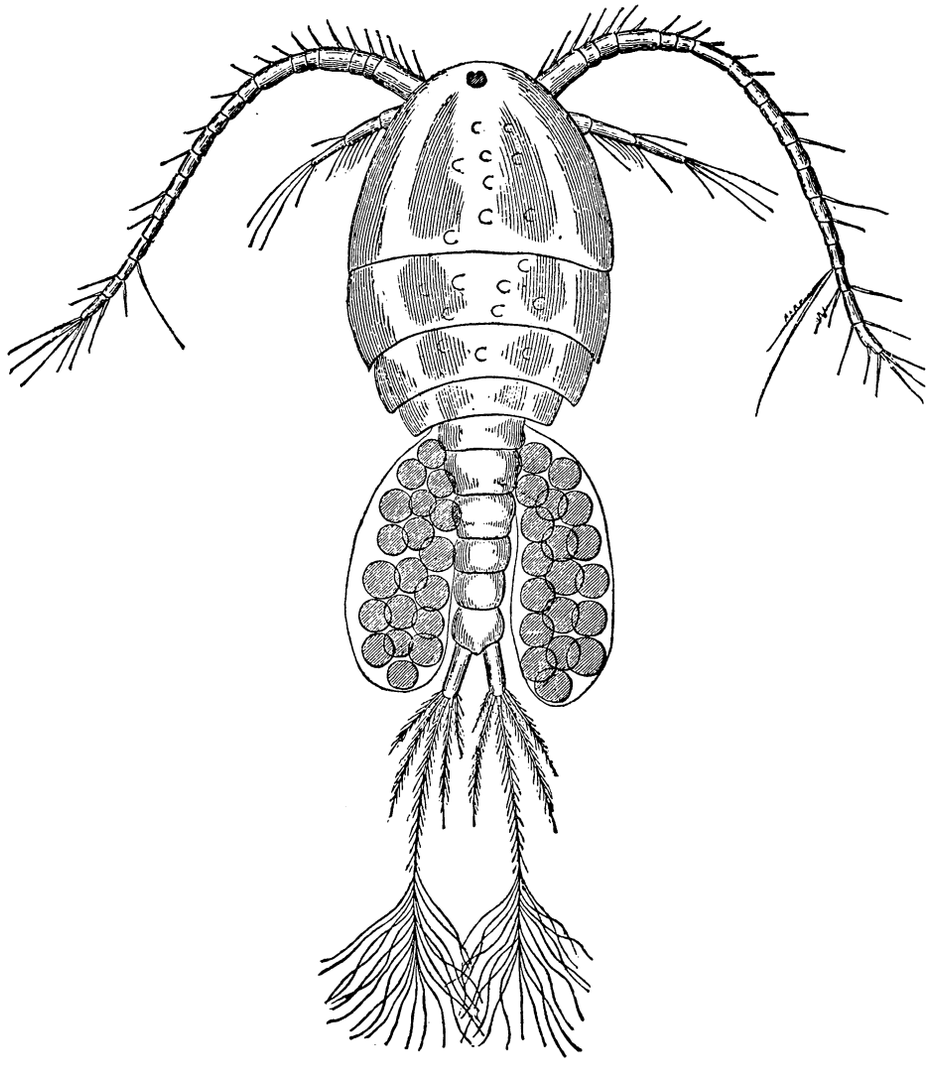
Marinho, P., D. L. Bang, I. Vidigal, and A. A. Giaretta. 2022. A new cryptic species of *Boana* (Hylinae: Cophomantini) of the *B. polytaenia* clade from the Brazilian Atlantic Forest. Journal of Herpetology 56:278–293.

Orrico, V. G., I. Nunes, C. Mattedi, A. Fouquet, A. W. Lemos, M. Rivera Correa, M. L. Lyra, et al. 2017. Integrative taxonomy supports the existence of two distinct species within *Hypsiboas crepitans* (Anura: Hylidae). Salamandra 53:99–113.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Schneider, R. G., D. E. Cardozo, F. Brusquetti, F. Kolenc, C. Borteiro, C. Haddad, N. G. Basso, et al. 2019. A new frog of the *Leptodactylus fuscus* species group (Anura: Leptodactylidae), endemic from the South American Gran Chaco. PeerJ 7:e7869.

Associate Editor: Joel W. McGlothlin
Editor: Volker H. W. Rudolf



"The Cyclops has received more than its share of names, owing to the great difference between different stages of its existence." From "Fresh-Water Entomostraca" by C. L. Herrick (*The American Naturalist*, 1879, 13:620–628).