Received: 5 April 2024 | Accepted: 26 September 2024

DOI: 10.1111/2041-210X.14438

Methods in Ecology and Evolution

RESEARCH ARTICLE

Extending phylogenetic regression models for comparing within-species patterns across the tree of life

Dean C. Adams[1] | Michael L. Collyer[2]

[1]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, USA

[2]Department of Science, Chatham University, Pittsburgh, Pennsylvania, USA

Correspondence
Dean C. Adams
Email: dcadams@iastate.edu

Funding information
Division of Biological Infrastructure, Grant/Award Number: 1902511 and 1902694; Division of Environmental Biology, Grant/Award Number: 2140720 and 2146220

Handling Editor: Gustavo Burin

Abstract

1. Evolutionary biologists characterize macroevolutionary trends of phenotypic change across the tree of life using phylogenetic comparative methods. However, within-species variation can complicate such investigations. For this reason, procedures for incorporating nonstructured (random) intraspecific variation have been developed.
2. Likewise, evolutionary biologists seek to understand microevolutionary patterns of phenotypic variation within species, such as sex-specific differences or allometric trends. Additionally, there is a desire to compare such within-species patterns across taxa, but current analytical approaches cannot be used to interrogate within-species patterns while simultaneously accounting for phylogenetic non-independence. Consequently, deciphering how intraspecific trends evolve remains a challenge.
3. Here we introduce an extended phylogenetic generalized least squares (E-PGLS) procedure which facilitates comparisons of within-species patterns across species while simultaneously accounting for phylogenetic non-independence.
4. Our method uses an expanded phylogenetic covariance matrix, a hierarchical linear model, and permutation methods to obtain empirical sampling distributions and effect sizes for model effects that can evaluate differences in intraspecific trends across species for both univariate and multivariate data, while conditioning them on the phylogeny.
5. The method has appropriate statistical properties for both balanced and imbalanced data. Additionally, the procedure obtains evolutionary covariance estimates that reflect those from existing approaches for nonstructured intraspecific variation. Importantly, E-PGLS can detect differences in structured (i.e. microevolutionary) intraspecific patterns across species when such trends are present. Thus, E-PGLS extends the reach of phylogenetic comparative methods into the intraspecific comparative realm, by providing the ability to compare within-species trends across species while simultaneously accounting for shared evolutionary history.

KEYWORDS
comparative analysis, evolutionary biology, linear models, macroevolution, statistics

1 | INTRODUCTION

Deciphering patterns of phenotypic variation is a mainstay of evolutionary research, yet the analysis of phenotypic data remains distinctly different enterprises between microevolutionary and macroevolutionary studies. At contemporary timescales, microevolutionary investigations strive to identify the relationship between patterns of phenotypic variation across individuals or populations, and the environmental factors that underlie selection and the generation of phenotypic change. These studies leverage phenotypic information from many individuals (typically hundreds or more), and across multiple populations, to elucidate the ecological factors that associate with patterns of phenotypic diversity within species (e.g. Beausoleil et al., 2023; Collyer et al., 2015; Reyes-Puig et al., 2023; Stroud et al., 2023; Tejero-Cicuéndez et al., 2023). In contrast to within-species studies, macroevolutionary studies utilize a phylogenetic perspective to examine trait associations across species and in light of evolutionary history. Here, phylogenetically informed statistical procedures are leveraged to evaluate trait associations across species while accounting for their phylogenetic non-independence (Felsenstein, 1985; Harmon, 2019; Harvey & Pagel, 1991; Martins & Hansen, 1997; O'Meara, 2012). As with microevolutionary analyses, selective forces are thought to play an important role in shaping phenotypic diversification across taxa and clades, as well (Burns et al., 2024; Hunt et al., 2023; Sidlauskas, 2008). However, while it is generally thought that phenotypic patterns resulting from contemporary processes may 'scale up' to macroevolutionary patterns found across species and clades, evaluating this prediction is challenging, as it requires the ability to interrogate microevolutionary patterns in a phylogenetic context (i.e. account for the non-independence among species when evaluating within-species trends).

In macroevolutionary studies, phylogenetic comparative methods (PCMs) are a set of statistical tools that enable biologists to condition the data on the phylogeny under a particular model of evolutionary change (Brownian motion, Ornstein–Uhlenbeck, etc.). Patterns of covariation between traits may then be evaluated across species while accounting for the evolutionary relationships among them (Felsenstein, 1985; Harmon, 2019; Harvey & Pagel, 1991; Martins & Hansen, 1997; O'Meara, 2012). Since their advent in the late 20th century, the development of PCMs has progressed at a rapid pace, resulting in myriad analytical tools that enable evolutionary biologists to evaluate a wide array of biological hypotheses describing how ecological and evolutionary forces shape patterns of phenotypic diversification (Adams, 2014a, 2014b; Beaulieu et al., 2012; Blomberg et al., 2003; Butler & King, 2004; Clavel et al., 2015; Collyer et al., 2022; Gaboriau et al., 2024; Khabbazian et al., 2016; Mitov et al., 2019; O'Meara et al., 2006; Revell & Harmon, 2008; Uyeda & Harmon, 2014).

Phylogenetic comparative analyses typically use species means as trait data, and thus implicitly assume that intraspecific variation is negligible relative to interspecific variation (Felsenstein, 2008; Garamszegi, 2014; Ives et al., 2007). This is not always the case. Additionally, because species means are obtained from finite samples of measured individuals, estimation error will be present (Felsenstein, 2008; Ives et al., 2007; Lynch, 1991). Together, this uncertainty can result in inaccuracy during PCM analyses, which can have a profound effect on biological inferences. For example, within-species variation can instigate bias in parameter estimates (Felsenstein, 2008; Ives et al., 2007), can generate elevated type I error rates for hypothesis testing procedures (Felsenstein, 2008; Harmon & Losos, 2005), can result in low statistical power of such tests (Harmon & Losos, 2005) and can cause model misspecification when comparing the fit of alternative evolutionary models to phenotypic data (Silvestro et al., 2015). To mitigate these effects, several procedures have been developed that incorporate intraspecific variation into the PCM analytical pipeline (Felsenstein, 2008; Gaboriau et al., 2020; Hadfield & Nakagawa, 2010; Hansen & Bartoszek, 2012; Ives et al., 2007; Kostikova et al., 2016; Revell & Reynolds, 2012; Villemereuil & Nakagawa, 2014). In general, these procedures consider the totality of within-species variation as measurement error, which is represented as nonstructured (random) sampling error around the mean phenotype. Macroevolutionary trends across species may be evaluated while taking into consideration both within-species variation, as well as the phylogenetic non-independence among taxa. Importantly, the variation among sampled individuals, populations or other units within species all contribute to the overall within-species variation, but without regard to the underlying sources that might influence that variation.

Nonetheless, intraspecific variation is often structured in biologically meaningful ways that convey additional information that empiricists wish to investigate, particularly at microevolutionary scales. For instance, differences between the sexes (i.e. sexual dimorphism) or phenotypic differences that vary with body size (e.g. allometric trends) represent *structured* patterns of intraspecific variation and are frequently of interest to investigate and compare among taxa (Juarez & Adams, 2022; Reyes-Puig et al., 2023; Tejero-Cicuéndez et al., 2023). Clearly, when comparing such patterns across species, one must also account for their phylogenetic non-independence, but current analytical approaches cannot interrogate and compare microevolutionary trends sampled within species, while simultaneously accounting for phylogenetic non-independence based on relatedness defined at the species level. Thus at present, we lack tools that are capable of bridging this particular microevolutionary—macroevolutionary divide.

In this work, we introduce a novel phylogenetic comparative method for comparing within-species patterns among species while accounting for their phylogenetic non-independence. Our approach may be used to evaluate trends in univariate or multivariate phenotypic data, and extends phylogenetic generalized least squares procedures (E-PGLS) by fitting individual-level data in a hierarchical linear model while conditioning on an expanded phylogenetic covariance matrix. Permutation procedures are developed which generate empirical sampling distributions and effect sizes for model effects, allowing the statistical evaluation of both species-level patterns and for comparing within-species trends among species across the tree of life. As demonstrated below, the E-PGLS procedure

displays appropriate statistical properties, with acceptable type I error and high statistical power. Additionally, when intraspecific variation is nonstructured (random), E-PGLS obtains evolutionary covariance estimates that are comparable to those found from existing approaches (Felsenstein, 2008). However, a unique contribution of our approach is that when within-species variation is structured, E-PGLS is capable of detecting differences in intraspecific patterns across taxa when such trends are present. Thus, E-PGLS facilitates the comparison of within-species (microevolutionary) patterns across species while accounting for among-species variation, as well as the phylogenetic non-independence of those species. As such, E-PGLS is capable of revealing macroevolutionary insights regarding the evolution of intraspecific phenotypic trends in a manner that is currently beyond the reach of standard PCMs.

2 | CONCEPTUAL DEVELOPMENT

2.1 | Model construction

E-PGLS combines a hierarchical linear model and an expanded phylogenetic covariance matrix to interrogate variation in one or more phenotypic traits while conditioning the individual-level data on the phylogeny. Let n represent the total number of individuals sampled across N species, with n_i individuals sampled for the ith species, such that $\sum_{i=1}^{N} n_i = n$. It is not assumed that species have the same n_i. As shown below, the approach may be used with imbalanced data (which occurs frequently due to sampling limitations when utilizing museum collections or field samples). Phenotypic values, $\mathbf{Z} = f(\mathbf{Y})$, comprise an $n \times p$ matrix for n individuals, with p measured traits per individual. $\mathbf{Z}$ may represent mean-centred data ($\mathbf{Y} - \overline{\mathbf{Y}}$), or a set of transformed variables, such as Procrustes-aligned coordinates representing geometric morphometric shape data (Adams et al., 2013; Rohlf & Slice, 1990). Next, an $n \times (N + k)$ linear model design matrix, $\mathbf{X}$ can be constructed that describes both species values and within-species covariates or sub-species factors. $\mathbf{X}$ is a partitioned matrix, that is $\mathbf{X} = \mathbf{X}_S \mid \mathbf{X}_X$, with $\mathbf{X}_S$ containing species classifiers and $\mathbf{X}_X$ corresponding to factors (e.g. sex, population, ecotype) or covariates (e.g. size, performance) measured at the individual (within-species) level. The hierarchical linear model describing patterns of phenotypic variation is then defined as:

$$\mathbf{Z} = \mathbf{X}_S\hat{\mathbf{B}}_S + \mathbf{X}_X\hat{\mathbf{B}}_X + \mathbf{E} \quad (1)$$

where $\mathbf{X}_S$ is an $n \times N$ species design matrix of dummy variables (0s and 1s), whose N columns designate the species to which each individual belongs, and $\mathbf{X}_X$ is an $n \times k$ design matrix containing one or more explanatory variables. Model coefficients are found in the $N \times p$ and $k \times p$ matrices $\hat{\mathbf{B}}_X$ and $\hat{\mathbf{B}}_S$ respectively, and $\mathbf{E}$ is an $n \times p$ matrix of residual error, distributed as: $\mathbf{E} \sim \mathcal{MN}(0, \Omega_n)$. Here, $\mathbf{E}$ follows a matrix normal distribution (a generalization of a multivariate normal distribution), with expected covariances that express the correlations among species as a result of their shared phylogenetic history (Blomberg et al., 2012; Garland & Ives, 2000; Grafen, 1989).

Coefficients in Equation 1 are estimated jointly, not iteratively. The equation could be written as, $\mathbf{Z} = \mathbf{X}\hat{\mathbf{B}} + \mathbf{E}$, recognizing that $\mathbf{X} = (\mathbf{X}_S \mid \mathbf{X}_X)$ and $\hat{\mathbf{B}} = \begin{pmatrix} \hat{\mathbf{B}}_S \\ \hat{\mathbf{B}}_X \end{pmatrix}$. Equation 1 conveniently demonstrates that coefficients and estimates from them can be separated into species and within-species effects. However, this should not imply that species effects are estimated first and within-species effects second, in a sequence. Rather, the linear model design matrix and matrix of coefficients are concatenated from the submatrices shown separately in Equation 1.

In standard PGLS implementations that utilize one individual per species (e.g. the species mean), phylogenetic correlations are expressed in Ω, which is an $N \times N$ phylogenetic covariance matrix obtained from the phylogeny under a particular model of evolutionary change (Brownian motion, Ornstein-Uhlenbeck, etc.). The residual error of E-PGLS follows an analogous construction, but where Ω_n is an $n \times n$ expanded phylogenetic covariance matrix, describing the expected covariation among individuals across species due to phylogenetic non-independence. This matrix is found by pre- and post-multiplying the species-level phylogenetic covariance matrix Ω by $\mathbf{X}_S$:

$$\Omega_n = \mathbf{X}_S\Omega\mathbf{X}_S^T \quad (2)$$

where $\mathbf{X}_S$ is the species design matrix defined above. As with standard implementations of PGLS, Ω in Equation 2 may be obtained under a Brownian motion model, or using some other model of evolutionary change, such as an Ornstein-Uhlenbeck model. Thus, Ω_n describes the evolutionary covariation among individuals as is expected under the specified model of evolutionary divergence found in Ω (additional scaling options for Ω_n are described in the Supporting Information). Using Equation 2, individuals sampled from the same species display the same expected covariances when compared to individuals from other species. Note that Equation 2 provides a mathematically compact means of obtaining the expanded phylogenetic covariance matrix in López-Fernández et al. (2014).

An important attribute of E-PGLS is the ability to compare within-species (microevolutionary) patterns across taxa while accounting for phylogenetic non-independence. To accomplish this, species-specific within-species patterns (cross-species effects) are modelled by including interactions between model main effects and the species design matrix (e.g. $\mathbf{X}_{C_X} = \mathbf{X}_X \times \mathbf{X}_S$). Note that if $\mathbf{X}_X$ contains more than one independent variable, cross-species trends could be considered for multiple factors or covariates in $\mathbf{X}_X$. Incorporating cross-species trends in Equation 1 results in the full E-PGLS hierarchical model:

$$\mathbf{Z} = \mathbf{X}_S\hat{\mathbf{B}}_S + \mathbf{X}_X\hat{\mathbf{B}}_X + \mathbf{X}_{C_X}\hat{\mathbf{B}}_{C_X} + \mathbf{E}. \quad (3)$$

In general, these three sets of coefficients—species ($\hat{\mathbf{B}}_S$), full-clade association of a trait with an independent variable ($\hat{\mathbf{B}}_X$) and cross-species comparisons ($\hat{\mathbf{B}}_{C_X}$)—should allow an expansion of hypothesis tests compared to standard PGLS (which provides $\hat{\mathbf{B}}_X$, only.) First, it could be possible to test for a species effect (based on $\hat{\mathbf{B}}_S$) because of replicated observations within species. Second, it would be possible

to test for heterogeneity of microevolutionary patterns ($\hat{\mathbf{B}}_{C_X}$) and possibly infer the association of microevolutionary patterns with respect to macroevolutionary events (e.g. subclade distinction of microevolutionary relationships). Hypothesis tests first require, however, reliable estimates of micro- and macroevolutionary parameters.

2.2 | Parameter estimation

Once the components of the linear model have been constructed, the data are conditioned on the phylogeny using phylogenetic transformation (Adams, 2014b; Adams & Collyer, 2015, 2018b, 2022; Garland & Ives, 2000), and model parameters are obtained. Phylogenetic transformation is an implementation of the well-known statistical transformation that reformulates generalized least squares (GLS) problems to ordinary least squares (OLS) problems (see Johnston & DiNardo, 1997; Judge et al., 1985; Kariya & Kurata, 2004; Rencher, 2000). Following Garland and Ives (2000), the phylogenetic transformation matrix is found as,

$$\mathbf{P} = \left(\mathbf{U}\Lambda^{-1/2}\mathbf{U}^{T}\right)^{-1}. \tag{4}$$

Here, $\mathbf{U}$ and Λ are matrices of eigenvectors and eigenvalues of Ω_n. Note that Ω_n is not of full rank, because its rows and columns are replicated for each n_i within each species. Thus, the number of eigenvectors and eigenvalues used in Equation 4 is restricted to the rank of Ω_n. Additionally, because Ω_n contains many zero elements, sparse-matrix methods and enhanced analytical algorithms (*linear model checkers*, sensu Baken et al., 2021) can be used to increase computational efficiency.

Using $\mathbf{P}$, the data matrix is conditioned on the phylogeny via matrix projection (i.e. phylogenetic transformation): $\tilde{\mathbf{Z}} = \mathbf{PZ}$. Next, the model design matrices are concatenated (i.e. $\mathbf{X}_{\text{tot}} = \mathbf{X}_S \mid \mathbf{X}_X \mid \mathbf{X}_{C_X}$), and the model design matrix is also conditioned on the phylogeny via matrix projection: $\tilde{\mathbf{X}}_{\text{tot}} = \mathbf{PX}_{\text{tot}}$. Model parameters are then obtained using OLS procedures:

$$\hat{\mathbf{B}}_{\text{tot}} = \left(\tilde{\mathbf{X}}_{\text{tot}}^{t}\tilde{\mathbf{X}}_{\text{tot}}\right)^{-1}\tilde{\mathbf{X}}_{\text{tot}}^{T}\tilde{\mathbf{Z}} \tag{5}$$

The linear model design matrix and coefficients in Equation 5 represent a condensed nomenclature for concatenated matrices, which are used for joint estimation of species, within-species, and cross-species coefficients and effects. The partitioning of matrices can be illustrated as

$$\mathbf{X}_{\text{tot}} = \left(\mathbf{X}_S|\mathbf{X}_X|\mathbf{X}_{C_X}\right),$$

and

$$\hat{\mathbf{B}}_{\text{tot}} = \begin{pmatrix} \hat{\mathbf{B}}_S \\ \hat{\mathbf{B}}_X \\ \hat{\mathbf{B}}_{C_X} \end{pmatrix}.$$

$\hat{\mathbf{B}}_S$ represents coefficients (parameter estimates) that describes macroevolutionary differences among species mean phenotypes, while $\hat{\mathbf{B}}_{C_X}$ describe differences in species-specific (i.e. microevolutionary) trends. On the other hand, $\hat{\mathbf{B}}_X$ can represent either microevolutionary or macroevolutionary patterns, depending upon the study design. For example, coefficients of $\hat{\mathbf{B}}_X$ representing covariates in $\mathbf{X}_X$ that segregate across species encode for patterns that differ among species, and thus describe macroevolutionary trends across the phylogeny. By contrast, coefficients of $\hat{\mathbf{B}}_X$ representing factors in $\mathbf{X}_X$ that assign individuals within species to different groups (e.g. population or sex) describe microevolutionary trends within species. Together, the set of coefficients in $\hat{\mathbf{B}}_{\text{tot}}$ are used in combination with the set of model design matrices in $\mathbf{X}_{\text{tot}}$ to estimate the variance components of the phylogenetic linear model, and to conduct permutation-based hypothesis testing procedures that evaluate model terms statistically.

2.3 | Variance components

E-PGLS requires that within-species variation is taken into consideration while evaluating other effects in the model. This mathematical requirement dictates the manner in which sums of squares and cross-products (SSCP) matrices and variance components are calculated.

For the species effect, E-PGLS uses marginal (type III) sums of squares, which estimates the SSCP between a model containing all terms and a model lacking the species term. Here, the full and reduced models are found as $\mathbf{X}_F = \mathbf{X}_S \mid \mathbf{X}_X \mid \mathbf{X}_{C_X}$ and $\mathbf{X}_R = \mathbf{X}_0 \mid \mathbf{X}_X \mid \mathbf{X}_{C_X}$, which differ only by species parameters ($\mathbf{X}_0$ is a vector of 1s: the "intercept" model). Model parameters are then estimated as above, and the difference in estimations between models is

$$\begin{aligned} \Delta\hat{\mathbf{Z}}_S &= \hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R = \mathbf{X}_F\hat{\mathbf{B}}_F - \mathbf{X}_R\hat{\mathbf{B}}_R \\ &= \left(\mathbf{X}_S|\mathbf{X}_X|\mathbf{X}_{C_X}\right)\begin{pmatrix} \hat{\mathbf{B}}_S \\ \hat{\mathbf{B}}_X \\ \hat{\mathbf{B}}_{C_X} \end{pmatrix}_F - \left(\mathbf{X}_0|\mathbf{X}_X|\mathbf{X}_{C_X}\right)\begin{pmatrix} \hat{\mathbf{B}}_0 \\ \hat{\mathbf{B}}_X \\ \hat{\mathbf{B}}_{C_X} \end{pmatrix}_R \\ &= \left(\mathbf{X}_S\hat{\mathbf{B}}_{s,F} + \mathbf{X}_X\hat{\mathbf{B}}_{X,F} + \mathbf{X}_{C_X}\hat{\mathbf{B}}_{C_X,F}\right) - \left(\mathbf{X}_0\hat{\mathbf{B}}_{0,R} + \mathbf{X}_X\hat{\mathbf{B}}_{X,R} + \mathbf{X}_{C_X}\hat{\mathbf{B}}_{C_X,R}\right). \end{aligned} \tag{6}$$

From this the SSCP for the species effect while holding constant all other effects in the model is found as

$$\mathbf{S}_S = \left(\hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R\right)\mathbf{P}^{T}\mathbf{P}\left(\hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R\right) = \left(\hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R\right)^{T}\Omega_n^{-1}\left(\hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R\right). \tag{7}$$

The variances and covariances associated with the species effect is then found by dividing $\mathbf{S}_S$ by $N - 1$.

For all remaining effects in the model, E-PGLS uses conditional (type II) sums of squares to obtain variance components. Here, intraspecific variation is restricted within species blocks, ensuring that the other model effects are not confounded by inter-species effects. This is accomplished by comparing full and reduced models for the effect of interest, but still containing the species effect, $\mathbf{X}_S$. For instance, the difference between fitted values for X can be found as

$$\begin{aligned}\Delta\hat{\mathbf{Z}}_X &= \hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R = \mathbf{X}_F\hat{\mathbf{B}}_F - \mathbf{X}_R\hat{\mathbf{B}}_R \\ &= \left(\mathbf{X}_S\hat{\mathbf{B}}_{S,F} + \mathbf{X}_X\hat{\mathbf{B}}_{X,F}\right) - \left(\mathbf{X}_S\hat{\mathbf{B}}_{S,R}\right).\end{aligned} \tag{8}$$

This ensures that the effect of X is conditioned on the estimation of species means, which are also conditioned on the non-independence of those means because of phylogenetic relatedness. SSCP matrices for each model effect are then obtained in a manner analogous to Equation 7, except that the reduced and full models both contain $\mathbf{X}_S$.

Likewise, the cross-species X_{C_X} SSCP matrix is estimated based on a difference between fitted values from the previous full model, now the reduced model, and a full model containing $\mathbf{X}_{C_X}$,

$$\begin{aligned}\Delta\hat{\mathbf{Z}}_{C_X} &= \hat{\mathbf{Z}}_F - \hat{\mathbf{Z}}_R = \mathbf{X}_F\hat{\mathbf{B}}_F - \mathbf{X}_R\hat{\mathbf{B}}_R \\ &= \left(\mathbf{X}_S\hat{\mathbf{B}}_{S,F} + \mathbf{X}_X\hat{\mathbf{B}}_{X,F} + \mathbf{X}_{C_x}\hat{\mathbf{B}}_{C_X,F}\right) - \left(\mathbf{X}_S\hat{\mathbf{B}}_{S,R} + \mathbf{X}_X\hat{\mathbf{B}}_{X,R}\right).\end{aligned} \tag{9}$$

Type II SSCP estimation means that for multiple cross-species comparisons, effects are conditioned on phylogenetically adjusted species means and the common trends of the multiple variables.

Finally, the residuals of the model are obtained by contrasting the data and fitted values of the E-PGLS model containing all terms,

$$\begin{aligned}\mathbf{E} &= \mathbf{Z} - \hat{\mathbf{Z}}_F = \mathbf{Z} - \mathbf{X}_F\hat{\mathbf{B}}_F \\ &= \mathbf{Z} - \left(\mathbf{X}_S\hat{\mathbf{B}}_{S,F} + \mathbf{X}_X\hat{\mathbf{B}}_{X,F} + \mathbf{X}_{C_x}\hat{\mathbf{B}}_{C_X,F}\right).\end{aligned} \tag{10}$$

The residual SSCP, $\mathbf{S}_{\text{Residuals}} = (\mathbf{PE})^T\mathbf{PE} = \mathbf{E}^T\mathbf{P}^T\mathbf{PE} = \mathbf{E}^T\Omega_n^{-1}\mathbf{E}$, is then calculated as in Equation 7. Additional details of the E-PGLS procedure may be found in the Supporting Information.

2.4 | Inferential statistics and permutation procedures

To evaluate model effects, E-PGLS uses a combination of robust statistical summary measures and the randomization of residuals in a permutation procedure (RRPP: Collyer et al., 2015, 2022; Adams & Collyer, 2018b, 2022). Generally, linear model summary statistics are found from SSCP matrices for various terms in the model relative to the residual covariance matrix, $\hat{\Sigma}_{\text{Residuals}} = (n-1)^{-1}\mathbf{S}_{\text{Residuals}}$. This forms the basis of inferential tests. Using this approach, the residual SSCP may be leveraged to obtain MANOVA statistics based on the Eigen decomposition of a matrix product that generalizes the ratio of term SSCP to residual SSCP,

$$\mathbf{U}\Lambda\mathbf{U}^T = \mathbf{S}_{\text{Residuals}}^{-1}\mathbf{S}_{\text{Term}}. \tag{11}$$

Note that this decomposition is often necessary, as the data ($\mathbf{Z}$) may be rank-deficient; in which case, the data are first subjected to a principal components analysis, and all calculations above are based on the set of non-trivial dimensions (i.e. those whose $\lambda > 0$). From Equation 11, MANOVA statistics (e.g. Roy's maximum root ($\lambda_{\max}$), Wilks' Λ, etc.) may then be obtained as a summary of multivariate dispersion (for additional information see Supporting Information).

Additionally, one may obtain a series of multivariate test statistics using likelihood. Here, the determinant of $\mathbf{U}\Lambda\mathbf{U}^T$ (i.e. $\det\left(\mathbf{S}_{\text{Residuals}}^{-1}\mathbf{S}_{\text{Term}}\right)$) is a ratio of determinants. This value is equivalent to a likelihood ratio and may be obtained for each term in the model. Then, a permutation test evaluating these values is commensurate with a non-parametric likelihood ratio test (for additional information see Supporting Information).

Alternatively, an F-statistic can be calculated based on the traces of the SSCPs, which therefore does not require matrix inversion. Here the summary statistic is calculated as

$$F = \frac{\text{tr}\left(\mathbf{S}_{\text{Term}}\right)/\text{df}_{\text{Term}}}{\text{tr}\left(\mathbf{S}_{\text{Residuals}}\right)/\text{df}_{\text{Residuals}}}, \tag{12}$$

where tr represents the trace of the matrix (i.e. the sum of its diagonal elements), and $\mathbf{S}_{\text{Term}}$ and $\mathbf{S}_{\text{Residuals}}$ are the SSCP matrices as above. Note that for univariate data, Equation 12 is the standard F-statistic, though it may be used for data of any dimension, including those containing singular dimensions (e.g. when $p > N$: see Adams, 2014b; Adams & Collyer, 2022). For this reason, F-statistics are commonly obtained in this manner for high-dimensional data (e.g. Anderson & Robinson, 2001; Goodall, 1991; Klingenberg & McIntyre, 1998). However, unlike the MANOVA statistics derived from Equation 11, this F-statistic does not incorporate the covariances among variables in its calculation.

For any of these summary statistics, permutation procedures are then used to generate empirical sampling distributions. Here, residuals from the reduced model for each model effect are permuted, as they represent the appropriate exchangeable units for the RRPP procedure (for justification see Supporting Information, and also Collyer et al., 2015, 2022; Adams & Collyer, 2018b, 2022). In addition, for all model effects except $\mathbf{X}_S$, RRPP is restricted to within species so that intraspecific effects are not confounded by interspecific effects (for details, see Supporting Information).

Finally, effect sizes may be calculated for each term in the model to provide a means of evaluating the strength of signal of each effect,

$$Z = \frac{\theta_{\text{obs}} - \mu_\theta}{\sigma_\theta}, \tag{13}$$

where μ and σ are the mean and standard deviation of the normalized empirical sampling distribution of the test statistic (e.g. $\theta = f(F)$), respectively.

3 | SIMULATION METHODS AND RESULTS

To verify the performance of E-PGLS, we used a series of stochastic sampling experiments to examine its behaviour under various conditions. Below we report on the two primary simulation experiments. Additional simulations under a wider set of conditions may be found in the Supporting Information.

3.1 | PGLS with nonstructured (random) intraspecific variation

First, when intraspecific variation is nonstructured, it may be treated as random sampling error around the mean phenotype. Thus, under these conditions, we would expect that E-PGLS should perform similarly to existing methods that can incorporate within-species variation (e.g. Felsenstein, 2008; Ives et al., 2007). To verify whether this was the case, we conducted a series of stochastic sampling experiments. Our design was as follows: phenotypic values were simulated on a phylogeny under Brownian motion (BM) with a known covariation between traits, and individuals within species were generated by adding nonstructured (random) variation to the mean phenotypic values for that species. We then estimated the evolutionary covariation between traits using species means, within-species independent contrasts (Felsenstein, 2008) and E-PGLS.

To implement this test, we simulated 1000 pure birth phylogenies for $N = 100$ species. Two phenotypic traits were simulated on each phylogeny under Brownian motion (BM), using an input covariance matrix (Σ) whose trait covariance was $\rho = 0.4$. For each dataset, a random number of individuals ($n_i = 1 \rightarrow 10$) was selected for each species, and phenotypic values for each individual were generated by adding nonstructured (random) variation to the mean phenotypic values for that species ($\mathbf{Y}_{n_{ij}} = \overline{\mathbf{Y}}_{N_i} + \mathbf{E}_{\mathcal{N}(0,0.2)}$). This procedure generated nonstructured intraspecific variation among individuals within species in a manner similar to the simulations performed by Felsenstein (2008). For each dataset, we then estimated the covariation between the two traits using E-PGLS, within-species independent contrasts (Felsenstein, 2008), and using species means. Computer code and additional details are found in the Supporting Information.

3.1.1 | Results

Our analyses revealed that when intraspecific variation was nonstructured (random), estimates of covariation between traits obtained from E-PGLS and from within-species contrasts closely matched the input value, and displayed considerably less variation across datasets as compared with estimates obtained using species means (Figure 1). Additionally, across simulations there was a very strong correlation between estimates from E-PGLS with those from within-species contrasts ($r = 0.94$). Thus, when intraspecific variation was nonstructured, we verified that E-PGLS provided estimates of evolutionary trait covariation commensurate with those obtained using prior PCMs that incorporate within-species variation as random error into the analysis (Felsenstein, 2008).

3.2 | Comparing within-species patterns: Structured intraspecific variation

One important advance that E-PGLS enables is the ability to compare within-species (i.e. microevolutionary) patterns across species, while simultaneously conditioning this analysis on the phylogeny. Accomplishing this amounts to an evaluation of structured patterns of intraspecific variation, where those patterns may differ across species. In this section, we evaluated whether E-PGLS was capable of detecting differences in structured patterns of within-species variation across the phylogeny. To verify this, we performed stochastic sampling experiments similar to those described above but where the multivariate phenotypes also contained structured intraspecific variation. The structured intraspecific variation represented phenotypic differences between groups of individuals within species, as may be found between sexes, localities or other biological factors of interest. Additionally, structured variation was simulated to be species-specific, where some species displayed within-species patterns while others did not. Simulations were also performed where the degree of structured intraspecific variation differed in its intensity across simulations. We then used E-PGLS to evaluate phenotypic variation attributed to potential effects via the model, $\mathbf{Z} = \mathbf{X}_S\hat{\mathbf{B}}_S + \mathbf{X}_X\hat{\mathbf{B}}_X + \mathbf{X}_{Gp}\hat{\mathbf{B}}_{Gp} + \mathbf{X}_{C_{Gp}}\hat{\mathbf{B}}_{C_{Gp}} + \mathbf{E}$, where the subscript, $_{Gp}$ refers to a grouping factor with two levels.

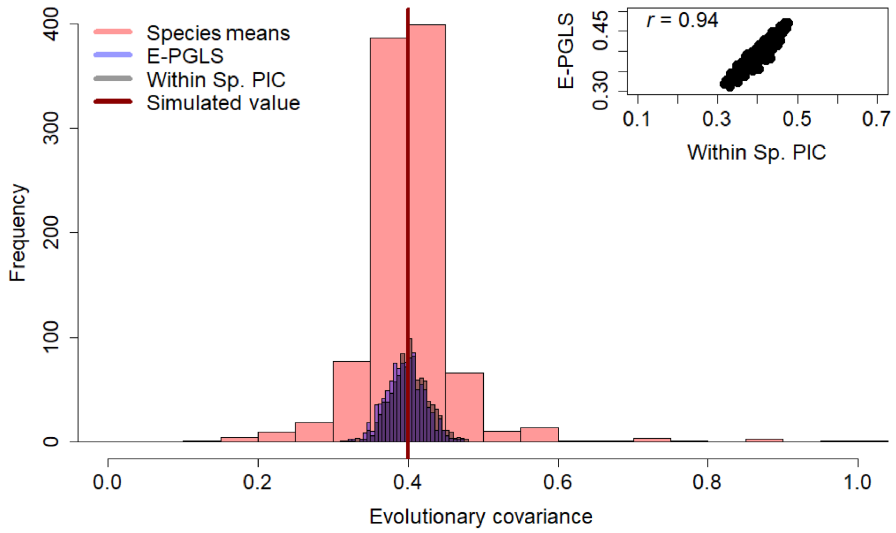

FIGURE 1 Histogram of evolutionary covariance estimates for 1000 simulated datasets. Estimates of evolutionary covariance were obtained using (1) species means, (2) within-species independent contrasts and (3) E-PGLS. Inset is a dot plot between estimates from within-species independent contrasts and E-PGLS for all 1000 datasets.

To implement this test, we first simulated 1000 pure birth phylogenies for $N = 100$ species. Three phenotypic traits ($\mathbf{Y}_{(1,2)}$ and X_1) were simulated on each phylogeny under Brownian motion (BM), using an input covariance matrix (Σ) with all trait covariances set to $\rho = 0.5$. For each dataset, $n_i = 4$ individuals per species were generated (i.e. balanced within-species sampling), and individuals within species were randomly assigned to one of two groups. Additional simulations were also conducted with $n_i = 10$ individuals per species, as well scenarios where the number of individuals within species was imbalanced ($n_i = 2,4,6,8$ or 10 individuals per species). These simulations were performed to evaluate the effect of varying levels of within-species sampling on the approach (see Supporting Information). For each simulation, phenotypic values for individuals were obtained by adding multivariate normal variation $\mathcal{N}(0, \Sigma)$ to the mean phenotypic values for that species (following Felsenstein, 2008). Group-specific differences were then incorporated into $\mathbf{Y}$ by adding $X_{Gp}\beta_{X_{Gp}}$ to all individuals. Here, X_{Gp} was a design matrix of 0s and 1s indicating group association for all individuals. For half of the species (randomly selected), we set the values of X_{Gp} to zero, so that no group-specific variation was added to these species. $\beta_{X_{Gp}}$ then modelled the strength of structured within-species group differences for a given simulation and varied progressively across simulations as $\beta_{X_{Gp}} = (0.0,0.2,0.4,0.6,0.8,1.0)$, the choice of which depended on conditions for that stochastic sampling experiment. When combined, $X_{Gp}\beta_{X_{Gp}}$ generated no structured patterns of intraspecific variation (when $\beta_{X_{Gp}} = 0.0$), or increasing levels of structured patterns of intraspecific variation (when $\beta_{X_{Gp}} > 0.0$). This enabled the evaluation of type I error and statistical power respectively. Empirical sampling distributions for these effects were also scrutinized, and compared to parametric distributions (theoretical density functions). Finally, additional simulations were performed under a wider set of conditions and using other types of X variables, such as a covariate rather than a grouping factor (see Supporting Information). Computer code and additional details are found in the Supporting Information.

3.2.1 | Results

Our analyses revealed that E-PGLS was capable of detecting patterns of structured variation, and did so with increasing power when stronger within-species patterns were present (Figure 2a). The method also had appropriate type I error when no structured variation was included. Additionally, E-PGLS detected cross-species trend differences, again with appropriate type I error rates and high power (Figure 2a). Unsurprisingly, statistical power increased with an increasing number of species (results not shown), and with greater within-species sampling (n_i) and with balanced data; though statistical power was acceptable under all scenarios examined (see Supporting Information). Finally, empirical sampling distributions for model effects closely matched the parametric F-distributions (Figure 2b,c), indicating that significance levels obtained from these distributions were appropriate. Overall, these analyses verified that E-PGLS was capable of discerning patterns of structured intraspecific variation, and interrogating differences in those patterns across species while accounting for phylogenetic non-independence (i.e. phylogenetically comparing microevolutionary trends across taxa). These results were consistent when using other types of X variables, such as a covariate rather than a grouping factor (see Supporting Information).

4 | EMPIRICAL EXAMPLE: SEXUAL DIMORPHISM IN PUPFISHES

To illustrate the utility of E-PGLS on empirical data, we deployed the approach to interrogate evolutionary patterns of sexual dimorphism in body shape in pupfishes (*Cyprinodon*). Our data were part of a larger study investigating phenotypic diversification in the genus *Cyprinodon*. Briefly, from museum collections we obtained a total 969 individuals representing both males and females from each of seven species. An average of $\mu = 69$ individuals for each Species × Sex combination were measured (range: 35 – 122).

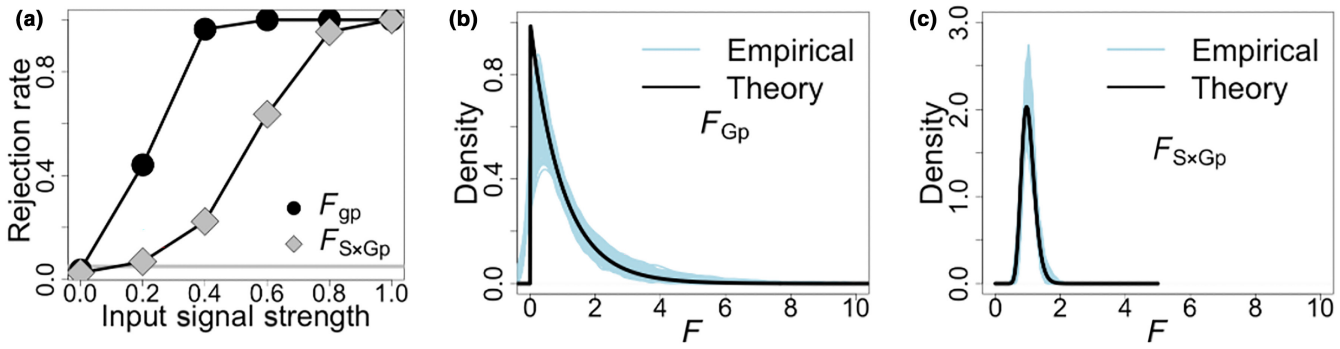

FIGURE 2 Results from stochastic sampling experiments of multivariate phenotypic data containing structured patterns of intraspecific variation (with balanced data and $n_i = 4$) that differed across species and were thus species-specific. (a) Plot of rejection rates for the intraspecific group effect (F_{Gp}) and the species-specific intraspecific group effect ($F_{S\times Gp}$). Rejection rates represent type I error when input signal = 0, and statistical power when input signal >0. Empirical sampling distributions for F-statistics obtained via RRPP for 1000 simulations representing (b) intraspecific group effects (F_{Gp}), and (c) species-specific intraspecific group effects ($F_{S\times Gp}$).

From each individual the locations of 12 landmarks and 74 semilandmarks were then digitized (Figure 3a), and phenotypic data (body shape) were characterized using geometric morphometric methods (Adams et al., 2013; Bookstein, 1991; Mitteroecker & Schæfer, 2022). Here, a generalized Procrustes analysis (Rohlf & Slice, 1990) with minimum bending energy sliding of semilandmarks (Bookstein, 1997) was performed with 10 sliding iterations to eliminate non-shape variation (Figure 3a), and the first 40 PCs, which together represented 99.4% of the total shape variation, were treated as a set of shape variables for all subsequent statistical analyses (see Supporting Information for additional details).

Next, a fossil-calibrated molecular phylogeny was obtained from fishtree (Chang et al., 2019; Rabosky et al., 2018), from which we generated an expanded phylogenetic covariance matrix (Ω_n) for the 969 individuals. We then compared patterns of sexual dimorphism across species conditioned on the phylogeny (under Brownian motion) using the E-PGLS model: $\mathbf{Z} = \mathbf{X}_{sp}\mathbf{B}_{sp} + \mathbf{X}_{sex}\mathbf{B}_{sex} + \mathbf{X}_{sp:sex}\mathbf{B}_{sp:sex} + \mathbf{X}_{pop}\mathbf{B}_{pop} + \mathbf{E}_{\mathcal{N}(0,\Omega_n)}$, where species ($\mathbf{X}_{sp}$), sex ($\mathbf{X}_{sex}$), and their interaction ($\mathbf{X}_{sp:sex}$) were the primary model effects. Because shape variation among populations within species was prevalent, population ($\mathbf{X}_{pop}$) was also included in the model. The likelihood ratio statistic was used for calculation of effect sizes and *p*-values (see Supporting Information for details). To further interrogate within-species patterns of sexual dimorphism (SD) across species a phylogenetically-informed trajectory analysis was performed (Adams & Collyer, 2009; Collyer & Adams, 2013), where multivariate phenotypic vectors representing patterns of sexual dimorphism within each species were characterized, and the degree of SD (magnitude) and direction of SD in morphospace were statistically compared in a phylogenetic context using RRPP. All analyses were performed using the R-packages geomorph 4.0.8 (Adams et al., 2024; Baken et al., 2021) and RRPP 2.0.3 (Collyer & Adams, 2018, 2024), which incorporate sparse-matrix algorithms to assure fast computational results.

4.1 | Results

Our analyses revealed that (1) species variation was significant, though with a weak effect size, and perhaps exceeded what would be expected with a Brownian motion model of evolutionary divergence (although not accounting for phylogeny would produce a larger species effect; see Supporting Information); and (2) within-species patterns of sexual dimorphism were not concordant among species, as indicated by a strong Species × Sex interaction (Table 1). Across the phylogeny there was considerable evolutionary lability

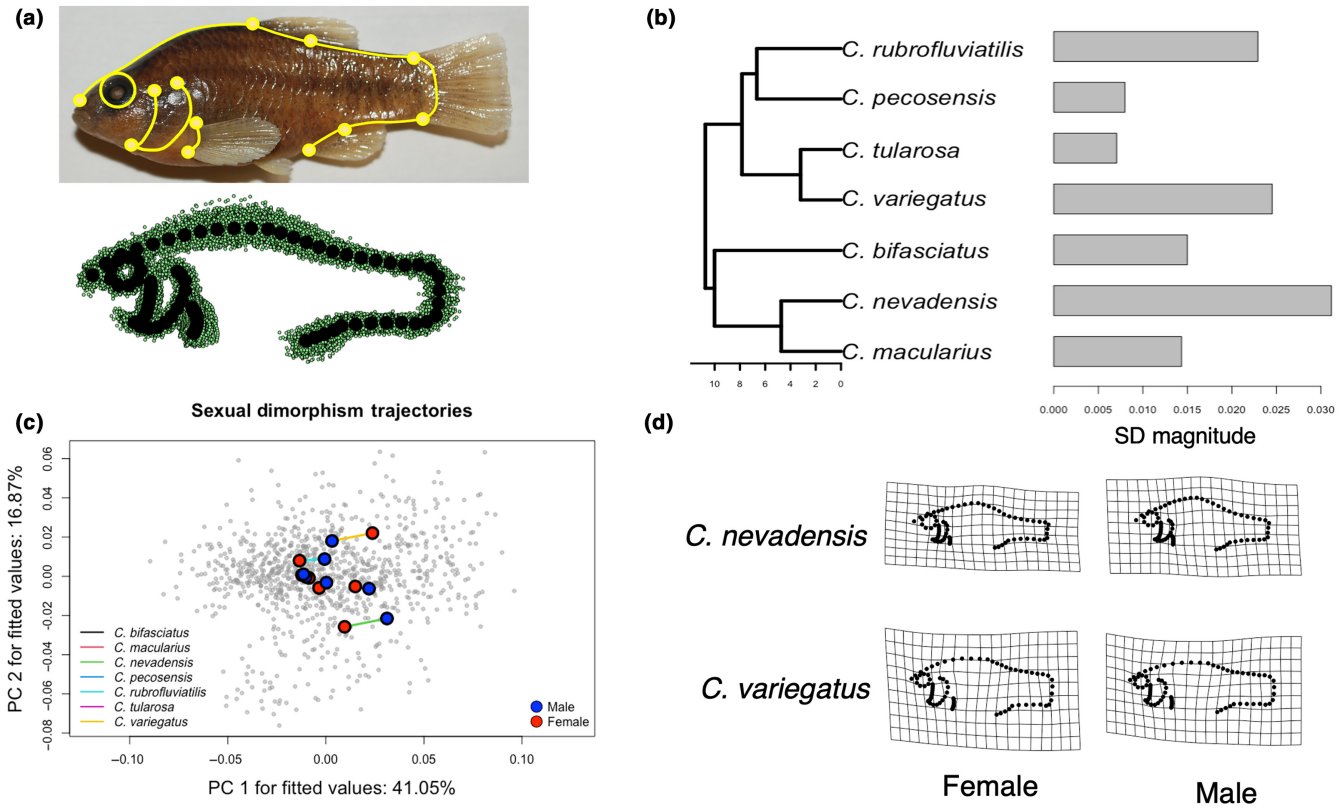

FIGURE 3 Results of comparisons of multivariate phenotypic vectors of sexual dimorphism in *Cyprinodon* pupfish. (a) One example of landmarks and Bézier curves from which semilandmarks were sampled, plus the aligned coordinates for the entire data set, representing shape variation. (b) Time-calibrated phylogeny for the seven species. Horizontal bars represent the magnitude of sexual dimorphism displayed by each species. (c) PC plot of shape variation. Evolutionary trajectories of sexual dimorphism for all species are superimposed in the plot. (d) Thin-plate spline deformation grids describing shape differences between males and females in *C. nevadensis* and *C. variegatus* (shape differences displayed at 2× magnification to facilitate visual interpretation of observed patterns).

in the magnitude of sexual dimorphism (Figure 3b), with some species displaying no detectable levels of sexual dimorphism in body shape (*C. bifasciatus*, *C. pecosensis*, *C. tularosa*), while other species exhibited large and significant patterns of sexual dimorphism (*C. nevadensis*: $Z = 3.77$, $p = 0.001$; *C. variegatus*: $Z = 3.42$, $p = 0.001$; *C. rubrofluviatilis*: $Z = 2.60$; $p = 0.006$). Analysis of evolutionary trajectories further revealed that the mode of within-species patterns differed among taxa. Notably, the species exhibiting the largest levels of sexual dimorphism did so in opposing directions in morphospace (Figure 3c). On one hand, the sexual shape dimorphism in *C. nevadensis* and *C. rubrofluviatilis* was characterized as a contrast between relatively deeper bodied males and more slender females (Figure 3d). Such patterns have previously been observed in *Cyprinodon* (Collyer et al., 2005, 2007, 2015), and may indicate selection for deeper, more laterally compressed body shapes associated with the defence of breeding territories (Collyer et al., 2015). By contrast, sexual dimorphism in *C. variegatus* displayed the reverse trend, where females were relatively deeper bodied and males exhibited a more slender phenotype (Figure 3d). Such patterns may indicate predator-mediated selection, where the presence of predators generates selection on males for enhanced predator avoidance, and thus a more streamlined body form (Collyer et al., 2015; Langerhans et al., 2004; Vinterstare et al., 2022). Finally, the variation observed in sexual dimorphism across the phylogeny (Figure 3c) suggests the hypothesis that multiple selection factors, stemming from local environmental conditions, and distinct predatory influences, interact with one another, and concomitantly affect male and female body shapes, resulting in dynamic evolutionary shifts in the magnitude and direction of sexual dimorphism across species. Interestingly, *C. variegatus* is a coastal species with a distribution spanning the Atlantic coast and Gulf of Mexico in North America. All other species in this study have small, limited, inland distributions in North America. Diversification of inland species is believed to have followed vicariance after Pleistocene formations of large, pluvial lakes that connected with the Gulf of Mexico, allowing colonization by *C. variegatus*-like ancestors (Echelle et al., 2005). The results here suggest that vicariance in non-marine environments could spur an evolutionary reversing of the typical marine sexual dimorphism in body shape in *Cyprinodon*. Overall, our analyses provided strong evidence that evolutionary trends in sexual dimorphism were species-specific in *Cyprinodon* and demonstrated that E-PGLS was capable of revealing differences in intraspecific trends across the tree of life, thereby yielding new macroevolutionary insights.

TABLE 1 MANOVA table of variance components from E-PGLS.

	Wilks	LR	*Z*	*p*-value
Species	0.648	420.769	2.221	0.020
Sex	0.927	73.864	3.132	0.002
Population	0.049	2912.715	22.501	0.001
Species:Sex	0.540	597.515	9.656	0.001
Full.Model	0.016	4014.514	4.062	0.001

Note: Multivariate dispersion is summarized using the likelihood ratio statistic (LR), which is a transformation of the Wilks' Λ statistic. All variance components are obtained while conditioning the individual-level data on the phylogeny (via Ω_n), as per Equations (1) and (2).

5 | DISCUSSION

Evolutionary biologists routinely describe phenotypic trends across species using phylogenetic comparative methods (PCMs), which condition the analysis on the phylogeny to account for the evolutionary non-independence among taxa (Felsenstein, 1985; Harmon, 2019; Harvey & Pagel, 1991; Martins & Hansen, 1997; O'Meara, 2012). However, while these approaches afford empiricists the ability to decipher macroevolutionary trends across species means, or across species variances (e.g. Gaboriau et al., 2020, 2024; Kostikova et al., 2016), it has not been possible to compare within-species (microevolutionary) patterns across species in an analogous phylogenetic context. In this article, we developed an extended phylogenetic generalized least squares (E-PGLS) procedure for characterizing within-species patterns of phenotypic variation, and for comparing these patterns among species while accounting for their phylogenetic non-independence. The approach utilizes a hierarchical linear model, permutation procedures, and an expanded phylogenetic covariance matrix (Ω_n) to enable the analysis of individual-level patterns across the tree of life. In so doing, E-PGLS furthers the capabilities of phylogenetic comparative analyses in several important evolutionary arenas.

E-PGLS facilitates the characterization of macroevolutionary trends in univariate, multivariate, or high-dimensional phenotypic traits, while properly accounting for variation among multiple individuals within species. Previous approaches that incorporate intraspecific variation in phylogenetic comparative analyses were largely developed considering single-valued (univariate) phenotypes (e.g. Hansen & Bartoszek, 2012; Ives et al., 2007; Kostikova et al., 2016; Revell & Reynolds, 2012; Villemereuil & Nakagawa, 2014), though some explicitly considered multivariate phenotypic data (Felsenstein, 2008). While most could be extended (in theory) to the multivariate case for estimating model coefficients, statistical evaluation of those models using parametric techniques based on maximum likelihood (or some variant) would be limited to those cases where $N >> p$. The reason is that for cases where the data are rank-deficient (because of linear dependencies among phenotypic variables or where $p > N$), statistical summaries cannot be completed without some form of covariance regularization or dimension reduction, as obtaining the likelihood requires both the determinant and the inverse of the $p \times p$ trait covariance matrix, and neither can be computed (Adams, 2014b; Adams & Collyer, 2018a, 2018b, 2019b). Additionally, even when such adjustments to the data are made, as with other PCMs evaluated parametrically, their type I error rates will still be elevated as the number of trait dimensions (p) increases (Adams, 2014b; Adams & Collyer, 2018a). This limitation also precludes the parametric

evaluation of high-dimensional data (where $p > > N$), which are increasingly used in evolutionary biology. By contrast, E-PGLS does not suffer these restrictions, as model evaluation is based on summary statistics that are robust to these issues, and empirical sampling distributions are obtained using permutation (RRPP) rather than integrating parametric probability density functions, only for cases that have more observations than variables. Thus, E-PGLS is capable of extending comparative analyses that incorporate within-species variation into the realm of high-dimensional multivariate phenotypes.

One important consideration when conducting macroevolutionary comparisons of microevolutionary trends is sampling effort. Clearly, such an analysis requires that multiple individuals per species are quantified so that within-species trends are adequately characterized. Therefore, when possible, it is paramount that empiricists maximize within-species sampling effort, so that a sufficient number of individuals per species is examined. However, this can sometimes be problematic, as sampling limitations due to lack of specimens from museum collections, or uneven field sampling across species, can result in imbalanced datasets. Fortunately, our simulation results revealed that E-PGLS was capable of detecting differences among within-species trends when they were present, even at low sample sizes and with imbalanced data. Predictably, our simulations revealed that as the number of species sampled increased, or as the within-species sample size increased, statistical power to detect known patterns also increased. Likewise, balanced data exhibited slightly higher power than imbalanced data (see Supporting Information). However, E-PGLS retained solid performance even at sub-optimal sampling levels. Thus, while in an ideal world researchers should strive to maximize the number of individuals measured per species, attain balanced data between species (i.e. the same n_i per species), and include as many species as is possible, our work has shown that E-PGLS is capable of characterizing and interrogating biological patterns with high statistical power, even when intraspecific sample sizes are imbalanced and low.

With E-PGLS, empiricists may compare within-species patterns across species while taking into consideration their lack of independence due to phylogenetic relatedness. One interpretation of this approach is that E-PGLS provides a means of comparing microevolutionary patterns across the phylogeny. Viewed from this perspective, one can conceptualize the within-species (microevolutionary) pattern as a 'trait', and consider its evolutionary change through time. Indeed, our macroevolutionary interrogation of the pupfish dataset revealed that patterns of sexual dimorphism within species were not constant, but rather evolved considerably across the phylogeny. While our focus with that example was to determine how patterns of sexual dimorphism differed across taxa, other evolutionary investigations could be envisioned. For instance, because species share common ancestry, one might expect that closely related species display similar within-species patterns. In this case, within-species (microevolutionary) trends would appear clumped on the phylogeny, and may thus exhibit phylogenetic signal (sensu Blomberg et al., 2003; Collyer et al., 2022). Conceptually, one could envision a formal test of phylogenetic signal of these patterns, provided the within-species trend for each species could be adequately characterized by an appropriate summary measure (i.e. an effect size, see Adams & Collyer, 2016, 2019a). Additionally, E-PGLS could serve as a means of revealing such patterns, if for instance, groups of closely related species were found to differ in their microevolutionary patterns from other species on the phylogeny. As such, E-PGLS opens the door for further studies that may reveal not only how patterns of within-species variation differs among species and clades, but why such patterns evolve.

Fundamentally, E-PGLS has two primary purposes. First, it enables the implementation of phylogenetic least-squares estimation when more than one individual per species is found in the dataset. Second, E-PGLS facilitates the comparison of within-species patterns across species, while accounting for among-species variation, as well as accounting for the phylogenetic non-independence of those species. For both scenarios, it is imperative that within-species variation can be quantified and potentially assessed despite variation among species. With respect to the former, E-PGLS provides a means of evaluating comparative trends across species while accounting for variation among individuals within species. Here, variation among individuals is treated as random (nonstructured) measurement error around the mean phenotype. Our simulation results revealed that in such cases, estimates of evolutionary trait covariance obtained from E-PGLS are comparable with those obtained from existing approaches that incorporate intraspecific variation (Felsenstein, 2008).[1] Thus, E-PGLS is equally appropriate for the analysis of such data, and may be used to accommodate multiple individuals per species in a phylogenetic comparative analysis whose focus is on macroevoluationary trends across the tree of life.

With respect to the latter, when within-species variation is structured, E-PGLS enables the comparison of within-species (microevolutionary) trends across taxa, and does so while accounting for their phylogenetic non-independence. Prior approaches that incorporated intraspecific variation in PCMs (e.g. Felsenstein, 2008; Gaboriau et al., 2020, 2024; Ives et al., 2007; Revell & Reynolds, 2012) did not treat such variation as structured, and were thus incapable of determining whether microevolutionary trends within species differed across species or across the phylogeny. By contrast, our findings demonstrate that E-PGLS is more flexible, and can not only evaluate comparative trends while incorporating intraspecific variation, but can also interrogate evolutionary scenarios where structured intraspecific variation may be modelled and compared among species. This latter capability of E-PGLS enables biologists to determine whether microevolutionary within-species patterns *change* across the phylogeny, or whether groups of taxa display similar within-species trends; all while taking phylogenetic non-independence into

[1]Note that treating structured variation as random variation has no effect on the performance of prior approaches (results not shown).

consideration. Thus when used in this manner, E-PGLS is the interrogation of microevolutionary data sampled within species with a generalized least-squares method of estimation based on relatedness defined at the species level. This novel application opens the door to evaluating how microevolutionary patterns evolve at macroevolutionary scales. As such, E-PGLS provides the ability to decipher complex macroevolutionary dynamics of microevolutionary trends, and discern how those trends evolve across the tree of life: something that is inaccessible via standard phylogenetic comparative approaches. This represents an important advance in phylogenetic comparative methods: the ability to compare within-species trends across species while simultaneously accounting for their non-independence due to shared evolutionary history.

AUTHOR CONTRIBUTIONS

Dean C. Adams and Michael L. Collyer contributed equally to developing the approach for comparing within species patterns phylogenetically, for implementing and testing the procedure, and contributed to all portions of this manuscript. All authors approve of the final product and are willingly accountable for any portion of the content.

ACKNOWLEDGEMENTS

This work was sponsored in part by National Science Foundation Grants DBI-1902511 and DEB-2140720 (to D.C.A.) and DBI-1902694 and DEB-2146220 (to M.L.C.). Empirical data used collections from the Museum of Southwestern Biology, Albuquerque, NM, and the University of Michigan Museum of Zoology, Ann Arbor, MI.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data available via the Dryad Digital Repository https://doi.org/10.5061/dryad.np5hqc00h (Adams & Collyer, 2024). R-scripts used for simulation tests are provided as Supporting Information.

ORCID

Dean C. Adams https://orcid.org/0000-0001-9172-7894
Michael L. Collyer https://orcid.org/0000-0003-0238-2201

REFERENCES

Adams, D. C. (2014a). A generalized kappa statistic for estimating phylogenetic signal from shape and other high-dimensional dultivariate data. *Systematic Biology*, *63*, 685–697. https://doi.org/10.1093/sysbio/syu030

Adams, D. C. (2014b). A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution*, *68*, 2675–2688. https://doi.org/10.1111/evo.12463

Adams, D. C., & Collyer, M. L. (2009). A general framework for the analysis of phenotypic trajectories in evolutionary studies. *Evolution*, *63*, 1143–1154. https://doi.org/10.1111/j.1558-5646.2009.00649.x

Adams, D. C., & Collyer, M. L. (2015). Permutation tests for phylogenetic comparative analyses of high-dimensional shape data: What you shuffle matters. *Evolution*, *69*, 823–829. https://doi.org/10.1111/evo.12596

Adams, D. C., & Collyer, M. L. (2016). On the comparison of the strength of morphological integration across morphometric datasets. *Evolution*, *70*, 2623–2631.

Adams, D. C., & Collyer, M. L. (2018a). Multivariate phylogenetic comparative methods: Evaluations, comparisons, and recommendations. *Systematic Biology*, *67*(1), 14–31.

Adams, D. C., & Collyer, M. L. (2018b). Phylogenetic ANOVA: Group-clade aggregation, biological challenges, and a refined permutation procedure. *Evolution*, *72*(6), 1204–1215. https://doi.org/10.1111/evo.13492

Adams, D. C., & Collyer, M. L. (2019a). Comparing the strength of modular signal, and evaluating alternative modular hypotheses, using covariance ratio effect sizes with morphometric data. *Evolution*, *73*(12), 2352–2367.

Adams, D. C., & Collyer, M. L. (2019b). Phylogenetic comparative methods and the evolution of multivariate phenotypes. *Annual Review of Ecology, Evolution, and Systematics*, *50*, 405–425.

Adams, D. C., & Collyer, M. L. (2022). Consilience of methods for phylogenetic analysis of variance. *Evolution*, *76*, 1406–1419. https://doi.org/10.1111/evo.14512

Adams, D. C., & Collyer, M. L. (2024). Data from: Extending phylogenetic regression models for comparing within-species patterns across the tree of life. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.np5hqc00h

Adams, D. C., Collyer, M. L., Kaliontzopoulou, A., & Baken, E. K. (2024). *Geometric morphometric analyses of 2D and 3D landmark data, version 4.0.8*. R Foundation for Statistical Computing. Retrieved from https://cran.r-project.org/package=geomorph

Adams, D. C., Rohlf, F. J., & Slice, D. E. (2013). A field comes of age: Geometric morphometrics in the 21st century. *Hystrix*, *24*, 7–14. https://doi.org/10.4404/hystrix-24.1-6283

Anderson, M. J., & Robinson, J. (2001). Permutation tests for linear models. *Australian & New Zealand Journal of Statistics*, *43*, 75–88.

Baken, E. K., Collyer, M. L., Kaliontzopoulou, A., & Adams, D. C. (2021). Geomorph 4.0 and gmShiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods in Ecology and Evolution*, *12*, 2355–2363. https://doi.org/10.1111/2041-210X.13723

Beaulieu, J. M., Jhwueng, D.-C., Boettiger, C., & O'Meara, B. C. (2012). Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution*, *66*, 2369–2383. https://doi.org/10.1111/j.1558-5646.2012.01619.x

Beausoleil, M.-O., Carrión, P. L., Podos, J., Camacho, C., Rabadán-González, J., Richard, R., Lalla, K., Raeymaekers, J. A. M., Knutie, S. A., De León, L. F., Chaves, J. A., Clayton, D. H., Koop, J. A. H., Sharpe, D. M. T., Gotanda, K. M., Huber, S. K., Barrett, R. D. H., & Hendry, A. P. (2023). The fitness landscape of a community of darwin's finches. *Evolution*, *77*, 2533–2546. https://doi.org/10.1093/evolut/qpad160

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, *57*, 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Blomberg, S. P., Lefevre, J. G., Wells, J. A., & Waterhouse, M. (2012). Independent contrasts and PGLS regression estimators are equivalent. *Systematic Biology*, *61*, 382–391. https://doi.org/10.1093/sysbio/syr118

Bookstein, F. L. (1991). *Morphometric tools for landmark data: Geometry and biology*. Cambridge University Press.

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Medical Image Analysis*, *1*, 225–243. https://doi.org/10.1016/s1361-8415(97)85012-8

Burns, M. D., Friedman, S. T., Corn, K. A., Larouche, O., Price, S. A., Wainwright, P. C., & Burress, E. D. (2024). High-latitude ocean habitats are a crucible of fish body shape diversification. *Evolution Letters*, *8*, 669–679. https://doi.org/10.1093/evlett/qrae020

Butler, M. A., & King, A. A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American Naturalist*, *164*, 683–695. https://doi.org/10.1086/426002

Chang, J., Rabosky, D. L., Smith, S. A., & Alfaro, M. E. (2019). An R package and online resource for macroevolutionary studies using the ray-finned fish tree of life. *Methods in Ecology and Evolution*, *10*, 1118–1124. https://doi.org/10.1111/2041-210x.13182

Clavel, J., Escarguel, G., & Merceron, G. (2015). MvMORPH: An R package for fitting multivariate evolutionary models to morphometric data. *Methods in Ecology and Evolution*, *6*, 1311–1319. https://doi.org/10.1111/2041-210x.12420

Collyer, M. L., & Adams, D. C. (2013). Phenotypic trajectory analysis: Comparison of shape change patterns in evolution and ecology. *Hystrix, The Italian Journal of Mammalogy*, *24*, 75–83. https://doi.org/10.4404/hystrix-24.1-6298

Collyer, M. L., & Adams, D. C. (2018). RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution*, *9*, 1772–1779. https://doi.org/10.1111/2041-210X.13029

Collyer, M. L., & Adams, D. C. (2024). *RRPP: Linear model evaluation with randomized residuals in a permutation procedure, version 2.0.3*. R Foundation for Statistical Computing. Retrieved from https://cran.r-project.org/package=RRPP

Collyer, M. L., Baken, E. K., & Adams, D. C. (2022). A standardized effect size for evaluating and comparing the strength of phylogenetic signal. *Methods in Ecology and Evolution*, *13*, 367–382. https://doi.org/10.1111/2041-210x.13749

Collyer, M. L., Novak, J. M., & Stockwell, C. A. (2005). Morphological divergence of native and recently established populations of white sands pupfish (*Cyprinodon tularosa*). *Copeia*, *2005*, 1–11. https://doi.org/10.1643/cg-03-303r1

Collyer, M. L., Sekora, D. J., & Adams, D. C. (2015). A method for analysis of phenotypic change for phenotypes described by high-dimensional data. *Heredity*, *115*, 357–365. https://doi.org/10.1038/hdy.2014.75

Collyer, M. L., Stockwell, C. A., Adams, D. C., & Reiser, M. H. (2007). Phenotypic plasticity and contemporary evolution in introduced populations: Evidence from translocated populations of white sands pupfish (*Cyrpinodon tularosa*). *Ecological Research*, *22*, 902–910. https://doi.org/10.1007/s11284-007-0385-9

de Villemereuil, P., & Nakagawa, S. (2014). General quantitative genetic methods for comparative biology. In *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 287–303). Springer.

Echelle, A. A., Carson, E. W., Echelle, A. F., Van Den Bussche, R., Dowling, T. E., & Meyer, A. (2005). Historical biogeography of the new-world pupfish genus cyprinodon (Teleostei: Cyprinodontidae). *Copeia*, *2005*(2), 320–339.

Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, *125*, 1–15. https://doi.org/10.1086/284325

Felsenstein, J. (2008). Comparative methods with sampling error and within-species variation: Contrasts revisited and revised. *The American Naturalist*, *171*(6), 713–725. https://doi.org/10.1086/587525

Gaboriau, T., Mendes, F. K., Joly, S., Silvestro, D., & Salamin, N. (2020). A multi-platform package for the analysis of intra- and interspecific trait evolution. *Methods in Ecology and Evolution*, *11*, 1439–1447. https://doi.org/10.1111/2041-210x.13458

Gaboriau, T., Tobias, J. A., Silvestro, D., & Salamin, N. (2024). Exploring the macroevolutionary signature of asymmetric inheritance at speciation. *Systematic Biology*, *73*, syae043. https://doi.org/10.1093/sysbio/syae043

Garamszegi, L. Z. (2014). Uncertainties due to within-species variation in comparative studies: Measurement errors and statistical weights. In *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 157–199). Springer.

Garland, T., & Ives, A. R. (2000). Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. *The American Naturalist*, *155*, 346–364. https://doi.org/10.1086/303327

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society: Series B: Methodological*, *53*(2), 285–321.

Grafen, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *326*, 119–157. https://doi.org/10.1098/rstb.1989.0106

Hadfield, J. D., & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, *23*, 494–508. https://doi.org/10.1111/j.1420-9101.2009.01915.x

Hansen, T. F., & Bartoszek, K. (2012). Interpreting the evolutionary regression: The interplay between observational and biological errors in phylogenetic comparative studies. *Systematic Biology*, *61*, 413–425. https://doi.org/10.1093/sysbio/syr122

Harmon, L. J. (2019). *Phylogenetic comparative methods: Learning from trees*.

Harmon, L. J., & Losos, J. B. (2005). The effect of intraspecific sample size on type i and type II error rates in comparative studies. *Evolution*, *59*, 2705–2710. https://doi.org/10.1554/05-224.1

Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford University Press.

Hunt, E. S. E., Felice, R. N., Tobias, J. A., & Goswami, A. (2023). Ecological and life-history drivers of avian skull evolution. *Evolution*, *77*, 1720–1729. https://doi.org/10.1093/evolut/qpad079

Ives, A. R., Midford, P. E., & Garland, T. (2007). Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology*, *56*, 252–270. https://doi.org/10.1080/10635150701313830

Johnston, J., & DiNardo, J. (1997). *Econometric methods*. McGraw Hill Publishing.

Juarez, B. H., & Adams, D. C. (2022). Evolutionary allometry of sexual dimorphism of jumping performance in anurans. *Evolutionary Ecology*, *36*, 717–733. https://doi.org/10.1007/s10682-021-10,132-x

Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T.-C. (1985). *The theory and practice of econometrics*. John Wiley & Sons.

Kariya, T., & Kurata, H. (2004). *Generalized least squares*. John Wiley & Sons.

Khabbazian, M., Kriebel, R., Rohe, K., & Ané, C. (2016). Fast and accurate detection of evolutionary shifts in ornstein–uhlenbeck models. *Methods in Ecology and Evolution*, *7*, 811–824. https://doi.org/10.1111/2041-210x.12534

Klingenberg, C. P., & McIntyre, G. S. (1998). Geometric morphometrics of developmental instability: Analyzing patterns of fluctuating asymmetry with procrustes methods. *Evolution*, *52*(5), 1363–1375.

Kostikova, A., Silvestro, D., Pearman, P. B., & Salamin, N. (2016). Bridging inter- and intraspecific trait evolution with a hierarchical bayesian approach. *Systematic Biology*, *65*, 417–431. https://doi.org/10.1093/sysbio/syw010

Langerhans, R. B., Layman, C. A., Shokrollahi, A. M., & DeWitt, T. J. (2004). Predator-driven diversification in *Gambusia affinis*. *Evolution*, *58*, 2305–2318. https://doi.org/10.1111/j.0014-3820.2004.tb01605.x

López-Fernández, H., Arbour, J., Willis, S., Watkins, C., Honeycutt, R. L., & Winemiller, K. O. (2014). Morphology and efficiency of a specialized foraging behavior, sediment sifting, in neotropical cichlid fishes. *PLoS One*, *9*, e89832. https://doi.org/10.1371/journal.pone.0089832

Lynch, M. (1991). Methods for the analysis of comparative data in evolutionary biology. *Evolution*, *45*, 1065–1080. https://doi.org/10.1111/j.1558-5646.1991.tb04375.x

Martins, E. P., & Hansen, T. F. (1997). Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist*, *149*, 646–667. https://doi.org/10.1086/286013

Mitov, V., Bartoszek, K., & Stadler, T. (2019). Automatic generation of evolutionary hypotheses using mixed gaussian phylogenetic models. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 16921–16926. https://doi.org/10.1073/pnas.1813823116

Mitteröcker, P., & Schæfer, K. (2022). Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. *American Journal of Biological Anthropology*, *178*, 181–210. https://doi.org/10.1002/ajpa.24531

O'Meara, B. C. (2012). Evolutionary inferences from phylogenies: A review of methods. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 267–285. https://doi.org/10.1146/annurev-ecolsys-110,411-160,331

O'Meara, B. C., Ané, C., Sanderson, M. J., & Wainwright, P. C. (2006). Testing for different rates of continuous trait evolution using likelihood. *Evolution*, *60*, 922–933. https://doi.org/10.1111/j.0014-3820.2006.tb01171.x

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*, 392–395. https://doi.org/10.1038/s41586-018-0273-1

Rencher, A. C. (2000). *Linear models in statistics*. John Wiley & Sons.

Revell, L. J., & Harmon, L. J. (2008). Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. *Evolutionary Ecology Research*, *10*, 311–331.

Revell, L. J., & Reynolds, R. G. (2012). A new Bayesian method for fitting evolutionary models to comparative data with intraspecific variation. *Evolution*, *66*, 2697–2707. https://doi.org/10.1111/j.1558-5646.2012.01645.x

Reyes-Puig, C., Adams, D. C., Enriquez-Urzelai, U., & Kaliontzopoulou, A. (2023). Rensch's rule: Linking intraspecific to evolutionary allometry. *Evolution*, *77*, 2576–2589. https://doi.org/10.1093/evolut/qpad172

Rohlf, F. J., & Slice, D. E. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, *39*, 40–59. https://doi.org/10.2307/2992207

Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: A phylomorphospace approach. *Evolution*, *62*, 3135–3156. https://doi.org/10.1111/j.1558-5646.2008.00519.x

Silvestro, D., Kostikova, A., Litsios, G., Pearman, P. B., & Salamin, N. (2015). Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*, *6*, 340–346. https://doi.org/10.1111/2041-210x.12337

Stroud, J. T., Moore, M. P., Langerhans, R. B., & Losos, J. B. (2023). Fluctuating selection maintains distinct species phenotypes in an ecological community in the wild. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(42), e2222071120. https://doi.org/10.1073/pnas.2222071120

Tejero-Cicuéndez, H., Menéndez, I., Talavera, A., Mochales-Riaño, G., Burriel-Carranza, B., Simó-Riudalbas, M., Carranza, S., & Adams, D. C. (2023). Evolution along allometric lines of least resistance: Morphological differentiation in *Pristurus* geckos. *Evolution*, *77*, 2547–2560. https://doi.org/10.1093/evolut/qpad166

Uyeda, J. C., & Harmon, L. J. (2014). A novel bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology*, *63*, 902–918. https://doi.org/10.1093/sysbio/syu057

Vinterstare, J., Brönmark, C., Nilsson, P. A., Langerhans, R. B., Chauhan, P., Hansson, B., & Hulthén, K. (2022). Sex matters: Predator presence induces sexual dimorphism in a monomorphic prey, from stress genes to morphological defences. *Evolution*, *77*, 304–317. https://doi.org/10.1093/evolut/qpac030

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting information S1. Supporting Information contains additional analytical details of the method and additional simulations verifying its statistical properties.

How to cite this article: Adams, D. C., & Collyer, M. L. (2024). Extending phylogenetic regression models for comparing within-species patterns across the tree of life. *Methods in Ecology and Evolution*, *15*, 2234–2246. https://doi.org/10.1111/2041-210X.14438