BIOLOGICAL REVIEWS
Cambridge Philosophical Society
Biol. Rev. (2025), **100**, pp. 1294–1316.
doi: 10.1111/brv.70001

Multi-response phylogenetic mixed models: concepts and application

Ben Halliwell[1,2,*], Barbara R. Holland[1,2] and Luke A. Yates[1,2]

[1] *School of Natural Sciences, Private Bag 55, University of Tasmania, Hobart, Tasmania, Australia*
[2] *ARC Centre of Excellence for Plant Success in Nature and Agriculture, Private Bag 55, University of Tasmania, Hobart, Tasmania, Australia*

ABSTRACT

The scale and resolution of trait databases and molecular phylogenies is increasing rapidly. These resources permit many open questions in comparative biology to be addressed with the right statistical tools. Multi-response (MR) phylogenetic mixed models (PMMs) offer great potential for multivariate analyses of trait evolution. While flexible and powerful, these methods are not often employed by researchers in ecology and evolution, reflecting a specialised and technical literature that creates barriers to usage for many biologists. Here we present a practical and accessible guide to MR-PMMs. We begin with a review of single-response (SR) PMMs to introduce key concepts and outline the limitations of this approach for characterising patterns of trait coevolution. We emphasise MR-PMMs as a preferable approach for analyses involving multiple species traits, due to the explicit decomposition of trait covariances. We discuss multilevel models, multivariate models of evolution, and extensions to non-Gaussian response traits. We highlight techniques for causal inference using graphical models, as well as advanced topics including prior specification and latent factor models. Using simulated data and visual examples, we discuss interpretation, prediction, and model validation. We implement many of the techniques discussed in example analyses of plant functional traits to demonstrate the general utility of MR-PMMs in handling complex real-world data sets. Finally, we discuss the emerging synthesis of comparative techniques made possible by MR-PMMs, highlight strengths and weaknesses, and offer practical recommendations to analysts. To complement this material, we provide online tutorials including side-by-side model implementations in two popular R packages, *MCMCglmm* and *brms*.

Key words: evolutionary ecology, trait evolution, phylogenetic comparative methods, multivariate statistics, variance partitioning, generalised linear mixed models.

CONTENTS

* Author for correspondence (Tel.: +61 477 908 879; E-mail: benjamin.halliwell@utas.edu.au).

I. INTRODUCTION

The motivation for developing multivariate phylogenetic comparative methods is now broadly appreciated (Adams & Collyer, 2018; Uyeda, Caetano & Pennell, 2015; Uyeda, Zenil-Ferguson & Pennell, 2018; Garamszegi, 2014). Modern methods aim to move beyond simple phylogenetic regression to analyses capable of evaluating the strength, direction, and conservatism of relationships within networks of continuous and discrete variables (Westoby *et al.*, 2023; Haba & Kutsukake, 2019; Hadfield, 2010). These multivariate techniques are applicable to a broad range of species traits, from morphology, physiology, and behaviour to environmental tolerance limits and niche characteristics. Furthermore, as the tide of 21st-century data collection erodes historical constraints on the scale and complexity of analyses, opportunities to apply these methods are increasing. Advances in climate modelling, remote-sensing technology, and collaborative data projects are accelerating the availability and resolution of suitable data sets (Green *et al.*, 2022; Herberstein *et al.*, 2022; Falster *et al.*, 2021; Kattge *et al.*, 2020); the genomic revolution continues to provide more accurate and complete phylogenies (Young & Gillung, 2020; Laumer *et al.*, 2019; Smith & Brown, 2018; Yeates *et al.*, 2016); and the computational resources necessary to pursue fully Bayesian analyses are now accessible to most researchers. Despite this coalescence of opportunities, challenges in implementing modern methods create barriers to usage for many biologists.

A robust literature developing statistical models of multivariate trait evolution has existed for some time. However, new techniques are often slow to enter research practice, especially where they demand a technical or conceptual leap from users. Development of the generalised least-squares framework helped to bring phylogenetic regression and correlative analyses of trait evolution into the mainstream (Rohlf, 2001; Martins & Hansen, 1997; Lynch, 1991; Grafen, 1989; Felsenstein, 1985). Subsequent extensions to phylogenetic generalised linear mixed models (PMMs) provided additional benefits, including support for non-Gaussian traits, flexibility in the specification of hierarchical group effects, and a growing familiarity with mixed models among researchers in ecology and evolution (Ives & Helmus, 2011; Ives & Garland Jr, 2010; Bolker *et al.*, 2009; Housworth, Martins & Lynch, 2004). Currently, phylogenetic mixed modelling also supports multi-response (MR) analyses, facilitating the estimation (and decomposition) of covariances between discrete and continuous species traits (Hadfield, 2010; Hadfield & Nakagawa, 2010).

The development of user-friendly software for fitting MR-PMMs (Bürkner, 2017; Hadfield, 2010) has prompted researchers from diverse scientific fields to implement the method: in anthropology to examine coevolution between climate and cranial form among neolithic humans (Katz, Grote & Weaver, 2016); in animal behaviour to study the evolution of multivariate behavioural repertoires (Hernández *et al.*, 2021) and coevolution between group size and social complexity (Downing, Griffin & Cornwallis, 2020); in epidemiology to understand relationships between growth rate, transmission mode, and virulence among pathogens (Leggett *et al.*, 2017); in disease ecology to examine covariance in macro- and micro-parasite species richness

among host species (Gutiérrez, Piersma & Thieltges, 2019); and in evolutionary ecology to study the multivariate evolution of species functional traits, for example, coevolution among milk macro-nutrient concentrations (Blomquist, 2019), climate, life history and vital rates (Kelly *et al.*, 2021), and plant hydraulic traits (Sanchez-Martinez *et al.*, 2020).

Early applications of MR-PMMs to comparative biology involved critical tests of theory made possible by the estimation of phylogenetic covariances (see Section IV.1). Examples include the effects of sexual selection on brain and body size evolution (García-Peña *et al.*, 2013), whether character displacement is driven by trait divergence in allopatry or sympatry (Tobias *et al.*, 2014), and whether low rates of extra-pair paternity, long lifespans, and tolerance of harsh environments represent causes or consequences of transitions to cooperative breeding (Cornwallis *et al.*, 2010; Downing, Cornwallis & Griffin, 2015; Cornwallis *et al.*, 2017). More recently, MR-PMMs have been highlighted as a powerful framework for eco-evolutionary studies incorporating spatio-temporal random effects (Gomes *et al.*, 2023), trait-based models of community assembly (Gallinat & Pearse, 2021), and analyses of function-valued traits (e.g. parameters of species reaction norms) (Pottier *et al.*, 2024), as well as for disentangling correlations between response variables that manifest at different levels within hierarchical data sets (Westoby *et al.*, 2023; also see Downs & Dochtermann, 2014). These strengths of MR-PMMs echo a growing consensus that multivariate statistical approaches are often valuable even when response correlations are not of specific interest; fitting multiple correlated response variables allows for the inclusion of partially missing data and may improve predictive accuracy due to the sharing of information across common grouping variables (see Section V; also see Riley *et al.*, 2017; Pottier *et al.*, 2024).

Despite these benefits, and the example of pioneering researchers, MR-PMMs remain underutilised, even for data sets to which they would be well suited, such as recent large-scale comparative analyses of species traits (e.g. Cássia-Silva *et al.*, 2020; Grossnickle, 2020; Bruelheide *et al.*, 2018; Díaz *et al.*, 2016). An important series of publications highlighting widespread misconceptions about phylogenetic comparative methods, erroneous statistical practices, and the potential for spurious inference, confirm the need for more translational research (Uyeda *et al.*, 2015, 2018; Cooper, Thomas & Fitzjohn, 2016; Revell, 2010; Freckleton, 2009; Revell, Harmon & Collar, 2008). In particular, practical guidance for biologists wanting to apply multivariate methods that provide a more meaningful decomposition of trait relationships is needed to advance the field beyond the limits of univariate approaches (Westoby *et al.*, 2023).

This review aims to bridge the gap between theory and practice in mixed model analyses of trait evolution. We focus on Bayesian implementations of MR-PMMs, because this model class provides a highly flexible framework with general utility for comparative biology, although maximum likelihood implementations do exist for special cases (Mitov *et al.*, 2020; Butler *et al.*, 2023). The capacity for MR-PMMs to provide a more informative analysis of trait (co)evolution compared with single-response PMMs (SR-PMMs), is well appreciated (Housworth *et al.*, 2004; Hadfield & Nakagawa, 2010). However, despite a long history of use among quantitative geneticists (Meyer, 1991), we believe a lack of familiarity, doubts about data requirements, and practical barriers to implementation continue to prevent common usage of MR-PMMs among ecologists and evolutionary biologists. Thus, we begin (Section II) with a brief conceptual introduction to mixed models. We define the basic SR-PMM, the different parameterisations used to quantify phylogenetic signal, and the implementation of alternative models of evolution (i.e. other than pure Brownian motion). After outlining the limitations of SR models for characterising patterns of coevolution between traits, we shift focus to MR-PMMs.

In particular, we highlight the different components of trait covariance that are modelled in MR-PMMs (e.g. phylogenetic and residual), how these covariance structures are specified using covariance matrices, and how these matrices are parameterised to estimate correlations between species traits at different levels in the model hierarchy. We discuss extensions to the basic MR-PMM (Section III) including multilevel models, multivariate models of trait evolution, non-Gaussian response traits, and techniques for causal inference using graphical models. In Section IV we discuss interpretation and explore data simulated from MR-PMMs (Fig. 1) to clarify the distinction between phylogenetic and non-phylogenetic components of trait correlation. In Section V we cover predictive assessment, including methods based on posterior predictive distributions and leave-one-out (LOO) cross validation (CV). In Section VI, we present an example analysis on leaf functional traits in Eucalyptus from the AusTraits database (Falster *et al.*, 2021). In Section VII, we explore a number of extended topics to showcase emerging techniques, including prior specification for shrinkage of target parameters, as well as latent factor methods for dimension reduction. Finally, in Section VIII we discuss the strengths and weaknesses of MR-PMMs, highlight emerging syntheses facilitated by the method, and provide practical advice for analysts.

To support our exposition, we provide online tutorials, including annotated code and mathematical appendices. These tutorials demonstrate how to simulate multivariate trait data containing phylogenetic structure, and fit corresponding MR-PMMs in two popular R packages, *MCMCglmm* (Hadfield, 2010) and *brms* (Bürkner, 2017). They also cover many common tasks and challenges in the MR-PMM workflow, including data cleaning and manipulation, *via* example analyses of real-world data sets.

II. MODELS

(1) Mixed models

Biologists are familiar with structured sampling designs: these apply when each observation from a survey or experiment is

a member of one or more recognised groups. Acknowledging structure in data is important because observations within a group will often be more similar than can be explained by available predictors. In a simple case, the hierarchical structure may be limited to a single random (or group-level) intercept. In the R programming language, a common syntax to express this model would be:

$$\texttt{y} \sim \texttt{x} + (\texttt{1}|\text{group}), \tag{1}$$

where `y` is the observed response, the measured covariate `x` specifies a fixed effect, and (`1`|group) specifies a random intercept (a random effect that adjusts the baseline mean) at the group level. In other words, we model y as a linear combination of effects, including a random effect that accounts for the fact that observations belong to distinct groups.

A more general mathematical form for mixed models is given by,

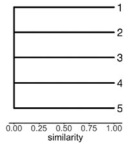
$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{2}$$

where $\mathbf{y}$ is a vector of observations. The design matrices $\mathbf{X}$ and $\mathbf{Z}$ relate fixed and random predictors to the data, the corresponding parameter vectors $\boldsymbol{\beta}$ and $\mathbf{u}$ contain the fixed and random effects to be estimated, and $\mathbf{e}$ is a vector of residual errors.

For the coded model formula (Equation 1), the random intercept (`1`|group) assumes the distribution of the random effect $\mathbf{u}$ is characterised by an identity matrix $\mathbf{I}$ (a matrix with 1 s along the diagonal and 0 s in all off-diagonals) equal in dimension to the number of groups (where $\mathbf{I}_{\text{Ngroup}}$ is of dimension $\text{N}_{\text{group}} \times \text{N}_{\text{group}}$), and scaled by a group-level variance, σ^2_{group}.

For example, for a study including data on five different species,

$$\mathbf{u} \sim \mathcal{N}(0, \sigma^2_{\text{species}}\mathbf{I}_{\text{N}_{\text{species}}}) \rightarrow \mathbf{I}_{\text{N}_{\text{species}}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3}$$



where the notation $\sim$ means "is distributed as" and $\mathcal{N}(\mu,\boldsymbol{\Sigma})$ denotes a (multivariate) normal distribution with mean μ and (co)variance $\boldsymbol{\Sigma}$. The off-diagonal elements of $\mathbf{I}_{\text{Nspecies}}$ are all zero which implies independence between different levels of the grouping factor, such that no pair of species is expected to produce more similar observations than any other. Graphically, an identity matrix can be represented as a comb where tips (group levels) have no shared edges (Equation 3). The residual errors $\mathbf{e}$ are also assumed to be independent *a priori*, capturing observation-specific deviations from the modelled mean,

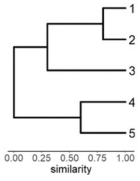
$$\mathbf{e} \sim \mathcal{N}\left(0, \sigma^2_{\text{res}}\mathbf{I}_{\text{N}_{\text{obs}}}\right). \tag{4}$$

Mixed models permit the specification of complex variance structures using multiple random effects which may be given as nested (e.g. populations within species), crossed to act independently, or encoded with additional structure such as spatial or temporal effects. This is achieved by adjusting our assumptions about the distribution of random effects $\mathbf{u}$, for example, by substituting the identity matrix with a covariance matrix derived from the pairwise distances between sampling sites. Structured random effects are employed in this way to model dependencies arising from a range of data-generating processes, including evolutionary processes of species diversification.

(2) Phylogenetic mixed models (PMMs)

In analyses of inter-species data, dependence manifests as phylogenetic signal: the tendency for closely related species to resemble each other due to the hierarchical evolutionary history of life (Felsenstein, 1985; Pagel, 1999). This phylogenetic structure is often represented by a phylogenetic correlation matrix $\mathbf{C} = (c_{ij})$ – an $\text{N}_{\text{species}} \times \text{N}_{\text{species}}$ matrix derived from an ultrametric phylogeny (all paths from root to tip are of equal length). $\mathbf{C}$ encodes our expectation of similarity among species phenotypes, where for a pair of species i and j, the expected correlation is given by the off-diagonal element c_{ij}. In a PMM, $\mathbf{C}$ enters the model as the expected correlation of a species-level random effect scaled by the phylogenetic variance σ^2_{phy}. For example, given a phylogeny of five species, we may write,

$$\mathbf{u} \sim \mathcal{N}(0, \sigma^2_{\text{phy}}\mathbf{C}) \rightarrow \mathbf{C} = \begin{bmatrix} 1 & 0.8 & 0.3 & 0 & 0 \\ 0.8 & 1 & 0.3 & 0 & 0 \\ 0.3 & 0.3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.6 \\ 0 & 0 & 0 & 0.6 & 1 \end{bmatrix}. \tag{5}$$



Importantly, $\mathbf{C}$ is derived by assuming a particular model of evolutionary trait change along branches of the phylogeny [Equation 5 assumes Brownian motion (BM), where c_{ij} represents the shared path length of species i and j, but see Section II.2.*b*]. $\mathbf{C}$ is typically supplied by the user, although joint inference of the phylogenetic tree and the phylogenetic mixed model is possible in some instances, for example using BEAST (Hassler *et al.*, 2023; Bouckaert *et al.*, 2019). In a Bayesian setting, phylogenetic uncertainty can also be incorporated by combining posterior samples across a set of candidate topologies (de Villemereuil *et al.*, 2012; also see Nakagawa & de Villemereuil, 2019).

To summarise, a PMM is a special case of a linear mixed model (Equation 2) that incorporates phylogenetic information *via* an associated correlation matrix, $\mathbf{C}$. Species trait values $\mathbf{y}$ are modelled as the sum of fixed effects, a phylogenetic random effect, and residual error (see Section III.3 for non-Gaussian extensions). Additional random effects are permitted to model other sources of dependence among observations, such as study effects and within-species replication (see Section III.1).

(*a*) *Phylogenetic signal*

A common objective of phylogenetic comparative analyses is to quantify phylogenetic signal in species traits. Numerous

metrics have been developed for this purpose, each with different assumptions and applications (Münkemüller *et al.*, 2012). In a PMM, phylogenetic signal λ describes the proportion of variance in $\mathbf{y}$ attributable to the phylogenetic structure and is expressed in terms of the estimated variance components as

$$\lambda = \frac{\sigma^2_{\text{phy}}}{\sigma^2_{\text{phy}} + \sigma^2_{\text{res}}}. \quad (6)$$

Additional variance components may be included in the denominator of Equation (6) when they capture meaningful components of within- and between-species phenotypic variance, for example trait (co)variance owing to (shared) environmental effects (Kruuk & Hadfield, 2007). When fixed or random variance components are excluded from the denominator, Equation (6) expresses phylogenetic signal conditional on those factors. This is generally desirable for covariates, which are often introduced to account for known biases or aspects of experimental design and thus do not contribute to the variation we seek to decompose (de Villemereuil *et al.*, 2018). For example, when analysing data on morphometric measurements of preserved museum specimens, we may include storage duration as a fixed effect covariate to account for differences in tissue shrinkage among specimens in the collection. Such sources of variation may be considered extraneous to the estimation of λ, and thus excluded from the denominator of Equation (6).

(*b*) *Models of evolution*

Values of $\lambda < 1$ imply a component of variation in $\mathbf{y}$ that cannot be explained by phylogenetic dependence, conditional on a phylogenetic tree and the specified model of evolution. The default model of evolution in a PMM is BM – a continuous-time stochastic process characterised by a random walk of Gaussian distributed increments. The process branches at each node in the phylogeny with the property that trait variance increases linearly with time, meaning that shared branch lengths are proportional to the expected (co) variances of trait values among species. Thus, for BM, there is a direct and simple translation from phylogeny to covariance matrix (Ives, 2018; Felsenstein, 1985).

In the original presentation of phylogenetic generalised least squares (PGLS) (Grafen, 1989), a scalar multiple of the phylogenetic correlation matrix is taken to represent the total covariance among observations, $\boldsymbol{\Sigma} = \sigma^2\mathbf{C}$, which assumes traits have evolved *via* BM on the tree with no additional contributing sources of variation. This assumption is equivalent to fixing $\sigma^2_{\text{res}} = 0$ in Equation (6). With subsequent developments, $\mathbf{C}$, or equivalently the branch lengths of the corresponding phylogenetic tree, is transformed as a function of one or more additional parameters $\boldsymbol{\theta}$, allowing different evolutionary models to be fitted to the data (Symonds & Blomberg, 2014; Harmon *et al.*, 2008). For a model with no additional random effects, this transformed phylogenetic correlation matrix $\mathbf{C}(\boldsymbol{\theta})$, is then multiplied by the scalar variance parameter σ^2 to define

$$\text{var}(\mathbf{u}) = \sigma^2\mathbf{C}(\boldsymbol{\theta}). \quad (7)$$

where var denotes the covariance matrix of a random vector. A common instance is when $\mathbf{C}(\boldsymbol{\theta})$ is a weighted sum of $\mathbf{C}$ and an identity matrix $\mathbf{I}$ (equal in dimension to $\mathbf{C}$) characterised by a single parameter, $\theta = \lambda$ (Pagel, 1999),

$$\mathbf{C}(\lambda) = \lambda\mathbf{C} + (1-\lambda)\mathbf{I}. \quad (8)$$

In this popular formulation of PGLS, the additive terms express the relative proportions, scaled by λ, of the total variance attributable to phylogenetic and residual effects. For the special case of a single observation per species (e.g. a species mean), Equations (7) and (8) can be combined to obtain equivalence with the model given by Equations (5) and (4) for $\sigma^2 = \sigma^2_{\text{phy}} + \sigma^2_{\text{res}}$ and $\lambda = \frac{\sigma^2_{\text{phy}}}{\sigma^2_{\text{phy}} + \sigma^2_{\text{res}}}$ (also see Housworth *et al.*, 2004; Cinar, Nakagawa & Viechtbauer, 2022). This equivalence clarifies the relationships to other common methods: The basic PMM is equivalent to the original presentation of PGLS when $\boldsymbol{\Sigma} = \sigma^2_{\text{phy}}\,\mathbf{C}$ (i.e. $\sigma^2_{\text{res}} = 0$) and to ordinary least squares when $\boldsymbol{\Sigma} = \sigma^2_{\text{res}}\,\mathbf{I}$ (i.e. $\sigma^2_{\text{phy}} = 0$) (Westoby *et al.*, 2023; Blomberg *et al.*, 2012).

In addition to Equation (8), many alternative models for the evolution of continuous traits have been proposed, including those with directional trends, stabilising selection, or changes in evolutionary rates over time (Harmon *et al.*, 2010; Butler & King, 2004; Blomberg, Garland & King, 2003; Pagel, 1999; Hansen, 1997). Phylogenetic mixed modelling supports several popular models [e.g. δ, κ, early burst (EB), Ornstein-Uhlenbeck (OU), accelerating-decelerating (ACDC)] by appropriate transformations to $\mathbf{C}$, as in Equation (7). This extends the estimation of phylogenetic random effects to a range of evolutionary models besides BM (see Tung Ho & Ané, 2014; Blomberg *et al.*, 2003), for which model selection techniques can be used to discriminate among alternative evolutionary hypotheses (Revell & Harmon, 2022; Goolsby, Bruggeman & Ané, 2017). Unfortunately, however, neither *MCMCglmm* nor *brms* currently provide options for estimating the parameters of evolutionary models other than BM.

(*c*) *Limitations of single response models*

PMMs provide an elegant solution to the statistical problem of phylogenetic dependence in inter-species data. However, univariate implementations have important limitations that arise from treating one species trait, often arbitrarily, as the response variable and other traits as predictor variables. In particular, the validity of inferences can be compromised when predictor traits themselves contain phylogenetic signal (Westoby *et al.*, 2023).

Suppose we observe a phenotypic correlation between two traits, $\mathbf{y}_1$ and $\mathbf{y}_2$, across a number of species and that both

traits display phylogenetic signal; a common occurrence in empirical data sets (Adams & Collyer, 2019; Blomberg *et al.*, 2003; Freckleton, Harvey & Pagel, 2002). Part of the covariance between $\mathbf{y}_1$ and $\mathbf{y}_2$ across species is due to a phylogenetically conserved relationship between these traits [e.g. phylogenetic niche conservatism (Losos, 2008; Wiens *et al.*, 2010)], with the remainder due to causes that are independent of phylogeny (e.g. within-species trait covariance). If we model these data using a single-response model (Equation 2), treating $\mathbf{y}_2$ as a fixed predictor of $\mathbf{y}_1$ reduces this composite of correlations operating at the phylogenetic and non-phylogenetic levels to a single-slope parameter. Not only are the separate (co)variance components not recoverable from the model, but phylogenetic signal in the predictor trait $\mathbf{y}_2$ is confounded with (phylogenetic components of) the residual covariance structure [Warton, 2022; also see Wilson (2008), Rausher (1992) and Marques, Kneib & Klein (2022) for analogous issues in quantitative genetics and spatial statistics]. In a worst-case scenario, these two correlation components may cancel each other out, leading to the erroneous conclusion that the traits in question are not meaningfully related (Westoby *et al.*, 2023; also see Figs 4B and 5).

An alternative approach, which reframes this analysis as a multivariate statistical problem, is to treat both traits $\mathbf{y}_1$ and $\mathbf{y}_2$ as response variables. That is, move all traits to the left-hand side of the model equation (Equation 10), into a stacked column-vector $\mathbf{y}=(\mathbf{y_1}^\top,\mathbf{y_2}^\top)^\top=\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}$.

This shift to a multivariate framework has several distinct benefits: (*i*) it avoids often arbitrary decisions about which trait should be considered the response variable and which the predictors; (*ii*) it allows phylogenetic signal to be modelled in all traits simultaneously; (*iii*) it permits a partitioning of trait correlations into phylogenetic and non-phylogenetic components; and (*iv*) it exploits correlations between response traits to improve prediction of missing and/or partially observed new data. Importantly, there are also challenges associated with treating all traits as responses variables. This approach increases model complexity due to the higher-dimensional covariance matrices, which can be problematic when data are limited. Additionally, phylogenetic variance components are typically estimated with greater uncertainty.

(3) Multi-response phylogenetic mixed models (MR-PMMs)

In a MR-PMM, all species traits are modelled jointly as response variables, which means both the random effects $\mathbf{u}$ and residual errors $\mathbf{e}$ associated with each response must also be modelled jointly. For Gaussian response traits, this is achieved directly using multivariate normal distributions (but see Section III.3 for non-Gaussian response traits). Due to the compact form of the mathematical notation, the multi-response model is written identically to the single-response case,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}+\mathbf{e}, \tag{9}$$

where $\mathbf{y}=(\mathbf{y_1}^{\mathrm{T}},\mathbf{y_2}^{\mathrm{T}},\ldots,\mathbf{y}_{\mathrm{N_{traits}}}^{\mathrm{T}})^{\mathrm{T}}$ contains observations for all traits and all species, stacked in a single column vector. To illustrate, we represent the linear predictor for the bivariate case by placing each trait in a separate row,

$$\mathbf{y}=\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}=\begin{pmatrix}\mathbf{X}_1\boldsymbol{\beta}_1+\mathbf{Z}_1\mathbf{u}_1+\mathbf{e}_1\\ \mathbf{X}_2\boldsymbol{\beta}_2+\mathbf{Z}_2\mathbf{u}_2+\mathbf{e}_2\end{pmatrix} \tag{10}$$

with

$$\mathbf{u}=(\mathbf{u}_1^\top,\mathbf{u}_2^\top)^\top\sim\mathcal{N}(0,\boldsymbol{\Sigma}^{\mathrm{phy}}\otimes\mathbf{C}) \tag{11}$$

$$\mathbf{e}=(\mathbf{e}_1^\top,\mathbf{e}_2^\top)^\top\sim\mathcal{N}(0,\boldsymbol{\Sigma}^{\mathrm{res}}\otimes\mathbf{I}_{\mathrm{N_{obs}}}), \tag{12}$$

where $\mathbf{X}_1\boldsymbol{\beta}_1$ and $\mathbf{Z}_1\mathbf{u}_1$ are the linear predictors for the fixed and random effects, respectively, for response trait $\mathbf{y}_1$, and similarly $\mathbf{X}_2\boldsymbol{\beta}_2$ and $\mathbf{Z}_2\mathbf{u}_2$ for $\mathbf{y}_2$. The notation adopted in Equations (11) and (12) emphasises the joint distribution of effects for each response trait, where between-trait covariances across both phylogenetic $\mathbf{u}$ and residual $\mathbf{e}$ components are modelled with the trait covariance matrices

$$\boldsymbol{\Sigma}^{\mathrm{phy}}=\begin{pmatrix}\left(\sigma_1^{\mathrm{phy}}\right)^2 & \sigma_1^{\mathrm{phy}}\sigma_2^{\mathrm{phy}}\rho_{12}^{\mathrm{phy}}\\ \sigma_2^{\mathrm{phy}}\sigma_1^{\mathrm{phy}}\rho_{21}^{\mathrm{phy}} & \left(\sigma_2^{\mathrm{phy}}\right)^2\end{pmatrix}$$

and

$$\boldsymbol{\Sigma}^{\mathrm{res}}=\begin{pmatrix}\left(\sigma_1^{\mathrm{res}}\right)^2 & \sigma_1^{\mathrm{res}}\sigma_2^{\mathrm{res}}\rho_{12}^{\mathrm{res}}\\ \sigma_2^{\mathrm{res}}\sigma_1^{\mathrm{res}}\rho_{21}^{\mathrm{res}} & \left(\sigma_2^{\mathrm{res}}\right)^2\end{pmatrix}, \tag{13}$$

where the off-diagonal matrix elements of $\boldsymbol{\Sigma}^{\mathrm{phy}}$ and $\boldsymbol{\Sigma}^{\mathrm{res}}$ are scalar covariance components between traits *i* and *j*, containing the phylogenetic (ρ_{ij}^{phy}) and residual (ρ_{ij}^{res}) components of trait correlation, respectively. The Kronecker product $\otimes$ is a type of matrix multiplication which, using the phylogenetic component as an example, can be written as

$$\boldsymbol{\Sigma}^{\mathrm{phy}}\otimes\mathbf{C}=\begin{pmatrix}\left(\sigma_1^{\mathrm{phy}}\right)^2\mathbf{C} & \sigma_1^{\mathrm{phy}}\sigma_2^{\mathrm{phy}}\rho_{12}^{\mathrm{phy}}\mathbf{C}\\ \sigma_2^{\mathrm{phy}}\sigma_1^{\mathrm{phy}}\rho_{21}^{\mathrm{phy}}\mathbf{C} & \left(\sigma_2^{\mathrm{phy}}\right)^2\mathbf{C}\end{pmatrix}. \tag{14}$$

The Kronecker product has the effect of introducing phylogenetic dependence, *via* $\mathbf{C}$, into the covariance structure of each trait and each pairwise trait relationship (see tutorial in Section X for fully worked examples). Despite the large dimension of the Kronecker product matrices, the number of estimated (co)variance parameters in a MR-PMM is $\mathrm{N_{trait}}$ ($\mathrm{N_{trait}}$ + 1) (e.g. only six parameters for the bivariate case).

(*a*) *Implications*

Moving from a SR to a MR model facilitates more refined inferences, improves prediction, and allows users to include

all available data, not just complete cases. In terms of inference, the estimated quantity $\widehat{\rho}_{ij}^{\text{phy}} \neq 0$ provides evidence for a component of correlation between traits i and j that is conserved over evolutionary time (i.e. a portion of trait covariance that is associated with phylogeny), while $\widehat{\rho}_{ij}^{\text{res}} \neq 0$ provides evidence for a component of correlation that is independent of phylogeny (see Fig. 1 for an illustration). By contrast, the SR model fails to account for signal in "predictor" traits, leading to a confound between phylogenetic and residual effects (Westoby *et al.*, 2023; also see Section II.2.*c*).

For predictive goals, a MR-PMM has the potential to be both more precise and more accurate than univariate linear models due to sharing of information across levels that accounts for phylogenetic relationships. Prediction of missing trait values leverages all observed data across traits and species, weighting their contributions by the estimated phylogenetic and non-phylogenetic covariances (see Section V). Meaningful reductions in predictive variance compared to SR models may require that traits are at least moderately correlated (Riley *et al.*, 2017; Jackson, Riley & White, 2011), especially on the phylogenetic level, where estimates of covariance generally have higher uncertainty. However, conserved trait correlations are not uncommon in nature (Westoby *et al.*, 2023), with comparative studies often investigating hypotheses about coordination (e.g. trade-offs, co-selection) among species traits that display a strong phylogenetic signal (e.g. Kelly *et al.*, 2021; Sanchez-Martinez *et al.*, 2020; Blomquist, 2019). For example, recent work highlights the utility of MR-PMMs for analysing function-valued traits such as

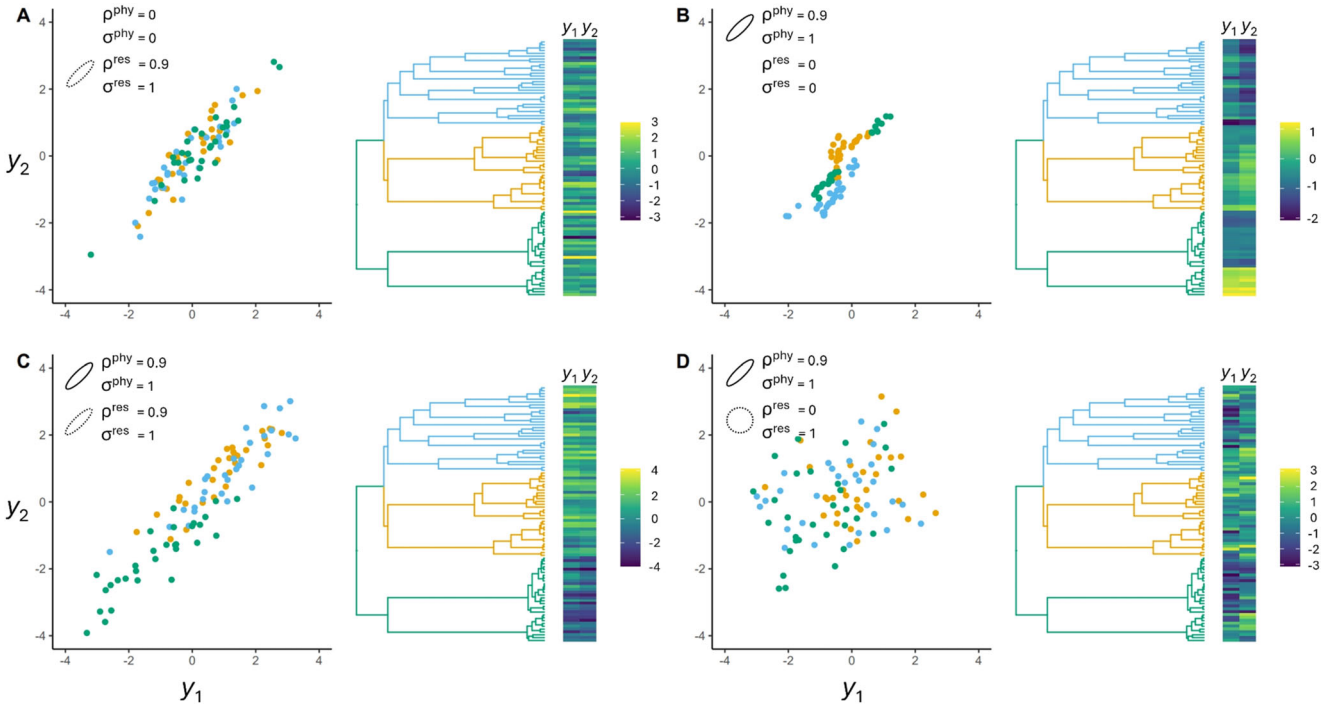


Fig. 1. Bivariate trait data (y_1, y_2) simulated from a basic multi-response phylogenetic mixed model (MR-PMM) (see Equation 10) containing different levels of phylogenetic (ρ^{phy}) and residual (ρ^{res}) correlation (A–D). Simulation conditions for each panel are inset at the top left of each scatterplot, with ellipses providing a visual representation of the strength of correlation within each variance component. Data are plotted in scatterplots, as well as heatmaps arranged against the generating phylogeny. In A, y_1 and y_2 have no phylogenetic signal ($\sigma^{\text{phy}} = 0$), but a strong positive residual correlation ($\rho^{\text{phy}} = 0$, $\rho^{\text{res}} = 0.9$). Clades overlap completely in the scatter plot and the heatmap shows bands of colour across y_1 and y_2 that appear random with respect to phylogeny. B shows the opposing situation, y_1 and y_2 are positively correlated but entirely with respect to phylogeny ($\rho^{\text{phy}} = 0.9$, $\rho^{\text{res}} = 0$), with no residual variation in either trait ($\sigma^{\text{res}} = 0$). The scatter plot shows clearly distinguishable clades and a tendency for both within- and between-clade correlation. The extent of between-clade correlation (the tendency for clades to arrange along a positive slope) depends on the topology of the tree, with deep splits promoting separation of clades along the major axis of covariation. The heatmap shows phylogenetic structure weakening across clades from green, to orange to blue as the topology becomes more deeply nested, that is as subclades become less clearly separated in evolutionary time. In C, y_1 and y_2 have equal phylogenetic and residual variances, and a strong positive correlation operating on both levels. This scenario, where phylogenetic and residual correlations are similar in sign and magnitude, is likely to be common for many biological traits. In D, both traits contain phylogenetic and residual variance, but correlation is only present on the phylogenetic level. This shows how easily conserved correlations are obscured when residual sources of variation contribute substantially to trait variance. Notably, D represents a set of conditions for which a single-response PMM (SR-PMM), such as PGLS, will typically fail to detect a significant association between y_1 and y_2 (Westoby *et al.*, 2023).

performance curves, growth trajectories, or reaction norms (Pottier *et al.*, 2024). Parameters that describe these functions (e.g. intercepts, slopes) are often correlated, and the strength and direction of these correlations may be biologically meaningful, providing opportunities to test hypotheses about the functional form of trait relationships (e.g. Pettersen *et al.*, 2023; Kontopoulos *et al.*, 2020).

Finally, the structure of the MR model allows us to flexibly include partially observed data. This means that users can retain observations where only a subset of the response traits have been observed, which improves inference [e.g. observations reporting values for just two response traits out of the complete set improve the estimation and decomposition of the (co)variances for these two traits]. Including incomplete records usually means that more data (observations and species) can be included in analyses, especially as the number of included response variables increases and incomplete cases may comprise a substantial fraction of the data.

III. MR-PMM – EXTENSIONS TO THE BASIC MODEL

In this section, we explore several useful extensions to the basic MR-PMM that expand its utility for comparative biology.

(1) Multilevel models

Data for phylogenetic comparative analyses are typically compiled from multiple sources. This often yields multiple observations per species, with individual studies contributing observations on only a subset of focal traits for a subset of species. Mixed models like MR-PMM offer a general framework for dealing with these complex, cross-classified effect structures *via* multilevel models. Benefits include the ability to partition variation among multiple hierarchical levels (e.g. between-species, within-species, within-individual), account for variation in sampling effort, and fit models to data containing a mixture of record types [McNeish, 2021; Nakagawa & Santos, 2012; also see Fig. 4 in Nakagawa *et al.* (2017b) for a helpful visualisation]. Furthermore, accounting for multilevel structure may be necessary for valid inference of phylogenetic (co)variances, which are likely to be confounded with other sources of between-species variation (Cinar *et al.*, 2022; Garamszegi & Møller, 2017).

Equation (10) can be extended to account for multilevel structure in the data by modelling additional random effect components in $\mathbf{u}$ that capture the sampling hierarchy. For example, to account for study ID effects, we include a random effect at the study level $\mathbf{u}^{\text{study}}$, which models variability between studies. To account for multiple observations per species, we include an additional (unstructured) random effect at the species level $\mathbf{u}^{\text{species}}$, which models variability between species that is independent of phylogeny [generalising the univariate presentation of phylogenetic and non-phylogenetic variance components given in Equation (8)]. Each component term of $\mathbf{u}$ is associated with a (co)variance structure that specifies dependence within that component term (Equation 15). However, component terms are assumed to be independent *a priori* [i.e. the off-diagonal blocks of var($\mathbf{u}$) are 0],

$$\mathbf{u}=\begin{bmatrix}\mathbf{u}^{\text{phy}}\\ \mathbf{u}^{\text{species}}\\ \mathbf{u}^{\text{study}}\end{bmatrix}\sim\mathcal{N}\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}\mathbf{\Sigma}^{\text{phy}}\otimes\mathbf{C} & 0 & 0\\ 0 & \mathbf{\Sigma}^{\text{species}}\otimes\mathbf{I}_{N_{\text{species}}} & 0\\ 0 & 0 & \mathbf{\Sigma}^{\text{study}}\otimes\mathbf{I}_{N_{\text{species}}}\end{bmatrix}\right), \tag{15}$$

where $\mathbf{\Sigma}^{\text{study}}=\text{diag}\left(\left(\sigma_1^{\text{study}}\right)^2,\left(\sigma_2^{\text{study}}\right)^2,\ldots,\left(\sigma_{N\text{traits}}^{\text{study}}\right)^2\right)$. In *brms* syntax, this model is specified using independent additive random effects for each grouping factor:

$$\texttt{y}\sim\cdots+(\texttt{1}|\text{gr}(\text{phy},\text{cov}=\text{C}))+(\texttt{1}|\text{species})+(\texttt{1}|\text{study}). \tag{16}$$

Modelling these additional components in $\mathbf{u}$ has important consequences for the way trait variation is partitioned, affecting the interpretation of model parameters. For example, the random effects specified in Equation (15) represent a shift from modelling all non-phylogenetic sources of trait correlation in the residual error $\mathbf{e}$ as in Equation (10), to modelling non-phylogenetic between-species trait correlation explicitly *via* $\mathbf{u}^{\text{species}}$, study effects *via* $\mathbf{u}^{\text{study}}$, and any remaining trait (co)variance as residual error in $\mathbf{e}$. Residual correlation for such a model represents within-species trait correlation, which may require large data sets to be robustly estimated (Zhou, Cieraad & van Bodegom, 2022). Alternatively, users may suppress residual correlation (i.e. constrain $\mathbf{\Sigma}^{\text{res}}$ to a diagonal matrix) by setting `rescor=F` and `idh(trait):units` in *brms* and *MCMCglmm*, respectively, noting that this will have the effect of forcing any residual correlation into correlations at higher levels.

Additional components could, of course, be added to $\mathbf{u}$ (Equation 15) to model other sources of covariance between observations, such as spatial or temporal effects (e.g. Markovski *et al.*, 2023; Gomes *et al.*, 2023; also see Freckleton & Jetz, 2009). This flexibility of multi-level models makes MR-PMMs well suited for large-scale comparative studies including complex hierarchically structured data.

(2) Multivariate models of evolution

In recent years, there have been considerable advances in software implementation and efficient computation of multivariate evolutionary models (Bartoszek *et al.*, 2023; Blomberg, Rathnayake & Moreau, 2020; Mitov *et al.*, 2020; Clavel, Escarguel & Merceron, 2015; Bartoszek *et al.*, 2012). Models

of evolution comprise both stochastic and deterministic components and are naturally expressed as stochastic differential equations where the rate of change of trait values over time may depend on current values. Solutions (i.e. the distribution of trait values) are obtained by integrating these differential equations over the branches of the phylogeny. For many multivariate models (e.g. multivariate BM, OU, EB and ACDC), the trait distributions are Gaussian and therefore characterised by their means and covariances. In the multivariate BM case, the solutions are more easily obtained (although still present convergence issues even for the most sophisticated algorithms; Butler *et al.*, 2023) since the covariance is proportional to $\mathbf{C}$ (Equation 14) up to a re-scaling of the off-diagonal elements of the total variance, equivalent to a simple branch-length transformation. For more complex models of evolution, the covariance matrix cannot generally be obtained *via* branch-length transformation, which appears to limit the scope of MR-PMMs in accommodating alternative evolutionary models. However, in certain cases, analytic expressions for the total covariance have been derived (Bartoszek *et al.*, 2012). Perhaps the mostly commonly used alternative to BM is the multivariate OU model,

$$\mathrm{d}\mathbf{y}(t) = \mathbf{A}(\mathbf{y}(t) - \boldsymbol{\eta})\mathrm{d}t + \boldsymbol{\Lambda}\mathrm{d}\boldsymbol{W}(t), \tag{17}$$

where, $\mathbf{y}$ is an N_{traits} vector of trait values, $\mathbf{A}$ is a $N_{traits} \times N_{traits}$ strength of selection (or rate of adaptation) matrix, η is an N_{traits} vector of trait optima values, and $\boldsymbol{W}$ is an N_{traits} dimensional Brownian process with diffusion matrix $\boldsymbol{\Lambda}$. Viewed in discrete time as a vector autoregressive model, the differential equation (Equation 17) provides a rule to update trait values based on their current values together with a (Brownian) stochastic step. Each off-diagonal element of $\mathbf{A}$ models the deterministic effect of one trait value on the update of another trait, permitting a test of Granger causality – the notion that historical values of one trait improve the prediction of another (Shojaie & Fox, 2022). Although Granger causality provides a more direct inference of causation than a correlation coefficient, especially given $\mathbf{A}$ in Equation (17) need not be symmetric, it is not necessarily more meaningful than the trait correlations that we can derive from $\boldsymbol{\Lambda}$ (e.g. ρ^{phy}_{ij}). Indeed, estimated correlations in the stochastic step may also capture causal relationships, just not those associated with adaptation towards trait optima.

In principle, analytic computation of the expected covariance between two traits $\mathbf{y}_i$ and $\mathbf{y}_j$ under different models of evolution extends MR-PMMs to a broad class of multivariate Gaussian evolutionary models. In practice, fitting such models may be difficult and constraints on model parameters may be necessary to ensure identifiability. For the multivariate OU model, many common variants are characterised by strong constraints on the form of $\mathbf{A}$ which may simplify the evaluation of $\mathrm{cov}(\mathbf{y}_i,\mathbf{y}_j)$, for example, when attraction to optima is assumed to be equal and independent for all traits; that is when $\mathbf{A}$ is a multiple of the identity matrix (Goolsby *et al.*, 2017; Tung Ho & Ané, 2014). However, options to implement models other than BM are yet to be integrated into popular software packages such as *MCMCglmm* and *brms*, despite being well developed in other modelling contexts (Bouckaert *et al.*, 2019; Mitov *et al.*, 2020; Clavel *et al.*, 2015; Goolsby *et al.*, 2017). A valuable direction for future work would therefore be to explore the repertoire of evolutionary models compatible with MR-PMMs and integrate them into standard software; together with algorithms that exploit the tree-structured nature of multivariate likelihoods (Hadfield, 2010; Mitov *et al.*, 2020; Bastide *et al.*, 2021; Hassler *et al.*, 2022). Progress in this direction would yield a powerful modelling framework, capable of synthesising model selection among alternate multivariate evolutionary models with the generalisable multilevel framework of MR-PMMs.

(3) Non-Gaussian response traits

One advantage of MR-PMMs compared to other methods (e.g. Goolsby *et al.*, 2017; Clavel *et al.*, 2015; Bartoszek *et al.*, 2012) is the capacity to include both continuous and discrete response traits. To include different response types simultaneously, a latent-variable formulation is used. This approach models between-trait covariance in the usual way, *via* a multivariate Gaussian distribution, while at the same time permitting trait-specific probability distributions and associated link functions (Hadfield, 2010). For traits $i = 1, \ldots, N_{trait}$, the latent-variable formulation is written

$$\boldsymbol{l} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \tag{18}$$

where $\boldsymbol{l} = \left(\boldsymbol{l}_1^{\mathrm{T}}, \boldsymbol{l}_2^{\mathrm{T}}, \cdots, \boldsymbol{l}_{N\mathrm{traits}}^{\mathrm{T}}\right)^{\mathrm{T}}$ are latent variables associated with the observed traits and the corresponding (latent-trait) covariances are given by matrices $\boldsymbol{\Sigma}^{\mathrm{phy}}$ and $\boldsymbol{\Sigma}^{\mathrm{res}}$ which characterise the distributions of $\mathbf{u}$ and $\mathbf{e}$ as in Equations (11) and (12); the latter are now "pseudo" rather than true residuals introduced as observation-level random effects and represent additive overdispersion. A model for the observations $\mathbf{y}_i$ is given in terms of a trait-specific probability distribution f_i and link function η_i

$$\mathbf{y}_i \sim f_i\left(\eta_i^{-1}(\boldsymbol{l}_i), \boldsymbol{\phi}_i\right), \tag{19}$$

where $\boldsymbol{\phi}_i$ includes any distribution-specific parameters (if required). Depending on the choice of distribution f_i, additional constraints may need to be imposed on associated variance terms to ensure parameter identifiability (Sorensen & Gianola, 2002).

Variance partitioning to define an analogue of λ (Equation 6) is possible for the latent-variable models given by Equations (18) and (19), but is more nuanced because there is a distinction between latent and observation-scale partitioning where distribution-specific variances must be taken into account (Nakagawa, Johnson & Schielzeth, 2017a). Variance terms are usually more easily interpreted when they are expressed on the scale of trait observations

rather than that of the latent variables (de Villemereuil, 2018). For multi-response models involving two or more distribution types, there are generally no closed-form solutions for the observation-level variances and associated statistics, however numerical methods are readily applied given point estimates or a set of posterior samples of the model parameters (de Villemereuil *et al.*, 2016). A further complication for interpretation, however, is that the non-linear transformations required to transform covariances from the latent to the observation scale can change the bounds of the correlation coefficient. For example, the minimum correlation of two exponentiated random normal variables is approximately −0.37 and the correlation bounds of a random normal and an exponentiated random normal variable are approximately ±0.76. Similar issues may complicate inference on causal associations between Gaussian and non-Gaussian response variables (see Section III.5), because correlations are modelled on the latent, rather than observation, scale.

(4) Considerations for fixed effects

There are three common reasons for including fixed effects in a MR-PMM. The first is to model the influence of predictors that are unlikely to display strong phylogenetic signal. The second is to account for sampling biases, such as differences in data-collection methods or laboratory protocols. The third reason is to explore relationships between responses conditional on covariates, typically to evaluate hypotheses about the mechanistic basis of trait correlations. For example, Sanchez-Martinez *et al.* (2020) used a MR-PMM to partition correlations between a range of plant hydraulic traits into phylogenetic and non-phylogenetic contributions, then refitted models with relevant climate covariates included as fixed effects. Whether or not trait correlations remained after accounting for climate effects was used to evaluate hypotheses about trade-offs and integration between these traits in the evolution of plant hydraulic systems. Equivalent procedures are used to control for environmental variation in multivariate species co-occurrence models (Ovaskainen *et al.*, 2017; Warton *et al.*, 2015). However, in a MR-PMM, a preferable approach will often be to treat biological covariates as response traits and assess conditional dependencies *via* partial correlations.

(5) Partial correlations

The practice of assessing conditional relationships between traits (i.e. relationships after controlling for relevant covariates) is well established for univariate models. For example, it is common for researchers to use multiple regression to estimate partial regression coefficients for each fixed effect predictor. In a MR-PMM, we estimate conditional relationships from a joint model where partial correlations are derived from elements of the inverse trait covariance matrix, known as the precision matrix $\mathbf{\Omega} = (\Omega_{ij})$; that is,

$$\mathbf{\Omega}^{\text{phy}} = \left(\mathbf{\Sigma}^{\text{phy}}\right)^{-1} \text{and } \mathbf{\Omega}^{\text{res}} = \left(\mathbf{\Sigma}^{\text{res}}\right)^{-1}, \tag{20}$$

for which the corresponding partial correlation coefficients are $-\Omega_{ij}/\sqrt{\Omega_{ii}\Omega_{jj}}$.

The estimation of partial correlations *via* precision matrices casts MR-PMMs in the framework of Gaussian graphical models (Popovic *et al.*, 2019; Epskamp *et al.*, 2018; Yuan & Lin, 2007; Magwene, 2001), which have great potential for clarifying assumptions of phylogenetic comparative methods (Uyeda *et al.*, 2018). Elements of these precision matrices relate directly to the existence of edges in a graphical causal network (Fig. 2). For example, suppose we obtain data on

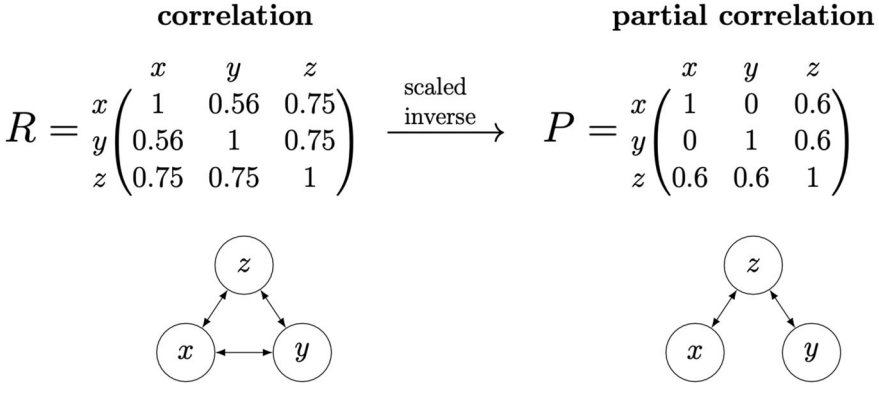


Fig. 2. Correlation $R = \left(\frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}\right) = \left(\rho_{ij}\right)$ and partial correlation $P = \left(\frac{-\Omega_{ij}}{\sqrt{\Omega_{ii}\Omega_{jj}}}\right) = \left(\rho_{ij|kl\ldots}\right)$ matrices between species traits x, y and z, where $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ and $kl\ldots$ indexes all variables other than i and j. The undirected network graphs below provide a qualitative representation of trait relationships, where the absence of an edge between two traits corresponds to a zero off-diagonal element in the matrix above. To evaluate whether the relationship between x and y can be explained by a covariate z, we compute partial correlations from the precision matrix of trait covariances, which quantifies the linear relationship between x and y when controlling for z. In this example, x and y are strongly correlated, however this is fully explained by their relationships to the covariate z, that is x and y are independent, conditional on z.

three co-varying species traits x, y and z. In the precision view (Fig. 2, right panel), the absence of an edge between x and y signifies conditional independence: that x and y are independent, given z. This does not imply that the correlation between x and y is zero (Fig. 2, left panel), only that the partial correlation between these traits is zero. Precision matrices are often sparser than their corresponding covariance matrices (i.e. contain more off-diagonal elements that are effectively zero), which focuses our inference on a reduced set of candidate causal relationships. For example, Halliwell *et al.* (2024) used a MR-PMM to disentangle direct climate effects on the evolution of social behaviour from indirect effects driven by adaptations to climate (i.e. traits correlated with climate) that go on to promote the evolution of social behaviour.

Outside of controlled experiments, there is always the strong possibility that partial correlations could be explained by missing variables. Nonetheless, precision matrices are a powerful tool for generating candidate causal hypotheses from observational data (Popovic *et al.*, 2019; Pearl, 1995; also see Section X for tutorial), and are readily obtained from a fitted MR-PMM.

IV. INTERPRETATION

The biological interpretation of phylogenetic models and their utility for addressing specific ecological and evolutionary hypotheses has been a subject of lively debate (Freckleton *et al.*, 2002; Westoby, Leishman & Lord, 1995; Björklund, 1997; Harvey, Read & Nee, 1995). Reflecting on this constructive body of work, we believe MR-PMMs emerge as a flexible, pluralistic framework for analyses of multiple species traits. However, consensus on interpretation is critical to realise the full, operational potential of the method.

(1) Phylogenetic (co)variances

Phylogenetic (co)variances estimate the conserved component of species phenotypes, given a model of evolution and a hypothesis of the phylogenetic relationships between taxa. The extent to which these conserved effects result from stochastic processes, or reflect constraints and adaptive responses to selection, will vary for different traits, between clades, and across phylogenetic scales. These forces are also likely to fluctuate throughout the course of evolutionary history, complicating the interpretation of trait covariance (Revell *et al.*, 2008; Losos, 2008, 2011). As Housworth *et al.* (2004, p. 93), point out "[PMM] envisions phenotypic evolution as being the result of a complex of forces including some that are retained over long periods of time, forming patterns in trait variation that reflect the underlying phylogenetic structure". Examples include variation arising from genetic differences between clades that have accumulated over evolutionary time, as well as non-genetic (e.g. spatial, environmental or developmental) effects that are phylogenetically structured for one reason or another.

(2) Pattern *versus* process

There are inherent limitations to the inferences we can make about evolutionary processes from data observed only in the present day. Indeed, many have argued that the interpretation of phylogenetic comparative analyses should be limited to patterns of variation, rather than the explicit processes that generated them (Revell *et al.*, 2008; Losos, 2008, 2011; Ives, 2018). Others have suggested that integration of fossil evidence, path analyses, ancestral state reconstruction, and simulation studies may extend our epistemic reach to hypothesis tests of the processes generating observable variation (Thorson & van der Bijl, 2023; Uyeda *et al.*, 2018; Quental & Marshall, 2010; Slater, Harmon & Alfaro, 2012; Uyeda & Harmon, 2014).

Because phylogenetic and non-phylogenetic correlations can arise from multiple distinct causal processes, we argue for caution around mechanistic interpretations. Two traits may be phylogenetically correlated if, for example, conserved genes underlying a set of traits show linkage or pleiotropy that constrains the evolutionary potential for certain trait combinations, or because traits form part of a coordinated life-history strategy that involves phylogenetic niche conservatism (Westoby *et al.*, 2023; Wiens *et al.*, 2010). Such hypotheses should not be considered mutually exclusive (Losos, 2011; Revell *et al.*, 2008). Deriving partial correlations can help evaluate evidence for alternative mechanistic hypotheses in a MR-PMM (see Section III.5). Ultimately, however, it is rarely the pleasure of a comparative biologist to declare causation, but rather to uncover meaningful patterns of variation, test predictions from theory, and generate hypotheses for future comparative and experimental research.

(3) Visualisations

Visualisations of simulated data are powerful heuristics for biologists, as they provide tangible representations of the abstract covariance structures we aim to partition with a MR-PMM. To compare and contrast phylogenetic and non-phylogenetic (co)variance components, and clarify their biological interpretation, we simulated data from a simple MR-PMM for two Gaussian response traits $\mathbf{y}_1$ and $\mathbf{y}_2$, as in Equation (10). Figure 1 shows scatter plots of simulated data together with the tree used to derive the phylogenetic correlation matrix $\mathbf{C}$. A heat map of trait data plotted against the phylogeny highlights the distinct signature each source of covariance leaves in the data (see Fig. 1 legend for details). For more technical coverage of simulating multivariate data containing phylogenetic and non-phylogenetic covariance structure, we include detailed examples and R code in the tutorial (Section X).

(4) Relevance for biological hypotheses

We see the flexible framework and nuanced interpretation of MR-PMMs as having at least four strong usage cases for biologists: (*i*) to provide a more meaningful decomposition of trait (co)variances, including phylogenetic and non-phylogenetic components, among both continuous and discrete response traits; (*ii*) to test for partial correlations, consistent with functional relationships between species traits, in a rigorous phylogenetic framework; (*iii*) to test theory and generate new hypotheses about the ecological and evolutionary drivers of trait variation; and (*iv*) to predict species responses and vulnerability to environmental change based on phylogenetically structured trait–trait and trait–environment relationships. With an emerging synthesis of evolutionary, ecological and genomic approaches to comparative analyses (James *et al.*, 2021; Smith *et al.*, 2020), we concur with recent work recognising the considerable potential of MR-PMMs to advance our understanding of trait evolution, niche conservatism, and community assembly at numerous scales (Pottier *et al.*, 2024; Westoby *et al.*, 2023; Gallinat & Pearse, 2021; also see Ovaskainen & Abrego, 2020).

V. PREDICTION

While the estimation of trait covariance, rather than trait prediction *per se*, is the focus of a MR-PMM, the estimated covariance structure can be fully exploited for predictive goals. In the presence of strong phylogenetic signal and trait correlations, this approach improves on ordinary multiple regression for prediction as it makes use of phylogenetic structure in the predictors. In particular, when data for focal species are incomplete, a MR-PMM leverages information from closely related species to impute missing values. The predictive distribution of a fitted model is also useful for model validation, such as posterior-predictive checks, or model comparison using predictive assessment such as cross validation. We now review conceptual aspects of predictive distributions in MR-PMMs, and refer the reader to the accompanying tutorial (Section X) for worked examples of how these predictive methods can be implemented in R.

(1) Predicting new or missing data

Missing data are a perennial problem in comparative biology (Freckleton, 2009; Nakagawa & Freckleton, 2008). For many traits of interest, data collection remains too expensive, technically demanding, or logistically challenging to keep pace with the expansion of species phylogenies and functional trait databases. For example, plant hydraulic traits are central to water stress tolerance and therefore key metrics for evaluating species vulnerability to climate change (Brodribb *et al.*, 2020; Choat *et al.*, 2012). Hydraulics data are scarce because measurements require specialised equipment and are labour intensive. However, because hydraulics forms part of a coordinated growth strategy involving conserved trait networks (Blackman *et al.*, 2024; Sanchez-Martinez *et al.*, 2020; Skelton *et al.*, 2021; Liu *et al.*, 2024), phylogenetically structured trait correlations can enhance our ability to predict them where sufficient data on correlated functional traits are available (Sanchez-Martinez *et al.*, 2024). Predicting species conservation status provides a more general example. Vulnerability to extinction is a complex trait integrating species biology, biogeography, climate and land-use practices (Lee & Jetz, 2011; González-del Pliego *et al.*, 2019), with risk assessment of data-deficient species a persistent challenge (Borgelt *et al.*, 2022). Commonly measured traits that show strong phylogenetic signal (e.g. body size, plant height, age at maturity) have been linked to extinction risk *via* species physiology and demographic processes such as generation time (González-del Pliego *et al.*, 2019; Kelly *et al.*, 2021; Cardillo, 2003; Jetz & Freckleton, 2015; Lee & Jetz, 2011). Leveraging phylogenetically conserved trait–trait and trait–environment relationships may provide tractable avenues to predict these crucial, yet costly to measure traits, and the real-world outcomes that depend on them.

(2) Model validation using posterior predictive checks

In Bayesian settings, posterior predictive checks use simulated data from the fitted model to test both the adequacy of the fit to the data and the plausibility of the model predictions (Gelman *et al.*, 2013). These checks are typically visual plots that rely on qualitative assessments (Gabry *et al.*, 2019). To test for adequacy of fit, one option is to superimpose the observed data onto a plot of the distribution of the simulated data. For a PMM, this type of check could be performed using a separate plot for each trait with the observed data point for each species plotted on top of a five-point summary of the predicted distribution (see Fig. 3 for an example). For assessment of model plausibility, the current state of knowledge should be used to evaluate model predictions within ecologically plausible but perhaps unobserved ranges of the included covariates. For example, do predictions for $\mathbf{y}_1$ make sense across the entire range of plausible $\mathbf{y}_2$ values for a given set of taxa and/or region of interest?

(3) Predictive assessment using cross validation

Cross validation is the use of data splitting to estimate the predictive performance of one or more statistical models, usually for the purpose of model comparison, validation, or selection (Yates *et al.*, 2022). Model selection is used when discrete decisions must be made about model structure, for example, whether or not to include various fixed effects (i.e. variable selection) or the choice of probability distribution (e.g. Poisson or negative binomial for count data). Predictive assessment tools such as cross validation are also useful to quantify or simply visualise how well a model can predict new data (e.g. new taxon–trait pairs in a PMM) which is distinct from typical assessments of model adequacy which concern prediction of data to which a model was fitted.

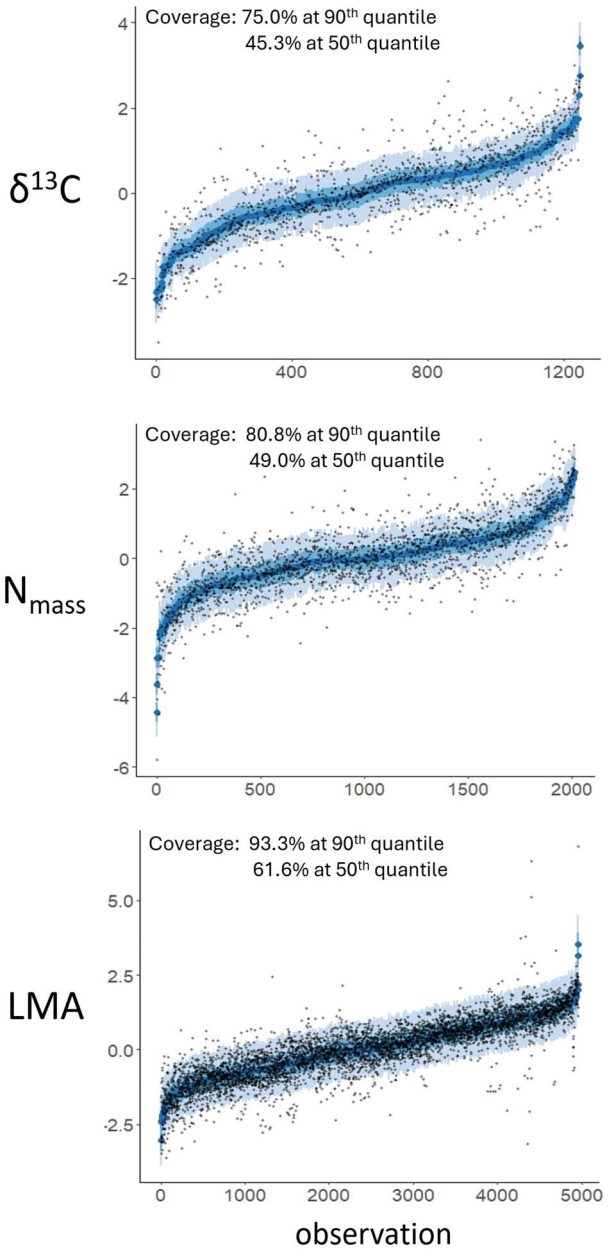


Fig. 3. Posterior predictive checks from a fitted multi-response phylogenetic mixed model (MR-PMM) of leaf traits across 457 species of *Eucalyptus*. For each trait, black points represent observed values, while five-point summaries show the median (dark blue points), 0.5 CI (blue bars), and 0.95 CI (light blue bars) of the posterior predictive distribution for each observation. Predictions are made by conditioning on all random effects, with observations ordered by the predictive mean. $\delta^{13}C$, carbon isotope ratio in leaf tissue; N_{mass}, nitrogen content per dry mass of leaf tissue; LMA, leaf mass per unit area.

Cross validation works by fitting each model to a subset of the available data and then assessing the models' predictive capacities on the remaining data. The splitting procedure is systematically iterated to select different test data and the overall predictive performance is summarised as a cross-validation score (Arlot & Celisse, 2010). When the measure of predictive performance is the log likelihood of the test data, then the predictive assessment is said to be information theoretic. Information criteria such as Akaike's Information Criterion (AIC; Akaike, 1973), or for Bayesian analyses the Widely Applicable Information Criterion (WAIC; Watanabe, 2010), approximate predictive log likelihood without data splitting by adding a bias correction to the log likelihood of the full data, meaning each model has to be fitted only once. Information criteria are therefore faster to compute than cross-validation scores, however the latter are often preferred as they are less sensitive to violations of model assumptions and are readily combined with techniques to mitigate overfitting (Yates, Richards & Brook, 2021). For a Bayesian MR-PMM estimated using Monte Carlo sampling, model fitting may be too slow to permit the use of ordinary cross validation. However, recently developed approximate methods provide a rapidly computed and accurate alternative (Vehtari, Gelman & Gabry, 2017; Bürkner, Gabry & Vehtari, 2021) that is easily applied to *brms* model objects *via* the *loo* R package.

VI. EXAMPLE ANALYSIS – LEAF TRAITS IN *EUCALYPTUS*

To demonstrate potential applications of a MR-PMM, we present an example analysis using data on leaf traits for 457 species of *Eucalyptus* from the AusTraits database (Falster *et al.*, 2021). The data contain partially missing values for each trait and have a complex multilevel structure with observations grouped by species and study. Our intentions for this analysis were to: (*i*) decompose the variance in each trait and evaluate the phylogenetic signal; (*ii*) estimate phylogenetic and non-phylogenetic between-species trait correlations while accounting for multilevel structure in the data; (*iii*) evaluate partial correlations; and (*iv*) evaluate model predictive performance. Full details of this analysis are provided in the tutorial (Section X).

(1) Methods

We derived the phylogenetic correlation matrix **C** from the maximum likelihood time-calibrated eucalypt phylogeny "ML1" presented in Thornhill *et al.* (2019). We focused on three target leaf traits: leaf mass per unit area (LMA); nitrogen content per dry mass of leaf tissue (N_{mass}); and the ratio of carbon isotopes 12 and 13 in leaf tissue ($\delta^{13}C$). All responses were log-transformed, zero-centred and scaled to unit variance prior to analysis. The random effects were specified as in Equation (15), to estimate phylogenetic and non-phylogenetic between-species trait correlations while accounting for within-species replication and study effects.

For model validation, we performed posterior predictive checks. The proportion of the data falling in the

predictive intervals was close to the nominal quantiles, indicating that the model is adequately calibrated (although the proportions for $\delta^{13}C$ were a little low), and visual inspection of the plotted distributions verified the capacity of the model to generate plausible data (Fig. 3). The lower coverage of the posterior predictive intervals for $\delta^{13}C$ may indicate that the model is missing relevant covariates. Indeed, study effects accounted for the largest proportion of variance in $\delta^{13}C$, suggesting that methodological differences between studies or local environmental factors may be a contributing factor to poor predictive performance. Further, the chosen probability distribution (Gaussian), or the log-transformation, may be less well suited to this trait. For predictive assessment using approximate LOO-CV (Bürkner *et al.*, 2021), see the tutorial (Section X).

(2) Results

Phylogenetic signal, and hence the tendency for similar values among closely related species (Fig. 4), varied considerably between traits (Fig. 5A; posterior median for $\widehat{\lambda}$: LMA = 0.73, N_{mass} = 0.51, $\delta^{13}C$ = 0.23).

The traits selected for this analysis are tightly linked to resource-use strategies and leaf economics (Reich, 2014;

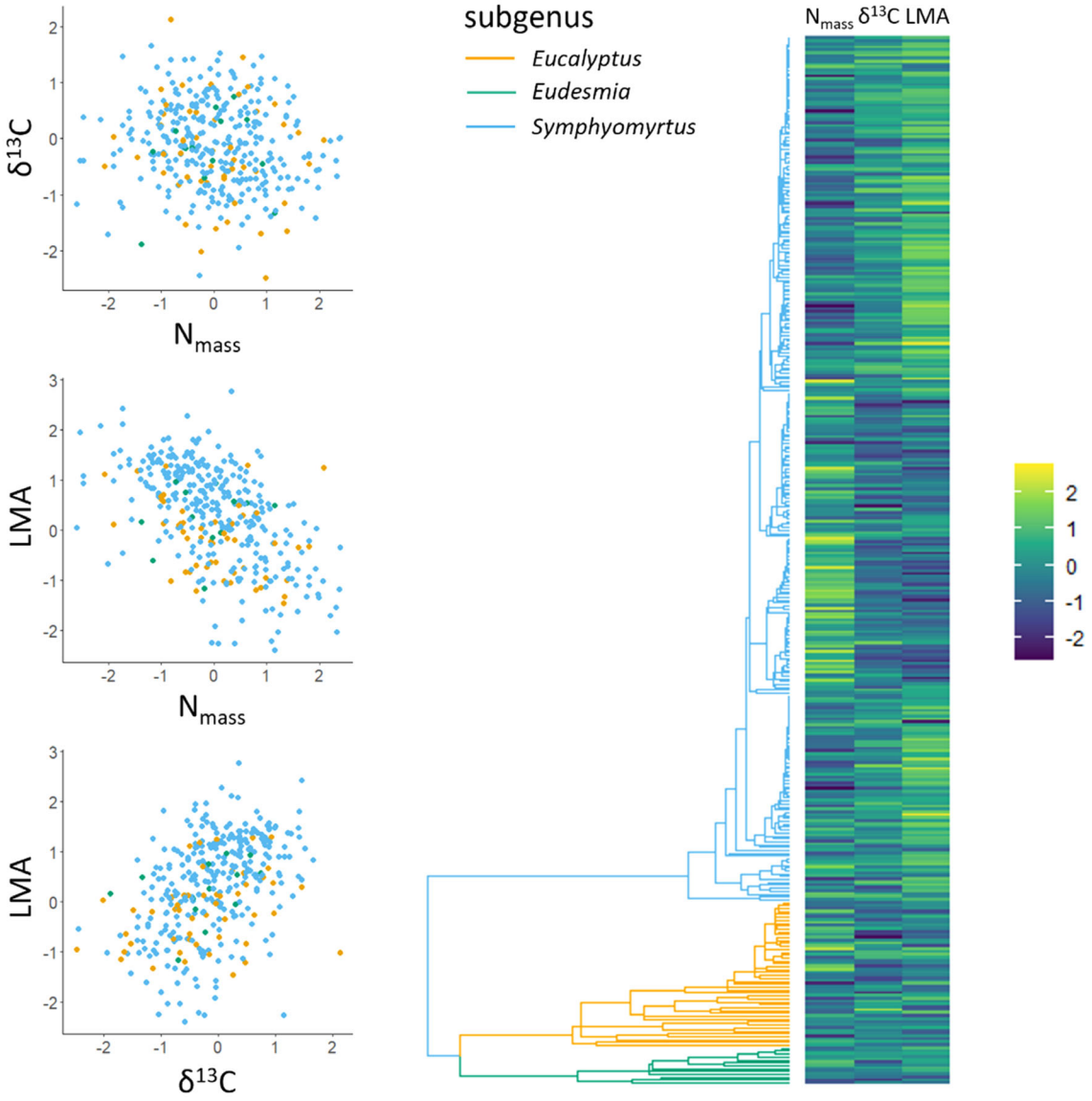


Fig. 4. Scatterplots (left) and heat-maps (right) of three leaf traits across 361 species of *Eucalyptus* (data filtered to complete cases for plotting). Trait values have been log-transformed and scaled. For heat maps (right), trait values are aligned with the corresponding species in the phylogeny (centre). For $\delta^{13}C$ *versus* N_{mass} (top left), opposing phylogenetic and non-phylogenetic correlations reported by the model (Fig. 5) are obscured at the level of species phenotypes. $\delta^{13}C$, carbon isotope ratio in leaf tissue; N_{mass}, nitrogen content per dry mass of leaf tissue; LMA, leaf mass per unit area.

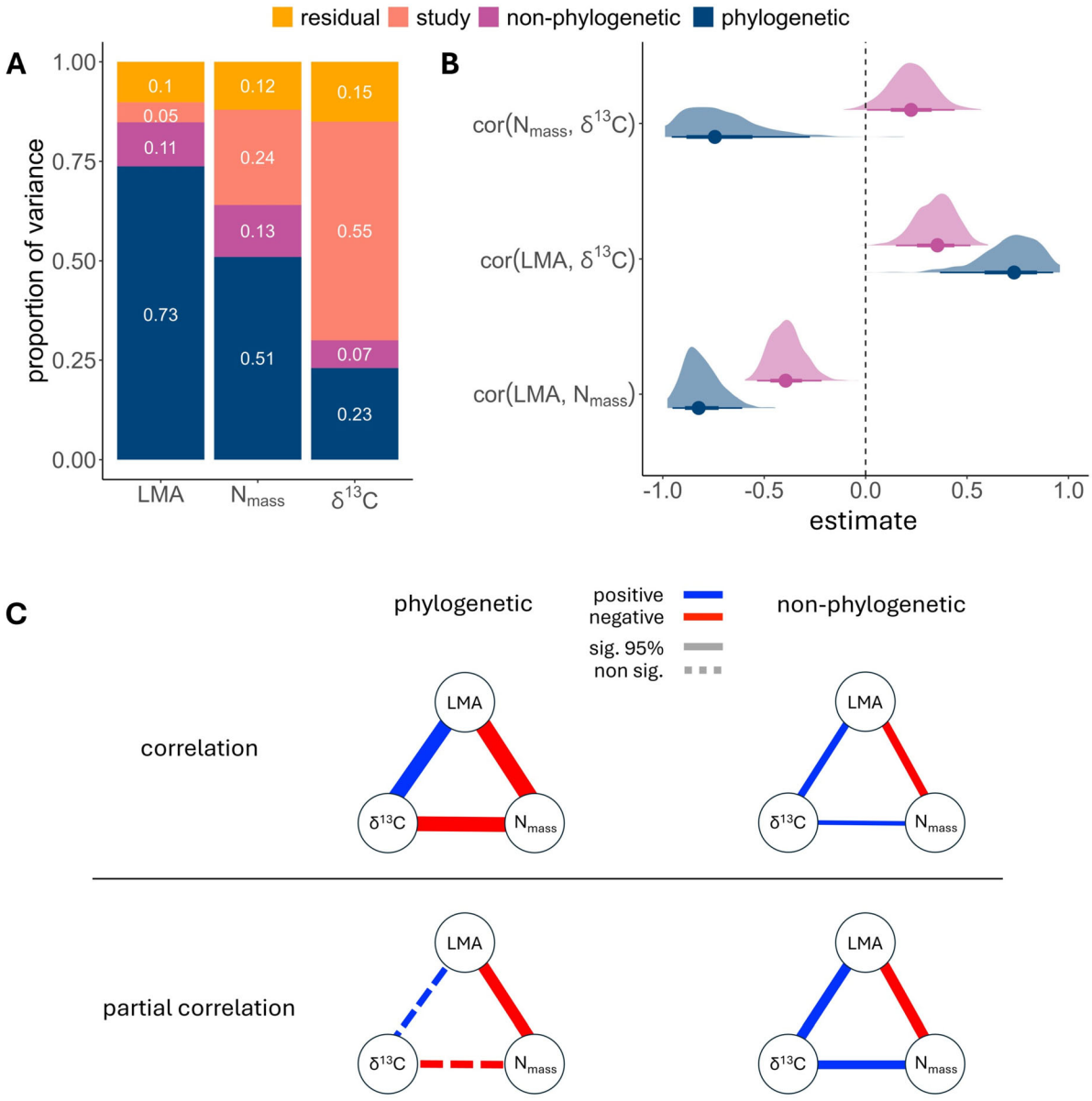


Fig. 5. Results from a multi-response phylogenetic mixed model (MR-PMM) of leaf traits across 457 species of *Eucalyptus*. (A) Variance decomposition reveals large differences between response traits in phylogenetic signal, as well as the relative contributions of different sources of variance. (B) Phylogenetic and non-phylogenetic between-species correlation coefficient estimates. Points represent the posterior mean for each estimate, with 50% and 95% confidence intervals (CIs) represented by heavy and light wicks respectively. (C) Correlations and partial correlations between traits represented as network graphs. Edge widths are proportional to the posterior mean of each coefficient estimate, with line type indicating significance at the 95% CI. $\delta^{13}C$, carbon isotope ratio in leaf tissue; N_{mass}, nitrogen content per dry mass of leaf tissue; LMA, leaf mass per unit area.

Wright *et al.*, 2004). For example, Prieto *et al.* (2018) found that plant species in a Mediterranean woodland showing more resource acquisitive strategies (low LMA and high N_{mass}), were associated with lower water use efficiency (low $\delta^{13}C$). This is consistent with predictions from leaf economic theory (Wright *et al.*, 2004), that LMA, $\delta^{13}C$ and N_{mass} represent components of a coordinated life-history strategy which trade off predictably as different regions of niche space are explored during species diversification. These coevolutionary relationships are clearly supported by the fitted model (Fig. 5B), which reports a positive phylogenetic correlation between LMA and $\delta^{13}C$, and negative phylogenetic correlations between these traits and N_{mass}. Thus, in line with theory and empirical observations from other plant groups, these traits appear to co-vary predictably over evolutionary time in *Eucalyptus*.

A notable result is that for N_{mass} and $\delta^{13}C$, the non-phylogenetic correlation is positive while the phylogenetic

correlation is negative (Fig. 5B). This situation highlights an important strength of MR-PMM: the capacity to disentangle trait relationships operating at different levels in the model hierarchy. In line with simulations, consistent correlations across levels produce clear trends in inter-species data, while opposing correlations tend to obscure these relationships (Figs 4 and 1D; see tutorial in Section X).

Finally, while all non-phylogenetic trait relationships were retained as partial correlations, only the negative relationship between LMA and $\mathcal{N}_{mass}$ was significant on the phylogenetic level (Fig. 5C). LMA and $\mathcal{N}_{mass}$ also showed the strongest relationships overall (Fig. 5B), suggesting a deeply conserved functional integration between these traits.

VII. EXTENDED TOPICS

The following section explores technical, yet important aspects of fitting MR-PMMs. These topics are particularly relevant in the case of high-dimensional and/or sparse trait data, where issues such as prolonged fitting time, model convergence difficulties, or parameter identifiability require the application of specialised computational methodologies.

(1) Priors for multivariate normal distributions

As discussed in Section III.5, trait covariance matrices and their inverses are often targets for inference; for example, to estimate conditional dependencies between traits or compute phylogenetic signal. However, it can be challenging to estimate these matrices as they generally have a large number of parameters; there are $\frac{1}{2}\mathcal{N}(\mathcal{N}+1)$ parameters in each of $\mathbf{\Sigma}^{phy}$ and $\mathbf{\Sigma}^{species}$, where $\mathrm{N} = \mathrm{N}_{trait}$ denotes the number of traits throughout this section. In the MR-PMM framework, the covariance matrices $\mathbf{\Sigma}^{phy}$ and $\mathbf{\Sigma}^{species}$ each parameterise a multivariate normal distribution which are priors for the vectors $\mathbf{u}^{phy}$ and $\mathbf{u}^{species}$, respectively [see Equation (15)]. In a fully Bayesian setting, it is also necessary to specify priors for the covariance matrices themselves, for which there are many different options to choose from depending on the way the model is parameterised. The choice of prior impacts on the model-fitting strategy and efficiency, as well as the degree and type of shrinkage.

The classical choices of parameterisation for the multivariate normal are the covariance and precision forms. To each of these choices, there is a conjugate prior, the inverse Wishart and the Wishart distributions, respectively, characterised by a target matrix and degree of belief parameter. Conjugate priors are an algebraically convenient choice because they provide analytic solutions for certain steps of the model fitting which can greatly reduce estimation time in a Gibbs sampling approach (the covariance form with inverse Wishart prior is implemented in *MCMCglmm*). Departing from conjugacy, regularising priors such as the zero-centred Laplacian (two-sided exponential) and Gaussian distributions can be used to "shrink" the off-diagonal elements of the covariance or precision matrix. In particular, the use of the Laplacian prior for the precision matrix, called the *graphical lasso*, is able to shrink matrix elements all the way to zero, generating a sparse network of conditional dependencies. Further examples of sparsity-inducing priors include the adaptive lasso (Zou, 2006) and the horseshoe (Carvalho, Polson & Scott, 2009) which contain global terms to control the total amount of shrinkage and local terms to allocate shrinkage flexibly on a per-parameter basis. These types of sparsity-inducing priors have been used extensively for the general problem of variable selection, and hold enormous potential for inferring trait relationships in the MR-PMM setting, especially when data are sparse relative to the total number of candidate correlations.

A shortcoming of specifying priors for the covariance (or the precision) matrix is that the variances and correlations (or the precisions and partial correlations) are not treated independently. Indeed, while the inverse-Wishart prior with degrees of freedom equal to the number of response traits (or marginally greater) is a weakly informative prior for variance components, its marginal prior for correlations concentrates probability mass away from zero, potentially inflating false positive rates (see tutorial in Section X). In practice, we may wish to apply shrinkage to the correlation coefficients while maintaining an uninformative prior on the variances, or *vice versa*. To address this issue, the covariance matrix can be decomposed as $\mathbf{\Sigma} = \mathbf{SRS}^{T}$ where $\mathbf{S} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{\mathcal{N}})$ is the diagonal matrix of standard deviations and $\mathbf{R}$ is a correlation matrix (Barnard, McCulloch & Meng, 2000). This decomposition allows priors for the standard deviations to be specified separately from those of the correlation matrix. The matrix $\mathbf{R}$ can be further factored into the form $\mathbf{R} = \mathbf{LL}^{T}$ where $\mathbf{L}$ is a lower-triangular matrix called the Cholesky factor. The latter parameterisation is implemented in *brms* where the default priors are half Student-*t* for the standard deviations and the Lewandowski–Kurowicka–Joe (LKJ; Lewandowski, Kurowicka & Joe, 2009) prior for $\mathbf{L}$. The LKJ prior has a single parameter η which for $\eta = 1$ specifies a uniform distribution in the space of correlation matrices and for $\eta > 1$ regularises estimates towards the identity where each correlation coefficient ρ_{ij} has marginal density $\mathrm{Beta}\left(\eta + \frac{\mathcal{N}-2}{2}, \eta + \frac{\mathcal{N}-2}{2}\right)$.

Use of the Cholesky factor is computationally advantageous for multivariate normal models, as it simplifies evaluation of the likelihood and allows an unconstrained parameterisation of $\mathbf{R}$ in terms of correlation angles *via* a certain geometric representation (Forrester & Zhang, 2020). The latter are efficient to sample and allow rapid evaluation of the LKJ prior while guaranteeing the validity of the corresponding correlation matrix.

When available, we recommend priors that separate correlations from variances rather than conjugate (inverse) Wishart distributions. In practice, however, the choice of prior is often constrained by the software being used. In *MCMCglmm*, the use of conjugate priors together with efficient use of the tree

structure can make it the preferred software choice when fitting time is an issue, at least for Gaussian-distributed traits. The package allows parameter-expanded versions of the prior which address some of the issues arising from inverse Wishart for small variances. The separation strategy is implemented by default in *brms*, although there is very little restriction on the choice of prior which allows the user to experiment freely with alternatives. For smaller data sets, it may be worth trying η values in the LKJ prior above the default of 1 (e.g. 5 or 10) to regularise correlations.

(2) Latent factor methods for dimension reduction

Although regularising priors can reduce the effective number of parameters for an estimated covariance matrix, for large numbers of traits it can be necessary to reduce the actual number of parameters, however regularised they may be. A commonly used technique for dimension reduction in high-dimensional multivariate statistics is latent factor analysis (also called factor analytic methods), which is based on a rank-reduced representation of a covariance or correlation matrix. The key idea of a latent factor model is that an $n \times n$ covariance $\mathbf{\Sigma}$ can be modelled using a rectangular $n \times k$ matrix $\mathbf{\Lambda}$, where $k < n$, and an $n \times n$ diagonal matrix $\mathbf{\Psi}$, such that

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}^{\top} + \mathbf{\Psi}. \quad (21)$$

The choice of k determines the rank of the approximation with smaller values yielding less-complex models. Fitting latent factor models requires attention to various technical details regarding parametrisation (Butler *et al.*, 2023), regularisation (Runcie & Mukherjee, 2013), and the identifiability of the matrix elements in $\mathbf{\Lambda}$ (Hassler *et al.*, 2022).

Latent factor models provide an effective means to reduce the complexity of high-dimensional multivariate models. In the MR-PMM setting, these models can be fitted with the R package *asreml* (Butler *et al.*, 2023) using residual maximum likelihood or using the Julia package *PhylogeneticFactorAnalysis* (Hassler *et al.*, 2022) which acts as an interface for the Bayesian phylogenetic inference software BEAST (Suchard *et al.*, 2018). Both packages offer parameterisation options and the latter pair will automatically select the number of factors k using cross validation. Neither *MCMCglmm* nor *brms* are able to fit latent factor models at this stage, although this class of models can be coded explicitly in the Stan (Stan Development Team, 2024) statistical modelling language, which is the platform underpinning *brms*.

VIII. DISCUSSION

(1) Summary

Phylogenetic mixed models are a familiar, flexible, and scalable framework for comparative analyses of species traits. Phylogenies are used to derive an expectation of similarity among trait values, which enters the model as a phylogenetic correlation matrix. Different evolutionary models are fitted *via* transformations of this matrix, or else by deriving expected pairwise phylogenetic correlations under the chosen model. For a single response trait, modelling phylogenetic structure means the proportion of variance attributable to phylogeny (i.e. phylogenetic signal) can be estimated. However, single-response models with fixed effects may confound an ensemble of trait correlations if predictor and response variables both show phylogenetic signal. MR-PMM provides a solution *via* the explicit decomposition of trait covariances. This allows correlations between response traits to be partitioned into phylogenetic and non-phylogenetic contributions, providing a more detailed and informative analysis.

(2) Strengths and weaknesses

One limitation of MR-PMMs, as currently implemented in the popular R packages *MCMCglmm* and *brms*, is the restriction to a multivariate BM model of evolution. Despite its statistical elegance, BM is clearly a gross oversimplification of the varied and dynamic processes generating phenotypic diversity over macro-evolutionary time (Uyeda *et al.*, 2018; Losos, 2011; Freckleton & Harvey, 2006). Most notably, BM does not explicitly account for adaptive evolution (e.g. stabilising and directional selection). At first glance, this appears to be a fatal weakness. However, phylogenetic signal consistent with BM is, in fact, also consistent with adaptation *via* niche conservatism; these processes can produce indistinguishable phylogenetic patterns in ecological data (reviewed in Westoby *et al.*, 2023; Wiens *et al.*, 2010). Brownian phylogenetic (co)variance is therefore relevant for understanding divergence in ecological strategy between clades, a major contributor to the functional diversity of ecosystems.

Looking beyond PMMs, there exists a broad class of processes expressible in terms of stochastic partial differential equations that characterise more complex and dynamic multivariate evolutionary models (Blomberg *et al.*, 2020; Bartoszek *et al.*, 2012; Butler & King, 2004). These include fully parameterised multivariate OU models (i.e. unstructured $\mathbf{A}$ matrices), as well as multiple optima and shift models, examples of which are implemented in *mvSLOUCH* (Bartoszek *et al.*, 2012), *mvMORPH* (Clavel *et al.*, 2015), *Rphylopars* (Goolsby *et al.*, 2017), *PCMbase* (Mitov *et al.*, 2020), and *phylosem* (Thorson & van der Bijl, 2023). Like MR-PMM, many of these methods accommodate missing data and within-species variation but offer more detailed and (potentially more) realistic evolutionary models. On the other hand, model selection approaches to identify the correct structure (e.g. for $\mathbf{A}$) are non-trivial, and may require very large comparative data sets (Bartoszek *et al.*, 2023). Furthermore, these methods are restricted to continuous traits (although see Hassler *et al.*, 2022) and do not permit an arbitrary multilevel decomposition of trait covariances, which is a desirable and informative simplification for many biological questions.

Another important limitation of the MR-PMM specified in Equation (10), is that it assumes the chosen evolutionary model is both constant through time and homogeneous across the tree, making it vulnerable to misspecification under certain evolutionary scenarios, for example rare events (Uyeda *et al.*, 2018; Clavel *et al.*, 2015). Some alternative methods allow for shifts, meaning parameters of the evolutionary model (e.g. σ^2_{phy}), even the model itself (BM, OU, etc), are allowed to vary between clades and across time (Pagel, O'Donovan & Meade, 2022; Mitov *et al.*, 2020; Mitov, Bartoszek & Stadler, 2019; Clavel *et al.*, 2015), including (when furnished with fossil data) changes in the trait-level covariance matrix itself through time (e.g., $\mathbf{\Sigma}^{phy} = \mathbf{\Sigma}^{phy}(\boldsymbol{\theta}, t)$) (Blomberg *et al.*, 2020, 2024). While methods to infer shifts in a PMM framework are likely to be hindered by computational constraints (i.e. involve fitting too many alternative models), it is possible to estimate separate phylogenetic random effects for pre-defined clades (see tutorial, Section X). However, other techniques will be more appropriate when trying to infer the timing or phylogenetic position of rare evolutionary events (Pagel *et al.*, 2022; Uyeda & Harmon, 2014; Slater, 2013).

One strength of the mixed model approach is its extension to generalised linear models (Nakagawa & Schielzeth, 2010; Hadfield, 2010). Many evolutionary hypotheses invoke causal relationships between continuous and discrete traits. However, currently MR-PMM is one of few methods available to estimate correlations between continuous and discrete traits in a fully phylogenetic framework (see also Haba & Kutsukake, 2019). Despite complications arising from non-linear latent variable transformations (see Section III.3), this has enabled several long-standing evolutionary hypotheses to be addressed in recent years (e.g. Downing *et al.*, 2020; Cornwallis *et al.*, 2017). A second strength of MR-PMMs is the capacity to specify multi-level models that appropriately capture correlated hierarchical effects (Cinar *et al.*, 2022; Nakagawa & Santos, 2012; Hadfield & Nakagawa, 2010). Integrating data from multiple studies to draw general conclusions is a common practice in comparative biology. It enables researchers to increase the power and complexity of analyses and is simply necessary to make full use of growing public data repositories. The capacity to account for multi-level structure is therefore crucial for methods to scale effectively with modern data compilations (e.g. Falster *et al.*, 2021; Kattge *et al.*, 2020).

(3) Recommendations and future directions

One challenge many analysts face when attempting a phylogenetic comparative analysis is that not all available species with trait data are featured in a suitable published phylogeny. A powerful workaround is to use the best available species-level phylogeny to derive a phylogeny for a higher taxonomic rank (e.g. series, section, or genus). This approach has two tangible benefits: (*i*) computations involving the phylogenetic correlation matrix $\mathbf{C}$ are a major bottleneck in the evaluation of the model likelihood. Therefore, reducing the dimension of $\mathbf{C}$ by refocusing to a higher taxonomic rank comes with substantial reductions in model fitting time. (*ii*) It is often possible to include data from more species, as those missing from species-level phylogenies can instead be treated as replicates of a higher taxonomic rank that does feature in the phylogeny. This method has the obvious drawback of ignoring phylogenetic structure close to the tips (e.g. sister species relationships), modelling only those effects owing to deeper splits within the tree. However, in cases such as the eucalypts (see Section VI), where taxonomic series often represent freely hybridising species groups (Pfeilsticker *et al.*, 2023; Larcombe *et al.*, 2015), modelling phylogenetic effects above the species level may in fact be desirable (see tutorial, Section X), as it relaxes the assumption that recent divergences can be expressed in terms of a strictly bifurcating evolutionary process.

A more general pitfall in phylogenetic comparative analyses is the failure to account adequately for multilevel structure. Assuming sufficient replication, the preferred approach is always to model species traits at the lowest possible level (e.g. observations on individual organisms), using a multilevel model. Partitioning variation across the model hierarchy utilises all available information, while facilitating evaluation of more interesting biological questions, for example, separation of within- and between-group variance components (e.g. species, population, individual) (Garamszegi & Møller, 2017; Westneat, Wright & Dingemanse, 2015); separation of phylogenetic and spatial effects (Gomes *et al.*, 2023; Freckleton & Jetz, 2009); and the estimation of within-species residual (co)variances (i.e. patterns of species-specific phenotypic plasticity and trait co-variation) (Goolsby *et al.*, 2017). Detailed variance partitioning may, in fact, be necessary to identify drivers of selection in natural populations. For example, partitioning variance in anti-predator behaviour of 254 bird species supported the view that within- and between-population variances are driven by different selective forces (Garamszegi & Møller, 2017). Similarly, within-species variances account for a considerable proportion of community-level variation in plant functional traits, particularly for whole-plant (e.g. plant height), rather than organ-level (e.g. leaf size), traits (Siefert *et al.*, 2015).

Integrating multilevel MR-PMMs with sophisticated evolutionary models remains a considerable methodological challenge (see Section III.2). In particular, software implementations are limited and modern comparative analyses including many thousands of taxa pose genuine computational barriers to more complex, parameter-rich evolutionary models. Of course, such challenges do not obviate the need to account for structural dependencies in comparative data, rather they call for a more tractable model. Of the available options, BM is easily the most tractable, a virtue repeated so often it risks being overlooked (Blomberg *et al.*, 2020). Nonetheless, large-scale studies, particularly in the field of plant functional ecology, continue to be published without directly accounting for phylogeny (e.g. Joswig *et al.*, 2022; Bruelheide *et al.*, 2018; Díaz *et al.*, 2016; also see Koricheva & Gurevitch, 2014). At the very least, phylogenetic models assuming BM evolution should feature in the analysis of such data sets, and MR-PMMs offer a robust, flexible

and informative framework to do so (Westoby *et al.*, 2023). However, with the inexorable expansion of biological trait databases (Falster *et al.*, 2021; Kattge *et al.*, 2020), even models assuming BM can become computationally burdensome as the number of species and traits included in analyses grows. Frequentist approaches such as ASReml offer a powerful alternative in such cases, especially for Gaussian response variables (Butler *et al.*, 2023). Future work could address computational constraints by developing approximate inference methods for MR-PMMs, such as mean field variational Bayes algorithms [Ormerod & Wand (2010), although the accuracy of covariance components for non-Gaussian response traits remains a challenge, see Hughes, García-Fiñana & Wand, 2023]. A particularly powerful approach would emerge from integrating computationally efficient algorithms for multivariate evolutionary models beyond BM (e.g. Mitov *et al.*, 2020) into the generalisable, multi-level framework of MR-PMMs.

Another clear direction for future research is a more comprehensive assessment of statistical power and sensitivity in MR-PMMs. In particular, it would be instructive to have clearer expectations around inferential power with respect to the number of response traits, the combination of response distributions, the relative magnitude of phylogenetic and non-phylogenetic (co)variances, as well as phylogenetic uncertainty and sample sizes. In general, MR-PMM is well suited to data sets with many species n relative to the number of model parameters p (i.e. high n:p ratios). Reduction of model complexity, for example *via* shrinkage (Section VII.1) or rank-reduction of the covariance matrix (Section VII.2), may be required for data sets with low n:p ratios (e.g. geometric morphometrics), for which specialised methods are also available [Clavel & Morlon, 2020; Adams & Collyer, 2019, 2018; Clavel, Aristide & Morlon, 2018; Adams, 2014; Hassler *et al.*, 2022; also see Runcie *et al.* (2021) for methods relevant to genomic prediction]. Even relatively small data sets may potentially yield useful inferences by using sparsity-inducing priors (Section VII.1) to identify a small subset of parameters that escape shrinkage to zero (Piironen & Vehtari, 2017; Li, Craig & Bhadra, 2019). However, a formal examination of power and sensitivity, provided by simulation studies, is necessary to clarify the inherent limitations of estimation and inference in an MR-PMM framework. Fortunately, the ability to deal with partially observed records often means that more data can be included in analyses.

IX. CONCLUSIONS

(1) Species phenotypes are the product of a complex network of causal processes that promote co-selection of traits, over both short and long timescales. For some traits, (co)variation is conserved over evolutionary time, creating patterns in inter-species data that reflect phylogenetic relationships (Section II).
(2) Distinguishing the influence of these conserved effects from those that are decoupled from phylogenetic history is a fundamental objective of evolutionary ecology, because it is the balance of these forces that defines whether constraints, trade-offs and coordinated strategies should be understood in terms of deep evolutionary integration or labile responses to prevailing conditions (Section IV).
(3) Univariate approaches to cross-species analysis are not parameterised to distinguish between these components of trait correlation (Section II), leading to potential confounds and inferential limitations. By contrast, MR-PMMs, in which correlations between traits can be partitioned into phylogenetic and non-phylogenetic components, offer a more informative analysis of trait evolution under many circumstances (Sections IV and VI).
(4) The capacity to fit multilevel models (Section III.1), impose different models of trait evolution (Section III.2), include continuous and discrete response variables (Section III.3), estimate partial correlations (Section III.5), and employ conditional predictions for imputation and model validation (Section V), makes MR-PMM a robust and unifying approach to many open questions in comparative biology.
(5) In particular, MR-PMM is well-equipped to advance an emerging synthesis of genetic, ecological, and evolutionary approaches to the study of trait variation. Data sets will soon exist that allow the partitioning of within- and between-species trait variances into additive genetic, phylogenetic, spatial, environmental, and residual components. We argue that careful parameterisation of MR-PMM (Sections VII.1 and VII.2) may offer a scalable framework for analysing these large complex data sets.
(6) Currently, software implementations of the various methods discussed are available piecemeal across several different applications. In practice, we found this led to different use cases for the two R packages we explored for fitting MR-PMMs. The analytic solutions leveraged by a Gibbs sampling approach give *MCMCglmm* a considerable computational advantage when all responses variables are Gaussian, with the restriction that conjugate priors be used. By contrast, the Hamiltonian Monte Carlo sampler used by *brms* is likely to outperform for non-Gaussian error distributions and allows flexible prior specification. *brms* also benefits from a user-friendly syntax and integrates well with other useful packages such as *tidyverse* (Wickham *et al.*, 2019) and *loo* (Vehtari *et al.*, 2024).
(7) We expect that more researchers in ecology and evolution will adopt MR-PMMs as barriers to implementation for non-specialists are broken down. To this end, we provide extensive tutorial materials, including annotated code and applied examples.

X. TUTORIAL

Tutorial material relevant to this article can be found at https://github.com/Benjamin-Halliwell/MR-PMM.

XI. ACKNOWLEDGEMENTS

We thank Ian Wright, Mark Westoby, and Simone Blomberg for helpful discussions during preparation of the

manuscript. We thank Shinichi Nakagawa for editorial suggestions, and Jarrod Hadfield and Pierre de Villemereuil for detailed reviews that greatly improved the manuscript. We thank Rachael Gallagher for consultation on the AusTraits database. This work was funded by The Australian Research Council Centre of Excellence for Plant Success in Nature and Agriculture (CE200100015). Open access publishing facilitated by University of Tasmania, as part of the Wiley - University of Tasmania agreement via the Council of Australian University Librarians.

XII. DATA AVAILABILITY STATEMENT

The data set utilised for this research is "AusTraits.v6.0.0" available at https://zenodo.org/records/11188867.

XIII. REFERENCES

Adams, D. C. (2014). A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution* **68**, 2675–2688.

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

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

Akaike, H. (1973). *Tsahkadsor, 1971*, pp. 267–281. Academiai Kiado, Budapest.

Arlot, S. & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys* **4**, 40–79.

Barnard, J., Mcculloch, R. & Meng, X. L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica* **10**, 1281–1311.

Bartoszek, K., Fuentes-Gonz'Alez, J., Mitov, V., Pienaar, J., Piwczyn'Ski, M., Pucha Lka, R., Spalik, K. & Voje, K. L. (2023). Model selection performance in phylogenetic comparative methods under multivariate Ornstein–Uhlenbeck models of trait evolution. *Systematic Biology* **72**, 275–293.

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

Bastide, P., Ho, L. S. T., Baele, G., Lemey, P. & Suchard, M. A. (2021). Efficient Bayesian inference of general Gaussian models on large phylogenetic trees. *The Annals of Applied Statistics* **15**, 971–997.

Björklund, M. (1997). Are "comparative methods" always necessary? *Oikos* **80**, 607–612.

Blackman, C. J., Halliwell, B., Hartill, G. E. & Brodribb, T. J. (2024). Petiole xla (xylem to leaf area ratio) integrates hydraulic safety and efficiency across a diverse group of eucalypt leaves. *Plant, Cell & Environment* **47**, 49–58.

Blomberg, S. P., Garland Jr, T. & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution; International Journal of Organic Evolution* **57**, 717–745.

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.

Blomberg, S. P., Muniz, M., Bui, M. N. & Janke, C. (2024). Multivariate trait evolution: models for the evolution of the quantitative genetic g-matrix on phylogenies. bioRxiv 2024–10.

Blomberg, S. P., Rathnayake, S. I. & Moreau, C. M. (2020). Beyond Brownian motion and the Ornstein-Uhlenbeck process: stochastic diffusion models for the evolution of quantitative characters. *The American Naturalist* **195**, 145–165.

Blomquist, G. E. (2019). Adaptation, phylogeny, and covariance in milk macronutrient composition. *PeerJ* **7**, e8085.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H. & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* **24**, 127–135.

Borgelt, J., Dorber, M., Høiberg, M. A. & Verones, F. (2022). More than half of data deficient species predicted to be threatened by extinction. *Communications Biology* **5**, 679.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., Du Plessis, L., *et al.* (2019). Beast 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **15**, e1006650.

Brodribb, T. J., Powers, J., Cochard, H. & Choat, B. (2020). Hanging by a thread? Forests and drought. *Science* **368**, 261–266.

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., Botta-Dukát, Z., Chytry, M., Field, R., Jansen, F., Kattge, J., Pillar, V. D., Schrodt, F., Mahecha, M. D., Peet, R. K., *et al.* (2018). Global trait–environment relationships of plant communities. *Nature Ecology & Evolution* **2**, 1906–1917.

Bürkner, P.-C. (2017). Brms: an R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1–28.

Bürkner, P.-C., Gabry, J. & Vehtari, A. (2021). Efficient leave-one-out cross-validation for Bayesian non-factorized normal and student-t models. *Computational statistics & data analysis* **36**, 1243–1261.

Butler, D., Cullis, B., Gilmour, A., Gogel, B. & Thompson, R. (2023). *ASReml-R Reference Manual Version 4.2*. VSN International Ltd, Hemel Hempstead, HP2 4TP, UK.

Butler, M. A. & King, A. A. (2004). Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist* **164**, 683–695.

Cardillo, M. (2003). Biological determinants of extinction risk: why are smaller species less vulnerable? In *Animal Conservation Forum* (Volume **6**), pp. 63–69. Cambridge University Press, Cambridge, UK.

Carvalho, C. M., Polson, N. G. & Scott, J. G. (2009). Handling sparsity via the horseshoe. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, 5 of Proceedings of Machine Learning Research* (eds D. van Dyk and M. Welling), pp. 73–80. Hilton Clearwater Beach Resort, Florida, USA.

Cássia-Silva, C., Freitas, C. G., Lemes, L. P., Paterno, G. B., Dias, P. A., Bacon, C. D. & Collevatti, R. G. (2020). Higher evolutionary rates in life-history traits in insular than in mainland palms. *Scientific Reports* **10**, 1–10.

Choat, B., Jansen, S., Brodribb, T. J., Cochard, H., Delzon, S., Bhaskar, R., Bucci, S. J., Feild, T. S., Gleason, S. M., Hacke, U. G., Jacobsen, A. L., Lens, F., Maherali, H., Martinez-Vilalta, J., Mayr, S., *et al.* (2012). Global convergence in the vulnerability of forests to drought. *Nature* **491**, 752–755.

Cinar, O., Nakagawa, S. & Viechtbauer, W. (2022). Phylogenetic multilevel meta-analysis: a simulation study on the importance of modelling the phylogeny. *Methods in Ecology and Evolution* **13**, 383–395.

Clavel, J., Aristide, L. & Morlon, H. (2018). A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys brain evolution. *Systematic Biology* **68**, 93–116.

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.

Clavel, J. & Morlon, H. (2020). Reliable phylogenetic regressions for multivariate comparative data: illustration with the MANOVA and application to the effect of diet on mandible morphology in phyllostomid bats. *Systematic Biology* **69**, 927–943.

Cooper, N., Thomas, G. H. & Fitzjohn, R. G. (2016). Shedding light on the "dark side" of phylogenetic comparative methods. *Methods in Ecology and Evolution* **7**, 693–699.

Cornwallis, C. K., Botero, C. A., Rubenstein, D. R., Downing, P. A., West, S. A. & Griffin, A. S. (2017). Cooperation facilitates the colonization of harsh environments. *Nature Ecology & Evolution* **1**, 1–10.

Cornwallis, C. K., West, S. A., Davis, K. E. & Griffin, A. S. (2010). Promiscuity and the evolutionary transition to complex societies. *Nature* **466**, 969–972.

de Villemereuil, P. (2018). Quantitative genetic methods depending on the nature of the phenotypic trait. *Annals of the New York Academy of Sciences* **1422**, 29–47.

de Villemereuil, P., Morrissey, M. B., Nakagawa, S. & Schielzeth, H. (2018). Fixed-effect variance and the estimation of repeatabilities and heritabilities: issues and solutions. *Journal of Evolutionary Biology* **31**, 621–632.

de Villemereuil, P., Schielzeth, H., Nakagawa, S. & Morrissey, M. (2016). General methods for evolutionary quantitative genetic inference from generalized mixed models. *Genetics* **204**, 1281–1294.

de Villemereuil, P., Wells, J. A., Edwards, R. D. & Blomberg, S. P. (2012). Bayesian models for comparative analysis integrating phylogenetic uncertainty. *BMC Evolutionary Biology* **12**, 1–16.

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bonisch, G., Westoby, M., Poorter, H., Reich, P. B., *et al.* (2016). The global spectrum of plant form and function. *Nature* **529**, 167–171.

Downing, P. A., Cornwallis, C. K. & Griffin, A. S. (2015). Sex, long life and the evolutionary transition to cooperative breeding in birds. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20151663.

Downing, P. A., Griffin, A. S. & Cornwallis, C. K. (2020). Group formation and the evolutionary pathway to complex sociality in birds. *Nature Ecology & Evolution* **4**, 479–486.

Downs, C. & Dochtermann, N. (2014). Testing hypotheses in ecoimmunology using mixed models: disentangling hierarchical correlations. *Integrative and Comparative Biology* **54**, 407–418.
Epskamp, S., Waldorp, L. J., Mõttus, R. & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research* **53**, 453–480.
Falster, D., Gallagher, R., Wenk, E. H., Wright, I. J., Indiarto, D., Andrew, S. C., Baxter, C., Lawson, J., Allen, S., Fuchs, A., Monro, A., Kar, F., Adams, M. A., Ahrens, C. W., Alfonzetti, M., *et al.* (2021). Austraits, a curated plant trait database for the Australian flora. *Scientific Data* **8**, 1–20.
Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist* **125**, 1–15.
Forrester, P. J. & Zhang, J. (2020). Parametrising correlation matrices. *Journal of Multivariate Analysis* **178**, 104619.
Freckleton, R. (2009). The seven deadly sins of comparative analysis. *Journal of Evolutionary Biology* **22**, 1367–1375.
Freckleton, R. P. & Harvey, P. H. (2006). Detecting non-Brownian trait evolution in adaptive radiations. *PLoS Biology* **4**, e373.
Freckleton, R. P., Harvey, P. H. & Page, M. (2002). Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist* **160**, 712–726.
Freckleton, R. P. & Jetz, W. (2009). Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. *Proceedings of the Royal Society B: Biological Sciences* **276**, 21–30.
Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **182**, 389–402.
Gallinat, A. S. & Pearse, W. D. (2021). Phylogenetic generalized linear mixed modeling presents novel opportunities for eco-evolutionary synthesis. *Oikos* **130**, 669–679.
Garamszegi, L. Z. (2014). *Modern Phylogenetic Comparative Methods and their Application in Evolutionary Biology: Concepts and Practice*. Springer, Berlin, Heidelberg.
Garamszegi, L. Z. & Møller, A. P. (2017). Partitioning within-species variance in behaviour to within- and between-population components for understanding evolution. *Ecology Letters* **20**, 599–608.
García-Peña, G., Sol, D., Iwaniuk, A. & Székely, T. (2013). Sexual selection on brain size in shorebirds (Charadriiformes). *Journal of Evolutionary Biology* **26**, 878–888.
Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2013). *Bayesian Data Analysis*, 3rd Edition. Chapman and Hall/CRC, London, UK.
Gomes, L. D. M., Garcia, G. S., Cordeiro, C. A., Gouveia, N. A., Ferreira, C. E., Bender, M. G., Longo, G. O., Quimbayo, J. P. & Gherardi, D. F. (2023). Complex phylogenetic origin and geographic isolation drive reef fishes' response to environmental variability in oceanic islands of the southwestern Atlantic. *Ecography* **2023**, e06559.
González-Del Pliego, P., Freckleton, R. P., Edwards, D. P., Koo, M. S., Scheffers, B. R., Pyron, R. A. & Jetz, W. (2019). Phylogenetic and trait-based prediction of extinction risk for data-deficient amphibians. *Current Biology* **29**, 1557–1563.
Goolsby, E. W., Bruggeman, J. & Ané, C. (2017). Rphylopars: fast multivariate phylogenetic comparative methods for missing data and within-species variation. *Methods in Ecology and Evolution* **8**, 22–27.
Grafen, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **326**, 119–157.
Green, S. J., Brookson, C. B., Hardy, N. A. & Crowder, L. B. (2022). Trait-based approaches to global change ecology: moving from description to prediction. *Proceedings of the Royal Society B* **289**, 20220071.
Grossnickle, D. M. (2020). Feeding ecology has a stronger evolutionary influence on functional morphology than on body mass in mammals. *Evolution* **74**, 610–628.
Gutiérrez, J. S., Piersma, T. & Thieltges, D. W. (2019). Micro-and macroparasite species richness in birds: the role of host life history and ecology. *Journal of Animal Ecology* **88**, 1226–1239.
Haba, Y. & Kutsukake, N. (2019). A multivariate phylogenetic comparative method incorporating a flexible function between discrete and continuous traits. *Evolutionary Ecology* **33**, 751–768.
Hadfield, J. & 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.
Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* **33**, 1–22.
Halliwell, B., O'connor, E. A., Uller, T., Meiri, S., Holland, B., Cornwallis, C. K. & While, G. (2024). Reproductive and ecological adaptations to climate underpin the evolution of sociality in lizards. bioRxiv 2024–05.
Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution; International Journal of Organic Evolution* **51**, 1341–1351.
Harmon, L. J., Losos, J. B., Davies, T. J., Gillespie, R. G., Gittleman, J. L., Jennings, W. B., Kozak, K. H., McPeek, M. A., Moreno-Roark, F., Near, T. J., Purvis, A., Ricklefs, R. E., Schluter, D., Schulte, J. A. I. I., Seehausen, O., *et al.* (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution; International Journal of Organic Evolution* **64**, 2385–2396.
Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. & Challenger, W. (2008). Geiger: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131.
Harvey, P. H., Read, A. F. & Nee, S. (1995). Why ecologists need to be phylogenetically challenged. *Journal of Ecology* **83**, 535–536.
Hassler, G. W., Gallone, B., Aristide, L., Allen, W. L., Tolkoff, M. R., Holbrook, A. J., Baele, G., Lemey, P. & Suchard, M. A. (2022). Principled, practical, flexible, fast: a new approach to phylogenetic factor analysis. *Methods in Ecology and Evolution* **13**, 2181–2197.
Hassler, G. W., Magee, A. F., Zhang, Z., Baele, G., Lemey, P., Ji, X., Fourment, M. & Suchard, M. A. (2023). Data integration in Bayesian phylogenetics. *Annual Review of Statistics and Its Application* **10**, 353–377.
Herberstein, M. E., Mclean, D. J., Lowe, E., Wolff, J. O., Khan, M. K., Smith, K., Allen, A. P., Bulbert, M., Buzatto, B. A., Eldridge, M. D., Falster, D., Fernandez Winzer, L., Griffith, S. C., Madin, J. S., Narendra, A., *et al.* (2022). Animaltraits – a curated animal trait database for body mass, metabolic rate and brain size. *Scientific Data* **9**, 265.
Hernández, D. G., Rivera, C., Cande, J., Zhou, B., Stern, D. L. & Berman, G. J. (2021). A framework for studying behavioral evolution by reconstructing ancestral repertoires. *eLife* **10**, e61806.
Housworth, E. A., Martins, E. P. & Lynch, M. (2004). The phylogenetic mixed model. *The American Naturalist* **163**, 84–96.
Hughes, D. M., García-Fiñana, M. & Wand, M. P. (2023). Fast approximate inference for multivariate longitudinal data. *Biostatistics* **24**, 177–192.
Ives, A. R. (2018). Mixed and phylogenetic models: a conceptual introduction to correlated data. https://leanpub.com/correlateddata.
Ives, A. R. & Garland, T. Jr. (2010). Phylogenetic logistic regression for binary dependent variables. *Systematic Biology* **59**, 9–26.
Ives, A. R. & Helmus, M. R. (2011). Generalized linear mixed models for phylogenetic analyses of community structure. *Ecological Monographs* **81**, 511–525.
Jackson, D., Riley, R. & White, I. R. (2011). Multivariate meta-analysis: potential and promise. *Statistics in Medicine* **30**, 2481–2498.
James, M. E., Wilkinson, M. J., Bernal, D. M., Liu, H., North, H. L., Engelstädter, J. & Ortiz-Barrientos, D. (2021). Phenotypic and genotypic parallel evolution in parapatric ecotypes of *Senecio*. *Evolution; International Journal of Organic Evolution* **75**, 3115–3131.
Jetz, W. & Freckleton, R. P. (2015). Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140016.
Joswig, J. S., Wirth, C., Schuman, M. C., Kattge, J., Reu, B., Wright, I. J., Sippel, S. D., Rüger, N., Richter, R., Schaepman, M. E., Van Bodegom, P. M., Cornelissen, J. H. C., Díaz, S., Hattingh, W. N., Kramer, K., *et al.* (2022). Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nature Ecology & Evolution* **6**, 36–50.
Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., *et al.* (2020). TRY plant trait database–enhanced coverage and open access. *Global Change Biology* **26**, 119–188.
Katz, D. C., Grote, M. N. & Weaver, T. D. (2016). A mixed model for the relationship between climate and human cranial form. *American Journal of Physical Anthropology* **160**, 593–603.
Kelly, R., Healy, K., Anand, M., Baudraz, M. E., Bahn, M., Cerabolini, B. E., Cornelissen, J. H., Dwyer, J. M., Jackson, A. L., Kattge, J., Niinemets, Ü., Penuelas, J., Pierce, S., Salguero-Gómez, R. & Buckley, Y. M. (2021). Climatic and evolutionary contexts are required to infer plant life history strategies from functional traits at a global scale. *Ecology Letters* **24**, 970–983.
Kontopoulos, D.-G., Van Sebille, E., Lange, M., Yvon-Durocher, G., Barraclough, T. G. & Pawar, S. (2020). Phytoplankton thermal responses adapt in the absence of hard thermodynamic constraints. *Evolution* **74**, 775–790.
Koricheva, J. & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology* **102**, 828–844.
Kruuk, L. E. & Hadfield, J. D. (2007). How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology* **20**, 1890–1903.
Larcombe, M. J., Holland, B., Steane, D. A., Jones, R. C., Nicolle, D., Vaillancourt, R. E. & Potts, B. M. (2015). Patterns of reproductive isolation in eucalyptus—a phylogenetic perspective. *Molecular Biology and Evolution* **32**, 1833–1846.
Laumer, C. E., Fernández, R., Lemer, S., Combosch, D., Kocot, K. M., Riesgo, A., Andrade, S. C., Sterrer, W., Sørensen, M. V. & Giribet, G. (2019). Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proceedings of the Royal Society B* **286**, 20190831.

Lee, T. M. & Jetz, W. (2011). Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings of the Royal Society B: Biological Sciences* **278**, 1329–1338.

Leggett, H. C., Cornwallis, C. K., Buckling, A. & West, S. A. (2017). Growth rate, transmission mode and virulence in human pathogens. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**, 20160094.

Lewandowski, D., Kurowicka, D. & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* **100**, 1989–2001.

Li, Y., Craig, B. A. & Bhadra, A. (2019). The graphical horseshoe estimator for inverse covariance matrices. *Journal of Computational and Graphical Statistics* **28**, 747–757.

Liu, H., Ye, Q., Lundgren, M. R., Young, S. N., Liu, X., Luo, Q., Lin, Y., Ye, N. & Hao, G. (2024). Phylogeny and climate explain contrasting hydraulic traits in different life forms of 150 woody fabaceae species. *Journal of Ecology* **112**, 741–754.

Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* **11**, 995–1003.

Losos, J. B. (2011). Seeing the forest for the trees: the limitations of phylogenies in comparative biology: (American society of naturalists address). *The American Naturalist* **177**, 709–727.

Lynch, M. (1991). Methods for the analysis of comparative data in evolutionary biology. *Evolution* **45**, 1065–1080.

Magwene, P. M. (2001). New tools for studying integration and modularity. *Evolution; International Journal of Organic Evolution* **55**, 1734–1745.

Markovski, J., Callaghan, C. T., Cornwell, W. K. & Nakagawa, S. (2023). A global analysis reveals the dynamic relationship between sexual selection and population abundance in space and time.

Marques, I., Kneib, T. & Klein, N. (2022). Mitigating spatial confounding by explicitly correlating Gaussian random fields. *Environmetrics* **33**, e2727.

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.

McNeish, D. (2021). Specifying location-scale models for heterogeneous variances as multilevel SEMs. *Organizational Research Methods* **24**, 630–653.

Meyer, K. (1991). Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. *Genetics Selection Evolution* **23**, 67–83.

Mitov, V., Bartoszek, K., Asimomitis, G. & Stadler, T. (2020). Fast likelihood calculation for multivariate Gaussian phylogenetic models with shifts. *Theoretical Population Biology* **131**, 66–78.

Mitov, V., Bartoszek, K. & Stadler, T. (2019). Automatic generation of evolutionary hypotheses using mixed Gaussian phylogenetic models. *Proceedings of the National Academy of Sciences* **116**, 16921–16926.

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jomba, T., Schiffers, K. & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution* **3**, 743–756.

Nakagawa, S. & de Villemereuil, P. (2019). A general method for simultaneously accounting for phylogenetic and species sampling uncertainty via Rubin's rules in comparative analysis. *Systematic Biology* **68**, 632–641.

Nakagawa, S. & Freckleton, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends in Ecology & Evolution* **23**, 592–596.

Nakagawa, S., Johnson, P. C. D. & Schielzeth, H. (2017a). The coefficient of determination r^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* **14**, 20170213.

Nakagawa, S., Noble, D. W., Senior, A. M. & Lagisz, M. (2017b). Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biology* **15**, 1–14.

Nakagawa, S. & Santos, E. S. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology* **26**, 1253–1274.

Nakagawa, S. & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews* **85**, 935–956.

Ormerod, J. T. & Wand, M. P. (2010). Explaining variational approximations. *The American Statistician* **64**, 140–153.

Ovaskainen, O. & Abrego, N. (2020). *Joint Species Distribution Modelling: With Applications in R. Ecology, Biodiversity and Conservation*. Cambridge University Press, Cambridge, UK.

Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin, T. & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters* **20**, 561–576.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884.

Pagel, M., O'Donovan, C. & Meade, A. (2022). General statistical model shows that macroevolutionary patterns and processes are consistent with Darwinian gradualism. *Nature Communications* **13**, 1–12.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika* **82**, 669–688.

Pettersen, A. K., Feiner, N., Noble, D. W., While, G. M., Uller, T. & Cornwallis, C. K. (2023). Maternal behavioral thermoregulation facilitated evolutionary transitions from egg laying to live birth. *Evolutionary letters* **7**(5), qrad031. https://doi.org/10.1093/evlett/qrad031.

Pfeilsticker, T. R., Jones, R. C., Steane, D. A., Vaillancourt, R. E. & Potts, B. M. (2023). Molecular insights into the dynamics of species invasion by hybridisation in Tasmanian eucalypts. *Molecular Ecology* **32**, 2913–2929.

Piironen, J. & Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics* **11**, 5018–5051.

Popovic, G. C., Warton, D. I., Thomson, F. J., Hui, F. K. & Moles, A. T. (2019). Untangling direct species associations from indirect mediator species effects with graphical models. *Methods in Ecology and Evolution* **10**, 1571–1583.

Pottier, P., Noble, D. W., Seebacher, F., Wu, N. C., Burke, S., Lagisz, M., Schwanz, L. E., Drobniak, S. M. & Nakagawa, S. (2024). New horizons for comparative studies and meta-analyses. *Trends in Ecology & Evolution* **39**, 435–445.

Prieto, I., Queretja, J. I., Segrestin, J., Volaire, F. & Roumet, C. (2018). Leaf carbon and oxygen isotopes are coordinated with the leaf economics spectrum in Mediterranean rangeland species. *Functional Ecology* **32**, 612–625.

Quental, T. B. & Marshall, C. R. (2010). Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology & Evolution* **25**, 434–441.

Rausher, M. D. (1992). The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution; International Journal of Organic Evolution* **46**, 616–626.

Reich, P. B. (2014). The world-wide "fast–slow" plant economics spectrum: a traits manifesto. *Journal of Ecology* **102**, 275–301.

Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution* **1**, 319–329.

Revell, L. J. & Harmon, L. J. (2022). *Phylogenetic Comparative Methods in R*. Princeton University Press, Princeton, NJ.

Revell, L. J., Harmon, L. J. & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology* **57**, 591–601.

Riley, R. D., Jackson, D., Salanti, G., Burke, D. L., Price, M., Kirkham, J. & White, I. R. (2017). Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *British Medical Journal* **358**, j3932. https://doi.org/10.1136/bmj.j3932.

Rohlf, F. J. (2001). Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution; International Journal of Organic Evolution* **55**, 2143–2160.

Runcie, D. E. & Mukherjee, S. (2013). Dissecting high-dimensional phenotypes with Bayesian sparse factor analysis of genetic covariance matrices. *Genetics* **194**, 753–767.

Runcie, D. E., Qu, J., Cheng, H. & Crawford, L. (2021). Megalmm: mega-scale linear mixed models for genomic predictions with thousands of traits. *Genome Biology* **22**, 1–25.

Sanchez-Martinez, P., Ackerly, D. D., Martínez-Vilalta, J., Mencuccini, M., Dexter, K. G. & Dawson, T. E. (2024). A framework to study and predict functional trait syndromes using phylogenetic and environmental data. *Methods in Ecology and Evolution* **15**, 666–681.

Sanchez-Martinez, P., Martínez-Vilalta, J., Dexter, K. G., Segovia, R. A. & Mencuccini, M. (2020). Adaptation and coordinated evolution of plant hydraulic traits. *Ecology Letters* **23**, 1599–1610.

Shojaie, A. & Fox, E. B. (2022). Granger causality: a review and recent advances. *Annual Review of Statistics and Its Application* **9**, 289–319.

Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W., Baraloto, C., Carlucci, M. B., Cianciaruso, M. V., Dantas, V. L., De Bello, F., Duarte, L. D. S., Fonseca, C. R., Freschet, G. T., *et al.* (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters* **18**, 1406–1419.

Skelton, R. P., Anderegg, L. D., Diaz, J., Kling, M. M., Papper, P., Lamarque, L. J., Delzon, S., Dawson, T. E. & Ackerly, D. D. (2021). Evolutionary relationships between drought-related traits and climate shape large hydraulic safety margins in western North American oaks. *Proceedings of the National Academy of Sciences* **118**, e2008987118.

Slater, G. J. (2013). Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the cretaceous-Paleogene boundary. *Methods in Ecology and Evolution* **4**, 734–744.

Slater, G. J., Harmon, L. J. & Alfaro, M. E. (2012). Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution: International Journal of Organic Evolution* **66**, 3931–3944.

Smith, S. A. & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany* **105**, 302–314.

Smith, S. D., Pennell, M. W., Dunn, C. W. & Edwards, S. V. (2020). Phylogenetics is the new genetics (for most of biodiversity). *Trends in Ecology & Evolution* **35**, 415–425.

Sorensen, D. & Gianola, D. (2002). *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics. Statistics for Biology and Health*. Springer, New York.

Stan Development Team (2024). Stan modeling language users guide and reference manual. Version 2.32.6.

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J. & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4**, vey016.

Symonds, M. R. & Blomberg, S. P. (2014). A primer on phylogenetic generalised least squares. In *Modern Phylogenetic Comparative Methods and their Application in Evolutionary Biology*, pp. 105–130. Springer, Berlin, Heidlelberg.

Thornhill, A. H., Crisp, M. D., Külheim, C., Lam, K. E., Nelson, L. A., Yeates, D. K. & Miller, J. T. (2019). A dated molecular perspective of eucalypt taxonomy, evolution and diversification. *Australian Systematic Botany* **32**, 29–48.

Thorson J. T. van der Bijl W. Phylosem: a fast and simple R package for phylogenetic inference and trait imputation using phylogenetic structural equation models *Journal of Evolutionary Biology* 2023 **36** 1357 1364

Tobias, J. A., Cornwallis, C. K., Derryberry, E. P., Claramunt, S., Brumfield, R. T. & Seddon, N. (2014). Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. *Nature* **506**, 359–363.

Tung Ho, L. S. & Ané, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology* **63**, 397–408.

Uyeda, J. C., Caetano, D. S. & Pennell, M. W. (2015). Comparative analysis of principal components can be misleading. *Systematic Biology* **64**, 677–689.

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.

Uyeda, J. C., Zenil-Ferguson, R. & Pennell, M. W. (2018). Rethinking phylogenetic comparative methods. *Systematic Biology* **67**, 1091–1109.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T. & Gelman, A. (2024). Loo: efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 2.8.0.

Vehtari, A., Gelman, A. & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432.

Warton, D. I. (2022). *Eco-Stats: Data Analysis in Ecology: From t-Tests to Multivariate Abundances*. Springer Nature, Berlin, Germany.

Warton, D. I., Blanchet, F. G., O'hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C. & Hui, F. K. (2015). So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution* **30**, 766–779.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross-validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* **11**, 3571–3594.

Westneat, D. F., Wright, J. & Dingemanse, N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews* **90**, 729–743.

Westoby, M., Leishman, M. R. & Lord, J. M. (1995). On misinterpreting the 'phylogenetic correction'. *Journal of Ecology* **83**, 531–534.

Westoby, M., Yates, L., Holland, B. & Halliwell, B. (2023). Phylogenetically conservative trait correlation: quantification and interpretation. *Journal of Ecology* **111**, 2105–2117.

Wickham, H., Averick, M., Bryan, J., Chang, W., Mcgowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., *et al.* (2019). Welcome to the tidyverse. *Journal of Open Source Software* **4**, 1686.

Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., Damschen, E. I., Davies, T. J., Grytnes, J.-A., Harrison, S. P., Hawkins, B. A., Holt, R. D., Mccain, C. M. & Stephens, P. R. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* **13**, 1310–1324.

Wilson, A. (2008). Why h^2 does not always equal Va/Vp? *Journal of Evolutionary Biology* **21**, 647–650.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., *et al.* (2004). The worldwide leaf economics spectrum. *Nature* **428**, 821–827.

Yates, L. A., Aandahl, Z., Richards, S. A. & Brook, B. W. (2022). Cross validation for model selection: a review with examples from ecology. *Ecological Monographs* **93**, e1557.

Yates, L. A., Richards, S. A. & Brook, B. W. (2021). Parsimonious model selection using information theory: a modified selection rule. *Ecology* **102**, e03475.

Yeates, D. K., Meusemann, K., Trautwein, M., Wiegmann, B. & Zwick, A. (2016). Power, resolution and bias: recent advances in insect phylogeny driven by the genomic revolution. *Current Opinion in Insect Science* **13**, 16–23.

Young, A. D. & Gillung, J. P. (2020). Phylogenomics—principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology* **45**, 225–247.

Yuan, M. & Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika* **94**, 19–35.

Zhou, J., Cieraad, E. & Van Bodegom, P. M. (2022). Global analysis of trait–trait relationships within and between species. *New Phytologist* **233**, 1643–1656.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association* **101**, 1418–1429.

(*Received 11 January 2024; revised 20 January 2025; accepted 24 January 2025; published online 7 April 2025*)