

# A Guided Tour of Phylogenetic Comparative Methods for Studying Trait Evolution

Charlie K. Cornwallis<sup>1</sup> and Ashleigh S. Griffin<sup>2</sup>

<sup>1</sup>Department of Biology, Lund University, Lund, Sweden; email: charlie.cornwallis@biol.lu.se

<sup>2</sup>Department of Biology, University of Oxford, Oxford, United Kingdom



[www.annualreviews.org](http://www.annualreviews.org)

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Ecol. Evol. Syst. 2024. 55:181–204

First published as a Review in Advance on August 6, 2024

The *Annual Review of Ecology, Evolution, and Systematics* is online at [ecolsys.annualreviews.org](http://ecolsys.annualreviews.org)

<https://doi.org/10.1146/annurev-ecolsys-102221-050754>

Copyright © 2024 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.



## Keywords

macroevolution, macroecology, coevolution, comparative genomics, ancestral reconstruction, phenotypic evolution

## Abstract

Phylogenetic comparative methods are important tools in biology, providing insights into the way traits evolve. There are many technical resources describing how these methods work. Our aim here is to complement these with an overview of the types of biological questions that can be addressed by different methods and to outline potential pitfalls and considerations when embarking on comparative studies. First, we introduce what comparative methods are and why they are important. Second, we outline how they can be used to understand when, where, and how frequently traits evolve. Third, we examine how the coevolution of traits within and between species can be studied, along with patterns of causality. Finally, we discuss how to approach comparative analyses and the ways in which different types of data, such as published relationships, omic, and remote sensing data, can be integrated.

## 1. INTRODUCTION

Phylogenetic comparative analyses have enabled inferences about the evolution of diverse biological phenomena from the colonization of land by plants to the prevalence of cancer (Alves et al. 2019, Bowles et al. 2020, Harvey & Pagel 1991, Martins 1996, Nunn 2011). Over the last 30 years, there have been important developments in the techniques and the types of data available for conducting phylogenetic comparative analyses (Garamszegi 2014, Harmon 2018, O'Meara 2012, Revell & Harmon 2022). These developments have expanded the scope of the questions that can be addressed, and there are numerous excellent technical reviews and online tutorials that provide more detailed how-to guides (Garamszegi 2014, Harmon 2018, Nunn 2011, Revell & Harmon 2022). It can, however, be difficult to navigate why particular methods are used to study specific questions and what the latest developments are. Our aim in this review is to provide researchers with a guide to identifying the questions they can ask with phylogenetic comparative methods and what to consider when setting off on such adventures.

Phylogenetic comparative methods can be used to study a range of biological processes including the evolution of traits (Adams & Collyer 2019, Hernández et al. 2013, Martins 1996), population dynamics (Germain et al. 2023), species diversification (Tobias et al. 2020), cospeciation between different organisms, such as hosts and microbes (Dismukes et al. 2022, Hadfield et al. 2014), biogeography (Ronquist & Sanmartín 2011, Sukumaran & Knowles 2018), and the structure of communities (Basu et al. 2023, Mouquet et al. 2012, Webb et al. 2002). Each of these areas requires an in-depth treatment, many of which have recently been covered. We focus on trait evolution, where recent advances have expanded the questions that can be studied and where further synthesis may help researchers take up such methods.

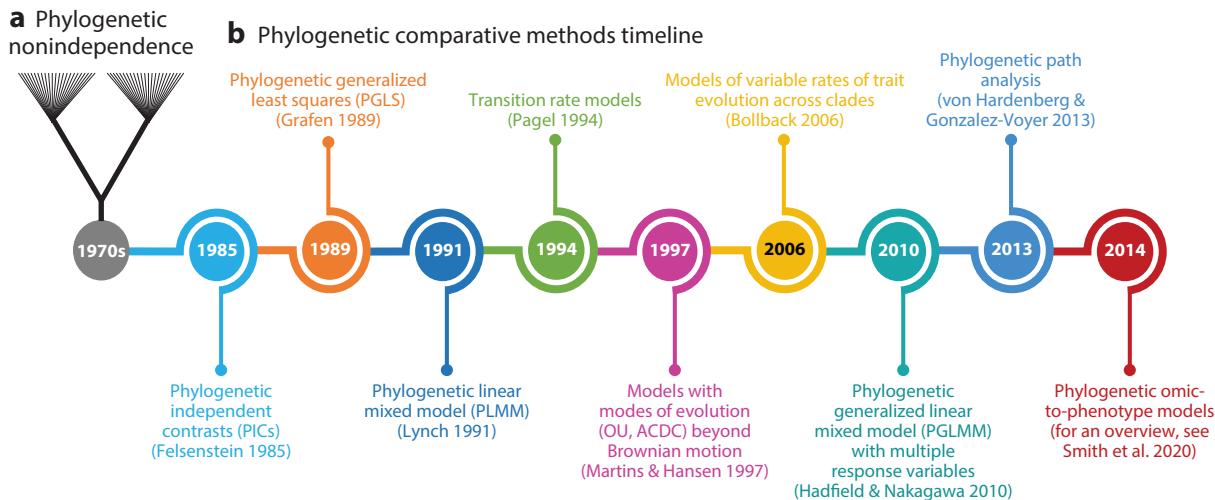
To help readers that might be new to the field, we first give a brief overview of the history of phylogenetic comparative methods and why they were developed (Section 1). Second, we examine the approaches available for understanding: (a) when in evolutionary time, where geographically, and how frequently traits have evolved (Section 2); (b) the evolutionary relationships between traits (Section 3); and (c) the way causality and the order of trait evolution can be estimated (Section 4). Third, we provide a roadmap for how to approach comparative analyses (Section 5). Finally, we discuss how comparative approaches can be used in conjunction with field studies and how they can incorporate different types of molecular, physiological, behavioral, and environmental data (Section 6).

## 2. A SHORT INTRODUCTION TO PHYLOGENETIC COMPARATIVE METHODS

### 2.1. What Are Comparative Methods and Why Are They Useful?

Phylogenetic comparative analyses require two main ingredients: information about traits across evolutionarily distinct units (usually species but sometimes others such as populations) and estimates of the ancestry between those biological units (a phylogeny). This basic setup enables variation in traits across species to be investigated over timescales that are well beyond the grasp of observational or experimental studies, providing unique insight into how evolution operates. As a result, phylogenetic comparative methods have become an indispensable toolkit for answering questions in ecology and evolution (Garamszegi 2014, Harmon 2018, Harvey & Pagel 1991, Martins 1996, Nunn 2011, Revell & Harmon 2022) (see Section 2.2).

There are many reasons why phylogenetic comparative analyses are useful, but two feature most prominently (Frumhoff & Reeve 1994, Losos 2011). First, species cannot be assumed to be independent data points (**Figure 1a**). When comparisons are made across species, relationships between traits can arise spuriously because closely related species can have more



**Figure 1**

The problem of phylogenetic nonindependence and some of the developments it stimulated. (a) Felsenstein's (1985) stark presentation of the problem of phylogenetic nonindependence. In this phylogeny, he illustrates the worst-case scenario for a phylogeny where 40 species represent a single independent comparison. (b) A powerful set of comparative methods have been developed since the problem of phylogenetic nonindependence was realized in the 1970s, catalyzed by Felsenstein (1985). Abbreviations: ACDC, accelerating–decelerating model; OU, Ornstein–Uhlenbeck model.

similar phenotypes; conversely, relationships can be hidden by large differences between clades (Felsenstein 1985, Losos 2008) (**Figure 1a**). For example, a comparison of 566 trait relationships across 65 studies showed that when species were treated as independent data points, stronger relationships and lower  $p$  values were found compared to when phylogenetic relationships were accounted for (Carvalho et al. 2006). Additionally, the evolution of some traits, for example, limb number, can be examined only by comparing across species, as there is no variation within species. Second, by integrating information on the traits of contemporary species with knowledge of their ancestry, it is possible to estimate how traits have evolved over time (Felsenstein 1985, Garamszegi 2014, Harmon 2018, Nunn 2011, Pagel 1994, Revell & Harmon 2022). In such cases, phylogenetic comparative analyses do not just statistically control for nonindependent data but can be used to explicitly study evolutionary processes from current biological patterns (Frumhoff & Reeve 1994, Losos 2011). They can infer the probable order of evolutionary events and tell us about the most likely direction of causality. In this sense, phylogenetic methods are like statistical time machines, transporting us back in time to estimate the changes in ancestral lineages that occurred as they diverged. Comparative analyses are therefore tools that can help us describe and understand diversity among species as well as test specific hypotheses about how and why traits evolve.

## 2.2. A Brief History of the Phylogenetic Comparative Method

Biologists have been using the comparative method for over a century; it was Darwin's favorite method, and in the 150 years since *The Origin of Species*, biologists have been using differences between species to test hypotheses. However, it was not until the 1970s and 1980s that statistical rigor was introduced into the process of testing evolutionary hypotheses using comparative methods. The issue of the nonindependence of species as data points first appeared in the literature as a source of concern, not from experts in statistics or systematics, but from behavioral ecologists testing hypotheses about trait adaptation in animals (**Figure 1**).

Clutton-Brock & Harvey (1977) used the comparative method to understand variation in the social organization of primate species. The discussion of nonindependence is restricted to a small section of the results where they point out that the correlation between social organization and ecology could be due to evolutionary inertia—shared ancestry of different species leading to a correlation between traits of interest—rather than cause and effect. They lament, “there is no simple answer to this problem” (Clutton-Brock & Harvey 1977, p. 8), but they were the first to try to address it by using genera, rather than species, as independent data points in their analysis.

Baker & Parker (1979) ran up against the same problem 2 years later, when analyzing bird coloration. They include a section on the problem of nonindependence in their analyses, acknowledging that their regressions are vulnerable to false positives due to some genera being more represented than others in their dataset. They attempted to address this by testing their hypotheses both within and between genera. These papers are fine examples of scientists articulating a problem in their study and offering a solution that they acknowledge is inadequate, a reminder that papers do not have to be perfect to make a lasting contribution.

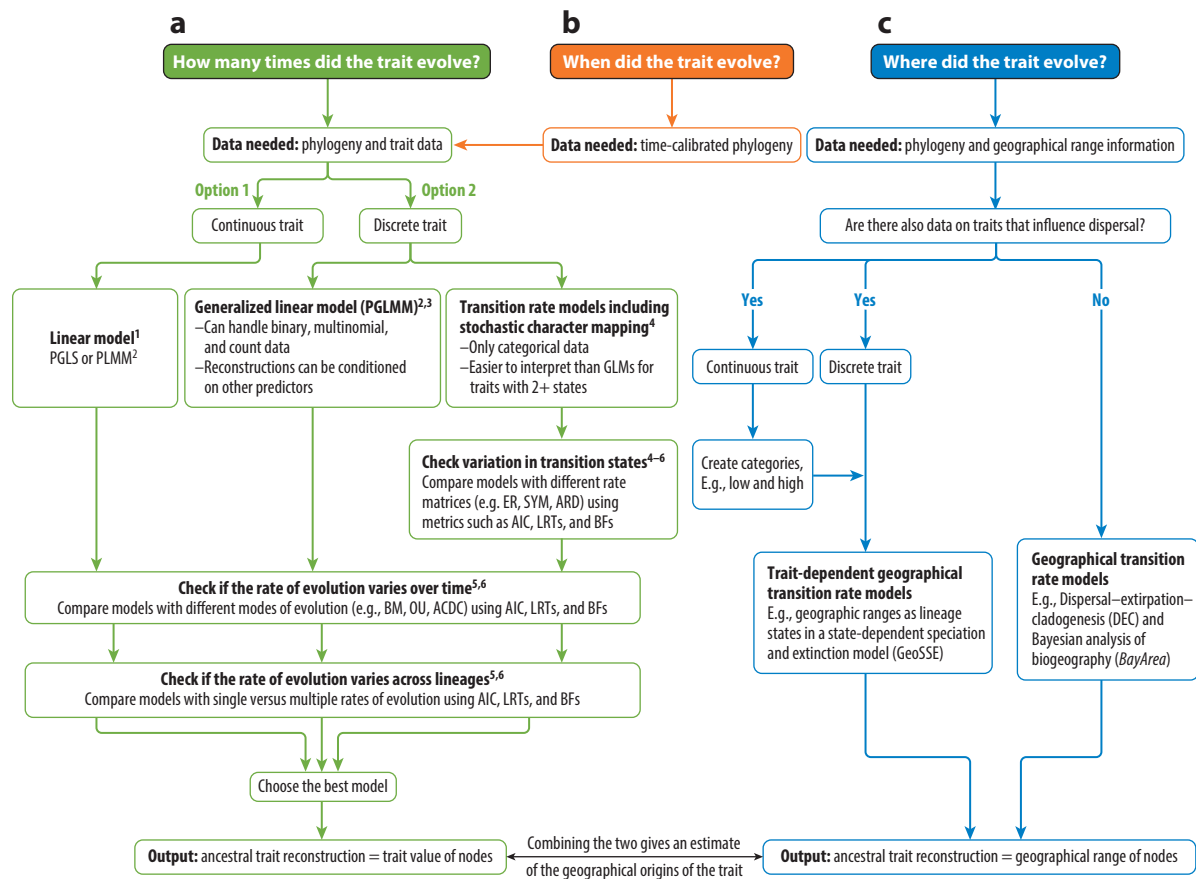
The publication of these papers did not prevent evolutionary biologists from treating species as independent data points in comparative analyses. It was not until the publication of Joseph Felsenstein’s (1985) paper “Phylogenies and the Comparative Method,” where he introduced his method of phylogenetic independent contrasts (PICs), that it really became unacceptable to ignore species ancestry (**Figure 1**). In the revealing, penultimate paragraph, Felsenstein directly addresses the paper’s reviewers who accused him of being “rather nihilistic” (Felsenstein 1985, p. 14) by insisting on the need to take phylogenetic history into account. It is worth remembering that the reviewers were writing at a time before genome sequencing, when confidence in detailed phylogenies was still a distant dream. Was Felsenstein’s logic spelling doom for a method of inference used by evolutionary biologists since Darwin? Maybe, he says in his response, but “nevertheless, efforts to cope with the effects of the phylogeny will have to be made. Phylogenies are fundamental to comparative biology; there is no doing it without taking them into account” (Felsenstein 1985, p. 14).

Felsenstein’s method was developed to take into account when phylogenies were not perfectly resolved, and computer packages were developed to implement these methods and make them accessible for any “nihilistic” PhD student who understood the need for them (Grafen 1989, Harvey & Pagel 1991, Pagel 1992) (**Figure 1**). These contributions, along with parallel advances in phylogeny construction (Baum & Smith 2012, Felsenstein 2003), made comparative methods the textbook standard and led to them being treated with the same respect as experimental data as a way of testing hypotheses in evolution (Davies & Krebs 2009, Maynard-Smith & Holliday 1979).

### 3. WHEN, WHERE, AND HOW OFTEN DO TRAITS EVOLVE?

Characterizing when, where, and how often traits evolve is crucial to explaining biological diversity. Answering these questions requires the reconstruction of ancestral trait values (**Figure 2**). Even the original debates surrounding the validity of Darwin’s theory of evolution by natural selection relied upon ancestral trait values, so-called missing links (e.g., the feathers of Archaeopteryx).

The biological insights that can be gained from reconstructing likely patterns of past evolution are exemplified by research on plants and fungi. Approximately 300 million years ago (mya), the formation of coal declined, but the reasons for this were unclear. Coal is produced by dead plant matter, especially lignin, as it is resistant to degradation. Phylogenetic studies of fungi showed that the evolution of lignin-degrading enzymes, called peroxidases, evolved approximately 290 mya in white-rot fungi, which coincided with the cessation of coal production (Ayuso-Fernández et al. 2018, Floudas et al. 2012) (**Figure 3**). Ancestral reconstructions have indicated that these

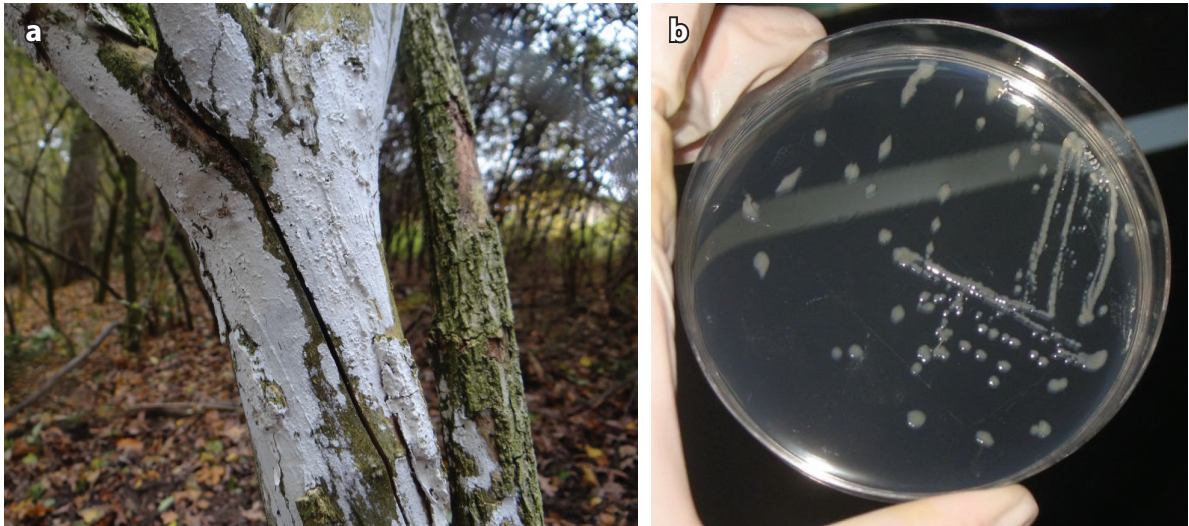


**Figure 2**

A roadmap to reconstructing the evolutionary origins of traits. Data on continuous and discrete traits can be combined with data on phylogenetic relationships to estimate when, where, and how often traits evolve using a variety of statistical models, such as PGLS, PLMM, and GeoSSE. Abbreviations: ACDC, accelerating–decelerating model; AIC, Akaike information criterion; ARD, all rates different; BF, Bayes factor; BM, Brownian motion; ER, equal rates; GeoSSE, geographic state speciation and extinction; GLM, generalized linear model; LRT, log-likelihood ratio test; OU, Ornstein–Uhlenbeck model; PGLS, phylogenetic least squares; PLMM, phylogenetic linear mixed models; SYM, symmetrical rates.

enzymes arose in two separate lineages in a remarkable example of convergent evolution that dramatically reduced global levels of undegraded plant matter (Ayuso-Fernández et al. 2018). It is also possible to use phylogenetic comparative techniques to estimate ancestral amino acid sequences (reviewed in Arenas 2022). This enabled researchers to resurrect ancestral versions of lignin-degrading enzymes at different points back in time. In the lab, they tested the efficiency of the ancestral enzymes at breaking down soft and hardwood and related this to the availability





**Figure 3**

The evolutionary history of lignin-degrading enzymes and the cessation of coal formation. (a) White-rot fungi, such as this *Phanerochaete* species, evolved lignin-degrading capabilities approximately 290 mya. Photo by Jerzy Opiola (CC BY-SA 4.0).

(b) Predicted ancestral lignin-degrading enzymes were expressed in *Escherichia coli*, and their lignin degradation capabilities were tested in the laboratory. Photo by Kelsey M. Flanagan (CC BY 3.0).

of plant hosts across time, which showed that the evolution of lignin-degrading enzymes coincided with the proliferation of angiosperm species (Ayuso-Fernández et al. 2018, 2019) (**Figure 3**).

This research demonstrates how phylogenetic comparative methods can be used in conjunction with (a) genomic information to characterize species relationships and physiological traits; (b) the fossil record to ground truth the appearance of organisms, such as white-rot fungi, as well as the cessation of geochemical cycles, such as coal production; and (c) experimental work to test the functional aspects of ancestral traits (for additional discussion of integrating different approaches, see Section 6).

### 3.1. Where Do Traits Evolve?

Phylogenies can also be used to identify the geographical areas or habitats where traits originated. For example, plants have colonized desert environments across the globe, but many of the lineages that tolerate such conditions appeared before desertification (Axelrod 1972, Lichter-Marck & Baldwin 2023). As a result, preadaptations to dry microhabitats, before deserts emerged, are thought to have played an important role in their colonization (Axelrod 1972). Phylogenetic reconstructions of ancestral biogeographic regions and biomes of rock daisies (tribe *Perityleae*) support this idea, showing that adaptation to bare rocky habitats preceded and facilitated the invasion of deserts across North America (Lichter-Marck & Baldwin 2023).

There is now a range of techniques that can be used for reconstructing the ancestral geographical ranges of species (Ronquist & Sanmartín 2011, Sukumaran & Knowles 2018). These techniques can address questions about the evolution of species biogeography and the origins of their traits across clades that have been applied to different organisms from protists (Pinseel et al. 2020) to fish (Cassemiro et al. 2023). They can also be used to trace evolutionary events at quite different scales. For example, in cancer biology, phylogenetic biogeography models have helped reconstruct the spread of cancerous cell lineages within specific patients (Alves et al. 2019).

By building phylogenies of cell lineages from different tumors, it is possible to identify the primary site of cancer and track how subsequent lineages colonize different parts of the body during metastasis (Alves et al. 2019).

### 3.2. Ancestral Reconstruction: Pitfalls and Considerations

Many of the analyses we have described so far rely on an accurate estimate of time. In some cases, time may be directly measured, such as in microevolutionary experiments or when tracking the spread of cancer cells (Alves et al. 2019). However, in other cases, we must rely on estimates from molecular clocks and/or fossil evidence. Molecular clocks offer a tractable way of estimating evolutionary time in phylogenies, especially in clades with poor fossil records (dos Reis et al. 2016). The accuracy of molecular clocks is nevertheless still controversial, even when fossil data are available. Calibrating node ages in molecular phylogenies using fossil data is clearly important, but the accuracy of the estimates is heavily dependent on the quality of fossil evidence and how this information is incorporated into the tree-building process. This is not trivial (Budd & Mann 2023, dos Reis et al. 2016). The development of new modeling techniques that incorporate different types of information (molecular, fossil, and geological) continues to offer more accurate timing estimates (Wright et al. 2022). Although accurate estimates of absolute time are particularly important for examining when and where traits evolved, the questions about trait relationships covered in later sections rely more on relative divergence times across species.

Another important issue related to time is how fast traits evolve and whether this varies across clades (Gingerich 2019). Often, rates of trait evolution are modeled as a Brownian motion process, where trait divergences across species occur randomly and continuously through time at a constant rate (Felsenstein 1985, 1988; Harmon 2018; Martins 1994; Revell & Harmon 2022). In many cases, Brownian motion appears to fit well, but rates of change in trait values can vary over time and clades. Estimating changes in the rates of trait evolution can lead to interesting insights; for example, brain neuron density rapidly evolved independently in birds and mammals (Kverková et al. 2022), there was a burst of morphological innovation in bills early on in the radiation of birds (Cooney et al. 2017), and the rate of warning-signal evolution (aposematism) dictated opportunities for nondefended mimics in Western Ghat butterfly communities (Basu et al. 2023). But even if variation in the rates of trait evolution is not the focus of one's research, failing to take them into account can lead to erroneous conclusions about how traits evolve (Freckleton 2009). For example, live birth has evolved numerous times across squamate reptiles but at different rates in different clades (King & Lee 2015). Failing to account for this rate variation leads to the ancestors of reptiles being estimated to be live bearers and many live-bearing lineages showing reversals to egg laying (Pyron & Burbrink 2014), both of which are thought to be biologically unlikely (King & Lee 2015). Using techniques that model different rates of trait evolution across clades, for example, hidden Markov models (Boyko & Beaulieu 2021, Garamszegi 2014), leads to the opposite conclusions: The ancestor of squamate reptiles was egg laying and reversals from live bearing to egg laying are rare (King & Lee 2015). Overviews of the different models of evolution, such as Ornstein–Uhlenbeck (OU), accelerating–decelerating (ACDC), and variable rates, have been published previously (Harmon 2018, Hernández et al. 2013, Pagel 1999, Powell et al. 2020).

## 4. STUDYING THE EVOLUTIONARY RELATIONSHIPS BETWEEN TRAITS

### 4.1. How Do Traits Evolve?

Many hypotheses in biology involve testing predictions about the effect of one variable on another, for example, how trade-offs between traits influence feeding specialization (Vamosi et al.

2014), whether accelerated climate change alters the timing of reproduction (Hällfors et al. 2020), and whether females prefer to reproduce with males with larger ornaments (McLean et al. 2012). These types of questions can be addressed using Felsenstein's method of independent contrasts, but this generally restricts analyses to two continuous traits. Grafen (1989) pointed out that independent contrasts are a form of linear model and that phylogenetic nonindependence of data could be accounted for in a similar way using multiple regression analyses that are more flexible. Grafen called this the phylogenetic regression, and with some modifications, it later became known as phylogenetic generalized least squares (PGLS) (Martins & Hansen 1997, Pagel 1997).

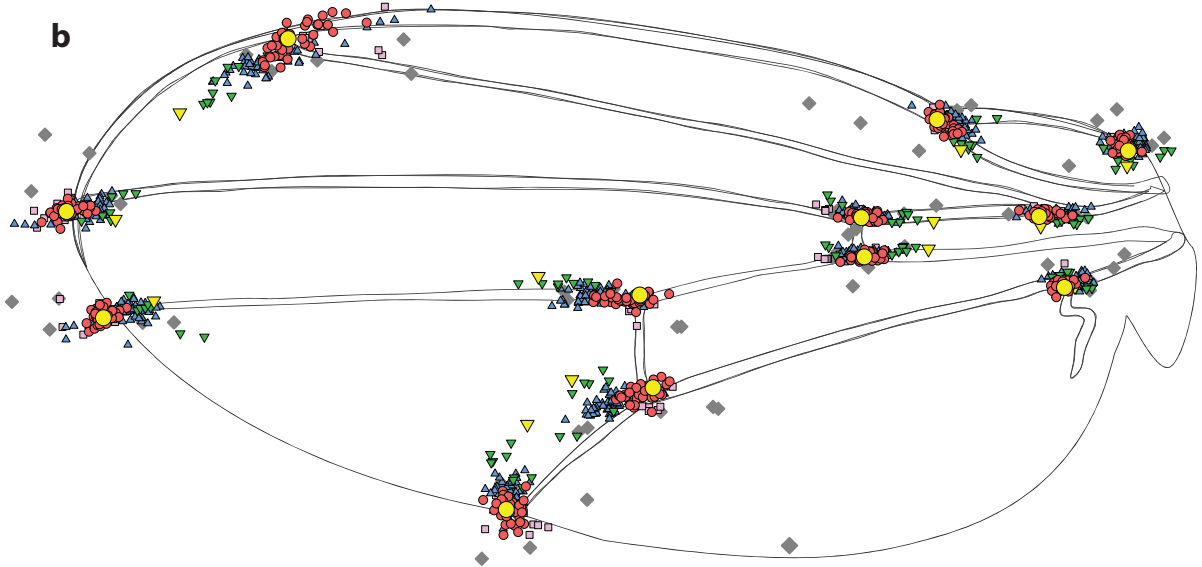
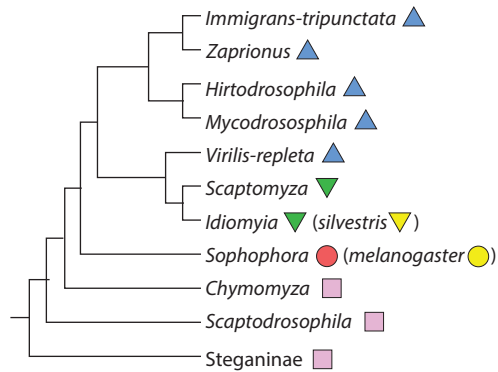
PGLS enables multiple variables to be analyzed together, allowing more nuanced tests of competing hypotheses. For example, germline mutation rates have an important influence on rates of adaptation and have been predicted to decline with effective population size and increase with traits associated with slower life histories. As effective population sizes are often correlated with life-history traits, these variables need to be analyzed together. Bergeron et al. (2023) sequenced parents and offspring across 68 vertebrate species and used PGLS to show that effective population size, generation time, age at maturity, and number of offspring produced per generation all independently influenced germline mutation rates.

While PGLS has allowed a greater variety of predictor variables to be modeled, it is restricted to analyzing variation in continuous traits (response variable), and it cannot deal with more complicated sources of nonindependent data, such as repeated measurements of species (Garamszegi 2014). These issues have been resolved by phylogenetic generalized linear mixed models (PGLMMs), aided by developments in quantitative genetics (Felsenstein 1988, Hadfield & Nakagawa 2010, Housworth et al. 2004, Lynch 1991). Multiple sources of nonindependent data can be modeled using random effects (e.g., repeated measures of species, geographical region) in PGLMMs, and new algorithms have enabled the analysis of a much larger range of response variables (e.g., binary, count, multinomial, zero-truncated traits) (Bürkner 2017, Hadfield & Nakagawa 2010). For example, discrete data, such as eye color, and count data, such as the number of fins, can be analyzed using PGLMMs.

**4.1.1. The advantages of phylogenetic mixed models.** The flexibility of PGLMMs has expanded the range of questions that can be answered using comparative analyses. One example comes from research asking whether evolutionary processes over generations (microevolution) predict long-term evolutionary patterns (macroevolution), providing insight into why some clades are more evolvable than others. The links between micro- and macroevolution have been debated, partly because of challenges in analyzing data at these different scales together (Gingerich 2019). PGLMMs have been used to show how population-level characteristics, such as mutation rates and genetic correlations between traits, can predict the divergence of traits across species over millions of years (Houle et al. 2017, Opedal et al. 2023) (**Figure 4**). More broadly, such approaches allow questions relating within-species variation (contemporary evolution) to variation across species (deeper evolutionary time). This has included analyses of how the ecology of the human gut determines levels of relatedness and cooperation of gut bacteria (Simonet & McNally 2021), how genetic diversity within populations influences the spread of disease and whether this is predicted by life-history differences across hosts (Bensch et al. 2021), and how gene expression networks influence the evolution of venom across snakes (Barua & Mikhayev 2019).

While PGLMMs can handle data collected at different biological scales (e.g., individuals, populations, species), they can also be used for meta-analyses (Adams 2008, Hadfield & Nakagawa 2010, Nakagawa et al. 2023). Meta-analytical techniques have been developed to quantify the strength of the relationships between variables (effect sizes) from summary statistics such as regression coefficients or from means and standard deviations. Effect sizes can be analyzed with PGLMMs to





**Figure 4**

Mutation rates within species of (a) *Drosophila busckii* were associated with (b) differences in standing genetic variation and wing divergence across species. By measuring 50,000 wing landmarks from 117 taxa representing 40 million years of evolution (shown in different colors), Houle et al. (2017) used phylogenetic mixed models to examine how microevolutionary parameters relate to long-term evolutionary patterns (*inset* shows the phylogeny of the major groups included in the study). Panel a photo by Charlie Cornwallis; panel b adapted with permission from Houle et al. (2017).

examine how relationships evolve across species while controlling for variation across studies, for instance, due to sample size and study design differences. Phylogenetic meta-analysis methods now enable questions related to differences in means and differences in variation to be analyzed (e.g., Bensch et al. 2021). For example, it was shown that megafauna (species of animals over 45 kg) have important effects on ecosystems across six continents, reducing the average abundance of small mammals and increasing heterogeneity in plant cover and biomass (Trepel et al. 2024).

**4.1.2. Phylogenetic regressions: pitfalls and considerations.** PGLS and PGLMMs have features that are important to highlight. First, PGLS and PGLMMs can be used for reconstructing ancestral values of traits (Goolsby 2015, Hadfield & Nakagawa 2010, Martins & Hansen 1997) (Figure 2) by predicting the response variable at internal nodes [reported as residuals from the

intercept that are also known as best linear unbiased predictors (BLUPs) (Hadfield & Nakagawa 2010, Lynch 1991)]. It is not uncommon to see studies that examine relationships between traits using PGLS or PGLMMs to then use different modeling techniques for ancestral reconstructions, which seems unnecessary. Also, variance in the tree tip and node estimates (BLUPs) relative to the total trait variance is a measure of the strength of the phylogenetic effects, sometimes referred to as phylogenetic heritability, that can circumvent separate tests of phylogenetic signal (Hadfield & Nakagawa 2010, Lynch 1991, Nakagawa et al. 2017).

Second, the need to consider the rate of trait evolution across time and across different clades (Section 2) also applies to analyzing relationships between traits. By default, most implementations of PGLS and PGLMMs assume response variables evolve via Brownian motion. It is therefore important that this assumption is checked before using linear models, and if not met, there are a range of options for analyzing traits whose rates of evolution deviate from Brownian motion (Harmon 2018, Hernández et al. 2013, Pagel 1999, Powell et al. 2020).

Third, regression models make assumptions about predictor variables that may not necessarily hold: They are assumed to causally influence the response variable, be measured without error, and consist of phylogenetically independent observations [only phylogenetic dependence in the response variable is modeled (Westoby et al. 2023)]. However, predictor variables are often traits that are evolving, and it can be more appropriate to estimate the phylogenetic correlations between traits (Westoby et al. 2023) (**Figure 5**). In more drastic cases, ignoring phylogenetic dependence in predictor variables can lead to relationships being both over- and underestimated (Sanchez-Martinez et al. 2020, Westoby et al. 2023).

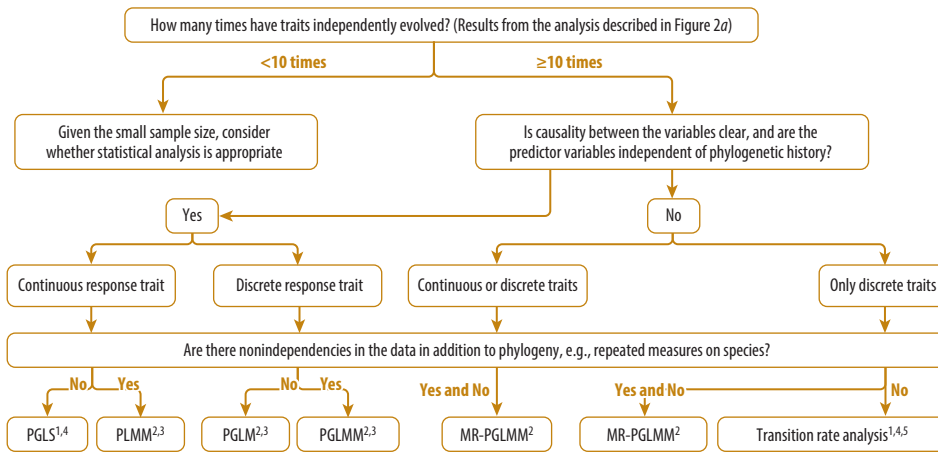
## 4.2. How Do Traits Coevolve?

Questions about the way traits coevolve are similar to the questions discussed in Section 3.1 that were concerned with the relationships between traits. However, studies that explicitly estimated the coevolutionary relationships between traits historically examined discrete characters, whereas regression-based methods have more often been applied to continuous traits (Felsenstein 1985, Maddison 1990, Pagel 1994). Models of discrete trait evolution have focused on estimating rates of change between character states over evolutionary time (transition rates), such as changes in DNA bases (Jukes & Cantor 1969) and the evolution of mating strategies (Pagel 1994, Ridley 1983).

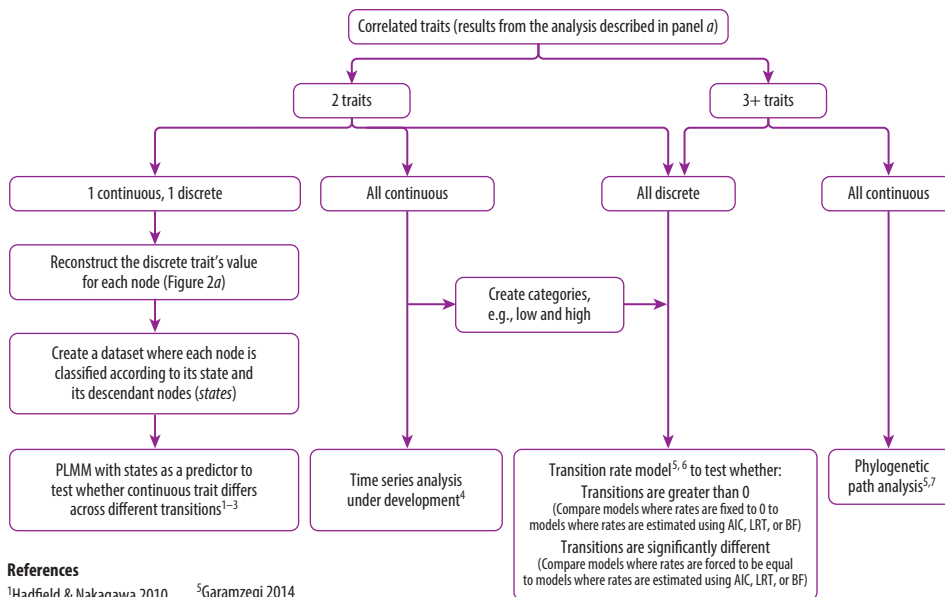
Transition rate models were extended to analyze correlated evolution between multiple discrete characters across phylogenies (Maddison 1990, Pagel 1994). They compare the explanatory power of models where changes in one trait depend on the states of other traits (coevolutionary models) to models where rates vary independently across traits (independent evolution models) (Pagel 1994). For example, transition rate models have been used to understand variation in the structure of human societies. Human ritual sacrifice has been proposed to evolve as a mechanism to control stratification in wealth and maintain social class structure (Winkelman 2014). This idea has been controversial in the absence of analyses controlling for nonindependence among human cultures. Watts et al. (2016) used models of discrete trait evolution to estimate whether rates of change in social stratification between egalitarian, moderately stratified, and highly stratified societies differed according to the presence and absence of ritual sacrifice (Watts et al. 2016) (**Figure 6**). They found that models where states of social stratification were allowed to coevolve with the presence of human ritual sacrifice best explained the data (Watts et al. 2016).

The possibilities for analyzing coevolutionary relationships between traits have also been expanded by developments in linear modeling. The need to estimate genetic correlations between traits using pedigree data in animal and plant breeding programs led to the development

## a How traits coevolve



## b Causality between traits and the order of trait evolution

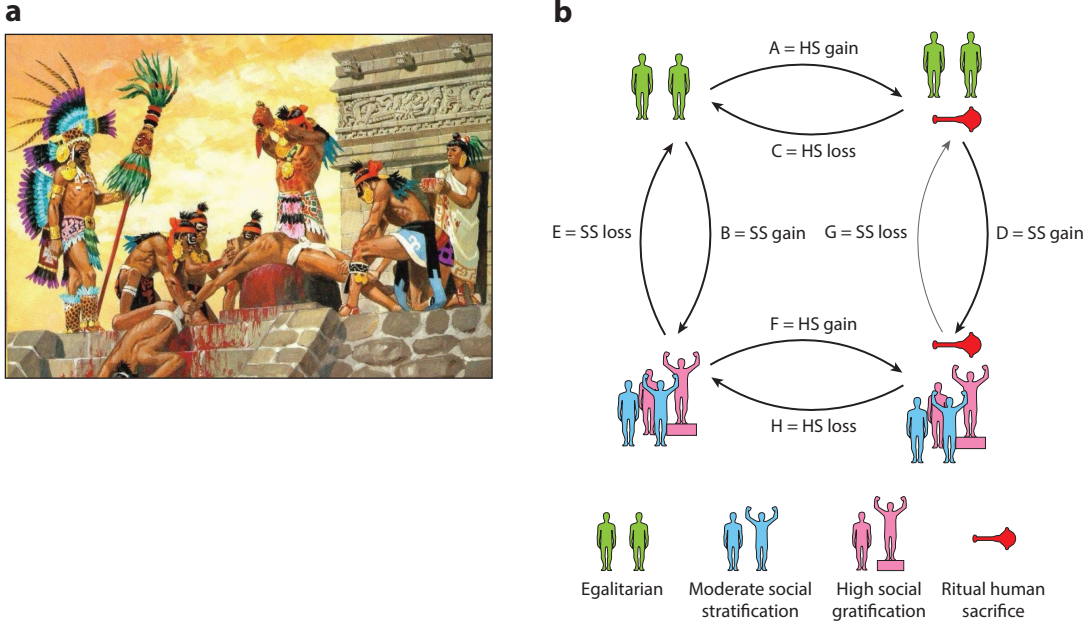


### References

- <sup>1</sup>Hadfield & Nakagawa 2010    <sup>5</sup>Garamzegi 2014  
<sup>2</sup>Gallinat & Pearse 2020    <sup>6</sup>Revell & Harmon 2022  
<sup>3</sup>Hansen et al. 2022    <sup>7</sup>van Hardenberg 2013  
<sup>4</sup>Sheenan et al. 2023

### Figure 5

A roadmap to using phylogenetic comparative methods to (a) address how traits evolve and (b) infer causality and the order of trait evolution. Abbreviations: AIC, Akaike information criterion; BF, Bayes factor; LRT, log-likelihood ratio test; MR-PGLMM, multi-response PGLMM; MR-PLMM, multi-response PLMM; PGLM, phylogenetic generalized linear model; PGLMM, phylogenetic generalized linear mixed model; PGLS, phylogenetic generalized least squares; PLMM, phylogenetic linear mixed model.



**Figure 6**

Coevolution of human societal traits and the order they evolved in. (a) Human ritual sacrifice evolved as a way to maintain social class structure and stratification of wealth. (b) A transition rate analysis examining the coevolution of human sacrifice (HS) and high social stratification (SS). Capital letters (A–H) represent the rates at which the indicated transitions occur, and the thicknesses of the arrows are proportional to these rates of change between states. The results of the analysis show that HS prevents the loss of SS, i.e., the rate of SS loss is lower in societies in which HS is present (G) compared to that in societies where HS is not present (E). Panel a created by Daphne Breeman (CC BY-SA 4.0); panel b adapted with permission from Watts et al. (2016).

of multi-response generalized linear mixed models (MR-GLMMs) (Sorensen & Gianola 2002, Walsh & Lynch 2018). MR-GLMMs analyze variation in multiple response traits and allow their covariances to be estimated at different levels, such as individuals, cohorts, and genotypes, as specified by random effects. When phylogeny is fitted as a random effect (MR-PGLMMs), multiple-response models can be used to estimate phylogenetic correlations between traits.

As MR-GLMMs can model both continuous and discrete response variables together, it is possible to examine the coevolutionary relationships between a much greater range of traits. For example, species that reproduce early in life are predicted to evolve behaviors that quickly increase resource and mating acquisition. Eckerström-Liedholm et al. (2019) repeatedly measured the behavior and metabolic rates of 20 species of killifish and used MR-PGLMMs to analyze correlations between traits within species and across the phylogeny (Eckerström-Liedholm et al. 2019). This revealed that trade-offs between current and future reproductive success were more strongly related to mating behaviors compared to trade-offs between traits involved in energy acquisition (Eckerström-Liedholm et al. 2019).

### 4.3. How Do Traits of Different Species Coevolve?

Coevolution between different species, such as mutualisms, host–parasite interactions, and predation, have featured prominently in ecological and evolutionary studies (Dismukes et al. 2022, Douglas 2022, Mouquet et al. 2012). As a result, a range of phylogenetic methods are available to test for cospeciation between different groups of organisms (Dismukes et al. 2022,

Hadfield et al. 2014, Hembry & Weber 2020). While tests of cospeciation reveal whether combinations of species have been associated over time, they do not test whether traits in one group of species (e.g., hosts) coevolve with the traits of another group of species (e.g., parasites).

Extensions to PGLMMs can be used to analyze how the traits of different groups of organisms coevolve (Gallinat & Pearse 2021, Hadfield et al. 2014). For example, in studies of hosts and parasites, the effects of host phylogenetic history, parasite phylogenetic history, and their interaction can be fitted as different random effects to understand their effect on different processes, such as variation in microbe abundances across hosts and how the traits of hosts relate to the traits of parasites, in addition to patterns of cospeciation (Hadfield et al. 2014). PGLMMs have the advantage that additional variables can be included in models to examine patterns at different scales [e.g., geographical region can be fitted as a random effect (Gallinat & Pearse 2021, Hadfield et al. 2014)] and control for variation in data quality [e.g., sequencing depth fitted as a fixed effect (Pankey et al. 2022)].

#### 4.4. Phylogenetic Correlations: Pitfalls and Considerations

An important distinction between phylogenetic regressions and phylogenetic correlations is that missing data are handled differently. In regression analyses, missing data are not permitted in predictor variables. As a result, datasets can be greatly reduced in size when multiple predictor variables are included in analyses and each has nonoverlapping missing values, which can be common in comparative datasets (Molina-Venegas 2023, Nakagawa & de Villemereuil 2019). MR-PGLMMs permit missing data in response variables, allowing correlations between traits to be estimated without pruning datasets down to complete information. Missing values in response variables are predicted based on the effects in the models (fixed and random), such as phylogeny (Nakagawa & de Villemereuil 2019, Thorson et al. 2023). Studies have shown that the relationships between variables can be accurately estimated with quite high levels of missing data (Hadfield & Nakagawa 2010, Nakagawa & de Villemereuil 2019). Such models can be used to generate predictions for species that lack trait information across different contexts, such as different regions or different habitats, which can be used to direct subsequent field or experimental data collection or inform conservation practices (Arnold & Nunn 2010, Debastiani et al. 2021, Goberna & Verdú 2016, Mouquet et al. 2012).

### 5. INFERRING CAUSALITY AND THE ORDER OF TRAIT EVOLUTION

Comparative analyses provide correlative evidence, but we often want to gain insight into the causal links between traits. For example, cooperative breeding has been predicted to be more likely to evolve when females mate monogamously, as this increases both relatedness within families and the genetic benefits of helping parents raise more siblings (Boomsma 2007). Consistent with this idea, monogamy is positively correlated with cooperative breeding across birds (Cornwallis et al. 2010). However, living in family groups may reduce the opportunities for multiple mating, leading to the opposite direction of causality. How do we tell whether this correlation is more likely due to monogamy and higher relatedness affecting the evolution of cooperative behavior or to cooperative behavior influencing monogamy (Cornwallis et al. 2010, Dillard & Westneat 2016)?

An important way to test the effect of one trait on another is to complement comparative analyses with follow-up experiments (see Section 6) (Arnold & Nunn 2010). Such experiments can, however, be very challenging or even impossible to conduct. Analytical techniques that provide insights into the most likely causal relationships between traits are therefore important tools, enriching our knowledge of the evolutionary process.



## 5.1. Discrete Traits

Inferences about the causal relationships between discrete traits can be gained using transition rate analyses, such as those used to investigate ritual sacrifice and the structure of human societies (Watts et al. 2016). By comparing whether transitions from egalitarian to highly stratified human societies occur more frequently from a state of ritual sacrifice, versus a state of no sacrifice, it was deduced that sacrificial practices stabilize stratified social class systems (**Figure 6**). In the case of cooperative breeding and monogamy, low levels of polyandry were found to precede the evolution of cooperative breeding, rather than levels of polyandry changing after cooperative breeding evolved (Cornwallis et al. 2010). Similar types of analyses have been used to investigate causality across a wide range of biological scenarios, including the effect of internal fertilization on sperm morphology across animals (Kahrl et al. 2021), whether antibiotic tolerance evolves before resistance in cystic fibrosis infections (Ghoul et al. 2023), and whether brood parasitism evolves in response to cooperative breeding across birds (Wells & Barker 2017).

Transition rate analyses are restricted to discrete traits, but methods are available to infer how discrete traits influence rates of continuous trait evolution (Hansen et al. 2022, May & Moore 2020). For example, many insects that feed on nutritionally imbalanced food sources rely on microbial symbionts to supplement their diets (Douglas 2022). Whether such relationships occur because symbionts allow nutrient-imbalanced resources to be colonized, or whether insects feeding on such resources recruit symbionts to supplement their diets, has been unclear. Reconstructing the presence of symbionts (discrete trait) and using this to predict changes in nutrients (continuous traits) showed that symbiosis evolved before the invasion of feeding niches that are depauperate in B vitamins (Cornwallis et al. 2023).

## 5.2. Continuous Traits

Estimating the causal relationships between continuous traits has been more challenging, but developing path analysis to include phylogenies has helped (von Hardenberg & Gonzalez-Voyer 2013). Path analysis works by establishing a set of potentially plausible biological hypotheses that specify how traits affect each other. Evidence for these different hypotheses can be evaluated using a series of phylogenetic regressions in which the proposed causal relationships between traits are dictated by which variables are predictors and which variables are response variables (von Hardenberg & Gonzalez-Voyer 2013). For example, the increase in global urbanization has led to great interest in the traits that enable species to adapt to urban environments (Szulkin et al. 2020). This has been challenging to address because many behavioral, ecological, and life-history traits are intercorrelated. Santini et al. (2019) used phylogenetic path analysis to estimate the causal relationships between traits that influenced the presence of different groups of mammals in urban environments. Their analyses showed that litter sizes consistently increase in urban environments but that changes in other traits, such as body size and diet diversity, differ across mammal groups (Santini et al. 2019).

## 5.3. Inferring Causality: Pitfalls and Considerations

The lack of techniques for investigating causal relationships between continuous traits has meant that researchers have often categorized continuous traits (e.g., high versus low) and used transition rate models to investigate causality. How traits are categorized can have an influence on the outcome of such analyses, so it is important to examine the sensitivity of results to these decisions. While phylogenetic path analysis can avoid the pitfalls of categorizing continuous traits, it is not appropriate for all situations: It requires relatively large sample sizes, and it cannot be used to assess causality between just two variables.

Examining the causality between continuous traits is an area that would benefit from further development. Adopting techniques used for time-series analysis may offer some options. The branch lengths of phylogenetic trees provide an estimate of evolutionary time that can then be used to test whether changes in one continuous trait precede or follow the changes in a second continuous trait. For example, across human societies, religion and politics are recognized to be historically intertwined, but how they have coevolved has been less clear. Sheehan et al. (2023) recently used phylogenetic methods based on time-series analysis to test whether continuous estimates of religious authority predict subsequent changes in continuous estimates of political authority and vice versa across 97 human Austronesian societies.

## 6. HOW TO APPROACH A PHYLOGENETIC COMPARATIVE ANALYSIS

How does one go about designing a phylogenetic comparative study to address questions about trait evolution? Not all datasets are suitable for addressing the question of interest, so how do we proceed? What we say here is not new, but checking the suitability of datasets is often forgotten with the enthusiasm for running new analyses. It can be useful to assess the potential limitations of datasets by going through a series of questions (**Figures 2 and 6**), which can also help clarify what analytical techniques are most appropriate for particular studies (for a list of analytical techniques with example studies and the software, see **Table 1**).

### 6.1. How Large of a Sample Size Do I Need?

If the aim of the study is to examine where, when, or how often traits evolve, then the answer is relatively straightforward. The set of species that is chosen depends on what biological insight the investigators want to gain. This may require thousands of species, for example, if the aim is to address how many times flight has been lost in birds, whereas only a few species may be needed for other questions, such as whether the ancestor of birds was flightless or not (Sackton et al. 2019).

For studies examining the evolutionary relationships between traits, the question of how many species are needed may be more difficult to answer. Estimating the sample size required to detect an effect can be challenging for any statistical analysis (Nakagawa et al. 2024), but for comparative analyses, there are extra considerations (Boettiger et al. 2012). As Felsenstein pointed out in his nihilistic worst-case scenario (Section 2.2), the power of statistical analyses depends on how traits are distributed across the phylogeny (Boettiger et al. 2012, Felsenstein 1985, Frumhoff & Reeve 1994, Uyeda et al. 2018). This is easily demonstrated for discrete characters, where a single origin among hundreds of sampled species prevents any statistical tests regarding that trait's evolution. The problem arising from treating species as replicates also applies to continuous traits but can be less obvious to spot. Plotting differences in trait values from the mean across the phylogeny can be useful to assess the number of independent increases and decreases in traits across the sample. A crucial first step in determining the limitation of any comparative dataset, or in selecting which species to sample, is therefore estimating how frequently traits evolve or change over evolutionary time (**Figures 2 and 6**).

### 6.2. At What Scale Do Traits Vary?

If the rate of trait evolution is too fast or too slow across the sampled set of species, the inferences that can be made from comparative analyses can be limited (Frumhoff & Reeve 1994). Basic measurements of how much traits vary between closely and more distantly related species can be very informative. Relatively little difference between groups of species that vary in their divergence times is an indication that traits are evolving slowly and wider taxonomic sampling may be needed to capture deeper evolutionary timescales. Relatively large differences, even between

Table 1 Tools for implementing phylogenetic comparative methods

Response traits	Number of traits that can be analyzed as response variables	Explanatory variables	Number of traits that can be analyzed as explanatory variables	Additional nonindependent data <sup>a</sup>	Methods	Example software	Example references
Cont	1	Cont & Cat	1+	No	Phylogenetic generalized least squares	<i>nlme</i> <i>Caper</i> <i>phylolm</i>	Basu et al. 2023, Bergeron et al. 2023, Garcia et al. 2020
Cont or Cat	1	Cont & Cat	1+	No	Phylogenetic generalized linear models	<i>phylolm</i>	Ives & Garland 2014, Garamszegi 2014 (chapter 9)
Cont	1	Cont & Cat	1+	No	Phylogenetic regression with variation in rates of evolution across clades	<i>RRphylo</i> <i>bayou</i> <i>BayesTraits</i> <sup>b</sup>	Basu et al. 2023, Cooney et al. 2017, Kverková et al. 2022
Cont or Cat	1	Cont & Cat	2+	No	Phylogenetic path analysis	<i>phylopath</i>	Santini et al. 2019, von Hardenberg & Gonzalez-Voyer 2013
Cont & Cat	1+	Cont & Cat	1+	Yes	Phylogenetic generalized linear mixed model	<i>phyr</i> <i>brms</i> <i>MCMCglmm</i>	Barua & Mikhayev 2019, Cornwallis et al. 2023, Eckerström-Liedholm et al. 2019, Houle et al. 2017, Opedal et al. 2023
Depends on software	Depends on software	Cont & Cat	1+	Yes	Phylogenetic meta-analysis	<i>Metafor</i> <i>brms</i> <i>MCMCglmm</i>	Bensch et al. 2021, McLean et al. 2012
Depends on software	Depends on software	NA	NA	Depends on software	Phylogenetic correlations	<i>MCMCglmm</i> <i>corHMM</i> <i>ratematrix</i> <i>BayesTrait</i> <sup>b</sup> <i>RevBayes</i> <sup>b</sup>	Cooney et al. 2017, Cornwallis et al. 2023, Eckerström-Liedholm et al. 2019
Cat	1+	NA	NA	No	Trait transition rate analyses	Stochastic character mapping ( <i>RevBayes</i> , <sup>b</sup> <i>Phytools</i> ) <i>BayesTraits</i> <sup>b</sup> <i>corHMM</i>	Floudas et al. 2012, Ghoul et al. 2023, Kahrl et al. 2021, King & Lee 2015, Warrs et al. 2016, Wells & Barker 2017
Cat	1	NA	NA	No	Geographical transition rate models (biogeographical models)	<i>GeoGeoBears</i> <i>Diversitree</i> <i>RevBayes</i> <sup>b</sup> <i>BayArea</i>	Alves et al. 2019, Cassemiro et al. 2023, Lichter-Marcik & Baldwin 2023, Pinseel et al. 2020

There are numerous techniques to address comparative questions, and the range of software available to implement these techniques is vast and continues to grow. Deciphering where to start can be overwhelming. Here, we list some key methods and how they can be implemented, centered mainly on R due to the ease of use. This is by no means exhaustive, so practitioners should use this only as a starting point.

<sup>a</sup>Can nonindependent data (in addition to phylogenetic history) be analyzed?

<sup>b</sup>Standalone software; all others are available in R.

Abbreviations: Cat, categorical; Cont, continuous; NA, not applicable.

closely related species, can be a sign of the opposite problem—traits are evolving fast, eroding any phylogenetic signal. In such cases, if the aim is to gain insights into the evolutionary process, for example, by reconstructing ancestral trait values, then sampling more closely related species may be needed, although traits may be variable for other reasons, such as measurement error. Low phylogenetic signal is, of course, not a problem for studies where estimating evolutionary effects is not the focus.

### 6.3. What if Data Are Missing for Some Species?

As mentioned above, there are phylogenetic comparative methods (e.g., MR-PGLMM) that can cope with some degree of missing data, but how does one decide if this is too much? Some key considerations include how missing data are distributed across the phylogeny, the degree to which missing data can be predicted by phylogenetic history, and the strength of relationships with other traits. This is an area of active development, and techniques are available for assessing each of these points and evaluating the accuracy with which missing data can be predicted (Debastiani et al. 2021, Molina-Venegas 2023, Nakagawa & de Villemereuil 2019).

### 6.4. What if the Quality of Data Varies Across Species?

The quality of data for each species may vary; for instance, different numbers of individuals may have been studied for different species. Phylogenetic meta-analysis has been developed to account for differences in the accuracy of data by weighting data points by their estimated variances [e.g., studies with fewer individuals generally produce more variable data (Adams 2008, Hadfield & Nakagawa 2010, Nakagawa et al. 2023)]. It is also worth considering whether there are other factors that can influence the accuracy of data and how they can be incorporated into models. For example, some species may have been studied for longer periods of time or studied in more places than others, and projects that include genome data might want to include measures of sequencing quality.

### 6.5. What About Uncertainty in Phylogenetic Relationships?

With the increased availability of sequence data, the accuracy of phylogenies has greatly improved, but uncertainty around relationships between lineages remains. Tree building techniques document this uncertainty in different ways, such as bootstrap values in maximum likelihood analyses or by producing sets of the most likely trees in Bayesian analyses (Baum & Smith 2012). Such estimates of uncertainty can be integrated into comparative analyses. For example, Bayesian comparative methods use algorithms that iteratively estimate effects (e.g., the regression coefficient of trait  $x$  on trait  $y$  or the amount of variance in trait  $y$  explained by the phylogeny). Uncertainty in phylogenetic relationships can be incorporated into such models by inputting a different tree from a sample of likely trees across iterations (Garamszegi 2014, Pagel & Meade 2006, Revell & Harmon 2022).

## 7. DEEPENING OUR UNDERSTANDING OF HOW ORGANISMS EVOLVE BY INTEGRATING DIFFERENT SCIENTIFIC APPROACHES

### 7.1. The Feedback Between Comparative Studies and Data Collection

There can be tension between researchers with muddy boots collecting data in the field and those in front of computers reducing a lifetime's worth of data to a single data point in comparative analyses. In reality, comparative analyses cannot be done without empirical data on each species, and comparative analyses can be informative for directing future empirical work [e.g.,

phylogenetic targeting (Arnold & Nunn 2010)]. They are complementary approaches that provide us with different insights into the ecology and evolution of species. Here, we discuss some of the potential misconceptions concerning the extent to which field and experimental studies can be integrated with comparative studies.

**7.1.1. Traits that are highly variable within species are not suitable for comparative analyses.** Comparative studies in ecology and evolution have focused on between-species variation and characterizing traits at the species level (e.g., average body size). This has sometimes led to the belief that comparative methods are not suited to analyzing the evolution of traits that vary substantially within species, but this depends on the sources of within-species variation and the aims of the study (Garamszegi 2014).

First, it is necessary to determine the extent to which within-species variation is due to measurement error, using repeatability analyses (Nakagawa et al. 2017). Low repeatability due to high measurement error is a potential problem for all studies, not just comparative analyses. Solving such issues often requires more accurate measurements or more thorough methodology (e.g., more measurements per individual). Second, it is important to establish whether traits change across environments and whether this is relevant to the question being asked. If trait changes are not directly relevant to the study, then data should be collected while controlling for such environmental effects (e.g., all individuals are measured under the same environmental condition). If trait changes are relevant to the study, then responses need to be quantified by repeatedly measuring individuals across environments (sometimes referred to as the evolution of reaction norms or random slopes) (Goolsby 2015). Under such circumstances, comparative analyses can be used to understand how the relationships within species evolve across species (Bensch et al. 2021, Eckerström-Liedholm et al. 2019, Hällfors et al. 2020, Houle et al. 2017, Opedal et al. 2023). This can be done by calculating the effect size of the relationship for each species (Nakagawa et al. 2023) or by analyzing raw data with PGLMMs. For example, variation in how much traits change across environments due to species differences and phylogenetic history can be quantified by fitting random slopes at the species and phylogenetic level in PGLMMs (Gallinat & Pearse 2021).

Third, is the goal of the study to gain insight into how traits evolve, for example, using ancestral reconstructions? If so, then it is important to quantify whether within-species variation swamps any signal of between-species variation (low phylogenetic signal). If this is the case, it is not possible to use phylogenies to predict differences between species and their ancestral nodes. Low phylogenetic signal in average trait values does not, however, always mean there is low phylogenetic signal in how traits change with environments. For example, average body temperature across a group of animals may be similar and have low phylogenetic signal, but the way animals respond to changes in temperature (reaction norm) may be quite different and depend on their evolutionary history. If the study is not concerned with estimating patterns of past evolution, then greater within-species variation relative to that between species may not be a problem.

**7.1.2. Species complexities are not captured by the simplistic measures used in comparative studies.** The biology of individual species is extremely complicated, and these complexities are often averaged away in summary measures when constructing datasets for comparative analyses—this is one of the ways comparative biologists can upset the researchers collecting the data! Whether this causes problems for drawing conclusions from the analyses depends on the aim of one's study and the way it was done. For example, comparative methods have been extensively used to study the evolution of brains and cognitive abilities (Logan et al. 2018). Brain size has been used as a proxy for cognitive ability and intelligence, as these are difficult to quantify across large numbers of species. However, this approach has received criticism as cognition is a



highly complex trait (Healy & Rowe 2006, Logan et al. 2018). Other, more accurate measures of cognitive ability may be more appropriate, such as the density, morphology, and structure of brain neurons (the connectome), which are now being characterized for greater numbers of species (Assaf et al. 2020). The usefulness of comparative studies therefore depends on how hypotheses are converted into comparative tests (e.g., how brains evolve versus how cognition evolves) and whether researchers accurately capture biological processes by what they measure (e.g., brain size versus the connectome).

**7.1.3. Comparative analyses reduce the value of collecting more data.** The success of comparative analyses in summarizing patterns sometimes means they are considered a topic killer. If data have been collated across 5,000 species, what is the point in collecting data on just one more species? An alternative viewpoint is that comparative analyses help show where more data need to be collected and which species may offer the greatest new insights. Following on with the example of brain size, comparative studies have documented which species have relatively large and small brain sizes. If we want to find out more about the causes and consequences of evolving brains of different sizes, these are the species that should be targeted. Without examining phylogenetic patterns of brain size while controlling for correlated factors, such as body size, species with unusually large or small brains are difficult to identify. We can also take inspiration from research on the lignan-degrading capacities of fungi that illustrates how comparative genomics enables experimental tests across past evolutionary time points (Ayuso-Fernández et al. 2019). The utility of comparative analyses in guiding experimental studies is nevertheless underutilized.

## 7.2. Integrating Different Types of Data

It is an exciting time for comparative ecological and evolutionary research. Published information is accumulating at a breakneck speed across a wide range of species, and new types of data are continually becoming available. Extensive data on the geographical distributions and abundances of species, through remote sensing and logging technology, offer exciting possibilities for understanding the biogeography and conservation requirements of species (Debastiani et al. 2021, Mouquet et al. 2012, Whitford & Klimley 2019). This has led to developments in ecological niche modeling (also known as species distribution models) that can now be integrated with phylogenies to gain greater insight into the way traits influence the evolution of species ecologies and vice versa (Alvarado-Serrano & Knowles 2014).

Data on genomes, transcriptomes, proteomes, metabolomes, connectomes, and microbiomes, and other omic data, offer new opportunities to gain a deeper, more mechanistic understanding of the biology of species (Nagy et al. 2020, Smith et al. 2020). The challenges associated with gaining omic data previously restricted comparisons to just a few species. This often led to the  $N = 1$  problem, where only one or a few species were compared, yet multiple independent origins of trait evolution are required to statistically test for associations between traits, which can have important effects on the conclusions that are drawn (e.g., Dewar et al. 2021). However, higher-throughput omic data collection combined with comparative methods is resulting in research of greater scope (Nagy et al. 2020, Smith et al. 2020). For example, phylogenetic comparative analyses of 5,397 bacterial genomes showed that species that inhabit multiple environments (generalists) have expanded genetic machineries for producing extracellular proteins, which have important effects on resource acquisition and species interactions (Garcia et al. 2020). Phylogenetic comparative methods can now even be used to reconstruct ancestral genomes, providing novel insights into how genomes form (Muffato et al. 2023).

As more studies are published and more data become available, synthesizing information becomes an increasing challenge. Reproducibility and transparency in how information is collected

has never been more important, and many useful guides are out there (e.g., Foo et al. 2021). Techniques that integrate phylogenetic information with meta-analytic methods are now well established and provide a powerful quantitative framework for synthesizing research (Adams 2008, Hadfield & Nakagawa 2010, Nakagawa et al. 2023).

## 8. CONCLUSIONS

Anyone can learn comparative methods, and it is a rewarding process. Sometimes the comparative approach is thought of as a quick and easy option, as data are already available. In our experience, this is not the case, and comparative studies can be complicated and arduous. The generality of the insights that emerge makes the effort worthwhile, and some questions can be addressed only by taking a comparative approach. We encourage one to dive in. Maximizing the impact of one's study starts at the very beginning of the process: Any one of the statistical and analytical techniques we discuss in this review is of limited use unless the biological question being addressed is defined clearly and care is taken to avoid mining the ever-growing data mountain for correlations (Freckleton 2009).

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We are very grateful to Stuart West, Philip Downing, and Matilda Brindle for comments on the manuscript and to the Knut and Alice Wallenberg Foundation (Wallenberg Academy fellowship number 2018.0138) for funding.

## LITERATURE CITED

- Adams DC. 2008. Phylogenetic meta-analysis. *Evol. Int. J. Org. Evol.* 62(3):567–72
- Adams DC, Collyer ML. 2019. Phylogenetic comparative methods and the evolution of multivariate phenotypes. *Annu. Rev. Ecol. Evol. Syst.* 50:405–25
- Alvarado-Serrano DF, Knowles LL. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. *Mol. Ecol. Resour.* 14(2):233–48
- Alves JM, Prado-López S, Cameselle-Teijeiro JM, Posada D. 2019. Rapid evolution and biogeographic spread in a colorectal cancer. *Nat. Commun.* 10:5139
- Arenas M. 2022. Methodologies for microbial ancestral sequence reconstruction. In *Environmental Microbial Evolution: Methods and Protocols*, ed. H Luo, pp. 283–303. New York: Springer
- Arnold C, Nunn CL. 2010. Phylogenetic targeting of research effort in evolutionary biology. *Am. Nat.* 176(5):601–12
- Assaf Y, Bouznach A, Zomet O, Marom A, Yovel Y. 2020. Conservation of brain connectivity and wiring across the mammalian class. *Nat. Neurosci.* 23(7):805–8
- Axelrod DI. 1972. Edaphic aridity as a factor in angiosperm evolution. *Am. Nat.* 106(949):311–20
- Ayuso-Fernández I, Rencoret J, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT. 2019. Peroxidase evolution in white-rot fungi follows wood lignin evolution in plants. *PNAS* 116(36):17900–5
- Ayuso-Fernández I, Ruiz-Dueñas FJ, Martínez AT. 2018. Evolutionary convergence in lignin-degrading enzymes. *PNAS* 115(25):6428–33
- Baker RR, Parker GA. 1979. The evolution of bird coloration. *Philos. Trans. R. Soc. B* 287:1018:63–130
- Barua A, Mikheyev AS. 2019. Many options, few solutions: Over 60 My snakes converged on a few optimal venom formulations. *Mol. Biol. Evol.* 36(9):1964–74
- Basu DN, Bhaumik V, Kunte K. 2023. The tempo and mode of character evolution in the assembly of mimetic communities. *PNAS* 120(1):e2203724120

- Baum DA, Smith SD. 2012. *Tree Thinking: An Introduction to Phylogenetic Biology*. Greenwood Village, CO: Roberts
- Bensch HM, O'Connor EA, Cornwallis CK. 2021. Living with relatives offsets the harm caused by pathogens in natural populations. *eLife* 10:e66649
- Bergeron LA, Besenbacher S, Zheng J, Li P, Bertelsen MF, et al. 2023. Evolution of the germline mutation rate across vertebrates. *Nature* 615(7951):285–91
- Boettiger C, Coop G, Ralph P. 2012. Is your phylogeny informative? Measuring the power of comparative methods. *Evol. Int. J. Org. Evol.* 66(7):2240–51
- Bollback JP. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinform.* 7:88
- Boomsma JJ. 2007. Kin selection versus sexual selection: why the ends do not meet. *Curr. Biol.* 17(16):R673–83
- Bowles AMC, Bechtold U, Paps J. 2020. The origin of land plants is rooted in two bursts of genomic novelty. *Curr. Biol.* 30(3):530–36.e2
- Boyko JD, Beaulieu JM. 2021. Generalized hidden Markov models for phylogenetic comparative datasets. *Methods Ecol. Evol.* 12(3):468–78
- Budd GE, Mann RP. 2023. Two notorious nodes: a critical examination of relaxed molecular clock age estimates of the bilaterian animals and placental mammals. *Syst. Biol.* 73(1):223–34
- Bürkner P-C. 2017. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* 80:1–28
- Carvalho P, Diniz-Filho JAF, Bini LM. 2006. Factors influencing changes in trait correlations across species after using phylogenetic independent contrasts. *Evol. Ecol.* 20(6):591–602
- Casemiro FAS, Albert JS, Antonelli A, Menegotto A, Wüest RO, et al. 2023. Landscape dynamics and diversification of the megadiverse South American freshwater fish fauna. *PNAS* 120(2):e2211974120
- Clutton-Brock TH, Harvey PH. 1977. Primate ecology and social organization. *J. Zool.* 183(1):1–39
- Cooney CR, Bright JA, Capp EJR, Chira AM, Hughes EC, et al. 2017. Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature* 542(7641):344–47
- Cornwallis CK, van 't Padje A, Ellers J, Klein M, Jackson R, et al. 2023. Symbioses shape feeding niches and diversification across insects. *Nat. Ecol. Evol.* 7:1022–44
- Cornwallis CK, West SA, Davis KE, Griffin AS. 2010. Promiscuity and the evolutionary transition to complex societies. *Nature* 466(7309):969–72
- Davies NB, Krebs JR. 2009. *An Introduction to Behavioural Ecology*. Hoboken, NJ: John Wiley & Sons. 3rd ed.
- Debastiani VJ, Bastazini VAG, Pillar VD. 2021. Using phylogenetic information to impute missing functional trait values in ecological databases. *Ecol. Inform.* 63:101315
- Dewar AE, Thomas JL, Scott TW, Wild G, Griffin AS, et al. 2021. Plasmids do not consistently stabilize cooperation across bacteria but may promote broad pathogen host-range. *Nat. Ecol. Evol.* 5(12):1624–36
- Dillard JR, Westneat DF. 2016. Disentangling the correlated evolution of monogamy and cooperation. *Trends Ecol. Evol.* 31(7):503–13
- Dismukes W, Braga MP, Hembry DH, Heath TA, Landis MJ. 2022. Cophylogenetic methods to untangle the evolutionary history of ecological interactions. *Annu. Rev. Ecol. Evol. Syst.* 53:275–98
- dos Reis M, Donoghue PCJ, Yang Z. 2016. Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17(2):71–80
- Douglas AE. 2022. *Insects and Their Beneficial Microbes*. Princeton, NJ: Princeton Univ. Press
- Eckerström-Liedholm S, Sowersby W, Morozov S, van der Bijl W, Rowiński PK, et al. 2019. Macroevolutionary evidence suggests trait-dependent coevolution between behavior and life-history. *Evolution* 73(11):2312–23
- Felsenstein J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125(1):1–15
- Felsenstein J. 1988. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* 19:445–71
- Felsenstein J. 2003. *Inferring Phylogenies*. Oxford, UK: Oxford Univ. Press
- Floudas D, Binder M, Riley R, Barry K, Blanchette RA, et al. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–19
- Foo YZ, O'Dea RE, Koricheva J, Nakagawa S, Lagisz M. 2021. A practical guide to question formation, systematic searching and study screening for literature reviews in ecology and evolution. *Methods Ecol. Evol.* 12(9):1705–20

- Freckleton RP. 2009. The seven deadly sins of comparative analysis. *J. Evol. Biol.* 22(7):1367–75
- Frumhoff PC, Reeve HK. 1994. Using phylogenies to test hypotheses of adaptation: a critique of some current proposals. *Evolution* 48(1):172–80
- Gallinat AS, Pearse WD. 2021. Phylogenetic generalized linear mixed modeling presents novel opportunities for eco-evolutionary synthesis. *Oikos* 130(5):669–79
- Garamszegi LZ, ed. 2014. *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*. Berlin: Springer
- García M, Theunissen F, Sèbe F, Clavel J, Ravignani A, et al. 2020. Evolution of communication signals and information during species radiation. *Nat. Commun.* 11(1):4970
- Germain RR, Feng S, Chen G, Graves GR, Tobias JA, et al. 2023. Species-specific traits mediate avian demographic responses under past climate change. *Nat. Ecol. Evol.* 7(6):862–72
- Ghoul M, Andersen SB, Marvig RL, Johansen HK, Jelsbak L, et al. 2023. Long-term evolution of antibiotic tolerance in *Pseudomonas aeruginosa* lung infections. *Evol. Lett.* 7(6):389–400
- Gingerich PD. 2019. *Rates of Evolution: A Quantitative Synthesis*. Cambridge, UK: Cambridge Univ. Press
- Goberna M, Verdú M. 2016. Predicting microbial traits with phylogenies. *ISME J.* 10(4):959–67
- Goolsby EW. 2015. Phylogenetic comparative methods for evaluating the evolutionary history of function-valued traits. *Syst. Biol.* 64(4):568–78
- Grafen A. 1989. The phylogenetic regression. *Philos. Trans. R. Soc. B* 326(1233):119–57
- Hadfield JD, Krasnov BR, Poulin R, Nakagawa S. 2014. A tale of two phylogenies: comparative analyses of ecological interactions. *Am. Nat.* 183(2):174–87
- Hadfield JD, Nakagawa S. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* 23(3):494–508
- Hällfors MH, Antão LH, Itter M, Lehtikoinen A, Lindholm T, et al. 2020. Shifts in timing and duration of breeding for 73 boreal bird species over four decades. *PNAS* 117(31):18557–65
- Hansen TF, Bolstad GH, Tsuboi M. 2022. Analyzing disparity and rates of morphological evolution with model-based phylogenetic comparative methods. *Syst. Biol.* 71(5):1054–72
- Harmon L. 2018. *Phylogenetic Comparative Methods: Learning from Trees*. Scotts Valley, CA: CreateSpace Indep. Publ. Platf.
- Harvey PH, Pagel MD. 1991. *The Comparative Method in Evolutionary Biology*. Oxford, UK: Oxford Univ. Press
- Healy SD, Rowe C. 2006. A critique of comparative studies of brain size. *Proc. R. Soc. B* 274(1609):453–64
- Hembry DH, Weber MG. 2020. Ecological interactions and macroevolution: a new field with old roots. *Annu. Rev. Ecol. Evol. Syst.* 51:215–43
- Hernández CE, Rodríguez-Serrano E, Avaria-Llatureo J, Inostroza-Michael O, Morales-Pallero B, et al. 2013. Using phylogenetic information and the comparative method to evaluate hypotheses in macroecology. *Methods Ecol. Evol.* 4(5):401–15
- Houle D, Bolstad GH, van der Linde K, Hansen TF. 2017. Mutation predicts 40 million years of fly wing evolution. *Nature* 548(7668):447–50
- Housworth EA, Martins EP, Lynch M. 2004. The phylogenetic mixed model. *Am. Nat.* 163(1):84–96
- Ives AR, Garland T Jr. 2014. Phylogenetic regression for binary dependent variables. In *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*, ed. LZ Garamszegi, pp. 231–61. Berlin: Springer
- Jukes T, Cantor CR. 1969. Evolution of protein molecules. In *Mammalian Protein Metabolism*, Vol. 3, ed. HN Munro, pp. 21–132. New York: Academic
- Kahrl AF, Snook RR, Fitzpatrick JL. 2021. Fertilization mode drives sperm length evolution across the animal tree of life. *Nat. Ecol. Evol.* 5(8):1153–64
- King B, Lee MSY. 2015. Ancestral state reconstruction, rate heterogeneity, and the evolution of reptile viviparity. *Syst. Biol.* 64(3):532–44
- Kverková K, Marhounová L, Polonyiová A, Kocourek M, Zhang Y, et al. 2022. The evolution of brain neuron numbers in amniotes. *PNAS* 119(11):e2121624119
- Lichter-Marck IH, Baldwin BG. 2023. Edaphic specialization onto bare, rocky outcrops as a factor in the evolution of desert angiosperms. *PNAS* 120(6):e21214729120
- Logan CJ, Avin S, Boogert N, Buskell A, Cross FR, et al. 2018. Beyond brain size: uncovering the neural correlates of behavioral and cognitive specialization. *Comp. Cogn. Behav. Rev.* 13:55–89

- Losos JB. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.* 11(10):995–1003
- Losos JB. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *Am. Nat.* 177(6):709–27
- Lynch M. 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45(5):1065–80
- Maddison WP. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* 44(3):539–57
- Martins EP. 1994. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.* 144(2):193–209
- Martins EP. 1996. *Phylogenies and the Comparative Method in Animal Behavior*. Oxford, UK: Oxford Univ. Press
- Martins EP, Hansen TF. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* 149(4):646–67
- May MR, Moore BR. 2020. A Bayesian approach for inferring the impact of a discrete character on rates of continuous-character evolution in the presence of background-rate variation. *Syst. Biol.* 69(3):530–44
- Maynard-Smith J, Holliday R. 1979. *The Evolution of Adaptation by Natural Selection: A Royal Society Discussion Meeting*. London, UK: R. Soc.
- McLean MJ, Bishop PJ, Nakagawa S. 2012. Male quality, signal reliability and female choice: assessing the expectations of inter-sexual selection. *J. Evol. Biol.* 25(8):1513–20
- Molina-Venegas R. 2023. How to get the most out of phylogenetic imputation without abusing it. *Methods Ecol. Evol.* 15(3):456–63
- Mouquet N, Devictor V, Meynard CN, Munoz F, Bersier L-F, et al. 2012. Ecophylogenetics: advances and perspectives. *Biol. Rev.* 87(4):769–85
- Muffato M, Louis A, Nguyen NTT, Lucas J, Berthelot C, Roest Crollius H. 2023. Reconstruction of hundreds of reference ancestral genomes across the eukaryotic kingdom. *Nat. Ecol. Evol.* 7(3):355–66
- Nagy LG, Merényi Z, Hegedűs B, Bálint B. 2020. Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. *Nucleic Acids Res.* 48(5):2209–19
- 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. *Syst. Biol.* 68(4):632–41
- Nakagawa S, Johnson PCD, Schielzeth H. 2017. The coefficient of determination  $R^2$  and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* 14(134):20170213
- Nakagawa S, Lagisz M, Yang Y, Drobniak SM. 2024. Finding the right power balance: Better study design and collaboration can reduce dependence on statistical power. *PLoS Biol.* 22(1):e3002423
- Nakagawa S, Yang Y, Macartney EL, Spake R, Lagisz M. 2023. Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environ. Evid.* 12(1):8
- Nunn CL. 2011. *The Comparative Approach in Evolutionary Anthropology and Biology*. Chicago: Univ. Chicago Press
- O'Meara BC. 2012. Evolutionary inferences from phylogenies: a review of methods. *Annu. Rev. Ecol. Evol. Syst.* 43:267–85
- Opedal ØH, Armbruster WS, Hansen TF, Holstad A, Pélabon C, et al. 2023. Evolvability and trait function predict phenotypic divergence of plant populations. *PNAS* 120(1):e2203228120
- Pagel M. 1992. A method for the analysis of comparative data. *J. Theor. Biol.* 156(4):431–42
- Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. Biol. Sci.* 255(1342):37–45
- Pagel M. 1997. Inferring evolutionary processes from phylogenies. *Zool. Scr.* 26(4):331–48
- Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401(6756):877–84
- Pagel M, Meade A. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* 167(6):808–25
- Pankey M, Plachetzki DC, Macartney KJ, Gastaldi M, Slattery M, et al. 2022. Cophylogeny and convergence shape holobiont evolution in sponge–microbe symbioses. *Nat. Ecol. Evol.* 6(6):750–62
- Pinseel E, Janssens SB, Verleyen E, Vanormelingen P, Kohler TJ, et al. 2020. Global radiation in a rare biosphere soil diatom. *Nat. Commun.* 11:2382



- Powell S, Price SL, Kronauer DJC. 2020. Trait evolution is reversible, repeatable, and decoupled in the soldier caste of turtle ants. *PNAS* 117(12):6608–15
- Pyron RA, Burbrink FT. 2014. Early origin of viviparity and multiple reversions to oviparity in squamate reptiles. *Ecol. Lett.* 17(1):13–21
- Revell LJ, Harmon LJ. 2022. *Phylogenetic Comparative Methods in R*. Princeton, NJ: Princeton Univ. Press
- Ridley M. 1983. *The Explanation of Organic Diversity: The Comparative Method and Adaptations for Mating*. Oxford, UK: Clarendon
- Ronquist F, Sanmartín I. 2011. Phylogenetic methods in biogeography. *Annu. Rev. Ecol. Evol. Syst.* 42:441–64
- Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364(6435):74–78
- Sanchez-Martinez P, Martínez-Vilalta J, Dexter KG, Segovia RA, Mencuccini M. 2020. Adaptation and coordinated evolution of plant hydraulic traits. *Ecol. Lett.* 23(11):1599–610
- Santini L, González-Suárez M, Russo D, Gonzalez-Voyer A, von Hardenberg A, Ancillotto L. 2019. One strategy does not fit all: determinants of urban adaptation in mammals. *Ecol. Lett.* 22(2):365–76
- Sheehan O, Watts J, Gray RD, Bulbulia J, Claessens S, et al. 2023. Coevolution of religious and political authority in Austronesian societies. *Nat. Hum. Behav.* 7(1):38–45
- Simonet C, McNally L. 2021. Kin selection explains the evolution of cooperation in the gut microbiota. *PNAS* 118(6):e2016046118
- Smith SD, Pennell MW, Dunn CW, Edwards SV. 2020. Phylogenetics is the new genetics (for most of biodiversity). *Trends Ecol. Evol.* 35(5):415–25
- Sorensen D, Gianola D. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*. New York: Springer
- Sukumaran J, Knowles LL. 2018. Trait-dependent biogeography: (re)integrating biology into probabilistic historical biogeographical models. *Trends Ecol. Evol.* 33(6):390–98
- Szulkun M, Munshi-South J, Charmanier A, eds. 2020. *Urban Evolutionary Biology*. Oxford, UK: Oxford Univ. Press
- Thorson JT, Maureaud AA, Frelat R, Mérigot B, Bigman JS, et al. 2023. Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. *Methods Ecol. Evol.* 14(5):1259–75
- Tobias JA, Ottenburghs J, Pigot AL. 2020. Avian diversity: speciation, macroevolution, and ecological function. *Annu. Rev. Ecol. Evol. Syst.* 51:533–60
- Trepel J, le Roux E, Abraham AJ, Buitenwerf R, Kamp J, et al. 2024. Meta-analysis shows that wild large herbivores shape ecosystem properties and promote spatial heterogeneity. *Nat. Ecol. Evol.* 8(4):705–16
- Uyeda JC, Zenil-Ferguson R, Pennell MW. 2018. Rethinking phylogenetic comparative methods. *Syst. Biol.* 67(6):1091–109
- Vamosi JC, Armbruster WS, Renner SS. 2014. Evolutionary ecology of specialization: insights from phylogenetic analysis. *Proc. R. Soc. B* 281(1795):20142004
- von Hardenberg A, Gonzalez-Voyer A. 2013. Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis: phylogenetic path analysis. *Evolution* 67(2):378–87
- Walsh B, Lynch M. 2018. *Evolution and Selection of Quantitative Traits*. Oxford, UK: Oxford Univ. Press
- Watts J, Sheehan O, Atkinson QD, Bulbulia J, Gray RD. 2016. Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature* 532(7598):228–31
- Webb CO, Ackerly DD, McPeck MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33:475–505
- Wells MT, Barker FK. 2017. Big groups attract bad eggs: Brood parasitism correlates with but does not cause cooperative breeding. *Anim. Behav.* 133:47–56
- Westoby M, Yates L, Holland B, Halliwell B. 2023. Phylogenetically conservative trait correlation: quantification and interpretation. *J. Ecol.* 111(10):2105–17
- Whitford M, Klimley AP. 2019. An overview of behavioral, physiological, and environmental sensors used in animal biotelemetry and biologging studies. *Anim. Biotelemetry* 7(1):26
- Winkelman M. 2014. Political and demographic-ecological determinants of institutionalised human sacrifice. *Anthropol. Forum.* 24(1):47–70
- Wright AM, Bapst DW, Barido-Sottani J, Warnock RCM. 2022. Integrating fossil observations into phylogenetics using the fossilized birth–death model. *Annu. Rev. Ecol. Evol. Syst.* 53:251–73