CHAPTER 10

PHYLOGENETIC APPROACHES FOR RESEARCH IN COMPARATIVE COGNITION

Evan L. MacLean and Charles L. Nunn

Comparative studies have the potential to address a wide range of questions about how, when, and why different traits have evolved. The comparative approach refers to studies in which variation in traits of different species (or different populations) is used to test specific hypotheses or to generate new hypotheses about evolutionary phenomena. In many cases, comparative methods are used to investigate how two or more traits covary. Comparative methods are also used to reconstruct evolutionary history or to assess how traits influence patterns of diversification (speciation and extinction; Nunn, 2011). In the context of comparative psychology, these methods have been used, for example, to investigate associations between aspects of neuroanatomy and executive function (MacLean et al., 2014; Shultz & Dunbar, 2010; see also Chapter 24, this volume), to explore how life history traits influence temporal decision making (Stevens, 2014; see also Volume 2, Chapter 24, this handbook), and to make inferences about how life in complex societies has shaped primate cognitive evolution (Amici, Aureli, & Call, 2008; Burkart et al., 2014; MacLean et al., 2013; see also Chapters 12 and 13, this volume).

Understanding evolutionary relationships—or phylogeny—among the species of interest is critical for effective inference in comparative research. In a very real sense, phylogeny is the scaffolding on which one investigates evolutionary change in traits and the factors that lead to these changes. This is important conceptually, and it is also critically relevant in statistical analyses. For example, in the context of studying how traits covary, data on different species cannot be treated as independent observations in statistical analyses because of patterns of inheritance on a bifurcating phylogeny (Felsenstein, 1985; Martins & Garland, 1991). To deal with this nonindependence, comparative biologists incorporate information about the evolutionary relationships between species. Phylogenetic comparative methods are a set of statistical approaches designed for exactly this purpose (Garamszegi, 2014; Garland, Bennett, & Rezende, 2005; Nunn, 2011; Rezende & Diniz-Filho, 2012), a toolkit that is becoming increasingly useful in the field of comparative psychology. In addition to being applied in a broader range of evolutionary contexts, comparative methods themselves are evolving, and many classical techniques are rapidly being replaced by newer and more flexible approaches.

In this chapter, we introduce the reader to a range of phylogenetic comparative methods that can be used to address fundamental questions in comparative psychology, including some recent methodological advances that will create powerful opportunities for future research. We illustrate these methods by using a combination of simulated data and analyses of published datasets. Given the rapid growth of research on comparative cognition, we highlight the utility of comparative methods for the study of cognitive evolution. However, the concepts and statistical approaches described in this chapter can be similarly applied in other areas of comparative psychology.

At the outset, we want to emphasize that some traits are measured on a quasi-continuous scale (e.g., percentage of correct responses on an

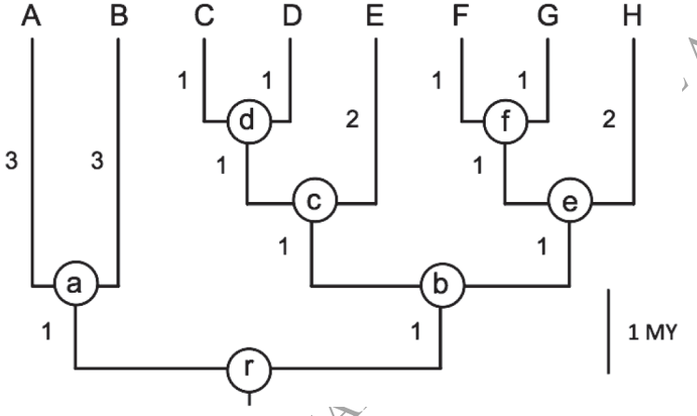
http://dx.doi.org/10.1037/XXXXX-XXX
APA Handbook of Comparative Psychology: Vol. 1. Basic Concepts, Methods, Neural Substrate, and Behavior, J. Call (Editor-in-Chief)

experimental task), and other traits are measured on a discrete scale (e.g., presence or absence of mirror self-recognition). These two types of data often require different statistical approaches. For brevity, we present a single variant of each method below, yet we also direct the reader to relevant literature on other approaches throughout and to several recent reviews (Garamszegi, 2014; Garland et al., 2005; Nunn, 2011; Rezende & Diniz-Filho, 2012). We begin with a brief introduction to additional key terms and concepts that are foundational to all phylogenetic comparative methods.

PHYLOGENETIC TREES

As just noted, phylogenies represent the evolutionary relationships between taxa and are frequently visualized as trees with a branching pattern (see Figure 10.1A). Phylogenetic trees consist of nodes and branches. Nodes indicate speciation events where an

(A)



(B)

	A	B	C	D	E	F	G	H
A	**4**	1	0	0	0	0	0	0
B	1	**4**	0	0	0	0	0	0
C	0	0	**4**	3	2	1	1	1
D	0	0	3	**4**	2	1	1	1
E	0	0	2	2	**4**	1	1	1
F	0	0	1	1	1	**4**	3	2
G	0	0	1	1	1	3	**4**	2
H	0	0	1	1	1	2	2	**4**

FIGURE 10.1. A phylogenetic tree and variance–covariance matrix representing the evolutionary relationships between species. A: The root of the tree is at the bottom with the tips (terminal branches) extending upward. Uppercase letters refer to extent taxa at the tips of the tree. Nodes are indicated by lowercase letters enclosed in circles, and branch lengths are shown to the left of each branch in the tree. The scale bar indicates 1 million years (MY). B: The variance–covariance matrix represents the total age of the phylogeny along the diagonal, with the extent of shared evolutionary history between pairs of species indicated on the off-diagonals.

ancestral lineage gave rise to two (or more) descendent species. Nodes are connected by branches, which are typically drawn to be proportional to evolutionary time. Figure 10.1A shows a phylogenetic tree for eight species with circles enclosing the internal nodes. Node r is located at the root of the phylogeny and represents the oldest bifurcation in the tree, which in this example occurred 4 million years ago. The internal branches of the tree (i.e., branches connecting nodes, rather than ending at a tip) represent the time that species have shared evolutionary history, whereas the terminal branches (leading to a tip) represent the time that each extant species has evolved independently of other taxa in the tree (of course keeping in mind that there are often many extinct lineages that are not represented on a phylogeny of extant species; see Nunn, 2011).

The extent of shared evolutionary history between pairs of species in a phylogeny can be represented as a variance–covariance matrix (see Figure 10.1B; Cunningham, Omland, & Oakley, 1998; Freckleton, Harvey, & Pagel, 2002). The diagonal of the matrix (gray background) represents the variances, or the total time from the root to the tips of the tree (4 million years). The off-diagonals of the matrix represent the covariances, or the amount of time that pairs of species have shared evolutionary history since the root of the tree. In this example, the covariance between species C and D is 3 million years, reflecting the time between nodes r and d, whereas the covariance between species C and H is 1, reflecting the time between nodes r and b. Variance–covariance matrices play an important role in many phylogenetic comparative methods; hence, we revisit this concept throughout the chapter.

A first step in most comparative analyses is to obtain a phylogeny for the species of interest. Fortunately, digital phylogenies are widely available and can be downloaded from sites such as 10ktrees (http://10ktrees.fas.harvard.edu; Arnold, Matthews, & Nunn, 2010), TreeBase (http://treebase.org; Sanderson, Donoghue, Piel, & Eriksson, 1994), and the Open Tree of Life (http://blog.opentreeoflife.org). Garamszegi and Gonzalez-Voyer (2014) provided an overview of the steps for obtaining a phylogeny for comparative research.

PHYLOGENETIC SIGNAL

As a result of shared evolutionary history, closely related species tend to resemble one another more so than less closely related taxa; this tendency is termed *phylogenetic signal* (Blomberg & Garland, 2002; Blomberg, Garland, & Ives, 2003). Interestingly, the extent to which traits are associated with phylogeny varies widely from one trait to another, with morphological traits tending to exhibit higher levels of phylogenetic signal than behavioral, cognitive, or ecological variables (Blomberg et al., 2003; Kamilar & Cooper, 2013; MacLean et al., 2012). Quantitative estimates of phylogenetic signal in continuous traits can be obtained by using a variety of different approaches (for reviews, see Kamilar & Cooper, 2013; Münkemüller et al., 2012; Nunn, 2011). Here we focus on one commonly used metric, Pagel's lambda (Freckleton et al., 2002; Pagel, 1999a).

Lambda is a continuous parameter that ranges from 0 to 1, with a value of 0 indicating that trait covariances are independent of phylogeny and a value of 1 indicating that variation among species approximates expectations from a Brownian motion model of evolution (i.e., a random walk in which trait variance accumulates proportionally to evolutionary time). The lambda parameter scales the internal branches of a phylogeny by multiplying the internal branch lengths (the off-diagonals in the variance–covariance matrix) by lambda while retaining the original variances along the diagonal. Thus, when $\lambda = 0$, the internal structure of the tree is entirely eliminated (all internal branch lengths are 0, yielding a "star phylogeny"; see Felsenstein, 1985). In contrast, the internal branch lengths (off-diagonals in the variance–covariance matrix) remain unchanged when $\lambda = 1$. Lambda may take any value between 0 and 1 (and even slightly higher than 1, subject to constraints determined by characteristics of the tree). The lambda value for a given trait is typically estimated using maximum likelihood to find the lambda transformation that makes covariances between species most likely under a Brownian motion model of evolution.

To illustrate the concept of lambda and phylogenetic signal, Figure 10.2 shows the values of two

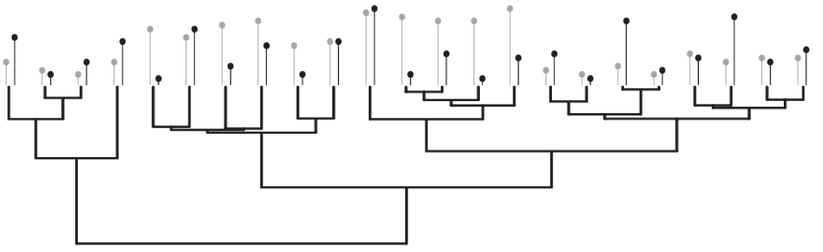


FIGURE 10.2. Two traits (indicated in gray and black) with differing levels of phylogenetic signal plotted on a phylogeny. The vertical positions of the points are proportional to the trait values. The gray trait has strong phylogenetic signal (closely related species have more similar trait values), and the black trait has weak phylogenetic signal (trait variance is independent of phylogeny).

continuous traits that were simulated on a phylogeny to have different levels of phylogenetic signal. The maximum likelihood estimate of lambda for the gray trait is 0.99, indicating that closely related species have highly similar trait values (notice, e.g., that sister species have gray bars of approximately the same magnitude). In contrast, the maximum likelihood estimate of lambda for the black trait is ~0, indicating that phylogeny is a weak predictor of species-level variance (hence, one sees no pattern of sister species having black bars of more similar magnitude, compared with other species on the tree). More important, these branch-length transformations are not a revised estimate of the actual species divergence times (the best estimate of which was known before analysis). Instead, the transformation estimates the extent to which phenotypic covariance matches the expected covariance on the basis of phylogeny and a Brownian motion model of evolution. For tests of phylogenetic signal in discrete traits, we direct the reader to Maddison and Slatkin (1991), Abouheif (1999), and Fritz and Purvis (2010).

Assessing phylogenetic signal is an important first step in comparative analyses because it provides information about the extent of nonindependence in the data. To date, only a few studies have investigated phylogenetic signal in measures of cognition across species. Yet the initial findings have suggested that several cognitive measures are characterized by substantial phylogenetic signal relative to behavioral or ecological traits. For example, in a comparative analysis of primate data from an intertemporal choice task, Stevens found that the maximum likelihood estimate of lambda for indifference points was ~1 (see Volume 2, Chapter 24, this handbook). Similarly, in tests of self-control administered to 36 species, including primates, carnivores, rodents, and birds, MacLean et al. (2014) reported lambda values of 0.76 across the entire sample of species (for the average score across cognitive tasks) and 0.86 within primates. Thus, as with many other traits, phenotypic variation in cognition is not independent of phylogeny.

One common misconception about testing for phylogenetic signal in individual traits is that the results of these analyses determine whether phylogenetic approaches are needed for subsequent analyses. For example, if none of the individual traits being studied display phylogenetic signal, researchers may be tempted to use these findings as a justification to forego additional phylogenetic analyses. However, as we show in the next section, it is phylogenetic signal not in individual traits but in the error variance of a statistical model that is relevant in comparative analyses of correlated evolution in two or more traits. Of course, a lack of phylogenetic signal should also lead one to be cautious about applying phylogenetic methods to reconstruct ancestral nodes or to assess how traits influence speciation and extinction rates across the tree.

"HOW" AND "WHY" QUESTIONS: TESTING ADAPTIVE HYPOTHESES

Many of the most interesting questions about cognitive evolution concern when, how, and why particular aspects of cognition evolved. "When" questions

often involve assessing whether a particular character state is ancestral in a clade or evolved independently. These questions embrace Tinbergen's (1963) question about evolutionary history. "How" questions are typically questions about the proximate mechanisms (Tinbergen, 1963) underlying phenotypic variance. For example, do differences in life history strategies, gene expression in the brain (see Chapter 22, this volume), or the volume of brain regions relate to species differences in cognition (see Chapters 12 and 24, this volume)? "Why" questions are typically questions about the selective pressures that have favored cognitive evolution. For example, does living in complex societies, relying on spatiotemporally distributed resources, or facing high levels of niche competition favor particular cognitive adaptations?

One of the most powerful approaches for investigating how and why questions involves tests of correlated evolution. In brief, the rationale for such tests is that correlations between traits across phylogeny suggest that these traits may be functionally linked (Harvey & Purvis, 1991; Nunn, 2011; Nunn & Barton, 2001). However, as we discovered in our review of phylogenetic signal, the nonindependence of species-level data is an important consideration when testing such hypotheses because many traits are correlated at the species level through inheritance from a common ancestor, not independent evolutionary origins.

General linear models provide a unified statistical framework for predicting a continuous dependent measure as a function of an intercept and a linear combination of predictor variables weighted by regression coefficients (see Chapter 8, this volume). General linear models assume that errors (model residuals) are normally and independently distributed with common variance (homoscedasticity). Comparative data frequently violate this assumption because the residuals of closely related species tend to be more similar than those of more distantly related species, and thus errors are not independently distributed (i.e., they exhibit phylogenetic signal). More important, this scenario can arise even in cases when neither the predictor nor the response variable has high levels of phylogenetic signal, and thus the absence of phylogenetic signal in X or Y alone is insufficient justification for ignoring phylogeny in regression models using these variables (Revell, 2010).

Generalized least squares (GLS) is an extension of the general linear model that can accommodate nonindependence in the data, as well as non-Gaussian error distributions (see Chapter 8, this volume). Phylogenetic generalized least squares (PGLS) is a regression model that is based on the GLS framework (Grafen, 1989; Martins & Hansen, 1997; Pagel, 1999a). It is similar to GLS in that it generates regression coefficients, confidence intervals, and significance estimates for one or more predictors (which can be continuous or discrete) of a continuous response variable. PGLS accommodates the nonindependence of species-level data by incorporating the variance–covariance matrix (described above) into the error term of the model to account for covariance in model residuals given the phylogeny and the extent of phylogenetic signal in the data (Symonds & Blomberg, 2014).

The flexibility of PGLS lies in its ability to adjust for statistical nonindependence on the basis of the actual error structure of the data (based on estimating lambda or other scaling parameters). This is a critical difference between PGLS and independent contrasts, an earlier approach that is used to obtain statistically independent comparisons within a phylogeny (for a review, see Felsenstein, 1985; Nunn & Barton, 2001). Specifically, independent contrasts are calculated by generating differences—or contrasts—between pairs of lineages throughout the tree. The method assumes that phylogeny accurately predicts the model's error structure and thus standardizes contrasts by dividing them by the sum of their branch lengths to control for heteroskedasticity. An alternative is to estimate the extent of phylogenetic signal in the data and adjust the model accordingly (Symonds & Blomberg, 2014). To do so, PGLS is often combined with estimating branch-length transformations (such as Pagel's lambda) that accommodate diverse evolutionary models and different levels of phylogenetic nonindependence. As in the case for estimating phylogenetic signal in a single trait, maximum likelihood can be used to find the branch-length transformation that optimizes the error structure of the residuals in a regression model

(Revell, 2010). When lambda is estimated to be 1, the results of the PGLS model will be identical to those from independent contrasts, and when lambda is estimated to be 0, PGLS results will mirror those from a general linear model (or GLS with a Gaussian distribution) that does not include phylogeny. However, in many cases the maximum likelihood estimate of lambda will lie intermediate to these two extremes.

A common misconception about phylogenetic models is that they are more conservative than nonphylogenetic approaches and thus guard only against Type I statistical errors (see Chapter 8, this volume). However, by meeting the assumptions of the underlying statistical model, phylogenetic approaches not only reduce false positives by correcting for pseudoreplication but also have increased power to detect correlated evolution in cases in which a general linear model does not do so (Type II errors). Both of these scenarios are illustrated in hypothetical examples exploring the relationship between an ecological predictor and a cognitive dependent measure in Figure 10.3. Figure 10.3A depicts the classic problem of pseudoreplication (see Felsenstein, 1985). Data from three clades (monophyletic groups) are shown, with clade membership indicated by the shading of points. Within each clade, no significant association exists between X and Y, as shown in the clade-specific dashed regression lines. Between clades, however, grade shifts occur in both X and Y. When treating species as independent data points (i.e., when lambda is forced to equal 0), GLS strongly overestimates the relationship between X and Y (gray regression line; $R^2 = .74$, $p < .001$). In contrast, PGLS accounts for the nonindependence in the model and correctly fails to reject the null hypothesis ($\lambda = 0.72$, $R^2 = .05$, $p = .07$).

Figure 10.3B shows a different situation in which PGLS again leads one to a correct interpretation of scaling relationships. In this case, there is a grade shift in X but not in Y, which when analyzed by GLS (gray regression line) obscures the strong relationships between X and Y within each of the three clades (dashed regression lines). In this case, GLS fails to detect the correlation between X and Y ($R^2 = .05$, $p = .10$), whereas PGLS has increased power to do so ($\lambda = 0.84$, $R^2 = .60$, $p < .001$).

As noted previously, PGLS can accommodate both continuous and discrete predictors of a continuous response variable, but it is inappropriate to use PGLS for a discrete response variable (i.e., logistic regression). For a review of methods appropriate for categorical response variables, we direct the reader to Ives and Garland (2014).

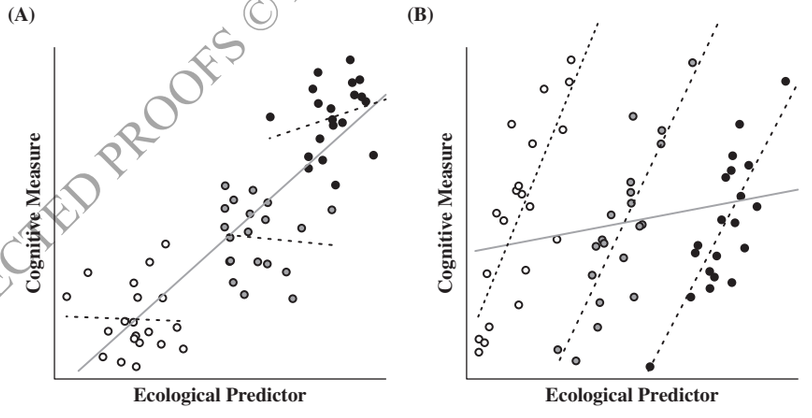


FIGURE 10.3. Examples of scenarios in which phylogenetic generalized least squares avoids Type I (A) and Type II (B) statistical errors. Clade membership is represented by the color of points. The solid gray regression line is from a nonphylogenetic model that fails to account for the nonindependence of species-level data. The dashed black lines show the clade-specific associations between the ecological predictor and cognitive dependent measure. See text for details.

PGLS has been incorporated in several recent studies of comparative cognition. For example, using PGLS with a sample of 13 primate species, Stevens (2014) identified links between a range of allometric variables, including body mass, brain volume, life span, and home range size, and species differences in a delayed reward task. In another study, MacLean et al. (2014) found that absolute brain volume and dietary breadth, but not frugivory or species-typical social group size, predicted differences on two self-control tasks among 36 species of mammals and birds. Interestingly, when a smaller subset of species was tested on both a social cognitive measure and one of these self-control tasks, MacLean et al. (2013) found that social group size was a significant predictor of species differences on the social but not the self-control task. Last, in a recent comparative study of 15 primate species (including humans), Burkart et al. (2014) found that extensive allomaternal care was the best predictor of proactive prosociality, implicating cooperative breeding as a possible selective pressure for human hypercooperation (see Chapter 13, this volume). We return to this example below to assess whether humans show substantially more prosociality than other primates given their level of allomaternal care (see Phylogenetic Prediction section).

Thus, PGLS has allowed researchers to test a wide range of hypotheses about cognitive evolution. Unlike studies using proxy variables for cognition, such as brain size, a combination of experimental and phylogenetic methods has led to unique insights about the selective pressures that have favored cognitive evolution in specific cognitive domains. Collectively, these types of studies highlight an exciting new direction for research in comparative cognition, one that has potential to lead to major insights about long-standing questions regarding how and why cognition evolves.

"WHEN" QUESTIONS: RECONSTRUCTING THE COGNITIVE PAST

Comparative psychologists study the behavior and cognition of extant species, yet many questions in comparative psychology involve inferences about the past. For example, did the last common ancestor of humans and great apes have a theory of mind? When, and how many times, has social learning evolved? Do vertebrate species share a primitive number sense?

The answers to these types of questions rely on the ability to make inferences about the past using information from the present and thus capture the when questions identified above. Several comparative methods are designed for ancestral state reconstructions. Here, we focus on maximum likelihood approaches for inferences about discrete traits, although similar methods can be used with continuous variables (for review, see MacLean et al., 2012). Maximum likelihood methods differ from parsimony reconstructions in a number of important ways (Pagel, 1999b). Reconstructions based on parsimony attempt to minimize the number of transitions on the tree, and they ignore information about branch lengths (i.e., changes are assumed to be equally likely to occur along the shortest and the longest branch on the tree). Maximum parsimony methods also do not provide easily interpretable statistical support measures for reconstructed states (Nunn, 2011). More important, parsimony lacks an underlying evolutionary model and is being driven to extinction by model-based approaches, such as those based on maximum likelihood. By using a specific evolutionary model, maximum likelihood approaches integrate information about branch lengths (the amount of time that lineages have evolved independently of other taxa in the tree) and rates of evolutionary change, and they can provide estimates of statistical support for each reconstructed node in the tree. These models can also accommodate specific types of trait evolution, such as asymmetric transition rates in which the probability of moving from state A to state B differs from that of moving from state B to state A.

To illustrate a maximum likelihood reconstruction using discrete variables, we present an example using data on social learning, extractive foraging, and tool use taken from Reader, Jager, and Laland (2011). This dataset includes counts of published reports for each of these categories in a large sample of primates, along with the total number of citations for each species at the time data were compiled. To create discrete traits based on these data, we

first restricted the dataset to species with at least 20 citations in the Zoological Record database, effectively limiting the data to a subset of reasonably well-studied species. Species-level data were then collapsed to a genus-level summary (44 genera) to allow a concise example for illustrative purposes. Finally, within each genus we created three discrete variables reflecting whether social learning, extractive foraging, and tool use had been observed in any member of the genus (each of these coded as yes [1] or no [0]). Ancestral states were then estimated using maximum likelihood in a model with equal transition rates between the character states for each variable, meaning that the rates of state yes transitioning to state no were equivalent to rates of state no transitioning to state yes.

The results of this analysis are plotted on the genus-level tree in Figure 10.4. The tips of the tree depict the

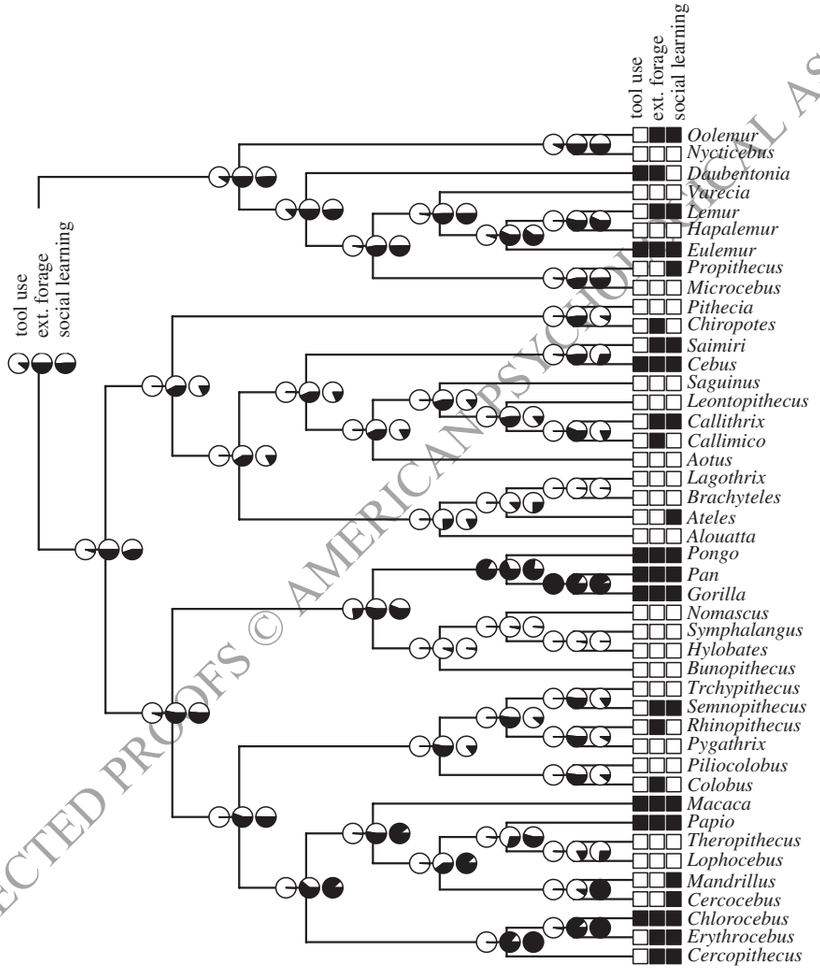


FIGURE 10.4. Ancestral state reconstruction of tool use, extractive foraging, and social learning modeled as discrete traits in primate genera. Squares at the tips of the phylogeny represent the character states for extant species (shaded = trait present, unshaded = trait absent). The circles at the internal nodes represent the scaled likelihood that a trait was present at that node (proportion shaded = likelihood a trait was present). The leftmost circles and squares correspond to tool use. The center circles and squares correspond to extractive foraging. The rightmost circles and squares correspond to social learning. See text for details.

character states for extant genera using shaded (trait present) or unshaded (trait absent) squares (left, tool use; center, extractive foraging; right, social learning). The circles at the internal nodes of the phylogeny indicate the scaled likelihood (range = 0–1) that each of the three traits was present at that node.

This analysis indicates that the last common ancestor of the extant great apes used tools (see Volume 2, Chapter 30, this handbook), practiced extractive foraging, and learned socially (see Volume 2, Chapter 19, this handbook), effectively answering a when question about these behaviors. At the root of the phylogeny, much more uncertainty exists about the character states of all three variables; this is a common circumstance, which often requires that additional outgroups be added to the tree to better resolve the state of the character at the deeper nodes. Visualizing the trends across the phylogeny, one sees that tool use is estimated to be a recently evolved trait (compared with social learning and extractive foraging), and a trait that appears only after both social learning and extractive foraging are well established. Thus, one plausible explanation is that tool use is adaptive primarily in species already feeding on difficult-to-access or embedded foods and that social learning may be required for tool use to become widespread (i.e., it is acquired through social learning). This hypothesis could be evaluated using a test for correlated evolution of discrete traits, in which the state of one variable influences the probability of change in another (Pagel, 1994). However, a recent article urged caution in using this and some other methods of correlated evolution for discrete characters (Maddison & FitzJohn, 2015).

Historically, estimating ancestral states has not been a common approach in comparative psychology. However, we expect that with growing comparative databases on animal cognition, these types of analysis will become increasingly common. For empirical examples of ancestral state estimation for continuous cognitive traits, we direct the reader to MacLean et al. (2014) and Reader et al. (2011).

PHYLOGENETIC PREDICTION

Evolutionary singularities—or uniquely derived traits—present a challenge for comparative methods that rely on convergence to infer adaptation. However, comparative methods can be used to determine whether a particular species deviates from what one expects given the phylogeny and ~~given~~ the relationships between a set of predictor and response variables measured in other taxa (Nunn & Zhu, 2014; Organ, Nunn, Machanda, & Wrangham, 2011). For example, human brains may have more neurons than any other primate species, but this trait may not be extraordinary given knowledge of human brain volumes and the general cellular scaling rules of primate brains (Herculano-Houzel, 2009, 2012).

Many long-standing questions in comparative psychology relate to whether aspects of cognition are qualitatively different in humans and nonhuman animals (Penn, Holyoak, & Povinelli, 2008; Shettleworth, 2012). Often, these questions are investigated by comparing the performance of humans and nonhuman animals on a set of cognitive measures with the aim of identifying areas of similarity and difference (e.g., Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007; Inoue & Matsuzawa, 2007). However, even in cases in which humans do differ substantially from other species, these differences may be predictable within a phylogenetic framework and, in this sense, consistent with broader evolutionary patterns. Methods for phylogenetic prediction allow statistical inferences about the extent to which an apparent outlier is truly remarkable given phylogeny and a set of relevant predictor variables.

As an example of phylogenetic prediction, we consider whether humans are unique in their prosocial tendencies (see Chapters 13 and 44, this volume). Our analysis here builds on recent research by Burkart et al. (2014), who focused on accounting for variation in data from prosociality experiments in 15 primates species. Across species, the extent of allomaternal care was the best predictor of proactive prosociality. Humans exhibited very high levels of proactive prosociality, and they deviated from the nonhuman primate regression line by less than 1 standard deviation, although this latter analysis does not account for phylogeny in assessing whether this residual is extreme. Here, we use more powerful methods to assess whether humans are different from other primates, accounting for variation in

allomaternal care and phylogenetic placement of humans. We also use this opportunity to introduce Bayesian methods, which are becoming increasingly common in phylogenetic methodology.

A number of approaches are available for assessing whether a species (or set of species) exhibits an exceptionally large (or small) state of a continuously varying trait, such as body mass, height, or even proactive prosocial tendencies. Focusing on humans, we could investigate whether the lineage leading to humans shows a higher rate of evolution in prosocial behavior compared with other lineages on the phylogeny. This can be achieved using independent contrasts (McPeek, 1995) or by fitting models with two different rates for different lineages and assessing whether that two-rate model improves on a model that fits a single rate (O'Meara, Ané, Sanderson, & Wainwright, 2006; Revell, 2008). Other, more sophisticated variants on these methods are now available (Revell, Mahler, Peres-Neto, & Redelings, 2012). One could also examine residuals from a regression of prosocial tendencies on another trait, such as allomaternal care, with the expectation that humans would have an exceptionally high positive residual (i.e., they are more prosocial than one might expect given the degree of allomaternal care observed in humans). To incorporate phylogeny, this latter approach can be implemented with independent contrasts or PGLS (Garland & Ives, 2000).

Here, we use a new approach to phylogenetic prediction that involves one or more predictor variables in a PGLS framework but runs the analysis using a Bayesian approach implemented with Markov chain Monte Carlo, or MCMC (Nunn & Zhu, 2014). Briefly, at each iteration, the procedure evaluates a new candidate set of parameter values that are accepted (replacing the current values in the chain) or rejected (maintaining the current values in the chain). In general, the algorithm is designed to select values that make the data more likely, but less likely solutions may be accepted probabilistically in relation to the likelihood (i.e., parameter sets that make the data less likely are less likely to be accepted). A record of selected parameter values is recorded at predefined intervals (e.g., every 100 iterations). The frequency with which particular values are sampled ultimately approximates a probability distribution. Thus, in addition to generating a point estimate of particular parameters (e.g., the mean value across saved iterations), uncertainty about these values is reflected in the distribution of values that are recorded ("sampled"). In a Bayesian framework, this is referred to as a posterior (as opposed to a prior) probability distribution.

Using MCMC in a phylogenetic regression, one can sample the posterior probability distribution of regression coefficients, the intercept, and lambda that are central to prediction in the PGLS model while also sampling from across a set of trees that reflect phylogenetic uncertainty (Nunn, 2011; Pagel & Lutzoni, 2002). The analysis takes into account all these sources of uncertainty in coefficients, phylogeny, and scaling parameters; the resulting set of parameters is then used to generate a posterior probability prediction of, for example, the state of prosocial behavior in humans, predicted on the basis of human levels of allomaternal care and humans' relationship to other primates. This posterior probability distribution is a natural way to incorporate uncertainty in our prediction. If the true human value of prosocial behavior falls outside the 95% credible interval of the posterior probability distribution for predicted prosocial behavior, we would say that human prosocial behavior is exceptional relative to other primates, meaning that humans are an evolutionary outlier (Nunn, 2011; Nunn & Zhu, 2014; Organ et al., 2011).

We used the program BayesModelS to run the MCMC analysis (Nunn & Zhu, 2014). The program takes input data on prosocial behavior and allomaternal care, which we obtained from Burkart et al. (2014), and a posterior probability distribution of trees, which we obtained from Version 3 of 10kTrees (Arnold et al., 2010). The user must also decide how large of a posterior probability sample to obtain (i.e., how many samples to save) and how often to sample from the chain of parameters that is produced. Sampling can too often lead to a high correlation among the estimates that are used in the posterior probability distribution, whereas sampling too little can result in a longer run time. The user must also ensure that the sample has reached a steady state of estimates (i.e., that it is a post–"burn-in" set of parameters during which

likelihoods increase rapidly across iterations as the model parameters are initially tuned to the data). On the basis of initial runs of the model, we decided to sample 8,000 estimates of the parameters, with 50 generations separating each saved set of parameters and an initial burn-in of 100 generations (in which no samples were retained).

We also estimated both lambda and kappa, where the latter parameter scales branches by raising them to the power kappa. We allowed BayesModelS to use an MCMC procedure to decide whether to estimate lambda or kappa, with the proportion of samples estimating one or the other giving a measure of relative support for one transformation over the other. Likewise, we allowed the model to estimate the regression coefficient for the effect of allomaternal care on prosocial behavior or to set it to zero; the proportion of time that the coefficient was included in the model and estimated, along with the credible interval on the estimate, gives a sense of how strongly prosocial behavior covaries positively with allomaternal care. Finally, we transformed the data. We used the logit transformation to rescale the prosocial behavior from a truncated distribution in percentage or proportional terms to negative infinity to positive infinity (for extremes of 0 and 1 proportion of successes, respectively). For two species (*Macaca silenus* and *Varecia variegate*), values of zero for prosocial behavior were set to 0.001 before applying the logit to avoid negative infinity values. For allomaternal care, we used a $\log_{10}$ transformation.

The first stage of our analysis involves estimating the association between prosocial behavior and allomaternal care. The MCMC model reached stationarity (relatively consistent likelihood estimates after the initial burn-in) and showed a good distribution of likelihoods (see Figure 10.5A), with most of the chain sampling parameters that made the data more likely but also sampling some less likely parameter combinations (less often). We found strong evidence that allomaternal care should be included in the statistical model with prosocial behavior as a response variable, with 98.2% of the models including allomaternal care, and a generally higher likelihood in models that included allomaternal care (see Figures 10.5B and 10.5C). The posterior probability distribution of the regression coefficient relating allomaternal care to prosocial behavior is shown in Figure 10.5D. It can be seen that the vast majority of the coefficients are greater than zero, which is strongly indicative of a positive association. Therefore, the model using nonhuman primates suggests that allomaternal care is one reason *why* proactive prosociality may have increased in certain lineages (Burkart et al., 2014). Finally, we found support for both lambda and kappa transformations, with most of the chain favoring estimation of lambda rather than kappa, and a relatively flat distribution of kappa (see Figures 10.5E and 10.5F).

With posterior probability distributions of coefficients, kappa or lambda, and intercepts, we then predicted the value of prosocial behavior in humans, using our phylogenetic placement and value of allomaternal care (3.56, untransformed) for each of the 8,000 stored coefficients and parameters. This analysis produced 8,000 predictions, summarized by the posterior probability distribution for logit-transformed data shown in Figure 10.5G, which can be compared with the distribution of logit-transformed prosocial behavior for all primates (see Figure 10.5H). As can be seen, humans are not exceptionally different from other primates (shown as the black vertical line in Figure 10.5G), as might be expected given the small residual noted above. Thus, although the value of prosocial behavior in humans is slightly higher than expected, it is clearly not an outlier given human levels of allomaternal care and humans' placement in the primate phylogeny.

CONCLUDING REMARKS

The comparative methods reviewed in this chapter provide an essential tool kit for investigating a wide range of questions about when, why, and how cognition evolves. Coupled with high-quality datasets on species differences in cognition, this approach has the potential to catalyze an exciting revolution in comparative psychology. More important, the success of this endeavor will require datasets covering larger comparative samples than are currently common in this field. For example, several of the methods we reviewed are known to be highly sensitive to sample size, having limited statistical power in samples

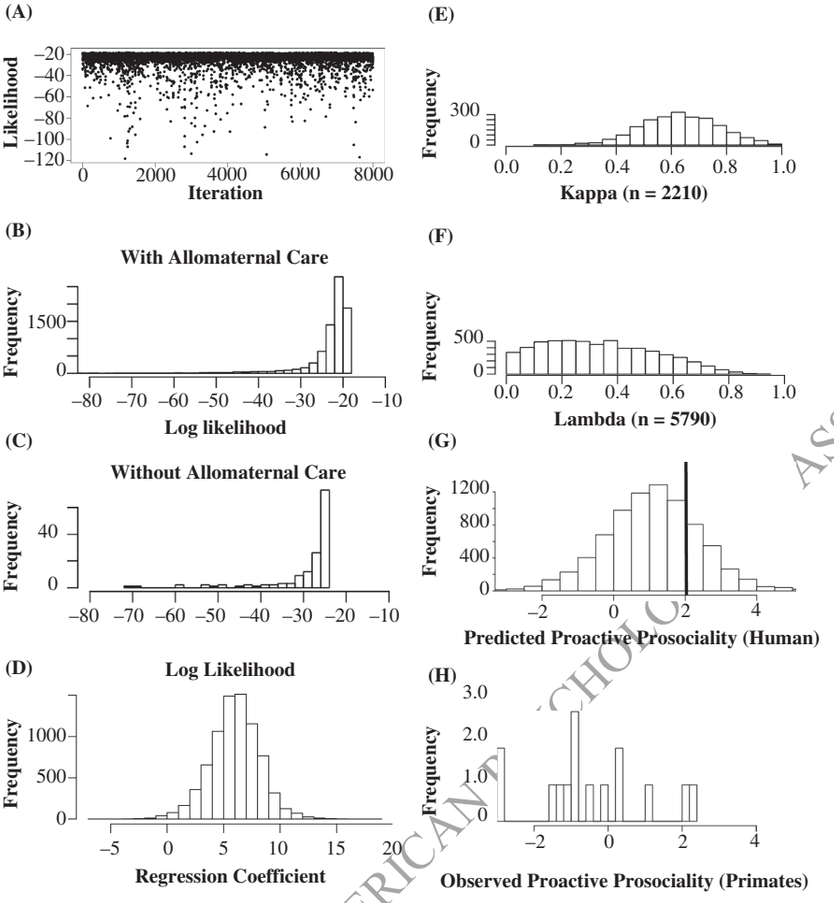


FIGURE 10.5. Phylogenetic prediction of proactive prosociality in humans. A: Distribution of likelihoods based on the Markov chain Monte Carlo model. B: Likelihoods for models including allomaternal care. C: Likelihoods for models excluding allomaternal care. D: Posterior probability distribution of regression coefficients relating allomaternal care to proactive prosociality. E and F: Distributions of kappa and lambda across model fits. G and H: Posterior probability distribution for human proactive prosociality compared with observed variation in nonhuman primates. The vertical black line (G) indicates the observed value for humans. See text for details.

smaller than 20 to 30 species (Kamilar & Cooper, 2013; Münkemüller et al., 2012). Thus, phylogenetic approaches for the study of comparative psychology will require not only special statistical methods but also datasets that are appropriate for the implementation of these techniques. In this regard, we expect that large-scale collaboration will also be an essential ingredient for success (MacLean et al., 2012, 2014).

For readers interested in learning more about comparative methods, a wide variety of freely available resources now exist that cover both basic and advanced topics (see Table 10.1). Although comparative methods can be implemented in many different software packages, which range in the extent of their flexibility, the amount of coding required by the user, and documentation, the vast majority of comparative methods (including all of those presented in this chapter) are available in R (http://www.r-project.org), a free and open-source language for statistical computing. Table 10.1 provides a list of useful packages for phylogenetic analysis using R, and tutorials including example code for the methods presented in

TABLE 10.1

Online Resources for Phylogenetic Comparative Methods

Resource	URL	Reference
	Websites	
AnthroTree	http://www.anthrotree.info	Nunn (2011)
10kTrees Project	http://10ktrees.fas.harvard.edu	Arnold, Matthews, and Nunn (2010)
TreeBASE	http://treebase.org	Sanderson, Donoghue, Piel, and Eriksson (1994)
Open Tree of Life	http://tree.opentreeoflife.org	Hinchliff et al. (2014)
	Programming language	
R Project for Statistical Computing	http://www.r-project.org	R Core Team (2014)
	R packages	
APE	http://cran.r-project.org/package=ape	Paradis, Claude, and Strimmer (2004)
CAPER	http://cran.r-project.org/package=caper	Orme et al. (2011)
Phytools	http://cran.r-project.org/package=phytools	Revell (2011)
Geiger	http://cran.r-project.org/package=geiger	Harmon et al. (2009)

this chapter can be accessed online at the AnthroTree website (http://www.anthrotree.info).

In conclusion, phylogenetic comparative methods will play a key role in unraveling the natural history of cognition and inferring the historical processes that have shaped the minds of the species we study today. We hope the topics reviewed in this chapter will stimulate further interest in phylogenetic approaches to the study of comparative psychology, and we look forward to future discoveries emerging from this approach.

References

Abouheif, E. (1999). A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research*, *1,* 895–909.

Amici, F., Aureli, F., & Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Current Biology*, *18,* 1415–1419. http://dx.doi.org/10.1016/j.cub.2008.08.020

Arnold, C., Matthews, L. J., & Nunn, C. L. (2010). The 10k trees website: A new online resource for primate phylogeny. *Evolutionary Anthropology*, *19,* 114–118. http://dx.doi.org/10.1002/evan.20251

Blomberg, S. P., & Garland, T., Jr. (2002). Tempo and mode in evolution: Phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, *15,* 899–910. http://dx.doi.org/10.1046/j.1420-9101.2002.00472.x

Blomberg, S. P., Garland, T., Jr., & 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. http://dx.doi.org/10.1111/j.0014-3820.2003.tb00285.x

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., . . . van Schaik, C. P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications*, *5,* 4747. http://dx.doi.org/10.1038/ncomms5747

Cunningham, C. W., Omland, K. E., & Oakley, T. H. (1998). Reconstructing ancestral character states: A critical reappraisal. *Trends in Ecology and Evolution*, *13,* 361–366. http://dx.doi.org/10.1016/S0169-5347(98)01382-2

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

Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *American Naturalist*, *160,* 712–726. http://dx.doi.org/10.1086/343873

Fritz, S. A., & Purvis, A. (2010). Selectivity in mammalian extinction risk and threat types: A new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, *24,* 1042–1051. http://dx.doi.org/10.1111/j.1523-1739.2010.01455.x

Garamszegi, L. Z. (Ed.). (2014). *Modern phylogenetic comparative methods and their application in evolutionary biology*. http://dx.doi.org/10.1007/978-3-662-43550-2

Garamszegi, L. Z., & Gonzalez-Voyer, A. (2014). Working with the tree of life in comparative studies: How to build and tailor phylogenies to interspecific datasets. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 19–48). http://dx.doi.org/10.1007/978-3-662-43550-2

Garland, T., Jr., Bennett, A. F., & Rezende, E. L. (2005). Phylogenetic approaches in comparative physiology. *Journal of Experimental Biology*, *208,* 3015–3035. http://dx.doi.org/10.1242/jeb.01745

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

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

Harmon, L., Weir, J., Brock, C., Glor, R., Challenger, W., & Hunt, G. (2009). Geiger: Analysis of evolutionary diversification [Computer software]. Retrieved from http://cran.r-project.org/package=geiger

Harvey, P. H., & Purvis, A. (1991). Comparative methods for explaining adaptations. *Nature*, *351,* 619–624. http://dx.doi.org/10.1038/351619a0

Herculano-Houzel, S. (2009). The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, *3,* 31. http://dx.doi.org/10.3389/neuro.09.031.2009

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences, USA, 109*(Suppl. 1), 10661–10668. http://dx.doi.org/10.1073/pnas.1201895109

Herrmann, E., Call, J., Hernàndez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317,* 1360–1366. http://dx.doi.org/10.1126/science.1146282

Hinchliff, C. E., Smith, S. A., Allman, J. F., Burleigh, J. G., Chaudhary, R., Coghill, L. M., . . . Gude, K. (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences, USA, 112*, 12764–12769. http://dx.doi.org/10.1073/pnas.1423041112

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17,* R1004–R1005. http://dx.doi.org/10.1016/j.cub.2007.10.027

Ives, A. R., & Garland, T. J. (2014). Phylogenetic regression for binary dependent variables. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 231–261). http://dx.doi.org/10.1007/978-3-662-43550-2

Kamilar, J. M., & Cooper, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 368,* 20120341. http://dx.doi.org/10.1098/rstb.2012.0341

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., . . . Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences, USA, 111*, E2140–E2148. http://dx.doi.org/10.1073/pnas.1323533111

MacLean, E. L., Matthews, L. J., Hare, B. A., Nunn, C. L., Anderson, R. C., Aureli, F., . . . Wobber, V. (2012). How does cognition evolve? Phylogenetic comparative psychology. *Animal Cognition, 15,* 223–238. http://dx.doi.org/10.1007/s10071-011-0448-8

MacLean, E. L., Sandel, A. A., Bray, J., Oldenkamp, R. E., Reddy, R. B., & Hare, B. A. (2013). Group size predicts social but not nonsocial cognition in lemurs. *PLOS ONE, 8,* e66359. http://dx.doi.org/10.1371/journal.pone.0066359

Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology*, *64,* 127–136. http://dx.doi.org/10.1093/sysbio/syu070

Maddison, W. P., & Slatkin, M. (1991). Null models for the number of evolutionary steps in a character on a phylogenetic tree. *Evolution: International Journal of Organic Evolution*, *45,* 1184–1197. http://dx.doi.org/10.2307/2409726

Martins, E. P., & Garland, T. (1991). Phylogenetic analyses of the correlated evolution of continuous characters: A simulation study. *Evolution: International Journal of Organic Evolution*, *45,* 534–557. http://dx.doi.org/10.2307/2409910

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. *American Naturalist*, *149,* 646–667. http://dx.doi.org/10.1086/286013

McPeek, M. A. (1995). Testing hypotheses about evolutionary change on single branches of a phylogeny using evolutionary contrasts. *American Naturalist*, *145,* 686–703. http://dx.doi.org/10.1086/285763

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, *3,* 743–756. http://dx.doi.org/10.1111/j.2041-210X.2012.00196.x

Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. http://dx.doi.org/10.7208/chicago/9780226090009.001.0001

Nunn, C. L., & Barton, R. A. (2001). Comparative methods for studying primate adaptation and allometry. *Evolutionary Anthropology*, *10,* 81–98. http://dx.doi.org/10.1002/evan.1019

Nunn, C. L., & Zhu, L. (2014). Phylogenetic prediction to identify "evolutionary singularities.". In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 481–514). http://dx.doi.org/10.1007/978-3-662-43550-2

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

Organ, C., Nunn, C. L., Machanda, Z., & Wrangham, R. W. (2011). Phylogenetic rate shifts in feeding time during the evolution of *Homo*. *Proceedings of the National Academy of Sciences, USA*, *108,* 14555–14559. http://dx.doi.org/10.1073/pnas.1107806108

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., & Isaac, N. (2011). CAPER: Comparative analyses of phylogenetics and evolution in R (Version 0.4/r64) [Computer software]. Retrieved from http://cran.r-project.org/package=caper

Pagel, M. (1994). Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 255,* 37–45. http://dx.doi.org/10.1098/rspb.1994.0006

Pagel, M. (1999a). Inferring the historical patterns of biological evolution. *Nature*, *401,* 877–884. http://dx.doi.org/10.1038/44766

Pagel, M. (1999b). The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology*, *48,* 612–622. http://dx.doi.org/10.1080/106351599260184

Pagel, M., & Lutzoni, F. (2002). Accounting for phylogenetic uncertainty in comparative studies of evolution and adaptation. In M. Lässig & A. Valleriani (Eds.), *Biological evolution and statistical physics* (pp. 148–161). http://dx.doi.org/10.1007/3-540-45692-9_8

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20,* 289–290. http://dx.doi.org/10.1093/bioinformatics/btg412

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31,* 109–130. http://dx.doi.org/10.1017/S0140525X08003543

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366,* 1017–1027.

Revell, L. J. (2008). On the analysis of evolutionary change along single branches in a phylogeny. *American Naturalist*, *172,* 140–147. http://dx.doi.org/10.1086/588078

Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution*, *1,* 319–329. http://dx.doi.org/10.1111/j.2041-210X.2010.00044.x

Revell, L. J. (2011). Phytools: Phylogenetic tools for comparative biology (and other things) (Version 0.0-8) [Computer software]. Retrieved from http://cran.r-project.org/package=phytools

Revell, L. J., Mahler, D. L., Peres-Neto, P. R., & Redelings, B. D. (2012). A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution: International Journal of Organic Evolution*, *66,* 135–146. http://dx.doi.org/10.1111/j.1558-5646.2011.01435.x

Rezende, E. L., & Diniz-Filho, J. A. F. (2012). Phylogenetic analyses: Comparing species to infer adaptations and physiological mechanisms. *Comprehensive Physiology*, *2,* 639–674.

Sanderson, M., Donoghue, M., Piel, W., & Eriksson, T. (1994). Treebase: A prototype database of phylogenetic analyses and an interactive tool for browsing the phylogeny of life. *American Journal of Botany*, *81,* 183.

Shettleworth, S. J. (2012). Modularity, comparative cognition and human uniqueness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367,* 2794–2802. http://dx.doi.org/10.1098/rstb.2012.0211

Shultz, S., & Dunbar, R. I. M. (2010). Species differences in executive function correlate with hippocampus volume and neocortex ratio across nonhuman primates. *Journal of Comparative Psychology*, *124,* 252–260. http://dx.doi.org/10.1037/a0018894

Stevens, J. R. (2014). Evolutionary pressures on primate intertemporal choice. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 281,* 20140499.

Symonds, M. R., & Blomberg, S. P. (2014). A primer on phylogenetic generalised least squares. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 105–130). http://dx.doi.org/10.1007/978-3-662-43550-2

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20,* 410–433. http://dx.doi.org/10.1111/j.1439-0310.1963.tb01161.x