

A Bayesian extension of phylogenetic generalized least squares: Incorporating uncertainty in the comparative study of trait relationships and evolutionary rates

Jesualdo A. Fuentes-G.,^{1,2}  Paul David Polly,³  and Emília P. Martins⁴ 

¹Department of Biological Sciences, The University of Alabama, Tuscaloosa, Alabama

²E-mail: jafuentesgonzalez@ua.edu

³Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, Indiana

⁴School of Life Sciences, Arizona State University, Tempe, Arizona

Received November 29, 2017

Accepted November 19, 2019

Phylogenetic comparative methods use tree topology, branch lengths, and models of phenotypic change to take into account nonindependence in statistical analysis. However, these methods normally assume that trees and models are known without error. Approaches relying on evolutionary regimes also assume specific distributions of character states across a tree, which often result from ancestral state reconstructions that are subject to uncertainty. Several methods have been proposed to deal with some of these sources of uncertainty, but approaches accounting for all of them are less common. Here, we show how Bayesian statistics facilitates this task while relaxing the homogeneous rate assumption of the well-known phylogenetic generalized least squares (PGLS) framework. This Bayesian formulation allows uncertainty about phylogeny, evolutionary regimes, or other statistical parameters to be taken into account for studies as simple as testing for coevolution in two traits or as complex as testing whether bursts of phenotypic change are associated with evolutionary shifts in intertrait correlations. A mixture of validation approaches indicates that the approach has good inferential properties and predictive performance. We provide suggestions for implementation and show its usefulness by exploring the coevolution of ankle posture and forefoot proportions in Carnivora.

KEY WORDS: Bayesian statistics, Carnivora, evolutionary regimes, phylogenetic comparative methods, phylogenetic generalized least squares, uncertainty.

The results of phylogenetic comparative methods are only as accurate as the phylogenies and models of trait evolution assumed in calculating those results. Both phylogeny reconstruction and the phylogenetic comparative method have come a long way since Felsenstein (1985) proposed his simple independent contrasts procedure, and evolutionary biologists can now easily build complex statistical models to address a wide variety of sophisticated questions using interspecific data (Garamszegi 2014). However, applying these methods involves making assumptions about a myriad of sources of error and uncertainty in the data (e.g., Hansen and Bartoszek 2012), phylogeny, and the models used to reconstruct trait evolution along those phylogenies (e.g., Pennell et al. 2015).

Although randomization and other early approaches to incorporating phylogenetic uncertainty can be useful with simple phylogenetic comparative methods, Bayesian approaches offer a modern and more natural way to incorporate uncertainty in complex models (Currie and Meade 2014). Here, we illustrate the process of developing a Bayesian extension of a complex interspecific model, and its impact on our understanding carnivoran limb evolution.

Traditionally, questions about phenotypic disparity (understood as the degree of dissimilarity in a given trait for a group of species) and evolutionary trait relationships have been addressed separately. Shifts in phenotypic disparity have been mainly

studied by focusing on evolutionary rates (e.g., O'Meara et al. 2006; Thomas et al. 2006), as in Collar et al.'s (2010) study which found that dragon lizards living in terrestrial and semiarboreal habitats have experienced higher disparity in their limbs and body-forms than have dragon lizards living in rock-dwelling and arboreal habitats. Other researchers have focused on group differences (e.g., Díaz-Uriarte and Garland 1998; Lindenfors 2006) and trait covariation (e.g., Garland et al. 1993; Butler et al. 2000), finding, for example, that the relative size of the small intestine (relationship between intestine and overall body size) in birds and nonflying mammals is associated with diet (Lavin et al. 2008). However, these two types of questions are not mutually exclusive and important pieces of information can be concealed to methods that focus exclusively on either phenotypic disparity or trait relationships (Fuentes-G. et al. 2016). More recent comparative methods account for both phenomena simultaneously by fitting multiple evolutionary rate matrices to a phylogenetic tree (Caetano and Harmon 2018b; Revell and Collar 2009; Caetano and Harmon 2017), or by estimating changes in phenotypic variances as well as regression coefficients of specific lineages (Barton and Venditti 2014; Fuentes-G. et al. 2016).

Despite the advantages of accounting for both phenotypic disparity and trait covariation, the aforementioned examples are affected by several sources of uncertainty. The most obvious source of uncertainty for any comparative method is the phylogeny (Harvey and Pagel 1991), which is usually assumed to be fully known and completely correct (Huelsenbeck and Rannala 2003). But since phylogenies are inferred (not observed), they are working hypotheses of common ancestry and are subject to error (Pagel and Lutzoni 2002; Garland et al. 2005; Blomberg et al. 2012; Hernández et al. 2013). Any single phylogeny can be affected by topological errors, incorrect branch lengths, and unresolved nodes (Rezende and Diniz-Filho 2012). Ignoring phylogenetic uncertainty can produce biased estimates of the residual variance as well as artificially narrow confidence intervals, leading to a false perception of precision (de Villemereuil et al. 2012).

Several phylogenetic comparative methods also require knowing the ancestral states of one or more traits. For example, tests of unequal evolutionary rates (e.g., O'Meara et al. 2006; Thomas et al. 2009; Fuentes-G. et al. 2016) and early versions of the adaptation-inertia model (Hansen 1997; Butler and King 2004) require specifying trait values for each branch of the phylogeny, which is usually accomplished via ancestral state reconstruction, thereby introducing another source of error (Pagel and Harvey 1988; Ronquist 2004; Ng and Smith 2014). This practice requires specifying the branches on which the trait might change or equally unlikely assumptions such as shifts taking place only at speciation events (O'Meara et al. 2006). These assumptions can be relaxed by estimating the shifts themselves (e.g., Eastman et al. 2011; Barton and Venditti 2014; Rabosky et al. 2014) or adding extra

parameters in the model (e.g., g in Fuentes-G. et al. 2016), at the expense of model complexity.

Both phylogeny and ancestral state reconstructions directly influence the expected covariance between species, which will have implications for parameter estimation (Thomas et al. 2006). But being unknown quantities, parameter estimates themselves are subject to uncertainty (Bernardo 2003), a condition that can affect modern phylogenetic comparative methods relying on specific models of phenotypic change (Boettiger et al. 2012). Brownian motion (BM), as well as models incorporating constraining forces (Felsenstein 1988; Martins 1994) or **some measure of phylogenetic signal** (Grafen 1989; Lynch 1991; **Pagel 1999**; Blomberg et al. 2003), relies on evolutionary parameters that are critical for explicit interpretations of the results, but are subject to uncertainty just as any parameter. Moreover, explicit models of phenotypic change are advantageous because they link the results with evolutionary processes such as genetic drift and fluctuating directional selection (Martins and Hansen 1996). We can learn about these processes by comparing sets of phenotypic models (Hernández et al. 2013), but it is important to keep in mind that the models themselves are mere approximations to reality (Donoghue and Ackerly 1996; Burnham and Anderson 2004). A given model may be selected because it fits the data better than other candidates, but this does not guarantee that the preferred option is correct, nor that the other candidates contain meaningless information (Box and Draper 1987; Boettiger et al. 2012). Because a true model cannot be known, there is uncertainty about the adequacy of competing models, and parameter estimation should reflect such uncertainty (Hoff 2009).

The Bayesian framework provides formal tools for accommodating all these sources of uncertainty, because all quantities in the analysis are treated as random variables (Huelsenbeck et al. 2000; Huelsenbeck and Rannala 2003). This treatment is particularly relevant in the analysis of interspecific data where a unique history has produced the trait distributions we observe today, but there is uncertainty about the details of such unique history (Schultz and Churchill 1999). The success of the Bayesian paradigm in dealing with such historical accounts is shown by a long list of methods dealing with a variety of topics such as phylogenetic reconstruction (e.g., Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003; Drummond and Rambaut 2007), ancestral state estimation (Pagel et al. 2004), correlated evolution (Pagel and Meade 2006), and even for making joint inferences of different evolutionary questions (Caetano and Harmon 2018b; Huelsenbeck and Rannala 2003). Recent Markov chain Monte Carlo (MCMC) developments freed comparative methods from using specific phylogenetic partitions based on ancestral state reconstructions to estimate shifts in trait optima (Uyeda and Harmon 2014; Uyeda et al. 2017), species diversification (Rabosky 2014), phenotypic disparity (Eastman et al. 2011;

Venditti et al. 2011; Revell et al. 2012), and even combinations of these processes (Rabosky et al. 2013; Rabosky et al. 2014). In general, these methods use reversible-jump MCMC (Green 2003; Sisson 2005) to estimate the location of evolutionary shifts, accounting for the lack of knowledge about character state transitions on the phylogenetic tree. For example, Barton and Venditti (2014) presented a similar phylogenetic regression framework as we do here, but using a reversible-jump variable rates model to avoid specifying the evolutionary history of shifts in advance.

These types of reversible-jump models are sound when the shifts are not linked to specific hypotheses of historical events, and when the phylogeny is known. However, the phylogeny is rarely known without error (Martins 1996), and there are cases in which we would like the historical events to be informed by transitions in factors such as habitat, diet, or behavior (as these can be part of the hypotheses being tested). On these lines, we adopt an alternative solution by using stochastic character mapping (SCM; Nielsen 2002; Huelsenbeck et al. 2003) under the Bayesian PGLS formulation proposed by Blomberg et al. (2012) and de Villemereuil et al. (2012). SCM is a Bayesian method that samples multiple descriptions of character histories on phylogenies in proportion to their posterior probability. By considering multiple phylogenetic hypotheses and character mappings, comparative methods do not have to assume that the history of transitions nor the phylogenies in which they are reconstructed are known (Ronquist 2004; Thomas et al. 2006). Although we show here how this can be accomplished for models estimating unequal phenotypic rates (e.g., O'Meara et al. 2006; Thomas et al. 2006; Fuentes-G. et al. 2016), the same principle can be used for methods dealing with shifts in trait optima (e.g., Hansen 1997; Butler and King 2004), or a combination thereof (e.g., Beaulieu et al. 2012).

In summary, we present a Bayesian PGLS extension that evaluates temporal shifts in trait relationships and relative phenotypic rates, while accounting for uncertainty in parameter estimation, model selection, phylogenetic relationships, and reconstructions of historical events. We show the main properties of the approach by exploring a coevolutionary trend between forefoot proportions and hindlimb posture in the order Carnivora. We also provide scripts implementing the approach as freely available software.

How the Method Works

The likelihood of a PGLS model can be specified through a multivariate normal distribution:

$$\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \mathbf{V} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}) \quad (1)$$

where \mathbf{Y} is the response variable as a column matrix, \mathbf{X} is a matrix that includes the explanatory variables, $\boldsymbol{\beta}$ is the vector of partial regression coefficients, and \mathbf{V} is the phylogenetic covariance matrix

which also incorporates a model of evolutionary change (Hansen and Martins 1996). Under BM, the model is characterized by the phenotypic variance (γ) and the divergence times of the phylogenetic tree. It is also possible to measure the extent to which the covariance structure of the residuals deviates from a Brownian process by estimating a phylogenetic signal parameter (λ) to this model (Pagel 1999; Freckleton et al. 2002). This lambda model (LM) comprises both BM and the case when phylogeny is not explicitly identified (independent and identically distributed residuals: IID).

The use of both categorical and continuous predictors allows detection of how different states of a factor generate changes in the association between a phenotypic trait and a covariate (e.g., Smaers and Rohlf 2016). Tests for bursts of phenotypic change can be incorporated into this framework by assuming rates of evolution that are dependent on the trait of the discrete predictor, with a separate parameter (γ) for each value of that discrete variable (Fuentes-G. et al. 2016). Below we show how a Bayesian version of this model can account for uncertainty in phylogenetic relationships, divergence times, character mapping, model selection, and parameter estimates.

First, the Bayesian formulation requires quantitative representations of parameter uncertainty, which can be explicitly incorporated by using prior distributions on model parameters. The specification of prior distributions for regression models can be challenging because parameter values do not have direct equivalents to observed measures. For example, it is more obvious to determine the expected mean for height or weight (which can be directly measured) than for a regression coefficient, but one way to work around this complication is by using diffuse priors (Hoff 2009; Kruschke and Liddell 2016). In this paper, we used diffuse conjugate univariate normal priors for regression coefficients (Blomberg et al. 2012): $\boldsymbol{\beta} \sim N(0, 10^6)$. It is common in Bayesian analysis to parameterize variances in terms of precision, so a uniform distribution can be applied on the inverse of the phenotypic variance to specify a diffuse prior (Gelman 2006): $\gamma^{-1} \sim U(0, 100)$. When trees are ultrametric and scaled to a height of one, a bounded uniform prior can be specified for λ in models estimating phylogenetic signal (de Villemereuil et al. 2012): $\lambda \sim U(0, 1)$.

When using a single topology with a particular set of branch lengths, it is assumed that the true phylogenetic covariance matrix (\mathbf{V}) is known. To account for this uncertainty, an empirical prior distribution of trees with multiple sets of branch lengths can be incorporated, which can be written generally as (de Villemereuil et al. 2012): $\mathbf{V} \sim \Pi(\xi)$, where Π represents any relevant distribution with parameters ξ . More specifically, the distribution represents a collection of trees derived, for example, from Bayesian phylogenetic inference (e.g., a posterior distribution of trees) or bootstrap analysis (Felsenstein 1988; Ronquist 2004;

Rabosky 2015). By using MCMC sampling, we can integrate over the collection of trees weighted by the probability that each is correct (Huelsenbeck et al. 2000; Blomberg et al. 2012). To incorporate the uncertainty of ancestral state reconstructions, we generate stochastic character maps for the trees in the empirical prior of phylogenies so that it represents probability distributions of topologies, branch lengths, and mappings of the factor. By integrating parameter estimates over this distribution, the different sources of uncertainty are taken into account.

It is common to fit different models (e.g., IID, BM, LM) under PGLS analyses to better assess the influence of phylogeny on phenotypic associations, drawing conclusions from those candidates that fit best the data (e.g., Lavin et al. 2008; Slater and Pennell 2014; Fuentes-G. et al. 2016; Zúñiga-Vega et al. 2017). Here, we explore an alternative that turns this model selection problem into a parameter estimation procedure. Given that both IID and BM can be seen as specific subsets of LM, we formulated a simple mixture model (MM) by specifying an index for each component with a hyperprior of equal probability (Bernardo and Smith 2009). The values for the index can be generated by a generalized Bernoulli distribution (Kruschke 2011): $M_i \sim \Pi(\{IID, BM, LM\}, [p_{IID}, p_{BM}, p_{LM}])$, where p_i correspond to prior probabilities for each component (with $\sum_{i=1}^{k=3} p_i = 1$); In this way, the Markov chain samples models of phylogenetic relatedness in direct proportion to their posterior probability. Besides estimating a probability for each type of model given the data: $p(M_i | D)$, this approach facilitates the use of Bayesian model averaging, in which parameter estimation is conducted under a set of models of high probability (Raftery et al. 1997).

An Applied Example: Correlation between Forelimb and Hindlimb Evolution

CARNIVORAN LIMB PROPORTIONS

We use the Bayesian approach to explore limb coevolution. Mammals have different ways of maximizing locomotor performance, such as speed (Garland and Janis 1993; Polly 2007). One of the best-known ways is by increasing the overall length of the limbs, because longer limbs produce longer strides, which results in increased speed (Gregory 1912; Garland and Janis 1993). The relative proportion of the metapodials to other limb elements is generally associated with stride length, stance, and running speed (Polly 2007; Wang and Tedford 2007; Samuels et al. 2013). Specializations in the fore and hindlimbs can, in principle, be independent of one another (e.g., Meachen-Samuels and Van Valkenburgh 2009; Bell et al. 2011; Martín-Serra et al. 2014). Forelimb elongation can be manifested through shifts in foot proportions, where long metacarpals and short digits (e.g., hyenas, jackals) are associated

with increased speed (Van Valkenburgh 1985). In the hindlimbs, limb elongation can be manifested through shifts in posture, where animals that walk with elevated heels (digitigrade posture in dogs and cats, for example) use the metatarsus as an additional limb segment and thus increase running ability (Polly 2007).

If running speed is the major driving force in both limbs, we would expect a coevolutionary trend between forefoot proportions and posture, with relatively short digits showing a consistent association with the use of the metapodials as an additional limb segment (Van Valkenburgh 1985). Further, morphological specialization under running speed can lead to the expectation of a tight coevolutionary trend (Martín-Serra et al. 2015), under which species with a digitigrade posture can only exhibit a restricted range of forefoot proportions. This is, forefoot proportions would exhibit low phenotypic disparity under digitigrady, because large deviations from those digit and metacarpal lengths that maximize speed could compromise locomotor performance (Gregory 1912). But as other evolutionary forces (e.g., bearing loads or manipulating food) can also act on limbs (Andersson 2004; Polly 2007; Meachen-Samuels and Van Valkenburgh 2009), changes in forefoot proportions and hindlimb posture do not have to concomitantly favor running ability, weakening or obliterating the association between the two traits. Here, we test for coevolution in hindlimb posture and forelimb proportions in fissiped (terrestrial) families of the order Carnivora.

Forefoot proportions have sometimes been captured by the metacarpal/phalanx ratio, a morphological indicator of locomotor behavior (Van Valkenburgh 1985; Meachen-Samuels and Van Valkenburgh 2009; Samuels et al. 2013). The variables involved are the lengths of the third metacarpal and the proximal phalanx, which were obtained from Samuels et al. (2013). Rather than assuming the nature of the association between these two morphological traits (Albrecht et al. 1993; Jasieński and Bazzaz 1999; Packard and Boardman 1999), here we explore their relationship explicitly by incorporating them as individual variables in a multiple regression setup (García-Berthou 2001; Freckleton 2002). Hindlimb posture was coded as a dichotomous variable where plantigrade animals walk with their heels always or frequently touching the ground (e.g., red pandas, skunks), and digitigrade animals always have their heels elevated during normal locomotion (e.g., dogs and cats). We started with the categorizations from Polly (2010), adding and amending categorizations as necessary based on the aforementioned criteria and observations of videos of animals in motion (character state assignments and further details about the categorization are available in the example data and description files in the Dryad Digital Repository: <https://doi.org/10.5061/dryad.9kd51c5ct>; Fuentes-G. et al. 2019). We incorporated forelimb phalanx length as response variable (**Y**), whereas the matrix of predictors (**X**) included four elements: a unit vector to account for the global

intercept, metacarpal length as covariate, hindlimb posture as factor, and the interaction between metacarpal length and posture.

The null hypothesis under this variable setup would be the lack of coevolutionary pattern, reflected in indistinguishable regression lines (between forelimb phalanx and metacarpal lengths) for plantigrades and digitigrades. A coevolutionary pattern favoring running speed would be reflected in separate regression lines, with digitigrady showing shorter phalanxes for equivalent metacarpal lengths (i.e., regression line for plantigrady above regression line for digitigrady). If such coevolutionary pattern leads to morphological specialization, phenotypic rates linked to the digitigrade line would be lower than those for plantigrady, indicating a tighter association for the former. The opposite outcome in terms of phenotypic disparity (i.e., larger rates for digitigrady) would be indicative of a coevolutionary pattern that promotes running ability without constraining forefoot proportions, allowing fore- and hindlimbs to respond to different functional demands.

DATA AND METHOD DETAILS

We used a set of 1000 phylogenies with branch lengths obtained from the *10kTrees* project (Arnold et al. 2010) under version 1 (Fig. 1). The tree distribution obtained from this resource was generated under Bayesian inference with 14 mitochondrial and 15 autosomal genes analyzed under different schemes according to a reversible-jump MCMC that used specific proportion of invariable sites and rate variation for each marker. Node ages were inferred using the mean molecular branch lengths from the Bayesian search and 16 fossil calibration points. We only used taxa that were identified to the species level, and for which both phenotypic data and trees were available, resulting in a dataset of 102 species.

For each tree in the distribution, we generated 10 stochastic maps of posture using SIMMAP Version 1.5 (Bollback 2006) under the Mk model (Lewis 2001) with a beta distribution prior for the bias parameter ($\alpha = 3.91$, $k = 31$) and a gamma distribution prior for the rate parameter ($\alpha = 3.76$, $\beta = 0.53$, $k = 60$). These priors were configured under a Bayesian procedure (Schultz and Churchill 1999) using scripts made available by Bollback (2009) for R (R Development Core Team 2016), along with the packages MASS (Venables and Ripley 2002) and TeachingDemos (Snow 2016). Briefly, this procedure accounts for uncertainty in the selection of correct priors by considering ranges of parameter values that are weighted by their probability given the data. This was implemented by running a MCMC under loose priors (for the bias parameter: $\alpha = 1.0$, $k = 31$; for the rate parameter: $\alpha = 1.25$, $\beta = 0.25$, $k = 60$) for 100,000 generations on the consensus of the posterior distribution of trees, sampling every 200 steps and discarding the first 10,000 as burn-in. The best fitting beta and gamma distributions of this initial spectrum of values informed the priors for the stochastic mappings, allowing the parameters

of the process of character evolution to be estimated rather than fixed. The resulting 10,000 reconstructions were randomly sampled without replacement to generate an empirical prior with 1000 mappings of equal probability. Using the entire distribution of reconstructions would be ideal but also computationally limiting (see de Villemereuil et al. 2012 for details). Reducing the set of phylogenies for the empirical prior, and/or the number of stochastic maps for each tree, can make the problem computationally feasible but at the same time can make the estimation more susceptible to sampling errors. The random sampling scheme adopted here constitutes a reasonable compromise between computational feasibility and the advantages of using a comprehensive posterior probability distribution of both trees and ancestral reconstructions (Collar et al. 2010; de Villemereuil et al. 2012). Still, if phylogenetic information is highly variable and the sample is small, it is less likely that the empirical prior will represent the true probability distribution of historical reconstructions, and the results will only partially account for this source of uncertainty.

Phenotypic rates of evolution of forefoot proportion were estimated under mappings for specific character states according to their own prior (γ_1 for plantigrady, γ_2 for digitigrady), using partitioned phylogenetic covariance matrices: $\mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2$. In principle, the diffuse priors defined above should have little influence on the posterior distribution so that relationships among biological variables are dominated by the likelihood function (Huelsenbeck and Rannala 2003). However, there are situations in which such priors do not actually conform to this expectation (e.g., Yang and Rannala 2005). We tested if this was the case for our prior specifications by running the MCMC without data and conducting sensitivity analysis (Supporting Information A). The former allowed us to determine the effects of priors alone on the posterior distributions, and the latter allowed us to assess the robustness of results to different prior specifications.

We ran three chains for a total of 150,000 generations, without thinning (Link and Eaton 2012). We used 1000 steps to tune the samplers and excluded the first 15,000 generations as burn-in. We evaluated the behavior of the chain in different ways, using the R packages stats, graphics, and coda (Plummer et al. 2006). First, we evaluated white noise through trace plots, under which the mixing of the chain can be diagnosed. We also diagnosed the mixing of the chain through autocorrelation plots showing self-similarity of the samples in the chain. We computed the effective sample sizes (N_e) for each parameter; acceptable behavior was determined when $N_e > 1000$. Finally, we confirmed convergence by applying stationarity and half-width tests (Heidelberger and Welch 1981; Heidelberger and Welch 1983) with $\alpha = 0.05$ and $\varepsilon = 0.1$.

We assessed changes in phenotypic relationships by exploring the posterior distribution of regression coefficients. First, we obtained the 95% highest density interval (HDI), which provides

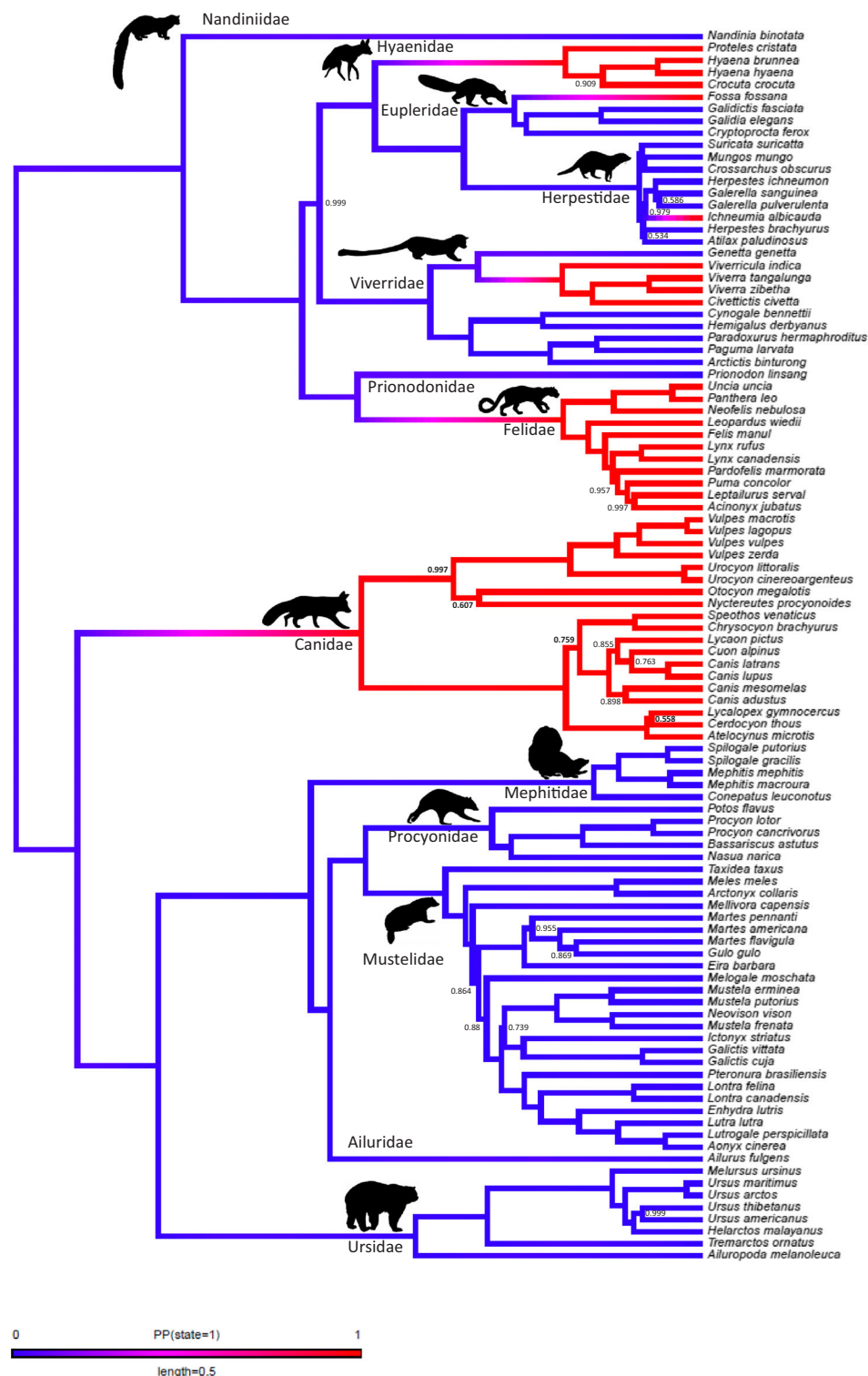


Figure 1. Consensus for the posterior distribution of 1000 trees showing a density map for posture and highlighting the **carnivorous families included in the study** (silhouettes from PhyloPic; see Acknowledgments). Nodal support provided (not shown when the $pp = 1$). The density map was built with a sample of 500 mappings in phytools (Revell 2012), and is consistent with the entire set (10,000 stochastic character mappings) on which the maximum likelihood analysis was based. Branch lengths are in units of relative amount of change expected in the phenotypic traits; the scale bar provided at the bottom left (with total length of 0.5) shows a color scheme of the posterior probabilities of character states for posture at every point in the tree (information about codification in the main text), with blue corresponding to high probability of plantigrady (0), red to high probability of digitigrady (1), and intermediate colors (purple) reflecting degrees of uncertainty in the reconstruction.

values of highest posterior probability that include the 95% most probable values of the parameter, given the data (Kruschke and Liddell 2016). Second, we determined the posterior probability (pp) of parameter values being greater or less than zero, by determining the area under the probability density function falling in a particular range (de Villemereuil et al. 2012). Changes in phenotypic disparity were assessed by exploring the posterior distribution of the difference among rates; these distributions provided the posterior probability that a specific rate was greater than another, as well as its credibility through the 95% HDI. As a confirmatory procedure, we compared the single-rate and the multirate models through the deviance information criterion (DIC), which corresponds to a Bayesian analogue of the Akaike information criterion and has similar interpretation (Spiegelhalter et al. 2002). We also conducted simulations to evaluate the power of our approach for detecting shifts in rates and trait relationships (Supporting Information B).

Besides the posterior probability estimated under MM, we used the posterior distribution of λ to assess the uncertainty in the influence of phylogeny, with deviations from 0.5 reflecting unbalanced influence of evolutionary or ahistorical components (e.g., Hsiang et al. 2016). To contrast the behavior of the MM with information criteria, we also ran Bayesian analyses under specific models (i.e., IID, BM, LM) and compared them by means of DIC (as with the phenotypic rate inferences outlined above, the DIC estimation was mostly conducted for comparative purposes). We also conducted posterior predictive checks to determine if the full model (MM) was properly accounting for the data (Supporting Information C), and if its extra complexity was compromising predictive properties when compared with a simpler alternative (LM). For most parameters, we obtained estimates through the arithmetic mean of the posterior distribution after discarding the burn-in (when the chain is supposed to be at stationarity). In the case of λ , we used the median instead of the mean as point estimate, because this parameter often produces asymmetric distributions that are not well represented by the mean (Steel and Kammeyer-Mueller 2008), especially when the phylogenetic signal is very high or very low (Revell and Graham Reynolds 2012).

We approximated the posterior distribution of these parameters using MCMC as implemented in JAGS Version 4.2.0 (Plummer 2003) associated with R through the package rjags (Plummer 2016). JAGS uses MCMC samplers such as Gibbs, slice, and the Metropolis-Hastings algorithm, all of which theoretically approach the posterior distribution with enough number of generations. The current implementation (deposited, along with the data and example files in the Dryad Digital Repository: <https://doi.org/10.5061/dryad.9kd51c5ct>; Fuentes-G. et al. 2019) extends the common models developed by de Villemereuil et al. (2012) and the codes built by Kruschke (2011).

For comparison, we analyzed the same dataset under maximum likelihood using the R scripts provided by Fuentes-G. et al. (2016). We used the consensus of the posterior distribution of trees described above for this purpose (particularly the 50% majority rule consensus tree), with branch lengths averaged over all the topologies containing the clade. This consensus tree was also generated by the *10kTrees* project (Arnold et al. 2010). For the specification of the evolutionary regime (i.e., fractions of branch lengths assigned to phylogenetic partitions characterized by specific states of the factor), we generated 10,000 SCMs for posture in SIMMAP using the same prior configuration explained earlier (Fig. 1). Character state transitions were assigned to those branches in which the posterior probability of a gain or loss in posture exceeded 0.5 (e.g., Schultz et al. 2016). We estimated regression coefficients with equal and unequal rates for each type of model (i.e., IID, BM, LM), and compared the resulting candidates using the second-order bias correction version of Akaike information criterion (AIC_c) (Akaike 1974, 1992). Lower AIC_c values are indicative of better fit and a contrast greater than two between them ($\Delta AIC_c > 2$) can be interpreted as a substantial difference between models (Burnham and Anderson 2002).

RESULTS

There is evidence for phenotypic disparity in carnivoran limb coevolution (Table 1), with the phenotypic rate of forefoot proportions under digitigrady being credibly larger than under plantigrady ($pp > 0.95$, zero outside HDI). The coevolutionary trend seems uncertain given that both group effect (β_2) and interaction (β_3) include zero within the HDI. The posterior distribution of regression lines shows high overlap for small metacarpal lengths (Fig. 2), consistent with a non-credible group effect (β_2). However, the lines diverge more clearly for carnivorans with larger metacarpals (Fig. 2), and in fact a negative interaction effect (β_3) has high posterior probability, albeit lower than 0.95 (Table 1). To explore this further, we computed and plotted the expected mean differences between responses under different levels of the factor for the same covariate value (Fig. 3). The differences are credible and negative at the upper end of the covariate, indicating that digitigrady is associated with decreases in phalanx length but not for carnivorans with short metacarpals, that is, below 3.3 in logarithmic scale (Fig. 3). This makes sense considering that there are no digitigrade carnivorans with metacarpal length (\ln) smaller than 3.1 (Fig. 2).

The relevance of a historical component on this pattern can be evidenced in different ways. First, phylogenetic signal is high and credibly larger than 0.5 (Table 1), indicating a strong influence of phylogeny. Second, the pattern is credibly better explained by models involving phylogenetic effects ($pp > 0.95$, Table 1).

Overall, the Bayesian results suggest that there is strong phenotypic disparity influenced by phylogeny, and although weaker,

Table 1. Results under Bayesian model.

Parameter	Estimate	HDI		<i>pp</i>
		Lower	Upper	
Plantigrady intercept (β_0)	−0.16	−0.57	0.258	0.223
Plantigrady slope (β_1)	0.89	0.779	1	1
Digitigrady group effect (β_2)	0.27	−0.49	1.01	0.765
Digitigrady interaction (β_3)	−0.14	−0.341	0.074	0.096
Plantigrady phenotypic rate (γ_1)	0.06	0.028	0.086	1
Digitigrady phenotypic rate (γ_2)	0.2	0.052	0.382	1
Phylogenetic signal (λ)	0.88	0.662	0.99	1

Parameter estimates under Bayesian model averaging are presented with credible ranges (lower and upper margins of HDI) and posterior probabilities (*pp*) of being larger than zero. Estimates correspond to posterior means for all parameters except λ , for which the median is reported when MM is under LM (see text for details). The parameter is credibly larger than 0.5 (*pp* = 0.992), with the HDI completely above this value. The MM spent most of its time in models accounting for phylogeny: $p(\text{IID}|\text{D}) = 0$, $p(\text{BM}|\text{D}) = 0.19$, $p(\text{LM}|\text{D}) = 0.81$. The posterior phenotypic rate for digitigrady (γ_2) is larger than for plantigrady (γ_1), with the HDI of the two parameters showing little overlap. Indeed, the difference between the two phenotypic rates is credible (HDI = 0.01, 0.317), with digitigrady being larger than plantigrady (*pp* = 0.996).

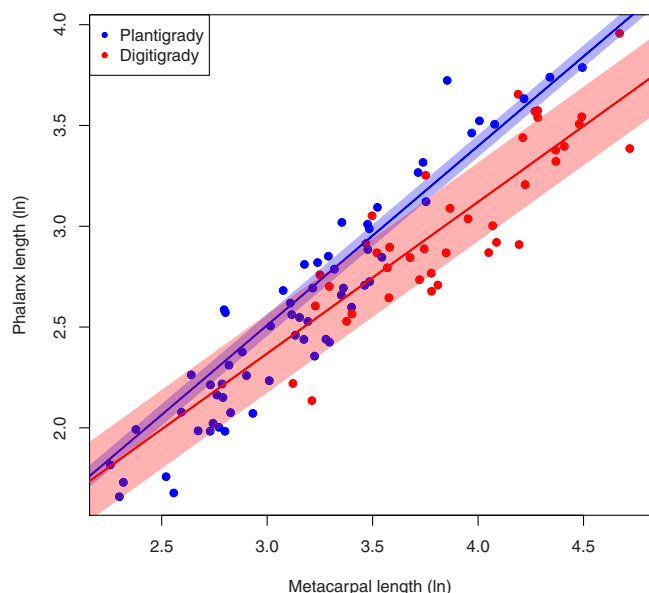


Figure 2. Phalanx versus metatarsal length (both originally reported in mm) for species of the order Carnivora. Posture as indicated in the legend. Approximated regression lines are plotted in blue for plantigrady and in red for digitigrady, with transparencies showing the scatter of the respective phenotypic rate. Posterior regression lines overlap near the intercept and diverge for high values of the covariate, consistent with a possible interaction (β_3) but an absent group effect (β_2). The scatter for the digitigrady line (γ_2) is wider than for the plantigrady line (γ_1), reflecting the burst in phenotypic disparity.

there is also a coevolutionary pattern involving carnivorans with long metacarpals. Individual models contrasted through DIC provide consistent results but more so in regards to phenotypic disparity and phylogenetic effects, as the coevolutionary pattern is less

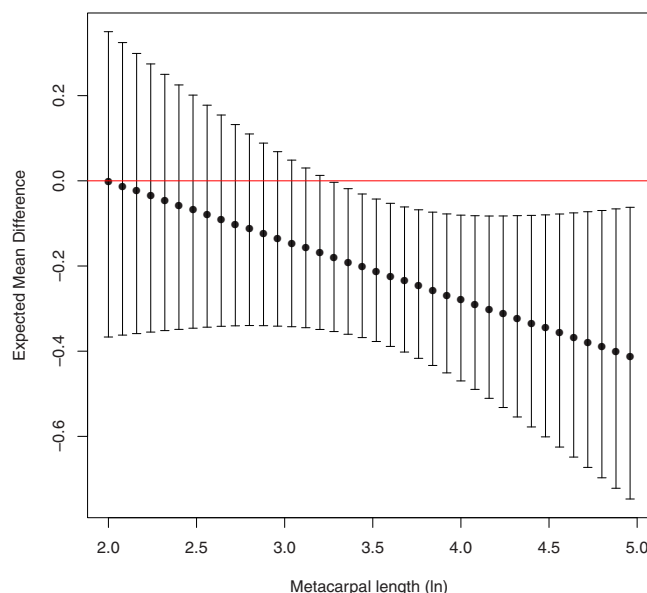


Figure 3. Posterior distribution of the effect of digitigrady over plantigrady in relative phalanx length for species the order Carnivora: $\beta_2 + \beta_3 * \ln(\text{metacarpal length})$. The differences (dots) are credible (95% credible intervals excluding the red horizontal line representing zero) and negative (below the red horizontal line) when metacarpal length (\ln) is longer than 3.3, approximately.

conclusive under this approach (Supporting Information D). All the estimates are reliable as the chains were well behaved, showing moderate autocorrelation, good mixing, stationarity, large effective sample sizes ($N_e > 1000$), as well as passing both stationarity and half-width tests for all parameters (Supporting Information E). In addition, our validation procedures indicated that priors (Supporting Information A), statistical inferences (Supporting

Information B), and predictive performance (Supporting Information C) behaved as expected.

The pattern under maximum likelihood, which relies exclusively on the consensus, is less clear (Table 2). Two models receive substantial support ($\Delta AIC_c < 2$): LM with equal and unequal rates. In terms of the type of model, both maximum likelihood and Bayesian approaches provide consistent results by giving more relevance to phylogenetic history (λ estimates under both approaches are also high). However, the two highly supported models represent a challenge to conclude anything about the relationship between ankle posture and forefoot proportions. One model (single-rate LM) suggests that there is a difference in slopes (significant interaction) but not in evolutionary rates (γ estimated for the entire tree). The other model (multirate LM) suggests that there is a difference in evolutionary rates (γ_1 for plantigrady and γ_2 for digitigrady), but not in slopes (nonsignificant interaction). Thus, something is clearly happening with ankle posture, but we cannot specify what under maximum likelihood.

Discussion

The approach outlined here facilitates the exploration of shifts in both trait relationships and phenotypic rates when using a mixture of continuous and discrete variables. It could be used to study the effect of, for example, behavioral complexity on brain size (Maximino 2009), reproductive mode on pelvic girdle morphology (Oufiero and Gartner 2014), or diet on craniodental features (Muñoz-Durán and Fuentes 2012). Its Bayesian formulation allows incorporating different sources of uncertainty in the analysis, as described in more detail below.

Assessing uncertainty in parameter estimation is important in comparative analyses, but usually involves conducting extra procedures to obtain confidence intervals for either regression parameters (e.g., Garland and Ives 2000) or phylogenetic signal (e.g., Boettiger et al. 2012). Such split procedures are not necessary under Bayesian techniques that are concerned with the estimation of distribution of parameters and, therefore, provide point estimates with associated credible ranges simultaneously (Steel and Kammeyer-Mueller 2008; Blomberg et al. 2012). Measures of uncertainty are important for comparative data because parameters derived from phylogenies (such as λ) are noisy, and our ability to estimate them can affect the evolutionary conclusions that result from the analyses (Boettiger et al. 2012). Moreover, the intervals used in our Bayesian approach (HDI in Table 1) include information about the shape of the distribution that is lacking in the classical confidence interval (Kruschke and Liddell 2016).

The distributional properties of parameters conferred clear benefits in the analysis of the carnivoran data. Parameters obtained under individual models obscured the detection a coevolutionary trend, as both group (β_2) and interaction effects (β_3)

were inconclusive under both maximum likelihood and Bayesian approaches (Tables 1 and 2, and Supporting Information D). But this way of thinking does not necessarily reflect the biological significance of the regression coefficients (Burnham and Anderson 2002; O'Meara et al. 2006), which can be better studied by assessing the broader importance of variables (Freckleton 2009). The posterior probability of the interaction (Table 1) as well as the significance of the parameter under single-rate LM (Table 2) suggested a coevolutionary trend, albeit uncertain. Exploration of the parameters using the joint posterior distribution of regression coefficients through the expected mean differences helped to understand the nature of such uncertainty. Differences in phalanx length do in fact exist, but only for carnivorans in a specific range of metacarpal lengths (Figs. 2 and 3).

The results from our example show support for a coevolutionary trend maximizing running performance by mixing digitigrady with decreased relative phalanx length (Van Valkenburgh 1985), but indicate that such trend does not apply for carnivorans with short metacarpals (Fig. 2). The uncertainty for low values of the covariate was more reflective of the lack of digitigrades with short metacarpals than a weak effect of posture on relative phalanx length (Fig. 3). Most likely this is a scaling phenomenon, as carnivorans with short metacarpals will be also the smallest ones in the order. Small carnivorans generally engage in a wide range of activities (e.g., digging, climbing, swimming) without specialized limb modifications, giving them more overall dexterity in the limbs (Andersson 2003; Andersson 2004; Meloro et al. 2013). But the cost of supporting their weight affects locomotor efficiency, so it is expected that large animals will morphologically respond differently to the problem of running economically (Kram and Taylor 1990).

Now, if speed was the only functional demand acting on limbs, we could expect a constrained trend (Martín-Serra et al. 2015) in which increases in metacarpal length are always accompanied by proportional shortening on phalanx length. However, the phenotypic rate comparison shows that this is not the case. Shifts in posture are also associated with increased residual variance in phalanx length (Table 1, Fig. 2). This indicates that digitigrade carnivorans exhibit more variability in forefoot surface area, which makes sense because morphological features enhancing cursorial lifestyles compromise other locomotor strategies such as borrowing, climbing, and swimming (Andersson 2004; Polly 2007; Meachen-Samuels and Van Valkenburgh 2009). Among digitigrades, there are not only cursorial pursuit hunters (e.g., African wild dog, gray wolf, spotted hyaena), but also climbers (e.g., clouded leopard, snow leopard, gray fox), generalist hunters that pounce or chase prey (e.g., large Indian civet, African civet, crab-eating fox), ambush hunters that grapple prey (e.g., lion, puma, serval), and carnivorans that walk on snow (e.g., Canada lynx), where more variability in forefoot surface area can be

Table 2. Results under maximum likelihood models.

Model	β_0	β_1	β_2	β_3	γ_1	γ_2	g	λ	AIC _c
IID	−0.66 (0.149)*	1.02 (0.047)*	0.57 (0.303)	−0.22 (0.082)*	0.03	—	—	—	−47.08
	−0.66 (0.137)*	1.02 (0.043)*	0.57 (0.323)	−0.22 (0.086)*	0.03	0.04	—	—	−46.62
BM	−0.13 (0.283)	0.88 (0.072)*	0.28 (0.353)	−0.14 (0.101)	0.16	—	—	—	−35.31
	−0.13 (0.193)	0.88 (0.048)*	0.21 (0.383)	−0.12 (0.106)	0.07	0.32	2.42×10^{-8}	—	−58.95
LM	−0.31 (0.206)	0.93 (0.059)*	0.49 (0.312)	−0.20 (0.088)*	0.05	—	—	0.7	−61.95
	−0.21 (0.191)	0.90 (0.052)*	0.30 (0.345)	−0.15 (0.096)	0.05	0.19	0.12	0.87	−60.8

Parameter estimates (descriptors as in Table 1 except for g which is presented below) and model fit (AIC_c) under maximum likelihood are presented assuming different models (IID, BM, LM). For each type of model, a single-rate version is presented (top row) with a single parameter estimated for the entire tree (γ), and a multi-rate version is presented (bottom row) with different phenotypic rates for plantigrady (γ_1) and digitigrady (γ_2). Standard errors for regression coefficients shown in parenthesis; an asterisk indicates whether the parameter estimate is significant under a t -test at the 0.05 level. The proportions of transitional branches (g) tend to be low (especially for BM), suggesting that shifts to digitigrady happen at the end of the branch.

advantageous (Van Valkenburgh 1985; Meachen-Samuels and Van Valkenburgh 2009; Samuels et al. 2013). By increasing the range of relative phalanx lengths, digitigrades can avoid locomotor constraints while keeping the advantages of increased speed.

The results discussed above rely on a distribution of trees, rather than on any given topology with a specific set of branch lengths. In many comparative analyses, phenotypic data have to be excluded due to poor knowledge of the phylogenetic relationships of the taxa. A good example of this comes from Andersson (2004), who had to exclude the herpestids and the fossa from his functional study on carnivoran elbow-joint morphology. Indeed, the nodes associated with these lineages (as well as others) in our phylogenetic tree exhibited low support (Fig. 1), thus reflecting phylogenetic uncertainty. Still, none of these lineages had to be removed from our analysis because the Bayesian approach integrated over a distribution of trees that reflected such uncertain phylogenetic placement.

Accounting for phylogenetic uncertainty does not mean that we do not need to seek robust phylogenetic trees, but instead relieves the necessity of assuming that a single tree is entirely correct (Harvey and Pagel 1991; Huelsenbeck et al. 2000). The consensus tree used for the carnivoran example provides a summary of the empirical prior of trees, but does not reflect the uncertainty embedded in the phylogenetic hypotheses (Pagel and Lutzoni 2002). Ignoring such uncertainty can lead to bias (de Villemereuil et al. 2012), which can explain why our Bayesian and maximum likelihood results were not equivalent. The differences between the two approaches could also reflect that the consensus represents a summary of trees, rather than optimally inferred phylogeny in itself (it is thus possible that the maximum likelihood method conducted under a phylogenetic tree estimated in turn by maximum likelihood could reduce the mismatches of the two approaches). Nevertheless, by using the PGLS model in multiple trees of the MCMC sample, the combined results can be interpreted as independent of a given underlying phylogeny, regardless of how it was

obtained (Pagel and Lutzoni 2002; Arnold et al. 2010; Hernández et al. 2013).

Accounting for phylogeny is important, but it is also important to determine the degree with which the residual error correlates with shared patterns of common ancestry (Hernández et al. 2013). We achieved this in our analysis by estimating phylogenetic signal, which also helped in accounting for model selection uncertainty. Model selection should not be conducted with the idea of finding the true model, but with the idea of informing what inferences are supported by the data (Burnham 2002; Burnham and Anderson 2002). Instead of giving us a winning model, our approach informed about a substantial contribution of historical effects on the pattern, admitting that both LM and BM had relevant attributes for explaining the data (especially the former). These attributes are accounted for in parameter exploration through multimodel inference, so that our interpretation about the evolutionary pattern is not conditioned by any specific model (Burnham 2002). By allowing several models to inform inferences, the uncertainty in model selection is taken into account. Note that the relevance of means and variances was mutually exclusive under the best supported models of the maximum likelihood comparison. If we use merely the lowest AIC_c to draw a conclusion, the pattern would be explained in terms of the means (regression coefficients) but not the variances (phenotypic rates). Bayesian models contrasted under DIC offered a somewhat opposite result, by suggesting that the pattern was only strong for the variances. But after assessing the relative contribution of each competing model under the Bayesian MM, and incorporating such contribution in parameter estimation through multimodel inference (Burnham 2002; Burnham and Anderson 2002), we were able to confirm the importance of both means and variances in the pattern. Multimodel inference is also possible under maximum likelihood by using tools like Akaike weights, under which parameter estimates can be weighed based on how specific models fit the data (Burnham and Anderson 2004). In fact, the same type of weights can be computed for

DIC as a Bayesian analogue of AIC (Burnham 2002; Spiegelhalter et al. 2002). The conceptual appeal in MM is in the way it addresses a model comparison problem from a parameter estimation perspective, economizing Markov chains while accounting for uncertainty in model selection and avoiding controversies associated with measures of fit such as DIC (e.g., Brooks 2002; Meng 2002; Smith 2002) and Bayes factors (e.g., Spiegelhalter et al. 2002; Yang and Rannala 2005).

In methods that specify evolutionary regimes for estimating different parameters across a tree (e.g., Hansen 1997; O'Meara et al. 2006; Revell and Collar 2009), errors in ancestral state reconstructions can result in the incorrect partitioning of the phylogenetic covariance matrix (Thomas et al. 2006). For this purpose, we integrate PGLS with SCM, which accounts for the timing and placement of character transitions, as well as the duration of different states in the tree (Bollback 2006). In this way, the uncertainty of the evolutionary regime not only involves character states at nodes, but also the placement of shifts in the internodes. Moreover, the conditions of character change are unique for each branch and transition type (increase or decrease in rate) at no cost for model complexity (such as the g parameter in Table 2). Although using SCM does not involve modeling the mechanisms of rate change (i.e., instantaneous or gradual), this uncertainty is accounted for by including several mappings in the analysis. The density map (Fig. 1) shows that some portions of the tree involve more uncertainty (in purple). Although every single SCM is in essence instantaneous, high uncertainty in the transition of one state to the other in the evolutionary regime will be reflected in long portions of the internode characterized by intermediate probabilities of character states. Therefore, after accounting for uncertainty (by integrating over the distribution mappings that generate those intermediate probabilities), these areas of rate transition will behave more as gradual than instantaneous, because phenotypic rates will be estimated according to the posterior probability of each state of the evolutionary regime in such areas. In this way, SCM relaxes the assumptions about the placement and mechanisms of rate change, without losing phylogenetic information (as the “censored” approach in O'Meara et al. 2006).

The integration of phylogenetic and mapping samples with tests of trait relationships and phenotypic disparity is also available through the estimation of Bayesian evolutionary rate matrices (Caetano and Harmon 2018b; Caetano and Harmon 2017), implemented in the R package *ratematrix* (Caetano and Harmon 2018a). This approach could have been used to explore the carnivoran dataset presented above under a similar concept, albeit with some fundamental differences in interpretation and implementation. We report here the evolutionary rates of forefoot proportions (i.e., residual variance of regression lines between phalanx and metatarsal lengths), whereas the rates reported by

ratematrix refer to each independent continuous variable (i.e., different rates of trait evolution for phalanx and metatarsal lengths). Because *ratematrix* operates outside a regression framework and allows the inclusion of many continuous variables, its multivariate exploratory capabilities are advantageous. However, the regression framework outlined here is nevertheless useful in situations where a specific set of continuous variables is relevant as a scaling or other type of correction factor, and where rates are calculated as a function of the entire model describing relationships between variables. A good example of this difference comes about with the interpretation of regression coefficients. Although the slopes reported here can be conceptually compared to the covariances reported by *ratematrix*, no true equivalent exists in the latter for intercept estimates (e.g., β_0 and β_2 in the example), which can be relevant in allometric studies (Albrecht et al. 1993; Packard and Boardman 1999; Uyeda et al. 2017). Also, the dependency of *ratematrix* on BM is relaxed here by estimating the degree of phylogenetic relatedness (λ) and implementing a MM that can contrast BM with other alternatives.

The way in which *ratematrix* and our approach deal with mapping uncertainty holds some promise for other methodologies relying on the specification of evolutionary regimes. For example, the adaptation-inertia framework (Hansen 1997; Escudero et al. 2012) has a strong theoretical foundation that makes it ideal to address questions about natural selection (Ho and Ané 2014). However, a known weakness of this framework is its strong dependence on fixed ancestral state reconstructions to specify selective regimes (Hansen 2014). In a similar way as Barton and Venditti (2014) solve the issue of the unknown location of the evolutionary rate changes, reversible-jump MCMC approaches have been formulated to avoid the specification of selective regimes for Ornstein-Uhlenbeck (OU) models (Uyeda and Harmon 2014; Uyeda et al. 2017). These approaches are useful in that they estimate the location of shifts, but are less helpful for testing specific hypothetical regimes (e.g., migration), or to account for the prevalent problem of phylogenetic uncertainty (Harvey and Pagel 1991; Huelsenbeck et al. 2000; Venditti et al. 2011). SCM can be integrated with the adaptation-inertia framework to account for phylogenetic uncertainty and relax the strong dependence on specific ancestral state reconstructions. By incorporating unequal stationary variances in the lines of the approach presented here, a powerful tool to explore both adaptation and phenotypic disparity would open new opportunities to address important questions in evolutionary biology.

AUTHOR CONTRIBUTIONS

JAF-G and PDP compiled the data and designed the analyses. JAF-G and EPM developed the method and conceived the study. All authors conducted the analyses and wrote the manuscript.

ACKNOWLEDGMENTS

We are very grateful to Daniel Manrique-Vallier and John Kruschke whose lessons and coding input made this work possible. Thanks to Elizabeth Housworth and Ellen Ketterson for valuable discussions, as well as Jason Pienaar for assistance with the simulations. We also thank the contributors of PhyloPic (<http://phylopic.org/>) for making available the silhouettes that enriched Fig. 1, particularly Margot Michaud, Steven Traver, David Orr, Birgit Lang, and Mathieu Basille. This material is based on work supported by fellowships to JAF-G from the Colombian COLCIENCIAS (Becas Caldas 497–2009) and the Indiana University Center for the Integrative Study of Animal Behavior. The work was also supported by the U.S. National Science Foundation through Grants IOS-1257562 to EPM and EAR-1338298 and 0843935 to PDP. Authors do not have a conflict of interest to declare.

DATA ARCHIVING

Data package is available on the Dryad Digital Repository: <https://doi.org/10.5061/dryad.9kd51c5ct>

LITERATURE CITED

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19:716–723.
- Akaike, H. 1992. Information theory and an extension of the maximum likelihood principle. Pp. 610–624 in S. Kotz and N. L. Johnson, eds. *Breakthroughs in statistics*. Springer-Verlag, New York, NY.
- Albrecht, G. H., B. R. Gelvin, and S. E. Hartman. 1993. Ratios as a size adjustment in morphometrics. *Am. J. Phys. Anthropol.* 91:441–468.
- Andersson, K. 2003. Locomotor evolution in the Carnivora (Mammalia): evidence from the elbow joint. Pp. 49. Department of Earth Sciences, Historical Geology and Paleontology. Uppsala University, Uppsala, Sweden.
- . 2004. Elbow-joint morphology as a guide to forearm function and foraging behaviour in mammalian carnivores. *Zool. J. Linn. Soc.-Lond.* 142:91–104.
- Arnold, C., L. J. Matthews, and C. L. Nunn. 2010. The 10kTrees website: a new online resource for primate phylogeny. *Evol. Anthropol.* 19:114–118.
- Barton, R. A., and C. Venditti. 2014. Rapid evolution of the cerebellum in humans and other great apes. *Curr. Biol.* 24:2440–2444.
- Beaulieu, J. M., D.-C. Jhwueng, C. Boettiger, and B. C. O'Meara. 2012. Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution* 66:2369–2383.
- Bell, E., B. Andres, and A. Goswami. 2011. Integration and dissociation of limb elements in flying vertebrates: a comparison of pterosaurs, birds and bats. *J. Evol. Biol.* 24:2586–2599.
- Bernardo, J. M. 2003. Bayesian statistics. Pp. 1–46 in R. Viertl, ed. *Probability and statistics*. UNESCO, Oxford, U.K.
- Bernardo, J. M., and A. F. M. Smith. 2009. *Bayesian theory*. John Wiley & Sons, Chichester, U.K.
- Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.
- Blomberg, S. P., J. G. Lefevre, J. A. Wells, and M. Waterhouse. 2012. Independent contrasts and PGLS regression estimators are equivalent. *Syst. Biol.* 61:382–391.
- Boettiger, C., G. Coop, and P. Ralph. 2012. Is your phylogeny informative? Measuring the power of comparative methods. *Evolution* 66:2240–2251.
- Bollback, J. P. 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* 7:88.
- . 2009. SIMMAP Software. Available at <http://www.simmap.com/>. Accessed February 23, 2017.
- Box, G. E. P., and N. R. Draper. 1987. *Empirical model-building and response surfaces*. John Wiley & Sons, Oxford, U.K.
- Brooks, S. P. 2002. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *J. R. Stat. Soc. B Stat. Methodol.* 64:616–618.
- Burnham, K. P. 2002. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *J. R. Stat. Soc. B Stat. Methodol.* 64:629.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference: A practical information-theoretic approach*. Springer-Verlag, New York.
- . 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Method. Res.* 33:261–304.
- Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164:683–695.
- Butler, M. A., T. W. Schoener, and J. B. Losos. 2000. The relationship between sexual size dimorphism and habitat use in Greater Antillean *Anolis* Lizards. *Evolution* 54:259–272.
- Caetano, D., and L. Harmon. 2018a. ratematrix: Bayesian estimation of the evolutionary rate matrix. Ver. 1.1.
- Caetano, D. S., and L. J. Harmon. 2017. ratematrix: an R package for studying evolutionary integration among several traits on phylogenetic trees. *Methods Ecol. Evol.* 8:1920–1927.
- . 2018b. Estimating correlated rates of trait evolution with uncertainty. *Syst. Biol.* 68:412–429.
- Collar, D. C., J. A. Schulte, B. C. O'Meara, and J. B. Losos. 2010. Habitat use affects morphological diversification in dragon lizards. *J. Evol. Biol.* 23:1033–1049.
- Currie, T. E., and A. Meade. 2014. Keeping yourself updated: Bayesian approaches in phylogenetic comparative methods with a focus on Markov Chain models of discrete character evolution. Pp. 263–286 in L. Z. Garamszegi, ed. *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*. Springer, Berlin.
- de Villemeureuil, P., J. A. Wells, R. D. Edwards, and S. P. Blomberg. 2012. Bayesian models for comparative analysis integrating phylogenetic uncertainty. *BMC Evol. Biol.* 12:102.
- Díaz-Uriarte, R., and T. Garland. 1998. Effects of branch length errors on the performance of phylogenetically independent contrasts. *Syst. Biol.* 47:654–672.
- Donoghue, M. J., and D. D. Ackerly. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philos. T. Roy. Soc. B.* 351:1241–1249.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Eastman, J. M., M. E. Alfaro, P. Joyce, A. L. Hipp, and L. J. Harmon. 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* 65:3578–3589.
- Escudero, M., A. L. Hipp, T. F. Hansen, K. L. Voje, and M. Luceño. 2012. Selection and inertia in the evolution of holocentric chromosomes in sedges (*Carex*, Cyperaceae). *New Phytol.* 195:237–247.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125:1–15.
- Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* 19:445–471.
- Freckleton, R. P. 2002. On the misuse of residuals in ecology: regression of residuals vs. multiple regression. *J. Anim. Ecol.* 71:542–545.
- . 2009. The seven deadly sins of comparative analysis. *J. Evolution. Biol.* 22:1367–1375.
- Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* 160:712–726.
- Fuentes-G., J. A., E. A. Housworth, A. Weber, and E. P. Martins. 2016. Phylogenetic ANCOVA: estimating changes in evolutionary rates as well as relationships between traits. *Am. Nat.* 188:615–627.

- Fuentes-G., J. A., P. D. Polly, and E. P. Martins. 2019. Data from "A Bayesian extension of phylogenetic generalized least squares (PGLS): incorporating uncertainty in the comparative study of trait relationships and evolutionary rates". Dryad, <https://doi.org/10.5061/dryad.9kd51c5ct>. Available at <https://doi.org/10.5061/dryad.9kd51c5ct>. Accessed December 25, 2019.
- Garamszegi, L. Z. ed. 2014. Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice. Springer, Berlin.
- García-Berthou, E. 2001. On the misuse of residuals in ecology: testing regression residuals vs. the analysis of covariance. *J. Anim. Ecol.* 70:708–711.
- Garland, T., and A. R. Ives. 2000. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *Am. Nat.* 155:346–364.
- Garland, T., and C. M. Janis. 1993. Does metatarsal/femur ratio predict maximal running speed in cursorial mammals? *J. Zool.* 229:133–151.
- Garland, T., A. F. Bennett, and E. L. Rezende. 2005. Phylogenetic approaches in comparative physiology. *J. Exp. Biol.* 208:3015–3035.
- Garland, T., A. W. Dickerman, C. M. Janis, and J. A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* 42:265–292.
- Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* 1:515–534.
- Grafen, A. 1989. The phylogenetic regression. *Philos. T. Roy. Soc. B.* 326:119–157.
- Green, P. J. 2003. Trans-dimensional Markov chain Monte Carlo. Pp. 179–198 in P. J. Green, N. L. Hjort, and S. Richardson, eds. Highly structured stochastic systems. Oxford Univ. Press, Oxford, U.K.
- Gregory, W. K. 1912. Notes on the principles of quadrupedal locomotion and on the mechanism of the limbs in hoofed animals. *Ann. NY Acad. Sci.* 22:267–294.
- Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51:1341–1351.
- . 2014. Use and misuse of comparative methods in the study of adaptation. Pp. 351–379 in L. Z. Garamszegi, ed. Modern phylogenetic comparative methods and their application in evolutionary biology. Springer, Berlin.
- Hansen, T. F., and K. Bartoszek. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. *Syst. Biol.* 61:413–425.
- Hansen, T. F., and E. P. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50:1404–1417.
- Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford, U.K.
- Heidelberger, P., and P. D. Welch. 1981. A spectral method for confidence interval generation and run length control in simulations. *Commun. ACM* 24:233–245.
- . 1983. Simulation run length control in the presence of an initial transient. *Oper. Res.* 31:1109–1144.
- Hernández, C. E., E. Rodríguez-Serrano, J. Avaria-Llatureo, O. Inostroza-Michael, B. Morales-Pallero, D. Boric-Bargetto, C. B. Canales-Aguirre, P. A. Marquet, and A. Meade. 2013. Using phylogenetic information and the comparative method to evaluate hypotheses in macroecology. *Methods Ecol. Evol.* 4:401–415.
- Ho, L. S. T., and C. Ané. 2014. Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. *Methods Ecol. Evol.* 5:1133–1146.
- Hoff, P. D. 2009. First course in Bayesian statistical methods. Springer, New York, NY.
- Hsiang, A. Y., L. E. Elder, and P. M. Hull. 2016. Towards a morphological metric of assemblage dynamics in the fossil record: a test case using planktonic foraminifera. *Philos. T. Roy. Soc. B.* 371:1–24.
- Huelsenbeck, J. P., and B. Rannala. 2003. Detecting correlation between characters in a comparative analysis with uncertain phylogeny. *Evolution* 57:1237–1247.
- Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
- Huelsenbeck, J. P., B. Rannala, and J. P. Masly. 2000. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* 288:2349–2350.
- Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic mapping of morphological characters. *Syst. Biol.* 52:131–158.
- Jasieński, M., and F. A. Bazzaz. 1999. The fallacy of ratios and the testability of models in biology. *Oikos* 84:321–326.
- Kram, R., and C. R. Taylor. 1990. Energetics of running: a new perspective. *Nature* 346:265–267.
- Kruschke, J. K. 2011. Doing Bayesian data analysis. A tutorial with R and BUGS. Academic Press/Elsevier, Burlington, MA.
- Kruschke, J. K., and T. M. Liddell. 2016. The Bayesian new statistics: hypothesis testing, estimation, meta-analysis, and planning from a Bayesian perspective. *Psychon Bull Rev* 25:178–206.
- Lavin, S. R., W. H. Karasov, A. R. Ives, K. M. Middleton, and T. Garland, Jr. 2008. Morphometrics of the avian small intestine compared with that of nonflying mammals: a phylogenetic approach. *Physiol. Biochem. Zool.* 81:526–550.
- Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–925.
- Lindenfors, P. 2006. A method for calculating means and variances of comparative data for use in a phylogenetic analysis of variance. *Evol. Ecol. Res.* 8:975–995.
- Link, W. A., and M. J. Eaton. 2012. On thinning of chains in MCMC. *Methods Ecol. Evol.* 3:112–115.
- Lynch, M. 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45:1065–1080.
- Martín-Serra, A., B. Figueirido, and P. Palmqvist. 2014. A three-dimensional analysis of the morphological evolution and locomotor behaviour of the carnivoran hind limb. *BMC Evol. Biol.* 14:13.
- Martín-Serra, A., B. Figueirido, J. A. Pérez-Claros, and P. Palmqvist. 2015. Patterns of morphological integration in the appendicular skeleton of mammalian carnivores. *Evolution* 69:321–340.
- Martins, E. P. 1994. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.* 144:193–209.
- . 1996. Conducting phylogenetic comparative studies when the phylogeny is not known. *Evolution* 50:12–22.
- Martins, E. P., and T. F. Hansen. 1996. A microevolutionary link between phylogenies and comparative data. Pp. 273–288 in P. H. Harvey, A. J. Leigh Brown, J. Maynard Smith, and S. Nee, eds. New uses for new phylogenies. Oxford Univ. Press, Oxford, U.K.
- Maximino, C. 2009. A quantitative test of the thermogenesis hypothesis of cetacean brain evolution, using phylogenetic comparative methods. *Mar. Freshw. Behav. Phys.* 42:1–17.
- Meachen-Samuels, J., and B. Van Valkenburgh. 2009. Forelimb indicators of prey-size preference in the Felidae. *J. Morphol.* 270:729–744.
- Meloro, C., S. Elton, J. Louys, L. C. Bishop, and P. Ditchfield. 2013. Cats in the forest: predicting habitat adaptations from humerus morphometry in extant and fossil Felidae (Carnivora). *Paleobiology* 39:323–344.
- Meng, X.-L. 2002. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *J. R. Stat. Soc. B Stat. Methodol.* 64:633.
- Muñoz-Durán, J., and J. A. Fuentes. 2012. Evolución de la socialidad, estrategias alimentarias y anatomía craneodental en la subfamilia Caninae. *Acta Biol. Colombiana* 17:173–200.
- Ng, J., and S. D. Smith. 2014. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. *J. Evol. Biol.* 27:2035–2045.

- Nielsen, R. 2002. Mapping mutations on phylogenies. *Syst. Biol.* 51:729–739.
- O'Meara, B. C., C. Ané, M. J. Sanderson, and P. C. Wainwright. 2006. Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60:922–933.
- Oufiero, C. E., and G. E. A. Gartner. 2014. The effect of parity on morphological evolution among phrynosomatid lizards. *J. Evol. Biol.* 27:2559–2567.
- Packard, G. C., and T. J. Boardman. 1999. The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? *Comp. Biochem. Physiol. A.* 122:37–44.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401:877–884.
- Pagel, M., and F. Lutzoni. 2002. Accounting for phylogenetic uncertainty in comparative studies of evolution and adaptation. Pp. 148–161 in M. Lässig and A. Valleriani, eds. *Biological evolution and statistical physics*. Springer, Berlin.
- Pagel, M., and A. Meade. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* 167:808–825.
- Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* 53:673–684.
- Pagel, M. D., and P. H. Harvey. 1988. Recent developments in the analysis of comparative data. *Q. Rev. Biol.* 63:413–440.
- Pennell, M. W., R. G. FitzJohn, W. K. Cornwell, and L. J. Harmon. 2015. Model adequacy and the macroevolution of angiosperm functional traits. *Am. Nat.* 186:E33–E50.
- Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vienna.
- . 2016. *rjags: Bayesian graphical models using MCMC*. Ver. 4–6.
- Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6:7–11.
- Polly, P. D. 2007. Limbs in mammalian evolution. Pp. 245–268 in B. K. Hall, ed. *Fins into limbs: evolution, development, and transformation*. Univ. of Chicago Press, Chicago.
- . 2010. Tiptoeing through the tropics: geographic variation in carnivoran locomotor ecomorphology in relation to environment. Pp. 374–410 in A. Goswami, and A. Friscia, eds. *Carnivoran evolution: new views on phylogeny, form, and function*. Cambridge Univ. Press, Cambridge, U.K.
- R Development Core Team. 2016. *R: a language and environment for statistical computing*. Ver. 3.3.2. R Foundation for Statistical Computing, Vienna, Austria.
- Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9:e89543.
- . 2015. No substitute for real data: a cautionary note on the use of phylogenies from birth–death polytomy resolvers for downstream comparative analyses. *Evolution* 69:3207–3216.
- Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* 4:8.
- Rabosky, D. L., S. C. Donnellan, M. Grundler, and I. J. Lovette. 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. *Syst. Biol.* 63:610–627.
- Raftery, A. E., D. Madigan, and J. A. Hoeting. 1997. Bayesian model averaging for linear regression models. *J. Am. Stat. Assoc.* 92:179–191.
- Revell, L. J. 2012. *phytools: an R package for phylogenetic comparative biology (and other things)*. *Methods Ecol. Evol.* 3:217–223.
- Revell, L. J., and D. C. Collar. 2009. Phylogenetic analysis of the evolutionary correlation using likelihood. *Evolution* 63:1090–1100.
- Revell, L. J., and R. Graham Reynolds. 2012. A new Bayesian method for fitting evolutionary models to comparative data with intraspecific variation. *Evolution* 66:2697–2707.
- Revell, L. J., D. L. Mahler, P. R. Peres-Neto, and B. D. Redelings. 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution* 66:135–146.
- Rezende, E. L., and J. A. F. Diniz-Filho. 2012. Phylogenetic analyses: Comparing species to infer adaptations and physiological mechanisms. *Compr. Physiol.* 2:639–674.
- Ronquist, F. 2004. Bayesian inference of character evolution. *Trends Ecol. Evol.* 19:475–481.
- Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- Samuels, J. X., J. A. Meachen, and S. A. Sakai. 2013. Postcranial morphology and the locomotor habits of living and extinct carnivorans. *J. Morphol.* 274:121–146.
- Schultz, N. G., M. Lough-Stevens, E. Abreu, T. Orr, and M. D. Dean. 2016. The baculum was gained and lost multiple times during mammalian evolution. *Integr. Comp. Biol.* 56:644–656.
- Schultz, T. R., and G. A. Churchill. 1999. The role of subjectivity in reconstructing ancestral character states: a Bayesian approach to unknown rates, states, and transformation asymmetries. *Syst. Biol.* 48:651–664.
- Sisson, S. A. 2005. Transdimensional Markov Chains: a decade of progress and future perspectives. *J. Am. Stat. Assoc.* 100:1077–1089.
- Slater, G. J., and M. W. Pennell. 2014. Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. *Syst. Biol.* 63:293–308.
- Smaers, J. B., and F. J. Rohlf. 2016. Testing species' deviation from allometric predictions using the phylogenetic regression. *Evolution* 70:1145–1149.
- Smith, J. 2002. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *J. R. Stat. Soc. B Stat. Methodol.* 64:619–620.
- Snow, G. 2016. *TeachingDemos: demonstrations for teaching and learning*. Ver. 2.10.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde. 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B Stat. Methodol.* 64:583–639.
- Steel, P. D. G., and J. Kammeyer-Mueller. 2008. Bayesian variance estimation for meta-analysis: quantifying our uncertainty. *Organ. Res. Methods* 11:54–78.
- Thomas, G. H., R. P. Freckleton, and T. Székely. 2006. Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. *P. Roy. Soc. B-Biol. Sci.* 273:1619–1624.
- Thomas, G. H., S. Meiri, and A. B. Phillimore. 2009. Body size diversification in *Anolis*: novel environment and island effects. *Evolution* 63:2017–2030.
- Uyeda, J. C., and L. J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Syst. Biol.* 63:902–918.
- Uyeda, J. C., M. W. Pennell, E. T. Miller, R. Maia, and C. R. McClain. 2017. The evolution of energetic scaling across the vertebrate tree of life. *Am. Nat.* 190:185–199.
- Van Valkenburgh, B. 1985. Locomotor diversity within past and present guilds of large predatory mammals. *Paleobiology* 11:406–428.
- Venables, W. N., and B. D. Ripley. 2002. *Modern applied statistics with S*. Springer, New York, NY.

- Venditti, C., A. Meade, and M. Pagel. 2011. Multiple routes to mammalian diversity. *Nature* 479:393–396.
- Wang, X., and R. H. Tedford. 2007. Evolutionary history of canids. Pp. 3–20 in P. Jensen, ed. *The behavioural biology of dogs*. CABI International, Oxford, U.K.
- Yang, Z., and B. Rannala. 2005. Branch-length prior influences Bayesian posterior probability of phylogeny. *Syst. Biol.* 54:455–470.
- Zúñiga-Vega, J. J., J. A. Fuentes-G., J. G. Zamora-Abrego, U. O. García-Vázquez, A. Nieto-Montes de Oca, and E. P. Martins. 2017. Evolutionary patterns in life-history traits of lizards of the genus *Xenosaurus*. *Herpetol. J.* 27:346–360.

Associate Editor: J. Light
Handling Editor: M. R. Servedio

Supporting Information

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

Figure A1. Posterior distribution of regression coefficients (β_0 , β_1 , β_2 , β_3), evolutionary rates (γ_1 , γ_2), and phylogenetic signal (λ) after running the MCMC without data.

Figure A2. Comparison of posterior distributions for regression coefficients (β_0 , β_1 , β_2 , β_3), evolutionary rates (γ_1 , γ_2), and phylogenetic signal (λ) after sensitivity analysis.

Table A1. Posterior probability for each model of phylogenetic relatedness under different prior specifications.

Table B1. Main simulations results.

Figure C1. Boxplot comparing the predicted means of MM and LM.

Table D1. Results of Bayesian analyses under specific models.

Table D2. Bayesian results under MM with a single rate.

Figure D1. Posterior distribution of the effect of digitigrady over plantigrady in phalanx length for multirate LM, after accounting for metacarpal length.

Figure E1. Autocorrelation function (ACF) showing similarity of observations between time lags for all parameters.

Figure E2. Trace plots showing the evolution of the MCMC samples as time series (same parameters as in Fig. E1).

Table E1. Effective sample sizes (N_e) and results for the stationarity and half-width tests under each of the three chains.