Challenges and advances for phylogenetic comparative models of trait evolution

A Dissertation
Presented in Partial Fulfilment of the Requirements for the
Degree of Doctorate of Philosophy
with a
Major in Bioinformatics and Computational Biology
in the
College of Graduate Studies
University of Idaho
by

Daniel Caetano da Silva

Major Professor: Luke Harmon, Ph.D.
Committee Members: David Tank, Ph.D.; Jack Sullivan, Ph.D.; Paul Hohenlohe, Ph.D.
Department Administrator: Eva Top, Ph.D.

August 2017

Authorization to Submit Dissertation

This dissertation of Daniel Caetano da Silva, submitted for the degree of Doctorate of Philosophy with a major in Bioinformatics and Computational Biology and titled "Challenges and advances for phylogenetic comparative models of trait evolution," has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____Date_____
Luke Harmon, Ph.D.

Committee Members: _____Date_____
David Tank, Ph.D.

_____Date_____
Jack Sullivan, Ph.D.

_____Date_____
Paul Hohenlohe, Ph.D.

Department Administrator: _____Date_____
Eva Top, Ph.D.

Abstract

A phylogenetic tree is a hypothesis of evolutionary relationships among lineages. The branching pattern of such tree tells us the history of species groups derived from their most recent common ancestors. The length of each branch of this tree represent time, a critical information to study evolutionary processes. Such a tree is the starting point for phylogenetic comparative studies, which aim is to use phylogenies to test hypotheses about macroevolution. One of the most interesting questions on macroevolution is how morphological traits evolved to gave origin to the impressive biodiversity we see today. In this dissertation, I use and develop statistical models to ask questions such as the association between morphological differentiation and lineage diversification and patterns of evolutionary correlation among several traits. The present dissertation is divided into four chapters. The first tests whether the diverse coloration patterns associated with Neotropical false-coral snakes of the family Dipsadidae have a positive effect on the rates of diversification of the group. In the second chapter I focused on evolutionary modularity and developed a novel Bayesian approach to study patters of evolutionary correlation among several continuous traits using Markov-chain Monte Carlo. The third chapter is composed by the implementation of the new method described on Chapter 2 as an R package named 'ratematrix'. In this same chapter, I developed and implemented an extension of Felsenstein's pruning algorithm applied when multiple independent multivariate Brownian-motion rate regimes are fitted to the same phylogenetic tree. The fourth chapter introduces the use of mathematical functions to describe the heterogeneity in the rate of evolution of a continuous trait across the branches of a phylogeny as predicted by a different (also continuous) trait. In summary, in this dissertation I visit different topics associated with phylogenetic comparative models of trait evolution. Each chapter focus on a current challenge in phylogenetic comparative studies; the association of trait with rates of diversification on chapter one, simultaneous study of several traits on chapters two and three and the correlation between rates of evolution and a potential predictor trait on chapter four.

Acknowledgements

It is not easy to forget the process of moving from Sao Paulo (Brazil) to Moscow (USA). Such a decision represented for me the best move I could ever make in my career. For the American reader I want to ask you to look around and imagine that your profession of choice is conducted in another language. Now also try to think that the next big conference in your field will be overseas, the next also, and the other too. You now also have to worry if a reviewer will ask whether a native speaker reviewed your writing before sending your manuscript to review. That is reality on the majority of countries around the world that do science and everyday concerns for both students and faculty. So, yes, coming to US to pursue my doctorate degree was a big thing.

Fortunately, all the anxiety that I felt after realizing the size of Moscow, ID, went away when I started working with the fantastic community of evolutionary biologists at the University of Idaho. The frequent interactions with students and faculty makes this place warm to the heart, more than a necessity given the extreme low temperatures during the dark afternoons of winter. The constant open conversation and clear focus on scientific arguments rather than academic hierarchies makes any student feel as part of something bigger. This environment makes us want to grow and to develop as far as we can. The whole department should be proud of the energy that it transmits.

This is also a place of routine, but in a good way. The schedule of the week is quite clear: work, PEES, IBEST Lunch, meet the speaker, and, of course, the Friday: PURGE, check email, Biology seminar, and... is already 5PM, so beers. I am guilty of complaints about so many things to do sometimes. However, I am sure that I am going to miss all this so much! Among all the things, PURGE will have the longest mark on me, both as a person and as a professional. The opportunity of discussing published articles with fellow students and professors weekly is just fantastic! I learned so much with everyone. Learned that it is good to know and even better to know you don't know. Learned to listen and make yourself listened. And, of course, the occasional practice of talking about a manuscript you failed to read will, I am sure, be handy some day.

The colleagues, collaborators and friends that made me company on this five years journey also deserve many thanks. Professors Luke Harmon, Dave Tank and Jack Sullivan were always there to discuss any topic and for this I am hugely grateful. The Harmon lab, my home, is a group of big ideas and bold projects. I learned a lot from all lab-mates. Thanks to Matt Pennell for helping me with all the first steps as a graduate student and to inspire (also provoke) me to pursue big accomplishments. Thanks Denim Jochimsen for helping me with all the snake-things. Many thanks to Rafael Maia and Eliot Miller, was super cool to

have you in the lab and bring much more diversity of topics to the table. Special thanks to Rosana Zenil-Ferguson for raising the level of the statistical work of everyone in the lab and for the patience to explain all the things. Another super special thanks to Josef Uyeda. I am happy to say that Josef is a model of the scientist I hope one day to be and I am super proud of having shared a lot of ideas and beers with you. Finally, my big thanks to Luke Harmon! I am still impressed with how luck I was to have the opportunity to join your team. Being your student was an amazing experience and a very important turning point in my academic career. I will aim high and do all the things Luke, I promise!

Many thanks to the Tank lab, my second home, and all the plant people that effectively help the Harmon lab to keep things real. Since Dave Tank moved to the Biology Department it has been great to have him around. You helped me a lot Dave, thanks! Thanks to Hannah Marx and Simon Uribe-Convers, both PhDs for some time already. Many thanks also to Diego Morales-Briones and Sarah Jacobs for making me company in the office, discussing all the science and also having all the fun. Many thanks also to Megan Ruffley, Ian Gilman and Sebastian Mortimer, you are all great and wish all the good things for your next steps!

Thanks also for the members of my committee Luke Harmon, Dave Tank, Jack Sullivan and Paul Hohenlohe for all the help with brainstorming ideas and career advices.

Of course, many thanks to the Moscow crew! Too many names to list here but, hey, you all know who you are. See you at BBBBBBBBBB!!

Finally I want to thanks all the funding agencies that made this work possible. Coordenação de Pessoal de Nível Superior (CAPES: 1093/12-6) supported me with a full scholarship for four years and the Bioinformatics and Computational Biology program of the University of Idaho awarded me with a scholarship for an additional year. All the computational resources were provided by a grant from the National Science Foundation (award DEB-1208912).

Dedication

Para meus pais por todo o esforço para me dar muito mais do que eles tiveram na vida.
Tudo o que consegui até hoje foi por causa do seu apoio, paciência e amor.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: HAVE CORAL SNAKE MIMICS DIVERSIFIED MORE THAN NON-MIMICS?

1.1 Abstract

Dipsadidae is one of the most diversified family of snakes, composed of species showing an impressive variety of color patterns. Some species are cryptic whereas others have contrasting patterns comprised by bright colors alternated with darker shades, including particular combinations of vivid colors characteristic of coral snakes (Elapidae). Species with such patterns are thought to be mimics of coral snakes based on their color pattern similarity, predator avoidance of such patterns in field experiments, and the geographical concordance between models and mimics. Here we test whether color patterns associated with coral snake mimicry and contrasting color patterns in general influenced the diversification dynamics of the group. We compile a large database of color patterns with color descriptions for about 80% of the known diversity of the group (594 species). We used trait-dependent diversification models along with extensive simulations to deal with the recently described statistical bias associated with such methods. We also tested whether color patterns are associated with trait-independent estimates of diversification. Despite the apparent survival advantage associated with coral snake mimicry, we show that there is no detectable influence of color types in the dynamics of diversification in Dipsadidae and discuss insights into the potential functions of color patterns.

1.2 Introduction

Colors play an important role in avoiding predation. Patterns similar to the background environment make prey difficult for the predators to detect and recognize (Merilaita and Lind, 2005; Stevens and Merilaita, 2009). On the other hand, bright and contrasting colors displayed by unpalatable, toxic or venomous animals (i.e., aposematic patterns) serve as warning signals that are often avoided by visually oriented predators (Wallace, 1867; Mappes et al., 2005; Speed and Ruxton, 2005). However, such conspicuous colors can also be displayed by mimics, which gain protection by deceiving predators that avoid their false warning signals. Strong evidence from field experiments shows that mimicry of warning signals decreases predation pressure when compared to cryptic color patterns (Jeffords et al., 1979; Brodie III, 1993; Brodie III and Janzen, 1995; Pfennig et al., 2001; Pinheiro, 2011; Pfennig et al., 2015). Such reduction in predation pressure may also have positive impacts

on habitat use by aposematic lineages and their mimics. Cryptic animals are to some degree restricted to backgrounds which their color patterns match and may only be active at certain times because movement is often antithetic to good crypsis (Speed et al., 2010; Stevens and Ruxton, 2012). In contrast, such restrictions may be weaker in aposematic or mimic lineages, which could promote more opportunities to exploit habitat resources (Speed et al., 2010).

In contrast with aposematism, the survival advantage of Batesian mimicry is dependent on the relationship between the model and the mimetic organism because predators need to associate the unpalatability or hazard of the model with the warning signals of the deceiver. Once this association is broken, a mimicry breakdown occurs and the mimic phenotype might become maladaptive since warning signals can make individuals more conspicuous to predators (Mallet and Joron, 1999; Pfennig et al., 2001, 2015). Mimicry breakdown can be caused by allopatry between mimic and model populations as a result of population expansion of the mimic or local extinction of the model (Pfennig and Mullen, 2010). Allopatric mimics are conspicuous to naïve predators that might not avoid their deceptive warning signals and this may result in higher predation rates and eventual extinction of the mimic population (Pfennig et al., 2015). On the other hand, population expansion or migration of mimics can create opportunities for local adaptation to novel aposematic models. This process could result in selection against intermediate hybrids followed by decreased gene flow among populations and eventually promote reproductive isolation (Mallet and Joron, 1999; Pfennig et al., 2015). Over longer time scales such processes might have a positive effect on rates of diversification of mimetic lineages. Previous studies show that aposematic lineages are more species-rich than cryptic ones (Santos et al., 2003; Przeczek et al., 2008), suggesting that the evolution of the aposematic condition may even represent a key innovation (Speed et al., 2010). This key innovation hypothesis could be extended to mimicry; however, the potential effects of mimicry evolution on lineage diversification have yet to be investigated.

Among snakes, groups of relatively harmless or mildly venomous species showing color patterns similar to those of venomous coral snakes (Elapidae) have instigated a long debate on whether such patterns are mimetic (see a comprehensive review in Pough, 1988). Reports often rely on the similarity of color patterns between mimics and models to argue in favor of mimicry relationships (Dunn, 1954; Hecht and Marien, 1956; Greene and McDiarmid, 1981; Savage and Slowinski, 1992). Additional evidence come from parallel geographic variation of coral snakes and their putative mimics (e. g., Hecht and Marien, 1956; Zweifel, 1960; Greene and McDiarmid, 1981; Marques and Puorto, 1991; Rabosky et al., 2016) and from field studies using replicas of coral snakes and other similar color patterns (Smith, 1975; Brodie III, 1993; Brodie III and Janzen, 1995; Hinman et al., 1997; Pfennig et al., 2001; Buasso et al., 2006).

Some authors pointed to the possibility that contrasting colors, including the stereotypical banded pattern observed in almost all coral snakes, could serve a disruptive function (Gadow, 1908; Thayer, 1909; Dunn, 1954; Brattstrom, 1955). Those reports suggested that the alternate pattern of bands could blend to the background environment and break the outline of the snake body, making recognition by visually oriented predators difficult. Recently, Titcomb and colleagues 2014 showed that the contrasting ringed pattern of coral snake mimics can create an illusory effect when the individuals are moving fast. The effect, called flicker-fusion, can give advantage to snakes against avian predators independent of mimicry. Despite its protective effect, the plausible disruptive function of the contrasting bands do not invalidate the existence of a mimicry complex between elapids and snakes from other families, since the same color pattern can perform both functions (Titcomb et al., 2014).

The family Dipsadidae (Zaher et al., 2009, *sensu* Grazziotin et al., 2012) is a diverse group of snakes, with ca. 700 species occurring from Central to South America (Grazziotin et al., 2012; Uetz et al., 2014), and is characterized by an impressive variety of color patterns (see Martins and Oliveira, 1998, for some examples). Some dipsadids have color patterns similar to those of coral snakes, and have long been suggested as cases of mimicry of New World coral snakes (family Elapidae; Wallace, 1867; Greene and McDiarmid, 1981; Sazima and Abe, 1991; Savage and Slowinski, 1992; Martins and Oliveira, 1993; Pough, 1988; Almeida et al., 2014). The contrasting coloration found in dipsadid snakes always includes bright colors but is not restricted to ringed patterns. In general, species can vary from the coral snake pattern of black, red and yellow rings or bands to a less colorful homogeneous red body with a single black or cream band on the neck (nuchal collar). Besides contrasting color patterns, the family also shows a diverse array of cryptic color patterns, characterized by blotches and shades of brown, gray, or green. Included in the latter are species whose dorsum is cryptic and whose venter has a plain bright color and even a coral snake pattern. Mimetic and cryptic patterns can be found both within and among genera and make dipsadid snakes an ideal study system to investigate the possible effects of such distinct color types on macroevolutionary patterns.

Herein we test whether distinct color patterns have an influence on the diversification of the family Dipsadidae. We investigate whether color patterns similar to coral snakes (and contrasting color patterns in general) show diverging macroevolutionary patterns when compared to non-mimic and cryptic lineages, respectively. We show that there is no detectable influence of supposedly mimic or contrasting color patterns in the dynamics of diversification.

1.3 Materials and Methods

1.3.1 Phylogenetic reconstruction

We used sequence data for Dipsadidae and outgroup species previously analyzed by Grazziotin and colleagues (2012, see GenBank accession numbers in their Appendix S1). We aligned sequences using MAFFT (Katoh et al., 2005) under the G-INS-i strategy and selected models of molecular evolution for each of the eight gene sequences using a decision theory framework in DT-ModSel (Minin et al., 2003). We concatenated the alignments and set four partitions; one partition for each nuclear gene (bdnf, c-mos, and rag2) and a single partition with the mitochondrial genes (12S, 16S, cytb, nd2, and nd4). We used phyutility (Smith and Dunn, 2008) to trim down all sites with 75% or more missing data and inferred a Maximum Likelihood (ML) tree using GARLI 2.0 (Zwickl, 2011). We used the resulting ML phylogeny as the starting tree for three independent searches in BEAST 1.8 (Drummond et al., 2012) for 270 million generations with a thinning interval of 1500 generations each. Since there are sequences available for only few species of each genera we set an incomplete sampling birth-death tree prior (Stadler, 2009) and an uncorrelated relaxed clock model to estimate relative branching times. We checked each run for convergence using Tracer 1.6 (Drummond et al., 2012) and excluded 50% of the posterior chain as burnin. We then combined the posterior from the three BEAST searches and randomly sampled 100 trees, rescaled all trees to a total depth of 1, and retained only one randomly selected species of each genera while pruning the rest. When the original tree had paraphyletic genera we selected the most inclusive monophyletic clade representing each group and kept a single species to represent each of those genera. Then, we used the resulting pool of genus-level phylogenetic trees to perform all subsequent comparative analyses. The 100 sampled trees and the BEAST xml file comprising the data matrix, selected models of molecular evolution, starting tree and prior parameters is available in FigShare (http://dx.doi.org/10.6084/m9.figshare.831493). We also deposited the configuration and log files for GARLI 2.0. Figures 1.5 and 1.6 show the resulting maximum clade credibility (MCC) tree and respective posterior probability support values.

1.3.2 Color patterns

To understand the evolution of colors and its effect on diversification we compiled a large database of coloration patterns for dipsadid snakes. We searched several information sources such as comprehensive taxonomic reviews (e.g., Downs, 1967), published articles and books containing photographs of identified individuals (e.g., Savage, 2002; Campbell and Lamar,

2004), trusted on-line photo repositories (e.g., CalPhotos - http://calphotos.berkeley.edu/ and Reptile Database – Uetz et al., 2014), photographs of live individuals, and museum specimens. We excluded invalid taxa or names presenting nomenclatural problems that are still appearing in the literature or online databases. We avoided subspecific ranks for coding the currently recognized taxa (with the exception of four subspecies of *Alsophis antillensis*) because terminals in available phylogenies correspond to species only and less than 10% of the members of the family Dipsadidae present valid subspecies to date.

While color diversity makes the family Dipsadidae interesting for studies focusing on the evolution of color patterns such as ours, this is also the most challenging characteristic of the system. Since it is not possible to consider all diversity of color patterns for comparative analyses, we used categories that are directly related to the hypotheses tested. For that, we describe below two distinct classifications of the color patterns, each one including a different important aspect of the biology of the group. We repeated all comparative analyses with each of these color pattern classifications in order to access how distinct interpretations of color diversity in the group affect our macroevolutinary conclusions.

First, we compared coral-mimics with non-mimics. We call coral-mimics species that resemble the color pattern of any New World coral snake species (see Roze, 1996; Campbell and Lamar, 2004). Species included in this category can show the coral-mimic pattern throughout the dorsum (e.g., *Simophis rhinostoma*) or restricted to the anterior portion of the body (e.g., *Pseudoboa coronata*). On the other hand, all species not defined as potential mimics of coral snakes, independent of whether their color pattern was better described as contrasting or cryptic, were included in the category of non-mimics. As a result, the non-mimic category comprise species with cryptic color patterns and others with bright coloration but not resembling any known lineage of New World coral snake. Second, we compared contrasting with cryptic species. We defined as contrasting species that show brightly colored patterns in general, independent of whether the color pattern was similar to those of coral snakes. The category coral-mimic is a subset of the contrasting category; every coral-mimic lineage is among the species defined as contrasting, but the reverse is not true. On the other hand, we defined as cryptic all color patterns lacking contrasting colors. Examples of such patterns are blotches with hues of brown, reddish brown, gray, and other combinations of dark colors. We also considered as cryptic species whose dorsum is homogeneously green, since individuals of those species are usually found among leaves of trees and bushes (e.g., *Uromacer*).

There are three other important aspects of the color patterns in the group that could potentially affect our study; the presence of color polymorphism, ontogenetic changes in coloration and contrasting coloration restricted to the venter of the body. In the case of

species with color polymorphism, some populations may show cryptic patterns occasionally associated to thermoregulation (Tanaka, 2005, but see Lorioux et al., 2008). Alternatively, contrasting colors in polymorphic populations can be due to increasing sexual dichromatism in the course of the reproductive season (Forsman, 1995; Lindell and Forsman, 1996), related to non-selective processes, such as migration and dispersal (King and Lawson, 1995) or genetic drift in local (Brakefield, 1990) or island populations (Bittner and King, 2003). Independent of the potential sources of the polymorphic color patterns we assigned species to the non-mimic and cryptic categories every time a cryptic morph was described among color types. Some Pseudoboini snakes (*sensu* Zaher et al., 2009) show ontogenetic changes in color pattern in which juveniles are brightly colored but become cryptic when adults (Martins and Oliveira, 1998). We included those species into the contrasting category since juveniles correspond to the life stage most threatened by predation (Bonnet et al., 1999) and thus their defensive tactics are fundamental for individuals to reach sexual maturity. Finally, some species show cryptic color patterns in the dorsum but have contrasting patterns restricted to the venter. The distinct patterns in the dorsum and venter are usually associated with a threatening display in which individuals twist the body and expose the bright colors when disturbed (Martins and Oliveira, 1993; Sawaya et al., 2008; Tozetti et al., 2009). Hence, we classified those cases into the contrasting category instead of following their dorsal coloration. We provide additional data with the list of species included in each of these cases. We also repeated comparative analyses with alternative color categories defined to accommodate such cases, but results showed no appreciable difference when compared to the coral-mimic vs. non-mimic or contrasting vs. cryptic categories.

1.3.3 Comparative analyses

In order to test whether color patterns influence the dynamics of diversification of Dipsadidae snakes we performed a series of complimentary phylogenetic comparative analyses using the pool of 100 trees sampled from the posterior distribution and the three distinct color pattern classifications defined above. We used the Binary State Speciation and Extinction model (BiSSE – Maddison et al., 2007; FitzJohn et al., 2009; FitzJohn, 2012) and the Hidden State Speciation and Extinction model (HiSSE – Beaulieu and O'Meara, 2016) followed by model adequacy simulations. Both BiSSE and HiSSE are joint models of trait and trees (Beaulieu and O'Meara, 2016) because model parameters reflect diversification rates conditioned on the state of the traits and trait transitions throughout the tree. In addition, we estimated trait-independent diversification rates using BAMM (Rabosky et al., 2014) and tested whether the distribution of color patterns across dipsadid genera is associated with the diversification

dynamics of the group. We describe details for each analyses below.

Since a species-level tree is not available for the Dipsadidae family, we used the pool of genus-level trees as terminally unresolved trees (*sensu* FitzJohn et al., 2009) to estimate trait-dependent diversification using BiSSE. We used Markov chain Monte Carlo (MCMC) to estimate the posterior distribution of the parameters for the unconstrained BiSSE model (six free parameters) and the constrained model, in which λ and μ are not related to color types (four free parameters). We used an exponential prior distribution with rate parameter equal to 0.3 for all BiSSE parameters and a starting point equal to the maximum likelihood estimate (MLE) for the model. We ran 10,000 generations of MCMC (chain length was based on preliminary analyses), discarded 50% of generations as burnin and checked convergence using the 'coda' package (Plummer et al., 2006). To test whether the full model explained the data better than the trait-independent model we performed model selection using the Bayesian Deviance Information Criteria (DIC – Gelman et al., 2013). In order to account for the behavior of the BiSSE model when its assumptions are violated (Rabosky and Goldberg, 2015; Beaulieu and O'Meara, 2016), we did posterior predictive checks for model adequacy to test whether data simulated under the trait-dependent and trait-independent BiSSE models are similar to the observed data. For each model, we simulated traits using the 'tree.bisse' function in the package 'diversitree' (FitzJohn, 2012) with parameters drawn from the joint posterior distribution resulting from the BiSSE analysis and constraining simulations to have a tree depth equal to 1 (identical to the empirical trees). Then, we compared the number of species and the relative frequencies of each trait generated by the simulations with the empirical dataset. If models are adequate, simulated phylogenies should produce both diversity and frequency of states similar to the observed data.

Recently, Beaulieu and O'Meara (2016) described the HiSSE model that introduces hidden states associated with each of the traits and helps to accommodate the rate heterogeneity observed on empirical phylogenetic trees (Rabosky and Goldberg, 2015; Beaulieu and O'Meara, 2016). The implementation of HiSSE integrates incomplete sampling using the proportion of known species associated with each trait across the whole tree, but does not include the option of terminally unresolved phylogenetic trees such as 'diversitree' (FitzJohn, 2012). We fitted the 20 trait-dependent (including BiSSE and HiSSE models) and 4 trait-independent models to each of the 100 genus-level phylogenetic trees and performed model choice among models fitted for each tree using the Akaike Information Criteria (AIC).

Finally, we used BAMM (Rabosky et al., 2014) to estimate rates of diversification and shift locations in the tree independent of the trait data. Different from the BiSSE and HiSSE analyses, we used species-level phylogenies by randomly resolving the relationships within each genera using a constant rate Birth-Death model (Kuhn et al., 2011). This approach

assumes that diversification rates are homogeneous within each genera but should not bias the estimation of diversification rates for the backbone tree, since BAMM fits distinct rate regimes to different parts of the tree. Additionally, we set the BAMM model with constant rates through time within each regime (similar to MEDUSA – Alfaro et al., 2009), since this is a more adequate model given that diversification rates within genera were constrained to a homogeneous clock model. We set two independent MCMC chains for each of the 10 phylogenetic trees randomly sampled from the BEAST analysis. We ran each chain for 5 million generations, discarded 20% of generations as burnin and checked convergence using the 'coda' package (Plummer et al., 2006). Finally, we tested for a linear relationship between diversification rates and the proportion of each color pattern among species of each genera. Since there is large variation in species richness across the genera, we performed two analyses; the first including all genera with at least two species and the second excluding all genera with less than 10 species (total of 54 and 14 genera, respectively).

Recently, Moore and colleagues 2016 pointed out some important issues with BAMM's implementation; that the likelihood function does not incorporate diversification shifts on unobserved branches of the tree, that posterior estimates of the number of shifts is sensitive to the prior specification and that use of the compound Poisson process (CPP) prior make it difficult to differentiate many rate shifts of small effect from fewer rate shifts of large effect. Shortly after, Rabosky and colleagues 2017 responded to the critiques by showing that not incorporating shifts on unobserved branches has only minor impacts on inferences and argued that Moore et al. (2016) assessment of prior sensitivity on posterior estimates of the number of shifts was problematic. In this study, we use only relative diversification rates compared across genera (i.e., the slope of the regression) and our analyzes do not focus on the posterior distribution of the number of rate shifts across the tree. Thus, we believe that most concerns associated with BAMM should not have major impacts on our biological conclusions. Nevertheless, we repeated BAMM analyses with different prior distributions for the expected number of shifts ($\gamma = 0.1$, 1, or 10) to test whether our results could be affected by the prior sensitivity.

We repeated all trait-dependent analyses with each of the three distinct color pattern categories. R scripts, BAMM files, phylogenetic trees and coloration data to reproduce all analyses are available in the github repository `https://github.com/Caetanods/Dipsadidae_color_evolution` .

1.4 Results

We compiled a large report of coloration descriptions for 594 species of dipsadid snakes covering over 80% of the known diversity of the group. We were able to get detailed color descriptions for most species, but for some the information available was incomplete or limited (i.e., taxa known from a single specimen). Although those cases were not suitable for definition as coral-mimic or non-mimic lineages (i.e., state unknown or 'NA'), we managed to classify those as either contrasting or cryptic for all but a few exceptions. Among all data sources, museum specimens are the most difficult to categorize since colors fade after preservation and only light and dark hues remain. Bright colors such as yellow, orange, pink, and red (derived from carotenoid pigments) fade completely over time turning into cream on preservative fluid. In contrast, dark pigmentation is preserved and sometimes turns into shades of black or dark brown. As a result, we assigned museum specimens with alternate bright and dark bands (or with a distinct nuchal collar) as contrasting and homogeneous light or dark patterns as cryptic. We provide the species list and their color patterns under the distinct categorization schemes in the online data repository (`https://github.com/Caetanods/Dipsadidae_color_evolution`).

We made analyses using BiSSE based on the two different categorizations of the same color description dataset (see Material and Methods). When we compared coral-mimic to non-mimic patterns, net diversification estimates associated with the coral-mimic trait ($\lambda_1 - \mu_1$) were in average two times higher than non-mimics (Figure 1.1). In contrast, when diversification rates are constrained to be independent of color types we recovered intermediate net diversification values relative to the trait-dependent model. These results are qualitatively similar to the analyses based on the contrasting and cryptic alternative categorization (Figure 1.7). Both trait-independent and trait-dependent BiSSE models estimated strongly asymmetrical transition rates with changes from the coral-mimic to the non-mimic state (q_{10}) in average two to three times more frequent than the reverse (q_{01}). The median of the posterior distribution of transition rates are also qualitatively similar independent of color categorization (Figures 1.1 and 1.7). These results were consistent across all sampled trees and the trait-dependent diversification model was the one preferred by the DIC model selection criteria independent of the phylogeny and the color categorization used in the analysis (Figures 1.2 and 1.8).

We performed posterior predictive simulations to evaluate whether the BiSSE trait-dependent diversification model is adequate to explain the observed data. Since there is no appreciable difference in parameter estimates and DIC results across the categorizations of color types, we performed simulations based only in the contrasting vs. cryptic cate-

gory. The posterior predictive simulations using both trait-dependent and trait-independent BiSSE models produced trees much smaller than the 594 species provided as the observed total diversity of the group (see left column of Figure 1.3). Since the stopping criteria for simulations were tree depth, the number of species in each simulated tree was free to vary. Trees simulated under the full model had on average 211 species and only 9% of those showed more species than the observed data. Similarly, trees simulated using the trait-independent model had on average 186 species, of which only 5% were larger than the observed data. With respect to trait frequency, both models simulated datasets biased towards higher frequencies of the cryptic color type. On average the full model had 67% and the trait-independent model 84% of the simulations showing frequencies of the cryptic color type higher than observed in the empirical data (see right column of Figure 1.3). Therefore, our results show that both BiSSE models have similar biases and the characteristics of the data not satisfactorily explained by the models are independent to whether rates of diversification are associated with color types or not.

The analyses comparing 16 HiSSE, 4 BiSSE, and 4 trait-independent models estimated using maximum likelihood resulted in only two variants of the BiSSE model selected as the best model across the sample of 100 phylogenetic trees (Table 1). These results are also robust to different color categorizations (Table S1). These variants of the BiSSE model only differ on whether the extinction fraction is conditioned on the trait or not. Since BiSSE models implemented on 'hisse' (Beaulieu and O'Meara, 2016) only differ from the ones in 'diversitree' (FitzJohn, 2012) due to the use of orthogonal transformations of speciation and extinction for the purpose of more efficient parameter estimation (Beaulieu and O'Meara, 2016), we focus our attention to the posterior distribution of parameter estimates and posterior predictive simulations resulted from 'diversitree' BiSSE models.

Finally, we estimated trait-independent diversification rates across a pool of 10 Birth-Death resolved trees and tested whether there is a correlation between the proportion of mimic species and estimates of net diversification among genera. Results are consistent across all trees and show no evidence for an increase in net diversification rates associated to genera with higher proportion of mimics independent of the color category utilized (Figure 1.4 and 1.9). We repeated analyses using different parametrization of BAMM's prior on the expected number of shifts, but results show no influence of the prior on the outcome of the correlation tests applied to the posterior estimates of diversification rates (Figure 1.10).

1.5 Discussion

1.5.1 Model selection and (in)adequacy

We fitted the trait-dependent and trait-independent BiSSE models to the data and performed model selection using the Deviance Information Criteria (DIC). DIC showed strong support for the trait-dependent model for all three color pattern categorizations. On the other hand, our posterior predictive simulations showed that both trait-dependent and trait-independent BiSSE models are equally inadequate in explaining the observed data and share similar biases. Unfortunately, unlike other models of trait evolution (Pennell et al., 2015), it is not clear which set of summary statistics can be used to assess the adequacy of BiSSE models. More studies are needed to better understand the scenarios under which this and other models of the xxSSE family are prone to misbehave and elect a set of informative summary statistics for predictive posterior checks.

Results from comparing 24 different models across the pool of phylogenetic trees randomly sampled from the posterior distribution of the BEAST analyses consistently favored trait-dependent BiSSE models over all alternative models. This outcome is surprising because previous studies suggest that HiSSE models are often preferred over BiSSE models (e.g., Beaulieu and O'Meara, 2016; Frédérich et al., 2016), likely due to the accommodation of heterogeneity in both rates of transition between traits and rates of trait-dependent diversification. It is important to note that models estimated using the 'hisse' package (Beaulieu and O'Meara, 2016) apply a global sampling frequency approach that is agnostic to the distribution of unsampled species (and their trait data) among genera. Although it is plausible that this approach can influence parameter estimates, it is unclear whether a potential bias would favor BiSSE models over HiSSE and trait-independent models.

The BiSSE models implemented on 'diversitree' (FitzJohn, 2012) and 'hisse' (Beaulieu and O'Meara, 2016) packages are identical and, therefore, results from the posterior predictive simulations show that both are equally inadequate. Our additional analyses using results from trait-independent estimates of diversification among genera and testing whether there is an association with the proportion of mimic species also further corroborates the lack of an effect of coloration on the dynamics of diversification of the group. Based on the important limitations to explain the variation observed in the data shared by both trait-dependent and trait-independent BiSSE models and the disassociation between rates of diversification and the proportion of different color types across genera, we can conclude the more plausible macroevolutionary scenario is not of a trait-dependent mode of diversification. The choice of the trait-dependent BiSSE model over a trait-independent model based on Bayesian DIC

and as best model when comparing 24 alternative models using AIC are poorly justified by the data and might be the result of statistical artifacts.

1.5.2 Color patterns have no effect on the diversification of dipsadid snakes

The function of coral-mimic and contrasting color patterns in snakes and other groups has been intensely debated since the 50s (Dunn, 1954; Hecht and Marien, 1956). However, only a few studies have investigated the evolution of snake color patterns using an explicit phylogenetic approach (e.g., Pyron and Burbrink, 2009; Rabosky et al., 2016) despite recent advances in comparative methods. Our results show no consistent effect of the color types in diversification rates despite the impressive diversity of color patterns found in the group. Estimates of diversification rates independent of color types also give support to the trait-independent model. Overall, our results suggest that the ecological functions of coral-mimic and non-mimic color types, as well as contrasting and cryptic colors, in dipsadid snakes have no distinguishable effect on the macroevolution of the group.

The hypotheses of mimicry based on color similarity between mimics and models have received strong support from field experiments and the survival advantage of mimicry has been often demonstrated (Smith, 1975; Brodie III, 1993; Brodie III and Janzen, 1995; Hinman et al., 1997; Pfennig et al., 2001; Buasso et al., 2006). If mimicry explains the color types of dipsadid lineages that are similar to coral snakes, the protection attributed to the warning signal can have a positive effect on diversification, similar to the effect of aposematism (Mallet and Joron, 1999; Przeczek et al., 2008). However, increased survival has no necessary link to the generation of new species and our results show that distinct color types are not associated with appreciable differences in macroevolutionary patterns in dipsadids, independent of their putative ecological function.

Mimetic and contrasting patterns share bright colors that might be significantly more conspicuous to visually oriented predators than cryptic patterns. Field experiments using plasticine replicas provide evidence that the stereotyped alternating, bright colored bands (bearing shades of red, yellow, black and/or white) found among New World elapid species are avoided by predators (Brodie III, 1993; Brodie III and Janzen, 1995; Hinman et al., 1997; Pfennig et al., 2001; Buasso et al., 2006), even when these patterns are simplified to an extent that only a single ring occurs on the neck or head and the remaining of the body is plain red (Brodie III, 1993; see also Hinman et al., 1997). However, contrasting patterns are not restricted to such alternating, bright colored bands. In the absence of mimicry, such patterns might serve as disruptive coloration and deceive predators by creating optical illusions or hindering prey recognition. Indeed, it is plausible that coral-like patterns might

function both as mimetic and disruptive coloration (Dunn, 1954; Titcomb et al., 2014). The disruptive function (or flicker-fusion effect) can prevent or mitigate maladaptation caused by allopatric distribution of mimics and models, since it provides protection even in the presence of naïve predators. The double function of the contrasting pattern as mimetic and illusionary may help explain the impressive diversity of colors and patterns found among dipsadid snakes.

1.6 Concluding remarks

Herein we compiled from primary sources and made available a database of color patterns for 594 species of dipsadid snakes, the largest compilation of color descriptions for reptiles to date. We found that coral-mimic or contrasting patterns have no significant effect on rates of diversification when compared to non-mimic or cryptic color types. This is an intriguing contrast with the fact that aposematic clades are more species-rich than their cryptic sister groups (Przeczek et al., 2008). Speciation or extinction of mimetic lineages are theoretically linked to the relationship with their models. However, this dependence can be loosened if the mimetic trait is associated with a secondary protective function. Both eventual extinction events caused by allopatry with models and speciation as a result of local adaptation to novel models can be 'buffered' by the secondary function of the trait. The protection by illusion might be a precursor for both the remarkable convergence of snake lineages to coral-like forms and the maintenance of mimicry despite the supposed likelihood of mimicry breakdowns.

It is naïve to think that a unique set of traits, such as color patterns, can reflect all relevant factors that drive the dynamics of diversification of any group. A more detailed analysis of our questions could be accomplished by overlapping evolutionary patterns of the Dipsadidae with those of New World coral snakes, for example. However, both phylogenetic data and suitable comparative models are not yet available. Further appreciation of transitions between contrasting and cryptic color patterns within dipsadid genera can shed light on whether disruptive colors can serve as a pre-adaptation to mimicry and help insert new pieces into the coral snake mimicry puzzle. Understanding under which phylogenetic and ecological scenarios mimicry is likely to evolve is a key factor to explain the patterns of phenotypic convergence observed among distantly related lineages across the tree of life.

Table 1.1: Results from the three best models ranked among 24 trait-dependent and trait-independent models using Akaike Information Criteria (AIC) across a pool of 100 phylogenetic trees. Model BiSSE I has trait-dependent transition rates, turnover, and extinction fraction. Model BiSSE II is similar to BiSSE I but has trait-independent extinction fraction. Model HiSSE I has different turnover and extinction fraction for each state (including hidden states), but transition rates among states (including hidden states) are constrained to be the same. Model HiSSE II is similar to HiSSE I, but has trait-independent extinction fraction. Model HiSSE III is similar to HiSSE I, but turnover and extinction fraction are linked for all states except the hidden state associated with the coral-mimic category (state 1).

	BiSSE I	BiSSE II	HiSSE I	HiSSE II	HiSSE III
First best	17	83	–	–	–
Second best	83	17	–	–	–
Third best	–	–	63	1	36

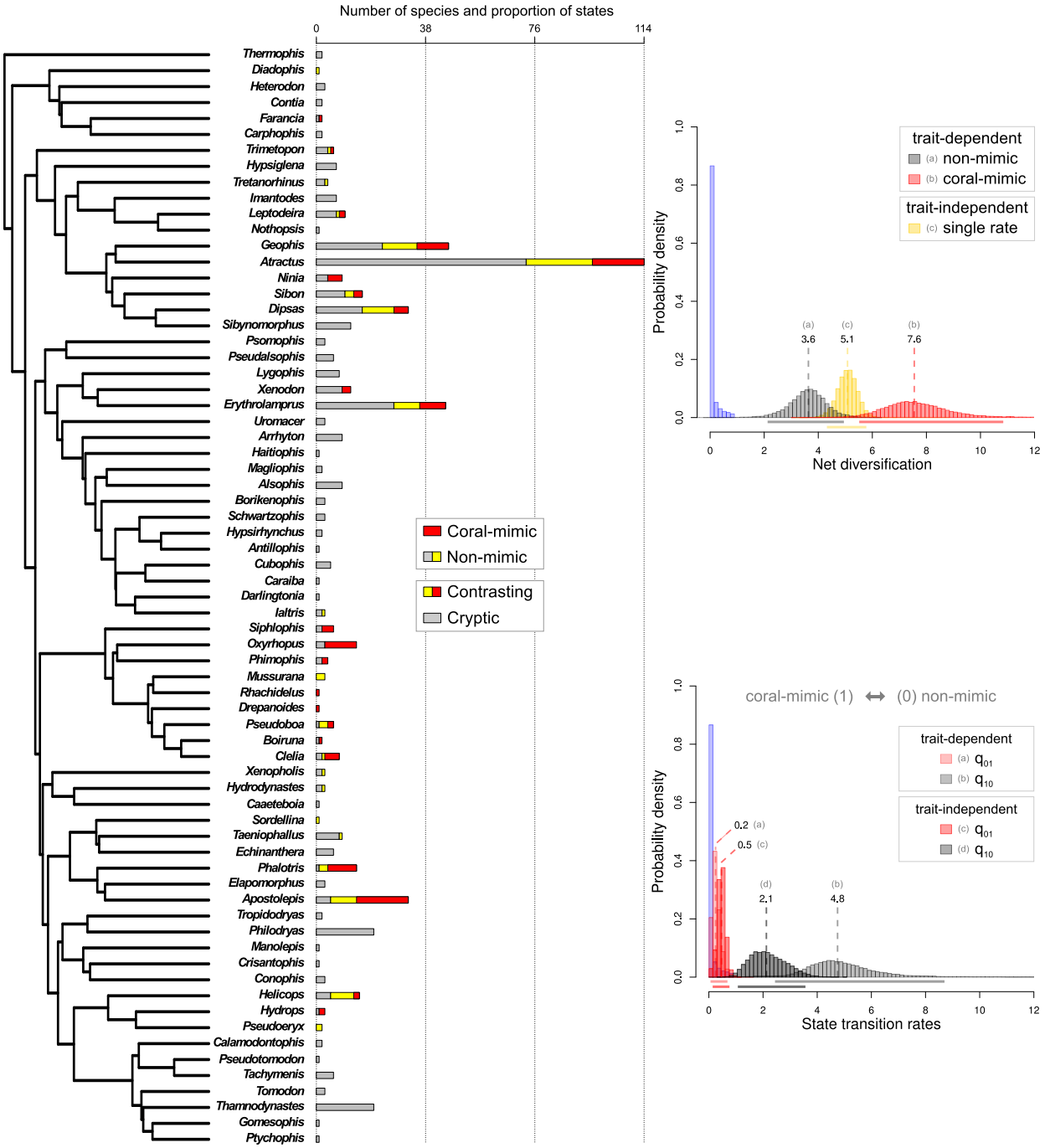


Figure 1.1: Genus-level maximum clade credibility (MCC) tree of the family Dipsadidae showing the number of species assigned to each color category and posterior distribution of parameter estimates for the BiSSE model. Each color of the stacked bar chart (center) correspond to different elements of the color patterns in the group; species that do not show contrasting color patterns (i.e., cryptic coloration) are shown in gray, species with contrasting color patterns but that are not supposed mimics of coral snakes are shown in yellow, and species that show contrasting color patterns and are considered mimics of coral snakes are shown in red. (*Continue on next page.*)

Figure 1.1: (*Continued.*) The legend in the center of the plate applies different combinations of gray, yellow, and red to show the number of species assigned to each color category: coral-mimic (red) versus non-mimic (gray and yellow) and contrasting (yellow and red) versus cryptic (gray). Support for the genera relationships are provided in Figures 1.5 and 1.6. Top-right plot shows the posterior distributions of net diversification rates under the trait-dependent (coral-mimic vs. non-mimic) and trait-independent BiSSE models. Bottom-right plot shows the posterior distributions of transition rates under the trait-dependent and trait-independent BiSSE models. Estimates are the combined posterior distribution from MCMC BiSSE runs across a pool of 100 phylogenetic trees. Prior distributions for the MCMC searches are shown in blue (unmarked distributions), the horizontal lines below each posterior distribution represent the 95% confidence interval, and the vertical hashed lines show median values.

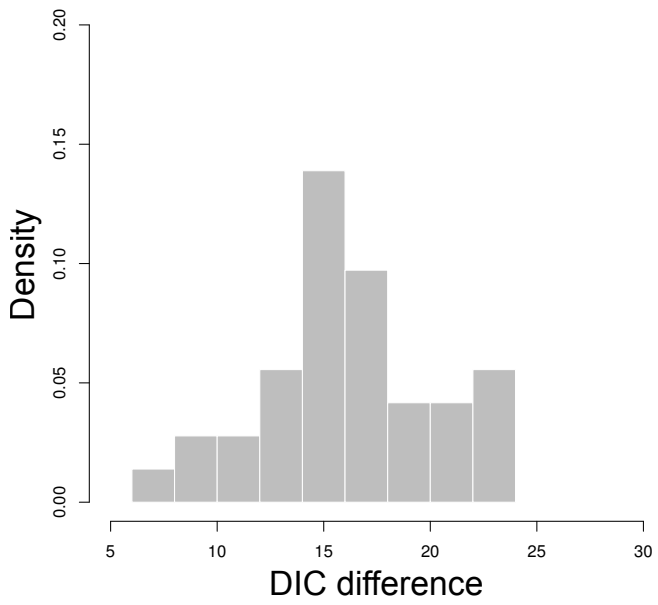


Figure 1.2: Results of model selection between trait-dependent and trait-independent BiSSE models using the Bayesian Deviance Information Criteria (DIC) for the coral-mimic versus non-mimic category. Plot shows the DIC scores for the trait-dependent model (full model) subtracted from the scores for the trait-independent model (constrained model). DIC values were calculated across a pool of 100 phylogenetic trees. Large values (larger than 4 units as a rule of thumb) are expected if the trait-dependent model is to be preferred over the trait-independent model.

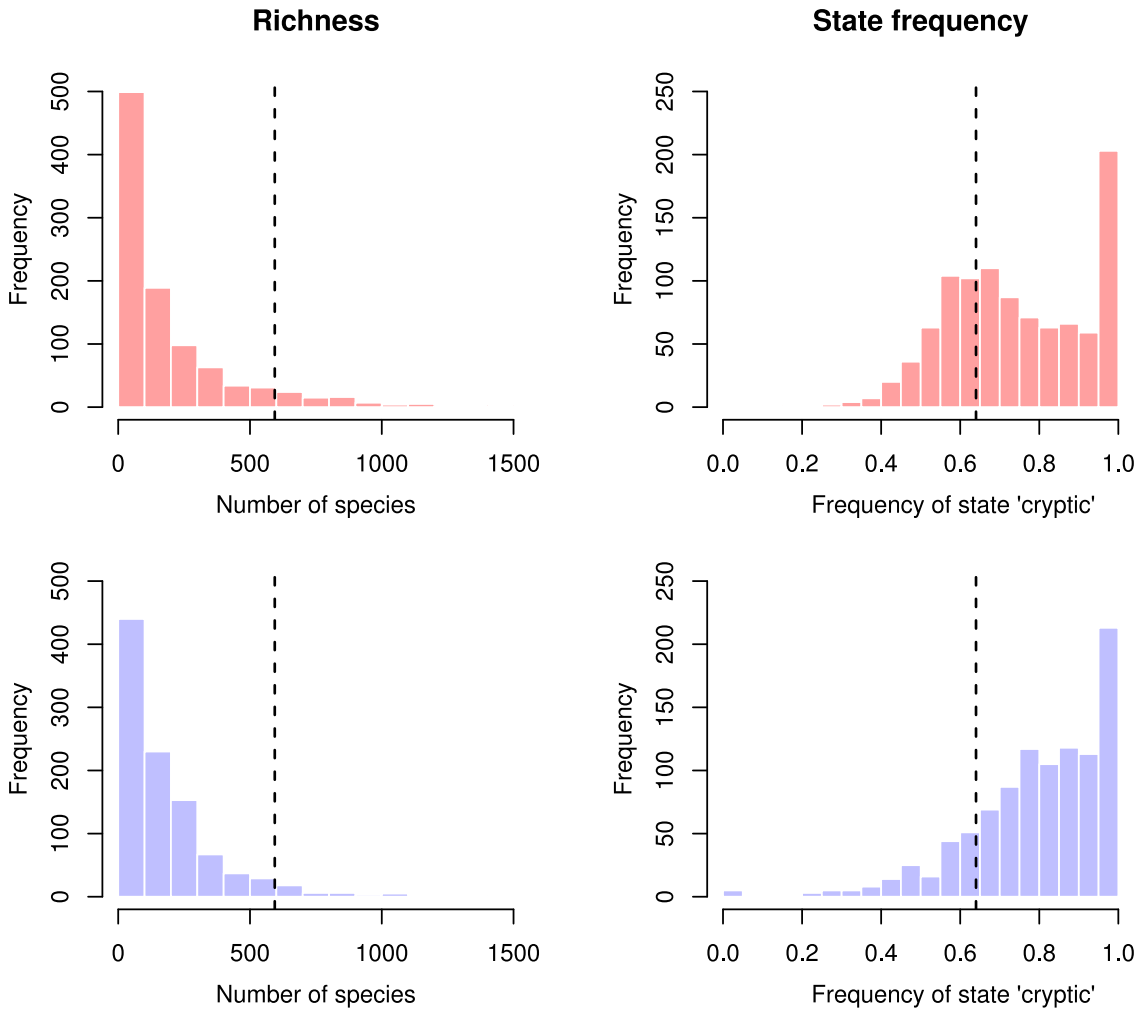


Figure 1.3: Results from the posterior predictive simulations for BiSSE under the trait-dependent (top row - red) and the trait-independent models (bottom row - blue). Simulation parameters were drawn from the joint posterior distribution of each BiSSE model (contrasting versus cryptic color category) and phylogenetic trees. At each replicate a phylogeny was simulated under the BiSSE model until the sum of branch lengths from root to tip of the tree was equal to 1 (same as the empirical phylogeny). Left column: Total number of species in the resulting phylogenies. Vertical dashed lines show the number of species used in our analysis (594 spp.). Right column: Relative frequency of the state 0 (cryptic). Vertical dashed lines show the observed frequency of cryptic species in the data (0.64). Note that both trait-dependent and trait-independent BiSSE models show similar deviations from the observed data.

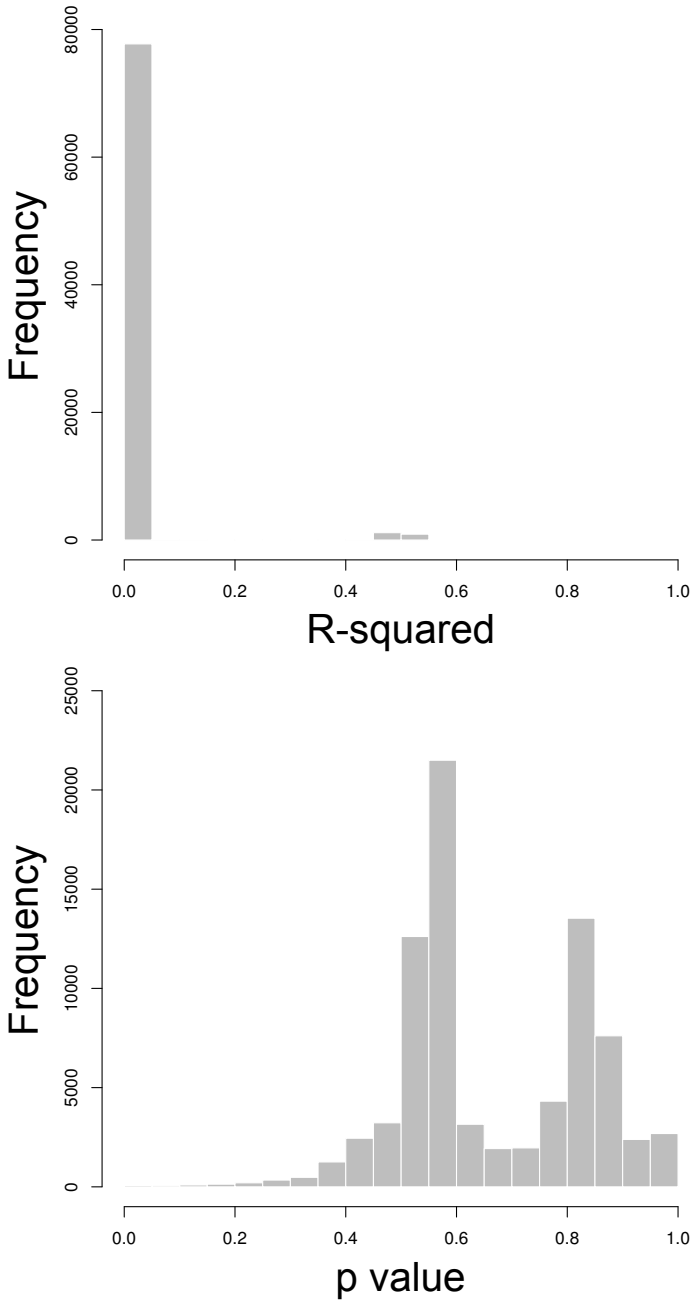


Figure 1.4: Results from linear regressions between net diversification rates estimated using BAMM (prior on expected number of shifts $\gamma = 1$) and the proportion of the category coral-mimic for each dipsadid genera. Plots show a distribution of linear regression tests generated by repeating the analyses for each sample of the posterior distribution of the BAMM diversification estimates combined across 10 randomly sampled phylogenetic trees. Top plot shows the resulting distribution of R-squared and bottom plot the distribution of p values.

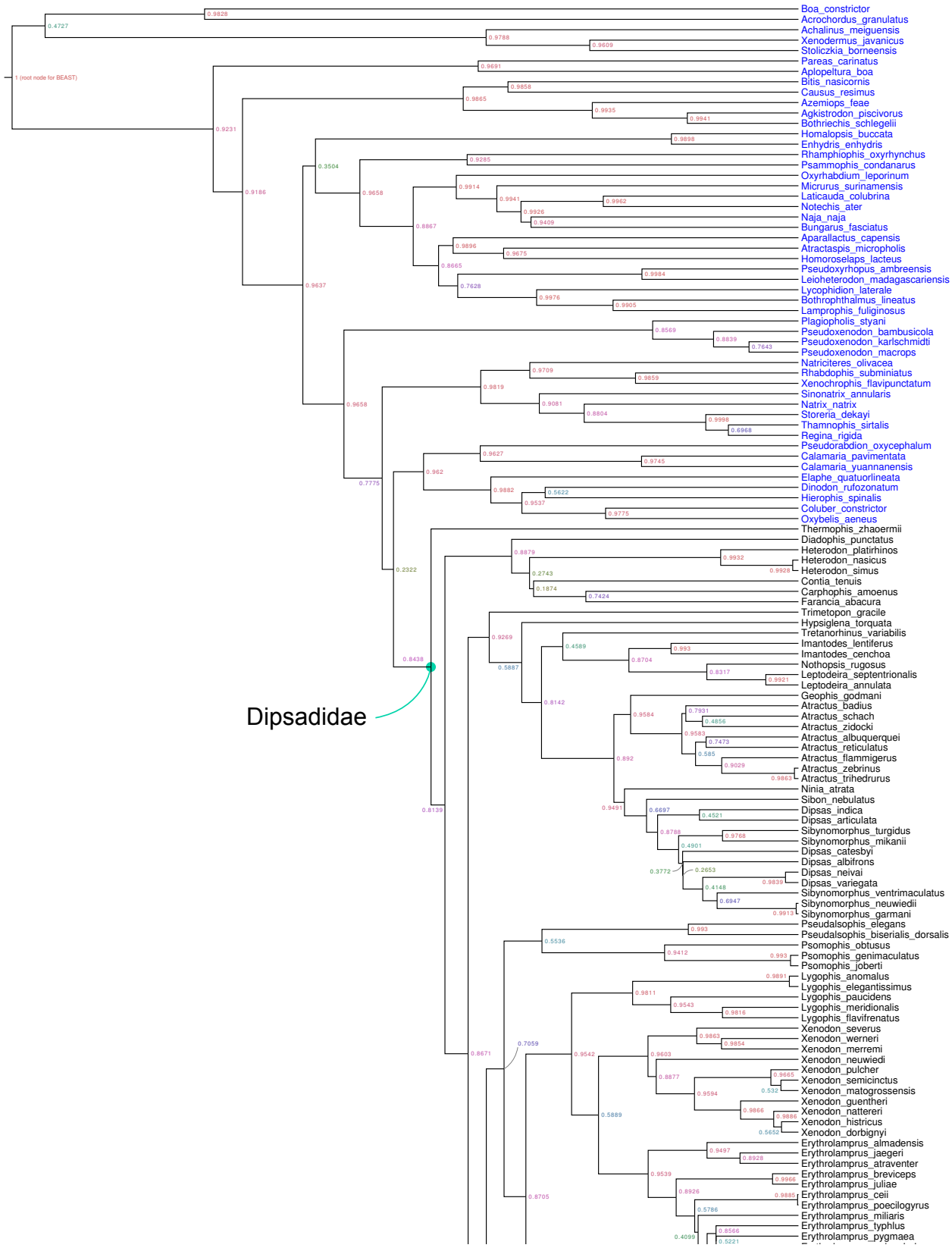


Figure 1.5: Maximum clade credibility tree with all species. Posterior probabilities are only visible on digital format. Outgroup species highlighted in blue. Continues on Figure 1.6.

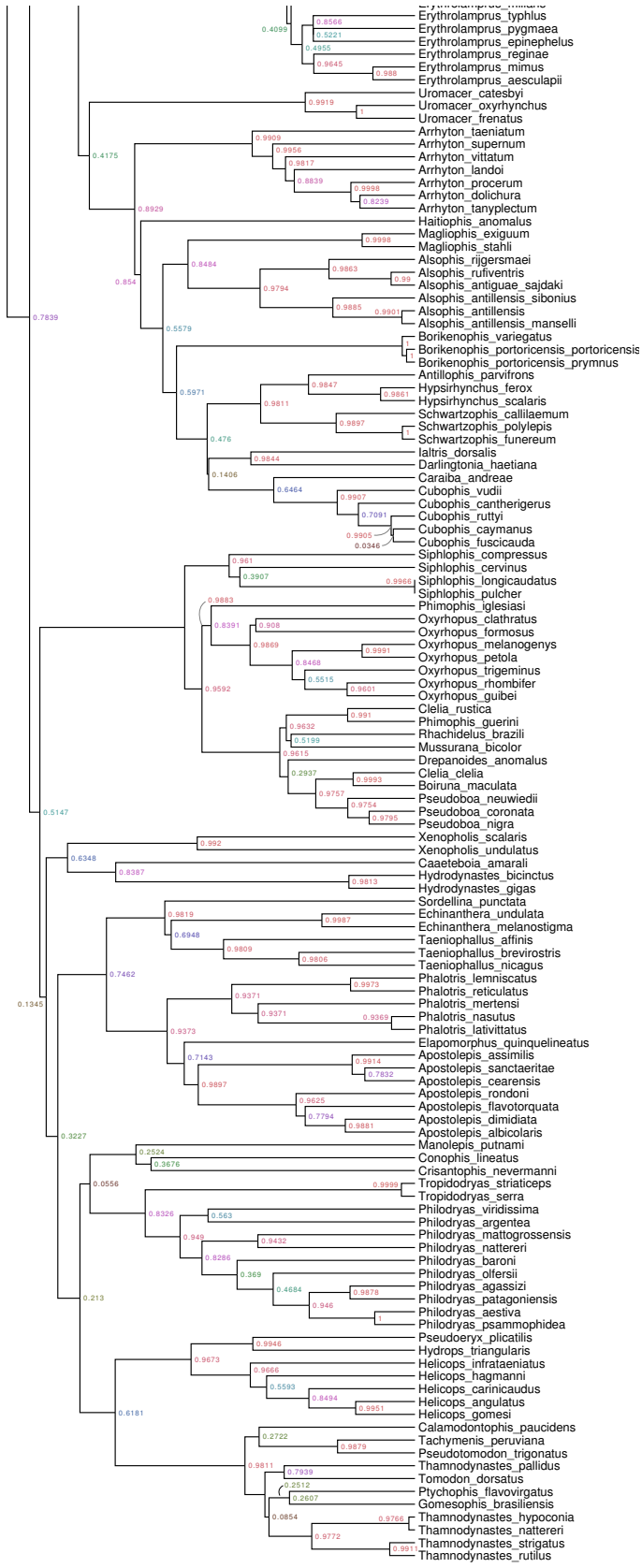


Figure 1.6: Maximum clade credibility tree with all species. Posterior probabilities are only visible on digital format. Continuation for Figure 1.5.

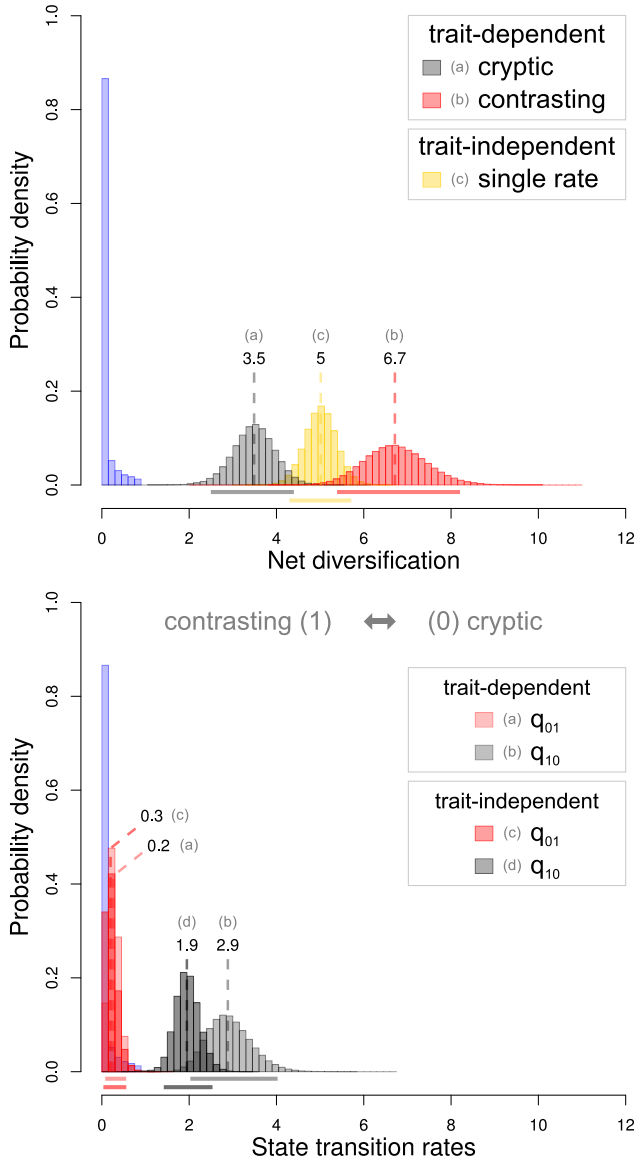


Figure 1.7: Parameter estimates for the BiSSE model under the contrasting versus cryptic color category. Top plot shows the posterior distributions of net diversification rates under the trait-dependent and trait-independent BiSSE models. Bottom plot shows the posterior distributions of transition rates under the trait-dependent and trait-independent BiSSE models. Estimates are the combined posterior distribution from MCMC BiSSE runs across a pool of 100 phylogenetic trees. Prior distributions for the MCMC searches are shown in blue (unmarked distributions), the horizontal lines below each posterior distribution represent the 95% confidence interval, and the vertical hashed lines show median values.

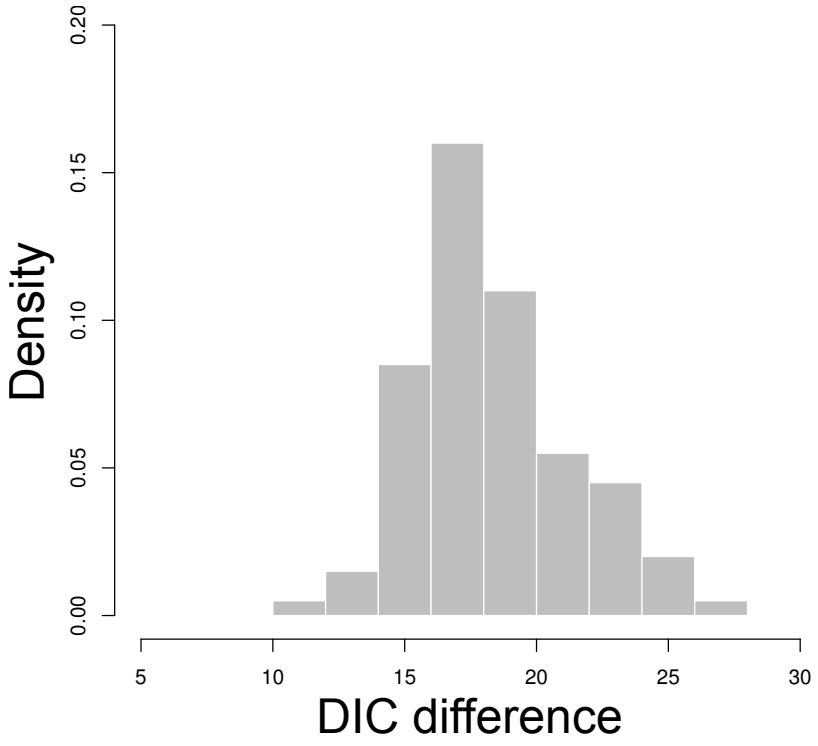


Figure 1.8: Results of model selection between trait-dependent and trait-independent BiSSE models using the Bayesian Deviance Information Criteria (DIC) for the contrasting versus cryptic category. Plot shows the DIC scores for the trait-dependent model (full model) subtracted from the scores for the trait-independent model (constrained model). DIC values were calculated across a pool of 100 phylogenetic trees. Large values (larger than 4 units as a rule of thumb) are expected if the trait-dependent model is to be preferred over the trait-independent model.

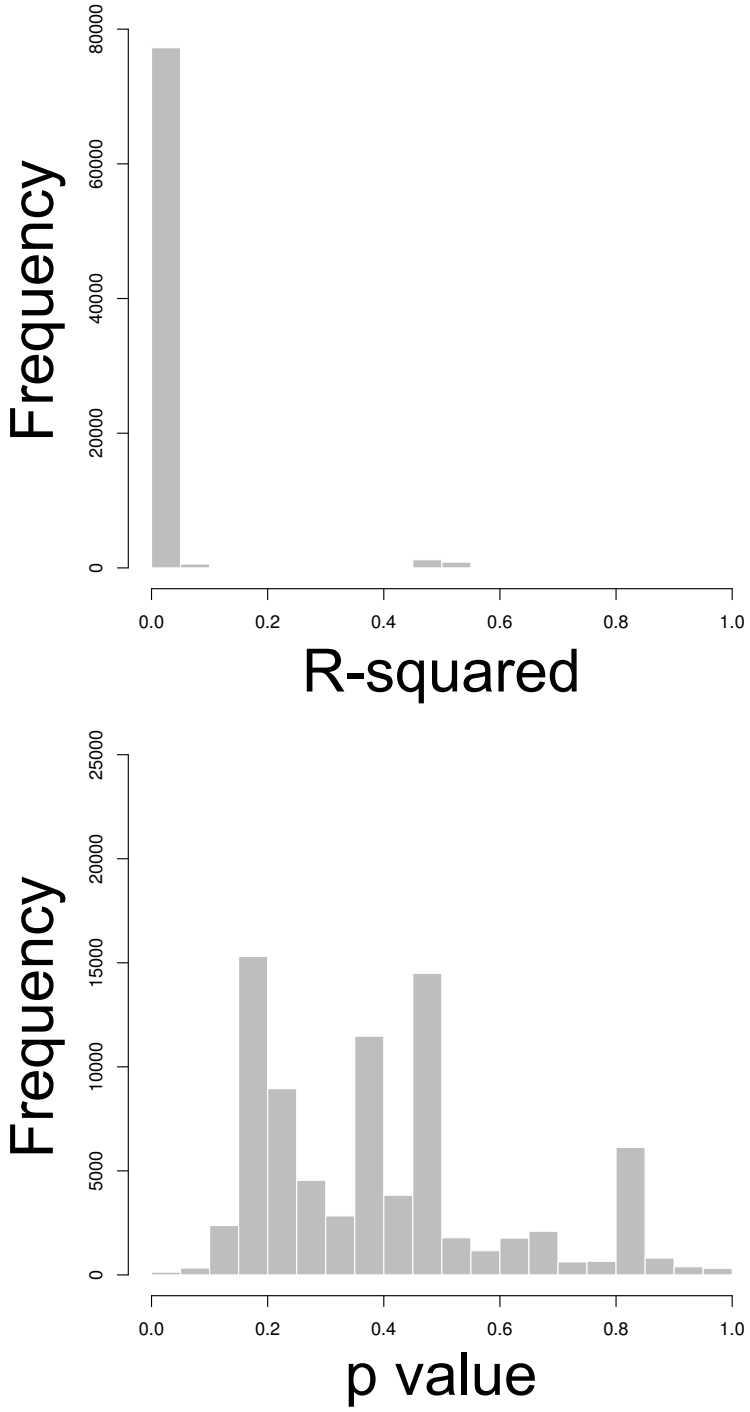


Figure 1.9: Results from linear regressions between net diversification rates estimated using BAMM (prior on expected number of shifts $\gamma = 1$) and the proportion of the category contrasting for each dipsadid genera. Plots show a distribution of linear regression tests generated by repeating the analyses for each sample of the posterior distribution of the BAMM diversification estimates combined across 10 randomly sampled phylogenetic trees. Top plot shows the resulting distribution of R-squared and bottom plot the distribution of p values.

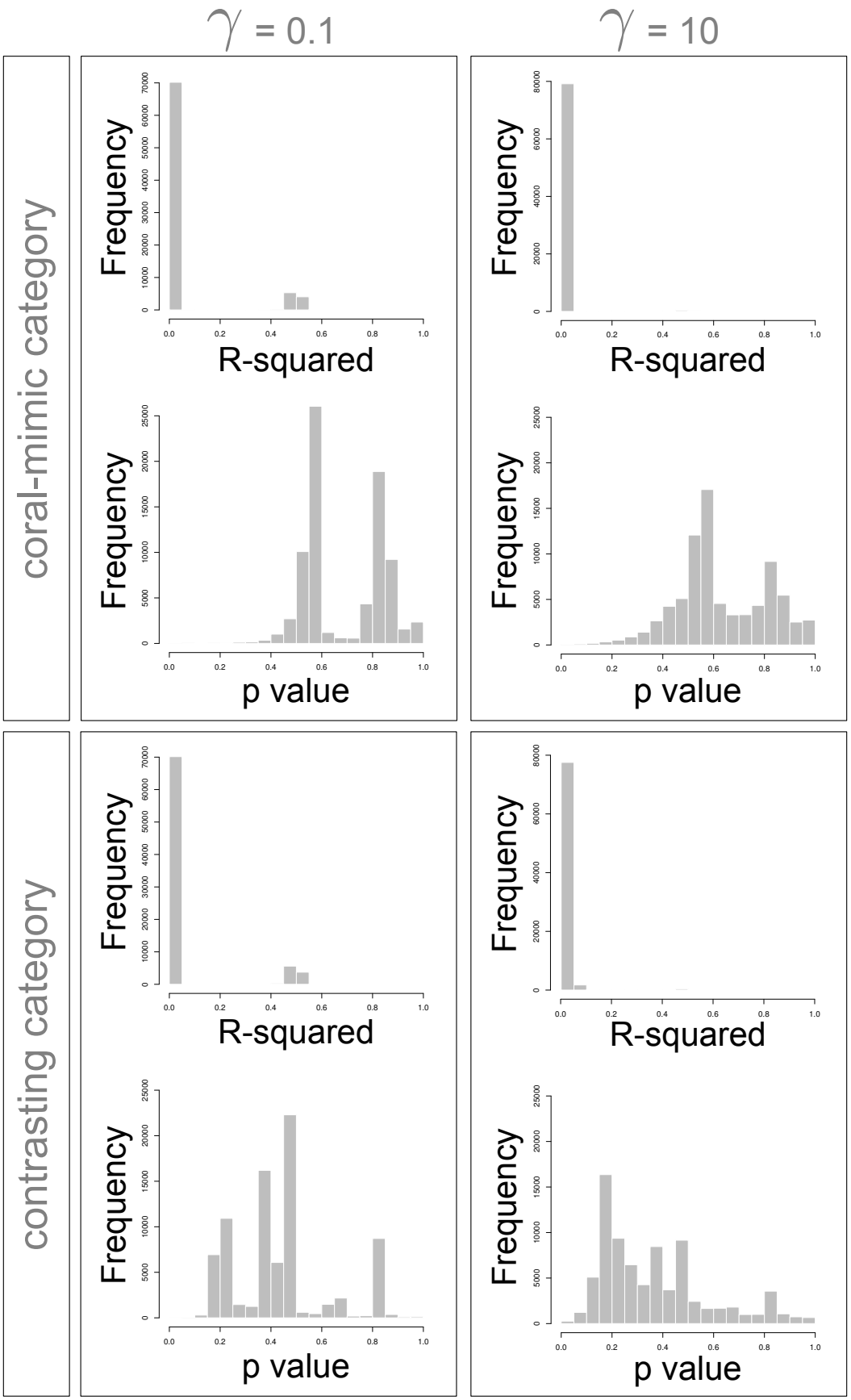

γ = 0.1
γ = 10
coral-mimic category
contrasting category
Frequency
R-squared
p value

Figure 1.10: Results from linear regressions between net diversification rates estimated using BAMM and the proportion of both color categories coral-mimic and contrasting for each dipsadid genera under different prior distributions on the number of expected rate shifts. Each plot show a distribution of results from linear regression tests generated by repeating the analyses for each sample of the posterior distribution of BAMM diversification estimates combined across 10 randomly sampled phylogenetic trees. The top plot within each treatment show the resulting distribution of R-squared and bottom plot the distribution of p values. Top row are results using the color category coral-mimic and bottom row using category contrasting. Results on the left column were based on a prior distribution for the expected number of shifts (γ) equal to 0.1 and on the right column equal to 10.

CHAPTER 2: ESTIMATING CORRELATED RATES OF TRAIT EVOLUTION WITH UNCERTAINTY

2.1 Abstract

Correlated evolution among traits can happen due to genetic constraints, ontogeny, and selection and have an important impact on the trajectory of phenotypic evolution. Thus, shifts in the pattern of evolutionary integration may allow the exploration of novel regions of the morphospace by lineages. Here we use phylogenetic trees to study the pace of evolution of several traits and their pattern of evolutionary correlation across clades and over time. We use regimes mapped to the branches of the phylogeny to test for shifts in evolutionary integration. Our approach incorporates the uncertainty related to phylogeny, ancestral state estimates and parameter estimates to produce posterior distributions using Bayesian Markov chain Monte Carlo. We implemented the use of summary statistics to test for regime shifts based on a series of attributes of the model that can be directly relevant to biological hypotheses. In addition, we extend Felsenstein's pruning algorithm to the case of multivariate Brownian motion models with multiple rate regimes. We performed extensive simulations to explore the performance of the method under a series of scenarios. Finally, we provide two test cases; the evolution of a novel buccal morphology in fishes of the family Centrarchidae and a shift in the trajectory of evolution of traits during the radiation of anole lizards to the Caribbean islands.

2.2 Introduction

Correlated evolution among traits, known as evolutionary integration, is ubiquitous across the tree of life and can have an important impact on the trajectory of phenotypic evolution (Olson and Miller, 1958; Klingenberg and Marugán-Lobón, 2013; Armbruster et al., 2014; Klingenberg, 2014; Goswami et al., 2014, 2015; Melo et al., 2016). Genetic constraints, ontogeny, and selection have pivotal roles in the development and maintenance of morphological integration over time (Arnold, 1992; Arnold et al., 2001; Hansen and Houle, 2004; Goswami et al., 2015; Melo et al., 2016). When the additive genetic covariance between traits is strong, then evolutionary correlation is likely due to genetic factors. In contrast, traits might not show strong genetic covariance and still be evolutionarily integrated due to correlated selection, which can be a result of distinct factors, such as anatomical interactions during growth or coordinated function (Armbruster and Schwaegerle, 1996; Armbruster

et al., 2014). For instance, correlated evolution can be favored by selection to maintain a cohesive pattern of variation among traits with a shared function, but evolution can be hindered if the genetic covariance is not aligned with the selection gradient (Lande, 1979; Schluter, 1996; Villmoare, 2012; Goswami et al., 2014). Alternatively, when evolutionary correlation is mainly a result of correlated selection then the morphospace occupied by lineages can be restricted by the strength and direction of the selection gradient (Felsenstein, 1988; Armbruster and Schwaegerle, 1996). Shifts in the pattern of evolutionary integration among traits over macroevolutionary scales, due to changes in the genetic architecture or selection gradient, may play a fundamental role in the exploration of novel regions of the morphospace (Young and Hallgrímsson, 2005; Goswami, 2006; Revell and Collar, 2009; Monteiro and Nogueira, 2010; Hallgrímsson et al., 2012; Claverie and Patek, 2013).

Macroevolutionary transitions in morphospace evolution have been associated with both increases and decreases in the evolutionary integration among traits. In centrarchid fishes, for example, the evolution of a novel mouth morphology was followed by a rapid differentiation of feeding habits. More specifically, the increase in the evolutionary correlation between two morphological features of the suction-feeding mechanism in species of *Micropterus* is associated with a specialization towards consumption of larger prey (Collar et al., 2005; Revell and Collar, 2009). In contrast, the once strong developmental integration between the fore- and hindlimbs of early tetrapods underwent a dramatic change allowing the limbs to respond to diverging selective pressures and leading to the evolution of bipedalism and flight (Young and Hallgrímsson, 2005; Young et al., 2010; Dececchi and Larsson, 2013). These examples show the role of shifts in evolutionary integration associated with the evolution of novel morphologies. However, stable patterns of evolutionary integration over long time scales can be responsible for the constraint of lineages to limited regions of the morphospace and might be a plausible mechanism associated with patterns of stasis observed in the fossil record (Hansen and Houle, 2004; Bolstad et al., 2014; Goswami et al., 2015). Thus, evolutionary trait correlations are central to the maintenance of form and function through time, but can either drive or slow morphological differentiation.

Despite the prevalent role of evolutionary integration, most of what we know about the tempo and mode of trait evolution come from studies of single traits (e.g., Harmon et al., 2010; Hunt et al., 2015, among others). Even when multiple traits are the object of investigation, studies often use principal component axes (or phylogenetic PCA; Revell, 2009) to reduce the dimensionality of the data so that univariate methods can be applied (Harmon et al., 2010; Mahler et al., 2013; Klingenberg and Marugán-Lobón, 2013, see Uyeda et al. 2015 for more examples). This is most likely a reflection of the phylogenetic comparative models of trait evolution available for use, since few are focused on two or more traits (but see

Revell and Harmon, 2008; Hohenlohe and Arnold, 2008; Revell and Collar, 2009; Bartoszek et al., 2012; Adams, 2012, 2014b; Clavel et al., 2015). However, studying one trait at a time eliminates the possibility of identifying patterns of evolutionary correlation, while principal component axes does not allow testing for evolutionary shifts in integration because the orientation of the PC axes are homogeneous across the branches of the phylogenetic tree. Furthermore, it also has been shown that PCA can influence our biological interpretation about the mode of evolution of the data (Uyeda et al., 2015) because the first PC axes are consistently estimated as early bursts of differentiation whereas the last axes store a strong signal of stabilizing selection, independent of the true model of evolution of the traits. As a result, we need models that apply to multivariate data as such in order to better understand macroevolutionary patterns of evolutionary integration.

One way to model multivariate trait evolution using phylogenetic trees is through the evolutionary rate matrix (Hohenlohe and Arnold, 2008; Revell and Harmon, 2008; Revell and Collar, 2009; Adams and Felice, 2014). This is a variance-covariance matrix that describes the rates of trait evolution under Brownian motion in the diagonals and the evolutionary covariance among traits (i.e., the pattern of evolutionary integration) in the off-diagonals (Huelsenbeck and Rannala, 2003; Revell and Harmon, 2008). The evolutionary rate matrix is ideal for studying patterns of evolutionary integration because it allows for simultaneous estimate of the individual rates of evolution of each trait as well as the evolutionary covariance between each pair of traits. It is also a flexible model, since any number of evolutionary rate matrix regimes can be fitted to the same phylogenetic tree (Revell and Collar, 2009). The contrast between evolutionary rate matrices independently estimated in different regions of the tree can inform us about the magnitude and direction of shifts in the pattern of evolutionary integration.

One of the challenges of working with rate matrices is that covariances can be hard to estimate, especially when the number of species (observations) is small relative to the number of traits (parameters) in the model. As the number of parameters in a model increases, the amount of data required for proper estimation also increases and it becomes crucial to directly incorporate uncertainty in parameter estimates when interpreting results. However, the majority of studies to date have relied on point estimates of the evolutionary rate matrix by maximum likelihood (Revell and Harmon, 2008; Revell and Collar, 2009; Clavel et al., 2015; Goolsby, 2016, but see Huelsenbeck and Rannala, 2003 and Dines et al., 2014 for exceptions). Although the confidence interval around the maximum likelihood estimate can be used as a measure of uncertainty, this quantity is rarely reported (Revell and Harmon, 2008; Revell and Collar, 2009; Adams, 2012; Immler et al., 2012; Adams and Felice, 2014; Collar et al., 2014). Furthermore, the uncertainty in parameter estimates does not take direct

part in model selection using likelihood ratio tests or AIC (Burnham and Anderson, 2003), which can lead researchers to erroneous conclusions about their models. Besides the possible uncertainty in parameter estimates, there is an important computational burden associated with the evaluation of the likelihood function of the multivariate Brownian motion model due to the computation of matrix inversions and determinants (Felsenstein, 1973; Hadfield and Nakagawa, 2010; Freckleton, 2012). Thus, computational time can become a limitation when performing a large number of likelihood evaluations, such as in simulation based approaches. Recently, Adams (2014b) described a method to estimate the rate of evolution under Brownian motion of traits defined by several dimensions (high-dimensional data), even when the number of trait dimensions exceeds the number of lineages in the phylogeny. This method was extended to a plethora of variations based on the same general framework (Adams and Felice, 2014; Adams, 2014a; Denton and Adams, 2015, see also Goolsby (2016) for a different implementation). These methods work with high-dimensional data as a result of the use of distance matrices rather than covariance matrices, since the later becomes singular if the number of variables is larger than the number of observations. However, by avoiding the calculation of the covariance among trait dimensions (Adams, 2014b), such suite of methods assume a homogeneous rate of evolution shared by all dimensions of a trait (the σ^2_{mult}). Thus, σ^2_{mult} is ideal for high-dimensional traits such as shape data, but it has limitations for the study of evolutionary integration among multiple traits.

In order to ask questions about the evolution of integration using phylogenetic trees we need a computationally efficient method that can estimate evolutionary rate matrices while incorporating uncertainty in parameter estimates. Here we implement a Bayesian estimate for the evolutionary rate matrix using Markov chain Monte Carlo (MCMC) to provide a direct assessment of the uncertainty associated with parameter estimates in the form of a posterior distribution. Our implementation also allows for multiple regime configurations and/or phylogenetic trees to be incorporated in the MCMC chain, thus integrating the uncertainty associated with ancestral state estimates and phylogenetic reconstruction to the analysis. In order to increase the performance of the likelihood evaluation, we implemented Felsenstein's (1973) pruning algorithm. We also derive a new version of the pruning algorithm that is suitable for the special case when several rate regimes of the multivariate Brownian motion model are fit to different branches of the same phylogenetic tree. We apply our new approach to two biological examples: the fast evolution of morphology associated with the radiation of *Anolis* lizards from mainland South America to the Caribbean islands (Pinto et al., 2008; Mahler et al., 2013; Moreno-Arias and Calderón-Espinosa, 2016) and the shift of feeding habits driven by the change in mouth morphology in Centrarchidae fishes (Revell and Collar, 2009). We show that there is no detectable shift in the evolutionary integration among

morphological traits during the anole radiation and that there is significant uncertainty in estimates of evolutionary correlation associated with the Centrarchidae mouth traits. We also provide results from extensive simulations showing that our approach has good performance under diverse scenarios of correlated evolution.

2.3 Methods

2.3.1 A new pruning algorithm for multivariate Brownian motion with multiple regimes

To test for shifts in the pattern of evolutionary integration among traits we need to estimate the rates of evolution for the individual traits and their evolutionary covariation, i.e. by estimating the evolutionary rate matrix (**R**; Revell and Harmon, 2008). Revell and Collar (2009) derived a general form of the likelihood function for the model that allows for several independent matrices assigned to different branches of the phylogenetic tree.

$$L_p = \frac{\exp[-(\mathbf{y}-\mathbf{D}\mathbf{a}^T)^T(\sum_{k=1}^{p}\mathbf{R}_k\otimes\mathbf{C}_k)^{\frac{-1(\mathbf{y}-\mathbf{D}\mathbf{a}^T)}{2}}]}{\sqrt{(2\pi)^{nr}\mid\sum_{k=1}^{p}\mathbf{R}_k\otimes\mathbf{C}_k\mid}} \tag{2.1}$$

Where $\mathbf{y}$ is a vector of length $n\cdot r$ derived by concatenating the columns of a n by r matrix of trait values for n tips and r traits; $\mathbf{D}$ is a $n\cdot r$ by r design matrix composed of 1 for each (i,j) entry that satisfies $(j-1)\cdot n < i \leq j\cdot n$ and 0 otherwise; $\mathbf{a}$ is a vector with r root values for the tree (or the phylogenetic mean); $\mathbf{R}_k$ is the kth evolutionary rate matrix with size r. Each of the $\mathbf{C}_k$ matrices has only the sum of branch lengths which were assigned to the respective evolutionary rate matrix. Thus, $\sum_{k=1}^{p}\mathbf{C}_k$ is equal to the phylogenetic covariance matrix ($\mathbf{C}$) for the whole tree. The elements of $\mathbf{C}$ are composed by the sum of branch lengths shared by each pair of taxa (Felsenstein, 1973). Finally, p is the number of $\mathbf{R}$ matrix regimes fitted to the tree. When p is equal to 1, equation (2.1) reduces to the likelihood function for a single $\mathbf{R}$ matrix (Revell and Harmon, 2008).

The likelihood function for the evolutionary rate matrix as shown requires the matrix inversion and determinant to be computed for the sum of the Kronecker product between each $\mathbf{R_k}$ and $\mathbf{C_k}$ matrices. However, the matrices resulted from this product can be very large because each $\mathbf{R}$ has dimension equal to the number of traits in the data whereas $\mathbf{C}$ is as large as the number of tips in the phylogeny. Some methods can be used to speed up the computation in the case of multiple rate regimes applied to the tree. For instance,

the 'rpf' method avoids the explicit computation of the matrix inversion and determinant by applying Cholesky factors (Gustavson et al., 2010; Clavel et al., 2015) whereas Goolsby (2016) recently introduced the use of pairwise composite likelihoods, which consists of the product of the pairwise likelihoods computed for all combinations of traits. These methods reduce the computational time for the evaluation of the likelihood but are still more time consuming than the pruning algorithm (Felsenstein, 1973; Freckleton, 2012; Caetano and Harmon, 2017b). Here, we expand the pruning algorithm applied to the multivariate Brownian motion model (Felsenstein, 1973; Freckleton, 2012) to compute the likelihood even when multiple evolutionary rate matrices are fitted to different branches of the phylogenetic tree. This algorithm is implemented in the R package `ratematrix` (Caetano and Harmon, 2017b).

2.3.2 Computing the likelihood for the multivariate Brownian motion model with multiple regimes using the new pruning algorithm

The pruning algorithm explores the property that trait changes in each of the branches can be modelled independently and applies a multivariate normal density to compute the likelihood of evolutionary changes at each branch assuming a Brownian motion model (Felsenstein, 2004; Freckleton, 2012). When multiple rate regimes are fitted to a phylogeny, the likelihood is often computed by scaling the branch lengths by the rates (e.g., Eastman et al., 2011). However, this procedure is not applicable to the multivariate case, since the product of the length of a branch and the BM rate is a matrix. We derived the pruning algorithm for multiple rate regimes by following the same procedures described by Felsenstein (1973, 2004), but assuming that all rates are multivariate, that rates are different at each branch and that branches can have more than one rate regime (after Revell and Collar, 2009). This algorithm completely avoids the calculation of the matrix inverse and the determinant of the phylogenetic covariance matrix ($\mathbf{C}$) or the Kronecker product between $\mathbf{R}$ and $\mathbf{C}$ matrices. However, the inverse of the $\mathbf{R}$ matrix, which will have dimensions equal to the number of traits in the data set, is still required.

In this extension of the algorithm, each branch of the phylogeny can be assigned to one or more evolutionary rate matrix ($\mathbf{R}$) regimes and the sum of the portions of the branch assigned to each regime need to be equal to the total length of that branch (Revell and Collar, 2009). We demonstrate that the algorithm yields the same likelihood as in Felsenstein (1973) and Freckleton (2012) by showing that all calculations converge when a single regime is fitted to tree. The pruning algorithm works by visiting the tips and going down node by node. At each step the contrast between two tips is computed and a new "phenotype" value replaces

the two original tips, becoming the new tip. The likelihood of the contrast is calculated and we move to the next contrast until we reach the root node. From here on we will refer to Figure 2.1 as an example of phylogenetic tree. Where $\mathbf{x}_i$ is a vector with r trait values for tip i and v_i is the branch length leading to tip or node i. We will refer to the node representing the common ancestor of tips 1 and 2 as the node 4 and the node representing the common ancestor of all tips as the root node. The method works as following:

1. **Calculate the contrast.** Choose a pair of tips i and j with a unique and exclusive common ancestor k. In our example, the selected species are 1 and 2. Compute the contrast $\mathbf{u}_{ij} = \mathbf{x}_i - \mathbf{x}_j$.

2. **Compute the log-likelihood.** Use the vector of contrasts ($\mathbf{u}_{ij}$), the number of traits in the data (r), the branch lengths (v_i and v_j), and the length of the branches assigned to each of the k evolutionary rate matrix regimes (Revell and Collar, 2009) to compute the log-likelihood:

$$L = -\frac{1}{2}\left(r\log(2\pi) + \log|\mathbf{S_i} + \mathbf{S_j}| + \mathbf{u}_{ij}^{\intercal}\left(\mathbf{S_i} + \mathbf{S_j}\right)^{-1}\mathbf{u}_{ij}\right)$$

where

$$\begin{aligned}\mathbf{S_i} &= \mathbf{R_1}\, v_{1i} + \mathbf{R_2}\, v_{2i} + \ldots + \mathbf{R_k}\, v_{ki} \\ \mathbf{S_j} &= \mathbf{R_1}\, v_{1j} + \mathbf{R_2}\, v_{2j} + \ldots + \mathbf{R_k}\, v_{kj}\end{aligned} \tag{2.2}$$

and

$$\begin{aligned}v_i &= v_{1i} + v_{2i} + \ldots + v_{ki} \\ v_j &= v_{1j} + v_{2j} + \ldots + v_{kj}\end{aligned}$$

If we assume a single evolutionary rate matrix is fitted to the whole tree, equation 2.2 reduces to equation 10 in Freckleton (2012):

Let

$$\mathbf{R} = \mathbf{R_1} = \mathbf{R_2} = \ldots = \mathbf{R_k}$$

then

$$\begin{aligned}\mathbf{S_i} &= \mathbf{R_1}\, v_{1i} + \mathbf{R_2}\, v_{2i} + \ldots + \mathbf{R_k}\, v_{ki} \\ &= \mathbf{R}\, v_{1i} + \mathbf{R}\, v_{2i} + \ldots + \mathbf{R}\, v_{ki} \\ &= \mathbf{R}\sum_{l=1}^{k} v_{li}\end{aligned}$$

We know, from equation 2.2, that the sum of the portions of the branch length assigned to each regime is equal to the total length of the branch. Then:

$$\mathbf{S_i} = \mathbf{R}v_i \quad \text{as well as} \quad \mathbf{S_j} = \mathbf{R}v_j$$

and

$$\mathbf{S_i} + \mathbf{S_j} = \mathbf{R}(v_i + v_j)$$

Substituting into equation 2.2, we have:

$$\begin{aligned} L &= -\frac{1}{2}\left(r\log(2\pi) + \log|\mathbf{R}(v_i+v_j)| + \mathbf{u}_{ij}^{\intercal}\left(\mathbf{R}(v_i+v_j)\right)^{-1}\mathbf{u}_{ij}\right) \\ &= -\frac{1}{2}\left(r\log(2\pi) + \log|\mathbf{R}| + r\log(v_i+v_j) + \frac{\mathbf{u}_{ij}^{\intercal}\left(\mathbf{R}\right)^{-1}\mathbf{u}_{ij}}{(v_i+v_j)}\right) \end{aligned} \tag{2.3}$$

Which is the same as equation 10 in Freckleton (2012)[1].

3. **Calculate the new phenotype vector $\mathbf{x}_n$ for the node n.** This quantity is originally calculated as the weighted average of the vector of species means for species i and j with weights equal to the length of the branches v_i and v_j. For the case of a single trait, $\mathbf{x}_{1i}$ and $\mathbf{x}_{1j}$, we would have:

$$\mathbf{x}_{1n} = \frac{v_i\,\sigma^2_{1i}}{v_i\,\sigma^2_{1i} + v_j\,\sigma^2_{1j}}\,\mathbf{x}_{1j} + \frac{v_j\,\sigma^2_{1j}}{v_i\,\sigma^2_{1i} + v_j\,\sigma^2_{1j}}\,\mathbf{x}_{1i} \tag{2.4}$$

When $\sigma^2_i = \sigma^2_j$, equation 2.4 becomes equivalent to equation 7 in Felsenstein (1973) and the rates of each branch can be omitted. However, here we assume that rates are different in every branch, that the evolutionary covariance among traits are non-zero and that more than one rate regime can be assigned to the same branch. As a result, the rates need to be represented as variance-covariance matrices ($\mathbf{R_1}, \mathbf{R_2}, \ldots, \mathbf{R_k}$) and the sum of the product between the portions of each branch and their rate regimes is given by the matrices $\mathbf{S_i}$ and $\mathbf{S_j}$ (see equation 2.2). By expanding equation 2.4, we have:

$$\mathbf{x}_n = \mathbf{S_i}\left(\mathbf{S_i}+\mathbf{S_j}\right)^{-1}\mathbf{x}_j + \mathbf{S_j}\left(\mathbf{S_i}+\mathbf{S_j}\right)^{-1}\mathbf{x}_i \tag{2.5}$$

[1]Note that the published equation in Freckleton (2012) has a printing error. The corrected form is $L = -\frac{1}{2}\left(k\log(2\pi) + \log|\mathbf{C}| + k\log V_i + \frac{\mathbf{u}_i^t\mathbf{C}^{-1}\mathbf{u}_i}{V_i}\right)$. The correct form can be appreciated in the function 'clikGeneral' on line 393 of the Supporting Information file `MEE3_220_sm_demo.R` available online (Freckleton, 2012).

In the first step of our example, we calculate the phenotype value for the node 4 ($\mathbf{x}_4$). Then, we prune the tips 1 and 2 from the tree, leaving only the tip 3 and the new tip 4 with vector of trait values $\mathbf{x}_4$. The next contrast will be calculated between $\mathbf{x}_4$ and $\mathbf{x}_3$.

4. **Compute the variance of $\mathbf{x}_n$.** After computing the vector of trait values for the node n, we need to calculate the variance associated with the uncertainty in the estimation of $\mathbf{x}_n$. This uncertainty is added to the variance of the branch immediately bellow the node n. For a single trait and we would have:

$$var[\mathbf{x}_{1n}] = \frac{v_i\sigma^2_{1i}\, v_j\sigma^2_{1j}}{v_i\sigma^2_{1i} + v_j\sigma^2_{1j}} + v_n\sigma^2_{1n} \tag{2.6}$$

Where $m, \ldots, n$ are the indexes for the branches that connect the root to the node n of the tree. Again, when a single rate regime is fitted to the tree, equation 2.6 is equivalent to equation 9 in Felsenstein (1973). For the multivariate case, this quantity becomes a variance-covariance matrix which is added to $\mathbf{S_n}$ (the branch length below the node n multiplied by the rate regimes; see equation 2.3) and can be calculated as:

$$var[\mathbf{x}_n] = \left((\mathbf{S_i})^{-1} + (\mathbf{S_j})^{-1}\right)^{-1} + \mathbf{S_n} \tag{2.7}$$

The equivalence between equations 2.6 and 2.7 can be easily verified by checking the computation of the harmonic mean of matrices. For two scalar quantities the harmonic mean is given by $\frac{2\,a\,b}{a+b}$ whereas for matrices we have $((\mathbf{A})^{-1} + (\mathbf{B})^{-1})^{-1}$.

5. **Repeat.** Steps 1 to 4 are repeated until only two tips remains. The root node will have a zero contrast. The variance associated with the root node is computed as:

$$var[\mathbf{root}] = \left((\mathbf{S_i})^{-1} + (\mathbf{S_j})^{-1}\right)^{-1} \tag{2.8}$$

6. **Compute the final log-likelihood.** The final log-likelihood conditioned on the phylogenetic tree, rate regime and trait data is computed as the sum of all partial (node-by-node) log-likelihoods computed in step 2.

2.3.3 MCMC prior densities and sampling strategy

We have developed and implemented a Bayesian method to estimate one or more evolutionary rate matrices from phylogenetic comparative data. Our primary objective is to provide a framework to incorporate uncertainty in the estimates of $\mathbf{R}$ as well as to build a flexible

model to study shifts in evolutionary integration across clades and over time. Our method requires a phylogenetic tree with branch lengths, continuous data for two or more traits for each tip species, and it uses Metropolis-Hastings Markov chain Monte Carlo (MCMC, Metropolis et al., 1953; Hastings, 1970).

We model the prior density for the vector of root values (**a**) as an uniform or normal distribution and we use an uniform sliding window proposal density to sample the root value for every trait simultaneously. In contrast, the prior density and sampling scheme for the evolutionary rate matrix requires more elaboration because variance-covariance matrices are positive definite and are relatively hard to be estimated. We model **R** with two independent distributions; one for the vector of standard deviations and another for the correlation matrix (Barnard et al., 2000; Zhang et al., 2006). This method allows the prior density for the rates (vector of standard deviations) to be parametrized independently of the evolutionary integration (correlation matrix). Under this parametrization, one can assign any distribution of positive real values to the vector of standard deviations (here we use an uniform or a exponential density) and the correlation matrix is modelled as the Cholesky decomposition of variance-covariance matrices sampled from an inverse-Wishart distribution (Zhang et al., 2006). This parameter extension approach is named 'separation-strategy' (Barnard et al., 2000; Zhang et al., 2006) because it relies on the independent modelling of the vector of standard deviations and the correlation matrix that together compose the evolutionary rate matrix. The advantage of the separation-strategy is twofold; it allows for intuitive modelling of rates of evolution and evolutionary integration and it is an efficient proposal scheme, because matrices are guaranteed to be positive definite at every draw (Barnard et al., 2000; Zhang et al., 2006).

2.3.4 Incorporating uncertainty in regime configurations and phylogenetic trees

Our approach can integrate any number of evolutionary rate matrix regimes fitted to the same phylogenetic tree. A regime is often dictated by some categorical data which states are hypothesized to be associated with shifts in the tempo and mode of evolution of the traits under study. Regimes are often 'painted' to the phylogenetic tree using stochastic mapping simulations (Huelsenbeck et al., 2003) and analyses are repeated over a sample of stochastic maps. In order to facilitate incorporation of uncertainty in both ancestral state estimates and phylogenetic inference, such as multiple phylogenetic trees sampled from a posterior distribution, we implemented a MCMC that integrates over multiple rate regime configurations and/or phylogenetic trees. At each step of the MCMC chain one phylogenetic

tree is randomly sampled from a pre-determined pool of trees and used to evaluate the likelihood of the model. The approach assumes that each phylogeny or regime configuration in the pool has equal chance to be sampled, but one can also assign the frequency of sampling as a result of a previous analysis. This pool is assumed to be gathered *a priori*, as a result of stochastic mapping simulations, samples from a posterior distribution of trees or other similar analyses. Although this method does incorporate the uncertainty related to alternative regime configurations, different topologies and set of branch lengths, it is not a joint estimation of the tree and the model because the MCMC only applies proposal steps to the vector of root values and the evolutionary rate matrices.

2.3.5 Testing for shifts between rate regimes

A useful criterion to perform model selection in a Bayesian framework is the Bayes factor (Kass and Raftery, 1995), which is a ratio between the marginal likelihoods of the competing models. However, the estimation of marginal likelihoods is a computationally expensive and contentious task. One of the most accurate methods to estimate the marginal likelihood is the stepping stone approach. This method consists of taking samples from a series of weighted posterior distributions by scaling the likelihood of the model so that a continuum between the prior and the posterior is created (Fan et al., 2011; Xie et al., 2011; Uyeda and Harmon, 2014). However, the stepping stone method adds significantly to the computation burden of the analysis, because each step of the continuum represents a complete MCMC chain and a large number of steps are required to produce a sufficient approximation of the marginal likelihood (Uyeda and Harmon, 2014).

Here, we do not use Bayes factor to compare models, although implementation is feasible for future work. We focus our interpretation of results on the distribution of posterior parameter estimates, and quantify the amount of uncertainty and the magnitude of the difference between components of the evolutionary rate matrices fitted to different regimes of the tree. We implemented summary statistics that provide a framework to decide whether there is enough signal in the data to support a model comprised by multiple $\mathbf{R}$ matrix regimes. First we check the difference between untransformed $\mathbf{R}$ matrices (ss-overall), then we contrast the vector of standard deviations (ss-rates) and the difference between correlation matrices (ss-correlation) derived from these $\mathbf{R}$ matrices. The first quantity check for overall changes in the evolutionary rate matrix whereas the later two quantities check for a shift in the rates of evolution of each individual trait and the structure of evolutionary integration among traits. We perform tests by calculating the percentile of the 0 value with respect to the distribution of the difference between summary statistics computed from the joint posterior distribution

of parameter estimates. If the 0 value is within 95% of the density, then there is significant overlap between the posterior distribution of the corresponding parameter estimates and we cannot reject that the rate regimes are likely samples from the same distribution.
The approach using summary statistics described here is justified by the fact that the models are nested. This means that it is possible to collapse the posterior distribution of evolutionary rate matrices fitted to different regions of the same phylogenetic tree to produce a single distribution if enough overlap is detected. For example, a model with three evolutionary rate matrix regimes can be reduced into a model with two regimes and so on. Thus, the summary statistics tests help us decide whether the posterior distribution between two parameters show enough overlap to justify their collapse into a single one. This argument extends to different attributes of the **R** matrix, such that one can collapse the rates of evolution of the traits into a single regime while accepting a shift in the pattern of evolutionary correlation among traits.

2.3.6 Simulation study

We performed simulations to test the performance of our Bayesian MCMC estimates and the use of summary statistics under different scenarios of correlated and uncorrelated evolution. For each simulation we used rejection sampling to generate a phylogenetic tree with 200 tips and at least one monophyletic clade containing 50 tips under a birth-death model. Then, we simulated data using a multivariate Brownian motion model for three continuous traits with two evolutionary rate matrix regimes, one for the 50 tips clade and another for the background group (Fig. 2.2). We performed four simulation scenarios: no shift (equal matrices), shift of orientation (positive versus negative evolutionary correlations), shift of rates (same evolutionary correlation but varying rates of evolution), and shift of integration (same rates but different degrees of evolutionary correlation). We applied two treatments for the scenario of shift of rates and shift of integration by varying the magnitude of the shifts. Figure 2.3 shows the total number of simulation treatments and their true parameter values. For all simulations we used a uniform prior for the vector of standard deviations, a marginally uniform prior for the correlation matrix (Barnard et al., 2000), and a multivariate normal prior for the vector of phylogenetic means centered in the mean of the tip data for each trait and with standard deviation equal to two times the standard deviation of the tip data (Fig. 2.4). We chose an informative prior for the phylogenetic mean in order to facilitate the convergence of the MCMC chains, since the root values are not the primary focus of this set of simulations. Nevertheless, we repeated a subset of the simulations using a uninformative prior assigned to the root values to show that the MCMC also performs

well under this scenario. For each simulation treatment we performed 100 replicates, each replicate composed by two independent MCMC chains of 500,000 generations. The initial state of every MCMC chain was set to a random draw from its prior distribution. We checked for convergence using the Gelman and Rubin (1992) test applied to each parameter of the model (each element of the root values, standard deviation vector, and correlation matrix was considered a separate parameter). We plotted the distribution of the percentiles of the true parameter values for the simulations compared to the posterior distributions to show the proportion of MCMC estimates that contained the true value of the simulation within the 95% highest probability density (HPD) interval. We simulated phylogenies, traits and mapped regimes using the R package `phytools` (Revell, 2012) and performed all parameter estimates with the package `ratematrix` (Caetano and Harmon, 2017b).

In order to check for congruence between our approach and maximum likelihood estimators, we used the R package `mvMORPH` (Clavel et al., 2015) to find the best model using likelihood ratio tests (one regime versus two regimes) for all simulated scenarios. We compared the results from the MCMC with the maximum likelihood estimates by calculating the percentile of the MLE estimates for the two regimes model with respect to the posterior distributions and checked whether the model favored by the likelihood ratio test also showed support when relying on the summary statistics computed from the posterior distribution of parameter estimates. The comparison between the likelihood ratio test and our posterior check approach is not a formal evaluation of model test performance, since the two approaches are fundamentally distinct. On the other hand, this serve as a pragmatic comparison to show whether we can adopt the use of summary statistics calculated from the posterior distribution to make reliable choices between models with direct incorporation of uncertainty in parameter estimates while retaining the explanation power of a more formal model testing approach.

2.3.7 Empirical examples

We use two examples to show the performance of the approach with empirical datasets and to further explore the impact of the direct incorporation of uncertainty in parameter estimates and model comparison. The first example tests for a shift in the evolutionary integration among anoles traits during the Caribbean radiation. Then, we repeat the analysis from Revell and Collar (2009) study on the evolution of buccal traits in Centrarchidae fishes.

Anoles are small lizards that live primarily in the tropics. There are nearly 400 anole species with diverse morphology and they have become a model system for studies of adaptive radiation and convergence (Losos, 2009; Mahler et al., 2013, and references therein). The ancestral

distribution of the genus is in Central and South America and the history of the clade includes island dispersion and radiation as well as dispersal back to the mainland (Nicholson et al., 2005; Losos, 2009). The adaptive radiation of anoles to the Caribbean islands and the repeated evolution of ecomorphs are the main focus of evolutionary studies in the genus (Mahler et al., 2010; Losos, 2009; Mahler et al., 2013). However, mainland anoles are distributed from the north of South America to the south of North America and show more species (60% of all species) than island anoles and equally impressive morphological diversity (Losos, 2009). Mainland and island anole species form distinct morphological clusters (Pinto et al., 2008; Schaad and Poe, 2010; Moreno-Arias and Calderón-Espinosa, 2016), but rates of trait evolution have been shown not to be consistently different (Pinto et al., 2008). Island ecomorphs can be readily distinguished by body size and the morphology of limbs, head and tail (Losos, 2009; Mahler et al., 2013). Thus, it is plausible that a shift in the structure of evolutionary integration among those traits associated with the radiation to the islands played an important role on the exploration of novel regions of the morphospace and allowed the repeated evolution of specialized morphologies. Herein we test this hypothesis by fitting two evolutionary rate matrix regimes, one for mainland and other for island anole lineages. We compiled data for snout-vent length (SVL), tail length (TL), and head length (HL) of 125 anole species (99 Caribbean and 26 mainland species) made available by Mahler et al. (2013) and Moreno-Arias and Calderón-Espinosa (2016). We chose this set of traits because they are important for niche partitioning among anoles (Pinto et al., 2008; Losos, 2009; Mahler et al., 2013) and also provided the best species coverage given the data currently available. We use Gamble et al. (2014) maximum clade credibility tree for all comparative analyses, but we trimmed the phylogeny to include only the species that we have morphological data. To map the different **R** matrix regimes to the phylogenetic tree we classified species as 'island' or 'mainland' and used the package `'phytools'` (Revell, 2012) to estimate the transition rates between the states in both directions using an unconstrained model (e.g., the 'all rates different' model) and to perform 100 stochastic mapping simulations. We set the model to estimate one **R** matrix for each mapped state ('island' or 'mainland') and we used the pool of 100 stochastic maps in the MCMC to take into account the uncertainty associated with ancestral state estimation. We ran four independent MCMC chains of 2 million generations each and used a random sample from the prior as the starting point of each chain. We set a uniform prior for the phylogenetic mean, a marginally uniform prior on the correlation matrices and a uniform prior on the vector of standard deviations for the **R** matrices. We discarded 25% of each MCMC chain as burn-in and checked for convergence using the potential scale reduction factor (Gelman and Rubin, 1992). In order to test the influence of the root state for the rate regimes, we repeated the analyses by setting the

root state for the stochastic mapping simulations as a random sample between 'island' and 'mainland' and with the ancestral distribution fixed as 'mainland'.

In addition to the analyses of mainland and island anoles lizards, we replicated the study by Revell and Collar (2009) as an exercise to contrast the inference of evolutionary rate matrices in the presence of a direct estimate of uncertainty provided by the posterior densities. Revell and Collar (2009) showed that the evolution of a specialized piscivorous diet in fishes of the genus *Micropterus* is associated with a shift towards a stronger evolutionary correlation between buccal length and gape width (see Fig. 1 in Revell and Collar, 2009). This tighter integration might have allowed *Micropterus* lineages to evolve better suction feeding performance. For this analysis we used the same data and phylogenetic tree made available by the authors. We set prior distributions using the same approach for the analysis of anole lizards described above. We also ran four MCMC chains starting from random draws from the prior for 1 million generations and checked for convergence using the potential scale reduction factor (Gelman and Rubin, 1992).

We provided scripts to reproduce simulations and analyses of the empirical data as Supplementary Material.

2.4 Results

2.4.1 Performance of the method

We ran a total of 1,200 Markov chain Monte Carlo chains to check the performance of the model under six different scenarios of correlated evolution among traits. All chains finished without errors, showed good convergence after 500,000 generations and results were congruent both with the true simulation parameters and with maximum likelihood estimates (Table 2.1 and Fig. 2.5). Figures 2.3 and 2.8 show examples of the posterior distribution of evolutionary rate matrices and root values for each simulation scenario. Changing the prior distribution for the vector of root values from multivariate normal to uniform showed no detectable bias in the posterior distribution, however the MCMC required approximately twice as many iterations to converge (Fig. 2.9). The distribution of percentiles for both the MLE and the true value for the simulations with respect to the posterior distribution of parameter estimates were, on average, within the 95% highest posterior density interval (Fig. 2.5). The likelihood ratio tests supported the two rates model about as often as our test based on summary statistics across all simulation scenarios (Table 2.1). When data was simulated with a single evolutionary rate matrix across the tree but tested for two regimes, both the likelihood ratio tests and the summary statistics (ss-overall) resulted in less than

5% of the 100 replicates with support for the wrong model.

Alternatively, one might be interested on shifts in some of the attributes of the evolutionary rate matrices more than others, such as specific hypotheses about the change in the pattern of evolutionary integration without *a priori* expectations about shifts in the rates of trait evolution. Table 2.1 shows the results of the summary statistics approach with respect to different attributes of the evolutionary rate matrices fitted to the data. These results are congruent with the simulation scenarios and show that the approach using summary statistics calculated from the posterior distribution of parameter estimates is a reliable and flexible way to identify changes in rates of trait evolution (ss-rates) or shifts in the pattern of evolutionary integration (ss-correlation).

2.4.2 Empirical examples

The biogeographic reconstruction using a more recent anole phylogeny (Gamble et al., 2014) is mostly congruent with previous studies (Glor et al., 2005; Nicholson et al., 2005; Losos, 2009). There are multiple radiations from mainland South America to the Caribbean islands and a single radiation from the islands back to mainland South America (Fig. 2.6). In contrast, the Jamaican clade (*A. reconditus* + *A. grahami*), that previous results have shown to be sister to the clade that dispersed from the Caribbean islands back to mainland (Losos, 2009), is now nested within this secondary radiation. These results are maintained when we used all species from Gamble et al. (2014) instead of the trimmed tree (see Fig. 2.10). The **R** matrix estimates for each regime show no difference in the structure of integration but the rates of evolution for the Caribbean anole lineages are twice as fast as mainland lineages (Fig. 2.6, see also figures 2.11 for the posterior of root values and 2.12 for trace plots). In other words, the evolutionary rate matrix for the two regimes are proportional (ss-overall=0.004, ss-rates=0.0002, and ss-correlation=0.4). Setting the root state as 'mainland' does not influence the posterior distribution of parameter estimates. Head length and tail length are positively correlated along the phylogeny and also show a strong positive evolutionary correlation with body size.

In the case of the Centrarchidae fishes, there is a clear distinction between the results of the maximum likelihood point estimate and the Bayesian estimate of the evolutionary rate matrix regimes. Under MLE, we found a significant difference between the **R** matrix regimes using likelihood ratio tests. There is a stronger evolutionary correlation between the gape width and the buccal length of the *Micropterus* clade (r=0.83) when compared with other lineages (r=0.36). In contrast, the direct incorporation of uncertainty in parameter estimates reveal an important overlap between the posterior densities for the **R** matrices estimated for

each regime (Fig. 2.7, see also figures 2.13 for the posterior of root values and 2.14 for trace plots). The posterior density does not show evidence of a shift towards stronger evolutionary correlation between gape width and buccal length in *Micropterus* (ss-correlation=0.46) and the overall overlap between the posterior of evolutionary rate matrix fitted to each regime is pronounced (ss-overall=0.58). Thus, after taking the uncertainty in parameter estimates into account, it is unlikely that a shift on the pattern of evolutionary correlation happened in the *Micropterus* clade.

2.5 Discussion

Here we implemented a Bayesian Markov chain Monte Carlo estimate of the evolutionary rate matrix. Our approach allows multiple regimes to be fitted to the same phylogenetic tree and integrates over a sample of trees or regime configurations to account for uncertainty in ancestral state estimates and phylogenetic inference. We also implement summary statistics to compare the posterior distribution of parameter estimates for different regimes. We show that our approach has good performance over a series of different scenarios of evolutionary integration and is congruent with parameter estimates using maximum likelihood. The use of maximum likelihood estimate is definitely faster, since the MCMC chain requires many more evaluations of the likelihood function. However, our new extension of Felsenstein (1973) pruning algorithm applied when multiple $\mathbf{R}$ matrices are fitted to the same tree reduces the computation time of the likelihood for the model significantly. The integration of uncertainty in parameter estimates provided by the posterior distribution and the use of summary statistics to describe patterns in the data that can be directly relevant to our biological predictions are significant rewards for the longer time invested in data analysis.
The use of summary statistics to evaluate the overlap between the posterior distributions of parameter estimates from different regimes is a intuitive and reliable framework to make decisions of whether or not the data show a strong signal for multiple regimes. Our simulations showed that results from this approach are, in average, congruent with the likelihood ratio test. More importantly, summary statistics computed from the posterior distribution can recognize meaningful discrepancies between distinct evolutionary rate matrix regimes across a series of simulation scenarios. In this study we focused on the evolutionary rates for each trait (ss-rates) and the evolutionary correlation among traits (ss-correlation), but any other summary statistics computed over the posterior distribution of parameter estimates and representing an attribute of the model relevant for a given question could be implemented. For example, characteristics of the eigen-structure of the matrices or more formal tests such as the Flury hierarchy (Phillips and Arnold, 1999) could be also implemented.

This framework is flexible, does not require an estimate of the marginal likelihood and can be easily tailored towards specific biological predictions of the study system. On the other hand, it is important to note that the use of summary statistics does not constitute a formal model test, but instead asks the question of whether the parameter estimates for the regimes are distinct enough for us to accept the hypothesis of heterogeneity in the tempo and mode of trait evolution.

Point estimates such as the maximum likelihood can generate a false impression of certainty that may limit our biological interpretations if not accompanied by estimates of the variance. It is possible to calculate the confidence interval around the MLE and use this interval to check for overlaps in the parameter estimates (i.e., using the Hessian matrix). The disadvantage of this approach is that the confidence interval provides only the percentiles of the density around the MLE. As a result, it is not possible to calculate summary statistics, incorporate uncertainty in downstream analyses, or to provide a visualization of the distribution of parameter estimates such as in this study (see Figures 2.3 and 2.7). Maximum likelihood estimate of models of trait evolution are commonly reported without any estimate of the variance, most likely because the focus are often on the results of model tests and p values rather than in our ability to reliably estimate and interpret the parameters of a model (see discussion in Beaulieu and O'Meara, 2016 on a related issue). Furthermore, model tests such as the likelihood ratio test and the Akaike information criteria (AIC) do not incorporate any measure of the variance of estimates in their calculations. This is problematic when parameters can be hard to estimate and models are challenged by reduced sample sizes, which is a common issue in phylogenetic comparative methods analyses in general.

The analysis of mouth shape evolution in function of diet in Centrarchidae fishes (Revell and Collar, 2009) is an interesting example of the impact of uncertainty in parameter estimates on our biological conclusions. The likelihood ratio test showed a strong support for a shift in the structure of evolutionary correlation associated with the evolution of piscivory in the *Micropterus* clade. In contrast, the summary statistics computed from the posterior distribution did not show strong evidence for the same scenario of macroevolution. When we contrast the result from the MLE with the posterior distribution (Fig. 2.15), we can visualize the origin of the incongruence. The likelihood ratio test focus on the relative fit of the constrained model (one regime) compared with the full model (two regimes) whereas the summary statistics compute whether our posterior knowledge about the model reflects a strong signal for a shift between regimes using the overlap between the posterior distribution of parameter estimates. Furthermore, the same trend can be shown by computing the confidence interval around the MLE estimates, since there is an important overlap between the $\mathbf{R}$ matrices fitted to each regime.

The results from the test of whether mainland and island anole species differ in the pattern of evolutionary integration among traits are intriguing. The posterior distribution of evolutionary rate matrices fitted to each regime show a constant pattern of evolutionary integration whereas rates of trait evolution are faster on island anole lineages. The radiation of anole lizards on the Caribbean islands is one of the most striking examples of adaptive radiation in evolutionary biology. It is natural to expect that a shift in the trajectory of evolution of morphological traits associated with ecomorphs would occur, since mainland and island anole species are known to occupy different regions of the morphospace (Pinto et al., 2008; Schaad and Poe, 2010; Moreno-Arias and Calderón-Espinosa, 2016). Surprisingly, our results suggest that ecomorphs evolved under a constant pattern of evolutionary integration among traits when compared with mainland lineages but differ due to faster rates of trait evolution. One hypothesis is that the evolutionary correlation among traits, which determine the major axes of morphological evolution in the group, do not act as a constraint to the exploration of the morphospace by the lineages. Thus, island and mainland anole lineages are not distinct in their potential to explore the morphospace and ecomorphs might be special in the sense of repetitive radiations and not due to exclusive morphological evolution when compared to their mainland counterparts. This explanation has some support by the fact that a few mainland species are morphologically similar to island ecomorphs (Schaad and Poe, 2010). In contrast, higher rates of evolution is most likely a reflection of the rapid morphological differentiation observed on the Caribbean anole lineages and associated with the ecological opportunity posed by the new island habitats coupled to the reduction in predation risk. Our results corroborate the idea that ecomorphs might also have evolved among mainland species since there is no detectable shift in the trajectory of evolution among morphological traits. However, efforts to understand anole biodiversity, ecology and evolution have been strongly focused on island systems and still relatively very little is known about mainland lineages.

2.6 Conclusion

Most of what we know about the tempo and mode of trait evolution come from studies of individual traits, but evolutionary integration is ubiquitous across the tree of life. Recently we have seen an increase in comparative tools aimed to deal with the challenges posed by high-dimensional traits, such as shape data. However, the discipline is still in need of better models to deal with multiple traits, such as the examples explored in this study. Our framework is aimed primarily on the test of shifts in the structure of evolutionary integration among traits across clades and over time. However, the implementation of summary statistics

make it feasible to extend such tests to be focused on any attribute of the evolutionary rate matrix that might fit the biological predictions of a specific study. Another important advantage of simulation based approaches, such as the Bayesian MCMC, is that proposals can be modified to integrate over different number of regime configurations, distinct models of trait evolution, and even simultaneously estimate parameters for the trait evolution model and the phylogenetic tree. Thus, our implementation lays the groundwork for future advancements towards flexible models to explore multiple facets of the evolution of integration over long time scales using phylogenetic trees. Integration among traits is a broad and yet fundamental topic in evolutionary biology. Understanding the interdependence among traits over the macroevolutionary scale can be key to tie together our knowledge about the genetic basis of traits, development, and adaptive shifts in the strength or direction of evolutionary correlation.

Table 2.1: Proportion of simulation replicates showing support for two **R** matrix regimes under likelihood ratio test (LRT) and using summary statistics computed from the posterior distribution of parameter estimates. The 'ss-overall' summary statistics compares the entire evolutionary rate matrix, 'ss-rates' refers to the rates of evolution for the individual traits and 'ss-correlation' represents only the structure of evolutionary correlation among traits. Simulations were performed with no shift (Single), shift of orientation (Orient), weak shift of rates (Rates I), strong shift of rates (Rates II), weak shift of integration (Integ I), and strong shift of integration (Integ II). Figure 2.3 show the true value for each simulation and a plot of the posterior distribution of one simulation replicate and 2.8 show the posterior distribution of root values.

	Single	Orient	Rates I	Rates II	Integ I	Integ II
LRT	0.04	1	1	1	0.25	0.98
ss-overall	0.02	1	0.85	1	0.28	0.84
ss-rates	0.01	0.04	0.98	1	0.03	0.02
ss-correlation	0.01	1	0.06	0.03	0.26	1

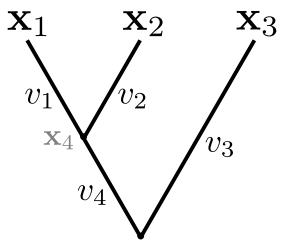


Figure 2.1: Example of phylogeny used to compute the likelihood of a multivariate Brownian-motion model using the new pruning algorithm.

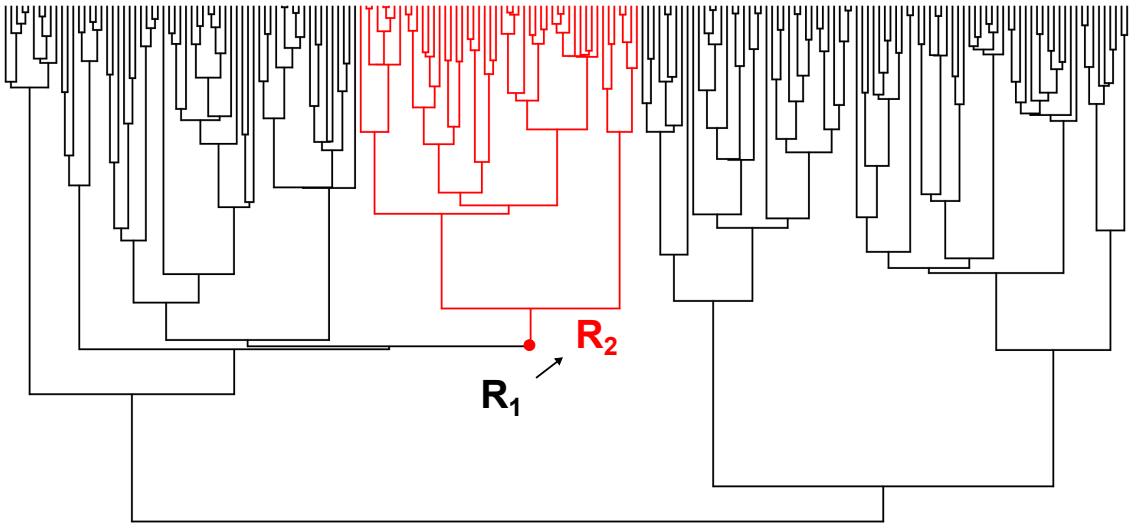


Figure 2.2: Example of phylogeny used for the simulation study. We simulated phylogenies with 200 tips using a homogeneous birth-death model. Then, we randomly selected one node with exact 50 daughter tips to set the location of the transition between the background rate regime $\mathbf{R_1}$ and the focus clade regime $\mathbf{R_2}$ showed in red.

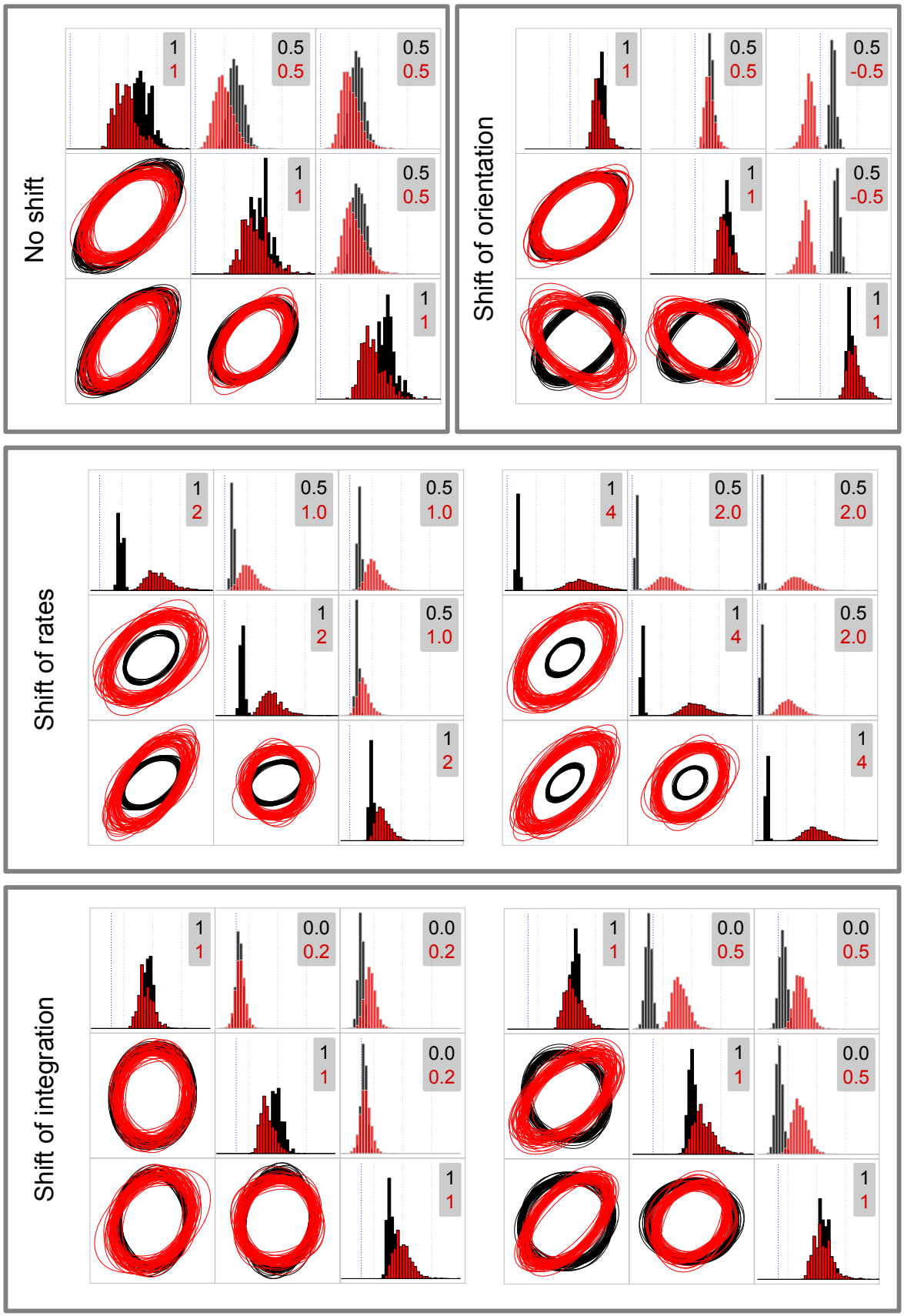

No shift
1
1
0.5
0.5
0.5
0.5
1
1
0.5
0.5
1
1
0.5
0.5
1
1
Shift of orientation
1
1
0.5
0.5
0.5
-0.5
1
1
0.5
-0.5
1
1
Shift of rates
1
2
0.5
1.0
0.5
1.0
1
2
0.5
1.0
1
2
1
4
0.5
2.0
0.5
2.0
1
4
0.5
2.0
1
4
Shift of integration
1
1
0.0
0.2
0.0
0.2
1
1
0.0
0.2
1
1
1
1
0.0
0.5
0.0
0.5
1
1
0.0
0.5
1
1

Figure 2.3: Example of posterior distribution for the six simulation treatments with three traits each. Top-left plot shows the results with no shift in the evolutionary rate matrix regime and top-right shows the results with a shift in the orientation of **R**. Middle row are results with a shift in the rates of evolution of each trait and bottom row shows the results when the strength of the evolutionary correlation shifts between regimes. Estimates for the background regime are showed in black and for the focus regime in red (see Fig. 2.2). For each plot: diagonal histograms show evolutionary rates (variances) for each trait, upper-diagonal histograms show pairwise evolutionary covariation (covariances), and lower-diagonal ellipses are samples from the posterior distribution showing the 95% confidence interval of each bivariate distribution. Numbers in the top left of histograms are the true value used for each simulation; background rate regimes are showed in black and focus clade regimes in red. Table 2.1 shows the aggregate results for each simulation replicate: 'Single' and 'Orient' correspond to top-left and top-right plots. 'Rates' I and II are middle row left and right plots. 'Integ' I and II are bottom row left and right plots. The two replicates in the middle and bottom rows differ in the strength of the shift between regimes, left is weak and right is strong shift.

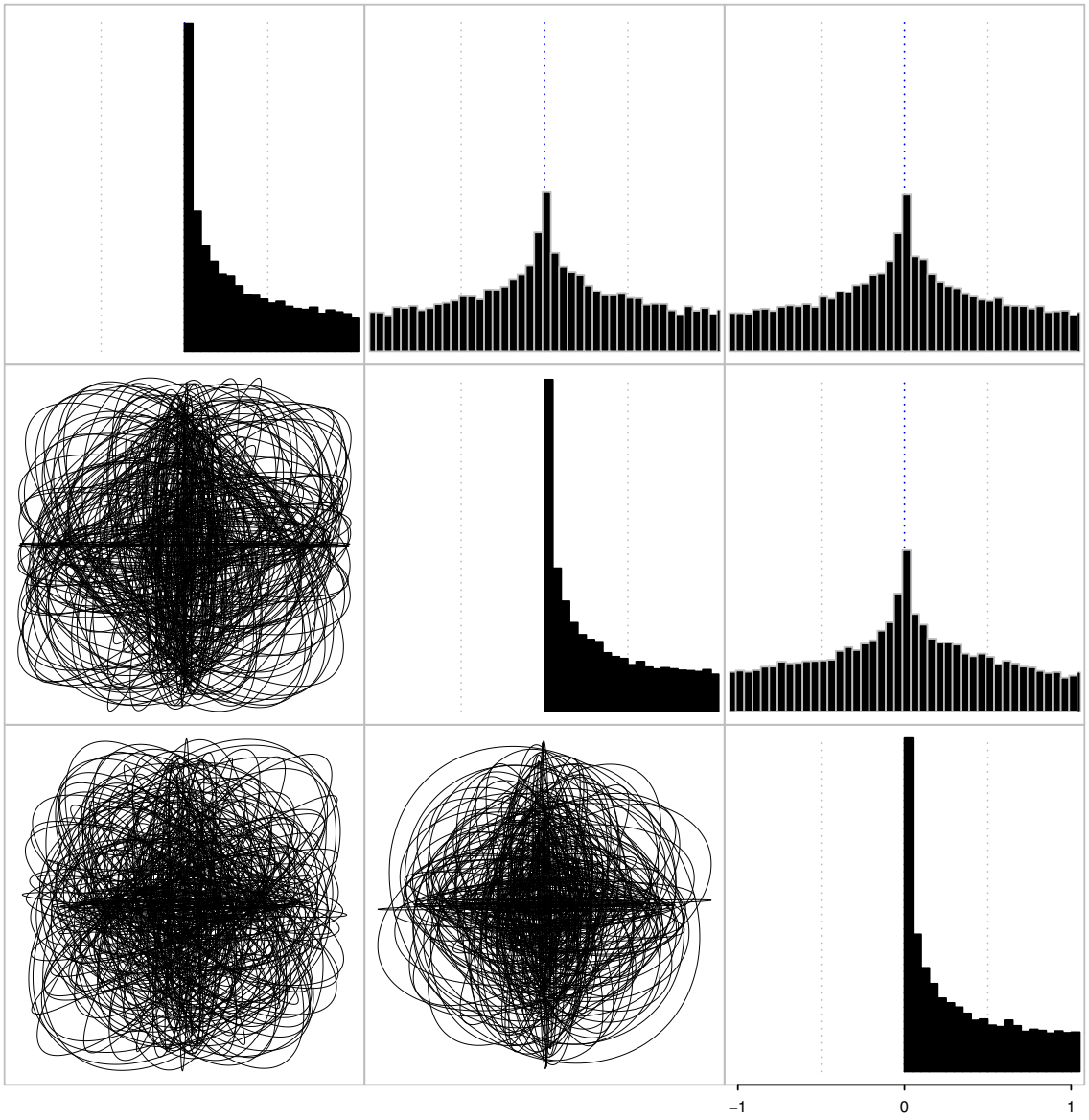


Figure 2.4: Prior distribution for the evolutionary rate matrix (**R**) used for all analyses. Plate shows samples in the interval between -1 and 1 from the prior for a model with three traits. Diagonal plots represent the prior for evolutionary rates (variances) for each trait, upper-diagonal plots show pairwise evolutionary covariation (covariances), and lower-diagonal are samples from the posterior distribution of ellipses showing the 95% confidence interval of each bivariate distribution.

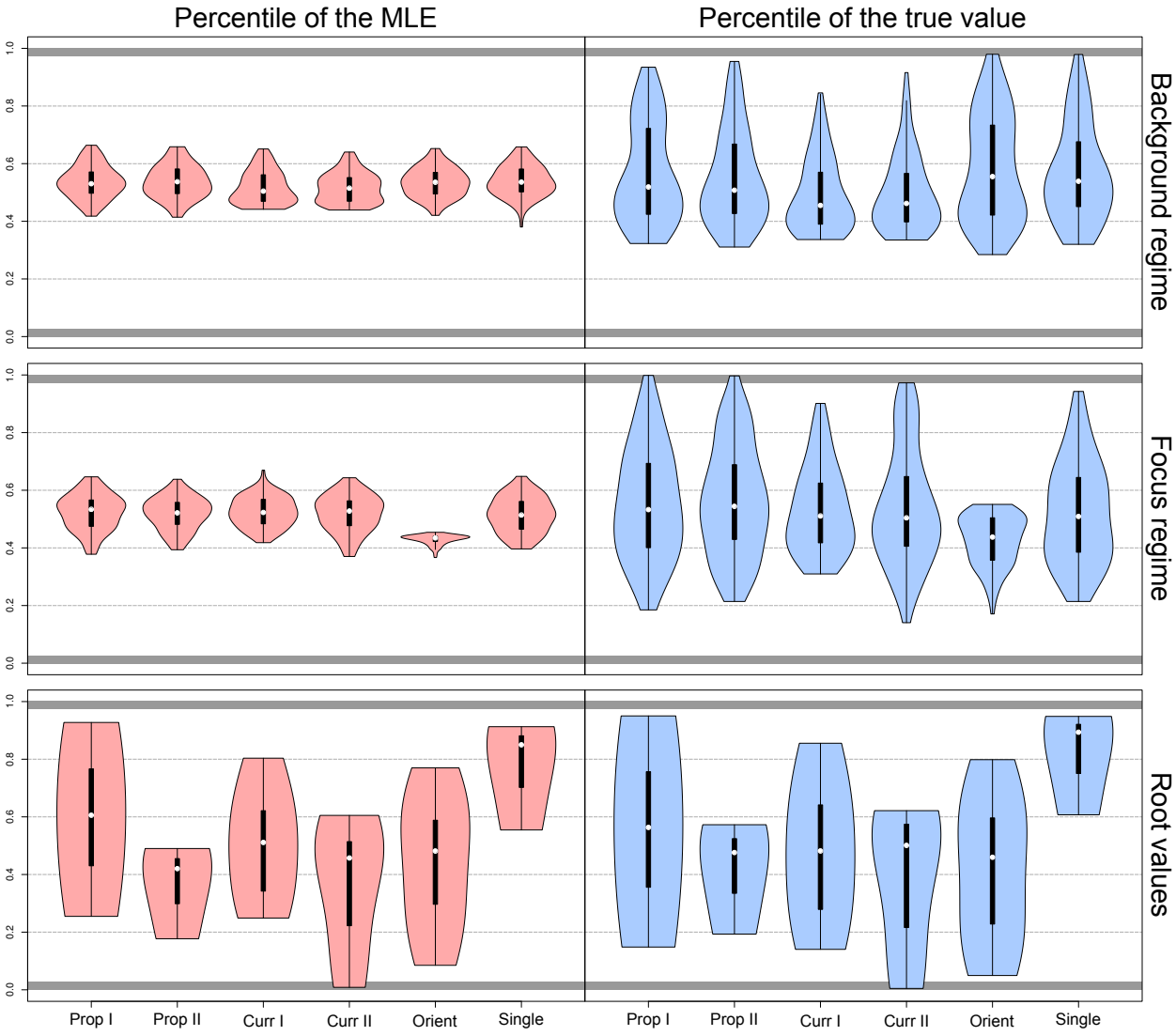


Figure 2.5: Distribution of percentiles for the maximum likelihood estimate (MLE) of the full model and for the true value of the simulations with respect to the posterior distribution of each simulation replicate. Plots to the left (pink) show the percentiles for the MLE whereas plots to the right (blue) show the percentiles for the true value of the simulations. Most of the density across all simulation scenarios and parameters is within the 95% HPD of the posterior distribution.

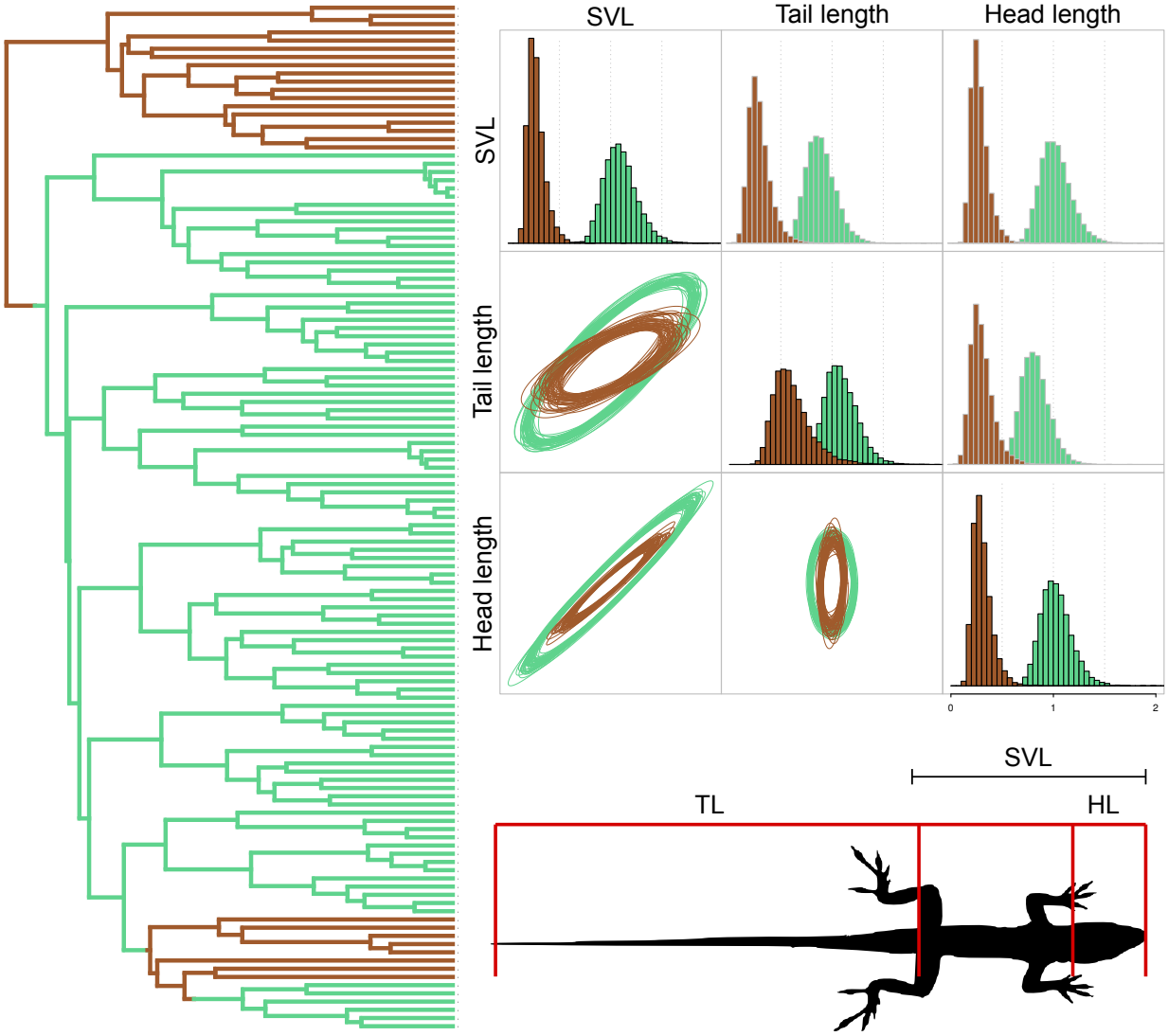


Figure 2.6: Posterior distribution of the **R** matrix regimes fitted to the island anole (green) and mainland anole (brown) lineages. Left figure shows the maximum clade credibility tree (MCC) from Gamble et al. (2014) with only the taxa used in this study. State reconstruction for the branches was performed with a stochastic map simulation using 'mainland' as the root state for the genus. Right upper plot shows the posterior distribution of parameter estimates for the evolutionary rate matrices. Diagonal plots show evolutionary rates (variances) for each trait, upper-diagonal plots show pairwise evolutionary covariation (covariances), and lower-diagonal are samples from the posterior distribution of ellipses showing the 95% confidence interval of each bivariate distribution. Right bottom figure shows a representation of each trait (TL: tail length; HL: head length; SVL: snout-vent length).

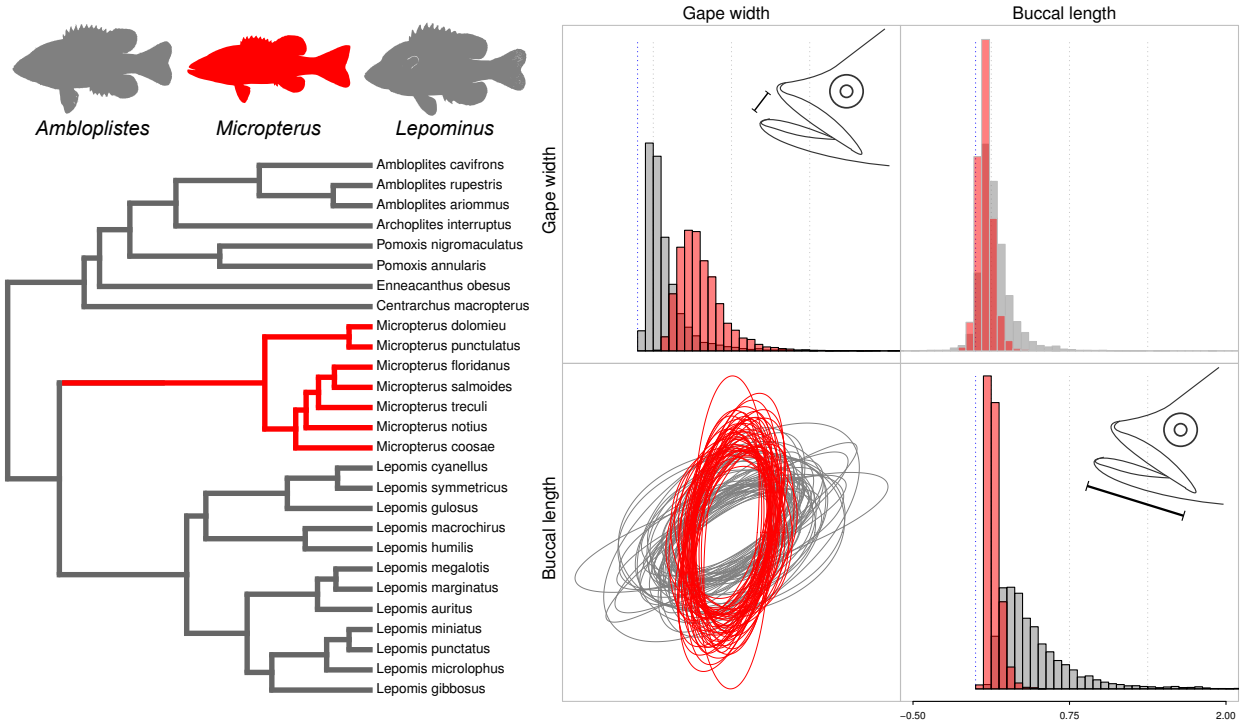


Figure 2.7: Posterior distribution of the **R** matrix regimes fitted to the background group (gray) and to the *Micropterus* clade (red). Left figure shows the phylogeny from (Revell and Collar, 2009) and the silhouette of some representatives of the Centrarchidae genera. Right plot shows the posterior distribution of parameter estimates for the evolutionary rate matrices. Diagonal plots show evolutionary rates (variances) for each trait, upper-diagonal plots show pairwise evolutionary covariation (covariances), and lower-diagonal are samples from the posterior distribution of ellipses showing the 95% confidence interval of each bivariate distribution.

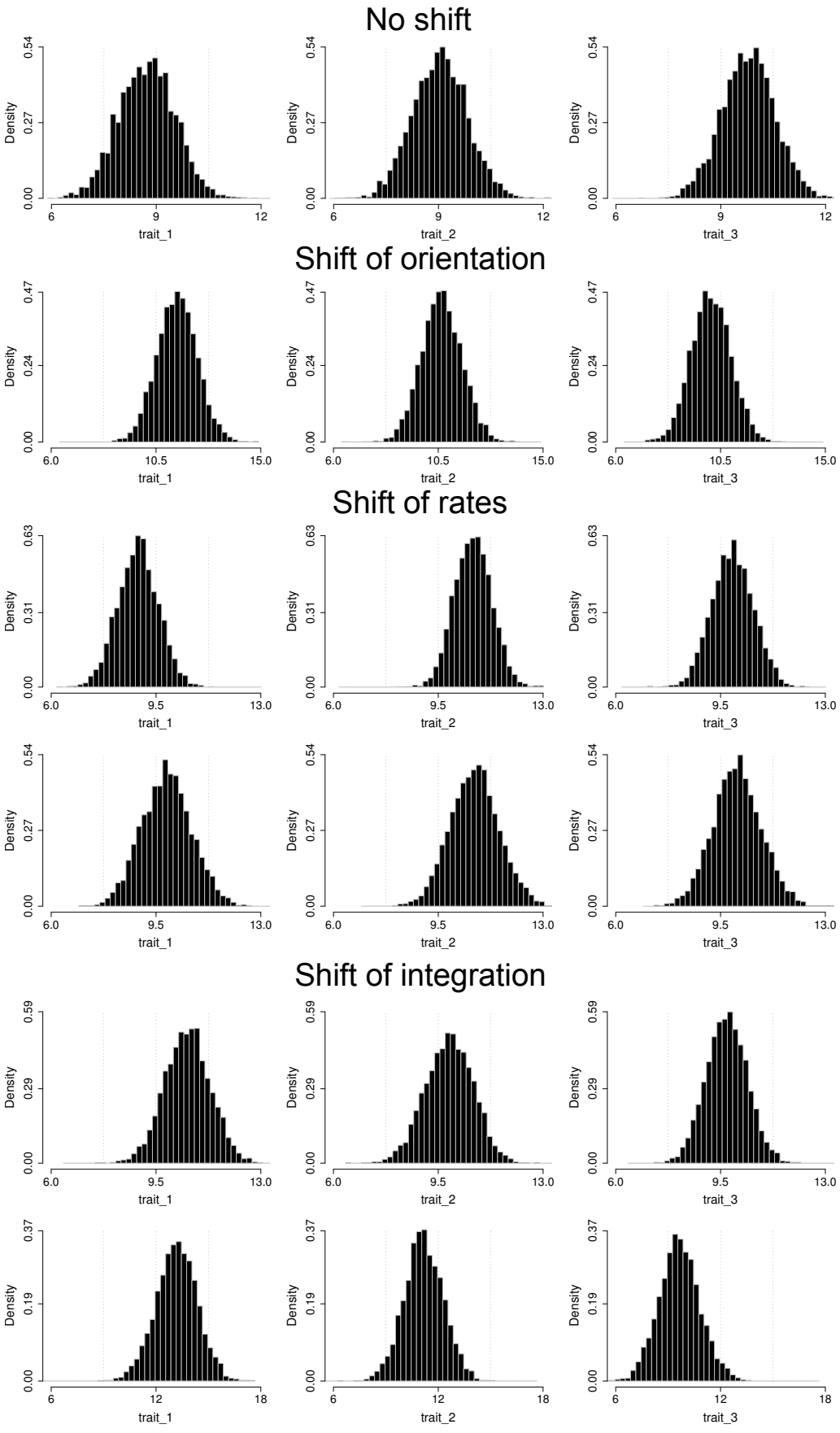

No shift
Shift of orientation
Shift of rates
Shift of integration
Density
trait_1
trait_2
trait_3

Figure 2.8: Example of posterior distribution of root values for the six simulation treatments with three traits each. Simulation treatments are the same as showed on Figure 2.3. Top and bottom plots for 'Shift of rates' and 'Shift of integration' treatments correspond to the left and right plots of the same treatments on Figure 2.3, respectively. The true value for the ancestral state of each trait in all simulations was equal to 10.

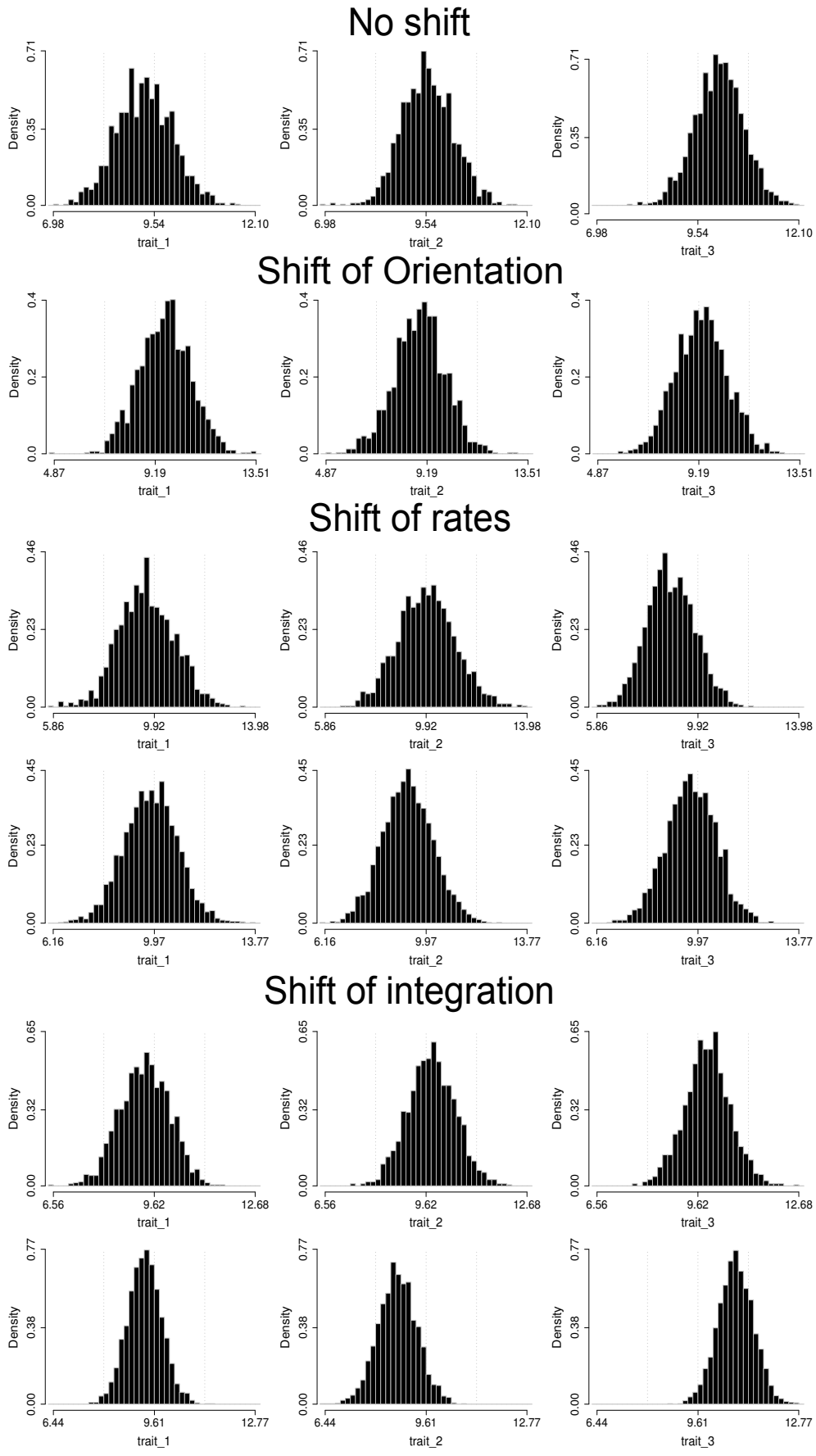

No shift
Density
0.00
0.35
0.71
6.98
9.54
12.10
trait_1
trait_2
trait_3
Shift of Orientation
0.0
0.2
0.4
4.87
9.19
13.51
Shift of rates
0.23
0.46
5.86
9.92
13.98
0.45
6.16
9.97
13.77
Shift of integration
0.32
0.65
6.56
9.62
12.68
0.38
0.77
6.44
9.61
12.77

Figure 2.9: Example of posterior distribution of root values for the six simulation treatments with three traits each using a uniform prior for the vector of root values. Simulation treatments are the same as showed on Figure 2.3 and Figure 2.8. The true value for the ancestral state of each trait in all simulations was equal to 10.

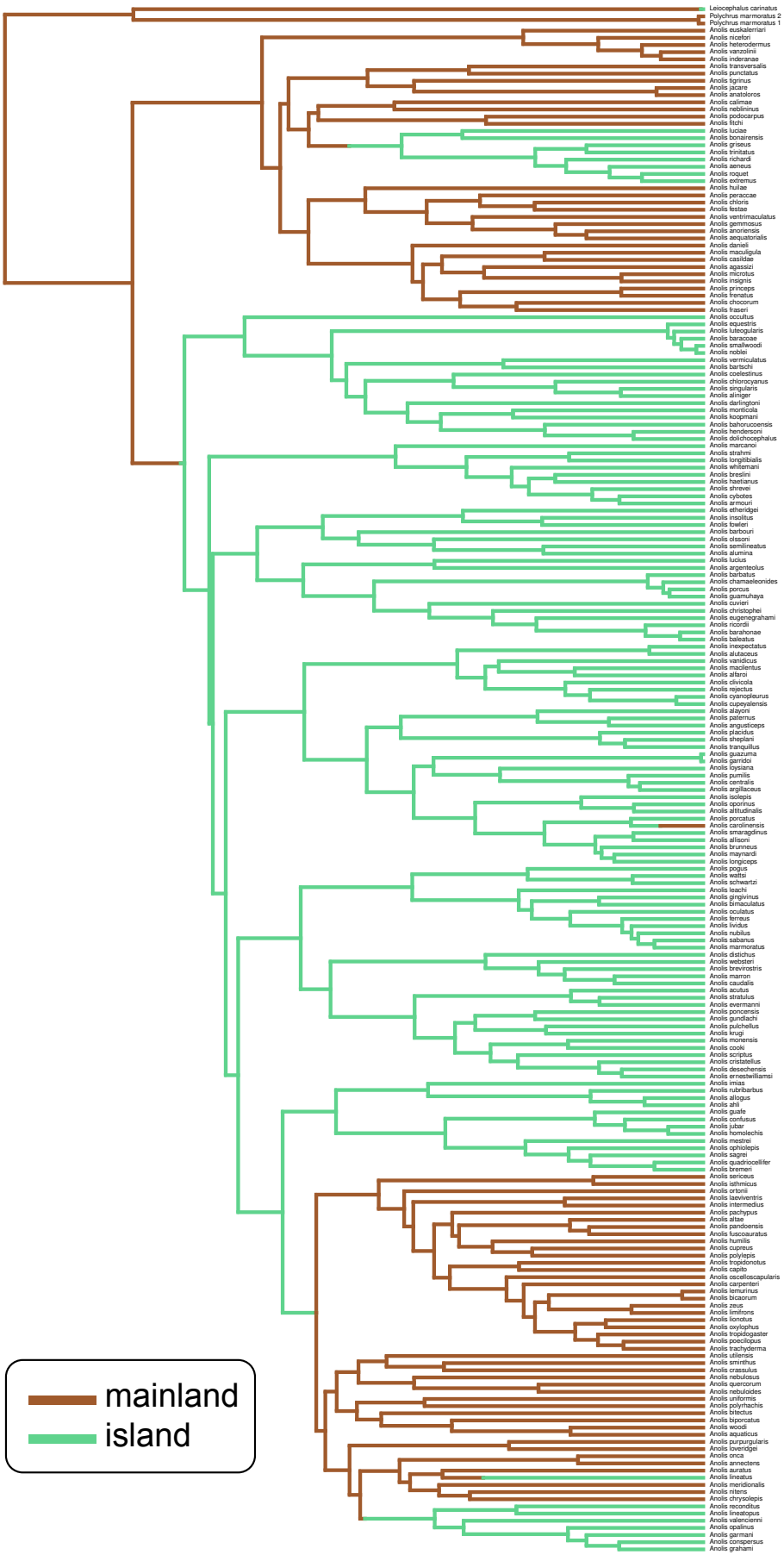

Leiocephalus carinatus
Polychrus marmoratus 2
Polychrus marmoratus 1
Anolis euskalerriari
Anolis nicefori
Anolis heterodermus
Anolis vanzolinii
Anolis inderenae
Anolis transversalis
Anolis punctatus
Anolis tigrinus
Anolis lacaze
Anolis anatoloros
Anolis calimae
Anolis neblininus
Anolis podocarpus
Anolis litchi
Anolis luciae
Anolis bonairensis
Anolis griseus
Anolis trinitatus
Anolis richardi
Anolis aeneus
Anolis roquet
Anolis extremus
Anolis huilae
Anolis peraccae
Anolis chloris
Anolis festae
Anolis ventrimaculatus
Anolis gemmosus
Anolis ancoriensis
Anolis aequatorialis
Anolis danieli
Anolis maculigula
Anolis casildae
Anolis agassizi
Anolis microtus
Anolis insignis
Anolis princeps
Anolis frenatus
Anolis chocorum
Anolis fraseri
Anolis occultus
Anolis equestris
Anolis luteogularis
Anolis baracoae
Anolis smallwoodi
Anolis noblei
Anolis vermiculatus
Anolis bartschi
Anolis coelestinus
Anolis chlorocyanus
Anolis singularis
Anolis aliniger
Anolis darlingtoni
Anolis monticola
Anolis koopmani
Anolis bahorucoensis
Anolis hendersoni
Anolis dolichocephalus
Anolis marcanoi
Anolis strahmi
Anolis longitibialis
Anolis whitemani
Anolis breslini
Anolis haetianus
Anolis shrevei
Anolis cybotes
Anolis armouri
Anolis etheridgei
Anolis insolitus
Anolis fowleri
Anolis barbouri
Anolis olssoni
Anolis semilineatus
Anolis alumina
Anolis lucius
Anolis argenteolus
Anolis barbatus
Anolis chamaeleonides
Anolis porcus
Anolis guamuhaya
Anolis cuvieri
Anolis christophei
Anolis eugenegrahami
Anolis ricordii
Anolis barahonae
Anolis baleatus
Anolis inexpectatus
Anolis alutaceus
Anolis vanidicus
Anolis macilentus
Anolis alfaroi
Anolis clivicola
Anolis rejectus
Anolis cyanopleurus
Anolis cupeyalensis
Anolis alayoni
Anolis paternus
Anolis angusticeps
Anolis placidus
Anolis sheplani
Anolis tranquillus
Anolis guazuma
Anolis garridoi
Anolis loysiana
Anolis pumilis
Anolis centralis
Anolis argillaceus
Anolis isolepis
Anolis sporinus
Anolis altitudinalis
Anolis porcatus
Anolis carolinensis
Anolis smaragdinus
Anolis allisoni
Anolis brunneus
Anolis maynardi
Anolis longiceps
Anolis pogus
Anolis wattsi
Anolis schwartzi
Anolis leachi
Anolis gingivinus
Anolis bimaculatus
Anolis oculatus
Anolis ferreus
Anolis lividus
Anolis nubilus
Anolis sabanus
Anolis marmoratus
Anolis distichus
Anolis websteri
Anolis brevirostris
Anolis marron
Anolis caudalis
Anolis acutus
Anolis stratulus
Anolis evermanni
Anolis poncensis
Anolis gundlachi
Anolis pulchellus
Anolis krugi
Anolis monensis
Anolis cooki
Anolis scriptus
Anolis cristatellus
Anolis desechensis
Anolis ernestwilliamsi
Anolis imias
Anolis rubribarbus
Anolis allogus
Anolis ahli
Anolis guafe
Anolis confusus
Anolis jubar
Anolis homolechis
Anolis mestrei
Anolis ophiolepis
Anolis sagrei
Anolis quadriocellifer
Anolis bremeri
Anolis sericeus
Anolis isthmicus
Anolis ortonii
Anolis laeviventris
Anolis intermedius
Anolis pachypus
Anolis altae
Anolis pandoensis
Anolis fuscoauratus
Anolis humilis
Anolis cupreus
Anolis polylepis
Anolis tropidonotus
Anolis capito
Anolis oscelloscapularis
Anolis carpenteri
Anolis lemurinus
Anolis bicaorum
Anolis zeus
Anolis limifrons
Anolis lionotus
Anolis oxylophus
Anolis tropidogaster
Anolis poecilopus
Anolis trachyderma
Anolis utilensis
Anolis smithius
Anolis crassulus
Anolis nebulosus
Anolis quercorum
Anolis nebuloides
Anolis uniformis
Anolis polyrhachis
Anolis bitectus
Anolis biporcatus
Anolis woodi
Anolis aquaticus
Anolis purpurgularis
Anolis loveridgei
Anolis onca
Anolis annectens
Anolis auratus
Anolis lineatus
Anolis meridionalis
Anolis nitens
Anolis chrysolepis
Anolis reconditus
Anolis lineatopus
Anolis valencienni
Anolis opalinus
Anolis garmani
Anolis conspersus
Anolis grahami
mainland
island

Figure 2.10: Maximum clade credibility tree from Gamble et al. (2014) study showing the distribution of 'mainland' and 'island' anole species. The species 'sp_nov_1', 'sp_nov_2', and 'sp_nov_3' were excluded from the phylogenetic tree. Data for the distribution of anole species and outgroups were compiled from Nicholson et al. (2005), Losos (2009), Thomas et al. (2009), Reptile database (reptile-database.org) and GBIF (gbif.org). Ancestral state reconstruction was performed using stochastic mapping with the 'all rates different' model and the root state set as 'mainland' (Nicholson et al., 2005; Losos, 2009).

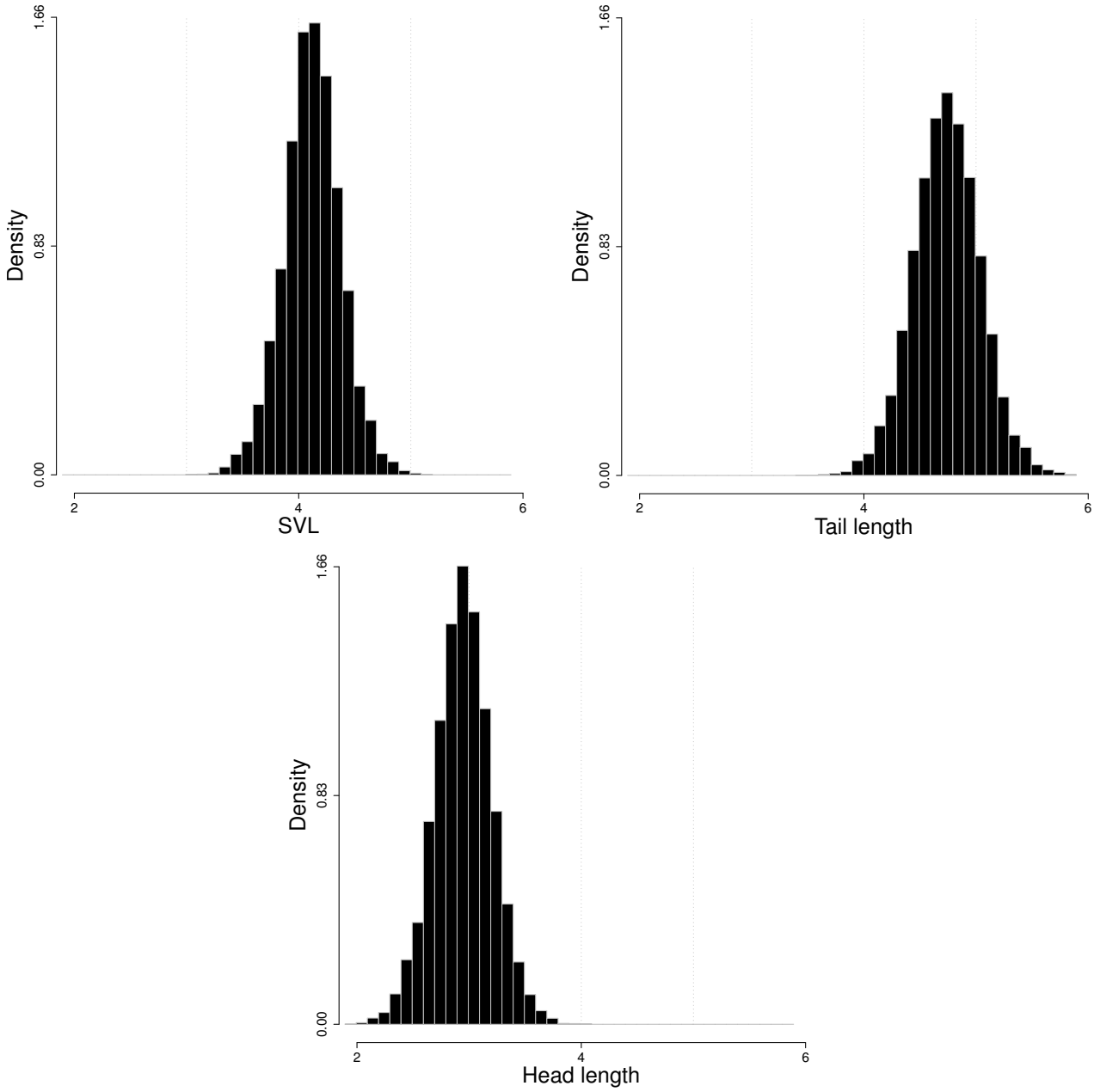


Figure 2.11: Posterior distribution of root values fitted to the island and mainland anole lineages (SVL: snout-vent length).

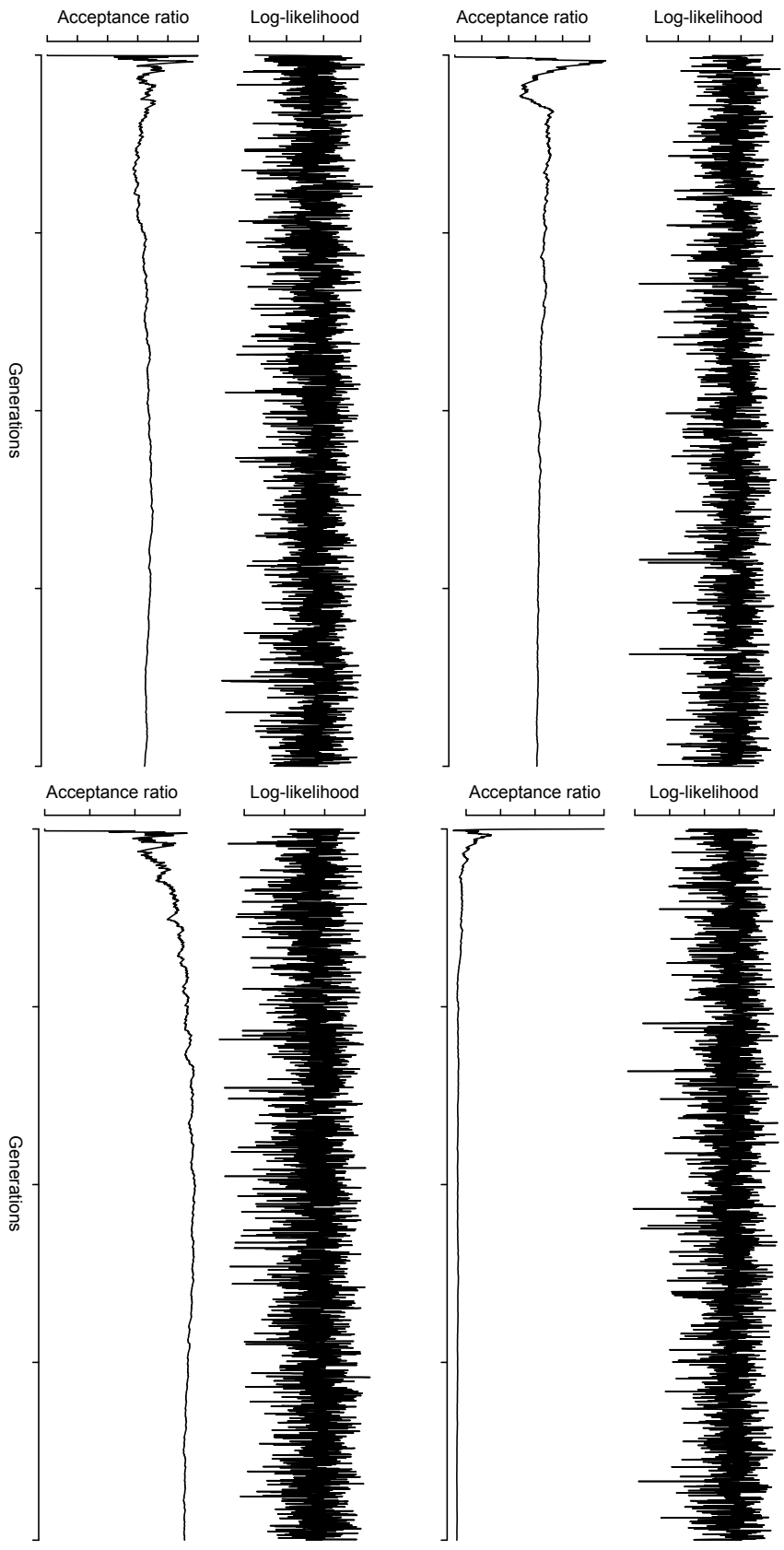


Figure 2.12: Trace plots of the log-likelihood and the acceptance ratio for the four independent MCMC chains of the island and mainland anole analysis.

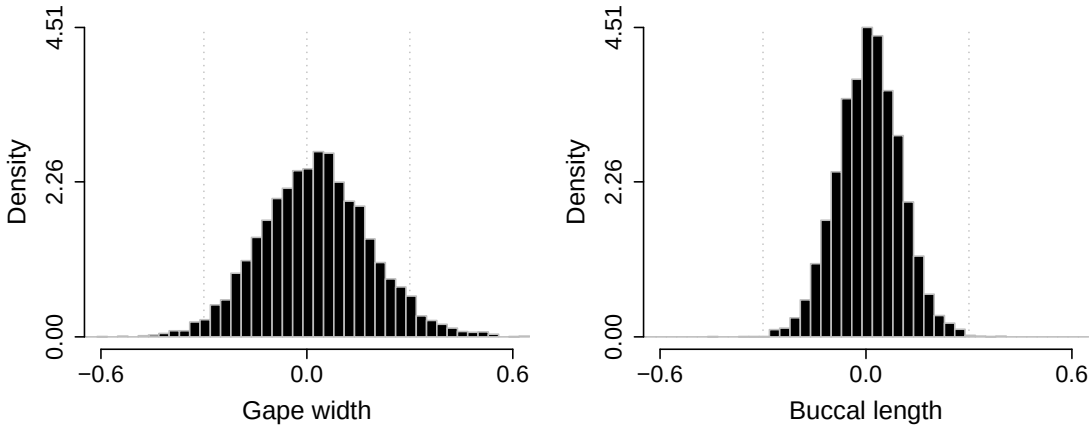


Figure 2.13: Posterior distribution of root values fitted to the Centrarchidae fishes.

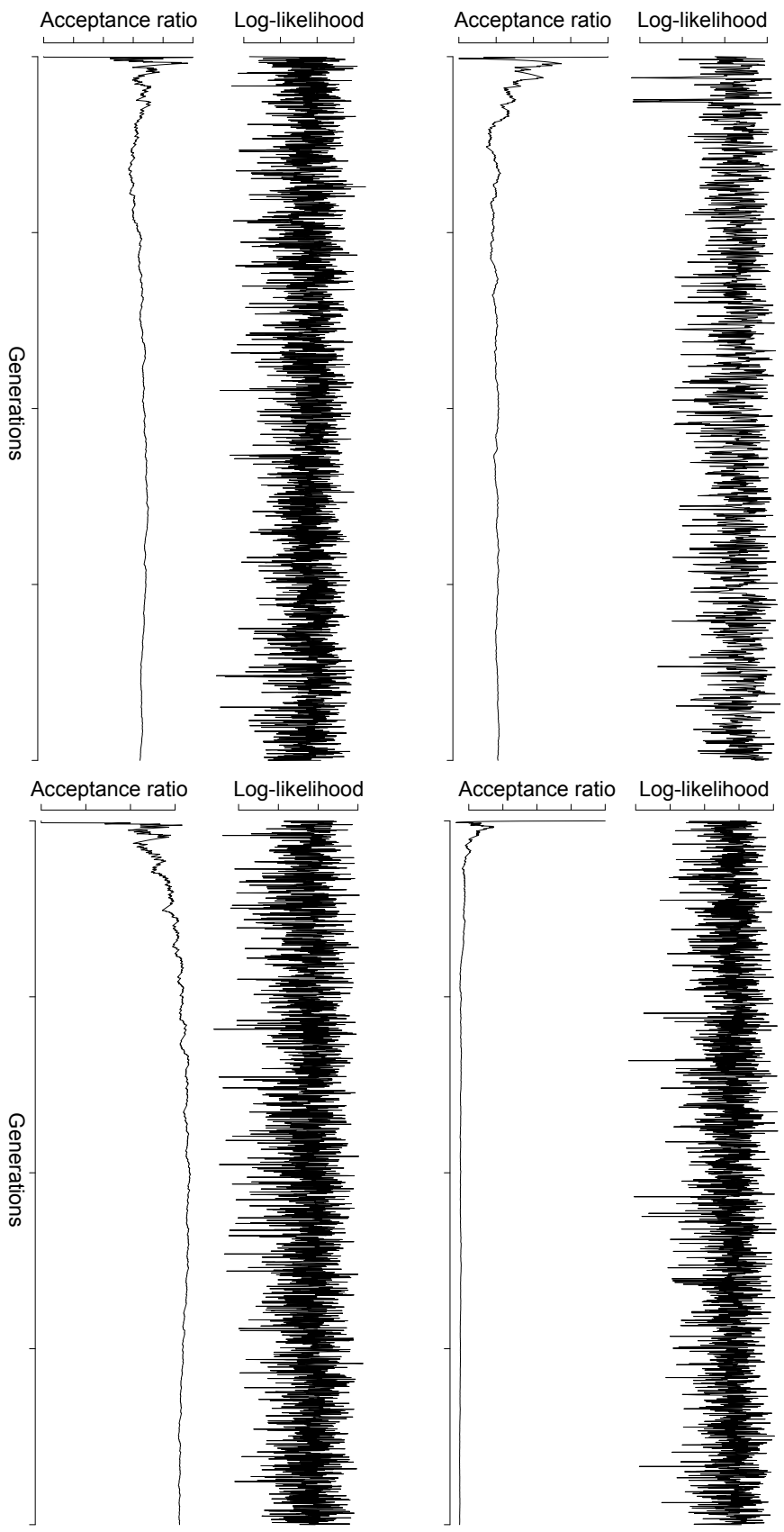


Figure 2.14: Trace plots of the log-likelihood and the acceptance ratio for the four independent MCMC chains of the Centrarchidae fishes analysis.

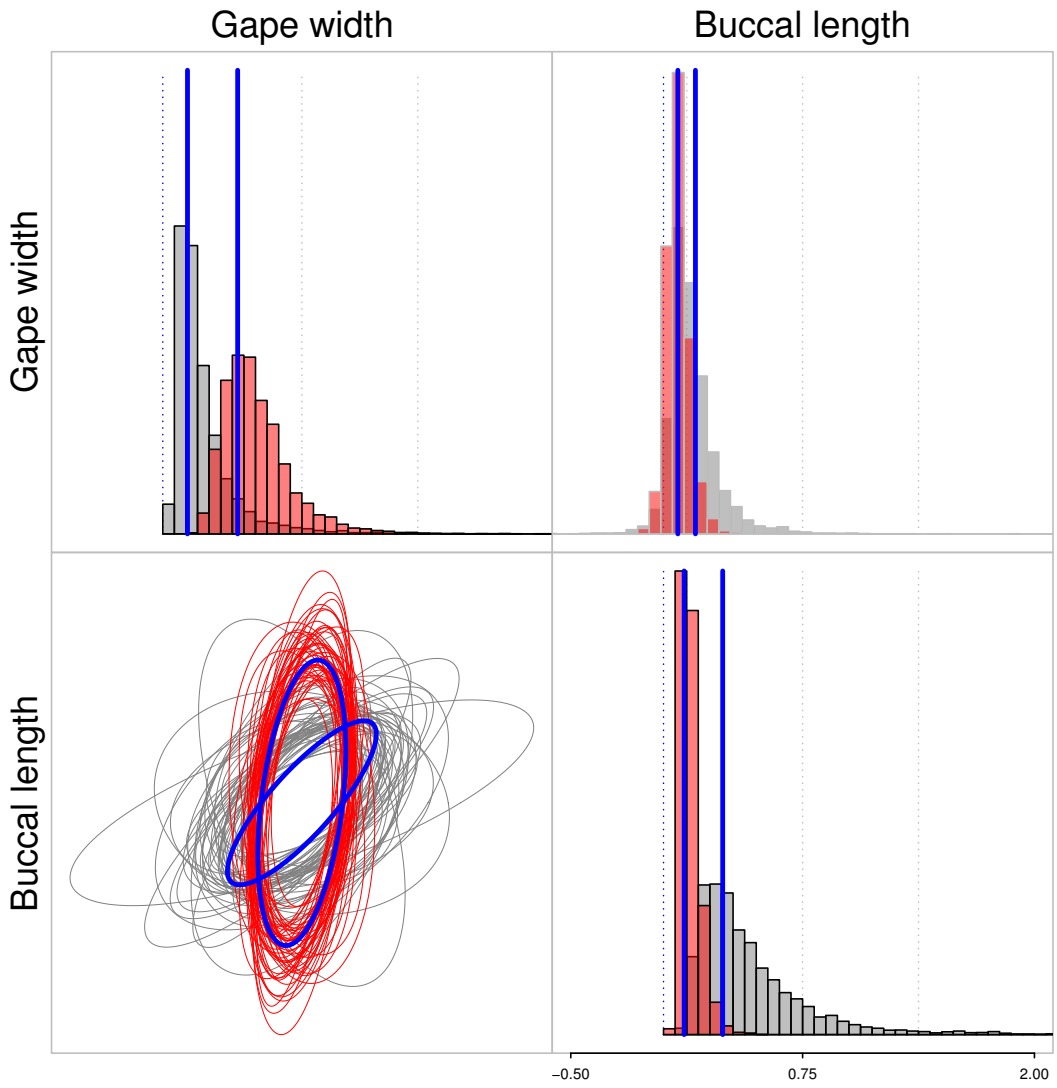


Figure 2.15: Posterior distribution of the **R** matrix regimes fitted to the background group (gray) and to the *Micropterus* clade (red). Maximum likelihood estimate for the same data and phylogenetic tree showed in blue lines. Plot shows the posterior distribution of parameter estimates for the evolutionary rate matrices. Diagonal plots show evolutionary rates (variances) for each trait, upper-diagonal plot show pairwise evolutionary covariation (covariances), and lower-diagonal plot shows samples from the posterior distribution of ellipses showing the 95% confidence interval of each bivariate distribution.

CHAPTER 3: AN R PACKAGE FOR STUDYING EVOLUTIONARY INTEGRATION AMONG SEVERAL TRAITS ON PHYLOGENETIC TREES

3.1 Abstract

Evolutionary integration occurs when two or more phenotypes evolve in a correlated fashion. Correlated evolution among traits can happen due to genetic constraints, ontogeny, and selection and have an important impact on the trajectory of phenotypic evolution. Phylogenetic trees can be used to study such pattern on macroevolutionary time scales by estimating the strength of evolutionary covariance among traits through time and across clades. However, only few applications implement models to conduct comparative analyses of evolutionary integration. We introduce a Bayesian Markov chain Monte Carlo approach to estimate the evolutionary correlation among two or more traits using the evolutionary rate matrix (**R**). **R** is a covariance matrix that represents both the rates of evolution of each trait and the structure of evolutionary correlation among traits. Here we present the R package `ratematrix`, a resource to test hypotheses of evolutionary integration using multivariate data and phylogenetic trees. `ratematrix` provides a flexible framework allowing for any number of evolutionary rate matrix regimes fitted to the same phylogenetic tree and it incorporates the uncertainty associated with parameter estimates, ancestral state reconstruction and phylogenetic estimation in the analyses. The `ratematrix` package uses a novel pruning algorithm that significantly improve computational time. We also provide specific functions that facilitate users to conduct long MCMC analysis when computational resources are limited.

3.2 Introduction

Evolutionary changes in one trait are often associated with changes in other traits, such that species traits often do not vary independently of each other (Olson and Miller, 1958). This pattern can be observed in the covariation among traits both within and among populations (Arnold, 1992; Arnold et al., 2001; Revell and Harmon, 2008; Revell and Collar, 2009). The pattern of correlated evolutionary changes among two or more traits is known as evolutionary integration and can be a result of genetic constraints (e.g., pleiotropy), ontogenetic integration, or correlated selection (Arnold, 1992; Arnold et al., 2001; Hansen and Houle, 2004; Goswami et al., 2015; Melo et al., 2016). Although evolutionary integration is ubiquitous

across the tree of life, only few comparative methods and associated software applications to date implement models that can estimate evolutionary correlations among traits using phylogenetic trees (Revell and Harmon, 2008; Hohenlohe and Arnold, 2008; Revell and Collar, 2009; Bartoszek et al., 2012; Adams and Otárola-Castillo, 2013; Clavel et al., 2015; **?**).
Here we describe the R package `ratematrix`, which implements a Bayesian estimate of evolutionary rate matrices (**R**; Revell and Harmon, 2008) fitted to phylogenetic trees and trait data using Markov chain Monte Carlo (as described in Caetano and Harmon, 2017a). The **R** matrix is a variance-covariance matrix that describes the rates of trait evolution under Brownian motion in the diagonals and the evolutionary covariance among traits (i.e., the pattern of evolutionary integration) in the off-diagonals (Revell and Harmon, 2008; Revell and Collar, 2009; Adams and Felice, 2014). With such a matrix we are able to simultaneously investigate the pace of evolution and the structure of evolutionary integration among two or more continuous traits evolving along the branches of a phylogenetic tree. We can also fit multiple **R** matrices to the same tree in order to test hypothesis of shifts in the evolutionary integration of these traits across clades on the tree.
The **R** matrix can be estimated using current R packages, however all available implementations rely on point estimates using maximum likelihood. In contrast, the use of a Bayesian framework, as presented here, allows for direct incorporation of uncertainty in parameter estimates in the form of a posterior distribution (Caetano and Harmon, 2017a). This is especially important because covariances can be hard to estimate when the number of observations is small relative to the number of parameters in the model, which is commonplace among phylogenetic comparative studies in general.

3.3 The model and MCMC implementation

To study the pattern of correlated evolution among two or more continuous traits we use the model described by Revell and Harmon (2008), which consists of a multivariate Brownian motion model with rate equal to the **R** matrix and root value equal to the vector **a** (see Equations 2 and 3 in Revell and Collar, 2009). Our implementation allows for multiple independent rate regimes fitted to different branches of the phylogenetic tree (as in Revell and Collar, 2009). Rate regimes can be either fixed *a priori* or a collection of multiple regime configurations can be included in the analysis. For example, multiple regimes applied to the same analysis could be samples from a stochastic character mapping (Huelsenbeck et al., 2003), alternative reconstructions due to missing data, or other plausible hypotheses. However, all regimes need to share the same data at the tips of the tree and same number of rate matrices fitted to the tree. The `ratematrix` package implements Metropolis-Hastings

Markov chain Monte Carlo (MCMC) to estimate the posterior distribution of each **R** matrix fitted to the tree and the vector of phylogenetic root values (**a**).

Here we detail the proposal distribution used for each set of parameters as well as the options of prior densities currently implemented in the package. At each step of the MCMC chain we choose between the vector of root values and the set of one or more **R** matrices by drawing from a binomial distribution. The probability that each of these two sets of parameters will be updated is fixed throughout the chain, but can be determined by the user (see function `ratematrixMCMC`). Every time that the set of rate matrices is chosen, only one **R** matrix is updated but all **R** matrices fitted to the tree are equally likely to be updated. In contrast, once chosen, the root value for every trait is updated simultaneously. Updates are performed with different configurations of sliding window proposal distributions.

We implemented a uniform distribution with width controlled by the parameter `'w_mu'` as the proposal distribution for each element of the vector of phylogenetic means. In contrast, the proposal of **R** matrices requires a more elaborate scheme, since variance-covariance matrices are constrained to be positive definite. Furthermore, **R** matrices describe both the rate of evolution of the traits and their pattern of evolutionary correlation, so the proposal distribution needs to provide good mixing for both the variance of each trait and the correlation structure of such matrices. We implemented a separation strategy (Barnard et al., 2000; Zhang et al., 2006; Liu et al., 2016) proposal scheme which consists of making updates to the vector of standard deviations and the correlation matrix derived from the variance-covariance matrix in separate steps (Figure 3.1). The vector of standard deviations can be updated directly using a sliding window proposal distribution. On the other hand, the proposal scheme for the correlation matrix requires two steps: first, draw covariance matrices from an inverse-Wishart distribution; and second, we derive correlation matrices from this sample. Finally, we recompose the evolutionary rate matrix in order to calculate the likelihood of the model. Figure 3.1 shows a diagram that describes the procedure. Note that the vector of standard deviations generated by decomposing the variance-covariance matrix is not evaluated by the likelihood of the model (Zhang et al., 2006). As a result of this parameter-extension approach, we need to correct the acceptance ratio for the transformation from variance-covariance matrix to the corresponding correlation matrix and vector of variances (Figure 3.1). This proposal scheme allows for independent priors and proposals for the rates of evolution and the evolutionary integration among traits. The `ratematrix` package allows control over the width of the uniform sliding window proposal for the vector of standard deviations (`'w_sd'`) as well as the degrees of freedom of the inverse-Wishart (`'v'`) used to sample correlation matrices.

The prior densities for the model naturally follow the proposal scheme implemented. Prior densities are determined for the vector of root values, the vector of standard deviations and the correlation matrix (which is sampled by a transformation from the inverse-Wishart distribution). The separation between standard deviations and correlation matrix enable users to translate their biological intuition about the pace and mode of evolution into model parameters in a straightforward manner. The user can set independent priors (options are uniform, normal or log-normal) for both the vector of root values and the vector of standard deviations of the **R** matrices. For the correlation matrix that, together with the standard deviation, will constitute the **R** matrix the user can set the degrees of freedom (ν) and the scale matrix (Ψ) of the inverse-Wishart distribution. Small values for ν make the distribution wider (Figure 3.2, top) whereas larger values reduce the variance of the distribution, so samples will be closer to Ψ and, as a result, the prior will be more informative (Figure 3.2, bottom). Note from Figure 3.2 that a change in the parameters of the inverse-Wishart prior on the correlation matrix will not change the prior distribution of variances. Since the inverse-Wishart will only be used to sample correlation matrices, Ψ can be set as any correlation matrix or variance-covariance matrix.

3.4 Description of the `ratematrix` R package

The package `ratematrix` offers a plethora of functions to allow flexible choices of prior distributions for all parameters in the model, customizable MCMC chains, plots, and robust analyses of convergence (Table 3.1). The package can be installed from our `github` repository using the R package `devtools`:

```
devtools::install_github("Caetanods/ratematrix", build_vignettes =
TRUE)
 library(ratematrix)
```

The option `build_vignettes` will make the package vignettes available after installation. A list of vignettes can be accessed using:

```
browseVignettes("ratematrix")
```

We will use the same data from Caetano and Harmon (2017a) on mainland and island anole lizards as a demonstration of the package. In this study we test whether the radiation of anole lizards from Central and South America to the Caribbean islands was associated with a shift in the pattern of evolutionary integration among morphological

traits (head length, tail length and snout-vent length). Here the phylogenetic tree (object `anoles$phy`) is a stochastic map with two regimes produced with the package `phytools`, one rate regime for island and other for mainland species (see Figure 3.5). Both the trait data and phylogenetic tree are included in the `ratematrix` package.

```
data(anoles) # Load trait data and phylogeny.
```

3.4.1 Estimating rates of correlated evolution

After loading the package and data, we choose the prior distributions for the MCMC chain. We set a uniform prior for the vector of root values and variances and a marginally uniform prior for the covariance matrices, following Barnard et al. (2000) (see also documentation for `makePrior`). The marginally uniform prior produces uniform distributions for each of the covariance terms ($\sigma^2_{i,j}$ for $i \neq j$) of the variance-covariance matrix after integrating over the uncertainty of the other parameters (i.e., the marginal distribution for $\sigma^2_{i,j}$). Many characteristics of the Markov chain Monte Carlo can be customized (see documentation for `ratematrixMCMC`). We encourage users to run short preliminary chains in order to adjust the width of the proposal distributions for each set of parameters in function of the acceptance ratio (see function `logAnalyzer`) prior to a full MCMC chain analysis. This procedure can improve the mixing of the chains, which might decrease the number of generations required until convergence and increase the effective sample size (ESS) of the posterior distribution. The following lines of code will run only a short example, starting with a random sample from the prior distribution. The package also provide results from previous MCMC analyses with this same data as examples.

```
estimateTimeMCMC(data=anoles$data[,1:3], phy=anoles$phy, gen=10000)
handle <- ratematrixMCMC(data=anoles$data[,1:3], phy=anoles$phy
, prior="uniform", gen=10000)
```

The `estimateTimeMCMC` function estimates the time for the MCMC chain whereas `ratematrixMCMC` runs it. The MCMC function writes one file with the parameter samples and another with the log information for each generation. Both files are marked with a unique identifier that prevents multiple chains of overwriting each other. The `handle` object is a list containing detailed information about the MCMC chain and is required in order to read the posterior distribution from files, analyze the log information and continue an unfinished MCMC chain. Bellow we show an example of how to read the posterior

distribution from the files. Then, we make plots (Figure 3.3) and calculate summary statistics based on results of a converged MCMC chain provided as example data.

```
( short_chain <- readMCMC(handle, burn=0.25, thin=1) )
logAnalyzer(handle, burn=0.25, thin=1) # Log information for the chain.
plotRatematrix(chain=short_chain) # Plots the short chain.
plotRootValue(chain=short_chain) # Plots the short chain.

data(anolesPost) # Load example of posterior distribution.
plotRatematrix(chain=anolesPost$chain1)
plotRootValue(chain=anolesPost$chain1)
checkConvergence(anolesPost$chain1, anolesPost$chain2)
testRatematrix(chain=anolesPost$chain1, par="correlation")
testRatematrix(chain=anolesPost$chain1, par="rates")
```

The `logAnalyzer` function calculates the acceptance ratio for each parameter of the model and plots the trace of the log-likelihood for the MCMC chain. The plot functions show the posterior distribution of parameter estimates. `plotRatematrix` produces a plate with histograms for the rate of evolution of each trait (diagonal) and the pairwise evolutionary covariation among traits (upper-diagonal). The lower-diagonal plots show ellipses for the 95% confidence interval of the bivariate distribution between each pair of traits. Different from the histograms, the ellipses are only a sample from the posterior distribution (see documentation for `plotRatematrix`).

Here we applied the test of convergence for the chains using Gelman and Rubin (1992) potential scale factor analysis (see function `checkConvergence`). This convergence test requires two or more independent MCMC chains and compares the variance of parameter estimates between chains and within each chain.

The `testRatematrix` function calculates a series of summary statistics based on the pairwise degree of overlap among the posterior distribution of $\mathbf{R}$ matrices fitted to same phylogenetic tree (Caetano and Harmon, 2017a). If this overlap exceeds 5%, then we can conclude that the difference between posterior parameter estimates is not strong enough to support the hypothesis that regimes are representations of distinct macroevolutionary patterns. When we compare the posterior distribution for the evolutionary correlation (`par="correlation"`; overlap of 0.4) and the rates of evolution for each trait (`par="rates"`; overlap of 0.0002) there is no evidence for a shift in the pattern of evolutionary integration but island anole lineages show faster rates of trait evolution when compared to mainland lineages (Figure 3.3,

see also Caetano and Harmon, 2017a). We refer readers to Caetano and Harmon (2017a) for an extensive simulation study of the performance of the method as well as the use of summary statistics under a diverse set of scenarios of correlated evolution.

3.4.2 Integration of uncertainty in regime configurations

One of the advantages of Bayesian implementations is that analyses can integrate uncertainty from different sources. Rate regimes, for example, are used to map the set of nodes and branch lengths that will be assigned to each **R** matrix fitted to a phylogenetic tree. Such regimes are usually determined *a priori* by ancestral estimate, since we often are interested in the association between some characteristic and a possible shift in the pattern of evolutionary integration among traits. However, ancestral state estimates can be uncertain and alternative reconstructions are often possible, specially when the states of the characteristics under study are polymorphic or of dubious interpretation. Nevertheless, most comparative methods are implemented to estimate the parameters of the model with a single regime configuration and users need to perform multiple independent analyses in order to incorporate uncertainty associated with ancestral state estimates.

The package `ratematrix` offers a different approach by allowing a pool of phylogenetic trees to be directly incorporated in the MCMC. This pool can comprise repeated simulations from a stochastic mapping analysis, equally parsimonious ancestral state reconstructions or even a random sample of trees from the posterior distribution of a phylogenetic inference analysis. In order to incorporate the pool of trees, we randomly sample one phylogenetic tree each time the likelihood function of the model is evaluated. Note that this procedure is not the same as a joint Bayesian MCMC estimate of the trait model and the phylogenetic tree because the posterior distribution of rate regimes or phylogenetic trees is not sampled as part of the MCMC. However, the posterior distribution of parameter estimates for the model is sampled conditioned on the pool of trees provided by the user. Bellow we show how to create an analysis based on a pool of stochastic mapping simulations:

```
    library(phytools) # Using Revell (2012).
    state <- setNames(anoles$data$Location, rownames(anoles$data))
    phy_map <- make.simmap(anoles$phy, x=state, nsim=100)
    handle_map <- ratematrixMCMC(data=anoles$data[,1:3], phy=phy_map
, prior="uniform", gen=10000, outname="phy.pool")
    logAnalyzer(handle_map, burn=0.25, thin=1)
```

When a pool of phylogenetic trees is provided for the MCMC, `logAnalyzer` returns the acceptance ratio for each of the phylogenetic trees. Relative low acceptance ratio for a given tree means that proposals were rejected more often than with the rest of trees from the same pool. This might be due to a low likelihood score for the multivariate Brownian-motion model given the tree as a result of an unlikely topology, branch lengths or rate regime configuration. If this is the case, one can check if the tree has some particular attributes that make it distinct from other trees in the pool, since such patterns might carry important biological information. Furthermore, we recommend that an independent analysis is performed with such a tree (or trees) so that one can test whether results are significantly different than the former analysis using the entire pool.

3.4.3 Continuing unfinished chains or adding extra iterations

The `ratematrix` package allows for continuing an unfinished MCMC analysis or to append additional generations to the previous MCMC chain. This is an essential feature given the computational burden associated with any Bayesian simulation approach. In both cases, the user needs to provide the `handle` object returned by the `ratematrixMCMC` function (or saved to the working directory). The following example will add iterations to the previous MCMC chain:

```
handle_map_add <- continueMCMC(handle_map, add.gen=1000)
```

One can use `continueMCMC` alongside `checkConvergence` to add generations to the MCMC until the chain(s) pass the convergence test. This is especially relevant given that the number of generations required for acceptable convergence is dependent on the data and the configuration of the sampler.

3.5 New pruning algorithm improves computational time

The package `ratematrix` implements a novel algorithm to evaluate the likelihood function of a multivariate Brownian motion model when two or more **R** matrix regimes are fitted to the same phylogenetic tree (Caetano and Harmon, 2017a). In previous implementations, a large matrix composed by the multiplication between the phylogenetic variance-covariance matrix, with dimension equal to the number of species in the tree, and the evolutionary rate matrix, with dimension equal to the number of traits, needed to be computed. However, any operation with such large matrices can become very computationally intensive. Recently, Caetano and Harmon (2017a) implemented an extension of Felsenstein (1973) pruning algorithm that

avoids such calculations. As a result, matrix operations need only to be performed with the evolutionary rate matrix ($\mathbf{R}$), which is usually a fairly small matrix. Figure 3.4 shows the computational time for the likelihood function under different approaches. Computation using the full inverse and determinant of the matrices is the approach that scales worst with number of traits and size of the phylogeny. Although the 'rpf' method (Gustavson et al., 2010), which avoids the computation of the full inverse and determinants, shows a significant improvement, the pruning algorithm has the best performance. With respect to the asymptotic upper bounds (O), all approaches scale equally with the number of traits in the analysis when the size of the phylogeny is held constant, but there is a remarkable improvement with the scaling in function of the number of tips in the phylogeny. The pruning algorithm scales with $O(n + r^3)$ whereas the other methods scale with $O(n^3 + r^3)$, where n is the number of tips in the phylogeny and r is the number of traits. The reduction of time to evaluate the likelihood of the model is fundamental to the implementation of simulation based approaches such as the Bayesian Markov chain Monte Carlo estimates performed by the package `ratematrix` (Caetano and Harmon, 2017a).

3.6 Resources

`ratematrix` is an open-source R package that can be installed from the github repository `https://github.com/Caetanods/ratematrix`. A series of tutorials are available at `https://github.com/Caetanods/ratematrix/wiki`.

Table 3.1: Principal functions available in `ratematrix`.

Function	Description
`checkConvergence`	Perform tests of convergence with one or multiple MCMC chains.
`continueMCMC`	Continue or add generations to a MCMC chain.
`estimateTimeMCMC`	Estimate the time that a MCMC chain will take to run.
`likelihoodFunction`	Compute the log-likelihood of the multivariate Brownian-motion model.
`logAnalyzer`	Compute acceptance ratio for parameters and pool of phylogenies and make trace plots for the log-likelihood and acceptance ratio.
`makePrior`	Create prior densities for the model.
`makeStart`	Create starting point for the MCMC chain.
`mergePosterior`	Merge multiple chains from the same data into a single chain.
`mergeSimmap`	Merge rate regimes mapped to a phylogenetic tree.
`plotPrior`	Plot the prior distribution for the parameters of the model.
`plotRatematrix`	Plot the posterior distribution of evolutionary rate matrices ($\mathbf{R}$).
`plotRootValue`	Plot the posterior distribution of root values (phylogenetic mean).
`ratematrixMCMC`	Make the Bayesian Markov chain Monte Carlo analysis.
`readMCMC`	Read MCMC samples from the `ratematrixMCMC` output file.
`samplePrior`	Draw samples from the prior density created by `makePrior`.
`simRatematrix`	Simulate data given a phylogenetic tree, evolutionary rate matrix and root values.
`testRatematrix`	Use summary statistics to test for shifts between rate regimes.

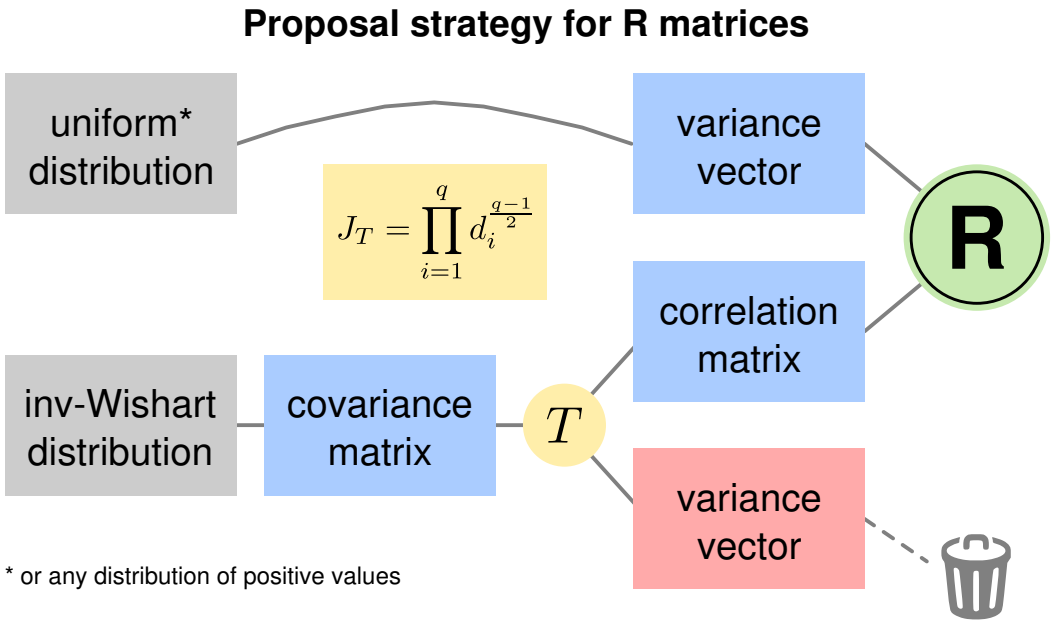


Figure 3.1: Diagram of the separation strategy proposal (Barnard et al., 2000). Boxes in grey show the proposal distributions for the variance vector and correlation matrix that compose the evolutionary rate matrix (**R**). Boxes in blue show the elements that are directly (or indirectly, in the case of the covariance matrix) evaluated in the acceptance step of the MCMC. The yellow circle shows the transformation (T) required to decompose the variance-covariance matrix sampled from a inverse-Wishart into a correlation matrix and the variance vector. The yellow square shows the formula for the Jacobian correction due to T (Zhang et al., 2006), where d_i stands for the variance of traits 1 to q. The red square demonstrates that a additional variance vector is produced in the process, but it is discarded. The green circle is a representation of the **R** matrix as a product between the variance vector and the correlation matrix.

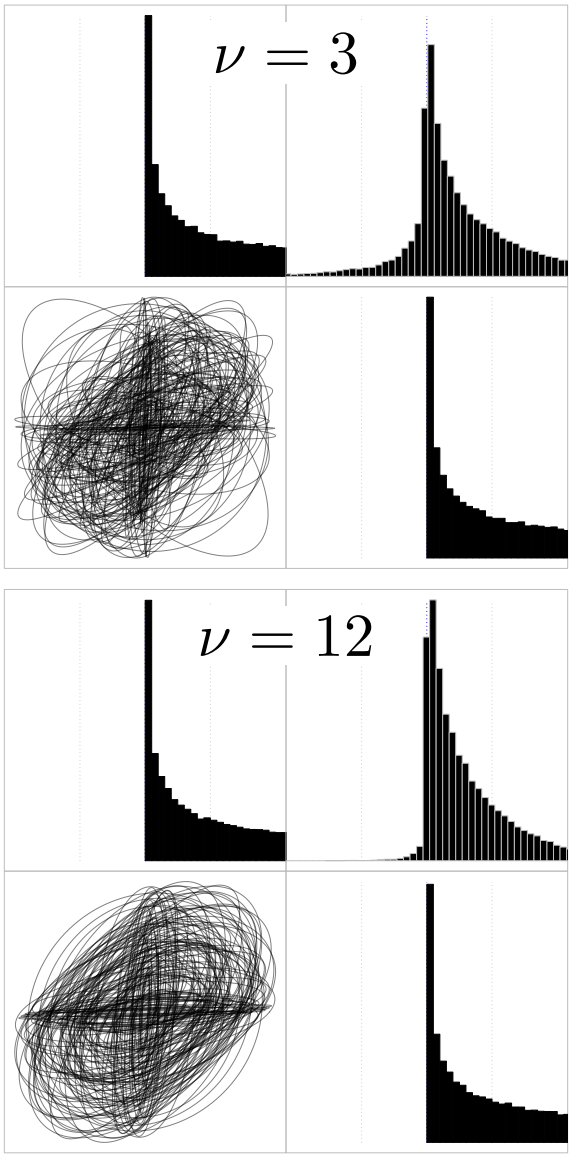


Figure 3.2: Samples from the prior of the evolutionary rate matrix (**R**) for two simulated traits using the separation strategy (Barnard et al., 2000). Standard deviation was modelled as a uniform distribution between 0 and 10 and the correlation matrix was derived from a inverse-Wishart centered on a scale matrix with positive correlation (corr = 0.5). Top figure shows a weak prior with small value for the degrees of freedom parameter ($\nu = 3$) whereas bottom figure are draws from a more informative prior ($\nu = 12$). In each figure, the plots in the diagonal show evolutionary rates for each trait whereas the upper-diagonal plot shows the evolutionary covariation. Lower-diagonal plot show 150 randomly sampled ellipses representing the 95% quantile for the bivariate distribution. Although both priors share the same scale matrix, when ν is small the prior distribution has more variance than when ν is larger. The diagonal plots are held constant since the prior distribution for the standard deviation is the same in both figures. Note that ellipses show both positive and negative correlation when the prior is weak (top figure) and positive or no correlation when the prior is more informative (bottom figure).

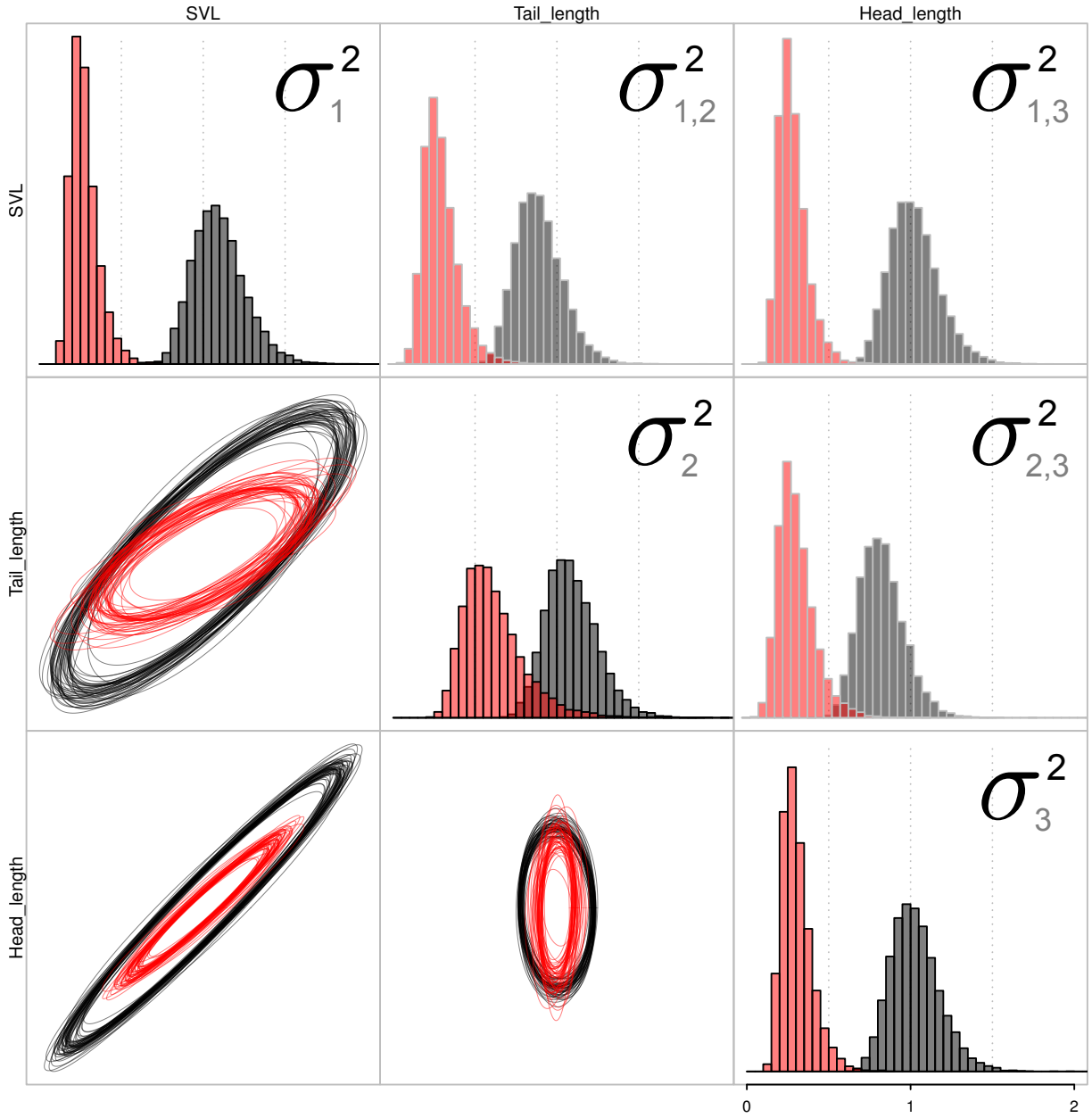


Figure 3.3: Posterior distribution of the evolutionary rate matrix (**R**) regimes fitted to the island anole (gray) and mainland anole (pink) lineages. A different **R** matrix were jointly estimated for each regime. The plots in the diagonal show evolutionary rates (variances) for each trait; σ^2_1 for SVL, σ^2_2 for tail length, and σ^2_3 for head length. Upper-diagonal plots show pairwise evolutionary covariation (covariances); $\sigma^2_{1,2}$ between SVL and tail length, $\sigma^2_{1,3}$ between SVL and head length, and $\sigma^2_{2,3}$ between tail length and head length. The ellipses in the lower-diagonal plots represent the 95% confidence interval of each bivariate distribution for 50 randomly sampled **R** matrices from the posterior. The order of the ellipse plots is a mirror reflection from the upper-diagonal evolutionary covariance plots. Ellipses are only a sample of the posterior because a very large number of lines can become hard to visualize, however the user can set any number of samples (or the entire posterior).

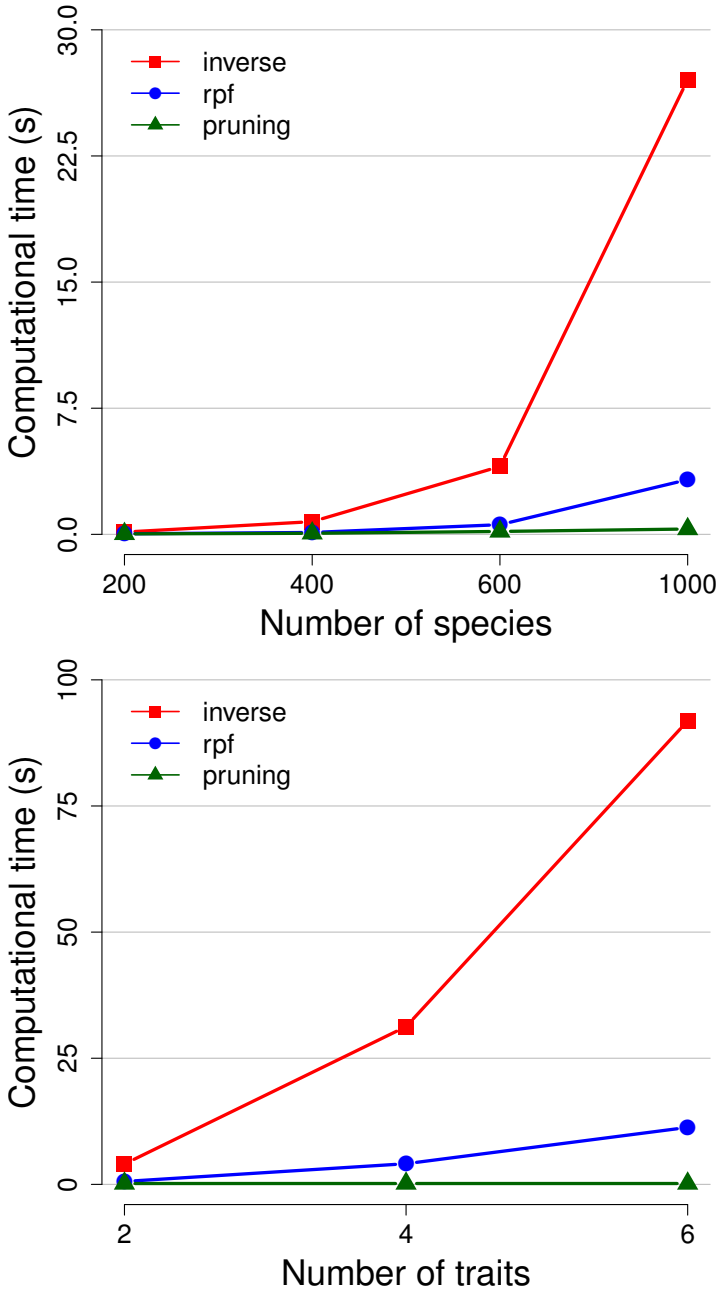


Figure 3.4: Time in seconds spent to compute the likelihood function using different approaches. Top figure shows computational time for two traits and a phylogeny of different number of species. Bottom figure shows computational time with a phylogeny of 400 species and increasing number of traits. Both plots show a comparison among three approaches: 'inverse' uses the full inverse and determinant of matrices as implemented in `phytools` (Revell, 2012), 'rpf' uses the rectangular full-packed format algorithm as implemented in `mvMORPH` (Clavel et al., 2015), and 'pruning' uses Felsenstein (1973) pruning algorithm as implemented in `ratematrix`.

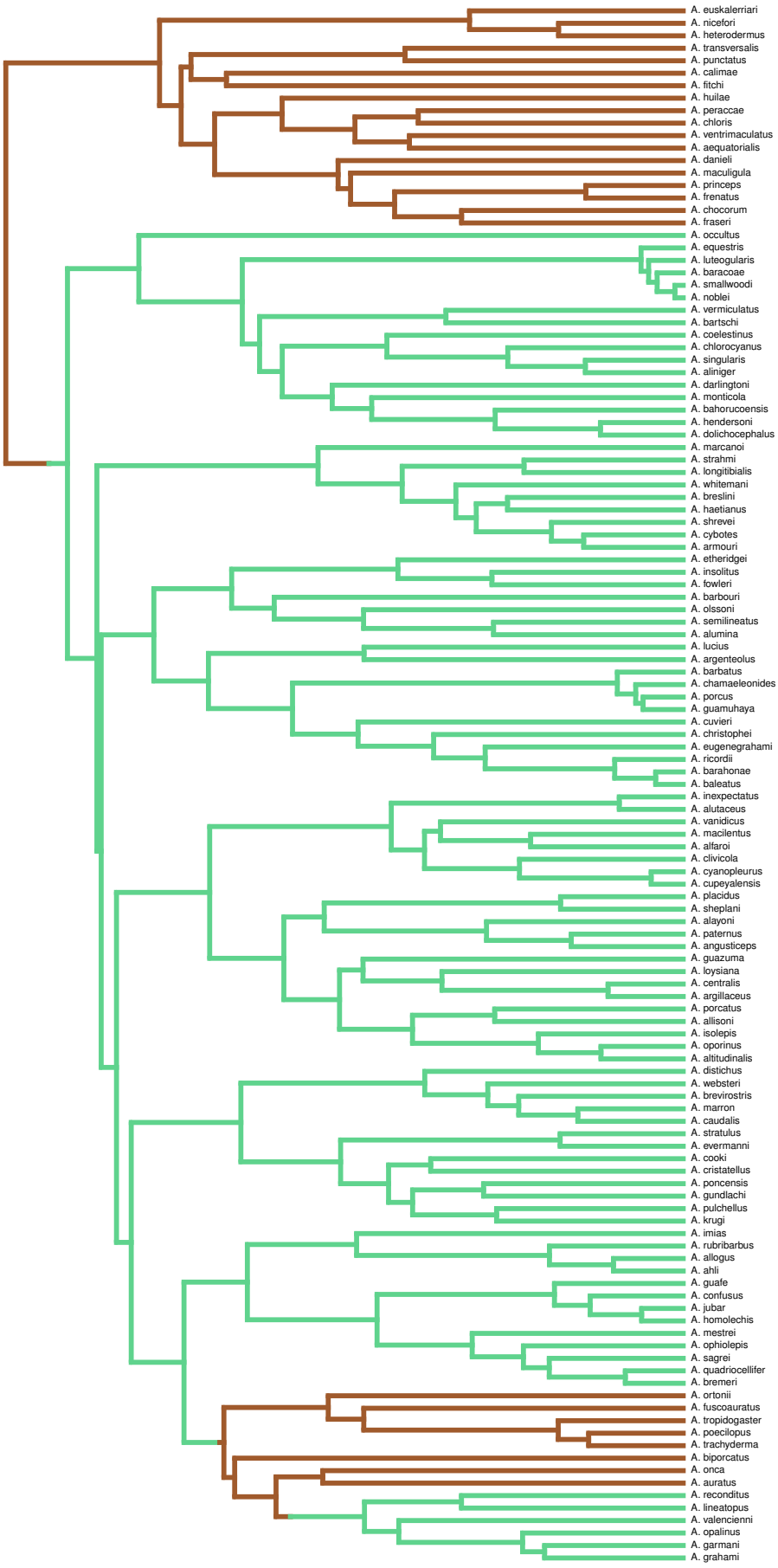

A. euskalerriari
A. nicefori
A. heterodermus
A. transversalis
A. punctatus
A. calimae
A. fitchi
A. huilae
A. peraccae
A. chloris
A. ventrimaculatus
A. aequatorialis
A. danieli
A. maculigula
A. princeps
A. frenatus
A. chocorum
A. fraseri
A. occultus
A. equestris
A. luteogularis
A. baracoae
A. smallwoodi
A. noblei
A. vermiculatus
A. bartschi
A. coelestinus
A. chlorocyanus
A. singularis
A. aliniger
A. darlingtoni
A. monticola
A. bahorucoensis
A. hendersoni
A. dolichocephalus
A. marcanoi
A. strahmi
A. longitibialis
A. whitemani
A. breslini
A. haetianus
A. shrevei
A. cybotes
A. armouri
A. etheridgei
A. insolitus
A. fowleri
A. barbouri
A. olssoni
A. semilineatus
A. alumina
A. lucius
A. argenteolus
A. barbatus
A. chamaeleonides
A. porcus
A. guamuhaya
A. cuvieri
A. christophei
A. eugenegrahami
A. ricordii
A. barahonae
A. baleatus
A. inexpectatus
A. alutaceus
A. vanidicus
A. macilentus
A. alfaroi
A. clivicola
A. cyanopleurus
A. cupeyalensis
A. placidus
A. sheplani
A. alayoni
A. paternus
A. angusticeps
A. guazuma
A. loysiana
A. centralis
A. argillaceus
A. porcatus
A. allisoni
A. isolepis
A. oporinus
A. altitudinalis
A. distichus
A. websteri
A. brevirostris
A. marron
A. caudalis
A. stratulus
A. evermanni
A. cooki
A. cristatellus
A. poncensis
A. gundlachi
A. pulchellus
A. krugi
A. imias
A. rubribarbus
A. allogus
A. ahli
A. guafe
A. confusus
A. jubar
A. homolechis
A. mestrei
A. ophiolepis
A. sagrei
A. quadriocellifer
A. bremeri
A. ortonii
A. fuscoauratus
A. tropidogaster
A. poecilopus
A. trachyderma
A. biporcatus
A. onca
A. auratus
A. reconditus
A. lineatopus
A. valencienni
A. opalinus
A. garmani
A. grahami

Figure 3.5: Maximum clade credibility phylogenetic tree for anole lizards made available by Gamble et al. (2014). Only anole species included in the analysis are shown. Branches painted in brown represent mainland lineages whereas branches in green are island lineages. Regimes were mapped to the tree using stochastic mapping with root state set to mainland, as implemented in the package `phytools` (Revell, 2012). Please refer to Caetano and Harmon (2017a) for more information.

CHAPTER 4: PREDICTING RATES OF EVOLUTION FOR ONE TRAIT USING A CONTINUOUS GRADIENT OF ANOTHER TRAIT

4.1 Abstract

Shifts in the tempo and mode of trait evolution are often interpreted as instantaneous changes between different regimes. Such regimes are defined as the portion of the phylogenetic tree related to the presence or absence of a given discrete variable. Then, phylogenetic comparative models are used to estimate patterns associated with these regimes and test whether there is a significant difference among them. However, often times the variable that defines evolutionary regimes is not essentially discrete. Many important factors that influence rates of trait evolution are continuous such as climate, degree of species interactions, range size, habitat specialization, among others. As a result, one needs to create arbitrary bounds and fit such traits into discrete categories in order to perform comparative analyses, since, up to date, there is no method able to use continuous variables as predictors in analyses of rates of trait evolution across the branches of a phylogenetic tree. Here we describe a new comparative method which aims to use a continuous predictor variable in order to predict the variation of evolutionary rates for a continuous response variable. For this we fit different continuous mathematical functions that map predictor trait values to evolutionary rates for the response trait. Our approach is not to be confused with PGLS analyses, since the former relate predictor trait values to rates of evolution for the response trait whereas the latter tests for the correlation between the values of two or more continuous traits. We use simulations to show that the new approach is able to correctly estimate parameter values for three distinct predictive functions. The use of mathematical functions as predictive models for rates of evolution is a powerful approach and might contribute to the development and test of more general processes driving trends of phenotypic evolution across the tree of life.

4.2 Introduction

There is a striking degree of phenotypic diversity among groups of species across the tree of life. Many are the factors that influence the tempo and mode of phenotypic evolution (e.g., Adams et al., 2009; Cooper and Purvis, 2009; Hipsley et al., 2014). Furthermore, shifts in the pace of trait evolution, which can disconnect phenotypic disparity from clade age, are common throughout the history of several groups (Eastman et al., 2011; Benson and

Choiniere, 2013; Mahler et al., 2013; Rabosky et al., 2013; Slater, 2013; Rabosky et al., 2014; Uyeda and Harmon, 2014). These shifts are often associated with important events on the evolutionary history of lineages, such as mass extinctions (Slater, 2013), the evolution of a new function for an existent trait (Benson and Choiniere, 2013; Dececchi and Larsson, 2013) or even changes in the structure of evolutionary correlation among traits affecting the ability of lineages to explore novel regions of the morphospace (Revell and Collar, 2009; Caetano and Harmon, 2017a).

Changes in the tempo and mode of evolution of lineages are often interpreted as instantaneous shifts in the history of the group. On a phylogenetic context, discrete evolutionary events are mapped to specific locations on branches or nodes of the tree following ancestral reconstruction analyses (Schluter et al., 1997; Pagel, 1999; Huelsenbeck et al., 2003). These positions then divide the phylogenetic history of the group into different regimes defined by the presence or absence of the events under study (Butler and King, 2004; O'Meara et al., 2006; Revell and Collar, 2009; Eastman et al., 2011; Caetano and Harmon, 2017a). This framework is the most used to study macroevolutionary patterns of trait evolution using available phylogenetic comparative methods (PCMs) and is grounded in the concept of 'key innovations' (Simpson, 1953; Heard and Hauser, 1995) as the driver of trends in phenotypic evolution. Interestingly, most macroevolutionary studies are focused on the influence of rare events on the history of the groups even though the reduced number of observations is prone to bias results due to the problem of phylogenetic pseudo-replication (Maddison and FitzJohn, 2015). Whether such focus is a result of an inordinate fondness for unique, and often extraordinary, events across the tree of life or a paradigm associated to the history of comparative studies is hard to understand.

On the other hand, not all traits acting as important driving forces of phenotypic evolution among lineages are essentially discrete. Important factors such as climate (Hunt et al., 2015; Clavel and Morlon, 2017), degree of species interactions (Galetti et al., 2013; Thompson et al., 2013), range size (Diniz-Filho and Balestra, 1998; Pigot et al., 2012), and habitat specialization (Bonetti and Wiens, 2014; Hardy and Otto, 2014) are inherently continuous. Continuous traits are not naturally interpreted on a phylogenetic context by the use of discrete transitions or regimes. Furthermore, while discrete 'key innovations' are commonly associated with shifts in the tempo and mode of phenotypic evolution (e.g., Revell and Collar, 2009; Claverie et al., 2011; Benson and Choiniere, 2013; Dececchi and Larsson, 2013; Maddison and FitzJohn, 2015), continuous factors are often addressed as predictors on regression models that use phylogenetic information only as a means to address the lack of independence among species traits (Felsenstein, 1985; Grafen, 1989; Blomberg et al., 2012). However, continuous traits can also have important impacts on the tempo and mode of

phenotypic evolution and, similar to regimes mapped for a discrete trait, the gradient of a continuous variable across the branches of the phylogenetic tree could be used as a predictor for the dynamics of phenotypic evolution.

Most phylogenetic comparative models of trait evolution focus on the macroevolutionary patterns of single traits (Butler and King, 2004; O'Meara et al., 2006; Eastman et al., 2011; Rabosky et al., 2014; Uyeda and Harmon, 2014). Evolutionary rates for traits are estimated based on species mean values, the distribution of branch lengths, and the topology of the tree whereas rate shifts are estimated in function of the evolutionary variance for the trait estimated on different regions of the tree (O'Meara et al., 2006; Eastman et al., 2011). Some methods allow rate regimes to be mapped *a priori* to the phylogeny following another discrete variable (Butler and King, 2004; O'Meara et al., 2006) while others automatically find the location of shifts (Eastman et al., 2011; Rabosky et al., 2014; Uyeda and Harmon, 2014). More recently, models have incorporated continuous varying rates of trait evolution across the branches of the tree, but such rates vary independently of any predictor variable (Rabosky et al., 2014). In the case that two or more continuous traits are incorporated into the analysis, models use variance-covariance matrices to estimate both the rate of evolution for each trait and their evolutionary covariance (Revell and Collar, 2009; Bartoszek et al., 2012; Clavel et al., 2015; Caetano and Harmon, 2017a). However, up to date, no comparative model is able to estimate rates of trait evolution for a continuous trait as predicted by another continuous variable mapped to the same tree.

We present a novel framework that uses mathematical functions to describe the dynamics of rates of evolution of a continuous trait with respect to the values of a predictor variable mapped to the phylogenetic tree. For this we model rate variation alongside the branches of the phylogeny using a semi-continuous approach. At each branch, rates of trait evolution are predicted using some continuous mathematical function mapping to the values of the predictor variable. We show that the new approach estimates meaningful patterns of rate variation throughout the tree using different mathematical functions and provide a approach to test the predictive power of gradients to explain rates of trait evolution across phylogenetic trees.

4.3 Methods

4.3.1 Description of the model

Here we use continuous mathematical functions to model the rates of evolution of a trait (i.e., the response trait) under a Brownian-motion model (BM) with respect to the values

of another trait (i.e., the predictor trait). The model aims to study the association between the macro-evolutionary patterns of one trait evolving on a phylogenetic tree and the rates of evolution of another trait on the same tree. Figure 4.2 shows the fundamental concepts of the model. The predictor and response traits are continuous traits for the same set of species. The linear function represented by the line on Figure 4.2 describes the variation of evolutionary rates for the response trait ($\sigma^2_{response}$) in function of the trait values of the predictor trait. The difference between the present model and other phylogenetic comparative models of trait evolution with varying rates, such as AUTEUR (Eastman et al., 2011), BAMM (Rabosky et al., 2014), and bayOU (Uyeda and Harmon, 2014), is that these methods estimate rates of trait evolution informed only by the distribution of trait values and the phylogenetic tree. In contrast, here we introduce the use of a second trait to inform the variation of rates of evolution across the branches of the tree.

The model can be sub-divided into two components; ancestral values for the predictor trait are mapped to the branches of the phylogenetic tree and evolutionary rate regimes for the response trait are assigned to the phylogeny in function of the map of ancestral predictor trait values. First we divide the predictor trait into k ordered categories defined by k - 1 equidistant breakpoints across the range of the predictor trait values. Then we estimate the evolutionary rate transitions among the k categories using a Markov process (Pagel, 1994) that restricts transition rates to happen only between neighbouring states. For example, a transition from k_i to a larger trait value category k_{i+3} need to be preceded by transitions to and from the intermediary categories k_{i+1} and k_{i+2}. The number of categories reflects how fine-grained is the model with respect to the macro-evolutionary patterns of the (continuously distributed) predictor trait; larger k produce a more fine detailed model whereas smaller k yield a more coarse model (Boucher and Démery, 2016).

The transition rates between the k categories of the predictor trait can be estimated using a meristic Markov model (MKn) with a single homogeneous rate or multiple rates. Homogeneous rate can be described by constraining all transition rates between neighbouring categories to be equal whereas more complex evolutionary patterns can be described by allowing transition rates to and from each category to vary. Of course, one can compare and choose the model that best describes the evolutionary history of the predictor trait across the branches of the phylogenetic tree. As the number of categories (k) increases the single rate Mkn model becomes equivalent to a single rate Brownian-motion (BM) model whereas the unconstrained Mkn model spans trait evolution models with heterogeneous rates such as multiple rate BM (O'Meara et al., 2006), BM with a directional trend (Hunt, 2006) and Ornstein–Uhlenbeck (OU – Butler and King, 2004) (Boucher and Démery, 2016). However, the Mkn model, as used here, fits a single transition matrix to the whole phylogenetic tree

and, as a result, can only be equivalent to time-dependent models, such as the accelerated/decelerated (ACDC – Blomberg et al., 2003; Uyeda et al., 2015) and early burst (EB – Harmon et al., 2010), if the predictor map shows a time-dependent structure.
In order to map ancestral values of the predictor trait to the branches of the phylogenetic tree we generate multiple stochastic histories for the trait categories using the estimated Mkn transition matrix (Huelsenbeck et al., 2003). Each of these maps associate the branches of the phylogeny with the k categories for the ancestral values of the predictor trait. Then we use a mathematical function (see Figures 4.1 and 4.2) to map these predictor trait regimes to evolutionary rate regimes for the response trait and compute the likelihood of a multiple rates Brownian-motion model (O'Meara et al., 2006) with the response trait as the tip data.

4.3.2 Mathematical functions and model choice

Virtually any mathematical function can be used to map values of the predictor trait categories to evolutionary rate regimes of the response trait. Different functions can be fit to the data using maximum likelihood (ML) and compared using standard model choice approaches such as Likelihood-ratio tests (LRT) for nested or the Akaike information criterion (AIC) for non-nested models (Burnham and Anderson, 2003). Using model choice criteria that penalizes for the number of parameters (such as AIC – Burnham and Anderson, 2003) is desirable, since distinct functions of varying complexity can produce identical maps. Some mathematical functions are commonly applied across a series of biological disciplines and are likely to be erected as *a priori* hypotheses for the macro-evolutionary association between a plethora of traits. Figure 4.1 shows a collection of functions that describe patterns commonly observed in biological data, especially in studies of trait evolution using phylogenetic trees.
One of the advantages of our approach is that the number of parameters varies with respect to the chosen mathematical function rather than the number of categories used to describe the predictor trait (see Figure 4.1). This is a result of maximizing the likelihood of the data with respect to the parameters of the mathematical function rather than to the rate regimes directly. Thus, we can use a large number of rate regimes in order to model a trait with (semi-)continuously varying rates of evolution across the phylogeny without increasing the number of parameters of the model. This aspect of the model is similar to the strategy implemented in BAMM (Rabosky et al., 2014), which uses an exponential function to describe the continuous decrease or increase of rates of trait evolution through time.

4.3.3 Model implementation

We implemented the model as a R package named `'phylofx'`. The package offers a simple interface to fit continuous mathematical functions to model the variation of evolutionary rates of a response trait in function of a predictor trait across the branches of a phylogenetic tree. All mathematical functions showed on Figure 4.1 are available in the package and can be chosen from a simple menu. The package also have options for users to define their own mathematical functions.

4.3.4 Performance simulations

To check the performance of the method we will focus in three very common, but distinct, nested models: a constant relationship, with homogeneous evolutionary rates (σ^2=0.5); a step function, with two distinct rates separated by an instantaneous transition step (σ^2_{left}=1, σ^2_{right}=0.5, break point=mean of predictor trait); and a linear function, with a continuous relationship between the predictor and the response traits (see Figure 4.1). For the linear function we defined β_1=0.5, the set of predictor trait values at the tip as X, and computed the intercept such that:

$$\beta_0 = \beta_1 \; minX - 0.1 \tag{4.1}$$

We generated a phylogenetic tree with 200 species using a pure-birth model (tree height=1). Then we simulated a predictor trait following a single rate BM model (σ^2=0.5) and divided it into 10 trait categories mapped to the branches of the tree. We will refer to the result of this simulation as the true mapped tree. We used the true mapped tree to assign evolutionary rate regimes to the branches following one of the mathematical functions described above and simulated the response trait under a multi-rate BM model. We repeated this process in order to produce 100 datasets for each of the three mathematical functions used in the simulations tests.

In order to fit the models to the generated data, we estimated a meristic Mkn transition matrix for the predictor trait with k=5 and equal transition rates. We produced 10 stochastic mapping histories based on this transition matrix and performed a maximum likelihood estimate for each of the mathematical functions under each of the stochastic mapping histories. We chose the best model by comparing the mean pairwise AIC values across all stochastic maps with threshold equal to 4 ΔAIC units (Burnham and Anderson, 2003). We repeated model fit and model test for each of the 300 simulated datasets. We computed error rates as the frequency in which the model used to generate the data was not selected as the best

model.

Preliminary tests showed that it is often difficult to find the global maximum likelihood for the parameters of the model using minimization algorithms. Thus, we applied three distinct strategies to generate the starting point for the searches. First we generated starting points by drawing from a flat distribution with a large range of parameter values (from -400 to 400). Starting points generated using this approach are unlikely to be close to the global maximum but provide a wide sample of the parameter space. We refer to this approach as 'wide'. Second we used a more informed strategy by optimizing the parameters of the mathematical functions to produce evolutionary rates equal to σ^2 estimated for a homogeneous BM model with the response trait as the tip data. Then we defined a narrow range of parameter values around the best fit (by adding -10 and 10) and used this distribution to draw starting points. This starting point strategy, which we refer as 'narrow', makes sure that starting parameter values for the functions yield evolutionary rates that are congruent with the observed data and avoid possible problems resulting from starting at a flat region of the likelihood surface. Finally, we applied the most informative strategy by setting the starting point of the ML searches as the true parameter value for the models. When estimating a model different from the one that generated the data we set the parameters of the mathematical function as to minimize the distance relative to the evolutionary rates predicted by the true model for each trait category. We defined this search strategy as 'fixed'. We compare and discuss results among the different search strategies.

4.3.5 Likelihood surface for the linear function

Results from simulations showed that the linear model estimated with starting points draw from a wide range of the parameter space (the 'wide' strategy) often have worse fit than estimates with starting points informed by the data (the 'narrow' strategy). To investigate whether this pattern is associated with the shape of the likelihood surface rather than a systematic bias due to the implementation of a restricted pool of starting points when using the 'narrow' strategy we computed summary statistics from the results of maximum likelihood searches based on the simulated data.

First we calculated the number of times that the multiple independent searches for the same data reached the maximum likelihood score among all the searches, we defined this quantity as 'hits'. To compute the number of 'hits' we aggregated the maximum likelihood score returned by each of the independent searches. Then we counted the number of searches for which the Δlog-likelihood with respect to the maximum log-likelihood among all searches was less than the 0.0001 threshold. A large number of 'hits' indicates that multiple searches

starting from different points in the parameter space converged towards the same likelihood score and, therefore, the same combination of parameter values. Such result support the hypothesis that this point is the global maximum log-likelihood for the model. Following, we investigated whether there is an association between the number of 'hits' and how close is the model estimate to the true model that generated the data. For this we only used results based on the same models that generated the data.

4.4 Results

4.4.1 Performance simulations

First we computed the proportion of times that each of the generating models was recovered by model selection using the Akaike information criterion (AIC) while integrating the uncertainty associated with the use of stochastic mapped histories and comparing different approaches used to define the starting points for the maximum likelihood searches. Table 4.1 shows the number of times among all 100 simulated datasets that each model had the best AIC score, computed as the mean pairwise AIC over stochastic mapped histories. Then we used ΔAIC scores, also based on the pairwise AIC across stochastic mapped histories, to record support for (Table 4.2) and against (Table 4.3) the model that generated the data using a threshold of 4 ΔAIC units. For each simulation replicate and starting point strategy we plotted the distance between σ^2 values for the predictor trait categories estimated under the models and the true value for the simulation (Figures 4.3, 4.4 and 4.5). Below we describe results when data are simulated with the constant, linear, and step functions.

The majority of replicates, independent of search strategy, recovered the true model as the model with best AIC scores when data was generated under a constant evolutionary rate throughout the tree (Table 4.1). However, for most cases, there is a lack of support in favor (Table 4.2) or against (Table 4.3) the constant model when compared with alternative models. These results suggest that ΔAIC cannot substantiate any difference between models when data is simulated under a constant rate of evolution. Indeed, the distance between rates estimated for each model and the true rates is comparable among most of the replicates, starting point strategies, and models (Figure 4.3). Only the linear function estimated using the 'wide' starting point strategy can be rejected as a good model to explain the data when contrasted with the true model. Figure 4.3 (top row, middle column) shows that rates estimated under this model are much larger than reasonable values given the tree and the trait data, which suggests that the estimator failed to find the global maximum for the likelihood of the model.

Data generated under the linear function, where rates of evolution of the response trait show a positive correlation with predictor trait values, show results with the true model having the highest AIC on the majority of simulation replicates for both the 'narrow' and 'fixed' starting point approaches (Table 4.1). However, this result changes in the case of the 'wide' distribution of starting points, because the step function shows the best AIC scores even though the data was simulated under a linear function. When we compute the support for the model that generated the data both under the 'narrow' and 'fixed' starting points, we see that there is support for the linear model against the constant model on the majority of replicates and about half of the replicates show support in favor of the linear model when compared with the step model (Table 4.2). On the other hand, when using the 'wide' starting point distribution, ΔAIC scores point to both alternative models as better models than the true model that generated the data (Table 4.3). Parameter estimates for these models show that results from the linear model under the 'wide' starting point strategy produce overly large evolutionary rates compared to competing models (Figure 4.4 - top row, middle column), which is likely the reason for the poor fit to the data. Using either the 'narrow' and 'fixed' starting point produce better parameter estimates for the linear function (Figure 4.4 - middle and bottom rows) that reflect as more support for the true model based on ΔAIC scores (Table 4.2).

The step model is the most parameter-rich among the tested models, with three parameters (see Figure 4.1). When data was simulated under this model, only results from the 'wide' starting point distribution show the true model as the model most frequently with the highest AIC scores (Table 4.1). Under the other starting point strategies the step model is as frequent the best AIC scoring model as the linear model (Table 4.1). This pattern is also present in the results of the support for the step model over alternative models. When using the 'wide' starting point distribution the step model is favored over the linear model in the majority of replicates whereas there is no difference between these models in the majority of replicates when using the other starting point approaches (Tables 4.2 and 4.3). Parameter estimates for the linear function under the 'wide' starting point distribution show the same pattern as in the previous results: rates are too fast and too distant from the parameters that generated the data (Figure 4.5 - top row, middle column). On the other cases the rates estimates for the linear and the step models show good fit to the data.

4.4.2 Likelihood surface for the linear function

The maximum likelihood estimate (MLE) for the parameters of the different mathematical functions show a trend that the 'wide' approach to sample starting points often result in worse

fit than the other starting point strategies (see Figures 4.3, 4.4, and 4.5). When we compute the number of 'hits' among all replicates, each with 500 independent searches, for the 'wide' and 'narrow' starting point distributions we can see that the same trend is reflected in how often independent searches return the same maximum likelihood point. Figure 4.6 (middle row) shows the distribution of frequencies of number of 'hits' for the linear function using the 'wide' and 'narrow' starting schemes. When using the 'wide' starting point distribution, almost all replicates returned only a single 'hit' whereas the frequency of 2 or more 'hits' increase significantly under the 'narrow' search strategy. Interestingly, the effect of the 'wide' distribution of starting point on the number of 'hits' is not so pronounced with the other models in the set (Figure 4.6).

The association between a small number of 'hits' and poor parameter estimates for the model is made clear by the results shown in Figure 4.7. The mean absolute distance between the parameter estimates and the true value for the model is larger when the search result returned a single 'hit' than when 2 or more 'hits' occur. Furthermore, the variance of this distance is much larger when the search resulted in a single 'hit', meaning that one is likely to get a maximum likelihood estimate far from the parameter values that generated the data. Surprisingly, there is no gradual improvement in estimates with the number of 'hits', searches with 2 or more 'hits' show similar mean and variance of the distance between estimated and true parameter values for the model (Figure 4.7).

4.5 Discussion

The models described herein test for an association between a continuous trait and the evolutionary rates of a second trait evolving on the same phylogenetic tree. We tested the performance of three nested models using simulated datasets: the constant, linear and step functions. However, the same analysis framework used here can be easily extended to any mathematical function, nested or non-nested, that relates the values of a predictor trait to rates of evolution of a response trait under a Brownian-motion model. For instance, the R package `phylofx` provides six functions that can be easily implemented (Figure 4.1) besides the option to create any other function that better describes the hypothesis of rate variation observed in the data.

Our simulations show that it is difficult to recover the true model that generated the data as the best model using the Akaike information criterion (AIC) because, on average, there is not enough distinction among the ΔAIC scores of competing models (Table 4.2). However, parameter estimates for all models generally produced evolutionary rates congruent with rates used to generate the data even when the true model was not significantly different

than the alternative models (Figures 4.3, 4.4, 4.5). When data was generated under the simplest model, for example, the slope of the linear function (mean=0.01, sd=0.22) and the difference between rates of the step function (mean=0.01, sd=0.29) were on average centered on 0. This means that the gradient of evolutionary rates predicted across the tree was fairly similar across all models and suggests that any of the models in this set would yield similar conclusions about the macroevolution of the traits. Such results reinforce the idea that when performing multi-model inference it is important to focus on the parameter estimates for each model rather than turning our attention only to the model selection criteria or the arbitrary threshold statistics that may rank one model above the others (Caetano and Harmon, 2017a).

The performance of the linear model, that describes a linear relationship between the predictor trait and the rates of evolution of the response trait, was found to be dependent on the strategy used to draw random starting points for the maximum likelihood search. The likelihood surface for this model has multiple local optima which increase the chance that multiple independent searches will not return the global optima for the likelihood when starting from distant regions of the parameter space (See Figures 4.6 and 4.7). For this reason, searches using a very wide starting point region (i.e., the 'wide' strategy) will often result in poor parameter estimates for the linear function. However, using the evolutionary rate for a homogeneous Brownian-motion model estimated for the response trait to define a parameter space region where to draw starting points from (i.e., the 'narrow' strategy) is a reliable approach to set starting points that are within reach of the global maximum despite the presence of various local optima. We suggest that empirical studies implementing the linear function, or other parameter-rich mathematical functions, use a similar starting point strategy for model estimation.

The step model, different from the linear model, does not describe a continuous relationship between the predictor and response traits. This model applies a threshold value that is dependent on the predictor trait, then regions of the phylogenetic tree reconstructed as below the predictor threshold share a single rate that is independent of the rate for regions of the phylogeny reconstructed as above the threshold. For this reason, the contrast with the linear function is important to understand how the transition from discrete regime-oriented models to continuous functions can improve our inferences of macroevolutionary patterns. Our simulations show that when the linear model is properly estimated (using the 'narrow' starting point strategy) it is difficult to differentiate the regime-like step model from the continuous linear model (Tables 4.2 and 4.3). Two attributes of the models and inference might be driving such results: model complexity and sensibility to discretization.

There is a balance between the number of parameters that constitute a model and how well

the model fits to the data or, more precisely, how likely is the observed data to be generated under the model when implementing model selection using the Akaike information criterion (AIC). For instance, the linear model has one fewer parameter than the step model and, all else equal, it should be preferred over the latter. Indeed, we found the linear model to be preferred over the step model on 30% of our simulations whereas the reverse only occurred once (Table 4.2). The linear model predicts a distinct rate value for each predictor trait category following a strict relationship whereas the step model has two independent rate regimes. When the number of categories used to discretize the model is low, it is possible that two independent rate regimes fit to the data better even when the generating model was linear. Thus, support for the continuous model might increase if a higher number of rate categories is used. However, this is a hypothesis to be explored in future studies.

4.5.1 Future directions

Our simulations explored different relationships between the predictor trait and the response trait. Results suggest that we are able to estimate correctly the pattern of rate variation across the branches of the phylogenetic tree if rates are constant, positively or negatively associated with the predictor trait. However, differentiating between competing models using ΔAIC scores remains a challenge. Herein we used phylogenetic trees with 200 species and we discretized continuous functions using 5 rate categories (k). It is plausible, and likely, that both the number of species and k might influence the power to differentiate among models using AIC or other model choice criteria.

To perform such tests one need to increase both the number of rate categories used to discretize the models and the number of branches that stochastic mapped histories need to be simulated on. However the current implementation of stochastic mapping reconstructions (Huelsenbeck et al., 2003) offered by the R package `phytools` (Revell, 2012) does not scale well with the number of traits in the data and creates an important challenge for such tests. Fortunately, Irvahn and Minin (2014) have developed a framework[1] that drastically improves computational time for generating stochastic mapped histories. Incorporating this framework on our package `phylofx` will make it possible to perform analyses with a large number of species and rate categories for the predictor trait.

[1]This is not that much recent, I know. The issue is that the performance barrier associated with stochastic mapped histories increases faster in function of the number of traits than the number of species in the phylogeny. This happens because the $\mathbf{Q}$ matrix need to be exponentiated at every branch. However, studies often have a reasonable number of traits and the time to perform stochastic mappings is commonly bearable. This is not the case for our approach. For example, for 200 species and 15 traits a *single* stochastic map simulation using `phytools` can take up to 10 hours!

4.6 Concluding remarks

Interspecific trait variation plays a fundamental role in our understanding of evolutionary processes happening on multiple scales. Macroevolutionary analyses of trait evolution, for instance, focus on describing patterns and testing hypotheses about the phenotypic differentiation of lineages over time. However, we are still far from comprehending the general processes that may have shaped the diversity of traits on the majority of groups. Most of what we know is restricted to clades that show some kind of evolutionary 'key innovation' or the rare, but well-studied, examples of adaptive radiation. Evolutionary innovations attract studies that test whether such events are or not responsible for shifts in how lineages changed over time, unfortunately, these are unique events in the evolutionary history of a group and often provide results only applicable to the clade under study or their close relatives. Thus, if a general trend of phenotypic evolution across the tree of life is to be ever established, our focus need to shift from rare events to more commonly occurring predictors that can be applied to very disparate groups.

Different from trait evolution, lineage diversification studies have made significant advances towards testing large scale hypotheses derived from processes that might be acting on the dynamics of lineage accumulation over time. These are often based on gradients acting on ecological and evolutionary levels and, most importantly, have predictor factors common to most groups, such as latitude, elevation, ecological specialization and competition. The generality of such factors is the key element for the predictive power of lineage diversification hypotheses and allow the inclusion of distantly related clades into the same framework. Unfortunately, phenotypic evolution studies often focus on much more group-specific questions. The lack of suitable phylogenetic comparative methods to implement predictor gradients might play a role on the perceivable disparity of approaches when comparing diversification with trait evolution studies. In this light, our new framework expands the comparative method toolkit to incorporate the use of continuous predictor variables that describe rates of trait evolution varying across the branches of a phylogenetic tree using parameters for a series of mathematical functions. The introduction of mathematical functions allow for the estimate of powerful predictive models that can be used to test hypothesis regarding the gradient of general factors across the tree of life and, hopefully, help us understand more about general patterns and processes of trait evolution.

Table 4.1: Number of times each model showed the best Akaike information criterion (AIC) score across 100 simulations using different search strategies. For each simulation AIC was computed as the mean AIC score across stochastic mapped histories. 'Wide' denotes searches in which starting points were randomly draw from a wide uniform distribution of parameter values. 'Narrow' and 'Fixed' use more informed starting points for the MLE searches: the first draw from an uniform distribution around the parameter values after setting the function to produce rates of evolution equal to the rate estimated using a single rate Brownian motion model; the second sets the starting point as close as possible to the rates that generated the data.

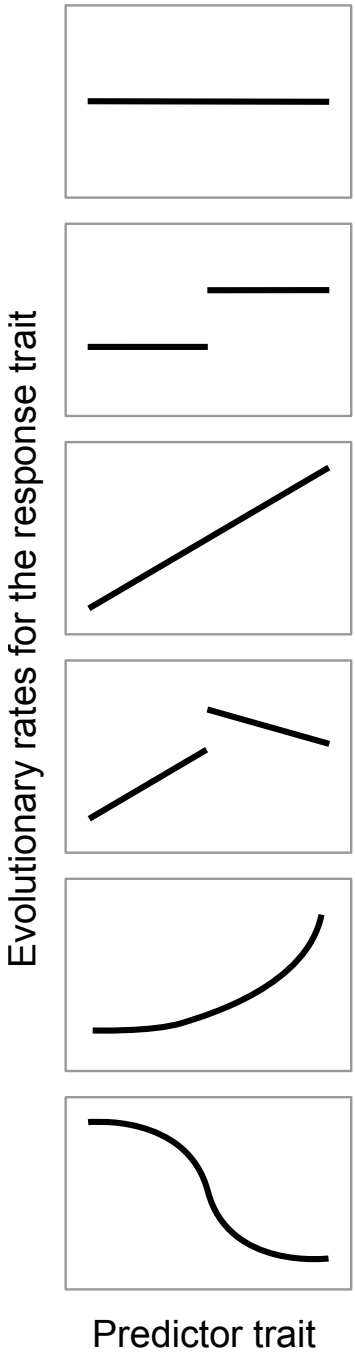
True model	**Search strategy**	**Constant**	**Linear**	**Step**
Constant	Wide	84	1	15
Constant	Narrow	75	11	14
Constant	Fixed	89	2	9
Linear	Wide	9	13	78
Linear	Narrow	3	84	13
Linear	Fixed	9	64	27
Step	Wide	17	8	75
Step	Narrow	16	54	30
Step	Fixed	20	39	41

Table 4.2: Number of times that mean pairwise ΔAIC across stochastic mapping histories for each model was larger than 4 units in favor of the model that generated the data. See main text and Table 4.1 for details on the different search strategies. Column 'N' shows the number of times that the true model had the best AIC score among all other models.

True model	Search strategy	N	Constant	Linear	Step
Constant	Wide	84	—	94	0
Constant	Narrow	75	—	1	0
Constant	Fixed	89	—	13	0
Linear	Wide	13	25	—	2
Linear	Narrow	84	78	—	30
Linear	Fixed	64	67	—	38
Step	Wide	75	38	79	—
Step	Narrow	30	26	1	—
Step	Fixed	41	24	18	—

Table 4.3: Number of times that mean pairwise ΔAIC across stochastic mapping histories for each model was larger than 4 units in favor of the alternative model when compared to the true model. See main text and Table 4.1 for details on search strategies. Column 'N' shows the number of times that the true model failed to show the best absolute AIC score across all other models.

True model	Search strategy	N	Constant	Linear	Step
Constant	Wide	16	—	0	2
Constant	Narrow	25	—	2	3
Constant	Fixed	11	—	1	0
Linear	Wide	87	65	—	77
Linear	Narrow	16	0	—	0
Linear	Fixed	46	5	—	7
Step	Wide	25	0	1	—
Step	Narrow	70	0	5	—
Step	Fixed	59	0	0	—



σ^2	A homogeneous rate throughout the tree.
$\sigma^2_{left} \longrightarrow x \leq break$ $\sigma^2_{right} \longrightarrow x > break$	Two homogeneous rate regimes with shifts ocurring at a threshold value of the predictor trait.
$f(x) = \beta_1 x + \beta_2$	A linear model. Rate changes linearly with the value of the predictor trait.
$f(x) = \beta_{1a} x + \beta_{2a} \longrightarrow x \leq break$ $f(x) = \beta_{1b} x + \beta_{2b} \longrightarrow x > break$	A two regimes linear model. Rate changes linearly with the value of the predictor trait, but the linear function shifts at a threshold
$f(x) = c\ x^p$	A power function model. Rate changes over a power function of the value of the predictor trait.
$f(x) = c_1 + \frac{c_2}{1+e^{\beta_1\ (\beta_2 - x)}}$	A sigmoid function model. This is the most flexible function implemented by default on the 'phylofx' package.

Figure 4.1: List of mathematical functions implemented in the R package `phylofx`. Plots on the left show regression lines between the values of the predictor trait and the rates of evolution of the response trait ($\sigma^2_{response}$). Middle column shows the mathematical function(s) associated with each plot and right column provides a brief description of each function. These functions are easily selected from a menu when using the `phylofx`, however any mathematical function that associates trait predictor trait values to $\sigma^2_{response}$ can be implemented on `phylofx`.

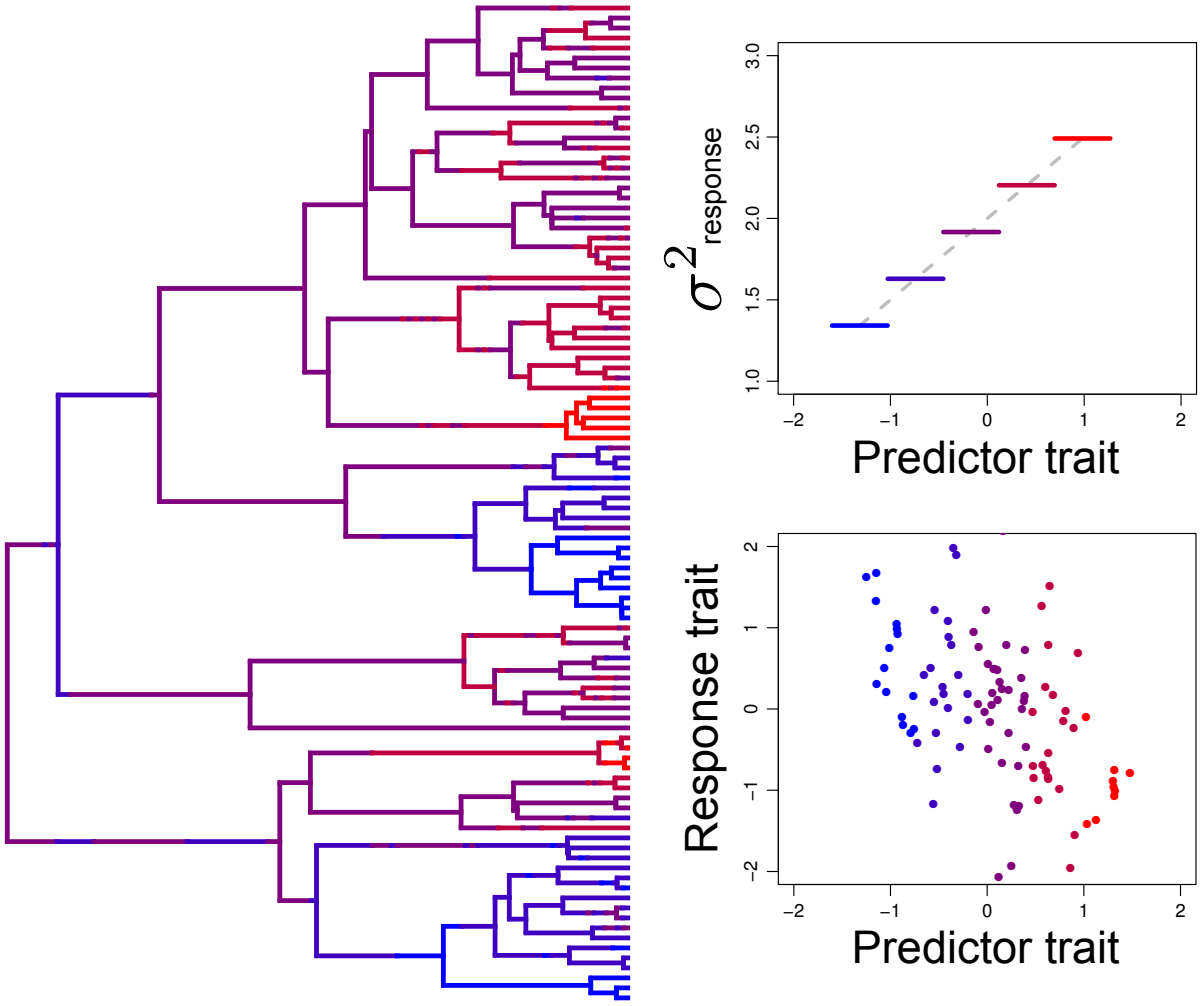


Figure 4.2: Example of phylogeny showing changes in σ^2 for the response trait across the branches in function of the predictor trait values following a linear model with positive slope. Cold and warm colors represent slower and faster evolutionary rates, respectively. All colors match among plots. Each portion of a branch showed in the phylogenetic tree on the left is mapped to a value of $\sigma^2_{response}$ following the ancestral reconstructed value for the predictor trait. The relationship between the rate of evolution and the predictor trait value is shown on the upper right plot. The dashed gray line represents the linear function that generated the data (see main text for information on parameter values). Each horizontal line is one of the five rate categories (k) used to discretize the continuous gradient of rates. All predictor trait values on the extent of these horizontal lines are mapped to a respective $\sigma^2_{response}$ value. Bottom right plot show a scatter plot between the predictor and response trait values for the tips of the tree. See that there is no clear correlation between the trait values, since the association here is between the predictor trait values and the rates of trait evolution of the response trait.

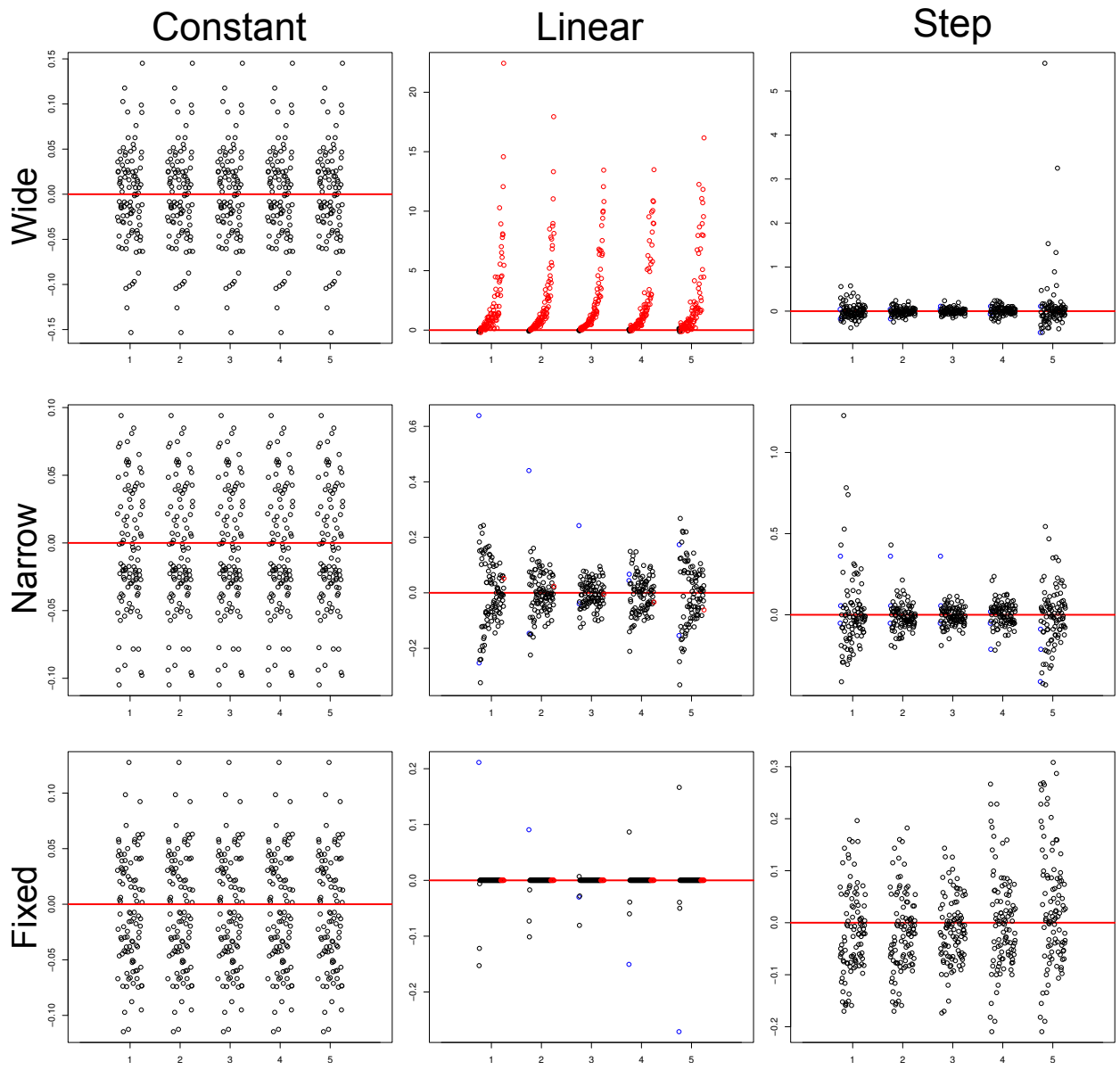


Figure 4.3: Results from performance simulations using datasets generated with a constant evolutionary rate. Each plot shows the distance between the true value for the evolutionary rate and the estimated value for each of the 5 predictor trait categories used in the analyses. Columns show parameter estimates under different models and rows correspond to three search strategies. The x axes are the predictor trait categories from 1 to 5 while y axes show $\sigma^2_{response}$ associated with each category. The horizontal red line marks 0, which correspond to parameter estimates equal to the true value used to generate the data. The color of the points mark if the model is significantly better (threshold of 4 ΔAIC units) than the model that generated the data (blue), worse than the generating model (red), or show no significant difference (black). Each point is the mean parameter estimate across stochastic mapped histories for each of the 100 simulations. Points were slightly dislocated horizontally for better visualization.

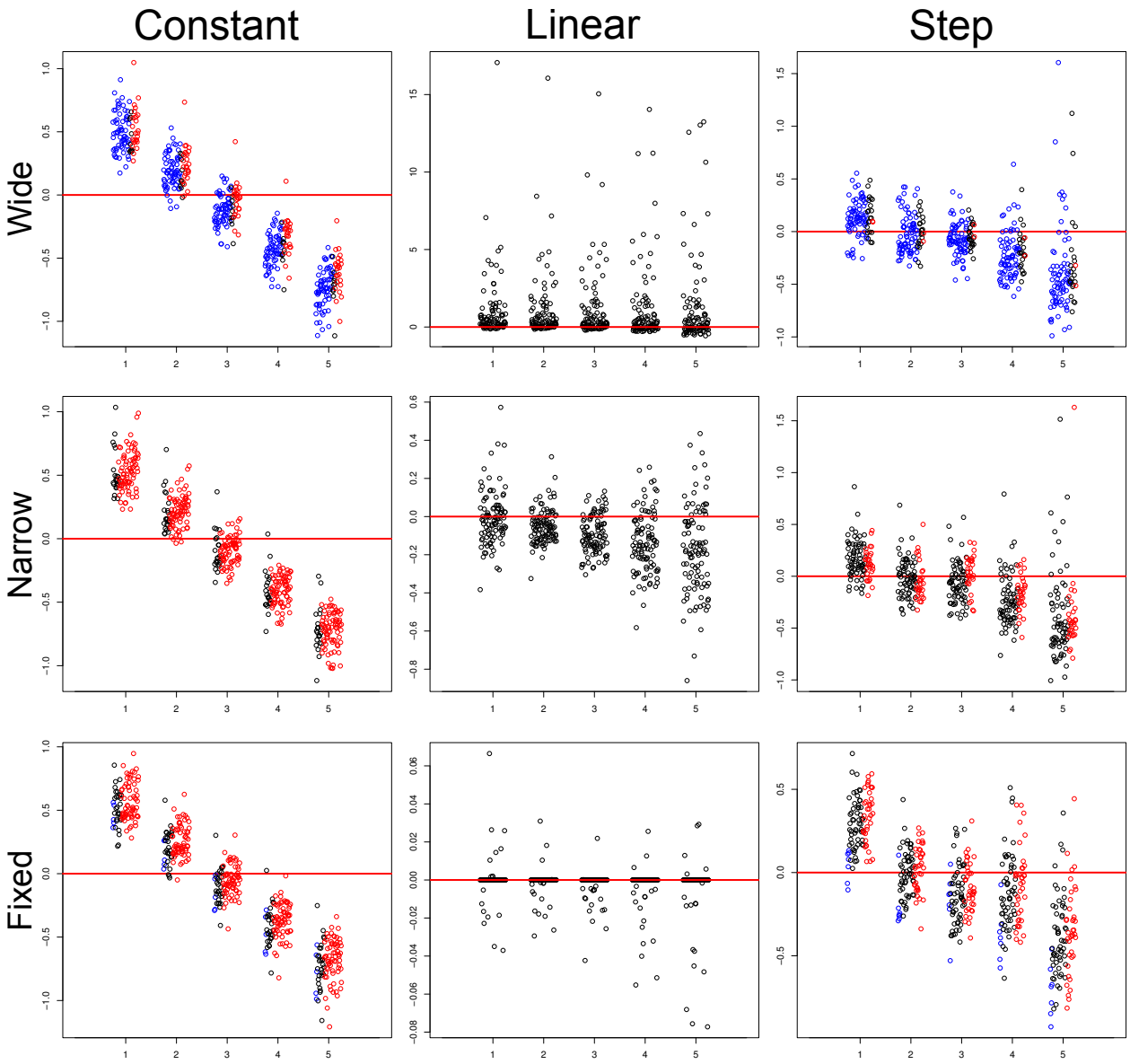


Figure 4.4: Results from performance simulations using datasets generated with a linear function between predictor trait values and rates of evolution of the response trait. Each plot shows the distance between the true value for the evolutionary rate and the estimated value for each of the 5 predictor trait categories used in the analyses. Columns show parameter estimates under different models and rows correspond to three search strategies. The x axes are the predictor trait categories from 1 to 5, y axes show σ^2 associated with each category. The horizontal red line marks 0, which correspond to parameter estimates equal to the true value used to generate the data. The color of the points mark if the model is significantly better (threshold of 4 ΔAIC units) than the model that generated the data (blue), worse than the generating model (red), or show no significant difference (black). Each point is the mean parameter estimate across stochastic mapped histories for each of the 100 simulations. Points were slightly dislocated horizontally for better visualization.

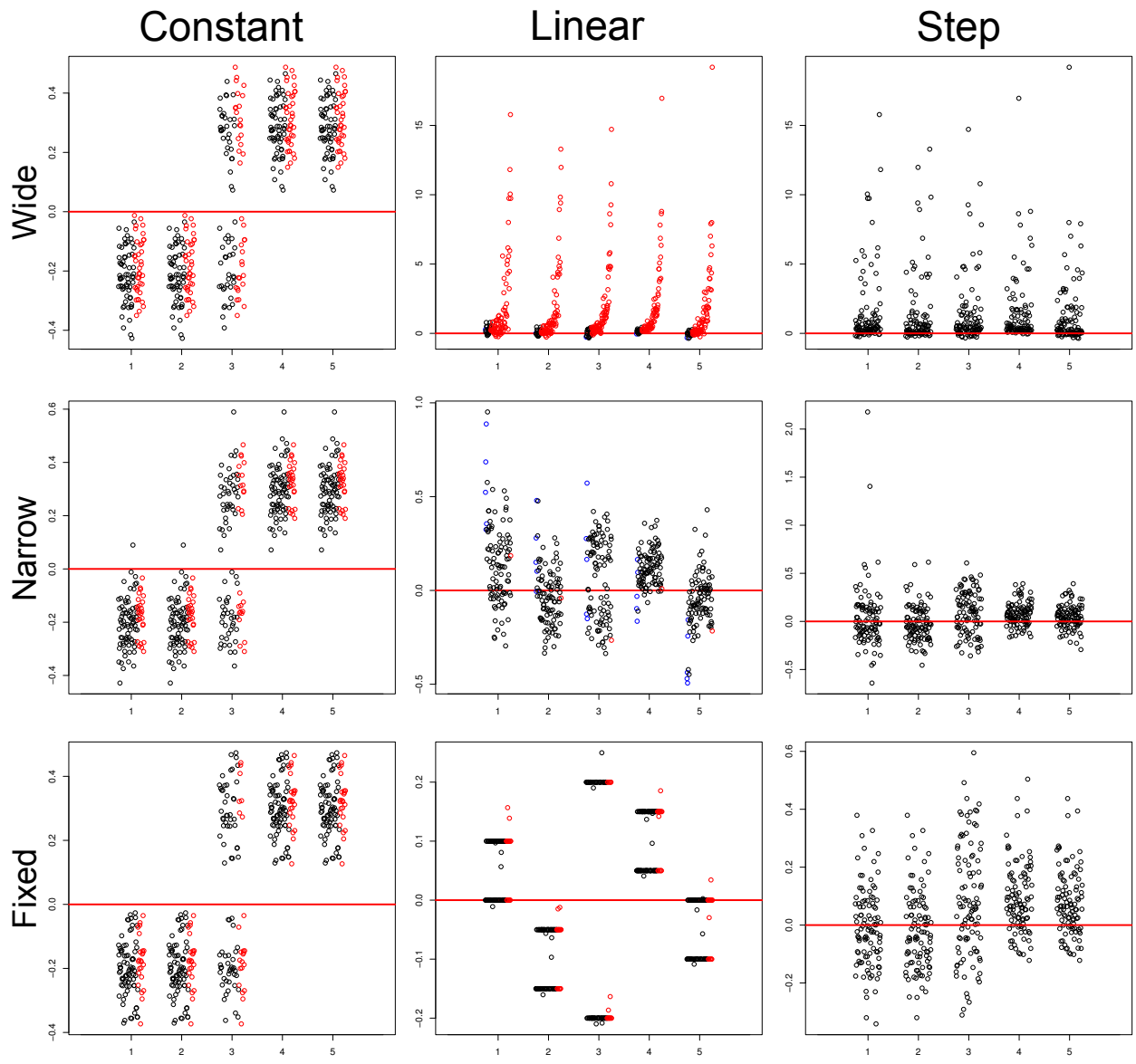


Figure 4.5: Results from performance simulations using datasets generated with a step function between predictor trait values and rates of evolution of the response trait. Each plot shows the distance between the true value for the evolutionary rate and the estimated value for each of the 5 predictor trait categories used in the analyses. Columns show parameter estimates under different models and rows correspond to three search strategies. The x axes are the predictor trait categories from 1 to 5, y axes show σ^2 associated with each category. The horizontal red line marks 0, which correspond to parameter estimates equal to the true value used to generate the data. The color of the points mark if the model is significantly better (threshold of 4 ΔAIC units) than the model that generated the data (blue), worse than the generating model (red), or show no significant difference (black). Each point is the mean parameter estimate across stochastic mapped histories for each of the 100 simulations. Points were slightly dislocated horizontally for better visualization.

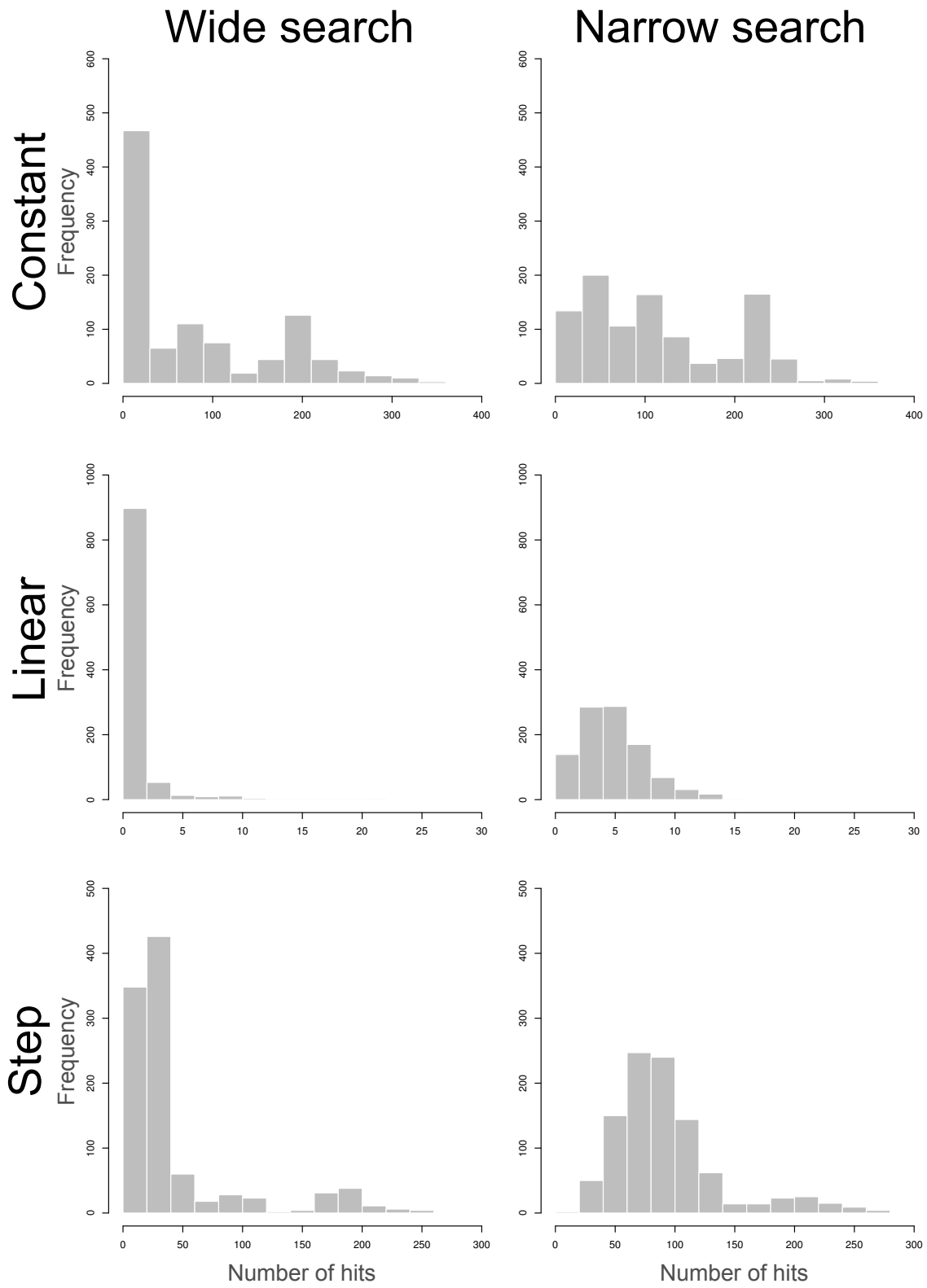


Figure 4.6: Number of 'hits' computed for 500 independent searches across each of the stochastic mapped histories for each simulation replicate and model. Large number of 'hits' means that multiple independent searches converged to the same log-likelihood score.

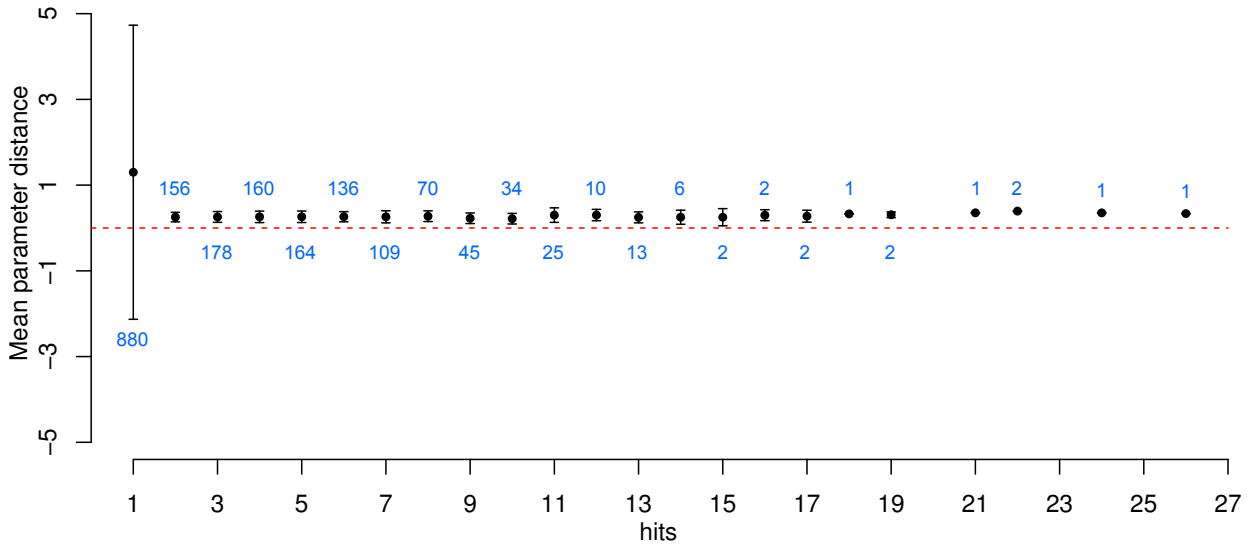


Figure 4.7: Relationship between the mean absolute distance of parameter estimates from the parameter values that generated the data and the number of 'hits'. Plot show the results for 2000 datasets generated using the linear function and estimated using the same model. For each replicate we performed 500 independent searches for the best log-likelihood score starting from random points in the parameter space (using both the 'wide' and 'narrow' starting point strategies). Points show the mean absolute distance between the estimated and the true parameter values for each parameter of the model (i.e., slope and intercept). Bars around the points show standard deviations and sample size numbers are shown above or below each point. The dashed red line marks when estimated parameter values are identical to the parameter values that generated the data.

Bibliography

Adams, D. C. 2012. Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. Syst. Biol. 62:181–192.

Adams, D. C. 2014a. A generalized K statistic for estimating phylogenetic signal from shape and other high-dimentional multivariate data. Syst. Biol. 63:685–697.

Adams, D. C. 2014b. Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. Syst. Biol. 63:166–177.

Adams, D. C., C. M. Berns, K. H. Kozak, and J. J. Wiens. 2009. Are rates of species diversification correlated with rates of morphological evolution? Proc. R. Soc. B. Biol. Sci. 276:2729–2738.

Adams, D. C. and R. N. Felice. 2014. Assessing trait covariation and morphological integration on phylogenies using evolutionary covariance matrices. PLoS One 9:e94335.

Adams, D. C. and E. Otárola-Castillo. 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Method. Ecol. Evol. 4:393–399.

Alfaro, M. E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, and L. J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc Natl Acad Sci USA 106:13410–13414.

Almeida, P. C., D. T. Feitosa, P. Passos, and A. L. C. Prudente. 2014. Morphological variation and taxonomy of Atractus latifrons (Günther, 1868) (Serpentes: Dipsadidae). Zootaxa 3860:64.

Armbruster, W. S., C. Pélabon, G. H. Bolstad, and T. F. Hansen. 2014. Integrated phenotypes: Understanding trait covariation in plants and animals. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369:20130245.

Armbruster, W. S. and K. E. Schwaegerle. 1996. Causes of covariation of phenotypic traits among populations. J. Evol. Biol. 9:261–276.

Arnold, S. J. 1992. Constraints on phenotypic evolution. Am. Nat. 140:S85–S107.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112-113:9–32.

Barnard, J., R. McCulloch, and X.-L. Meng. 2000. Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. Stat. Sinica 10:1281–1312.

Bartoszek, K., J. Pienaar, P. Mostad, S. Andersson, and T. F. Hansen. 2012. A phylogenetic comparative method for studying multivariate adaptation. J. Theor. Biol. 314:204–215.

Beaulieu, J. M. and B. C. O'Meara. 2016. Detecting Hidden Diversification Shifts in Models of Trait-Dependent Speciation and Extinction. Syst. Biol. 65:583–601.

Benson, R. B. J. and J. N. Choiniere. 2013. Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. Proc. R. Soc. Lond. B. Biol. Sci. 280:20131780.

Bittner, T. D. and R. B. King. 2003. Gene flow and melanism in garter snakes revisited: a comparison of molecular markers and island vs. coalescent models. Biol. J. Linn. Soc. 79:389–399.

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.

Bolstad, G. H., T. F. Hansen, C. Pélabon, M. Falahati-Anbaran, R. Pérez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. Phil. Trans. R. Soc. B 369:20130255.

Bonetti, M. F. and J. J. Wiens. 2014. Evolution of climatic niche specialization: a phylogenetic analysis in amphibians. Proc. R. Soc. Lond. B. Biol. Sci. 281:20133229.

Bonnet, X., G. Naulleau, and R. Shine. 1999. The dangers of leaving home: dispersal and mortality in snakes. Biol. Cons. 89:39–50.

Boucher, F. C. and V. Démery. 2016. Inferring Bounded Evolution in Phenotypic Characters from Phylogenetic Comparative Data. Syst. Biol. 65:651–661.

Brakefield, P. M. 1990. Genetic drift and patterns of diversity among colour-polymorphic populations of the homopteran *Philaenus spumarius* in an island archipelago. Biol. J. Linn. Soc. 39:219–237.

Brattstrom, B. H. 1955. The Coral Snake 'Mimic' Problem and Protective Coloration. Evolution 9:217–219.

Brodie III, E. 1993. Differential Avoidance of Coral Snake Banded Patterns by Free-Ranging Avian Predators in Costa Rica. Evolution 47:227–235.

Brodie III, E. D. B. and F. J. Janzen. 1995. Experimental Studies of Coral Snake Mimicry: Generalized Avoidance of Ringed Snake Patterns by Free-Ranging Avian Predators. Func. Eco. 9:186.

Buasso, C. M., G. C. Leynaud, and F. B. Cruz. 2006. Predation on snakes of Argentina: Effects of coloration and ring pattern on coral and false coral snakes. Stud. Neot. Fauna and Env. 41:183–188.

Burnham, K. P. and D. R. Anderson. 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd edition ed. Springer, New York.

Butler, M. A. and A. A. King. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. Am. Nat. 164:683–695.

Caetano, D. S. and L. J. Harmon. 2017a. Estimating rates of trait evolution with uncertainty. bioRxiv doi: `https://doi.org/10.1101/102939`.

Caetano, D. S. and L. J. Harmon. 2017b. ratematrix: An R package for studying evolutionary integration among several traits on phylogenetic trees. Method. Ecol. Evol. .

Campbell, J. and W. Lamar. 2004. The Venomous Reptiles of the Western Hemisphere. Cornell University Press, Ithaca, NY.

Clavel, J., G. Escarguel, and G. Merceron. 2015. mvmorph: An R package for fitting multivariate evolutionary models to morphometric data. Method. Ecol. Evol. 6:1311–1319.

Clavel, J. and H. Morlon. 2017. Accelerated body size evolution during cold climatic periods in the Cenozoic. Proc. Natl. Acad. Sci. 114:4183–4188.

Claverie, T., E. Chan, and S. N. Patek. 2011. Modularity and scaling in fast movements: power amplification in mantis shrimp. Evolution 65:443–61.

Claverie, T. and S. N. Patek. 2013. Modularity and rates of evolutionary change in a power-amplified prey capture system. Evolution 67:3191–3207.

Collar, D. C., T. J. Near, and P. C. Wainwright. 2005. Comparative analysis of morphological diversity: Does disparity accumulate at the same rate in two lineages of centrarchid fishes? Evolution 59:1783–1794.

Collar, D. C., P. C. Wainwright, M. E. Alfaro, L. J. Revell, and R. S. Mehta. 2014. Biting disrupts integration to spur skull evolution in eels. Nature Communications 5:5505.

Cooper, N. and A. Purvis. 2009. What factors shape rates of phenotypic evolution? A comparative study of cranial morphology of four mammalian clades. J. Evol. Biol. 22:1024–1035.

Dececchi, T. A. and H. C. E. Larsson. 2013. Body and limb size dissociation at the origin of birds: Uncoupling allometric constraints across a macroevolutionary transition. Evolution 67:2741–2752.

Denton, J. S. S. and D. C. Adams. 2015. A new phylogenetic test for comparing multiple high-dimensional evolutionary rates suggests interplay of evolutionary rates and modularity in lanternfishes (Myctophiformes; Myctophidae). Evolution 69:2425–2440.

Dines, J. P., E. Otárola-Castillo, P. Ralph, J. Alas, T. Daley, A. D. Smith, and M. D. Dean. 2014. Sexual selection targets cetacean pelvic bones. Evolution 68:3296–3306.

Diniz-Filho, J. A. F. and R. Balestra. 1998. Hierarchical effects on body size evolution and macroecology of south american rainforest mammals. Ecol. Aust. 8:23–30.

Downs, F. 1967. Intrageneric relationships among colubrid snakes of the genus *Geophis* wagler. Misc. Publ. Mus. Zool. Univ. Mich. 131:1–188.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29:1969–1973.

Dunn, E. R. 1954. The Coral Snake 'Mimic' Problem in Panama. Evolution 8:97–102.

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.

Fan, Y., R. Wu, M.-H. Chen, L. Kuo, and P. O. Lewis. 2011. Choosing among partition models in Bayesian phylogenetics. Mol. Biol. Evol. 28:523–532.

Felsenstein, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Hum. Genet. 25:471–492.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Felsenstein, J. 1988. Phylogenies and quantitative characters. Annu. Rev. Eco. Sys. 19:445–471.

Felsenstein, J. 2004. Brownian motion and gene frequencies. chap. 23, Pages 391–414 *in* Inferring Phylogenies. Sinauer.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Met. Eco. Evo. 3:1084–1092.

FitzJohn, R. G., W. P. Maddison, and S. P. Otto. 2009. Estimating Trait-Dependent Speciation and Extinction Rates from Incompletely Resolved Phylogenies. Syst. Biol. 58:595–611.

Forsman, A. 1995. Opposing fitness consequences of colour pattern in male and female snakes. J. Evol. Biol. 8:53–70.

Freckleton, R. P. 2012. Fast likelihood calculations for comparative analyses. Method. Ecol. Evol. 3:940–947.

Frédérich, B., G. Marramà, G. Carnevale, and F. Santini. 2016. Non-reef environments impact the diversification of extant jacks, remoras and allies (Carangoidei, Percomorpha). Proc. R. Soc. B 283:20161556.

Gadow, H. 1908. Through southern Mexico. Witherby and Co. Press, London, UK.

Galetti, M., R. Guevara, M. C. Côrtes, R. Fadini, S. V. Matter, A. B. Leite, F. Labecca, T. Ribeiro, C. S. Carvalho, R. G. Collevatti, M. M. Pires, P. R. Guimarães, P. H. Brancalion, M. C. Ribeiro, and P. Jordano. 2013. Functional Extinction of Birds Drives Rapid Evolutionary Changes in Seed Size. Science 340:1086–1090.

Gamble, T., A. J. Geneva, R. E. Glor, and D. Zarkower. 2014. *Anolis* sex chromosomes are derived from a single ancestral pair. Evolution 68:1027–1041.

Gelman, A., J. Hwang, and A. Vehtari. 2013. Understanding predictive information criteria for Bayesian models. Statistics and Computing 24:997–1016.

Gelman, A. and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7:457–472.

Glor, R. E., J. B. Losos, and A. Larson. 2005. Out of Cuba: Overwater dispersal and speciation among lizards in the *Anolis carolinensis* subgroup. Mol. Ecol. 14:2419–2432.

Goolsby, E. W. 2016. Likelihood-based parameter estimation for high-dimensional phylogenetic comparative models: Overcoming the limitations of "distance-based" methods. Syst. Biol. 65:852–870.

Goswami, A. 2006. Cranial modularity shifts during mammalian evolution. Am. Nat. 168:270–280.

Goswami, A., W. J. Binder, J. Meachen, and F. R. O'Keefe. 2015. The fossil record of phenotypic integration and modularity: A deep-time perspective on developmental and evolutionary dynamics. Proc. Natl. Acad. Sci. 112:4891–4896.

Goswami, A., J. B. Smaers, C. Soligo, and P. D. Polly. 2014. The macroevolutionary consequences of phenotypic integration: From development to deep time. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369:20130254.

Grafen, A. 1989. The Phylogenetic Regression. Phil. Trans. R. Soc. B 326:119–157.

Grazziotin, F. G., H. Zaher, R. W. Murphy, G. Scrocchi, M. A. Benavides, Y.-P. Zhang, and S. L. Bonatto. 2012. Molecular phylogeny of the New World Dipsadidae (Serpentes: Colubroidea): a reappraisal. Cladistics 28:437–459.

Greene, H. W. and R. W. McDiarmid. 1981. Coral Snake Mimicry: Does It Occur? Science 213:1207–1212.

Gustavson, F. G., J. Waśniewski, J. J. Dongarra, and J. Langou. 2010. Rectangular full packed format for Cholesky's algorithm: Factorization, solution and inversion. ACM T. Math. Software 37:1–33.

Hadfield, J. D. and S. Nakagawa. 2010. General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. J. Evol. Biol. 23:494–508.

Hallgrímsson, B., H. A. Jamniczky, N. M. Young, C. Rolian, U. Schmidt-Ott, and R. S. Marcucio. 2012. The generation of variation and the developmental basis for evolutionary novelty. J. Exp. Zool. B Mol. Dev. Evol. 318:501–517.

Hansen, T. F. and D. Houle. 2004. Evolvability, Stabilizing Selection, and the Problem of Stasis. *in* Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes (M. Pigliucci and K. Preston, eds.). Oxford University Press.

Hardy, N. B. and S. P. Otto. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. Proc. R. Soc. Lond. B. Biol. Sci. 281:20132960.

Harmon, L. J., J. B. Losos, T. Jonathan Davies, R. G. Gillespie, J. L. Gittleman, W. Bryan Jennings, K. H. Kozak, M. A. McPeek, F. Moreno-Roark, T. J. Near, A. Purvis, R. E. Ricklefs, D. Schluter, J. A. Schulte II, O. Seehausen, B. L. Sidlauskas, O. Torres-Carvajal, J. T. Weir, and A. Ø. Mooers. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64:2385–2396.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Heard, S. B. and D. L. Hauser. 1995. Key evolutionary innovations and their ecological mechanisms. Hist. Biol. 10:151–173.

Hecht, M. K. and D. Marien. 1956. The coral snake mimic problem: A reinterpretation. J. Morp. 98:335–365.

Hinman, K. E., H. L. Throop, K. L. Adams, A. J. Dake, K. K. McLauchlan, and M. J. McKone. 1997. Predation by Free-Ranging Birds on Partial Coral Snake Mimics: The Importance of Ring Width and Color. Evolution 51:1011–1014.

Hipsley, C. A., D. B. Miles, and J. Muller. 2014. Morphological disparity opposes latitudinal diversity gradient in lacertid lizards. Biol. Lett. 10:20140101–20140101.

Hohenlohe, P. A. and S. J. Arnold. 2008. MIPoD: A hypothesis-testing framework for microevolutionary inference from patterns of divergence. Am. Nat. 171:366–385.

Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic mapping of morphological characters. Syst. Biol. 52:131–158.

Huelsenbeck, J. P. and B. Rannala. 2003. Detecting correlation between characters in a comparative analysis with uncertain phylogeny. Evolution 57:1237–1247.

Hunt, G. 2006. Fitting and Comparing Models of Phyletic Evolution: Random Walks and beyond. Paleobiology 32:578–601.

Hunt, G., M. J. Hopkins, and S. Lidgard. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. Proc. Natl. Acad. Sci. 112:4885–4890.

Immler, S., A. Gonzalez-Voyer, and T. R. Birkhead. 2012. Distinct evolutionary patterns of morphometric sperm traits in passerine birds. Proc. R. Soc. Lond. B. Biol. Sci. 279:4174–4182.

Irvahn, J. and V. N. Minin. 2014. Phylogenetic stochastic mapping without matrix exponentiation. J. Comp. Biol. 21:676–690.

Jeffords, M. R., J. G. Sternburg, and G. P. Waldbauer. 1979. Batesian Mimicry: Field Demonstration of the Survival Value of Pipevine Swallowtail and Monarch Color Patterns. Evolution 33:275–286.

Kass, R. E. and A. E. Raftery. 1995. Bayes Factors. J. Am. Stat. Ass. 90:773–795.

Katoh, K., K.-i. Kuma, H. Toh, and T. Miyata. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nuc. A. Res. 33:511–518.

King, R. B. and R. Lawson. 1995. Color-Pattern Variation in Lake Erie Water Snakes: The Role of Gene Flow. Evolution 49:885–896.

Klingenberg, C. and J. Marugán-Lobón. 2013. Evolutionary covariation in geometric morphometric data: Analyzing integration, modularity, and allometry in a phylogenetic context. Syst. Biol. 62:591–610.

Klingenberg, C. P. 2014. Studying morphological integration and modularity at multiple levels: Concepts and analysis. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369:20130249.

Kuhn, T. S., A. Ø. Mooers, and G. H. Thomas. 2011. A simple polytomy resolver for dated phylogenies. Met. Eco. Evo. 2:427–436.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. Evolution 33:402–416.

Lindell, L. E. and A. Forsman. 1996. Sexual dichromatism in snakes: support for the flicker-fusion hypothesis. Can. J. Zoo. 74:2254–2256.

Liu, H., Z. Zhang, and K. J. Grimm. 2016. Comparison of Inverse Wishart and Separation-Strategy Priors for Bayesian Estimation of Covariance Parameter Matrix in Growth Curve Analysis. Structural Equation Modeling: A Multidisciplinary Journal 23:354–367.

Lorioux, S., X. Bonnet, F. Brischoux, and M. De Crignis. 2008. Is melanism adaptive in sea kraits? Amp.-Rep. 29:1–5.

Losos, J. B. 2009. Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles. University of California Press.

Maddison, W. P. and R. G. FitzJohn. 2015. The Unsolved Challenge to Phylogenetic Correlation Tests for Categorical Characters. Syst. Biol. 64:127–136.

Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a Binary Character's Effect on Speciation and Extinction. Syst. Biol. 56:701–710.

Mahler, D. L., T. Ingram, L. J. Revell, and J. B. Losos. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. Science 341:292–295.

Mahler, D. L., L. J. Revell, R. E. Glor, and J. B. Losos. 2010. Ecological opportunity and the rate of morphological evolution in the diversification of greater Antillean anoles. Evolution 64:2731–2745.

Mallet, J. and M. Joron. 1999. Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. Ann. Rev. Eco. Syst. 30:201–233.

Mappes, J., N. Marples, and J. Endler. 2005. The complex business of survival by aposematism. Trend. Ecol. Evol. 20:598–603.

Marques, O. A. V. and G. Puorto. 1991. Padrões cromáticos, distribuição e possível mimetismo em *Erythrolamprus aesculapii* (Serpentes, Colubridae). Mem. Inst. But. 53:127–134.

Martins, M. and M. E. Oliveira. 1993. The snakes of the genus *Atractus* Wagler (Reptilia: Squamata: Colubridae) from the Manaus region, central Amazonia, Brazil. Zool. Mede. 67:21–40.

Martins, M. and M. E. Oliveira. 1998. Natural history of snakes in forests of the Manaus region, Central Amazonia, Brazil. Herp. Nat. Hist. 6:78–150.

Melo, D., A. Porto, J. M. Cheverud, and G. Marroig. 2016. Modularity: Genes, development, and evolution. Annu. Rev. Ecol. Evol. Syst. 47:463–486.

Merilaita, S. and J. Lind. 2005. Background-matching and disruptive coloration, and the evolution of cryptic coloration. Proc. R. Soc. London B 272:665–670.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21:1087–1092.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-Based Selection of Likelihood Models for Phylogeny Estimation. Syst. Biol. 52:674–683.

Monteiro, L. R. and M. R. Nogueira. 2010. Adaptive radiations, ecological specialization, and the evolutionary integration of complex morphological structures. Evolution 64:724–744.

Moore, B. R., S. Höhna, M. R. May, B. Rannala, and J. P. Huelsenbeck. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. Proc. Nat. Ac. Sci. Page 201518659.

Moreno-Arias, R. A. and M. L. Calderón-Espinosa. 2016. Patterns of morphological diversification of mainland *Anolis* lizards from northwestern South America. Zool. J. Linn. Soc. 176:632–647.

Nicholson, K. E., R. E. Glor, J. J. Kolbe, A. Larson, S. Blair Hedges, and J. B. Losos. 2005. Mainland colonization by island lizards: Mainland colonization by *Anolis*. J. Biogeogr. 32:929–938.

Olson, E. and R. Miller. 1958. Morphological integration. Univ. of Chicago Press, Chicago, IL.

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.

Pagel, M. 1994. Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. Proc. R. Soc. Lond. B. Biol. Sci. 255:37–45.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst. Biol. 48:612–622.

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.

Pfennig, D. W., C. K. Akcali, and D. W. Kikuchi. 2015. Batesian mimicry promotes pre- and post-mating isolation in a snake mimicry complex. Evolution Pages 1085–1090.

Pfennig, D. W., W. R. Harcombe, and K. S. Pfennig. 2001. Frequency-dependent Batesian mimicry. Nature 410:323–323.

Pfennig, D. W. and S. P. Mullen. 2010. Mimics without models: causes and consequences of allopatry in Batesian mimicry complexes. Proc. R. Soc. B 277:2577–2585.

Phillips, P. C. and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.

Pigot, A. L., I. P. F. Owens, and C. D. L. Orme. 2012. Speciation and Extinction Drive the Appearance of Directional Range Size Evolution in Phylogenies and the Fossil Record. PLoS Biol. 10:e1001260.

Pinheiro, C. E. G. 2011. On the evolution of warning coloration, Batesian and Müllerian mimicry in Neotropical butterflies: the role of jacamars (Galbulidae) and tyrant-flycatchers (Tyrannidae). J. Avi. Biol. 42:277–281.

Pinto, G., D. L. Mahler, L. J. Harmon, and J. B. Losos. 2008. Testing the island effect in adaptive radiation: Rates and patterns of morphological diversification in Caribbean and mainland *Anolis* lizards. Proc. R. Soc. B 275:2749–2757.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

Pough, F. H. 1988. Mimicry and related phynomena. Pages 154–234 *in* Biology of Reptilia, Ecology B, Defence and Life History (C. Gans and R. Huey, eds.) vol. 16. Allan R. Liss inc., New York, NY.

Przeczek, K., C. Mueller, and S. M. Vamosi. 2008. The evolution of aposematism is accompanied by increased diversification. Int. Zoo. 3:149–156.

Pyron, R. and F. T. Burbrink. 2009. Body size as a primary determinant of ecomorphological diversification and the evolution of mimicry in the lampropeltinine snakes (Serpentes: Colubridae). J. Evol. Biol. 22:2057–2067.

Rabosky, A. R. D., C. L. Cox, D. L. Rabosky, P. O. Title, I. A. Holmes, A. Feldman, and J. A. McGuire. 2016. Coral snakes predict the evolution of mimicry across New World snakes. Nature Comm. 7:11484.

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.

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.

Rabosky, D. L. and E. E. Goldberg. 2015. Model Inadequacy and Mistaken Inferences of Trait-Dependent Speciation. Syst. Biol. Page syu131.

Rabosky, D. L., J. S. Mitchell, and J. Chang. 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. Syst. Biol. .

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. Nature Communications 4.

Revell, L. J. 2009. Size-correction and principal components for interspecific comparative studies. Evolution 63:3258–3268.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Method. 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 L. J. Harmon. 2008. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evol. Ecol. Res. 10:311–331.

Roze, J. 1996. Coral snakes of the Americas: Biology, identification, and venoms. Krieger Publishing Company, Malabar, FL.

Santos, J. C., L. A. Coloma, and D. C. Cannatella. 2003. Multiple, recurring origins of aposematism and diet specialization in poison frogs. Proc. Nat. Ac. Sci. 100:12792–12797.

Savage, J. and J. Slowinski. 1992. The colouration of the venomous coral snakes (family Elapidae) and their mimics (families Aniliidae and Colubridae). Biol. J. Linn. Soc. 45:235–254.

Savage, J. M. 2002. The amphibian and reptiles of Costa Rica: Herpetofauna between two continents, between two seas. Chicago Univ. Press, Chigago, IL.

Sawaya, R. J., O. A. V. Marques, and M. Martins. 2008. Composition and natural history of a Cerrado snake assemblage at Itirapina, São Paulo state, southeastern Brazil. Bio. Neot. 8:0–0.

Sazima, I. and A. S. Abe. 1991. Habits of five Brazilian snakes with coral-snake pattern, including a summary of defensive tactics. Stud. Neot. Fauna Env. 26:159–164.

Schaad, E. W. and S. Poe. 2010. Patterns of ecomorphological convergence among mainland and island *Anolis* lizards. Biol. J. Linnean Soc. 101:852–859.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.

Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of Ancestor States in Adaptive Radiation. Evolution 51:1699–1711.

Simpson, G. G. 1953. The major features of evolution. Columbia University Press, New York.

Slater, G. J. 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. Methods Ecol. Evol. 4:734–744.

Smith, S. A. and C. W. Dunn. 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24:715–716.

Smith, S. M. 1975. Innate recognition of coral snake pattern by a possible avian predador. Science 187:759–760.

Speed, M. P., M. A. Brockhurst, and G. D. Ruxton. 2010. The dual benefits of aposematism: Predator avoidance and enhanced resource collection. Evolution 64:1622–1633.

Speed, M. P. and G. D. Ruxton. 2005. Aposematism: what should our starting point be? Proc. R. Soc. B 272:431–438.

Stadler, T. 2009. On incomplete sampling under birth–death models and connections to the sampling-based coalescent. J. Theor. Bio. 261:58–66.

Stevens, M. and S. Merilaita. 2009. Animal camouflage: current issues and new perspectives. Phil. Trans. R. Soc. London B 364:423–427.

Stevens, M. and G. D. Ruxton. 2012. Linking the evolution and form of warning coloration in nature. Phil. Trans. R. Soc. London B 279:417–426.

Tanaka, K. 2005. Thermal aspects of melanistic and striped morphs of the snake Elaphe quadrivirgata. Zoo. Sci. 22:1173–1179.

Thayer, G. 1909. Concealing-coloration in the animal kingdom: An exposition of the laws of disguise through color and pattern. 2 ed. The MacMillan Co., New York, NY.

Thomas, G. H., S. Meiri, and A. B. Phillimore. 2009. Body size diversification in *Anolis*: Novel environment and island effects. Evolution 63:2017–2030.

Thompson, J. N., C. Schwind, P. R. Guimarães, and M. Friberg. 2013. Diversification through multitrait evolution in a coevolving interaction. Proc. Natl. Acad. Sci. 110:11487–11492.

Titcomb, G. C., D. W. Kikuchi, and D. W. Pfennig. 2014. More than mimicry? Evaluating scope for flicker-fusion as a defensive strategy in coral snake mimics. Curr. Zoo. 60:123–130.

Tozetti, A. M., R. B. d. Oliveira, and G. M. F. Pontes. 2009. Defensive repertoire of *Xenodon dorbignyi* (Serpentes, Dipsadidae). Bio. Neot. 9:157–163.

Uetz, P., J. Hosek, and (eds.). 2014. The reptile database. http://www.reptile-database.org/.

Uyeda, J. C., D. S. Caetano, and M. W. Pennell. 2015. Comparative analysis of principal components can be misleading. Syst. Biol. 64:677–689.

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.

Villmoare, B. 2012. Morphological integration, evolutionary constraints, and extinction: A computer simulation-based study. Evol. Biol. 40:76–83.

Wallace, A. R. 1867. Mimicry, and other protective resemblances among animals. Alfred Russel Wallace Classic Writings 8:1–27.

Xie, W., P. O. Lewis, Y. Fan, L. Kuo, and M. H. Chen. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst. Biol. 60:150–160.

Young, N. M. and B. Hallgrímsson. 2005. Serial homology and the evolution of mammalian limb covariation structure. Evolution 59:2691–2704.

Young, N. M., G. P. Wagner, and B. Hallgrímsson. 2010. Development and the evolvability of human limbs. Proc. Natl. Acad. Sci. USA 107:3400–3405.

Zaher, H., F. G. Grazziotin, J. E. Cadle, R. W. Murphy, J. C. d. Moura-Leite, and S. L. Bonatto. 2009. Molecular phylogeny of advanced snakes (Serpentes, Caenophidia) with an emphasis on South American Xenodontines: a revised classification and descriptions of new taxa. Pap. Avul. Zoo. 49:115–153.

Zhang, X., W. J. Boscardin, and T. R. Belin. 2006. Sampling correlation matrices in Bayesian models with correlated latent variables. J. Comp. Graph. Stat. 15:880–896.

Zweifel, R. G. 1960. Results of the Puritan-American Museum of Natural History Expedition to Western Mexico. 9, Herpetology of the Tres Marías Islands. Bulletin of the AMNH ; v. 119, article 2 .

Zwickl, D. J. 2011. GARLI 2.0. https://www.nescent.org/wg_garli/main_page.