



#### ANNUAL REVIEWS **Further**

Click [here](#) to view this article's online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

# Integrating Paleontological and Phylogenetic Approaches to Macroevolution

Gene Hunt<sup>1</sup> and Graham Slater<sup>2</sup>

<sup>1</sup>Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560; email: [hunte@si.edu](mailto:hunte@si.edu)

<sup>2</sup>Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637; email: [gslater@uchicago.edu](mailto:gslater@uchicago.edu)

Annu. Rev. Ecol. Evol. Syst. 2016. 47:189–213

First published online as a Review in Advance on August 19, 2016

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

This article's doi:  
10.1146/annurev-ecolsys-112414-054207

Copyright © 2016 by Annual Reviews.  
All rights reserved

## Keywords

diversification, fossil record, molecular phylogeny, phylogenetic comparative methods, trait evolution

## Abstract

With proliferation of molecular phylogenies and advances in statistical modeling, phylogeneticists can now address macroevolutionary questions that had traditionally been the purview of paleontology. Interest has focused on three areas at the intersection of phylogenetic and paleontological research: time-scaling phylogenies, understanding trait evolution, and modeling species diversification. Fossil calibrations have long been crucial for scaling phylogenies to absolute time, but recent advances allow more equal integration of extinct taxa. Simulation and empirical studies have shown that fossil data can markedly improve inferences about trait evolution, especially for models with heterogeneous temporal dynamics and in clades for which the living forms are unrepresentative remnants of their larger clade. Recent years have also seen a productive cross-disciplinary conversation about the nature and uncertainties of inferring diversification dynamics. Challenges remain, but the present time represents a flowering of interest in integrating these two views on the history of life.

## 1. INTRODUCTION

Traditionally, large-scale evolutionary dynamics and trends were investigated through study of the fossil record (Simpson 1944, Jepsen et al. 1949, Stanley 1979). The recent explosion of phylogenetic data for living species and the simultaneous development of phylogenetic comparative methods to infer macroevolutionary dynamics from them has opened up this field of research to evolutionary biologists. In some cases this has led to substantially divergent interpretations of clade dynamics, and the increasing sophistication of methods and models that can be applied to phylogenetic data has led some to question whether fossil data are even needed (e.g., Monroe & Bokma 2010, Venditti et al. 2011, Baker et al. 2015). Despite these conflicts, the paleontologist and biologist investigate the same tree of life, and our shared understanding of macroevolution should be enriched by integrating the views obtained by looking backward from the present with the help of trees and synoptically through time via the fossil record. Although a consensus on this perspective has begun to emerge (Fritz et al. 2013, Pennell & Harmon 2013, Benton 2015), practical challenges remain and it is unclear under which circumstances paleontological or phylogenetic approaches should be most informative.

In this review, we attempt to summarize the current state of fields we deem most relevant to the ongoing integration of phylogenetic and paleontological approaches to macroevolution. Rather than discuss the mechanics of methods that have been extensively reviewed elsewhere (e.g., O'Meara 2012, Morlon 2014), we focus on three main areas of overlapping concern to both phylogenetic and paleontological communities: integrating fossil information in time-scaled phylogenies, inferring the tempo and mode of morphological evolution, and understanding speciation and extinction dynamics through time. We end by discussing some areas that seem particularly promising for a more complete integration.

## 2. TREES WITH FOSSILS

The most obvious route for integrating phylogenetic and paleontological studies of macroevolution is to insert fossil taxa in phylogenetic trees of extant species. In principle, this is straightforward: If characters can be coded for both extant and fossil taxa, then analyses can be run on the combined data set. There are more challenges than it might initially seem, however. In this section, we review two main areas of ongoing development: How do we put fossil species into trees of living species, and how do we time-scale those trees so they can be used for downstream analysis?

### 2.1. Putting Fossils into Trees

Except in rare cases, morphology provides the sole means for assessing the relationships of fossil species to extant taxa. The default methodology for morphological phylogeneticists has traditionally been maximum parsimony. Under the parsimony criterion, we should prefer phylogenetic hypotheses that minimize the number of character state changes implied for a given character matrix. Felsenstein (1978) has pointed out that under certain conditions parsimony is statistically inconsistent, meaning that as the amount of data increases, an incorrect answer is returned with increasing probability. This artifact occurs most commonly when one or more branches exhibit more character state change than expected under a homogenous rate of change, and it results in taxa sitting at the ends of these long branches being attracted to one another due to increasingly probable convergent acquisition of similar character states. Fossil taxa, if available, can play a pivotal role in parsimony analyses by breaking up long branches and informing polarity of character state transitions (Gauthier et al. 1988, Huelsenbeck 1991). However, if rates of morphological

evolution are too rapid along internal branches of the underlying phylogeny, parsimony may still do a poor job of recovering the correct topology, regardless of whether fossils are sampled or not (Wright & Hillis 2014).

In molecular systematics, where long branches cannot often be broken, the realization that evolutionary rates are heterogeneous across sites, even within the same gene (e.g., transitions are more frequent than transversions; third codon positions evolve faster than first or second positions), led to the development of probabilistic phylogenetics based on more realistic models of sequence evolution. Lewis (2001) suggested a corresponding model for the evolution of morphological characters based on a Markov transition ( $Mk$ , where  $k$  refers to the number of character states) model with equal rates of change among states. The use of probabilistic models for phylogenetic inference from morphological data has proved philosophically contentious for some (Spencer & Willberg 2013) but yields statistically uncontroversial results. For example, simulations show that for data sets consisting of extant and fossil taxa, Bayesian phylogenetic inference using an  $Mk$  model yields lower percentage topological error than parsimony when evolutionary rates are fast or heterogeneous (Wright & Hillis 2014). Model-based inference using paleontological data sets can also be improved by partitioning characters according to anatomical region (Clarke & Middleton 2008) or by sampling rates from probability distributions (Wagner 2012, Harrison & Larsson 2015), just as is done with molecular data. Indeed, the use of model-based inference for morphological data may allow better integration of fossils in molecular phylogenies of extant taxa by providing an appropriate statistical framework for the analysis of concatenated data sets that evolved under heterogeneous processes (Bull et al. 1993). Recent work has suggested that processes related to fossil preservation may remove characters nonrandomly such that character modifications are more readily lost, causing fossil taxa to appear more primitive than they really are (see “stemward slippage” in Sansom & Wills 2013). Comparison of parsimony and model-based methods with artificially degraded data suggests that the latter may do a better job than parsimony of avoiding stemward slippage (Pattinson et al. 2014, Wright & Hillis 2014), though further work is required to explore the circumstances in which we should prefer either approach.

Although a lot of effort has gone into showing that missing data per se are not an overwhelming issue in model-based phylogenetics, the structure of missing data can be problematic for integrating extinct and extant organisms in the same phylogeny, regardless of inference method. Intuition might suggest that neontological morphologists have focused data collection efforts disproportionately on resolving higher-level relationships and, at least for mammals, this appears to be true (Guillerme & Cooper 2015). This poses a significant barrier for the integration of fossils in phylogenies of extant taxa, as a complete molecular but partial morphological data set for extant taxa provides little information to refine the placement of fossil taxa represented only by morphology. In turn, this results in unstable topological placements (Guillerme & Cooper 2016) and may skew downstream macroevolutionary analyses in the same way as integrating over random resolutions of branching patterns (Rabosky 2015). Topology searches could be fixed or heavily constrained to recover fossil taxa with inferred extant relatives that have no associated character data (Ronquist et al. 2012, Slater 2013), but this is similarly likely to lead to biased results. Perhaps the most significant barrier to complete phylogenies of living and extinct taxa remains the need for increased morphological data collection on the neontological side to guide the placement of fossils (Guillerme & Cooper 2015).

## 2.2. Time-Scaling Phylogenies with Fossil Taxa

The notion that time is an informative component of tree inference and thus is part of total evidence has a long and controversial history in the paleobiological literature. We will not review

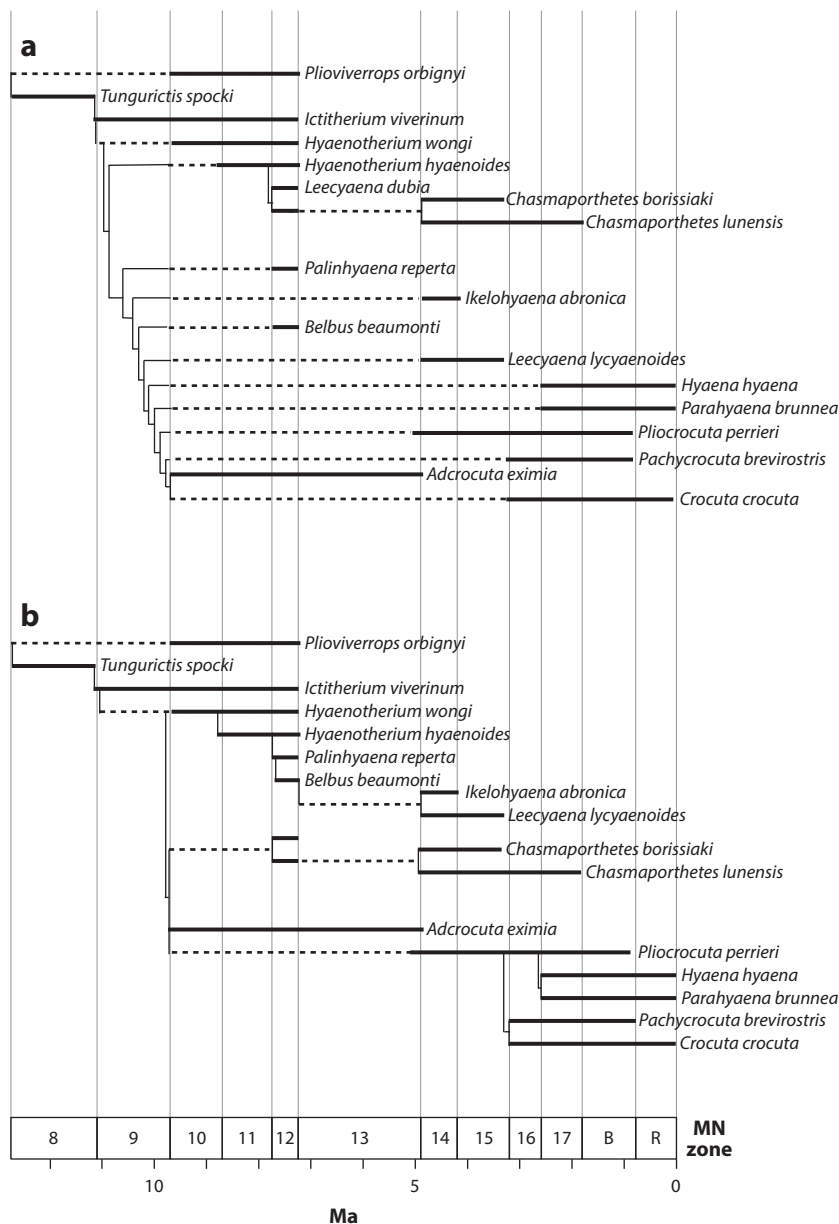
this vast literature here; a slightly outdated but nonetheless informative summary of viewpoints can be found in a *Nature* debate at <http://www.nature.com/nature/debates/fossil/>. Instead, we recognize two main approaches that can be taken to time-scaling phylogenies of living and extinct taxa: simultaneous analysis of characters with stratigraphic data, and post hoc scaling.

**2.2.1. Simultaneous inference.** The basic argument for combining character and stratigraphic data directly in phylogenetic analysis is that, all things being equal, the order in which taxa appear in the fossil record should tell us something about phylogeny. Given a set of competing phylogenetic hypotheses, we should therefore prefer those that provide a consensus best fit for both the character data and the distribution of fossil ages. The concept of stratigraphic debt (Huelsenbeck 1994, Clyde & Fisher 1997) arose from this thinking as a way of quantifying the amount of missing fossil record implied by a phylogenetic tree, typically in terms of discrete time bins such as stages or land mammal ages (Figure 1). Stratocladistics (Fisher 2008), though not strictly speaking a time-scaling approach, was the first to reconcile the observation that minimum-length trees inferred from parsimony might be rejected in favor of slightly longer trees after accounting for stratigraphic data (Clyde & Fisher 1997). Huelsenbeck & Rannala (1997) suggested a maximum likelihood approach to infer speciation and extinction times from phylogenies of fossil taxa using stratigraphic information only (i.e., without character data). Wagner (1998) extended this model with a maximum likelihood approach based on a model of fossil sampling intensity and extinction rate (Foote 1997a) combined with an approximation of character likelihood using parametric bootstrapping of parsimony scores. Applying this method to the phylogeny of Hyaenidae, he recovered a maximum likelihood estimate of phylogeny that required 11 additional character state changes compared to a parsimony analysis but implied 39 fewer units of stratigraphic debt (Figure 1).

Until recently, the primary role of fossil data in molecular phylogenetics has been to serve as minimum age calibrations for nodes uniting extant taxa. A full account of node dating methods is beyond the scope of this review (see Donoghue & Benton 2007, Ksepka et al. 2011), except to note that the idea here is that if a fossil can be confidently assigned to a particular clade, then its age provides a minimum bound for the age of that clade. However, because there is often considerable uncertainty as to which specific branch within the clade the fossil is associated with (it could have diverged early in the history of the clade or relatively late), a probability distribution is typically used to express a plausible range of older ages for the clade, and numerical methods are used to obtain a sample of ages based on this prior distribution and the molecular data.

More recently, a new literature has begun to emerge that, like earlier paleobiological approaches, is focused on integrating fossil data simultaneously in both phylogenetic inference and time calibration through a process referred to as tip dating (Pyron 2011, Ronquist et al. 2012). Tip dating typically assumes that morphological character data are available for fossil and extant taxa and that they evolve in a clocklike fashion, either strict or relaxed, in the same way as molecular data (although this is not strictly necessary; see Heath et al. 2014). Because fossil taxa have associated occurrence times, it subsequently becomes possible to calibrate this morphological clock and, in turn, to infer the divergence times for extinct and extant members of the clade. Although tip dating does not directly penalize tree topologies invoking higher stratigraphic debt in the way stratocladistic or stratolikehood approaches do, the same effect should, in principle, remain; tip-dated topologies that erroneously group morphologically similar but stratigraphically distant taxa should be less likely than alternative topologies that minimize stratigraphic debt, because they result in less likely distributions of branch-specific clock rates.

Tip dating has been applied to a diversity of clades, from mammals (Slater 2013) to plants (Grimm et al. 2014, Larson-Johnson 2016), using combined molecular and morphological data



**Figure 1**

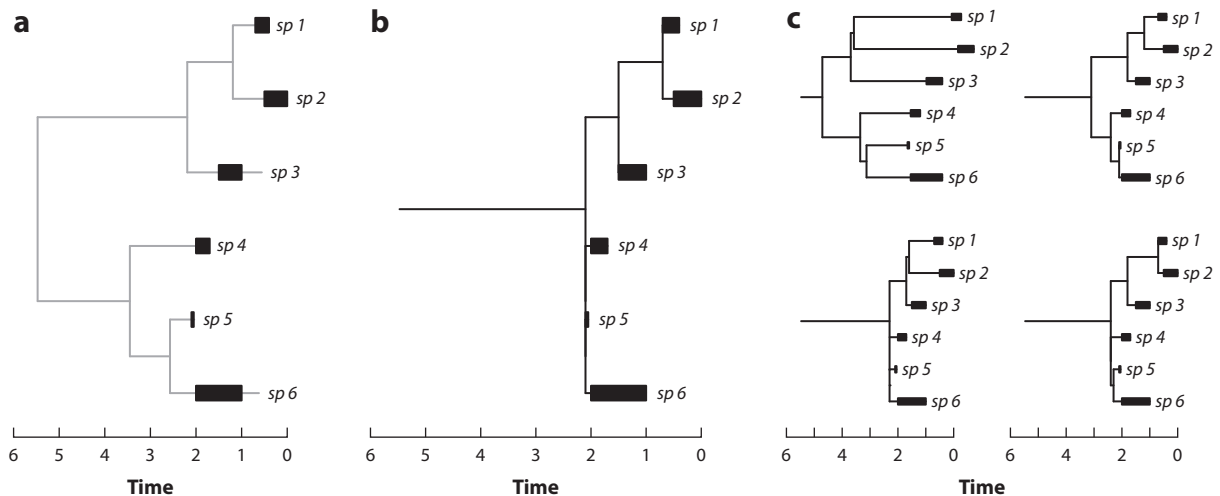
Accounting for stratigraphic debt can result in different optimal tree topologies. The hyaenid phylogeny depicted in panel *a* is the most parsimonious tree based on analysis of 19 morphological characters for 18 species and results in 47 units of stratigraphic debt after allowing for ancestor–descendant relationships. The tree in panel *b* is 10 steps longer (50 versus 60 morphologic steps) but has only 11 units of stratigraphic debt, making it the more parsimonious phylogenetic hypothesis overall. Stratigraphic debt is computed as the number of stratigraphic bins, shown here in units of European land mammal (MN) zones, for which a taxon is not observed even though the phylogeny implies its existence. All units are treated as equivalent, so a gap in the relatively long MN13 is considered equivalent to a gap in the much shorter MN14. Abbreviation: Ma, millions of years ago. Figure modified from Wagner (1998) with permission.

sets as well as morphology alone (Slater 2015). Although theoretically better than node calibration, early iterations of tip-dated trees tend to exhibit inflated divergence time estimates, sometimes dramatically and implausibly so, for nodes lying deeper in the tree (O'Reilly et al. 2015). Some authors have worked around this issue by placing strong age priors on nodes for which additional information is available (Slater 2013, 2015), though such an approach is far from ideal. The basis for this behavior remains to be definitively diagnosed but may lie in the priors used to weight the expected distributions of branching events in time-scaled trees of living and fossil taxa, sometimes referred to as tree priors in Bayesian phylogenetics. Early implementations of tip dating (e.g., Ronquist et al. 2012) used a uniform tree prior, considering any relative distribution of branching times as equally probable a priori. It is well appreciated that extinction tends to result in reconstructed trees (i.e., trees of extant taxa) with more tipward-biased divergence events (Pybus & Harvey 2000). In cases where fossil species are used to calibrate phylogenies of extant taxa, uniform tree priors provide equal weight to trees that we know a priori are less likely (though fossil representatives of extant species would not influence our prior belief in the same way).

Heath et al. (2014) suggested an alternative model that they called the fossilized birth-death (FBD) process that jointly considers speciation, extinction, and sampling rates. Consideration of sampling rates is essential when dealing with fossil data, as poor fossil recovery can mimic extinction (Foote 1997a,b), in turn biasing parameter estimates in birth-death models that do not consider them. Phylogenies inferred using the FBD process tend to result in more reasonable divergence time estimates than those using uniform tree priors, particularly for the problematic, early-branching nodes (Grimm et al. 2014, Heath et al. 2014, Arcila et al. 2015), and the model has recently been further extended to allow for inference of direct ancestor–descendant relationships between taxa from different time periods (Foote 1996b; Gavryushkina et al. 2014, 2015).

Tip dating has some benefits over stratocladistics and related approaches. Phylogenetic uncertainty can be easily accommodated in Bayesian tip dating using Markov chain Monte Carlo methods, whereas stratocladistics tend to depend on a step-by-step comparison of tree topologies to the stratigraphic record (but see the StrataPhy program; Marcot & Fox 2008). Although exceedingly promising, further considerations remain. For example, many paleontologists would argue that budding cladogenesis (i.e., the origin of a new taxon from a progenitor that itself continues after speciation) is a common pattern in the fossil record (Bapst 2014a), but it is difficult to detect using phylogenetic methods. Gavryushkina et al. (2015) suggest that this kind of pattern might be identifiable if multiple, temporally distinct populations of the progenitor were coded as separate terminals in analyses. Similarly, the assumption that fossil sampling rates (or, for that matter, speciation and extinction dynamics) are consistent across the phylogeny is undoubtedly naive. Extensions exist to the FBD model that allow for temporal shifts in these key parameters, but more realistic models can also consider geographic variation in sampling probabilities (e.g., Wagner & Marcot 2013). Method development is moving rapidly in this area and we anticipate much additional progress in the near future.

**2.2.2. Post hoc scaling.** Although combined analysis of topology and branch lengths remains ideal, for many data sets it is impractical. Paleobiologists often lack a character matrix, for example, where using supertrees assembled from multiple, smaller source trees (e.g., Betancur-R et al. 2015) or where insufficient character data are available to generate a character matrix but a tree can be drawn from a literal reading of a densely sampled fossil record (e.g., Aze et al. 2011). In these cases, post hoc scaling methods are useful. In their simplest form, post hoc methods scale branches so that the age of a particular node is determined by the first appearance date of its oldest descendant species (Norell 1992). This simple approach can nevertheless cause problems. For example, if the oldest fossil belonging to a clade is not positioned as the earliest diverging (i.e., it is nested several



**Figure 2**

Sampling can have significant impacts on time-scaling of paleontological trees. (a) Stratigraphic ranges (black bars) for terminal taxa in the underlying phylogeny (gray). Species (sp) 5 has a stratigraphic range that is older than sp 4, even though it is nested one node above it. As a result, (b) scaling the tree based on first appearance dates only (the “basic” method) results in a polytomy for the lower clade. Furthermore, because no stratigraphic data are available for internal edges, the age of the entire clade is reconstructed to be the same as the first appearance of sp 5. These issues can be somewhat overcome by (c) considering speciation, extinction, and sampling rates using the *cal3* method. Here, we show four stochastic realizations of branching times. Although some trees (e.g., the bottom row) look similar to those produced via the basic method, others (e.g., the top row) come closer to matching the true set of branching times. These outcomes emphasize the need to sample multiple realizations of branching times using the *cal3* approach. Analyses were done using functions in the *paleotree* package (Bapst 2012); time was simulated in arbitrary units.

nodes up; **Figure 2a**), then one or more zero-length branches result (**Figure 2b**). A number of different solutions have been proposed to deal with such situations, such as arbitrarily adding some small amount of time to all zero-length branches or redistributing time over zero-length branches (e.g., Brusatte et al. 2008), perhaps with reference to degree of character change (Ruta et al. 2006).

Bapst (2013) has proposed a different approach, which he calls *cal3*, that attempts to stochastically sample branching times and branch lengths based on a model that requires estimates of speciation, extinction, and sampling rates. Bapst’s approach works in a stepwise fashion. The tree is initially scaled using the coarse approach described above based on oldest first appearances. The algorithm then visits each node, from the root up, and a node age is sampled. Descendant branches are then rescaled to extend to their last appearance times in the case of tips or to a new, stochastically sampled age in the case of internal nodes. Bapst’s (2013) use of a rate-calibrated model to estimate node ages is key here as it provides an estimate of how much missing evolutionary time we might reasonably expect between the divergence of a taxon from its sister lineage and its first appearance in the fossil record (**Figure 2c**). The *cal3* approach is thus similar in spirit to that of Wagner (1998) in that it does not necessarily attempt to minimize the difference between divergence and first appearance.

As phylogenetic comparative methods (PCMs) are increasingly used to explore macroevolutionary dynamics in large (>100 taxa) trees consisting exclusively of extinct taxa, post hoc scaling methods will likely become more prominent. The *cal3* approach appears to result in time-scaled trees that yield better results when fitting macroevolutionary models than simpler



scaling methods, but performance is not perfect (Bapst 2014a). This is a field that would benefit from increased attention in the near future.

### 3. TRAIT EVOLUTION

The paleontologist and PCM practitioner both seek to understand the patterns, processes, and drivers of trait evolution. Moreover, relative to the population geneticist, their data often have similar limitations: coarse temporal resolution and scant information directly relevant to microevolutionary processes. They, like all scientists, also deal with incomplete information. The fossil record is famously incomplete for a variety of geological, biological, and sampling reasons, all of which can interfere with evolutionary inference (Kidwell & Holland 2002). The information available to the phylogeneticist is not filtered through the geologic record, but it too is incomplete; modern diversity is usually incompletely sampled, often nonrandomly so (Garamszegi & Moller 2011), and of course lacks direct evidence from lineages that did not survive to the present day.

Phenotypic traits are almost endless in their variety, but an important practical distinction is to be made between those coded with qualitatively discrete states such as the presence/absence of features versus those such as body size that vary on a continuous scale. These data types require different models and analytical procedures (O'Meara 2012). Analysis of discrete states relies mostly on the *Mk* model described in Section 2.1. There is a broader variety of models for continuous traits, though most are variants of Brownian motion (BM) or Ornstein–Uhlenbeck (OU) models. BM is an unbounded diffusion process with a drift parameter that specifies directional change and a variance parameter that specifies the rate of diffusion. Usually, PCMs focus on the special case of nondirectional evolution, in which the drift parameter is zero and the diffusion variance controls the pace at which the trait evolves. Common elaborations of BM involve varying this rate over time or across branches of a tree. OU models have BM-like dynamics in combination with an attracting optimum that pulls traits toward it like a rubber band. For the purposes of this review, we intermingle examples regardless of data type because the benefits and challenges of integrating paleontological and phylogenetic approaches are similar.

#### 3.1. Estimating Ancestral States

Much early effort in PCMs focused on estimating the phenotypes of unsampled ancestors to better understand evolutionary transitions. The uncertainty associated with these estimates can be quite broad, even under uniform evolutionary models such as BM (with no directional drift) and the equal-rates *Mk* model (Schluter et al. 1997). Incorporating fossils in such analyses can greatly improve ancestral state estimates because uncertainty increases as one tries to reach farther back in time from sampled taxa. Extinct taxa are closer in time to the ancestral nodes that are to be estimated, and these shorter path lengths can greatly reduce estimation uncertainty (Polly 2001, Finarelli & Flynn 2006).

The payoff for incorporating fossils is even greater when temporal heterogeneity or selective extinction makes the living members of a clade unrepresentative of ancestral forms. For example, Webster & Purvis (2002) and Finarelli & Flynn (2006) demonstrate the inaccuracy of ancestral body size estimates from living members in planktonic foraminifera and caniform mammals, respectively, because of a systematic trend of increasing body size in each group. In the latter example, analysis of living caniforms alone suggests an ancestral body size of 23 kg, whereas including fossil taxa yields estimates <2 kg, in line with the body size of the earliest known fossils in the group (see also Albert et al. 2009, Pyron & Burbrink 2012, Slater et al. 2012a). Another recent example comes from Betancur-R et al. (2015), who investigated the ancestral habitat (marine



versus freshwater) for ray-finned fishes. Extant-only analyses favored freshwater ancestry, but the addition of many fossil taxa flipped the result, strongly supporting a primitive marine habitat for the clade. In this case, the modern fauna mislead because several basal groups of fishes with marine ancestry experienced selective extinction in marine environments and are now restricted to freshwater.

### 3.2. Fitting Evolutionary Models: The Benefits of Fossil Taxa

Although ancestral state estimates can be informative for evaluating specific evolutionary scenarios (e.g., In what habitat did ray-finned fishes originate?), the PCM field has shifted emphasis from ancestral states to explicit models of trait change. This shift emphasizes evaluating evolutionary hypotheses as specified models and then leveraging the power of modern statistical tools such as likelihood, information criteria, and Bayesian methods to fit and compare competing explanations for observed trait distributions.

The mechanics of a phylogenetically informed analysis of trait evolution are little different whether or not fossil taxa are incorporated. One needs a tree or set of trees, usually with branch lengths scaled in units of time by one of the methods discussed in Section 2. Terminal taxa are characterized by their values for one or more traits and are usually evaluated by reference to a suitable model of evolutionary change. One complication that arises is that trait evolution methods are sometimes implemented assuming trees are ultrametric, an assumption that will be violated when fossils are integrated in analysis of extant taxa (Slater 2014). Assuming an ultrametric tree can simplify programming, but it is not always obvious to the end user when this assumption is being made (Slater 2014). Whenever practical, this practice should be avoided.

As with estimation of ancestral states, benefits of adding fossil taxa to fit models accrue disproportionately when evolutionary dynamics are not uniform through time. Slater et al. (2012a) simulated trait evolution under five models of continuous character change: BM, BM with a trend, OU, decelerating evolution (also called the early-burst model), and accelerating evolution. They found that model support for the correct model was not affected by the addition of fossil taxa under BM and OU models, but performance was greatly improved by the addition of fossil taxa where the expected trait value or evolutionary rate varied as a function of time. The increased propensity to support the correct model was not simply the result of adding more taxa to an analysis; when the authors swapped extant for extinct taxa one-for-one, the performance benefit from fossil taxa remained. Thus, on a per taxon basis, fossil taxa can be more informative about evolutionary mode than extant ones (Slater et al. 2012a).

The trend model for continuous data, in particular, relies on fossil information in order to be estimated. The signal of directionality is a systematic relationship between the trait value and elapsed time (Hunt 2006). Terminal taxa in ultrametric trees are equidistant from the root node and thus contain no information about the strength of a trend, at least for the uniform BM model with a trend (see Section 3.5). When extinct taxa are present in a tree, the resulting range of taxon ages permits trends to be estimated and, indeed, the tendency to support a trend model when it is true increases sharply when even a few extinct taxa are sprinkled into an extant-only tree (Slater et al. 2012a; for an example of directionality in discrete character traits with and without fossils, see Klopstein et al. 2015).

A good illustration of this effect comes from attempts to test for a trend of increasing body size (Cope's rule) in mammals. Paleontologists have long suggested that mammal lineages tend to get larger-bodied over time, which was supported by Alroy's (1998a) study of over 1,500 species from the well-sampled North American fossil record. Yet, Monroe & Bokma (2010) found little support for Cope's rule in a phylogenetic analysis of >3,000 extant mammal species. A trend model was

estimable in that study because their model allowed for evolution to be concentrated at speciation events. This effectively makes the tree nonultrametric because the (inferred) number of speciation events from the root varies across terminal taxa. Bokma et al. (2015) updated this approach by incorporating body size estimates for over 500 extinct mammal species that were used to inform priors on internal nodes of the tree of extant species. The new, integrated fossil-modern analysis found overwhelming support for Cope's rule. Other than the fossil information, the two studies used very similar data and methods, indicating that nearly all the signals of directionality reside in the data from fossil mammals. However, two recent studies of extant taxa have been able to produce estimates of ancestral body mass in mammals that are concordant with fossil estimates, even in the face of a body size trend (Baker et al. 2015, Puttick & Thomas 2015). Both studies allowed for variation in evolutionary rate across branches of the tree, which evidently allowed this dynamic to be inferred without fossils (though when added, fossils greatly narrowed confidence intervals around ancestral size estimates; Puttick & Thomas 2015).

In addition to indicating a trend of increasing body size, Alroy's (1998a) compilation supported the traditional view that mammal body size disparity jumped markedly after the Cretaceous/Paleogene mass extinction, presumably reflecting a pulse of origination into the niche space left vacant by extinction of nonavian dinosaurs. This effect is visually obvious: One could pick out the extinction boundary on Alroy's figure even if the time axis were unlabeled. Nevertheless, analysis of a large phylogeny of extant mammals, though finding ample evidence for variability in rate across branches, found no signal of an increase in the rate of body size evolution associated with the mass extinction (Venditti et al. 2011). In contrast, an analysis that integrated body size in extant and extinct mammals (Slater 2013, 2014) found strong evidence for a shift in evolutionary mode from constrained (OU) divergence in the Cretaceous to unbounded (BM) dynamics after the mass extinction. This is likely a case in which the signal of a strong evolutionary pattern in the distant past is eroded by subsequent extinction and thus not easily detectable from the living fauna.

### 3.3. Does Speciation Lead to Pulses of Phenotypic Change?

In a now famous paper, Eldredge & Gould (1972) suggested that most species do not change much over their lifetimes, a phenomenon they called stasis, and that most phenotypic changes are instead associated with speciation events. This punctuated equilibrium model sparked much controversy among paleontologists (Gingerich 1985) and evolutionary biologists (Charlesworth et al. 1982, Pennell et al. 2014), motivating scientists of disparate fields to gather evidence to evaluate its claims.

Although hundreds of paleontological case studies were compiled to assess the stasis part of punctuated equilibrium (Hunt et al. 2015), tests of speciation (or cladogenetic) change were more difficult to come by because speciation is so rarely captured in fossil sequences. Alan Cheetham and colleagues (Cheetham 1986, Jackson & Cheetham 1990) recognized that adequately testing punctuated equilibrium requires stitching together ancestor–descendant sequences into a phylogenetic framework so that within-lineage (anagenetic) and between-lineage (cladogenetic) components of change could be estimated separately. When the authors applied this approach to the rich Neogene fossil record of the bryozoan *Metrarabdotos*, they found that evolutionary changes were predominantly associated with speciation and that species overwhelmingly experienced stasis after their first appearance in the fossil record.

Bokma (2002) developed a model of mixed cladogenetic and anagenetic evolution that allowed for punctuated equilibrium–like dynamics to be fit to phylogenetic trees of extant species (see also Bokma 2008, Bokma et al. 2015). This approach requires inferring “missing” speciation events, that is, splits that are not visible in the extant tree, because extinction has pruned the lineages

needed to detect them. Moreover, it can be difficult to differentiate anagenetic and cladogenetic components with modern data alone; confidence intervals of parameter estimates often include solutions that are consistent with both strongly anagenetic and strongly cladogenetic change. Fossil data have the potential to help here because ancestor–descendant sequences of fossil populations provide information about anagenetic change that is uncontaminated by changes associated with cladogenesis. However, when the Bokma model was estimated jointly over a phylogenetic tree and fossil ancestor–descendant sequences in an ostracode clade (Hunt 2013), considerable uncertainty remained about the relative importance of anagenetic and cladogenetic components of trait evolution, even with the fossil constraints on anagenesis.

Moving from phenotypic to molecular divergence, Ezard et al. (2013) found that the number of nodes between root and tip on a phylogeny of living and fossil macroperforate foraminifera better predicted the species-specific rate of molecular evolution than the same count on the extant-only phylogeny or any ecological or environmental factors. This result provides a stronger and more convincing test of the hypothesis that speciation is associated with a pulse of molecular evolution (e.g., Pagel et al. 2006) than could be obtained from a phylogeny of extant taxa alone.

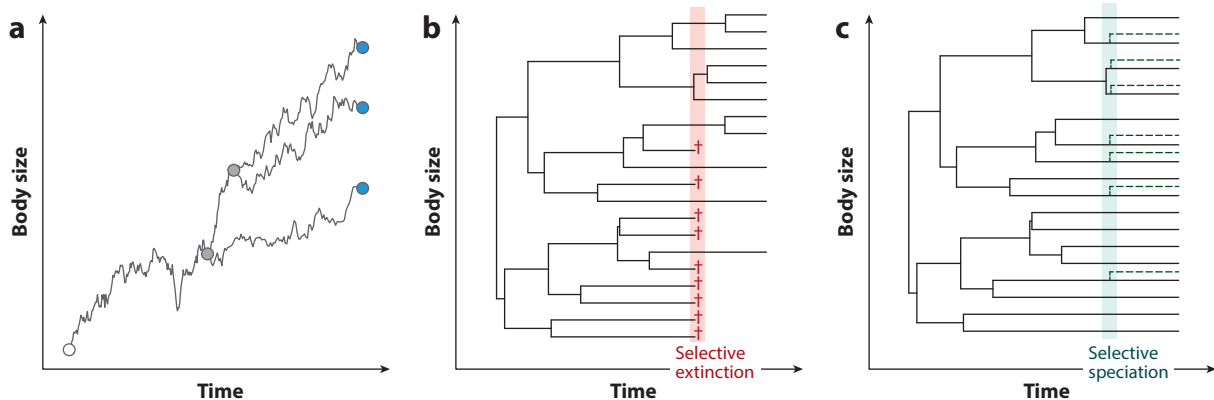
### 3.4. Disparity and Early Bursts of Morphological Change

Studies of disparity, or morphological diversity, arose in paleontology as an attempt to test the classic idea that as major groups radiate, early lineages spread rapidly to colonize most of the available morphological and ecological space, with subsequent lineages filling in the interstices (Simpson 1944, Valentine 1980). Foote (1990, 1991) introduced the approach of quantifying morphological disparity through time, usually computed as a multivariate variance (or sometimes as the range or volume of morphospace occupation, or as the number of unique character combinations). Many empirical studies of different clades and traits followed, yielding a variety of disparity trajectories (Foote 1997b), among which an early peak in disparity was quite common (Hughes et al. 2013).

Early peaks in disparity are consistent with rates of trait evolution that are initially high but then decrease, but the same patterns could also be caused by selective origination and extinction or even evolution proceeding at a constant rate but within a bounded morphospace (Foote 1996a). Phylogenetic information is required to distinguish among these rather different macroevolutionary processes. Initially rapid exploration of morphospace can be modeled by extending BM to allow its rate parameter to decrease over time (Blomberg et al. 2003). When applied to a compilation of trees of extant taxa, however, this early-burst model was seldom favored as an explanation for body size or shape evolution (Harmon et al. 2010).

What explains the apparent discrepancy between extinct and extant support for the early-burst model? Simulations show that the power to detect an early-burst model can be quite limited but improve when fossils are added (Slater & Pennell 2014). Thus, one might expect that this pattern could go undetected in phylogenies of extant species, especially if they were of modest size. Another factor might be the phylogenetic or taxonomic scale of the studies: Most paleontological disparity work characterizes the diversification of broad higher taxa, but the compilation by Harmon et al. (2010) captured mostly relatively small clades. Studies with broader taxonomic scope have favored early-burst and similar models (e.g., Cooper & Purvis 2010), especially in paleontological trees (Wagner 1995, Close et al. 2015, Hopkins & Smith 2015), although whether the burst coincides with the initiation of the clade can depend on which node is considered the starting point.

The jury remains out as to whether early bursts in evolutionary rate are rare or common, though there is certainly ample evidence of variation in rate across clades (Eastman et al. 2011, Rabosky et al. 2013). Disparity is still usefully employed in fossil comparative studies, but it is best to couple it with analyses that are model-based and phylogenetically grounded. Studies of



**Figure 3**

Three different mechanisms that can cause trends are illustrated with examples of Cope's rule, which postulates a pattern of increasing body size over time. In all panels, body size is indicated by the vertical axis and time proceeds from left to right. (a) Trend in a small clade caused by directional evolutionary changes occurring in the clade's constituent species. Lines show the trajectory of trait increase, here modeled as Brownian motion with a trend. Shown are a root node (*white*), internal nodes (*gray*), and terminal taxa (*blue*) at the end of the short simulation. (b) Clade that evolves without extinction until an interval (shaded *red*) during which a size-selective extinction occurs; daggers mark species going extinct. As extinction has preferentially removed small-bodied species, it increases the mean body size of the clade. (c) Similar scenario to that of panel *b*, except that size-selective speciation rather than extinction occurs during the same interval (shaded *green*). The differential proliferation of large-bodied species increases their proportion in the clade, increasing the mean body size of the clade's members. In both panels *b* and *c*, changes in body size are shown to occur only at speciation, but the mechanism operates regardless of how changes are apportioned between cladogenesis and anagenesis. Only a single interval of selective speciation or extinction is shown in panels *b* and *c*; to generate a sustained trend rather than a one-time shift, such selectivity must be repeated or sustained over time.

higher-level clades combined with a focus on where rather than if bursts of elevated evolution occur would be a fruitful avenue for future exploration in phylogenetic studies.

### 3.5. Paleontological Time-Series of Traits

The fundamental advantage of fossil information is its temporal scope, so it is perhaps not surprising that paleontologists regularly construct time-series to understand patterns of trait evolution. When points in a time-series correspond to populations interpreted to have direct ancestor-descendant relationships, then a time-series is the correct way to represent phylogenetic dependence among them. For these anagenetic lineages, many of the same models of trait evolution in PCMs can be fit with similar, likelihood-based methods (Hunt 2006, Hunt et al. 2015) and the results directly compared to those from phylogenetic data.

Of more direct relevance to this review are examples in which time-series are used to summarize trait evolution, especially trends, over entire clades rather than within species-level lineages (Stanley 1973, McShea 1994, Novack-Gottshall & Lanier 2008, Finarelli & Goswami 2013, Heim et al. 2015). Although time-series methods can account for temporal autocorrelation among time intervals, they do not properly account for the phylogenetic dependence among species. Moreover, time-series can reflect both trait microevolution within lineages and the sorting of species from differential speciation or extinction with respect to the focal trait (i.e., species selection). For example, a time-series of body size can show an increasing trend (Cope's rule) from any combination of three mechanisms (Figure 3): (a) widespread directional evolution within the lineages that constitute the focal clade, (b) preferential speciation among large-bodied species,

or (c) preferential extinction in small-bodied species. It can be difficult to distinguish among these mechanisms from a time-series or time-scaled phylogeny, but additional paleontological observations can help. Trends caused by widespread anagenesis can be detected by documenting directionality within the individual fossil species that constitute the clade (e.g., Hunt & Roy 2006), and species selection is implicated when originations and extinctions are disproportionately concentrated in taxa with certain trait values (e.g., Hansen 1982, Finnegan et al. 2012).

## 4. DIVERSIFICATION

Life has diversified from a single lineage several billion years ago to the millions of species alive today. The trajectory of this increase, at least for readily fossilized forms, has long been a central concern of paleontology (Simpson 1944, Sepkoski et al. 1981). How to best estimate fossil diversity in the face of an incomplete fossil record has been much researched (reviewed in Alroy 2010), and we will not focus on it here. Instead, we consider the origination and extinction rates that underpin diversity trajectories, as this is an area of active interest from both phylogenetic and paleontological perspectives. In studies of trait evolution, extinct and extant taxa are used in basically the same way: as terminal taxa on a time-scaled tree. This is much less true in diversification analyses; models can be shared across phylogenetic and paleontological approaches, but the data and mechanics of analysis are dissimilar enough to warrant a brief review.

### 4.1. Estimating Rates from Fossils

The fundamental observation used to estimate origination and extinction rates from the fossil record is the occurrence of a taxon in a fossil deposit. This taxon may be resolved to species level, but genera are more commonly used in large-scale fossil studies (Hendricks et al. 2014; see Section 5.1.1). The ages of fossil-bearing deposits are generally known only to the resolution of the discrete time bins, usually geological stages or biostratigraphic zones, to which they can be correlated. These bins have durations that typically range from several hundred thousand to several million years, placing practical limits on the temporal resolution of fossil data (Bapst 2014b). Newer, high-resolution methods of correlation have the potential to generate much higher value paleontological data sets than previously possible (e.g., Crampton et al. 2016).

Paleontological studies of diversification were initially based on counts of the first and last time bins in which a taxon had been observed. Methods were developed to estimate origination, extinction, and sampling rates from these data constituting only the range end points of taxa (Foote 2000, 2003). The increasing popularity of occurrence-based databases such as the Paleobiology Database (<http://paleobiodb.org>) has shifted interest to methods that exploit occurrences of taxa within their stratigraphic ranges, not just the range end points. These internal occurrences are important because they provide direct information about sampling intensities. When a taxon is not observed during an interval between its first and last occurrences, this absence must be a failure of sampling (barring errors in identifying and dating fossils). Sampling rates can be estimated by aggregating this information over many taxa. This is crucial because sampling heterogeneity strongly affects the observed occurrences of taxa and the rates computed from them (e.g., Foote 2000, Alroy 2014). For example, a time bin with poor sampling can cause a pulse of taxon last appearances in the immediately preceding interval, causing a spurious peak in extinction rate.

Methods that make use of occurrence data to constrain sampling probabilities include Alroy's (2014) moving window approach, capture-mark-recapture (CMR) methods adapted from wildlife biology (Liow & Nichols 2010), and a recently developed Bayesian approach called PyRate (Silvestro et al. 2014). All these approaches estimate origination and extinction while accounting

for fossil record incompleteness via jointly estimated sampling rates. Moreover, CMR and PyRate offer rich statistical frameworks for modeling variation in rates across taxa and time. These frameworks allow extrinsic variables to influence rates, so that one can test if speciation or extinction rates depend on climate, morphology, or other variables. These capabilities have only recently started to be exploited in empirical studies (e.g., Liow et al. 2008, Liow & Finarelli 2014, Pires et al. 2015, Silvestro et al. 2015).

## 4.2. Estimating Rates from Phylogenies

Phylogenies capture a clade's net diversification from its most recent common ancestor to its present day species richness. Less obviously, the branching times of a time-scaled tree retain separate information about speciation and extinction rates, even if the tree includes no fossil taxa (Nee et al. 1994). Initially, applications to molecular phylogenies assumed a constant-rate birth-death process that applied uniformly to all branches in a tree. This model was seen as unrealistic, and in practice it often produced near-zero extinction rates (Morlon et al. 2011), which is at odds with a fossil record littered with extinct species. The past 10 years have seen an explosion of work to generalize this approach to account for realistic sources of heterogeneity. As these developments have been recently reviewed (Pyron & Burbrink 2013, Morlon 2014), we summarize them only briefly before moving on to the interplay between these phylogenetic methods and those derived from the fossil record.

Initial results indicated that a signal of declining diversity was common in empirical phylogenies. This pattern was linked to speciation slowdowns as radiations increasingly filled niche space (Phillimore & Price 2008, Rabosky 2013; but see also Moen & Morlon 2014). A metric that captured this slowdown, the  $\gamma$  statistic, was developed (Pybus & Harvey 2000) and widely applied. It was followed by approaches that explicitly modeled temporal variation in evolutionary rates as discrete shifts to new regimes (Rabosky 2006, Stadler 2011), smooth functions of time (Rabosky & Lovette 2008b, Morlon et al. 2011, Rabosky 2014), or even functions of time-varying covariates such as climate proxies (Condamine et al. 2013). Some methods also allow for variation in speciation and extinction rates among clades (Alfaro et al. 2009, Morlon et al. 2011, Rabosky 2014), which must be increasingly important as workers analyze larger and larger clades. Whereas temporally declining diversification was initially treated as an indication of diversity-dependent rates, some other methods explicitly modeled diversity dependence (Rabosky & Lovette 2008a, Etienne et al. 2012).

Another class of model postulates that speciation and extinction rates are influenced by phenotypic traits. The original version, called binary-state speciation and extinction (BiSSE), applied to phenotypic traits with two discrete states (Maddison et al. 2007), each of which could be associated with different speciation and/or extinction rates. This approach was soon extended to apply to multistate and quantitative traits (FitzJohn 2012) and situations in which trait change is concentrated at speciation events (Goldberg & Igic 2012, Magnuson-Ford & Otto 2012).

## 4.3. Integrated Diversification Analyses

Even if one accepts that much is to be gained by combining fossil and phylogenetic data, it can be a challenge to do so because few taxa combine a rich fossil record with a robust molecular phylogeny and many current methods do not allow for analysis of phylogenies containing fossil taxa. Nevertheless, three recent studies compared diversification analyses across paleontological and phylogenetic approaches to the same clade. Simpson et al. (2011) examined coral diversification over the past 200 million years and showed that paleontological and phylogenetically derived curves



shared many features, including diversification peaks in the mid-Jurassic, Late Cretaceous, and Neogene and generally low rates otherwise. Similarly, Cantalapiedra et al. (2014) subjected the fossil record and phylogeny of ruminants to an extensive comparative analysis. Paleontological and phylogenetic rate estimates were significantly correlated, although their correspondence was strongest during relatively recent time intervals and when only the crown members of living groups were considered. This study also revealed substantial differences between paleontological methods (Alroy's rate metrics versus PyRate modeling) and between sets of time-scaled phylogenies drawn from different studies. Finally, paleontological and phylogenetic perspectives on Fagales (oaks, beeches, and their allies) were less easily reconciled (Xing et al. 2014). Lineage accumulation in the molecular phylogeny indicated an acceleration of diversification toward the present that was not supported by fossil analyses. This discrepancy may have been caused by lower completeness of the fossil record in regions with the most actively radiating clades.

In addition to individual studies that compare phylogenetic and paleontological approaches, there are several data sets that have been subjected to iterative analyses from different perspectives. Quental & Marshall (2010) documented a complex history of origination and extinction leading to waxing and waning diversity in fossil whales, a pattern that they argued would not be recoverable with phylogenetic methods available at the time. Morlon et al. (2011) reconsidered this example and found that the decline in whale diversity could be recovered without fossils by considering phylogenetic models with rates that varied over time and across branches (see also Etienne et al. 2012, Rabosky 2014). Analysis of the extremely well-sampled fossil record of planktonic foraminifera found evidence for the influence of climate, species ecology, and diversity dependence on diversification (Ezard et al. 2011), and a subsequent phylogenetic analysis corroborated the diversity-dependent component of these dynamics (Etienne et al. 2012).

For clades with less rich fossil records, even the integration of relatively little fossil data can improve the inference of diversification (e.g., Pyron & Burbrink 2012). One common approach uses phylogenetic information to adjust the observed stratigraphic ranges of taxa (Norell 1992). If sister taxa can be assumed to be of the same age, then observed differences in their first appearances can be attributed to fossil record incompleteness, and thus the first appearance of the younger taxon can be pulled backward in time to coincide with its older sister taxon. These adjusted ranges can then be used to reconstruct diversity histories (for some limitations to this approach, see Lane et al. 2005, Bapst 2014a). In a series of papers, Jablonski, Roy, and other collaborators combined information from the rich fossil record of bivalves with modern occurrence data and partial phylogenetic data to address a variety of macroevolutionary issues, such as the effect of climate on latitudinal diversity gradients (Huang et al. 2014), degree to which extinction rates are phylogenetically conserved and therefore similar among close relatives (Roy et al. 2009), and loss of evolutionary history by extinction in phylogenetic and fossil approaches (Huang et al. 2015).

## 5. THE CHALLENGE AND PROMISE OF INTEGRATION

It is an auspicious time to be a scientist interested in macroevolution; not since the punctuated equilibrium debates of the 1970s and 1980s has there been such sustained and productive back-and-forth between biologists and paleontologists. Biologists who work on living faunas are engaging in a serious way with insights from the fossil record (e.g., Pennell & Harmon 2013, Rabosky 2013), and phylogenetically minded paleontologists are rapidly adopting newly developed, tree-based methods. Nevertheless, challenges yet remain to integrate paleontological and phylogenetic approaches to macroevolution and there are promising veins that have yet to be fully mined.



## 5.1. Challenges to Resolve

A variety of impediments can get in the way of a complete integration of paleontological and phylogenetic approaches to macroevolution. In this section, we focus on issues of data comparability between fossil and extant-only studies as well as concerns about method robustness to data heterogeneity and violations in model assumptions.

**5.1.1. Genus- versus species-level fossil data.** Fossil analyses of diversification often use genera rather than species as the analytical units. This choice reflects practical advantages of working at the genus level. Genera are more inclusive than species, and thus their fossil record is more complete. In addition, some fossil groups can be difficult to resolve to species, especially when preservation varies, and there is a perception that genera are more reliably identified than species (Hendricks et al. 2014). Support for genera as useful proxies for species comes from simulations (Sepkoski & Kendrick 1993) and empirical studies that have evaluated patterns at both levels (e.g., Simpson et al. 2011, Liow & Finarelli 2014; but see Xing et al. 2014). However, even if genera are good proxies for species, the magnitudes of origination and extinction rates will not be comparable across these taxonomic levels. Future analytical developments may allow for the translation of processes at the two levels (see Foote 2011), and the development of more fossil data sets at the species level would be welcome even though such data sets have their own difficulties (Section 5.1.3).

**5.1.2. How fragile are phylogenetic comparative methods when model assumptions are violated?** Some workers have expressed concern that PCMs may not be robust to model violations, especially when extinction is non-negligible. Quental & Marshall (2009) demonstrated that the  $\gamma$  statistic correctly identified diversification slowdowns only under certain parameter combinations for speciation and extinction (see also Liow et al. 2010, Quental & Marshall 2010). Rabosky (2010) similarly showed that unmodeled heterogeneity in diversification rate could compromise estimates of extinction (but see Beaulieu & O'Meara 2015). Problems induced by model violations can be mitigated by developing more complex models that can handle violations of the constant-rate, birth-death model (Morlon 2014). There are, however, many ways in which reality can deviate from simple diversification models, and it is a challenge to assess all of them, especially in combination. Other concerns about PCM robustness have recently surfaced for methods with discrete categorical variables, including those in the BiSSE family discussed above. These methods can perform quite poorly in some situations, at least partly because of model violations (Maddison & FitzJohn 2015, Rabosky & Goldberg 2015). Furthermore, some authors have expressed concern regarding the sensitivity of continuous trait models to error in topology, branch lengths, and trait measurements, all of which can lead to reduced phylogenetic signal in traits and strong but incorrect support for OU models (Cooper et al. 2016). This is of particular concern in studies based on paleontological phylogenies, where coarsely resolved or inappropriately time-scaled trees show exactly this pattern (Bapst 2014a, Soul & Friedman 2015). Iterative rounds of method development and testing via simulation are likely needed here.

**5.1.3. How robust are paleontological approaches?** Many similar concerns apply to paleontological approaches, though they may be robust to some violations that cause problems for extant-only analysis (Liow et al. 2010). Variation over time is almost always explicitly addressed in paleontological studies, but temporal variation in fossil recovery can still be a serious concern (e.g., Foote 2000, Alroy 2014). Variation across taxa can be handled with existing methods but is not always evaluated. Current methods that are appropriate for occurrence data (Alroy's metrics,

CMR, and PyRate) can, in principle, account for these kinds of heterogeneities. But they are just starting to be widely used and would benefit from simulations testing their performance under different kinds of process variation. Parallel concerns can be raised about methods for post hoc time-scaling of paleontological trees, the performance of which under known conditions has only recently been explored (Bapst 2014a).

In addition to issues with model assumptions, large-scale paleontological data sets raise important concerns about data quality and consistency. Such analyses often rely on the Paleobiology Database or similar community resources. Although undoubtedly a boon for the discipline, such repositories host data that can be very heterogeneous. In addition to the usual concerns about temporal and spatial coverage, the quality of identifications can be very mixed. Ideally, taxonomic experts would vet and harmonize identifications for a particular study, although this is a daunting task at broad temporal, spatial, and taxonomic scales. Also welcome would be simulations of data contaminated by misidentifications, in order to assess the seriousness of this problem for particular questions (Plotnick & Wagner 2006).

## 5.2. Further Promise of Integration

Much progress has been made in establishing a framework that allows for fossil and modern data to contribute seamlessly to macroevolutionary inference. As discussed in this section, even greater gains may be possible through better feedback about data and models between the two disciplines and from additional development of methods that can handle data sets that are less than ideal.

**5.2.1. Cross-pollinating phylogenetic and paleontological models.** Although diversification models are shared across phylogenetic and paleontological approaches, the two fields often emphasize different kinds of variation. This can be important because the empirical findings of one discipline inform the modeling choices of the other, and vice versa. For example, paleontologists sometimes find evidence that origination and extinction rates are correlated across groups and time intervals (e.g., Stanley 1990, Xing et al. 2014). There are exceptions, and the correlation is not perfect—otherwise, there would be no differential diversification—but this correlation may be a general feature that is not presently integrated in phylogenetic modeling approaches.

Another example concerns how temporal variation is modeled. Paleontologists often focus on interval-to-interval variation in rates, reflecting the field's interest in what happens during mass extinctions and other critical intervals of earth history. PCMs, in contrast, focus more on smooth curves of rate variation. Each approach has its merits, but there may be benefits to exporting each view to the opposite field. Paleontological studies often reveal time-series of rates that look like fluctuations with occasional large excursions into both origination and extinction (e.g., Alroy 1998b, Foote 2003, Peters & Ausich 2008), a dynamic that does not receive much attention in the phylogenetic literature outside of limited modeling of mass extinction (Stadler 2011, Condamine et al. 2013). On the other hand, paleontology's focus on interval-to-interval changes results in highly parameterized models, often with separate origination, extinction, and preservation rates for each interval, although this is less true for the PyRate approach. For some data sets and questions, using simpler models with fewer parameters might improve parameter estimates and clarify interpretation.

**5.2.2. Environmental variables as inputs to models.** It is common to suppose that macroevolutionary processes may depend on environmental factors such as climate. Recently, Condamine et al. (2013) developed a model in which speciation and extinction rates vary as a function of a

proxy for environmental conditions, which itself varies over time. They used this approach to fit a model in which cetacean speciation rates were influenced by temperature as measured by a proxy for global climate. This kind of approach (which can also be implemented with fossil-only data using CMR and PyRate) allows for integrative analyses not previously possible. A focus on environmental drivers also invites collaboration between phylogeneticists and earth scientists in the selection and interpretation of geological proxies.

**5.2.3. Methods for poor fossil records and incomplete phylogenies.** Bapst (2014b) emphasized that the fossil records of different taxa can be qualitatively different in terms of what they can contribute to macroevolutionary inference. Examples reviewed here that thoroughly integrate paleontology with PCMs tend to be drawn from taxa, such as mammals, with good phylogenetic data and a rich fossil record. Most clades alive today have sparse records, and so there is much potential for methods that can leverage more modest amounts of fossil data. For trait evolution, Slater et al. (2012a) developed an approach that incorporates fossil phenotypes as priors on internal nodes of extant-only trees. This approach matches the common situation in which one or a few fossils are known but are too incomplete to incorporate formally in the phylogenetic analysis. In addition, methods that infer diversification dynamics from clade richness in incompletely resolved phylogenies can readily be extended to incorporate paleodiversities (Alfaro et al. 2009, Etienne & Apol 2009, Slater et al. 2012b). Although total diversities can be estimated from incomplete fossil records (e.g., Liow & Nichols 2010), such estimates can be quite uncertain. Accordingly, one might want to incorporate fossil diversities as minimum estimates or, instead, make use of the relative diversities in different time periods or taxa to constrain diversification histories.

Many fossil groups have rich records but have received little phylogenetic attention. In such cases, genealogical information in the form of putative ancestor–descendant species pairs (Alroy 1998a) can help with analyses of trait evolution, as can methods that use trees that are not fully resolved (noted above) or taxonomy combined with stratigraphy as a source of phylogenetic information (Soul & Friedman 2015).

#### SUMMARY POINTS

1. Increasing sophistication of the models employed by phylogenetic comparative methods has made phylogenies of extant species a viable source for understanding macroevolutionary processes, even those occurring in the distant past. With these developments has come increasing motivation to compare and integrate phylogenetic insights to those from the fossil record.
2. The most basic way to integrate fossil and phylogenetic data is through incorporating extant and extinct taxa in a combined, time-scaled tree. Recently developed methods permit extinct taxa to be treated equivalently to extant ones as dated terminal taxa (i.e., through tip dating). For trees composed of extinct taxa, a variety of post hoc tree scalings are useful for converting cladograms to time-scaled phylogenies.
3. Ancestral trait estimation, parameter estimation, and model inference can be improved by the addition of a modest amount of data from fossils, especially for evolutionary dynamics that are heterogeneous over time.
4. Although earlier phylogenetic methods for inferring diversification dynamics often failed to produce reasonable parameter estimates, especially for extinction, recent approaches

perform much better in the face of deviations from a constant-rate, birth-death model. These analyses have in some cases reproduced patterns known independently from the fossil record, suggesting that they may be fruitfully applied to taxa with poor fossil records.

5. Important challenges to an integrated macroevolution remain, including assessing the robustness of phylogenetic and paleontological methods, having the empirical results of each discipline feed back to the modeling choices of the other, and further developing methods that harness the geological record of environmental change.

## DISCLOSURE STATEMENT

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

## ACKNOWLEDGMENTS

We are grateful to D. Bapst, L. Soul, and S. Otto for thoughtful comments about this manuscript. G.S. was supported by a Peter Buck Fellowship of the National Museum of Natural History.

## LITERATURE CITED

- Albert JS, Johnson DM, Knouft JH. 2009. Fossils provide better estimates of ancestral body size than do extant taxa in fishes. *Acta Zool.* 90:357–84
- Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, et al. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *PNAS* 106:13410–14
- Alroy J. 1998a. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* 280:731–34
- Alroy J. 1998b. Equilibrial diversity dynamics in North American mammals. In *Biodiversity Dynamics: Turnover of Populations, Taxa and Communities*, ed. ML McKinney, JA Drake, pp. 232–87. New York: Columbia Univ. Press
- Alroy J. 2010. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. In *Quantitative Methods in Paleobiology*, ed. J Alroy, G Hunt, pp. 55–80. New Haven, CT: Paleontol. Soc.
- Alroy J. 2014. Accurate and precise estimates of origination and extinction rates. *Paleobiology* 40:374–97
- Arcila D, Pyron RA, Tyler JC, Ortí G, Betancur-R R. 2015. An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Mol. Phylogenet. Evol.* 82:131–45
- Aze T, Ezard THG, Purvis A, Coxall HK, Stewart DRM, et al. 2011. A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. *Biol. Rev.* 86:900–27
- Baker J, Meade A, Pagel M, Venditti C. 2015. Adaptive evolution toward larger size in mammals. *PNAS* 112:5093–98
- Bapst DW. 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods Ecol. Evol.* 3:803–7
- Bapst DW. 2013. A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. *Methods Ecol. Evol.* 4:724–33
- Bapst DW. 2014a. Assessing the effect of time-scaling methods on phylogeny-based analyses in the fossil record. *Paleobiology* 40:331–51
- Bapst DW. 2014b. Preparing paleontological datasets for phylogenetic comparative methods. In *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*, ed. LZ Garamszegi, pp. 515–44. Berlin: Springer-Verlag
- Beaulieu JM, O'Meara BC. 2015. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution* 69:1036–43

- Benton MJ. 2015. Exploring macroevolution using modern and fossil data. *Proc. R. Soc. B* 282:20150569
- Betancur-R R, Orti G, Pyron RA. 2015. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol. Lett.* 18:441–50
- Blomberg SP, Garland T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioural traits are more labile. *Evolution* 57:717–45
- Bokma F. 2002. Detection of punctuated equilibrium from molecular phylogenies. *J. Evol. Biol.* 15:1048–56
- Bokma F. 2008. Detection of “punctuated equilibrium” by Bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. *Evolution* 62:2718–26
- Bokma F, Godinot M, Maridet O, Ladevèze S, Costeur L, et al. 2015. Testing for Depéret’s rule (body size increase) in mammals using combined extinct and extant data. *Syst. Biol.* 65:98–108
- Brusatte SL, Benton MJ, Ruta M, Lloyd GT. 2008. Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. *Science* 321:1485–88
- Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL, Wadell PJ. 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* 42:384–97
- Cantalapiedra JL, FitzJohn RG, Kuhn TS, Fernandez MH, DeMiguel D, et al. 2014. Dietary innovations spurred the diversification of ruminants during the Caenozoic. *Proc. R. Soc. B* 281:20132746
- Charlesworth B, Lande R, Slatkin M. 1982. A neo-Darwinian commentary on macroevolution. *Evolution* 36:474–98
- Cheetham AH. 1986. Tempo of evolution in a Neogene bryozoan: rates of morphological change within and across species boundaries. *Paleobiology* 12:190–202
- Clarke JA, Middleton KM. 2008. Mosaicism, modules, and the evolution of birds: results from a Bayesian approach to the study of morphological evolution using discrete character data. *Syst. Biol.* 57:185–201
- Close RA, Friedman M, Lloyd GT, Benson RBJ. 2015. Evidence for a mid-Jurassic adaptive radiation in mammals. *Curr. Biol.* 25:2137–42
- Clyde WC, Fisher DC. 1997. Comparing the fit of stratigraphic and morphologic data in phylogenetic analysis. *Paleobiology* 23:1–19
- Condamine FL, Rolland J, Morlon H. 2013. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16:72–85
- Cooper N, Purvis A. 2010. Body size evolution in mammals: complexity in tempo and mode. *Am. Nat.* 175:727–38
- Cooper N, Thomas GH, Venditti C, Meade A, Freckleton RP. 2016. A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biol. J. Linn. Soc.* 118:64–77
- Crampton JS, Cooper RA, Sadler PM, Foote M. 2016. Greenhouse-icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *PNAS* 113:1498–503
- Donoghue PCJ, Benton MJ. 2007. Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol. Evol.* 22:424–31
- Eastman JM, Alfaro ME, Joyce P, Hipp AL, Harmon LJ. 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* 65:3578–89
- Eldredge N, Gould SJ. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In *Models in Paleobiology*, ed. TJM Schopf, pp. 82–115. San Francisco: Freeman, Cooper
- Etienne RS, Apol MEF. 2009. Estimating speciation and extinction rates from diversity data and the fossil record. *Evolution* 63:244–55
- Etienne RS, Haegeman B, Stadler T, Aze T, Pearson PN, et al. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. R. Soc. B* 279:1300–9
- Ezard THG, Aze T, Pearson PN, Purvis A. 2011. Interplay between changing climate and species’ ecology drives macroevolutionary dynamics. *Science* 332:349–51
- Ezard THG, Thomas GH, Purvis A. 2013. Inclusion of a near-complete fossil record reveals speciation-related molecular evolution. *Methods Ecol. Evol.* 4:745–53
- Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Biol.* 27:401–10
- Finarelli JA, Flynn JJ. 2006. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Syst. Biol.* 55:301–13

- Finarelli JA, Goswami A. 2013. Potential pitfalls of reconstructing deep time evolutionary history with only extant data, a case study using the Canidae (Mammalia, Carnivora). *Evolution* 67:3678–85
- Finnegan S, Heim NA, Peters SE, Fischer WW. 2012. Climate change and the selective signature of the Late Ordovician mass extinction. *PNAS* 109:6829–34
- Fisher DC. 2008. Stratocladistics: integrating temporal data and character data in phylogenetic inference. *Annu. Rev. Ecol. Evol. Syst.* 39:365–85
- FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 3:1084–92
- Foote M. 1990. Nearest-neighbor analysis of trilobite morphospace. *Syst. Zool.* 39:371–82
- Foote M. 1991. Morphological and taxonomic diversity in a clade's history: the blastoid record and stochastic simulations. *Contrib. Mus. Paleontol. (Univ. Mich.)* 28:101–40
- Foote M. 1996a. Models of morphological diversification. In *Evolutionary Paleobiology*, ed. D Jablonski, DH Erwin, JH Lipps, pp. 62–86. Chicago: Univ. Chicago Press
- Foote M. 1996b. On the probability of ancestors in the fossil record. *Paleobiology* 22:141–51
- Foote M. 1997a. Estimating taxonomic durations and preservation probability. *Paleobiology* 23:278–300
- Foote M. 1997b. The evolution of morphological diversity. *Annu. Rev. Ecol. Syst.* 28:129–52
- Foote M. 2000. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* 26:74–102
- Foote M. 2003. Origination and extinction through the Phanerozoic: a new approach. *J. Geol.* 111:125–48
- Foote M. 2011. Evolutionary dynamics of taxonomic structure. *Biol. Lett.* 8:135–38
- Fritz SA, Schnitzler J, Eronen JT, Hof C, Bohning-Gaese K, Graham CH. 2013. Diversity in time and space: wanted dead and alive. *Trends Ecol. Evol.* 28:509–16
- Garamszegi LZ, Moller AP. 2011. Nonrandom variation in within-species sample size and missing data in phylogenetic comparative studies. *Syst. Biol.* 60:876–80
- Gauthier J, Kluge AG, Rowe T. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* 4:105–209
- Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D, Drummond AJ. 2015. Bayesian total evidence dating reveals the recent crown radiation of penguins. arXiv:1506.04797 [q-bio.PE]
- Gavryushkina A, Welch D, Stadler T, Drummond AJ. 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLOS Comput. Biol.* 10:e1003919
- Gingerich PD. 1985. Species in the fossil record: concepts, trends, and transitions. *Paleobiology* 11:27–41
- Goldberg EE, Igic B. 2012. Tempo and mode in plant breeding system evolution. *Evolution* 66:3701–9
- Grimm GW, Kapli P, Bomfleur B, McLoughlin S, Renner SS. 2014. Using more than the oldest fossils: dating Osmundaceae with three Bayesian clock approaches. *Syst. Biol.* 64:396–405
- Guillerme T, Cooper N. 2015. Assessment of cladistic data availability for living mammals. bioRxiv:022970
- Guillerme T, Cooper N. 2016. Effects of missing data on topological inference using a Total Evidence approach. *Mol. Phylogenet. Evol.* 94:146–58
- Hansen TA. 1982. Modes of larval development in Early Tertiary neogastropods. *Paleobiology* 8:367–77
- Harmon LJ, Losos JB, Davies TJ, Gillespie RG, Gittleman JL, et al. 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64:2385–96
- Harrison LB, Larsson HCE. 2015. Among-character rate variation distributions in phylogenetic analysis of discrete morphological characters. *Syst. Biol.* 64:307–24
- Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *PNAS* 111:E2957–66
- Heim NA, Knope ML, Schaal EK, Wang SC, Payne JL. 2015. Cope's rule in the evolution of marine animals. *Science* 347:867–70
- Hendricks JR, Saupe EE, Myers CE, Hermesen EJ, Allmon WD. 2014. The generification of the fossil record. *Paleobiology* 40:511–28
- Hopkins MJ, Smith AB. 2015. Dynamic evolutionary change in post-Paleozoic echinoids and the importance of scale when interpreting changes in rates of evolution. *PNAS* 112:3758–63
- Huang DW, Goldberg EE, Roy K. 2015. Fossils, phylogenies, and the challenge of preserving evolutionary history in the face of anthropogenic extinctions. *PNAS* 112:4909–14
- Huang S, Roy K, Jablonski D. 2014. Do past climate states influence diversity dynamics and the present-day latitudinal diversity gradient? *Glob. Ecol. Biogeogr.* 23:530–40



- Huelsenbeck JP. 1991. When are fossils better than extant taxa in phylogenetic analysis? *Syst. Biol.* 40:458–69
- Huelsenbeck JP. 1994. Comparing the stratigraphic record to estimates of phylogeny. *Paleobiology* 20:470–83
- Huelsenbeck JP, Rannala B. 1997. Maximum likelihood estimation of phylogeny using stratigraphic data. *Paleobiology* 23:174–80
- Hughes M, Gerber S, Wills MA. 2013. Clades reach highest morphological disparity early in their evolution. *PNAS* 110:13875–79
- Hunt G. 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology* 32:578–601
- Hunt G. 2013. Testing the link between phenotypic evolution and speciation: an integrated palaeontological and phylogenetic analysis. *Methods Ecol. Evol.* 4:714–23
- Hunt G, Hopkins MJ, Lidgard S. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. *PNAS* 112:4885–90
- Hunt G, Roy K. 2006. Climate change, body size evolution, and Cope's rule in deep-sea ostracodes. *PNAS* 103:1347–52
- Jackson JBC, Cheetham AH. 1990. Evolutionary significance of morphospecies: a test with cheilostome Bryozoa. *Science* 248:579–82
- Jepsen GL, Mayr E, Simpson GG, eds. 1949. *Genetics, Paleontology, and Evolution*. Princeton, NJ: Princeton Univ. Press
- Kidwell SM, Holland SM. 2002. The quality of the fossil record: implications for evolutionary analysis. *Annu. Rev. Ecol. Syst.* 33:561–88
- Klopfstein S, Vilhelmsen L, Ronquist F. 2015. A nonstationary Markov model detects directional evolution in hymenopteran morphology. *Syst. Biol.* 64:1089–103
- Ksepka DT, Benton MJ, Carrano MT, Gandolfo MA, Head JJ, et al. 2011. Synthesizing and databasing fossil calibrations: divergence dating and beyond. *Biol. Lett.* 7:801–3
- Lane A, Janis CM, Sepkoski JJ. 2005. Estimating paleodiversities: a test of the taxic and phylogenetic methods. *Paleobiology* 31:21–34
- Larson-Johnson K. 2016. Phylogenetic investigation of the complex evolutionary history of dispersal mode and diversification rates across living and fossil Fagales. *N. Phytol.* 209:418–35
- Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–25
- Liow LH, Finarelli JA. 2014. A dynamic global equilibrium in carnivoran diversification over 20 million years. *Proc. R. Soc. B* 281:20132312
- Liow LH, Fortelius M, Bingham E, Lintulaakso K, Mannila H, et al. 2008. Higher origination and extinction rates in larger mammals. *PNAS* 105:6097–102
- Liow LH, Nichols JD. 2010. Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: capture-mark-recapture (CMR) approaches. In *Quantitative Methods in Paleobiology*, ed. J Alroy, G Hunt, pp. 81–94. New Haven, CT: Paleontol. Soc.
- Liow LH, Quental TB, Marshall CR. 2010. When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Syst. Biol.* 59:646–59
- Maddison WP, FitzJohn RG. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Syst. Biol.* 64:127–36
- Maddison WP, Midford PE, Otto SP. 2007. Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* 56:701–10
- Magnuson-Ford K, Otto SP. 2012. Linking the investigations of character evolution and species diversification. *Am. Nat.* 180:225–45
- Marcot JD, Fox DL. 2008. StrataPhy: a new computer program for stratocladistic analysis. *Palaeontol. Electron.* 11:1–16
- McShea DW. 1994. Mechanisms of large-scale evolutionary trends. *Evolution* 48:1747–63
- Moen D, Morlon H. 2014. Why does diversification slow down? *Trends Ecol. Evol.* 29:190–97
- Monroe MJ, Bokma F. 2010. Little evidence for Cope's rule from Bayesian phylogenetic analysis of extant mammals. *J. Evol. Biol.* 23:2017–21
- Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17:508–25



- Morlon H, Parsons TL, Plotkin JB. 2011. Reconciling molecular phylogenies with the fossil record. *PNAS* 108:16327–32
- Nee S, May RM, Harvey PH. 1994. The reconstructed evolutionary process. *Philos. Trans. R. Soc. B* 344:305–11
- Norell MA. 1992. Taxic origin and temporal diversity: the effect of phylogeny. In *Extinction and Phylogeny*, ed. MJ Novacek, QD Wheeler, pp. 89–118. New York: Columbia Univ. Press
- Novack-Gottshall PM, Lanier MA. 2008. Scale-dependence of Cope's rule in body size evolution of Paleozoic brachiopods. *PNAS* 105:5430–34
- O'Meara BC. 2012. Evolutionary inferences from phylogenies: a review of methods. *Annu. Rev. Ecol. Evol. Syst.* 43:267–85
- O'Reilly JE, dos Reis M, Donoghue P. 2015. Dating tips for divergence-time estimation. *Trends Genet.* 31:637–50
- Pagel M, Venditti C, Meade A. 2006. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* 314:119–21
- Pattinson DJ, Thompson RS, Piotrowski AK, Asher RJ. 2014. Phylogeny, paleontology, and primates: Do incomplete fossils bias the tree of life? *Syst. Biol.* 64:169–86
- Pennell MW, Harmon LJ. 2013. An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Ann. N.Y. Acad. Sci.* 1289:90–105
- Pennell MW, Harmon LJ, Uyeda JC. 2014. Is there room for punctuated equilibrium in macroevolution? *Trends Ecol. Evol.* 29:23–32
- Peters SE, Ausich WI. 2008. A sampling-adjusted macroevolutionary history for Ordovician–Early Silurian crinoids. *Paleobiology* 34:104–16
- Phillimore AB, Price TD. 2008. Density-dependent cladogenesis in birds. *PLOS Biol.* 6:e71
- Pires MM, Silvestro D, Quental TB. 2015. Continental faunal exchange and the asymmetrical radiation of carnivores. *Proc. R. Soc. B* 282:20151952
- Plotnick RE, Wagner PJ. 2006. Round up the usual suspects: common genera in the fossil record and the nature of wastebasket taxa. *Paleobiology* 32:126–46
- Polly PD. 2001. Paleontology and the comparative method: ancestral node reconstructions versus observed node values. *Am. Nat.* 157:596–609
- Puttick MN, Thomas GH. 2015. Fossils and living taxa agree on patterns of body mass evolution: a case study with Afrotheria. *Proc. R. Soc. B* 282:20152023
- Pybus OG, Harvey PH. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B* 267:2267–72
- Pyron RA. 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.* 60:466–81
- Pyron RA, Burbrink FT. 2012. Trait-dependent diversification and the impact of palaeontological data on evolutionary hypothesis testing in New World ratsnakes (tribe Lampropeltini). *J. Evol. Biol.* 25:497–508
- Pyron RA, Burbrink FT. 2013. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends Ecol. Evol.* 28:729–36
- Quental TB, Marshall CR. 2009. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution* 63:3158–67
- Quental TB, Marshall CR. 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25:434–41
- Rabosky DL. 2006. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60:1152–64
- Rabosky DL. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64:1816–24
- Rabosky DL. 2013. Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 44:481–502
- Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLOS ONE* 9:e89543
- Rabosky DL. 2015. No substitute for real data: a cautionary note on the use of phylogenies from birth–death polytomy resolvers for downstream comparative analyses. *Evolution* 69:3207–16
- Rabosky DL, Goldberg EE. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* 64:340–55

- Rabosky DL, Lovette IJ. 2008a. Density-dependent diversification in North American wood warblers. *Proc. R. Soc. B* 275:2363–71
- Rabosky DL, Lovette IJ. 2008b. Explosive evolutionary radiations: Decreasing speciation or increasing extinction through time? *Evolution* 62:1866–75
- Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, et al. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* 4:1958
- Ronquist F, Klopstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Syst. Biol.* 61:973–99
- Roy K, Hunt G, Jablonski D. 2009. Phylogenetic conservatism of extinctions in marine bivalves. *Science* 325:733–37
- Ruta M, Wagner PJ, Coates MI. 2006. Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc. R. Soc. B* 273:2107–11
- Sansom RS, Wills MA. 2013. Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Sci. Rep.* 3:2545
- Schluter D, Price T, Mooers A, Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699–711
- Sepkoski JJ, Bambach RK, Raup DM, Valentine JW. 1981. Phanerozoic marine diversity and the fossil record. *Nature* 293:435–37
- Sepkoski JJ, Kendrick DC. 1993. Numerical experiments with model monophyletic and paraphyletic taxa. *Paleobiology* 19:168–84
- Silvestro D, Antonelli A, Salamin N, Quental TB. 2015. The role of clade competition in the diversification of North American canids. *PNAS* 112:8684–89
- Silvestro D, Schnitzler J, Liow LH, Antonelli A, Salamin N. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Syst. Biol.* 63:349–67
- Simpson C, Kiessling W, Mewis H, Baron-Szabo RC, Muller J. 2011. Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. *Evolution* 65:3274–84
- Simpson GG. 1944. *Tempo and Mode in Evolution*. New York: Columbia Univ. Press
- Slater GJ. 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous–Palaeogene boundary. *Methods Ecol. Evol.* 4:734–44
- Slater GJ. 2014. Correction to ‘Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous–Palaeogene boundary’, and a note on fitting macroevolutionary models to comparative paleontological data sets. *Methods Ecol. Evol.* 5:714–18
- Slater GJ. 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *PNAS* 112:4897–902
- Slater GJ, Harmon LJ, Alfaro ME. 2012a. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66:3931–44
- Slater GJ, Harmon LJ, Wegmann D, Joyce P, Revell LJ, Alfaro ME. 2012b. Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate Bayesian computation. *Evolution* 66:752–62
- Slater GJ, Pennell MW. 2014. Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. *Syst. Biol.* 63:293–308
- Soul LC, Friedman M. 2015. Taxonomy and phylogeny can yield comparable results in comparative paleontological analyses. *Syst. Biol.* 64:608–20
- Spencer MR, Wilberg EW. 2013. Efficacy or convenience? Model-based approaches to phylogeny estimation using morphological data. *Cladistics* 29:663–71
- Stadler T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *PNAS* 108:6187–92
- Stanley SM. 1973. An explanation for Cope’s rule. *Evolution* 27:1–26
- Stanley SM. 1979. *Macroevolution: Pattern and Process*. Baltimore, MD: Johns Hopkins Univ. Press
- Stanley SM. 1990. The general correlation between rate of speciation and rate of extinction: fortuitous causal linkages. In *Causes of Evolution: A Paleontological Perspective*, ed. RM Ross, WD Allmon, pp. 103–27. Chicago: Univ. Chicago Press
- Valentine JW. 1980. Determinants of diversity in higher taxonomic categories. *Paleobiology* 6:444–50

- Venditti C, Meade A, Pagel M. 2011. Multiple routes to mammalian diversity. *Nature* 479:393–96
- Wagner PJ. 1995. Testing evolutionary constraint hypotheses with early Paleozoic gastropods. *Paleobiology* 21:248–72
- Wagner PJ. 1998. A likelihood approach for evaluating estimates of phylogenetic relationships among fossil taxa. *Paleobiology* 24:430–49
- Wagner PJ. 2012. Modelling rate distributions using character compatibility: implications for morphological evolution among fossil invertebrates. *Biol. Lett.* 8:143–46
- Wagner PJ, Marcot JD. 2013. Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. *Methods Ecol. Evol.* 4:703–13
- Webster AJ, Purvis A. 2002. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc. R. Soc. B* 269:143–49
- Wright AM, Hillis DM. 2014. Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological data. *PLOS ONE* 9:e109210
- Xing Y, Onstein RE, Carter RJ, Stadler T, Peter Linder H. 2014. Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. *Evolution* 68:2821–32