

Testing Character Evolution Models in Phylogenetic Paleobiology 1

1 Introduction

Historically, drawing macroevolutionary inferences from phylogenetic trees has been a two-step process (Harvey and Pagel, 1991). First, a researcher would estimate a phylogenetic tree from a matrix of phylogenetic characters (typically morphological characters or molecular sequence characters). Then, they would use that tree (or a set of trees, such as a posterior sample) to fit a macroevolutionary model. Over the past decade, models that blend macroevolutionary inference with phylogenetic inference have become increasingly common. For example, the fossilized birth–death process is used to estimate dated phylogenetic trees (Stadler, 2011; Heath et al., 2014). This process is usually implemented as a Bayesian hierarchical model, in which one model describes the process of character change for phylogenetic characters, one describes the distribution of evolutionary rates over the tree, and one describes the process of speciation, extinction, and sampling that led to the observed tree (see Warnock and Wright in this issue for a more complete discussion of this). In this Element, we describe an approach to fitting complex hierarchical models using a focal dataset of cinctan echinoderms.

We can divide macroevolutionary hypotheses into two non-mutually exclusive groups: those making predictions about origination and extinction dynamics, and those making predictions about rates and modes of trait evolution. The latter group includes hypotheses about shifts in rates of anatomical change and hypotheses about driven trends in which particular character states become more (or less) common over time. Hypotheses predicting such patterns come both from macroecological theory and from evolutionary-developmental theory, and thus span a range of basic issues including developmental, ecological, and physical constraints, and selection (Valentine et al., 1969; Valentine, 1980). Research programs dedicated to assessing shifts in rates and modes of anatomical evolution have been a staple of quantitative paleobiology since the early 1990s. Accordingly, anatomical character evolution models that describe the predictions of these different macroevolutionary hypotheses have important theoretical implications for these endeavors.

Phylogeneticists have long been interested in the same sorts of character evolution models, albeit for very different reasons. Hypotheses of phylogenetic relationships make exact predictions about character state evolution among taxa given observed data and models of character change (e.g., Kimura, 1980; Felsenstein, 1981; Hasegawa et al., 1985; Tavaré, 1986). The most common phylogenetic model for morphology makes the assumption of time-invariant models with no biases in the rate of character acquisition and loss (Lewis, 2001). The expectations of character evolution, of associations of characters

with one another, and disparities between taxa are quite different when rates of acquisition and loss vary among characters and with time. This is particularly true when we include divergence times as part of phylogenetic hypotheses (Huelsenbeck et al., 2000; Sanderson, 2002; Drummond et al., 2006), but it is still true if we worry only about general cladistic relationships (i.e., which taxa are most closely related to each other; see Felsenstein (1981); Nylander et al. (2004); Wright et al. (2016)). In other words, many of the conceptual mice that paleobiologists seek to capture are the conceptual mousetraps that systematists seek to use to capture phylogenies.

Many readers' first instincts will be that this presents paleobiological phylogeneticists with a quandary: which comes first, the character evolution models or the phylogenetic inference? Part of the dilemma here stems from a historical view that we should treat phylogenetic analysis and macroevolutionary analysis as two separate endeavors (e.g., Harvey and Pagel, 1991). When we estimate a phylogenetic tree, we typically need to make simplifying assumptions about the evolution of our phylogenetic characters for tractability of the analysis. For our comparative methods, we are often using a smaller subset of the data to explore more complex models, possibly even seeking to falsify those same simplifying assumptions. Here, we advocate a very different approach that stems from hierarchical Bayesian phylogenetic approaches. That is, we should not view phylogenetic analysis and macroevolutionary analysis as two independent projects, but instead as two parts of the same endeavor of unraveling the evolutionary history of fossil taxa. These evolutionary histories include when clades and lineages diverged, the consistency of character change rates, biases in state acquisition, the process of diversification that led to the observed tree, and (of course) exactly how taxa were related to each other. Along the same lines, we have to accept and even embrace the fact that there will always be some degree of uncertainty in all of these things. These uncertainties are not a reason to abandon the endeavor as hopeless; on the contrary, it will mean that those conclusions that we can reach do not assume that specific historical details are true.

In this work, we will provide an example of the approach that we are advocating using a series of analyses of the *Cincta*, an extinct clade of “carpoid” echinoderms from the middle Cambrian. We will detail how paleobiologists can adapt different clock models and character state evolution models initially devised to accommodate uncertainties in molecular evolution to represent and model macroevolutionary hypotheses. In doing so, we will also outline protocol that paleontologists can replicate to conduct analogous analyses on other clades. We will emphasize how the combination of Markov Chain Monte Carlo analyses and stepping-stone tests allow us to marginalize specific details of

Testing Character Evolution Models in Phylogenetic Paleobiology 3

character evolution models and phylogenetic relationships in order to generate the best joint summary of a clade's evolutionary history. Because there are innumerable possible models that one might consider, we will draw attention to existing methods with which paleontologists might already be familiar that should be useful for suggesting particular models as worthy of consideration. Finally, we will briefly outline other theoretical and methodological areas that remain for paleobiologists and systematists to resolve and unite in the future.

2 Taxonomic Background and Data

2.1 Cincta: An Enigmatic Clade of Cambrian Echinoderms

Echinoderms are a diverse phylum of marine animals represented today by more than 7,000 living species (Brusca and Brusca, 2003) distributed among five extant classes, including sea stars, brittle stars, echinoids, sea cucumbers, and crinoids. However, the spectacular diversity of extant echinoderms, measured by both species richness and anatomical variety, represents a paltry fraction of their prodigious evolutionary history recorded in the fossil record. Because echinoderms possess a mineralized endoskeleton made of high-magnesium calcite (calcium carbonate) and occur in virtually all habitats across the spectrum of marine depositional environments, the echinoderm fossil record is spectacularly complete and reveals approximately 30 clades distributed among 21 taxonomic classes spanning the entire Phanerozoic Eon (Sprinkle and Kier, 1987; Sumrall, 1997; Sumrall and Waters, 2012; Zamora and Rahman, 2014; Wright et al., 2017; Sheffield and Sumrall, 2019). Unlike familiar echinoderms inhabiting modern oceans, such as sea stars and sea urchins (echinoids), Cambrian lineages comprise an unfamiliar, taxonomically and morphologically diverse assemblage of predominately stem-group taxa exhibiting a diversity of body plans, life modes, and ecological traits unseen in extant lineages (Sprinkle, 1973; Zamora et al., 2013a; Zamora and Rahman, 2014).

Perhaps no group of early echinoderms has received greater attention and controversy than the carpoids (Rahman, 2009). Sometimes called homalozoans or calcichordates in the literature, carpoids comprise a heterogeneous assemblage of extinct echinoderms including ctenocystoids, cinctans (*Homostealea*), solutes (*Homoiostealea*), and stylophorans. Although carpoids possess unique skeletal features that unambiguously identify their echinoderm affinities (David et al., 2000; Bottjer et al., 2006; Rahman, 2009; Zamora et al., 2020), they lack other traits considered synapomorphies of crown-group echinoderms. For example, all extant echinoderms exhibit pentaradial symmetry in adults and possess a water vascular system, unique to the phylum, used for locomotion,

respiration, and excretion (Nichols, 1972). In contrast, ‘carpoid’ taxa exhibit bilateral to asymmetrical forms, and it’s debated whether some possessed a water vascular system (Smith, 2005; Lefebvre et al., 2019). Although the phylogenetic position of carpoids has long been contentiously debated (see Rahman, 2009, and Rahman et al., 2009, and articles cited therein), only recently have computer-based phylogenetic analyses played a major role in evaluating alternative hypotheses (Sumrall, 1997; Smith and Zamora, 2013; Zamora and Rahman, 2014), and only one previous study tested phylogenetic hypotheses using stratigraphic data (Rahman et al., 2009). Crucially, the character matrices constructed for these analyses have greatly benefited from recent improvements to identifying homologous characters across morphologically disparate early echinoderm lineages, often arising from new fossil discoveries (e.g., Zamora et al., 2012; Smith and Zamora, 2013). Taxonomic controversy remains (David et al., 2000), though both recent computational phylogenetic analyses and stratigraphic congruence metrics support the hypothesis that carpoids comprise a paraphyletic assemblage of stem-group echinoderms (Rahman et al., 2009; Smith and Zamora, 2013; Zamora and Rahman, 2014). If this view is correct, then carpoids help document the radical transition in echinoderm evolution from an ancestral, bilaterian body plan to the pentaradial symmetry characteristic of crown-group forms that have dominated marine ecosystems since the close of the Cambrian. Regardless of their specific branching relationships in the echinoderm tree of life, it is nevertheless clear that understanding the distribution of character combinations and patterns of trait evolution in these enigmatic, pre-radial echinoderm lineages is critical to deciphering the early evolution of the phylum.

Cinctans are a significant group of non-radiate, carpoid echinoderms temporally restricted to the middle Cambrian (Miaolingian 509–497 Ma, with occurrences of cinctans from 506.6–297 Ma) and paleogeographically restricted to western Gondwana, Avalonia, and Siberia. Cinctans are generally small (i.e., 1 to 10 mm in length), flattened, symmetrical to irregularly shaped fossils resembling a tennis racquet or badly formed pancake, generally interpreted as an adaptation to an epibenthic, suspension-feeding lifestyle (Rahman and Zamora, 2009; Rahman et al., 2015). Like all echinoderms, cinctans have a complex, multielement, calcitic endoskeleton, which makes them particularly amenable for coding discrete, phylogenetic characters in fossil taxa (Smith and Zamora, 2009). The main body, called the theca, is surrounded by a series of rigid, stout, marginal plates (called the cinctus), which surrounds a central body of smaller, tessellated integument plates on both dorsal and ventral sides. The mouth is a circular opening located at the end of a narrow food groove (or pair of grooves) on the right anterior side of the theca. A posterior appendage,

Testing Character Evolution Models in Phylogenetic Paleobiology 5

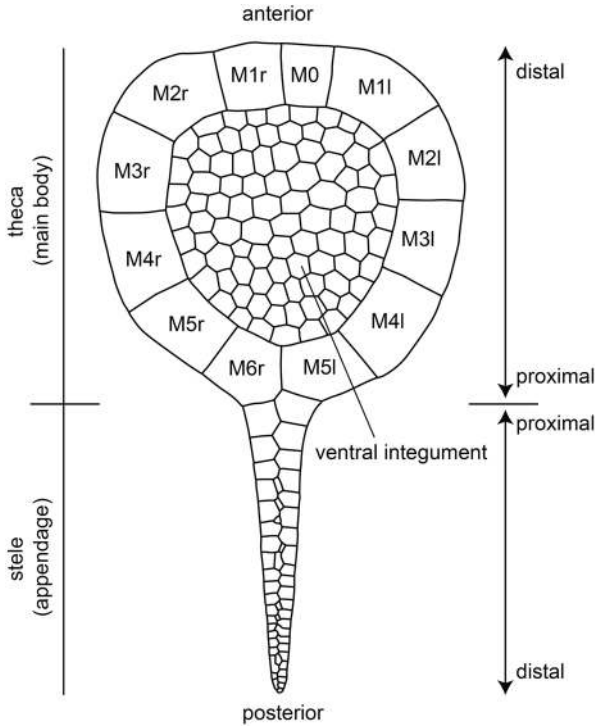


Figure 1 Generalized diagram of cinctan morphology based on the ventral side of *Trochocystites*. Marginal plates (i.e., comprising the cinctus) are labeled from anterior to posterior following Friedrich (1993); “l” and “r” refer to the left and right sides of the theca in dorsal view. See Smith and Zamora (2009) and Rahman (2016) for additional views of fossils and their anatomical reconstructions.

called the stele, forms a rigid structure extending from the cinctus, commonly subequal in length to the theca (Figure 1).

Despite their diminutive size, geological antiquity, and narrow paleogeographic and stratigraphic ranges, the significance of cinctans to understanding early echinoderm evolution, as well as their evolutionary implications surrounding ancestral character states in ancient deuterostomes (Smith and Swalla, 2009), has led to a substantial amount of interest to decipher their paleobiology. Recent advances in cinctan paleobiology include efforts to better understand patterns of taxonomic diversity (Zamora and Álvaro, 2010), ontogeny and development (Smith, 2005; Zamora et al., 2013b), life mode and feeding ecology (Rahman et al., 2009, 2015; Zamora and Rahman, 2015), convergence and adaptive evolution (Zamora and Smith, 2008), and phylogenetic relationships (Friedrich, 1993; Sdzuy, 1993; Smith and Zamora, 2009; Zamora et al., 2013b). In this Element, we combine morphological data with fossil age information to

reevaluate phylogenetic relationships and evolutionary dynamics among cinctan lineages using hierarchical Bayesian phylogenetic models, and provide a phylogenetic template for future systematic and macroevolutionary studies.

2.2 Character Data

We use the character data initially published by Smith and Zamora (2009) and subsequently augmented by Zamora et al. (2013b). The analyzed matrix includes 22 cinctan species plus one outgroup taxon (*Ctenocystis*, represented by *Ctenocystis utahensis*). An additional four cinctan species are excluded due to inadequate material for coding. We refer the readers to the papers cited above for additional information concerning the character data.

Both disparity analyses of these data conducted and arguments pertinent to early echinoderm evolution in the literature (e.g., Smith et al., 2013) suggest that cinctans might exhibit “Early Burst”-type dynamics (Figure 2A), in which a broader range of anatomies appears early in clade history than expected if rates of change were reasonably consistent over time. Standing disparity versus log-standing richness patterns deviate strongly from expectations given constant rates of change (Jablonski, 2020; Wright, 2017a), but this reflects in part the clade decreasing in richness through its later history. The same relationship with cumulative disparity (i.e., disparity among all species known by some date) shows a weaker trend toward higher disparity than expected during the first half of clade evolution (Figure 2C). This in turn suggests that rates of change might have been higher early in clade history (Foote, 1996b).

2.3 Chronostratigraphic Data

Our chronostratigraphic data come from 221 occurrences of Cambrian echinoderm species from 143 localities last downloaded from the Paleobiology Database (PBDB) on January 1, 2020 (Wagner, 2021).

The locality and occurrence data came from 81 references with the seven biggest sources including Nardin et al. (2017), Zamora (2009), Chlupac et al. (1998), Sprinkle and Collins (2006), Termier and Termier (1973), and Sprinkle (1973). We ourselves entered 108 of those occurrences and 72 of those localities, and updated 51 of the remaining localities for the purpose of this Element. After accounting for synonymies and variant spellings, the localities represent 55 different rock units (i.e., formations and formations+members). The accepted names of the species occurrences reflect 361 taxonomic opinions, 155 of which we entered for the purposes of this Element.

The Paleobiology Database returns ages based only on the entered interval, which usually is a stage. Here, nearly every cinctan-bearing locality is assigned

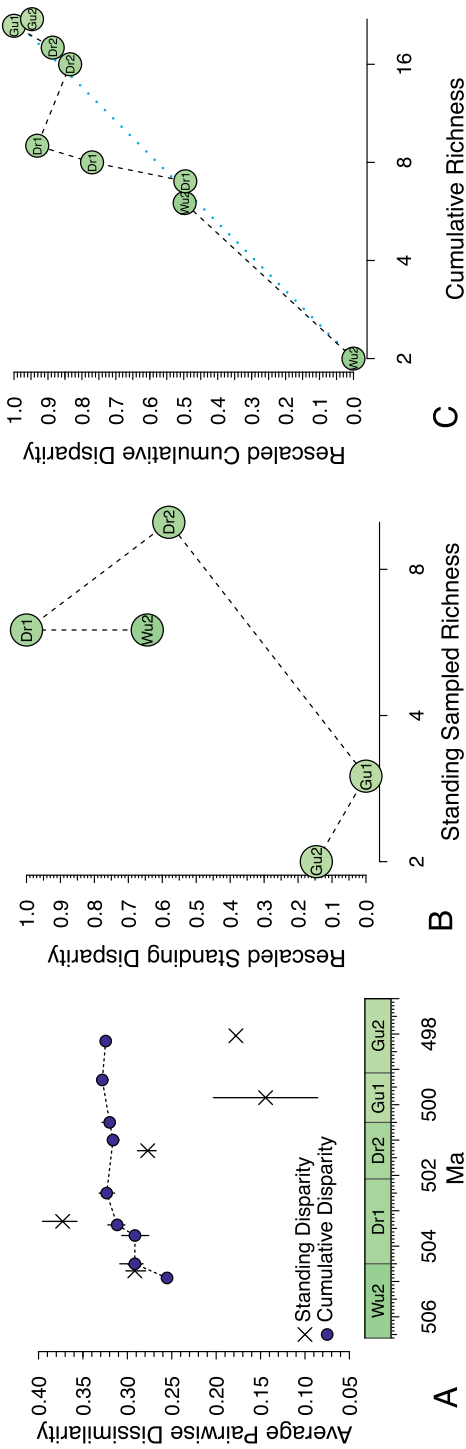


Figure 2 Disparity patterns for cinctans. (A) Disparity over time based on average pairwise dissimilarities among taxa (Foote, 1992). Vertical bars represent 90-percentile error bars from bootstrapping pairwise comparisons (Foote, 1993). X gives “traditional” standing disparity, which reflects only species extant during a stage slice (see, e.g., Hughes et al., 2013). Note also that a minimum of two species must be present. Circles give the cumulative disparity among all members of the clade sampled in rocks that age or older, regardless of whether they are still extant, and thus depict the total range of anatomical types derived within the clade (Wagner and Estabrook, 2015). (B) Rescaled standing disparity versus log-taxonomic richness. The dashed line here and in (C) gives the expected change given continuous rates of morphological innovation. Because cinctans begin with relatively high richness and then decline over time, the curve despite no increase in standing richness. (C) Rescaled cumulative disparity given cumulative richness. The scale here is finer as it reflects disparity evolved through points in time rather than extant during stage slices. The rapid rise in disparity from the Wuliuan to the early Drumian in (A) and (B) now reflects the appearance of very anatomically disparate species in the early Drumian while earlier species introduce disparity comparable to that introduced by late Drumian and later species.

to the middle Cambrian and thus receives a possible age of 513–501 Ma. However, PBDB provides information allowing much more exact ages. For example, PBDB collection 67775 is one of four including *Trochocystites bohemicus*. This collection is assigned to the Middle Cambrian and thus is dated by the PBDB as 513–501 Ma. However, this collection represents the Skryje Shale, which is known to span four trilobite zones that restrict the age to 505.2–500.7 Ma (based on correlations among trilobite zones by Geyer (2019) to trilobite zone ages in Gradstein et al. (2012)). Thus, if we had no further information, then those would be the oldest and youngest possible ages for this collection. However, PBDB collection 67775 also is assigned to the *Eccaparadoxides pusillus* trilobite zone, which further restricts the age to 505.2–504.5 Ma. We use a database of Paleozoic rock units and faunal zones compiled by one of us (PJW) as a thesaurus to provide more exact earliest and latest possible ages for each cinctan-bearing collection. In addition to refining dates, the database also effectively updates the chronostratigraphic unit to which a locality is assigned if current ideas about the age of a trilobite zone have changed since the paper(s) providing the original data. The typical locality now can be restricted to a 0.7 million year window. Prominent sources for the information relevant to our Element and for interregional correlations of rock units and trilobite zones include Alvaro et al. (2001); Liñán et al. (2004); Geyer and Landing (2006); Geyer and Shergold (2000); Alvaro et al. (2007); and Geyer (2019). The overall timescale is that of Gradstein et al. (2012). (Note that the more recent timescale of Gradstein et al. (2020), which was published after we conducted these analyses, provides nearly identical dates for the relevant Middle Cambrian trilobite zones and thus should not generate markedly different results.)

We use the refined dates to put lower and upper bounds on the possible first-appearance (FA) dates of the cinctan species (Table 1). For species known from only single intervals or trilobite zones, the lower and upper bounds for both first and last appearances are necessarily identical. This is not the case for species spanning multiple intervals. For example, *Gyrocystis platessa* occurs in rocks as old as the *Badulesia granieri* trilobite zone (504.9–504.5 Ma given Geyer (2019) and Gradstein et al.'s timescale) and also occurs in rocks as young as the *Solenopleuropsis marginata* trilobite zone (501.0–499.3 Ma). Here, the latest first possible appearance is 504.5 Ma. We choose the widest possible uncertainty. For example, *Gyrocystis erecta* occurs in rocks belonging to the *Solenopleuropsis* zone (503.1–501.0 Ma) but also in rocks dated more specifically to the *Solenopleuropsis thorali* subzone (501.6–501.0 Ma). Because the former set of occurrences might be as old as 503.1 Ma (given existing information), we set the possible lower and upper bounds on the FA for *G. erecta* at $FA_{LB}=503.1$ and $FA_{UB}=501.0$ Ma.

Testing Character Evolution Models in Phylogenetic Paleobiology 9

Table 1 Chronostratigraphic information for analyzed taxa based on occurrences in the Paleobiology Database. FA and LA denote first and last appearance dates, with LB and UB giving the oldest and youngest possible FA and LA given the finest chronostratigraphic resolution possible (e.g., a trilobite zone or local chronostratigraphic unit). “N_S” gives number of sites (= collections or localities) that a species occupies. “N_R” gives number of rock units (formations or members) that a species occupies. Dates for *Ctenocystis* represent the entire genus; however, the coded species (*C. utahensis*) is also the oldest known *Ctenocystis*.

Taxon	FA _{LB}	FA _{UB}	LA _{LB}	LA _{UB}	N _S	N _R
<i>Ctenocystis</i>	506.6	506.5	501.0	500.5	4	3
<i>Gyrocystis platessa</i>	504.9	504.5	501.0	499.3	13	4
<i>Gyrocystis testudiformis</i>	503.1	502.5	503.1	502.5	4	1
<i>Gyrocystis cruzae</i>	503.1	501.0	503.1	501.0	1	1
<i>Gyrocystis badulesiensis</i>	503.1	501.0	503.1	501.0	2	1
<i>Gyrocystis erecta</i>	503.1	501.0	501.6	501.0	2	1
<i>Progyrocystis disjuncta</i>	503.1	501.0	503.1	501.0	1	1
<i>Protocinctus mansillaensis</i>	506.6	505.4	506.6	505.4	1	1
<i>Elliptocinctus barrandei</i>	501.6	501.0	501.6	499.3	6	3
<i>Elliptocinctus vizcainoi</i>	504.5	503.4	504.5	503.4	1	1
<i>Sucocystis theronensis</i>	501.6	501.0	501.6	501.0	2	2
<i>Sucocystis bretoni</i>	501.0	500.5	501.0	500.5	1	1
<i>Lignacycystis barriosensis</i>	501.6	501.0	501.6	501.0	3	1
<i>Undatacinctus undata</i>	501.0	499.3	501.0	499.3	1	1
<i>Sucocystis acrofera</i>	499.3	498.2	499.3	498.2	2	1
<i>Undatacinctus quadricornuta</i>	501.0	499.3	501.0	499.3	1	1
<i>Undatacinctus melendezi</i>	501.0	499.3	498.2	497.0	9	2
<i>Asturicycystis jaekeli</i>	504.9	504.5	504.9	504.5	1	1
<i>Sotocinctus ubaghsi</i>	505.4	504.9	505.4	504.5	2	2
<i>Trochocystites bohemicus</i>	505.2	504.5	503.0	502.2	4	3
<i>Trochocystoides parvus</i>	504.5	503.7	504.5	503.7	1	1
<i>Ludwigicinctus truncatus</i>	501.6	500.5	501.6	500.5	1	1
<i>Graciacycystis ambigua</i>	504.9	504.5	503.7	503.1	3	1

3 Methods

3.1 Estimating Starting Values for Sampling and Diversification Rates

We use the Paleobiology Database occurrences described above to derive initial estimates of origination, extinction, and sampling for Cambrian echinoderms. Although this Element focuses on just cinctans, other echinoderms represent a

taphonomic control for initial estimates of sampling: rocks from environments permitting other identifiable fossils of other Cambrian echinoderms are those in which there is some probability > 0 that we would be able to sample cinctans if they had lived in those environments (Bottjer and Jablonski, 1988); in contrast, fossiliferous localities lacking identifiable echinoderms might represent environments in which echinoderms (cinctan or otherwise) lived, but that no longer represent sampling opportunities. Other echinoderms also provide a much larger sample size for initial estimates of origination and extinction rates than do cinctans. Although diversification rates likely varied among and within echinoderm clades as well as over time, paleontological data long have suggested that different major clades are typified by general diversification rates (Sepkoski, 1981). The larger sample sizes afforded by all echinoderms reduce the chance that our *initial* estimates will be wildly inaccurate as an artifact of sample size. We use a modified version of the Three-Timer method (Alroy, 2015) that uses lognormal distributions for sampling rates per stage slice rather than a single value (Wagner and Marcot, 2013). Note that we use both the diversification and sampling rates to seed the prior distribution with feasible starting values from which to generate new proposals for diversification and sampling parameters in MCMC generations and not as fixed values. Our rationale for adding this extra step is simply that, like all other search algorithms, MCMC analyses beginning with “realistic” parameters should converge on “correct” parameters faster than those analyses beginning with unrealistic parameters; and as diversification and sampling are among the parameters being varied from one iteration to the next in each analysis, this should make it easier for the analyses to achieve convergence for all of the parameters. We did estimate two parameters that remain static in MCMC analyses directly from PBDB data. One is the probability of taxon sampling for the youngest species. This parameter is separate from the general sampling parameter because, unlike this analysis, many analyses include extant taxa that reflect a fundamentally different sampling regime. The other is the earliest possible divergence date for the clade. We estimate this using the cal-3 metric (Bapst, 2013), with the lower bounds set at $p=0.003$ (i.e., $1 - 0.05^4$).

3.2 Models

The fossilized birth–death is a hierarchical model, meaning that different model subcomponents explain the evolution of the phylogenetic characters (the morphological evolution model), the distribution of evolutionary rates across the tree (the clock model), and the model that describes the speciation (λ), extinction (μ), and fossil sampling intensity (ψ) leading to the tree (the tree model).