



**Cite this article:** Crouch NMA, Edie SM, Collins KS, Bieler R, Jablonski D. 2021

Calibrating phylogenies assuming bifurcation or budding alters inferred macroevolutionary dynamics in a densely sampled phylogeny of bivalve families. *Proc. R. Soc. B* **288**: 20212178. <https://doi.org/10.1098/rspb.2021.2178>

Received: 1 October 2021

Accepted: 10 November 2021

**Subject Category:**

Evolution

**Subject Areas:**

evolution, palaeontology

**Keywords:**

ghost lineage, mass extinction, lineage-through-time, time calibration, Mollusca

**Author for correspondence:**

Nicholas M. A. Crouch

e-mail: ncrouch@uchicago.edu

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5718844>.

# Calibrating phylogenies assuming bifurcation or budding alters inferred macroevolutionary dynamics in a densely sampled phylogeny of bivalve families

Nicholas M. A. Crouch<sup>1</sup>, Stewart M. Edie<sup>2</sup>, Katie S. Collins<sup>3</sup>, Rüdiger Bieler<sup>4</sup> and David Jablonski<sup>1</sup>

<sup>1</sup>Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA

<sup>2</sup>Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA

<sup>3</sup>The Natural History Museum, Cromwell Road, London SW7 5BD, UK

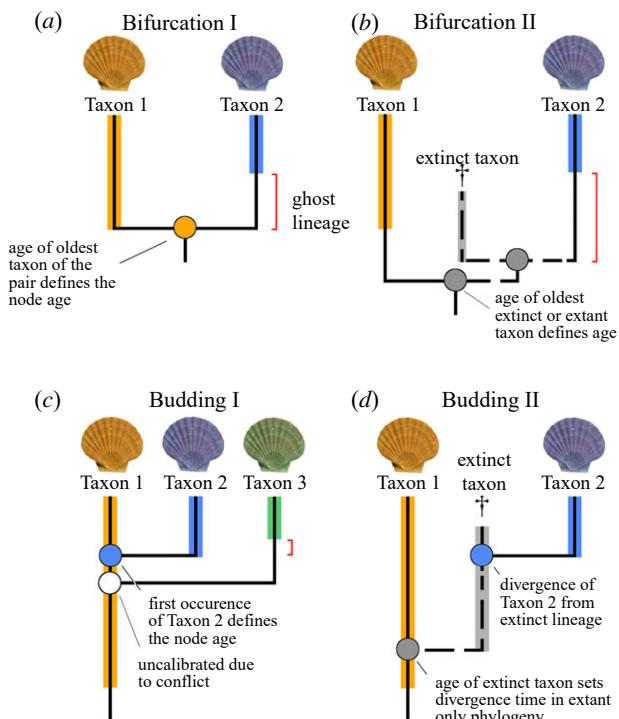
<sup>4</sup>Negaunee Integrative Research Center, Field Museum of Natural History, Chicago, IL 60605, USA

NMAC, 0000-0002-3504-8245; SME, 0000-0003-2843-7952; KSC, 0000-0002-3379-4201; RB, 0000-0002-9554-1947; DJ, 0000-0001-9163-8824

Analyses of evolutionary dynamics depend on how phylogenetic data are time-scaled. Most analyses of extant taxa assume a purely bifurcating model, where nodes are calibrated using the daughter lineage with the older first occurrence in the fossil record. This contrasts with budding, where nodes are calibrated using the younger first occurrence. Here, we use the extensive fossil record of bivalve molluscs for a large-scale evaluation of how branching models affect macroevolutionary analyses. We time-calibrated 91% of nodes, ranging in age from 2.59 to 485 Ma, in a phylogeny of 97 extant bivalve families. Allowing budding-based calibrations minimizes conflict between the tree and observed fossil record, and reduces the summed duration of inferred 'ghost lineages' from 6.76 billion years (Gyr; bifurcating model) to 1.00 Gyr (budding). Adding 31 extinct paraphyletic families raises ghost lineage totals to 7.86 Gyr (bifurcating) and 1.92 Gyr (budding), but incorporates more information to date divergences between lineages. Macroevolutionary analyses under a bifurcating model conflict with other palaeontological evidence on the magnitude of the end-Palaeozoic extinction, and strongly reduce Cenozoic diversification. Consideration of different branching models is essential when node-calibrating phylogenies, and for a major clade with a robust fossil record, a budding model appears more appropriate.

## 1. Introduction

The choice of branching model—whether lineages split through budding or bifurcation—can dramatically affect how macroevolutionary dynamics are inferred from phylogenies by changing estimated ages of lineage origination [1–7], and thus the inferred pace of evolution. In the commonly used bifurcating model, the daughter lineage with the older first occurrence in the fossil record constrains the minimum age of a given node (figure 1a). This tends to shift nodes deeper in time relative to a budding model, where the geologically younger of the two lineages dates the split between them. Palaeontological studies make different assumptions about lineage branching [9], allowing budding evolution when sister clades have overlapping stratigraphic ranges but first occur at different points in geologic time, the consequences of which are only now being addressed. Virtually all current comparative methods require a bifurcating topology as input with the topology dated accordingly, so understanding the effects of this assumption is essential. Here, we use a newly time-calibrated phylogeny for a model macroevolutionary system,



**Figure 1.** In a purely bifurcating model of evolution (a) the age of a node is determined by both daughter lineages, which may include extinct lineages (b). However, this model can create ghost lineages if the first recorded occurrence of one daughter is younger than the other (see also [8]). Assuming a budding model (c) removes this ghost lineage as the node age is determined from the youngest age of the younger lineage. If the stratigraphic range of a taxon is younger than a more derived lineage, this will create a ghost lineage whose minimum duration is the difference in age between the two taxa. In some instances (d), the node age would be at the split of the paraphyletic stem lineage to Taxon 2, if such extinct groups are included in analyses, as we do in some of the analyses presented here. The branches are highlighted by the temporal range of the fossil record of each lineage's fossil record. (Online version in colour.)

the molluscan class Bivalvia, chosen for its rich and well-documented fossil record, to evaluate how different models of lineage branching impact the interpretations of macroevolutionary dynamics.

Most phylogenetic inferences have implemented a bifurcating model in dating topologies. However, budding, as indicated by partially overlapping stratigraphic ranges of sister taxa and the persistence of genetic and/or morphological characteristics of the ancestral lineage following its split with a descendant, is also considered a likely evolutionary pattern [8] manifesting in analyses at the species level and for higher taxa [10–12]. Newer approaches to phylogenetic time calibration jointly estimate tree topology and branching times while allowing both bifurcating and budding processes to occur across the phylogeny, which increases the congruence between the fossil record and the tree topology. Relative to a budding model, bifurcation commonly creates more extensive 'ghost lineages'—the temporal gaps determined by the difference in first stratigraphic occurrences of sister clades [13,14] (figure 1). However, ghost lineage is not a perfect metric for estimating the fit of phylogenies to the fossil record, as taphonomic biases may mask the true first occurrences of lineages, misleadingly supporting a Budding I scenario over a more appropriate Budding II or Bifurcating I.

Such cases may result from true sampling failure, but become increasingly unlikely with greater quality of the fossil record of the target group.

An empirical approach to analysing the effects of assuming a bifurcating branching model on inferred macroevolutionary dynamics requires a system with both a taxonomically broad phylogenetic hypothesis and a rich and well-documented fossil record. Bivalves are exceptional for such analyses because their calcium carbonate shells and relatively high abundances have produced a dense fossil record. Imperfections and uncertainties in this record remain, but these issues are sufficiently well understood that biases can generally be factored out [15–17]. Despite their good fossil record, analysing bivalves comes with a cost: like many other major clades, they originated in the Cambrian explosion [18,19] and diversified during the Ordovician [17,20,21], and this extreme crown age makes resolution of higher relationships difficult. However, deeper nodes have recently been resolved with greater confidence using data from high-throughput sequencing, such as transcriptomes [22].

Here, we address the bifurcation-versus-budding issue using 86 vetted time-calibrations (91% of 95 internal nodes) across the family-level phylogeny of extant Bivalvia ( $n = 97$  families). We assign calibrations to nodes based on the stratigraphic ranges of their descendant taxa under individual and joint assumptions of budding and bifurcating models, and use paraphyletic stem groups to more closely capture the divergence times of the extant families. We predict that a budding model will date lineages such that their topological ages more closely match their observed stratigraphic ranges in the fossil record and will align more closely with previous studies documenting the diversity history of the marine biota. We use these calibrated phylogenies to compare different inferred macroevolutionary histories.

## 2. Material and methods

We first generated a family-level phylogenetic topology of 97 bivalve families using a combination of molecular and morphological studies, with branch lengths rescaled to geologic time using node calibrations spanning the taxonomic and temporal history of the group. We then use lineage-through-time (LTT) plots and Bayesian analyses of diversification rates to compare the effects of branching model process on inferred macroevolutionary dynamics of the class.

### (a) Phylogenetic topology

The transcriptome-based phylogeny of Lemer *et al.* [22] includes 98 ingroup bivalve species, with notably higher confidence in the topology of the deeper branches than in previous analyses ([23] and references therein). We chose one representative species per family from this phylogeny to create a family-level topology ( $n = 69$ ; see electronic supplementary material, text for details). To this backbone, we added 13 families whose positions had been assessed using the five-gene alignment of Combosch *et al.* [23], which was sufficient to resolve more recent divergence events. One family, Chamidae, is placed in a polytomy within order Venerida, due to the inability of transcriptomics [22] and other methods [24] to resolve its placement more finely. We then added five families whose molecular data have not been evaluated in the context of the entire clade (e.g. Neoleptonidae) and, finally, placed 10 families on the backbone that have not been analysed genetically but were evaluated from morphological affinities used in bivalve systematics (e.g. Pholadomyidae).

Placement of families based solely on morphological characteristics is feasible within bivalves given the strong congruence between molecular and morphological-based phylogenetics [24–26]. We provide full documentation of how families were placed in the electronic supplementary material, text and figure 3.

### (b) Time calibration

Time-scaling phylogenetic data assigns ages to tree nodes using observed fossil occurrences of lineages. The ages used here were from a database of first and last occurrences of all living and extinct bivalve genera used in previous analyses (updated version from that in [27–29]). Each age is derived from a published record of a bivalve taxon ( $n = 517$  sources). Electronic supplementary material, text provides the citations associated with each age, containing further information on taxonomy, geography and stratigraphic occurrences. Each node in the topology is assigned a minimum and maximum age, rather than a point value, to incorporate uncertainty in the age of that node (as required for the dating method described below). Here, we took a conservative approach, assigning age ranges as the upper and lower time points of the stratigraphic interval in which a taxon is first observed in the fossil record. These stratigraphic intervals vary in duration, but are all very small compared to the age of crown Bivalvia: the mean stratigraphic age bracket (i.e. the difference between a minimum and maximum age for a node) in our data is 4.1 Myr, with the crown age of Bivalvia being approximately 475 Ma (the largest age bracket is 18.9 Myr for a fossil occurrence unresolved within the Norian Stage). Thus, the uncertainty in any estimated node age is negligible compared to the crown age and almost all nodes within the phylogeny can be assigned an age prior (see below), indicating that most of the variation in inferred diversification times will come from the branching model under which the calibrations were derived.

We generated four calibration ages for each node based on different combinations of branching models and sampled taxa. These were: (i) bifurcating, with nodes dated by the oldest known member (base age) of each extant sister pair (Bifurcating I, standard procedure for calibrating molecular phylogenies; figure 1a), (ii) bifurcating, incorporating the ages of extinct paraphyletic lineages (Bifurcating II, figure 1b), (iii) budding among extant lineages, dating the node with the younger occurrence in the fossil record (Budding I, figure 1c) and (iv) budding incorporating the ages of extinct paraphyletic lineages (Budding II, figure 1d). Placement of extinct paraphyletic taxa interposed between extant lineages (as in figure 1b,d), which occur in the classification of many major groups having an extensive fossil record [30], is based largely on the classification provided by the revised *Treatise on Invertebrate Paleontology* [31] (for their approach to paraphyletic stem taxa see [32]; see electronic supplementary material, text for details). Phylogenetic positions of stem lineages were used to infer node ages (figure 1; electronic supplementary material, figure S1 and text), but these taxa were omitted from the final phylogeny of extant families, where branch lengths are time-scaled to reflect millions of years. Node ages ranged from the Cambrian stem (age range approx. 519–525 Ma) to the crown (age range approx. 483–485 Ma) to the most recently branched family (age range 2.59–3.60 Ma) (electronic supplementary material, dataset S1).

We time-scaled the phylogeny to convert branch lengths to geologic time using *treePL* [33], which uses penalized likelihood to calibrate topologies using minimum and maximum age constraints on node ages (see electronic supplementary material, text for discussion of alternative methods, i.e. fossilized birth-death range). We ran the *treePL* analysis specifying the *thorough* option to ensure the analysis continued until convergence was reached, and a random subsample and replicate cross-validation procedure (which calculates error in evolutionary rates and node ages across a random selection of terminal nodes [33]) to

determine the smoothing parameter (which assigns a penalty for rate variation across the phylogeny [33]) by specifying the *randomcv* option (final value = 1000). We repeated this procedure across the budding and bifurcating models defined in figure 1.

We also time-scaled the family topology using *cal3* [13,34], implemented in the R package *paleotree* [35]. *cal3* uses stochastic sampling of node ages under a birth–death-sampling model calibrated via the observed branching pattern, information on clade diversification and the probability of having sampled lineages in the fossil record. Thus, the method requires a cladogram describing the relationships between taxa and three parameters: origination rate, extinction rate and sampling rate. These can either be single values or lineage-specific. We used lineage-specific values for these parameters: *per capita* origination and extinction rates of families from Huang *et al.* [28] and Collins *et al.* [36], and sampling rates using the distribution of stratigraphic ranges of genera within families using the *make\_durationFreqCont* function in *paleotree*. Using these parameter values, we generated 100 time-scaled phylogenies using *cal3*. LTT plots of all calibrated phylogenies were calculated using the R [37] package *ape* [38].

### (c) Quantifying ghost lineages

We quantified durations of ghost lineages, apparent stratigraphic sampling gaps in the fossil record implied by phylogenetic relationships ([14] and references therein), depending on the assumed process of lineage origination. For bifurcation, ghost lineage is the difference in first fossil occurrence between two sister taxa (figure 1), and for budding, it is the temporal gap between the first fossil occurrence of a lineage and a topologically derived, yet geologically older, sister lineage (figure 1). For both branching models, the summed duration of ghost lineages across all nodes provides an indication of how closely the phylogeny matches the observed fossil record, although this value must be evaluated with caution if there is a strong prior expectation that taphonomic biases prevent the observed stratigraphic ranges from accurately capturing origination times of lineages. We provide ghost lineage values associated with every node in electronic supplementary material, dataset S1.

We also plotted the ghost lineage associated with each node in the tree against the node's age to assess temporal variation in the discordance between the phylogeny and fossil record. First, we quantified the cumulative sum of ghost lineage per node from the root of the phylogeny to the tips, including nodes subtending extinct families for Bifurcating II and Budding II models. Second, we quantified a normalized ghost lineage per 5 Myr interval, where the summed duration of ghost lineages was divided by the number of nodes in an interval to account for increasing sample size towards the present and for increased number of nodes in Bifurcating II and Budding II.

### (d) Simulation-based sensitivity testing

We used simulations to quantify the effect of two potential sources of error in our analyses: taxonomic uncertainty and preservation biases. Families that were placed using systematics (i.e. morphological affinities) may have been incorrectly placed, which might affect the time calibration of the phylogeny and assessment of ghost lineages. Therefore, we randomly shuffled the positions of these 10 families within their respective order or major clade (see electronic supplementary material, figure S1 for family assignments) to assess how time calibration of the phylogeny and summed ghost lineage totals are sensitive to their placement (incorrect placement of families is increasingly improbable above the order and major clade levels).

Preservation biases manifest as sampling failure which should change the shape of the LTT plots and the summed ghost lineage. If they are systematic (i.e. the preservation rates on lineages are drawn from the same distribution), preservation biases will shift

divergence estimates closer to the present, which would artificially increase inferred diversification rates. In terms of ghost lineage, such systematic biases should reduce absolute amounts, but the proportional differences between branching models should be constant. To test this, we simulated and sampled a fossil record of taxa at four frequencies to quantify whether a single, clade-wide sampling rate would systematically alter summed ghost lineage values. In reality, sampling is not likely to be approximately constant across a clade with diverse preservation potentials. Thus, as a first step towards understanding this bias and in addition to the simulation work, we evaluated the relative preservation potential of the two taxa at each node in our empirical phylogeny using the prevailing shell mineralogy of each family. While shelled animals have generally high preservation potential, those with aragonite shells are more susceptible to taphonomic loss than calcitic or bimimetic ones [16,39,40]. Thus, we would be less confident in a budding hypothesis at a node if the apparently younger taxon had poorer preservation potential (i.e. was aragonitic) than the apparently older one (i.e. was bimimetic or calcitic).

### (e) Quantifying macroevolutionary dynamics

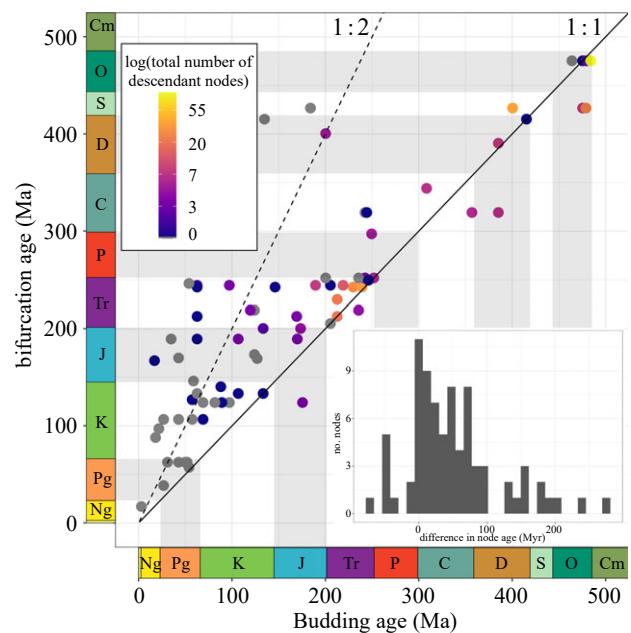
The effects of time calibration using different branching models on macroevolutionary dynamics were evaluated, first, as qualitative comparisons of their inferred diversification histories from resulting LTT plots, and second, as the quantitative differences in the timing of shifts in diversification rate. These shifts were detected using the CoMET model in the R package TESS [41], with empirical hyperpriors for the origination and extinction rate parameters calculated by short MCMC analyses of the data run under a constant-rate birth-death process. Analyses were run for 10 000 generations, discarding the first 20% as burnin.

## 3. Results

### (a) Comparison of calibrated phylogenies

Different branching models produced dramatically different age calibrations (figure 2). Differences in node ages between Bifurcating I (the most commonly implemented method of calibrating phylogenies) and Budding II (incorporating the most information from the fossil record; see figure 3) range from –66 to 280 Myr (figure 2). The distribution of differences is right-skewed (median 38 Myr, mean 51 Myr), with 34% of the nodes differing by greater than 65 Myr (equivalent to the duration of the entire Cenozoic Era), and with 21 nodes more than doubling in age from Budding II to Bifurcating I (figure 2). Nodes are younger when calibrated in Bifurcating I compared to Budding II in 10 instances, owing to the use of paraphyletic stems to date nodes in the Budding II model (electronic supplementary material). For example, the age of the order Cardiida increases under both Bifurcating II and Budding II models relative to Bifurcating I when including the extinct family Palaeocarditidae (electronic supplementary material, figure S1). Because of the differences in node ages, the Bifurcating I and Budding II also give dramatically different branch lengths (figure 4). Although these differences occur throughout the phylogeny, the biggest differences are concentrated in the Euheterodontia (=Anomalodesmata + Imparidentia) at nodes subtending relatively few families (figure 4).

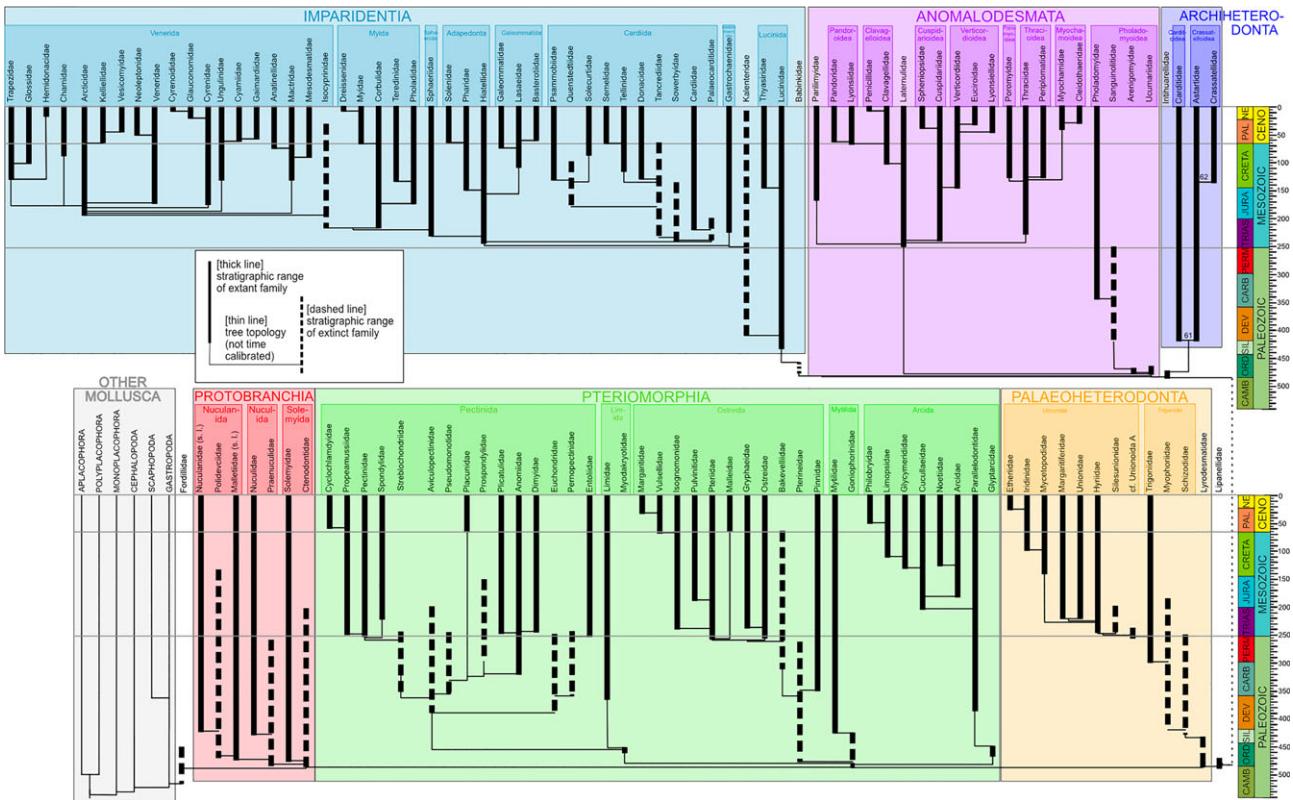
Nine nodes in the phylogeny are uncalibrated (figure 4) due to inconsistencies between the topology and the fossil record. Five of these conflicts are in a single sequence of nodes within the order Venerida, primarily because the fossil record of the Arcticidae reportedly extends far deeper than its sister families



**Figure 2.** Comparison of node ages when assuming Budding II (including extinct paraphyletic lineages) or Bifurcation I (excluding paraphyletic lineages). Points are coloured by the number of nodes between the target node and the tips of the phylogeny. Grey nodes are those subtending only terminal branches. The 1:1 line shows where the bifurcation and budding ages are equal, and the 1:2 line shows where the bifurcation age is twice that of the budding age. Inset shows the distribution of differences between Bifurcation I and Budding II ages (residuals to the 1:1 line). (Online version in colour.)

(electronic supplementary material, text). The currently monospecific Arcticidae have a high preservation probability because of their large, robust shells, and were taxonomically, morphologically and geographically far more extensive in the past than today. Another conflict is at the split between the orders Adapedonta and Galeommatida. This conflict probably results from the poor fossil record of the Galeommatida, owing to their small size, fragile shells and frequent commensalism [16]. Taxonomic challenges within the group also hinder the association of older fossil material with the family. The remaining two conflicts are within the initial diversification of the group and involve relatively modest uncertainties deriving from the sparse Early Ordovician fossil record: one of approximately 10 Myr (node 79), and one involving an Early Tremadocian/Late Tremadocian discrepancy (approx. 5 Myr; node 91).

The summed duration of ghost lineages is remarkably higher in the bifurcating model, with 6.76 Gyr in Bifurcating I and 7.86 Gyr in Bifurcating II versus 1.00 Gyr in Budding I and 1.92 Gyr in Budding II (Budding II estimates drop to 1.86 Gyr if Konduriidae should be interposed between Babinkidae and Kalenteridae and 1.875 Gyr if Kalenteridae range to the Late Ordovician; see electronic supplementary material, text and figure S2A). Budding I and Budding II give identical values wherever extant families are linked without an intervening extinct family (e.g. Kelliellidae and Vesicomyidae, electronic supplementary material, figure S1), or if the stratigraphic ranges of sisters overlap (e.g. Nuculanidae and Mallettiidae, electronic supplementary material, figure S1). When stems are interposed between extant families in the 'II' models, ghost lineages increase, apparently because they create more opportunities for conflicts and gaps in the fossil record, particularly in the less completely preserved Early Palaeozoic portions of bivalve history (electronic supplementary material, figure S1 and



**Figure 3.** Budding II model for bivalve phylogeny, incorporating selected extinct families; see electronic supplementary material, figure S1 and text, for details and the other three models. (Online version in colour.)

dataset S1). Bifurcating models have a more rapid increase in ghost lineage totals through time than budding models, especially following the end-Permian mass extinction (electronic supplementary material, figure S2A). Accounting for the different number of nodes in the different models, the pattern is largely unchanged, but with greater disparity between the Bifurcating I and Bifurcating II models (electronic supplementary material, figure S2B), suggesting that the topology better aligns with a budding model.

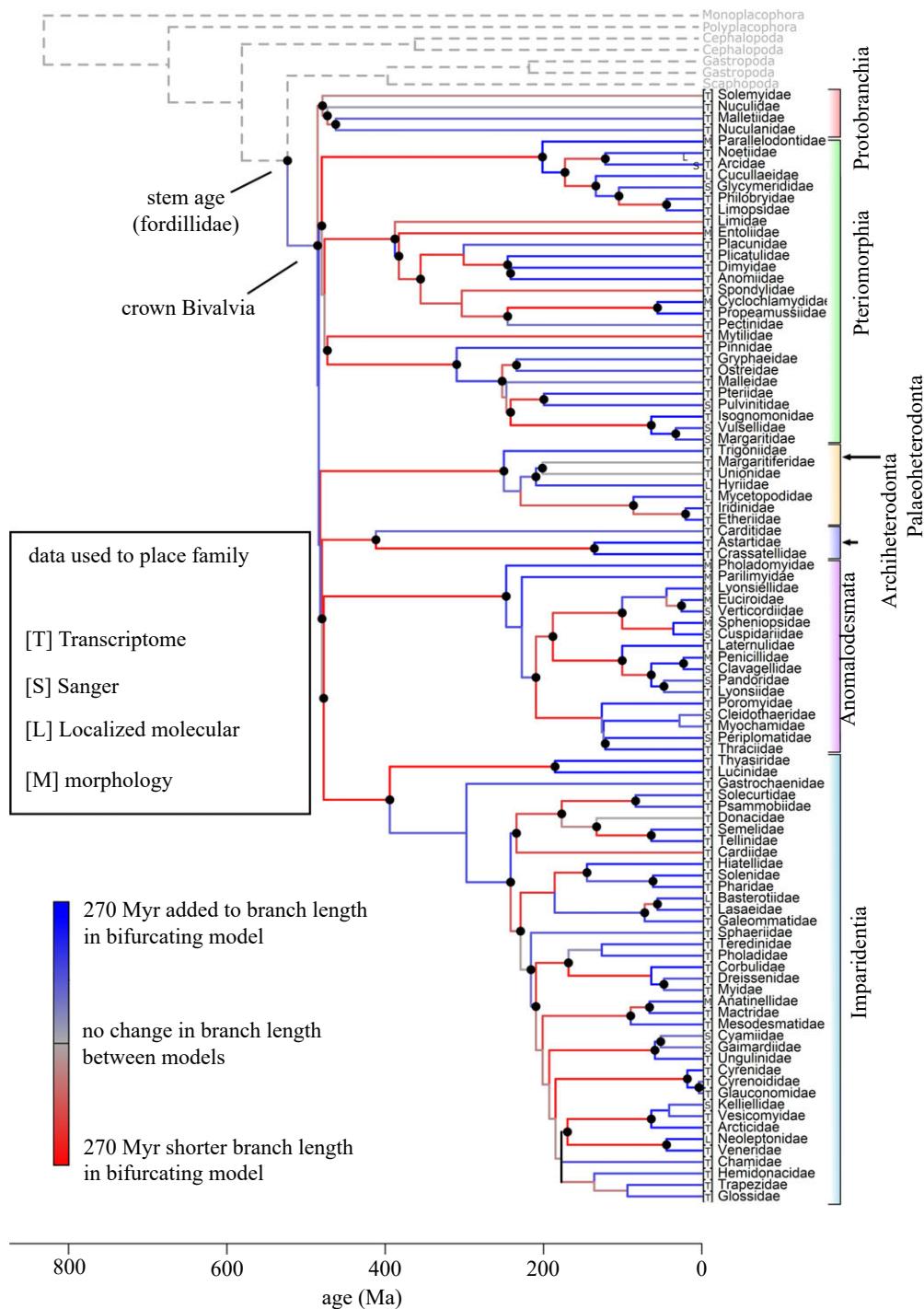
The LTT plot for the cal3 model mostly fell between the bifurcating and budding lines (figure 5). However, through the Mesozoic and Cenozoic the LTT trajectory generally follows the shape of both budding curves. It falls above these curves due to a greater amount of inferred bifurcation in the Palaeozoic; in particular, the cal3 analysis pulls more of the Ostreida deep into the Palaeozoic. This adjustment may reflect an apparently genuine acceleration in origination and extinction rates—which enter into sampling estimates—that occurs in the Early and Mid-Mesozoic when oysters and their relatives repeatedly invaded soft sediments [42].

### (b) Effect of branching model on inferred macroevolutionary history

In the LTT plots, Bifurcating II and Budding II show a rapid upward shift immediately following the appearance of crown bivalves in the earliest Ordovician (figure 5). This pulse is caused by extremely short internode distances between the major lineages or their directly associated stem lineages, reflecting pronounced diversification during the Ordovician radiation (as long recognized from fossil material of extant and extinct groups; e.g. [20,21,43]). Following this initial radiation, the LTT plots derived from bifurcation dates are

consistently above the lines derived from budding dates, as the nodes are consistently older, and thus diversity accrues earlier in the clade's history. Incorporating the ages of extinct families (Bifurcating II and Budding II) increases the number of lineages prior to the end-Permian mass extinction. All plots show upticks in diversity following the end-Permian and end-Cretaceous mass extinctions, but the pace of the longer-term diversification (i.e. the slope of the LTT line) differs between the two branching models. Specifically, the bifurcating models exaggerate the diversification following the end-Permian extinction and diminish the one following the end-Cretaceous relative to the fossil record of extant and extinct bivalve genera (e.g. [17,44]). In such cases, we believe the fossil record accurately captures historical processes because bivalve taxa from a range of preservation categories are present through these intervals (e.g. participation by both more resistant calcitic and more dissolution-prone aragonitic forms [39], as well as large-bodied and small-bodied forms), suggesting the observed pattern to be comparatively free of taphonomic bias. Nevertheless, the phylogenetic topology can provide an additional check on the quality of the fossil record and/or our phylogenetic understanding of extinct taxa—for example, the pronounced stratigraphic conflict around the Pholadomyidae warrants further investigation.

LTT plots are minimally affected by uncertainty in the placement of extant families that lacked molecular data (electronic supplementary material, figure S3). Both Budding I and Bifurcating I showed high congruence between the empirical and simulated phylogenies, with the Bifurcating I simulations showing greater variance in the Palaeozoic. However, simulations for Budding II and Bifurcating II, which incorporated uncertainty in the placement of the stem families, showed much lower congruence with the

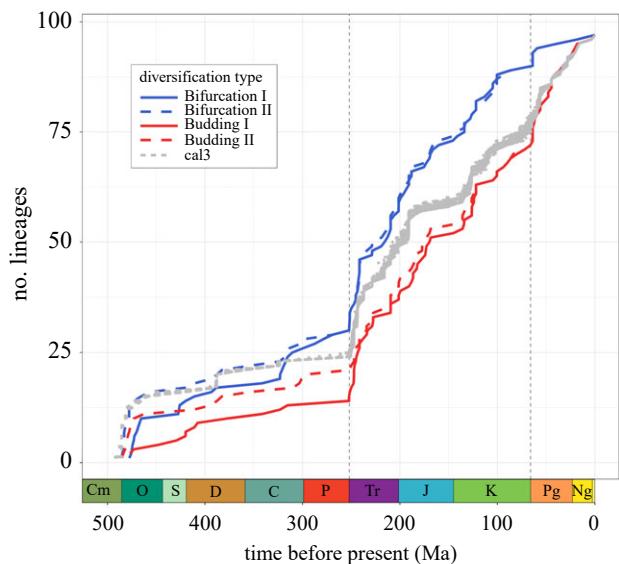


**Figure 4.** Time-scaled family-level bivalve phylogeny dated assuming a budding model of evolution (with nodes scaled to include paraphyletic stem families, Budding II) where branch colours show the difference in branch length (in Myr) relative to the tree when scaled assuming a purely bifurcating model of evolution (excluding paraphyletic lineages, Bifurcating I). Blue branches are those lengthened by the older ages of the parent node. Absolute values of branch lengths were used in log-transforming the branch lengths for the colour scale. Highlighted nodes are those used to calibrate the phylogeny ( $n = 86$ ). (Online version in colour.)

empirical data; simulations fall consistently above the empirical line. This was a consequence of including extinct families hypothesized to lie phylogenetically between extant ones; these extinct families tend to be concentrated in the Palaeozoic, significantly closer to the root (electronic supplementary material, figure S4). These families are much more likely to fall higher on the tree if their placement is randomized and thus, because of their age, pull more lineages deeper in time, causing the simulated LTT lines to accumulate diversity earlier than the empirical analyses. The Budding II model was slightly less affected by this bias, showing lower variance in simulations and more congruence to the empirical data than the Bifurcating II model. Protecting the monophly

of extant orders caused the simulated plots to be slightly more coincident with the empirical data, again with less variance under the Budding model.

Randomizing the location of families placed using morphological data still resulted in a bifurcating branching model having higher summed ghost lineage (electronic supplementary material, figure S5). None of the randomizations for the budding model exceeded, or came close to, the observed empirical value. Additionally, the variance in summed ghost lineage was lower for the budding branching model due to its inherent minimization of ghost lineage. The summed ghost lineage values from the randomizations were higher than the empirical value for both models, suggesting



**Figure 5.** Lineage-through-time plots generated from phylogenetic data, excluding outgroups, scaled assuming budding and bifurcating models of evolution (figure 1) plus 100 estimates from cal3. (Online version in colour.)

that these families are likely placed correctly, or at least close to their true location, within the phylogeny.

Phylogenies calibrated using dates inferred from a budding model produced origination rate estimates from the CoMET analysis more consistent with other analyses of the fossil record (electronic supplementary material, figure S6). The two budding topologies produced nearly identical patterns with a largely constant rate through the Palaeozoic, a sharp increase following the end-Permian mass extinction, and a second shift in the Late Jurassic indicating a decrease in the origination rate following recovery from the end-Permian extinction. Recovery from the end-Cretaceous mass extinction is only present as a small upward inflection in the 95% confidence interval. Conversely, the phylogenies calibrated assuming a bifurcating model show rate decreases towards the present in a stepwise manner, with no obvious signal of recovery from any mass extinction events; a result in sharp contrast to other analyses of the fossil record [16,17].

## 4. Discussion

### (a) Comparing calibrations to the fossil record

Using a budding model to derive calibrations improves two potential measures of alignment to the observed fossil record: stratigraphic conflicts and summed duration of ghost lineages. We identified only nine instances of modest conflict—where the fossil record of the putative daughter lineage extended deeper in time than its parent branch—on the time scale of the phylogeny using either of our two budding models (11 for bifurcating models), indicating very high consistency between the fossil record and the hypothesized topology. Conflicts may result from many sources, such as omission of stem lineages (which also probably accounts for some of the stratigraphic gaps, e.g. between *Intihuarellidae* and the extant Archiheterodonta), uncertainty in the molecular topology, incomplete preservation of one lineage (e.g. poor preservation potential: small and fragile, as noted above, and/or restriction to the less commonly preserved deep-sea and freshwater environments [15,16]). All of these

are testable, in any clade. For example, taphonomic controls [3,45,46] can assess the credibility of a family's age; the likelihood that a family's earliest fossil has been sampled increases if other lineages with similar preservation potential are recorded in the preceding geologic time interval.

The summed duration of ghost lineages—the inferred evolutionary history missing from the fossil record—also suggests that the budding model is better aligned with the palaeontological data (1.92 Gyr for Budding II versus 7.86 Gyr for Bifurcating II). Bifurcation assumes that lineages diverge simultaneously from a shared ancestor, almost always resulting in a stratigraphic gap leading to the daughter lineage with the younger stratigraphic range (figure 1; top two panels in electronic supplementary material, figure S1). A budding model derives the daughter lineage from the parent lineage (or through anagenesis), and thus can more directly reflect the fossil record of the study system. It is worth noting that summed ghost lineages for a bifurcating model will be accentuated by preservation biases if the biases are asymmetric between sister lineages (i.e. one lineage has a much more complete record than its sister), but not if these biases are random across the tree (as shown in the simulations in electronic supplementary material, figure S7). We doubt asymmetric biases are pervasive across the bivalve phylogeny as we only identified one node where the two families used to date the node had differing shell mineralogy (electronic supplementary material, figure S8), which is one determinant of bivalve preservation potential [15,16,39]. Therefore, while our data may be less susceptible to one source of bias, pronounced ghost lineage totals remain for the budding model. The branch-specific ghost lineage values provided here, along with the identified conflicts, give additional insights into the incompleteness of sampling and/or taxonomic understanding of the bivalve fossil record, helping to pinpoint parts of the phylogeny needing closer attention (for example, around the *Intihuarellidae*).

Including extinct families in either branching model increases ghost lineage totals. Although greater taxon sampling could be expected to increase overall fit to the fossil record, paraphyletic stem lineages are mainly recorded in older, less complete parts of the fossil record, creating stratigraphic gaps. Extinct families are likely to fill or reduce those gaps as phylogenetic analyses continue for some of the more ancient parts of the bivalve record. With continued molecular sampling and increased resolution of deep splits, and further evaluation of extinct families, genera and species, we expect that the total sum of ghost lineage branch lengths will decrease as current gaps in the stratigraphic record are narrowed and topological conflicts are reduced.

The results shown here suggest that bifurcation should not be the default evolutionary model for time-calibrating trees; however, a budding model also has its limitations. First, the quality of the fossil record varies across the Tree of Life—some major lineages require exceptional preservation to enter the fossil record at all. Such groups may only have material to date just one daughter lineage in a pair, and at just a few nodes in a molecular phylogeny. Second, incomplete preservation can make divergence dates considerably younger than their true age (e.g. lineages in poorly or patchily recorded environments), indicating the need for future work assessing ages using taphonomic controls. This may be the case for freshwater bivalves [15], producing the ghost lineage associated with *Mycetopodidae*

and other freshwater unionids, and the Sphaeriidae. Finally, phylogenetic uncertainty, either in molecular topologies or assignment of fossil genera to families, can contribute to incorrect or problematic divergence dates (e.g. the stubbornly uncertain position of the Chamidae). Considering these factors when incorporating budding models into analyses of other clades should provide more accurate inferences of evolutionary dynamics.

Here, we have used simple metrics such as summed ghost lineage to evaluate budding as an appropriate branching model, as has been suggested for other clades [13]. Budding is most frequently discussed in the palaeontological literature, where the persistence of a parent lineage can be readily observed following the derivation of a new species [5]. However, budding is also frequently inferred in neontological studies, although not by this name. For example, any form of speciation in which the parent population does not immediately go extinct involves budding speciation. It is worth noting that the amount of budding speciation probably varies between clades [9], and its inference can depend on the type of data used in phylogenetic reconstructions [12]. Nevertheless, the continued development of methods that allow lineages to originate by budding will be valuable for a wide range of future studies.

### (b) Calibrations and macroevolutionary dynamics

Budding calibrations produce macroevolutionary patterns that not only reduce conflicts and ghost lineages, but are more concordant with the history of metazoan biodiversity. Including stem lineages in the evolutionary models indicates that the major branches of bivalve evolution were established as part of the well-known Ordovician radiations around 480 Ma [18,20,21]. Excluding those stem families shifts the age of the nodes approximately 50 Myr later, an offset nearly equivalent to the duration of the Cenozoic (figure 5). On the other hand, the bifurcating calibrations artificially pull lineages deeper in time, inflating the number of extant lineages present in the Late Palaeozoic and Mesozoic (see strong deviations from 1:1 line in figure 2). This bifurcation-based shift has dramatic consequences for the inferred diversification through the Cenozoic, pulling the origination ages of 14 families and seven lineages leading to pairs of sister taxa back into the Mesozoic. In all of these instances, the budding model yields a diversity history for bivalve families that is more consistent with other lines of palaeontological evidence, such as local and regional time series documenting extinction events and diversifications [44,47,48], as did CoMET analyses of phylogenetic data. More generally, although a budding model can still be subject to the same taphonomic biases as a bifurcating model, our

results show how bifurcating models can exhibit a directional bias on the evolutionary history of the clade by shifting diversification earlier in time.

## 5. Conclusion

Quantifying the timing of lineage branching is the basis for many evolutionary analyses. However, most studies assume a pure bifurcating model of evolution since this is the structure currently constraining molecular topologies. By analysing a study system with an exemplary fossil record under two branching models, we show that a bifurcating model pushes bivalve diversity considerably deeper in time, reducing the magnitude of the group's response to mass extinction events and making Cenozoic diversification less pronounced. Relaxing this assumption to allow budding appears to better align with the observed fossil record of bivalves by diminishing summed duration of ghost lineages, and produces macroevolutionary dynamics more consistent with local and regional analyses of the fossil record. When fossil data are too sparse to distinguish between budding and bifurcating calibration models, understanding the biases that calibration can have on downstream evolutionary analyses becomes a critical step—our results show that bifurcation tends to lengthen terminal branches relative to budding, thereby reducing the magnitude of later diversification. Thus, the potential consequences of different evolutionary branching models for calibrating phylogenetic data, and suggest that budding models, a common pattern in palaeontological data, should not be ruled out *a priori* in many analyses.

**Data accessibility.** All data and code necessary to reproduce analyses and figures are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.1vhmgqsd> [49]. Electronic supplementary material provides details on constructing and time calibrating the phylogeny.

**Authors' contributions.** N.M.A.C.: conceptualization, formal analysis, visualization, writing—original draft, writing—review and editing; S.M.E.: conceptualization, data curation, validation, writing—review and editing; K.S.C.: conceptualization, data curation, validation, writing—review and editing; R.B.: funding acquisition, investigation, validation, writing—review and editing; D.J.: conceptualization, data curation, funding acquisition, validation, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Competing interests.** We declare we have no conflicts of interest.

**Funding.** This work was supported by grants from the National Science Foundation (grant nos. EAR-0922156 and DEB-2049627 to D.J., DEB-0732854 to R.B.) and the National Aeronautics and Space Administration (grant no. EXOB08-0089 to D.J.).

**Acknowledgements.** We thank N. J. Morris, J. D. Taylor, T. R. Waller, T. E. Yancey and three anonymous reviewers for their insightful advice and discussion, and M. Terezow and N. Hudson for providing images of type material for comparison. Final decisions on calibration dates are solely the responsibility of the authors.

## References

1. Benton MJ, Pearson PN. 2001 Speciation in the fossil record. *Trends Ecol. Evol.* **16**, 405–411. (doi:10.1016/s0169-5347(01)02149-8)
2. Huang D, Goldberg EE, Roy K. 2015 Fossils, phylogenies, and the challenge of preserving evolutionary history in the face of anthropogenic extinctions. *Proc. Natl. Acad. Sci. USA* **112**, 4909–4914. (doi:10.1073/pnas.1409886112)
3. Marshall CR. 2019 Using the fossil record to evaluate timetree timescales. *Front. Genet.* **10**, 1049. (doi:10.3389/fgene.2019.01049)
4. Wagner PJ. 2019 On the probabilities of branch durations and stratigraphic gaps in phylogenies of fossil taxa when rates of diversification and sampling vary over time. *Paleobiology* **45**, 30–55. (doi:10.1017/pab.2018.35)
5. Foote M. 1996 On the probability of ancestors in the fossil record. *Paleobiology* **22**, 141–151. (doi:10.1017/S0094837300016146)
6. Stadler T, Gavryushkina A, Warnock RCM, Drummond AJ, Heath TA. 2018 The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes. *J. Theor. Biol.* **447**, 41–55. (doi:10.1016/j.jtbi.2018.03.005)

7. Bapst DW. 2014 Assessing the effect of time-scaling methods on phylogeny-based analyses in the fossil record. *Paleobiology* **40**, 331–351. (doi:10.1666/13033)

8. Wagner PJ. 2000 The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diversity. *Syst. Biol.* **49**, 65–86. (doi:10.1080/10635150050207393)

9. Silvestro D, Warnock RCM, Gavryushkina A, Stadler T. 2018. Closing the gap between palaeontological and neontological speciation and extinction rate estimates. *Nat. Commun.* **9**, 5237. (doi:10.1038/s41467-018-07622-y)

10. Smith AB. 1994 *Systematics and the fossil record: documenting evolutionary patterns*, pp. 130–133. Oxford, UK: Blackwell Scientific.

11. Wagner PJ, Erwin DH. 1995 Phylogenetic tests of speciation hypotheses. In *New approaches to studying speciation in the fossil record* (eds DH Erwin, RL Anstey), pp. 87–122. New York, NY: Columbia University Press.

12. Parins-Fukuchi C. 2021 Morphological and phylogeographic evidence for budding speciation: an example in hominins. *Biol. Lett.* **17**, 20200754. (doi:10.1098/rsbl.2020.0754)

13. Bapst DW, Hopkins MJ. 2017 Comparing cal3 and other a posteriori time-scaling approaches in a case study with the pterocephaliid trilobites. *Paleobiology* **43**, 49–67. (doi:10.1017/pab.2016.34)

14. Donoghue PCJ, Yang Z. 2016 The evolution of methods for establishing evolutionary timescales. *Phil. Trans. R. Soc. B* **371**, 20160020. (doi:10.1098/rstb.2016.0020)

15. Harper EM. 1998 The fossil record of bivalve molluscs. In *The adequacy of the fossil record* (eds SK Donovan, CRC Paul), pp. 243–267. Chichester, UK: Wiley.

16. Valentine JW, Jablonski D, Kidwell S, Roy K. 2006 Assessing the fidelity of the fossil record by using marine bivalves. *Proc. Natl Acad. Sci. USA* **103**, 6599–6604. (doi:10.1073/pnas.0601264103)

17. Mondal S, Harries PJ. 2016 The effect of taxonomic corrections on Phanerozoic generic richness trends in marine bivalves with a discussion on the clade's overall history. *Paleobiology* **42**, 157–171. (doi:10.1017/pab.2015.35)

18. Fang Z-J, Sánchez T. 2012 Part N, revised, volume 1, chapter 16: origin and early evolution of the Bivalvia. *Treatise Online* **43**, 1–21. (doi:10.17161/to.v0i0.4275)

19. Zhuravlev AY, Wood RA. 2018 The two phases of the Cambrian explosion. *Sci. Rep.* **8**, 16656. (doi:10.1038/s41598-018-34962-y)

20. Cope JCW. 2002 Diversification and biogeography of bivalves during the Ordovician Period. *Geol. Soc. Lond. Spec. Publ.* **194**, 25–52. (doi:10.1144/GSL.SP.2002.194.01.04)

21. Miller AI. 2012 The Ordovician radiation: macroevolutionary crossroads of the Phanerozoic. In *Earth and life* (ed. JA Talent), pp. 381–394. Dordrecht, The Netherlands: Springer.

22. Lemer S, Bieler R, Giribet G. 2019 Resolving the relationships of clams and cockles: dense transcriptome sampling drastically improves the bivalve tree of life. *Proc. R. Soc. B* **286**, 20182684. (doi:10.1098/rspb.2018.2684)

23. Combosch DJ *et al.* 2017 A family-level Tree of Life for bivalves based on Sanger-sequencing approach. *Mol. Phylogenet. Evol.* **107**, 191–208. (doi:10.1016/j.ympev.2016.11.003)

24. Bieler R *et al.* 2014 Investigating the Bivalve Tree of Life—an exemplar-based approach combining molecular and novel morphological characters. *Invertebr. Syst.* **28**, 32–115. (doi:10.1071/is13010)

25. Jablonski D, Finarelli JA. 2009 Congruence of morphologically-defined genera within molecular phylogenies. *Proc. Natl Acad. Sci. USA* **106**, 8262–8266. (doi:10.1073/pnas.0902973106)

26. Sharma PP *et al.* 2012 Phylogenetic analysis of four nuclear protein-encoding genes largely corroborates the traditional classification of Bivalvia (Mollusca). *Mol. Phylogenet. Evol.* **65**, 64–74. (doi:10.1016/j.ympev.2012.05.025)

27. Roy K, Hunt G, Jablonski D. 2009 Phylogenetic conservatism of extinctions in marine bivalves. *Science* **325**, 733–737. (doi:10.1126/science.1173073)

28. 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**, 530–540. (doi:10.1111/geb.12153)

29. Edie SM, Jablonski D, Valentine JW. 2018 Contrasting responses of functional diversity to major losses in taxonomic diversity. *Proc. Natl Acad. Sci. USA* **115**, 732–737. (doi:10.1073/pnas.1717636115)

30. Budd GE, Mann RP. 2020 The dynamics of stem and crown groups. *Sci. Adv.* **6**, eaaz1626. (doi:10.1126/sciadv.aaz1626)

31. Carter JG *et al.* 2011 A synoptical classification of the Bivalvia (Mollusca). *Univ. Kans. Paleontol. Contrib.* **4**, 1–47. (doi:10.17161/PC.1808.8287)

32. Carter JG *et al.* 2015 The paracladistic approach to phylogenetic taxonomy. *Paleontol. Contrib.* **12**, 1–9. (doi:10.17161/PC.1808.17551)

33. Smith SA, O'Meara BC. 2012 treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690. (doi:10.1093/bioinformatics/bts492)

34. Bapst DW. 2013 A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. *Methods Ecol. Evol.* **4**, 724–733. (doi:10.1111/2041-210X.12081)

35. Bapst DW. 2012 paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods Ecol. Evol.* **3**, 803–807. (doi:10.1111/j.2041-210X.2012.00223.x)

36. Collins KS, Edie SM, Hunt G, Roy K, Jablonski D. 2018 Extinction risk in extant marine species integrating palaeontological and biodistributional data. *Proc. R. Soc. B* **285**, 20181698. (doi:10.1098/rspb.2018.1698)

37. R Core Team. 2019 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.

38. Paradis E, Schliep K. 2018 ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528. (doi:10.1093/bioinformatics/bty633)

39. Kidwell SM. 2005 Shell composition has no net impact on large-scale evolutionary patterns in mollusks. *Science* **307**, 914–917. (doi:10.1126/science.1106654)

40. Foote M, Crampton JS, Beu AG, Nelson CS. 2015 Aragonite bias, and lack of bias, in the fossil record: lithological, environmental, and ecological controls. *Paleobiology* **41**, 245–265. (doi:10.1017/pab.2014.16)

41. Höhna S. 2013 Fast simulation of reconstructed phylogenies under global time-dependent death–death processes. *Bioinformatics* **29**, 1367–1374. (doi:10.1093/bioinformatics/btt153)

42. Seilacher A. 1984 Constructional morphology of bivalves: evolutionary pathways in primary versus secondary soft-bottom dwellers. *Palaeontology* **27**, 207–237.

43. Pojeta Jr J. 1978 The origin and early taxonomic diversification of pelecypods. *Phil. Trans. R. Soc. B* **284**, 225–246. (doi:10.1098/rstb.1978.0065)

44. Ros S, De Renzi M, Damborenea SE, Márquez-Aliaga A. 2011 Coping between crises: early Triassic–early Jurassic bivalve diversity dynamics. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **311**, 184–199. (doi:10.1016/j.palaeo.2011.08.020)

45. Bottjer DJ, Jablonski D. 1988 Paleoenvironmental patterns in the evolution of post-Paleozoic benthic marine invertebrates. *Palaios* **3**, 540–560. (doi:10.2307/3514444)

46. Kidwell SM, Holland SM. 2002 The quality of the fossil record: implications for evolutionary analyses. *Annu. Rev. Ecol. Syst.* **33**, 561–588. (doi:10.1146/annurev.ecolsys.33.030602.152151)

47. Aberhan M, Kiessling W. 2015 Persistent ecological shifts in marine molluscan assemblages across the end-Cretaceous mass extinction. *Proc. Natl Acad. Sci. USA* **112**, 7207–7212. (doi:10.1073/pnas.1422248112)

48. Sepkoski Jr JJ. 1996 Patterns of Phanerozoic extinction: a perspective from global data bases. In *Global events and event stratigraphy in the Phanerozoic* (ed. OH Walliser), pp. 35–51. Berlin, Germany: Springer.

49. Crouch NMA, Edie SM, Collins KS, Bieler R, Jablonski D. 2021 Calibrating phylogenies assuming bifurcation or budding alters inferred macroevolutionary dynamics in a densely sampled phylogeny of bivalve families. Dryad Digital Repository. (doi:10.5061/dryad.1vhmgqsd)