

Two Notorious Nodes: A Critical Examination of Relaxed Molecular Clock Age Estimates of the Bilaterian Animals and Placental Mammals

GRAHAM E. BUDD^{1,*} AND RICHARD P. MANN²

¹Department of Earth Sciences, Palaeobiology Programme, Uppsala University, Villavägen 16 SE 75236, Sweden, and

²Department of Statistics, School of Mathematics, University of Leeds, Leeds LS2 9JT, UK

*Correspondence to be sent to: Department of Earth Sciences, Palaeobiology Programme, Uppsala University, Villavägen 16 SE 75236, Sweden;

E-mail: graham.budd@pal.uu.se.

Received 3 February 2023; reviews returned 30 August 2023; accepted 8 September 2023

Associate Editor: Jeremy Beaulieu

Abstract.—The popularity of relaxed clock Bayesian inference of clade origin timings has generated several recent publications with focal results considerably older than the fossils of the clades in question. Here, we critically examine two such clades: the animals (with a focus on the bilaterians) and the mammals (with a focus on the placentals). Each example displays a set of characteristic pathologies which, although much commented on, are rarely corrected for. We conclude that in neither case does the molecular clock analysis provide any evidence for an origin of the clade deeper than what is suggested by the fossil record. In addition, both these clades have other features (including, in the case of the placental mammals, proximity to a large mass extinction) that allow us to generate precise expectations of the timings of their origins. Thus, in these instances, the fossil record can provide a powerful test of molecular clock methodology, and why it goes astray, and we have every reason to think these problems are general. [Cambrian explosion; mammalian evolution; molecular clocks.]

The apparent clash between molecular clock estimates of the time of origins of clades versus what the fossil record might be thought to say has been a significant source of anxiety in the last two decades (e.g. [dos Reis and Yang 2013](#); [Bromham et al. 2018](#); [Bromham 2019](#); [Klopfstein 2021](#)). Nevertheless, the fossil record clearly shows a significant structure that would be difficult to reconcile with such *general* unreliability ([Benton et al. 2000, 2009](#)). In addition, our ability to map rapid post-extinction radiations ([Budd and Jensen 2003](#)) such as at least ordinal-level placental mammals strongly suggests that, in some cases, the fossil record must be giving us a more or less accurate picture of temporal evolution.

Molecular clock estimates, which have sometimes been argued to have essentially taken over from the fossil record as the ultimate measures of evolutionary time ([Warnock et al. 2012](#)), conversely, often give much larger gaps between first appearances of clades and the fossil record. To give some recent examples of the estimated gaps between the first recognized fossil of a clade and the 95% credible interval of molecular clock estimates: beetles (55–134 myr older than the oldest fossil ([Cai et al. 2022](#))); sand dollar and sea biscuit echinoids (40–65 myr ([Mongiardino Koch et al. 2022](#))); cheilostome bryozoans (132–238 myr ([Orr et al. 2022](#))); angiosperms (23–121 myr ([Barba-Montoya et al. 2018](#))); and bilaterians (50–143 myr ([dos Reis et al. 2015](#))). Thus, one ends up in the curious situation where the fossil record is taken seriously in providing a prior for the ages of nodes, but in retrospect is discovered to be wildly unreliable. The possible reasons for these sorts of large gaps have been much discussed, but here we wish to focus on how calibration

ranges are converted to priors, and the influence that those priors then have on the outcome. Whilst these issues have been the subject of considerable discussion in the literature ([Warnock et al. 2015](#); [Barba-Montoya et al. 2017](#)), we do not believe that these discussions have been followed through to their logical conclusion, which, as we argue here, is that relaxed clocks do not provide good evidence for large gaps between clade origins and their fossil record. As a focal point for our investigation, we take two rather notorious nodes: the origin of the animals ([dos Reis et al. 2015](#)), and the place of the bilaterians within them, and the origin of the placental mammals ([Álvarez-Carretero et al. 2022](#)); both of which have generated considerable controversy. We have chosen these two studies, first because they are of widespread interest in their own right, and indeed are widely taken (as we shall argue below, incorrectly) to show that the fossil record of the early parts of these radiations is unreliable; and secondly, because of the broader implications that can be drawn from them; that is, we believe the problems we discuss are inherent in relaxed clock analyses in general.

In the following, we use the following terminology: *calibration density*: the user-defined distribution of plausible ages of a tree node based on (in these cases) the fossil record; *effective prior*: the prior probability distribution that is constructed from the calibration densities by enforcement of known topological constraints. We note that while calibration densities are defined independently for each node, the effective priors of all the nodes are interdependent to some degree since they must obey the topological ordering of the tree.

MATERIALS AND METHODS

In the following, we reproduced the MCMCTree molecular clock analyses of [dos Reis et al. \(2015\)](#) and [Álvarez-Carretero et al. \(2022\)](#) using PAML4.8 and PAML4.9 ([Yang 2007](#)) respectively, and the molecular datasets associated with each. In both cases, we used the efficient approximation method ([dos Reis and Yang 2011](#)) as in the original studies, that is, an estimation of the Hessian and gradient was first calculated with BASEML that was then fed into a final MCMCTree inference. For the metazoan study, we used the LG+Γ4+F substitution model (as in [dos Reis et al. \(2015\)](#)) and the independent rate relaxed clock, and we used the two partition dataset. Other priors were set as in the original paper (i.e. mean rate prior of $G(2,40)$ and for $\sigma^2 G(1,10)$, both with a gamma-Dirichlet distribution; and parameters for the birth–death priors for uncalibrated nodes of $\lambda = \mu = 1$ and $\rho = 0$). For the mammal study, we used the HKY + Γ5 substitution model in the calculation of the Hessian and gradient in BASEML for the 72 taxa dataset, with both independent and autocorrelated rates, and the aligned molecular dataset in 4 partitions of the original publication. As in the original publication, we set the birth–death process for uncalibrated nodes to $\lambda = \mu = 1$ and $\rho = 0.1$; the mean rate prior was set to $G(2,40)$ and the σ^2 prior to $G(1,10)$. Multiple runs were made to check for convergence in both cases ([Supplementary Figs. S1 and S2](#)). In the case of the mammals, large samples (60,000) were made of the posterior (thinned at one sample every 800 generations) and the first 15,000 were then manually discarded. To sample the effective priors in all cases, we ran the analyses without data. To produce [Figure 5](#), we used the methods and code of [Budd and Mann \(2020b\)](#) to evaluate the probability of crown-group origins and expected lineage diversification through lineage-through-time (LTT), setting the background extinction rate to be 0.5/myr following [Alroy \(1999\)](#), and the speciation rate was set so as to produce an expected 4000 species in the present, commensurate in scale with the number of known placental species. To produce [Figure 6](#), which shows both stem- and crown-group diversity through time, we further used the calculations provided by [Budd and Mann \(2020b\)](#) in order to obtain the probability distribution of these diversities, including the effects of the mass extinction.

ANIMAL ORIGINS

The comprehensive study of animal origins by [dos Reis et al. \(2015\)](#) used 4 different calibration strategies to generate priors for analyses that were run under autocorrelated and uncorrelated rate conditions, and with a variety of rate partitions in the data. Under this wide range of conditions, they recovered dates for the origin of the animals between 833 and 650 Ma, and bilaterians between 688 and 596 Ma. In contrast, the oldest generally accepted animal fossils are younger

than about 575 Ma ([Matthews et al. 2021](#)). It is widely accepted that trace fossils younger than 560 Ma represent at least total-group bilaterians ([Budd and Jensen 2000](#); [Martin et al. 2000](#); [Budd and Mann 2020a](#)); and that body (and small carbonaceous) fossils close to the base of the Cambrian represent bilaterian clades such as total-group chaetognaths (i.e. protoconodonts) or scalidophorans ([Moczyłowska, et al. 2015](#); [Willman and Slater 2021](#)). It has also been argued that the trace fossil *Treptichnus pedum*, which marks the base of the Cambrian, but drifts below it, was made by a priapulid-like animal ([Vannier et al. 2010](#); [Kesidis et al. 2019](#)). A fossil-based date of c. 555–540 Ma for the origin of the crown-group bilaterians would thus not be unreasonable. In addition, there is a compelling congruence between the order of appearance of fossils in the record and phylogeny ([Cunningham et al. 2017](#); [Budd and Mann 2020a](#)). A considerable gap, therefore, exists between even the youngest molecular clock estimates and the fossil record, and the stratigraphic congruence with phylogeny makes this gap even more problematic.

Although [dos Reis et al. \(2015\)](#) examine a range of different calibration strategies, we will only examine the uniform density ones here, for clarity of illustration. We note, however, that the basal node of the metazoans as a whole was in each of their 4 cases calibrated with a uniform density. Uniform calibration densities are in addition extended by “soft” margins at both ends in order to account for mistakes in fossil assignment or oldest age estimation ([Yang and Rannala 2006](#)).

We have used the same dates as [dos Reis et al. \(2015\)](#) in our analysis. The soft maximum age used for the Metazoa was the maximum age of the Bitter Springs exceptionally preserved biota at c. 833 Ma ([Benton et al. 2015](#)); and the soft maximum for the other 3 nodes was the Lantian biota at c. 635 Ma. Since [dos Reis et al. \(2015\)](#) was published, the Lantian has been related to be younger than c. 602 ± 8 myr ([Yang et al. 2022](#)). In addition, we do not accept that *Kimberella*, used as the oldest protostome, is in fact a protostome ([Budd and Jensen 2017](#)). Conversely, we would accept the oldest bilaterian-style trace fossils as the oldest possible crown-group bilaterians at c. 560 Ma ([Budd and Mann 2020a](#)). For the sake of simplicity again, we are choosing to bracket possible dates from other putative early metazoans such as the 890 Ma sponge described by [Turner \(2021\)](#), although we do not find these compelling.

As noted by [dos Reis et al. \(2015\)](#), the logical necessity of clades being younger than the clades they are nested within means that initial calibration densities must be truncated to form an effective prior. In the case of the basal nodes of the animals, this truncation is striking ([Fig. 1](#)).

As can be seen, in each of the Metazoa, Bilateria, Protostomia and Deuterostomia, the effective prior has a much older, and, perhaps as importantly, narrower age range than the calibration range. Indeed, in all cases apart from the Protostomia, the prior assigns less than 5% probability of an age within 20

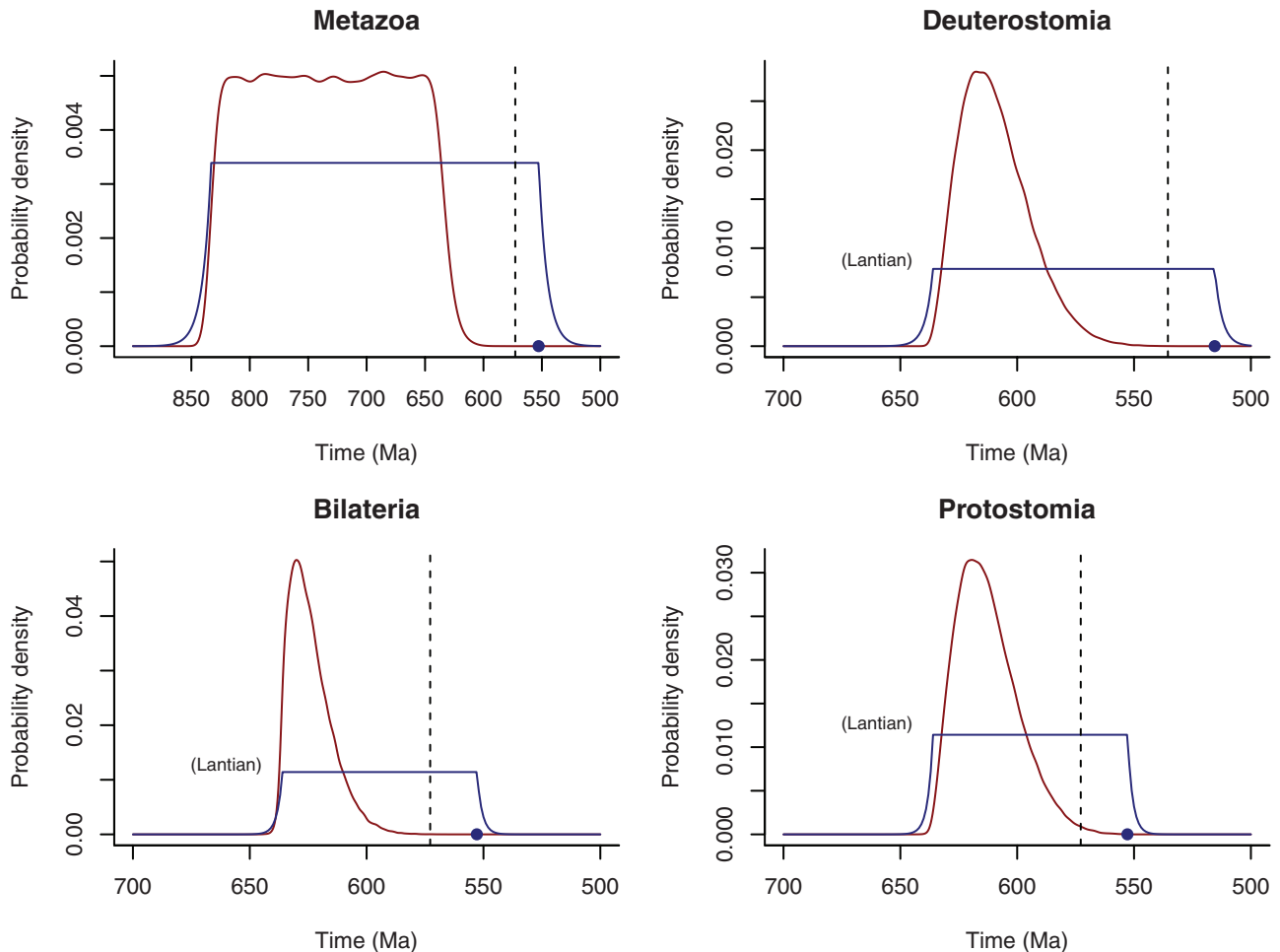


FIGURE 1. The uniform calibration densities (blue) and subsequent effective priors (red) were used in [dos Reis et al. \(2015\)](#). The oldest fossil used for setting the youngest soft limit is marked with a blue dot; the dashed line is 20 myr older than this fossil in each case. Note that the Lantian biota, used as a soft maximum age for three of these nodes, has recently been dated to be c. 602 ± 8 Ma ([Yang et al. 2022](#)). Note that in each case, the effective prior typically assigns little probability to dates near the oldest fossil, or even to dates 20 myr older than it.

myr of the oldest fossil (dashed lines in [Fig. 1](#)). We would thus argue that their analyses do not provide a convincing test of the hypothesis of the fossil record more or less faithfully recording the origin of the animals; instead, the effective prior structure more or less rules out this hypothesis *a priori*, that is, before any molecular data are introduced.

Prior–Posterior Plots: When the Priors and Posteriors Agree

A further reason for finding the above formation of effective priors problematic is the known heavy correlation of posterior age estimates with them (e.g. [Brown and Smith 2018](#)), which we illustrate here by complete prior–posterior plots (with the mean and 95% credible interval for the effective prior and posterior for each node in an analysis) ([Fig. 2](#); c.f. [Warnock et al. 2015](#); [Parins et al. 2017](#); [Chazot et al. 2019](#)). In the plot of the analysis of the focal result (i.e. with independent rates) of [dos Reis et al. \(2015\)](#) for uniform priors ([Fig. 2a](#)), it can be seen that most nodes lie close to or even on the

line corresponding to a 1:1 match between the prior and posterior, and have similar credible intervals. In addition, those few nodes that lie far from this line typically have extremely broad posterior ranges, ranging over hundreds of millions of years. The posteriors thus either simply reproduce the priors, or are essentially uninformative.

We then re-ran the focal analysis of [dos Reis et al. \(2015\)](#) but changing a single number, the soft maximum age of the Metazoa, to reflect the diversification seen in the fossil record, that is, to 580 Ma ([Fig. 2b](#); c.f. [Budd and Mann 2020a, 2020b](#)). Because of the tree topology, the maximum ages of the effective priors on other basal nodes will also be truncated to the same age. As can be seen, and as expected, the analysis accommodates this shift simply by bunching up the earliest nodes, including the bilaterians. Once again, there is a close prior–posterior correspondence.

We thus find, analogous to [Brown and Smith \(2018\)](#), that the effective priors, especially those over the maximum age, exert an overwhelming influence on the

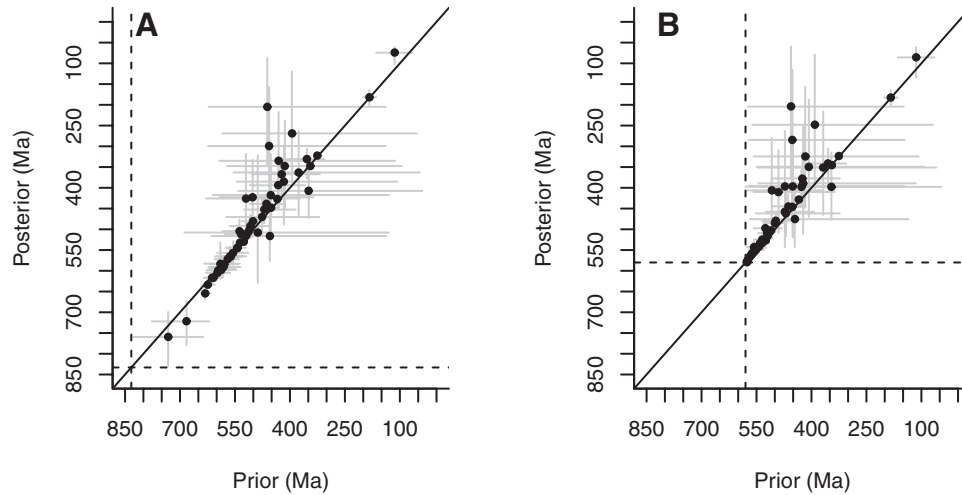


FIGURE 2. Prior–posterior plots for the analysis of [dos Reis et al. \(2015\)](#) on the origin of the animals, using uniform calibration ranges with independent rates. Each point represents a node in the phylogeny with its associated 95% HDP range. The dashed lines in this case represent the soft maximum age for the Metazoa (i.e. the root of the tree). a) The focal analysis of [dos Reis et al. \(2015\)](#) setting the soft maximum age at c. 833 Ma. b) A reanalysis setting the soft maximum age at 580 Ma, that is, just before the oldest Ediacaran assemblages ([Matthews et al. 2021](#)).

posteriors. A corollary to this is that the molecular data are compatible with a wide range of true clade ages, including ones faithful to the fossil record, as we demonstrate here because clock techniques can only ever estimate rate \times time and not time independently of other parameters (e.g. [Britton 2005](#); [dos Reis et al. 2016](#)).

WHEN THE PRIORS AND POSTERiors DISAGREE: THE CASE OF THE PLACENTAL MAMMALS

The origin of the placental mammals has been examined by, among others ([Álvarez-Carretero et al. 2022](#)). This analysis used a huge phylogenomic dataset and concluded with strong probability that the crown placental mammals emerged c. 83.3–77.6 Ma, that is, considerably before the K-Pg boundary at c. 66 Ma. This is in stark contrast to analyses such as that of [O’Leary et al. \(2013\)](#), which concluded that the crown placentals emerged rapidly after the K-Pg boundary.

In the [Álvarez-Carretero et al.](#) analysis, a uniform calibration for the placentals from the oldest fossil they selected at 61.66–162.5 Ma was used, derived, at least partly, from previous molecular clock results ([Benton et al. 2015](#)).

Once again, the calibration to effective prior conversion in this case is problematic ([Fig. 3a](#)), with an extreme bias toward older dates of the calibration. Sampling the effective prior 1 million times generated no dates younger than c. 69 Ma ([Fig. 3b](#)). Thus, the effective prior assigns a vanishingly low probability to the placentals emerging after the K-Pg boundary, that is, much less than one in a million (c.f. [Brown and Smith \(2018\)](#)). Nevertheless, this is, after all, what the fossil record suggests: there are thousands of stem-group fossils known from before the boundary and no

uncontroversial crown-group forms; yet a large number of unambiguous crown group forms appear soon after the boundary ([Prasad et al. 2010](#); [Springer et al. 2013](#); [Davies et al. 2017](#)). The prior–posterior plot in this case is of particular interest because the posterior of the placentals and nodes near it plot far from their respective priors ([Fig. 3c,d](#)). In this instance, the problem is not that the prior controls the posterior, but rather that the prior is so far from what the data suggest. In other words, the “signal” from the prior and the “signal” from the data are fundamentally different; and combining two very divergent signals in a Bayesian context ([Evans and Moshonov 2006](#)) should acknowledge the plausible scenario that one or other signal (or both) may be systematically erroneous ([Cheng et al. 2007](#)). In this instance, the effective prior assigns approximately 1 in 11,000 probability (0.009%) to ages younger than 83 Ma. That is, the age eventually presented as the focal estimate for the placental node is one that was considered highly unlikely according to the effective prior.

In addition to their focal analysis, [Álvarez-Carretero et al. \(2022\)](#) tested the influence of the shape of the fossil calibration density on the posterior by using two types of truncated Cauchy calibrations (their extended data [Fig. 2](#)), and showed that when a very sharply defined Cauchy calibration was used for the placentals, the mean age of appearance was only a few million years younger than that of the focal analysis, with the implication that choice of calibration had only limited influence. We consider this result to be problematic, however. Truncated Cauchy distributions have undefined means and variance, as they do not converge, which implies that they are very heavy tailed. Indeed, when comparing the focal prior to their “short” truncated Cauchy prior, one can see that the Cauchy prior starts assigning more probability than the focal prior as deep as c. 93 Ma ([Fig. 3e,f](#)). In addition, the gap between the placentals

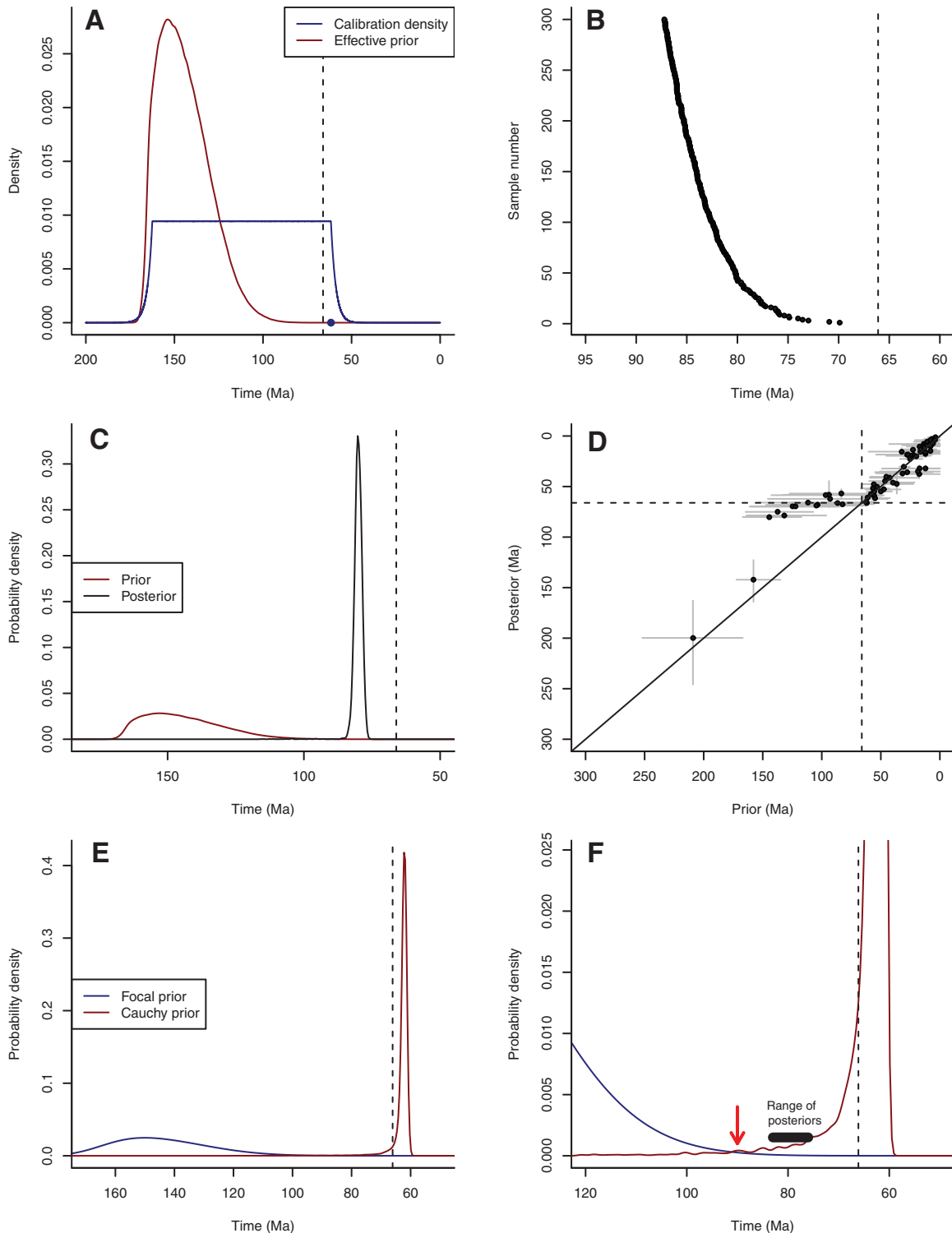


FIGURE 3. a) Calibration to effective prior conversion in the analysis of mammal origins of [Álvarez-Carretero et al. \(2022\)](#). Dashed line: K-Pg boundary; blue: uniform calibration range with soft limits; blue dot, oldest fossil; red: effective prior. b) The youngest 300 of one million samples of the effective prior, showing the vanishingly low probability of an appearance close to or after the K-Pg boundary. c) The effective prior and posterior distributions for the placental mammals from [Álvarez-Carretero et al. \(2022\)](#). Note that, in this case, the posterior lies far outside the prior distribution. d) The prior–posterior plot for the entire analysis (72 taxa). Note that the nodes close to the placentals trend off the prior–posterior equivalence line, a probable result of the very deep placental prior; the large gap between the Theria and the placentals, and the large number of nodes constrained by their priors to lie after the K-Pg boundary (dashed line). e) The “short” truncated Cauchy prior of [Álvarez-Carretero et al. \(2022\)](#) (red) plotted together with the prior of the focal analysis (blue). Note that when zoomed in f) the Cauchy prior can be seen to assign more probability than the focal prior for ages as deep as c. 93 Ma (red arrow), older than any of the posteriors.

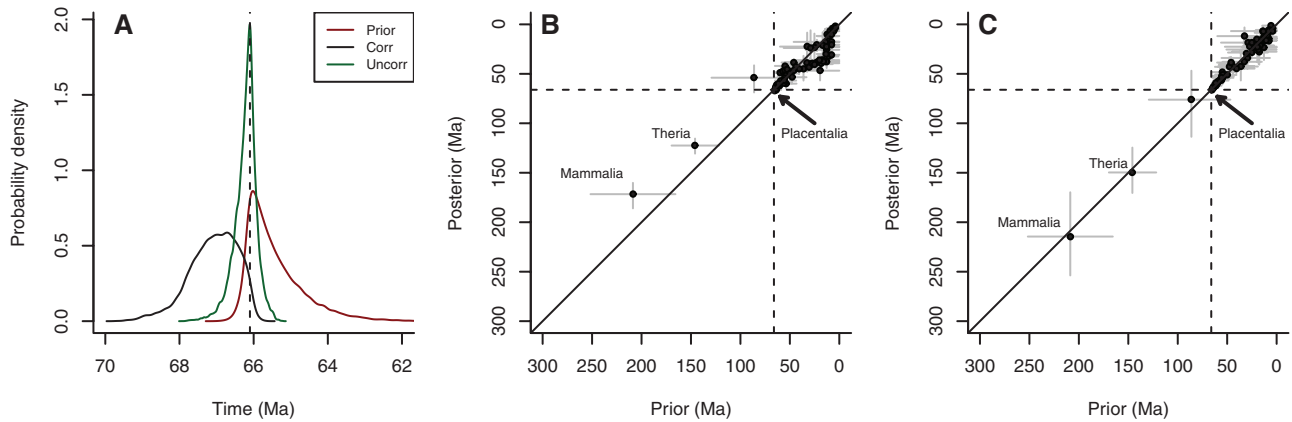


FIGURE 4. The focal analysis of [Álvarez-Carretero et al. \(2022\)](#), rerun with a uniform calibration density on the age of the placentals down to the K-Pg boundary, with soft limits. a) The effective prior and the posteriors for uncorrelated and autocorrelated runs for the placental node. b) prior–posterior plot for the mammals run with an autocorrelated rate model. c) Prior–posterior plot for the mammals run with an uncorrelated log-normal rates model. Note that under such conditions, the posteriors return to plotting broadly along the 1:1 prior–posterior line and that the molecular evidence provides no compelling evidence for a pre-K-Pg boundary emergence of the placentals.

and the next deepest node, the Theria, is considerable, so that the placentals are effectively the root of the clade. In such circumstances, the use of a Cauchy prior is hazardous because it is easy for the posterior to drift deep (c.f. [Brown and Smith 2018](#) who report similar effects with exponential priors).

Rather than change the *type* of prior, then, we re-ran the analysis using a uniform calibration for the placentals and other deep nodes within them from the age of their oldest fossil down to the K-Pg boundary, with soft bounds on both ages ([Fig. 4](#)). We note that similar calibration ranges are used for 11 other nodes in their analysis (e.g. Caniformia, Chiroptera, Carnivora, etc). Once again, construction of the placental effective prior to this calibration density led to a skewed distribution almost exactly on the boundary (mean c. 65.53 Ma, hdi 64.30–66.44 Ma). Under such conditions, the analysis recovered a posterior mean age very close to the age of the boundary (mean 66.47 Ma, 95% hdi 65.60–68.85 Ma for uncorrelated runs, and mean 66.55 Ma, 95% hdi 65.98–68.70 Ma for correlated runs), even when this would imply extremely rapid early radiation. In both cases, the posterior is compatible with a post-K-Pg origin of the placentals.

The case of the placental mammals is instructive because although their emergence before the K-Pg boundary is widely accepted, there are strong theoretical and empirical grounds for thinking that they really emerged *after* it. The first is provided by the lineage-through-time (LTT) plot ([Fig. 5a](#)). As can be seen, when the placental emergence is placed where [Álvarez-Carretero et al. \(2022\)](#) suggest it is (i.e. c. 80 Ma), the subsequent lineages rapidly increase across the K-Pg boundary with no apparent perturbation. However, it is clear that the mammals were subject to an extremely severe extinction at this point, with numbers up to 95% of species extinction being suggested ([Longrich et al. 2016](#)). The slope of the LTT plot is equal to the rate of speciation (λ) multiplied by the probability of survival

to the present. Under homogeneous conditions of constant speciation and extinction, the probability of survival to the present stays fairly constant until close to the present ([Budd and Mann 2018](#)). However, when these conditions are punctuated by mass extinctions, the survival probability is substantially greater after the mass extinction than before, typically leading to an increase in the rate of lineage creation at that point ([Budd and Mann 2020b](#)). For the LTT plot to pass unperturbed through the extinction would imply that there must be an immediate and sustained *decrease* in speciation rate after the mass extinction that exactly balances the increase in survival probability, such that lineages continue to be produced at the same rate. This is inherently implausible and unparsimonious, and the more logical expectation is that the break of slope that is clearly seen in the LTT plot of [Álvarez-Carretero et al. \(2022\)](#) should occur not just before, but *at* the mass extinction ([Budd and Mann 2020b](#)). Furthermore, the probability of the crown group forming is, in general, at its lowest just before a mass extinction. This is a direct consequence of the difficulty of coalescence past a mass extinction. Thus although an earlier MCMCTree analysis ([Liu et al. 2017](#)) reported “uninterrupted” diversification of placental mammals across the boundary, this is both theoretically implausible and contradicted by the fossil record.

The emergence of the crown group, and the effect of mass extinctions, can be modeled using birth–death modeling ([Lambert and Stadler 2013](#); [Budd and Mann 2020b](#)). A crown group can logically emerge either before or after a large mass extinction, but it is most likely to do so either early on in the history of the total group or just after the mass extinction ([Fig. 5b](#)). What is more if the crown group forms before the mass extinction we still expect to see a sharp increase in the rate of lineage creation immediately after the mass extinction, creating an inflection point in the LTT plot at 66 Ma ([Fig. 5c–e](#)), which is notably

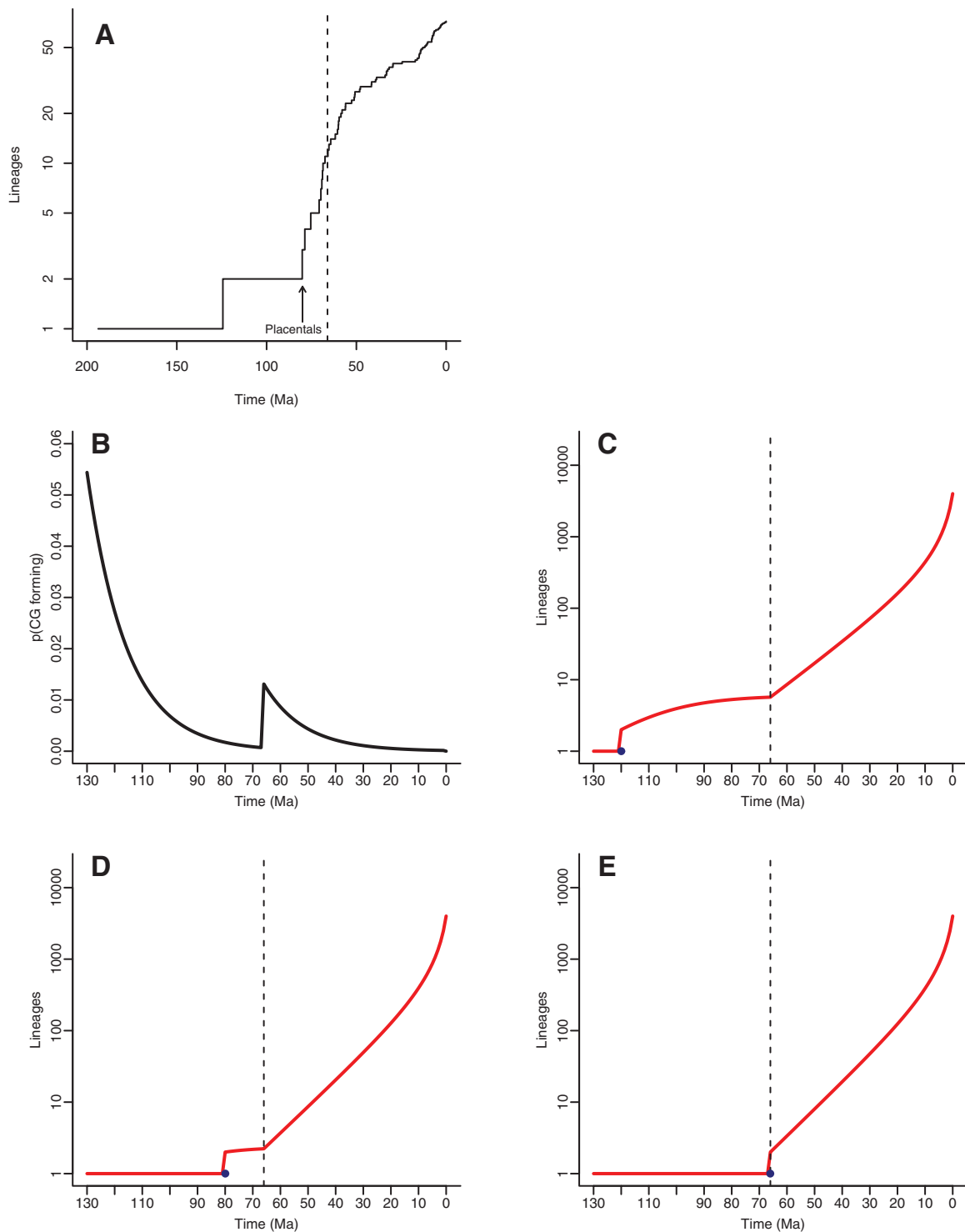


FIGURE 5. a) The central LTT plot of [Álvarez-Carretero et al. \(2022\)](#) for 72 taxa, showing the origin of the crown-group placentals. Note that in their dating, the LTT plot passes unperturbed through the K-Pg boundary, despite the very high rates of mammalian extinction at that point. (b–e) Birth–death models and mass extinctions. Here, a clade commencing at c. 130 Ma, and diversifying to give rise to c. 4000 living species with a background extinction rate of 0.5/myr passes through a mass extinction of 95% at 66 Ma. b) The probability of the crown group forming through time. c) The LTT plot when the crown group emerges at c. 125 Ma (the blue dot indicates the oldest possible crown-group fossil age). d) The same, but for an emergence at the focal time of [Álvarez-Carretero et al. \(2022\)](#). e) The same, but for crown-group emergence just after the mass extinction. Note that, in each case, the LTT plot has a distinct inflection at the point of extinction, reflecting the difficulty of coalescence back through a mass extinction.

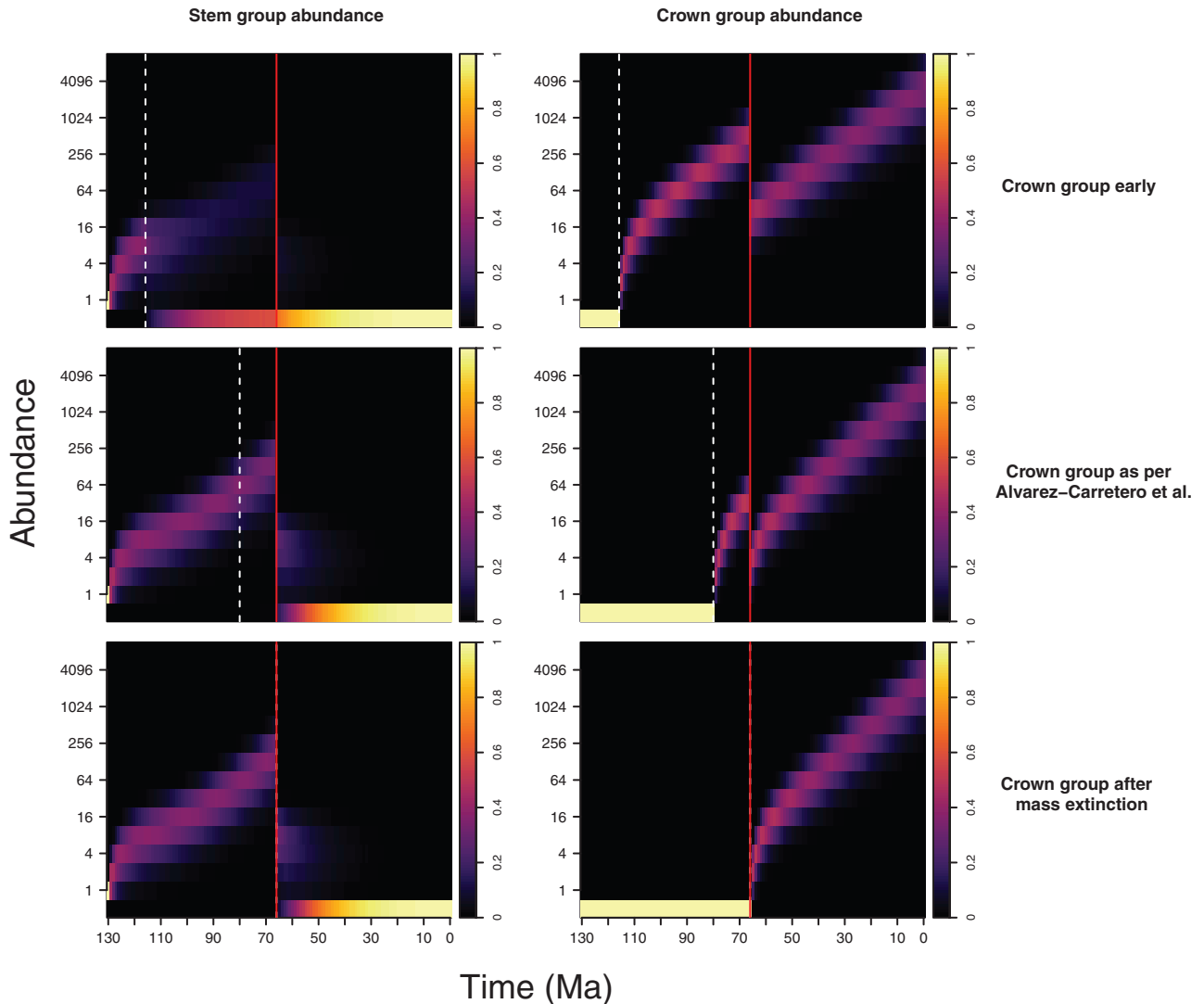


FIGURE 6. Heat maps for placental diversification for each of the three scenarios in Figure 5. White dashed line: time of crown-group origin; red dashed line, mass extinction. Note that in both the upper and middle scenarios, significant crown-group diversity should be expected in the Cretaceous, with, in the focal analysis of [Álvarez-Carretero et al. \(2022\)](#), approximately one third of total late-Cretaceous diversity being crown group, and with a total of several hundred species being expected during this interval of time. Conversely, the bottom scenario, with the crown group emerging just after the boundary, matches the known fossil record well.

absent in the results of [Álvarez-Carretero et al. \(2022\)](#). In simple terms, this is because, in retrospect, lineages reflect the routes to modern-day survivors; and clearly, a species living just after a mass extinction is much more likely to lead to modern-day survivors than one living just before.

As well as exploring the LTT plot morphology, it is also possible to produce heat maps using the method of [Budd and Mann \(2020b\)](#) (Fig. 6) for stem- and crown-group diversity through time for each of the scenarios of Fig. 5c–e. These are based on the calculation of the distribution of stem- and crown-group diversity at any point during a diversification for given speciation and extinction rates and can include the effects of mass extinctions. These reveal a further difficulty in reconciling the estimates of molecular

clocks with the established fossil record. As noted above, there are no uncontroversial crown-group placentals older than 66 Ma, but there are thousands of uncontroversial stem-group placental fossils before this K-Pg boundary. As shown in Figure 6, if the crown group of placentals emerged very early, or at c. 80 Ma, we would in both these cases expect a substantial diversification of the crown group before the K-Pg boundary, comparable to the stem diversity at the same time. Given the preservation of *stem-group* placentals during this interval, it is natural to ask where the crown-group fossils are ([Davies et al. 2017](#)). Conversely, if the crown-group forms immediately after the K-Pg mass extinction, the expected abundance of stem and crown groups through time closely matches the known distribution of fossils.

CALIBRATION DENSITY TRUNCATION AND ITS MEANING

In principle, a relaxed molecular clock should provide an unbiased (if broad) estimate of the age of a clade, if all the conditions of the model assumptions are met. However, one important such condition is the tree prior. In packages such as BEAST (Drummond et al. 2012) and FBDP in RevBayes (Höhna et al. 2016) (implementing the fossilized birth-death process), a birth–death process tree prior can be used. No tree prior is used in MCMCTree on nodes that have a fossil calibration. Nevertheless, there is an *implicit* model for how the tree should look, and this is reflected in the conversion of the calibration densities into the effective priors (for the mathematics of a simple case, see Inoue et al. 2010). When the oldest fossils of several nested deep clades are temporally close, and yet a deep maximum age is imposed both on the tree as a whole and on these nodes, then the resulting effective priors will be spread out evenly throughout the possible range. However, there are at least three good reasons for thinking that the bunching up of the nested clades (i.e. a rapid early radiation) as suggested by the fossils in both cases studied here is a real phenomenon: (i) modern taxa are typically under-sampled, either randomly or (especially) by “diversified sampling” (Höhna et al. 2011; Matschiner 2019), which has the effect of artificially lengthening the time between successive later nodes; (ii) true diversity-dependent diversification or other ecological effects will produce early lineages faster-than-expected compared to later ones (Moen and Morlon 2014); (iii) clades that are larger than expected given their diversification parameters (the “large clade effect”) will generate an essentially identical effect (Budd and Mann 2018, 2020b). In each case, the early nodes of a clade will appear faster than later ones. Rapid appearance of early fossils is *prima facie* evidence of such of at least points (ii) and (iii). In effect, the models used in molecular clock software are strongly biased against early rapid lineage radiations, even though we expect these to be common, either by construction (point (i)) above) or in reality (points (ii) and (iii)). While a strict molecular clock may be able to overcome this bias, with enough data (Budd and Mann 2020a) a relaxed clock clearly cannot. Indeed, our contention here is that if a clade really evolved quickly in this way, the clock could never recover this unless the priors were chosen to very closely reproduce the fossil record, which would largely obviate the need for the clock analysis.

DISCUSSION

In this contribution, we have considered two studies of well-known and controversial radiations, that of the animals and the placentals within the mammals (dos Reis et al. 2015; Álvarez-Carretero et al. 2022). In the first case, the posteriors largely reproduce the priors, and in the second, some are at complete odds with

them. In both cases, the studies concluded with a very high probability that the clades in question originated well before their fossil record commenced. However, we conclude from our analyses that relaxed molecular clocks in these cases do not provide a reliable way of estimating the ages of clades, because (i) deep nodes tend to have effective priors with mean ages much older than the raw calibration densities would suggest; (ii) such priors assign virtually no probability close to the fossil record, which implies that these clock analyses are not testing the hypothesis that the fossil record might be an accurate reflection of evolutionary events; (iii) simply changing the maximum age on the root calibration density can have very significant effects on basal node posterior age inference; (iv) in general, there is a strong dependence of the posterior estimates on the effective priors. In principle, the posterior of a Bayesian analysis should refine the prior, that is, lie within, but be narrower than the range of the prior, and this is not fully the case in either of our two examples here. In addition, the problematic way in which the effective priors are specified means that even if this were the case, the molecular clock analysis still would not give the correct answer. Overall, the clear implication is that the *molecular* part of the analysis does not allow us to distinguish between different times of origin of the clade, and thus does not contradict the general picture provided by the fossil record.

The explosive post-Cretaceous model for placental mammal origins has been deprecated, partly on the grounds of the alleged too-rapid rates of molecular evolution that must be inferred for the earliest branches (Springer et al. 2013). In addition, in order to fit in the several lineage-production events between the boundary and the first putative diagnosable in-group members of the placentals (Springer et al. 2013), several speciation events must have occurred in a few hundred of thousand years at a rate much higher than normally used in birth–death models. However, there is no good reason to think that speciation (or extinction) or indeed molecular substitutions are really Poisson distributed (Gillespie 1984; Rannala and Yang 2007; Lee et al. 2013), and it is entirely plausible that all these rates could be much higher than the long-term mean when short intervals are considered. This should be seen in the context that crown-group formation is inherently unlikely shortly before mass extinctions (Budd and Mann 2020b), and that no uncontroversial crown-group placental fossils have been found older than 66 Ma. Thus, it is unclear what the grounds really are for thinking that placentals emerged before the boundary; and there are several sound reasons for thinking they emerged after it.

Bayesian molecular clock analyses are popular because they can be run on large datasets and can take into account many uncertainties by setting priors over them. However, the calibrations, and, in particular, the priors that arise from them, are typically not biologically informed (e.g., in general, they do

not reflect what is known about the dynamics of radiations or the effect of mass extinctions), and we show here and elsewhere (e.g. [Budd and Mann 2020b](#)) that these factors are very important in understanding the timing of origins of clades, and thus should be considered as part of any “total evidence” approach ([Ronquist et al. 2016](#)). Dependence of posterior estimates on the priors, and the problems of range truncations have been shown for other clades ([Prum et al. 2015](#); [Moody et al. 2022](#)) and other software as well as the MCMCTree examples examined here ([Lee et al. 2013](#); [Field et al. 2020](#); [Sauquet et al. 2022](#)), strongly suggesting that the issues we raise here are widespread, and not confined to just MCMCTree analyses ([Chazot et al. 2019](#)). In other words, we believe that the problems we identified in the two studies we examined are typical of relaxed molecular clock analyses in general. We suggest that, at the minimum, complete prior–posterior plots should be included in molecular clock analyses as a matter of routine, so that the total contribution of the calibrations, priors, and molecular data can be readily seen. Even so, we believe that our results must cast severe doubt on all relaxed clock outcomes that substantially predate well-established fossil records, including those affected by mass extinctions.

As noted above, we regard the age of a clade as an empirical matter, and thus any estimate should and must be open to refinement or even refutation from evidence. Furthermore, we regard the fossil record in itself as providing evidence for such timings, albeit of highly varying quality. We would thus dissent, for example, from the recently expressed view about angiosperm origins that because molecular clock estimates are not reliable, we thus have no real idea about when they emerged ([Sauquet et al. 2022](#)). The ingrained idea that the fossil record can only start after a long and unpredictable “lag” period or “phylogenetic fuse” ([Cooper and Fortey 1998](#); [Cavalier-Smith 2017](#); [Marshall 2019](#)) has been criticized on various grounds ([Budd and Mann 2020b](#)), but seems to be the basis for the relative insouciance with which large gaps between clocks and fossils are accepted. When the fossil record is of high quality, with large numbers of fossils from geographically widespread localities, as is the case for the clades we are considering here, then to posit an origin much before the fossil record commences would require the addition of more empirical evidence. In the absence of evidence of this sort, though, the positing of large gaps in the fossil record, especially when they are well outside those suggested by stratigraphic confidence intervals ([Strauss and Sadler 1989](#); [Marshall 1997](#)), appears to be an unnecessary hypothesis that requires inevitable appeals to *ad hoc* explanations, which although common in the literature, cannot be general phenomena. In this light, molecular clock analyses should in our view attempt to test the fossil record derived date hypothesis, and they cannot do this if they assign vanishingly small *prior* probabilities to the eventuality that the fossil record is broadly correct even before the molecular

evidence is considered ([Brown and Smith 2018](#)); nor are they useful if they are compatible with both deep and shallow origins of clades, as we show here for the animals and placentals.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.rv15dv4b4>.

ACKNOWLEDGMENTS

We thank Richard Copley for some initial help with computing and Mario dos Reis Barros for kindly making some material from [dos Reis et al. \(2015\)](#) available.

FUNDING

This work was supported by UK Research and Innovation Future Leaders Fellowship grant MR/S032525/1 to RPM and by Swedish Research Council (VR) grant 2022-03522 to GEB.

CONFLICT OF INTERESTS

The authors have no competing interests to declare.

REFERENCES

- Alroy J. 1999. The fossil record of North American mammals: evidence for a paleocene evolutionary radiation. *Syst. Biol.* 48:107–118. [dataset] [Alvarez-Carretero S., Tamuri A., Battini M., Nascimento F.F., Carlisle E., Asher R., et al. 2021. Data for “A species-level timeline of mammal evolution integrating phylogenomic data”. Figshare. <https://doi.org/10.6084/m9.figshare.14885691.v1>.](#)
- [Álvarez-Carretero S., Tamuri A.U., Battini M., Nascimento F.F., Carlisle E., Asher R.J., Yang Z., Donoghue P.C.J., dos Reis M. 2022. A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* 602:263–267.](#)
- [Barba-Montoya J., dos Reis M., Yang Z. 2017. Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. *Mol. Phylogenet. Evol.* 114:386–400.](#)
- [Barba-Montoya J., Dos Reis M., Schneider H., Donoghue P.C.J., Yang Z. 2018. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous terrestrial revolution. *New Phytol.* 218:819–834.](#)
- [Benton M., Donoghue P.C.J., Asher R.J. 2009. Calibrating and constraining molecular clocks. In: Hedges S.B., Kumar S., editors. *The timetree of life*. New York: OUP. p. 35–86.](#)
- [Benton M.J., Donoghue P.C.J., Asher R.J., Friedman F., Near T.J., Vinther J. 2015. Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron.* 18\(1\):1–106.](#)
- [Benton M.J., Wills M.A., Hitchin R. 2000. Quality of the fossil record through time. *Nature* 403:534–537.](#)
- [Britton T. 2005. Estimating divergence times in phylogenetic trees without a molecular clock. *Syst. Biol.* 54:500–507.](#)
- [Bromham L. 2019. Six impossible things before breakfast: assumptions, models, and belief in molecular dating. *Trends Ecol. Evol.* 34:474–486.](#)

- Bromham L., Duchêne S., Hua X., Ritchie A.M., Duchêne D.A., Ho S.Y.W. 2018. Bayesian molecular dating: opening up the black box. *Biol. Rev.* 93:1165–1191.
- Brown J.W., Smith S.A. 2018. The past sure is tense: on interpreting phylogenetic divergence time estimates. *Syst. Biol.* 67:340–353.
- Budd G.E., Jensen S. 2000. A critical reappraisal of the fossil record of the bilaterian phyla. *Biol. Rev.* 75:253–295.
- Budd G.E., Jensen S. 2003. The limitations of the fossil record and the dating of the origin of the Bilateria. *Syst. Ass. Spec.* 66:166–189.
- Budd G.E., Jensen S. 2017. The origin of the animals and a “Savannah” hypothesis for early bilaterian evolution. *Biol. Rev.* 92:446–473.
- Budd G.E., Mann R.P. 2018. History is written by the victors: the effect of the push of the past on the fossil record. *Evolution* 72:2276–2291.
- Budd G.E., Mann R.P. 2020a. Survival and selection biases in early animal evolution and a source of systematic overestimation in molecular clocks. *Interface Focus*. 10:20190110.
- Budd G.E., Mann R.P. 2020b. The dynamics of stem and crown groups. *Sci. Adv.* 6(8):eaaz1626. doi: [10.1126/sciadv.aaz1626](https://doi.org/10.1126/sciadv.aaz1626).
- Cai C., Tihelka E., Giacomelli M., Lawrence J.F., Ślipiński A., Kundrata R., Yamamoto S., Thayer M.K., Newton A.F., Leschen R.A.B., Gimmel M.L., Lü L., Engel M.S., Bouchard P., Huang D., Pisani D., Donoghue P.C.J. 2022. Integrated phylogenomics and fossil data illuminate the evolution of beetles. *R. Soc. Open Sci.* 9(3):211771. doi: [10.1098/RSPS.20211771](https://doi.org/10.1098/RSPS.20211771).
- Cavalier-Smith T. 2017. Origin of animal multicellularity: precursors, causes, consequences—the choanoflagellate/sponge transition, neurogenesis and the Cambrian explosion. *Phil. Trans. R. Soc. B.* 372:372(1713):20150476. doi: [10.1098/RSTB.2015.0476](https://doi.org/10.1098/RSTB.2015.0476).
- Chazot N., Wahlberg N., Freitas A.V.L., Mitter C., Labandeira C., Sohn J.C., Sahoo R.K., Seraphim N., De Jong R., Heikkilä M. 2019. Priors and posteriors in Bayesian timing of divergence analyses: the age of butterflies revisited. *Syst. Biol.* 68:797–813.
- Cheng K., Shettleworth S.J., Huttenlocher J., Rieser J.J. 2007. Bayesian integration of spatial information. *Psychol. Bull.* 133:625–637.
- Cooper A., Fortey R. 1998. Evolutionary explosions and the phylogenetic fuse. *Trends Ecol. Evol.* 13:151–156.
- Cunningham J.A., Liu A.G., Bengtson S., Donoghue P.C.J. 2017. The origin of animals: can molecular clocks and the fossil record be reconciled? *Bioessays* 39:1–12.
- Davies T.W., Bell M.A., Goswami A., Halliday T.J.D. 2017. Completeness of the eutherian mammal fossil record and implications for reconstructing mammal evolution through the Cretaceous/Paleogene mass extinction. *Paleobiology* 43:521–536. doi: [10.1017/pab.2017.20](https://doi.org/10.1017/pab.2017.20).
- [dataset] Dos Reis M., Thawornwattana Y., Angelis K., Telford M.J., Donoghue P.C.J., Yang Z. 2015. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *figshare*. doi: [10.6084/m9.figshare.1525089.v1](https://doi.org/10.6084/m9.figshare.1525089.v1).
- Dos Reis M., Donoghue P.C.J., Yang Z. 2016. Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17:71–80. doi: [10.1038/nrg.2015.8](https://doi.org/10.1038/nrg.2015.8).
- Dos Reis M., Thawornwattana Y., Angelis K., Telford M.J., Donoghue P.C.J., Yang Z. 2015. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Curr. Biol.* 25:2939–2950.
- Dos Reis M., Yang Z. 2011. Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Mol. Biol. Evol.* 28:2161–2172.
- Dos Reis M., Yang Z. 2013. The unbearable uncertainty of Bayesian divergence time estimation. *J. Syst. Evol.* 51:30–43.
- Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29:1969–1973.
- Evans M., Moshonov H. 2006. Checking for prior-data conflict. *Bayesian Anal.* 1(4):893–914. doi: [10.1214/06-BA129](https://doi.org/10.1214/06-BA129).
- Field D.J., Berv J.S., Hsiang A.Y., Lanfear R., Landis M.J., Dornburg A. 2020. Chapter 5: timing the extant avian radiation: the rise of modern birds, and the importance of modeling molecular rate variation. *Bull. Am. Mus. Nat. Hist.* 440:159–181.
- Gillespie J. 1984. The molecular clock may be an episodic clock. *Proc. Natl. Acad. Sci. U.S.A.* 81:8009–8013.
- Höhna S., Landis M.J., Heath T.A., Boussau B., Lartillot N., Moore B.R., Huelsenbeck J.P., Ronquist F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* 65:726–736.
- Höhna S., Stadler T., Ronquist F., Britton T. 2011. Inferring speciation and extinction rates under different sampling schemes. *Mol. Biol. Evol.* 28:2577–2589.
- Inoue J., Donoghue P.C.J., Yang Z. 2010. The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Syst. Biol.* 59:74–89.
- Kesidis G., Slater B.J., Jensen S., Budd G.E. 2019. Caught in the act: priapulid burrowers in early Cambrian substrates. *Proc. R. Soc. B Biol. Sci* 286(1894):20182505. doi: [10.1098/rspb.2018.2505](https://doi.org/10.1098/rspb.2018.2505).
- Klopfstein S. 2021. The age of insects and the revival of the minimum age tree. *Austral Entomol.* 60:138–146.
- Lambert A., Stadler T. 2013. Birth-death models and coalescent point processes: the shape and probability of reconstructed phylogenies. *Theor. Popul. Biol.* 90:113–128. doi: [10.1016/j.tpb.2013.10.002](https://doi.org/10.1016/j.tpb.2013.10.002).
- Lee M.S., Soubrier J., Edgecombe G.D. 2013. Rates of phenotypic and genomic evolution during the Cambrian explosion. *Curr. Biol.* 23:1889–1895.
- Liu L., Zhang J., Rheindt F.E., Lei F., Qu Y., Wang Y., Zhang Y., Sullivan C., Nie W., Wang J., Yang F., Chen J., Edwards S.V., Meng J., Wu S. 2017. Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *PNAS*. 114:E7282–E7290.
- Longrich N.R., Sclerkeras J., Willis M.A. 2016. Severe extinction and rapid recovery of mammals across the Cretaceous–Palaeogene boundary, and the effects of rarity on patterns of extinction and recovery. *J. Evol. Biol.* 29:1495–1512.
- Marshall C.R. 1997. Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology*. 23:165–173.
- Marshall C.R. 2019. Using the fossil record to evaluate timetree timescales. *Front. Genet.* 10:1049.
- Martin M.W., Grazhdankin D.V., Bowring S.A., Evans D.A.D., Fedonkin M.A., Kirschvink J.L. 2000. Age of neoproterozoic bilaterian body and trace fossils, White Sea, Russia: implications for metazoan evolution. *Science* 288:841–845.
- Matschiner M. 2019. Selective sampling of species and fossils influences age estimates under the fossilized birth–death model. *Front. Genet.* 10:1064. doi: [10.3389/fgene.2019.01064](https://doi.org/10.3389/fgene.2019.01064).
- Matthews J.J., Liu A.G., Yang C., McIlroy D., Levell B., Condon D.J. 2021. A Chronostratigraphic framework for the rise of the Ediacaran Macrobiota: new constraints from mistaken point ecological reserve, Newfoundland. *Bull. Geol. Soc. Am.* 133:612–624.
- Moczyłowska M., Budd G.E., Agic H. 2015. Ecdysozoan-like sclerites among Ediacaran microfossils. *Geol. Mag.* 152:1145–1148. doi: [10.1017/S001675681500045X](https://doi.org/10.1017/S001675681500045X).
- Moen D., Morlon H. 2014. Why does diversification slow down? *Trends Ecol. Evol.* 29:190–197.
- Mongiardino Koch N., Thompson J.R., Hiley A.S., McCowin M.F., Armstrong A.F., Coppard S.E., Aguilera F., Bronstein O., Kroh A., Mooi R., Rouse G.W. 2022. Phylogenomic analyses of echinoid diversification prompt a re-evaluation of their fossil record. *Elife* 11:e72460. doi: [10.7554/ELIFE.72460](https://doi.org/10.7554/ELIFE.72460).
- Moody E.R.R., Mahendrarajah T.A., Dombrowski N., Clark J.W., Petitjean C., Offre P., Szöllösi G.J., Spang A., Williams T.A. 2022. An estimate of the deepest branches of the tree of life from ancient vertically evolving genes. *Elife* 11:e66695. doi: [10.7554/ELIFE.66695](https://doi.org/10.7554/ELIFE.66695).
- O’Leary M. 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 339:662–667.
- Orr R.J.S., Martino E.D., Ramsfjell M.H., Gordon D.P., Berning B., Chowdhury I., Craig S., Cumming R.L., Figuerola B., Florence W., Harmelin J.-G., Hirose M., Huang D., Jain S.S., Jenkins H.L., Kottenko O.N., Kuklinski P., Lee H.E., Madurell T., McCann L., Mello H.L., Obst M., Ostrovsky A.N., Paulay G., Porter J.S., Shunatova N.N., Smith A.M., Souto-Derungs J., Vieira L.M., Voje K.L., Waeschenbach A., Zágorský K., Warnock R.C.M., Liow L.H. 2022. Paleozoic origins of cheilostome bryozoans and their parental care inferred by a new genome-skipped phylogeny. *Sci. Adv.* 8:7452. doi: [10.1126/SCIADV.ABM7452](https://doi.org/10.1126/SCIADV.ABM7452).
- Parins-Fukuchi C., Brown J.W. 2017. What drives results in Bayesian morphological clock analyses? *BioRxiv*: 219048. doi: [10.1101/219048](https://doi.org/10.1101/219048).

- Prasad G.V.R., Verma O., Gheerbrant E., Goswami A., Khosla A., Parmar V., Sahni A. 2010. First mammal evidence from the Late Cretaceous of India for biotic dispersal between India and Africa at the KT transition. *C.R. Palevol* 9:63–71.
- Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573.
- Rannala B., Yang Z. 2007. Inferring speciation times under an episodic molecular clock. *Syst. Biol.* 56:453–466.
- Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using total-evidence dating. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150136.
- Sauquet H., Ramírez-Barahona S., Magallón S. 2022. What is the age of flowering plants? *J. Exp. Bot.* 73:3840–3853. doi: [10.1093/JXB/ERAC130](https://doi.org/10.1093/JXB/ERAC130).
- Springer M.S., Meredith R.W., Teeling E.C., Murphy W.J. 2013. Technical comment on “The placental mammal ancestor and the post-K-Pg radiation of placentals”. *Science* 341:613–613.
- Strauss D., Sadler P.M. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Math. Geol.* 21:411–427.
- Turner E.C. 2021. Possible poriferan body fossils in early neoproterozoic microbial reefs. *Nature* 596:87–91. doi: [10.1038/s41586-021-03773-z](https://doi.org/10.1038/s41586-021-03773-z)
- Vannier J., Calandra I., Gaillard C., Zylińska A. 2010. Priapulid worms: pioneer horizontal burrowers at the Precambrian-Cambrian boundary. *Geology* 38:711–714.
- Warnock R.C.M., Parham J.F., Joyce W.G., Lyson T.R., Donoghue P.C.J. 2015. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proc. R. Soc. B Biol. Sci.* 282:20141013. doi:[10.1098/rspb.2014.1013](https://doi.org/10.1098/rspb.2014.1013)
- Warnock R.C.M., Yang Z., Donoghue P.C.J. 2012. Exploring uncertainty in the calibration of the molecular clock. *Biol. Lett.* 8:156–159.
- Willman S., Slater B.J. 2021. Late Ediacaran organic microfossils from Finland. *Geol. Mag.* 158:2231–2244.
- Yang C., Li Y., Selby D., Wan B., Guan C., Zhou C., Li X.-H. 2022. Implications for Ediacaran biological evolution from the ca. 602 Ma Lantian biota in China. *Geology* 50:562–566.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586–1591.
- Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23:212–226.