

Age Estimation Under Fossilized Birth Death Models

Annotated Bibliography by Jenn Wagner & Benjamin Muddiman

Classic & Background

Nee et al 1994

Extinction Rates can be Estimated from Molecular Phylogenies. Sean Nee, Edward C. Holmes, Robert M. May and Paul H. Harvey. *Philosophical Transactions: Biological Sciences*. 1994.

Key points: This paper examines the effects of various processes on the shapes of molecular phylogenies. The overall takeaway is that the shapes of molecular phylogenies can reveal information about the underlying evolutionary parameters at play, but this is a complex relationship where you might have multiple combinations of parameters yielding similar shaped trees (and lineage-through-time [LTT] plots). In addition, some events that alter parameters are likely extremely difficult to estimate from the shape of the tree alone, e.g. evidence of a mass-extinction is terribly difficult to pick up from a molecular phylogeny alone.

Some details: The effects of “the push of the past” and “the pull of the present” are examined by comparing the shapes of lineage-through-time plots of actual vs. reconstructed phylogenies generated under a simple constant rate birth-death model. Both of these effects are noted to increase as the ratio of *deaths/births* (d/b) increases towards unity. Nee et al. also use a likelihood surface approach, in which they visualize the likelihood of parameters of the birth-death model--in this case $b-d$ and d/b --to show that, while an LTT plot may indicate a 0 extinction rate (d), the actual likelihood that $d \neq 0$ may be significant. A *Drosophila melanogaster* phylogeny is used in this example. They also examine the effect of incomplete sampling of a group on the expected shape of LTT plots. From there, they suggest that one could infer actual number of lineages from the LTT plot shape (HIV is used as an example here).

Foote 1996

On the probability of ancestors in the fossil record. Mike Foote. *Paleobiology*. 1996.

Even under the most pessimistic models of fossil preservation and recovery (homogeneous preservation/recovery), the probability of having Ancestor-Descendent pairs in the fossil record is 1%. Under more realistic models (e.g. including rates that vary between clades, through time, etc.), the probability is 10% or greater. Foote compares models of *budding*, *bifurcating* and *phyletic transformation* evolution in terms of their impacts on the recovery of A-D pairs. Overall, this paper highlights the importance of developing methods for identifying A-D pairs in our fossil collections. This paper is important to age estimation and phylogenetic models broadly, because non-Fossilized Birth Death models (node-based and tip-based) do not incorporate A-D pairs.

Rutschmann 2006

Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. Frank Rutschmann. *Diversity and Distributions*. 2006.

This paper reviews a wide variety of methods for estimating divergence time using models of molecular evolution. The methods are broken up into three main categories: (1) methods using molecular clock + one global substitution rate (= # substitutions/site/Myr); (2) methods that correct for rate heterogeneity; and (3) methods that try to incorporate rate heterogeneity. Methods in (1) use a strict clock, while (2) and (3) use forms of relaxed clocks. Methods in (2) attempt to correct for rate heterogeneity by using local rates (creating rate classes and then applying different rates across the tree) or by pruning branches. Within (3), a variety of methods for introducing rate heterogeneity across a tree are given, including: temporal autocorrelation in rates, the stationary process of rate change, and the compound Poisson process of rate change. Note that, because rates include a “per Myr” component, rates and divergence times are confounded, such that assumptions made about one will influence the other. The methods in (3) seek to manage this by estimating branch lengths while allowing rates to vary across the tree, and then to model divergence times and rates by “minimizing the discrepancies between branch lengths and the rate changes over the branches”. Within group (3), some methods require only a phylogram with branch lengths as input; some require a single “true” tree as input; and some use sequence data and model parameters to generate ultrametric trees, and are therefore able to incorporate topological uncertainty.

Stadler 2010

Stadler, T., 2010, Sampling-through-time in birth–death trees: *Journal of Theoretical Biology*, v. 267, p. 396–404, doi:10.1016/j.jtbi.2010.09.010.

This paper is relevant because it walks through the equations used to determine the prior that which can be used in Bayesian inference methods and demonstrates the density of a given tree with sampled extant and extinct individuals must be estimated when using data sets that ‘sample-through-time.’ For most of the paper, Stadler is explaining the equations used in estimating the tree density in order to use it as a prior in Bayesian inference methods. She discusses the method, relevance, and application of each metric. The probability density function provides the ability to infer maximum likelihood birth- and death-rates, speciation and extinction rates respectively, for given reconstructed trees by maximizing the density function over the parameters. This is important for reliability as the death estimates are not efficient when only present-day data is included (Rabosky, 2010). Tree densities are required for Bayesian analysis assuming a birth-death model in order to estimate posterior trees of present day on a calendar-timescale.

Quental & Marshall 2010

Quental, T.B., and Marshall, C.R., 2010, Diversity dynamics: molecular phylogenies need the fossil record: *Trends in Ecology & Evolution*, v. 25, p. 434–441, doi:10.1016/j.tree.2010.05.002.

Quental and Marshall suggest molecular phylogenies are only partially correct and the fossil record allows to model for substantially negative diversification rates. They argue that a wide range of processes can yield similar phylogenies and using a model that allows substantially negative diversification rates to account for the observed decreases in diversification.

The Cetacea example showed the molecular phylogenetic perspective was very different than the paleontological perspective. The molecular results suggested a non-significantly negative speciation value, and many analyses suggested multiple diversification events, implying exponential growth. It also implies an extinction rate of 0. The paleo data analysis involved using fossils, separated them based on genera, and assumed there was no difference in biological processes between Cetacea in the past and present. The results suggested the Cetacea had maintained a steady diversity of the last 12 million year, assuming a similar number of species; the fossil record probably underestimates the diversity during that time due to fossil preservation issues. Origination rates and extinction rates can be calculated using the fossil record; these results are very different than the molecular data- there is very high extinction. It supported the hypothesis that the Cetacea were in decline rather than increasing their diversity. The molecular differs from the fossil data due to assuming the average rate of diversification is related to the current or other past rates of diversification; this is unrealistic due to the assumption being held uniform. If the tree has stem-group taxa (predates the last common ancestor of extant), they don't need to be compared with the molecular estimates.

Node- & Tip-based (Total Evidence) Divergence Time Estimation

Ronquist et al. 2012

Ronquist, F., Klopstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., and Rasnitsyn, A.P., 2012, A Total-Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera: *Systematic Biology*, v. 61, p. 973–999, doi:10.1093/sysbio/sys058.

The goal of the paper is to compare node dating with using fossils and extant taxa via Bayesian **total-evidence analysis** using a Hymenoptera radiation event to characterize a phylogeny. Ronquist et al. address the shortcomings of node calibrations, specifically the issue of associating a fixed node with calibration data even though the nodes are not known with absolute certainty, the dating only extracts calibration info from the oldest fossil assigned to that group. The data from the fossil record (ages/morphology) must be switched to appropriate probability distributions; it is likely for a fossil to be on a side branch of an unknown length, not directly on the calibration node. The node dating used several calibration points based on the fossil record and the total-evidence dating used 343 morphological characters from 45 fossils and extant taxa. A uniform prior was created using fossils and was used to model speciation, extinction, sampling, and preservation as realistically as possible. They used a relaxed clock model in order to incorporate changing diversification rates. The posterior distributions on the

estimated divergence times are less affected by the prior assumptions, leading to a more accurate date than in node dating. Their results suggest that using total evidence dating is preferred, especially when there is access to the fossils and natural history of the organisms under study. They emphasize that the focus on divergence time estimates should be aimed at intense fossil morphological studies, understanding morphological evolution, and improving the models of rate variation over time. This study also shows the importance of modeling the sampling of extant taxa when looking at species-rich groups.

O'Reilly et al. 2015

Dating Tips for Divergence-Time Estimation. Joseph E. O'Reilly, Mario dos Reis, and Philip C.J. Donoghue. *Trends in Genetics*. 2015.

Key points: O'Reilly et al. point out that, currently, tip-based calibration is more sensitive to root time prior densities and yields less precise divergence time estimates than has been claimed (in part, due to propagation of uncertainty in fossil-dating), and has performed worse than node-calibration in many examples. Tip-based reconstructions also consistently yield older divergence dates than node-based methods. O'Reilly et al. also argue that node-based calibrations, such as biogeographic calibration, could be incorporated into tip-based analyses for improved results. (The components of tip-dating/total evidence are: "(i) the relaxed morphological clock, (ii) tip-calibration, and (iii) co-estimation of time and topology"--none of which are mutually exclusive with node-based dating methods.)

Some details: The paper first covers some known issues with node dating, namely: prior phylogenetic hypotheses (that place the fossils) are required; and methods for establishing *maxima* for clade are difficult, as these tend to require negative evidence (options include: birth-death models, statistical modeling based on stratigraphic distribution of fossils, and "taphonomic controls"). By comparison, tip dating does not require *ad hoc* placement of fossils; older fossils that could not be reasonably assigned to nodes *a priori* can be included in the analysis; age calibration can be based on all included fossils, rather than having specific fossils that calibrate specific nodes; tip calibrations only relate to the age of a single species, which means that you aren't interpreting the absence of a fossil of that species as negative evidence indicating the absence of that clade (as you do if you put a prior on a node-based calibration). Despite these benefits, calibrating tip ages of fossils is not straightforward: stratigraphic constraint, choice of density distribution for age estimation, and how to handle multiple occurrences of the same species are all complicating factors. In addition, there is the "stem-ward slippage" phenomenon in which non-random loss of phenotypic data of fossils tends to cause them to associate with "more primitive evolutionary grades", and the incorporating apparent morphological stasis into estimates of rates of morphological evolution.

Brown & Smith 2018

The Past Sure is Tense: On Interpreting Phylogenetic Divergence Time Estimates. Joseph W. Brown and Stephen A. Smith. *Systematic Biology*. 2018.

Brown & Smith examine issues with node-based dating through the lens of a phylogeny of angiosperms. They find that issues with the model itself may preclude convergence between paleontological data (oldest uncontroversial angiosperm fossils are Early Cretaceous) and model based age reconstructions (which tend to be older, ranging into the Triassic). Specifically, the joint prior (the tree prior + the calibration densities from the fossil-calibrated nodes) alone precludes the reconstruction of an Early Cretaceous origin. The fact that, despite having 24 calibrated fossils, they found almost no difference between marginal prior and posterior prior age distributions (meaning that the age reconstructions were informed almost entirely by the prior, and not updated during the model run, as one would hope they would be in a Bayesian framework) indicates that even a substantial amount of data was unable to “overcome” the prior. Brown & Smith advocate for prior/posterior comparisons, and point out that the issue raised here should be added to two other critical issues in estimating divergence times, raised by Beaulieu et al. (2015): “methods of divergence time inference can be misled by both (i) extreme state-dependent molecular substitution rate heterogeneity and (ii) biased sampling of representative major lineages.” The concern over the ability of fossil age information to “overcome” priors does not directly apply to Fossilized Birth Death models, as in these, ages are estimated differently, as a part of the overall model (see Heath et al. 2014).

Fossilized Birth Death

Heath, Huelsenbeck, & Stadler 2014

Heath, T.A., Huelsenbeck, J.P., and Stadler, T., 2014, The fossilized birth–death process for coherent calibration of divergence-time estimates: *Proceedings of the National Academy of Sciences*, v. 111, p. E2957, doi:10.1073/pnas.1319091111.

Heath et al. focus on using a ‘fossilized birth-death’ (FBD) process to calibrate estimates of species divergence dates. They discuss the efficiency of using FBD over the common calibration methods where a decent amount of the fossil data isn’t included which results in an arbitrarily chosen parametric distribution. The node calibration provides a valid minimum-age constraint for that node, but if the calibration node isn’t specified properly the conclusions may not be valid. Creates a probability distribution on the age of each calibrated node that is derived from the node specific calibration prior and tree wide prior on node ages which causes inconsistency in the results. The tree-wide prior distribution on branch times (MODEL FOR THE TREE AND SPECIATION TIMES) is inferred from a birth-death process of cladogenesis. Inconsistency can be avoided by separating the nodes and applying a birth-death process to uncalibrated nodes based on the calibrated nodes, which they express is preferred when using fossil data. Calibration priors aren’t created using information from biological processes or information from the fossil record, resulting in use of arbitrary criteria leading to arbitrary divergence times; in order to avoid this, included should be preservation history and the rate at which the fossils occur in the rock record. Heath et al. provide an alternative method for calibrating phylogenies using fossils to create parameters for a speciation-extinction-fossilization model (the FBD process). Including the speciation rate, extinction rate, fossil recovery rate, and proportion of sampled extant species generates the amount of uncertainty for each speciation event on the tree, which ultimately recognizes the extant and extinct individuals are observations from the same diversification process. Essentially, the FBD model yields reliable estimates of absolute divergence times, while node age estimates are prone to biased sampling strategies of extant and fossil taxa.

Zhang et al. 2015

Total-Evidence Dating under the Fossilized Birth–Death Process. Chi Zhang, Tanja Stadler, Seraina Klopstein, Tracy A. Heath, and Fredrik Ronquist. *Systematic Biology*. 2015.

Zhang et al. demonstrate that correctly accounting for sampling strategy has significant effects on age reconstruction under a Fossilized Birth Death (FBD) model. Specifically, they highlight the fact that taxon sampling has previously been modeled as being random, when most taxonomic sampling for phylogenetic reconstruction is actually “diversified”. (Diversified sampling refers to coverage that includes representatives chosen to span the clade(s) of interest, rather than a random subset of all taxa within the clade(s)). By explicitly modeling diversified sampling, Zhang et al. find that divergence times are pushed forward, leading to less of a “rocks vs. clocks” discrepancy. As an example, they reanalyze the Hymenoptera phylogeny used by Ronquist et al. (2012) and date that radiation as being end-Permian, only slightly older than the oldest known fossils, and significantly younger (~50 Myr younger) than the Carboniferous estimate given by Ronquist et al. In their modeling and model comparison, Zhang et al. utilize a piecewise constant version of the FBD model, in which speciation, extinction, and fossil recovery rates can differ from one interval to another. Zhang et al. use posterior/prior comparisons and find that their posteriors are informative, indicating that their model is able to pick up signal in the data. Additionally, Zhang et al. introduce or update some MCMC based methods for FBD trees, including: an add-branch move, a delete-branch move, a node-slider mechanism, and a subtree-pruning-and-regraft (SPR) moves mechanism.

Gavryushkina et al. 2016

Gavryushkina, A., Heath, T.A., Ksepka, D.T., Stadler, T., Welch, D., and Drummond, A.J., 2016, Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins: *Systematic Biology*, v. 66, p. 57–73, doi:10.1093/sysbio/syw060.

Gavryushkina et al. address the issue that model-based implementations of divergence time dating do not use models that address the diversification and fossilization process by using a ‘total evidence method’ through a Bayesian framework. The model incorporates stem fossil diversity in order to calibrate the nodes based on a probability distribution of the speciation/extinction events. They address previous use of the FBD method (Heath et al. 2014, Ronquist et al. 2012) and take it a step further by making the time of origin (T_{OR}) a model parameter rather than the age of the MRCA (the oldest splitting node leading to the extant species) leading to all fossils being assumed to be descendants of that node. Making it a model parameter allows for the oldest fossils to be a direct ancestor or sister lineage because the prior evidence does not rule that possibility out. The FBD method implemented in this model includes diversification, fossilization, and sampling but doesn’t assume descendants. Gavryushkina et al. applied this to extant and fossil penguin data using molecular data and fossil occurrence dates in addition to a model of morphological trait evolution. They estimated the origin of crown penguins to be ~12.7 Ma. Previous estimates include ~40.5 Ma (Baker et al. 2006), ~50 Ma (Brown et al. 2008), ~20 Ma (Subramanian et al 2013), and 23 Ma (Jarvis et al. 2014; Li et al. 2014). In addition to this, the results from this paper suggest

that most branching that lead to modern day penguins happened ~ 2 Ma, contradictory to what has previously been published. The estimate predicted in this paper is claimed to be more reliable because it includes stem fossil diversity which accounts for the possibility of ancestor-descendent relationships, generates a probability density that includes fossil, molecular, and morphological data and analyses it in tandem (rather than sequentially) to provide a dated tree estimate from ALL available data, and models for a diversification, fossilization, and sampling processes. They suggest future improvements for this model that include incorporating a changing fossil sampling rate, better modeling of morphological trait evolution, and models that permit lineage-dependent or age-dependent speciation and extinction rates while modeling fossil sampling.