Molecular clocks without rocks: new solutions for old problems

George P. Tiley[1], Jelmer W. Poelstra[1], Mario dos Reis[2], Ziheng Yang[3], Anne D. Yoder[1,*]

[1]Department of Biology, Duke University, Durham, NC 27708, USA

[2]School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, UK

[3]Department of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, UK

[*]Correspondence: anne.yoder@duke.edu

Keywords: Divergence Time Estimation, Multispecies Coalescent, Substitution Rate, Mutation Rate, Effective Population Size, Gene Tree Discordance

Abstract

Molecular data have been used to date species divergences ever since they were described as "documents of evolutionary history" in the 1960s. Yet, an inadequate fossil record and discordance between gene trees and species trees are persistently problematic. We examine how, by accommodating gene tree discordance and by scaling branch lengths to absolute time using mutation rate and generation time, multispecies coalescent (MSC) methods can potentially overcome these challenges. We find that time estimates can differ — in some cases, substantially — depending on whether MSC methods or traditional phylogenetic methods that apply concatenation are used, and whether the tree is calibrated with pedigree-based mutation rates or with fossils. We discuss the advantages and shortcomings of both approaches and provide practical guidance for data analysis when using these methods.

Divergence Time Estimation

Zukerkandl and Pauling [1] were the first to posit that genetic distances between organisms could be converted to absolute geological times, describing genomes as "documents of evolutionary history." The most commonly used **molecular clock** methods (See Glossary) estimate absolute times from genetic distances by calibrating the **species tree** with fossil data, assuming either a constant rate of evolution among lineages (the molecular clock) or variable rates (**relaxed clock models**) [2-4]). Recently, the **multispecies coalescent (MSC)** [5] is on the ascent as a method for estimating divergence times [6, 7] due at least in part to potential freedom from **fossil calibrations** [8, 9]. However, conflicts can arise in empirical studies between **traditional phylogenetic clock models** and MSC methods, raising the question of which method is more reliable for placing evolutionary events in a temporal context. Here, we examine the fundamental assumptions and analytical details of these two general methodological classes: 1) Traditional phylogenetic clock models that use **concatenation** of genetic **loci** and 2) MSC models that explicitly model **gene tree discordance** due to **incomplete lineage sorting (ILS)**. Both approaches can be used without fossil calibrations where *a priori* information on absolute rates of evolution are available, but some features of the MSC are ideal for estimating species divergence times by leveraging external ***de novo*** **mutation rate (μ)** estimates, typically measured from pedigree trios (Fig. 1; Key Figure). We conclude by describing conditions that influence the suitability of the two approaches and offer recommendations for the proper application of both.

The Allure of the Molecular Clock

Clock models and their applications have had enormous impacts on our understanding of the history of life on earth, including the timing of life history transitions [10], global ecological change in response to climate oscillations [11], the ancient origins of orders such as Lepidoptera [12], and even the origin of life or the last universal common ancestor (LUCA) after

the Moon-forming impact [13]. Calibration of the molecular clock has historically been performed using fossil ages [14] or geological events [15] though only a tiny fraction of phylogenetic lineages have reliable fossil records for appropriate calibration [16]. Thus, the lack of a detailed fossil record for many groups is a major constraint for investigating the evolutionary history of those lineages. For instance, both plant [17] and animal fossils [18] are difficult to characterize in tropical rainforests where available rock formations for fossilization are typically absent [19]. In some groups, such as grasses, calibrations based on phytolith microfossils are contentious because of ambiguous diagnostic characters that compromise accurate phylogenetic placement [20]. And for many groups, such as the glass frogs [21], fossils are entirely absent.

Thus, for the analysis of most clades across the tree of life, investigators must depend on fossil calibrations that are phylogenetically distant from the organisms of interest. As phylogenetic distance increases, the complexities of modeling rate variation among lineages also increases given the now extensive evidence that molecular rates change frequently across phylogeny. Finally, a growing body of literature suggests that by ignoring genetic polymorphism in ancestral species, divergence times may be systematically biased [6, 7, 9].

Relaxed Clock Models

When a calibration point can be placed with confidence within a given clade and close to its **most recent common ancestor (MRCA)**, it is possible to estimate **per-year substitution rates**. Using that rate, an investigator can then infer divergence times for other nodes in the phylogeny that do not have fossil calibrations. This assumes, however, that all lineages share a single rate of evolution: i.e., that there is a **strict molecular clock**. While this is not an unreasonable assumption for closely related species, the strict clock is typically violated when more distantly related species are included [22]. Such violations can arise not only from

differences in the molecular mechanisms that generate mutations [23], but also from variation in life history traits [24, 25]. For example, great apes have lower substitution rates compared to Old World and New World monkeys (the hominoid slowdown hypothesis [26]), a phenomenon that can largely be explained by differences in **generation time** among species [27]. Similar observations have been made in plants by comparing woody and herbaceous species [28, 29].

The clock can be "relaxed" by allowing for variable rates among branches on a phylogeny while maintaining computational tractability and statistical identifiability [2, 3, 30, 31]. The first relaxed clock methods that could leverage uncertainty across multiple calibrations were implemented with maximum likelihood and required *a priori* assumptions to partition branches into different rate groups (e.g., "local clocks" [32] or heuristic approaches [33, 34]). Recent **Bayesian methods** have incorporated uncertainty in calibrations and as well as in rates of evolution through the use of **prior distributions**. Different models of rate variation among branches are available, including autocorrelation among lineages [2, 35], uncorrelated rates [3, 30, 36], or a mixture of the two [37]. However, per-year substitution rates and divergence times are sensitive to prior distribution on node calibrations [38] and justifying informed node calibrations is not trivial [39]. Relaxed clock methods have recently been extended to account for uncertainty in fossil placement [40] by leveraging morphological data from both extant and fossil species [41-44]. These **total-evidence** [41] approaches include **tip-dating methods** that treat extinct fossil lineages as tips where fossil occurrence [40] or morphological characters from fossils [43] can calibrate rates of evolution to absolute time. They can also incorporate different speciation mechanisms that best suit an organismal group [45]. As with more traditional methods, however, these total-evidence tip-dating methods can only be applied to clades with an available fossil record [42] and therefore cannot solve the problem of poor or absent fossil records.

Tip-dating methods that use only molecular data [46, 47] offer one approach for overcoming an absence of fossil calibrations. These methods have been applied to viruses, where high substitution rates generate sufficient variation from contemporary samples to determine relative ages [48], as well as to cases wherein ancient DNA samples can calibrate the molecular clock such as for woolly mammoths [49] and humans [50]. Even so, **ancient DNA methods** are equally or even more restrictive than fossil-calibrated methods given that they can only be applied to a limited number of organisms for which well-preserved and relatively recent samples are available [51]. Most significantly, all of the above methods use **concatenation** of genetic loci**,** thereby making the fundamental assumption that the phylogenetic history of each locus matches the species tree. We here discuss how concatenation can be problematic, and how MSC methods overcome these problems.

The Multispecies Coalescent as a Backward Time Machine

Coalescent theory is a branch of population genetics that describes the genealogical histories of a sample of alleles in a population, going back from a sample of extant alleles to their most recent common ancestor [52]. Two alleles are said to coalesce when they share a common ancestor. The MSC is a simple extension of the single-population coalescent to multiple species [5] and accommodates the species phylogeny and the coalescent processes in both the extant and extinct species [53, 54]. The MSC jointly estimates **divergence times** and rates of evolution (Fig. 1) while explicitly modeling gene tree discordance due to incomplete lineage sorting (ILS, also known as deep coalescence). ILS occurs when sequences from different species fail to coalesce in their most recent ancestral species. The shorter the branch in coalescent units between two speciation events, the more likely is ILS to occur (Box 1). Short coalescent branch lengths can be caused not only by small time intervals between speciation events, but also by large ancestral **effective population size**.

131 It is now well accepted that gene trees do not consistently match species trees [55]. Though this
132 was initially considered to be a hindrance to the accurate reconstruction of phylogenies [56],
133 investigators are increasingly aware that these heterogeneities provide valuable information
134 about the timing and population dynamics of organismal lineages over their evolutionary history.
135 Described as a "backward time machine" [57], the MSC treats the stochastic variation of the
136 coalescent process over genes or genomic regions as a source of information rather than as
137 "mistakes" or "conflicts", and is thus uniquely suited to harness the power of many loci from
138 modern genomic data. Accordingly, the MSC is of increasing interest to investigators who seek
139 to place divergence events in a temporal context. The MSC makes a number of simplifying
140 assumptions including a lack of post-divergence **gene flow**, ILS as the only source of gene tree
141 discordance, no recombination within loci, and a lack of selection. Where high amounts of gene
142 flow among non-sister species are a concern, extensions to the MSC are available [58].
143
144 **Accounting for the Coalescent Process**
145 Traditional phylogenetic clock models equate species divergence (i.e., "split times") to sequence
146 divergence. This is problematic given that sequence divergence will always predate speciation
147 events in the absence of gene flow [59, 60] (Fig. 2). In contrast, coalescent methods explicitly
148 accommodate the differences between the two and directly estimate *species* divergence times
149 which are generally the evolutionary events of interest (Fig. 3; Box 2). Moreover, when fossil
150 calibrations are used, divergence time estimates can be strongly affected, with the direction of
151 the bias depending on the placement of the most precise calibrations. If these calibrations are
152 placed on young nodes within a phylogeny, divergence times will be underestimated across the
153 entire phylogeny, while if calibrations are placed on ancient nodes, the ages of young nodes are
154 likely to be overestimated. Accordingly, for phylogenies with complex mixtures of fossil
155 calibrations, both underestimation and overestimation of divergence times may occur across the
156 phylogeny – regardless of the analytic method applied.

157

158 Traditional phylogenetic analysis of concatenated sequences assumes that a single tree
159 topology with one set of divergence times underlies the multilocus sequence data, irrespective
160 of how rate variation is modeled among sites, loci, or branches. Gene tree discordance due to
161 ILS then appears as additional substitutions on branches in the species phylogeny [6, 61],
162 leading to overestimation of species divergence times when ILS is not accounted for [9]. In line
163 with these theoretical expectations, Stange et al. [7] showed that in cases of high gene tree
164 discordance, concatenation methods will overestimate ages of young nodes when ancient
165 nodes are constrained. Similarly, Fang et al. [62] found that recent species divergences were
166 correctly estimated to be more recent when using MSC methods. Simulations generally suggest
167 that the MSC can improve divergence time estimates when gene tree discordance is high [7, 9],
168 while comparable performance should be expected between concatenation and MSC methods
169 when gene tree discordance is low (Fig. 3).

170

171 Although empirical studies using MSC approaches have thus far focused on recent species
172 divergences (1-10 MYA) [7, 62, 63], the effects of discordant gene trees should also impact
173 divergence time estimates for older divergences where the coalescent branch length is short
174 and ILS is high [9]. These patterns are expected for rapid radiations that occurred deep in
175 evolutionary history, such as placental mammals [64], passerine birds [65], and lepidopterans
176 [12]. Divergence time estimates for these groups are important for interpreting species
177 biogeography and trait evolution, and as computational efficiency and resources continue to
178 improve, the evolutionary history of these groups should be re-evaluated with MSC models that
179 also leverage fossil calibrations. In angiosperms, reconciliation of molecular dates with those
180 interpreted from the fossil record has been the topic of vigorous debate even though molecular
181 data have largely been restricted to chloroplast sequences, which represent a single **gene tree**

[66-69]. As large multilocus nuclear datasets become increasingly available for plants [70], the benefits of fossil-calibrated MSC methods could be realized.

The Coalescent Time Unit

Because the average coalescence time between two randomly sampled sequences from a diploid population is $2N$ generations, it is convenient to scale branch lengths in the species tree in coalescent units, that is, to use $T = t/(2N)$ where t is the number of generations until the coalescent event. T can also be rescaled by mutations and represented as $\tau = \mu t$, where μ is the per-generation mutation rate, so that $T = \tau/(\theta/2)$. Here $\theta = 4N\mu$ is the population-scaled mutation rate; a fundamental parameter in population genetic models which represents the average number of mutations per site between two sequences randomly sampled from the population.

MSC programs like StarBEAST2 [6] and BPP [5, 71] use multilocus sequence alignments to estimate species trees as well as parameters in the MSC model including species divergence times (τ) and population sizes (BPP estimates θ and StarBEAST2 estimates $N\mu$), both measured by the expected number of mutations per site. If fossil calibrations or mutation rates are available to calibrate the tree, they can be used to convert genetic distance to absolute times and absolute rates. When a per-generation mutation rate is available, generation times are also necessary (Fig. 1) to convert to divergence times in years. This approach assumes that the per-generation mutation rate and generation time are constant throughout the species tree, which is a reasonable assumption for analyses of closely related species for which genetic divergences likely satisfy a strict clock [72, 73].

***de novo* Mutation Rate Estimates Provide Freedom from the Fossil Record**

In order to estimate absolute divergence times in the absence of fossil calibrations, direct estimates of the mutation rate estimates are needed. Recently, whole-genome sequencing data from **pedigree trios** have been used to estimate the ***de novo* mutation rate** for many animals [74-78] and parent-progeny pairs in plants [79]. Recent examples of divergence time estimation based on mutation rates and **coalescent age estimates** include the age of human migration events [80] and of domestication histories among agriculturally important species [81-83].

To estimate a *de novo* mutation rate, the father, mother, and offspring from a pedigree trio are sequenced and aligned to a reference genome. Variants detected in the child that are distinct from both the mother and father and do not match the reference are considered *de novo* mutations. Because the number of sequencing errors are more than an order of magnitude greater than the number of true mutations, strict filtering criteria in computational analysis must be applied to the called variants to avoid false positives. Also, mutations cannot be identified at all sites because of variable sequencing read depth and alignment uncertainty in repetitive regions. Thus, the number of **callable sites** needs to be estimated as the denominator to accurately estimate μ [76]. Ideally, the final estimate of μ is averaged over multiple pedigrees, as any single pedigree will yield few mutations. Best practices for reducing false positives and false negatives for inferred mutations are still being developed [84].

The availability of a reference genome can be a critical limitation for estimating *de novo* mutation rates in non-model organisms. Although high-quality reference genomes are anticipated for most Eukaryotic lineages in the near future [85], there will ultimately be barriers for some groups. In the absence of direct estimates of μ for a species of interest, distributions of μ can be developed based on studies of related organisms [72]. Generation time estimates must be considered as well given that mutation rates from pedigree studies are scaled by generation to recover absolute divergence times (Fig. 1).

Discrepancies between Concatenation and MSC methods for Divergence Time Estimates

Though empirical examples are as yet few, discrepancies between divergence dates estimated by fossil-calibrated concatenation and mutation rate-calibrated MSC methods are emerging (e.g., Fig 4). For the closely related species pair of human and chimpanzees, the mutation rate-calibrated MSC [9] and concatenated time estimation give similar results. Fossil-calibrated concatenation and fossil-calibrated MSC methods place the divergence between 5.7 and 10 MYA, typically near the center of the calibration density at 7.5 MYA [9, 86]. A mutation rate-calibrated MSC analysis that assumed the human mutation rate for both species recovered a posterior mean of 8.2 MYA [9]. Divergence time estimates calibrated directly with mutation rates but not using the MSC are also similar, but only after considering the difference between species and sequence divergence. In one such study, **pairwise sequence divergence** between chimp and human (t_{Seq}; Fig. 2) yielded a divergence time of 12.1 MYA assuming the human mutation rate [27], though subtracting $2N_{HC}$ (the effective population size for the human-chimpanzee common ancestor) yields a divergence time of 7.9 MYA. Thus, per-generation mutation rates can be used to estimate divergence times from concatenated data too, but the difference between species divergence and sequence divergence needs to be accommodated by some calculation of population size estimate (Fig. 2).

The sensitivity of these methods to the mutation rate estimate is keen. For example, when the human mutation rate was applied unilaterally across a primate phylogeny, a divergence time between Old world monkeys (*Macaca mulatta*) and humans of 62 MYA was recovered [87], in stark contrast to the 35 MYA age estimate indicated by fossil evidence [86]. The discrepancy is likely explained by a slower mutation rate in humans compared to Old World monkeys [88] and indicates that caution is needed when applying pedigree-based mutation rates to divergence time estimation, especially across large phylogenies. While one possible reason for

discrepancies across long times scales is that purifying selection may lead to lower substitution rates compared to mutation rates [89], as observed in mutation accumulation lines with *Arabidopsis thaliana* [90], the discrepancy in this case was in the opposite direction. Thus, given the small number of empirical examples at present, it is difficult to generalize the causes of disparities between substitution and mutation rates at present.

In one such empirical example, MSC methods produce significantly more recent age estimates than fossil-calibrated concatenation methods for mouse lemurs (genus *Microcebus*). Whereas a mutation rate-calibrated MSC analysis yields a MRCA for the genus of 1.5 MYA [63], previous analyses using fossil-calibrated concatenation methods yielded estimates of ~10 MYA [86, 91]. Though this discrepancy could, in part, be the consequence of a falsely elevated pedigree-based mutation rate estimate [91], the discrepancy would still be pronounced even if the true rate is only half of the estimated rate. Conversely, for the fossil-calibrated divergence time estimate using concatenation, phylogenetically distant, external calibrations [86, 91] were used by necessity given that there is a complete dearth of fossils within the lemuriform clade. As described above, the fossil-calibrated concatenation estimate is thus likely to overestimate divergence times for young nodes given the dependence on older fossil calibrations deeper in the phylogeny (Fig. 3; [9]). This is similar to cases of Stange et al. [7] and Fang et al. [62] where MSC methods using geological calibrations resulted in more recent divergence times compared with those found with concatenation – even when using the same calibrations. In summary, it is important to note that the differences in time estimates between the MSC and phylogenetic concatenation methods may be complex, depending on biases of mutation rate estimates and on the relative placement of calibrations within the phylogeny.

A New Frontier in Divergence Time Estimation

Divergence time estimates can fundamentally affect interpretations of trait evolution, biogeography, and the processes that underlie species radiations. Thus, the stakes for evolutionary studies are high. As an important step forward, future studies that leverage genomic data and fossil calibrations should consider comparing traditional phylogenetic clock models and the MSC to evaluate the effects of ILS on divergence time estimation. We further recommend that uncertainty in both mutation rates and generation times should be explicitly incorporated in analyses wherein coalescent units are converted to absolute time [63, 72]. This can be easily done by drawing mutation rates and generation times from prior distributions rather than relying on point estimates, given that variation in inferred mutation rates can be high among pedigrees [84], and mutation rates may change over time [27]. Moreover, estimating generation times can be problematic, especially for perennial plants given the lack of clear segregation in the germ line. The impact of the number, quality, and placement of fossil calibrations – as well as model choice on divergence time estimation using traditional phylogenetic concatenation methods – has been extensively studied [10, 38, 67, 69, 86]. Conversely, the careful evaluation of MSC methods for divergence time estimation is still in its infancy. We therefore predict that future studies that directly compare the two approaches are likely to identify as yet unrecognized though critical considerations for accurate divergence time analysis.

Concluding Remarks

We conclude by noting that despite its advantages, the MSC method involves a heavy computational burden and may not always be feasible for divergence time estimation on large phylogenies [92-94]. In such cases, traditional phylogenetic clock analyses that use concatenation may be the most practical approach [3, 30]. In particular, approximate likelihood calculation appears useful in estimating divergence times for large phylogenies or for very long alignments [86]. These models should not be seriously biased when divergence times are old

(Fig. 2) and ILS is low (Fig. 3). But given the prevalence of ILS across the tree of life, the applications of the MSC for divergence time estimation in both shallow and **deep phylogenies** will be of increasing interest and importance (see Outstanding Questions). It remains to be seen to what degree divergence time estimates will agree when both traditional phylogenetic clock models and mutation-rate calibrated MSC methods are applied within the same study systems.

Acknowledgements

We thank Priya Moorjani for her helpful comments on an early version. The editor and two anonymous reviewers provided comments that improved the manuscript. ADY gratefully acknowledges the John Simon Guggenheim Foundation and the Alexander von Humboldt Foundation for their support. ZY is supported by a Biotechnology and Biological Sciences Research Council grant (BB/P006493/1).

Glossary

Ancient DNA Methods: Sequence data is obtained from the remains of ancient specimens. The DNA is typically damaged and fragmented by absolute time and by exposure to damaging agents such as heat, oxidation, and UV irradiation.

Bayesian Methods: Bayes theorem is used to approximate the maximum likelihood estimates of a model and its parameters by sampling many estimates with proposal distributions and commonly implemented with a Markov chain Monte Carlo algorithm. Prior distributions are used to constrain the search space of parameters and may include *a priori* expectations for parameter estimates but are often left vague. Bayesian methods are used as a matter of computational convenience when maximum likelihood optimization is intractable.

Callable Sites: The number of sites where a *de novo* mutation should be detectable.

Concatenation: Multiple loci are treated as a single nonrecombining locus with a single underlying topology.

Coalescent: The stochastic process of lineage joining when one traces the genealogical history of a sample of sequences from a population backwards in time.

Coalescent Age Estimate: The divergence time for two sequences based on sampling theory and measured in the expected number of generations.

Coalescent Time Unit: The expected coalescent time for a pair of sequences, which is $2N$ generations for a diploid species with population size N.

***de novo* Mutation Rate**: The spontaneous germline mutation rate revealed in comparisons of whole genomes from both parents and their progeny (aka, pedigree trios).

Deep Phylogenies: Phylogenies that contain species with high sequence divergence. In such cases, substitutional saturation or "multiple hits," long-branch attraction, gene duplication and loss, and model misspecification can result in gene tree discordance. ILS can still be a substantial source of conflict between gene trees and species trees in deep phylogenies.

Divergence Times: The expected absolute age at which two species became isolated from each other.

Effective Population Size: The number of individuals that would produce the observed rate of genetic drift in an idealized Fisher-Wright population model.

Fossil Calibrations: Fossil evidence from morphological characters that can constrain the age of the crown group of a clade (e.g., with a hard minimum and a soft maximum).

Gene Flow: Exchange of alleles between two populations.

Gene Tree: The evolutionary history of a short, nonrecombining segment of the genome

Gene Tree Discordance: Difference in gene tree topology from the species tree, possibly caused by deep coalescence.

Generation Time: The average time between two generations which is often quantified as the average age of parents at birth, averaged over individuals.

Lineage sorting: The process by which gene lineages become fixed within a species such that all alleles within that species coalesce to a single ancestral allele within the species lineage.

Incomplete Lineage Sorting (ILS): Failure for two sequences from two species to coalesce in the most recent common ancestral species, also known as deep coalescence.

Loci: Orthologous non-recombining sequences. Each locus corresponds to an independent gene tree.

Markov chain Monte Carlo (MCMC): a simulation approach for sampling from a target distribution such as the posterior distribution of parameters in Bayesian inference.

Molecular Clock: Hypothesis that the rate of molecular evolution is constant over time and among lineages.

Most Recent Common Ancestor (MRCA): The most recent node on a phylogeny from which all individuals in a clade of interest are derived.

Multispecies Coalescent (MSC): The extension of the coalescent process to multiple species which accommodates the species tree as well as the coalescent within populations.

Pairwise Sequence Divergence: The evolutionary distance between a pair of sequences measured as the expected number of substitutions per site.

Pedigree Trio: A child and the two parents for whom whole genomes are sequenced and compared to identify the new mutations in the child.

Per-year Substitution Rates: The number of substitutions per-site per-year that are obtained when calibrating a phylogeny to absolute time with information at nodes or tips.

Reciprocal Monophyly: All alleles within a species coalesce with each other before the first coalescence with an allele from another species.

Relaxed Clock Models: An extension of the strict clock model to allow changes in evolutionary rate over branches in a phylogeny.

Species Tree: The evolutionary history of species, which is often estimated from many individual gene trees or loci.

Strict Molecular Clock: A single rate of molecular evolution is enforced for all branches in a phylogeny.

Tip-Dating Methods: Rates of molecular evolution are calibrated to absolute time by known sampling dates of individuals, whether extant or extinct, at the tips of a phylogeny.

Total-Evidence: Morphological characters for extinct (fossil) and extant tips and rates of morphological evolution are used to infer species divergence times jointly with molecular data.

Traditional Phylogenetic Clock Models: Models for divergence time estimation that assume one tree and one set of divergence times for all loci.

Box Legends

Box 1 – Incomplete Lineage Sorting on a Rooted Three-Taxon Species Tree.

Box 2 – Differences between Bayesian methods for divergence time estimation and programs for implementing them.

Figure Legends

Figure 1 – Overview of the MSC model and its use for estimating absolute divergence times with external mutation rate data. *Input Data* – The MSC requires aligned orthologous sequence data as input. There can be many individual loci, and each locus is assumed to be non-recombining and not under selection. The MSC allows for multiple alleles per species per locus, and sampling multiple alleles can improve parameter estimates. Although joint estimation of the species tree and MSC parameters is possible, using a fixed species tree is computationally more efficient. *Estimate MSC Parameters by MCMC* – The MSC estimates model parameters with Bayesian **Markov chain Monte Carlo (MCMC)**. This requires a prior distribution for all model parameters including population sizes (θ), species divergence times (τ), and possibly rates among loci (r). The MSC estimates a gene tree for each locus (1..n) and the coalescent times on those trees. Gene tees can be incongruent with the species tree. The distribution of gene trees and their coalescent age estimates are used to jointly estimate θ and τ on the species tree. *Calibrate with Mutation Rate* – A per-generation mutation rate (μ) is obtained from independent pedigree-based studies and a generation time (g) is estimated from the distribution of parent ages at the time of birth for offspring. μ and g can then be used to obtain an absolute rate of evolution by multiplying τ by g/μ. μ can also be used to rearrange the expression for θ and obtain absolute population sizes (N).

Figure 2 – Overestimation of species divergence times from genetic distances. a) Species A and B diverged t_{AB} generations ago. Sequences A and B coalesced further back in time at t_{Seq} generations ago, with a mean $t_{Seq} - t_{AB}$ of $2N_{AB}$ generations. b) Sequences were simulated under the MSC for a pair of species with constant $N_{AB} = 10^5$ and $\mu = 1 \times 10^{-8}$ per site per generation. c) The relative expected overestimation of species divergence times by $2N_{AB}$ becomes smaller as t_{AB} becomes larger, because $2N_{AB}$ contributes only a small proportion of time to the overall divergence time estimate.

Figure 3 – Effects of ILS and the MSC on divergence time estimation. Data are from Table 2 of Angelis and dos Reis [9]. a) Three-taxon species tree used for simulation. All data were simulated under the MSC with the Jukes Cantor model of molecular evolution. A μ of 1×10^{-8} per site per generation and a generation time of ten years was used, and the species tree root (node r) had an age of 10 MYA or τ of 0.01. Data were simulated with 4 different population sizes (N) that were constant along the species tree. The root was calibrated with a gamma distribution as one might in a fossil-calibrated divergence time analysis. b) Divergence time estimates of node s when using concatenation (MCMCTREE) or the MSC (BPP). Because the calibration is placed on the older root node, the younger node is overestimated when N or ILS is high by concatenation but not MSC methods. Points represent posterior means and error bars are the 95% credible intervals.

Figure 4 – Illustration of the consequences of differences between divergence time estimates. Pedigree symbols represent mutation rate-calibrated divergence times and probability distributions represent traditional phylogenetic clock model estimates. a) The MRCA of Madagascar's mouse lemurs. The mutation rate-calibrated MSC estimate yields a mean divergence time of 1.5 MYA whereas a traditional phylogenetic clock model with fossil

445 calibrations recovers a divergence time estimate of approximately 10 Ma. Though the position of
446 Madagascar relative to Africa is essentially the same at these two geological time points,
447 Madagascar's climate would have been very similar to that of today at 1.5 MYA whereas it
448 would have been much warmer and drier 10 MYA. b) The divergence between Old World
449 monkeys and apes. A mutation rate-calibrated divergence time estimate (though not with the
450 MSC) is 62 MYA whereas the traditional phylogenetic clock model yields a divergence time
451 estimate of approximately 35 MYA. There are striking differences in both continental
452 configuration and climate between these two time points. At 62 MYA, the earth was largely
453 tropical and sea levels were markedly high, isolating Africa from the northern continents. At 35
454 MYA, Africa has shifted northward, making contact with the northern continents and Antarctica
455 is partially glaciated indicating much cooler global temperatures. Global maps provided courtesy
456 of the Deep Time Maps project.

1. Zuckerkandl, E. and Pauling, L. (1965) Molecules as documents of evolutionary history. J Theor Biol 8 (2), 357-66.

2. Thorne, J.L. et al. (1998) Estimating the rate of evolution of the rate of evolution. MBE 15 (12), 1647-1657.

3. Drummond, A.J. et al. (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4 (5), e88.

4. dos Reis, M. et al. (2016) Bayesian molecular clock dating of species divergences in the genomics era. Nat Rev Genet 17 (2), 71-80.

5. Rannala, B. and Yang, Z. (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164, 1645-1656.

6. Ogilvie, H.A. et al. (2017) StarBEAST2 Brings Faster Species Tree Inference and Accurate Estimates of Substitution Rates. Mol Biol Evol 34 (8), 2101-2114.

7. Stange, M. et al. (2018) Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. Syst Biol 67 (4), 681-699.

8. Burgess, R. and Yang, Z. (2008) Estimation of hominoid ancestral population sizes under bayesian coalescent models incorporating mutation rate variation and sequencing errors. Mol Biol Evol 25 (9), 1979-94.

9. Angelis, K. and dos Reis, M. (2015) The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. Current Zoology 61 (5), 874–885.

10. Liu, L. et al. (2017) Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. Proc Natl Acad Sci U S A 114 (35), E7282-E7290.

11. Hackel, J. et al. (2018) Grass diversification in Madagascar: In situ radiation of two large C3 shade clades and support for a Miocene to Pliocene origin of C4 grassy biomes. Journal of Biogeography 45 (4), 750-761.

12. Kawahara, A.Y. et al. (2019) Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. Proc Natl Acad Sci U S A 116 (45), 22657-22663.

13. Betts, H.C. et al. (2018) Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat Ecol Evol 2 (10), 1556-1562.

14. Benton, M.J. and Donoghue, P.C. (2007) Paleontological evidence to date the tree of life. Mol Biol Evol 24 (1), 26-53.

15. De Baets, K. et al. (2016) Tectonic blocks and molecular clocks. Philos Trans R Soc Lond B Biol Sci 371 (1699).

16. Rannala, B. (2016) Conceptual issues in Bayesian divergence time estimation. Philos Trans R Soc Lond B Biol Sci 371 (1699).

17. Wing, S.L. et al. (2009) Late Paleocene fossils from the Cerrejon Formation, Colombia, are the earliest record of Neotropical rainforest. Proc Natl Acad Sci U S A 106 (44), 18627-32.

18. Bloch, J.I. et al. (2016) First North American fossil monkey and early Miocene tropical biotic interchange. Nature 533 (7602), 243-6.

19. Kidwell, S.M. (2001) Major biases in the fossil record. In Paleobiology II (Briggs, D.E.G. and Crowther, P.R. eds), pp. 297-303, Blackwell Science Ltd.

20. Kergoat, G.J. et al. (2018) Opposite macroevolutionary responses to environmental changes in grasses and insects during the Neogene grassland expansion. Nat Commun 9 (1), 5089.

21. Castroviejo-Fisher, S. et al. (2014) Neotropical diversification seen through glassfrogs. Journal of Biogeography 41 (1), 66-80.

22. Langley, C.H. and Fitch, W.M. (1974) An examination of the constancy of the rate of molecular evolution. JME 3, 161-177.

23. Lynch, M. et al. (2016) Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet 17 (11), 704-714.

24. Li, W.-H. et al. (1996) Rates of nucleotide substitution in primates and rodents and the generation–time effect hypothesis. Molecular Phylogenetics and Evolution 5 (1), 182-187.

25. Mello, B. and Schrago, C.G. (2019) The estimated pacemaker for great apes supports the hominoid slowdown hypothesis. Evolutionary Bioinformatics 15, 1-8.

26. Goodman, M. (1985) Rates of molecular evolution: the hominoid slowdown. Bioessays 3 (1), 9-14.

27. Moorjani, P. et al. (2016) Variation in the molecular clock of primates. Proc Natl Acad Sci U S A 113 (38), 10607-12.

28. Smith, S.A. and Donoghue, M.J. (2008) Rates of molecular evolution are linked to life history in flowering plants. Science 322 (5898), 86-9.

29. Lanfear, R. et al. (2013) Taller plants have lower rates of molecular evolution. Nat Commun 4, 1879.

30. Rannala, B. and Yang, Z.H. (2007) Inferring speciation times under an episodic molecular clock. Systematic Biology 56 (3), 453-466.

31. Ho, S.Y. and Duchene, S. (2014) Molecular-clock methods for estimating evolutionary rates and timescales. Mol Ecol 23 (24), 5947-65.

32. Yoder, A.D. and Yang, Z.H. (2000) Estimation of primate speciation dates using local molecular clocks. MBE 17 (7), 1081-1090.

33. Sanderson, M.J. (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. MBE 14 (12), 1218-1231.

34. Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. MBE 19 (1), 101-109.

35. Huelsenbeck, J.P. et al. (2000) A Compound Poisson Process for Relaxing the Molecular Clock. Genetics 154, 1879-1892.

36. Lepage, T. et al. (2007) A general comparison of relaxed molecular clock models. Mol Biol Evol 24 (12), 2669-80.

37. Lartillot, N. et al. (2016) A mixed relaxed clock model. Philos Trans R Soc Lond B Biol Sci 371 (1699).

38. Warnock, R.C. et al. (2015) Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. Proc Biol Sci 282 (1798), 20141013.

39. Parham, J.F. et al. (2012) Best practices for justifying fossil calibrations. Syst Biol 61 (2), 346-59.

40. Heath, T.A. et al. (2014) The fossilized birth-death process for coherent calibration of divergence-time estimates. Proc Natl Acad Sci U S A 111 (29), E2957-66.

41. Ronquist, F. et al. (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. Syst Biol 61 (6), 973-99.

42. Zhang, C. et al. (2016) Total-Evidence Dating under the Fossilized Birth-Death Process. Syst Biol 65 (2), 228-49.

43. Gavryushkina, A. et al. (2017) Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins. Syst Biol 66 (1), 57-73.

44. Alvarez-Carretero, S. et al. (2019) Bayesian Estimation of Species Divergence Times Using Correlated Quantitative Characters. Syst Biol 68 (6), 967-986.

45. Silvestro, D. et al. (2018) Closing the gap between palaeontological and neontological speciation and extinction rate estimates. Nat Commun 9 (1), 5237.

46. Rambaut, A. (2000) Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16 (4), 395-399.

47. Stadler, T. and Yang, Z.H. (2013) Dating phylogenies with sequentially sampled tips. Systematic Biology 62 (5), 674-688.

48. Worobey, M. et al. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455 (7213), 661-4.

49. Chang, D. et al. (2017) The evolutionary and phylogeographic history of woolly mammoths: a comprehensive mitogenomic analysis. Sci Rep 7, 44585.

609 50. Llamas, B. and al., e. (2016) Ancient mitochondrial DNA provides high-resolution time scale
610 of the peopling of the Americas. Sci. Adv.
611
612 51. Orlando, L. et al. (2013) Recalibrating Equus evolution using the genome sequence of an
613 early Middle Pleistocene horse. Nature 499 (7456), 74-8.
614
615 52. Kingman, J.F.C. (1982) The coalescent. Stochastic Processes and their Applications 13,
616 235-248.
617
618 53. Xu, B. and Yang, Z. (2016) Challenges in Species Tree Estimation Under the Multispecies
619 Coalescent Model. Genetics 204 (4), 1353-1368.
620
621 54. Rannala, B. et al. (2020) The multispecies coalescent model and species tree inference. In
622 Phylogenetics in the Genomic Era (Scoenavacca, C. et al. eds), pp. 1-20.
623
624 55. Maddison, W.P. (1997) Gene Trees in Species Trees. Syst Biol 46 (3), 523-536.
625
626 56. Rokas, A. et al. (2003) Genome-scale approaches to resolving incongruence in molecular
627 phylogenies. Nature 425, 798-804.
628
629 57. Yang, Z. (2014) Molecular Evolution: A Statisitical Approach, Oxford University Press.
630
631 58. Flouri, T. et al. (2020) A Bayesian Implementation of the Multispecies Coalescent Model
632 with Introgression for Phylogenomic Analysis. Mol Biol Evol 37 (4), 1211-1223.
633
634 59. Edwards, S.V. and Beerli, P. (2000) Perspective: gene divergence, population divergence,
635 and the variance in coalescence time in phylogeographic studies. Evolution 54 (6), 1839-1854.
636
637 60. Carstens, B.C. and Knowles, L.L. (2007) Estimating species phylogeny from gene-tree
638 probabilities despite incomplete lineage sorting: An example from Melanoplus grasshoppers.
639 Systematic Biology 56 (3), 400-411.
640
641 61. Mendes, F.K. and Hahn, M.W. (2016) Gene Tree Discordance Causes Apparent
642 Substitution Rate Variation. Syst Biol 65 (4), 711-21.
643
644 62. Fang, B. et al. (2020) Estimating uncertainty in divergence times among three-spined
645 stickleback clades using the multispecies coalescent. Mol Phylogenet Evol 142, 106646.
646
647 63. Schüßler, D. et al. (2019) Cryptic patterns of speciation in cryptic primates: microendemic
648 mouse lemurs and the multispecies coalescent. bioRxiv.
649
650 64. Liu, W. et al. (2019) An evolutionarily stable strategy to colonize spatially extended habitats.
651 Nature 575 (7784), 664-668.
652
653 65. Oliveros, C.H. et al. (2019) Earth history and the passerine superradiation. Proc Natl Acad
654 Sci U S A 116 (16), 7916-7925.
655
656 66. Li, H.T. et al. (2019) Origin of angiosperms and the puzzle of the Jurassic gap. Nat Plants 5
657 (5), 461-470.
658

67. Barba-Montoya, J. et al. (2018) Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. New Phytol 218 (2), 819-834.

68. Beaulieu, J.M. et al. (2015) Heterogeneous Rates of Molecular Evolution and Diversification Could Explain the Triassic Age Estimate for Angiosperms. Syst Biol 64 (5), 869-78.

69. Magallon, S. et al. (2015) A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 207 (2), 437-453.

70. Leebens-Mack and One Thousand Plant Transcriptomes, I. (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574 (7780), 679-685.

71. Flouri, T. et al. (2018) Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Mol Biol Evol 35 (10), 2585-2593.

72. Yoder, A.D. et al. (2016) Geogenetic patterns in mouse lemurs (genus Microcebus) reveal the ghosts of Madagascar's forests past. Proc Natl Acad Sci U S A 113 (29), 8049-56.

73. Federman, S. et al. (2018) Reconciling species diversity in a tropical plant clade (Canarium, Burseraceae). PLoS One 13 (6), e0198882.

74. Keightley, P.D. et al. (2015) Estimation of the spontaneous mutation rate in Heliconius melpomene. Mol Biol Evol 32 (1), 239-43.

75. Smeds, L. et al. (2016) Direct estimate of the rate of germline mutation in a bird. Genome Res 26 (9), 1211-8.

76. Pfeifer, S.P. (2017) Direct estimate of the spontaneous germ line mutation rate in African green monkeys. Evolution 71 (12), 2858-2870.

77. Thomas, G.W.C. et al. (2018) Reproductive longevity predicts mutation rates in primates. Curr Biol 28 (19), 3193-3197 e5.

78. Koch, E. et al. (2019) De novo mutation rate estimation in wolves of known pedigree. Mol Biol Evol.

79. Wang, L. et al. (2019) The architecture of intra-organism mutation rate variation in plants. PLoS Biol 17 (4), e3000191.

80. Pierron, D. et al. (2017) Genomic landscape of human diversity across Madagascar. Proc Natl Acad Sci U S A 114 (32), E6498-E6506.

81. Lawal, R.A. et al. (2020) The wild species genome ancestry of domestic chickens. BMC Biol 18 (1), 13.

82. Smith, O. et al. (2019) A domestication history of dynamic adaptation and genomic deterioration in Sorghum. Nat Plants 5 (4), 369-379.

83. Wu, D.D. et al. (2018) Pervasive introgression facilitated domestication and adaptation in the Bos species complex. Nat Ecol Evol 2 (7), 1139-1145.

84. Smith, T.C.A. et al. (2018) Large scale variation in the rate of germ-line de novo mutation, base composition, divergence and diversity in humans. PLoS Genet 14 (3), e1007254.

85. Lewin, H.A. et al. (2018) Earth BioGenome Project: Sequencing life for the future of life. Proc Natl Acad Sci U S A 115 (17), 4325-4333.

86. dos Reis, M. et al. (2018) Using Phylogenomic Data to Explore the Effects of Relaxed Clocks and Calibration Strategies on Divergence Time Estimation: Primates as a Test Case. Syst Biol 67 (4), 594-615.

87. Moorjani, P. et al. (2016) Human Germline Mutation and the Erratic Evolutionary Clock. PLoS Biol 14 (10), e2000744.

88. Scally, A. et al. (2012) Insights into hominid evolution from the gorilla genome sequence. Nature 483 (7388), 169-75.

89. Ho, S.Y. et al. (2011) Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Syst Biol 60 (3), 366-75.

90. Exposito-Alonso, M. et al. (2018) The rate and potential relevance of new mutations in a colonizing plant lineage. PLoS Genet 14 (2), e1007155.

91. Yang, Z.H. and Yoder, A.D. (2003) Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Systematic Biology 52 (5), 705-716.

92. Campbell, C.R. et al. (2019) Pedigree-based measurement of the de novo mutation rate in the gray mouse lemur reveals a high mutation rate, few mutations in CpG sites, and a weak sex bias. bioRxiv.

93. Zanne, A.E. et al. (2014) Three keys to the radiation of angiosperms into freezing environments. Nature 506 (7486), 89-92.

94. Jetz, W. et al. (2012) The global diversity of birds in space and time. Nature 491 (7424), 444-8.

95. Hudson, R.R. (1983) Testing the constant-rate neutral allele model with protein sequence data. Evolution 37 (1), 203-217.

96. Bouckaert, R. et al. (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Computational Biology 15 (4), e1006650.

97. Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol, 24 (8), 1586–1591.

98. Ronquist F. et al. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61 (3), 539-542.

99. Lartillot, N. and Philippe, H. A. (2004) A Bayesian mixture model for across-site heterogeneities in the amino acid replacement process. Mol Biol Evol, 21 (6), 1095-1109.

Highlights

- Molecular clock models using fossil calibrations have allowed investigators to estimate the age of speciation events.
- Theoretical and computational developments have relaxed the assumption of a molecular clock, thus improving the accuracy of divergence time estimation.
- Despite these advances, estimates can be biased when there is widespread incomplete lineage sorting (ILS).
- Increased understanding of gene tree heterogeneity has driven multispecies coalescent (MSC) methods to prominence, though the potential power of the MSC for divergence time estimation remains largely unexplored.
- Absolute times can be obtained by using mutation rates estimated from pedigrees, providing [some] freedom from the incomplete fossil record.
- Mutation-rate calibrated MSC methods and traditional phylogenetic clock-dating methods with fossil calibrations can yield strikingly different divergence times.

Outstanding Questions

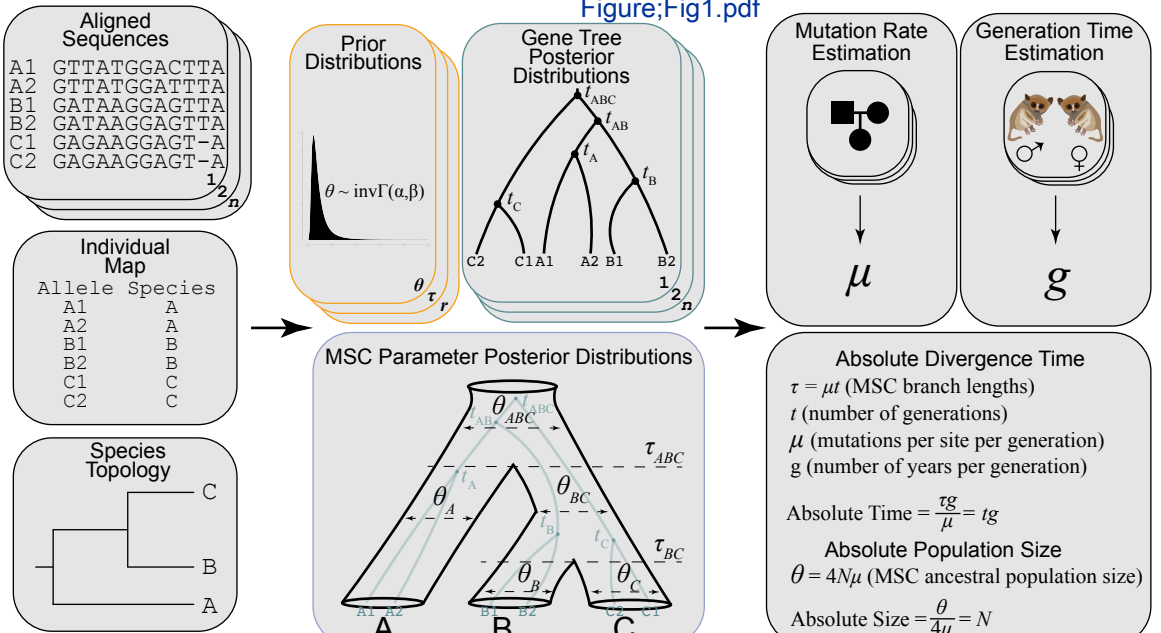
- To what extent is among-lineage rate variation modeled by relaxed clock methods due to gene tree discordance from ILS?
- Have divergence times throughout the tree of life been systematically overestimated in clades that rely on external, and typically older, calibrations?
- Do divergence time estimates based on per-generation mutation rates and per-year substitution rates yield similar results, especially if substitution rates are estimated from presumably neutral regions of the genome such as third codon positions?
- Will MSC estimates that leverage fossil calibrations bring new insights to contentious age estimates such as the origins of placental mammals or angiosperms?
- Should effective population size variation among species be a concern for divergence time estimation studies using concatenation?
- Can mutation-rate calibrated MSC methods that account for variable rates and generation times among branches improve divergence time estimation for clades that have rapid life history transitions?
- How can we develop standard operating procedures for evaluating the strength of evidence for divergence time estimates from traditional phylogenetic analyses versus ages inferred from MSC methods that rely on mutation rate and generation time estimates?
- Are there alternative ways forward for estimating the absolute age of clades with poor or non-existent fossil representation?
- To what extent do the methods (MSC versus concatenation) and the calibrations (mutation rate versus fossils) impact divergence time estimates?

Key Input Data

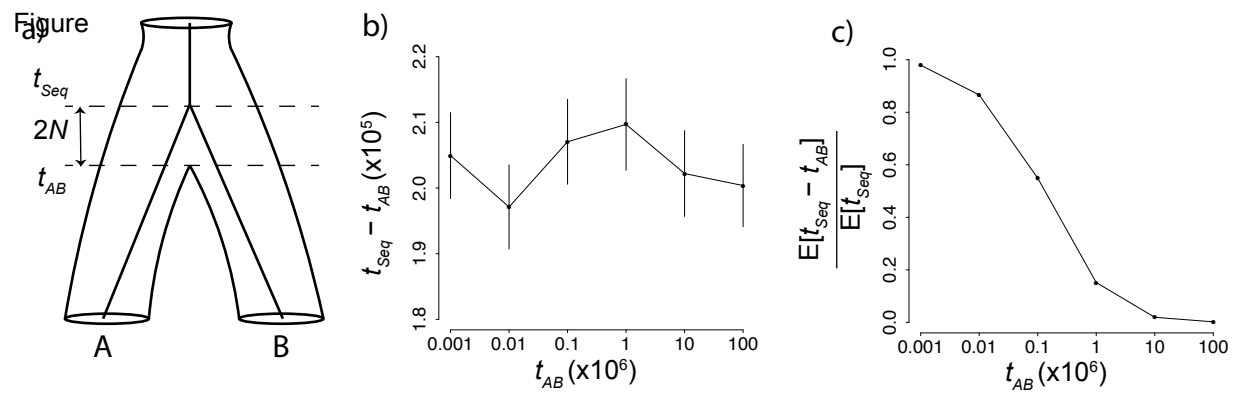
Estimate MSC Parameters by MCMC

Calibrate with Mutation Rate

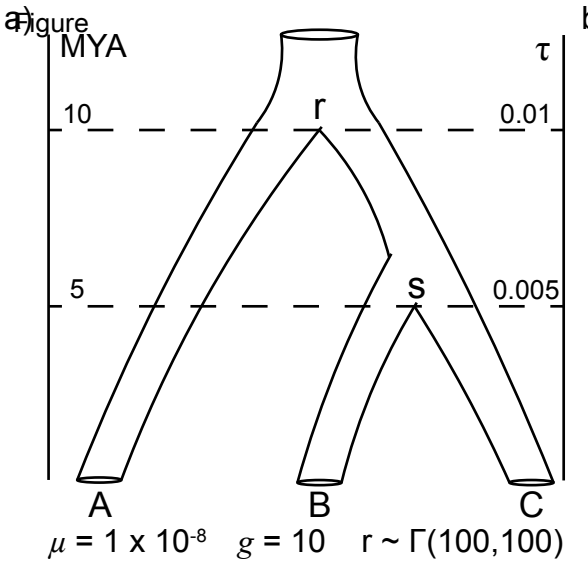
Click here to access/download Key Figure;Fig1.pdf



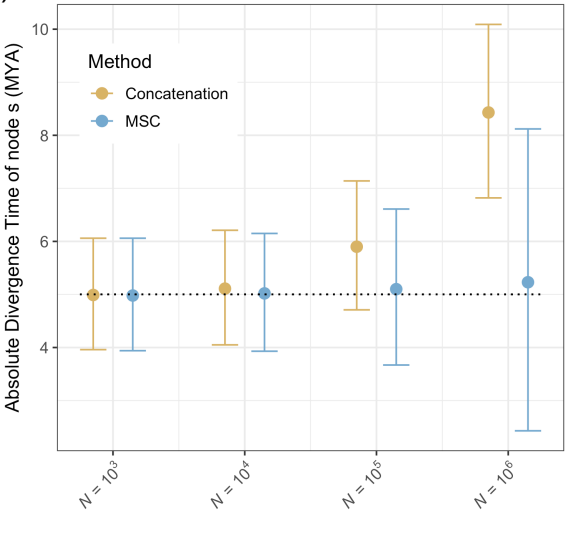
Figure


a)
t_{Seq}
2N
t_{AB}
A
B
b)
$t_{Seq} - t_{AB}$ (x10^5)
2.2
2.1
2.0
1.9
1.8
0.001
0.01
0.1
1
10
100
t_{AB} (x10^6)
c)
$\frac{E[t_{Seq} - t_{AB}]}{E[t_{Seq}]}$
1.0
0.8
0.6
0.4
0.2
0.0
0.001
0.01
0.1
1
10
100
t_{AB} (x10^6)

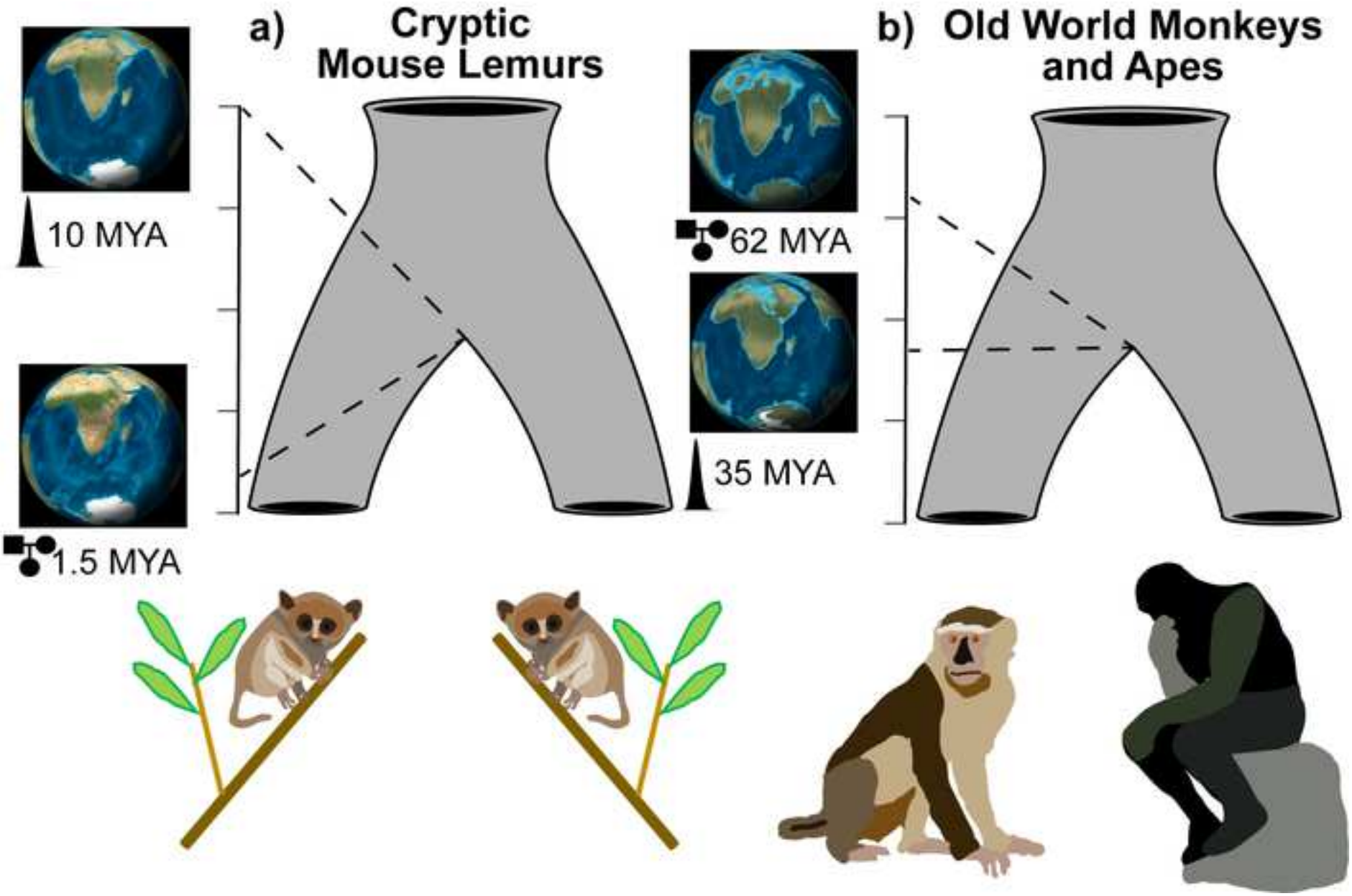
a) Figure


MYA
τ
10
r
0.01
5
s
0.005
A
B
C
μ = 1 x 10-8
g = 10
r ~ Γ(100,100)

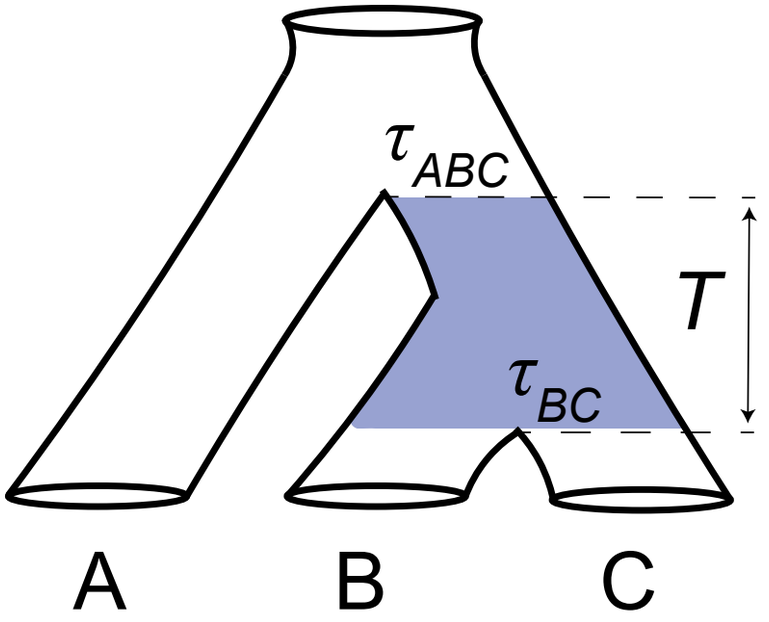
b)

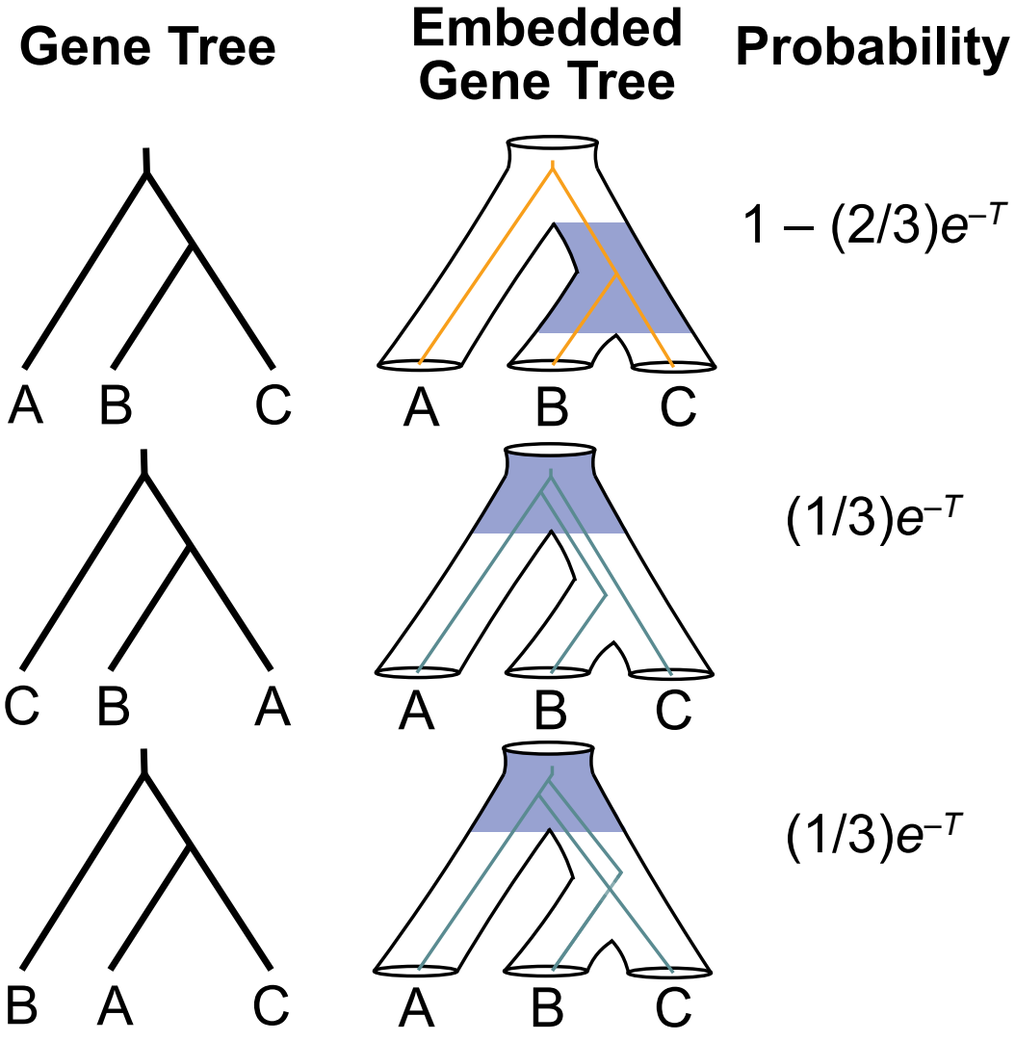

Absolute Divergence Time of node s (MYA)
10
8
6
4
Method
Concatenation
MSC
N = 10^3
N = 10^4
N = 10^5
N = 10^6

Figure

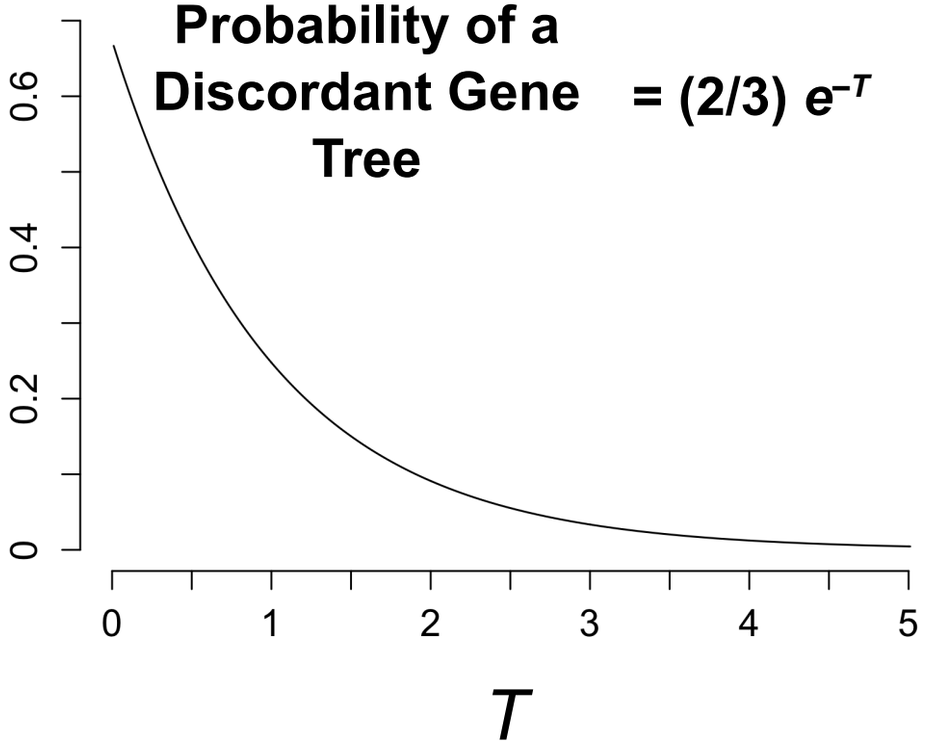


Figure

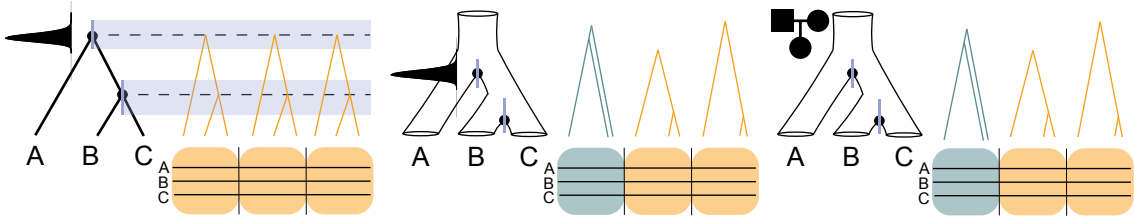




ILS occurs when B and C do not coalesce within T, the time between τ_{BC} and τ_{ABC} measured in $2N_{BC}$ generations. ILS can be identified visually by embedding gene trees within the species tree. The neutral coalescent provides expectations for the frequency with which ILS occurs that are dependent on T alone [95]. When T is 0, 2/3 of gene trees are expected not to match the species tree due to ILS. Less than 1% of gene trees are expected to be discordant around T of 5. The absolute divergence time does not affect T.



Figure



Strengths

Concatenation with Fossil Calibrations	MSC with Fossil Calibrations	Mutation Rate Calibrated MSC
Computationally efficient for large numbers of tips and loci	Considers discordance between gene trees and species trees	Does not require calibrations on nodes from external information such as fossils

Weaknesses

Concatenation with Fossil Calibrations	MSC with Fossil Calibrations	Mutation Rate Calibrated MSC
May produce biased estimates when ILS is high or when gene sequence divergence is far from species divergence	Increased computational complexity from averaging over gene trees to estimate species tree parameters	Requires external mutation rate estimates from sequenced pedigrees and potentially not appropriate for distant taxa

Common Programs

Concatenation with Fossil Calibrations	MSC with Fossil Calibrations	Mutation Rate Calibrated MSC
BEAST2 [96] MCMCTREE [97] MrBayes [98] PhyloBayes [99]	BPP [5,71] StarBEAST2 [6]	BPP [5,71] StarBEAST2 [6]