Investigating the reliability of molecular estimates of evolutionary time when substitution rates and speciation rates vary

ANDREW M. RITCHIE[1 *], XIA HUA[1 2], LINDELL BROMHAM[1]

1. *Research School of Biological Sciences, Australian National University, Canberra, ACT 2600, Australia*
2. *Mathematical Sciences Institute, Australian National University, Canberra, ACT 2600, Australia*

* Corresponding Author:

Andrew M. Ritchie

Research School of Biological Sciences, Robertson Building, 134 Linnaeus Way

Australian National University, Canberra, ACT 2600, Australia

Ph: +61 2 6125 5111

Email: Andrew.Ritchie@anu.edu.au

Acknowledgements

We would like to extend our gratitude to Marcel Cardillo and Alex Skeels for helpful conversations in the preparation of this manuscript.

Funding

This work was supported by the Australian Research Council grant DP190101916.

Abstract

Background

An accurate timescale of evolutionary history is essential to testing hypotheses about the influence of historical events and processes, and the timescale for evolution is increasingly derived from analysis of DNA sequences. But variation in the rate of molecular evolution complicates the inference of time from DNA. Evidence is growing for numerous factors, such as life history and habitat, that are linked both to the molecular processes of mutation and fixation and to rates of macroevolutionary diversification. However, the most widely used models of molecular rate variation, such as the uncorrelated and autocorrelated lognormal clocks, rely on idealised models of rate variation and molecular dating methods are rarely tested against complex models of rate change. One relationship that is not accounted for in molecular dating is the potential for interaction between molecular substitution rates and speciation, a relationship that has been supported by empirical studies in a growing number of taxa. If these relationships are as widespread as evidence indicates, they may have a significant influence on molecular dates.

Results

We simulate phylogenies and molecular sequences under three different realistic rate variation models – one in which speciation rates and substitution rates both vary but are unlinked, one in which they covary continuously and one punctuated model in which molecular change is concentrated in speciation events, using empirical case studies to parameterise realistic simulations. We test two commonly used “relaxed clock” molecular dating methods against these realistic simulations to explore the degree of error in molecular dates under each model. We find average divergence time inference errors ranging from 12%

of node age for the unlinked model when reconstructed under an uncorrelated rate prior, to up to 93% when punctuated simulations are reconstructed under an autocorrelated prior.

Conclusions

We demonstrate the potential for substantial errors in molecular dates when both speciation rates and substitution rates vary between lineages. This study highlights the need for tests of molecular dating methods against realistic models of rate variation generated from empirical parameters and known relationships.

1.Background

Understanding the timescale of evolution is critical to researching the processes that generate and shape the diversity of life on Earth. Analysis of DNA sequences offers the possibility of investigating the tempo and mode of evolution of all extant lineages, not just those that have a detailed fossil record. The earliest molecular dating analyses concluded that the average rate of molecular evolution was relatively constant over time in some proteins because of an apparently linear accumulation of amino acid differences when compared between species (1, 2). However, it has since become apparent that the average rate of molecular evolution can be influenced by a wide range of species traits, such as generation time, body size and longevity (3, 4), by macroevolutionary processes such as net diversification rate (5-10) and potentially also by environmental features like temperature (9-11). If the average rate of molecular evolution varies substantially and consistently between lineages, then we ideally need to incorporate rate variation in models of molecular evolution in order to accurately infer divergence times among species (12).

A wide variety of models have been developed to account for different patterns of among-lineage molecular rate variation (13-16). In the first instance, these models differ in how many different evolutionary rate parameters are estimated for different parts of the phylogeny. For example, a 'strict clock' model estimates only one rate parameter that applies to every branch in the tree, while a 'local clock' model estimates additional rate parameters for subsets of branches (17-19), and a 'relaxed clock' provides for a separate rate on every branch (20). Models also differ in how the rates on different branches are related to one another. Some models, notably in the penalized-likelihood framework, aim to minimize the overall variation of branch-specific rates or the differences between adjacent branches (21, 22). For Bayesian phylogenetic methods (23, 24), branch-specific rates are drawn from a prior probability distribution that determines which rates and patterns of relative rate variation are considered *a priori* more probable. We refer to this prior as the rate prior. Common variants include uncorrelated models, in which each branch-specific rate is an independent draw from a distribution with a common mean (20); and autocorrelated models, in which the mean of the distribution for each branch-specific rate depends on the rate drawn for its parent branch (25, 26).

Testing over the last two decades has overall supported the robustness of divergence time estimates under the Bayesian methodology to the choice of rate prior, provided that enough separate rate parameters are allowed and sufficient calibrating information is provided (27-33); but see (34). However, these studies rely heavily on simulated data or on specific and usually well-calibrated empirical examples, whereas simulations have indicated that different rate priors may give different results when calibrating information is poorer (30). Changing the rate prior in these circumstances could lead to differences in empirical dates and associated causal hypotheses. Examples of empirical studies where the choice of

rate prior appears to be a factor in differing age estimates include influenza viruses (35), the age of grasstrees (36) and the origins of Metazoa (37).

Currently used priors on rate variation are generally either uncorrelated, so that each lineage-specific rate is drawn independently from a common distribution, or autocorrelated so that lineage-specific rates depend evolve from the rates of their parent lineages according to a stochastic, random walk. However, the factors shaping variation in rate of molecular evolution are complex: in addition to random variation in rates between lineages, there is clearly also a systematic component influenced by many factors including species traits, population dynamics, environment and evolutionary history (38). Evolutionary rates among eukaryotes appear to be faster at lower elevations and latitudes (9, 39, 40) but slower in arid climates (41). Life history covaries with molecular rates across multiple axes, including body size (42-45), fecundity (46) generation time (47).and longevity (48, 49). Associations have also been observed between substitution rates and more complex species traits, such as parasitism (50), flightlessness (51) and sexual competition (52). Furthermore, substitution rates are the result of underlying microevolutionary processes, and may therefore be affected by population size (53-56).

In addition to covarying with other evolving traits, there is evidence that variation in molecular rates can be associated with macroevolutionary patterns. Relationships between the relative size of sister clades, representing their net diversification rates, and their molecular rates have been detected in several studies, include in reptiles (8, 57), birds (7, 58), flowering plants (6, 59, 60) and across Metazoa (61). Several studies have also noted relationships between the genetic divergence of species in a phylogeny and the number of intervening speciation events, which may also represent examples of this phenomenon (62-64).

The mechanism behind observed correlations between diversification and molecular evolution remains a matter of debate. One line of argument proposes that higher molecular rates could promote speciation by speeding the development of reproductive isolation (65). In this hypothesis, subpopulations become sporadically isolated over time and may accumulate alleles that may be effectively neutral on their own but deleterious in combination (66, 67). In this case, greater rates of molecular evolution could lead to faster development of reproductive isolation during periods of isolation, leading to a greater average rate of speciation (68, 69). Alternatively, higher molecular rates could reduce extinction rates by accumulating more standing genetic diversity which provides a buffer against changing environment by allowing more rapid selective response (70). Speciation rates and molecular rates could also be related through a common driving factor, such as climate or environmental productivity (5, 10). In the other causal direction, speciation itself could increase average molecular rates by producing rapid bursts of evolution (62, 63, 71). Suggested mechanisms for an increase in substitution rates during speciation involve an increase in fixation rates of mildly deleterious alleles and linked neutral mutations due to drift, or large-scale genomic divergence caused by an increase in the frequency of disruptive genomic events such as polyploidy or chromosomal rearrangements (71).

An association between substitution rates and diversification rates could affect the performance of existing divergence time estimation methods (e.g. 72). In particular, Bayesian methods divide their specification of prior beliefs about the evolutionary process between the rate prior, the prior on divergence times and tree topology (sometimes referred to as the tree prior or node time prior), and calibrating distributions on the ages of individual speciation events (23, 73). These priors are essential to estimating divergence time, because genetic

distances between sequences are the product of rate and time (the total number of character differences between related sequences). This means that even if genetic distances are correctly inferred on a phylogeny, rate and time are unidentifiable without additional constraints on joint values they can hold (74, 75). The rate prior, node time prior and calibrating distributions place either soft or hard constraints on these values, allowing researchers to infer evolutionary history on an absolute timescale. However, this means that inferred divergence times are always subject to the choice of rate, time and calibration priors even with large amounts of sequence data (76, 77). Most commonly used Bayesian molecular dating methods do not propose *a priori* relationships between the rate prior and the tree prior, so the joint distribution implied by these priors will assign relatively higher probability to long-lived, fast-evolving lineages and short-lived, slow-evolving lineages than would be expected when rates and times are related. It is therefore possible that an unmodelled relationship between rates and times could produce distributions of molecular rates and speciation times that differ substantially from those implied by common Bayesian priors, leading to widespread error in estimates of evolutionary dates.

Here, we investigate the potential influence of speciation-related rates of molecular evolution on the reliability of molecular dating studies. Rather than assuming only one specific model of the association between diversification rates and rates of molecular, we model a range of possible causal links. We include three alternative models of the association between rates of molecular evolution and speciation rates. In the first, both speciation rates and substitution rates vary over the phylogeny, but neither influences the other. In the second, we model continuous variation in a linked manner, which might represent either direct influence of one on the other (65) or an indirect association between speciation rates and substitution rates, for example if they are both influenced by environmental differences (5,

10). The third model is a punctuated model in which bursts of substitution are associated with speciation events (62, 63, 71, 78). The proportion of all substitutions associated with such bursts has recently been shown to impact Bayesian inference of clade crown ages under an uncorrelated lognormal rate prior (72).

In this study, we simulate the evolution of molecular sequences under all three models of the link between speciation and rate of molecular evolution – unlinked, continuous coevolution and a punctuated model - allowing for independent variation in molecular and speciation rates, and observe the average effects on errors in Bayesian divergence time estimates. In order to make our test have real-world significance, we also base our simulations, as far as possible, on empirically determined relationships between rates of molecular evolution and diversification rate. We parameterize our models using values taken from empirical inferences on bird data (79, 80) or which reproduce established empirical relationships between rates of molecular evolution and diversification rate in birds (81). Since the choice of rate prior may affect the results of phylogeny reconstruction (34), we also examine the effect of divergence time reconstruction under uncorrelated and autocorrelated rate priors. We reconstruct divergence times on these simulated data sets and examine the effects of different forms of correlated molecular and speciation rates on the performance of molecular divergence time reconstruction.

2. Results

2.1 Simulation design

We refer throughout to simulated phylogenies, which are known without error, as the 'simulated trees', and to the phylogenies inferred from analyzing the sequences evolved along each of these simulated phylogenies as the 'reconstructed trees'. Comparison of the reconstructed trees to the simulated trees allows us to evaluate the degree to which phylogenetic inference methods correctly infer the known evolutionary history of the simulated data.

Our overall design is summarised in Figure 1. For our simulation study, we first generate phylogenetic trees in units of absolute time with evolving rates of speciation and molecular evolution. We then simulate DNA sequences along these phylogenies. Finally, we infer dated trees from the alignment of these sequences using standard methods in Bayesian phylogenetics and compare the inferred tree to the simulated tree from the known evolutionary history from each simulation.

We simulate sets of 50 phylogenetic trees and nucleotide sequence alignments under each of three different models describing the association between speciation rates and rates of molecular evolution: one "Unlinked" model in which molecular and diversification rates each vary independently, as well as two alternative mechanisms for an association between rates of molecular evolution and speciation rates ("Punctuated" and "Continuous": see Table 1). We do not consider models in which molecular rates affect rates of extinction, so in our simulations, extinction rates remain constant over the tree. For all models, we first simulate the phylogenetic tree in continuous time, together with speciation and substitution rates that vary along the tree. From this, we derive branch lengths in expected genetic divergence (substitutions/site). Finally, we simulate the evolution of nucleotide sequences along these phylogenies.

231

232 **Table 1. Simulation design.**
233

Simulation type	Speciation rates covary with substitution rates?	Instantaneous covariance	Punctuated burst of substitutions at nodes?
Unlinked	*No*	0	*no*
Continuous	*Yes*	0.0044	*no*
Punctuated	*No*	0	*yes*

234 We lay out the conditions for each set of simulated trees. Each set consisted of 50 trees.
235 Lineage-specific speciation rates and substitution rates evolve continuously through time via
236 Brownian motion in all three studies. The three simulations differ in whether the covariance
237 between speciation and substitution rates is greater than zero in the Brownian motion and
238 whether an additional burst of substitutions occurs at nodes of the tree (speciation events).
239

240 In the Unlinked model, we simulate continuous variation in both substitution rate and
241 speciation rate through time, but neither affects the other. Substitution and speciation rates
242 evolve independently by Brownian motion, where each lineage continuously adds a separate
243 random increment to its substitution rate and speciation rate in each of many small time
244 intervals. The Unlinked model serves as a baseline to determine the level of uncertainty in
245 molecular dates caused by the random variation in substitution and diversification rate when
246 there is no mechanistic link between the two. In the Continuous model, speciation rates and
247 substitution rates coevolve with each other continuously over time. The coevolution is
248 simulated as a correlated Brownian motion, where the random increments in speciation and
249 molecular rates added over time are dependent on one another. Consequently, substitution
250 rate and speciation rate tend to increase or decrease together. For this model, we also ensured
251 realism simulated many data sets with different covariance parameters, and selected a set of
252 50 trees that reproduced the empirical correlation of (7) between clade size and branch length
253 (Supplementary Methods S1.1). In the Punctuated model, the genetic divergence along a
254 branch is the sum of a gradual accumulation of substitutions at a rate that evolves as in the

255 Unlinked model, and a burst component at each speciation event. A speciation event occurs
256 each time a new lineage is born, whether any of its descendants are observed in the present or
257 not. The amount of additional branch length due to the burst is independently and randomly
258 drawn for each speciation event. For this model, we condition our simulations to reproduce
259 the relationship observed in (62). There it was observed that the total number of substitutions
260 occurring along root-to-tip paths in their sample of trees correlates with the number of nodes
261 along the path. Although this relationship could also be produced by a mechanism similar to
262 the Continuous model (82-84), we model the mechanism proposed in (62) that the node-
263 associated substitutions are produced in bursts associated with speciation events.
264
265 We derived plausible speciation, extinction and substitution rates for our simulations
266 using the avian phylogeny in (85). We do not intend to make any inferences about the actual
267 evolution of birds or the accuracy of inferred divergence times. Birds are chosen only
268 because the relationships among traits, molecular rates and diversification rates are well
269 studied, so that we can use these to give us an empirical ground for choosing parameter
270 values. Branch-specific diversification rates have been estimated for many of the lineages in
271 this tree in (80), while substitution rates have been inferred in (79) assuming covariation of
272 molecular rates and life history. Speciation rate and molecular rate values taken from avian
273 data have been previously used to inform realistic simulations on ecological assemblies for
274 the purpose of estimating the effect of dating error on phylogenetic diversity estimates (86),
275 and we make use of similar procedures here. A relationship between molecular rates and net
276 diversification rate has also been established for birds (7, 8). We use this fact to select our
277 Continuous data set by simulating multiple data sets and varying the covariance parameter
278 until the relationship appears. Full details of how we used these studies to determine realistic
279 parameterisations for our simulations are available as Supplementary Methods S1.1.

2.2 Validation of tree simulations

The simulation procedure produced trees with average characteristics similar to the avian clades used to parameterize the study (80, supplementary data, 85). Across all data sets, the mean root date of simulated trees was 28.4 million years, compared to a range of 23.8-35.6 million years for bird clades with 70-80 taxa. Simulated lineage-specific speciation rates at the tips were on average slightly lower than the speciation rate of similar clades in the bird tree (80). The range of simulated speciation rates on log scale was 0.009-0.92 compared to a range of 0.034-2.25 average log lineages/million years for bird clades with 70-80 taxa, similar taxon number to our simulated trees. The range of speciation rates within each tree was also similar to the range across the avian clades with similar taxon number, with an average magnitude of 0.44 log lineages/million years compared to 0.32-2.21 for avian clades with 70-80 taxa.

2.3 Error in reconstructed trees

We used three metrics (3.2), the gamma statistic (3.3) and the sister pair analysis (3.4) to assess the accuracy of the reconstruction of trees simulated from the Unlinked model, the Continuous model, and the Punctuated model, under uncorrelated lognormal (UCLN) and autocorrelated lognormal (ACLN) models of among-lineage rate variation (Table 2).

305 **Table 2. Summary of results.**
306

Rate prior	Simulation	Node age error (%)	Topo. error	HPD width (%)	Gamma Statistic	Sister pair analysis
UCLN	*Unlinked*	6.3–30.9	0–14	43–29	-1.46–1.09	Null
UCLN	*Continuous*	6.7–30.3	0–16	35–605	-1.52–0.20	Null
UCLN	*Punctuated*	8.8–80.4	0-8	42–154	-2.16–3.92	Negative
ACLN	*Unlinked*	11–417	Fixed	84.9–217.8	-7.13–5.62	Positive
ACLN	*Continuous*	9.4–170	Fixed	73.5–319.4	-4.92–3.32	Positive
ACLN	*Punctuated*	13.7–397.8	Fixed	76.7 - 144.1	-4.95–2.88	Positive

307 Summary of reconstructed tree characteristics. Trees simulated under three models:
308 Unlinked , Continuous, and Punctuated. Molecular dates were then reconstructed under two
309 methods, the uncorrelated lognormal model (UCLN) and the autocorrelated lognormal model
310 (ACLN). We calculated three metrics for each tree and summarised these over each
311 simulation type. The metrics were median absolute % error of node ages, topological error
312 (Topo. Error) by number of true bipartitions not reconstructed, and uncertainty by mean
313 highest posterior density interval width as a proportion of reconstructed node age. We
314 compared the gamma statistics of the reconstructed trees to the simulated trees in order to
315 detect any systematic bias in node age error. We also performed sister pair analysis to test if
316 we can detect the positive correlation between speciation rate and substitution rate from the
317 reconstructed trees.
318

319 First, we examined the median absolute percentage error (MAPE) in the ages of all
320 nodes within each tree whose topology were correctly reconstructed (Figure 2). For trees
321 reconstructed under UCLN, the average (min–max) MAPE was 12.2% for Unlinked
322 simulations (6.3–30.9%). Errors for the Continuous simulation were similar at 14.2% (6.7–
323 30.3%). Average MAPE for the Punctuated simulation was higher than for Unlinked trees at
324 20.3% (8.8–80.4%). For trees reconstructed under the ACLN in PAML, MAPE values were
325 higher and range of errors were larger than for the UCLN in BEAST. Average MAPE for
326 Unlinked trees was 93.3% (11–417%), compared to 43.4% (9.4%–170%) for Continuous
327 trees and 84% (13.7–397.8%) for Punctuated trees.
328

329 Second, for trees reconstructed under the UCLN model in BEAST, we also
330 investigated the degree of topological error via the Robinson-Foulds distance between the
331 true and estimated tree (87). The median background topological error for the Unlinked
332 simulations was 6 bipartitions (0–14), compared to 5 (0–16) for the Continuous model
333 (Figure 3). Error was lower for the Punctuated model with a median distance of 2
334 bipartitions (0-8).
335
336 Third, we calculated the uncertainty of node age reconstruction as reported by the
337 method via highest posterior density intervals. Under UCLN (Figure 5), the average HPD
338 interval for the Unlinked trees had a width 60% of the reconstructed age of its associated
339 node (43 – 29%). For the Continuous trees this was 108% (35 – 605%), and for the
340 Punctuated trees it was 70% of node age (42 – 154%). Under ACLN, the average HPD width
341 for Unlinked trees was 120.7% of node age (84.9 – 217.8%); for Continuous trees, 112.3%
342 (73.5% – 319.4%); and for Punctuated trees, 107.0% (76.7 – 144.1%).
343
344 *2.4 Distribution of reconstructed vs simulated node ages*
345
346 If the difference in gamma values between reconstructed and true trees is negative,
347 the reconstructed tree has shorter internal branches and longer terminal branches than the true
348 tree, while if it is positive the reconstructed tree has longer internal branches and shorter
349 terminal branches than the true tree. Under the UCLN, Gamma for the reconstructed
350 Unlinked trees was distributed evenly around the gamma of the simulated tree with a mean
351 value of +0.11 (-1.46 – + 1.09; Figure 4). This distribution was skewed negative (shorter tips)
352 relative to the simulated tree for Continuous trees with a mean of -0.44 (-1.52 – +0.20). This
353 negative skew was more pronounced for the Punctuated trees at -0.94 (-2.16 – +3.92). Under

ACLN, the relative change in gamma was universally negative. For the Unlinked trees it was -2.87 (-7.13 – +5.62), while for the Continuous trees it was less negative at -1.33 (-4.92 – +3.32), and the Punctuated trees were again skewed negative with a mean of -2.74 (-4.95 – +2.88).

2.5 Efficacy of sister pair correlation on reconstructed branch lengths

In order to determine whether the simulated relationship between molecular rates and diversification rates could be recovered after tree reconstruction, we regressed log clade size contrasts against contrasts in log median branch length estimates in the data set of sister pair analysis generated from reconstructed trees under UCLN or ACLN for each simulation models (see 2.8). We present the p-value and sign of slope for each regression (Figure 6). Unlinked simulated trees show no significant relationship. Significant positive relationships are found for the Continuous and Punctuated simulations; note that this was a condition for selection of the Continuous data set (Supplementary Methods). When reconstructed under UCLN, the Unlinked data set and the Continuous data set show no significant relationship, while the Punctuated data set giving a significant negative relationship, opposite to what we expect. When reconstructed under ACLN, all three of the Unlinked, Continuous and Punctuated data sets have significant positive relationships between clade sizes and median branch length estimates.

3.Discussion

3.1 Errors associated with variation in speciation and substitution rate

379

380 The addition of rate-variable 'relaxed clock' models for use as rate priors represents a
381 significant extension to the sophistication of modern molecular dating methods (15). While
382 there is a wide variety of rate-variable molecular dating methods, all of them to date rely on
383 stochastic models of rate variation, such that rates of molecular evolution are expected to
384 evolve randomly over the phylogeny. However, it is clear that real substitution rates do not
385 always vary randomly. Instead, rates of molecular evolution are influenced by species life
386 history, niche and environment and macroevolutionary processes. It is important to examine
387 whether the stochastic models of rate variation can adequately model realistic biological
388 patterns of rate variation (3, 62, 65, 88). In the case of substitution rates, this means that we
389 must test molecular dating methods using rate simulations that are empirically based and
390 incorporate the effects of other biological processes. Here we have focused on testing the
391 impact of widely reported correlations between molecular rates and diversification on the
392 accuracy of molecular date reconstruction (5-8).

393

394 We used three models, Unlinked, Continuous and Punctuated, to reflect different
395 views on the possible links between substitution rates and diversification rates – either
396 varying independently; covarying continuously, linked directly by common mechanisms or
397 indirect through shared influences on rate variation; or linked by punctuated bursts of
398 increased substitutions related to speciation events. Rather than choosing parameters
399 arbitrarily, we parameterised our models using realistic values from the literature, or found
400 plausible values by tuning the models to produce data sets that reproduce empirical findings
401 (Supplementary Methods; 7, 79, 80). We used three different metrics (Table 2) to compare
402 the accuracy of phylogenetic inference using two population 'relaxed clock' molecular dating

methods, BEAST (using an uncorrelated rates model) and PAML (using an autocorrelated rates model).

3.2 Unlinked model

The Unlinked model describes the error that can be expected from independent variation in speciation and molecular rates. Inference on trees simulated under this model produced median absolute errors in node age inference of about 12% on average under the UCLN and a very high value of 93% under the ACLN (Figure 2). The largest errors were up to 30% for UCLN and the worst instances for ACLN had errors of over 400%. The Unlinked model produced lower errors under UCLN than the other two models, with an average error of 12% (Figure 2). Gamma differences for the Unlinked model are distributed evenly around zero under UCLN, so there is no overall tendency for nodes to be over- or under-estimated (Figure 4). The sources of error in the Unlinked model are most likely to be violation of the assumptions of rate priors and node time (tree) priors. Variation in speciation rates violates the assumptions of the birth-death node time priors used in both BEAST 2 and PAML, which assume universally constant rates of speciation and extinction (89). Speciation rate variation among lineages also induces more imbalanced trees on average (90), which can lead to bias and loss of precision in node age inferences (91). While the degree and scale on which speciation and extinction rates vary remains contested (92-95), there is increasing evidence that radiations and other forms of variation are a factor in the evolution of many taxa, so that errors of this magnitude may be widespread (96-99). Substitution rates that evolve in an autocorrelated fashion also violate the assumptions of the UCLN, which assumes that the evolutionary rate of each lineage is independent (20), and this may also induce errors in inference (30, 34).

The causes of the extremely high errors of the ACLN-based phylogenetic inference in PAML for the Unlinked trees (and throughout the study for the other models) are unclear. Examination of the output trees suggests that the most extreme errors are in trees with shallower calibrations, and that the ages of large uncalibrated clades are sometimes overestimated severely. The inferences in PAML and BEAST are not directly comparable due to numerous technical differences, such as the tree topology being fixed in PAML and BEAST placing prior distributions on speciation and extinction parameters, which are fixed in PAML. Nevertheless it is surprising that the performance of the ACLN in PAML was worse than UCLN in BEAST even though the underlying rate variation in these simulations is indeed autocorrelated. There are some indications that ACLN can be difficult to fit to data even when it is the correct model (91, 100). Another possibility is that the sensitivity of the prior to model violation in PAML was enhanced by our use of the approximate likelihood function (101). This is a second-order Taylor approximation to the likelihood evaluated at the maximum likelihood estimates of the branch lengths and other parameters. As discussed by dos Reis and Yang (101), while the approximation is very accurate under normal circumstances when the branch lengths proposed during Bayesian inference are very far from the maximum likelihood values the approximation becomes much less accurate. This can happen when the rate or node time (tree) priors are misspecified. Nevertheless, absolute error rates in other simulation studies have rarely been greater than 20% even with a deliberately misspecified rate prior and more limited calibrating information (30), so the extreme error rates we encountered could be down to analytical factors and not solely due to the rate prior itself.

3.3 Continuous model

The Continuous model has substitution rates and speciation rates that coevolve continuously rather than varying independently. Notably, we chose the degree of covariance that reproduced two well-established empirical phenomena – the correlation of clade size and branch length contrasts detected in a sister pairs analysis by (7), and the widespread relationship between branch lengths and tree nodes discovered by (62). This model produced absolute node time inference errors of 14% under UCLN, slightly higher than the Unlinked model, but the worst errors were up to 30% (Figure 2). Negatively-skewed gamma differences indicate that error generally leads to node ages being overestimated, pushing reconstructed nodes back towards the root (Figure 4). In addition to the error, which is the difference between reconstructed time estimates and the truth, we also examine how the reconstruction methods report their uncertainty about the truth through 95% highest posterior density intervals. Although not greatly different in terms of absolute error, the reported uncertainty of node time estimates was very high for this model under UCLN, with average HPD interval widths more than 107% of node height (Figure 5). This average is driven partly by some outlying trees with very high uncertainty, up to 600% of node height; the median is 75% of node height. Wide HPD intervals can be a sign of conflict between the prior and data, so it may be that covariance between diversification rates and substitution rates tends to produce patterns of branch lengths that are not well described by independent rate and time priors. The practical impact of this level of uncertainty could depend on how it is distributed within the tree and the nature of any downstream inferences; if the question at hand is the dating of major clades then high uncertainty in the age of tips may be tolerable, while the same level of uncertainty in internal branches would render the tree unusable. However, any macroevolutionary inference that relies on the distribution of nodes throughout the tree would likely be affected (59). Under ACLN, error rates were very high (mean of 43.5%; Figure 2)

but were lower for the Continuous simulations than they were for the Unlinked simulations. A reason for this could be that when substitution rate is positively correlated with speciation rate it is also negatively correlated with branch lengths in units of time. Then short terminal branches will tend to have faster rates and build up more substitutions. This in turn could enable lower variance in estimates of rate variation across the tips of the tree, which would aid in inferring the degree of autocorrelation.

3.4 Punctuated model

Levels of absolute error were highest under UCLN when speciation events were associated with large punctuational bursts of molecular evolution (mean of 20.3%, up to 80.4%; Figure 2). This implies that, if real evolutionary processes were to commonly involve such bursts of substitutions associated with speciation, or produce a similar distribution of evolutionary histories and rate variation through some other mechanism, median node age estimates reconstructed under the UCLN rate prior would be generally incorrect. Error rates were also extreme for ACLN (mean 80.4%, max 397.7%; Figure 2). The finding of high error rates in the Punctuated model corroborates a recent analysis (72), which showed that greater proportions of branch length associated with speciation events lead to greater crown age errors. Here, we show that the effect is strong enough to impact the median absolute error of all node ages, even when the crown node is calibrated. The very negative skew of Gamma away from the simulated value shows that these internal node ages are severely overestimated, leading to very distorted trees with long tips (Figure 4).

Despite having the highest degree of branch length error under UCLN, the Punctuated model had the lowest degree of topological error (a median of 2 bipartitions not correctly

503 reconstructed; Figure 3), comparing to the Unlinked and Continuous models (median of 5
504 and 6 respectively, Figure 3). Intuitively, this could be because rapidly diversifying clades
505 will normally have many short branches, making their internal relationships more uncertain.
506 The Punctuated model adds more substitutions to every branch regardless of its duration in
507 time, and so helps resolve these rapidly diversifying clades. The precise impact of topological
508 error is difficult to interpret, since our investigation does not distinguish among topological
509 errors at different levels and we do not designate any particular nodes as important. But the
510 practical upshot of systematically lower topological error in the simulation with the most
511 severe node age errors under UCLN is that these errors will be difficult to spot. Major
512 topological errors are easier to diagnose by checking against classical taxonomies or other
513 information, while little is usually known about the dates of uncalibrated nodes ahead of
514 time.
515
516 As an alternative method of diagnosis, we tried conducting a sister pairs analysis on
517 each set of simulated and reconstructed trees to see if an empirical relationship could be
518 recovered (Figure 6). However, tree reconstruction under either the UCLN or ACLN rate
519 prior disrupts the relationships between branch lengths and clade sizes displayed by the
520 underlying simulated trees (Figure 4). Under the UCLN they are even inverted, giving a
521 negative correlation between clade sizes and branch lengths. This could be related to the fact
522 that lineages with many speciation events have higher variance in branch lengths under
523 UCLN (102) which means their total length could be more strongly constrained by the node
524 time (tree) prior. Conversely, under ACLN we found significant positive relationships
525 between inferred branch lengths and clade size for all three models, even the Unlinked model
526 where speciation rates and substitution rates are not positively associated. A possible cause
527 of this inference of a positive assertion between diversification rates and substitution rate,

528 even in the absence of an underlying association between the two, is the node-density effect,
529 which results from systematic underestimation of branch lengths with high uncertainty (103).
530 The influence of the node density effect should be reduced by our procedure of selecting an
531 even number of samples in each sister clade, but may not be eliminated by it due to the
532 extremely high error rates under ACLN. The failure to recover the empirical correlation from
533 branch lengths reconstructed as part of a Bayesian dating analysis suggests that this is not a
534 useful way to test for the presence of interactions between substitution rates and speciation
535 rates. For some data sets it may be practical to check for such correlations by reconstructing
536 branch lengths in units of genetic divergence using a fast branch length reconstruction
537 method that does not impose a model of timing or rate variation (e.g. 104, 105-107).
538 However, even if such correlations can be diagnosed, there is currently no available method
539 for correcting any associated bias. This should be an area of active research.
540
541 The true role of punctuational bursts or coevolution of speciation and substitution
542 rates in macroevolution is currently a matter of debate. The primary empirical evidence is the
543 analysis of a large selection of empirical phylogenies in (63) and (62). These studies have
544 found that the length of root-to-tip paths in the phylogenies in substitutions per site is
545 frequently higher when there are more speciation events on the path. The relationship
546 remains even after removing the node-density effect that can generate spurious correlations
547 between branch lengths and node densities across clades (103, 108). Regardless of the
548 mechanisms, we have shown that if there were punctuated busts of substitutions associated
549 with speciation events, we could expect it to impact the accuracy of molecular date estimates,
550 since any mechanism that leads to similar patterns of rate variation and apparent correlation
551 between branch lengths and speciation events would produce similar effects.
552

553 *3.5 Molecular dating methodology*

554

555 Our study contrasts two common reconstruction methods, rather than varying aspects
556 of data choice. Many aspects of data choice and analytical method could exacerbate or reduce
557 the errors found in this analysis. For example, our alignments are shorter than many used in
558 actual Bayesian phylogenetic analyses but also likely more informative due to being
559 simulated by a simple method without invariant sites or many real sources of substitution
560 model misspecification. Shorter or less informative alignments could lead to greater or more
561 widespread error. Perhaps the most important element that we did not explore was the
562 influence of calibrations. Calibration choice and positioning, as well as the shape of the
563 calibrating distribution used, is often the most influential factor influencing divergence time
564 reconstruction (e.g. 32, 109) and is the source of many disagreements in published molecular
565 dating studies, including high-profile dating studies in birds (110-112) mammals (113, 114),
566 and insects (115-117). Simulation studies have suggested that calibrations can rescue
567 analyses strongly affected by rate prior misspecification (30), although others have found that
568 sufficiently misspecified rate priors cannot be rescued by more calibrations (34). The largest
569 and most data-rich molecular dating studies may have many more calibrations than we
570 implement (e.g. 139, 118). However our trees are smaller than in these studies, and our
571 calibrations are guaranteed to be correctly placed, have hard constraints, and are centered on
572 the true value. We also ensure the root is calibrated, which has been shown to have the largest
573 effect on accuracy in several studies (30, 34, 119). We therefore believe that our method
574 overall represents a balance between realism and conservatism in determining the amount of
575 calibrating information available to inform divergence time estimates. However, any
576 extension of the analyses undertaken in this study would certainly indicate investigation of
577 the effects of both alignment length and calibration numbers and positioning.

578

579 **4. Conclusions**

580

581 Our results here show the value of testing molecular phylogenetic and dating methods
582 against complex simulation models informed by empirically determined patterns of rate
583 variation. We have demonstrated that commonly applied Bayesian divergence time
584 estimation methods can experience average errors of 12% in node dates under a model in
585 which speciation and molecular rates vary independently, 14% when substitution rates are
586 linked to speciation rates and 20% under a punctuational model when analysed using an
587 ucorrelated rates prior (UCLN) in BEAST 2. Errors were much higher under all models when
588 trees were analysed with an autocorrelated rate prior (ACLN) in PAML. As demonstrated by
589 the negative skewness of gamma statistics, for trees with links between speciation and
590 molecular evolution these errors could lead to systematic overestimation of node ages.
591 Regardless of the mechanism generating the association between speciation rates and
592 substitution rates, we show that the potential for divergence time estimation error associated
593 with known empirical relationships between molecular evolution (manifest in substitution
594 rate) and macroevolution (manifest in the topology and branch lengths of phylogenetic trees)
595 can be significant and should be taken into account. Future work should see a broader range
596 of sophisticated divergence time estimation methods be tested against a wider variety of more
597 empirically realistic simulated data, including geographic and environmental biases,
598 concerted changes in population size, and directional changes in life history characteristics.
599 Ultimately the goal should be to produce new models that relax the assumption of
600 independence between priors on divergence times and priors on substitution rates. These will
601 lead to greater confidence in applications that depend on robust estimates from molecular
602 dating procedures.

5.Methods

5.2 Simulation trees and alignments

To simulate the phylogeny, we adopt a forward approach, starting at the root of the phylogeny and moving forward in time, as lineages divide into two daughter lineages, go extinct, and evolve in speciation and substitution rate as a continuous time Markov process. In each step, we draw the time until the next event from an exponential distribution with rate equal to the sum of the rates of all events in all lineages. Then, for every lineage l that is present at time t, one of the following three events may take place:

(i) The lineage gives rise to two daughter lineages with rate $\lambda(l, t)$, the current speciation rate of lineage l at time t. Both daughter lineages initially retain the same speciation rate and trait values as their parent lineages.

(ii) The lineage becomes extinct with rate μ, which is constant over the tree.

(iii) The lineage updates its speciation rate and substitution rate with rate q, which is constant throughout the tree. The update process adds an increment to the current speciation rate, which is a random draw from a bivariate normal distribution with mean 0 and an instantaneous variance-covariance matrix Σ multiplied by the time since last jump. Until one of these events happens, the lineage continues with the same speciation and molecular rates. High values of q make for longer simulations but produce a better approximation to Brownian motion. We found that q had no influence on observed speciation or substitution rate variance or average tree age for $q > 10$. For our final

simulations, we selected q = 50 to allow a significant margin of safety.

Specifically, the instantaneous variance-covariance matrix takes the form:

$$\Sigma = \begin{bmatrix} \sigma_\lambda^2 & \mathrm{Cov}(\lambda, r) \\ \mathrm{Cov}(\lambda, r) & \sigma_r^2 \end{bmatrix}$$

where σ_λ^2 describes how fast the speciation rate evolves and σ_r^2 describes how fast the substitution rate evolves. $\mathrm{Cov}(\lambda, r) = \rho\sigma_\lambda\sigma_r$, where ρ is the correlation coefficient between speciation rate and substitution rate, so $\rho = 0$ in the Unlinked Model and the Punctuated model. We employ the Generalised Sampling Approach (120) to ensure a correct distribution of trees with a given number of tips, which is set to 75, the approximate mean size of the avian clades taken from (80). Details of the approach is in Supplementary Methods S1.2.

In the Punctuated model, we simulate an initial tree as per the Unlinked sample, but without removing extinct taxa. We generate a burst of substitutions following each speciation event. We generate Punctuated trees to match the formula found by (62), in which 16 $\pm$ 5.4% of the tree length (sum of all branch lengths) results from node-associated bursts (Supplementary Methods S1.3).

Once we have simulated a tree, we simulate the evolution of nucleotide sequences along the tree, given that branch lengths in the tree represent the expected number of substitutions per site along the branch. The branch lengths in this tree represent the expected number of substitutions per site along the branch. We simulate alignments of 2000 bases, which is equivalent to a 6000-base coding alignment, since most substitutions in a coding alignment occur in the third codon sites. A 6000-base coding alignment is about the size of the concatenated coding sequences of a mitochondrial genome. Previous literature reviews have shown this to be a common length for phylogenetic data sets in published ecological

studies, and a tractable length for large Bayesian simulation studies (86). We parameterize our sequence simulations using realistic among-site rate variation values from birds (Supplementary Methods S1.4).

5.3 Calibrations

In addition to the trees and sequence alignments, we need calibrations to provide information on absolute time to allow substitution rates to be identifiable from absolute ages of nodes. For simulated trees of similar size to ours, diminishing accuracy returns for adding new calibrations are achieved between 1 and 5 calibrations, especially if one is at the root (30). Since our calibration dates are known with more accuracy than those in a real analysis, and we remove the possibility of incorrectly assigned calibrations, we choose to assign two calibrations per tree. For each tree, we choose calibrations at the root node and at one randomly chosen internal node. To ensure that the calibration is randomly chosen across the nodes of the simulated trees, we first decide the level of the calibrated node as a random integer between 1 and the largest number of nodes along any root-to-tip path in the tree. Then we randomly pick the calibrated node from all the nodes at that level.

For analyses in both BEAST 2 and PAML, we apply a hard-bounded uniform prior for the age of each calibrated node. Both calibrations have bounds at the true age ± 15%. Note that this is a conservative test of molecular dating because in the simulated datasets we have accurate knowledge of the true date of the speciation event. In reality, calibrations may be misplaced or incorrectly dated, and may only provide maximum or minimum bounds.

5.4 Phylogenetic inference

We generate a set of 50 trees under each of the Unlinked, Continuous and Punctuated simulation types. After sequence simulation, this gives us 50 sequence alignments for each of the three models.We infer divergence times from the simulated sequences using common Bayesian molecular dating procedures. First, we generate posterior distributions of tree topologies and divergence times using two common molecular dating methods: the uncorrelated lognormal rate prior (UCLN), implemented in the software package 'BEAST 2' (BEAST 2; 121), in which rates for each branch are independent draws from a common lognormal distribution, and the autocorrelated lognormal rate prior (ACLN), implemented in the 'mcmctree' program in the PAML software suite (105), in which the rate of each branch is lognormally distributed with a mean equal to the rate of its parent branch. To speed analysis, we used the approximate likelihood algorithm implemented in mcmctree. The PAML analyses differ from the BEAST analyses in that PAML requires a fixed tree topology, which in this case was the topology of the true simulated tree. Therefore only BEAST incorporates the possibility of topological inference errors. However, the three simulation models can be compared within each reconstruction method. Details of the inference method are available as Supplementary Methods (S1.5).

5.5 Error in reconstructed trees

For each reconstructed tree, we calculated accuracy in the estimation of node times using the median absolute percentage error (MAPE) of inferred node times for clades that are present in both true and reconstructed trees. The median was chosen rather than the mean to reduce the impact of extremely high values for the estimates of shallow nodes. Since these are only calculated for correctly reconstructed branches and do not capture errors in

reconstructing topology, we also calculate the Robinson-Foulds distance or number of non-shared bipartitions, which only reflects typographical error (87).

Model misspecification in Bayesian molecular dating methods can sometimes be associated with higher uncertainty in the reconstructed node ages, as measured by the width of Bayesian credible intervals. This is undesirable because the reconstructed history is less useful as an explanatory tool and less informative in downstream analyses, but it can also be a simple way to diagnose problems with the specified model. We examined the uncertainty with which node ages were reconstructed by the two Bayesian molecular dating methods. We scored uncertainty by averaging the width of the reconstructed 95% highest posterior density (HPD) intervals as a proportion of reconstructed node height across each tree, with a larger score indicating less precise estimates.

5.6 Distribution of reconstructed vs simulated node ages

We examine the direction of error by calculating the gamma statistic for each tree and comparing to the true tree (122). The gamma statistic is a description of how the times between branching events in a tree whose tips are all extant taxa change as they become closer to the present. It summarises information relating to the pattern of diversification, following a standard normal distribution for a tree with a constant rate of speciation and no extinction. Trees with higher values have a relatively greater proportion of their length in older branches, while those with lower values have relatively more of their length in younger branches. The statistic therefore tells us whether the ages of internal nodes within each tree tend to be older or younger on average than the true times while accounting for the different scale of this bias for older and younger nodes.

727

728 *5.7 Efficacy of sister pair correlation on reconstructed branch lengths*

729

730 We checked to see if we can use the reconstructed branch lengths, which are the
731 median of posterior branch lengths calculated by BEAST or mcmctee, to detect the
732 correlation between branch lengths and clade sizes found by sister pair analysis on the
733 simulated trees (see *Simulating realistic rate variation* above). This could be used as a rapid
734 post-hoc check to see whether published dated trees could be affected by reconstruction error
735 that is associated with substitution and speciation rate covariation. Data sets for the sister pair
736 analysis are generated from the 50 simulated trees under each of the three simulation models,
737 where each tree gives a pair of sister clades and the sister clades are the two clades that split
738 at the root. So, under each simulation model, we have a data set of 50 sister pairs. We then
739 regress contrasts in log clade sizes against contrasts in the log of the median branch length
740 estimates from BEAST or mcmctree.

741
742
743 **References**
744

745 1. Zuckerkandl E, Pauling L, Bryson V, Vogel H. Evolving genes and proteins. 1965.
746 2. Dickerson RE. The structure of cytochrome c and the rates of molecular evolution. J
747 Mol Evol. 1971;1(1):26-45.
748 3. Bromham L. Why do species vary in their rate of molecular evolution? Biol Lett.
749 2009;5:401-4.
750 4. Bromham L. The genome as a life-history character: why rate of molecular evolution
751 varies between mammal species. Phil Trans Roy Soc B. 2011;366(1577):2503-13.
752 5. Barraclough TG, Savolainen V. Evolutionary rates and species diversity in flowering
753 plants. Evolution. 2001;55:677-83.
754 6. Bromham L, Hua X, Lanfear R, Cowman P. Exploring the relationships between
755 mutation rates, life history, genome size, environment and species richness in flowering
756 plants. Am Nat. 2015;185:507-24.
757 7. Lanfear R, Ho SYW, Love D, Bromham L. Mutation rate is linked to diversification
758 in birds. Proc Natl Acad Sci U S A. 2010;107(47):20423-8.
759 8. Eo SH, DeWoody JA. Evolutionary rates of mitochondrial genomes correspond to
760 diversification rates and to contemporary species richness in birds and reptiles. Proc R Soc B.
761 2010 277:3587-92.

762 9. Gillman LN, Keeling DJ, Gardner RC, Wright SD. Faster evolution of highly
763 conserved DNA in tropical plants. J Evol Biol. 2010;23(6):1327-30.
764 10. Gillman LN, Wright SD. Species richness and evolutionary speed: the influence of
765 temperature, water and area. J Biogeogr. 2014;41(1):39-51.
766 11. Orton MG, May JA, Ly W, Lee DJ, Adamowicz SJ. Is molecular evolution faster in
767 the tropics? Heredity (Edinb). 2019;122(5):513-24.
768 12. Ho SYW. The changing face of the molecular evolutionary clock. Trends Ecol Evol.
769 2014;29(9):496-503.
770 13. Kumar S, Hedges SB. Advances in time estimation methods for molecular data. Mol
771 Biol Evol. 2016;33(4):863-9.
772 14. Kumar S. Molecular clocks: four decades of evolution. Nat Rev Genet.
773 2005;6(8):654-62.
774 15. dos Reis M, Donoghue PC, Yang Z. Bayesian molecular clock dating of species
775 divergences in the genomics era. Nat Rev Genet. 2016;17(2):71-80.
776 16. Ho SYW, Duchêne S. Molecular-clock methods for estimating evolutionary rates and
777 timescales. Mol Ecol. 2014;23(24):5947-65.
778 17. Heath TA, Holder MT, Huelsenbeck JP. A Dirichlet process prior for estimating
779 lineage-specific substitution rates. Mol Biol Evol. 2012;29(3):939-55.
780 18. Yoder AD, Yang Z. Estimation of primate speciation dates using local molecular
781 clocks. Mol Biol Evol. 2000;17(7):1081-90.
782 19. Drummond AJ, Suchard MA. Bayesian random local clocks, or one rate to rule them
783 all. BMC Biol. 2010;8(1):114.
784 20. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating
785 with confidence. PLOS Biol. 2006;4(5):e88.
786 21. Sanderson MJ. r8s: inferring absolute rates of molecular evolution and divergence
787 times in the absence of a molecular clock. Bioinformatics. 2003;19(2):301-2.
788 22. Smith SA, O'Meara BC. treePL: divergence time estimation using penalized
789 likelihood for large phylogenies. Bioinformatics. 2012;28(20):2689-90.
790 23. Sauquet H. A practical guide to molecular dating. Comptes Rendus Palevol.
791 2013;12(6):355-67.
792 24. Wilke T, Schultheiß R, Albrecht C. As time goes by: a simple fool's guide to
793 molecular clock approaches in invertebrates. Am Malac B. 2009;27(1/2):25-45.
794 25. Thorne JL, Kishino H, Painter IS. Estimating the rate of evolution of the rate of
795 molecular evolution. Mol Biol Evol. 1998;15(12):1647-57.
796 26. Kishino H, Thorne JL, Bruno WJ. Performance of a divergence time estimation
797 method under a probabilistic model of rate evolution. Mol Biol Evol. 2001;18(3):352-61.
798 27. Aris-Brosou S, Yang Z. Effects of models of rate evolution on estimation of
799 divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Syst
800 Biol. 2002;51(5):703-14.
801 28. Lepage T, Lawi S, Tupper P, Bryant D. Continuous and tractable models for the
802 variation of evolutionary rates. Math Biosci. 2006;199(2):216-33.
803 29. Sarver BAJ, Pennell MW, Brown JW, Keeble S, Hardwick KM, Sullivan J, et al. The
804 choice of tree prior and molecular clock does not substantially affect phylogenetic inferences
805 of diversification rates. PeerJ. 2019;7.
806 30. Duchêne S, Lanfear R, Ho SYW. The impact of calibration and clock-model choice
807 on molecular estimates of divergence times. Mol Phylogenet Evol. 2014;78:277-89.
808 31. Lepage T, Bryant D, Philippe H, Lartillot N. A general comparison of relaxed
809 molecular clock models. Mol Biol Evol. 2007;24(12):2669-80.

32. Foster CSP, Sauquet H, van der Merwe M, McPherson H, Rossetto M, Ho SYW. Evaluating the impact of genomic data and priors on Bayesian estimates of the angiosperm evolutionary timescale. Syst Biol. 2017;66(3):338-51.
33. Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood DTJ. Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata). Mol Biol Evol. 2006;23(10):1832-51.
34. Battistuzzi FU, Filipski A, Hedges SB, Kumar S. Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. Mol Biol Evol. 2010;27(6):1289-300.
35. Worobey M, Han G-Z, Rambaut A. A synchronized global sweep of the internal genes of modern avian influenza virus. Nature. 2014;508(7495):254-7.
36. Crisp MD, Hardy NB, Cook LG. Clock model makes a large difference to age estimates of long-stemmed clades with no internal calibration: a test using Australian grasstrees. BMC Evol Biol. 2014;14:263-.
37. dos Reis M, Thawornwattana Y, Angelis K, Telford Maximilian J, Donoghue Philip CJ, Yang Z. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. Curr Biol. 2015;25(22):2939-50.
38. Welch JJ, Bromham L. Molecular dating when rates vary. Trends in Ecology & Evolution. 2005;20(6):320-7.
39. Lin G, Huang Z, Wang L, Chen Z, Zhang T, Gillman LN, et al. Evolutionary rates of bumblebee genomes are faster at lower elevations. Mol Biol Evol. 2019;36(6):1215-9.
40. Lourenço JM, Glémin S, Chiari Y, Galtier N. The determinants of the molecular substitution process in turtles. J Evol Biol. 2013;26(1):38-50.
41. Goldie X, Gillman L, Crisp M, Wright S. Evolutionary speed limited by water in arid Australia. Proc R Soc B. 2010;277(1694):2645-53.
42. Berv JS, Field DJ. Genomic signature of an avian Lilliput effect across the K-Pg extinction. Syst Biol. 2018;67(1):1-13.
43. Barrera-Redondo J, Ramirez-Barahona S, Eguiarte LE. Rates of molecular evolution in tree ferns are associated with body size, environmental temperature, and biological productivity. Evolution. 2018;72(5):1050-62.
44. Qiu F, Kitchen A, Burleigh JG, Miyamoto MM. Scombroid fishes provide novel insights into the trait/rate associations of molecular evolution. J Mol Evol. 2014;78(6):338-48.
45. May JA, Feng Z, Orton MG, Adamowicz SJ. The effects of ecological traits on the rate of molecular evolution in ray-finned fishes: a multivariable approach. J Mol Evol. 2020;88(8-9):689-702.
46. Welch JJ, Bininda-Emonds ORP, Bromham L. Correlates of substitution rate variation in mammalian protein-coding sequences. BMC Evol Biol. 2008;8(1):53.
47. Thomas JA, Welch JJ, Lanfear R, Bromham L. A generation time effect on the rate of molecular evolution in invertebrates. Mol Biol Evol. 2010;27(5):1173-80.
48. Hua X, Cowman P, Warren D, Bromham L. Longevity is linked to mitochondrial mutation rates in rockfish: a test using Poisson regression. Mol Biol Evol. 2015;32(10):2633-45.
49. Galtier N, Blier PU, Nabholz B. Inverse relationship between longevity and evolutionary rate of mitochondrial proteins in mammals and birds. Mitochondrion. 2009;9(1):51-7.
50. Bromham L, Cowman PF, Lanfear R. Parasitic plants have increased rates of molecular evolution across all three genomes. BMC Evol Biol. 2013;13(1):126.
51. Mitterboeck TF, Adamowicz SJ. Flight loss linked to faster molecular evolution in insects. Proc R Soc B. 2013;280(1767):20131128.

52. Wong A. Covariance between testes size and substitution rates in primates. Mol Biol Evol. 2014;31(6):1432-6.
53. Woolfit M, Bromham L. Population size and molecular evolution on islands. Proc Biol Sci. 2005;272(1578):2277-82.
54. Lynch M, Ackerman MS, Gout J-F, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet. 2016;17(11):704-14.
55. Gossmann TI, Keightley PD, Eyre-Walker A. The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. Genome Biol Evol. 2012;4(5):658-67.
56. Charlesworth B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 2009;10(3):195-205.
57. Mindell DP, Sites Jr JW, Graur D. Speciational evolution: a phylogenetic test with allozymes in Sceloporus (Reptilia). Cladistics. 1989;5(1):49-61.
58. Iglesias-Carrasco M, Jennions MD, Ho SYW, Duchene DA. Sexual selection, body mass and molecular evolution interact to predict diversification in birds. Proc R Soc B. 2019;286(1899).
59. Duchêne DA, Hua X, Bromham L. Phylogenetic estimates of diversification rate are affected by molecular rate variation. J Evol Biol. 2017;30(10):1884–97.
60. Lancaster LT. Molecular evolutionary rates predict both extinction and speciation in temperate angiosperm lineages. BMC Evol Biol. 2010;10(1):162.
61. Fontanillas E, Welch JJ, Thomas JA, Bromham L. The influence of body size and net diversification rate on molecular evolution during the radiation of animal phyla. BMC Ecol Evol. 2007;7:95.
62. Pagel M, Venditti C, Meade A. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. Science. 2006;314(5796):119.
63. Webster AJ, Payne RJH, Pagel M. Molecular phylogenies link rates of evolution and speciation. Science. 2003;301:478.
64. Ezard THG, Thomas GH, Purvis A. Inclusion of a near-complete fossil record reveals speciation-related molecular evolution. Methods Ecol Evol. 2013;4(8):745-53.
65. Hua X, Bromham L. Darwinism for the genomic age: connecting mutation to diversification. Front Genet. 2017;8:12.
66. Dobzhansky T. Genetics and the origin of species. New York: Columbia University Press; 1937.
67. Muller H, editor Isolating mechanisms, evolution, and temperature. Biol Symp; 1942.
68. Gavrilets S. Fitness landscapes and the origin of species (MPB-41): Princeton University Press; 2004.
69. Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics. 1995;139(4):1805-13.
70. Frankham R. Genetics and extinction. Biol Cons. 2005;126(2):131-40.
71. Venditti C, Pagel M. Speciation as an active force in promoting genetic evolution. Trends Ecol Evol. 2010;25(1):14-20.
72. Janzen T, Bokma F, Etienne RS. Nucleotide substitutions during speciation may explain substitution rate variation. bioRxiv. 2021:2020.08.19.256891.
73. Bromham L, Duchêne S, Hua X, Ritchie AM, Duchêne DA, Ho SYW. Bayesian molecular dating: opening up the black box. Biol Rev. 2018;93(2):1165-91.
74. Warnock RCM, Parham JF, Joyce WG, Lyson TR, Donoghue PCJ. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. Proc Biol Sci. 2015;282(1798):20141013-.

908 75. Yang Z, Rannala B. Bayesian estimation of species divergence times under a
909 molecular clock using multiple fossil calibrations with soft bounds. Mol Biol Evol.
910 2005;23(1):212-26.
911 76. Zhu T, Dos Reis M, Yang Z. Characterization of the uncertainty of divergence time
912 estimation under relaxed molecular clock models using multiple loci. Syst Biol.
913 2015;64(2):267-80.
914 77. dos Reis M, Yang Z. The unbearable uncertainty of Bayesian divergence time
915 estimation. J Systemat Evol. 2013;51(1):30-43.
916 78. Manceau M, Marin J, Morlon H, Lambert A. Model-based inference of punctuated
917 molecular evolution. Mol Biol Evol. 2020;37(11):3308-23.
918 79. Nabholz B, Lanfear R, Fuchs J. Body mass-corrected molecular rate for bird
919 mitochondrial DNA. Mol Ecol. 2016;25(18):4438-49.
920 80. Maliet O, Hartig F, Morlon H. A model with many small shifts for estimating species-
921 specific diversification rates. Nat Ecol Evol. 2019;3(7):1086-92.
922 81. Lanfear R, Ho SYW, Love D, Bromham L. Mutation rate influences diversification
923 rate in birds. Proc Natl Acad Sci U S A. 2010;107(47):20423-8.
924 82. Rabosky DL. Positive correlation between diversification rates and phenotypic
925 evolvability can mimic punctuated equilibrium on molecular phylogenies. Evolution.
926 2012;66(8):2622-7.
927 83. Pennell MW, Harmon LJ, Uyeda JC. Is there room for punctuated equilibrium in
928 macroevolution? Trends Ecol Evol. 2014;29(1):23-32.
929 84. Pennell MW, Harmon LJ, Uyeda JC. Speciation is unlikely to drive divergence rates.
930 Trends Ecol Evol. 2014;29(2):72-3.
931 85. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds
932 in space and time. Nature. 2012;491:444.
933 86. Ritchie AM, Hua X, Cardillo M, Yaxley KJ, Dinnage R, Bromham L. Phylogenetic
934 diversity metrics from molecular phylogenies: modelling expected degree of error under
935 realistic rate variation. Divers Distrib. 2020;n/a(n/a).
936 87. Robinson DF, Foulds LR. Comparison of phylogenetic trees. Math Biosci.
937 1981;53(1):131-47.
938 88. Bromham L. The genome as a life-history character: why rate of molecular evolution
939 varies between mammal species. Phil Trans Roy Soc B. 2011;366(1577):2503-13.
940 89. Gernhard T. The conditioned reconstructed process. J Theor Biol. 2008;253(4):769-
941 78.
942 90. Stadler T. Recovering speciation and extinction dynamics based on phylogenies. J
943 Evol Biol. 2013;26(6):1203-19.
944 91. Duchêne D, Duchêne S, Ho SYW. Tree imbalance causes a bias in phylogenetic
945 estimation of evolutionary timescales using heterochronous sequences. Mol Ecol Res.
946 2015;15(4):785-94.
947 92. Hedges SBK, Sudhir The Timetree of Life. New York: Oxford University Press;
948 2009.
949 93. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. Tree of life reveals clock-like
950 speciation and diversification. Mol Biol Evol. 2015;32(4):835-45.
951 94. McPeek MA, Brown JM. Clade age and not diversification rate explains species
952 richness among animal taxa. Am Nat. 2007;169(4):E97-106.
953 95. Marin J, Rapacciuolo G, Costa GC, Graham CH, Brooks TM, Young BE, et al.
954 Evolutionary time drives global tetrapod diversity. Proc R Soc B. 2018;285(1872):20172378.
955 96. Cardillo M. Latitude and rates of diversification in birds and butterflies. Proc R Soc
956 B. 1999;266(1425):1221-5.

957 97. Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, et al. Nine
958 exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc
959 Natl Acad Sci U S A. 2009;106(32):13410-4.
960 98. Kozak KH, Wiens JJ. Testing the relationships between diversification, species
961 richness, and trait evolution. Syst Biol. 2016;65(6):975-88.
962 99. Scholl JP, Wiens JJ. Diversification rates and species richness across the Tree of Life.
963 Proc R Soc B. 2016;283(1838):20161334.
964 100. Ho SY, Duchêne S, Duchêne D. Simulating and detecting autocorrelation of
965 molecular evolutionary rates among lineages. Mol Ecol Resour. 2015;15(4):688-96.
966 101. dos Reis M, Yang Z. Approximate likelihood calculation on a phylogeny for Bayesian
967 estimation of divergence times. Mol Biol Evol. 2011;28(7):2161-72.
968 102. Guindon S. Rates and rocks: strengths and weaknesses of molecular dating methods.
969 Front Genet. 2020;11:526-.
970 103. Hugall AF, Lee MS. The likelihood node density effect and consequences for
971 evolutionary studies of molecular rates. Evolution. 2007;61(10):2293-307.
972 104. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
973 large phylogenies. Bioinformatics. 2014;30(9):1312-3.
974 105. Yang ZH. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol.
975 2007;24(8):1586-91.
976 106. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective
977 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol.
978 2014;32(1):268-74.
979 107. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New
980 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
981 performance of PhyML 3.0. Syst Biol. 2010;59(3):307-21.
982 108. Venditti C, Meade A, Pagel M. Detecting the node-density artifact in phylogeny
983 reconstruction. Syst Biol. 2006;55(4):637-43.
984 109. Sauquet H, Ho SYW, Gandolfo MA, Jordan GJ, Wilf P, Cantrill DJ, et al. Testing the
985 impact of calibration on molecular divergence times using a fossil-rich group: the case of
986 Nothofagus (Fagales). Syst Biol. 2011;61(2):289-313.
987 110. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, et al. Whole-genome analyses
988 resolve early branches in the tree of life of modern birds. Science. 2014;346(6215):1320-31.
989 111. Mitchell KJ, Cooper A, Phillips MJ. Comment on "Whole-genome analyses resolve
990 early branches in the tree of life of modern birds". Science. 2015;349(6255):1460-.
991 112. Cracraft J, Houde P, Ho SYW, Mindell DP, Fjeldså J, Lindow B, et al. Response to
992 Comment on "Whole-genome analyses resolve early branches in the tree of life of modern
993 birds". Science. 2015;349(6255):1460-.
994 113. Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, et al. Impacts
995 of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification.
996 Science. 2011;334(6055):521-4.
997 114. Phillips MJ. Geomolecular dating and the origin of placental mammals. Syst Biol.
998 2016;65(3):546-57.
999 115. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics
1000 resolves the timing and pattern of insect evolution. Science. 2014;346(6210):763-7.
1001 116. Tong KJ, Duchêne S, Ho SYW, Lo N. Comment on "Phylogenomics resolves the
1002 timing and pattern of insect evolution". Science. 2015;349(6247):487-.
1003 117. Kjer KM, Ware JL, Rust J, Wappler T, Lanfear R, Jermiin LS, et al. Response to
1004 Comment on "Phylogenomics resolves the timing and pattern of insect evolution". Science.
1005 2015;349(6247):487-.

118. Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, et al. An inverse latitudinal gradient in speciation rate for marine fishes. Nature. 2018;559(7714):392-5.
119. Wertheim JO, Sanderson MJ. Estimating diversification rates: how useful are divergence times? Evolution. 2011;65(2):309-20.
120. Hartmann K, Wong D, Stadler T. Sampling trees from evolutionary models. Syst Biol. 2010;59(4):465-76.
121. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Comp Biol. 2019;15(4):e1006650.
122. Pybus OG, Harvey PH. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc Biol Sci. 2000;267(1459):2267-72.

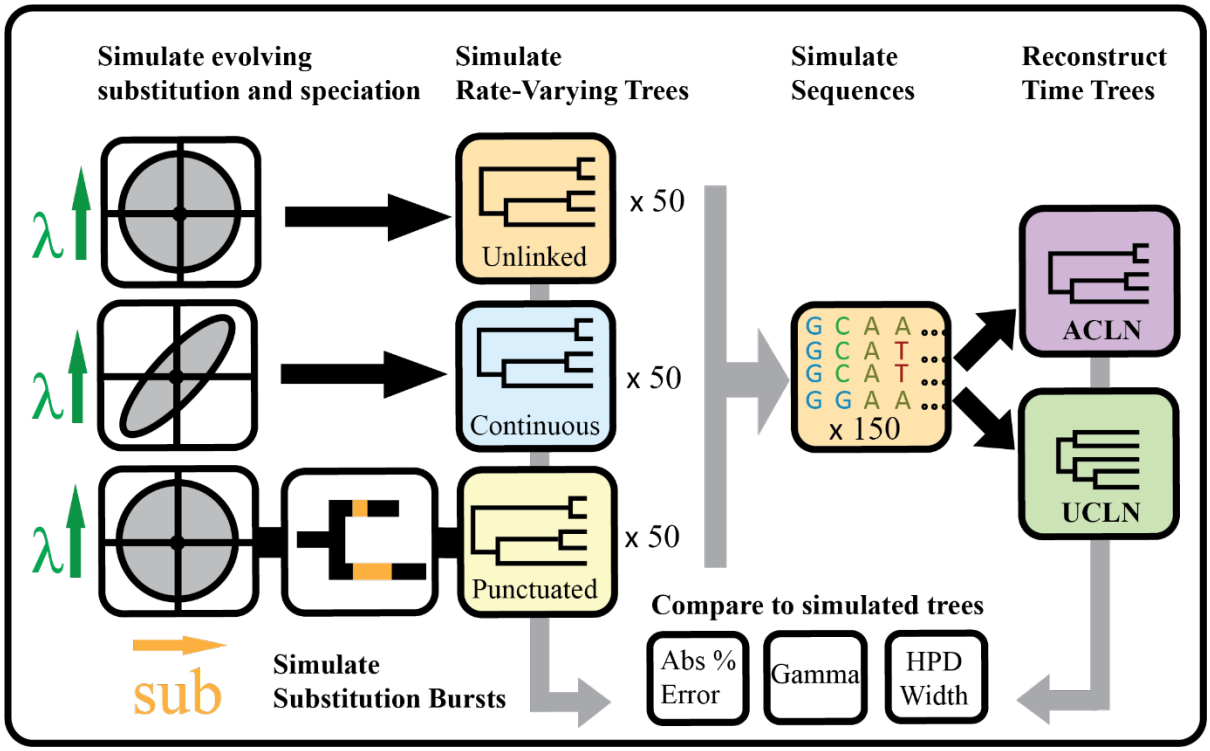
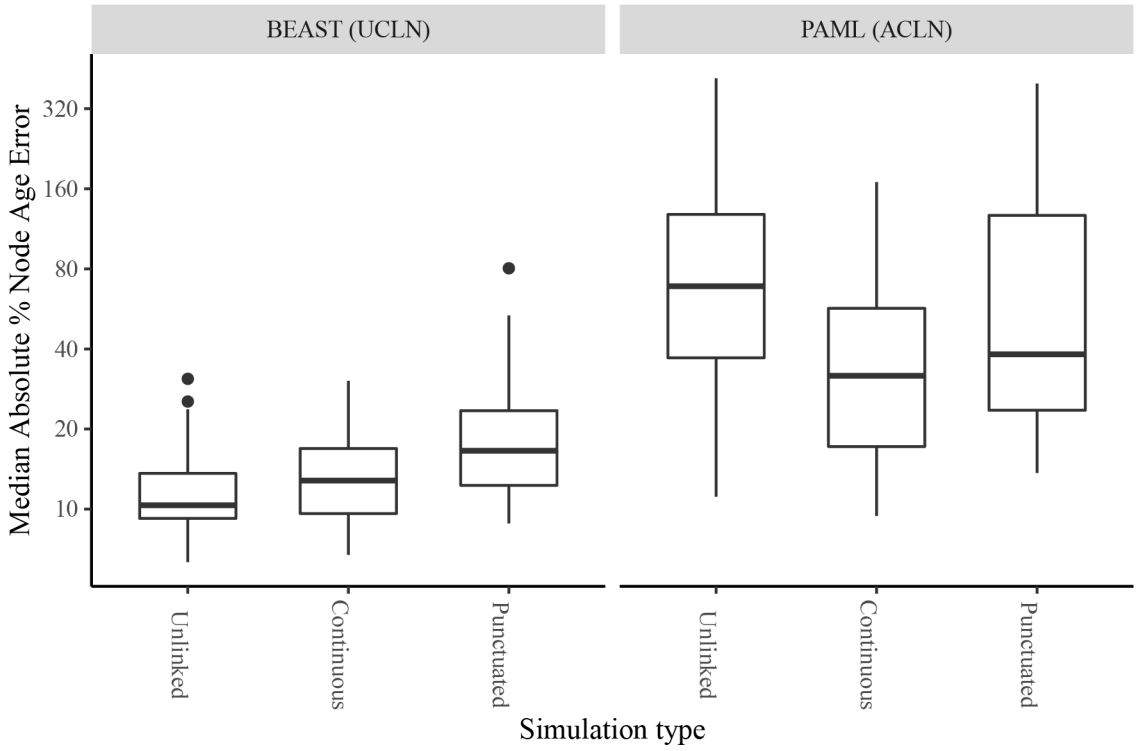


Figure 1. Flowchart of study design. Trees with branch lengths in units of time are simulated under three simulation conditions. For the Unlinked model, birth rates (λ)and substitution rates (sub) vary independently, while for the Continuous model they are correlated. For the Punctuated model, trees are simulated as for Unlinked, then some of the branch length is redistributed to form bursts of substitutions occurring upon speciation. Trees and branch lengths are then used to simulate molecular sequence alignments. Time trees are then reconstructed under two rate priors, UCLN and ACLN, and compared to the original, simulated trees to determine the error (Medain abs % error), uncertainty (HPD width) and bias towards over- or under-estimation (Gamma).



1045

1046 **Figure 2.** Median absolute percentage error of node times (on log scale) reconstructed from
1047 data sets simulated under three simulation types with different relationships between
1048 molecular evolution and speciation. The mean error is calculated only for branches whose
1049 bipartitions are shared between the true and inferred trees, and therefore ignores errors in
1050 topology which may occur in the BEAST analyses. The three simulation models are
1051 Unlinked (instantaneous covariance of molecular rates and speciation rates = 0), Continuous
1052 (instantaneous covariance = 0.0044), and Punctuated (instantaneous covariance = 0, bursts of
1053 substitutions added at speciation events). Topologies and node times were reconstructed
1054 using two different analytical methods, the uncorrelated lognormal 'relaxed clock' rate prior
1055 (UCLN) in BEAST 2, and the autocorrelated lognormal rate prior (ACLN) in PAML.

1056
1057

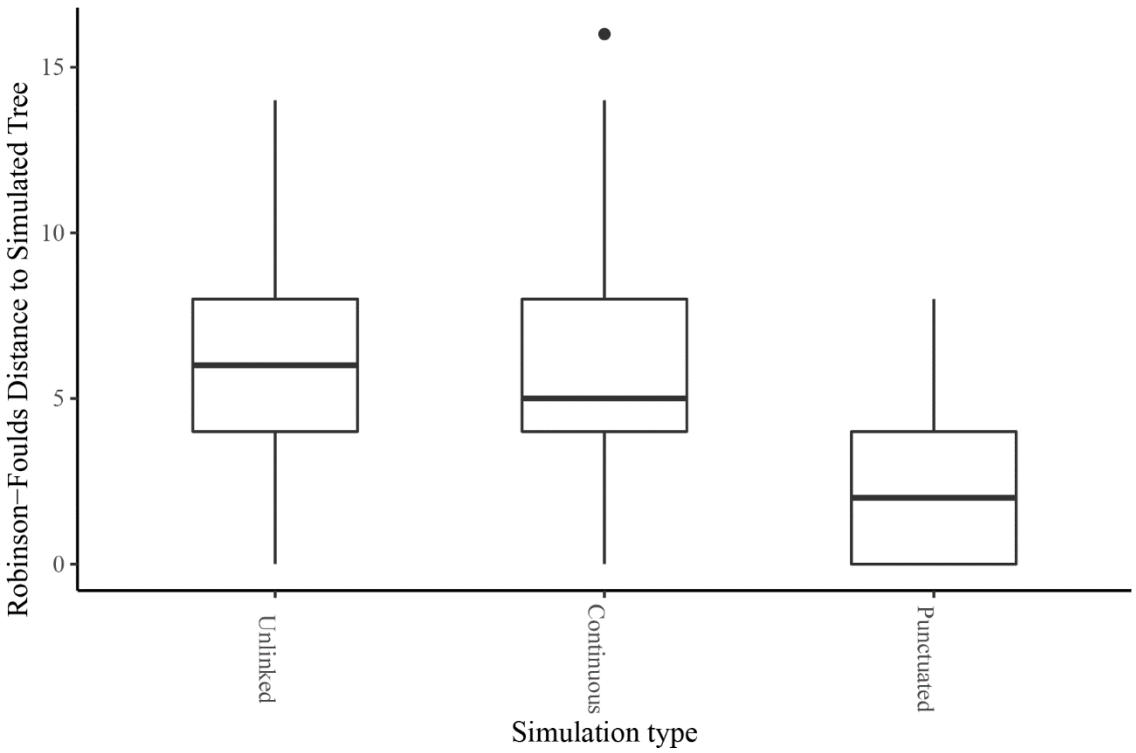
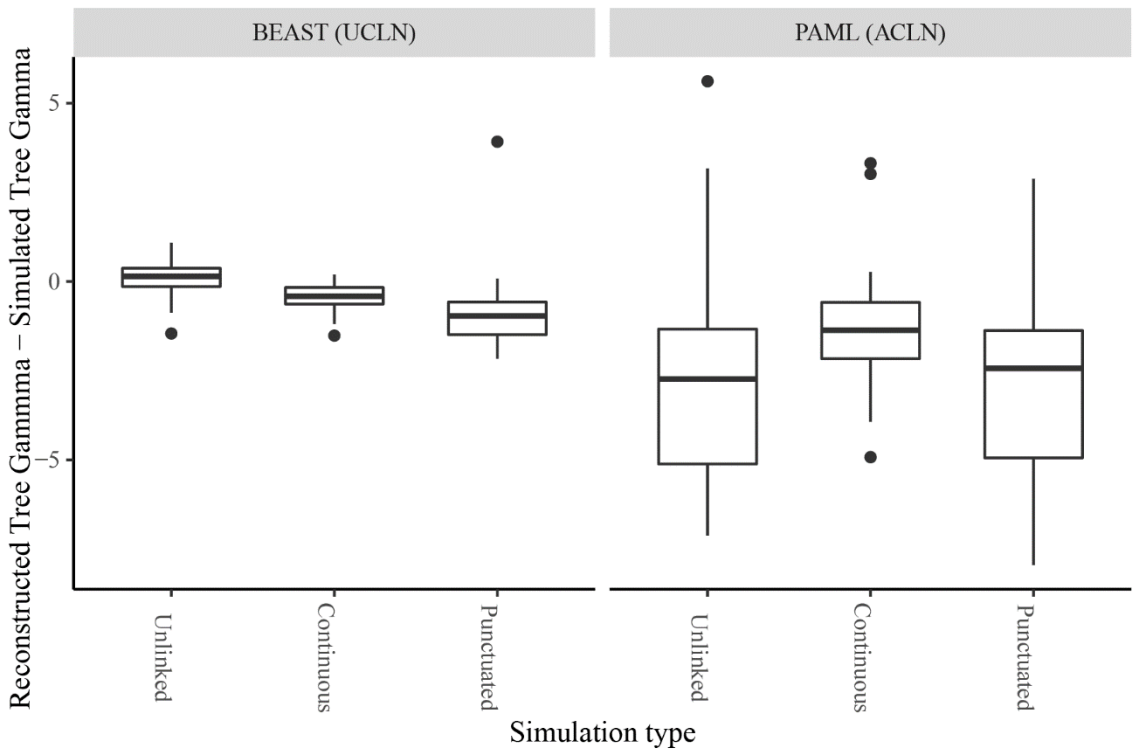


Figure 3. Topological error of trees reconstructed from data sets simulated under three simulation types with different relationships between molecular evolution and speciation. Topological error is given by the Robinson-Foulds distance (number of non-shared bipartitions). This distance incorporates both topological and branch time errors. The three simulation models are Unlinked (instantaneous covariance of molecular rates and speciation rates = 0), Continuous (instantaneous covariance = 0.0044), and Punctuated (instantaneous covariance = 0, bursts of substitutions added at speciation events). Topologies and node times were reconstructed using the uncorrelated lognormal 'relaxed clock' rate prior (UCLN) in BEAST 2. No results are available for the autocorrelated lognormal rate prior in PAML because the topology was fixed for the analysis.



1070

1071
1072

1073 **Figure 4.** Difference in gamma statistic values between reconstructed and simulated trees,
1074 where simulated trees were generated under three models with different relationships
1075 between molecular evolution and speciation. Negative scores indicate that reconstructed
1076 trees have a greater proportion of their branch times in tips vs internal branches ('tippier')
1077 than the simulated tree. The three simulation models are Unlinked (instantaneous covariance
1078 of molecular rates and speciation rates = 0), Continuous (instantaneous covariance = 0.0044),
1079 and Punctuated (instantaneous covariance = 0, bursts of substitutions added at speciation
1080 events). Topologies and node times were reconstructed using two different analytical
1081 methods, the uncorrelated lognormal 'relaxed clock' rate prior (UCLN) in BEAST 2, and the
1082 autocorrelated lognormal rate prior (ACLN) in PAML.

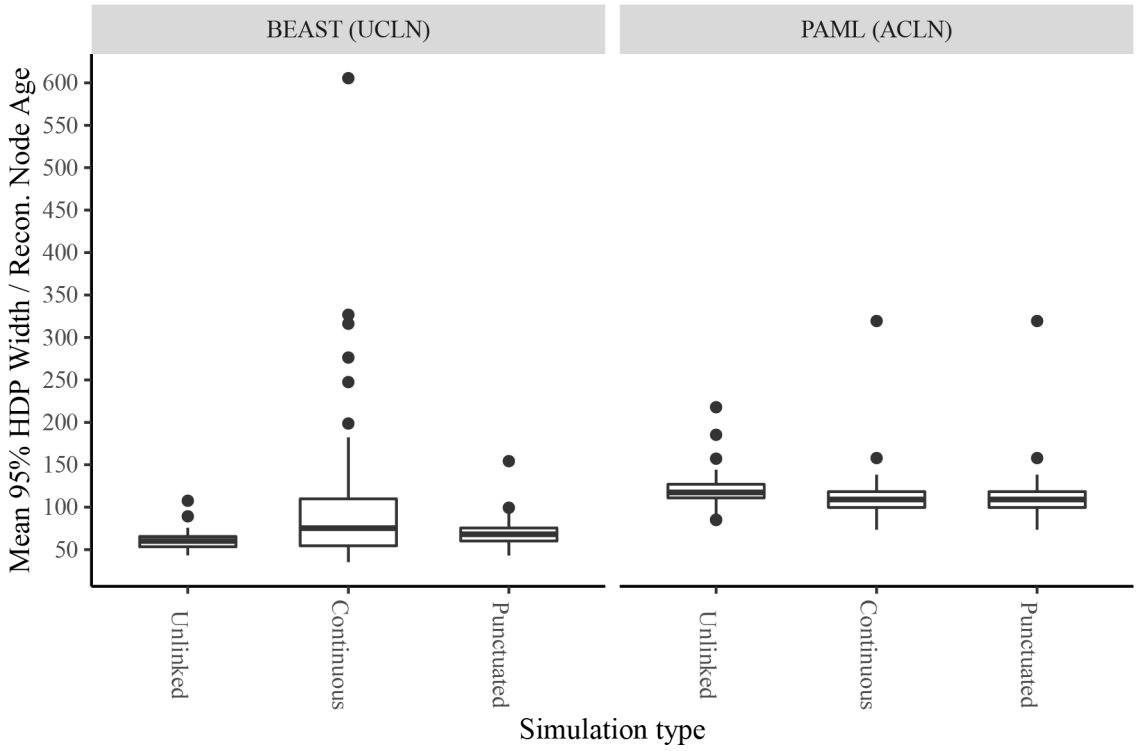


Figure 5. Uncertainty of trees reconstructed from phylogenies simulated under three models with different relationships between molecular evolution and speciation, in width as a percentage of reconstructed node age. Precision is measured by averaging the width of 95% highest posterior density (HPD) intervals as a proportion of reconstructed node height across each tree, with a larger score indicating less precise estimates. The three simulation models are Unlinked (instantaneous covariance of molecular rates and speciation rates = 0), Continuous (instantaneous covariance = 0.0044), and Punctuated (instantaneous covariance = 0, bursts of substitutions added at speciation events). Topologies and node times were reconstructed using two different analytical methods, the uncorrelated lognormal 'relaxed clock' rate prior (UCLN) in BEAST 2, and the autocorrelated lognormal rate prior (ACLN) in PAML.

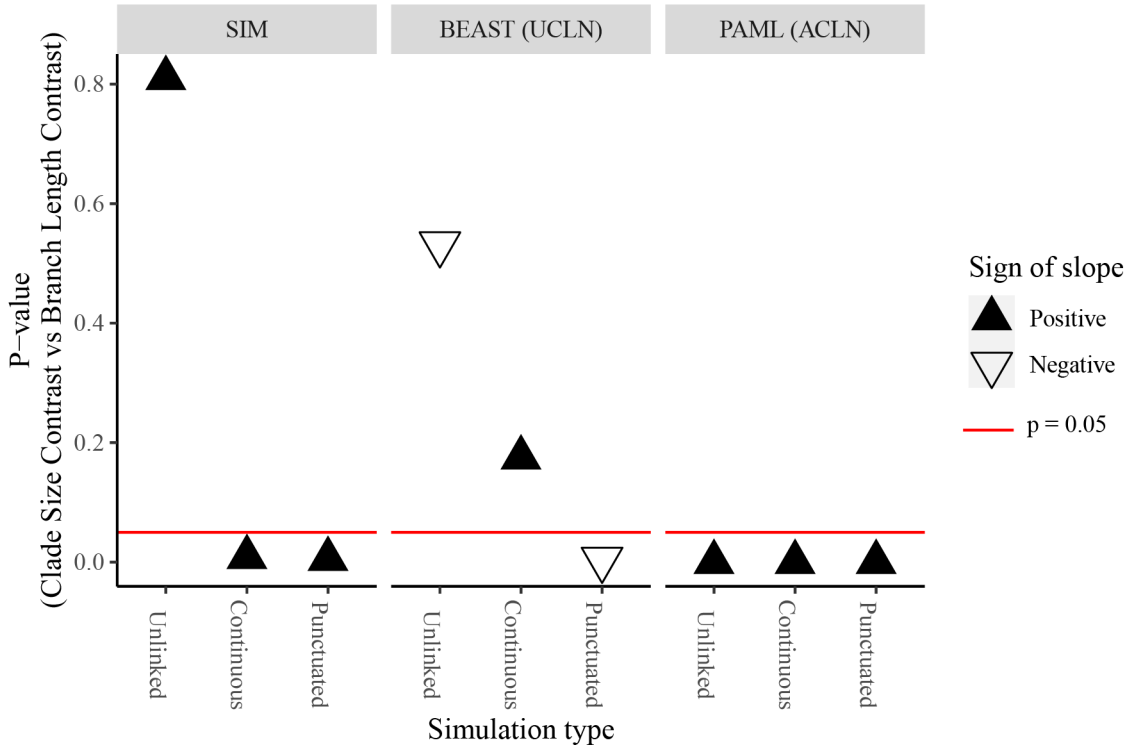
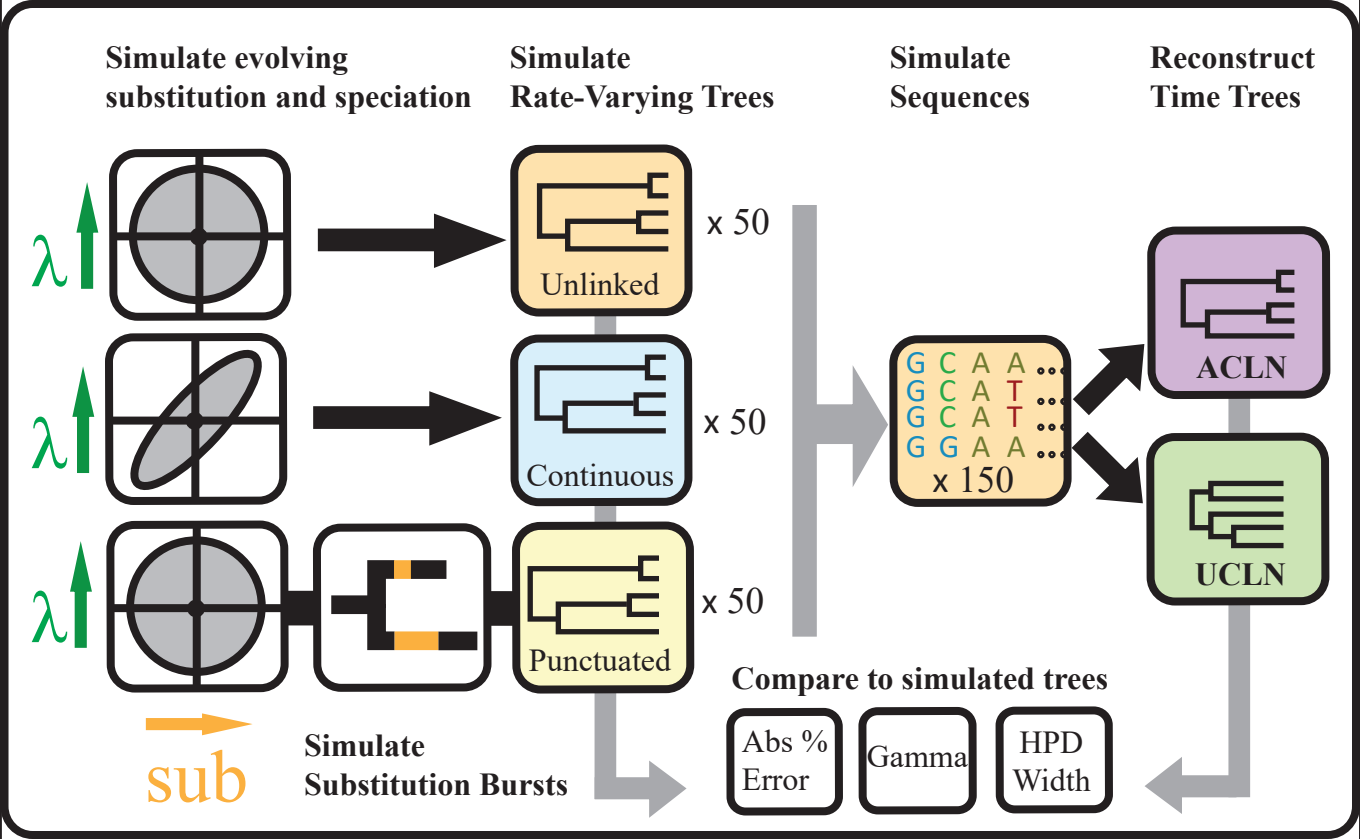
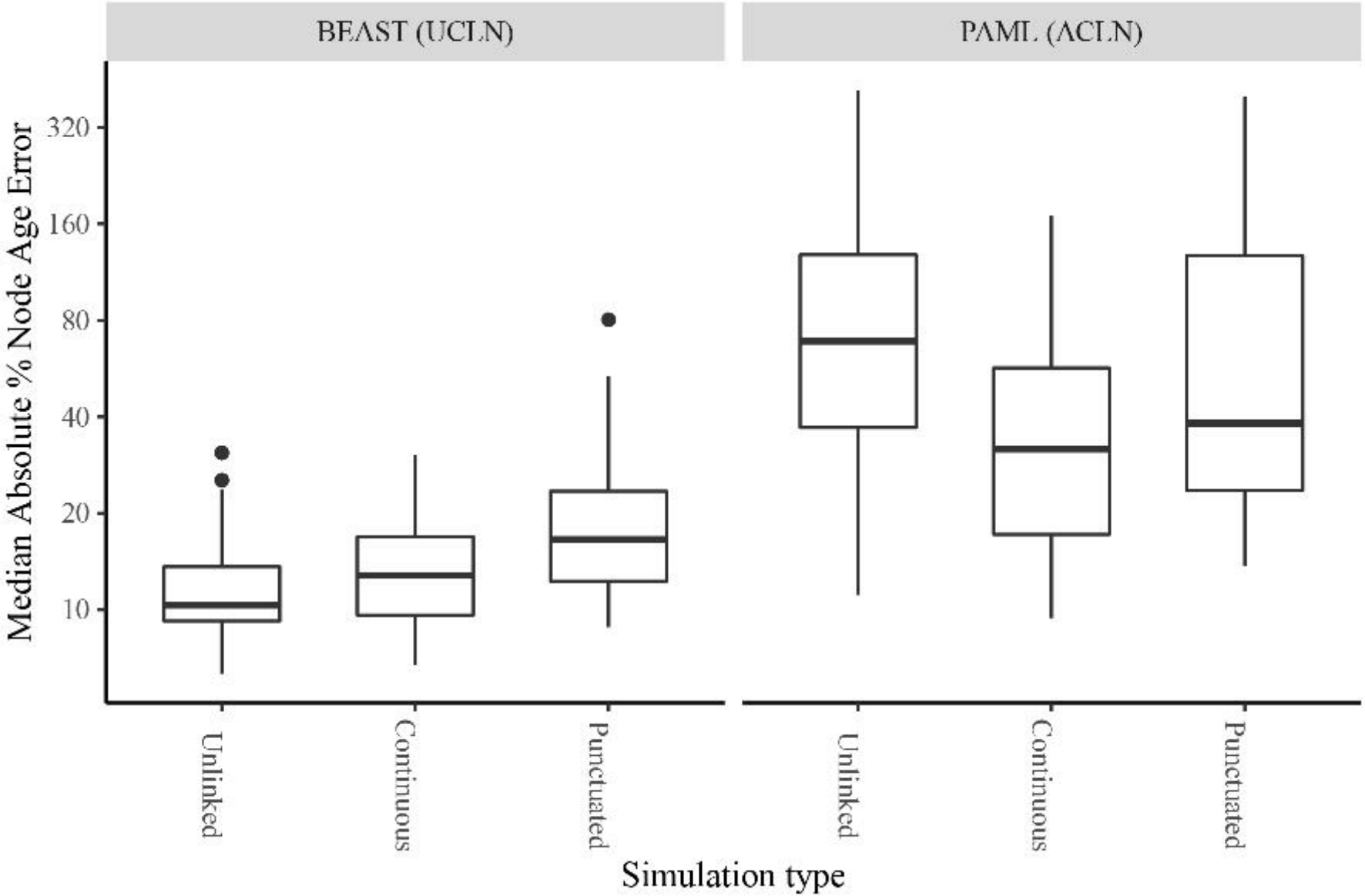
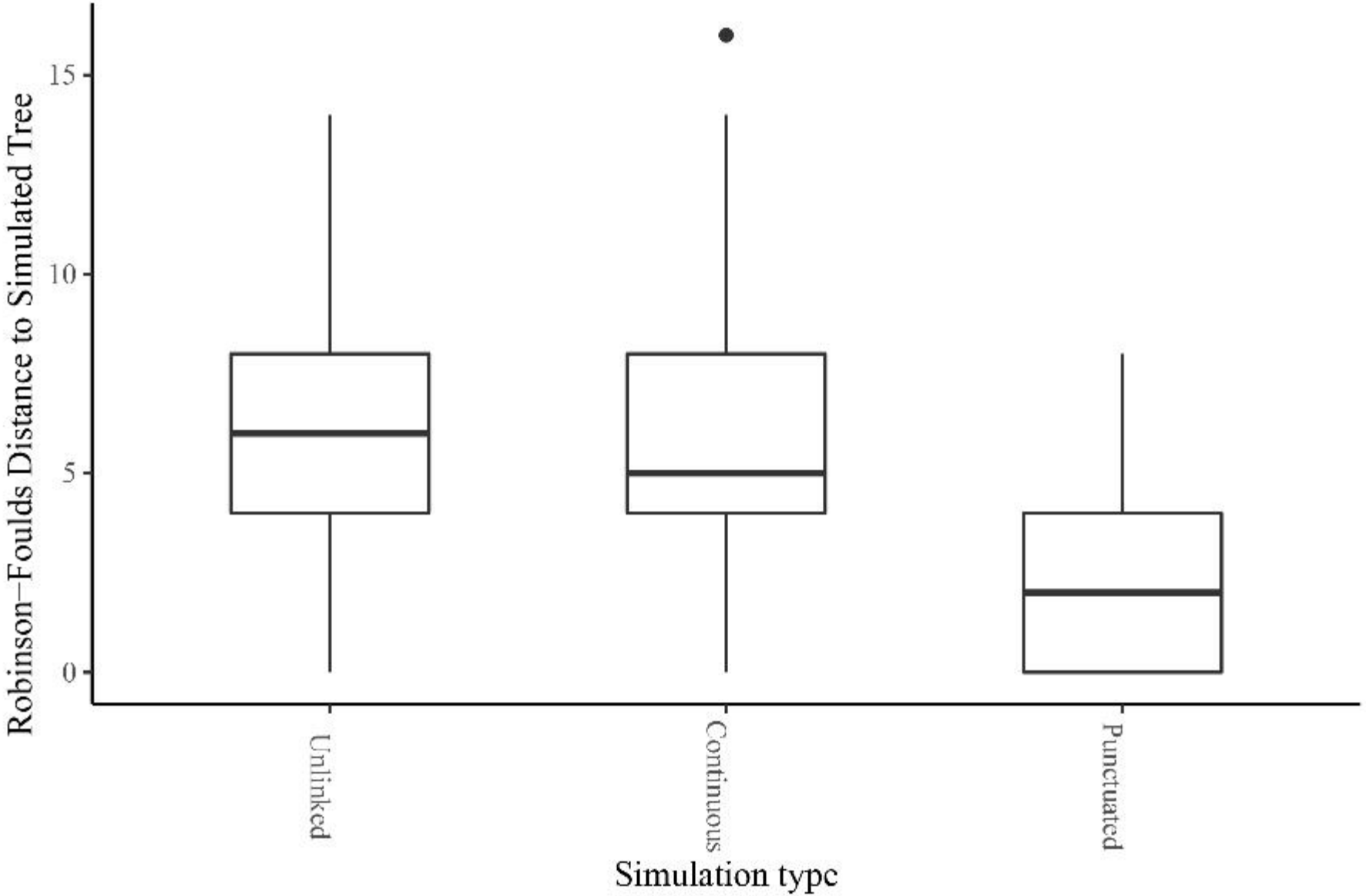
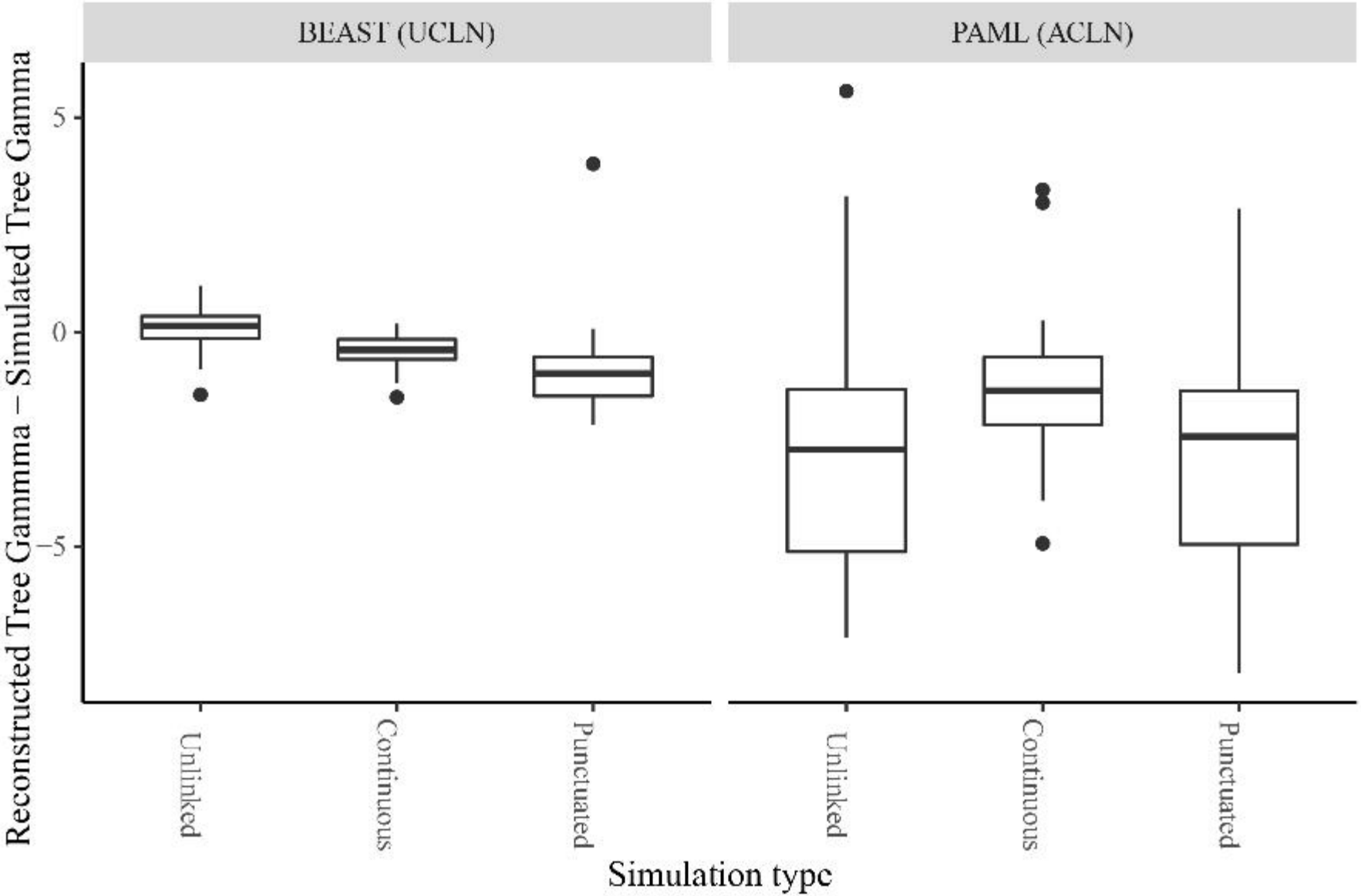


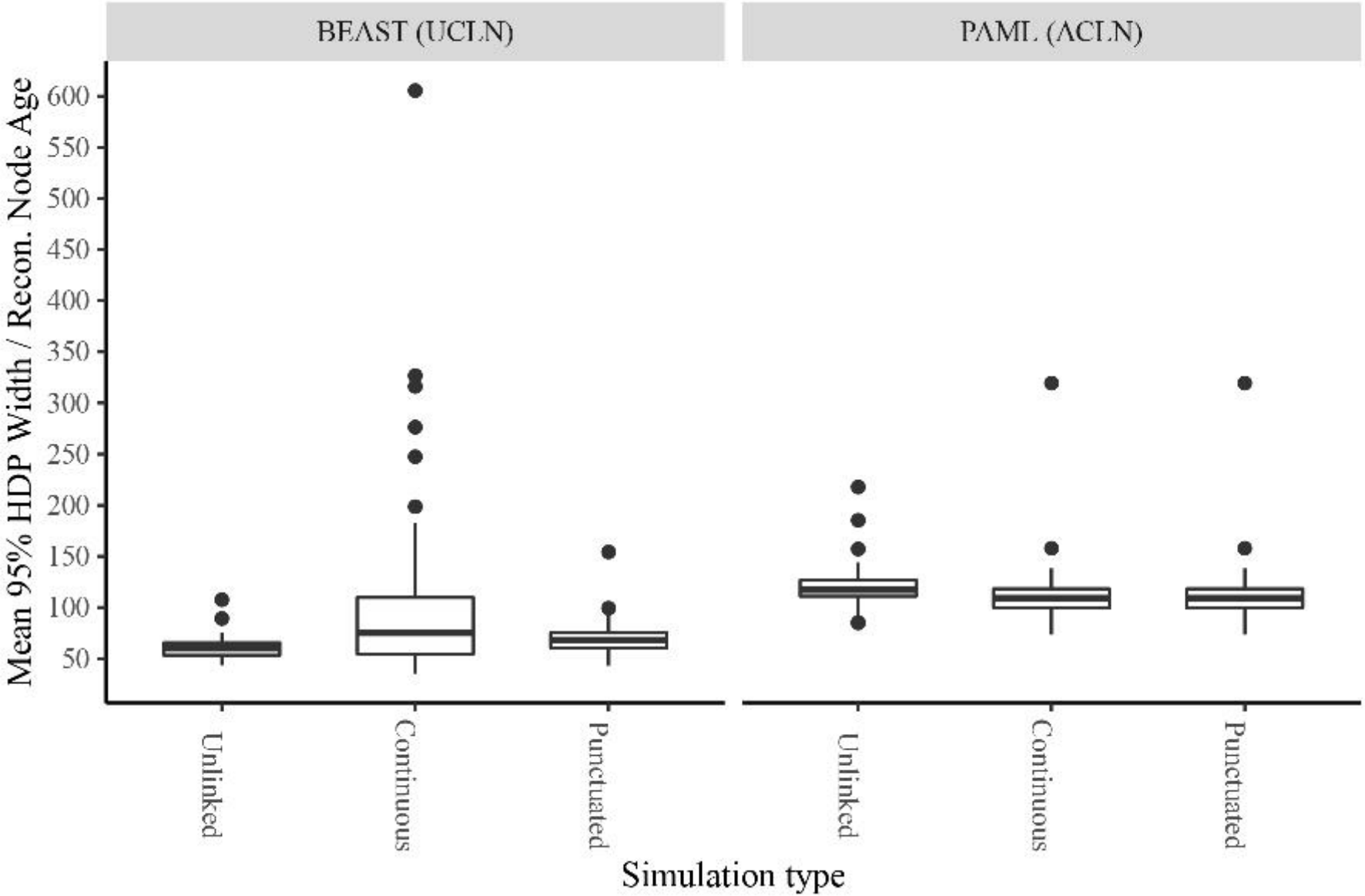
Figure 6. Slope and significance of mock sister pair analyses of clade size contrasts against contrasts in simulated branch lengths (SIM) or median branch length estimates reconstructed from simulated trees. Results are shown for branch lengths reconstructed under the uncorrelated lognormal clock (UCLN) or autocorrelated lognormal clock (ACLN). The three simulation models are Unlinked (instantaneous covariance of molecular rates and speciation rates = 0), Continuous (instantaneous covariance = 0.0044), and Punctuated (instantaneous covariance = 0, bursts of substitutions added at speciation events). Black triangles represent positive regression coefficients, while white inverted triangles represent negative coefficients. The red line indicates significance at the 0.05 level.

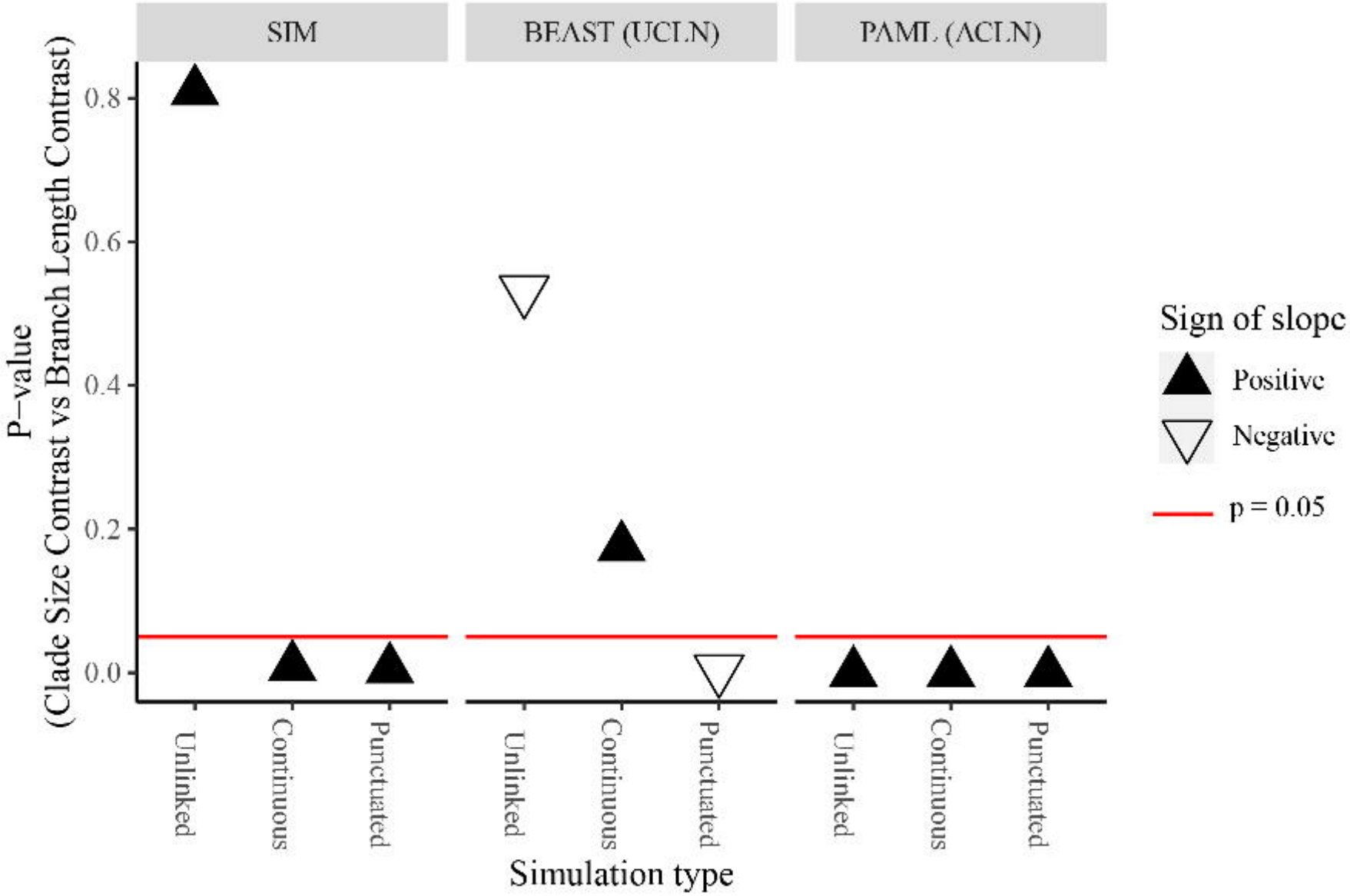

Simulate evolving substitution and speciation
Simulate Rate-Varying Trees
Simulate Sequences
Reconstruct Time Trees
λ
λ
λ
Unlinked
Continuous
Punctuated
x 50
x 50
x 50
G C A A ...
G C A T ...
G C A T ...
G G A A ...
x 150
ACLN
UCLN
sub
Simulate Substitution Bursts
Compare to simulated trees
Abs % Error
Gamma
HPD Width


BEAST (UCLN)
PAML (ACLN)
Median Absolute % Node Age Error
320
160
80
40
20
10
Unlinked
Continuous
Punctuated
Unlinked
Continuous
Punctuated
Simulation type


Robinson−Foulds Distance to Simulated Tree
15
10
5
0
Unlinked
Continuous
Punctuated
Simulation type


BEAST (UCLN)
PAML (ACLN)
Reconstructed Tree Gammma – Simulated Tree Gamma
5
0
−5
Unlinked
Continuous
Punctuated
Unlinked
Continuous
Punctuated
Simulation type


BEAST (UCLN)
PAML (ACLN)
Mean 95% HDP Width / Recon. Node Age
600
550
500
450
400
350
300
250
200
150
100
50
Unlinked
Continuous
Punctuated
Unlinked
Continuous
Punctuated
Simulation type


SIM
BEAST (UCLN)
PAML (ACLN)
P−value
(Clade Size Contrast vs Branch Length Contrast)
0.8
0.6
0.4
0.2
0.0
Unlinked
Continuous
Punctuated
Unlinked
Continuous
Punctuated
Unlinked
Continuous
Punctuated
Simulation type
Sign of slope
Positive
Negative
p = 0.05