Nucleotide substitutions during speciation may explain substitution rate variation

Thijs Janzen[1*], Folmer Bokma[2], Rampal S. Etienne[1]

[1] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Box 11103, 9700 CC Groningen, The Netherlands

[2] Center for Ecological and Evolutionary Synthesis (CEES) Department of BioSciences, University of Oslo, PO Box 1066, Blindern, 0316 Oslo, Norway

* corresponding author: t.janzen@rug.nl

ABSTRACT

Although molecular mechanisms associated with the generation of mutations are highly conserved across taxa, there is widespread variation in mutation rates between evolutionary lineages. When phylogenies are reconstructed based on nucleotide sequences, such variation is typically accounted for by the assumption of a relaxed molecular clock, which, however, is just a statistical distribution of mutation rates without any underlying biological mechanism. Here, we propose that variation in accumulated mutations may be partly explained by an elevated mutation rate during speciation. Using simulations, we show how shifting mutations from branches to speciation events impacts inference of branching times in phylogenetic reconstruction. Furthermore, the resulting nucleotide alignments are better described by a relaxed than by a strict molecular clock. Thus, elevated mutation rates during speciation potentially explain part of the variation in substitution rates that is observed across the tree of life.

Keywords: molecular clock, speciation, phylogenetic reconstruction, substitution rate variation

38 INTRODUCTION

39

40 Phenotypic diversification occurs at a higher rate in some clades than in others
41 (Simpson 1945; van Valen 1985; Ricklefs 2006; Rabosky et al. 2007; Jansson and
42 Davies 2008) and similarly, there is substantial variation across evolutionary
43 lineages in the rate of molecular evolution (King and Wilson 1975), such as that
44 of nucleotide sequences (Nabholz et al. 2008; Bromham 2011; Dowle et al. 2013;
45 Sung et al. 2016). As a consequence, studies attempting to reconstruct the
46 phylogeny of a clade often find that the sequence data do not support the
47 assumption of a strict molecular clock, i.e. constant substitution rates across
48 lineages. For such cases, phylogenetic inference software allows one to use a
49 relaxed molecular clock (Drummond et al. 2006; Lepage et al. 2007), which
50 assumes that the substitution rate varies between lineages according to a
51 statistical distribution such as a gamma or lognormal distribution. However, the
52 relaxed molecular clock thus introduces at least one additional degree of
53 freedom, namely the variance of the distribution of substitution rates. Moreover,
54 it is a rather ad-hoc solution with little underlying biological reasoning.

55

56 Here, we present a more explicit biological explanation for variation in
57 substitution rates across the tree of life; we posit that differences in apparent
58 substitution rates between lineages are due to processes acting exclusively or
59 predominantly during speciation. Such processes might include founder effects,
60 bottlenecks, inbreeding, hybridization, selection for an increased mutation rate,
61 divergent selection and local adaptation, all of which could lead to a short,
62 temporary surge in the substitution rate. Due to (effectively) random extinction

63 of lineages, different branches of a reconstructed phylogeny will differ in how
64 often they experienced such short episodes of accelerated substitution rates,
65 resulting in differences in apparent substitutions rates along these branches. It is
66 worthwhile to investigate and model this process, because empirical data
67 indicate that inferred substitution rates correlate with numbers of branching
68 events (Pagel et al. 2006; Eo and DeWoody 2010; Lanfear et al. 2010; Duchene
69 and Bromham 2013; Bromham et al. 2015 but see Goldie et al. 2011), indicating
70 that such a process might drive observed patterns.
71

METHODS

We propose a model where substitutions accumulate not only along the branches of a phylogeny, but also at speciation events including not only the internal nodes of the phylogeny but also those pruned from the phylogeny by extinction. We first make the standard assumption that gradual sequence evolution along a phylogenetic branch can be modeled as a time-homogeneous Markov process with substitution matrix:

$$\mathbf{Q} = \begin{bmatrix} -\mu_{AT}-\mu_{AC}-\mu_{AG} & \mu_{AT} & \mu_{AC} & \mu_{AG} \\ \mu_{TA} & -\mu_{TA}-\mu_{TC}-\mu_{TG} & \mu_{TC} & \mu_{TG} \\ \mu_{CA} & \mu_{CT} & -\mu_{CA}-\mu_{CT}-\mu_{CG} & \mu_{CG} \\ \mu_{GA} & \mu_{GT} & \mu_{GC} & -\mu_{GA}-\mu_{GT}-\mu_{GC} \end{bmatrix}$$

where μ_{ij} denotes the mutation rate from nucleotide i to nucleotide j. The transition probabilities of nucleotide substitutions after time t of gradual sequence evolution are then given by the matrix

$$\mathbf{P_a}(t) = \exp(\mathbf{Q}t)$$

where the subscript **a** indicates anagenetic change, i.e. gradual accumulation of substitutions over time. This matrix can be multiplied with an initial probability vector at time $t = 0$ to yield the probabilities for each of the four nucleotides at time t.

In addition to gradual sequence evolution over time, we assume that sequences may change rapidly during speciation. We can thus assume another matrix $\mathbf{P_c}$ (subscript c for "cladogenetic") that describes the nucleotide transition probabilities during a single speciation event.

The processes that may accelerate sequence evolution during speciation, such as founder effects, bottlenecks, inbreeding, hybridization, and adaptation to novel environments, may well result in different kinds of substitutions than those that take place over time in established species. However, for mathematical convenience we will here assume that we can write:

$$\mathbf{P}_{\mathrm{c}} = \exp\left(\mathbf{Q}\tau\right)$$

where τ is a parameter that measures the effect of substitutions during speciation. In other words, we assume that nucleotide sequence evolution is only accelerated during speciation events, but not qualitatively altered. The acceleration is then measured by parameter τ: a single speciation event causes as much sequence evolution as τ years of gradual evolution over time. Note that this interpretation requires that the μ_{ij} used in $\mathbf{P}_{\mathrm{c}}$ must be identical to that used in $\mathbf{P}_{\mathrm{a}}$. For $\tau = 0$, $\mathbf{P}_{\mathrm{c}}$ becomes the identity matrix, and our model reduces to the standard model of sequence evolution that only assumes substitutions along phylogenetic branches.

For simplicity we assume in our simulations that sequence evolution can be modeled as a Jukes-Cantor process, for which $\mathbf{Q}$ is given by:

$$\mathbf{Q} = \begin{pmatrix} -\frac{3\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & -\frac{3\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & -\frac{3\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & -\frac{3\mu}{4} \end{pmatrix}$$

Given a rooted phylogeny and a root sequence (e.g. some arbitrary sequence assumed to represent the ancestral sequence), several existing software packages (e.g. the R package phangorn, Schliep 2011; the python module pyvolve, Sipos *et al.* 2011; the R package phylosim, Spielman & Wilke 2015) provide algorithms to simulate sequence evolution along the branches of the phylogeny, by applying the transition matrix sequentially along the phylogenetic tree. Here, we extend this methodology to also include substitutions accumulated at the nodes of the phylogeny. We implemented this in the R package 'nodeSub' (https://github.com/thijsjanzen/nodeSub).

120 *Testing the impact of node substitution models using simulations*

121 To identify the amount of error in phylogenetic inference caused by
122 assuming a (relaxed) molecular clock when substitutions actually arise (in part)
123 during speciation, we simulated sequence evolution on known trees and then
124 reconstructed the phylogeny from the simulated sequences, assuming strict and
125 relaxed molecular clocks. We simulated sequence evolution with the node
126 substitution model introduced above, with various degrees of acceleration (τ) at
127 the nodes of the tree (including $\tau = 0$), and with various extinction rates. We then
128 compared the resulting trees with the original true tree using a number of
129 statistics: the gamma statistic (Pybus and Harvey 2000), the beta statistic
130 (Aldous 2001), the mean branch length (Faith 1992; Clarke and Warwick 2001),
131 tree height, the normalized Lineages Through Time (nLTT) statistic (Janzen et al.
132 2015) and the Jenson-Shannon distance metric comparing the Laplacian
133 spectrum (Lewitus and Morlon 2016).

134 To correct for any error in phylogenetic inference incurred by processes
135 independent of the node substitution model chosen (e.g. stochasticity in
136 substitution accumulation, stochasticity in phylogenetic tree creation), we also
137 inferred a phylogenetic tree for a *twin* (sensu Bilderbeek, Laudanno & Etienne
138 2020) alignment, where we simulated an alignment with the exact same number
139 of substitutions, using the same true tree, but using a standard, strict-clock,
140 substitution model. Then, for the distribution of trees resulting from this *twin*
141 alignment, we estimated the same summary statistics. The error introduced by
142 the node substitution model is then the difference between the error of the node
143 substitution posterior and the error in the *twin* posterior.

144 Phylogenetic reconstruction was performed with BEAST2 (Bouckaert et
145 al. 2019) using the R package *babette* (Bilderbeek and Etienne 2018). BEAST2
146 inference was performed using default priors with a birth-death prior as tree
147 prior (or a Yule prior if the extinction rate is zero), and assuming the JC69
148 nucleotide substitution model and a strict clock model. The BEAST chain was run
149 for 10 million steps, whilst sampling a tree every 5000 steps. After completion,
150 the first 10% of the chain was discarded as burn-in.
151
152 *Obtaining a twin alignment*
153 We generated a twin alignment conditional on a phylogeny, a node substitution
154 model, and a mutation rate. Because an alignment generated using a node
155 substitution model has accumulated substitutions at the nodes in addition to
156 those along the branches, generally the overall number of substitutions
157 accumulated is higher than an alignment simulated using a branch-only model.
158 Thus, in order to generate a *twin* alignment that contains the same amount of
159 information (substitutions) we need to increase the mutation rate. This can be
160 done by calculating the estimated time spent at the nodes, relative to the time
161 spent on the branches. This provides a correction factor which we used to make
162 a first educated guess as to which mutation rate is most likely to generate an
163 alignment with the branch-only model with the same number of substitutions.
164 We obtain:

$$\mu_{twin} = \mu\left(1 + \frac{\tau(2N + H)}{\sum t_{branch}}\right) \quad (1)$$

165 where μ is the mutation rate used in the node substitution model, τ is the time
166 spent on the node, N is the number of internal nodes in the tree, H is the number

167 of hidden nodes in the tree and $\sum t_{branch}$ is the total branch length of the tree. On
168 the assumption that the rates of speciation (b) and extinction (d) have been
169 constant throughout the tree, the number of hidden nodes H_{branch} on a branch
170 with endpoints t_0 and t_1 follows a Poisson distribution (Bokma et al. 2012;
171 Manceau et al. 2020):

$$H_{branch} = 2b(t_0 - t_1) + 2\ln\frac{b\exp\big((b-d)t_1\big) - d}{b\exp\big((b-d)t_0\big) - d} \quad (2)$$

172 We simulated sequences of 10 kb, with μ = 0.001. We used this educated guess as
173 a starting point to find the mutation rate that yields on average the same number
174 of mutations as observed in our alignment simulated with the node substitution
175 model (because the latter is just a single realization, which may be quite different
176 from the expected number of mutations). Then, using our corrected mutation
177 rate, we generated twin alignments until we obtained one exactly matching the
178 number of mutations of the alignment simulated with the node substitution
179 model.
180
181 *Simulation settings*
182 We generated birth-death trees with varying degrees of extinction rate d in [0,
183 0.1, 0.3, 0.5] and a single speciation rate of b = 1. Trees were simulated
184 conditional on 100 tips.
185
186 *Varying the time spent on the nodes relative to the crown age*
187 We varied τ in [0, 0.01, 0.05, 0.1, 0.2, 0.4] times the crown age (e.g. when the
188 crown age of the simulated tree is 3MY, τ = [0, 0.03, 0.15, 0.3, 0.6, 1.2] MY). Again,
189 for each combination of τ and extinction (d = [0, 0.1, 0.3, 0.5]) we simulated 100

trees and for each tree we generated one node substitution alignment and one *twin* alignment.

The impact of tree balance

In unbalanced trees, some terminal branches are connected by many more past branching events to the root of the tree than other terminal branches. Hence, we expect that balance of a tree might have a substantial effect on the error in phylogenetic inference: less balanced trees are expected to have higher error. To test this, we compared fully balanced (β = 10.0) with extremely unbalanced, caterpillar trees (β = -2). We did so by simulating the branching times of a birth-death tree, and assigning these to a fully balanced or fully unbalanced topology. Thus, the only difference between the trees is the topology. Then, for both the balanced and unbalanced tree a node substitution alignment was generated, with the same number of total substitutions, and setting τ as a function of crown age.

As an extra check, we also generated a node substitution alignment for the original birth-death tree from which the branching times were used. For all three alignments we inferred a phylogenetic tree as in the other scenarios, and compared the error in phylogenetic inference. Calculating the Laplacian Spectrum for caterpillar trees with extreme unbalance returned NA values, and hence we omitted the Laplacian Spectrum summary statistic in this analysis.

212 *Support for strict and relaxed clock models*

213 To test whether the alignment originating from a process with node
214 substitutions was better described by a relaxed than by a fixed clock model, we
215 repeated the analysis, but now inferred the marginal likelihood of the relaxed
216 clock and strict clock models using the "NS" package for BEAST2, which applies
217 Nested Sampling to obtain the marginal posterior likelihood for both models
218 (Russel et al. 2019). We used the function 'bbt_run' from the *babette* package
219 (Bilderbeek and Etienne 2018) in combination with the function
220 'create_ns_mcmc' from the *beautier* package (Bilderbeek and Etienne 2018). This
221 performs a Nested Sampling MCMC run using BEAST2, which runs until
222 convergence is detected. Then, we converted the obtained marginal likelihoods
223 to a relative weight for each model (by dividing both marginal likelihoods by
224 their sum) which allows for comparison of posterior support for each model
225 across parameter settings and trees.

226

RESULTS

Summary statistics

We compared summary statistics of trees inferred from alignments using the node substitution model, with summary statistics of *twin* trees inferred from alignments with identical information content, but generated without the node substitution model (e.g. with only substitutions along the branches, and a fixed clock rate). We find that particularly summary statistics influenced by branching times are affected (Figure 2, gamma and nLTT statistic). For these summary statistics, we find an increased difference as the node substitution rate τ increases. The differences in the value of these summary statistics manifest themselves across all extinction rates, although the effect tends to be somewhat smaller for higher extinction rates.

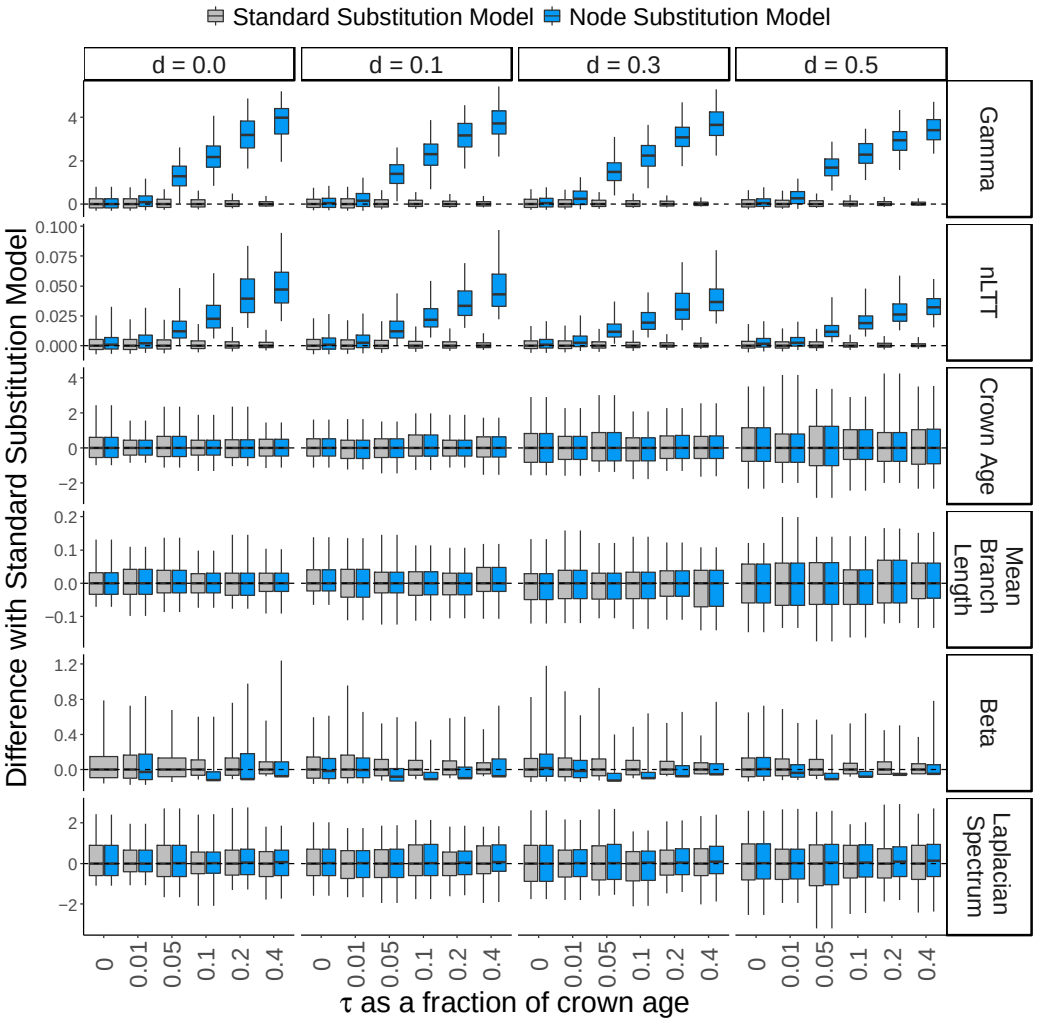

Figure 1. Difference in summary statistic values between trees inferred from an alignment generated with node substitutions (colored boxplots), and *twin* trees that were inferred from an alignment generated without node substitions (grey boxplots). Shown are the residual errors, reflecting distributions corrected by subtracting the median value of the *twin* trees. As a result, the distribution of the standard substitution model is always centered on 0. We explore τ as a fraction of crown age (horizontal axis), and the impact of extinction (d, columns). The dotted line indicates zero difference between the two substitution models. The summary statistics are the beta and gamma statistic, Laplacian spectrum, mean branch length, nLTT statistic, and crown age.

The impact of tree balance

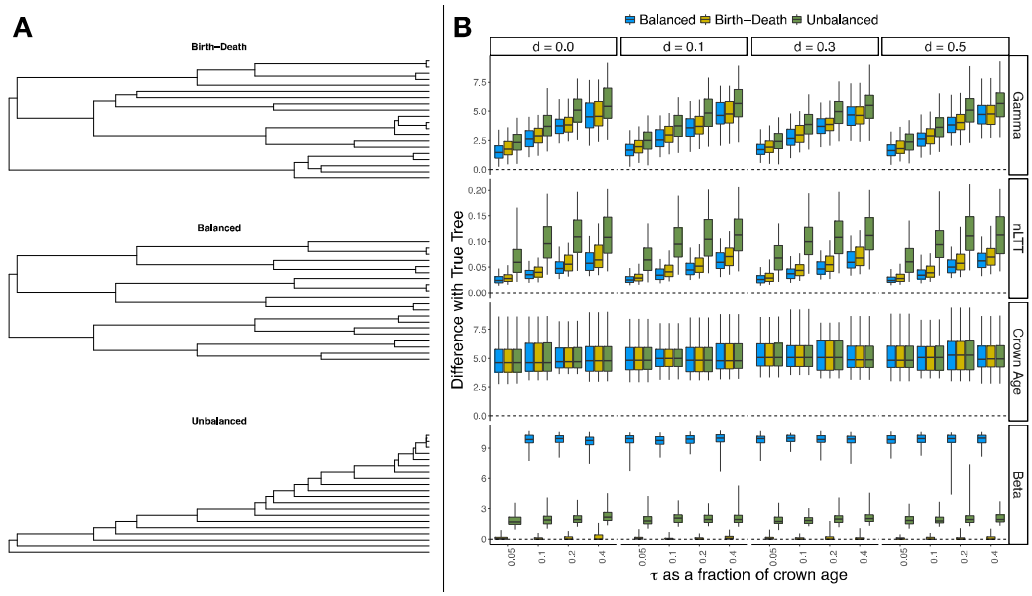

Figure 2. Effect of the node substitution model for phylogenies differing in balance **A**) example plots of a randomly generated birth-death tree (top), a fully balanced tree generated using the same branching times as the birth-death tree (middle) and a very unbalanced tree generated using the same branching times as the birth-death tree (bottom). Shown are trees with 20 tips for illustrative purposes, but results in **B** are from trees with 100 tips. **B)** Difference in summary statistic with the true birth-death tree for phylogenetic trees inferred from alignments generated using the node substitution model on either balanced, unbalanced or random trees. We explore τ as a fraction of crown age (horizontal axis), and the impact of extinction (*d*, columns). The dotted line indicates zero difference with the true tree. The summary statistics are the beta and gamma statistic, mean branch length, nLTT statistic and tree height.

Tree balance clearly influences the sensitivity of inference to node substitutions (Figure 2). The inference error is larger for unbalanced trees, again only for the gamma and the nLTT statistic. Fully balanced trees show slightly less error than birth-death trees. Overall, all three types of trees show an increased error when alignments are generated using the node substitution model. Errors are particularly large for the beta statistic, but that is expected because it measures topological features of the tree that we modified artificially.

Support for strict and relaxed clock models

We compare the relative support for each model, reflected by the relative weight of the marginal likelihood. With an increasing amount of time spent at the nodes τ, the median weight of the relaxed clock model increases for the node substitution alignment, with generally (across extinction rates) a higher weight than the strict clock model for values of τ that are equal or larger than 0.1 times the crown age (Figure 3). For the *twin* alignment, the strict clock model is preferred, as expected, because this is the generating model.

For low values of τ (smaller than 0.1 times the crown age), we do not find any effect of the balance of the tree on the marginal likelihood of the relaxed clock model (Figure 4), in line with our finding above. However, for intermediate values of τ (0.1 and 0.2), we find that unbalanced trees tend to have a higher marginal weight for the relaxed clock model. For high values of τ (0.4), we find that the marginal weight for the relaxed clock model is always higher, regardless of the balance of the tree.

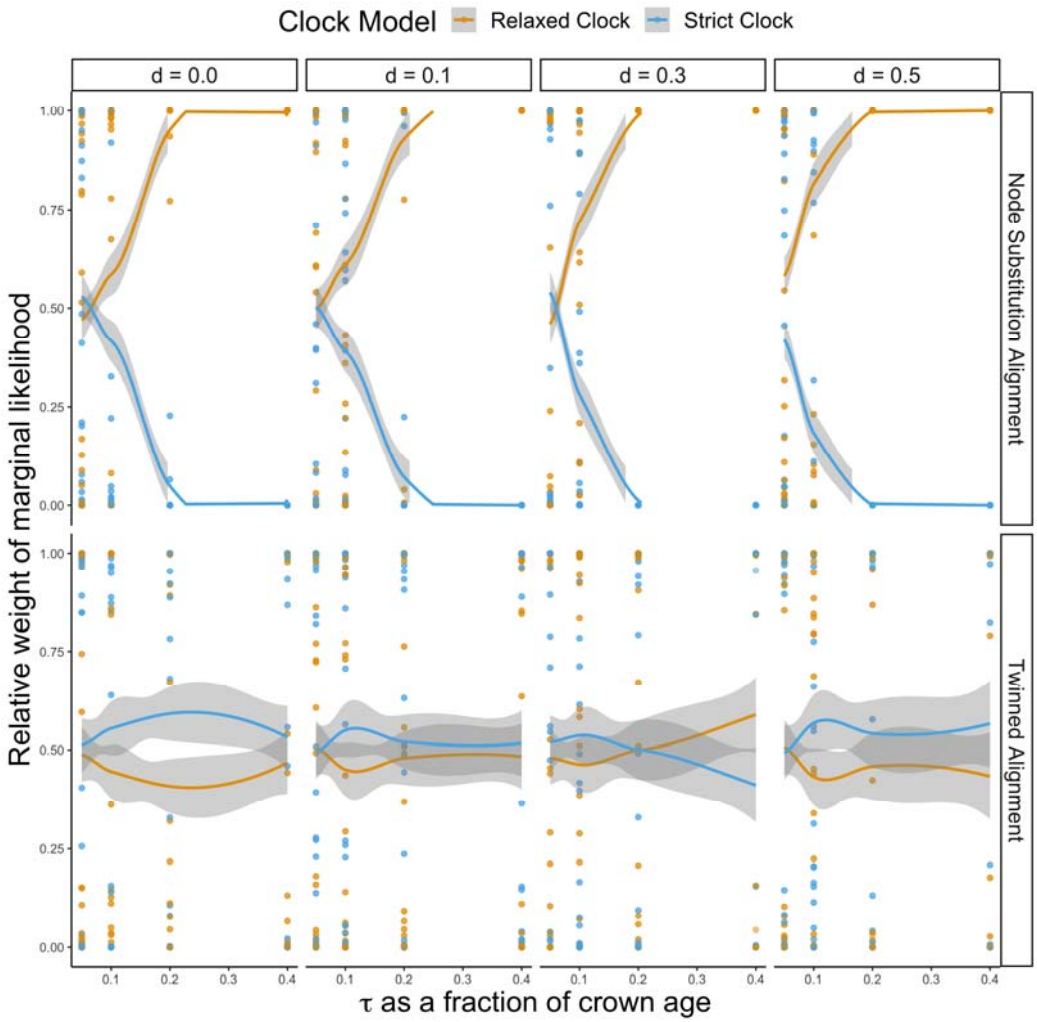

Figure 3. Marginal likelihood weight of the relaxed and strict clock model for varying time spent on the nodes (τ), where τ is chosen as fraction of the crown age. Alignments generated with a node substitution model (top row) are compared with alignments generated without node substitutions (bottom row). Per parameter combination, 100 replicate trees were analyzed. Because many dots are plotted on top of each other, we use solid lines to indicate the best fitting locally estimated scatterplot smoothing (LOESS), and the 95% Confidence interval (grey shaded area) of the LOESS curve. As the time spent on the nodes increases, posterior support for the relaxed clock model increases, but only if the alignment was generated with a node substitution model.

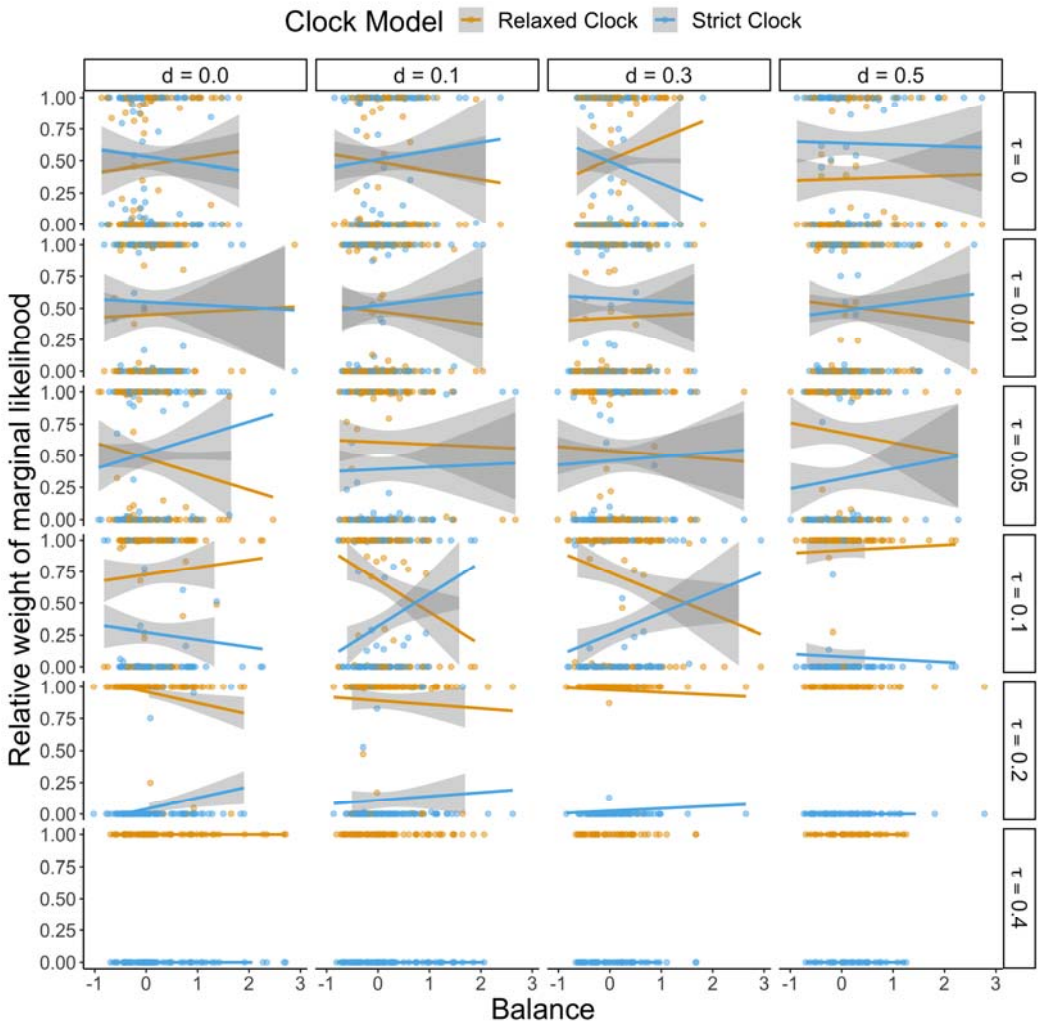

Figure 4. Marginal likelihood weight of the relaxed clock model for trees of varying balance, split out across different extinction rates (d = [0, 0.1, 0.3, 0.5]) and time spent on the nodes (τ), where τ is chosen as fraction of the crown age (e.g. $\tau = 0.1$ reflects a node time of 10% of the crown age). Per parameter combination, 100 replicate trees were analyzed. Because many dots are plotted on top of each other, we use solid lines to indicate the best fitting linear regression and the 95% confidence interval (grey shaded area) of regression.

312 Box 1

313 The general framework we have described here, allows for the exploration
314 of the impact of accumulation of substitutions during speciation events, which
315 we can vary in relation to the accumulation of substitutions over time between
316 speciation events. As the accumulation of substitutions during speciation events
317 becomes more prominent, the impact on the resulting tree becomes more and
318 more pronounced, in particular the impact on branching time estimates.

319 The next question is what proportion of substitutions in real data is the
320 result of processes acting during speciation, and what proportion arises
321 gradually over time between speciation events. Future work should employ a
322 likelihood framework able to infer the expected number of substitutions
323 accumulated during speciation events. The joint likelihood of the phylogenetic
324 tree and the sequence alignment can be computed using an SSE framework; one
325 needs to repeatedly apply equations (A1), (A2) and (A3) from Goldberg & Igic
326 (2012) to each nucleotide from tip to root of the phylogeny, where we assume
327 that the speciation and extinction rates do not depend on the nucleotide. The
328 equations then become (with $D_{\mathrm{N}i}$ the probability of the nucleotides at the focal
329 locus in subclade N being in their observed states when the nucleotide at time t
330 is equal to i, and E_i the probability of extinction of the clade when the nucleotide
331 at time t is i, and $D_{\mathrm{C}i}$ the probability just before an observed node in the
332 phylogeny):

333

$$\frac{dD_{Ni}}{dt} = -\left(\lambda + \sum_j r_{ij} + \mu\right) D_{Ni}$$

$$+ \sum_j r_{ij} D_{Nj} + \lambda \sum_{j,k} \left(P_{c,ij} D_{Nj} P_{c,ik} E_k + P_{c,ij} E_j P_{c,ik} D_{Nk}\right)$$

$$\frac{dE_i}{dt} = -\left(\lambda + \sum_j r_{ij} + \mu\right) E_i + \sum_j r_{ij} E_j + \mu + \lambda \sum_{j,k} \left(P_{c,ij} E_j P_{c,ik} E_k\right)$$

$$D_{Ci} = \lambda \sum_{j,k} \left(P_{c,ij} D_{Nj} P_{c,ik} D_{Mk} + P_{c,ij} D_{Mj} P_{c,ik} D_{Nk}\right)$$

This likelihood computation can be implemented in software able to accommodate joint estimation, such as RevBayes (Hohna et al. 2016).

339 DISCUSSION

340 We have shown that an increased substitution rate during speciation
341 events potentially provides a mechanistic explanation of variation in substitution
342 rates across the branches of phylogenetic trees. We have shown that trees
343 inferred from alignments generated with this substitution model differ
344 substantially from trees inferred from alignments generated with a standard
345 substitution model, especially concerning branching times. Furthermore, we find
346 that this new substitution model can potentially explain widespread support for
347 relaxed molecular clocks.

348 One might expect that a high extinction rate, by elevating numbers of
349 hidden nodes, leads to a greater impact of node substitutions. It may therefore be
350 counterintuitive that in our simulation study we did not find such an effect of
351 higher rates of extinction. However, we conditioned our alignments on having
352 the same total number of substitutions. Thus, with higher extinction and hence
353 more hidden nodes, fewer substitutions occur on the observed nodes, while the
354 number of hidden nodes is proportional to branch length (Eq.2) and can hence
355 be interpreted as substitutions on the branches. The result is that with higher
356 extinction, the signature of node substitutions decreases.

357 The model we propose here takes the simplest form, assuming a JC69
358 substitution matrix, identical substitution rates, identical substitution matrices
359 between nodes and branches, and constant birth-death rates over time. These
360 assumptions were made as a most basic starting point, but can be relaxed in
361 future analyses, for instance by introducing a different substitution matrix at the
362 nodes, or by studying the effect of node substitutions on trees that are generated
363 by density dependent speciation rates (Etienne et al. 2012). By starting with the

364 most tractable version of the node substitution model, we have provided a first
365 proof of concept of the potential impact of node substitutions without
366 overcomplicating matters.

367 Previous methods have applied rather ad-hoc corrections to account for
368 differences in substitution rates across different branches in the same
369 phylogeny, typically referred to as the 'relaxed clock' approach. These methods
370 provide satisfying statistical solutions to account for variation in substitution
371 rates, but refrain from providing biological explanations for this observed
372 phenomenon. The node substitution model we introduce here provides this
373 explanation: branches that have accumulated a number of 'hidden' branching
374 events, e.g. speciation events of species that have subsequently gone extinct,
375 have a higher number of accumulated substitution events during these 'hidden'
376 speciation events. When we compared the marginal likelihood of the relaxed
377 clock model versus the strict clock model for alignments generated with the
378 node substitution model, we found that marginal likelihoods for the relaxed
379 clock model are much higher. This indicates that our proposed process of
380 accumulating substitutions during speciation events can generate patterns in the
381 alignment that are picked up by phylogenetic methods as evidence for a relaxed
382 clock model, without actually using a relaxed clock model. This does not
383 necessarily mean that our model provides a perfect explanation, but we do want
384 to emphasize that our model provides a first potential mechanistic explanation
385 for the observed patterns.

386 Our model reflects many of the aspects of the substitution spikes
387 introduced by Manceau *et al.* (Manceau et al. 2020). In that model, Manceau *et al.*
388 introduce an alternative solution for punctuated equilibria-like patterns in

389 molecular evolution through the implementation of spikes of substitution, e.g.
390 moments in time at which there is an increased substitution rate. They let these
391 moments occur at speciation events, and also model the probability of such an
392 event happening at a speciation event (rather than assuming they always do, like
393 we did here). Furthermore, they provide a first foray into formulating a
394 likelihood for a node-associated substitution model, but only provide a
395 likelihood of the alignment, given the tree. Whereas their work modulates the
396 impact of substitutions at the nodes through modifying the substitution rate, our
397 work modulates this impact through modifying the time spent at the node. We
398 think these two models are highly complementary, and we embrace future
399 endeavors towards combining insights from both of them.

400 The notion of accelerated evolution during speciation events ties in
401 closely with the theory of punctuated equilibrium; where Eldridge and Gould
402 proposed that perhaps evolution is not a gradual process, but rather a process
403 with distinct bursts of phenotypic and morphological change. (Eldredge and
404 Gould 1972). Our framework provides a first step towards being able to test this
405 notion, where information on the estimated fraction of substitutions
406 accumulated at the nodes can directly inform us about whether the majority of
407 mutations is accumulated during periods of stasis (e.g. at the branches), or
408 during speciation (e.g. at the nodes).

409 With our introduction of the node substitution model, we hope to stimulate
410 discussion on the biological explanation of variation in substitution rates within
411 and across phylogenies. Furthermore, we hope to have set a first step in
412 improving our understanding of this variation, and improving phylogenetic
413 inference as a whole.

REFERENCES

Aldous D. 2001. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. Stat. Sci. 16:23–34.

Bilderbeek R.J.C., Etienne R.S. 2018. babette: BEAUti 2 , BEAST2 and Tracer for R. 2018:2034–2040.

Bilderbeek R.J.C., Laudanno G., Etienne R.S. 2020. Quantifying the importance of an inference model in Bayesian phylogenetics. bioRxiv.

Bokma F., van den Brink V., Stadler T. 2012. Unexpectedly many extinct hominins. Evolution. 66:2969–2974.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchene S., Fourment M., Gavryuskina A., Heled J., Jones G., Kuhnert D., De Maio N., Matschiner M., K. Mendes F., Muller N.F., Ogilvie H.A., du Plessis L., Popinga A., Rambaut A., Rasmussen D., Siveroni I., Suchard M.A., Wu C.H., Xie D., Zhang C., Stadler T., Drummond A.J. 2019. BEAST 2 .5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15:e1006650.

Bromham L. 2011. The genome as a life-history character: Why rate of molecular evolution varies between mammal species. Philos. Trans. R. Soc. B Biol. Sci. 366:2503–2513.

Bromham L., Hua X., Lanfear R., Cowman P.F. 2015. Exploring the Relationships between Mutation Rates , Life History , Genome Size , Environment , and Species Richness in Flowering Plants. 185:507–524.

Clarke K., Warwick R. 2001. A further biodiversity index applicable to species lists: variation in taxonomic distinctness. Mar. Ecol. Prog. Ser. 216:265–278.

Dowle E.J., Morgan-Richards M., Trewick S.A. 2013. Molecular evolution and the

439 latitudinal biodiversity gradient. Heredity (Edinb). 110:501–510.
440 Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics
441 and dating with confidence. PLoS Biol. 4:e88.
442 Duchene D., Bromham L. 2013. Rates of molecular evolution and diversification
443 in plants: Chloroplast substitution rates correlate with species-richness in
444 the Proteaceae. BMC Evol. Biol. 13:1.
445 Eldredge N., Gould S.J. 1972. Punctuated equilibria: an alternative to phyletic
446 gradualism. Models In Paleobiology. Freeman Cooper and Co. p. 82–115.
447 Eo S.H., DeWoody J.A. 2010. Evolutionary rates of mitochondrial genomes
448 correspond to diversification rates and to contemporary species richness in
449 birds and reptiles. Proc. R. Soc. B Biol. Sci. 277:3587–3592.
450 Etienne R.S., Haegeman B., Stadler T., Aze T., Pearson P.N., Purvis A., Phillimore
451 A.B. 2012. Diversity-dependence brings molecular phylogenies closer to
452 agreement with the fossil record. Proc. R. Soc. B Biol. Sci. 279:1300–1309.
453 Faith D.P. 1992. Conservation evaluation and phylogenetic diversity. Biol.
454 Conserv. 61:1–10.
455 Goldie X., Lanfear R., Bromham L. 2011. Diversification and the rate of molecular
456 evolution: No evidence of a link in mammals. BMC Evol. Biol. 11:286.
457 Hohna S., Landis M.J., Heath T.A., Boussau B., Lartillot N., Moore B.R.,
458 Huelsenbeck J.P., Ronquist F. 2016. RevBayes: Bayesian phylogenetic
459 inference using graphical models and an interactive model-specification
460 language. Syst. Biol. 65:726–736.
461 Jansson R., Davies T.J. 2008. Global variation in diversification rates of flowering
462 plants: Energy vs. climate change. Ecol. Lett. 11:173–183.
463 Janzen T., Höhna S., Etienne R.S.R.S. 2015. Approximate Bayesian Computation of

464 diversification rates from molecular phylogenies: introducing a new
465 efficient summary statistic, the nLTT. Methods Ecol. Evol. 6:566–575.
466 King M.C., Wilson A.C. 1975. Evolution at two levels in humans and chimpanzees.
467 Science. 188:107–116.
468 Lanfear R., Ho S.Y.W., Love D., Bromham L. 2010. Mutation rate is linked to
469 diversification in birds. Proc. Natl. Acad. Sci. U. S. A. 107:20423–20428.
470 Lepage T., Bryant D., Philippe H., Lartillot N. 2007. A general comparison of
471 relaxed molecular clock models. Mol. Biol. Evol. 24:2669–80.
472 Lewitus E., Morlon H. 2016. Characterizing and comparing phylogenies from
473 their laplacian spectrum. Syst. Biol. 65:495–507.
474 Manceau M., Marin J., Morlon H., Lambert A. 2020. Model-based inference of
475 punctuated molecular evolution. Mol. Biol. Evol.:msaa144.
476 Nabholz B., Glémin S., Galtier N. 2008. Strong variations of mitochondrial
477 mutation rate across mammals - The longevity hypothesis. Mol. Biol. Evol.
478 25:120–130.
479 Pagel M., Venditti C., Meade A. 2006. Large punctuational contribution of
480 speciation to evolutionary divergence at the molecular level. Science.
481 314:119–121.
482 Pybus O., Harvey P. 2000. Testing macro–evolutionary models using incomplete
483 molecular phylogenies. Proc. R. Soc. B Biol. Sci. 267:2267–72.
484 Rabosky D.L., Donnellan S.C., Talaba A.L., Lovette I.J. 2007. Exceptional among-
485 lineage variation in diversification rates during the radiation of Australia's
486 most diverse vertebrate clade. Proc. R. Soc. B Biol. Sci. 274:2915–2923.
487 Ricklefs R.E. 2006. Global variation in the diversification rate of passerine birds.
488 Ecology. 87:2468–2478.

Russel P.M., Brewer B.J., Klaere S., Bouckaert R.R. 2019. Model Selection and Parameter Inference in Phylogenetics Using Nested Sampling. Syst. Biol. 68:219–233.

Schliep K.P. 2011. phangorn: phylogenetic analysis in R. 27:592–593.

Simpson G.G. 1945. SECTION OF BIOLOGY: Tempo and Mode in Evolution. Trans. N. Y. Acad. Sci. 8:45–60.

Sipos B., Massingham T., Jordan G.E., Goldman N. 2011. PhyloSim - Monte Carlo simulation of sequence evolution in the R statistical computing environment. .

Spielman S.J., Wilke C.O. 2015. Pyvolve: A Flexible Python Module for Simulating Sequences along Phylogenies. :1–7.

Sung W., Ackerman M.S., Dillon M.M., Platt T.G., Fuqua C., Cooper V.S., Lynch M. 2016. Evolution of the insertion-deletion mutation rate across the tree of life. G3 Genes, Genomes, Genet. 6:2583–2591.

van Valen L.M. 1985. Why and how do mammals evolve unusually rapidly. Evol. Theory. 7:127–132.

507

508