bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

RelTime relaxes the strict molecular clock throughout the phylogeny

Fabia U. Battistuzzi[1,2], Qiqing Tao[3,4], Lance Jones[1], Koichiro Tamura[5,6], and Sudhir Kumar[3,4]

[1]Department of Biological Sciences, Oakland University, Rochester, Michigan

[2]Center for Data Science and Big Data Analytics, Oakland University, Rochester, Michigan

[3]Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, Pennsylvania

[4]Department of Biology, Temple University, Philadelphia, Pennsylvania

[5]Department of Biological Sciences, Tokyo Metropolitan University, Hachioji, Tokyo, Japan

[6]Research Center for Genomics and Bioinformatics, Tokyo Metropolitan University, Hachioji, Tokyo, Japan

Correspondence to:
Sudhir Kumar
Temple University
Philadelphia, PA 19122, USA
E-mail: s.kumar@temple.edu

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Abstract

The RelTime method estimates divergence times when evolutionary rates vary among lineages. It has been shown to perform well in the analysis of empirical and computer simulated datasets. In addition, theoretical analyses show that RelTime relaxes the strict molecular clock throughout a molecular phylogeny. However Lozano-Fernandez *et al.*'s (2017) suggested that RelTime imposes a strict molecular clock for deep divergences in an analysis of one dataset, whereas Bayesian methods do not. We show that their observation is caused primarily by the lack of equivalence between analysis conditions in the application of RelTime and Bayesian methods. We also found comparable differences even among Bayesian estimates when different root priors were used. Overall, less restrictive prior selections produced Bayesian results concordant with RelTime. Therefore, our results suggest that Bayesian estimates obtained under a range of priors need to be considered and presented in any discourse about molecular dating and comparisons with RelTime.

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Introduction

The RelTime method was developed to estimate timetrees from molecular sequence data when evolutionary rates vary among lineages (Tamura et al. 2017). It has been shown to be accurate in the analysis of computer simulated data that were generated with extensive rate heterogeneity throughout the tree (Tamura et al. 2012; Filipski et al. 2014; Tamura et al. 2017). In analyses of many large empirical datasets, RelTime estimated divergence times similar to those reported from Bayesian methods, as long as both methods were used under the same conditions (Mello et al. 2017). In addition, theoretical analyses clearly establish that a relative rate framework, which does not assume a strict molecular clock, forms the mathematical foundation of the RelTime method (Tamura et al. 2017).

Surprisingly, Lozano-Fernandez *et al.* (2017) recently presented some contradictory patterns: first, they suggested that RelTime functionally maintained a molecular clock for deep divergences in the analysis of a dataset containing 117 species and 2049 amino acids (Erwin et al. 2011); second, they surmised that this pattern is the cause of the curvilinear relationship between Bayesian and RelTime node age estimates observed by Battistuzzi *et al.* (2015); and, third, they reported that the standard errors (SE) of RelTime linearly increase in size with deeper node ages, unlike the Bayesian approaches. We investigated primary reasons for these observations, because we have previously found RelTime and Bayesian methods to generally produce similar results (Mello et al. 2017; Tamura et al. 2017). Our hypothesis was that the observed differences were caused by lack of full equivalence between Bayesian and RelTime analysis conditions. To evaluate this hypothesis, we re-analyzed the same dataset and estimated node ages and evolutionary rates by using RelTime in MEGA7 (Kumar et al. 2016) and Bayesian methods in Phylobayes (Lartillot et al. 2009). Phylobayes was selected because this Bayesian dating software was used in all the prior studies discussed here (Lozano-Fernandez et al. 2017; Battistuzzi et al. 2015; Erwin et al. 2011).

RelTime relaxes strict molecular clock in shallow as well as deep nodes

Lozano-Fernandez *et al.* (2017) found that RelTime does not relax the strict molecular clock in deep nodes because the (relative) rates reported by RelTime were close to 1 for many deep clades (note that RelTime rates are all relative to the rate of the ingroup root node, which is assigned a value of 1 for reference). The reason for the similarity of relative rates in these deep nodes is the use of the rate merging option by Battistuzzi *et al.* (2015) for RelTime analysis,

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

where rates between sister lineages are merged when they were not statistically different (Tamura et al. 2012). For the current dataset, the finding of similar relative rates for many deep nodes is an indication that the rates for many lineages are not significantly different. This is confirmed in RelTime analyses where node ages estimated with and without the rate merging option produced very similar divergence times (slope = 0.98; **Figure 1a**). The advantage offered by rate merging is that fewest rate parameters need to be estimated, which is a preferred statistical practice when final results are unaffected. In retrospect, we should have discussed the similarity of time estimates under these options in Battistuzzi *et al*. (2015), which would have prevented a misunderstanding that led to the primary conclusions of Lozano-Fernandez *et al.* (2017). We also wished that Lozano-Fernandez *et al.* (2017) had used RelTime to ensure the validity of their conclusions, as they would have quickly found that the distribution of lineage rates without using the rate merging option produces rates both higher and lower than the overall average rate throughout the tree (**Figure 1b**, black circles) in a pattern not too different from Bayesian analyses (**Figure 1b**, red circles). Therefore, RelTime relaxes the strict molecular clock in deep as well as shallow lineages. Because RelTime node ages with and without rate merging are very similar (**Figure 1a**), the relationship observed by Battistuzzi *et al.* (2015) between RelTime and Bayesian node ages cannot be caused by the rate relaxation issues, as stated by Lozano-Fernandez *et al.*, but rather by other factors such as calibration boundaries or other priors (e.g., speciation models). We explore these factors below.

Bayesian estimates with minimal sets of priors are inconsistent with each other

Because RelTime produces relative times and does not require the priors used in Bayesian methods, Bayesian analyses in Phylobayes that require the least number of parameters are needed to make proper comparisons. Such comparisons are actually useful because they allow us to determine the degree of similarity of results from two methods before many calibrations and uncertainty distributions are applied. This is particularly instructive because calibration boundaries are often not known with high degree of certainty and the underlying specification process can be difficult to know. While RelTime automatically produces relative times using a phylogeny and the corresponding sequence alignment, Phylobayes requires a minimal set of prior specifications: one for the root-age prior and the other for the pattern of speciation. The strategy chosen by Lozano-Fernandez *et al.* (2017) was to assign a root age prior of 1 as well as the default uniform distribution for modelling speciation. This is, however, only one of many choices. Phylobayes actually allows the root age prior to be omitted, where it will automatically

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

set it to 1000. This setting can be coupled with different priors for the patterns of speciation, such as the birth-death model. Because RelTime does not require any such priors, any difference observed between the results produced by two methods may simply be a function of the priors selected.

So, we explored the robustness of the Bayesian estimates to changes in priors and compared Bayesian estimates under different priors with RelTime estimates. We compared Phylobayes estimates from Lozano-Fernandez and colleagues (root age prior = 1 and uniform distribution) to new estimates where no root age prior was used (allowing it to default to 1000) and a birth-death speciation model was selected with default parameters (see **Methods** for details). This produced relative node ages that showed a curvilinear relationship between two sets of Bayesian estimates (**Figure 1c**). Therefore, Phylobayes estimates depend on priors selected, which is not unexpected (e.g., Warnock et al. 2017). Interestingly, Bayesian estimates using no root calibration and birth-death default prior showed a nearly linear relationship (slope = 0.97) with those from RelTime reported in Battistuzzi *et al.* (2015; **Figure 1d**). Therefore, Lozano-Fernandez *et al.*'s would have reached an opposite conclusion (i.e., full agreement between Bayesian and RelTime) if they had chosen a different set of priors. This highlights an enduring challenge faced by biologists when applying Bayesian methods, they need to select priors whose choice is rarely easy and comes down to personal experience and knowledge. We are of the opinion that Bayesian estimates from different combination of priors should be presented when objective choice of priors is not possible.

Deepest nodes are the most affected by prior choices

During our investigation above, we found that different reference nodes selected to scale times to generate relative ages after Bayesian analysis, for comparative purposes, also contributed to the strong curvilinear relationship reported by Lozano-Fernandez *et al.* The curvilinear trend obtained by using the age of the choanoflagellate Monosiga to normalize all ages for comparison (**Figure 1e**, red line) is very similar to that reported by Lozano-Fernandez *et al.* (2017) using the root node, but this trend becomes much less pronounced when Bayesian ages are normalized to the age of Metazoa, while keeping their root calibration to be 1.0 (**Figure 1e**, blue line). It becomes even more muted when the root calibration is not specified (**Figure 1e**, pink and green lines) and a birth-death model prior is used. These results show that the root and Monosiga nodes are the most affected by changes in priors most likely because they are the deepest nodes in the phylogeny and closest to the root.

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Bayesian and RelTime estimates of standard errors show similar trends

We also investigated if the priors used in Lozano-Fernandez *et al.* (2017) explain the difference reported between RelTime and Phylobayes estimates of standard errors (SEs) of node ages. While credibility/confidence intervals would usually be compared, here we present SEs to ensure that our results are directly comparable to those presented by Lozano-Fernandez *et al.* (2017). In order to generate results under comparable conditions, we fixed the ingroup root calibration to be 1 in RelTime. As expected, the relationship of node age and SEs for RelTime showed a trend in which SEs first rise and then decrease with increasing node age (**Figure 1f**, black circles). This pattern is similar to that produced in the Bayesian analysis (see Figure 1f in Lozano-Fernandez *et al.* (2017)). This means that the imposition of a root node calibration of 1.0 strongly affected SE estimates, because this calibration clips all confidence intervals in order to avoid predating the calibration constraint. Upon omitting the root node calibration, we found that the Bayesian SE estimates increased with time (**Figure 1f**, red circles), which is a trend similar to that observed for RelTime without any constraints. Therefore, the pattern reported in Lozano-Fernandez *et al.* (2017) was caused by the comparison of results obtained under different conditions.

Priors and calibrations that impact absolute molecular dates

Battistuzzi *et al.* (2015) used data from Erwin *et al.* (2011) as an example because its analysis clearly showed that (a) the two maxima and the root prior have a very large impact on molecular time estimates obtained via Bayesian methods and that (b) different combinations of (maximum) calibrations and priors produce very different time estimates. Results similar to Battistuzzi *et al.* (2015) were also obtained by Lozano-Fernandez *et al.* (2017) (see their Fig. 3), but were discounted because they judged the induced effective prior to be overly diffused. These results actually highlight a well-known attribute of Bayesian analyses: the user-specified parameters and priors can be very different from the induced prior distributions because of complicated interdependencies of the parameters and relatively arbitrary approaches to truncation of the prior distributions (Warnock et al. 2015; Barba-Montoya et al. 2017; Eme et al. 2014). In the dataset analyzed here, it is clear that the Bayesian relative and absolute time estimates are strongly affected by the selection of priors. In the absolute dating analysis, the root prior used by both Erwin *et al.* (2015) and Lozano-Fernandez *et al.* (2017) is stricter than that explored in Battistuzzi *et al* (2015). While the best prior cannot be unequivocally identified, a recent study on molecular timing of eukaryotes adds some new information. This study obtained a Bayesian

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

divergence time estimate of ~1375 million years ago for Opisthokonta (Animals+Fungi) in an analysis of a dataset with 116 taxa and 2166 amino acids (Gold et al. 2017). Using a dataset of comparable size, this study produced a divergence time estimate at the upper end of the marginal prior distribution used by Erwin *et al.* (2011) and Lozano-Fernandez *et al.* (2017), but well within the distribution of Battistuzzi *et al.* (2015). Use of a wider prior distribution, thus, improves Bayesian analyses, as the improvement of time estimates associated with greater prior accuracy justifies a sacrifice in precision (i.e. using a prior distribution with higher uncertainty).

DISCUSSION

We have presented a number of results above that showed that the conclusions drawn by Lozano-Fernandez *et al.* (2017) about the relationship of RelTime rates and node ages with each other and with Bayesian methods are the result of comparing rates and dates obtained under analysis settings that were not equivalent and that produced inconsistent results among Bayesian analyses. This is because Bayesian estimates depend strongly on the calibrations and priors used and their interactions (dos Reis et al. 2016; Inoue et al. 2010; Parham et al. 2012; Warnock et al. 2012, 2015, 2017; Hug & Roger 2007; Barba-Montoya et al. 2017; Eme et al. 2014). Battistuzzi *et al.* (2015) provided a protocol using RelTime to identify calibrations and priors that have a large effect on inferred rates and dates because of their inconsistency with the divergence time signal coming from molecular data.

In addition to calibration priors, it is also a challenge to select other priors (e.g., speciation models) in Bayesian analysis. Based on analyses presented here, we suggest that the application of RelTime approach could help to inform about priors and distributions that are not likely to exert undue impact on molecular divergence times inferred using Bayesian methods, because, under some conditions, RelTime and Bayesian methods produced similar results for the data analyzed in Lozano-Fernandez *et al.* (2017) and because there is often no way to objectively select correct priors for Bayesian analyses. Of course, because Bayesian methods can become very time consuming for bigger datasets, and RelTime speed is orders of magnitude faster (Tamura et al. 2012, 2017), one may simply choose RelTime as a practical and theoretically sound alternative to Bayesian methods (Tamura et al. 2017).

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

MATERIAL AND METHODS

The dataset consists of 117 species and 2049 aligned amino acids (Erwin et al. 2011). All analyses were conducted with RelTime (Tamura et al. 2012) and Phylobayes v. 4.1f (Lartillot et al. 2009). Phylobayes was selected because it was the Bayesian dating software used in all the previous studies discussed here (Erwin et al. 2011; Battistuzzi et al. 2015; Lozano-Fernandez et al. 2017). For the RelTime analyses, no calibration times were used. All analyses were conducted with MEGA 7, with the exception of analyses employing the rate merging option. These analyses were conducted with MEGA 6 (Tamura et al. 2013), the software used by Battistuzzi *et al.* (2015). Note that RelTime produces relative lineage rates. Phylobayes produces branch-specific rates, which were obtained with two sets of parameters: one without any root calibration and another one calibrating the root node to 1 (as done in Lozano-Fernandez *et al.* (2017)). Phylobayes automatically scales node ages to 1000 when no root calibration is specified. Therefore, we divided estimated node ages by 1000 to make direct comparisons between RelTime and Phylobayes estimates. All rates were normalized to their mean rate estimate over the whole tree (Figure 1b). We also tested the effects of using uniform *vs.* birth-death prior distributions on divergence times and narrower (10^{-6}) or more diffuse (10^{-3}) values for the birth-death hyperprior. Along with the prior distributions, hyperprior distributions are also not known in advance and have strong effects on posterior estimates. The final options used in this study were –cat –gtr –cir –bd with default hyperprior (10^{-3}) and –cat –gtr –cir –bd with hyperprior = 10^{-6}, which produces results equivalent to those from a uniform distribution. All analyses in Phylobayes were run for at least 20,000 generations. While full convergence is expected to take many months, time estimates from these truncated analyses appear reliable, because relative time estimates remained stable at 2500, 6000, 8500, 12500, and 20000 generations. This convergence approach produced times identical to those obtained by Lozano-Fernandez *et al.* under the same analysis conditions and validated using Tracecomp (Lozano-Fernandez et al. 2017).

ACKNOWLEDGEMENTS

We thank Heather Rowe and Beatriz Mello for comments and Davide Pisani for sharing the results from Lozano-Fernandez *et al.* (2017). We thank Nicolas Lartillot for guidance about the best ways to evaluate parameters in Phylobayes. We also wish to thank Davide Pisani for useful discussions and critical comments on the previous version of this manuscript, which does not imply his agreement with our conclusions. This work was supported by grants from

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

National Aeronautics and Space Administration [NNX16AJ30G to F.U.B.], National Science Foundation [DBI 1356548 to S.K.], and Tokyo Metropolitan University [DB105 to KT].

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

References

Barba-Montoya J, dos Reis M, Yang Z. 2017. Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. Mol. Phylogenet. Evol. 114:386–400. doi: 10.1016/j.ympev.2017.07.005.

Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S. 2015. A Protocol for Diagnosing the Effect of Calibration Priors on Posterior Time Estimates: A Case Study for the Cambrian Explosion of Animal Phyla. Mol. Biol. Evol. 32:1907–1912. doi: 10.1093/molbev/msv075.

Eme L, Sharpe SC, Brown MW, Roger AJ. 2014. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb. Perspect. Biol. 6. doi: 10.1101/cshperspect.a016139.

Erwin DH et al. 2011. The Cambrian Conundrum: Early Divergence and Later Ecological Success in the Early History of Animals. Science. 334:1091–1097. doi: 10.1126/science.1206375.

Filipski A, Murillo O, Freydenzon A, Tamura K, Kumar S. 2014. Prospects for Building Large Timetrees Using Molecular Data with Incomplete Gene Coverage among Species. Mol. Biol. Evol. 31:2542–2550. doi: 10.1093/molbev/msu200.

Gold DA, Caron A, Fournier GP, Summons RE. 2017. Paleoproterozoic sterol biosynthesis and the rise of oxygen. Nature. 543:420–423. doi: 10.1038/nature21412.

Hug LA, Roger AJ. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. Mol. Biol. Evol. 24:1889–1897.

Inoue J, Donoghue PCJ, Yang Z. 2010. The impact of representation of fossil calibrations on Bayesian estimation of species divergence times. Syst. Biol. 59:74–89.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33:1870–1874.

Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics. 25:2286–2288. doi: 10.1093/bioinformatics/btp368.

Lozano-Fernandez J, dos Reis M, Donoghue PCJ, Pisani D. 2017. RelTime Rates Collapse to a Strict Clock When Estimating the Timeline of Animal Diversification. Genome Biol. Evol.

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

9:1320–1328. doi: 10.1093/gbe/evx079.

Mello B, Tao Q, Tamura K, Kumar S. 2017. Fast and Accurate Estimates of Divergence Times from Big Data. Mol. Biol. Evol. 34:45–50. doi: 10.1093/molbev/msw247.

Parham JF et al. 2012. Best Practices for Justifying Fossil Calibrations. Syst. Biol. 61:346–359. doi: 10.1093/sysbio/syr107.

dos Reis M, Donoghue PCJ, Yang Z. 2016. Bayesian molecular clock dating of species divergences in the genomics era. Nat. Rev. Genet. 17:71–80. doi: 10.1038/nrg.2015.8.

Tamura K et al. 2012. Estimating divergence times in large molecular phylogenies. Proc. Natl. Acad. Sci. U. S. A. 109:19333–19338. doi: 10.1073/pnas.1213199109.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30:2725–2729. doi: 10.1093/molbev/mst197.

Tamura K, Tao Q, Kumar S. 2017. Theoretical foundation of the relTime method for estimating divergence times from variable evolutionary rates. bioRxiv. doi: 10.1101/180182.

Warnock RCM, Parham JF, Joyce WG, Lyson TR, Donoghue PCJ. 2015. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. Proc. R. Soc. B Biol. Sci. 282:20141013. doi: 10.1098/rspb.2014.1013.

Warnock RCM, Yang Z, Donoghue P c. j. 2017. Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. Proc. R. Soc. B Biol. Sci. 284:20170227.

Warnock RCM, Yang Z, Donoghue PCJ. 2012. Exploring uncertainty in the calibration of the molecular clock. Biol. Lett. 8:156–159. doi: 10.1098/rsbl.2011.0710.

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

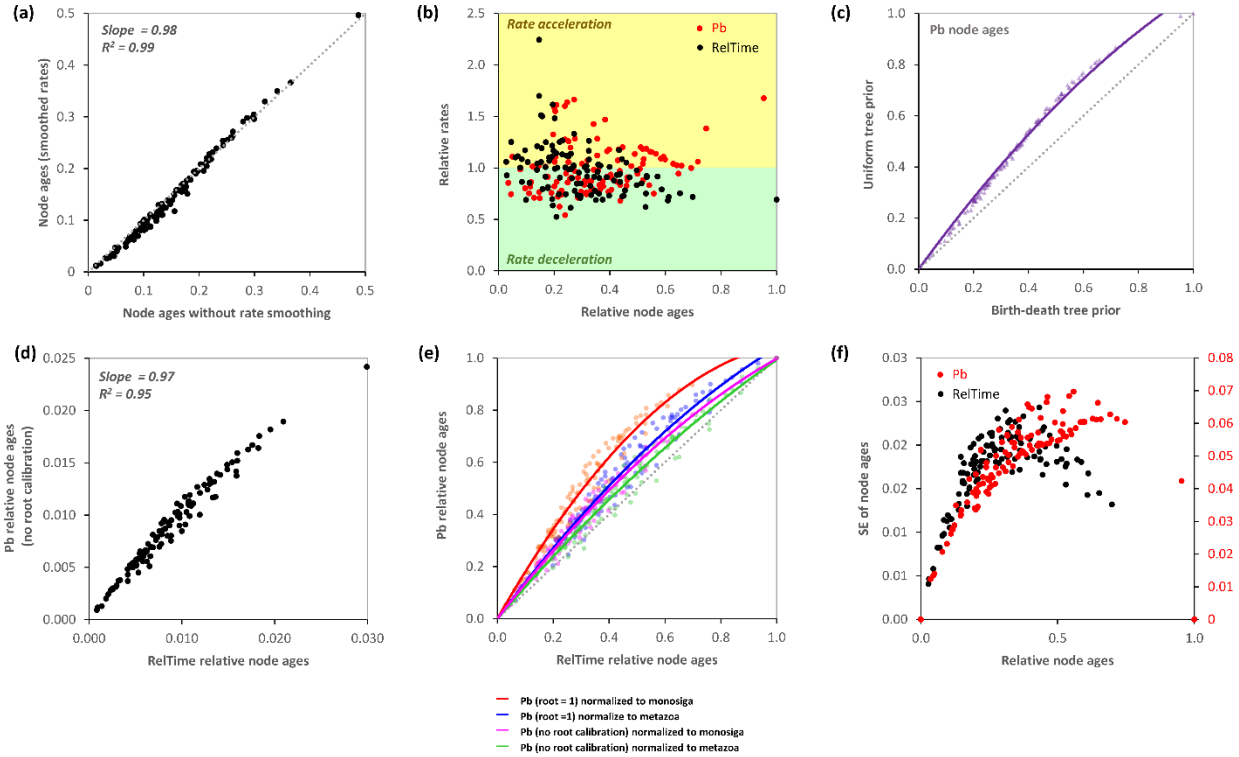

Figure 1 Comparisons of rates, dates, and standard errors from Bayesian and RelTme analyses. (**a**) RelTime estimates of node ages calculated with and without the rate merging option. The linear slope and R^2 value are shown. Dotted gray line represents 1:1 relationship. (**b**) Normalized relative rates for nodes at different time depths, with rates greater than the average, above 1.0, showing acceleration and those below 1.0 showing a slow-down (yellow and green backgrounds, respectively). Black circles: RelTime; Red circles: Phylobayes. In Phylobayes, rates were estimated without the root calibration and with a birth-death prior. (**c**) Relationship of Phylobayes node estimates under birth-death tree prior without root prior and uniform tree prior with root age constraint at 1. Solid line shows the polynomial fit and dotted gray line represents 1:1 relationship. (**d**) Relationship of RelTime and Phylobayes node ages obtained without root calibration and birth-death prior. All node ages were normalized to the sum of ingroup node ages. The linear slope and R^2 value are shown. (**e**) Relationship of RelTime estimates with Phylobayes with and without specified root calibration, and normalized to either Monosiga (Choanoflagellate) or Metazoa. Solid lines show polynomial fit for each comparison and dotted gray line represents 1:1 relationship. The R^2 values for the polynomial fit are all greater than 0.94. (**f**) Standard errors (SEs) of node ages produced by RelTime and Phylobayes under different calibration constraints. Black circles: RelTime estimates of SE of node ages when the ingroup root node is constrained at 1. Red circles: Phylobayes estimates of standard errors of node ages without the root

bioRxiv preprint first posted online Nov. 16, 2017; doi: http://dx.doi.org/10.1101/220194. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

calibration; Phylobayes estimates were divided by 1000 for direct comparisons because root calibration is automatically set to 1000 when no root calibration is specified.