Reviewer reports:  
  
Reviewer 1: This paper by Zhong and Drummond describe a new class of operators designed to improve the performance of BEAST2 inferences under a relaxed clock model.  
The performance of MCMC inferences is a sticking point in many empirical analyses, and the problem is only going to get worse with larger datasets and more complex models. Thus I'm happy to see more research in that area.   
The operators proposed by this manuscript are well described and the figures make the steps of the algorithm very clear.  
The Results section, on the other hand, is much less polished and the writing is confusing in places, as detailed in the comments below.   
Provided this section is improved, I recommend this paper for approval.  
  
General comments  
- p2, l7-8 It is not clear here if better estimates of divergence times have been obtained in the referenced papers or if this is simply an expectation.  
- p6, l44-45 On a dataset with sampling-through-time, the leaf node does not have to be the younger child. Does the operator still work in that case or is it a requirement that tO > tY ?  
- There is a mismatch between Figures 12 and 13 which show results for S3 and Table 5 which contains the standard deviation of the clock, which according to Figure 11 is S1. Similarly p9, l26 it is unclear which of S1 or S3 was actually measured.  
- p9 l33-34 I suggest adding some details on why configuration 1 was chosen as a comparison.  
- p10, l16 I don't understand what "the sampled trees were filtered by the shared common ancestor of each taxa" means. Later l22 it appears that the authors discarded the trees which did not match the reference topology, but it's unclear if that is the same thing as the previously mentioned filtering or an entirely different step of the process.  
- p10, l37-42 Is it really unexpected that the results obtained using the new operators would be consistent with the correlations enforced by these new operators ? Overall it's unclear what the conclusions from this section are.  
- p11, l8-23 It is very unclear what was actually done in this section. For instance, were the assigned divergence times and rates fixed in the analysis ? If so, how were they calculated, and if not, what are the values shown in figure 17 ? How were the summary trees obtained ?  
- Figures 12 and 13 are identical.  
- Figure 16b I suggest adding a legend showing the range of plotted values.  
- I would put a time axis on figure 17, as the credible intervals shown are not very meaningful without it.  
  
Minor comments  
- p1, l30 It is noticed that -> Note that  
- p2, l39 Prelimiaries -> Preliminaries  
- p4, l37 max(tj, tk) -> max(tL, tR)  
- p4, l53 the total genetic distance dL and dR -> the total genetic distance dL + dR  
- p6, l54 what happens if tO' < min(tG1, tG2) ?  
- p9, l26 MCMC simulation -> MCMC inference  
- p10, l48 becomes -> become  
- Figure 1 remove the box which only contains =  
- Figure 1 H3,R and H4,R should be H3,L and H4,L  
- Figure 11 substituion tree -> substitution tree  
- Figure 16 compare -> comparison  
  
  
  
Reviewer 2: My overall impression of this paper is positive. It addresses a large problem in the field (compute times) and makes a noteworthy improvement. I am satisfied with the work done and my concerns are mainly with the clarity of the manuscript and some important details that should be included.  
  
Summary  
This paper proposes a new MCMC operator (composed of multiple sub-operators) to address the issue of long compute times required for time-calibrated phylogenetics. An operator on internal node ages and branch rates is developed and three operators on root branch rates are also developed, and in all cases the Jacobian of the operators is derived. Simulations show that these operators work as expected, and performance on real and simulated data shows a substantial increase in computational efficiency when using the new operator.  
  
Paper Strengths  
-This paper includes a number of very helpful figures that make it much easier to understand the operators.  
-I appreciate the thoroughness with which the algorithms for the operator are laid out.  
-The operators themselves produce an impressive increase in MCMC performance. I expect that employing these operators will become standard practice for phylogenetic MCMC, in BEAST and beyond.  
  
Major Issues  
-More background on the operator is required.  
--It is necessary to explain why preserving genetic distances is the goal of the operator. That is, to point out that the transition probability matrix for a branch is exp(Qrt) so holding d = rt constant does not change the likelihood along that branch, and thus requires no re-computation of any partial likelihoods, speeding up MCMC.  
--Similarly, a brief introduction to the notion of an underlying unrooted phylogenetic tree would be useful for understanding the Pulley operators.  
  
-Small Pulley and Big Pulley can only be used on reversible CTMC models where unrooted trees can be used in inference. This is not a huge limitation in practice, but it should be mentioned.  
  
-The description of the asymmetric case in Big Pulley appears to assume that the younger child is a tip, but this is only a given if the tree has no heterochronous samples (which are increasingly common in real datasets). At a quick glance, it appears that the move could still work in this case but would require t\_Y < t\_O' < t\_X' and not just t\_O' < t\_X'.  
  
-Additional information is required about the simulation study.  
--What priors were used for inference? Especially important is the prior on the root age.  
--What other operators were used on the tree and the branch rates? These are the only operators that can change the underlying unrooted phylogeny, which makes them crucial to performance.  
  
-More information is needed when discussing the performance of the new operator.  
--What were p and q (from Figure 1), the proportion of root operations for Simple Distance and Small Pulley?  
--Without discussing operator weights, it is difficult to interpret the change in run time cost due to the Constant Distance operator. Discussing time required per operator may be clearer still, allowing comparison directly between node age proposals.  
  
-Figures 12 and 13 appear to be completely identical, it would appear that the 20-taxon figure was duplicated.  
  
Minor Issues  
-In the preliminaries, there are some issues with switching between parameterizations in terms of node times, t, and in terms of the tree, g.  
--The change from Pr(g) in equation 1 to Pr(t | Phi) in equation 2 is a bit jarring and equation 2 is less general. Pr(t | Phi) assumes independence between tree topology and divergence times, which is not always the case (for example the model of Barido-Sottani et al. (2018)).  
--Page 3, lines 52-54 refer to proposing a tree g', whereas page 3 line 38 states the operator works on times.  
--Readers will have an easier time if one parameterization is used consistently. I personally see no strong argument in favor of Pr(t | Phi), Pr(g | Phi) still allows the use of the vector of node times, t.  
  
-In Small Pulley there are some issues with clarity.  
--The statement "Small Pulley proposes a new genetic distance of a branch on one side of the root" is somewhat misleading, as it in fact proposes new distances on both sides of the root (by proposing a single number and using it to change both).  
--It would help to introduce D = d\_L + d\_R around page 4 line 53 and then state that d\_R will be adjusted simultaneously so as to preserve D.  
  
-In Big Pulley there are some issues with clarity.  
--Explaining Exchange() before the moves is important, but the sentence "Firstly, a method called Exchange is designed to propose a new tree topology" is confusing when in fact calling Exchange() is step 3.  
--The description of symmetric tree step 3 (page 6 lines 6-7) is confusing, as 50% of the time we will apply the method to L and either child of R.  
--In equation 8, presumably d\_1 is d\_H1, but this is not stated. Equation 10 uses d\_G1 instead of d\_1, which seems more clear.  
  
-In the section, "Correlation analysis of rates and node times," there are some issues.  
--A statement of motivation for this section is needed: what purpose does this experiment serve?  
--The statement, "With full length genomes now available, this limiting case might be approached in some data sets," ignores the complexities involved in inferring trees from genomes and requires assuming both a single topology across all loci in a genome (ignoring, for example, incomplete lineage sorting) and shared branch lengths at different loci (which need not be the case partitioning the dataset for analysis, see for example Lanfear et al. (2012)).  
--The current comparison scheme is difficult to interpret. The rate-to-rate and age-to-age correlations do not seem to be important, but take up more of the figure than the important comparisons. It would be simpler to directly compare branch lengths to the rates of those branches, perhaps by taking the Pearson correlation coefficient of length and rate across the posterior. Branches could be matched across trees much as they currently are. The results could be presented as a histogram or a heatmap as is currently done.  
  
-In the appendix there are some issues with clarity.  
--The relationship between son/dau and L/R is unclear. This makes understanding Algorithm 1 difficult.  
--The section on sampling from the prior needs an overview to explain, briefly, the motivation, design, and goals of the experiments.  
  
-The numbering on the figures and tables is perplexing. A number of tables and figures are only referenced from the appendix but have lower numbers than main-text figures and tables. This makes it seem as if one has accidentally skipped portions of the manuscript when reading through it.  
  
-The proposal to infer unrooted trees and then use those as data is interesting. Some discussion of related approaches (see below) is in order.  
--Thorne and Kishino (1998), Guindon (2010), and dos Reis and Yang (2011) perform a pre-MCMC step to approximate the likelihood surface of the underlying unrooted phylogeny, bypassing the need for the pruning algorithm but allowing for changes to the genetic distances.  
--Non-Bayesian methods such as TreeTime (Sagulenko et al. 2018), r8s (Sanderson 2003), and LSD (To et al. 2015) use an unroted phylogeny as data to estimate the time tree.  
  
Typos and Other Minor Comments  
-While the operators as discussed in this paper are, to my knowledge, novel, others have used operators similar to the proposal on internal node heights (e.g. <https://github.com/revbayes/revbayes/blob/master/src/core/moves/compound/RateAgeBetaShift.cpp>)  
-The proposed operator is discussed in the context of uncorrelated clock models, but it should also be applicable to autocorrelated models like that of Thorne and Kishino (1998).  
-The choice of kappa in the simulation study is somewhat strange, as usually the transition-transversion rate-ratio is expected to be above 1.  
-It is somewhat perplexing that fewer of the 120-taxon simulations had the mean rate in the 95% CI.  
-I wonder if there may be efficiency gains by employing proposals other than a uniform, such as a bactrian proposal (Yang and Rodriguez 2013)  
-Page 2 lines 7-8: The sentence "By allowing rates" is somewhat unclear as currently phrased.  
-Page 2 line 24, the statement "since each step in the chain requires a likelihood calculation" is somewhat misleading, with cached partial likelihoods many moves only require parts of the likelihood to be re-evaluated.  
-In "Simple Distance" (page 4 line 38), t\_i and t\_j should be t\_R and t\_L.  
-Page 5 line 20 should "rooted" be "unrooted"?  
-Page 10 line 15 states "After analyzing the ratite dataset," but this dataset has not been previously mentioned.  
-Page 10 line 18, taxa should be taxon  
-The axis label "number of runs" for Figures 12 and 13 might be more clear as something like "replicate" or "simulation number."  
-In figures 14 and 15, the same color scheme is used but the meanings of the colors are different. It would be easier to follow if different colors were used in these figures.  
  
Citations  
Barido-Sottani, Joëlle, Timothy G. Vaughan, and Tanja Stadler. "Detection of HIV transmission clusters from phylogenetic trees using a multi-state birth-death model." Journal of The Royal Society Interface 15.146 (2018): 20180512.  
Guindon, Stéphane. "Bayesian estimation of divergence times from large sequence alignments." Molecular Biology and Evolution 27.8 (2010): 1768-1781.  
Lanfear, Robert, et al. "PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses." Molecular biology and evolution 29.6 (2012): 1695-1701.  
dos Reis, Mario, and Ziheng Yang. "Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times." Molecular Biology and Evolution 28.7 (2011): 2161-2172.  
Sagulenko, Pavel, Vadim Puller, and Richard A. Neher. "TreeTime: Maximum-likelihood phylodynamic analysis." Virus evolution 4.1 (2018): vex042.  
Sanderson, Michael J. "r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock." Bioinformatics 19.2 (2003): 301-302.  
Thorne, Jeffrey L., Hirohisa Kishino, and Ian S. Painter. "Estimating the rate of evolution of the rate of molecular evolution." Molecular biology and evolution 15.12 (1998): 1647-1657.  
To, Thu-Hien, et al. "Fast dating using least-squares criteria and algorithms." Systematic biology 65.1 (2015): 82-97.  
Yang, Ziheng, and Carlos E. Rodríguez. "Searching for efficient Markov chain Monte Carlo proposal kernels." Proceedings of the National Academy of Sciences 110.48 (2013): 19307-19312.  
  
  
  
  
  
  
%ATTACH\_FOR\_AUTHOR\_DEEP\_LINK  
  
  
If improvements to the English language within your manuscript have been requested, you should have your manuscript reviewed by someone who is fluent in English. If you would like professional help in revising this manuscript, you can use any reputable English language editing service. We can recommend our affiliates Nature Research Editing Service (<http://bit.ly/NRES_BS>) and American Journal Experts (<http://bit.ly/AJE_BS>) for help with English usage. Please note that use of an editing service is neither a requirement nor a guarantee of publication. Free assistance is available from our English language tutorial (<https://www.springer.com/gb/authors-editors/authorandreviewertutorials/writinginenglish>) and our Writing resources (<http://www.biomedcentral.com/getpublished/writing-resources>). These cover common mistakes that occur when writing in English.   
  
  
--------------------  
Editorial Policies  
--------------------  
Please read the following information and revise your manuscript as necessary. If your manuscript does not adhere to our editorial requirements, this may cause a delay while this is addressed. Failure to adhere to our policies may result in rejection of your manuscript.  
  
In accordance with BioMed Central editorial policies and formatting guidelines, all manuscript submissions to BMC Evolutionary Biology must contain a Declarations section which includes the mandatory sub-sections listed below.  Please refer to the journal's Submission Guidelines web page for information regarding the criteria for each sub-section (<https://bmcevolbiol.biomedcentral.com/>  
).  
  
Where a mandatory Declarations section is not relevant to your study design or article type, please write "Not applicable" in these sections.  
  
For the 'Availability of data and materials' section, please provide information about where the data supporting your findings can be found.  We encourage authors to deposit their datasets in publicly available repositories (where available and appropriate), or to be presented within the manuscript and/or additional supporting files.  Please note that identifying/confidential patient data should not be shared.  Authors who do not wish to share their data must confirm this under this sub-heading and also provide their reasons.  For further guidance on how to format this section, please refer to BioMed Central's editorial policies page (see links below).  
  
Declarations  
-     Ethics approval and consent to participate  
-     Consent to publish  
-     Availability of data and materials  
-     Competing interests  
-     Funding  
-     Authors' Contributions  
-        Acknowledgements  
  
Further information about our editorial policies can be found at the following links:  
Ethical approval and consent:  
<http://www.biomedcentral.com/about/editorialpolicies#Ethics>  
Availability of data and materials section:  
<http://www.biomedcentral.com/submissions/editorial-policies#availability+of+data+and+materials>  
  
Recipients of this email are registered users within the Editorial Manager database for this journal. We will keep your information on file to use in the process of submitting, evaluating and publishing a manuscript. For more information on how we use your personal details please see our privacy policy at <https://www.springernature.com/production-privacy-policy>. If you no longer wish to receive messages from this journal or you have questions regarding database management, please contact the Publication Office at the link below.  
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: <https://www.editorialmanager.com/evob/login.asp?a=r>). Please contact the publication office if you have any questions.