**Response to the editor and reviewer**

We greatly appreciate the editor and the reviewer for the efforts and the valuable suggestions and hope that deficiencies pointed out in the original submission are overcome in the revised version. Our responses of the Referee’s Report are given below.

--------------------------------------------------

**General comments**

*1. The terminology around taxa with sampling times in the past is confusing.*

*(1) In algorithm 2, the use of extant/extinct does not match any common usage. The definitions on page 7 line 33 for Y and O are clearer and avoid this confusion. I suggest removing extant/extinct entirely.*

*(2)* *It seems that the authors are using "sampled ancestor" to mean "a tip whose sampling time is in the past," or more compactly "a heterochronous tip." This is not the standard usage of sampled ancestor, which generally means "a sampled taxon who has descendants which are also sampled taxa."*

*(3)* *On page 6 (l 24-25), it is stated that a sampled ancestor has no descendants (is of degree 1, having only a parent). This suggests that "sampled ancestor" is being used to mean "heterochronous tip."*

*(4) Does the "Big Pulley" algorithm apply if O is a node of degree 2 (having a parent and a single child)? It seems like it should, but I do not know if it would fall under the symmetric or asymmetric case. The authors are free to leave this case to future work, so long as the terminology surrounding what O represents is made clearer.*

**Author’s Response:**

Thank you for your professional comments and suggestions.

(1) We have removed the notation of “extant/extinct” in the revised manuscript. The two child nodes of the root are denoted by O (having two child nodes) and Y (having no child nodes).

(2) – (3) We have removed the usage of “sampled ancestor” in the revised manuscript. In the trees where the Big Pulley operator works on, node Y means a heterochronous tip.

(4) In the revised manuscript, we have made it clear that O is the node with two child nodes. The situation where O has a parent and a single child node will be handled in our future work.

*2. Figure 10 and Table 7 appear contradictory about the models used to analyze RSV2 and HIV-1. In the table it is stated that there is an operator on the population size, but in the figure there are efficiencies listed for birth and death rates instead.*

**Author’s Response:**

Thank you for pointing out this mistake.

In the latest manuscript, we have corrected the inconsistent parameters in Figure 10 and Table 7. To be specific, the Anolis data set has a birth-death tree prior in our model and two corresponding parameters (birth rate and death rate) are sampled in the analysis. For RSV2 and HIV-1 data sets, we used coalescent model as the tree prior, where the parameter population size (pop.size) is sampled. Moreover, we also sampled clock mean (ucld.mean) for RSV2 and HIV-1 because we specified sampled dates at the tips. Therefore, in Figure 7, birth.rate and death.rate are compared in Anolis data set, pop.size and ucld.mean are compared in RSV2 and HIV-1. The rest of parameters are the same in the three data sets analysis.

**Typos and other minor comments:**

*1. The comparison of ESS for the clock standard deviation between "cons" and "categories" doesn't seem exactly fair, since there is an entirely new operator for the standard deviation of the clock in "cons."-It may be useful to mention the "nocons" tests were run in previous analyses. This would assuage any concerns that the difference between "categories" and "cons" is about the change from discretized to continuous branch rates, rather than about the use of the constant distance operator.*

**Author’s Response:**

Thank you for your professional comments

In the revised manuscript, we have added a new subsection to discuss the “NoCons” configuration in Appendix Section 4. The details are as follows.

In the comparison of ESS for the clock standard deviation (denoted by ucld.stdev in Figure 10), we specified a normal scale operator in "Category" configuration. In "Cons" configuration, the UcldstdevScaleOperator is used to sample the clock standard deviation of continuous rates. To avoid the concern that the difference between "Category" and "Cons" is a result of how rates are parameterised (i.e. discrete or continuous), we set another configuration where continuous rates are sampled without using the ConstantDistance operator (denoted by "NoCons" configuration). The weights of the operators in "NoCons" are the same as those in "Category" which is detailed in Table 7. We ran the analysis using the three real data sets (Anolis, RSV2 and HIV-1) and the efficiency comparison between "Category", Cons" and "NoCons" is summarised in Figure 25. The figure shows ESS per hour in log\_10 space of ucld.stdev in 20 independent MCMC chains. As can be seen, "Cons" configuration has the largest ESS per hour. More specifically, on average, ESS per hour in "Cons" is a little bit larger than that in "Category", which is consistent with Figure 10. Moreover, ESS per hour is improved a lot in "Cons" when comparing with "NoCons", where both continuous rates are sampled. However, we also noticed that the rate parameterisation does have an effect on the mixing performance in MCMC chains. In the future, we will further investigate how to parameterise branch rates to get better performance when using the proposed operators.

*2. Table 7: "Substituion model"*

**Author’s Response:**

In the revised manuscript, it has been replaced by “Substitution model”.

*3. p26 l28: the simulations are not definitive proof, but rather a convincing demonstration*

**Author’s Response:**

In the revised manuscript, the inappropriate statement “Thus, Simple Distance is proved to be correct.” is replaced by “Thus, Simple Distance samples the root time and two branch rates correctly.”

**We greatly appreciate the reviewer for the valuable suggestions. We try our best to overcome the deficiencies pointed out in the original submission. If there are any problems in the revised version, please do not hesitate to point out. We will revise the submission according to reviewer’s suggestions.**