The authors conducted an interesting study in exploring the issues affecting dating single gene trees using molecular clock approaches. They appear to have a good understanding of molecular clock literature. The perspective is very novel and makes much sense. Strictly speaking, there are many issues to address some of which are even not covered in my following comments. However, to encourage open science and the new academic publishing way using PCI, I would suggest giving the authors an opportunity to revise.

In addition, reviewing this technically interesting but complex ms requires much expertise much of which may be beyond my knowledge. So my apologizes in advance for any misunderstanding and please correct me if that happens.

Major points:

1. In Fig. 1, the authors mentioned the ultimate goal, but this seems to me to make the present ms a bit confusing as duplication is not mentioned and dating gene trees with duplications also involve other issues for example the accelerated rate due to relaxed purifying selection, loss of duplicates in some lineages, and many others. In the current analysis, the authors removed all genes that showed duplications, which is clearly indicated. So, it seems a bit strange that Fig. 1 is presented in this way. Instead, I was wondering if it's better to move anything related to duplication to the end or SI.
2. The use of TimeTree to set calibrations should be very careful. See below.
    a. L156: The authors set the calibration at the specified node based on the recorded dates in TimeTree database but there might be some to address. Looking at the corresponding method description in L419-L431), the authors chose to fit the 95% CI and the point estimation using a gamma distribution. However, as some suggest, the times in TimeTree should be interpreted very carefully and they simply do not like the idea of using that as a reference. Hence, if the authors are going to take this advice, they might want to choose one or a few alternative calibrations for analysis.
    b. Also, setting the calibrations at other nodes or adding a few calibrations would also help. The reason is simple: only when people do not have good calibrations would setting a single calibration at the root makes much sense. So in this regard, the results are very interesting but the practical significance seems to me not very clear. This at least should be discussed.
    c. Speaking as a researcher not believing the above point so much, even if the calibration used by the authors is widely used by many people, it is almost always suggested to run the same analysis using alternative calibrations.
    d. Further, the detailed parameters of the Gamma distribution used as the calibration should be made clearer. Is it the following in L428? If so that corresponds to Gamma(4.6,0.656)

but I see that α and β are the same for the two calibrated nodes, aren't they? Also, the specified Gamma distribution above has a mean at $\frac{4.6}{0.656}$ which is apparently smaller than 70.8 and 40.9. Am I misunderstanding anything?

95% interval: location 70.8 My, α=4.6 and β = 0.656 for ~~Primates~~, and location 40.9 My, α=4.6 and β = 0.656 for Simiiformes. ~~The MCMC was run in one chain of 50,000,000~~

e. Another way to rely on TimeTree is to fit all recorded dates of the node of interest into a gamma distribution instead of basing the analysis on only a mean value and the CI. This particular subpoint is simply a suggestion so the authors can well take the liberty to accommodate it or not and actually I see no need to accommodate it.

f. Perhaps more importantly, the authors gave me a "wrong" impression that the dates recorded in TimeTree were somehow taken as a "reference" although this is neither case for setting the calibration nor when comparing their obtained posterior dates with TimeTree. I think this is partly because the authors mentioned at the very beginning the use of TimeTree and throughout the manuscript. I suggest the authors change this way of writing which can easily mislead readers like me.

3. L207: In Fig. 3, why was the intercept set to zero? I think it is the result of OLS i.e., the second round of your regression analysis, isn't it? Also for Fig. 3, are these 9 items the only ones that were significant? Was there any multiple testing?
4. I think the use of simulations are worth mentioning in the abstract. This part also seems to me to make more sense since by that people know the "true" values of the parameters to estimate. So I suggest expanding the simulation part more.
5. One important thing to explore for single gene dating might be the impact of the time priors on the time posteriors. Some suggest that single gene trees contain not much information so their dating results are not trustable. This seems to me to be important, but I also understand this may be beyond the scope of the study. Nevertheless, I was wondering if the authors can discuss this a bit.
6. The ms involves many methods. In general, I would suggest referring to Methods when they are mentioned in Results. For example, L281: the authors should mention "see Methods", and the same applies to elsewhere. This would greatly help readers understand the work.

Minor points:

1. L74: per site?
2. L89: I could be wrong but I don't think heterotachy is related to across-site difference. It in my memory specifically refers to the heterogeneity among branches.

3. L100: I suggest citing the corresponding papers of the two models here, although I see the papers I would cite are cited later in the para.
4. L122: The authors mention the following "the most critical limitations are the difficulty of characterizing the type of clock relaxation and the uncertainty in calibration points themselves (dos Reis et al. 2015)". So, I was wondering how they were dealt with in the present ms, particularly the choice of the clock model (as I see the authors chose to use an independent rate model assuming the rate across lineages $\overset{i.i.d.}{\sim}$ log-normal). This could be a bit challenging but if not possible to test at least the authors should discuss this limitation.
5. L164-L165: the statement is not wrong, but because the authors used a single calibration on some 5000 single gene trees, I would tend to believe those estimated by other studies. So in my view this sentence is not very meaningful to be mentioned here.
6. L203: any reference?
7. L215-L219: my perhaps not accurate understanding is that even for uncalibrated nodes, they have time priors derived from a birth-death process (in the present ms I think the authors mentioned it's a Yule process), and also there are effective priors caused by the truncation effects both of which could lead to time priors not as flat as a uniform distribution (see Barba-Montoya et al. 2017 MPE). That said, given proper priors, I am unsure if the non-identifiability mentioned by the authors holds. Again, I am unsure.
8. L223: I think you may want to cite https://link.springer.com/article/10.1007/BF00160154.
9. L415-L417: the authors used the codon alignments if I did not misunderstand anything. It should be made clearer here if so.
10. L421: What's the single parameter for the Yule process to specify the priors for uncalibrated nodes?
11. L428: what software did the authors use for assessing the ESS?
12. L452: Isn't $\kappa = 4$ a bit large?
13. L452-L453: If I interpreted it correctly, the authors did not involve an across-rate heterogeneity in the analysis by setting the α parameter. Is that correct? This is fine as this part of analysis is not the main one, but the authors should clearly indicate this.
14. L509: do you mean "across-branch rate heterogeneity"? If so, I'd also suggest showing the equation that calculates v. I also do not quite understand the equation in L511. I wondered if the authors mind showing how it is derived.
15. L516: grammar mistake
16. The figures are duplicated at the end. Already inserted when they are first mentioned. This is not an error and could be due to the requirement of bioRxiv in uploading the files so actually no need to address.