One thing I thought might help “sell” the paper to an phylogenetically unfamiliar audience of paleontologists would be to highlight more empirical case studies mostly or entirely comprised of fossil data (to help point readers to examples on groups having fossil records like the one they know best). That said, I wouldn’t necessarily advocate any additional section(s) be added to compensate--just a thought.

Although the joint estimation of relationships and node ages is emphasized in the manuscript (and in the literature), I think it’s worth briefly mentioning the possibility of dating taxonomy-based phylogenies of fossil taxa. Although this should be done with caution, this approach could benefit paleontological studies of trait evolution, which typically do not incorporate information on relationships at all. This issue is discussed in Soul and Friedman (2015) and I think it might be helpful to some paleo-focused readers if it were mentioned (only a sentence or two would be needed).

Done.

Specific comments

\*\*\*Note the manuscript does not have line numbers, so my comments reference both the page number in the upper left/right side of the page and the opposite page number. For example, “Page 5 of 41” appears on the left of the same page as 3 on the right.

Page 3 of 41, 1:

2nd paragraph: There’s a reference to a missing figure here??

Done.

Page 5 of 41, 3

3rd paragraph: Would it be helpful to include a figure showing how the shape of a distribution (e.g., gamma, beta) changes with different combinations of parameter values? It might help readers understand what’s meant by a “flexible” prior. I guess this is sort of addressed in figure 4 though.

4th paragraph: I don’t see a box labeled ‘The Likelihood’. Do you mean the box labeled “The likelihood the prior and the posterior” on page 22?

Correct, changed.

Page 6 of 41, 4

2nd paragraph: In the third to final sentence, “the analysis [is/was?] bad”.

Correct, changed.

3rd paragraph: Maybe change “equivalent” to “analog”, since credible intervals are exactly the conceptually as confidence intervals.

Done.

In the discussion on summary trees, it might be good to reference Heled and Bouckaert’s (2013) paper.

Done.

Page 7 of 41, 5

3rd paragraph: “…put strong prior information on the average substitution rate and speciation times”. Do you mean simultaneously? Or should this say “substitution rate or speciation times”?

The second. Corrected.

Page 10 of 41, 8:

Under clock models: To my reading, it seems like the comma is awkwardly placed in the first sentence here.

Corrected.

Page 15 of 41, 13:

1st paragraph: I think “probably” is supposed to say “probability”?

Corrected.

Page 25 of 41, 23:

Maybe change “best likelihood” to something like “greatest” or “highest” likelihood?

Corrected.

Page 35 of 41:

Figure 1:

For the undated topology, do the branch lengths have units or is this just meant to be a cladogram illustrating topology? The text in the caption implies it has units of character change (e.g., # of expected changes per character). If so, should the tree be ultrametric (implying identical amounts of change from the root-to-tips)? Or would it be clearer to readers if the tree was non-ultrametric but had a scale bar depicting some unit amount of change? Either is okay but I think it would help to make it clear.

Figure 2:

No suggestion—just want to say this is a great figure!

References

Heled, J. and Bouckaert, R.R., 2013. Looking for trees in the forest: summary tree from posterior samples. BMC evolutionary biology, 13(1), p.221.

Soul, L.C. and Friedman, M., 2015. Taxonomy and phylogeny can yield comparable results in comparative paleontological analyses. Systematic Biology, 64(4), pp.608-620.

Reviewer 2

Fig. 2. I know this suggestion is a bit picky, but here the binary information illustrated to represent phylogenetic characters might only be representative of morphological characters. Would it be possible to also add three lines next to this for molecular data illustration in this box "The data" such as:

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

0101... ATTG...

1101... AND/OR TTGC...

0100... ATTC...

Phylogenetic

characters

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

As you are mentioning this, maybe you could also introduce the need for acquisition/ascertainment bias correction here due to the lack of these characters in morphological matrices?