



Tony Jhwueng <tony.jhwueng@gmail.com>

Fwd: Systematic Biology - Decision on Manuscript ID USYB-2015-180

1 message

Brian O'Meara <bomeara@utk.edu>
To: Tony Jhwueng <tony.jhwueng@gmail.com>

Wed, Oct 28, 2015 at 10:08 AM

Brian O'Meara
Associate Professor
Dept. of Ecology & Evolutionary Biology
U. of Tennessee, Knoxville
<http://www.brianomeara.info>

Postdoc collaborators wanted: <http://nimbios.org/postdocs/>
Calendar: <http://www.brianomeara.info/calendars/omeara>

----- Forwarded message -----

From: <feander@siu.edu>
Date: Tue, Oct 27, 2015 at 11:58 AM
Subject: Systematic Biology - Decision on Manuscript ID USYB-2015-180
To: bomeara@utk.edu
Cc: sysbio.editorialoffice@oup.com, lukeh@uidaho.edu

27 October 2015

Dear Dr. O'Meara,

Decision on USYB-2015-180, Trait Evolution on Phylogenetic Networks:

Reject; resubmission encouraged

Thank you for your Systematic Biology submission. It has been reviewed by Associate Editor Dr. Luke Harmon and an anonymous reviewer. Their comments are listed at the end of this letter. The reviewers and the AE provide some excellent constructive suggestions that I am sure you will appreciate.

I second Luke's apology in the delay in returning a decision to you. I think we all agree that this is a great idea, and also quite timely, as network methods are slowly (finally) starting to see more use in systematics. The reviewer and Luke have a few major concerns, and I agree with them that until those are addressed, the manuscript is not ready for acceptance.

I have one substantive concern that may require some new analyses, and several minor points/questions about the writing.

My only substantive point:

1) Lines 249-253: This is really clunky, both in writing and analytical details. How many sequences are we talking about here? You need to state that information. How many genes? Which genes? How many taxa? (ah...some of this information shows up later, but see my comment #13 below) Unless there were thousands of individuals/species and lots of genes (which does not appear to be the case), surely you could get a great estimate of the ML tree and branch lengths using Garli? What is the basis for using the GTR+G model here? Is a clock model justified? What does "We filtered for best" mean? This section needs substantial clarification, and it seems like some analytical shortcuts may have been taken that probably weren't really necessary and aren't justified.

Minor points:

1) Lines 20-21: "...of Brownian motion, as well as..." – Redundant with first part of sentence, so you can delete

“of Brownian motion”

2) Line 21: Probably better as “at the tips”?

3) Line 25: You can obviously call your package whatever you want, but why “BMhyd”? Wouldn’t “BMhyb” be better? I guess I’m mostly just curious about why you avoided what I think is the obvious choice here.

4) Line 71: “two empirical data sets of cichlid and *Nicotiana*” might be better as “two empirical data sets for cichlids and *Nicotiana*”.

5) In Figure 1, t1, t2 and t3 seem to refer time spans (e.g., t1 is the time between the basal split and the split between D and X), but in the caption and the manuscript text, they seem to refer to time points (e.g., t1 is the time point at which D and X diverged). This needs to be clarified on the figure.

6) Line 128: Should be “there exists widespread heterosis”

7) Line 159: I think I would write “non-hybrid species” here instead of “usual species”.

8) Line 190: Should be “off-diagonal elements”

9) Line 199: “the proportion of flow of hybrid inherited from the parents” – This needs to be rephrased. How about “the proportion of genes inherited from each parent”?

10) Line 217: I think I might write “and across different methods of summarizing parameters” instead of “and how a summary of the parameters could be calculated”, but I guess this is just a matter of taste.

11) Line 218: No need to hyphenate “model averaging”.

12) Line 247: “trim ends of sequences for only a small subset of taxa” – What does this mean? Why trim the ends? Why only for a subset of taxa?

13) Line 257-261: I think this information (as well as the number of loci you studied) needs to be presented before you present the analytical details.

14) Line 275-279: Was relative seedling growth the only character you could use? Are there others that are not proportions?

15) Line 281: I urge authors to avoid sentences that start with clauses like “Figure 2 represent[s]...” or “...are shown in Fig. 3”. I think you should just cite Figure 2 when you mention the cichlid tree, and cite it again when you talk about *Nicotiana*.

16) Line 305: Should be “Edwards (1992)”.

17) Line 355-356: This is somewhat awkward. It might be better as “having them as the result of ten independent hybridization events” or something like that.

18) Line 362-363: See my comment above about sentences like this. Adjust it if you can – it just makes for smoother reading.

19) Line 363-366: Could some (or all) of this information just be presented in the figure caption?

20) Line 367-368: See comment above; modify “...are shown in Supp. Fig. 1” sentence.

21) Line 388: I’m sorry the range of estimates wasn’t narrow, but this range made me chuckle. Hopefully the reviewer’s correction to the covariance matrix (if [s]he is correct that the matrix needs correction) can help tighten up some of your inferences.

22) Line 395-396: This sentence should end with a question mark.

23) Line 397: “Results are shown in Figure 4” – See my comments above about sentences like these.

24) Table 1 caption: Should be “The model-averaged parameter estimates...are reported”

25) Figure 5 caption: Should be “gray or black dot.”

If you choose to submit an entirely re-worked, new paper on this topic to Systematic Biology please address each point made by the Editor, AE and reviewers. Include the responses in the Manuscript Central field under "Response to Decision Letter." The best way to address each point would be to copy this file, and insert your comments after each point made. Please do not change the order of or delete any of the comments because this makes it difficult to review again and would slow the review process. Be sure to clearly distinguish between your comments and the reviewers' comments. Feel free to argue your case with careful documentation if you disagree with any of the suggestions.

PLEASE NOTE: when you are ready to submit a new version, please do so as a resubmission. To do this, log into Manuscript Central, go to your Author Center, and click "Manuscripts with Decisions." You should find the previous version listed there, with the option "create a resubmission." Please use this link. Include any data or other supplemental files when you upload the resubmission, because they will not automatically transfer over from the original version. If you have any questions or problems please email us at sysbio.editorialoffice@oup.com. Author instructions are available online at [HTTP://SYSBIO.OXFORDJOURNALS.ORG](http://SYSBIO.OXFORDJOURNALS.ORG).

Would you please acknowledge receipt of the reviews by email to sysbio.editorialoffice@oup.com and let us know if you plan to submit a new paper on this topic?

Thank you very much for your submission.

Sincerely,

Prof. Frank (Andy) Anderson
Editor in Chief, Systematic Biology
feander@siu.edu

Associate Editor: Dr. Luke Harmon
Comments to author
Recommendation #1: Reject; encourage resubmission
Associate Editor: Harmon, Luke
Comments to the Author:
Dear Drs. Jhwueng and O'Meara,

I apologize for the delay - for some time now I have had one review of your manuscript in hand and been waiting for another. But I think we have all waited long enough! Unfortunately this means that I will be returning just one review along with some of my own comments. Hopefully this is still helpful enough, and the one review is substantive.

I think that the method you have developed is important. We need comparative methods for networks, and your paper is a key step in that direction. However, I am also persuaded by the reviewers argument that the diagonal elements in your VCV are incorrect. I am convinced by the worked example, along with the intuitive idea that the variance of hybrids should be lower than non-hybrids in a way that depends on m and t. Thankfully the reviewer also suggests a solution from the pop-gen literature. I am not sure what effect this will have on overall inference from the simulations, but you could see a bigger effect looking at the real data.

The reviewer also questions the interpretation of m given that species' traits are being considered on a log scale. I think that this question is worth considering, although it seems possible to me that one might be able to predict phenotypes of hybrids by calculating weighted averages on a log-scale. After all, quantitative genetics often uses log-transformed traits to calculate heritability and so on - and one would see strange things transforming back to the original trait scale. To me, the argument that saves the authors' interpretation is to consider the trait as a product of a large number of independent factors, so that the trait is a sum of the large number of factors on a log-scale (this argument is attributed to Galton 1879 in Lynch and Walsh p 295). In that sense I think the use of m makes good sense, and one can even interpret $x_1^m * x_2^{(1-m)}$. In any case, this whole discussion is worth mentioning in the paper.

I have three comments on reading the paper.

My first comment is that I found the paper, overall, clear but got stuck a bit in the complexity of Figure 1. The main issue is, I think, introducing both the main model and the role of extinct species at the same time. I recommend a two-panel figure, one with a more "standard" situation of hybridization among sampled lineages, and one involving unsampled lineages (I note, also, that this is a bit different from what is described in the cichlid example as "extinct" versus "extant" hybridization, as I think what matters is not whether an ancient lineage is extant or not but whether it is included in the tree). I would also like a little bit of a discussion about when and how one can infer the presence and timing of hybridization involving unsampled lineages. I think for this method you need to infer two times to get the right VCV, so this all seems worthy of discussion.

I also wonder a bit about the interpretation of vH. I agree that a large vH will capture transgressive segregation - but the question of whether or not hybrids deviate from the range of the parents depends on the difference in means and vH together. I don't think that what you say in the paper is incorrect, per se, but I do think that people might misinterpret vH given what is written.

Finally, I am worried about the role of SE in analyzing the empirical data. This comes from my own experience trying to do this in a totally different context, but we have found that trying to estimate SE has unpredictable effects on estimating other parameters. Also, empirical biologists almost never measure just one individual and can give reasonable estimates of SE that are really hard to get from the comparative approach. I suppose what I would like to see is a re-analysis of the empirical data using some fixed SE to see how much that matters.

As a more minor point, I think the results section could be re-written a bit. It reads a bit dire right now - which might be fine - but the positive aspects of parameters that you can estimate well don't come in until the discussion.

Again I apologize for the single review, and the continued critique of what I think is a worthwhile and important paper.

Sincerely,
Luke Harmon

Reviewer(s)' Comments to Author:

Reviewer: 1

Recommendation: Reject; resubmission encouraged

Comments:

Thank you for the opportunity to review your paper. Please find my review in the attached pdf.

Additional Questions:

Do you wish to remain anonymous?: Yes

How significant is this work?: Moderately

Is the author aware of the background and source material to the problems set forth?: No

Are the conclusions justified by the evidence presented and the assumptions involved?: No

Are the illustrations and tables clear and understandable?: Yes

In number are they: Sufficient



review.pdf
39K