September 22nd, 2017

Department of Ecology and Evolutionary Biology

Yale University

New Haven, CT 06520-8106

Dear Dr. Near

Editor-in-Chief of *Systematic Biology:*

We are grateful to the Associate Editor and a reviewer for their thoughtful and constructive criticisms. Below are our responses to their comments. The reviewer’s comments are in bold, and our responses are in plain text underneath. Attached is the revised version of the manuscript that has incorporated these revisions. Thank you for your consideration.

Sincerely,

Dwueng-Chwuan Jhwueng

Brian O’Meara

**Decision on USYB­2015­180, 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 249­253: 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.**

We have clarified the phylogenetic analysis. Our goal was to use the single gene from that empirical study to reconstruct their tree of just 27 taxa, which was not deposited anywhere. Our goal was not to derive a new hypothesis of cichlid (or Nicotiana) relationships, just get as close to the tree used by the original authors as possible. This is in flavor similar to doing congruification in Geiger, but using empirical branch lengths as starting branch lengths and treePL, which is more robust, albeit slower, than pathd8.

**Minor points:**

1. **Lines 20­21: “...of Brownian motion, as well as...” – Redudant with first part of sentence, so you can delete “of Brownian motion”**

We have made the editorial change.

1. **Line 21: Probably better as “at the tips”?**

We have made the editorial change.

1. **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.**

We have changed our package name to BMhyb and it is currently on CRAN.

<https://cran.r-project.org/web/packages/BMhyb/index.html>

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

We have made the editorial change.

1. **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.**

We have made these editorial changes. We assumed that the root occurred at time point t=0. We then refer t1 as the time point where D and X diverged, and t1+t2 as the time where R and Y diverged. Under this circumstance, the time spans t1= t1-0 is equivalent to the time between the basal split and the split between D and X. And t1+t2 = t1+t2 – 0 is equivalent to the time between the basal split and the split between C and R. And [t1, t1+t2] is the time interval that the changes occurring on the X branch (Figure 1) that cannot be shared with the hybrid B.

1. **Line 128: Should be “there exists widespread heterosis”**

We have made the editorial change.

1. **Line 159: I think I would write “non­hybrid species” here instead of “usual species”.**

We have made the editorial change.

1. **Line 190: Should be “off­diagonal elements”**

We rephrase the sentence without using “off-diagonal elements”.

1. **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”?**

We now use equally proportion of genes inherited from each parent for our new simulations.

1. **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.**

We now look at the deviation between the observed and true parameters across different methods of summarizing parameters.

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

We have made the editorial change.

**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?**

We have removed this paragraph in our manuscript.

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

We have prepared this information and included it in the manuscript.

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

There are other possible characters which are not proportion, our current search use the the relative seedling growth data in proportion provided in Komori et al. (2000). Note that Grotkopp et al. (2002) suggests the relative growth rate analysis is a powerful tool for understanding life-history trait as it combines aspects of species’ anatomy, morphology, and physiology.

**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.**

We have made these editorial changes and rephrase the sentences.

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

We have made the editorial change.

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

We have made the editorial change. We rewrite the sentence.

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

We have made the editorial change.

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

We have made these editorial changes.

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

We have made the editorial change.

**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.**

We have corrected our covariance matrix and redo the analysis and simulations. Our results show that the estimates for hybrid vigor often centered on the true value, while the estimates of the hybrid variation parameter improved from previous analysis but was still inaccurate in particular for fairly large taxa over 100 taxa with 10 hybrids.

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

We have made the editorial change.

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

We have made the editorial change.

**24) Table 1 caption: Should be “The model­averaged parameter estimates...are reported”**

We have made the editorial change.

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

We have made the editorial change.

**If you choose to submit an entirely re­worked, 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.**

**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**](mailto: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 non­hybrids in a way that depends on m and t. Thankfully the reviewer also suggests a solution from the pop­gen 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 log­scale. After all, quantitative genetics often uses log­transformed 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 log­scale (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 x1^m \* x2^(1­m). 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 two­panel figure, one with a more “standard” situation of hybridization among sampled lineages, and one involving unsampled lineages**

We have now done this as a three panel figure: standard hybridization with sampled parents, a middle panel showing the true history if there’s an unsampled parent, and a final panel showing how the hybridization appears with incomplete sampling.

**(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).**

Indeed: we have modified the text appropriately.

**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.**

We have added a short discussion of this. Note that our method doesn’t infer the timing of hybridization, but can use this information.

**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 re­analysis of the empirical data using some fixed SE to see how much that matters.**

Unfortunately, neither empirical dataset has SE. However, we have approximated SE for the fish dataset and now have analyses incorporating this*.*

**As a more minor point, I think the results section could be re­written 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 of USYB-2015-180**

**The bifurcating phylogenetic species tree requires that evolutionary traits are inherited ancestrally, but never horizontally between species as might occur through hybridization. Jhwueng and O’Meara extend the Brownian motion model to operate on phylogenetic hybridization networks, where hy- bridization may induce a burst of phenotypic variation (vH) or a rescaled (β) mixture of hybrid phenotypes (m). The authors approach this by adopting the multivariate normal representation of a tree-dependent Brownian mo- tion, but modifying how the covariance terms are defined in terms of hybrid related variation. Overall, I think the idea is a biologically reasonable one, but I have some concerns about the formulation of the model.**

**First, the covariance matrix definition may be incorrect. The simplest example showing this concern would be to consider two taxa X1 and X2 whose MRCA is the root. They hybridize at the present with m = 0.5, β = 1.0, and vH = 0.0 to create XH. X1 and X2 can be considered iid samples from a Normal distribution (i.e. Cov(X1,X2) = 0), but XH is the sum of weighted values, mX1 + (1 − m)X2.**

**Figure 1 shows that all contemporaneous extant species, since they have each evolved for the same amount of time since their common origin, have the same variance. Assuming σ2 = 1 can be suppressed, this shows that V ar(XH ) ̸= V ar(X1)**

**V ar(XH ) =Cov(XH , XH ) =m·m·Cov(X1,X1)+m·(1−m)·Cov(X1,X2)**

**(1−m)·m·Cov(X2,X1)+(1−m)·(1−m)·Cov(X2,X2) =0.5·0.5·1+0.5·0.5·0+0.5·0.5·0+0.5·0.5·1 =0.5**

**Or a small proof through simulation**

**> m=1.0;n=100000;var(m\*rnorm(n)+(1-m)\*(rnorm(n))) [1] 1.009986 > m=0.5;n=100000;var(m\*rnorm(n)+(1-m)\*(rnorm(n))) [1] 0.5010948**

**Now, if the hybridization event occurred immediately after X1 and X2 originated, then V ar(XH ) ≈ V ar(X1), since it would effectively manifest as another iid draw from the Normal.**

**This is a simple example (simpler than shown in Figure 1), and the im- plications of hybridization on covariance structure is more complicated in general. Since I did not see the work cited, the authors should be interested to learn of the work of Pickrell and Pritchard (2012) made available in the program TreeMix. They follow Cavalli-Sforza and Edwards (1967) in mod- eling allele frequency diffusion as a Brownian motion on a bifurcating tree, but allow admixed populations to inherit some proportion of alleles from two ancestral populations. In their Supporting Information document, equations 12 and 13 show how to compute the covariance for an arbitrary DAG. Using these equations, I believe V ar(R) = (m2 +(1−m)2)(t1 +t2 +t3)+2m(1−m)t3 for Figure 1, but the authors should verify this. The algorithm would be sim- ple enough to modify to include vH and β terms.**

We thank the reviewer for identifying this key flaw in our approach. People often complain about peer review, but this sort of thing is exactly why peer review is useful. We have modified the paper and equations, fixed the code, and re-run every analysis using the corrected code.

**From the Discussion, the authors conveyed some disappointment in the performance of their method. If they agree the covariance matrix is incorrect, they may find hope in the fact that fixing m = 0.5 maximizes the error, so they stand to gain the largest improvements with the correction (Line 295). Additionally, the correction will account for some missing covariance that arises when seemingly unrelated species covary due to complex admix- ture/covariance histories (Line 199 – 10 hybrids). On the other hand, if the covariance matrix is constructed identically for simulation and for inference, then performance may not improve by much. I am hopeful!**

**Aside from the covariance matrix, I worry about the log transformation of the data and its interpretation during hybridization events. The authors take some care with interpreting β in a log scale, but not m. When X1 and X2 are on a linear scale XH = mX1 + (1 − m)X2 is the m-weighted average of parent species and is intuitive. On the log scale, the mixture implies XH = exp(log(X )) = exp(m log(X ) + (1 − m) log(X )) = XmX(1−m), but it’s not clear what process this represents. Note, Pickrell and Pritchard (2012) don’t face this issue because they use a linear scale for allele frequencies and concede the model is poorly defined for boundary conditions (near 0 and 1). If the authors choose to remain in the log scale for traits, they must provide an interpretation for m.**

Minor note: following a suggestion by Cecile Ané, we have moved to the standard gamma rather than m, above. We do not follow the reasoning about difficulty interpreting gamma. At the extremes of 0 or 1, a trait is inherited completely from one parent or the other. At a value of 0.5, the value of the trait is the average of the log of the parent traits. For example, for a species with parents of size 10 cm and 20 cm, with a gamma of 0.5, its mean would be 14.1 cm, which seems reasonable.

**In light of what I view as fundamental flaws in the model, I cannot recommend the paper be published as is. The covariance matrix appears readily remedied, but I am not sure what to do about the hybridization parameter m other than to use a linear scale for Xi. If the authors choose to address these points, a fair portion of the manuscript will require rewriting. Although the flow and style were good, the content may need to change substantially, so I spent little time providing per-line corrections (typos, grammar, etc.).**

Thank you for the helpful comments.

**Thank you for the invitation to review this paper.**

**Pickrell and Pritchard (2012):**

**http://journals.plos.org/plosgenetics/article?id=10.1371/journal. pgen.1002967 Pickrell and Pritchard (2012) Supporting Info: http://journals.plos.org/plosgenetics/article/asset?unique&id=info: doi/10.1371/journal.pgen.1002967.s016**