**09-Nov-2017  
  
Dear Dr O'Meara,  
  
Decision on USYB-2017-184, Trait Evolution on Phylogenetic Networks:  
  
Reject; resubmission encouraged  
  
Thank you for your Systematic Biology submission. It has been reviewed by Associate Editor Dr Frank Burbrink and . 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.  
  
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**](mailto: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**](mailto: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,  
  
Dr Thomas Near  
Editor in Chief, Systematic Biology**[**thomas.near@yale.edu**](mailto:thomas.near@yale.edu) **Associate Editor: Dr Frank Burbrink  
Comments to author  
Recommendation #1: Reject; encourage resubmission  
  
Recommendation #2:  
Associate Editor: Burbrink, Frank  
Comments to the Author:  
The ms “Trait Evolution on Phylogenetic Networks “ by Jhwueng and O’meara for publication in the special issue of SysBio on phylogenetic reticulation have provided a much-needed method for examining trait evolution on a reticulating phylogeny using Brownian motion. The authors consider several different processes, transgressive segregation and basic introgression, as well as different network types (Fig.1) for examining the evolution of traits given reticulation. The paper is important for the obvious reason that the authors are suggesting and providing methods that will move forward the field of comparative phylogenetics to now include reticulating models; this will be more in line with new phylogenies using genomic data that show some influence of introgression or hybrid speciation. The paper is also important because it is likely to stimulate other areas of comparative phylogenetics (diversification, trait-driven diversification) to accommodate reticulation into new models.  
  
While both reviewers were quite enthusiastic about the paper, they have both converged on similar big problems. I summarize these issues here:  
  
Rev1 expresses concern that the worked version of the VCV matrix is inadequate; the authors provide no solution for this in the general case. Given that the authors have developed a package, BMhy, for estimating ancestral quantitative traits given reticulation, it seems they must have developed a method for the general case, however according to Rev1 this method in the package does not recover the correct solution. This reviewer has contacted the authors via github about this problem and has provided a simple script with their review demonstrating this problem. Additionally, Rev1 has provided an attachment of the ms with line-by-line edits.**

We appreciate the work done by the AE and both reviewers. We have revised the math and radically recoded the package, designed new simulations and run them, and reran the empirical analyses as well. The package performs MUCH better than the previous version, which we can put down to 1) no longer having incorrect math, 2) using better starting values.

Besides putting validation of the expected VCV in our unit tests, we have also done a search in PhyloNetworks in Julia (which was not available with our earlier submission) to validate the same results between that package and our independent implementation in R for the model that is common between the two approaches (we also have additional parameters; they have additional functionality (covariation between traits, ancestral state estimation). We include the relevant code here for both approaches. We put in gray the returned objects

Julia:

using PhyloNetworks

truenet = readTopology("((((D:0.4,C:0.4):4.8,((A:0.8,B:0.8):2.2)#H1:2.2::0.5):4.0,(#H1:0::0.5,E:3.0):6.2):2.0,O:11.2);")

vcv(truenet)

6×6 DataFrame

│ Row │ D │ C │ A │ B │ E │ O │

│ │ Float64 │ Float64 │ Float64 │ Float64 │ Float64 │ Float64 │

├─────┼─────────┼─────────┼─────────┼─────────┼─────────┼─────────┤

│ 1 │ 11.2 │ 10.8 │ 4.0 │ 4.0 │ 2.0 │ 0.0 │

│ 2 │ 10.8 │ 11.2 │ 4.0 │ 4.0 │ 2.0 │ 0.0 │

│ 3 │ 4.0 │ 4.0 │ 8.1 │ 7.3 │ 5.1 │ 0.0 │

│ 4 │ 4.0 │ 4.0 │ 7.3 │ 8.1 │ 5.1 │ 0.0 │

│ 5 │ 2.0 │ 2.0 │ 5.1 │ 5.1 │ 11.2 │ 0.0 │

│ 6 │ 0.0 │ 0.0 │ 0.0 │ 0.0 │ 0.0 │ 11.2 │

using DataFrames

dat = DataFrame(trait0 = [1, 2, 3, 4, 5, 6], tipNames=tipLabels(truenet))

using StatsModels

fitTrait1 = phyloNetworklm(@formula(trait0 ~ 1), dat, truenet)

fitTrait1

StatsModels.DataFrameRegressionModel{PhyloNetworkLinearModel,Array{Float64,2}}

Formula: trait0 ~ +1

Model: BM

Parameter(s) Estimates:

Sigma2: 0.498917

Coefficients:

───────────────────────────────────────────────────

Estimate Std.Error t value Pr(>|t|)

───────────────────────────────────────────────────

(Intercept) 4.18626 1.50908 2.77404 0.0392

───────────────────────────────────────────────────

Log Likelihood: -10.8993928821

AIC: 25.7987857642

**The R version:**

library(ape)

library(BMhyb)

phy.graph <- ape::read.evonet(text="((((D:0.4,C:0.4):4.8,((A:0.8,B:0.8):2.2)#H1:2.2::0.5):4.0,(#H1:0::0.5,E:3.0):6.2):2.0,O:11.2);")

traits <- sequence(6)

names(traits) <- phy.graph$tip.label

print(BMhyb:::PruneDonorsRecipientsFromVCV(ComputeVCV(phy.graph)))

D C A B E O

D 11.2 10.8 4.0 4.0 2.0 0.0

C 10.8 11.2 4.0 4.0 2.0 0.0

A 4.0 4.0 8.1 7.3 5.1 0.0

B 4.0 4.0 7.3 8.1 5.1 0.0

E 2.0 2.0 5.1 5.1 11.2 0.0

O 0.0 0.0 0.0 0.0 0.0 11.2

BMhyb:::BMhyb(phy.graph, traits, free.parameter.names=c("sigma.sq", "mu"))

sigma.sq mu AICc NegLogLik K

MLE 0.493941 4.176179 29.79914 10.89957 2

lower 0.1950513 0.9074645 NULL NULL NULL

upper 2.070029 7.444837 NULL NULL NULL

Key points are that the VCV matrices are identical, estimates of sigma.sq and mu are nearly identical (slight differences due to finite precision), estimates of log likelihood are also functionally the same (note that PhyloNetworks reports loglik, while we report negative log lik). We use AICc rather than their AIC, but were we to calculate AIC for ours, it comes to 25.79914, again subject to finite precision the same as their values.

**Rev2 points out that the method generally provides poor estimates of the parameters and asks the very simple question does this generate a better answer than simply using bifurcating trees? If the trait is better modeled on a bifurcating tree using OU or WN, then is it worth the trouble to attempt the solution provided here for a network? I would think that a reliable comparative method has to consider reticulation if the best phylogeny is reticulating. That clearly is the goal, but certainly must be better than a different model using an inaccurate representation of the tree of life. This reviewer is also justifiably concerned with the poor estimates of vh and beta parameters, and given the likelihood model and VCV matrix it doesn’t seem that the values for the vh parameter could be so inflated (other than via transgressive segregation gone wild). Rev2 also discusses an alternative to how transgressive segregation could be modeled given parental values (negatively related to the difference between parental values), though this is unlikely to adequately modeled from parental phenotypes.**

The new model works far better than the old, so this concern should be lessened, but we also empirically tested it. We compared the RMSE of parameter estimates between BMhyb and Geiger (which assumes a tree) for 1) all the simulation models and 2) only those models with beta of 1 and vH of 0 (so no bias or burst of variation for hybrids), since Geiger has no way to optimize those parameters. In both cases, BMhyb outperformed Geiger, suggesting that ignoring the network is not ideal.

Another advantage of our approach rather than tree approaches is the addition of our beta and vH parameters. These parameters address key questions in the hybridization literature. One could rig up tree approaches to somewhat get at these questions. As we include in our new discussion, one could fit an OU model with a different theta for hybrids vs nonhybrids to get at a shift, but this is not answering quite the same question as addressed with our model: it’s “do hybrids get pulled towards a higher mean [where they eventually stay around, once they reach equilibrium” under an OU, while we ask “does hybridization itself give you a kick to a higher [or lower] on average value, after which you continue evolving in the same way”. The models probably both reflect an underlying complex process to a coarse enough degree that if the true model is yes under either question, the other model would also say yes under its question, but it’s not an exact match (even ignoring the tree vs network aspect). A more exact match would be for v\_h: a burst of variation at the hybridization event could be modeled on a tree with sudden burst after a hybridization event (just painting on a faster regime post hybridization would not work, as the burst happens instantaneously, though one could approximate it with a short in time regime immediately after a hybridization). But our model and software finally perform well enough that just using the network directly is feasible and useful.

**I hope that the authors can fix these problems with VCV matrix and package in general. I envision a corrected version to be very important in our field.**

Thank you! **Reviewer(s)' Comments to Author:  
Reviewer: 1  
  
Recommendation: Reject; resubmission encouraged  
  
Comments:  
This is a pioneering work in dealing with the network structure of organisms in the study of quantitative traits, and the method implemented in R could be of great use to the community. The manuscript is clearly written, and I find the BM model very well presented and explained. However, my main concern is not about the main text of the manuscript, but about the companion R package "BMhyb" that implements the method. I could not understand it well, nor produce satisfying results with it on very simple examples. There are two facets to this problem, one theoretical, and one practical, as explained below.  
  
One of the key aspect of dealing with the network structure is to be able to compute the variance matrix it implies for measured traits at the tips. A simple example is worked out in the main text, but I could find no mention of an algorithm to deal with the general case. In my opinion, this is a major flaw of this work, as computing the variance matrix in the network case is not a straightforward extension of the tree case, and getting a correct variance structure is essential for the method to work. Hence, I think that the manuscript could highly benefit from a clear description of the method used to compute this variance matrix, either in the main text or in the appendices.  
  
If no algorithm is documented in the main text, one is implemented in the "BMhyb" package, for the method to work. However, I could not recover the correct variance matrix for several very simple networks when I tried it out. I provide a small R script showing my attempts and concerns about the results. These are not convoluted, special cases networks, as one of them is precisely the three tips network used in the manuscript. Without diving completely into the code, I browsed the function to compute the network induced variance matrix, and noticed some points that remain obscure to me. I pointed these out in the comments of the script provided. I've contacted the authors directly about these problems, and opened some public issues on GitHub. I think that they are currently working on a corrected version of the code.  
  
I discuss some other points of the manuscript in more details in the attached file. This file goes over the manuscript line by line, and contains some comments, suggestions, and questions about the text. However, for the two reasons pointed above (lack of theoretical explanation for a general case algorithm, and failure to get expected results on simple test cases), I am not convinced by the implementation the authors made of their method. It is my opinion that the two points raised above should be addressed before any other, and that they might have a strong impact on the rest of the results presented in the manuscript, that might need some extensive re-writing. (Among other consequences, the corrections that I think are needed in the variance matrix computation might improve the conditioning number of the matrices used, and hence maybe reduce the numerical stability problems that seems to be encountered by the authors (but this is just a guess)).  
  
Thank you for giving me the opportunity to review this interesting work. I hope that the comments I made will be useful to the authors.  
Paul Bastide**

We thank the reviewer for his insightful review and in saving us from making mathematical errors. We have changed nearly everything inside the package to correct the errors; we now include unit tests to validate that we’re getting the correct VCV. Package performance has also improved (correct math, plus better starting points). **Additional Questions:  
Directions for Reviewers: The authors will appreciate detailed comments on the manuscript. Please write comments for the authors in a separate file, numbering all items that should be addressed before the manuscript is acceptable for publication, and attach your file at the bottom of this form (if you’ve inserted comments on an electronic copy of the manuscript, please attach that file as well). Reviewers are reminded that Systematic Biology is interested in publishing well-written papers of high scientific quality and of general interest. Thus, in your review, please address both the appropriateness of the paper for the journal as well as its scientific strengths and weaknesses. Please note that our instructions for authors are available on our website,**[**systbiol.org**](http://systbiol.org/)**. Use the buttons above to access the manuscript files. The HTML and PDF buttons link to the entire manuscript. Individual submitted files (such as data files) are available under the “Supplementary Files” button. We encourage reviewers to make comments directly on an electronic copy of the ms. If you do not have software that would allow you to make comments on the pdf version, please check under “Supplementary Files” to see if a Word version is available.:   
  
Do you wish to remain anonymous?: No  
  
How significant is this work?: Very  
  
Is the author aware of the background and source material to the problems set forth?: Yes  
  
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  
  
The following people would be appropriate to review this MS further: Klaus Schliep  
  
  
Reviewer: 2  
  
Recommendation: Accept with major revisions  
  
Comments:  
The authors in this manuscript present a novel method for inferring Brownian trait evolution on phylogenetic networks. Overall, I find the paper interesting, novel and of interest to the readership of Systematic Biology. I also really like the adaptive confidence interval estimation. However, I have a few some suggestions for the authors that I hope will be implemented in a revision.  
  
Disappointingly, the estimation of hybridization-specific parameters ranges from not-very-good to terrible. This is useful to know, but suggests that this is not going to necessarily be the best avenue for estimating these parameters, unfortunately. However, one question is whether analyzing on a phylogenetic network is better than analyzing on a bifurcating phylogeny. Thus, I would like to see simulations where the data are analyzed twice—once using the proposed method and once using the true phylogenetic network. If analyzing with the phylogenetic network dramatically improves, for example, estimation of rate parameters, ancestral states, and phylogenetic covariances, then the method is definitely worth using. However, if we get the same answer regardless—then there is not too much justification for going through all of this trouble. I would suggest that authors should provide more guidance to the reader for the conditions that they should use their method, and when they shouldn’t bother. I suggest below, for example, that if the particular trait is better modeled by an OU model or even a white-noise model for the non-hybrid species, then hoping to do any better with model-fitting on a phylogenetic network is likely pretty futile.**

We believe the substantially better estimates in this version (due to fixing errors and better starting points, among other improvements) suggest this package is useful at estimating these parameters. **Second, I’d like it if the authors could figure out why estimation of beta and vh can become so terrible. They state that the estimated medians are near the simulated values, but it looks from Figure 3 like they are biased (hard to tell because of the scale) and get worse with more hybrids and more taxa. This is very strange, and counter to expectation. What is happening to the model when this parameter is misestimated? It seems erroneous to me. Given how the likelihood is constructed, I can’t see how such large values of vh (30 times the variance that comes from the tree?!), would ever be favored by the model. I would’ve predicted it had something to do with identifiability, but that doesn’t seem to be the case from the likelihood surfaces presented. Nor does it really make sense looking at the VCV of the model—I can’t imagine that in any of the simulations a hybrid species (e.g. species R in the toy example) could ever be so extreme so as to merit being 30 times sigma^2.... A deeper explanation and exploration of this problem would be beneficial in the current manuscript, and it may provide a path toward modifying the model to be better behaved and providing better inference on probably the most interesting parameter in the whole model! Without an explanation, it seems to me to likely represent an optimization error (for lack of anything else to understand it).**

This is now fixed; we in the paper now also discuss the proportion of variance coming from measurement error and from Brownian motion (in one of our empirical datasets, about 50:50, which seems reasonable).

**The treatment of possible transgressive segregation seems like a reasonable first step to me, and I’m certainly not going to require a more complex model be added to the package or manuscript (especially given the poor performance estimating this parameter).**

Thank you!

**However, I think that one of the explanations for transgressive segregation is not well described by simply adding extra variance (vh) to hybrid species, and I’d like to see some discussion of this and if the authors are interested, an additional model added to the package. Specifically, complementary gene action models predict that hybrids whose parents are quite divergent will tend to have intermediate hybrids, whereas hybrids whose parents are very similar phenotypically will tend to have extreme or transgressive offspring. This is easy to see if we consider the allelic effects of 3 haploid parents with 5 loci additively affecting a trait where variation is fixed within each species:  
species A: (2,2,2,2,2) = 10 ; species B: (2,-1,-1,-2,2) = 0; species C: (-2, 1, 1, 2, -2) = 0  
Hybrids between species A and species B will be at most: (2,2,2,2,2) = 10 and at smallest (2,-1,-1,-2,2) = 0; effectively bounding them between the two species. However, hybrids between species B and species C can be (2,1,1,2,2) = 8 and at a minimum (-2,-1,-1,-2,-2) = -8; certainly enabling a burst of evolution from transgressive segregation. I wonder then, if in real biological data, vh shouldn’t be a function that is negatively related to the difference between the two parents. While I doubt such a function could be well estimated from phylogenetic data alone, it does seem to me to be possible to obtain estimates or priors from additional hybridization experiments or quantitative genetic data that researchers applying this method may have information on that would make this additional model in the software of interest.**

Thank you – we have added this idea to the discussion. We credit “a reviewer” with the idea, as it’s not our own, though we agree with it. We’re happy to credit the reviewer by name, of course, though of course they’d have agree first (since it removes anonymity).

**`Figure 3 – What is bt? I think this is meant to be Beta. Also, it is not clear, especially for vh, why there are multiple clusters. I assume this is for different parameter values (as apparent for Beta). However, for vh, these all look the same because the scale is so off from the poor estimation of some values. I recommend in both the case of Beta and vh to specify in the caption/figure the true values for each cluster to draw attention to what these different clusters mean. Also, the caption says 110 taxa with 10 hybrids. It was not clear to me whether hybrids were “added” to the number of taxa, resulting in final trees with 101, 105, or 110 taxa, or if these hybrids were included in the original 100.**

We now use a table to present this information **Line 296-308 This seems like a reasonable procedure to extract an estimate of the standard deviation of body size, not standard error unless Fishbase bases its estimates on only 1 specimen. Regardless for clarity, the authors should distinguish standard error from standard deviation. However, I agree that a standard error of 10% seems reasonable (or “maybe slightly conservative but that’s ok”).**

Fixed. **Lines 521-527 Should mention here the alternative to measurement error—biological error that doesn’t follow Brownian motion, but rather an OU model (or at the extreme, white-noise). In fact, throughout the manuscript, it seems like a reasonable first step would be for researchers to analyze whether Brownian evolution is a reasonably good model for non-hybrid species before applying this method. It makes little sense to apply the model if an OU model is superior...as the authors found for cichlids (why waste all that effort optimizing when you can’t get at vh or beta anyway given that the variation is non-phylogenetic!).**

We have added a note about the similarity of OU and substantial measurement error. We’re a bit wary about proposing a preliminary test of BM vs OU on the non-hybrid taxa – there may be issues with reduced sample size, difficulty in distinguishing BM plus fitted measurement error from OU, and so forth that make this less informative than it seems. A better approach is to add OU hybrid models as we suggest (but sadly do not yet implement) later. Doing an OU1 tree scaling, and scaling the times of gene flows appropriately, is straightforward. We’re less sure about how this would affect covariance on the hybrid network, but this could be a rich area for future work. **Lines 540-542 Here is where the authors should discuss the other advantage of embedding this in a Bayesian context—some of these parameters could potentially be informed by experimental hybridization studies and genetic studies.**

Excellent point. Added. **Lines 547-551. This discussion of protracted hybridization makes me think of another interesting biological prediction from simple models of hybridization—the possibility that vh is negative. Long prolonged gene flow may have the effect of slowing divergence substantially between species and their hybrids.**

Negative vh was not something we have considered. After some thought, it may be more likely that such an effect is best modeled by a change in sigma-h, the rate of evolution after hybridization. We have added a note about the potential effect of long term hybridization on divergence. **Minor comments:   
Abstract, ln 19-20 “as well as…, as well as...”**

Fixed **ln 108-109 “it can look like flow forward in time” I don’t know what “flow forward in time” looks like.**

Fixed **Ln 112  “Under Brownian motion model” missing “a”.  Also, perhaps consider specifying what specific assumptions are in this model, since all of these are BM models. “Under a Brownian motion models without bursts or transgressive segregation...”**

We have specified that it is a time and taxon homogenous model. (Readers likely think Brownian models in general have this homogeneity, rather than bursts, etc., so this makes it clear that it’s canonical BM) **Ln 122 “trait values in species D” is better than “trait values at species D”**

Fixed **Ln 151 “calculated” not “calculate”**

Fixed **Ln 206 “The approach that seemed” not “we”**Fixed

**Ln 477-478  “suggesting that hybrid species may not statistically significantly have higher success rates as seedlings under drought conditions” is a confusing and convoluted sentence. I suggest “hybrid species may have higher success rates...but this result is not statistically significant” or something like that.**

Fixed **Ln 480 “Given the tree height...to be 1.1” Is this accounting for phylogenetic structure or just multiplying sigma^2 times the tree height? Either way is fine, just specify.**

Specified. **Table 1: Provide units  
  
Supplementary material seems admirably complete. Perhaps a tutorial on how to run the empirical examples would be helpful though?   
  
  
  
  
  
  
  
  
  
Additional Questions:  
Directions for Reviewers: The authors will appreciate detailed comments on the manuscript. Please write comments for the authors in a separate file, numbering all items that should be addressed before the manuscript is acceptable for publication, and attach your file at the bottom of this form (if you’ve inserted comments on an electronic copy of the manuscript, please attach that file as well). Reviewers are reminded that Systematic Biology is interested in publishing well-written papers of high scientific quality and of general interest. Thus, in your review, please address both the appropriateness of the paper for the journal as well as its scientific strengths and weaknesses. Please note that our instructions for authors are available on our website,**[**systbiol.org**](http://systbiol.org/)**. Use the buttons above to access the manuscript files. The HTML and PDF buttons link to the entire manuscript. Individual submitted files (such as data files) are available under the “Supplementary Files” button. We encourage reviewers to make comments directly on an electronic copy of the ms. If you do not have software that would allow you to make comments on the pdf version, please check under “Supplementary Files” to see if a Word version is available.:   
  
Do you wish to remain anonymous?: Yes  
  
How significant is this work?: Very  
  
Is the author aware of the background and source material to the problems set forth?: Yes  
  
Are the conclusions justified by the evidence presented and the assumptions involved?: Yes  
  
Are the illustrations and tables clear and understandable?: Yes  
  
In number are they: Sufficient  
  
The following people would be appropriate to review this MS further:**