

I have now assessed the revised version of the manuscript by Pirie and colleagues on the historical biogeography of the genus *Erica* in the Afrotemperate region. As stated in my original review, the writing of the ms is clear and I appreciate the robust hypothesis-testing setting of the study. I also think that the question(s) tackled here is important, especially in the era of statistical historical biogeography.

It is pleasing to see that the authors have taken into consideration and answered most of the concerns I had raised in my initial review, and I think that after these clarifications the manuscript would be publishable as is. I still have some comments that might help improve it further in my opinion.

In the Results section I still get the feeling that the two types of biogeographic hypotheses (distance/climate vs. specific biogeographic scenarios) are presented together, which might be a bit difficult for the readers to follow. The authors might want to improve that a bit still. The distinction between both kinds of biogeographic explanations is however made much clearer in the Discussion, which is very good and which I think improves a lot the clarity of the paper.

While I am still doubtful about the performance of '+j' models in Biogeobears (perhaps I need to see Matzke's rebuttal first so that I understand his line of defense), the authors' response to my critique of their use of these models is satisfying. They are taking some kind of a risk here, but they've investigated the question deeply and are aware of possible consequences. I appreciate the fact that they deal with this issue directly in the text and I also must say that the fact that most speciation events inferred are NOT founder event speciation makes me trust their results a lot: it is good choice to write it explicitly now in the revised version of the MS.

I would still recommend that the authors give the units of parameters j and d from the DEC+J model. I was asking for this in my first review but did not get an answer on this point. The parameter d is a probability of dispersal per unit time (Myrs probably here) and this is why I was asking about the typical branch length in the *Erica* phylogeny: since j is a probability of dispersal per node, its value would be better compared to $d \times \text{typical_Branch_length}$ I think.

Below are three more specific and minor comments, with line numbers referring to the tracked .docx file.

L. 78-79 : shall you briefly describe what both of these scenarios are for readers who are unfamiliar with them?

L. 96-100: I still find that the description of the 'niche similarity' hypothesis for dispersal is too short and vague, even though one sentence describing it has been added compared to the previous version of the ms.

L. 177: you should say here what is the region covered by PRECIS, most readers won't know it.

I hope that the authors will find these last comments helpful and I wish them good luck with their manuscript, which definitely deserves attention from the community.

Yours sincerely,

Florian C. Boucher