

The manuscript by Pirie and colleagues investigates the drivers of biogeographic movements in the large genus *Erica*. Using an already published phylogeny, the authors compare different hypotheses to explain dispersal scenarios in the genus across Africa.

The writing is clear and I appreciate the robust hypothesis-testing setting of the study. The article brings some interesting answers and I am confident that its main results hold true. However, I have some general comments that should help improve the study:

Something that is a bit unclear to me is that the five biogeographic scenarios (pictured in Figure 1) are not all on the same level. Two of them are general hypotheses to explain dispersal: distance vs. environment ; the three others are classic biogeographic hypotheses for the afrotemperate flora. Both classes of hypotheses are indeed treated differently in the Discussion, which is good, but more they are presented on the exact same level and statistically compared in the Results. I think the whole manuscript would largely benefit from clearly separating these two sets of hypotheses, or trying to integrate them.

The comparison of distance vs. climatic similarity in explaining dispersal probability is important and most welcome. However, I am wondering how much the climates of these different areas occupied by *Erica* differ compared to other climates between these areas. In my opinion it would be interesting to measure and report this: perhaps the differences in climate between these areas are minute compared to the climate throughout Africa and the Mediterranean.

I would also suggest that beyond these two alternative hypotheses the authors could include a hybrid one that makes much biological sense: combining climatic similarity and geographic distance into a single measure. Indeed, it seems to me that beyond contrasting these two alternatives, the Discussion suggests that both have played an important role (e.g. second paragraph of the Discussion). This could be done by creating some kind of resistance matrices...

Finally, I am concerned with all BioGeoBEARS analyses. The recent paper by Ree & Sanmartin (2018) that the authors cite clearly explains why we should not use '+J' models in BioGeoBEARS anymore. The present analysis is no exception, with the DEC+J model getting the best fit. By the way, stating that estimates of j were always lower than estimates of d needs clarification: what is the time unit in which d is expressed? How does this compare to the typical length of a branch? What is the 'effective d ' when you take into account the dispersal multiplier? The authors are probably not in a case where dispersal events are much more likely at nodes than on branches since in their stochastic mappings most speciation events occur within areas, but this is a general issue to look at. In summary, I think +J models should be removed from model comparisons.

In addition, I have some minor comments, either technical or typos:

Introduction, general : the two main alternatives for dispersal that the article proposes to test are interesting and important, but I found that while geographic proximity is well presented, there are too few details of the impact of niche similarity on dispersal. I suggest that the authors should beef up this part of their argument, possibly by giving some examples.

Page 2, last paragraph : 'This species-poor assemblage MOST LIKELY represents the ancestral area...'

Page 2, last paragraph : it is important to present the differences in the ecologies of *Erica* species, but it would be interesting as well to present their similarities, like their tolerance of poor/acidic soils if I'm not mistaken

Page 3, first paragraph : Why not having a more general argument and say that ‘ORGANISMS adapted to different habitats respond differently...’? A very convincing (plant) example of this has been published in Massatti & Knowles, *Molecular Ecology* (2016) 25, 3974–3986. There are other good phylogeographic examples from Lacey Knowles’ work on mammals or a recent paper by Burbrink et al. *Ecology Letters*, (2016) doi: 10.1111/ele.12695.

Page 3, second paragraph : shall you remove the second ‘so’ in ‘, so much so that clades...’?

Materials & Methods/Phylogenetic hypothesis : I understand that trees were not built for this article, but the authors still need to provide some basic details on how these trees were obtained. Which kind of tree inference was used : concatenation ? Species tree building ? How were the phylogenies dated ?

Materials & Methods/Defining ... : when curating GPS data, did you also collapse records that fell in the same pixel of your GPS layers ? I would consider them as duplicates for the purpose of these analyses.

Materials & Methods/Defining ... : I did not get which treatment was applied to species represented by a single record, i.e. the sentence starting by ‘However, since we...’

Materials & Methods/Defining ... : why do you use an altitude difference of 50m in your buffer ? Please specify.

Materials & Methods/Defining ... : I am not convinced by the linear transformation of distance into dispersal probability. A linear relationship would imply that the dispersal kernel is a triangle, while kernels often look more Gaussian (with a heavy tail). The authors need to address this issue, perhaps using an exponential transformation. BiogeoBEARS could do this if the ‘w’ parameter had not been fixed to 1 but freely estimated. I know from my own experience that this parameter is difficult to estimate, as stated by the authors, but there are solutions for this: the authors could try (i) a range of c. 5 fixed values or (ii) a range of different ML runs with ‘w’ as a free parameter but starting from different initial points for the optimization. In both cases, the run with the highest likelihood should be reported – the best case would be if two or more runs converge on the same (highest) likelihood.

Materials & Methods/Defining ... : Did you use the two first axes of the PCA only? This is fine but should be stated in the main text I think.

Materials & Methods/Ancestral ... : Again, which kind of ‘best tree’ did the authors use? The authors should explain if this is an MCC, a consensus, the ML tree dated using which method?

Materials & Methods/Ancestral ... : Here, the rationale for choosing 9 bootstrap trees should be explained I think. Appendix 4 gives the topologies, but branch lengths and thus divergence times might also differ between bootstrap trees, which will influence inference using biogeographic models. Since no details on this selection are given it is difficult to comment, but I suppose these trees were randomly selected, which is the best option. Why choosing a number of nine only then?

Materials & Methods/Ancestral ... : My apologies if I’m wrong but it seems that ‘the Drakensberg melting pot hypothesis’ has not been presented in the text at this point of the paper.

Results/last paragraph: when presenting the BSM results (which were run 50 times) I assume that 97% of within-area speciation events is the average you obtained across stochastic mappings. I

would recommend that you present the uncertainty around this estimate: what were the minimum and maximum frequencies of within-area speciation events across mappings.

Discussion: the first sentence of the discussion states that Erica is a model for other African plant groups but later on differences in dispersal patterns between Erica and other plant groups (the study of Galley et al.) are discussed. I think this should be harmonized.

Discussion: when the authors discuss the fact that ‘ecological distances’ are only calculated based on current ecological conditions, they could evoke the fact that exactly the same critique has been addressed to landscape genetics (Bolliger et al. 2014 Landscape Ecol., Epps & Keyghobadi 2015 Mol. Ecol.).

Figure 1: why are Ethiopian mountains not represented in these scenarios? It is especially troubling when one is looking at the ‘Cape to Cairo’ hypothesis but none of the migration routes goes through North-Eastern Africa.

References: be careful to write ‘Désamoré’, not ‘Désamore’