

Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotemperate by *Erica*

Our apologies that the revised version of this ms. has been so long in coming. We are very grateful for the constructive reviews, and have gone to considerable length to deal with all the comments, either incorporating the suggestions directly, or otherwise addressing the underlying issues by further clarification in the text.

I have uploaded a new version to bioRxiv, but to make it easier to see what we've done I will provide an additional tracked-changes version of the text, as well as this file with responses. The tracked-changes version will be a result of a compare/merge of submitted and revised versions, because tracking as we revised was getting too complicated to follow. As a result, there are big tracts of red text that have been replaced entirely, even though they were only modified. I hope this is still useful, and would add that this effect makes it look as though the revisions were far greater in effect than they really were: the main sweep is the same; the conclusions and abstract more or less identical. The devil was in the detail.

Responses to the comments follow below, inserted into the complete decision letter.

Thanks again for your time and efforts with our paper.

On behalf of the authors,
Michael Pirie

Needs a revision

All the reviewers now chimed in with their opinions and agreed on the interests of the manuscript. I commend the authors for all the work they have done, and I echo the reviewers. I think this study represents a nice piece of work investigating the factors and processes mediating dispersal of plant clades from Europe to Africa. I very much appreciated the thoroughness of the analyses and the clarity of the text.

I recommend the authors undertake a thorough revision based on the constructive comments of the reviewers, taking particular care to address the reviewer's methodological concerns. **I found particularly interesting the criticism of one reviewer regarding the conceptual problem of proposing hypotheses that are not mutually exclusive (or framed at different levels), as well as all methodological suggestions proposed by the three reviewers.**

I am, however, a bit more cautious than the reviewers regarding the results of this study. **In agreement with one reviewer, I regret the choice of biogeographic models. The DEC+J model has been shown at best to not be directly comparable to the DEC model, and at worst to present some statistical problems (Ree & Sanmartin 2018). I would recommend excluding this model from the comparisons and if the authors still wanting to present it, to do it on supplementary material.**

See detailed response below.

I am also suspicious with the DEC+* model of Massana & al. (2015). This model has been published as a preprint with no reviewer assessment on their quality/performance. I thus suggest excluding this model from the comparisons as well.

We have removed the model.

In addition, I have concerns on the validity of the results. In addition to the different dispersal matrices, you implemented an **adjacent area matrix to constraint the maximum number of areas allowed as ancestral states. I agree on this procedure to decrease model uncertainty, for example as you did by using the “maximum number of areas” command. However, I think the constrictions on the adjacent matrix you implemented are problematic and could have affected the results: 1) my apologies if I’m wrong, but I think that by using this adjacency matrix you force stepping stone dispersal to occur, while this is one of the main hypotheses you want to test.** For example, in the matrix on Appendix 3 you are impeding dispersal between Europe and the Cape region, the Drakensberg and Madagascar, while only allowing dispersal through tropical Africa. It is thus **not surprising that you found support for the “Drakensberg melting pot” stepping stone scenario in comparison with other long distance dispersal models. 2) In addition, by implementing this matrix you automatically disallow disjoint distributions at ancestral nodes, decreasing the likelihood of extinction to occur.**

We would argue that it is not biologically realistic to consider rare LDD events as sufficient to maintain gene flow between widely disjunct populations, and therefore that such reconstructions at ancestral nodes should ideally be explicitly disallowed. That said, this indeed cannot work under DEC without dictating stepping stone dispersal. Since the comparison of DEC and DEC+J is also controversial, to address both issues we have run additional analyses, comparing the results obtained under both DEC+J and DEC, and clarified the text.

Methods:

“Prior to comparing the different biogeographic hypotheses we tested whether an unconstrained model fitted the data better than a) restricting the maximum number of areas at nodes to two; and/or b) implementing an adjacent area matrix (Appendix 3; Results)..”

Results:

“Under DEC+J, models including an adjacent area matrix fitted the data better than those without constraint to dispersal. We additionally fixed the maximum number of ancestral areas to two, increasing the speed of the analyses without negatively impacting model fit. Under DEC, models with maximum areas at nodes restricted to two fitted the data better than those without constraint to ancestral ranges.”

An important part of the manuscript focuses on whether colonization was mediated by niche changes or occurred across similar habitats. On this regard one reviewer had concerns about the areas used for comparison. **I agree with him that differences between study areas could disappear when compared with other regions in Africa or Europe where Erica does not occur. I additionally regret you didn’t differentiate the Northern**

Hemisphere Mediterranean region from other Northern Hemisphere regions. My guess is that the low climatic similarity between southern African and Europe might most likely apply to the Eurosiberian region, but not to the Mediterranean one. I think this differentiation is important to test if long distance dispersals involved niche shifts.

Yes, the Mediterranean climate may be more similar to the South African than the Central European climate is. However, 1) we do not know the ancestral area of the African clade within Europe; 2) we are not comparing the climates of entire regions, we are estimating climatic similarity between ranges based on distribution data (whilst correcting for regional differences in 'available' climates); and 3) much wider comparison to areas that do not support *Erica* at all does not seem relevant to us.

Concerning model comparisons, in addition to AIC scores (e.g. on Appendix 8), I would like to see the differences in deltaAIC values, akaïke weights or any other metric that allows to evaluate model improvements and perform model choice. Generally, it is the differences between the likelihoods or AICs that matter, not their absolute values. That is, the larger difference in AIC indicates stronger evidence for one model over the other (Burnham and Anderson, 2002). Delta (AIC differences) within 0-2 has a substantial support for a suboptimal model; delta within 4-7 considerably less support and delta greater than 10 essentially no support.

This was an excellent suggestion. We have replaced the table and updated the appendices etc., adding deltaAIC values and all models scoring within deltaAIC of 2 to the table - also for the bootstrap trees - and deltaAIC values both overall and within stepping stone and distance model comparisons. Along with including more of the DEC analyses for comparison (see below), this represented rather a lot of additional results, and required us to pretty comprehensively re-write the results and parts of the methods, but the discussion and overall conclusions of the work were robust to the changes and required rather less adapting.

To conclude, apart from the nexus file I would suggest the author to include a figure with the most likely biogeographic reconstruction, maybe on SI.

Done: Appendix 13.

I leave this, and the reviewer's comments, for the authors to consider as they revise and improve the paper. I hope the authors will find that many of these will be helpful in improving the manuscript

Reviews

Reviewed by anonymous reviewer, 2018-04-24 15:10

This is an interesting manuscript exploring the interactions between geographical distance and ecological niche using the genus *Erica* as a model. The manuscript uses species occurrences and model testing to explore different biogeographic hypotheses. Although the manuscript does not introduce a novel idea, in general it provides new evidences regarding the colonization of new areas with subsequent niche change, and brings new evidences in

terms of historical African biogeography and the Erica genus. However, in my opinion, the manuscript would benefit by making some clarifications, highlighting better the hypotheses and the general argument. Next time I advise to introduce line numbers so that it is more comfortable to carry out the revision.

Apologies – done.

The study has a strong background in testing models. Although there are certain aspects of the models that I can not judge since they are outside of my expertise (I recommend that another reviewer or editor assess the robustness of the models), in their current state the models are not sufficiently clear to be reproducible. In the material and methods section there are several points where it is not clear which tools you are using to build the models (e.g. pg.4, paragraph "To incorporate in a solely distance-based biogeographic model"). It is necessary to clarify whether you have used a statistical program or you have programmed bioinformatic scripts. On the other hand, if any script have been programmed, it would be necessary to reference them in the text in order to have access to them and clarify the reproducibility of the study, since in its current state it is not sufficiently clear.

We are grateful for the review and have tried to re-write the methods in a way that it is easily understandable and reproducible, including additional references e.g. for calculation of distances etc.

For the section in question, it now reads:

"To incorporate a measure of geographic proximity among areas in a solely distance-based biogeographic model (the 'geographic distance' model; Fig. 1), we calculated the overall minimum geographic pairwise distances between the area ranges according to Meeus (1999) in WGS84 projection using the raster 2.3-33 package (Hijmans, 2015) in R (R Development Core Team, 2013). We converted geographic distances into dispersal rate multipliers as probabilities (0-1, whereby the largest distance has the smallest dispersal probability), and as distances that we scaled linearly (model intercept of 1 and a slope of -1.52^{-07}) and exponentially (-0.25, -1 and -2). "

pg. 3 - "which might apply to arid adapted plant groups for which past distributions have been more contiguous (Bellstedt et al., 2012)." Here I would recommend to give credits to recent studies of African groups, both arid and subtropical, which have provided new evidence regarding continuous past distributions in Africa. pg. 3 -"such as the more mesic temperate or alpine-like habitats of the "sky islands" of East Africa (Gehrke & Linder, 2009; Gizaw et al., 2013, 2016)." You are missing the relationships of the African continent with Macaronesia and I would recommend to introduce this concept in the text.

We have modified these sections with the addition of the Macaronesia theme and references.

Modified text: "Organisms adapted to different habitats respond differently to changing environmental conditions (Mairal, Sanmartín & Pellissier, 2017; Chala et al., 2017). Distribution patterns of arid-adapted plant groups, for which suitable habitats in Africa have been more contiguously distributed (Bellstedt et al., 2012), might thus be best described by biogeographic scenarios emphasising vicariance processes, such as for example the "Rand Flora" (Sanmartín et al., 2010; Pokorný et al., 2015), or the "African arid corridor" hypothesis (Verdcourt, 1969; White, 1983). Models that invoke concerted patterns of LDD might instead apply to plants adapted to habitats that remained largely isolated over time (Knox & Palmer, 1998; Galbany-Casals et al., 2014; Nürk et al., 2015; Míguez et al., 2017). Examples include the shared arid adapted elements of Macaronesia and adjacent North-West Africa and

Mediterranean (Kim et al., 2008; Fernández-Palacios et al., 2011; García-Aloy et al., 2017), and the more mesic temperate or alpine-like habitats of the “sky islands” of East Africa, in which, for example, multiple lineages originated from northern temperate environments (Gehrke & Linder, 2009; Gizaw et al., 2013, 2016).”

pg. 3 "One such scenario, inferred from Cape clades with distributions very similar to that of *Erica* involves dispersal north from the Cape to the East African mountains via the Drakensberg (“Cape to Cairo”; Galley & al., 2007). McGuire & Kron (2005) proposed a different scenario for *Erica*: southerly stepping stone dispersal." This is not written clearly enough for a reader not specialized in African biogeography. I understand what you mean, but it should be explained more clearly.

Changed to: “A more specific biogeographic scenario, inferred from Cape clades with distributions very similar to that of *Erica*, involves dispersal north from the Cape to the East African mountains via the Drakensberg (“Cape to Cairo”; Galley & al., 2007). McGuire & Kron (2005) proposed a different scenario for *Erica* instead: southerly stepping stone dispersal through the African high mountains to the Cape.”

pg. 3 - "clades of different ages (Pokorny et al., 2015) and/or origins, but with similar ecological tolerances, might show convergence to similar distribution patterns (Gizaw et al., 2016)." Here you reference to the main idea of the manuscript. However, previous work on this idea is not clearly introduced or disregarded. This was already explored in the manuscript of Mairal et al. 2017 in Journal of Biogeography, although you give credit to this manuscript elsewhere in the text, I miss that you introduce this idea with more details and you clearly establish a hypothesis.

Reference added in this revised section:

“The interplay between geographic distance and ecological suitability may be a decisive factor in the biogeographic history of plants (Donoghue, 2008; Carvajal-Endara et al., 2017), so much so that clades with similar ecological tolerances and origin might show convergence to similar distribution patterns (Gizaw et al., 2016; Mairal, Sanmartín & Pellissier, 2017)”

Pg. 3 - "we test five biogeographic hypotheses" - this is clear in the figure 1, however in the text you only refer to 4 hypotheses, please clarify.

We have re-written that part, it now states:

“Specifically, we test six biogeographic models, as illustrated in Fig. 1: Three that test the influence of geographic distance, climatic niche similarity, and the combination of both; and three differing stepping stone models that each imply geographical distance effects promoting dispersal predominantly between adjacent areas: northerly “Cape to Cairo”, “Southerly stepping stone” and a model that invokes elements of both, the “Drakenberg melting pot” hypothesis.”

It seems worrisome that within your hypotheses (figure 1) you have not included clearly the mountains of Eastern Africa (e.g. Harar plateau, Abyssinian plateau, Gregorian Rift ...). In these areas the genus *Erica* is highly diversified within each sky-island and these areas have served as stepping-stones for the colonization of eastern Africa from Europe and west Asia (e.g. *Lychnis* in Popp et al. 2008; *Cardueae* in Barres et al. 2013; *Hypericum* in Meseguer et al. 2013; *Canarina* in Mairal et al. 2015). Please clarify if you have had this area in consideration, and consequently, modify the figure or comment this bias clearly in the text.

Figure 1 shows the general hypotheses, without distribution data - we have the areas of the *Erica* distribution reasonably covered as illustrated in Fig. 2 - but this comment indicated

that we needed to shift the arrows to better reflect this - done.

Reviewed by Simon Joly, 2018-04-24 15:15

I enjoyed reading the manuscript of Pirie et al. entitled "Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotropics by *Erica*". It tests different hypotheses regarding the biogeography of the genus *Erica* present in Europe and in the South of Africa. Specifically, they compare previous hypotheses regarding plant dispersion with hypotheses based on distances alone or on bioclimatic niche similarity. The approach is interesting and the overall manuscript is clear and well written.

Thank you!

I see very little flaws with the manuscript, perhaps with one exception. The distance model that the author test assumes a negative linear relationship between geographical distance and dispersal probability. This seems quite inappropriate. Indeed, most studies on plant dispersion show that the relationship between distance and dispersal probability is not linear (see, for instance, Nathan 2006, *Science*; doi:10.1126/science.1124975). It is perhaps closer to an exponential function (or lognormal), where seeds have a larger probability to fall close to the plant and the probability to disperse far decreasing exponentially with distance. It seems to me that this is something that the authors should have considered to be thorough with their model testing. They could incorporate such non-linear relationships by using an exponential function with different alpha parameters to derive their dispersal probability, and check which one gives the best probability in their biogeographic model fitting. The same thing could probably be done with the niche model, although we probably know much less regarding the relationship between niche distance and dispersal probability.

We have included an exponential model for the physical distances and report the results in the MS. The models did not fit the data better. We have not done it for the environmental distances as this is, as already mentioned, not straightforward.

I also have a few minor comments:

1. In the methods, it would be nice if the authors list the genes used.

Done

2. The authors associate imprecise geographic coordinates to coordinates with less or equal to three decimals. But a coordinate could also be imprecise even with many decimal, such as when it is placed from general locality information instead of a GPS device (as when using Geolocate, for instance). Did the authors consider this type of uncertainty as well?

The cut-off at three decimals was particularly to remove centroids of QDS, which we were aware was the limitation of the precision of much of the PCECIS data. We carefully checked the other database derived occurrence data (especially data from GBIF) for accuracy by plotting occurrences on maps and

obviously erroneous locality data was removed. Over and above we did not further consider the source of or information on the precision of the geographical coordinates, because these are most often not stated in the database-derived occurrence records. Qualified in text as follows:

"We curated the species occurrence data by removing obviously erroneous locality data, duplicated records, and records with imprecise occurrence data (coordinates with ≤ 3 decimal places, many of which represented centroids of quarter degree squares which were originally represented in PRECIS), but did not further consider the source of or information on the precision of the geographical coordinates, because these are most often not stated in the database-derived occurrence records."

3. *"Prior to comparing the different biogeographic hypotheses we tested whether a model without constraint to dispersal or ancestral ranges fitted the data better than setting the adjacent area matrix and maximum areas at nodes to two (as would be implied by the present day distribution of the species, which never exceed two areas)."* It is a strange idea to test two parameters at once. Then you don't know if the difference is due to one or both.

We agree, our process was not explained well. We tested both independently. We adjusted the text accordingly.

Text modified:

"Prior to comparing the different biogeographic hypotheses we tested whether an unconstrained model fitted the data better than a) restricting the maximum number of areas at nodes to two; and/or b) implementing an adjacent area matrix (Appendix 3; Results)."

4. I feel that parsimony optimization is not necessary. But at the same time it is not problematic, so I leave the decision to keep it or not to the authors.

We would opt to keep these results, in particular in order to compare to the model-based methods (including +J) and show whether those are more or less parsimonious.

Review by Florian Boucher

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.

We have now presented the stepping stone dispersal scenarios and the distance based scenarios separately, whilst also maintaining a global model comparison. It is hopefully clear from the interpretation of the results that we do not treat these as in any way mutually exclusive.

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.

Doubtless the climatic differences are small in the context of the full range of areas that are uninhabitable for *Ericas*, but we would argue that that it is the differences between areas in which a given species can survive that is most relevant in this context.

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...

A very useful suggestion: we added an analysis combining environmental and physical distance, and the fit of this model was often good, confirming the importance of both factors.

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.

This is the point also emphasised by the handling editor, and we have gone to considerable lengths to address it in a way that will make the results and strengths of the conclusions clear, but without compromising our own principles on the matter: mostly by including more results (particularly DEC) and organising them differently instead of taking results out.

We have been in contact with Matzke since receiving these reviews, and perhaps unsurprisingly he takes issue with the conclusions of the Ree and Sanmartin paper. We're including a few detailed points at the end of this response, but do not believe that this ms. is the place to go into the detail of his argumentation; suffice to say, this is not the last word on the subject, a rebuttal is in prep., and we consider it premature to abandon either DEC+J or the model comparison approach in biogeoBEARS until the points have been properly debated.

Particularly pertinent to our results is that they do not exhibit obviously problematic phenomena identified by R&SM: Our j value for the best tree, best model, is low; it is lower than the value for d and similar to the d values for the DEC analysis; and the most common form of speciation inferred our dataset is instead within-area speciation – representing up to 97% of the events.

Clarification in text: “The vast majority of biogeographic events inferred using BSM under both DEC+J and DEC were within-area speciation (97.15 % and 96.26% respectively; Appendix 9). The values for range expansion (parameter d) were similar and low (0.0030 and 0.0027 respectively; Appendix 9). Under DEC+J, cladogenetic dispersal (parameter j) was 0.0024, i.e. lower than d and much lower than the maximum permitted value (3).”

Hence there is no indication of inappropriately high jump dispersal dominating the results. If we compare the results of DEC and DEC+J, we find that DEC is not the most parsimonious – unless you consider it more parsimonious to have widespread ancestry with extinction in one region for a lot of nodes, rather than a single jump dispersal event. The latter seems more plausible as well as more parsimonious given the rarity of widespread distributions at this scale in *Erica*. The corresponding section in the discussion is now slightly expanded:

“This suggests that species distributions were restricted throughout the evolution of the *Erica* African/Madagascan clade, and that the areas remained isolated during this period (i.e. the last c. 15 Ma; Pirie & al., 2016). We would also argue that it lends credibility to results obtained **under DEC+J, in which some range shifts were treated as cladogenetic dispersal events (instead of by inferring seemingly implausible widespread distributions)**, despite arguable drawbacks in the implementation of that model (Ree & Sanmartín, 2018)”

We would argue that DEC+J, despite potential flaws, is more biologically defensible in the case presented here, and we consider the model comparison an important element for the scientific rigour of the analyses. For those as yet unconvinced on either point, we have nevertheless reorganised our presentation of the results, including expanding reporting of DEC results, so that the strength of our conclusions can be judged given both DEC and DEC+J separately.

Further points concerning the R&SM critique:

One of the main arguments is, that the jump dispersal is time-independent and as such ignores a fundamental feature of all evolutionary models. This leads the R&S to the claim that the model should be abandoned in favour of DEC. To our knowledge not all models in evolutionary biology are time-dependent, e.g. the proportion of invariant sites parameter in DNA models. Furthermore, the jump dispersal are estimated on a speciation nodes of a dated phylogeny. As such it has a time component.

Overall we are not convinced by the arguments made by R&S using the two and four-taxon examples. Those examples seem to be not appropriate if we want to model assumptions with either 2 or 3 degrees of freedom (DEC and DEC+J respectively). Their argument that DEC+J models unparismonious pathways as shown in their Figure 3 is not convincing, as with such a small model, why should in-area speciation be more likely than range expansion. The more general and

important question should be, why a model which regularly models widespread ancestries over millions of years, as usually found with DEC models, should be more likely than a jump dispersal.

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?

This statement (expanded as above) was added to show that, contrary to the Ree and SanMartin (2018) example, j is not maximized in our DEC+J analysis. The maximum value for j in a DEC+J analysis is 3, in our best model under DEC+J, j is modelled as 0.0024. Comparison between the two best stepping-stone scenarios of DEC and DEC+J, d was estimated to be 0.003 and 0.0027 respectively, thus relatively similar values. Again, we don't want to get bogged down in this debate; hopefully in its current form anyone familiar - or not - with the issues raised in R&SM will be able to see that our analyses are not exhibiting obvious pathological behaviour.

According to our understanding, the values reported for d and j in Appendix 8 are the 'effective' values, incorporating the dispersal multipliers and 'w'. d and e depend on branch length and as such on the speciation rate, whilst j is independent of them. Appendices 10-12 present the different numbers of events as sampled in Biogeographic Stochastic Mapping. Those are mean values of events across the 50 runs. In total we found a mean value of 5.12 for range-expansion dispersal events and 3.06 jump dispersal events.

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.

The +J model is clearly controversial, but we believe that the method has merits and is worth reporting. According to the publication of Ree & SanMartin (2018), both models - DEC and DEC+J - are flawed as they assume a Yule-process and as such ignore lineage extinction and range-dependent speciation. They additionally argue in the paper, that +J is not parsimonious, but when comparing the results of the DEC and DEC+J analyses, for us it seems that DEC is not the most parsimonious (e.g. the dispersal/range extension from the base of the tree to the Cape species is modelled with DEC as E-ET-T-TC-C, while with DEC+J as E-ET-T-C). Furthermore, as above, even without model comparison between DEC and DEC+J we think that a model incorporating jump dispersal seems to make more sense regarding the distribution of *Erica*, than a pure vicariance-extinction model, as within *Erica* most species are regional endemics.

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.

Text modified as follows:

“The interplay between geographic distance and ecological suitability may be a decisive factor in the biogeographic history of plants (Donoghue, 2008; Carvajal-Endara et al., 2017), so much so that clades with similar ecological tolerances and origin might show convergence to similar distribution patterns (Merckx et al., 2015; Gizaw et al., 2016; Mairal, Sanmartín & Pellissier, 2017)”]

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

OK - done

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

Good idea - done.

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 Massatii & 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.

It seems clear that we needed to consider more relevant literature /examples here to support the general point. These particular papers may be very good, but they are arguably not the most pertinent examples - we have added reference to some additional African case studies

“Models that invoke concerted patterns of LDD might instead apply to plants adapted to habitats that remained largely isolated over time (Galbany-Casals et al., 2014; Nürk et al., 2015). Examples include the shared arid adapted elements of Macronesia and adjacent North-West Africa and Mediterranean (Kim et al., 2008; Fernández-Palacios et al., 2011; García-Aloy et al., 2017), and the more mesic temperate or alpine-like habitats of the “sky islands” of East Africa, in which, for example, multiple lineages originated from northern temperate environments (Gehrke & Linder, 2009; Gizaw et al., 2013, 2016).”

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

MP: This is English English and I’m clinging on to it.

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 ?

Done.

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.

Agree, records that fall into the same pixel of the GPS layer should be considered as duplicates when working with density estimates. Since we used the species occurrences only for defining the geographic range of the species (what we term 'area ranges') and only used the spatial points within these area ranges in the subsequent analyses, species records that fall into the same pixel of the GPS layer are already corrected for. This was not clearly stated, but has now been clarified in the text:

"We assessed the area ranges around the occurrences of all species in a biogeographic area using an arbitrarily set buffer of c. 1° lat/lon in radius and 50 m elevation around the individual occurrences (Europe 4667, Tropical Africa 42, Madagascar 70, Drakensberg 58, and Cape 1981 occurrences; Appendix 2) conservatively aiming at a representative approximation of spatial extent and ecological conditions of the species' distribution in a respective biogeographic area. Solely these area ranges, which include up to several thousands of spatial points, were used in the subsequent analyses to calculate geographical and ecological distances."

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...'

See answer above (the sentence has been reformulated to clarify this issue).

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

See answer above (the buffer was arbitrarily chosen, while considering Erica ecology. Sentence has been reformulated to clarify this issue).

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.

We have added some analyses with the physical distances transformed exponentially with different values. They are never better than the one with linear distances.

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.

Has been clarified: "Finally, the biogeographic model for the niche similarity hypothesis was defined using the pairwise Schoener's *D* values for the combined PCA axes 1 and 2 directly as dispersal rate multipliers between areas (for details see protocol in Appendix 2)." And in Appendix 2: "The first two principal components (PC axes) were selected for the subsequent analysis based on the broken-stick criterion (Jackson 1993)."

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?

Done

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?

We selected the 9 bootstrap trees according to topology. Most nodes that are relevant for the biogeographic reconstructions are well supported (most major clades), at nodes where this was not the case, we allowed the possible alternatives. This results in 9 different topologies, including the one also represented by the best tree. The actual tree with a certain topology was then taken by random and scaled to be ultrametric. We have added this more detailed information to the MS.

"These trees were selected to represent the possible resolutions of phylogenetic uncertainty between the geographically restricted major clades (Appendix 4), but were otherwise chosen randomly with respect to topologies and branch lengths."

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.

Yes, thanks for pointing that out. We have added the information when we introduce the three stepping stone hypotheses in the introduction.

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.

We added standard deviation to the text and Appendices.

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.

Hopefully addressed “we modelled shifts between biomes and dispersals over larger distances in the evolution of *Erica*, in order to test six hypotheses for the origins of Afrotemperate plant groups”

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.*).

Landscape ecology is certainly an interesting looking and potentially very relevant general topic, but in my superficial reading it seems to play out at more recent timescales and in a somewhat different (human influenced?) context - without really digging in (and we’re late enough with this revised ms.) I’m not really comfortable with introducing it here.

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.

See response above - arrows moved.

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

Done