I have decided not sending the paper again to reviewers in order to speed up the process, but a few important modifications are still necessary before the paper could be recommended.

Regarding the biogeographic models: I found satisfying that the results of the unconstrained model are presented together with the results in which jump dispersal was forced to occur between Europe and the Cape region. Although I am still sceptical with this assumption, I agree this approach could be valid if we assume that widespread ancestral reconstructions are biologically unlikely. I regret, however, the authors don't provide a more detailed justification of this assumption; if they argue in the introduction that ancestral continuous (widespread) distributions fragmented by vicariance probably explain disjoint patterns of (arid-adapted) African plant groups (e.g., Sanmartín et al., 2010; Pokorny et al., 2015; Bellstedt et al., 2012), why in this particular case this biogeographic scenario (an ancestral widespread distribution) does not make biological sense? I imagine it has something to do with the temperate adaptations of the genus and the fact that temperate conditions were never continuous in Africa, but this needs to be clarified (maybe in Appendix 3 or in the main text?).

Reviewer 3 was concerned for the potential correlation of the studied variables (i.e. geographic distances and niche similarity). In this new version, the authors evaluated their correlation and found what is generally considered a moderate to strong value of correlation (Kendall’s R = -0.64), but was this result taken into account at all? It has strong implications, as one of the main questions of the study (i.e. the relative importance of niche vs. distance in dispersal patterns) cannot be answered. I am convinced this limitation does not affect the main conclusions of the study since the preferred model was the combined “niche/distance” model. In addition, this is an interesting result “per se”. However, in light of this, I have the impression that comparisons of the “geographic distance” and the “niche similarity” models do not make sense anymore because these variables are not independent. Only comparisons of the null and the combined “niche/distance” model might remain informative. I leave to the authors the decision on whether they want to exclude these models from the model comparison table or not, but this limitation needs to be taken into account and the results/discussion sections modified accordingly.

Again, I hope you will find these last comments helpful and look forward to read the final version of the manuscript!

Best
Andrea