

# 1 Sensitivity of biogeographic 2 reconstructions to the use of differential 3 extinction rates

4 Jose L. Pinzón<sup>1</sup>

5 <sup>1</sup>Escuela de Biología, Universidad Industrial de Santander, Bucaramanga, Colombia.

6 Corresponding author:

7 Jose L. Pinzón<sup>1</sup>

8 Email address: jose.pinzon4@correo.uis.edu.co

## 9 ABSTRACT

10 Effects of differential extinction rates remain being an issue in biogeographic and evolutionary studies.  
11 Here, I use empirical examples and simulated datasets to asses how the specification of different  
12 extinction rates influences ancestral range estimation in historical biogeography. The results showed  
13 that variations in scale and asymmetry of extinction rates may have notorious effects in the accuracy of  
14 biogeographic inferences, specially when the rates of extinction are high. Further work may explore the  
15 behavior of current statistical methods of biogeographic inference with different estimates of extinction  
16 based on novel developments in this field.

## 17 INTRODUCTION

18 Statistical methods to ancestral range reconstructions are mainly focused in the estimation of two types  
19 of parameters: dispersal and extinction rates. Estimation of extinction rates is difficult, most of the  
20 current approaches rely on information about the timing of splits and branching patterns of molecular  
21 phylogenies (Sanmartín and Meseguer, 2016). A common assumption since the initial formulation of  
22 the Dispersal-Extinction-Cladogenesis by Ree et al. (2005) is that extinction rates are constant through  
23 lineages and areas (Ronquist and Sanmartín, 2011). More realistic scenarios may include different rates  
24 of speciation and extinction per region as implemented in the GeoSSE model (Goldberg et al., 2011) or  
25 the ideas of Webb and Ree (2012) inspired in the theory of island biogeography (MacArthur and Wilson,  
26 1967). However, these scenarios are not currently implemented in the most widely used packages for  
27 biogeographic inference.

28 While attempts to elucidate how uncertainty in dispersal routes and phylogenetic inference may alter  
29 reconstruction of biogeographic histories have been made (Nylander et al., 2008) (Chacón and Renner,  
30 2014), effects of the use of different extinction rates to this kind of data are limited. Due to these reasons,  
31 I test the accuracy of statistical methods to infer biogeographic reconstructions to varying extinction  
32 rates.

## 33 METHODS

34 I explore how using differential extinction rates may influence the ability of current methods to infer the  
35 ancestral area of lineages. This was done by two ways: empirical and simulated datasets.

36 Following a basic premise of island biogeography and former implemented in the software SHIBA  
37 (Webb and Ree, 2012), I use area size as a proxy for extinction rates. Although the use of geographical  
38 size may be inadequate, I use it for illustrative purposes.

39 As empirical example, I used the Indian Ocean Primulaceae dataset of Strijk et al. (2014), particularly  
40 the subclade inhabiting Madagascar and the Mascarene islands of Mauritius, Reunion and Rodrigues. This  
41 example was chosen due its particularity: each of these species is single-island endemic, and differences  
42 in island size are notorious.

43 Simulated data were generated in the R package diversitree (FitzJohn, 2012) as follow. First, I fitted  
44 the parameters of extinction, anagenetic change and extinction for the Primulaceae data to the MuSSE  
45 model (FitzJohn, 2012). The use of MuSSE rather than GeoSSE was preferred due the especial conditions  
46 mentioned above to match the number of possible states (4 in the empirical example). The root of each  
47 simulated scenario was assigned to state = 1, and 10 simulated datasets were generated with a limit of 30  
48 extant taxa to the present.

49 Ancestral range reconstruction analyses were carried in the Dispersal-Extinction-Cladogenesis model  
50 in a bayesian framework as implemented in Revbayes 1.0.1 (Höhna et al., 2016) in four different scenarios  
51 (See below). Maximum range size at the root of the trees was set to two in both empirical and simulated  
52 data. For simplicity reasons, I don't take into account temporal constraints of dispersal between regions,  
53 unlike the original study of Strijk et al. (2014).

54 Reconstructions Scenarios. First, extinction rates were held constant along the areas with a extinction  
55 rate of 1 (equalratesHigh); second, each area with its own extinction rate inversely proportional to the  
56 logarithm of the size of each island (diffratesHigh). More moderate scenarios were modeled scaling the  
57 previous two scenarios by a factor of 0.1, equalratesLow and diffratesLow, respectively.

58 Finally, I evaluated the behavior of the biogeographic reconstructions measuring the relative likelihood  
59 of the true state at the root of the empirical data and each of the simulated histories.

60 Figures were drawn with the R packages in ggplot2 (Wickham, 2009)

## 61 RESULTS AND DISCUSSION

### 62 Indian Ocean Primulaceae biogeography

63 Frequencies at the root of the tree in the case of family Primulaceae for each scenario were above 0.9 (See  
64 Table 1). Like the study of Strijk et al. (2014), most likely ancestral state was Madagascar. Likelihood  
65 values of the most frequent state at the root of the tree show a moderate decrease with augments in both  
66 asymmetry and higher extinction rates.

### 67 Simulated datasets

68 Results of simulated datasets yielded similar outcomes to the empirical examples, but severely marked.  
69 Figure 1 shows the effects of the use of differential extinction rates in biogeographic reconstruction. It's  
70 worth noting that relative likelihood of true state is high in both scenarios with low rates of extinction. A  
71 great variation is found in the inferred ancestral state when higher extinction rates are used.

72 Effects of varying dispersal rates shown that sensitivity of reconstruction of DEC is high. In a study  
73 with the plant family Hydatellaceae, Iles et al. (2014) conclude that the higher the extinction rates, the  
74 less accurate are the reconstructions in terms of the likelihood of the ancestral state, this is seen not only  
75 at the root, but also and each of the nodes of the tree. Similarities of their results to those here obtained,  
76 support the idea that confidence of statistical methods to reconstruct biogeographic histories are prone to  
77 varying extinction rates, even when extinction rates are high but not necessarily asymmetric.

78 Methodological issues, including the logic for the use of size as an initial estimate of extinction rates  
79 may be bias the results (See (Webb and Ree, 2012) for reasons). However, it seems more or less clear that  
80 a common pattern is found in this and other studies. The limited size of the simulated datasets doesn't  
81 allow to formulate proper conclusions, but it can be seen as an attempt to elucidate how ancestral range  
82 reconstruction are prone to the use of asymmetric and high spatial extinction rates.

## 83 CONCLUSIONS

84 Here, I have shown through empirical and simulated data how the use of different spatial extinction rates  
85 may influence the results of ancestral range reconstruction yielding to lower accuracy. Further work may  
86 explore more propitious estimates of extinction rates, including these from recent advances that take into  
87 account the fossil record and allow variation of rates through time.

## 88 SUPPLEMENTAL DATA

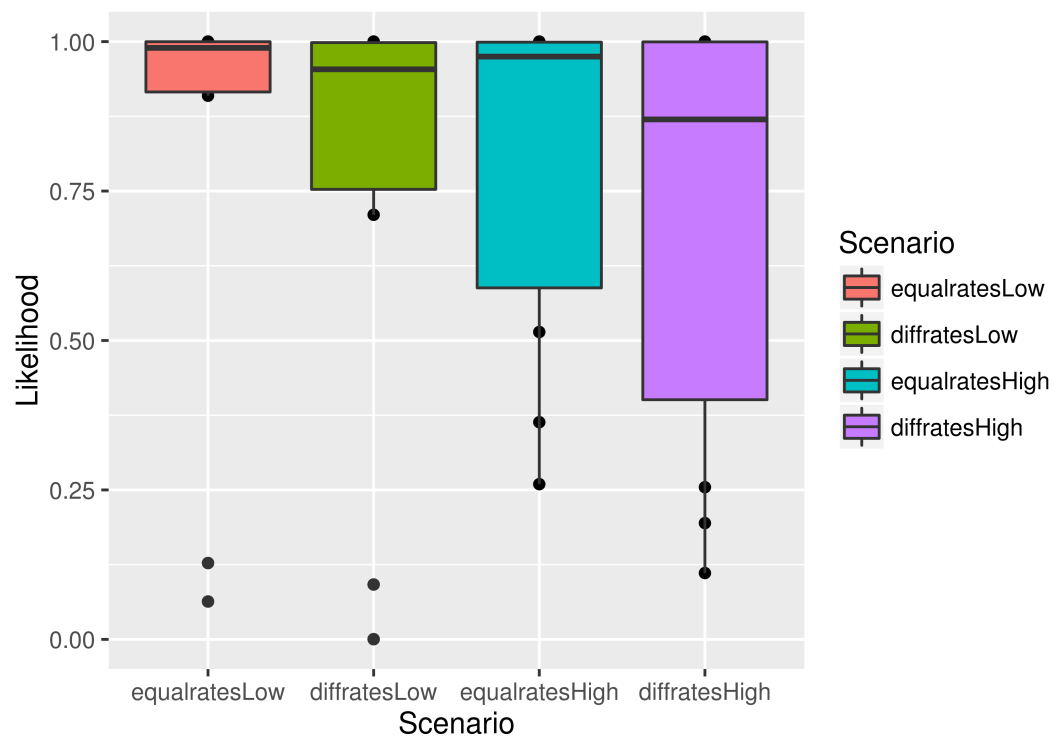
89 Supplementary results, data and scripts used to generate simulations are available at [http://github.com/jessop/Extinction-](http://github.com/jessop/Extinction-rates)  
90 rates.

91 **REFERENCES**

- 92 Chacón, J. and Renner, S. S. (2014). Assessing model sensitivity in ancestral area reconstruction using  
93 lagrange: a case study using the colchicaceae family. *Journal of biogeography*, 41(7):1414–1427.
- 94 FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in r. *Methods in*  
95 *Ecology and Evolution*, 3(6):1084–1092.
- 96 Goldberg, E. E., Lancaster, L. T., and Ree, R. H. (2011). Phylogenetic inference of reciprocal effects  
97 between geographic range evolution and diversification. *Systematic Biology*, 60(4):451–465.
- 98 Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., Huelsenbeck, J. P.,  
99 and Ronquist, F. (2016). Revbayes: Bayesian phylogenetic inference using graphical models and an  
100 interactive model-specification language. *Systematic Biology*, 65(4):726–736.
- 101 Iles, W. J., Lee, C., Sokoloff, D. D., Remizowa, M. V., Yadav, S. R., Barrett, M. D., Barrett, R. L.,  
102 Macfarlane, T. D., Rudall, P. J., and Graham, S. W. (2014). Reconstructing the age and historical  
103 biogeography of the ancient flowering-plant family hydatellaceae (nymphaeales). *BMC evolutionary*  
104 *biology*, 14(1):1.
- 105 MacArthur, R. H. and Wilson, E. O. (1967). Theory of island biogeography.(mpb-1). 1.
- 106 Nylander, J. A., Olsson, U., Alström, P., and Sanmartín, I. (2008). Accounting for phylogenetic uncertainty  
107 in biogeography: a bayesian approach to dispersal-vicariance analysis of the thrushes (aves: Turdus).  
108 *Systematic Biology*, 57(2):257–268.
- 109 Ree, R. H., Moore, B. R., Webb, C. O., and Donoghue, M. J. (2005). A likelihood framework for inferring  
110 the evolution of geographic range on phylogenetic trees. *Evolution*, 59(11):2299–2311.
- 111 Ronquist, F. and Sanmartín, I. (2011). Phylogenetic methods in biogeography. *Annual Review of Ecology,*  
112 *Evolution, and Systematics*, 42(1):441.
- 113 Sanmartín, I. and Meseguer, A. S. (2016). Extinction in phylogenetics and biogeography: From timetrees  
114 to patterns of biotic assemblage. *Frontiers in Genetics*, 7:35.
- 115 Strijk, J. S., Bone, R. E., Thébaud, C., Buerki, S., Fritsch, P. W., Hodkinson, T. R., and Strasberg, D.  
116 (2014). Timing and tempo of evolutionary diversification in a biodiversity hotspot: Primulaceae on  
117 indian ocean islands. *Journal of biogeography*, 41(4):810–822.
- 118 Webb, C. O. and Ree, R. (2012). Historical biogeography inference in malesia. *Biotic evolution and*  
119 *environmental change in Southeast Asia*, pages 191–215.
- 120 Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer Science & Business Media.

Scenario	Frequency of state
equalratesLow	0.9994
equalratesHigh	0.9805
diffratesLow	0.9774
diffratesHigh	0.9388

**Table 1.** Frequency of the most likely state at the root of the tree for the Indian Ocean Primulaceae data.



**Figure 1.** Effects of the use of differential extinction rates on ancestral range reconstruction for ten simulated datasets. Boxes represent the relative likelihood of the true state at the root of the tree under each scenario as inferred in revbayes.