Sensitivity of biogeographic reconstructions to the use of differential extinction rates

Jose L. Pinzón

Escuela de Biología, Universidad Industrial de Santander, Bucaramanga, Colombia.

Corresponding author:
Jose L. Pinzón

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

ABSTRACT

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

INTRODUCTION

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

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

METHODS

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

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

As empirical example, I used the Indian Ocean Primulaceae dataset of Strijk et al. (2014), particularly the subclade inhabiting Madagascar and the Mascarene islands of Mauritius, Reunion and Rodrigues. This example was chosen due its particularity: each of these species is single-island endemic, and differences in island size are notorious.
Simulated data were generated in the R package diversitree (FitzJohn, 2012) as follow. First, I fitted the parameters of extinction, anagenetic change and extinction for the Primulaceae data to the MuSSE model (FitzJohn, 2012). The use of MuSSE rather than GeoSSE was preferred due to the especial conditions mentioned above to match the number of possible states (4 in the empirical example). The root of each simulated scenario was assigned to state = 1, and 10 simulated datasets were generated with a limit of 30 extant taxa to the present.

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

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

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

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

RESULTS AND DISCUSSION

Indian Ocean Primulaceae biogeography

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

Simulated datasets

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

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

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

CONCLUSIONS

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

SUPPLEMENTAL DATA

Supplementary results, data and scripts used to generate simulations are available at http://github.com/jessop/Extinction-rates.
REFERENCES


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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Frequency of state</th>
</tr>
</thead>
<tbody>
<tr>
<td>equalratesLow</td>
<td>0.9994</td>
</tr>
<tr>
<td>equalratesHigh</td>
<td>0.9805</td>
</tr>
<tr>
<td>diffratesLow</td>
<td>0.9774</td>
</tr>
<tr>
<td>diffratesHigh</td>
<td>0.9388</td>
</tr>
</tbody>
</table>

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.