

Events recovery under different methods

Karen Méndez¹ and Daniel Miranda²

¹Universidad Industrial de Santander

²Universidad Industrial de Santander

Corresponding author:

First Author¹

Email address: karenmendez9404@gmail.com

ABSTRACT

To reconstruct the biogeographic history should be taken into account the processes involved in the current distribution of the biota, for example vicariance, dispersal, sympatry and extinction. Different methods have been created to answer this. From methods to detect only vicariant events to ancestral area analysis under different models where all events are considered to reconstruct the history at individual level and from which analyze if there are congruence with other taxonomic groups, methods based on cells, endemism areas, parsimony, parametric methods, in general there is a great variety of methods. All have a different approach to determine the amount and the location of the event in the phylogeny but all can give a resolution of how have been created this patterns of distribution.

INTRODUCTION

There are several biogeographic processes (vicariance, duplication, extinction, and dispersal) that contribute in shaping the current distribution patterns for organisms, exists different approaches to detect these events involved, one of them is based on identifying sister nodes with disjunct distributions (allopatric/vicariant), based on Hovenkamp's ideas (Hovenkamp, 1997, 2001), who suggests that the only evidence of a speciation process in a geographic context is allopatric distributions, takes distributional information and phylogeny without branch lengths and the optimality criterion seeks to find the best compromise between the maximum possible number of disjunct sister nodes (Arias et al., 2011).

Moreover the event-based biogeographic methods (Page, 1994; Ronquist, 1998) rely on models, reconstruct ancestral distributions and biogeographic events. One of the most common methods considers four types of events: vicariance, duplication, extinction and dispersal (Page, 1994), each event has a cost and assess the global cost for reconciling two trees and the optimal reconstruction minimizes the cost of the implied events (Ronquist 2002).

In Ancestral area analysis like the Dispersal-extinction cladogenesis model (DEC) specifies instantaneous transition rates between discrete states (ranges) along branches of the phylogeny and estimate likelihoods of ancestral states at cladogenesis events (Ree et al., 2005; Ree and Smith, 2008). Unlike existing methods for estimating ancestral areas, such as dispersal-vicariance analysis (DIVA (Ronquist, 1997), this approach incorporates information on the timing of both lineage divergences and the availability of connections between areas (dispersal routes).

The goal of the present analysis is determine the differences between the events recovered by different methods (like VIP, GEM, BioGeoBears and TreeFitter) for empirical and simulate data.

METHODS

For empirical data analysis were used two phylogenies obtained from a likelihood analysis in Phyl software (Guindon et al., 2010) of birds (Family: Capitonidae, representing approximately 86 % of all species, and Family: Ramphastidae, representing approximately 82 % of all species), which are distributed in the Amazon basin and carry out a simulation in the R package 'diversitree' (FitzJohn, 2012), which were simulated phylogenies with its respective areas of distribution (six areas in total) for 15 taxa. The georeferenced coordinates were get from the Global Biodiversity Information Facility (GBIF) (www.gbif.org).

To detect the disjunct distributions (vicariant, allopatric) among sister groups was used the Vicariance inference program (VIP) (Arias et al., 2011), which is based on the Hovenkamp ideas (Hovenkamp, 1997, 2001), this requires phylogenies without branch lengths and the coordinates for each terminal. Also was used a method of phylogenetic biogeography to infer events (like vicariance, full copy sympatry, point sympatry, and founder event) implemented in the Geographically explicit event model (GEM) (Arias, 2016: www.lillo.org.ar/phylogeny/GEM) also based on Hovenkamp ideas.

It carried out an ancestral area analysis for each taxon under the models: DEC (Ree et al., 2005) and DIVALIKE implemented in BioGeoBears (Matzke, 2013) with the same dispersal rates.

In TreeFitter (Ronquist, 2002) were assessed four different costs: the default model, reconciled trees (REC) and maximum codivergence (MC) considering four types of events: codivergence corresponding to vicariance in biogeography, duplication corresponding to speciation within a region, sorting corresponding to partial extinction and switching corresponding to dispersal. Each of these events within several models deal to explain the interaction between organism phylogenies (P-tree) and area cladogram (H-tree). To assess the significance an analysis of permutation was done, one when P-tree and H-tree and both were permuted to see how vary the events.

RESULTS AND DISCUSSION

For Capitonidae, both GEM and VIP had the same amount of vicariant events (4), and half of them were shared by both in the same node. GEM recovered the other half of events that were not shared by both methods like founder events and the rest of events were Vicariance V=4, Punctual sympatry P=4, Full sympatry S=3, Founder event F=2. For Ramphastidae the amount of vicariant events recovered by GEM were more than VIP (9 vs 4 respectively) and all events recovered by VIP were recovered by GEM. In GEM S=10, V=9, F=6, P=4. The comparison was done for vicariant events because is the event in common (Figure 1).

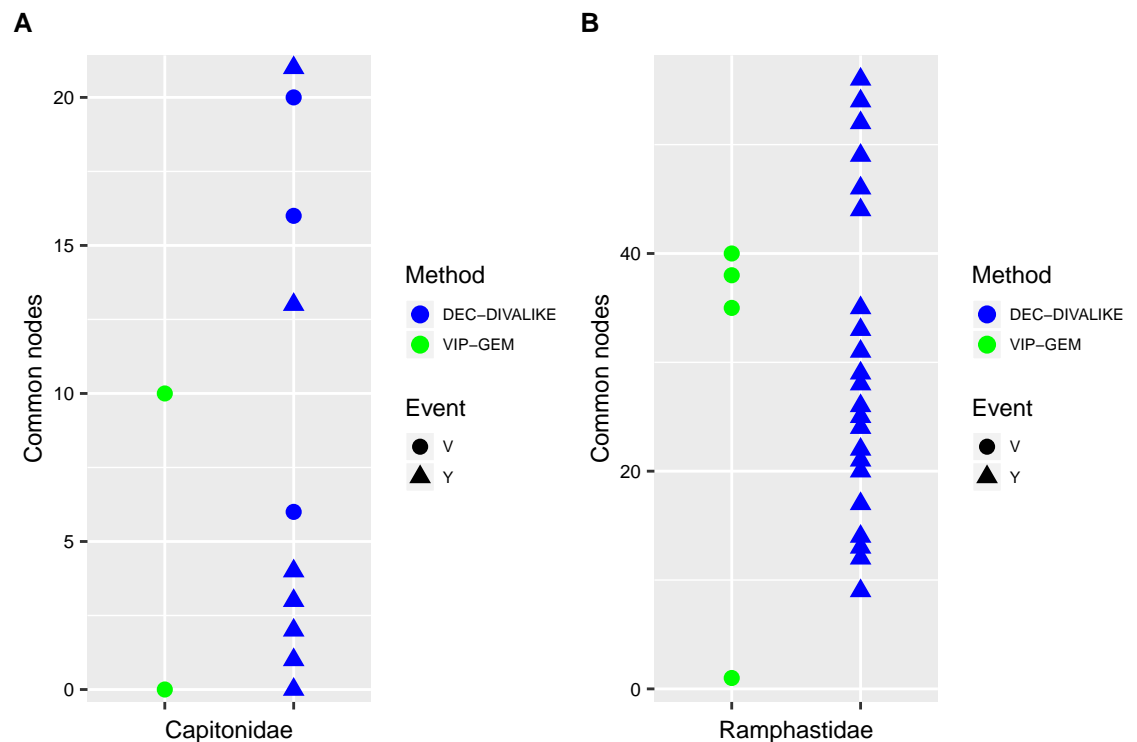


Figure 1. Common nodes recovery by methods based on cells and areas. Abbreviations: Y sympatry range copying, V vicariance

The percentage of common events for methods of cells were low for both, this could be related to cost management for overlapping or elimination distribution (Arias et al., 2011).

For Capitonidae the ancestral area analysis under the models DEC and DIVALIKE produced almost the same events on the same nodo, in terms of percentage were approximately shared 78 % and for Ramphastidae 90 % (Figure 1). In the analysis of simulated data the percentage the events shared by both DEC and DIVALIKE models is approximately 54 %. Although the simulation is an approximation to evaluate different things, this depends of several factors used to reach it, also because the simulated data are far away from reality.

The high congruence between the events recovered by both models in ancestral area analysis could be due to DIVALIKE is a modification of the DEC model (Matzke, 2013), so for that reason the events recovered by both models are almost the same.

In general, when VIP and GEM did not show more vicariant events than ancestral area analysis this yielded the same amount but never less, this can be because methods like VIP and GEM did not require a priori assumptions of areas, so the percentage of vicariant events can be more than methods like ancestral that use a priori areas (Ree et al., 2005).

The TreeFitter analyses showed that when the host terminals were permuted, the analysis detected more cospeciation events than expected by chance for the three costs, in other words the present day distribution of Ramphastidae and Capitonidae is a result of vicariance within Amazon basin and speciation within them. When parasite terminals and both the host and parasite terminals were permuted for default cost switch became significant and vicariance continued being significant for MC and REC (Table 1). For simulated data, the events that structure the association also codivergence, duplication and switch but never sorting.

Model	Permutation setting	Random cost<=observed	Event	Random<=observed	Random>=observed
Default	Hterm	0/1000*	Codivergence	994/1000	35/1000 *
MC	Hterm	18/1000*	Codivergence	1000/1000	18/1000*
REC	Hterm	1/1000*	Codivergence	1000/1000	1/1000*
			Duplication	1/1000*	1000/1000
Default	Pterm	0/1000*	Switch	4/1000*	997/1000
MC	Pterm	42/1000*	Codivergence	1000/1000	42/1000*
REC	Pterm	21/1000*	Codivergence	989/1000	21/1000*
			Duplication	21/1000*	989/1000
			Switch	37/1000*	974/1000
MC	HPterm	28/1000*	Codivergence	999/1000	28/1000*
REC	HPterm	27/1000*	Codivergence	986/1000	27/1000*
			Duplication	27/1000*	986/1000

Table 1. Events obtained through the analysis in TreeFitter. Only the significant events *

CONCLUSION AND FUTURE WORK

Our results emphasize the incongruence between methods that use cells and endemism areas to infer the events in the phylogeny, although this can not be the only factor is one of them.

It is necessary that the following study considers other variables like area size and cell size, the percentage of overlapping and the elimination of distributions in the recovery of events in all methods evaluated in the present analysis.

REFERENCES

- Arias, J. S., Szumik, C. A., and Goloboff, P. A. (2011). Spatial analysis of vicariance: a method for using direct geographical information in historical biogeography. *Cladistics*, 27(6):617–628.
- FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in r. *Methods in Ecology and Evolution*, 3(6):1084–1092.
- Hovenskamp, P. (1997). Vicariance events, not areas, should be used in biogeographical analysis. *Cladistics*, 13(1-2):67–79.
- Hovenskamp, P. (2001). A direct method for the analysis of vicariance patterns. *Cladistics*, 17(3):260–265.
- Matzke, N. J. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
- Page, R. D. (1994). Parallel phylogenies: reconstructing the history of host-parasite assemblages. *Cladistics*, 10(2):155–173.

- Ree, R. H., Moore, B. R., Webb, C. O., and Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, 59(11):2299–2311.
- Ree, R. H. and Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1):4–14.
- Ronquist, F. (1997). Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, 46(1):195–203.
- Ronquist, F. (1998). Three-dimensional cost-matrix optimization and maximum cospeciation. *Cladistics*, 14(2):167–172.
- Ronquist, F. (2002). Treefitter, version 1.2. *Software available from <http://morphbank.ebc.uu.se/TreeFitter>.*