

Spatial analysis of vicariance: a method for using direct geographical information in historical biogeography

J. Salvador Arias*, Claudia A. Szumik and Pablo A. Goloboff

CONICET, INSUE, Facultad de Ciencias Naturales e Instituto Miguel Lillo, Universidad Nacional de Tucumán, Miguel Lillo 205, S.M. de Tucumán (4000), Tucumán, Argentina

Accepted 26 January 2011

Abstract

Based on Hovenkamp's ideas on historical biogeography, we present a method for analysis of taxon history, spatial analysis of vicariance, which uses observed distributions as data, thus requiring neither predefined areas nor assumptions of hierarchical relations between areas. The method is based on identifying sister nodes with disjunct (allopatric/vicariant) distributions. To do this across the tree, internal nodes are assigned distributions (as the sum of the distributions of the descendant nodes). When distributions are less than ideal, ignoring the distribution of the problematic node(s) when assigning a distribution to their ancestors may allow us to consider additional sister nodes (i.e. those resulting from splits basal to the problematic node) as having disjunct distributions. The optimality criterion seeks to find the best (possibly weighted) compromise between the maximum possible number of disjunct sister nodes and the minimum number of eliminated distributions. The method can also take overlap into account. The methodology presented is implemented in VIP, a computer program available at <http://www.zmuc.dk/public/phylogeny/vip>.

© The Willi Hennig Society 2011.

In an important series of contributions, Hovenkamp (1997, 2001) proposed that the identification of barriers separating biotas is a goal more in line with the spirit of vicariance biogeography, rather than the discovery of relationships between predefined areas. This switches the focus from the relationships between areas to the identification (location and sequence) of barriers. Although Hovenkamp's papers have been widely cited, the crux of his argument—switching the focus from areas to barriers—has been ignored by most historical biogeographers.

Hovenkamp makes a clear distinction between studies focused on “Earth history” and those focused on “taxon history”. Methods focused on Earth history attempt to summarize the biogeographical information shared by multiple independent clades, i.e. a diverse taxon sampling; in this case, the information from multiple sources is used as evidence for vicariance events (e.g. “cladistic

biogeography” of Nelson and Platnick, 1981; Humphries and Parenti, 1999). In contrast, Hovenkamp defined “taxon history” methods as those that, given a phylogeny, try to explain the distribution of a particular taxon in a geological context (often in the form of an “area cladogram”; see Brooks, 1990; Page, 1994a; Ronquist, 2003). A good portion of taxon history methods are inspired by those studying coevolutionary host–parasite systems, considered as analogous to the area–taxon relationship (e.g. Brooks, 1981; Ronquist and Nylin, 1990; Page, 1994a; Ronquist, 1998).

According to Hovenkamp, failing to clearly distinguish these two goals is a frequent source of confusion in many biogeographical studies. Since Hovenkamp's (1997) interest is in Earth history methods, he does not discuss taxon history in depth, but many of the shortcomings he enumerates for Earth history methods also pervade taxon history approaches.

One of Hovenkamp's (1997, 2001, 2002) main criticisms of the standard approach in historical biogeography is that there is no reason to expect the history of Earth to be hierarchical. That is to say, taxa may well

*Corresponding author:

E-mail address: jsalarias@csnat.unt.edu.ar

not be transmitted vertically in an “area cladogram” as characters are in a phylogeny. Also, taxa can in theory move freely in any direction, as long as this is not prevented by a geological or ecological barrier. Thus, if there is no exact correspondence between the inhabited area and the barriers, then current areas of distribution need not be coincident, and need not reflect ancestral distributions.

The criticisms Hovenkamp levelled at Earth history methods apply equally to most taxon history methods, which assume an area cladogram as the geological context against which the taxon phylogeny must be examined. There have been some recent proposals for taxon history methods which do not require that the areas are hierarchically arranged (DIVA, Ronquist, 1997; DEC, Ree et al., 2005; Ree and Smith, 2008), but these methods continue to be based on discrete, predefined areas. The areas (e.g. continents or islands) may be evident in the specific examples used to propose those methods, but in the vast majority of real cases there are different degrees of overlapping among the distributions of taxa, and not all terminals show congruence between their distributions. Thus, the (prior) definition of the areas is far from trivial. Furthermore, within the framework of predefined areas, “widespread” taxa (here, we use “widespread” within quotations in the traditional sense of historical biogeography: occupying more than one predefined area; see Nelson and Platnick, 1981; Humphries and Parenti, 1999) continue being a problem without a natural solution (e.g. Nelson and Platnick, 1981; Axelius, 1991; Nelson and Ladiges, 1996; van Veller et al., 1999; Sanmartín and Ronquist, 2002). From the practical point of view, these taxon history methods use algorithms and implementations which allow only very limited numbers of areas, making them of little use for most current phylogenies (DIVA allows up to 16 areas, but only counts up to eight; so far, published studies with DEC use no more than ten areas). In most cases (e.g. Sanmartín, 2003; Nylander et al., 2008; Smith, 2009), very large and extensive areas are used, with arbitrary limits that explicitly ignore cases of overlap, or internally heterogeneous unit areas.

In contrast to those methods, Hovenkamp (1997, 2001) has proposed an approach which uses the spatial component of taxon distributions (Heads, 2010; provided discussion along similar lines). Although these methods were presented within the context of an Earth history approach, they include a first step that is also appropriate for taxon history methods. Based on Hovenkamp’s ideas, the goal of this paper is to present a taxon history method, spatial analysis of vicariance, which uses the observed distributions as data and therefore requires neither a set of predefined areas nor assumptions of hierarchical relations between areas. The method is already implemented in VIP, a computer

program available at <http://www.zmuc.dk/public/phylogeny/vip> (Arias, 2010).

The method

Background

The main goal in Hovenkamp’s method is to discover disjunct (allopatric, or vicariant) distributions among sister groups. In the context of taxon history methods, vicariance (i.e. a geological event or barrier which appeared and physically separated populations) is not necessarily implied by a disjunct distribution; a successful dispersal over a pre-existing barrier resulting in speciation (as in the colonization of an oceanic island) will also produce a disjunct distribution. Therefore, focusing on disjunct distributions does not imply that vicariance is the only process allowed by the analysis (Fattorini, 2008); rather, it recognizes that a disjunct distribution implies the existence of an effective barrier to dispersal (Hovenkamp, 1997, 2001).

Hovenkamp’s approach switches the focus of biogeographical analysis from relationships between predefined areas to barriers—i.e. to the space *between* the distributions. Even when using sophisticated models of dispersal (as in DEC), predefined areas may discard geographical information and impose an arbitrary scale, precluding the simultaneous study of phenomena occurring between and within “areas”, and the effects of boundary choice are largely unknown. As taxa can move in any direction, there is no reason to expect that the actual distributions of taxa are highly correlated and that areas are “natural” units. But dispersal *is* limited by barriers. As long as barriers are effective over time, they will maintain disjunct distributions. Thus, the search for vicariant distributions uses the most reliable data provided by taxon distributions (Hovenkamp, 1997, 2001).

Whereas Hovenkamp (1997, 2001) focuses on the search for traceable vicariance events, the present paper deals with analysis of disjunct distributions to emphasize that the focus of the analysis is on the distribution patterns of the taxon itself.

Spatial analysis of vicariance

The spatial analysis of vicariance takes as input a phylogenetic tree that includes spatial distributions on its terminals. Using an optimality criterion (outlined below), it chooses a set of distributional reconstructions in which cases of disjunction for sister nodes are maximized. These disjunct distributions can be displayed on a map and, by using Voronoi–Delaunay graphs (for a review see De Berg et al., 2008), it is possible to suggest potential barrier localizations.

Spatial distribution of terminals. The implementation of the method is based on a grid on which the spatial information is mapped; then, taxon distributions are taken from clouds of georeferenced physical points or distributions map. As in other grid-based methods, it is possible to use filling algorithms to correct bias in the grid origin, cell size, and gaps caused by low sampling effort (e.g. Szumik et al., 2002; Szumik and Goloboff, 2004; Aagesen et al., 2009).

Spatial distribution at internal nodes. The distribution of the internal tree nodes is calculated with a down-pass, as the sum or union (“ORing”) of the distributions of their descendants: a grid occupied in either or both descendants will be considered as “occupied” at the internal node. If the intersection in the distribution of the descendants is an empty set, then the node has descendants with disjunct distributions.

It is important to point out that distributions at internal nodes are statements similar to those used in taxonomic revisions to refer to the distribution of higher taxa, not necessarily based on the concept of ancestral area or ancestral distribution. The emphasis here is not on the distribution of the ancestors but instead on the disjunctions, or gaps, that correspond to speciations (Hovenkamp, 1997, 2001; Heads, 2010). In a sense, the distribution hypothesized at an internal node serves more to identify where the taxon is absent, rather than a positive statement of where it actually is.

Optimality criterion. So far, the method outlined does not differ substantially from Hovenkamp’s (1997, 2001), and requires no optimality criterion. However, the distribution of a taxon may (e.g. after dispersal) blur potentially disjunct distributions of sister nodes. According to Hovenkamp, these cases simply indicate that the evidence for those disjunctions has been destroyed (a similar problem may occur in the context of co-speciation when using ORing; Brooks, 1990; Page, 1994b).

When distributions are less than ideal, ignoring the distribution of the problematic node(s) may allow us to consider the splits basal to the node in question as disjunct distributions (Page, 1994a, b; implicit in Brooks, 1990). This, in turn, requires us to consider alternative combinations of possible distribution eliminations. This stage does require an optimality criterion to select the best combination of removal sets and the reconstructions of disjunction patterns they imply.

In the case of co-speciation, where a hierarchical pattern is a legitimate background, Page (1994b) proposed simply to maximize the number of co-speciations, or more generally co-divergences (Ronquist, 2003). In the present context (where hierarchy cannot be assumed), the exclusive use of this criterion may produce unsatisfactory solutions; Fig. 1 shows two possible

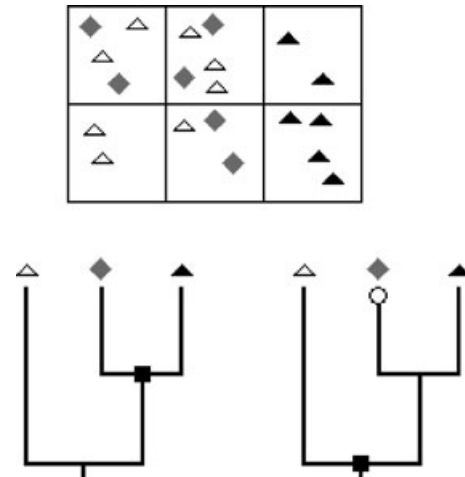


Fig. 1. Two reconstructions for a hypothetical distribution (shown in the grid). The only vicariant distribution is marked with a black box at the node. If only the minimization of disjunct nodes is taken into account, both reconstructions have a cost of 1. If distribution removal has a cost of 1, the reconstruction on the left is better (cost: 1 versus. 2).

reconstructions, both with the same number of overlapping sister nodes. Although both reconstructions imply the same number of disjunct sisters, one of them does not remove any distribution. Therefore, a count of the number of nodes with their distribution removed must also be included as part of the quantity to be minimized. Thus, the present approach seeks to find the best (possibly weighted) compromise between the maximum possible disjunct sister nodes and the minimum number of distribution eliminations.

The rationale for doing so is (as in phylogenetic parsimony) the minimization of *ad hoc* hypotheses. In the context of co-speciation (Page, 1994b; Ronquist, 2003), an elimination can be interpreted as a “host-shift”. In the present analysis, a node elimination might be interpreted as a “dispersal”, but this interpretation is problematic: there is no barrier (disjunct descendant distributions) associated with the node, and then it is impossible to know whether the dispersal occurred at the node, or the descendant. Then, an elimination cannot be univocally tied to a dispersal, and the number of eliminations does not provide an estimate of dispersal frequency. A distribution elimination is simply the ignoring of known spatial information, for the sole purpose of obtaining a better fit. This is an *ad hoc* strategy, with the sole purpose of maximizing the number of disjunct sisters. If the data can be explained equally without resorting to this strategy, that explanation must be preferred.

Formalization. Let a *j*-node be a node in which all-but-one descendant distributions are eliminated (Page, 1994b). In a binary tree, every distribution elimination

implies that the corresponding ancestor becomes a *j*-node. Every other node must lead to either disjunct or overlapping sisters (i.e. disjunction or overlap of descendants is a meaningless concept for a *j*-node). Let R_i be the cost of the reconstruction at node *i*, i.e. the cost of sub-tree *i*. Then the cost at node *m* will be:

$$R_m = \sum_{i = \text{desc}(m)} R_i + Y_m + X_m$$

where $\text{desc}(m)$ is the list of direct descendants of *m*, Y_m = the cost of distribution elimination (if *m* is an eliminated node) or $Y_m = 0$ (otherwise), and $X_m = 0$ (if *m* is a *j*-node or node leading to disjunct sisters) or X_m = cost of a node leading to overlapping sisters (otherwise).

The VIP implementation of the optimality criterion allows the user to provide the cost for distribution eliminations in relation to the cost (fixed at 1.00) of overlapping sisters; the final cost of the reconstruction is then given by the sum of both quantities. Then, the minimum cost of an elimination must be equal to the cost of overlapping sister nodes; otherwise, distribution eliminations of most terminals will always produce better reconstructions, even in the absence of disjunct sisters. In most of the examples used here, the cost of eliminations was set to 1.50.

Note that for a binary tree, setting the eliminations to 1.00 will produce the same result as just counting overlapping sisters. But this does not hold for multifurcating trees. So, in VIP, the user can set the “elimination cost” to 0.00, to allow us to explore the consequence of a pure maximization of disjunct sisters, regardless of the number of eliminations required.

Partial overlapping. There are many cases in which the distribution of two taxa is disjunct except for a small degree of overlap. Hovenkamp (1997) recognizes this problem, and even if he does not offer a formal solution, he proposes to take into account the degree of overlapping when deciding whether the overlapped distribution is to be accepted as vicariant. We formalize his suggestion here. Overlap is measured as the proportion of the distribution shared between two nodes (in number of shared cells) against the (number of cells of the) smaller distribution (Chesser and Zink, 1994; Barraclough and Vogler, 2000; Fitzpatrick and Turelli, 2006). If this proportion is below a user-defined overlap limit, then the distributions are accepted as disjunct, and no cost is assigned to the pair (or group) of sisters. Alternatively, it is possible to use the proportion of the overlap as the cost, i.e. to acknowledge that the sister distributions are not completely disjunct. For a 0% overlap limit, any overlap, however small, will rule out the sisters as disjunct. Given that a very large overlap casts doubts on the actual effectiveness of the barrier, it is advisable to use relatively small percentages of overlap (e.g. < 25%).

Detecting the overlap between immediate sister groups is often very simple, but overlap between more distantly related nodes may pose more subtle problems. For example, a taxon A may be perfectly disjunct with its sister group, while its distribution overlaps with another taxon B, which joins the tree at a more basal fork. In such a case, eliminating the distribution of taxon A results in losing at least one case of disjunct sisters, and it seems better to eliminate only part of the distribution of A (Fig. 2). In VIP, the user sets an upper limit to the proportion of the distribution that can be eliminated. The cost of a partial elimination is set by the user, and ranges from 0.00 to any value. Using 0.00 will allow several possibilities to be eliminated, so searches for optimal reconstructions will be greatly slowed down; on the other hand, using a value equal to or greater than the cost of elimination, few partial eliminations will be found. In most of the examples here, a fixed cost of partial eliminations is set to half an elimination (i.e. 0.75). It is also possible to use the proportion of the elimination as the cost (to acknowledge that only part of the distribution of the conflicting taxon is removed).

Two things must be noted at this point. First, we propose the use of a proportional overlap to make results more independent of grid size (i.e. scale). Second, we rescale the overlap relative to the smaller of the two distributions (Chesser and Zink, 1994; Barraclough and

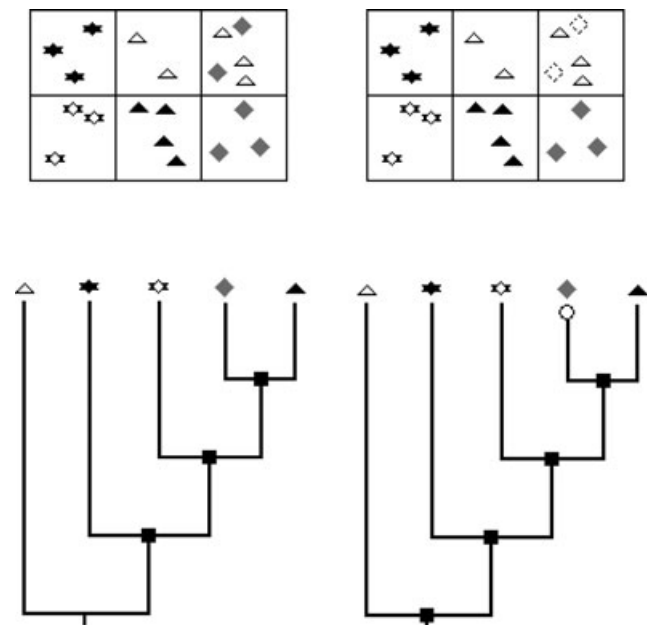


Fig. 2. The left reconstruction has only one case of overlapping sisters, because the distribution of black rhombus overlaps with the distribution of white triangles. If part of the distribution of black rhombus is removed (indicated as a white circle in the reconstruction), all sister nodes can be considered disjunct. If distribution removal has no cost, then the right reconstruction is one of the preferred reconstructions.

Vogler, 2000; Fitzpatrick and Turelli, 2006). Using the total overlap (i.e. the intersection of nodes A and B with respect to the union of A and B) would be problematic because the descendants might be considered as disjunct when one of them has a small distribution almost completely included within that of a hugely distributed sister node, when evidence of allopatry is in fact lacking in such a case.

A more technical description of the algorithms used by the method is given in Appendix S1.

Summarizing multiple reconstructions

For any given set of costs chosen, ambiguity may exist, i.e. there may be multiple equally optimal reconstructions. In most cases, the best way to summarize this ambiguity is by using a sort of conservative “consensus reconstruction”. In this way, a split is shown as disjunct if and only if this is the case in all optimal reconstructions (and in the case of polytomies, only if the descendants implied in the allopatry are always the same ones). The distribution of an ancestral node comprises all the cells present in all reconstructions for that node (i.e. intersection). A distribution elimination is accepted only if the distribution of that node is removed in all reconstructions. Of course, in highly ambiguous data sets, the consensus may have few or no nodes leading to unambiguously disjunct distributions—precisely the point of conservativeness.

Graphic representation of barriers

VIP uses the Voronoi–Delaunay tessellation to represent disjunct distributions (De Berg et al., 2008). This geometric method has the advantage that barriers can be expressed as lines (a Voronoi diagram), or as an area (a Delaunay triangulation) separating a set of points. This “barrier” is more a heuristic to help the user identify possible causal factors acting as barriers, rather than the barrier itself. Of course, as this procedure locates barriers on the midpoint, closer points will allow for a better delimitation of the barrier. In the case of overlapping, the points in the overlap are not taken into account to draw the barrier. Additional details of this graphic part will be described elsewhere (J.S. Arias, in preparation). In the end, as VIP is an open source program, interested users may always refer to the source code, and modify it if required.

Examples

Theoretical examples

Some simple theoretical examples may serve to illustrate the main properties of the method. Fig. 3

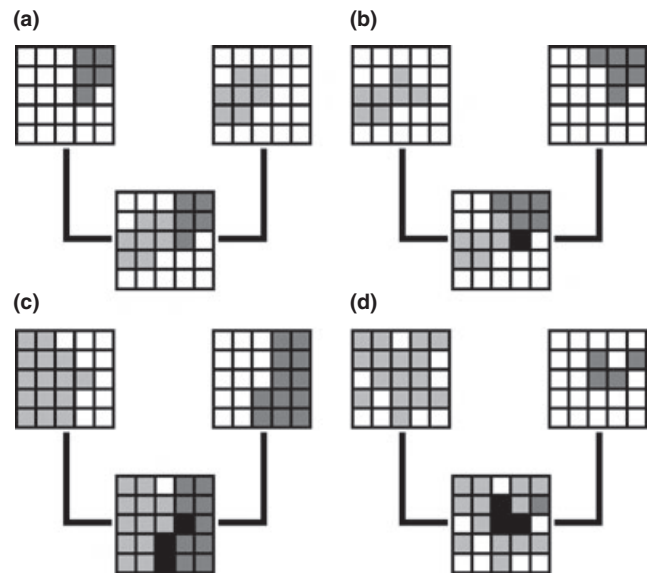


Fig. 3. Some examples of partial overlapping. (a) Without overlapping, the distribution of both descendants is disjunct; (b) a small overlap (16.7% relative to right descendant) is accepted if the overlap limit is set to 16.7%; (c) in this case the overlap is 25% relative to the right descendant and 20% relative to the left descendant, so that it is not counted as disjunct if the overlap limit is set to 25% (i.e. the greater value of overlap is always used to make the comparison); (d) in this case, the distribution of the right descendant is almost completely included within that of the left descendant, acceptable only for an overlap limit above 75%.

shows different degrees of overlap in different pairs of distributions. In this case, sisters are counted as disjunct or not, depending on the acceptable overlap (i.e. a user-set limit).

Unlike methods based on predefined areas, spatial analysis of vicariance is not affected by “widespread taxa”. As the distribution of each taxon is taken to be the set of cells where it has been observed (or assumed to exist) the vast majority of taxa will be “widespread” in terms of occupied grid cells (e.g. as in Hovenkamp, 1997), and even more so when the distributions are filled with a minimum polygon or similar methods. As long as the distributions are clearly allopatric, one can (at least in theory) increase the resolution of the analysis and continue obtaining similar results (as in Fig. 4).

It is true that when two or more taxa overlap, they may obscure other possible vicariant distributions. However, these taxa need not be “widespread”—they may even have smaller distributions than the taxa for which the disjunction is obscured (Fig. 5). In this case, eliminating the distribution detects those disjunctions. As pointed out by Hovenkamp, the “problems” that “widespread taxa” pose to the identification of barriers or vicariance events are no more than false problems, merely the consequence of focusing on relationships between predefined areas instead of focusing on the barriers themselves (Hovenkamp, 1997, 2001). The

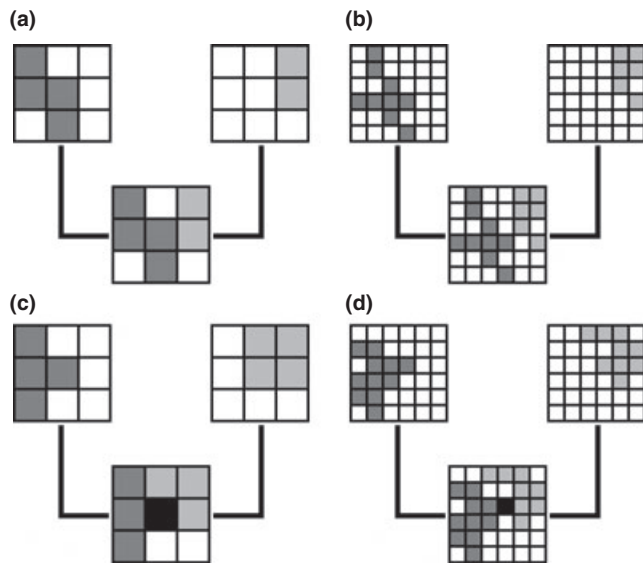


Fig. 4. Examples illustrating some scale effects. In the case of disjunct distributions (a, b), a finer grid (closer scale) does not change the reconstruction, maintaining the disjunction. When there is overlap, increased resolution allows a more detailed evaluation of the overlapping area; in this case, the overlap is far smaller than the overlap implied in the original scale (c).

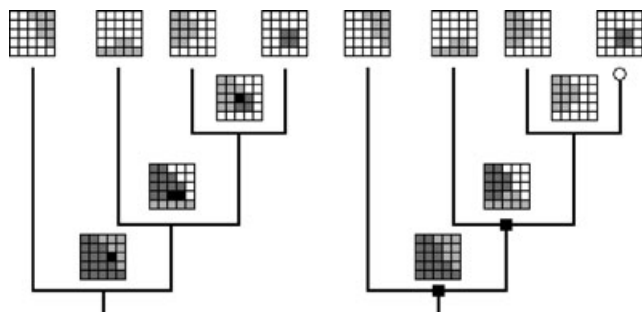


Fig. 5. Removal of node distributions. In cases of conflict it is possible to remove the distribution of a particular node and examine whether the number of disjoint distributions is increased. In the left reconstruction, it is not possible to detect any case of disjunct sister nodes for the complete data set, but removing the distribution of the right-most terminal, two cases of disjunct sister distributions become apparent. Note that the area of the terminal removed is smaller than the area of any other terminal, but under the context of predefined areas, this taxon would be considered as having a “widespread distribution”.

problem results actually not from larger areas of distribution, but instead from distributions with a high degree of overlap. As the number of such distributions increases (e.g. in deeper nodes), detection of disjunctions becomes more challenging or impossible.

An empirical example

To examine the results of the present method in a real data set, we have analysed here the biogeography of

Turdus, recently analysed by Nylander et al. (2008). Although we use the same taxon matrix from Nylander et al. (2008), we use more than 9000 distributional records (mostly from GBIF, see Appendix S1), instead of the eight predefined areas used by Nylander et al. Therefore, our results do not constitute a “re-analysis” of Nylander et al.’s data, as our data set is substantially different from theirs. This example was selected for two main reasons: Nylander et al.’s constitutes a “state of the art” biogeographical analysis, and it was based on a group for which it is possible to download good distributional data (i.e. actual point records) from the web. The general methodology and results are detailed in Appendix S1.

First, we wanted to know whether the limits of Nylander et al.’s areas are supported by the data. Our results found only 18 (out of 40) sister nodes with a disjunct distribution consistent with the predefined boundaries of Nylander et al., casting doubt on their a priori delimitation of areas (and the ancestral areas inferred) for the analysis of this particular group.

Nylander et al. argue that *Turdus* originated in Asia, with a subsequent migration to Africa. Although our results are not designed to show an ancestral area, we find that three of the basal splits of *Turdus* (including the root) present disjunct distributions associated with eastern Africa. This indicates that *Turdus* must already have been present in Africa at the beginning of its history. Results from DIVA ignore the fact that the distribution of *T. visciborus* and *T. philomelos* in Asia is allopatric with *T. mupinensis*. Nylander et al.’s optimization of an ancestor in Asia is an artefact of DIVA and the prior definition of areas, which assumes that the distributions of these taxa in Asia are the same, when in fact they are different. The same problem occurs in other places of the cladogram, for example in the most recent common ancestor of *Cichlherminia* (Lesser Antilles) and *T. jamaensis* + *T. swalesi* (Jamaica and Hispaniola), assigned to “Caribbean area”.

For node 12, Nylander et al. propose a South Asian origin with several dispersals to North Asia and Malaysia and Australasia. Only three nodes support the crossing barriers proposed by Nylander et al., whereas another 11 show partitions inside or across their predefined areas.

Discussion

Spatial information

Several recent methods have attempted to provide biogeographical studies with a quantitative framework (Ronquist, 1997; Ree et al., 2005). The development of optimality criteria to compare alternative reconstructions is an important step in quantitative biogeography.

However, it is equally important that these quantitative methods take into account the geographical nature of the data (i.e. what Szumik et al., 2002; and Szumik and Goloboff, 2004 have termed the “spatial component” in their method for endemism).

A distribution may or may not be congruent with the distribution of other taxa in the study, but ignoring these differences *a priori* by mixing distributions into a single “area” amounts to throwing away potentially informative data. The same holds for ignoring overlapping distributions completely.

One might argue (and some people have: see Dos Santos et al., 2008; with response by Casagrande et al., 2009) that the use of grids is as artificial as the use of predefined areas. But the distribution of an organism must be described as a surface, and a grid has provided a natural approximation to work with complex surfaces since the times of the Greeks. From a theoretical point of view, the grid should be as fine as possible, to provide the best estimation of the surface associated with the taxon. Of course, the practical problem with fine grids is that available data are usually very incomplete (and even if complete, the specimens are never distributed “infinitely” over the area!), so the data are full of holes. An alternative is, for example, the filling of empty cells based on a predefined range (such as the mean distance between points; e.g. Szumik and Goloboff, 2004; Aagesen et al., 2009; Casagrande et al., 2009), or some form of a minimum polygon.

Another huge advantage of using grids is that it makes working at several scales possible. That is, even if predefined areas correspond to real geographical entities, the use of differently sized grids could allow us to find phenomena at different scales. The direct use of grids is, in both theory and practice, the best way to deal with the available spatial information.

Despite the widespread use of predefined areas, few authors offer explicit reasons to prefer a predefined area over the actual areas of taxa. In most cases, this seems to be just for convenience:

[A]reas should be circumscribed according to the particular aspects of geographic history that are motivating the analysis: for example, if the history of dispersal between continents is of primary interest the continents may be specified as areas in the model (Ree et al., 2005, p. 2300).

As the *Turdus* example shows, even for seemingly simple problems such as transcontinental dispersal, there may be little relationship between predefined areas and the actual data. In most cases, the situation can be exceedingly complex. Instead of being contiguous, the distributions may have different degrees of overlap, in which case the definition of *a priori* barriers or corridors is not immediately evident. Even worse, the “motivation for the analysis” can be used to justify any kind of area, barrier or corridor.

Independently of the criterion used to delimit them, predefined areas have other drawbacks (Hovenkamp, 1997, 2001). Predefined areas produce several methodological and theoretical “problems” which do not belong to the realm of inferring biogeographical patterns, such as the value of factors or criteria to delimit the areas (e.g. Axelius, 1991; Ree et al., 2005; Ree and Smith, 2008; Ree and Sanmartín, 2009), the “wide-spread” and “redundant” taxa that lead to “biogeographical assumptions” (e.g. Nelson and Platnick, 1981; Engloff, 1996; Nelson and Ladiges, 1996; van Veller et al., 1999; Sanmartín and Ronquist, 2002), or attempts to give biological/epistemological meaning to the areas themselves (e.g. Morrone, 2001; Ree et al., 2005; Ree and Smith, 2008; Ree and Sanmartín, 2009).

On the frequency of allopatry

It might be argued that the spatial analysis of vicariance depends on how common allopatric speciation might be. Several studies (e.g. Chesser and Zink, 1994; Barraclough and Vogler, 2000; Fitzpatrick and Turelli, 2006) have investigated the amount of allopatric versus sympatric speciation using ideas similar to the ones developed here (assigning distributions to the nodes, and measuring the overlap among sister groups). The present method is substantially different for two reasons. First, the assignments on nodes are not necessarily to be seen as “ancestral” areas; and second, the number of occurrences of each form of speciation is not counted.

If most distributions are highly overlapped (to be expected if allopatric speciation is rare), the method will be unable to detect a particular geographical pattern, as few sister nodes will support disjunct distributions, even with the removal of several distributions. That is the reason to assign a cost to distribution elimination: eliminations are acceptable only if they substantially increase the number of disjunct sister nodes.

When sister nodes have overlapping distributions, that does not necessarily imply anything on how a particular speciation process caused the split. Rather, it means that geography alone cannot be used to explain the cladogenetic event.

Just as the maximization of homologous resemblance in phylogenetics does not assume homoplasy is rare (Farris, 1983), in spatial analysis of vicariance the maximization of disjunct nodes does not mean that allopatry is common.

Comparison with other methods

The ability to use direct geographical information is a critical difference between spatial analysis of vicariance and other methods. It could be argued that methods

working with predefined areas do so only to reduce computational burden and that those methods could be extended to grids, should huge computational power be available. At least for the current definition of those methods, this is simply incorrect, as operational aspects are intertwined in highly specific ways with the need to define areas *a priori*.

An example provided by Ree et al. (2005, p. 2302). shows DEC's deep reliance on predefined areas. Consider a disjunct distribution on two cells A and B. With the method proposed here, if A and B themselves are further divided in cells (A1, A2, B1, B2), the disjunct distribution continues being A1–A2/B1–B2 (and is identified as such by spatial analysis of vicariance). While Ree et al. (2005) accept the disjunction of A and B, the results of their method will change as resolution increases. In DEC, the area is the place where “speciation” occurs, and the “simultaneous” occurrence of a new species in two areas is forbidden. Therefore the DEC results for finer scales are highly artefactual; evidently aware of this, Ree and coauthors argue for reducing the number of areas in the study (Ree et al., 2005; Ree and Smith, 2008; Ree and Sanmartín, 2009).

This particular case is not problematic for DIVA (Ronquist, 1997), but others are. DIVA allows an ancestor to be widespread if their descendants are fully allopatric, but otherwise the ancestor is optimized in several equally costly ways. For example, if two descendants are present in the same cell A, the area A is assigned to their ancestor with a cost of 0. If the resolution increases and both descendants continue being sympatric (A1 and A2), there are now three possible reconstructions for its parent node (A1, A2 or A1A2), with a cost of 2. Thus from a single reconstruction with no cost, a decrease in grid size causes the method to switch to three alternative reconstructions with the same cost as the number of cells in each node (which casts doubts on Ronquist (1997 p. 196) of widespread ancestors being the “quintessence of the method”). The problem of dramatic changes in the reconstructed ancestral area with increased resolution is also present in the current implementation of DEC.

As Ronquist (1997) DIVA necessarily underestimates local extinction, dispersion and vicariance events (i.e. events within an area); the same applies to DEC. These two programs can count “events” only because they postulate arbitrary barriers; a “crossing” of such a barrier counts as an “event”. This is critical in the case of adjacent areas. For example, crossing an entire continent (e.g. South America) is “simpler” (less costly, without a particular likelihood under a DEC model) than changing areas (i.e. a passage from Central to South America). It is possible to find areas with such characteristics in almost any “event” based study using predefined areas (e.g. Sanmartín and Ronquist, 2004; Nylander et al., 2008; Santos et al., 2009; Smith, 2009). Given these problems

with arbitrarily defined areas, the subsequent calculation of dispersal versus vicariance frequencies (e.g. Sanmartín and Ronquist, 2004; Nylander et al., 2008) or dispersal and extinction rates (e.g. Santos et al., 2009; Smith, 2009) is absolutely unwarranted.

Other methods frequently used are character mapping (“areas as characters”) using Fitch optimization (Fitch, 1971) or stochastic mapping (Huelsenbeck et al., 2003). As far as we know, nobody has provided a theoretical argument for the use of such techniques in biogeography. Ronquist (1997) provided convincing criticism of these methods in general; equally strong criticisms can be applied to the specific application of these character mapping methods to grid-based data. In the biogeographical context, a drawback of these methods is (among others) that any overlap (however small) will be considered as a “non-empty intersection” and thus the area of overlap will be always considered as the “ancestral area”. In addition, if some form of filling algorithm is used, these character mapping methods often lead to the surprising conclusions that the ancestral area is an area where none of its descendants is found—just the middle of a filling polygon.

More recently, Lemmon and Lemmon (2008) proposed a method to calculate the location of an ancestor using geographical distributions and a diffusion model, intended (in their own words) as a sort of “genomic phylogeography”. Although this method does use direct geographical information, an important drawback is that for each terminal (and node), a single georeferenced point is assigned, which makes it impossible to use the spatial information contained in, for example, multiple museum specimens. In spatial analysis of vicariance, there is no requirement for the number of points of each terminal—it may be a single one, it may be many. In this way, even without inferring the single “ancestral point” of the taxa, patterns of disjunction among these points can be hypothesized. It is true that the goal of Lemmon and Lemmon is to detect patterns of dispersion for closely related populations, and the model they use might be reasonable within their conceptual paradigm. As much longer evolutionary times are considered, however, their method seems unsuited to actually detect barriers.

Potential criticisms

Many authors (e.g. Ronquist, 1997, 2003; Ree et al., 2005; Ree and Smith, 2008) dismiss methodologies not based on events as “pattern biogeography”. A brief rebuttal of those potential criticisms is offered here.

Widely distributed ancestors. Some authors interpret distribution at nodes as “widespread ancestors”. Depending on the stance, this may be intended positively (e.g. Heads, 2010) or negatively (e.g. Nylander

et al., 2008). Spatial analysis of vicariance has as its main goal the detection of disjunct distributions of sister nodes, not the goal of finding ancestral areas. The actual distribution of the ancestor may well have been much more limited or extended than the sum of the distributions of their descendants, but unless some of the descendants have been able to cross the original barrier (creating an overlap at the node), the dispersal abilities of the descendants, which enlarge the distribution of the node, are irrelevant (Hovenkamp, 1997, 2001). Considering the distributions assigned to nodes as ancestral distributions is an overinterpretation of the results, which goes well beyond using them simply to detect ancient barriers. As Hovenkamp (1997, pp. 70–71) puts it, focusing on vicariant distributions takes “the best part of the information contained in a [distribution]: the part where the [distribution] is constrained by an effective barrier”.

“*Ad-hoc*” interpretations. Unlike methods based on “events” (Ronquist, 1997, 2003), the present method is only capable of inferring disjunct distributions—as we are dealing with taxon history, we explicitly refrain from associating that pattern with a specific event or cause. Thus, the cause of the disjunction (e.g. a geophysical vicariance event or dispersal over an existing barrier) must be inferred from other lines of evidence. Ronquist (1997, 2003) and Page (1994b) criticize such kind of inferences as “*ad hoc*”. Although the main target of Ronquist’s and Page’s criticisms is BPA (e.g. Brooks, 1990), their criticism may be extended to any method not based on events. But event-based methods avoid these “*ad hoc* inferences” only by virtue of using arbitrarily defined areas. Most importantly, using independent sources of evidence to conclude dispersal does not make the inference as *ad hoc*—precisely the opposite! Even in event-based methods it is necessary to link the “event” with a geological and geographical space, so in this regard there is no real difference in the degree of *ad-hocness* of different methods.

A different type of interpretation may sometimes be warranted from elimination of distributions (as opposed to the interpretation of disjunctions); these may in some cases correspond to range expansions.

Branch lengths. Some likelihood-orientated authors (Ree et al., 2005; Nylander et al., 2008; Ree and Smith, 2008) have held that biogeographical methods should take into account branch lengths. Although the distinction is not always made clear, branch lengths are relevant only when interpreted purely as the time duration of a branch (i.e. in a dated tree), not as evolutionary rates. Although our current implementation does not consider branch lengths, it may be possible and desirable to let branch lengths influence the results. This influence may be less important than assumed by

Ree and coauthors: what matters the most is that the barrier has been effective during the existence of the branch (regardless of whether this time is short or long). More to the point, the length of the branch may affect whether vicariance or dispersal is a more reasonable explanation for an observed disjunction, not the fact that the disjunction exists.

Final comments

The spatial analysis of vicariance proposed here provides the first implementation for historical biogeography which can use spatial information in association with a cladogram, without using areas defined a priori. Although obviously at the initial stages of development, the method already allows: (i) inferring possible barriers based on georeferenced geographical and phylogenetic information; (ii) working without predefined areas, making it possible to detect phenomena at different spatial levels, and providing a more explicit statement on the scale of the analysis; (iii) comparing alternative hypotheses about the distribution of groups with an explicit optimality criterion (even allowing an analysis of sensitivity to different costs parameters). The methodology presented here is implemented in a computer program, VIP (Arias, 2010), available at <http://www.zmuc.dk/public/phylogeny/vip>.

Of course, there are many ways in which the method could be improved. The most important improvements would be, quite obviously, in using more sophisticated ways to measure the degree of overlapping and in developing appropriate measures of support.

Another possible addition is, when evaluating individual reconstructions, in penalizing not only the cases of overlapping sisters or eliminated distributions (as now done), but also adding a count of the extinctions or range contractions implied by the reconstruction (i.e. cases in which a descendant is absent from grid cells assigned to its ancestor). This requires solving the problem of how to distinguish these cases from obvious disjunctions in the descendants (e.g. when half of the distribution of the ancestor is missing in each of the descendants, but due to vicariance, not vanishment). Finally, although as presently described and presented, the method is conceived for analysis of taxon history, this procedure can be seen as a first step towards more sophisticated methods of Earth history.

Acknowledgments

We thank P. Hovenkamp for his comments and enthusiasm regarding this project. P. Sharma, P. Grandcolas and an anonymous reviewer also provided many useful suggestions. D. Casagrande, S. Catalano

and M. Mirande provided comments and encouragement at several steps of the research. We acknowledge the funding provided by CONICET and FONCYT (PICT 1314). The Willi Hennig society subsidized the version of TNT used for the phylogenetic analysis. VIP was written using GTK+ and Glib, and was compiled with GNU gcc, using CodeBlocks as IDE. J.S.A. received a Marie Stopes travel award from the Willi Hennig Society. The present paper (delivered by J.S.A.) received the Lars Brundin student award at the XXIX Hennig Meeting.

References

- Aagesen, L., Szumik, C.A., Zuloaga, F.O., Morrone, O., 2009. Quantitative biogeography in the South America highlands—recognizing the Altoandina, Puna and Prepuna through the study of Poaceae. *Cladistics* 25, 295–310.
- Arias, J.S., 2010. VIP: Vicariance Inference Program. Program, code, and documentation, available at <http://www.zmuc.dk/public/phylogeny/vip>.
- Axelius, B., 1991. Areas of distribution and areas of endemism. *Cladistics* 7, 197–199.
- Barracough, T.G., Vogler, A., 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* 155, 419–434.
- Brooks, D.R., 1981. Hennig's parasitological method: a proposed solution. *Syst. Zool.* 30, 229–249.
- Brooks, D.R., 1990. Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Syst. Zool.* 39, 14–30.
- Casagrandi, M.D., Arias, J.S., Goloboff, P.A., Szumik, C.A., Taher, L.M., Escalante, T., Morrone, J.J., 2009. Proximity, interpenetration, and sympatry networks: a reply to Dos Santos et al. *Syst. Biol.* 58, 271–276.
- Chesser, R.T., Zink, R.M., 1994. Modes of speciation in birds: a test of Lynch's method. *Evolution* 48, 490–497.
- De Berg, M., Cheong, O., van Kreveld, M., Overmars, M., 2008. *Computational Geometry, Algorithms And Applications*. Springer-Verlag, Berlin.
- Dos Santos, D.A., Fernandez, H.R., Cuezco, M.G., Dominguez, E., 2008. Sympatry inference and network analysis in biogeography. *Syst. Biol.* 57, 432–448.
- Engelhoff, H., 1996. Widespread taxa, sympatry, dispersal, and an algorithm for resolved area cladograms. *Cladistics* 12, 349–364.
- Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), *Advances in Cladistics 2: Proceedings of the Second Meeting of the Willi Hennig Society*. Columbia University Press, New York, pp. 7–36.
- Fattorini, S., 2008. Hovenkamp's ostracized vicariance analysis: testing new methods of historical biogeography. *Cladistics* 24, 611–622.
- Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20, 406–416.
- Fitzpatrick, B.M., Turelli, M., 2006. The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution* 60, 601–615.
- Goloboff, P.A., 1996. Methods for faster parsimony analysis. *Cladistics* 12, 199–220.
- Heads, M., 2010. Evolution and biogeography of Primates: a new model based on molecular phylogenetics, vicariance and plate tectonics. *Zool. Script.* 39, 107–127.
- Hovenkamp, P., 1997. Vicariance events, not areas, should be used in biogeographical analysis. *Cladistics* 13, 67–79.
- Hovenkamp, P., 2001. A direct method for the analysis of vicariance patterns. *Cladistics* 17, 260–265.
- Hovenkamp, P., 2002. Biogéographie de la vicariance: “mess” ou message? In: Deleporte, P., Silvain, J.-F., Hugot, J.-P. (Eds.), *Biosystema 20, Systématique Et Biogéographie*. Société Française De Systématique, Paris, France, pp. 15–26.
- Huelsenbeck, J.P., Nielsen, R., Bollback, J.P., 2003. Stochastic mapping of morphological characters. *Syst. Biol.* 52, 131–158.
- Humphries, C.J., Parenti, L.R., 1999. *Cladistic Biogeography*, 2nd edn. Oxford University Press, New York.
- Lemmon, A.R., Lemmon, E.M., 2008. A likelihood framework for estimating phylogeographic history on a continuous landscape. *Syst. Biol.* 57, 544–561.
- Morrone, J.J., 2001. Homology, biogeography and areas of endemism. *Divers. Distrib.* 7, 297–300.
- Nelson, G.J., Ladiges, P., 1996. Paralogy in cladistic biogeography and analysis of paralogy-free subtrees. *Am. Mus. Nov.* 3167, 1–58.
- Nelson, G., Platnick, N., 1981. *Systematics and Biogeography: Cladistics and Vicariance*. Columbia University Press, New York.
- Nylander, J.A.A., Olsson, U., Alstrom, P., Sanmartín, I., 2008. Accounting for phylogenetic uncertainty in biogeography: a bayesian approach to dispersal-vicariance analysis of the thrushes (Aves: *Turdus*). *Syst. Biol.* 57, 257–268.
- Page, R.D.M., 1994a. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. *Syst. Biol.* 43, 58–77.
- Page, R.D.M., 1994b. Parallel phylogenies: reconstructing the history of host-parasite assemblages. *Cladistics* 10, 155–173.
- Ree, R.H., Sanmartín, I., 2009. Prospects and challenges for parametric models in historical biogeographical inference. *J. Biogeogr.* 36, 1211–1220.
- Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57, 4–14.
- Ree, R.H., Moore, B.R., Webb, C.O., Donoghue, M.J., 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59, 2299–2311.
- Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46, 195–203.
- Ronquist, F., 1998. Three-dimensional cost-matrix optimization and maximum cospeciation. *Cladistics* 14, 167–172.
- Ronquist, F., 2003. Parsimony analysis of coevolving species associations. In: Page, R.D.M. (Ed.), *Tangled Trees: Phylogeny, Cospeciation, and Coevolution*. Chicago University Press, Chicago, pp. 22–64.
- Ronquist, F., Nylin, S., 1990. Process and pattern in the evolution of species associations. *Syst. Zool.* 39, 323–344.
- Sanmartín, I., 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *J. Biogeogr.* 30, 1883–1897.
- Sanmartín, I., Ronquist, F., 2002. New solutions to old problems: widespread taxa, redundant distributions and missing areas in event-based biogeography. *Anim. Biodivers. Conserv.* 25, 75–93.
- Sanmartín, I., Ronquist, F., 2004. Southern hemisphere biogeography inferred by event-based models: plant vs. animal patterns. *Syst. Biol.* 53, 216–243.
- Santos, J.C., Coloma, L.A., Summers, K., Caldwell, J.P., Ree, R., Canatella, D.C., 2009. Amazonian amphibian diversity is primarily derived from Late Miocene Andean lineages. *PLoS Biol.* 7, e1000056.
- Smith, S.A., 2009. Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifoliaceae. *J. Biogeogr.* 36, 2324–2337.
- Szumik, C.A., Goloboff, P.A., 2004. Areas of endemism: an improved optimality criterion. *Syst. Biol.* 53, 968–977.

Szumik, C.A., Cuezco, F., Goloboff, P.A., Chalup, A.E., 2002. An optimality criterion to determine areas of endemism. *Syst. Biol.* 51, 806–816.

van Veller, M.G., Zandee, M., Kornet, D., 1999. Two requirements for obtaining valid common patterns under different assumptions in vicariance biogeography. *Cladistics* 15, 393–406.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Details on the analysis of *Turdus*, including the dataset and results to be read with VIP.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Appendix 1. Pseudocode for the basic algorithms

The algorithms below are those used in VIP, the program which implements the spatial analysis of vicariance. Some of the algorithms are derived from those outlined by Hovenkamp (1996, 2001).

Data structures

A cellSet is a type that stores the distribution of a particular taxon. The cardinality (i.e. number of cells) of a cellSet C is indicated as |C|.

Apart from the basic topology information (ancestor and descendants), each node stores its distribution, its state (disjunct or overlapping descendants, removed distributions), and the accumulated cost of its subtree.

```
Node {
    anc, desc: pointer;
    C: cellSet;
    state;
    cost: real;
}
```

The tree struct contains a pointer to the root node, and to the list of nodes of the tree

```
Tree {
    root: pointer to a Node;
    nodeList: array of Node's;
}
```

Global variables

Global search parameters (i.e. available to all functions) are: the cost of a non-vicariant node (noVicCost), the cost of a distribution removal (removCost), the maximum acceptable overlap (maxOverlap), and maximum late overlap (lateOverlap).

REAL: GetNodeCost (actNode: Node)

This function calculates the cost of the specified node. It returns the cost of the evaluated node.

```
1. actNode fi cost ← 0;
2. if (actNode → state = = removed) then
3.   actNode → cost ← removCost;
4. end if
/* The node is a terminal */
5. if (actNode → desc = = NULL) then
6.   return actNode → cost;
7. end if
/* Get the cost of descendant nodes */
8. child ← actNode → desc;
9. left ← NULL;
10. right ← NULL;
11. while (child != NULL)
12.   actNode → cost ← actNode → cost + GetNodeCost (child);
13.   if (| child → C | != 0) then
14.     if (left = = NULL) then
15.       left ← child;
16.     else if (right = = NULL) then
17.       right ← child;
18.     end if
19.     actNode → C ← actNode → C union child → C;
20.   end if
21.   child ← GetNextDescendant (actNode);
22. End while
23. if ((left = = NULL) OR (right = = NULL)) then
24.   return actNode → cost;
25. end if
26. over ← left → C intersection right → C
27. if ((| over | / | left → C | < maxOverlap) AND (| over | / | right → C | < maxOverlap)) then
28.   actNode → state ← vicariant;
29.   return actNode fi cost;
30. end if
31. actNode → cost ← actNode → cost + noVicCost;
32. return actNode → cost;
```

END GetNodeCost

This is a basic implementation of this kind of function, and it is only used the first time the tree is optimized. To calculate the cost of modifying any node, it is possible to use incremental functions, such as those used for fast evaluation of rearrangements in a phylogenetic tree (e.g. Goloboff, 1996).

Note that, in contrast to phylogenetic analysis algorithms, just a downpass (to assign distributions to ancestral nodes, needed to deal with splits between internal nodes) is required, because two sister nodes are disjunct or not, regardless of the cells assigned to their ancestral nodes. Also note that ancestral nodes only get the cells (line 19) from nodes without a fully removed distribution (line 13).

cellSet: FindOverlap (actNode: Node)

This function looks for potential late overlap of the node actNode. If an overlap is found, then a cellSet containing the cells to remove is returned.

```
1. prevNode ← actNode → anc;
2. cells ← actNode → C intersection prevNode → C;
3. sister ← GetActiveSister (actNode);
4. if (sister != NULL)
5.   cells ← cells intersection (~ sister → C);
6. navNode ← prevNode → anc;
7. while (navNode != NULL)
8.   if ((navNode → state = = disjoint) OR (GetNumActive-
    Childs (navNode) != 2)) then
9.     cells ← cells intersection navNode → C;
```

```

10.     prevNode ← navNode;
11.     continue;
12. end if
13. sister ← GetActiveSister (prevNode);
14. over ← cells intersection sister;
15. if ((|over|/|actNode → C| < lateOverlap) AND (|over|/|
sister ← C| < lateOverlap)) then
16.     cells ← over;
17.     return cells;
18. end if
19. prevNode ← navNode;
20. navNode ← navNode → anc;
21. end while
22. return NULL;

```

END FindOverlap

The first ancestor of the node is skipped (i.e. updated before travelling towards the root, lines 1 and 2); if there is some overlap between actNode and its sister, it is removed from the cell set (line 5). This ensures that the overlap caused by the sister group must be evaluated by maxOverlap, and not by lateOverlap. For example, assume a tree and distribution (t1[ab] (t2[c] (t3[b] t4[bd]))). Without line 6, t1 and t4 share the cell [b], so cell b is removed; when subsequently checking the new reconstruction, (t3 t4) is now disjoint ([b] [d]), contradicting the maxOverlap value (if set below 50%).

The loop from lines 7 to 21 checks the descendants of each internal node in the route to the root. If nodes are polytomic, or already lead to disjunct descendants, they are ignored (lines 8–12), but the cellSet cells are updated (line 9). Otherwise, the overlapping is measured, if it is less than the lateOverlap value (line 15), the set of cells to remove (line 16) is returned (line 17).

The function GetActiveSister returns the sister of a node, if it is an active node (i.e. without a fully removed distribution) not belonging to a polytomy, or if it belongs to a polytomy where all the descendants but 2 are inactive.

REAL: NodeRemoval (actTree: Tree)

This function removes the distribution of a node one at a time, and evaluates whether such removal produces a better cost. The function returns the cost of the new reconstruction.

```

1. CopyArray (list, actTree fi nodeList);
2. foundBetter ← TRUE;
3. while (foundBetter) AND (j < list → size)
4.     foundBetter ← FALSE;
5.     bestCost ← tree → root → cost;
6.     bestPos ← 0;
7.     ShuffleList (list);

```

```

8.     for (j ← 1 to list → size)
9.         actNode ← list [j];
           /* Already removed */
10.        if (actNode → state = removed) then
11.            continue;
12.        end if
           /* Root node can not be removed */
13.        if (actNode → anc = NULL) then
14.            continue;
15.        end if
16.        tempState ← actNode → state;
17.        actNode → state ← removed;
18.        actCost ← GetCostNode (tree → root);
19.        if (actCost < bestCost) then
20.            foundBetter ← TRUE;
21.            bestCost ← actCost;
22.            bestPos ← j;
23.        end if
           /* Restores the original reconstruction */
24.        actNode → state ← tempState;
25.        GetCostNode (tree → root);
26.    end for
27.    if (foundBetter) then
28.        actNode → state ← removed;
29.        GetCostNode (tree → root);
30.    end if
31. end while
32. return tree → root → cost;

```

END NodeRemoval

As explained before, an incremental optimization provides an important time reduction during the search, because the distribution removal of each node requires two optimizations (lines 18 and 25). A potential improvement (not yet implemented) is to make a copy of all reconstructions and just reassign them in line 25 (when the new configuration is accepted, lines 27–30, the copy would have to be updated).

To evaluate partial distribution removals, the same function can be used, just adding a line to search the potential overlap and evaluate it, and another to restore the original distribution of the node.

This function runs until the cost cannot be improved. It is possible to change it to search for ties (not only improvements) of the cost. This could make the search somewhat slower, as it increases the time used by the function on sub-optimal solutions, but it could also be of great help for solutions that require simultaneous removals of distributions, or ambiguous data with several solutions.

To remove possible bias in the order of node evaluation, the list is randomized in each round before starting the distribution removal (line 7).