SPECIAL PAPER

Use of dispersal–vicariance analysis in biogeography – a critique

Ullasa Kodandaramaiah*

Department of Zoology, Stockholm University, Stockholm 106 91, Sweden

ABSTRACT

Aim Analytical methods are commonly used to identify historical processes of vicariance and dispersal in the evolution of taxa. Currently, dispersal–vicariance analysis implemented in the software DIVA is the most widely used method. Despite some recognized shortcomings of the method, it has been treated as error-free in many cases and used extensively as the sole method to reconstruct histories of taxa. In light of this, an evaluation of the limitations of the method is needed, especially in relation to several newer alternatives.

Methods In an approach similar to simulation studies in phylogenetics, I use hypothetical taxa evolving in specific geological scenarios and test how well DIVA reconstructs their histories.

Results DIVA reconstructs histories accurately when evolution has been simple; that is, where speciation is driven mainly by vicariance. Ancestral areas are wrongly identified under several conditions, including complex patterns of dispersals and within-area speciation events. Several potentially serious drawbacks in using DIVA for inferences in biogeography are discussed. These include the inability to distinguish between contiguous range expansions and across-barrier dispersals, a low probability of invoking extinctions, incorrect constraints set on the maximum number of areas by the user, and analysing the ingroup taxa without sister groups.

Main conclusions Most problems with inferences based on DIVA are linked to the inflexibility and simplicity of the assumptions used in the method. These are frequently invalid, resulting in spurious reconstructions. I argue that it might be dangerous to rely solely on DIVA optimization to infer the history of a group. I also argue that DIVA is not ideally suited to distinguishing between dispersal and vicariance because it cannot *a priori* take into account the age of divergences relative to the timing of barrier formation. I suggest that other alternative methods can be used to corroborate the findings in DIVA, increasing the robustness of biogeographic hypotheses. I compare some important alternatives and conclude that model-based approaches are promising.

*Correspondence: Ullasa Kodandaramaiah, Department of Zoology, Stockholm University, Stockholm 106 91, Sweden.
E-mail: ullasa.kodandaramaiah@zoologi.su.se

INTRODUCTION

One of the central aims of modern historical biogeography is to reconstruct the history of a particular taxon or a group of taxa based on phylogenetic hypotheses (Cox & Moore, 2005; Lomolino *et al.*, 2006). The advent of molecular datasets and improved analytical methods in phylogenetics has led to a proliferation of phylogenetic studies in the last two decades. Historical biogeography has been riding this molecular phylogenetic wave, and there has been a tremendous increase in the number of studies attempting to infer the history of various taxa. The increased reliability of available phylogenies,

the growing body of knowledge regarding the geological history of the Earth and techniques to date divergences between lineages have together opened the doors to unravelling the geographic context of diversification of life on Earth (Riddle *et al.*, 2008).

The processes that historical biogeography is concerned with are extinctions, range expansions and two kinds of allopatric speciation events – vicariance and dispersal (Futuyma, 1998). Speciation by vicariance is allopatric speciation *after the appearance* of a barrier, whereas speciation by dispersal is allopatric speciation following active colonization across a *pre-existing* barrier (Platnick & Nelson, 1978). Hence, in vicariance, colonization is assumed to have happened before the barrier formed, whereas in dispersal scenarios, colonization happens after barrier formation. I will hereafter use the term range expansion to refer to colonizations that lead to a widespread species and the term dispersal to refer to colonizations that result in speciation. Currently the majority of the studies attempting to identify vicariance, dispersal, range expansion and extinction events over the phylogeny do so using one of the many analytical methods available. One class of methods seeks to summarize area relationships based on general patterns among co-occurring clades and explain clade-specific deviations from general patterns as processes affecting individual clades (e.g. Brooks parsimony analysis: Brooks, 1990). 'Event-based' methods, on the other hand, analyse individual clades and attempt to identify historical events – vicariance, dispersal, extinction, and within-area speciation – that have affected their evolution (e.g. Ronquist & Nylin, 1990; Page, 1995; Ronquist, 1997; Ree *et al.*, 2005; Ree & Smith, 2008). Readers are referred to Crisci (2001) and Posadas *et al.* (2006) for a more detailed review of analytical methods in biogeography.

By far the most popular method in the past few years has been Ronquist's dispersal–vicariance analysis (Ronquist, 1997) implemented in the program DIVA (Ronquist, 1996). This method assigns a cost of one for colonizations and extinctions while assigning no cost for vicariance and within-area speciation events. The ancestral area with the least cost is deemed to be the most optimal. This relative weighting scheme implemented in DIVA favours hypotheses that maximize vicariance explanations while invoking only the minimum number of dispersal events, thus avoiding 'unparsimonious' dispersals. This is considered important since any hypothesis of origin can be explained by a suitable number of dispersals (Croizat *et al.*, 1974; Morrone & Crisci, 1995; Humphries & Parenti, 1999). The method allows for reticulate area history, and, in addition, the maximum number of areas at each node can be constrained prior to the analysis using the 'maxareas' option in the program. DIVA thus offers a simple quantitative technique for analysing distributions of extant taxa within a phylogenetic framework and estimating ancestral areas.

The method has gained rapid and widespread acceptance in the biogeographic community. The paper describing the algorithm has received 340 citations since it was published, and it continues to be a vital part of many biogeographic studies (53 citations in the year 2008; source: ISI Web of Knowledge, accessed in August 2009). Authors have used the ancestral areas reconstructed over their phylogenies to infer instances of vicariance and dispersal, in some cases without divergence times from dated nodes. Such inferences can be subject to various kinds of error, both in the reconstruction of ancestral areas by the method and in the interpretation of these areas by the user. Ronquist briefly mentions some pitfalls in dispersal–vicariance optimization in the documentation to the program and the paper describing it, with suggestions of how to overcome them. These have not always been followed. Deficiencies of the optimization process have been ignored in many cases, possibly with erroneous inferences made in several of these published studies. In this paper I highlight the potentially serious sources of error in this method, and include a more detailed discussion of some of the pitfalls pointed out by Ronquist.

Because the biogeographic history of any taxon is unknown it is impossible to test whether DIVA or any other method has correctly reconstructed its history. Simulation studies in phylogenetics have proved extremely useful for testing the accuracy of tree-reconstruction methods in recovering the known, albeit hypothetical, phylogeny of the simulated sequences (Poe & Swofford, 1999; Holland *et al.*, 2003; Ho & Jermiin, 2004). Here, I use a parallel approach to highlight circumstances under which inferences based on DIVA can be wrong. Hypothetical taxa are allowed to evolve in a specific geological scenario, and DIVA is used to estimate the ancestral areas.

SCENARIO

Assume four landmasses, A, B, C and D, that are connected to each other at some point in time. These landmasses separate following the advent of barriers between them (e.g. ocean), in the following sequence: A + (B + C + D) → A + B + (C + D) → A + B + C + D. Assume that a species *abcd'* was widespread in A, B, C and D and speciated into four species, *a'*, *b'*, *c'* and *d'*, as a result of allopatric speciation when the barriers formed between the areas (Fig. 1a). Let *mnop'* be found on A after the break-up of ABCD. It colonizes B across the barrier and speciates into *m'* and *nop'* in A and B, respectively. *nop'* colonizes C and splits into *n'* and *op'*.
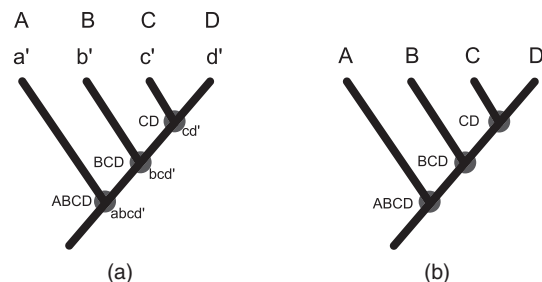
Figure 1 (a) History of clade abcd. (b) Reconstruction by DIVA, areas unconstrained. Dark circles indicate vicariance events.
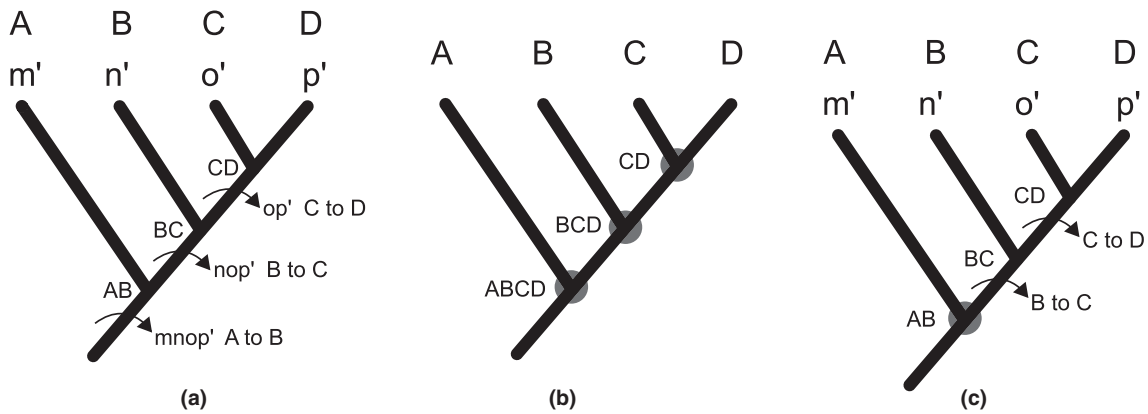
Figure 2 (a) History of clade mnop. (b) Reconstruction by DIVA, areas unconstrained. (c) Reconstruction with maxareas = 2. Filled circles indicate vicariance events. Arrows pointing to the right indicate across-barrier dispersals. The areas and species involved in the dispersal event are shown next to the arrows.

Similarly, op′ colonizes D and speciates, with the result that m′, n′, o′ and p′ are now found in A, B, C and D, respectively (Fig. 2a). Let ancestral species vwxzy′ be found on A before the break-up of ABCD. It speciates *in situ* in A to give v′ and wxyz′. wxyz′ colonizes BCD and eventually speciates when ABCD breaks up, such that w′, x′, y′ and z′ are now found in A, B, C and D, respectively, along with v′ in A (Fig. 3a).

DIVA correctly reconstructs the ancestral states for the clade abcd, and three instances of vicariance would be the 'most parsimonious' interpretation (Fig. 1b). For clade mnop, unless the maximum number of ancestral areas is constrained, DIVA reconstructs the same areas as for clade abcd, and three vicariance events would be inferred erroneously (Fig. 2b). If the maximum number of areas is constrained to two, DIVA recovers the areas correctly (Fig. 2c). The most parsimonious interpretation, however, would be vicariance between A and B followed by two dispersals. For clade vwxyz, an unconstrained DIVA analysis reconstructs the areas as shown in Fig. 3b. The most parsimonious inference would be one within-area speciation event and four vicariance events. If the maximum areas are constrained, DIVA wrongly reconstructs the areas in Fig. 3c and authors would infer at least two across-barrier dispersals. None of the possible colonization scenarios reflects the true history. If the clade wxyz is analysed without

constraints, DIVA would recover the areas as in Fig. 3a, and three vicariance events would be correctly inferred. If the maximum number of areas is constrained to two, areas are reconstructed as in Fig. 2c.

The above examples highlight several sources of error in DIVA. Four general points can be made, as follows.

1. *When speciation happens solely by vicariance, DIVA reconstructs ancestral areas without error.*

If speciation happened solely by vicariance, with shared patterns across different taxa, all methods would recover the history accurately (Wojcicki & Brooks, 2005). Biogeographic inference becomes confounding only when there have been extinctions, within-area speciation events, the formation of widespread species or speciation by dispersal. Simple vicariance scenarios are, however, not the rule; at the very least there must be range expansions for vicariance to act upon.

2. *DIVA does not distinguish between range expansion and across-barrier dispersals.*

DIVA cannot incorporate information about the time of barrier formation, and hence equates range expansion with speciation by dispersal (Brooks & McLennan, 2001). The two processes have different probabilities: jump-dispersals across a barrier that lead to new species have a lower likelihood of
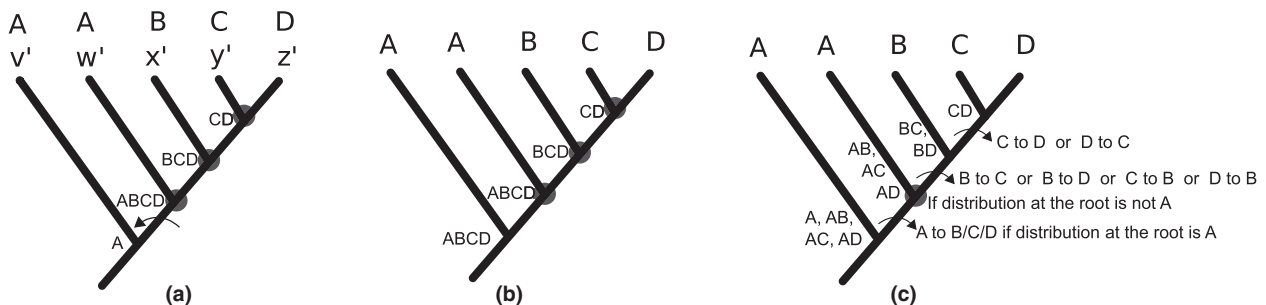
Figure 3 (a) History of clade vwxyz. (b) Reconstruction by DIVA, areas unconstrained. (c) Reconstruction with maxareas = 2. The arrow pointing to the left indicates three contiguous range expansions from A to occupy B, C and D. Filled circles indicate vicariance. Arrows pointing to the right indicate across-barrier dispersals. The areas and species involved in the dispersal event are shown next to the arrows.

occurrence than range expansions between connected areas. This can be a serious source of error in the optimization process, as illustrated in clade vwxyz. Here, if range expansions are assigned a lower cost, for example 0.2, and across-barrier dispersals 1, DIVA would reconstruct areas correctly. The basic DIVA protocol does not allow allocation of different costs to these two events. Even if differential costs can be allocated to the two events, it is necessary to introduce an additional assumption that geological events have the same effect on all lineages in the study group. Hence, taking into account geological events within DIVA would entail a much more complex model and a modification of the basic assumptions of the optimization protocol.

As a result of its failure to distinguish between range expansions and across-barrier dispersals, DIVA cannot by itself be used to distinguish between the two modes of speciation, even if the ancestral areas are reconstructed accurately and the colonization history inferred correctly. Some early studies attempted to distinguish between the modes of speciation without age estimates. With the increasing availability of methods to date phylogenies, a more common approach is to use the ancestral areas estimated by DIVA in combination with divergence-time estimates to detect vicariance events (e.g. Bremer, 2002; Van Bocxlaer et al., 2006; Kodandaramaiah & Wahlberg, 2009). Although this is preferable to completely ignoring age estimates, a priori incorporation of divergence times would improve the accuracy of optimization process significantly.

3. *DIVA is sensitive towards the exclusion of 'outgroup' taxa.*

Analysing the ancestral distribution of a clade without its sister group is biased towards inferring a widespread ancestor at the root (Ronquist, 1997). The example of clades wxyz and vwxyz illustrate this shortcoming. This leads to events of speciation by dispersal being erroneously considered as vicariance. Ronquist suggests including outgroups or restricting the maximum number of areas to two in the analysis to overcome this problem. Although the consequences of neglecting outgroups might be serious, relatively few studies include data from sister taxa to infer events within the ingroup. Many of these studies have inferred vicariance events at the base of the tree, where the true history could have been either vicariance or speciation by dispersal (e.g. Drovetski, 2003; Sanmartín, 2003; Bessega et al., 2006; Biswas & Pawar, 2006; Guo & Wang, 2007; Musilová et al., 2008).

4. *Wrongly constraining the number of maximum areas leads to spurious inferences of vicariance and dispersal.*

The lower the number of maximum areas, the higher the probability of inferring spurious dispersals, whereas the probability of spurious vicariance events increases when areas are unconstrained. Constraining the maximum number of areas is the most critical user-defined factor that can affect the results of the optimization process and it has been widely exploited (Ree et al., 2005); a perusal of 50 arbitrarily selected published works that have used DIVA showed that nearly half of them interpreted the results after constraining the maximum areas in their analysis, with various justifications for doing so. How many true vicariance events have been concealed by such analyses? The maximum number of areas a taxon can occupy is contingent on various factors, including the vagility of the taxon (Kodandaramaiah, 2009), the geological scenario, multiple ecological factors (e.g. predation pressure and abiotic factors) and chance events such as jump-dispersals. These factors change over time and affect the maximum range of a taxon differently at different times. How acceptable is the assumption that all nodes on the phylogeny have the same maximum number of areas occupied? I demonstrate elsewhere (Kodandaramaiah, 2009), with empirical examples, how varying the costs of different kinds of dispersal can lead to more realistic inferences with dispersal–vicariance optimization.

A related limitation of DIVA is its inability to restrict ancestral areas to one area. This often results in the ancestors being inferred as widespread (in two areas) even if most or all extant taxa in the analysis are restricted to a single area. By virtue of this and the relative weighting scheme favouring vicariance over dispersal, the method is on the whole biased towards vicariance.

I reanalyse a published dataset to illustrate this bias towards vicariance. Knapp et al. (2005) inferred the phylogeny of the plant genus *Nothofagus*, which is distributed disjunctly in South America, Australia and New Zealand (Fig. 2 in their paper). The three landmasses were connected to each other as part of Gondwana. New Zealand rifted from Gondwana c. 80 Ma, whereas South America and Australia remained in contact until c. 35 Ma, after which they separated (McLoughlin, 2001; and other references in Knapp et al., 2005). My DIVA analysis on their dataset indicates that the ancestor was widespread over the three landmasses and successive vicariance events drove divergence within the group. However, the phylogeny shows that the divergence between the Australian and New Zealand species happened after the split between the Australian and South American species. This sequence is incompatible with the sequence of Gondwanan break-up, and hence the hypothesis that vicariance is the sole mediator of speciation is refuted. Knapp et al. (2005) used a relaxed molecular clock approach to date their tree and found that the divergence between the Australian and New Zealand species happened in the Oligocene, leaving dispersal as the only viable alternative. Other authors have come to the same conclusion independently (Hill & Jordan, 1993; Martin & Dowd, 1993; Manos, 1997; Cook & Crisp, 2005a).

Other problems

The effects of extinction can be profound, but they are not commonly considered in biogeographic studies (Lieberman, 2002). Extinction is easily the most under-emphasized process in DIVA. Modelling extinction realistically is a daunting task owing to the complexity of ecological factors involved. DIVA, with its simplistic model, can lead to unrealistic inferences of vicariance and dispersal when the clade has experienced large-scale or non-random (with respect to areas) extinctions. Local

extinctions are probably as important as the three other processes in the evolution of distributions but are never inferred unless the areas are constrained (documentation to DIVA; Ronquist, 1996), and hence the inability to take into account extinctions can result in spurious colonization patterns being inferred. Indeed, to my knowledge extinctions have never been inferred in DIVA optimizations in any study. Lieberman (2002) discusses problems posed by extinctions with regard to general-area-cladogram-based methods and suggests some precautions to minimize the effects of extinction. Although DIVA is based on fundamentally different principles, these precautions are also applicable to its usage.

DIVA in its current form relies on a completely bifurcating tree. In the case of polytomous nodes and consensus trees, Ronquist suggests summarizing the results of individual DIVA analyses on all possible fully bifurcate trees and the trees from which the consensus is calculated, respectively. Nylander et al. (2008) offer a better solution to the problem in their new method, which they call BAYES–DIVA. This method exploits the fact that the Markov chain Monte Carlo run during Bayesian phylogenetic inference results in several hundreds of trees rather than a single most-optimal tree. DIVA analyses are run over a randomly chosen subset of these trees, and the frequencies of ancestral areas at each node are used to infer events.

DIVA has also been used in phylogeographic studies to infer the history of populations (e.g. Vidya et al., 2009). The conceptual underpinnings of DIVA are essentially invalidated at the population level as one cannot assume that population fragmentation and within-area population subdivision have a higher likelihood of occurrence than colonization or extinction of local populations. Hence, DIVA in its current form is not an optimal analytical tool for phylogeographic studies.

It must be stressed that the examples used here have been chosen expressly to highlight the kinds of scenarios for which DIVA analyses are potentially misleading. They do not necessarily represent the most common evolutionary scenarios. Moreover, any method that relies solely on information from extant distributions of members within a single clade will find it difficult to reconstruct the complex histories of the hypothetical clades described here. However, DIVA has several significant advantages over previously proposed methods (Sanmartín, 2007). It is simple to use and based on straightforward assumptions. It allows for reticulate area history, thus not enforcing a singular hierarchical area history (Bremer, 1992; Ronquist, 1994). Unlike the maximum-vicariance approach, colonizations are treated as integral components of evolution of distributions. Compared with character optimization methods, widespread species are treated more realistically. Furthermore, DIVA analyses on several co-occurring taxa can also be used to explore general area relationships (e.g. Sanmartín et al., 2001), especially when age estimates are available. These factors make DIVA an attractive option in historical biogeography, and this is evidenced by the huge popularity of the method.

Unfortunately, methodological complications have been ignored, and the output from DIVA analyses is frequently treated as error-free. Certain precautions can be taken to decrease the probability of incorrect inferences when using DIVA. It is imperative to include sister groups in the analysis, as suggested by Ronquist. Wherever possible, analyses with different levels of constraint on the number of areas should be considered, instead of simply using either the unconstrained analysis or the one with areas constrained to two. It is also necessary to take into account that a dispersal event immediately prior to the origin of the group being analysed (including sister groups) cannot be detected by DIVA and will be recovered as a vicariance event. Hence caution must be exercised when invoking vicariance events at the base of the tree. Because the DIVA optimization is especially sensitive to extinctions, it is arguably better to refrain from using DIVA when large-scale extinctions are suspected. Similarly, it might be dangerous to use DIVA in population-level analyses. Sanmartín & Ronquist (2002) discuss the problems presented by widespread species in DIVA analyses and suggest solutions to overcome them.

From a methodological perspective, several improvements can be made in order to obtain consistently more realistic inferences. One of the major limitations of the current version of the program is its rigidity with respect to relative costs. The option of specifying user-defined costs may significantly improve the realism of the reconstructions (de Queiroz, 2005). Perhaps information from phylogeographic data (Riddle & Hafner, 2006; Riddle, 2009) or metapopulation studies can be integrated, wherever it is available, to specify such costs. The a priori incorporation of age estimates and decoupling of contiguous range expansions from across-barrier dispersals are important. The option of restricting 'maxareas' to 1 is equally critical. Most importantly, the probability of extinctions needs to be modelled better.

One can have increased confidence in inferences if the results are supported by other analytical methods. A number of methods have been published since DIVA, and all methods have their strengths and weaknesses. A comprehensive overview of all methods is beyond the purview of this article, but I discuss some important alternative methods proposed after DIVA.

Hausdorf's weighted ancestral area analysis (WAAA; Hausdorf, 1998) is a parsimony-based method with several drawbacks in common with dispersal–vicariance analysis. Readers are referred to the original paper in which the author compares the two methods and discusses the differences in assumptions between them. Hovenkamp's vicariance analysis (Hovenkamp, 1997, 2001) is another parsimony-based method that places importance on the actual distributions of taxa rather than assigning each taxon to a set of predefined areas. Distribution maps of taxa are used to detect vicariant patterns at each node. Nodes at which vicariance events cannot be traced are ignored, as the method deals with the 'elucidation of Earth history and does not try to reconstruct all the events in the course of the history of each taxon' (Hovenkamp, 1997). Whenever DIVA

suggests a vicariance-only scenario, this method might be used to corroborate and refine the hypothesis.

General-area-cladogram-based methods rely on the assumption that vicariance affects all co-occurring taxa similarly, whereas jump dispersals, range expansions, extinctions and within-area speciation events result in clade-specific patterns. Distributional data from three or more clades are analysed to derive a general area cladogram, which is assumed to reflect the geological history of the areas (e.g. secondary Brooks parsimony analysis: Brooks, 1990; component analysis: Nelson & Platnick, 1981). Clade-specific processes can be inferred from this general area cladogram. Most methods introduce an additional constraint that the clades in question evolved contemporaneously (Donoghue & Moore, 2003), and ignore directionality of dispersals (Cook & Crisp, 2005b). PACT (phylogenetic analysis for comparing trees: Wojcicki & Brooks, 2005) is an exception in that it does not assume that the clades evolved together throughout their history. Using the rationale proposed by Lieberman (2000, 2003a,b), it also possible, in theory, to disentangle general patterns resulting from geodispersals (i.e. concordant colonizations) from those created by vicariance (see Folinsbee & Brooks, 2007, for an example). Although potentially every clade has an idiosyncratic evolutionary history, it is likely that geological and climatic events will have affected several clades in the same way. In light of this, seeking shared patterns amounts to searching for corroboration and support of a biogeographic hypothesis, especially vicariance and geodispersals. In a similar fashion, independent DIVA results from multiple taxa occurring in the area of interest can be compared to detect shared patterns of vicariance or geodispersals.

Sanmartín et al. (2008) developed a model-based method exclusively for scenarios in which dispersal is the main force behind speciation. This method can be used to describe general patterns in a set of areas as opposed to estimating specific processes for a clade, and is particularly suited for islands and other dispersal-mediated systems, which have long been neglected by vicariance-based methods. This approach can only be used when data are available from several clades, and cannot tell us much about clade-specific processes.

The approach used by Ree and colleagues (Ree et al., 2005; Ree & Smith, 2008) is promising for general usage in biogeographic studies when a molecular phylogeny is available for the taxon of study. It was first proposed in Ree et al. (2005) and later refined in Ree & Smith (2008) with modifications to the computational procedure. The method is based on a complex but flexible stochastic model, the DEC (dispersal extinction cladogenesis) model, in which evolution of geographic ranges both along a branch (anagenetic change) and at nodes (cladogenetic change) are modelled explicitly. In contrast to the area-cladogram approach used in DIVA, which tries to estimate the most likely ancestral distributional state of clades, this method attempts to estimate how ancestral areas were inherited by daughter lineages. Vicariance is not favoured over speciation by dispersal, in recognition of the fact that

colonization and vicariance are equally important to speciation.

The advantages of the DEC model over the DIVA model are its greater complexity and flexibility. A larger number of factors affecting the evolution of distributions, both physical and biological, are parameterized in the model. The physical component includes the set of areas and connections between them at different points in time. The biological component includes a time-calibrated phylogeny of the species in question, and information on rates of dispersal between areas and on extinction within an area. The flexibility allows the user to effectively incorporate available knowledge about the study group a priori. For instance, the relative probability of dispersal between any two areas at different points in time can be set, thereby overcoming the severe limitation of DIVA of assuming that the likelihood of dispersal is constant between all areas and at all times. Consequently, range expansions between contiguous areas can be assigned a higher probability than across-barrier dispersals. This also allows the user to incorporate information on the absence of the defined areas during certain points in time. For example, if dealing with a recently emerged volcanic island, we can infer that it was unavailable for occupation before its emergence. In this case, the prior probability of dispersal into and out of the area can be set to zero. Constraints can be imposed on the number of areas occupied by ancestors at different points in time, in contrast to the case for DIVA, in which only a fixed constraint across all lineages can be imposed. This is done in the form of prior probability distributions of ancestral range sizes. Ancestors can also be constrained to a single area, whereas they can only be constrained to two or more areas in DIVA.

The parameterized model is evaluated in a likelihood framework, and the ancestral range inheritance scenario with the best likelihood of observing the current species distributions is chosen as optimal. Other scenarios of inheritance at nodes are treated as plausible alternative explanations if their likelihood scores are not significantly lower than the most optimal scenario. This has been implemented in the software LAGRANGE (available at http://code.google.com/p/lagrange). The software is not as user-friendly as DIVA at the time of writing, but is under active development and it can be expected that it will become easier to use.

Because the DEC method is quite new, very few empirical studies have compared it with DIVA. Ree et al. (2005) compared the two methods on both hypothetical and empirical examples. They concluded that results from the two methods converged when the following parameter conditions were met: constant rates of dispersal and extinction through time, low dispersal rates, much lower extinction rates, and equal times between divergences. Whereas DIVA never inferred extinctions, DEC did so rarely. DIVA optimizations generally resulted in fewer reconstructions and were predictably unresponsive to varying prior expectations of dispersal and extinction. The DEC model allowed daughter species to inherit ancestral ranges, whereas DIVA always forced subdivision of the ancestral range. However, both methods

had a tendency to infer widespread ancestors. Clark *et al.* (2008) tested the two methods to reconstruct the biogeography of a plant genus on an island system. They observed that improbably wide ancestral areas were recovered more commonly in DIVA, despite the maximum number of areas being constrained to two. Clayton *et al.* (2009) reported similar findings. Further studies on systems with varying complexity will shed more light on how this method performs in practice, but it is promising in theory.

A drawback common to both methods is the assumption that anagenetic change along branches is independent of cladogenetic change at nodes (Ree *et al.*, 2005). Speciation links the two processes, and the mode of speciation should be expected to affect ancestor–descendant range. It is unclear to what extent ancestral range reconstruction scenarios are affected by the fact that the mode of speciation is ignored. Like DIVA, the DEC approach is prone to inferring widespread ancestors, and hence inclusion of the sister group to the clade of interest in analyses is encouraged. Another limitation at present is that DEC cannot take into account phylogenetic uncertainty. As Ree and colleagues acknowledge, model-based historical biogeography is still in its infancy, but the flexibility of the approach means that future models will be able to factor in explicitly the mode of speciation and phylogenetic uncertainty.

SUMMARY AND CONCLUSIONS

Currently, dispersal–vicariance analysis (DIVA) is the most widely used method in historical biogeography. I use hypothetical examples to highlight some scenarios in which reliance on DIVA can lead to spurious inferences of vicariance and dispersal. Although temporal information on lineage divergences and geology can be used *a posteriori* to distinguish between vicariance and dispersal, more complex models that can use them *a priori* in analyses are needed. I suggest that relying on a single DIVA analysis can be potentially dangerous, and it is hence important to compare the results with those obtained from other methods. Of all the alternatives, the maximum-likelihood approach based on the DEC model (Ree & Smith, 2008) is possibly the most promising.

ACKNOWLEDGEMENTS

REFERENCES

Bessega, C., Vilardi, J.C. & Saidman, B.O. (2006) Genetic relationships among American species of the genus *Prosopis* (Mimosoideae, Leguminosae) inferred from ITS sequences: evidence for long-distance dispersal. *Journal of Biogeography*, **33**, 1905–1915.

Biswas, S. & Pawar, S.S. (2006) Phylogenetic tests of distribution patterns in South Asia: towards an integrative approach. *Journal of Biosciences*, **31**, 95–113.

Bremer, K. (1992) Ancestral areas: a cladistic reinterpretation of the center of origin concept. *Systematic Biology*, **41**, 436–445.

Bremer, K. (2002) Gondwanan evolution of the grass alliance of families (Poales). *Evolution*, **56**, 1374–1387.

Brooks, D.R. (1990) Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Systematic Zoology*, **39**, 14–30.

Brooks, D.R. & McLennan, D.A. (2001) A comparison of a discovery-based and an event-based method of historical biogeography. *Journal of Biogeography*, **28**, 757–767.

Clark, J.R., Ree, R.H., Alfaro, M.E., King, M.G., Wagner, W.L. & Roalson, E.H. (2008) A comparative study in ancestral range reconstruction methods: retracing the uncertain histories of insular lineages. *Systematic Biology*, **57**, 693–707.

Clayton, J.W., Soltis, P.S. & Soltis, D.E. (2009) Recent long-distance dispersal overshadows ancient biogeographical patterns in a pantropical angiosperm family (Simaroubaceae, Sapindales). *Systematic Biology*, **58**, 395–410.

Cook, L.G. & Crisp, M.D. (2005a) Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2535–2544.

Cook, L.G. & Crisp, M.D. (2005b) Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *Journal of Biogeography*, **32**, 741–754.

Cox, C.B. & Moore, P.D. (2005) *Biogeography: an ecological and evolutionary approach*, 7th edn. Blackwell, Oxford, UK.

Crisci, J.V. (2001) The voice of historical biogeography. *Journal of Biogeography*, **28**, 157–168.

Croizat, L., Nelson, G. & Rosen, D.E. (1974) Centers of origin and related concepts. *Systematic Zoology*, **23**, 265–287.

Donoghue, M.J. & Moore, B.R. (2003) Toward an integrative historical biogeography. *Integrative and Comparative Biology*, **43**, 261–270.

Drovetski, S.V. (2003) Plio-Pleistocene climatic oscillations, Holarctic biogeography and speciation in an avian subfamily. *Journal of Biogeography*, **30**, 1173–1181.

Folinsbee, K.E. & Brooks, D.R. (2007) Miocene hominoid biogeography: pulses of dispersal and differentiation. *Journal of Biogeography*, **24**, 383–397.

Futuyma, D.J. (1998) *Evolutionary biology*. Sinauer Associates, Sunderland, MA.

Guo, X. & Wang, Y. (2007) Partitioned Bayesian analyses, dispersal–vicariance analysis, and the biogeography of Chinese toad-headed lizards (Agamidae: Phrynocephalus): a re-evaluation. *Molecular Phylogenetics and Evolution*, **45**, 643–662.

Hausdorf, B. (1998) Weighted ancestral area analysis and a solution of the redundant distribution problem. *Systematic Biology*, **47**, 445–456.

Hill, R.S. & Jordan, G.J. (1993) The evolutionary history of *Nothofagus* (Nothofagaceae). *Australian Systematic Botany*, **6,** 111–126.

Ho, S. & Jermiin, L. (2004) Tracing the decay of the historical signal in biological sequence data. *Systematic Biology*, **53,** 623–637.

Holland, B.R., Penny, D. & Hendy, M.D. (2003) Outgroup misplacement and phylogenetic inaccuracy under a molecular clock – a simulation study. *Systematic Biology*, **52,** 229–238.

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.

Humphries, C.J. & Parenti, L.R. (1999) *Cladistic biogeography: interpreting patterns of plant and animal distributions*, 2nd edn. Oxford University Press, Oxford.

Knapp, M., Stöckler, K., Havell, D., Delsuc, F., Sebastiani, F. & Lockhart, P.J. (2005) Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). *PLoS Biology*, **3,** 38–43.

Kodandaramaiah, U. (2009) Vagility: the neglected component in historical biogeography. *Evolutionary Biology*, **36,** 327–335.

Kodandaramaiah, U. & Wahlberg, N. (2009) Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) – patterns of colonization in the Holarctic. *Systematic Entomology*, **34,** 315–323.

Lieberman, B.S. (2000) *Paleobiogeography*. Plenum/Kluwer Academic, New York.

Lieberman, B.S. (2002) Phylogenetic biogeography with and without the fossil record: gauging the effects of extinction and paleontological incompleteness. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **178,** 39–52.

Lieberman, B.S. (2003a) Paleobiogeography: the relevance of fossils to biogeography. *Annual Review of Ecology, Evolution, and Systematics*, **34,** 51–69.

Lieberman, B.S. (2003b) Unifying theory and methodology in biogeography. *Evolutionary Biology*, **33,** 1–25.

Lomolino, M.V., Riddle, B.R. & Brown, J.H. (2006) *Biogeography*, 3rd edn. Sinauer Associates, Sunderland, MA.

Manos, P.S. (1997) Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. *American Journal of Botany*, **84,** 1137–1150.

Martin, P.G. & Dowd, J.M. (1993) Using sequences of *rbc*L to study phylogeny and biogeography of *Nothofagus* species. *Australian Systematic Botany*, **6,** 441–447.

McLoughlin, S. (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany*, **49,** 271–300.

Morrone, J.J. & Crisci, J.V. (1995) Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics*, **26,** 373–401.

Musilová, Z., Říčan, O., Janko, K. & Novák, J. (2008) Molecular phylogeny and biogeography of the Neotropical cichlid fish tribe Cichlasomatini (Teleostei: Cichlidae: Cichlasomatinae). *Molecular Phylogenetics and Evolution*, **46,** 659–672.

Nelson, G. & Platnick, N.I. (1981) *Systematics and biogeography: cladistics and vicariance*. Columbia University Press, New York.

Nylander, J.A.A., Olsson, U., Alström, P. & Sanmartín, I. (2008) Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal–vicariance analysis of the thrushes (Aves: *Turdus*). *Systematic Biology*, **57,** 257–268.

Page, R.D.M. (1995) Parallel phylogenies: reconstructing the history of host–parasite assemblages. *Cladistics*, **10,** 155–173.

Platnick, N.I. & Nelson, G. (1978) A method of analysis for historical biogeography. *Systematic Zoology*, **27,** 1–16.

Poe, S. & Swofford, D. (1999) Taxon sampling revisited. *Nature*, **398,** 299–300.

Posadas, P., Crisci, J.V. & Katinas, L. (2006) Historical biogeography: a review of its basic concepts and critical issues. *Journal of Arid Environments*, **66,** 389–403.

de Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution*, **2,** 68–73.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **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.

Riddle, B.R. (2009) What is modern biogeography without phylogeography? *Journal of Biogeography*, **36,** 1–2.

Riddle, B.R. & Hafner, D.J. (2006) A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments*, **66,** 435–461.

Riddle, B.R., Dawson, M.N., Hadly, E.A., Hafner, D.J., Hickerson, M.J., Mantooth, S.J. & Yoder, A.D. (2008) The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography*, **32,** 173–202.

Ronquist, F. (1994) Ancestral areas and parsimony. *Systematic Biology*, **43,** 267–274.

Ronquist, F. (1996) *DIVA, version 1.1*. Computer program and manual available by anonymous FTP from Uppsala University, Uppsala, Sweden. Available at: ftp.uu.se or ftp.systbot.uu.se.

Ronquist, F. (1997) Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46,** 195–203.

Ronquist, F. & Nylin, S. (1990) Process and pattern in the evolution of species associations. *Systematic Zoology*, **39,** 323–344.

Sanmartín, I. (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography*, **30,** 1883–1897.

Sanmartín, I. (2007) Event-based biogeography: integrating patterns, processes, and time. *Biogeography in a changing*

world (ed. by M.C. Ebach and R.S. Tangey), pp. 135–159. Taylor & Francis, London.

Sanmartín, I. & Ronquist, F. (2002) New solutions to old problems: widespread taxa, redundant distributions and missing areas in event-based biogeography. *Animal Biodiversity and Conservation*, **25,** 75–93.

Sanmartín, I., Enghoff, H. & Ronquist, F. (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, **73,** 345–390.

Sanmartín, I., Mark, P.V.D. & Ronquist, F. (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35,** 428–449.

Van Bocxlaer, I., Roelants, K., Biju, S.D., Nagaraju, J. & Bossuyt, F. (2006) Late Cretaceous vicariance in Gondwanan amphibians. *PLoS ONE*, **1,** e74.

Vidya, T.N.C., Sukumar, R. & Melnick, D.J. (2009) Range-wide mtDNA phylogeography yields insights into the origins of Asian elephants. *Proceedings of the Royal Society B: Biological Sciences*, **276,** 893–902.

Wojcicki, M. & Brooks, D.R. (2005) PACT: an efficient and powerful algorithm for generating area cladograms. *Journal of Biogeography*, **32,** 755–774.

BIOSKETCH

The author is currently part of the Nymphalidae Systematics Group (http://www.nymphalidae.utu.fi) and is interested in evolutionary patterns in butterflies. Apart from historical biogeography, he has studied wing pattern evolution and host–plant interactions. He has also undertaken experimental studies on these topics.

Editor: John Lambshead