Journal of Biogeography, **30**, 319–329

GUEST EDITORIAL

Cladistic and phylogenetic biogeography: the art and the science of discovery

ABSTRACT

All methods used in historical biogeographical analysis aim to obtain resolved area cladograms that represent historical relationships among areas in which monophyletic groups of taxa are distributed. When neither widespread nor sympatric taxa are present in the distribution of a monophyletic group, all methods obtain the same resolved area cladogram that conforms to a simple vicariance scenario. In most cases, however, the distribution of monophyletic groups of taxa is not that simple. *A priori* and *a posteriori* methods of historical biogeography differ in the way in which they deal with widespread and sympatric taxa. *A posteriori* methods are empirically superior to *a priori* methods, as they provide a more parsimonious accounting of the input data, do not eliminate or modify input data, and do not suffer from internal inconsistencies in implementation. When factual errors are corrected, the exemplar presented by M.C. Ebach & C.J. Humphries (*Journal of Biogeography,* 2002, **29**, 427) purporting to show inconsistencies in implementation by *a posteriori* methods actually corroborates the opposite. The rationale for preferring *a priori* methods thus corresponds to ontological rather than to epistemological considerations. We herein identify two different research programmes, cladistic biogeography (associated with *a priori* methods) and phylogenetic biogeography (associated with *a posteriori* methods). The aim of cladistic biogeography is to fit all elements of all taxon–area cladograms to a single set of area relationships, maintaining historical singularity of areas. The aim of phylogenetic biogeography is to document, most parsimoniously, the geographical context of speciation events. The recent contribution by M.C. Ebach & C.J. Humphries (*Journal of Biogeography,* 2002, **29**, 427) makes it clear that cladistic biogeography using *a priori* methods is an inductivist/verificationist research programme, whereas phylogenetic biogeography is hypothetico-deductivist/falsificationist. Cladistic biogeography can become hypothetic-deductive by using a *posteriori* methods of analysis.

Keywords

Historical biogeography, phylogenetic biogeography, cladistic biogeography, *a priori* and *a posteriori* methods, assumptions 0, 1, 2, inclusion, widespread taxa, sympatric taxa, paralogy, common pattern, evolutionary scenarios, vicariance, dispersal, extinction, ontology, epistemology.

INTRODUCTION

In a series of recent studies, Van Veller *et al.* (1999, 2000, 2001, 2002) investigated some of the empirical properties of different methods of historical biogeographical analysis. They developed a criterion by which the internal consistency of each method could be assessed, and were able to identify two categories of methods. They found that each method presented different problems with respect to internal consistency and proposed ways in which those shortcomings could be fixed. In doing this work, Van Veller *et al.* (1999, 2000, 2001, 2002) discovered that the various methods belonged to two distinct classes of methods, which they characterized as *a priori* and *a posteriori* methods. Van Veller & Brooks (2001) suggested that these methods represent different research programmes stemming from different ontological views of the world.

Ebach (2001) and Ebach & Humphries (2002) recently criticized all these efforts. We believe their commentary conflates epistemological and ontological considerations, and contains some factual errors. In this paper we hope to correct the errors and clearly distinguish between epistemological (methodological) considerations and ontological considerations to focus attention on the fundamental philosophical differences that have led to the emergence of two very different approaches to historical biogeography. We will suggest that points of mutual agreement can be viewed as benefiting both research programmes, and that points of disagreement reflect differences in research programmes rather than fundamental flaws in methods.

© 2003 Blackwell Publishing Ltd

EPISTEMOLOGICAL CONSIDERATIONS

Modern historical biogeography stems from a seminal paper by Nelson (1969) who proposed that the fundamental problem with historical biogeography was the lack of a method for discovering general patterns of geographical distributions, which he defined as area relationships supported by the geographical distributions of sister taxa representing multiple clades. Rosen (1978) was the first biogeographer to derive area cladograms from taxon phylogenies for inferring vicariant patterns and historical explanation of these patterns. He arrived at these area cladograms by replacing taxa in a taxon phylogeny of a monophyletic group by the areas in which these taxa are distributed. If each taxon is present in only a single area (no widespread taxa) and each area is inhabited by only a single taxon (no sympatric taxa), this replacement produces a taxon–area cladogram (TAC *sensu* Morrone & Carpenter, 1994; Enghoff, 1996) that only represents vicariant patterns and that can be completely explained by vicariant speciation (passive allopatric speciation: Brooks & McLennan 2002).

When sympatric or widespread taxa are present, however, replacing taxa by areas results in a taxon–area cladogram with either the same area at different terminal nodes (because of sympatric taxa) or more than one area at the same terminal node (due to a widespread taxon). It is in these cases that subsequent analyses are necessary in order to resolve the taxon–area cladogram into a resolved area cladogram (RAC *sensu* Morrone and Carpenter, 1994; Enghoff, 1996). Rosen (1978) choose to reduce the taxon–area cladograms by removing elements that represented the same area at different terminal nodes or more than one area at the same terminal node. Other biogeographers have since developed different methods [Component Analysis (CA) by Nelson & Platnick (1981); Page (1988); Reconciled Tree Analysis (RTA) by Page (1993a); Three Area Statement Analysis (TAS) by Nelson & Ladiges (1991a,b); Component Compatibility Analysis (CCA) by Zandee & Roos (1987); Brooks Parsimony Analysis (BPA) by Brooks (1981, 1985, 1990); Brooks *et al.* (2001); Wiley (1986, 1988a,b); Paralogy-Free Subtree Analysis (PSA) by Nelson & Ladiges, 1996] for obtaining resolved area cladograms from taxon–area cladograms that include widespread or sympatric taxa. Most of these different methods have been implemented in various computer programs [e.g. Component 1.5 by Page (1990a); Component 2.0 by Page (1993b); CAFCA by Zandee (1999); TAS by Nelson & Ladiges (1991c); TASS by Nelson & Ladiges (1995)].

It has generally been assumed that all these methods have been developed to implement the same research programme. If this is true, then one way to choose among them is to determine which one best optimizes the goals of the research programme. Van Veller *et al.* (2000) discovered that the different methods belong to two different categories based on the different protocols that they use for deriving area cladograms from taxon phylogenies and distributional data of the included taxa. The distinction between *a priori* and *a posteriori* methods is fully discussed in Van Veller *et al.* (2002). Briefly, *a priori* methods allow modification of data in the taxon–area cladogram by invoking certain assumptions to justify modifying input data in order to provide maximum fit of widespread and sympatric taxa to a single general area cladogram. *A posteriori* methods do not allow any modification of data in the taxon–area cladograms and deal with widespread and sympatric taxa in a parsimony analysis of the unmodified input cladograms. The most parsimonious depiction of all the data is selected as the general area cladogram and the taxa whose distributions conflict with that pattern are explained *a posteriori* as post-speciation dispersal or speciation by dispersal (peripheral isolates speciation). These explanations are obtained by optimizing the data of each monophyletic group in the general area cladogram (CCA and primary BPA) or by duplicating the areas in which the incongruent distributions occur (secondary BPA: for distinctions between primary and secondary BPA, see Brooks *et al.*, 2001; Van Veller & Brooks, 2001; Brooks & McLennan, 2002).

Once Van Veller *et al.* recognized that *a priori* and *a posteriori* methods have different optimality criteria (maximizing congruence with a general area cladogram vs. maximum parsimony), they proposed a third criterion that could be applied to all methods equally. Each of the different methods infers resolved area cladograms from taxon–area cladograms under one (or more) of the three different assumptions, called assumption 0 (A0), assumption 1 (A1) and assumption 2 (A2). Van Veller *et al.* (1999) connected these three assumptions to three sets of processes (A0: vicariance, A1: vicariance + extinction, A2: vicariance + extinction + dispersal) that can be invoked for explaining the presence of widespread or sympatric taxa. Based on these three sets of processes, assumed under A0, A1 and A2, Van Veller *et al.* (1999) developed two requirements for internal consistency for all methods with respect to their implementation of the assumptions and their capacity for finding general area cladograms for different monophyletic groups of taxa. Under requirement I inclusive sets of resolved area cladograms [i.e. solution sets *sensu* Van Veller *et al.* (1999)] should be derived under A0,

A1 and A2 when the different processes underlying these assumptions are *a priori* considered to be independent in their effect. Under requirement II, valid general area cladograms are obtained by comparison of sets of resolved area cladograms obtained for different monophyletic groups under the same assumption.

The *a priori* methods defended by Ebach & Humphries (2002) rely on A1 and A2. For dealing with widespread taxa A1 states (Nelson & Platnick, 1981; p. 421): '… whatever is true of the one occurrence is also true of the other occurrence' whereas A2 states (Nelson & Platnick, 1981; p. 432): '… whatever is true of the one occurrence might not be true of the other occurrence.' These two statements indicate an increase in the number arrangements that are considered to be true when moving from A1 towards A2 for dealing with widespread taxa. This increase in the number of arrangements corresponds with an increase in the number of resolved area cladograms in a solution set (Page, 1990b; Morrone & Crisci, 1995).

Van Veller & Brooks (2001; table 1, p. 6) showed that the number of resolved area cladograms obtained under A2 grows exponentially with an increase in the number of widespread taxa per taxon–area cladogram. This is caused by an increase in the number of possible arrangements under A2 for each widespread taxon that is added to the taxon–area cladogram. Together with the explosion in the number of resolved area cladograms (resulting in large solution sets) the chance of finding common patterns in the intersection of solution sets derived for different monophyletic groups under A2 approaches certainty, rendering the falsifiability of *a priori* methods, a point to which we will return below.

Different arrangements under A1 and A2 correspond to pruning or adding taxa to the taxon–area cladogram (Van Veller *et al.*, 1999, 2000, 2001). Ebach (2001) and Ebach & Humphries (2002) object to the claim that data are removed *a priori* under A2 in particular. They purport to show this by using examples that consist of taxon–area cladograms with paralogy in terminal nodes that branch off sequentially or terminal nodes that form a clade (Fig. 2 in Ebach, 2001; Figs 2 and 5 in Ebach & Humphries, 2002). Van Veller *et al.* (1999), however, showed that for such patterns of paralogy, inclusion or exclusion of redundant distributions does not change the topology of the area cladogram (see Fig. 1), so the exemplars by Ebach & Humphries (2002) are moot. More importantly, it is for the combination of paralogy (sympatric taxa *sensu* Van Veller *et al.*, 1999) and widespread taxa that A2 dictates *a priori* removal of data:

- For CA: 'The first step of assumption 2 is the removal of any occurrences of widespread taxa in areas that are occupied by endemic taxa.' (Page, 1988; p. 269)
- For RTA: 'If we delete from the range of each widespread taxon those areas about whose relationships the genera disagree (i.e. 3, 6 and 9), then the relationships of those areas will be determined by the relationships of the taxa endemic to those areas.' (Page, 1993b; p. 73)
- For TAS: 'For area D, the conflict is between widespread and endemic taxa. Assumption two resolves such conflict on behalf of endemics, eliminating area D (in parentheses) from association with areas BCG (Fig. 5j).' (Nelson & Ladiges, 1991b; p. 479)
- For PSA: 'If a node leads directly to one or more terminal taxon that is geographically widespread, and part of that distribution overlaps with that of another taxon, or taxa, then the widespread distribution is reduced to the nonoverlapping geographical element.' (Nelson & Ladiges, 1996; pp. 3, 4)

Ebach (2001) and Ebach & Humphries (2002) are thus factually incorrect on this issue.

Ebach & Humphries (2002) advocate PSA in combination with A2 for dealing with widespread taxa and reducing paralogy. They presented an exemplar for a taxon–area cladogram with two widespread

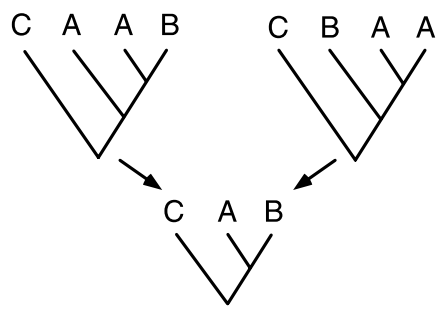

Figure 1 Solution of paralogy (caused by sympatric taxa) by collapsing nodes (A–C = areas).

and three sympatric taxa (reproduced in Fig. 2a). Under A0 and A2, Ebach and Humphries obtain two different resolved area cladograms (Fig. 2b,c, respectively). The resolved area cladogram that is obtained under A0 is not obtained under A2, leading Ebach and Humphries to conclude that requirement I (i.e. inclusive solution sets; Van Veller *et al.*, 1999) is violated by A0. This is not, however, a proper interpretation of their results. Ebach and Humphries neglected to present the resolved area cladograms that are obtained under A1. Under A1, with CA, one obtains the two resolved area cladograms that already have been obtained under A0 and A2. In addition, a third resolved area cladogram is obtained under A1 (Fig. 2d). In notation the solution sets have the following relationship: $S0 \subset S1 \not\subset S2$. As both S0 and S1 should be included in S2 (i.e. $S0 \subseteq S1 \subseteq S2$), non-inclusion is not caused by A0 but rather by A2. Van Veller *et al.* (2001) suggested a two-step procedure for dealing with widespread and sympatric taxa under A2. In the first step, the widespread taxa are dealt with under A2 by allowing the areas in which the widespread taxon occurs each in turn to float while the other area(s) in which the widespread taxon occurs stay in a fixed position. As a result of this first step, taxon–area cladograms are obtained with only a single area per terminal node. In the second step, paralogy can be reduced from these taxon–area cladograms thereby obtaining resolved area cladograms. When applying this two-step procedure to the taxon–area cladogram, the same solution set (consisting of three resolved area cladograms) obtained under A1 is obtained under A2 as well. The solution sets now have the following relationship: $S0 \subset S1 = S2$, and requirement I is met. This corroborates earlier findings (Van Veller *et al.*, 2000; Van Veller & Brooks, 2001) that when there is disagreement between *a priori* and *a posteriori* methods with respect to a general pattern, the result provided by the *a posteriori* methods is to be preferred, because the *a priori* methods are internally inconsistent on the issue of inclusion (in addition to providing less parsimonious results).

It is important to note that paralogy is not the only explanation for widespread and sympatric taxa. Invoking and then eliminating paralogy to obtain reconciliation of widespread and sympatric taxa requires particular combinations of lineage duplication and lineage sorting, because *a priori* methods do not allow post-speciation dispersal and speciation by dispersal (peripheral isolates speciation) as possible explanations. *A posteriori* methods, by contrast, include these possibilities and invoke them when such explanations provide a more parsimonious representation of all the data than would a combination of lineage duplication and extinction. As invoking one episode of dispersal is more parsimonious than invoking one episode of lineage duplication and, minimally, one episode of lineage sorting, the results of *a posteriori* analyses in such cases will always be more parsimonious than those provided by *a priori* methods. Thus, if the goal is maximum fit between actual data and hypothesis, one should use *a posteriori* methods. If, on the other hand, the goal is to maximize possible fit of data to a given hypothesis that excludes the possibility of areas having reticulated histories with respect to the taxa inhabiting them, then *a priori* methods should be preferred.

We have shown above (and in previous papers) that *a priori* approaches produce internally inconsistent results (do not produce inclusive solution sets) under a certain configuration of data. We have also shown above (and in previous papers) that when *a priori* and a *posteriori* methods give different results, those of the *a posteriori* methods are more parsimonious. Not surprisingly, advocates of *a priori* methods reject both parsimony (Platnick *et al.*, 1996) and inclusive solution sets (Ebach & Humphries, 2002) as criteria for assessing methods and results of historical biogeography. Rather, they

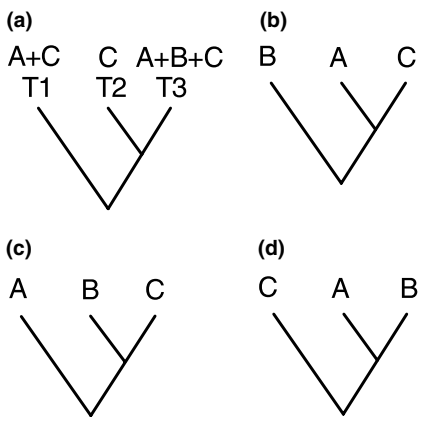

Figure 2 Taxon–area cladogram with widespread and sympatric taxa and area cladograms obtained under assumptions 0, 1 and 2. [(a) Taxon–area cladograms with widespread taxa and paralogy (caused by sympatric taxa), (b) area cladogram obtained under A0 and A1, (c) area cladogram obtained under A1 and A2, and (d) additional area cladogram obtained under A1] (T1–T3 = taxa; A–C = areas).

prefer to invoke episodes of lineage duplication and lineage sorting of a form and to the extent necessary to increase the apparent support for a single, simple pattern of area relationships. To us, such strongly held epistemological views indicate strongly held ontological views. Van Veller & Brooks (2001) suggested that the dispute between advocates of *a priori* and *a posteriori* methods has an underlying ontological basis. Others have also made this suggestion, either explicitly (Ronquist, 1997a,b) or implicitly (Hovenkamp, 1997), but it appears that the perspective needs clarification. Next we consider four categories of distinctions between *a priori* and *a posteriori* methods that we believe comprise strongly divergent ontological positions.

ONTOLOGICAL CONSIDERATIONS

Cladistic biogeography vs. Phylogenetic biogeography

Ebach (2001) and Ebach & Humphries (2002) describe what they do as *cladistic biogeography*, stating that the aim of this field is the search for biogeographical congruence among taxon–area cladograms for several monophyletic groups. Based on the congruent patterns in the taxon–area cladogram a general area cladogram, (or common pattern *sensu* Van Veller *et al.*, 1999) is derived. A general area cladogram '… provides insights into the geographical history of several unrelated groups' [Ebach & Humphries (2002; p. 428)]. We think this is true with respect to the geographical history that is common for different groups but not for the geographical history of speciation in general. Besides vicariant allopatric speciation, processes including sympatric speciation, speciation by dispersal (peripheral isolates speciation), post-speciation dispersal, or isolation without speciation (nonresponse to a vicariance event) are also responsible for the geographical distribution of taxa. However, except for vicariant speciation, these processes do not necessarily produce patterns congruent with the geographical history of several groups of taxa. Cladistic biogeographers discard unique patterns from cladistic geographical analysis as ambiguity.

Ebach & Humphries (2002) claim that *a posteriori* methods are methods of cladistic biogeography. This is not necessarily true, however. Following Hennig (1966), *phylogenetic biogeography* may be defined as the search for explanations of the geographical context of species formation (Wiley, 1981; Wiley & Mayden, 1985; Van Veller *et al.*, 1999, 2002; Brooks & McLennan, 2002). This search is made possible in large part by being able to distinguish general distribution patterns (area relationships supported by multiple clades) from unique elements, including those that are incongruent with the general pattern. Representing the unique elements may take the form of optimizing incongruent data onto the general area cladogram (primary BPA and CCA) or of duplicating areas that have a reticulate history with respect to the species inhabiting them (secondary BPA). Cladistic biogeographers (Platnick, 1988; Page, 1989; Ebach, 2001; Ebach & Humphries, 2002) in contrast, advocate maintaining the singularity of area history by invoking coupled lineage duplication events and lineage sorting events as the analogue of area duplications.

Phylogenetic biogeography is based on the assumption that cladograms produced by phylogenetic systematic analysis are hypotheses of speciation events, and can thus be used in studies of the evolution of species and of multispecies associations (communities or biotas, coevolutionary systems). It is assumed that evolution has been so historically contingent and complex that robust explanations require analysis of both the common and unique patterns, including the possibility of areas with reticulated histories relative to the species inhabiting them. This explicitly evolutionary perspective is ontological rather than epistemological, because if this assumption is not warranted, empirical studies using *a posteriori* methods will consistently fail to find such complexity and ambiguity, producing results completely congruent with those found using *a priori* methods.

Evolutionary vs. non-evolutionary historical biogeography

Ebach (2001) and Ebach & Humphries (2002; p. 427) argue that '… taking an evolutionary stance in biogeography leads to flaws in implementation.' With this statement, they make two mistakes. The first is that the evolutionary stance of phylogenetic biogeographers is not used *a priori* in implementing the methods, but *a posteriori* in explaining results, especially those incongruent with the general pattern.

The second is that cladistic biogeographers cannot escape an evolutionary stance. Ebach & Humphries (2002) look for common patterns among area relationships based on the assumption that vicariant allopatric speciation, surely an evolutionary concept, produces geographically congruent data. They assert that A1 and A2 are technical admonitions and have no evolutionary implications. It is

true that one may treat these assumptions in that manner, just as it is possible to treat A0 as purely technical. But it is also true that A1 and A2 have perfectly good evolutionary interpretations: A1 states that all widespread taxa are the result of sympatric speciation (*lineage duplication*) and extinction (*lineage sorting*), whereas A2 states that all apparent exceptions to vicariance are the result of sympatric speciation, extinction and post-speciation dispersal. Thus, it is in cladistic biogeography that evolutionary stances are used *a priori* in implementing the method of analysis, and those are used specifically to preserve the singularity of area history, even when it is more parsimonious to assert that one or more areas have a reticulated history with respect to the taxa inhabiting them.

Hennig's auxiliary principle vs. Hennig's auxiliary principle

The *a posteriori* methods stem directly from the principles of phylogenetic systematics. In phylogenetic systematics analyses are performed on characters, not taxa (Wiley, 1981; Kluge, 1999). Extending this reasoning to phylogenetic biogeography, BPA and CCA are not analyses of areas, but analyses of taxa and their relationships in a geographical context. It is for this reason that in phylogenetic biogeography dealing with widespread and sympatric taxa is performed via a taxon–relationship approach rather than an area-relationship approach. A species cannot have two different histories (hybrid species have two different ancestors, but still only a single history), therefore its occurrence in two different areas cannot have been caused by vicariance but must be the result of dispersal. Each such event produces a reticulated history of the area with respect to the species living in it. These historical reticulations can be depicted as homoplasy in character optimization (CCA, primary BPA) or as area duplications (secondary BPA).

Ebach (2001) and Ebach & Humphries (2002; see also Platnick, 1988; Page, 1989) advocate the area-relationship approach, in which the singularity of area history is maintained by invoking coupled lineage duplication events and lineage sorting events as the analogue of area duplications. As Ebach and Humphries do not consider taxon-relationships under A1 and A2, they claim that '… Hennig's (1966) Auxiliary Principle can never be violated in cladistic biogeography' (Ebach & Humphries 2002; p. 430).

This statement is remarkable. It suggests that their version of Hennig's Auxiliary Principle is 'Force all similarities to be homologous, no matter what the data say.' Or, as Ebach (1999; p. 388) wrote, 'Assumptions 1 and 2 (Nelson & Platnick, 1981) do allow for taxa and areas to be discarded, as homoplasy is assumed and not ruled out as in Hennig's auxiliary principle'. If Ebach and Humphries are serious, then by their own statements cladistic biogeography is unfalsifiable, and therefore inductivist/verificationist in nature. From that perspective, ambiguity must be treated as 'items of error' (Nelson & Platnick, 1981), rather than as 'items of falsification' (Van Veller & Brooks, 2001), and the use of A1 and A2 is justified. Ebach & Humphries (2002; p. 430) also defined Hennig's Auxiliary Principle as 'homoplasy should not be assumed beyond necessity'. If A1 and A2 are invoked maximally, homoplasy is assumed only to the extent necessary to fit ambiguous data to the general area cladogram, meaning that the result is as parsimonious *as possible* given the *a priori* dictum that the Auxiliary Principle cannot be violated. But because *a posteriori* methods obtain more parsimonious results for data requiring A1 and A2 intervention, *a priori* methods must be assuming homoplasy beyond necessity. Therefore, in treating all ambiguous data as congruence, *a priori* methods actually violate Hennig's Auxiliary Principle in an effort to maintain it.

In phylogenetic systematics, Hennig's Auxiliary Principle is defined as 'Never presume convergent or parallel evolution; always presume homology in the absence of contrary evidence' (Brooks & McLennan 2002; p. 36). This principle is not an *a priori* assumption of a model, but rather a technical presumption used to initiate the analysis; if we assumed that similarity was always the result of convergent or parallel evolution, we would never find any evidence of phylogeny. Hennig's Auxiliary Principle is violated each time we discover homoplasy (character incongruence), but it is violated *a posteriori*. Phylogeneticists expect this sort of character incongruence to be a common enough occurrence that one will need to resort to *a posteriori* invocations of logical parsimony in order to choose among competing hypotheses. In phylogenetic biogeography, Hennig's Auxiliary Principle is violated every time we discover dispersal. We discover homoplasy and dispersal *a posteriori* in the context of character congruence, and not of *a priori* character elimination, modification and fitting. *A posteriori* methods use only A0 precisely in order to prevent *a priori* violation of Hennig's Auxiliary Principle. Ebach & Humphries (2002) are thus factually incorrect when they claim that *a posteriori* methods treat ambiguity as congruence. *A posteriori* methods treat ambiguity as either incongruence or congruence as warranted by the data.

Area-history vs. Taxon-history perspectives

Cladistic biogeography can also be considered an *area-relationship* approach whereas phylogenetic biogeography is a *taxon-relationship* approach (Hovenkamp, 1997; see also Van Veller *et al*., 1999, 2000, 2001, 2002; Brooks *et al*., 2001; Van Veller & Brooks, 2001; Green *et al*., 2002; McLennan & Brooks, 2002). This distinction originated in the evolution of interpretations of A1 and A2 by cladistic biogeographers. Humphries (1982; pp. 453, 454) interpreted A1 to pertain to taxon relationships: '… the implications are that under assumption 1 the taxon occupying area CD [i.e. a taxon widespread in both areas C and D] will never be split into separate taxa.' Page (1989; p. 167), however, advocated that a crucial distinction must be drawn between the relationships of taxa and the relationships of areas and considered A1 and A2 to pertain to area relationships.

Cladistic biogeographers use the assumption of singularity of area history as a criterion for determining errors (of commission and omission) in the phylogenetic trees used to produce the area cladograms, but do not justify their *a priori* belief in the singularity of area histories relative to the evolution of biodiversity. For phylogenetic biogeographers, ambiguous data indicate complex and reticulated histories of areas with respect to the species inhabiting them. This is based on the assumption that each species has a single history and that without evidence we cannot postulate that one species is actually two species (lineage duplication), one of which has gone extinct (lineage sorting). Phylogenetic biogeographers maintain that if the input taxon phylogenies are incomplete or incorrect, they must be modified at the level of phylogenetic analysis of heritable traits of the taxa themselves, rather than at the level of biogeographical analysis. This is necessary to avoid the possibility of circular reasoning, i.e. 'Use taxon phylogenies to generate the general patterns, then use the general patterns to change the taxon phylogenies to fit the general pattern.'

CONCLUSIONS

We think that the above distinctions between cladistic biogeography and phylogenetic biogeography substantiate our claim that they have different ontological bases. Brooks *et al*. (2001), Van Veller & Brooks (2001) and Van Veller *et al*. (2002) discussed this view without expressing a preference for one over the other:

> 'In this paper we show how *a posteriori* and *a priori* methodologies differ. Each is capable of deriving general area cladograms consistently. If the two methodologies agree in their results, no dilemma arises. In cases where their results differ, additional groups of taxa should give a decisive answer on the processes that have resulted in the distribution of taxa of several monophyletic groups over the same areas.' (Van Veller *et al*., 2002).

Ebach & Humphries (2002; p. 432), therefore, are factually incorrect when they state that 'Cladistic biogeographers reading the papers of van Veller and colleagues would get the impression that the only worthwhile method in historical biogeography is BPA…', unless they subconsciously agree that *a posteriori* methods are superior to *a priori* methods for cladistic biogeography. The actual perspective by Van Veller & Brooks (2001) was that because ontological views are not subject to direct empirical testing, cladistic and phylogenetic biogeography may both be considered valid research programmes, each of which has its own particular focus of attention. Cases in which the alternative methods produce the same results should be cause for rejoicing, as empirical findings derived from the same evidence using different methods stemming from radically different views of the world must be robust indeed (see also Van Veller *et al*., 2002). Ebach & Humphries (2002) reject this rapprochement, preferring to misrepresent the published literature, including criticizing the contents of two papers ('van Veller, 2002a', 'van Veller, 2002b') that have not yet been written, appearing only as working titles on Van Veller's website.

A priori and *a posteriori* methods of historical biogeography are not different methods attempting to implement the same research programme, they are methods produced for implementing different research programmes. In addition, it appears from the presentation by Ebach & Humphries (2002) that these approaches also differ with respect to being inductivist / verificationist and deductivist / falsificationist, respectively. If one accepts Popper's distinction between induction and deduction, then Ebach and Humphries are correct in referring to cladistic biogeography as the *art* of discovery, as it is inductivist, in contrast to the *science* of discovery embodied in phylogenetic biogeography, as it is hypothetico-deductivist.

Confusion has resulted because both approaches start with the first-order explanation of vicariance and, for cases in which all taxon-history data most parsimoniously conform to a simple area-history scenario, both approaches give the same answer. The critical issue, however, is 'What do you do when the taxon-history data do not conform to a simple area-history scenario?' For cladistic biogeographers, the answer seems to be 'Modify input data to increase apparent fit of (reconcile) ambiguous data to a single simple area cladogram'. Phylogenetic biogeographers recognize that this situation occurs only when invoking post-speciation range expansion or peripheral isolates speciation would produce a more parsimonious description of the raw input data and inclusive solution sets of results under A0, A1 and A2, and thus choose to invoke post-speciation range expansion or peripheral isolates speciation. For all other configurations of data (vicariance, lineage duplication without lineage sorting and extinction), *a priori* and *a posteriori* methods produce the same results. This is the reason *a posteriori* methods can provide information relevant to cladistic biogeography in addition to forming the basis of phylogenetic biogeography. The following exemplar illustrates this point.

Consider the taxon–area cladogram (represented in Fig. 3a) with four geographical areas in place of five taxon names: (A^1,(B,(C,(D,A^2)))). What is the general area cladogram? *A priori* methods would attempt to induce a general pattern from this singular statement in which each area appears only once; in this example, the preferred general cladogram would be (B,(C,(D,A^2))), which requires three vicariance events and three sorting events (extinctions) to account for all the data, as opposed to (A^1,(B,(C,D))), which requires three vicariance events, three lineage duplication events (sympatric speciation) and eight lineage sorting events (extinction). *A posteriori* methods would assert that, given only this single clade, there are three possibilities (represented in Fig. 3b): (A^1,(B,(C,D))), (B,(C,(D,A^2))) or (A^1,(B,(C,(D,A^2)))). The issue of which of these is to be treated as the general set of area relationships can be deduced only by adding at least two additional clades to the analysis (the Three's Rule rule *sensu* Brooks *et al.*, 2001; Brooks & McLennan, 2002; Green *et al.*, 2002).

If the two (or more) additional clades correspond to (B,(C,(D,A^2))), then the general pattern is three vicariance events corresponding to (1) the isolation of B from C + D + A, (2) the isolation of C from D + A, and (3) the isolation of D and A. The species in A^1 would be explained as the result of peripheral isolates speciation by dispersal (and one would need to expand the scope of the study to determine the area from which it originated). Alternatively, if the two (or more) additional clades correspond to (A^1,(B,(C,D))), then the general pattern is also three vicariance events, corresponding to (1) the isolation of A from B + C + D, (2) the isolation of B from C +D, and (3) the isolation of C and D. In this case, the species in A^2 is the result of peripheral isolates speciation by dispersal from D. Finally, if the two (or more) additional clades correspond to (A^1,(B,(C,(D,A^2)))), then the general

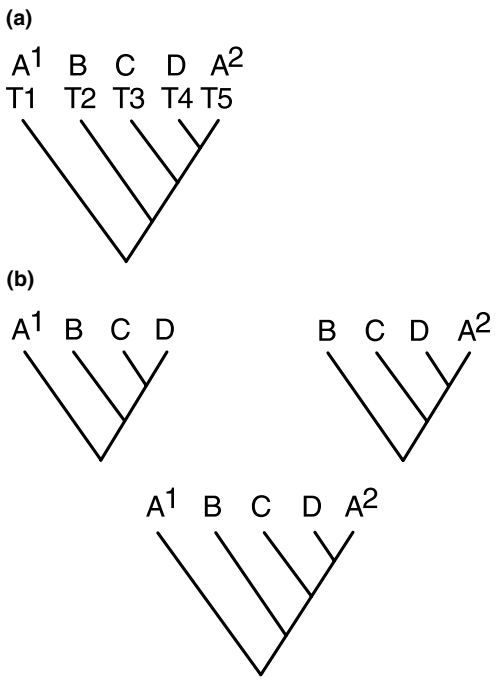

Figure 3 Taxon–area cladogram with two sympatric taxa (in area A) and area cladograms that may be obtained with *a posteriori* methods. [(a) taxon–area cladogram, and (b) area cladograms]. (T1–T5 = taxa; A–D = areas).

pattern comprises four vicariance events corresponding to (1) the vicariant isolation of A^1 from B + C + D, (2) the vicariant isolation of B from C + D, (3) the vicariant isolation of C and D, and (4) the vicariant isolation of D and A^2. For three recent empirical examples using this approach, see Brooks & McLennan (2001), Green *et al.* (2002) and McLennan & Brooks (2002).

We noted above that in previous papers we have emphasized the possibility of coexistence between cladistic biogeography and phylogenetic biogeography. If, however, Ebach and Humphries have characterized cladistic biogeography and *a priori* methods correctly as inductivist/verificationist, we, as deductivists/falsificationists, think a more accurate paraphrase of Ebach & Humphries (2002; p. 432) would be 'Cladistic biogeographers reading the papers of van Veller and colleagues [sh]ould get the impression that the only worthwhile method[s] in historical biogeography [are a *posteriori* methods]…'.

ACKNOWLEDGMENTS

We thank Professor Emeritus Philip Stott for giving us the opportunity to try to clarify these issues. DRB acknowledges funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

MARCO G. P. VAN VELLER[1,3]*
D. R. BROOKS[2] and
M. ZANDEE[3]

[1]*Biosystematics Group, Wageningen University, Wageningen, the Netherlands*
[2]*Department of Zoology, University of Toronto, Toronto, Ontario, Canada*
[3]*Section Theoretical Biology and Phylogenetics, Institute of Evolutionary and Ecological Sciences, Leiden University, Leiden, the Netherlands*
**Correspondence: Biosystematics Group, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, the Netherlands. E-mail: marco.vanveller@wur.nl*

REFERENCES

Brooks, D.R. (1981) Hennig's parasitological method: a proposed solution. *Systematic Zoology*, **30**, 229–249.

Brooks, D.R. (1985) Historical ecology: a new approach to studying the evolution of ecological associations. *Annals of the Missouri Botanical Garden*, **72**, 660–680.

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.

Brooks, D.R. & McLennan, D.A. (2002) *The nature of diversity: an evolutionary voyage of discovery*. University of Chicago Press, Chicago.

Brooks, D.R., Van Veller, M.G.P. & McLennan, D.A. (2001) How to do BPA, really. *Journal of Biogeography*, **28**, 343–358.

Ebach, M.C. (1999) Paralogy and the centre of origin concept. *Cladistics*, **15**, 387–391.

Ebach, M.C. (2001) Extrapolating cladistic biogeography: a brief comment on van Veller *et al.* (1999, 2000, 2001). *Cladistics*, **17**, 383–388.

Ebach, M.C. & Humphries, C.J. (2002) Cladistic biogeography and the art of discovery. *Journal of Biogeography*, **29**, 427–444.

Enghoff, H. (1996) Widespread taxa, sympatry, dispersal, and an algorithm for resolved area cladograms. *Cladistics*, **12**, 349–364.

Green, M.D., Van Veller, M.G.P. & Brooks, D.R. (2002) Assessing modes of speciation: range asymmetry and biogeographical congruence. *Cladistics*, **18**, 112–124.

Hennig, W. (1966) '*Phylogenetic systematics*'. University of Illinois Press, Urbana, IL.

Hovenkamp, P. (1997) Vicariance events, not areas, should be used in biogeographic analysis. *Cladistics*, **13**, 67–79.

Humphries, C.J. (1982) Vicariance biogeography in Mesoamerica. *Annals of the Missouri Botanical Garden*, **69**, 444–463.

Kluge, A.G. (1999) The science of phylogenetic systematics: explanation, prediction, and test. *Cladistics*, **15**, 429–436.

McLennan, D.A. & Brooks, D.R. (2002) Complex histories of speciation and dispersal: an example using some Australian birds. *Journal of Biogeography*, **29**, 1055–1066.

Morrone, J.J. & Carpenter, J.M. (1994) In search of a method for cladistic biogeography: an empirical comparison of Component Analysis, Brooks Parsimony Analysis, and Three-area Statements. *Cladistics*, **10**, 99–153.

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

Nelson, G. (1969) The problem of historical biogeography. *Systematic Zoology*, **18**, 243–246.

Nelson, G. & Ladiges, P.Y. (1991a) Standard assumptions for biogeographic analysis. *Australian Systematic Botany*, **4**, 41–58.

Nelson, G. & Ladiges, P.Y. (1991b) Three-area statements: standard assumptions for biogeographic analysis. *Systematic Zoology*, **40**, 470–485.

Nelson, G. & Ladiges, P.Y. (1991c) TAS: Three Area Statements, program and user's manual. Published by Authors, Melbourne, New York.

Nelson, G. & Ladiges, P.Y. (1995) TAX: MS DOS computer programs of systematics. Published by Authors, Melbourne, New York.

Nelson, G. & Ladiges, P.Y. (1996) Paralogy in cladistic biogeography and analysis of paralogy-free subtrees. *American Museum Novitates*, **3167**, 1–58.

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

Page, R.D.M. (1988) Quantitative cladistic biogeography: constructing and comparing area cladograms. *Systematic Zoology*, **37**, 254–270.

Page, R.D.M. (1989) Comments on component-compatibility in historical biogeography. *Cladistics*, **5**, 167–182.

Page, R.D.M. (1990a) *Component 1.5. Program and user's manual*. University of Auckland, Auckland.

Page, R.D.M. (1990b) Component analysis: a valiant failure? *Cladistics*, **6**, 119–136.

Page, R.D.M. (1993a) Genes, organisms, and areas: the problem of multiple lineages. *Systematic Biology*, **42**, 77–84.

Page, R.D.M. (1993b) *Component 2.0: tree comparison software for Microsoft Windows. Program and user's manual*. Natural History Museum, London.

Platnick, N.I. (1988) Systematics, evolution and biogeography: a Dutch treat. *Cladistics*, **4**, 308–313.

Platnick, N.I., Humphries, C.J., Nelson, G. & Williams, D.M. (1996) Is Farris optimization perfect? Three-taxon statements and multiple branching. *Cladistics*, **12**, 243–252.

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

Ronquist, F. (1997b) Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta*, **26**, 313–322.

Rosen, D.E. (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology*, **27**, 159–188.

Van Veller, M.G.P. & Brooks, D.R. (2001) When simplicity is not parsimonious: *a priori* and *a posteriori* methods in historical biogeography. *Journal of Biogeography*, **28**, 1–11.

Van Veller, M.G.P., Kornet, D.J. & Zandee, M. (2000) Methods in vicariance biogeography: assessment of the implementations of assumptions zero, 1 and 2. *Cladistics*, **16**, 319–345.

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

Van Veller, M.G.P., Kornet, D.J. & Zandee, M. (2002) *A posteriori* and *a priori* methodologies for testing hypotheses of causal processes in vicariance biogeography. *Cladistics*, **18**, 207–217.

Van Veller, M.G.P., Zandee, M. & Kornet, D.J. (2001) Measures for obtaining inclusive solution sets under assumptions zero, 1 and 2 with different methods for vicariance biogeography. *Cladistics*, **17**, 248–259.

Wiley, E.O. (1981) '*Phylogenetics: the theory and practice of phylogenetic systematics*'. Wiley-Interscience, New York.

Wiley, E.O. (1986) Methods in vicariance biogeography. 'Systematics and evolution: a matter of diversity' (ed. P. Hovenkamp), pp. 283–306. Utrecht University Press, Utrecht.

Wiley, E.O. (1988a) Vicariance biogeography. *Annual Review of Ecology and Systematics*, **19**, 513–542.

Wiley, E.O. (1988b) Parsimony analysis and vicariance biogeography. *Systematic Zoology*, **37**, 271–290.

Wiley, E.O. & Mayden, R.L. (1985) Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Annals of the Missouri Botanical Garden*, **72**, 596–635.

Zandee, M. (1999) *CAFCA versus 1.5j: a collection of APL functions for cladistic analysis. Program and user's manual*. Institute of Evolutionary and Ecological Sciences, Leiden University, Leiden.
Zandee, M. & Roos, M.C. (1987) Component-compatibility in historical biogeography. *Cladistics*, **3**, 305–332.

BIOSKETCHES

Marco G.P. van Veller received his PhD from Leiden University, working on the theoretical foundations of methods of historical biogeography. His research concentrates on the analogous application of phylogenetic methodology in (molecular) systematics, historical biogeography and studies of coevolutionary systems.

Daniel R. Brooks is Professor of Zoology, University of Toronto, specializing in the systematics and evolution of parasitic helminths. He is currently coordinating the inventory of eukaryotic parasites of vertebrates, Area de Conservacion Guanacaste, Costa Rica. He is co-author of *Phylogeny, ecology and behavior: a research programme in comparative biology* (1991), *Parascript: Parasites and the language of evolution* (1993) and *The nature of diversity: An evolutionary voyage of discovery* (2002).

Rino Zandee is Associate Professor of Theoretical Biology at Leiden University. His primary research focus is the methodology of phylogenetic systematics, historical biogeography and co-evolutionary systems in general.