

DISPERSAL–VICARIANCE ANALYSIS: A NEW APPROACH TO THE QUANTIFICATION OF HISTORICAL BIOGEOGRAPHY

FREDRIK RONQUIST

Department of Zoology, Uppsala University, Villavägen 9, SE-752 36 Uppsala, Sweden;
E-mail: fredrik.ronquist@zoologi.uu.se

Abstract.—Quantification in historical biogeography has usually been based on the search for a single branching relationship among areas of endemism. Unlike organisms, however, areas rarely have a unique hierarchical history. Dispersal barriers appear and disappear and may have different effects on different species. As a result, the biota of an area may consist of several components with separate histories, each of which may be reticulate rather than branching. In an attempt to address these problems, I present a new biogeographic method, dispersal–vicariance analysis, which reconstructs the ancestral distributions in a given phylogeny without any prior assumptions about the form of area relationships. A three-dimensional step matrix based on a simple biogeographic model is used in the reconstruction. Speciation is assumed to subdivide the ranges of widespread species into vicariant components; the optimal ancestral distributions are those that minimize the number of implied dispersal and extinction events. Exact algorithms that find the optimal reconstruction(s) are described. In addition to their use in taxon biogeography, the inferred distribution histories of individual groups serve as a basis for the study of general patterns in historical biogeography, particularly if the relative age of the nodes in the source cladograms is known. [Cladistic biogeography; comparative phylogeography; dispersal; extinction; historical biogeography; optimization; vicariance; widespread species.]

In historical biogeography, whether focused on patterns below the species level (comparative phylogeography) or above the species level, there is a need for quantitative methods to assess the likelihood of alternative hypotheses. Most methods used today are based on the assumption that there is a single branching pattern among areas of endemism caused by vicariance and that this pattern is common to many different groups of organisms (cladistic biogeography). Despite the simple basic idea, the treatment of dispersal and extinction has caused serious difficulties in the development of an analytical protocol. No method proposed so far has solved these problems satisfactorily. Most methods do not treat dispersal and extinction explicitly, which means that the analyses require a posteriori interpretation and/or a priori data manipulation (Wiley, 1988; Brooks, 1990; Page, 1990, 1994; cf. Ronquist and Nylin, 1990; Page, 1995). Some methods model these processes more realistically, but other difficulties remain (Ronquist & Nylin, 1990; Page, 1995; Ronquist, 1996b).

An even more serious problem with the

current approach, however, is the fundamental assumption of a single branching pattern. Dispersal barriers appear and disappear throughout evolutionary history and may have different effects on different species. As a result, the biota of an area may consist of several components with separate histories, each of which may be reticulate rather than branching. Such complicated area relationships should be considered as possible historical explanations of observed biogeographic patterns.

In the classic vicariance scenario, a widespread ancestor speciates by responding to successive subdivisions of its distribution range. This scenario implies that ancestral species were generally more widespread than their descendants, a paradox discussed by Brundin (1981) and more recently by Bremer (1992). We can escape this paradox by allowing successive dispersal events that counteract the decrease in distribution range caused by vicariance. Some dispersal is needed to explain the occurrence of widespread ancestors. Why should all dispersal events have occurred before the deepest split in the studied group, as required by a pure

vicariance scenario? Clearly, consideration of dispersal and extinction is essential in biogeographic reconstruction.

Theoretically, speciation is expected to be allopatric in most groups of organisms (Mayr, 1963; Futuyma and Mayer, 1980), and empirical studies support this view (Lynch, 1989; Asquith, 1993; Chesser and Zink, 1994). Thus, allopatric speciation, associated with geographical vicariance, may be accepted as the null model in historical biogeography. However, dispersal may also play an important role in speciation. For instance, in the case of successive colonization of an archipelago by a group of organisms, stepping-stone dispersal may be followed by vicariance (e.g., Hennig, 1966). This biogeographic pattern will be difficult to detect unless dispersal is modeled correctly.

Considering the problems with the hierarchical constraints in cladistic biogeography, we need new approaches to the reconstruction of the distribution history of individual groups (taxon biogeography) and in the search for general area relationships (area biogeography). Character optimization methods are interesting in this context because they allow reconstruction of ancestral distributions without constraining area relationships to hierarchical patterns (Bremer, 1992; Maddison et al., 1992; Ronquist, 1994). A central problem with these methods, however, is the treatment of widespread ancestors (Ronquist, 1995). Methods such as Fitch optimization normally restrict polymorphism (widespread species) to terminals, and ancestors are exclusively reconstructed as monomorphic (occurring in single areas). A simple way of including polymorphic ancestors using population-state coding was suggested by Ronquist (1994), but this technique assigns an unrealistically low cost (zero) to sympatric speciation in widespread ancestors and models vicariance as the result of sympatric speciation followed by extinction (Ronquist, 1995).

In this paper, I describe a quantitative method of analyzing problems in historical biogeography, *dispersal-vicariance analysis*, which is derived from character optimi-

zation methods. Like Fitch optimization, the method minimizes dispersal and extinction and allows multiple and reticulate relationships among areas. However, it is based on vicariance rather than on sympatric speciation. Widespread species cause no problems in dispersal-vicariance analysis; the treatment of widespread ancestors is the quintessence of the method.

SETTING THE SCENE

Assume that the distribution of a set of species (or terminal taxa) can be described in terms of a number of small unit areas (areas of endemism): A, B, C, etc. Each species occurs in one or several of these unit areas. Let D_i be the distribution of species i , that is, the set of areas where species i occurs. For instance, if species i occurs in areas A, B, and D, then $D_i = \{A, B, D\}$. It is assumed that the areas have been sampled sufficiently so that absence from an area is not likely to be caused by incomplete knowledge of that area.

The phylogenetic relationships among species, including the location of the root, is assumed to be known and to be fully bifurcate. It is possible to modify the method to take soft and hard polytomies into account, but this is not attempted here. The method reconstructs the distribution sets of the ancestral nodes in the cladogram according to a set of rules based on a simple biogeographic model.

THE RULES

Speciation is assumed to be caused by vicariance, i.e., by division of the ancestral species' distribution into two component parts. There are two possibilities. First, the ancestor may occur in a single unit area immediately before the speciation. In this case, allopatric speciation within that area is assumed, possibly coupled with small-scale dispersal. Immediately after the speciation event, both descendant lineages occur in the same area as their ancestor (Fig. 1a). The cost of this event is zero.

The second possibility is that the ancestor occurs in more than one unit area. In this case allopatric speciation leads to subdivision of the ancestral distribution into

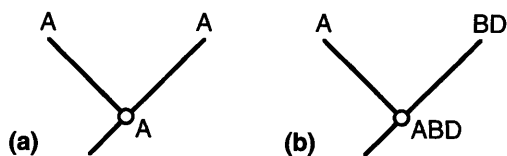


FIGURE 1. Expected biogeographic outcomes of speciation. (a) When the ancestral species occurs in a single area, the daughter species will be expected to occur in the same area. (b) When the ancestral species occurs in several areas, the daughter species will be expected to occur in mutually exclusive sets of areas.

two mutually exclusive sets of areas. For example, assume that the ancestor *a* has the distribution $D_a = \{A, B, D\}$. Immediately after the speciation, the daughter species *b* and *c* could have the distributions $D_b = \{A\}$ and $D_c = \{B, D\}$ (Fig. 1b), or $D_b = \{A, D\}$ and $D_c = \{B\}$, but not $D_b = \{A, B\}$ and $D_c = \{B, D\}$. A general definition of vicariance is that immediately after speciation $D_b \cup D_c = D_a$ and $D_b \cap D_c = \emptyset$. The cost of this event is zero.

To make the method complete, we only have to add dispersal and extinction. Dispersal is the addition of one or more unit areas to a distribution; the cost is 1 per area added. Extinction is the deletion of one or more unit areas from a distribution; the cost is 1 per area deleted. All species must occur in at least one unit area.

FINDING OPTIMAL RECONSTRUCTIONS

Finding the optimal reconstruction of ancestral distributions involves a search for the reconstruction with the minimum dispersal-extinction cost. Thus, optimality is determined by a parsimony criterion.

The reconstruction procedure is similar to optimization of a step-matrix character. However, an ordinary step matrix is two dimensional and specifies the cost for every possible transition from one state to another along an internode. In dispersal-vicariance analysis we need a three-dimensional matrix because the cost of an event depends on the particular combination of descendant distributions (Fig. 2).

The three-dimensional step matrix is constructed as follows. Assume that the distribution of the ancestral node is D_a , that of the left descendant node is D_l , and

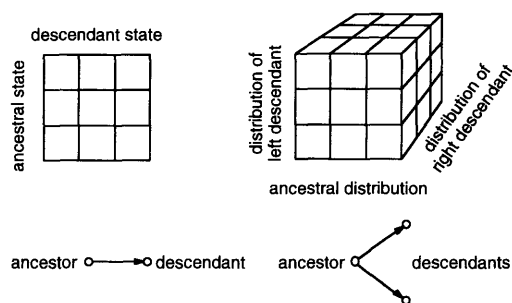


FIGURE 2. The difference between an ordinary step matrix and the cost matrix needed for the reconstruction of ancestral areas. An ordinary step matrix is two dimensional and specifies the cost of moving between states along an internode. The cost matrix used in dispersal-vicariance analysis is three dimensional and specifies the cost of combinations of ancestral, left descendant, and right descendant distributions.

that of the right descendant node is D_r . Let C be the dispersal-extinction cost associated with moving from D_a to D_l and D_r . For calculation purposes, define $D_u = D_l \cup D_r$ and $D_s = D_l \cap D_r$. Let $|D_i|$ be the number of elements in D_i . One can show that if $|D_a| > 1$, the cost is

$$C = |D_a| + |D_l| + |D_r| + |D_a \cap D_s| \\ - |D_a \cap D_l| - |D_a \cap D_r| \\ - |D_a \cap D_s|.$$

If $|D_a| = 1$, then the cost is

$$C = |D_a \cup D_l| - |D_a \cap D_l| + |D_a \cup D_r| \\ - |D_a \cap D_r|.$$

Once the cost matrix has been defined, the reconstruction of the ancestral distributions proceeds in a way similar to that of an ordinary step-matrix optimization (see Maddison and Maddison, 1992; Ronquist, 1996b). The first step is to assign the observed distribution of the terminal taxa to the terminal nodes in the cladogram. Then all internal nodes in the cladogram are assigned an array of all possible distributions, i.e., all possible combinations of unit areas. For instance, if the terminal taxa occur in the unit areas *A* and *B*, all internal nodes in the cladogram are assigned an array consisting of the possible distributions $\{A\}$, $\{B\}$, and $\{A, B\}$.

The possible distributions at each ances-

tral node are then assigned costs in a downpass, an uppass, and a final pass (Maddison and Maddison, 1992; Ronquist, 1996b). After the final pass, the optimal distribution at each ancestral node is the distribution associated with the minimum overall dispersal-extinction cost.

The cost array of a node specifies the dispersal-extinction cost of each possible distribution at that node. It may contain an infinite value for certain distributions, indicating that these distributions are not possible at the node. For instance, in the array of a terminal node, the cost of distributions not observed is set to infinity.

In the downpass of the optimization, the process goes from the terminals towards the root. In each step, the cost arrays of two descendant nodes are combined to yield the downpass cost array of their ancestral node. The downpass is completed when the root is reached. The total dispersal-extinction cost of the optimal reconstruction can be seen directly in the downpass array of the root node, which is equivalent to the final cost array of that node.

To calculate the final arrays for the other nodes, an uppass and a final pass are also needed. In the uppass, the process goes from the root towards the terminals. First, the uppass cost of all optimal distributions of the root node is set to 0, and the uppass array of the root node is combined with the downpass array of one of the descendant nodes to yield the uppass array of the other descendant node. Then, the uppass array of an ancestral node is successively combined with the downpass array of one of the descendant nodes to yield the uppass array of the other descendant node. The final array of a node is obtained by combining its uppass array with its downpass array.

There are two practical difficulties in implementing the procedure outlined above in a computer program. First, the size of the step matrix grows very rapidly with the number of different areas considered. For n unit areas there are $2^n - 1$ different distributions (unit area combinations), giving a cost matrix with $(2^n - 1)^2$ cells. Half

the cost matrix is redundant because the right and left descendants are equivalent, but the number of nonredundant cells still increases very rapidly with the number of areas. For instance, 10 different unit areas require a matrix with 535×10^9 nonredundant cost values. Storage of that many numbers will require more memory than most computers have.

Second, because of the large number of possible combinations of ancestral distributions that must be considered, the computational effort needed to solve even fairly small problems is enormous. Keeping the number of taxa constant, the required number of comparisons in the optimization increases exponentially with the number of different areas because the number of comparisons is directly proportional to the size of the cost matrix.

Fortunately, it is possible to address both of these problems in dispersal-vicariance optimization by using the following two rules, which hold generally for all ancestral nodes: (1) the optimal distribution of an ancestral node cannot contain a unit area not occupied by any descendant; and (2) the optimal distribution of an ancestral node must contain at least one unit area from the distributions of each of the two daughter nodes.

If we only consider ancestral distributions that could be optimal, the cost equations simplify considerably. When $|D_a| > 1$, the cost is

$$C = |D_l| + |D_r| - |D_a|.$$

When $|D_a| = 1$, the cost is

$$C = |D_l| + |D_r| - |D_a| - 1.$$

These equations are so simple that the costs can be calculated as they are needed. Thus, the cost matrix values need not be stored in computer memory.

By only considering ancestral distributions that could be optimal, it is also possible to reduce significantly the number of necessary comparisons in the tree traversals. Before the downpass, the set of possible distributions at each node is restricted according to the two rules described above. Once the root is reached in the

downpass, the cost of the optimal reconstruction is known. All ancestral node distributions with a downpass cost larger than the optimum can be omitted from further calculations.

The uppass and final pass are combined to speed up calculations further. As soon as an uppass array of a visited node is completed, its values are added to the corresponding values in the downpass array to yield the final array of the node. Only the optimal distributions in the final array need be considered during the next step in the uppass-final pass algorithm.

With the algorithms outlined here it is possible to find exact solutions to fairly large problems, but the computational complexity is strongly dependent on the data structure. When there are many widespread terminals, the algorithms will get bogged down by the large number of possible ancestral distributions that must be considered. If the cost of an optimal or near-optimal reconstruction is known before the calculations start, this value can be used as an upper bound to limit the number of alternative ancestral distributions in the basal nodes of the cladogram (Ronquist, 1996b). It is also possible to use a heuristic search that limits the number of different ancestral distributions kept at each step in the downpass.

The algorithms described above are implemented in the computer program DIVA (Ronquist, 1996a), which is available in MacOS (68K and PPC) and Win32 (Windows95 and WindowsNT) versions via FTP links from the author's home page (www.systbot.uu.se/personel/f_ronquist.html).

SOME SIMPLE EXAMPLES

Simple examples are useful because they demonstrate basic properties of methods. First, consider an area cladogram with two sister groups occurring in different areas (Fig. 1b). The optimal reconstruction has a cost of zero and implies that the pattern resulted from vicariant subdivision of a larger ancestral distribution. Whenever the observed pattern is fully consistent with

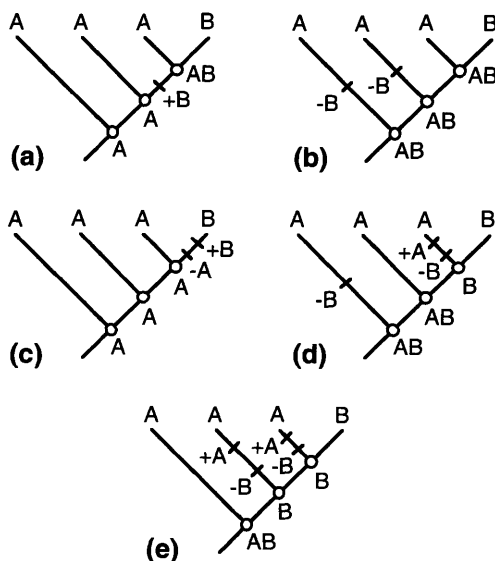


FIGURE 3. A hypothetical area cladogram. (a) Dispersal-vicariance analysis postulates dispersal into area B followed by vicariance. (b) Reconciliation with the only possible area cladogram suggests that two basal lineages went extinct in area B. This explanation is less parsimonious because, in addition to the extinction events, it also requires some dispersal to explain how the two earliest widespread ancestors both gave rise to two widespread descendants (duplications in the terminology of Page, 1995). (c) Brooks parsimony analysis indicates that the dispersal to area B occurred on the terminal branch. In a dispersal-vicariance framework, this reconstruction is suboptimal because it requires an extinction event in addition to the dispersal event. (d, e). Two of the three optimal solutions using the maximum cospeciation method proposed by Page (1995). The third solution is that given in cladogram b. All require more events than the dispersal-vicariance reconstruction because of the hierarchical constraint on area relationships.

such a classic vicariance scenario, this will be the optimal solution.

Second, consider a pattern suggesting dispersal (Fig. 3a). The optimal solution postulates dispersal into area B followed by allopatric speciation separating the con-specific populations in areas A and B. The dispersal must have occurred before speciation. Otherwise, one would have to assume that speciation occurred within area A and that one of the daughter species dispersed into area B and went extinct in area A (cf. Fig. 3c). Because this explanation requires a dispersal and an extinction event,

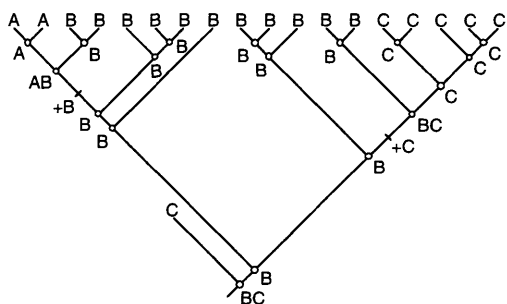


FIGURE 6. Optimal reconstruction of the biogeographic history of the chironomid subfamily Diame-sinae (cf. Fig. 5 and Table 1). A = Australia; B = South America; C = New Zealand.

bution must be the combination of those areas (from rule 2). This gives the optimal distributions for node 3 and node 15. For nodes 6, 7, and 16–18, we must consider several alternative distributions. For instance, both {B} and {A, B} need to be considered for node 6; but {A} cannot be the optimal distribution at this node because then the ancestor would not share a unit area with its right descendant (node 5), which only occurs in area B (rule 2). After the elimination procedure, relatively few possible ancestral distributions remain (Table 1).

The next step is the downpass, which starts with the ambiguous nodes nearest the terminals. Node 6 has descendants with the distributions $D_i = \{A, B\}$ and $D_r = \{B\}$. We need to calculate the costs of node 6 having the distributions {A, B} and {B}. For $D_a = \{A, B\}$, the cost added at this node would be $|D_i| + |D_r| - |D_a| = 2 + 1 - 2 = 1$. To obtain the downpass cost of the distribution {A, B} for node 6, $C_{d,6}(\{A, B\})$, we add the downpass costs from the descendants, in this case 0 for both descendants. Thus, $C_{d,6}(\{A, B\}) = 1$ (Table 1). Similarly, we obtain $C_{d,6}(\{B\}) = C_{d,3}(\{A, B\}) + C_{d,5}(\{B\}) + |\{A, B\}| + |\{B\}| - |\{B\}| - 1 = 0 + 0 + 2 + 1 - 1 - 1 = 1$ (Table 1). For node 7, we must consider the distributions {A, B} and {B}, and we must take two alternative distributions of the left descendant (node 6) into account. To find the downpass cost for the distribution {B}, for instance, we try all possible combinations

TABLE 1. Downpass, uppass, and final pass cost arrays for the ancestral nodes in the area cladogram of Figure 5. Ancestral nodes for which the reconstruction of the optimal distribution is trivial are not included. — = distribution excluded as possible optimal states prior to the downpass; * = values not needed because the distributions are excluded as possible optimal states after the downpass.

Node no.	Pass	Cost of ancestral distribution						
		A	B	C	AB	BC	AC	ABC
6	down	—	1	—	1	—	—	—
	up	—	1	—	2	—	—	—
	final	—	2 ^a	—	3	—	—	—
7	down	—	1	—	2	—	—	—
	up	—	1	—	5	—	—	—
	final	—	2 ^a	—	7	—	—	—
16	down	—	1	—	—	1	—	—
	up	—	1	—	—	2	—	—
	final	—	2 ^a	—	—	3	—	—
17	down	—	2	—	4	3	5	4
	up	—	0	—	*	*	*	*
	final	—	2 ^a	—	*	*	*	*
18	down	—	—	4	—	2	5	4
	up	—	—	*	—	0	*	*
	final	—	—	4	—	2 ^a	5	4

^a Optimal reconstruction (minimum overall dispersal-extinction cost for a given distribution).

of descendant distributions to find the one that gives the minimum downpass cost. In this case, the cost will be 1 if D_i is {B} but 2 if D_i is {A, B}. Thus, the downpass cost of {B} is 1 (Table 1).

When the root (node 18) is reached in the downpass, we know that the length of the optimal reconstruction is 2 and that the optimal distribution of the root node is {B, C} (Table 1). The uppass cost of this distribution is then set to 0, and the other distribution states at the root node are excluded from further calculations (Table 1). The uppass array of the next node, node 17, is obtained by combining the uppass cost of the optimal state at the root node with the downpass array of the left descendant node, in this case a terminal node with the observed distribution C and a downpass cost of 0 (Fig. 5). For node 17, we have to consider only the distribution {B} because the other possibilities have a downpass cost that exceeds the cost of the optimal reconstruction (Table 1). The uppass array of node 17 is added to the downpass array producing the final array of the node, in

this case consisting of only one possible state. When we have visited all nodes that could have different ancestral distributions as optimal, the reconstruction is complete. The optimal states of an ancestral node are those distributions having the minimum overall dispersal-extinction cost in the final cost array of the node (Table 1).

In the optimal reconstruction (Fig. 6), basally there is a vicariance event separating South America and New Zealand and later there is dispersal into Australia and New Zealand, respectively, followed by vicariance. This reconstruction agrees well with Brundin's interpretation of the biogeographic history of the group (Brundin, 1981), but it contrasts with the branching relationship expected in cladistic biogeography. The reconstruction mixes classic vicariance (the basal separation of elements occurring in New Zealand and South America) with dispersal followed by vicariance (from South America to New Zealand and Australia, respectively). Brundin (1966, 1981) relied heavily on the much criticized progression rule of Hennig (1966) in reconstructing biogeographic patterns, but it should be clear from this example that it is possible to justify his conclusions within a parsimony framework, without reference to the progression rule.

AREA BIOGEOGRAPHY

As described here, dispersal-vicariance analysis is primarily a method of reconstructing the historical biogeography of individual groups. However, it may also be a powerful tool in the search for general area relationships in at least two different ways. First, different hypotheses about area relationships could be compared by fitting area cladograms to them, the cost matrix in each case being adjusted according to the area hypothesis being tested. For instance, dispersal between areas postulated to have been connected at the time of dispersal should be less costly than dispersal between isolated areas. Three-dimensional cost matrices are extremely flexible in this context. If one is willing to assume hierarchical area relationships, for

example, three-dimensional matrix optimization could be used to find maximum vicariance reconstructions (the optimality criterion proposed by Page, 1995). Although logically satisfactory, however, this approach is computationally complex.

The second approach is to use the reconstructions of ancestral distributions obtained with simple dispersal-vicariance analysis as a heuristic tool in exploring general area relationships, rather like one might use Fitch optimization of character changes to investigate transition/transversion bias in the evolution of DNA sequences. Ideally, the nodes in the source cladograms should be associated with some relative time estimate, such as is common in comparative phylogeography (e.g., Zink, 1996). General vicariance events could then be identified as peaks in the frequency of different vicariance events plotted against time. Between general vicariance events, it should be possible to infer area connections and dispersal barriers by recording the frequency of different ancestral distributions and dispersal events in that time segment.

Even in the absence of dated nodes, dispersal-vicariance reconstructions can be used to examine general area relationships. Enghoff (1996) used hierarchical component analysis, among other methods, to examine the historical biogeography of the four major regions of the Holarctic based on area cladograms of nonmarine animals. He found considerable support for a hierarchical arrangement grouping the current continents, particularly in one of the examined data sets (Fig. 7a). This scenario postulates sister-group relationships between Nearctic and Holarctic groups. However, dispersal-vicariance optimizations indicate that the support for the hierarchical pattern is almost exclusively due to the presence of many terminals and near-terminal ancestors with continental (two area) distributions (Ronquist and Enghoff, in prep.). Holarctic distributions are rare in dispersal-vicariance optimizations and are often divided differently by vicariance than postulated by the hierarchical scenario. Dispersal-vi-

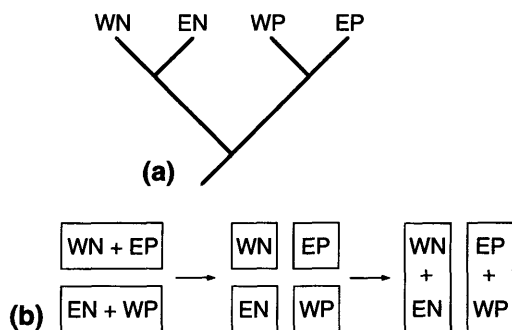


FIGURE 7. Alternative hypotheses of area relationships in the Holarctic (Engelhoff, 1996). Dispersal-vicariance reconstructions can be used to investigate whether the data fit the hierarchical or the reticulate scenario best. WN = West Nearctic; EN = East Nearctic; WP = West Palearctic; EP = East Palearctic. (a) Branching relationship suggested by hierarchical component analysis of a large set of taxon-area cladograms. (b) Reticulate relationship indicated by palaeogeographic data.

carian analysis suggests that continental distributions frequently are the result of dispersal from one of the component areas, which is more in agreement with the reticulate area relationships indicated by palaeogeographic data (Fig. 7b) than with the hierarchical arrangement postulated by component analysis.

CONCLUSIONS

Dispersal-vicariance analysis provides a powerful tool both in the reconstruction of the distribution history of individual groups and in the search for general area relationships. It will be particularly useful when general area relationships are not expected to conform to a hierarchical pattern.

ACKNOWLEDGMENTS

I thank Kåre Bremer, David Cannatella, D. R. Lindberg, Göran Nordlander, Lars Werdelin, Hans-Erik Wärntorp, Niklas Wikström, and an anonymous reviewer for comments, discussion, and valuable criticisms that helped improve early versions of the manuscript.

REFERENCES

ASQUITH, A. 1993. Patterns of speciation in the genus *Lopidea* (Heteroptera: Miridae: Orthotylinae). *Syst. Entomol.* 18:169–180.

BREMER, K. 1992. Ancestral areas: A cladistic reinterpretation of the center of origin concept. *Syst. Biol.* 41:436–445.

BROOKS, D. R. 1990. Parsimony analysis in historical biogeography and coevolution: Methodological and theoretical update. *Syst. Zool.* 39:14–30.

BRUNDIN, L. Z. 1966. Transantarctic relationships and their significance, as evidenced by chironomid midges. *K. Sven. Vetenskapsakad. Handl.* 11:1–472.

BRUNDIN, L. Z. 1981. Croizat's panbiogeography versus phylogenetic biogeography: A critique (G. Nelson and D. E. Rosen, eds.). Columbia Univ. Press, New York.

CHESSER, R. T., AND R. M. ZINK. 1994. Modes of speciation in birds: A test of Lynch's method. *Evolution* 48:490–497.

ENGELHOFF, H. 1996. Historical biogeography of the Holarctic: Area relationships, ancestral areas, and dispersal of non-marine animals. *Cladistics* 11:223–263.

FUTUYMA, D. J., AND G. C. MAYER. 1980. Non-allopatric speciation in animals. *Syst. Zool.* 29:254–271.

HENNIG, W. H. 1966. *Phylogenetic systematics*. Univ. Illinois Press, Urbana.

LYNCH, J. D. 1989. The gauge of speciation. Pages 527–553 in *Speciation and its consequences* (D. Otte and J. A. Endler, eds.). Sinauer, Sunderland, Massachusetts.

MADDISON, D. R., M. RUVOLO, AND D. L. SWOFFORD. 1992. Geographic origins of human mitochondrial DNA: Phylogenetic evidence from control region sequences. *Syst. Biol.* 41:111–124.

MADDISON, W. P., AND D. R. MADDISON. 1992. *MacClade: Analysis of phylogeny and character evolution, version 3.0*. Sinauer, Sunderland, Massachusetts.

MAYR, E. 1963. *Animal species and evolution*. Belknap Press, Cambridge, Massachusetts.

PAGE, R. D. M. 1990. Component analysis: A valiant failure? *Cladistics* 6:119–136.

PAGE, R. D. M. 1994. Maps between trees and cladistic analysis of relationships among genes, organisms, and areas. *Syst. Biol.* 43:58–77.

PAGE, R. D. M. 1995. Parallel phylogenies: Reconstructing the history of host-parasite assemblages. *Cladistics* 10:155–173.

RONQUIST, F. 1994. Ancestral areas and parsimony. *Syst. Biol.* 43:267–274.

RONQUIST, F. 1995. Ancestral areas revisited. *Syst. Biol.* 44:572–575.

RONQUIST, F. 1996a. DIVA, version 1.0. Computer program for MacOS and Win32. Available from www.systbot.uu.se/personel/f.ronquist.html.

RONQUIST, F. 1996b. Reconstructing the history of host-parasite associations using generalised parsimony. *Cladistics* 11:73–89.

RONQUIST, F., AND S. NYLIN. 1990. Process and pattern in the evolution of species associations. *Syst. Zool.* 39:323–344.

WILEY, E. O. 1988. Parsimony analysis and vicariance biogeography. *Syst. Zool.* 37:271–290.

ZINK, R. M. 1996. Comparative phylogeography in North American birds. *Evolution* 50:308–317.

Received 4 December 1994; accepted 13 September 1996
Associate Editor: David Cannatella