

A LIKELIHOOD FRAMEWORK FOR INFERRING THE EVOLUTION OF GEOGRAPHIC RANGE ON PHYLOGENETIC TREES

RICHARD H. REE,¹ BRIAN R. MOORE,² CAMPBELL O. WEBB,² AND MICHAEL J. DONOGHUE²

¹*Department of Botany, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605*

E-mail: rree@fieldmuseum.org

²*Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520*

Abstract.—At a time when historical biogeography appears to be again expanding its scope after a period of focusing primarily on discerning area relationships using cladograms, new inference methods are needed to bring more kinds of data to bear on questions about the geographic history of lineages. Here we describe a likelihood framework for inferring the evolution of geographic range on phylogenies that models lineage dispersal and local extinction in a set of discrete areas as stochastic events in continuous time. Unlike existing methods for estimating ancestral areas, such as dispersal-vicariance analysis, this approach incorporates information on the timing of both lineage divergences and the availability of connections between areas (dispersal routes). Monte Carlo methods are used to estimate branch-specific transition probabilities for geographic ranges, enabling the likelihood of the data (observed species distributions) to be evaluated for a given phylogeny and parameterized paleogeographic model. We demonstrate how the method can be used to address two biogeographic questions: What were the ancestral geographic ranges on a phylogenetic tree? How were those ancestral ranges affected by speciation and inherited by the daughter lineages at cladogenesis events? For illustration we use hypothetical examples and an analysis of a Northern Hemisphere plant clade (*Cercis*), comparing and contrasting inferences to those obtained from dispersal-vicariance analysis. Although the particular model we implement is somewhat simplistic, the framework itself is flexible and could readily be modified to incorporate additional sources of information and also be extended to address other aspects of historical biogeography.

Key words.—Ancestral-area estimation, dispersal, extinction, historical biogeography, speciation mode, vicariance.

Received March 29, 2005. Accepted September 13, 2005.

The geographic history of species and clades—their origin, persistence, dispersal, and extinction in space and time—is an important aspect of understanding the current distribution of biodiversity on Earth (e.g., Croizat 1958, 1964; MacArthur and Wilson 1967; Ricklefs and Schluter 1993; Brown and Lomolino 1998; Ricklefs 2004). However, this subject has not always figured prominently in the field of historical biogeography. Influenced by the principles and methodological rigor of phylogenetic systematics, historical biogeography was for many years focused on a different problem: the discovery of dichotomous area relationships (area cladograms) from organismal cladograms. Under this “cladistic” approach to biogeography, a plethora of quantitative inference methods were developed for this purpose (e.g., Platnick and Nelson 1978; Brooks 1981, 1985, 1990; Nelson and Platnick 1981; Page 1990, 1993, 1994; Humphries and Parenti 1999), and the area-cladogram paradigm has since remained dominant in the field. With respect to the geographic history of lineages, for example, the area cladogram is the conceptual basis for applying “tree-within-a-tree” inference methods to biogeographic studies (e.g., see Page and Charleston 1998; Ronquist 1998). These methods fit “contained” trees (organismal cladograms) to “container” trees (area cladograms) according to explicit optimality rules (e.g., Ronquist 2002).

In a notable departure from the area-cladogram paradigm, a pioneering series of papers shifted the focus of historical biogeography back toward inferring the geographic history of lineages (hereafter, “lineage geohistory”), emphasizing the reconstruction of their ancestral ranges (e.g., Bremer 1992, 1995; Ronquist 1994, 1995, 1996, 1997; Hausdorf 1998). This line of inquiry led to the development of dispersal-vicariance analysis (DIVA; Ronquist 1996, 1997), a

method for optimizing ancestral areas on a cladogram by minimizing the number of historical events (area vicariance, dispersal, and extinction) required to explain the geographic distribution of terminal taxa. DIVA has since come to be widely applied in studies of historical biogeography (e.g., Sanmartin and Ronquist 2004).

While novel in its approach to inferring lineage geohistory, DIVA nevertheless maintains strong conceptual ties to cladistic biogeography and the area cladogram paradigm. Most importantly, hypothetical area vicariance events maintain primary importance in explaining disjunct species distributions. Thus, the histories of areas and lineages are not conceptually separated within the DIVA framework. This general approach may be appropriate if nothing is known about area relationships other than the cladograms of the species inhabiting them. However, if independent information about them is available (particularly about the timing of geographic connections and other dispersal opportunities between areas), then methods for inferring lineage geohistory would profit from directly incorporating this information (Ronquist 1996, 1997), particularly if coupled with lineage durations and divergence times, estimates of which are increasingly available from calibrated rates of molecular evolution. At present, no biogeographic inference method incorporates such information.

Development of more integrative approaches to historical biogeography, using a broad spectrum of data types to address a greater diversity of questions, has recently been advocated (Donoghue and Moore 2003; Wiens and Donoghue 2004). In this paper, we depart from the area-cladogram paradigm and explore a new direction for historical biogeography: namely, parametric inference of lineage geohistory using explicit

models for geographic range evolution, both along phylogenetic internodes (anagenetic change), and at lineage branching events (cladogenetic change). For anagenetic range evolution, we introduce a probabilistic model for lineage dispersal and extinction in a set of areas through time that incorporates estimates of lineage divergence dates and paleogeographic information about the availability and timing of dispersal opportunities between areas. For cladogenetic range evolution, we describe a rationale for enumerating scenarios for the inheritance of geographic ranges by sister species from their ancestor. Importantly, our approach excludes ad hoc hypotheses of area vicariance; vicariant speciation scenarios are allowed, but they are neither explicitly associated with hypothetical area events, nor favored *a priori*.

We demonstrate how our approach can be used to calculate and compare likelihoods not only of alternative ancestral ranges for a given clade, but also of alternative scenarios of range inheritance at lineage branching points (speciation). The method therefore goes beyond asking the question “What was the ancestral geographic range of this clade?” to address “What was the geographic pattern of speciation in that ancestor?” and “How were ranges inherited by the daughter species?” These latter two questions are somewhat novel to methods for historical biogeography, and they focus attention on the geographic implications of speciation in a phylogenetic context. We illustrate the approach with hypothetical examples and a re-analysis of a previously published study of redbuds (Fabaceae: *Cercis*).

A MODEL FOR INFERRING THE GEOGRAPHIC HISTORY OF LINEAGES

The approach we develop here attempts to model ancestor-descendant change in the geographic range of lineages, with the goal of evaluating the likelihood of extant species ranges given information from phylogeny and paleogeography. In broad outline it is thus similar to the more familiar task of evaluating the likelihood of an organismal trait (character) on a phylogeny, given an evolutionary model for that character (e.g., Pagel 1994; Schluter 1995). Distinct geographic ranges in this case are equivalent to character states. In both cases, likelihoods are based on probabilities of ancestor-descendant change. In our model, the range of a lineage evolves as a result of two kinds of events, dispersal and local extinction, with each contributing to a stochastic process operating within the spatial and temporal constraints of discrete areas and the availability of dispersal routes between them. This process determines the transition probabilities associated with observing given ranges in an ancestor and its descendant. Unlike character analysis, in which speciation is assumed to produce descendants initially having the same state as their ancestor, our method enumerates the possible scenarios of ancestral range subdivision and inheritance that includes daughters inheriting nonidentical ranges. The global likelihood is then calculated using these transition probabilities and inheritance scenarios by adapting a well-known recursive algorithm.

Model Components

The model has two main components. The paleogeographic component consists of a set of areas with connections (dis-

persal routes) between pairs of areas. The biological component consists of a phylogeny relating the species of interest with branch lengths calibrated to yield lineage divergence times and parameters for rates of lineage dispersal between areas and extinction within an area.

Areas.—An area is a discrete geographic unit that is assumed to maintain its identity over the duration of interest. We define *A* to be the set of two or more areas that encompasses the total geographic range of the species under consideration, including any areas that were possibly inhabited by ancestors of those species. Although the geographic units recognized in the model may often correspond to areas of endemism (e.g., Hausdorf 2002), the use of rigorous methods to define such areas (e.g., Linder 2001; Mast and Nyffeler 2003) will not generally be necessary for our purposes. Instead, areas should be circumscribed according to the particular aspects of geographic history that are motivating the analysis: for example, if the history of dispersal between continents is of primary interest then continents may be specified as areas in the model.

Modeling areas as discrete units through time does not explicitly preclude plate tectonics or other mechanisms by which areas may become subdivided (area vicariance) or connected. Those phenomena may be incorporated into the model by parameterizing the history of dispersal opportunities between areas (see below).

Connections.—A connection represents a geographic link between two areas through which dispersal has been possible. It is essentially a function describing the relative probability of dispersal success through time. Formulating such a function ideally incorporates all relevant information available from geology, biology (e.g., dispersal mechanisms), paleoclimatology, and so on. Multiple dispersal routes may exist between two areas (e.g., the North Atlantic and Bering land bridges between North America and Eurasia), and each requires its own probability function (Fig. 1). Some types of connections, such as land bridges exposed by lowered sea levels, may have distinct peaks in probability representing times of high dispersal opportunity, while other types, such as prevailing winds or ocean currents between islands, may have probabilities that are more constant through time but are asymmetrical with respect to the direction of dispersal (Fig. 1).

Lineages.—A lineage corresponds to a branch (internode) on the phylogeny. A time-calibrated phylogeny provides estimates of the duration of each branch, meaning the interval from lineage origin (when it diverged from its most recent common ancestor) to when it either split into descendant branches or reached the present. We assign parameters to two inherent properties of a lineage that bear on inferring its geographic history: (1) its rate of dispersal to other areas (λ_D) and (2) its rate of extinction within an area (λ_E). Dispersal and extinction are modeled as Poisson processes with waiting times between events in an area distributed exponentially according to their respective rate parameters. For simplicity we assume that dispersal and extinction rates, while not constrained to be equal, are constant across lineages and through time.

The geographic range of a lineage, here denoted *R*, is the subset of areas in the model in which the lineage is distributed

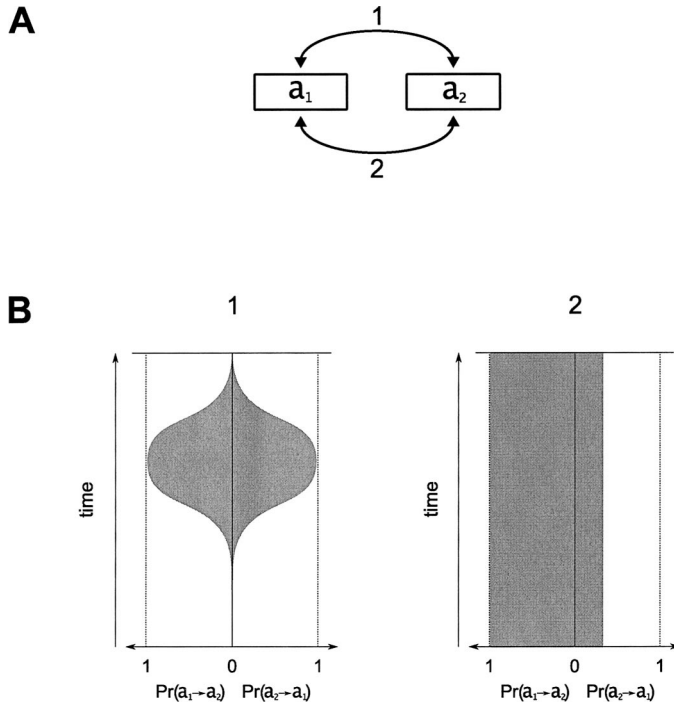


FIG. 1. (A) A hypothetical example of two areas, a_1 and a_2 , that persist as discrete entities over the duration of interest. The areas share two connections, each requiring specification of a dispersal function that describes the relative probability of dispersal success through time. (B) The probability of success across connection 1 varies through time, but is symmetric with respect to direction (as might be expected for dispersal across a land bridge exposed by lowered sea levels), whereas the probability of success across connection 2 is constant through time but favors dispersal from a_1 into a_2 (as might be expected for prevailing winds or currents between adjacent islands).

($R \subseteq A$) with $R(t)$ being the range at a specific time t . Let S denote the set of conceivable ranges a lineage may occupy. This set comprises all 2^n subsets of A , including the empty set, where n is the number of areas in A . For example, if $A = \{a, b\}$ then $S = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$. Often we are only interested in cases where a lineage is actually present (see below) so we define S_{trunc} as the truncated set of S that excludes the empty set.

Estimating Probabilities of Ancestor-Descendant Range Evolution

For each lineage (internode), we wish to know the matrix of transition probabilities from ranges in S_{trunc} at the time of lineage origin to ranges in S at the end of the lineage's duration (i.e., when it branches or reaches the present). We denote this matrix \mathbf{P} . It is analogous to the probability matrix of transitions between character states (e.g., nucleotide bases) with an important difference: the set of possible ranges at the end point includes the empty set, indicating global lineage extinction.

We do not attempt to derive \mathbf{P} analytically; instead, we estimate it by simulating range evolution from all starting points in S_{trunc} over the duration of each lineage. The basic procedure randomly generates a Markov chain of events through time based on dispersal and extinction being superimposed Poisson processes operating within the confines of the paleogeographic model (Fig. 2). Consider a lineage that starts at time t_0 and ends at time t_1 . We begin with the lineage having a range $R(t_0)$ at time t_0 . The time and place of the next event is predicted by randomly drawing a waiting time Δt for each area in R from the exponential distribution with mean $1/(\lambda_D + \lambda_E)$. The event occurs in the area with the minimum waiting time at time $t_0 + \Delta t$. The relative probability of the event being dispersal out of or extinction within

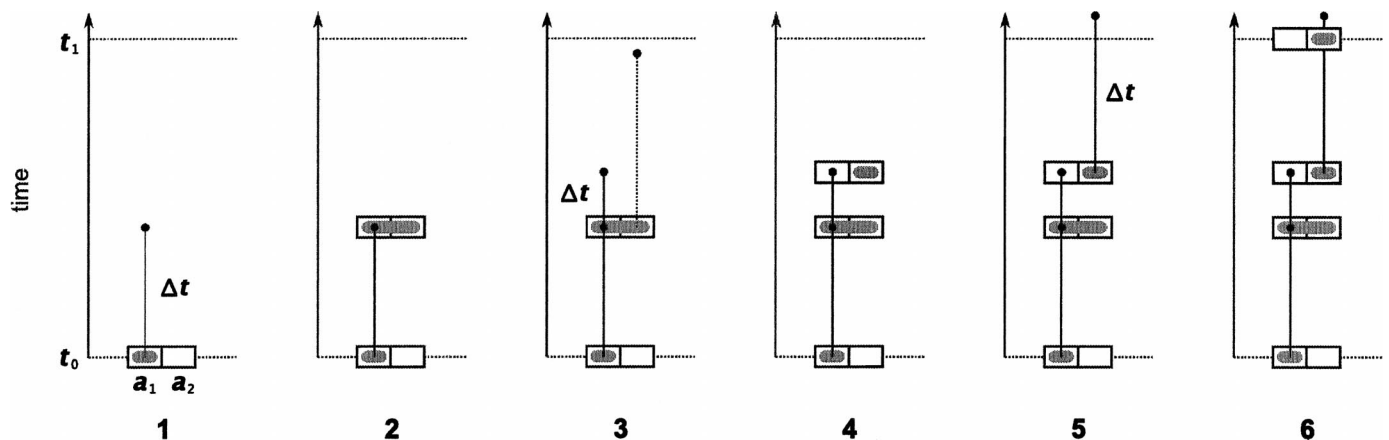


FIG. 2. Simulating the evolution of geographic range (R) along a lineage. Here there are two areas, a_1 and a_2 , respectively represented by the left and right halves of the horizontal rectangular boxes. (1) For each area occupied by the lineage at the start time t_0 (here $R = \{a_1\}$), the time to the next dispersal or extinction event, Δt , is randomly generated from a Poisson process. The probability of the event being dispersal or extinction is determined by their relative rates. The event, here dispersal, occurs in the area with the minimum waiting time. (2) A dispersal route (connection) is selected at random from those available, and dispersal success is stochastically determined by the probability function for that connection at that time. In this case, the lineage successfully disperses from a_1 into a_2 , and the latter area is added to its range ($R = \{a_1, a_2\}$). If dispersal had failed, the range would have remained unchanged. (3) Event waiting times are randomly generated for each area in R and a_1 has the minimum. (4) Extinction occurs in a_1 , removing it from R . (5) Waiting times and events are iteratively generated until the time of the next event exceeds the endpoint of the lineage (t_1), or the lineage has gone extinct in all areas. (6) The range of the lineage at time t_1 is recorded as the outcome of the simulation.

the area is determined by their relative rates, for example, $\Pr(\text{dispersal}) = \lambda_D / (\lambda_D + \lambda_E)$. If the event is dispersal, a dispersal route is chosen at random from those connecting the area to others and the probability of dispersal success is given by that connection's dispersal function at the time of the event. If dispersal is successful the destination area is added to R . If the event is extinction, the area of the event is removed from R . The procedure is iterated using updated values of $R(t)$ until the time of the next event exceeds the endpoint of the lineage ($t + \Delta t > t_1$) or the lineage has gone extinct in all areas ($R(t) = \emptyset$). The range of the lineage at this point is recorded as the outcome of the simulation.

Many simulations are run to generate a frequency distribution of outcomes, $R(t_1)$, given $R(t_0)$ and the model parameters. These frequencies estimate the transition probabilities $\Pr[R(t_1) | R(t_0)]$ where $R(t_i) \in S$. Iterating the procedure over all $R(t_0) \in S_{\text{trunc}}$ yields an estimate of the range transition matrix (\mathbf{P}). For a single lineage, the likelihood of observing $R(t_1)$ given $R(t_0)$ is

$$L[R(t_1) | R(t_0)] = \Pr[R(t_1)] \times \Pr[R(t_1) | R(t_0)] \times \Pr[R(t_0)] \quad (1)$$

where $\Pr[R(t_1) | R(t_0)]$ is given by \mathbf{P} , and $\Pr[R(t_0)]$ and $\Pr[R(t_1)]$ are the prior probabilities of the start and end of the lineage, respectively. We assume uniform priors for non-empty ranges: $\Pr(R) = 1/(2^n - 1)$ for all $R \in S_{\text{trunc}}$.

Given estimates of \mathbf{P} for all lineages, the next step in calculating the likelihood of observed species ranges involves evaluating the fractional likelihoods of ancestral ranges at internal nodes on the phylogeny, a step that requires consideration of how speciation affects the inheritance of geographic range.

Modeling Speciation and Range Inheritance

The ancestral range of an internal node on a phylogeny is the range of the lineage subtending the node immediately prior to branching. The fractional likelihood of a particular ancestral range is the product of its prior probability and the conditional likelihoods of each immediate descendant (daughter) lineage. To calculate the fractional likelihood, it is therefore critical to consider how an ancestral range is inherited by daughter lineages following speciation. Unlike characters, for which branching of an ancestral lineage that is fixed for a state will produce daughter lineages initially also fixed for that state (identical inheritance), geographic ranges are not always expected to be inherited identically. This stems from speciation being an inherently spatial event, as explained below.

Consider an ancestral species occurring in just a single area. In our model, speciation (cladogenesis) in that ancestor is constrained to occur in that area, because cladogenesis events are considered to be instantaneous and independent of dispersal and extinction, which are anagenetic processes. (Note that speciation within an area does not necessarily imply sympatric speciation, as areas are generally assumed to be large enough to encompass multiple populations; within an area, a population can become reproductively isolated in allopatry.) In this case both daughter species will inherit the same range as their parent (scenario 1, Fig. 3).

Now consider speciation in an ancestor that is widespread (i.e., its range spans more than one area). In this case, one or both daughters will inherit a range different from their parent, based on the following basic premise: that speciation is localized to a single area or, in other words, that one of the two daughter species originates in and inherits a range of only a single area. For each area in the ancestral range, there are two possible scenarios. The first is that the population(s) of the ancestor within the area diverge (by selective sweeps of mutations conferring reproductive isolation, stochastic coalescence, etc.) from those in other areas comprising the remainder of the ancestral range. As a result, one daughter species inherits a range of one area, while the other inherits the remainder of the ancestral range (scenario 2, Fig. 3). The second possibility is that speciation occurs within the area as described above, resulting in one daughter originating in that area and populations of the ancestor across the remainder of the range—including the area of speciation—continuing to exist as the other daughter. In this scenario, one daughter inherits a range of just the area where divergence occurred, while the other inherits the entire ancestral range (scenario 3, Fig. 3).

Given our assumption that speciation involves divergence either (1) between a single area and the remainder of the range or (2) within a single area, enumeration of possible range inheritance patterns is straightforward. We assume that each area in the range is equally likely to be the area involving speciation. We also assume that if the range includes more than one area then possibilities (1) and (2) are equiprobable for each area. In all cases, there are two ways that two ranges may be inherited by the daughter species. Enumeration of these scenarios is illustrated in Figure 3.

To clarify why we model speciation as an event involving only a single area, consider a hypothetical ancestral range comprising four areas, $\{a, b, c, d\}$, that becomes subdivided by speciation into daughter ranges $\{a, b\}$ and $\{c, d\}$. From the perspective of DIVA, this pattern could be explained by area vicariance if $\{a, b\}$ and $\{c, d\}$ represent subdivisions of a formerly contiguous set of areas $\{a, b, c, d\}$, but this explanation invokes a historical area event without considering the spatial and temporal context of that event, and we wish to avoid making this kind of ad hoc hypothesis. The least realistic scenario we can imagine for speciation in a widespread ancestor is that in which the ancestral range is inherited in its entirety by both daughters. Consider $\{a, b, c, d\}$ being inherited identically: this implies the simultaneous origin of both daughter species in each area in the range. This scenario could be explained by parallel allopolyploid origins (e.g., Soltis et al. 2004) or parallel adaptive divergence (e.g., Rajakaruna et al. 2003) followed by introgression across the range, but for present purposes we have not attempted to incorporate scenarios involving reticulate evolution into our model. Alternatively, this pattern could be due to erroneous species recognition (i.e., one apparent species is actually a unique, independently evolved lineage in each area) or erroneous area identification (i.e., the areas in the model in fact comprise just a single area within which speciation occurred).

Calculating the Likelihood of Observed Species Ranges

Consider a node x with descendant nodes y and z . Let $R_x(t_x)$ be the ancestral range at node x , where t_x is the divergence

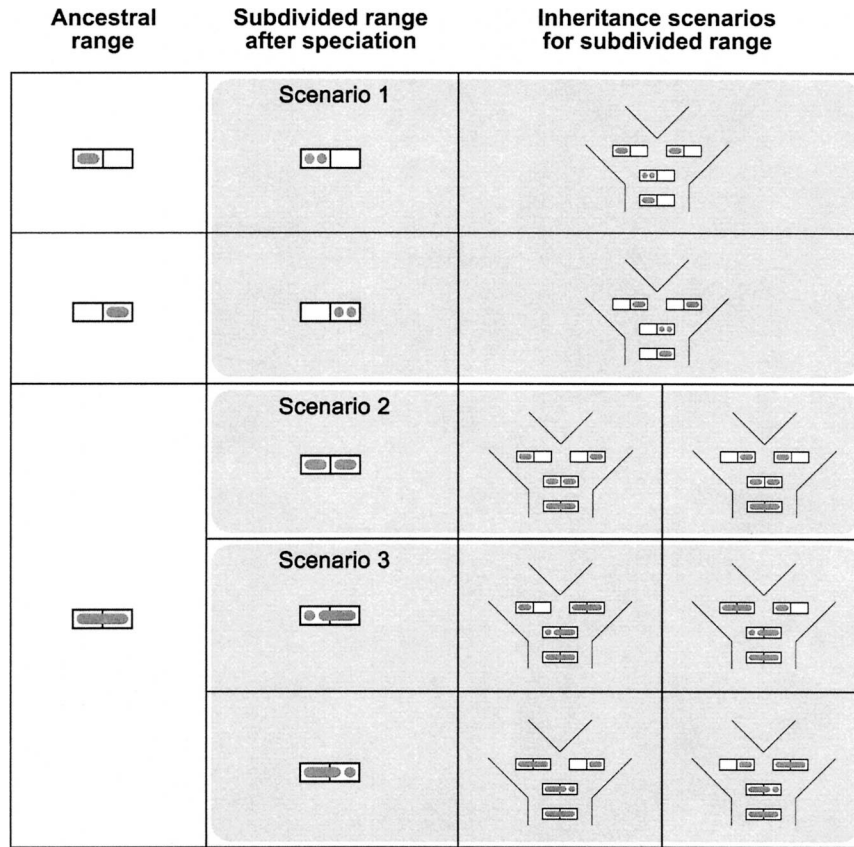


FIG. 3. Inheritance of geographic ranges following speciation is modeled as three distinct scenarios. Speciation is assumed to involve a single geographic area. When the ancestral range comprises a single area, that range is inherited identically by both descendant lineages (scenario 1). When the ancestor is widespread (i.e., its range spans > 1 area), one or both daughter species may inherit a range different from their ancestor. Geographically isolated populations may diverge, such that daughters inherit mutually exclusive ranges (scenario 2). Alternatively, when speciation occurs within an area, one daughter inherits a range of just the area where divergence occurred, while the other (the remainder of the ancestral lineage) inherits the entire ancestral range (scenario 3).

time. Following speciation, the ancestral range splits into ranges R_{x1} and R_{x2} , which are inherited by the lineages leading to nodes y and z , respectively. The prior probability of this scenario is the product of three terms, namely the priors associated with the ancestral range, how it becomes subdivided, and how the new ranges are inherited:

$$\Pr(X) = \Pr(R_x) \times \Pr(R_{x1}, R_{x2} | R_x) \times \Pr[R_y(t_x) = R_{x1}, R_z(t_x) = R_{x2}]. \quad (2)$$

The likelihood of the above scenario is the product of its prior probability and the conditional likelihoods of the descendant nodes y and z :

$$L(X) = \Pr(X) \times L[y | R_y(t_x) = R_{x1}] \times L[z | R_z(t_x) = R_{x2}], \quad (3)$$

where the conditional likelihood for node y is

$$L[y | R_y(t_x) = R_{x1}] = \sum_{R_y \in S_{\text{trunc}}} \Pr[R_y(t_y) | R_y(t_x) = R_{x1}] L[R_y(t_y)]. \quad (4)$$

with the same expression applying to node z and R_{x2} . The terms $\Pr[R_y(t_y) | R_y(t_x) = R_{x1}]$ and $\Pr[R_z(t_z) | R_z(t_x) = R_{x2}]$ are

range transition probabilities given by \mathbf{P}_y and \mathbf{P}_z , respectively. If y is a terminal node, then $L[R_y(t_y)] = \Pr(R)$ if R is the extant range of that species; otherwise $L[R_y(t_y)] = 0$. If y is an internal node, $L[R_y(t_y)]$ is the fractional likelihood of R_y at y prior to branching, which has already been calculated during the postorder traversal of the tree.

The fractional likelihood of an ancestral range at an internal node is thus obtained by summing over all its inheritance scenarios. Obtaining the total likelihood for the observed species ranges involves proceeding from the tips of the tree to the root, using the familiar pruning algorithm of Felsenstein (1981) to evaluate the fractional likelihoods of ancestral ranges at internal nodes. At the root, the total likelihood is calculated by summing the fractional likelihoods associated with each distinct ancestral range.

Comparing the Likelihoods of Alternative Range Inheritance Scenarios

Although it is interesting and worthwhile to calculate the global likelihood of observed species ranges given a phylogeny and paleogeographic model, our main objective at this point is not to optimize the free parameters (rates of lineage dispersal and extinction) against the data. Instead, we are

generally more interested in comparing the likelihoods of alternative biogeographic hypotheses given a set of parameter values. We focus here on comparisons of range inheritance scenarios (the rightmost column in Fig. 3).

Comparing alternative range inheritance scenarios at an internal node on the phylogeny is straightforward and simply involves fixing that node for each scenario in turn, each time recalculating the overall likelihood. This allows the likelihoods of alternative events at a particular node to be evaluated without conditioning on particular events at other nodes in the tree. This approach has been used in likelihood-based inference of ancestral character states (e.g., Schluter et al. 1997; Pagel 1999). The statistical significance of likelihood differences can be assessed using the conventional cutoff value of two log-likelihood units (Edwards 1992).

APPLICATION OF THE MODEL

To demonstrate use of the model, we begin with a simple hypothetical example that consists of a set of two areas, $A = \{a_1, a_2\}$, and a phylogeny of three species. Species s_1 is distributed in a_1 ; species s_2 and s_3 are both distributed in area a_2 . The areas share a single connection through which the probability of dispersal success is equal to one in both directions over the time interval of interest (0–20 million years, my). For this and subsequent analyses, the objective is to infer the most likely scenarios of range inheritance at all internal nodes. Inferences are made over a range of parameter values for lineage dispersal and extinction, including “high” rates ($\lambda_D + \lambda_E = 0.1$, corresponding to an average of one event per 10 my) and “low” rates ($\lambda_D + \lambda_E = 0.01$; one event per 100 my), for each combination of $\lambda_D > \lambda_E$, $\lambda_D = \lambda_E$, and $\lambda_D < \lambda_E$. For each internode and pair of parameter values, 10^5 simulations were run for each range in S_{trunc} to estimate the range transition matrices.

Results from the likelihood analysis of the three-species example, and the corresponding DIVA results, are shown in Figure 4. The highest overall likelihood is obtained at a low rate of dispersal and extinction where $\lambda_D > \lambda_E$. Magnifying the asymmetry of dispersal and extinction or lowering the overall rate even further does not increase the likelihood. Of all parameter combinations tested, only the high rate, $\lambda_D < \lambda_E$ combination is significantly less likely (i.e., more than two log-likelihood units lower than the maximum likelihood). Thus the phylogeny, model, and data in this example do not yield globally unequivocal inferences when model parameters are free to vary, but this is unsurprising given such a small phylogeny. At the root node, the most likely scenario is that of a widespread ancestor undergoing speciation in area a_2 , but it is also plausible that the ancestor was restricted to area a_2 because the difference in log-likelihood between the two scenarios is only 0.6. These ranges are equally optimal under DIVA. The ancestor of species s_1 and s_2 is inferred to have been widespread and undergone speciation between areas, a result that is robust over all low-rate parameter values, and is identical to the DIVA solution.

At low rates where $\lambda_D < \lambda_E$, a widespread ancestor at the root receives significant support, because a narrower range would require dispersal. At high rates, the most likely scenario at the root node is an ancestor restricted to a_2 , but in

contrast to low rates, the data and model have less power to distinguish alternative ancestral range scenarios. By contrast, DIVA captures only a small subset of this uncertainty and yields inferences that are insensitive to prior expectations about rates of dispersal and extinction.

To investigate whether additional data might narrow the range of plausible inferences, we modified the hypothetical example by adding another species basally to the tree, extending the age of the root another 10 my. This species is also currently distributed only in area a_2 . Repeating the likelihood analysis using this phylogeny yields the results shown in Figure 5A. As in the previous example, the highest likelihood is obtained with a low-rate, $\lambda_D > \lambda_E$ parameter combination, but in this case all other parameter values yield significantly lower likelihoods. A larger phylogeny has thus increased the statistical signal in historical range inferences. The ancestor of species s_1 and s_2 is again inferred to have been widespread. At the other internal nodes, the most likely scenario is an ancestor restricted to area a_2 , but a widespread ancestor undergoing speciation within a_2 (a scenario favored at the root in the previous example) remains plausible. In comparison, DIVA identifies only the maximum-likelihood ancestral ranges as the optimal solution, implying a dispersal event from a_2 into a_1 along the internal branch leading to $s_1 + s_2$ (Fig. 5B).

In the four-species case, DIVA yields a single reconstruction of ancestral ranges that is invariant in the face of information about divergence times and opportunities for dispersal. By contrast, inferences under the likelihood model are directly influenced by such information. For example, if the duration of the internal branch leading to $s_1 + s_2$ is long (i.e., the divergence of $s_1 + s_2$ was recent relative to the divergence of s_3), then dispersal is more likely along that branch and likelihood inferences give stronger support to the DIVA optimization (Fig. 6A). Conversely, if the duration of the internal branch is short and the divergence of s_1 and s_2 was farther back in time (Fig. 6B), dispersal becomes less likely along the short internal branch and more likely along the long branch leading to s_1 , yielding equivocal range inferences at internal nodes. Extinction is also more likely along the terminal branch, introducing the possibility that s_1 originated from an ancestor restricted to a_2 , a scenario requiring dispersal and extinction that is excluded from DIVA inferences. Finally, if dispersal from a_2 to a_1 is much less likely than in the reverse direction, ancestral range scenarios increasingly favor widespread ancestors at all internal nodes, with speciation occurring within a_2 at the root and in the immediate ancestor of s_3 (Fig. 6C).

An Empirical Example: *Cercis*

To demonstrate use of the model in an empirical case, we examine the biogeographic history of *Cercis* (Fabaceae). This angiosperm group is distributed throughout the temperate Northern Hemisphere and was studied by Davis et al. (2002). *Cercis* constitutes about 10 species distributed among four areas: *C. canadensis* grows in eastern North America (EN) and northeastern Mexico, *C. occidentalis* grows in western North America (WN), *C. siliquastrum* and *C. griffithii* (the latter unsampled) grow in west-central Eurasia (WE), and six

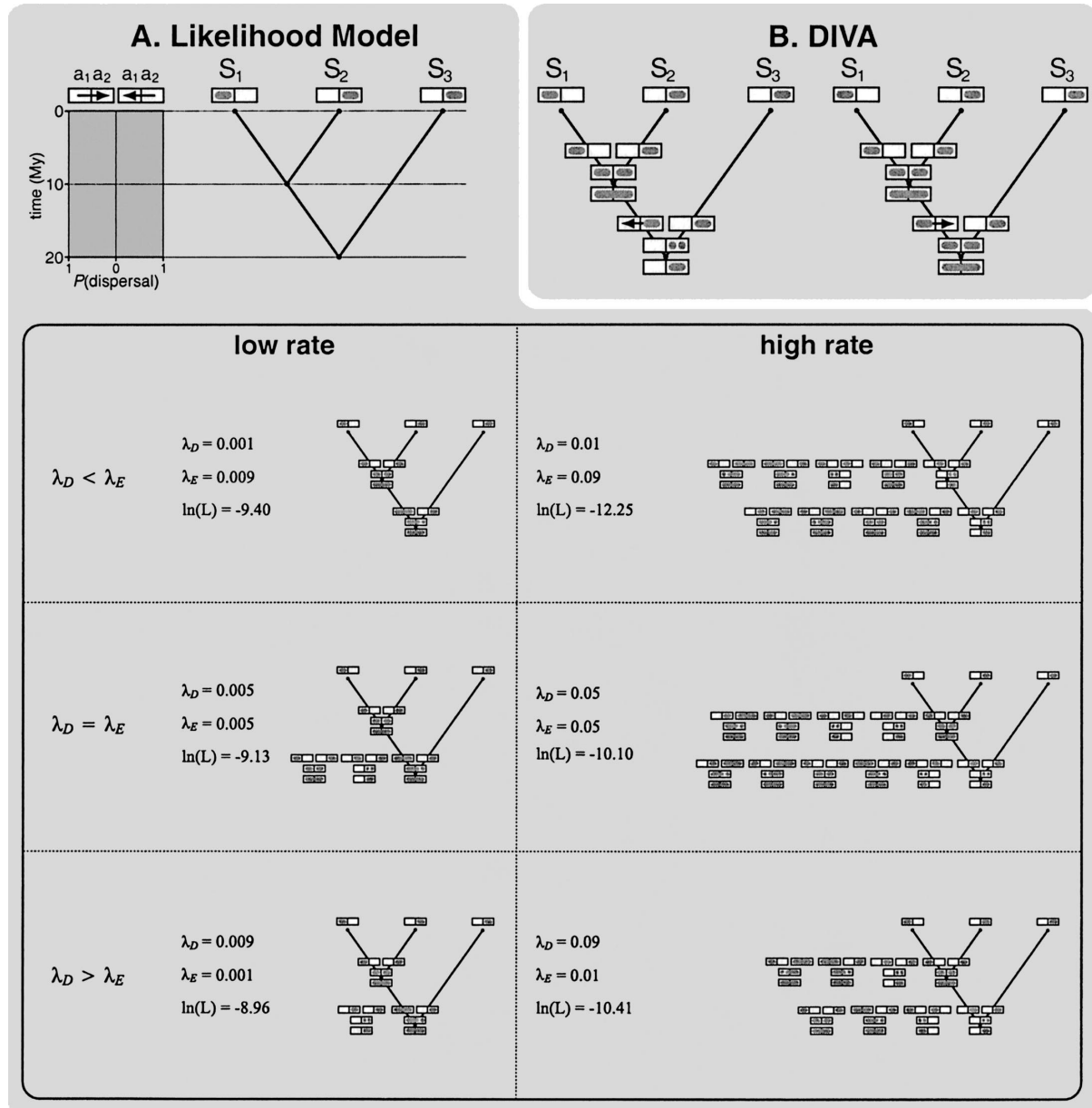


FIG. 4. (A) Likelihood inference of lineage geohistory for a hypothetical three-species phylogeny. The model includes two areas (a_1 and a_2) sharing a single connection, through which the probability of successful dispersal is constant over time and symmetrical with respect to direction. Likelihood estimates of ancestral ranges and subdivision-inheritance scenarios were obtained for different relative dispersal rates ($\lambda_D < \lambda_E$, $\lambda_D = \lambda_E$, and $\lambda_D > \lambda_E = 0.1$) and two overall rates, “low” ($\lambda_D + \lambda_E = 0.01$) and “high” ($\lambda_D + \lambda_E = 0.1$). Adjacent to every internal node are scenarios within two log-likelihood units of the maximum, ordered left to right from lowest to highest likelihood. The highest overall likelihood is obtained at a low rate with $\lambda_D > \lambda_E$. (B) Ancestral ranges and implied speciation scenarios inferred with dispersal-vicariance analysis (DIVA). The two equally most-parsimonious reconstructions correspond to the plausible scenarios inferred from the model parameters yielding the highest likelihood, although in the latter case a widespread ancestor at the root is favored.

species are endemic to eastern Eurasia (EE), of which Davis et al. (2002) sampled *C. occidentalis*, *C. gigantea*, and *C. chingii*. The fruits of *Cercis* are papery legumes likely dispersed by wind.

In applying our model to *Cercis* we used the maximum-likelihood time-calibrated molecular phylogeny from Davis et al. (2002) and combined the representatives of *C. canadensis* into a single terminal taxon (Fig. 7). We parameterized connections between the four areas as follows. Within con-

tinents (i.e., between WN and EN and between WE and EE) the probability of dispersal success was one over the time period under consideration (55 my). Between WN and EE (the Bering land bridge), the probability was one until 5 my ago, and zero thereafter (based on Hopkins 1967; Tiffney 1985, 2000). Between WE and EN (the North Atlantic land bridge), the probability was one until 25 my ago, then declined linearly to zero at 15 my ago (based on McKenna 1983; Manchester 1999; Tiffney 2000).

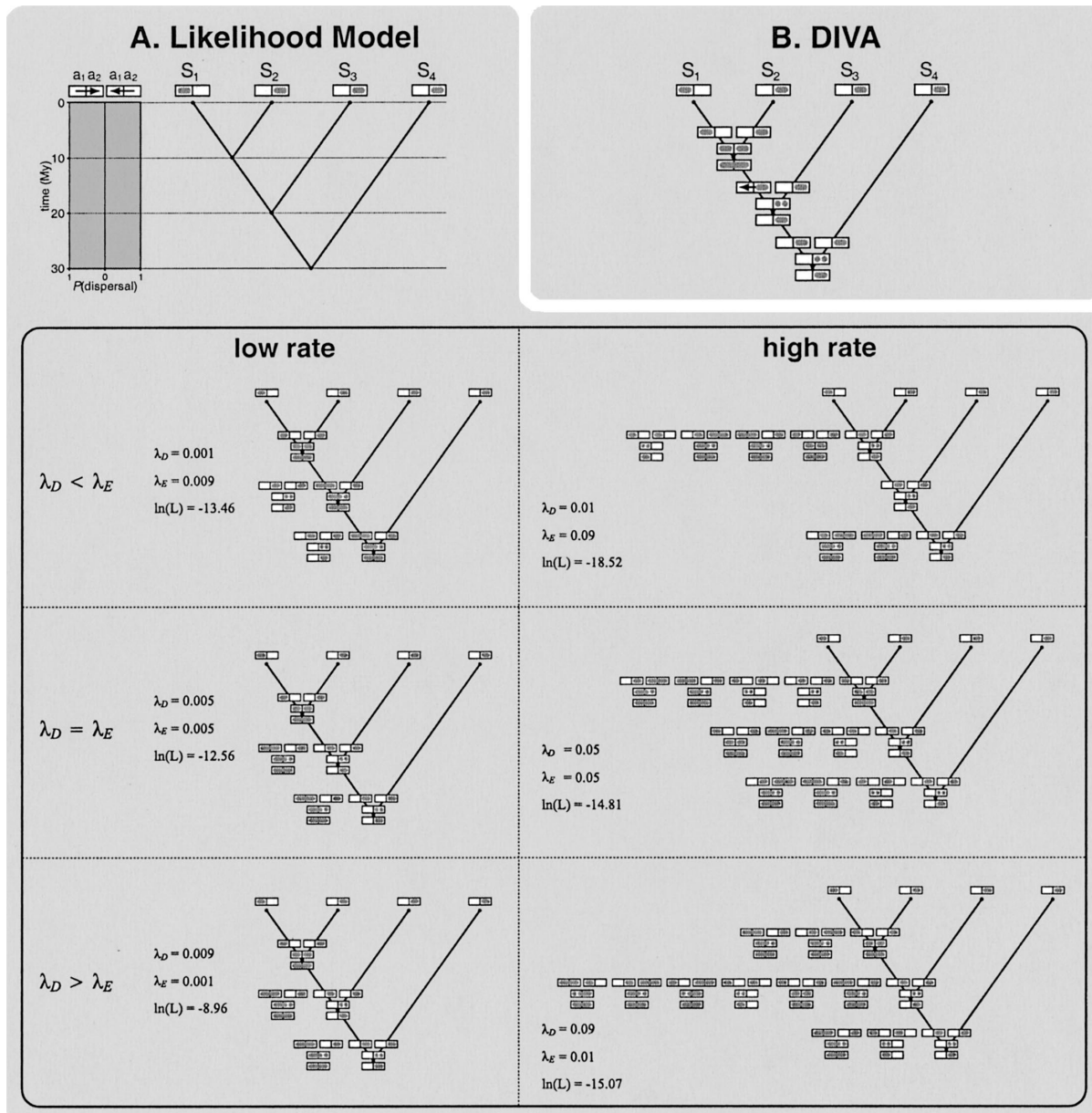


FIG. 5. (A) Likelihood inference of lineage geohistory for a hypothetical four-species example. Components of the geographic model and details of the analyses performed are the same as described previously (Fig. 4A). As in the three-species case, the highest overall likelihood is obtained at a low overall rate with $\lambda_D > \lambda_E$. (B) Ancestral ranges and implied speciation scenarios obtained using dispersal-vicariance analysis (DIVA). The single most-parsimonious reconstruction corresponds to the maximum-likelihood scenarios inferred in (A) but in the latter analysis, alternative scenarios cannot be statistically ruled out.

Using this paleogeographic model, likelihoods of the *Cercis* data at high rates ($\lambda_D + \lambda_E = 0.1$) are more than five log-likelihood units lower than those at low rates ($\lambda_D + \lambda_E = 0.01$); thus, we present only the latter results.

For parameter values where the rate of dispersal is less than, equal to, and greater than the rate of extinction, log-likelihoods are -41.79 , -40.75 , and -40.17 , respectively (Fig. 8A). As in the hypothetical examples, the highest likelihood is obtained at low overall rates with $\lambda_D > \lambda_E$. At the root (node 1 in Fig. 8A), the most likely ancestral range scenario is of a widespread ancestor, but several alternative scenarios that involve more restricted ancestral ranges are

also plausible. All scenarios inferred at the root place the origin of *C. chingii* in EE with most inferring a speciation event within EE (i.e., the ancestral range is inherited by the lineage leading to node 2 in Fig. 8A). At node 2, the most likely scenario of lineage splitting involves a widespread ancestral range being inherited from node 1 and vicariant speciation resulting in divergence of the lineage in EE from the rest. The clade of *C. chinensis* + *C. gigantea* is unequivocally inferred to originate from an ancestor that was restricted to EE and underwent speciation there. For the branch leading from node 2 to the clade containing *C. occidentalis*, the most likely range of origin is {WN, EN, WE}, but other more re-

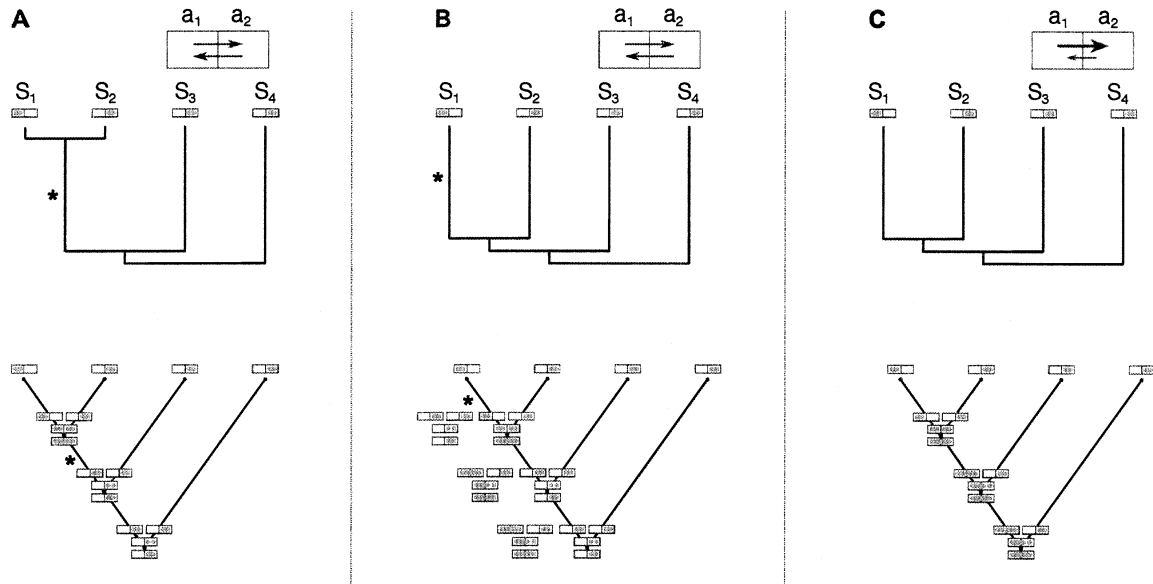


FIG. 6. The effect of divergence times and asymmetric dispersal probability on maximum-likelihood inferences of ancestral range scenarios on the phylogeny from Figure 5. (A) The long duration of the internal branch leading to $s_1 + s_2$ increases the probability of dispersal along it (indicated by the asterisk), consequently increasing the likelihood of ancestral ranges estimated by DIVA (Fig. 5B). (B) Conversely, a shorter duration for this internal branch and long durations of terminal branches increases the likelihood of both dispersal and extinction along the branch leading to s_1 , introducing the possibility that s_1 originated from an ancestor restricted to a_2 , a scenario excluded from DIVA inferences. Alternative range scenarios at internal nodes are almost equally likely. (C) If dispersal from a_2 to a_1 is unlikely (asymmetric arrows at top right), likelihood inferences shift toward widespread ancestors, with vicariant speciation between s_1 and s_2 and speciation within a_2 at other internal nodes.

stricted ranges are also plausible. By contrast, the range at the end of the branch, immediately prior to speciation, is strongly supported to be {WN,EN,WE} and the extant ranges of *C. canadensis*, *C. siliquastrum*, and *C. occidentalis* are unequivocally inferred to be the result of vicariant speciation between areas subdividing widespread ancestors, an inference robust to all alternative topological resolutions of those species. For this clade, statistical support for the vicariant pattern

is bolstered by the short branch durations separating internal nodes, because the opportunity for dispersal and extinction along those branches are reduced.

For the sake of comparison, ancestral areas for *Cercis* were also estimated by DIVA (Fig. 8B). As in the hypothetical examples, the ancestral areas inferred by DIVA were very similar to those estimated under the likelihood approach. However, in some cases the two methods infer different spe-

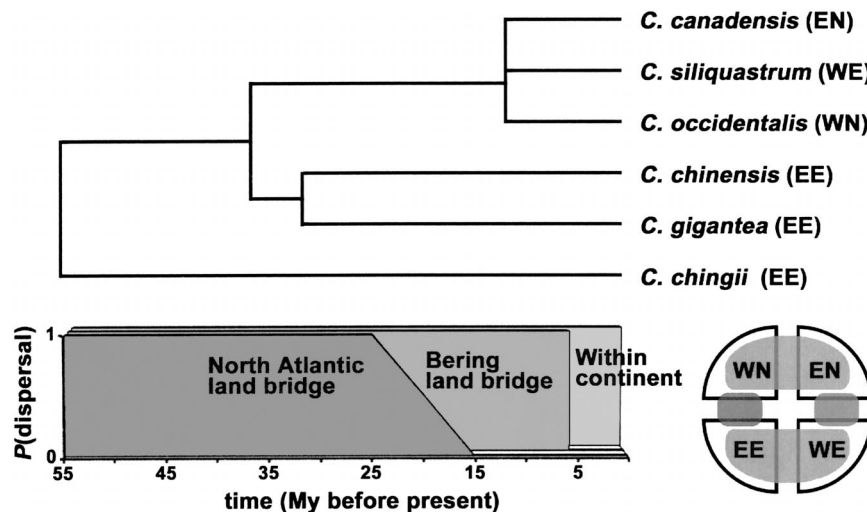


FIG. 7. Maximum-likelihood estimate of *Cercis* phylogeny with divergence times modified from Davis et al. (2002). The biogeographic model comprises four areas: eastern North America (EN), western North America (WN), eastern Eurasia (EE), and western Eurasia (WE). There are four connections between them: the North Atlantic land bridge (linking WE and EN), the Bering land bridge (linking EE and WN), and two intracontinental links (within North America and Eurasia). The probability of dispersal success through these connections (plotted through time) is symmetrical with respect to direction.

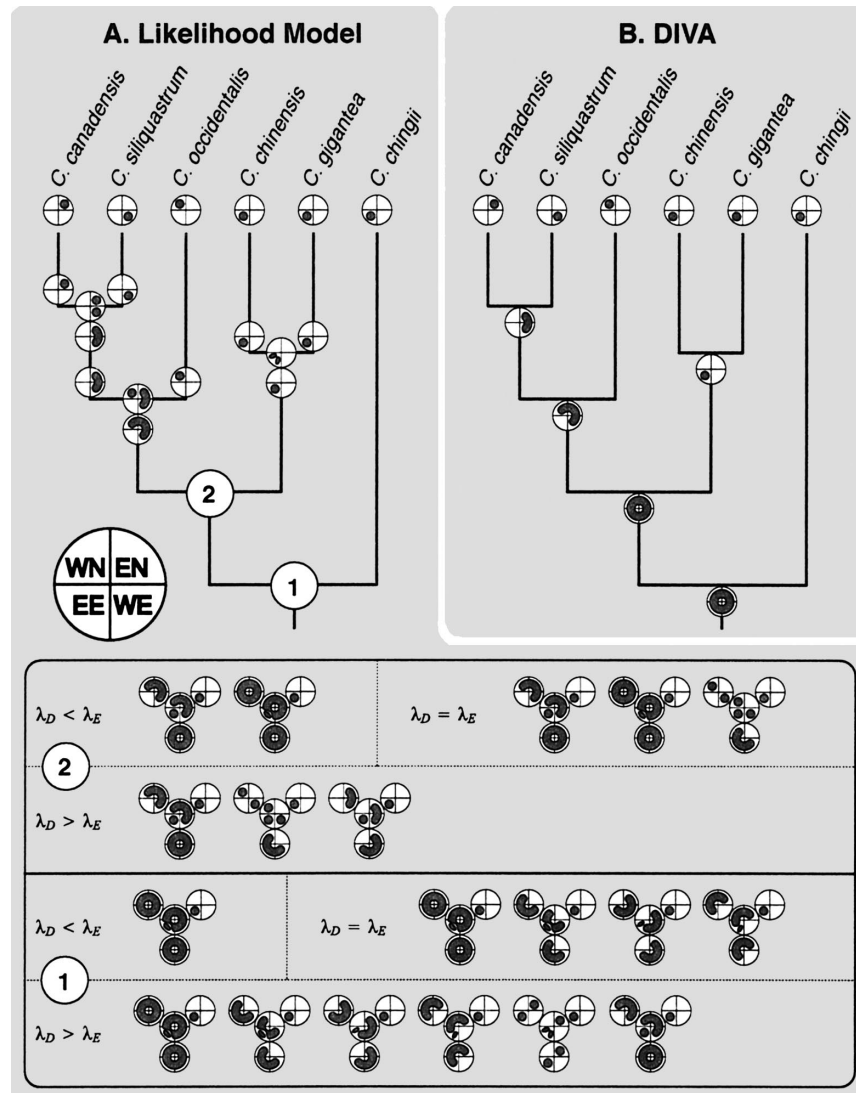


FIG. 8. (A) Likelihood estimates of ancestral ranges and subdivision-inheritance scenarios for *Cercis* under different relative dispersal and extinction rates, at a low overall rate ($\lambda_D + \lambda_E = 0.01$). For nodes 1 and 2, solutions within two log-likelihood units of the maximum are ordered, left to right, from highest to lowest likelihood in the corresponding cells of the table. The highest overall likelihood was obtained at a low overall rate with $\lambda_D + \lambda_E$. Magnifying the asymmetry of dispersal and extinction or lowering the overall rate further did not increase the likelihood. (B) Ancestral ranges obtained using dispersal-variance analysis (DIVA).

ciation scenarios for those ancestral areas. At the root of the *Cercis* tree (node 1 in Fig. 8B), the DIVA analysis optimized a widespread ancestral range that was subdivided by area vicariance to give rise to the lineage leading to *C. chingii* (restricted to EE), and the lineage leading to node 2 (having the range {WN,EN,WE}). Subsequent dispersal of the latter lineage into EE restored a widespread ancestral range at node 2. A second episode of vicariance gave rise to the ancestor of *C. chinensis* and *C. gigantea*, which was restricted to (and subsequently speciated within) EE and the lineage leading to node 3 (having the range {WN,EN,WE}). Vicariant subdivision of the ancestral range at node 3 gave rise to *C. occidentalis* (restricted to WN) and the ancestor of *C. canadensis* and *C. siliquastrum* (restricted to {EN,WE}). A final vicariance event is invoked to explain the origin of *C. canadensis* in EN and *C. siliquastrum* in WE.

DISCUSSION

The likelihood model introduced in this paper represents a substantial departure from the conceptual basis and implementation of other “event-based” methods for inferring lineage history. Unlike DIVA, our approach requires an explicit description (parameterization) of dispersal opportunities between areas as well as estimates of lineage divergence times, and models geographic range evolution along a lineage as the result of stochastic dispersal and local extinction. It extends inferences beyond ancestral ranges to explicitly consider range inheritance scenarios associated with speciation, including a wider range of speciation models. In this section, we attempt to clarify and expand upon various aspects of the approach, emphasizing how model parameterization relates to a parsimony framework, the significance of speciation models, and the problem of widespread ancestors.

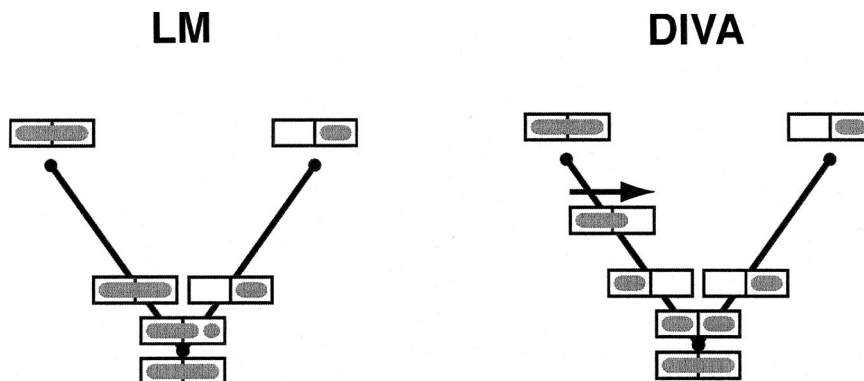


FIG. 9. Distinct range subdivision scenarios inferred by the likelihood method (LM) and implied by dispersal-vicariance analysis (DIVA), respectively, in the case where one daughter node is widespread and the other is restricted to a single area. The likelihood model allows speciation to occur within one area occupied by the widespread ancestor, with one daughter lineage inheriting the entire ancestral range and the other inheriting the area of origin; dispersal is not required. By contrast, DIVA invokes vicariant speciation (lineage divergence between areas), requiring subsequent dispersal to restore the range of the widespread daughter species.

Correspondence to DIVA.—We have shown how information on divergence times and dispersal opportunities can influence estimates of ancestral ranges and speciation scenarios (Fig. 6). Given that such information is not considered by DIVA, and that range estimates by both methods can be virtually identical (e.g., for ancestors not near the root of *Cercis* in Fig. 8), it is worthwhile to ask: Under what general conditions will this method yield results convergent with those of DIVA? Answering this question may be especially helpful in applying the method when reliable information on divergence times and dispersal opportunities is unavailable.

First, consider branch durations: DIVA treats all branches equally with respect to dispersal and extinction events. In the likelihood model, the expected number of events along a branch is the product of its duration and rates of dispersal and extinction. Assuming constant rates across the tree, the closest parametric approximation to DIVA is thus equal durations for branches, and these should be short if few events are expected to have occurred. Conversely, durations should be longer or rates higher on branches where the expectation of dispersal and/or extinction is greater (e.g., Fig. 6A). It perhaps makes more sense to specify branch-specific rates (i.e., relax the assumption of rate constancy) rather than arbitrarily adjust branch durations because the latter are generally easier to obtain from external data.

Second, consider dispersal: in the hypothetical three- and four-species cases, each DIVA reconstruction involved a single dispersal event on the trees, which have summed branch durations of 50 and 80 my, respectively. This result is roughly equal to the dispersal rate favored by the likelihood analyses ($\lambda_D = 0.009$), which yields the expectation of about one event per hundred million years.

Finally, consider extinction: in all the examples we considered, extinction is never required by DIVA—in fact, this is always the case if ancestral ranges are unconstrained (Ronquist 1996). Similarly, likelihoods were consistently highest at the lowest extinction rates tested ($\lambda_E = 0.001$), a result attributable to the effect of low rates reducing the probability of stochastic lineage extinction and increasing the likelihood of observing extant species ranges (with the side effect of favoring widespread ancestors; see below).

In summary, it appears that closest correspondence to parameter conditions implied by DIVA is obtained by assuming rate constancy, equal branch durations, low dispersal rates, and even lower extinction rates. In empirical applications of the likelihood approach, how might reasonable starting values for dispersal and extinction rates be generated? For practical purposes we suggest experimenting with a range of values, as demonstrated in the examples, but a useful rule of thumb may be to choose rates that correspond to prior expectations about the number of events on the phylogeny, given the summed branch durations.

Speciation scenarios.—Even in the case of identical ancestral range estimates, the speciation scenarios estimated by our method and those implied by DIVA may differ. When an ancestor and its descendants are restricted to a single area, both methods indicate the same scenario: speciation within the ancestral area (scenario 1, Fig. 3). When the ancestral range is widespread, the two methods may yield different scenarios. If the descendants are restricted to separate areas of the ancestral range, both methods indicate a scenario consistent with vicariant speciation (scenario 2, Fig. 3). However, if one descendant is widespread and the other is restricted to a single area, the two methods indicate different speciation scenarios (Fig. 8; scenario 3, Fig. 3). The likelihood model allows a daughter lineage to inherit the entire ancestral range, as might occur when one daughter species “buds off” from the lineage within one area and the remainder persists as its widespread sister lineage. By contrast, DIVA forces speciation to occur between areas of the ancestral range, such that the resulting daughter species are disjunct, requiring postspeciation dispersal to restore the range of the widespread daughter species (Fig. 9). By allowing a broader range of speciation models, and specifically by allowing the inheritance of a widespread ancestral range by one of the daughter lineages, our model tends to reduce the incidence of dispersal relative to DIVA.

Widespread ancestors.—In the so-called ‘ancestral area paradox’ (e.g., Bremer 1992, 1995; Ronquist 1997), extant species are typically restricted to a single area whereas ancestors are often inferred to be widespread; this result conflicts with the uniformitarian perspective that ancestral ranges

should be similar in extent to current ranges. This phenomenon is exhibited by both DIVA and our method. DIVA addresses the issue by allowing constraints on ancestral range size to be imposed, an option frequently exploited in empirical studies. In our parametric approach, prior probability distributions for ancestral ranges could be adjusted to condition on the sizes of extant ranges.

The tendency to infer widespread ancestors with the likelihood method is largely attributable to the fact that trees with longer branches require lower extinction rates. Recall that global extinction is a possible outcome of stochastic lineage persistence. Lineage survival along long basal branches (e.g., *Cercis chingii*; Fig. 7) thus becomes increasingly unlikely as extinction rate increases, shifting inferences toward low rates. This situation leads to widespread ancestors being favored, because if descendants are disjunct between two areas and their ancestor was restricted to one of those areas, one descendant must have dispersed out of and gone extinct in the ancestral area. Conversely, if the ancestor was widespread, vicariant speciation can explain the descendant ranges and extinction is not required.

The widespread ancestor problem may be due to a deficiency in the realism of both our model and DIVA, namely that the processes underlying anagenetic range evolution, particularly dispersal, are conceptually decoupled from the process of speciation. It is intuitive to view dispersal to a new area (range expansion) as an event that will increase the likelihood of subsequent vicariant speciation. Future work on models of range evolution that incorporate a stochastic coalescence component, such that widespread lineages are less likely to persist for long durations without splitting by vicariant speciation, should prove fruitful in addressing this deficiency.

Extending the Model

The specific realization of the approach presented here is very basic and could justifiably be criticized for being crude and incomplete. As currently formulated, our model incorporates lineage divergence times and dispersal opportunities between areas in estimating probabilities of geographic range evolution on phylogenetic branches. These are by no means the only sources of information relevant to the question, and more realistic inferences could conceivably be obtained if others were brought on board. We emphasize that the likelihood framework underlying our approach could readily be extended along several lines.

An important aspect of area history that could substantially enhance the realism of our model is habitat favorability through time (as might be provided by reconstructions of paleoclimate), as this bears directly on expectations of extinction rate. As currently described, persistence of a lineage within an area is determined solely by its intrinsic extinction rate, λ_E . Extinction is thus constant across lineages, areas, and through time. Modeling extinction rate as a function of both time and area would allow paleoecological data (e.g., on temperature or aridity) to be incorporated into historical inferences.

An important aspect of lineage history currently missing from the model is paleontological information. Fossils pro-

vide direct evidence for placing lineages in certain areas at certain times. It thus stands to reason that, to the extent that fossils can be assigned to specific phylogenetic branches, estimates of range transition probabilities along those branches should account for those observations—that is, simulations of lineage range evolution used to estimate \mathbf{P} must be consistent with the fossil data.

It is not hard to imagine other ways of incorporating external information into the basic framework we describe here, as illustrated by the following brief examples. Estimates of the size of areas through time might alter our initial assumption that within-area speciation (scenario 3, Fig. 3) is equiprobable across areas. Methods for locating shifts in diversification rate on phylogenies (e.g., Moore et al. 2004) could be used to specify lineage-specific expectations for extinction rate. Similarly, certain organismal traits (e.g., the winged wind-dispersed fruits of maples) might influence prior expectations about dispersal rate. Mapping those traits on a phylogeny could be used to justify lineage-specific values of λ_D .

Model Implementation

Python source code implementing the methods described here can be obtained from <http://www.phylodiversity.net/rree> or by contacting the corresponding author. Because of the current use of Monte Carlo methods to estimate \mathbf{P} , and because of speed limitations of the Python language, analyses are relatively slow (e.g., analysis of all parameter combinations for the *Cercis* example took about four hours on a 2.5 GHz processor). Run times increase exponentially with the number of areas in the model and linearly with the number of internodes in the phylogeny. For these reasons, analyses incorporating uncertainty in estimates of divergence times and/or phylogenetic relationships are currently impractical. However, these limitations are likely to be soon overcome: considerable speedups are possible with faster programming languages (e.g., Java; S. Smith, unpubl. data) and with methods that do not rely on simulation (R. H. Ree, unpubl. ms.).

Conclusions

Historical biogeography appears to be expanding its scope again after a period of focusing primarily on discerning area relationships using cladograms. At this juncture, traditional methods for extracting biogeographic patterns from tree topologies are too constraining, and new methods are needed for integrating more sources of relevant information (e.g., the timing of both lineage and geological events). We view the likelihood framework outlined here as a step in this direction—analogueous to the earliest probabilistic models of molecular evolution (e.g., Jukes and Cantor 1969). By explicitly incorporating temporal information and allowing a broader spectrum of speciation scenarios, we believe that our likelihood approach provides an improved basis for inferring ancestral ranges and how they are inherited following speciation. However, its greatest virtue may lie in its extensibility, not only with regard to integrating additional sources of data but also in its potential for addressing a wider variety of biogeographic questions. For example, although here we have highlighted ancestral area assessments within individual

clades, we envision applications involving multiple clades aimed at determining the frequency of alternative dispersal routes at different times or the assembly of regional species pools and ecological communities.

ACKNOWLEDGMENTS

We thank S. Otto and discussion group participants at the Field Museum for helpful feedback in developing these ideas and M. Moore for editorial advice. The manuscript was substantially improved by comments from two anonymous reviewers. MJD and COW were partially supported by National Science Foundation grant DEB-0212873.

LITERATURE CITED

- Bremer, K. 1992. Ancestral areas: a cladistic reinterpretation of the center of origin concept. *Syst. Biol.* 41:436–445.
- . 1995. Ancestral areas: optimization and probability. *Syst. Biol.* 44:255–259.
- Brooks, D. R. 1981. Henning's parasitological method: a proposed solution. *Syst. Zool.* 30:229–249.
- . 1985. Historical ecology: a new approach to studying the evolution of ecological associations. *Ann. Mo. Bot. Gard.* 72: 660–680.
- . 1990. Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Syst. Zool.* 39:14–30.
- Brown, J. H., and M. V. Lomolino. 1998. *Biogeography*. 2nd ed. Sinauer Associates, Sunderland, MA.
- Croizat, L. 1958. *Panbiogeography*. Privately published, Caracas, Venezuela.
- . 1964. *Space, time, form: the biological synthesis*. Privately published, Caracas, Venezuela.
- Davis, C. C., P. W. Fritsch, J. Li, and M. J. Donoghue. 2002. Phylogeny and biogeography of *Cercis* (Fabaceae): evidence from nuclear ribosomal ITS and chloroplast *ndhF* sequence data. *Syst. Bot.* 27:289–302.
- Donoghue, M. J., and B. R. Moore. 2003. Toward an integrative historical biogeography. *J. Integr. Comp. Biol.* 43:261–270.
- Edwards, A. W. F. 1992. *Likelihood*. Johns Hopkins Univ. Press, Baltimore.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17:368–376.
- Hausdorf, B. 1998. Weighted ancestral area analysis and a solution of the redundant distribution problem. *Syst. Biol.* 47:445–456.
- . 2002. Units in biogeography. *Syst. Biol.* 51:648–652.
- Hopkins, D. M. 1967. *The Bering land bridge*. Stanford Univ. Press, Stanford, CA.
- Humphries, C. J., and L. R. Parenti. 1999. *Cladistic biogeography*. Oxford Univ. Press, Oxford, U.K.
- Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 in H. N. Munro, ed. *Mammalian protein metabolism*. Academic Press, New York.
- Linder, H. P. 2001. On areas of endemism, with an example from the African Restionaceae. *Syst. Biol.* 50:892–912.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton Univ. Press, Princeton, NJ.
- Manchester, S. R. 1999. Biogeographical relationships of North American Tertiary floras. *Ann. Mo. Bot. Gard.* 86:472–522.
- Mast, A. R., and R. Nyffeler. 2003. Using a null model to recognize significant co-occurrence prior to identifying candidate areas of endemism. *Syst. Biol.* 52:271–280.
- McKenna, M. C. 1983. Cenozoic paleogeography of North Atlantic land bridges. Pp. 351–399 in M. H. P. Bott, S. Saxov, M. Talwani, and J. Thiede, eds. *Structure and development of the Greenland-Scotland ridge: new concepts and methods*. Plenum, New York.
- Moore, B. R., K. M. A. Chan, and M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pp. 483–533 in O. R. P. Bininda-Emonds ed. *Phylogenetic supertrees: combining information to reveal the Tree of Life*. Kluwer Academic, Dordrecht, The Netherlands.
- Nelson, G., and N. I. Platnick. 1981. *Systematics and biogeography: cladistics and vicariance*. Columbia Univ. Press, New York.
- Page, R. D. M. 1990. Temporal congruence and cladistic analysis of biogeography and cospeciation. *Syst. Zool.* 39:205–226.
- . 1993. COMPONENT: tree comparison software for Microsoft Windows. Ver. 2.0. Natural History Museum, London.
- . 1994. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. *Syst. Biol.* 43:58–77.
- Page, R. D. M., and M. A. Charleston. 1998. Trees within trees: phylogeny and historical associations. *Trends Ecol. Evol.* 13: 356–359.
- Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B* 255:37–45.
- . 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* 48:612–622.
- Platnick, N. I., and G. Nelson. 1978. A method of analysis for historical biogeography. *Syst. Zool.* 27:1–16.
- Rajakaruna, N., B. G. Baldwin, R. Chan, A. M. Desrochers, and B. A. Bohm. 2003. Edaphic races and phylogenetic taxa in the *Lasthenia californica* complex (Asteraceae: Heliantheae): an hypothesis of parallel evolution. *Mol. Ecol.* 12:1675–1679.
- Ricklefs, R. E. 2004. A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* 7:1–15.
- Ricklefs, R. E., and D. Schluter. 1993. Species diversity: regional and historical influences. Pp. 350–363 in R. E. Ricklefs and D. Schluter, eds. *Species diversity in ecological communities: historical and geographical perspectives*. Univ. of Chicago Press, Chicago.
- Ronquist, F. 1994. Ancestral areas and parsimony. *Syst. Biol.* 43: 267–274.
- . 1995. Ancestral areas revisited. *Syst. Biol.* 44:572–575.
- . 1996. DIVA. Ver. 1.1 Computer program and manual. Available by anonymous FTP (ftp.sysbot.uu.se) from Uppsala University, Uppsala, Sweden.
- . 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 45: 195–203.
- . 1998. Phylogenetic approaches in coevolution and biogeography. *Zool. Scripta* 26:313–322.
- . 2002. Parsimony analysis of coevolving species associations. Pp. 26–64 in R. D. M. Page, ed. *Tangled trees: phylogeny, cospeciation, and coevolution*. Univ. of Chicago Press, Chicago.
- Sanmartin, I., and F. Ronquist. 2004. Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* 53:216–243.
- Schluter, D. 1995. Uncertainty in ancient phylogenies. *Nature* 377: 108–109.
- Schluter, D., T. D. Price, A. O. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699–1711.
- Soltis, D. E., P. S. Soltis, J. C. Pires, A. Kovarik, J. A. Tate, and E. Mavrodiev. 2004. Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. *Biol. J. Linn. Soc. B* 82:485–501.
- Tiffney, B. H. 1985. The Eocene North Atlantic land bridge: its importance in tertiary and modern phytogeography of the northern hemisphere. *J. Arnold Arbor.* 66:243–273.
- . 2000. Geographic and climatic influences on the Cretaceous and Tertiary history of Euramerican floristic similarity. *Acta Univ. Carol. Geol.* 44:5–16.
- Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology, and species richness. *Trends Ecol. Evol.* 19:639–644.

Corresponding Editor: K. Crandall