

# Phylogenetic Methods in Biogeography

Fredrik Ronquist<sup>1</sup> and Isabel Sanmartín<sup>2</sup>

<sup>1</sup>Department of Biodiversity Informatics, Swedish Museum of Natural History, SE-10405 Stockholm, Sweden; email: fredrik.ronquist@nrm.se

<sup>2</sup>Department of Biodiversity and Conservation, Real Jardín Botánico, RJB-CSIC, ES-28014 Madrid, Spain; email: isanmartin@rjb.csic.es

Annu. Rev. Ecol. Evol. Syst. 2011. 42:441–64

First published online as a Review in Advance on  
August 29, 2011

The *Annual Review of Ecology, Evolution, and  
Systematics* is online at [ecolsys.annualreviews.org](http://ecolsys.annualreviews.org)

This article's doi:  
10.1146/annurev-ecolsys-102209-144710

Copyright © 2011 by Annual Reviews.  
All rights reserved

1543-592X/11/1201-0441\$20.00

## Keywords

vicariance, dispersal, diversification, speciation, extinction, parsimony, maximum likelihood, Bayesian inference, stochastic models

## Abstract

Phylogenetic approaches to biogeography are rapidly becoming more sophisticated. Four types of models are being explored in the literature: (a) diffusion models, (b) island models, (c) hierarchical vicariance models (HVMs), and (d) reticulate models. Diffusion models are used primarily for phylogeographic inference but can also offer insights into geographic discontinuities of significance on longer time scales. For the island models, statistical approaches are now well developed and typically offer more detailed insight than parsimony analysis. In contrast, parsimony may still be the best option for HVM analysis, because existing statistical approaches do not yet accommodate dispersal. The proposed probabilistic models of reticulate scenarios remain poorly understood, even though they may currently do the best job of integrating diversification processes into biogeographic analysis. Statistical approaches are gaining in popularity across the field, in part because of the flexibility of stochastic modeling. This allows investigators to address important related processes, such as ecological interactions and climate change, in biogeographic inference.

#### Area of endemism:

unit of analysis in discrete biogeographic methods, including pattern-based ones; circumscription criteria include a high degree of endemism (unique species)

## INTRODUCTION

The idea that phylogenetic relationships can provide information on past geographic distributions is as old as evolutionary biology itself; indeed, biogeography was central to Darwin's (1859) thinking about the origin of species. However, it was not until Hennig's (1966) cladistic method of reconstructing genealogical relationships was applied to biogeography (Brundin 1965, 1966) that the field of phylogenetic biogeography was born.

Since then, many methods have been proposed on the basis of the idea that taxa sharing similar phylogenetic and distributional patterns must have shared a common biogeographic history. The earliest methods were based on the parsimony principle and were aimed at finding patterns of relationships among areas of endemism (reviewed in Humphries & Parenti 1999, Morrone & Crisci 1995). Many of them were rather loosely tied to models of biogeographic processes, and their proponents emphasized the need to focus on patterns and not process during inference. This led Ronquist (1995b) to label the methods "pattern-based." Pattern-based methods establish connections between distributional patterns and evolutionary processes after the primary analytical results have been obtained (Brooks 2005, Parenti 2007). The classic pattern-based methods are Brooks parsimony analysis (Brooks 1985, Wiley 1988), component analysis (Nelson & Platnick 1981), and paralogy-free subtree analysis (Nelson & Ladiges 1996); more recent methods include three-area cladistics (Ebach et al. 2005a,b) and phylogenetic analysis for comparing trees (PACT; Wojciki & Brooks 2005).

Readers interested in a more detailed review of pattern-based approaches are directed elsewhere (Crisci et al. 2003, Humphries & Parenti 1999, Morrone 2009, Morrone & Crisci 1995, Parenti 2007). We restrict our attention entirely to event-based parsimony methods and more recent statistical approaches to biogeographic inference. Event-based parsimony methods rely on the parsimony principle, but unlike pattern-based methods they explicitly tie the costs that are minimized to events in biogeographic models (Ronquist 1995b, 2003). These models are quite similar to stochastic models; in our review we focus on the models and shift freely between event-based parsimony and statistical approaches to biogeographic inference.

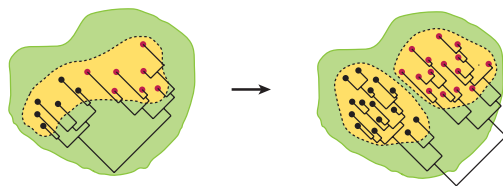
Even though the power of the parsimony approach still has not been fully explored, biogeographers have increasingly turned their attention to probabilistic techniques in recent years. This trend is due to several factors, including the explosive increase in available molecular data sets, the rapid development of maximum likelihood (ML) and Bayesian techniques, and the power of parametric statistics in addressing relevant processes (see sidebar, Expanding Biogeographic Models) and discriminating among alternative hypotheses. Because current and future development is likely to focus on parametric approaches, this is where we put our emphasis in this review.

## BIOGEOGRAPHIC MODELS

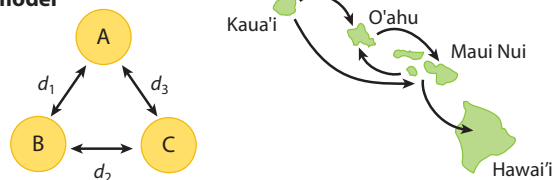
Four different classes of models have been considered in analytical biogeography (**Figure 1**). One of them represents distributions using continuous variables, and the others break distribution ranges into discrete units. In the continuous case, lineages are modeled as occupying points in a 2D or 3D landscape, and biogeographic processes correspond to movements in this landscape (diffusion models; **Figure 1a**). The geographically determined landscape can be combined with additional dimensions representing ecological or other factors. The landscape may be assumed constant or allowed to change over time.

The discrete models are based on geographic areas defined prior to analysis, and the processes correspond to events that change the biogeographic state, that is, the distribution of organisms. The discrete models ignore the movement of organisms within areas and instead focus on the

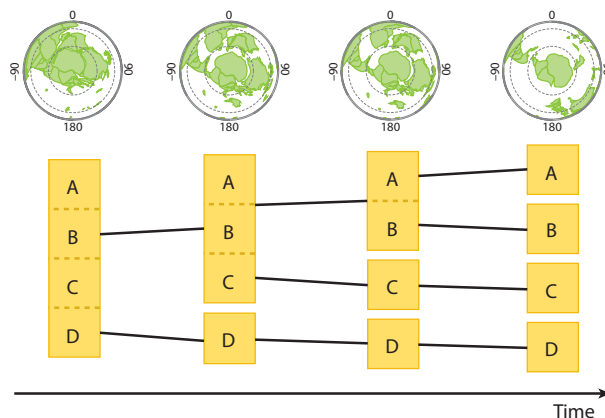
### a Diffusion model



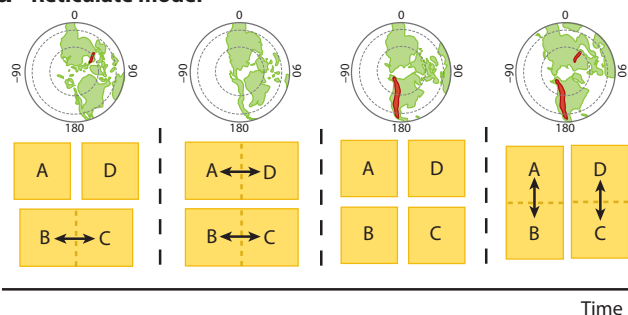
### b Island model



### c Hierarchical vicariance model (HVM)



### d Reticulate model



**Figure 1**

Model types considered in phylogenetic biogeography. Diffusion models are used primarily for phylogeographic analysis but may also be of interest in classical historical biogeography (*a*). Models that use discrete areas (A–D) fall into the categories of static island models (*b*), hierarchical vicariance models (HVMs) (*c*), and reticulate models (*d*). Static island models may allow variation over time in dispersal rates,  $d_1$  to  $d_3$ , but areas are constant. Both HVMs and reticulate models assume different area configurations over time. The empirical examples shown here are the Hawaiian archipelago (island model), Gondwana breakup (HVM), and Northern Hemisphere biogeography (reticulate model).

## EXPANDING BIOGEOGRAPHIC MODELS

Stochastic modeling allows biogeographers to integrate information from related disciplines in a rigorous statistical framework. Some interesting opportunities are listed below.

- Ecological niche modeling (ENM) uses the association between observed species occurrences and environmental variables (e.g., temperature, precipitation) to predict species ranges (Kozak et al. 2008). ENM can be used to infer distributional ranges when occurrence data are scarce (Pearson et al. 2007). Combined with paleoclimate data, ENM can project range predictions into the past (Evans et al. 2009, Graham et al. 2004, Smith & Donoghue 2010, Yesson & Culham 2006), which is clearly relevant for biogeographic reconstructions (Peterson et al. 1999, Smith & Donoghue 2010, Waltari et al. 2007, Weaver et al. 2006, Yesson & Culham 2006).
- Increasingly accurate and detailed paleogeographic maps (e.g., Hall 2009, Rosenbaum et al. 2002) are now becoming available. There also exist sophisticated online tools for paleogeography and paleoclimate reconstruction (Paleomap Project, <http://www.scotese.com>; Paleogeographic Atlas Project, <http://www.geo.arizona.edu/~rees/PGAPhome.html>), as well as for accessing and visualizing fossil (Paleobiology Database, <http://www.paleodb.org>), geological (GEON, <http://www.geongrid.org>), and lithological and sedimentary data (Paleointegration Project, <http://www.geongrid.org/index.php/projects/paleodata>).
- Climate research is producing high-resolution environmental layers that summarize temperature and precipitation (Hijmans et al. 2005; <http://www.worldclim.org>) as well as detailed paleoclimate reconstructions (Paleoclimate Model Output Visualization, <http://www.ncdc.noaa.gov/paleo/modelvis.html>; Paleoclimatic Modeling Intercomparison Project, <http://pmip.lscce.ipsl.fr>).

presumably much less frequent transitions between areas. If areas are appropriately defined, discrete models should be more powerful than continuous ones in reconstructing biogeographic patterns over long time scales. One reason for this can be illustrated by referring to the somewhat similar but more familiar problem of inferring phylogeny in the face of rate variation across sites in protein-coding sequences. If we treat all substitutions as equal (continuous biogeographic models), the frequent changes in the third codon position may swamp the rare, informative changes in the first and second positions. In particular, when the third position (movement within areas) is saturated, we do better by ignoring it completely than by treating all substitutions (all movement) as equal. For this and other reasons, historical biogeographers tend to prefer discrete models.

The discrete models come in three flavors: island models, hierarchical vicariance models (HVMs), and reticulate models. In island models, the focus is on the dispersal of organisms between areas (**Figure 1b**). The archetypal example is the analysis of movements among islands in an island group or island chain. However, island models are equally applicable to, for example, moving tectonic plates, habitat islands, isolated marine communities, or lowland regions separated by mountain ranges. Island models may allow dispersal rates to vary over time, and the islands themselves can appear (or possibly disappear) at different points in time.

An HVM describes the fragmentation of a contiguous ancient area by the successive appearance of dispersal barriers (**Figure 1c**). Each barrier splits an ancestral area into two, and it is assumed that this leads to frequent allopatric speciation, or vicariance, across the barrier. An HVM is often represented as a branching diagram of area relationships known as an area cladogram. The split-up of the Southern Hemisphere landmasses in the Mesozoic and Cenozoic is a classic example in which an HVM is considered appropriate.

In reticulate models, areas are allowed to undergo alternating cycles of fusion and splitting over time (**Figure 1d**). Splitting is associated with vicariance as in the HVM, whereas fusion leads

**Vicariance:** allopatric speciation in response to the appearance of a dispersal barrier; concordance across lineages may be implied, and peripatric patterns may be excluded

**HVM:** hierarchical vicariance model

to geodispersal, that is, range expansion across a previous dispersal barrier. Splitting and fusion can be assumed to affect all lineages in the same way, but such models are difficult to analyze. More commonly, it is assumed that lineages can assemble independently into ancestral ranges in any combination, with only the splitting and fusion probabilities (or costs) shared across lineages. The evolution of the Laurasian landmasses in the Mesozoic and Cenozoic is a good example of a system in which a reticulate model might be appropriate.

---

**Geodispersal:** range expansion in many lineages across a previous dispersal barrier that has disappeared

---

## INFERENCE METHODS

### Parsimony

Regardless of the model, biogeographic inference can be based on either parsimony or parametric statistics. Parsimony methods identify events and associate them with costs; inference is then based on finding the minimum-cost solutions. Randomization and permutation are frequently used for hypothesis testing, whereas bootstrapping and jackknifing are used to examine the robustness of estimates.

Parsimony methods used to dominate the field, but biogeographers are now rapidly shifting their focus to statistical approaches. Some reasons for abandoning parsimony include: (a) rates of processes and number of events are systematically underestimated; (b) many methods require difficult ad-hoc decisions about the relative costs of different events; (c) it is difficult to test alternative hypotheses; and (d) most methods ignore phylogenetic uncertainty and dating information, essential components in biogeographic analysis (Donoghue & Moore 2003, Sanmartín et al. 2008).

It is possible to address many of these shortcomings. For instance, phylogenetic uncertainty can be taken into account by averaging parsimony results over a tree set resulting from nonparametric bootstrapping (Micó et al. 2009, Ronquist 2003, Ronquist & Liljeblad 2001) or a Bayesian Markov chain Monte Carlo (MCMC) analysis (Huelsenbeck & Imennov 2002, Nylander et al. 2008). Nevertheless, parametric methods in statistics clearly offer a more flexible general framework for model building, parameter estimation, and hypothesis testing. Parsimony methods remain viable alternatives for some analyses mainly because of their speed and simplicity.

### Maximum Likelihood and Bayesian Inference

Statistical approaches to historical biogeography are based on stochastic models that describe biogeography in terms of rates of processes and probabilities of events. Model parameters are estimated using an appropriate data set and either ML or Bayesian inference (BI) principles (Felsenstein 2003, Holder & Lewis 2003, Ronquist & Deans 2010, Yang 2006); only a brief summary of these approaches is given here.

ML finds the parameter values with the maximum likelihood of generating the observed data, typically by using numerical optimization techniques. Confidence in ML results is often assessed using auxiliary methods such as nonparametric bootstrapping. In contrast, BI produces a probability distribution on parameter values (the posterior distribution) given the model, the observed data, and a specification of background knowledge (the prior probability distribution). The posterior distribution is almost always estimated using MCMC techniques. Credibility intervals (the BI equivalent of confidence intervals) are based entirely on the posterior distribution.

In practice, BI and ML are often quite similar, and it is possible to mix the two approaches, as in integrated likelihood or empirical Bayes methods. BI is easy to apply to complex models, although it may be tricky to obtain adequate MCMC samples (convergence) in a reasonable amount of time. ML is technically more challenging and can be quite a bit slower than BI, but clever optimization algorithms can easily push ML performance past that of BI MCMC implementations.

An advantage of BI is that it produces credibility estimates as a by-product, whereas ML typically uses repeated optimization to obtain its confidence intervals, which increases the computational complexity considerably. However, repeated optimization uses nonparametric techniques (nonparametric bootstrapping in particular) that make the confidence intervals more robust in some ways, whereas plain BI credibility intervals are more sensitive to the adequacy of the underlying model. Both BI and ML offer powerful approaches to model testing and model choice, including likelihood ratio and Bayes factor comparisons. Both approaches are currently used in biogeographic inference.

## DIFFUSION MODELS

### Stochastic Diffusion Analysis

Lemmon & Lemmon (2008) pioneered statistical biogeographic inference based on diffusion models. They modeled movement of lineages as a random walk (Brownian motion) in a 2D landscape. ML was then used to infer dispersal routes and locations of ancestors in a phylogeographic context. Lemey et al. (2010) later developed a BI MCMC approach; they found a better fit to empirical data when the dispersal rate was allowed to vary over time and across lineages in the tree rather than being constant as assumed originally.

The diffusion model is quite appropriate when the phylogeny represents a gene tree, because each lineage then traces the history of a single gene copy through the individuals that carried it. At any single point in time, therefore, a lineage corresponds to a single individual, whose distributional range is necessarily small. Continuous diffusion models are particularly relevant when the time scale is short, the focus is on populations rather than species, and sampling of individual lineages is dense. This is the domain of phylogeography and population genetics, the fields studying the processes that influence the spatial distribution and genetic variation within and among closely related taxa (Knowles 2009).

It is more questionable whether diffusion models also work well when sampling is sparse and the focus is on patterns covering larger spatial and temporal scales. Over time, local diffusion processes should become saturated, and the only events we can hope to reconstruct are those involving the rare traversal of significant dispersal barriers caused by geological, climatic, or biotic factors. In other words, continuous diffusion gradually fills in contiguous areas, which results in sharper definition of discontinuities over time (cf. **Figure 1**). Eventually, the location of lineages within contiguous areas becomes essentially random, whereas the historical signal is restricted to the pattern of movement between areas. This is an important reason why historical biogeographers tend to emphasize discrete models.

Another set of problems with diffusion models is related to the difference between gene trees and species trees. Because diffusion models deal with gene trees, it appears desirable to accommodate population genetic processes, such as mutation and selection, when using diffusion models to infer the biogeographic history of populations or species (Bloomquist et al. 2010). It might eventually be possible to simultaneously model both rapid local diffusion of individuals within populations and rare events at the population level or above, and thus increase the accuracy of inferences at both levels. Only time will tell whether the increased sampling effort and computational complexity of such studies will be worthwhile for historical biogeographers.

Clearly, biological dispersal is not a uniform spatial diffusion process; real dispersal rates are strongly affected by climate, habitat, and biotic factors. Some of these influences can be incorporated as additional landscape dimensions representing 3D spatial heterogeneity (Excoffier et al. 2009, Knowles & Alvarado-Serrano 2010) or present and past environmental



layers (Carsten & Richards 2007, Knowles 2009, Richards et al. 2007). Indeed, the ease with which such factors can be incorporated is one of the most attractive features of diffusion models.

## Identifying Areas of Endemism

Biogeographers often find it easy to identify suitable areas for discrete-model analysis, for instance, on the basis of geological features, habitat data, or the question under study. However, sometimes fuzzy boundaries and overlapping taxa make it difficult to circumscribe areas. In such cases, one can find appropriate distributional units for subsequent analysis by locating geographic discontinuities using methods based on continuous models. Stochastic diffusion analysis is one possibility (Bloomquist et al. 2010), but simpler parsimony techniques have also been proposed. For example, Szumik and colleagues (Szumik & Goloboff 2004, Szumik et al. 2002) developed an approach that minimizes the endemism score of each cell in an arbitrary grid. Another possibility is biotic element analysis (Hausdorf & Hennig 2003), which uses a grid representation and Monte Carlo simulations to test for nonrandom distributions of species ranges.

## ISLAND MODELS

### Fitch Parsimony and Related Methods

The oldest parsimony-based method in biogeography, Fitch parsimony, or Fitch optimization, treats distribution simply as an unordered character and then minimizes the number of character changes. Fitch parsimony implies an island model in which the cost of moving between distributional states is the same regardless of the states and directions involved (**Figure 2a**). A straightforward extension is to order the states into a sequence and allow dispersals only between adjacent states (Maddison et al. 1992). This is the stepping-stone model (**Figure 2b**), which is frequently discussed in the context of island chains (Funk & Wagner 1995, Sanmartín et al. 2008).

In the most general case, a cost matrix is used to represent the relative ease with which organisms can move between areas in the model (Bremer 1995; Ronquist 1994, 1995a), an approach that is extremely powerful. For instance, event costs can be made proportional to distances between areas or reflect that it is easier to move in one direction than in the other. If it is possible to set the matrix scores to the logarithm of the dispersal probabilities, then the method becomes similar to ML inference. However, cost-matrix methods are rarely used in practice because of the lack of general techniques for estimating the appropriate event costs.

Fitch parsimony has been and is still used frequently for biogeographic analysis, especially for inferring ancestral areas (e.g., Harbaugh et al. 2009, Hileman et al. 2001, Hovmöller et al. 2010, Moore et al. 2002). In contrast, there are few, if any, examples of empirical studies using the stepping-stone or cost-matrix approaches. The difficulty of finding the right sequence of states or the appropriate event costs may be one reason for this. In addition, the stepping-stone and cost-matrix models are typically more difficult to set up properly in most parsimony software packages.

### Probabilistic Dispersal Analysis

Probabilistic biogeographic models on discrete state spaces are based on continuous-time Markov chains (CTMCs), the same type of model as is used for the evolution of molecular sequences in statistical phylogenetics. CTMC models are most easily understood in terms of their instantaneous rate matrices, which describe the transition rates between states (**Figures 2a–c, bottom**).

---

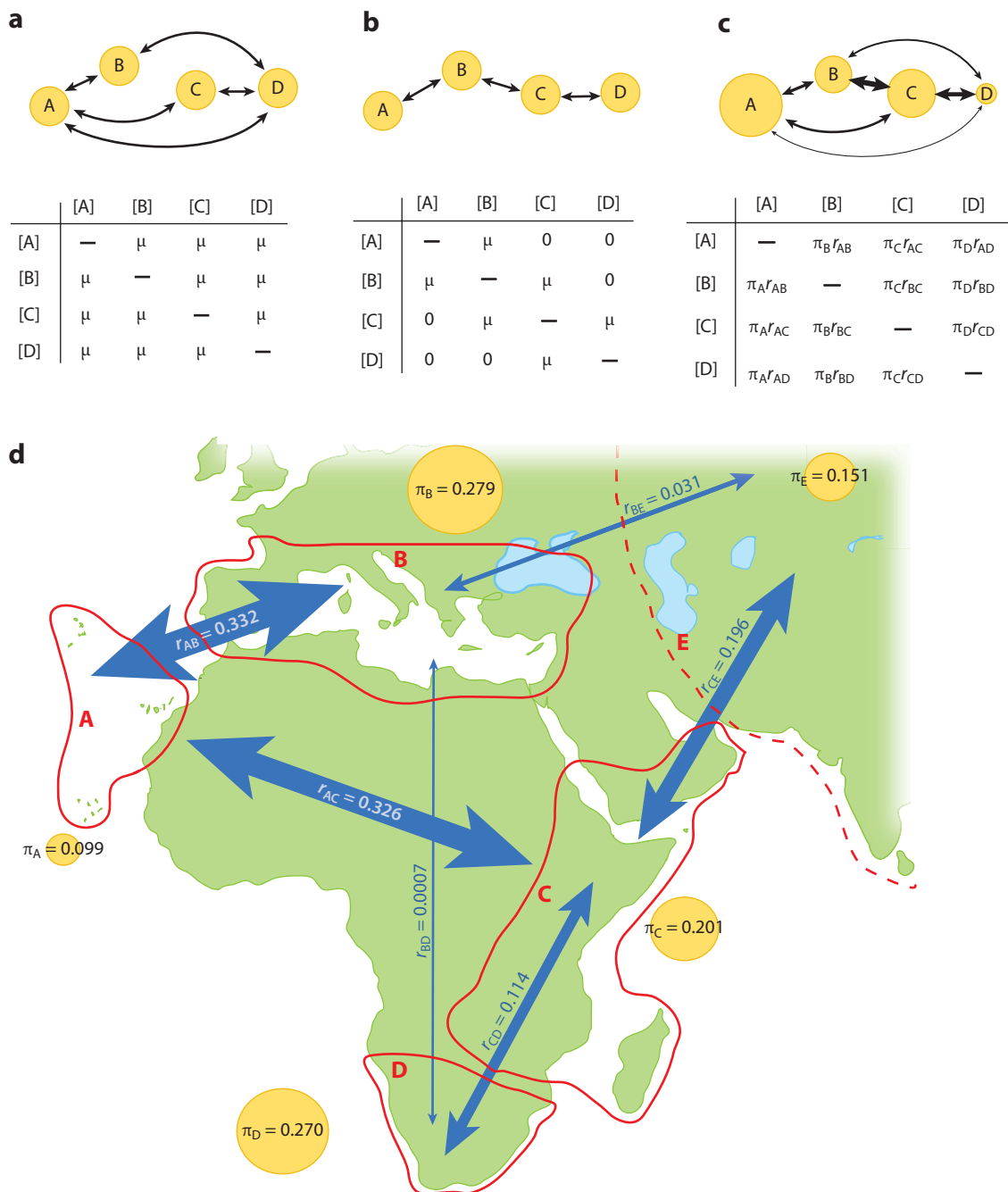
#### Continuous-time Markov chain (CTMC):

a stochastic, memory-less process used to model transitions between discrete states over continuous time

---

By exponentiating the rate matrix, it is possible to obtain the transition probabilities over a finite length of time. These probabilities, in turn, form the basic currency of ML and BI analyses.

In what appears to be the first statistical analysis in phylogenetic biogeography, Nepokroeff et al. (2003) used CTMC models and ML to infer patterns of colonization and dispersal in the Hawaiian tree genus *Psychotria* (Rubiaceae). The Hawaiian Islands were grouped into four biogeographic





states, and two CTMC models were used to infer ancestral distributions. In the first model, all transition rates were assumed to be the same (**Figure 2a**). This is analogous to the Jukes-Cantor (JC) model of nucleotide substitution, which can be understood as a probabilistic version of Fitch parsimony. The second model was a stepping-stone model in which transitions were allowed only between adjacent island groups in the instantaneous rate matrix. This is achieved by setting the nonadjacent rates to zero (**Figure 2b**), which forces all long-distance dispersal to pass through the intermediate islands in a stepping-stone fashion. When such a stepping-stone matrix is exponentiated and the time scale is short, each transition probability will largely reflect the product of the step probabilities. As time increases, the transition probabilities will tend toward equality because of the increasing frequency of multistep dispersals. Regardless of whether the JC or stepping-stone model was used for dispersals between islands, the basic rate of island colonization (the dispersal propensity) was allowed to differ from the rate of molecular substitution (Nepokroeff et al. 2003).

The past years have witnessed an increasing interest in using CTMC island models, combined with either ML or BI inference (Maddison & Maddison 2007, Pagel 1999, Pagel et al. 2004), to estimate the rate of dispersal between areas or the degree of dispersal asymmetry within single-organism groups (Havran et al. 2009, McGuire et al. 2007, Olsson et al. 2006, Renner et al. 2010, Vanderpoorten et al. 2010).

## Bayesian Island Biogeography

We appear to have been the first to propose the use of CTMC island models for statistical analysis of dispersal patterns across multiple groups (Sanmartín et al. 2008). When data from many groups are combined, statistical power will increase, and it should become possible to analyze more complex CTMC models.

When the CTMC transition rates are allowed to take on any values, the general time-irreversible (GTI) model results. Time irreversibility means that the direction of change makes a difference in calculating probabilities. However, evolutionary biologists often assume time reversibility, both for mathematical convenience and because there is an implicit assumption that evolution is a steady-state equilibrium process. In such cases, the probability is the same regardless of the direction of change. The time-reversibility constraints are captured in the general time-reversible (GTR) model by expressing the rates as the product of two factors, each of which comes from a set of values summing to 1; such a set is known as a simplex. One simplex,  $\pi$ , describes the stationary probabilities of the model states; over time, the state probabilities of the CTMC will converge onto  $\pi$  regardless of the starting point. The other simplex,  $r$ , describes the relative

**GTI:** general time-irreversible CTMC model

**GTR:** general time-reversible CTMC model

## Figure 2

Probabilistic analysis of island models; A–E indicate different areas or islands. (a) In the simplest case, all areas and dispersal rates are the same. (b) In the stepping-stone model, dispersal is allowed only between adjacent areas in a chain. (c) A general time-reversible (GTR) model allows relative carrying capacities ( $\pi$ ) and intensities of biotic exchange ( $r$ ) to differ among areas. The general time-irreversible (GTI) model is similar but also allows dispersal rates to be asymmetric (directional). Each model is represented by a diagram (*a–c, top*), in which the area of the circles is proportional to  $\pi$  and the thickness of the arrows proportional to  $r$ , and by its instantaneous rate matrix (*a–c, bottom*). (d) In a Bayesian inference analysis using the GTR model of the Rand Flora pattern, which involves disjunct floras along the margins of the African continent, clear differences are revealed between the endemic South African flora (area D; high carrying capacity, low rates of biotic exchange), and the dynamic western African flora (area A; low carrying capacity, high rates of exchange) (Sanmartín et al. 2010). The intense East-West exchange across the Sahara ( $r_{AC}$ ) suggests a previous connection between these floras. A, Macaronesia–northwestern Africa; B, Mediterranean region; C, eastern Africa–southern Arabia; D, southern Africa; E, western Asia–eastern Palearctic.

intensity of exchange between pairs of states (**Figure 2c, bottom**). In the biogeographic context,  $\pi$  corresponds to the relative carrying capacities of the areas and  $r$  to the intensity of biotic exchange between them.

A common misconception is that the GTI does not converge onto a set of stationary state probabilities. It does; we simply cannot express the rates in terms of  $\pi$  in the general case. Another common mistake is to think that rates are necessarily symmetric in a GTR model. They are not, but the rate ratio is constrained to be the same as the ratio of the stationary state probabilities. It is commonly assumed that a time-reversible CTMC is at equilibrium at the onset, such that the root states of a tree can be safely considered as drawn from  $\pi$ . In biogeography, however,  $\pi$  may not describe the initial colonization probabilities accurately. For a drastic example, consider the case in which all islands are not present or inhabitable at the time of colonization.

A CTMC dispersal model, whether time-reversible or not, can be represented graphically by a set of circles and arrows; the areas of the circles are proportional to the stationary state probabilities, and the thicknesses of the arrows are proportional to the exchangeability rates (or the transition rates in the time-irreversible case) (**Figure 2**).

To further illustrate the interpretation of the GTR model in the biogeographic context, it may be helpful to derive the number of expected migrations, in a small amount of time, from an area  $i$  to another area  $j$ . Because the vector  $\pi$  describes the proportion of the total number of lineages,  $N$ , expected in each area, the source area  $i$  will contain  $N\pi_i$  lineages. The per-lineage rate of migration from area  $i$  to area  $j$  is  $\pi_j r_{ij}$  (from the GTR rate matrix; **Figure 2c**). Finally, the number of migrants is affected by the relative dispersal propensity of the group, call it  $\mu$ . The total number of expected migrations is then equal to  $N\pi_i r_{ij} \pi_j \mu$ , that is, the total number of lineages in the source island ( $N\pi_i$ ) multiplied by the carrying capacity of the recipient islands ( $\pi_j$ ), the relative rate of biotic exchange between the two islands ( $r_{ij}$ ), and the group-specific dispersal rate ( $\mu$ ).

Simpler variants of the GTR model can be considered by introducing various constraints. One possibility is to allow unequal carrying capacities and exchangeability rates in the stepping-stone model; another is to allow unequal carrying capacities but to force exchangeability rates to be equal in the unconstrained model [analogous to the Felsenstein (1981) model for nucleotide characters].

We used BI and considered different CTMC models in analyzing dispersal patterns in the Canary Islands across a set of animal and plant phylogenies with a total of 954 terminals (Sanmartín et al. 2008). Each group was assumed to have its own characteristic dispersal rate, but other dispersal model parameters were shared across groups. The molecular data for each group were included in the analysis, along the distributional data, and MCMC sampling was performed across both phylogenetic and dispersal model parameters. Bayes factor analysis showed that the data favored the unconstrained GTR model over stepping-stone models and simpler unconstrained models. Biotic exchange occurred primarily between adjacent islands, whereas mainland-island exchange was rare and less dependent on geographic distance. The data were apparently not informative enough to allow effective inference of the parameters of the full GTI model.

We recently used the same approach in a continental setting to analyze the African Rand Flora, a phytogeographic pattern that involves related but disjunct floras along the margins of Africa and adjacent areas (Macaronesia, Horn of Africa–Southern Arabia, Eastern Africa, and South Africa) (**Figure 2d**). These subtropical floras are isolated from each other by ecological and climatic barriers such as the tropical lowlands of central Africa and the Saharan desert. The analysis differentiates the highly endemic South African flora (high carrying capacity, low biotic exchange) from the dynamic Macaronesian flora (low carrying capacity, high biotic exchange) and identifies intense past biotic exchange across the Sahara (Sanmartín et al. 2010).

Lemey et al. (2009) extended the approach in several important ways, perhaps most significantly by applying Bayesian stochastic variable selection, a technique for estimating which entries in the

CTMC transition matrix need to be nonzero to explain the data adequately. They also provide powerful tools for visualizing inferred scenarios on digital maps and demonstrate the utility of this approach in reconstructing the spread of viral pandemics.

Power is an important concern in statistical dispersal analysis. For example, consider a problem with ten trees, each with 100 tips and 10-kb sequence data. Inferring the parameters of a GTR rate matrix for the molecular data of each one of these trees can draw from a million observations of nucleotides (100 tips  $\times$  10,000 sites). Inference of dispersal model parameters, in contrast, relies on a single character (the distribution) scored across 10  $\times$  100 tips, that is, 1,000 observations. With only 0.1% of the data, we cannot expect to estimate dispersal model parameters with the same precision as substitution model parameters.

Statistical analysis of island models could be extended in several interesting ways. For instance, carrying capacities or exchangeability rates might be allowed to vary over time. Reticulate scenarios could be approximated by having the intensity of biotic exchange reflect the connectivity between areas (Ree & Sanmartín 2009). The change points could be fixed prior to analysis, or the number and position of change points could be estimated as part of the analysis (Christin et al. 2008, Green 1995). Another fascinating possibility is the incorporation of explicit birth-death models of diversification with island-specific rates of speciation and extinction.

#### Duplication:

diversification within a single area

**Sorting:** the absence of a lineage from part of the predicted range after a vicariance event

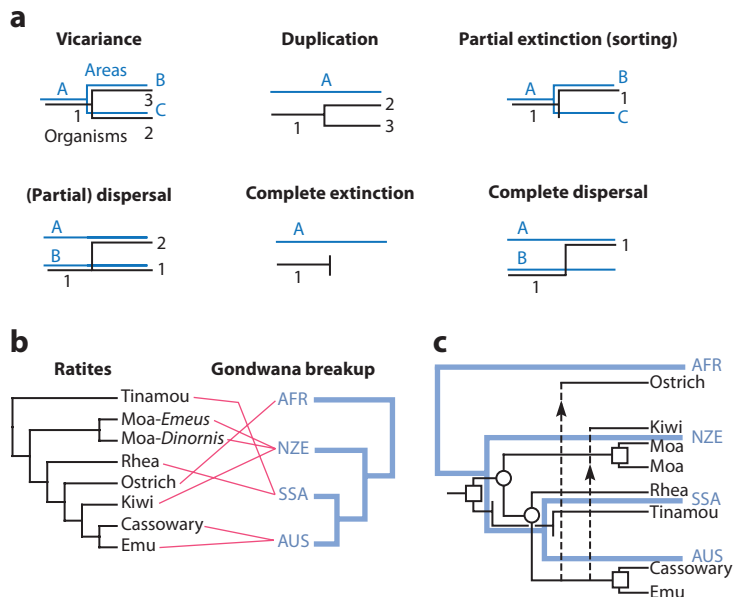
## HIERARCHICAL VICARIANCE MODELS

### Parsimony-Based Tree Fitting

Not long ago, the HVM seemed to be the only relevant model in phylogenetic biogeography. The explanation for this goes back to the origin of the field in the 1960s, when vague dispersal explanations dominated the scene. The emerging generation of phylogeneticists contrasted vicariance events, which left traces in phylogenies and therefore could be studied scientifically, with random dispersal (Humphries & Parenti 1999, Nelson & Platnick 1981). Later, it was realized that other models were also of interest in phylogenetic biogeography (Lieberman 2000, 2004; Sanmartín & Ronquist 2004), but the HVM remains the model of choice for many empirical problems.

The problem of fitting organism phylogenies to area cladograms is analogous to problems in identifying host-parasite cospeciation and in folding gene trees into species trees (Page & Charleston 1998). The broad applicability has stimulated methodological development. Six types of events have been considered in the proposed methods (**Figure 3a**): (a) vicariance, by which an organism lineage speciates in response to an area split; (b) duplication, which corresponds to diversification within an area; (c) partial extinction (or sorting), which results in the presence of an organism lineage in only one descendant area; (d) (partial) dispersal, which results in an organism lineage establishing a sister lineage in an unrelated area; (e) complete extinction, by which an organism lineage completely disappears from an area; and (f) complete dispersal, by which an organism lineage switches from one area to another. Simple parsimony methods cannot address the last two because of the limitations of cost optimization (Ronquist 2003), so they necessarily focus on the first four.

To assign costs to each of these events in a way that facilitates biogeographic analysis is not a simple task. If all events are assigned the same cost, then vicariance events never appear in optimal reconstructions (Ronquist 2003). A popular solution is to maximize vicariance, which is the same as assigning vicariance events a negative cost and all other events zero cost (Page 1995, Page & Charleston 1998). Unfortunately, maximizing vicariance may result in the postulation of an inordinate number of other events. For instance, with enough extinctions (at zero cost), every



**Figure 3**

Parsimony analysis of the hierarchical vicariance model (HVM). (*a*) Six different events could be considered in analyzing the HVM, but parsimony can retrieve only the first four (vicariance, duplication, partial extinction, and partial dispersal). (*b–c*) An example parsimony analysis of the HVM involving ratites and Southern Hemisphere biogeography (Pereira & Baker 2006). The organism tree, distributions, and HVM are displayed in a tanglegram (*b*). One possible reconstruction (*c*) is shown using standard symbols for parsimony-based tree fitting. It involves three duplications (*squares*), two vicariance events (*circles*), two sorting events (*short transverse lines* in the Tinamou lineage), and two dispersals (*arrows and dashed lines*). Abbreviations: AFR, Africa; AUS, Australia; NZE, New Zealand; SSA, Southern South America.

single dispersal can be explained as a vicariance event. A particularly important consequence of this is that a moderate number of duplications can make it impossible to detect significant fit to the HVM (Ronquist 2003).

Generalized parsimony (**Figure 3*b,c***), in which all events are assigned a positive cost, is more powerful (Ronquist 2003) than maximum vicariance, but one is left with the problem of estimating the exact costs to be used. Perhaps the best solution proposed thus far is to use random permutation tests, in which one randomly permutes the organism distributions in the phylogeny and compares the cost of fitting the permuted data sets to the HVM with the cost of fitting the original data. The optimal event costs are those that maximize the difference between the permuted and the original data (Ronquist 2003); that is, they maximize the possibility of detecting phylogenetically conserved distribution patterns. The optimal schemes typically assign low costs to duplication and vicariance and high costs to dispersal, relative to the cost of sorting events (Ronquist 2003). This makes sense in that both vicariance and duplication involve inheritance of distributional ranges, whereas dispersal and sorting events partly erase previous distributional history. Unfortunately, the method is computationally complex and can only suggest a range of optimal event costs, not pinpoint the best values.

Similar to island models, the HVM assumes that ancestors were restricted to single areas. However, these single distributional areas include not only the current areas but also the contiguous units they formed in the past according to the HVM. Distributional ranges comprising other

area combinations are not considered, and widespread extant taxa are therefore not necessarily compatible with the HVM. Sanmartín & Ronquist (2002) discuss this problem at length and suggest three alternative strategies for reconciling widespread taxa with the HVM.

There are few empirical examples of patterns that fit the HVM perfectly. Generalized parsimony methods often find a dominant pattern of duplication (within-area diversification) and few general vicariance events (Ahn et al. 2010, Bremer & Janssen 2006, Hosaka et al. 2008, Huttunen et al. 2008, Pirie et al. 2010). The breakup of southern Gondwana in the Mesozoic is perhaps the most famous HVM, but in a large analysis we showed that only animal distributions actually fit an HVM (Sanmartín & Ronquist 2004). Plant patterns, in particular the close relationship between the New Zealand and Australian flora, are better explained by concordant dispersal following the West Wind Drift, which moves waters and winds from east to west around Antarctica. These findings stirred a lively debate that involved discussions on the roles of prevailing winds and ocean currents in causing these patterns (Cook & Crisp 2005, De Queiroz 2005, Sanmartín et al. 2007, Vanderpoorten et al. 2010, Waters 2008), on the origins of New Zealand's present biota and the accuracy of current geological reconstructions (Giribet & Boyer 2010, Jolivet & Verma 2010, Wallis & Trawick 2009, Waters & Craw 2006), and on the appropriateness of parsimony-based tree fitting for inferring asymmetric dispersal patterns (Cook & Crisp 2005, Sanmartín et al. 2007).

Most parsimony-based analyses of the HVM ignore information about the timing of splits, even though such information is crucial in testing vicariance hypotheses. In recent years, it has become standard practice to complement HVM analyses with separate analyses of divergence times. In principle, timing information could also be incorporated into the parsimony analysis itself, but apparently this methodology remains to be developed in detail.

## Parametric Tree Fitting

The HVM still has not been treated statistically, but recent progress in addressing related models is promising. The model that is closest to the HVM is probably the one used in analysis of the evolution of gene families. As genomes evolve, individual genes are duplicated or lost through events that affect small or large regions of the genome, and in extreme cases the entire genome. Occasionally, foreign genes also become incorporated through the process of horizontal gene transfer (HGT).

To trace the evolution of a gene family (a set of related genes in the genome) inside a species tree (whole-genome tree), we need to consider the processes of duplication, loss, and HGT, which affect individual genes, and the processes of speciation and extinction, which affect the lineages carrying the genomes, that is, all genes. In the first statistical approach to this problem, Arvestad et al. (2003) used a birth-death model for gene duplication (birth) and gene loss (death), whereas speciation was assumed to result in identical copies of the ancestral genome. To simplify the problem, they did not consider HGT or sequence evolution. In subsequent papers (Åkerborg et al. 2009, Arvestad et al. 2004), many important computational advances were made, and the model was expanded to include sequence evolution and relaxation of the clock. Unfortunately, it still has not been possible to accommodate HGT.

The analogy to the HVM is straightforward: duplication corresponds to diversification within areas, loss to extinction, HGT to dispersal, and splits in the species tree to hierarchical vicariance events. The birth-death model is an appropriate choice for modeling diversification and extinction in the HVM. However, few investigators would be willing to analyze the HVM without considering dispersal, so incorporation of HGT is a prerequisite for gene family methods to be applicable to realistic biogeography problems. Nevertheless, phylogenetic biogeographers clearly should follow this field closely.

#### DIVA:

dispersal-vicariance  
analysis

#### Cladogenetic event:

an event occurring at a  
node in a phylogenetic  
tree

#### Anagenetic event:

an event occurring along  
a branch of a  
phylogenetic tree

Furthermore, there has been recent progress on the problem of fitting trees describing the evolution of gene copies, which we refer to as allele trees, into species trees. Both gene family trees and allele trees are referred to as gene trees in the literature, but they are quite different. Alleles occupy the same locus in the genome but on different copies of the chromosome, whereas gene family members occupy different loci on the same chromosome. In organisms with sexual reproduction, allele trees tend to differ from species trees because of lineage sorting, introgression, and hybridization. Incomplete lineage sorting is similar to partial extinction in the HVM, the origin of new alleles to duplication, and introgression and hybridization to dispersal. Modeling the evolution of alleles in populations is almost always based on the coalescence model, which is an approximation of basic population genetics models. Folding allele trees into species trees relies on extending the coalescent to multispecies contexts (Edwards et al. 2007, Heled & Drummond 2010). However, the coalescent is probably a poor fit to the diversification processes considered in biogeographic analysis, which makes allele tree models less interesting in phylogenetic biogeography. Introgression and hybridization also have not been included in these models, further restricting their utility in biogeographic analysis.

In addition, one might have expected recent progress on the statistical analysis of host-parasite cospeciation, yet another field using models analogous to the HVM. However, little appears to have happened after the pioneering paper by Huelsenbeck et al. (2000). The model considered in that paper does accommodate dispersal (host switches), but it restricts ancestral hosts to being inhabited by a single parasite lineage at a time. The analogy would be to restrict each area in a biogeographic analysis to occupation by a single representative of each studied organism group; it is difficult to come up with biogeography problems in which such an assumption would be defensible.

## RETICULATE MODELS

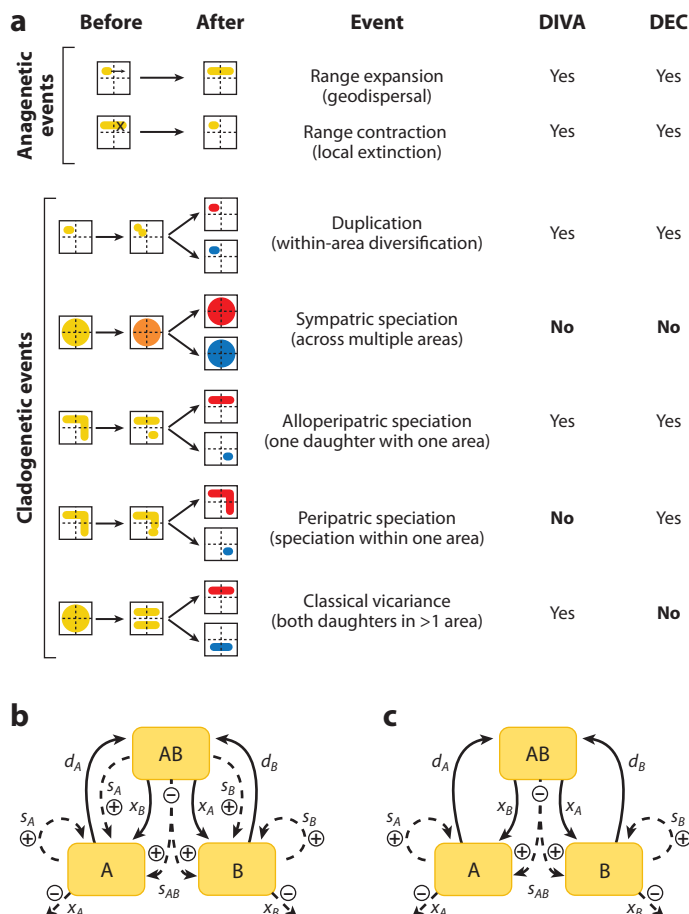
### Dispersal-Vicariance Analysis (DIVA)

In HVM and island models, each lineage is assumed to occupy a single area at any one point in time, but it is often plausible that some lineages are more widespread. Speciation in such widespread lineages presumably frequently results in subdivision of the original range through vicariance (allopatric speciation). These vicariance events could be shared across organism groups, but they could also be peculiar to individual lineages. For instance, the appearance of a desert might effectively isolate flightless forms but not fully winged ones of the same group. If one does not know how areas were connected in the past, it might be reasonable to allow all combinations of areas as potential ancestral distributions and all possible splits between areas as potential vicariance events.

Dispersal-vicariance analysis (DIVA) is a parsimony method that takes this approach (Figure 4a). Distributions are modeled in terms of unit areas; a lineage can occupy one or more of these in any combination. Speciation occurs either by vicariance, that is, the partitioning of an ancestral set of areas into two disjoint subsets, or by diversification within a single area, as in the island model. These cladogenetic (node) events are both assigned zero cost. On top of this, lineages can expand their range by adding unit areas to their distribution (cost one per unit area added), or their range can shrink through local extinction (cost one per unit area deleted); these are anagenetic (branch) events.

Despite their simplicity, DIVA reconstructions often approach intuitive analyses that consider multiple lines of evidence. Therefore, DIVA has become a popular technique for reconstructing the biogeographic history of organisms when it is not feasible or desirable to take evidence from other organisms or geological history into account (Barber & Bellwood 2005, Prieto-Marquez 2010) or when the geological history does not conform to the HVM





**Figure 4**

(a) Both dispersal-vicariance analysis (DIVA) (Ronquist 1997) and dispersal-extinction-cladogenesis (DEC) (Ree et al. 2005) are reticulate models that include anagenetic and cladogenetic events. The anagenetic events in both models are the same, but the models differ slightly in the cladogenetic events they allow. (b) The geological state speciation and extinction (GeoSSE) model combines DEC with an explicit model for speciation and extinction, thus providing a complete probabilistic description of DEC. The state transition graph for a GeoSSE model on two areas, A and B, describes the rates of speciation in A and B ( $s_A$  and  $s_B$ ), the rates of extinction in A and B ( $x_A$  and  $x_B$ ), the rate of between-area vicariance ( $s_{AB}$ ), and the rates of dispersal from A to B ( $d_A$ ) and from B to A ( $d_B$ ). Some transitions involve addition or subtraction of lineages (dashed arrows) and not regular state transitions, which makes it impossible to treat GeoSSE using the standard computational machinery for finite-state Markov chains. (c) A probabilistic version of DIVA for two areas is similar, except that transitions corresponding to peripatric speciation are absent.

(Mansion et al. 2008, Oberprieler 2005, Sanmartín 2003). Occasionally, the DIVA model is confused with the HVM (Biswas & Pawar 2006, Morrone 2009, Ree et al. 2005), which is understandable because the HVM is subsumed as a special case under DIVA. However, dispersal and extinction are modeled differently in DIVA and in the HVM (range expansion and contraction versus speciation-coupled dispersal and sorting; **Figures 2 and 4**).

A consequence of the simplicity of DIVA is a lack of power. For instance, extinction events never appear in optimal DIVA reconstructions (Ronquist 1996). Furthermore, DIVA tends to



## DEC:

dispersal-extinction-cladogenesis model

## Alloperipatric

**speciation:** in the DEC model, allopatric speciation that splits off one area from a wider ancestral range

## Peripatric speciation:

in the DEC model, speciation within one area of a wide ancestral range, as if by a temporary peripheral isolate

favor vicariance over dispersal when data are scarce, which results in association of deep nodes with wide distributions, whereas dispersals are biased toward terminal branches (Ree et al. 2005, Sanmartín 2007, Nylander et al. 2008, Kodandaramaiah 2010, Buerki et al. 2011). A possible solution is to constrain ancestral areas to contain a certain maximum number of unit areas. For example, one might be willing to assume that the maximum size of ancestral ranges was not larger than that of current ones (Nordlander et al. 1996, Sanmartín 2003).

The DIVA model itself does not force implied splits or fusions of areas to be shared across lineages. Nevertheless, DIVA reconstructions can be used to detect general biogeographic patterns by summarizing dispersal and vicariance frequencies across multiple groups (Donoghue & Smith 2004, Sanmartín et al. 2001). The significance of inferred patterns can be examined by randomly permuting distributions among tips in the tree and comparing observed frequencies of events with those expected by chance. By separating events into time bins, it is also possible to study how biogeographic processes vary over time (Sanmartín et al. 2001). For example, a decrease in dispersal frequency between two areas would suggest the formation of a new dispersal barrier, whereas an increase would indicate the mixing of biotas owing to area fusion. Other extensions of DIVA are conceivable but might be explored more productively within a parametric statistical context.

## The Dispersal-Extinction-Cladogenesis Model

A statistical approach to reticulate models was first proposed in a groundbreaking paper from Ree et al. (2005). Ree & Smith (2008) later developed the mathematical and computational foundations of ML inference under the method. The method is based on the dispersal-extinction-cladogenesis (DEC) model, which is similar to the DIVA model (**Figure 4**).

As in DIVA, the anagenetic events in DEC are range expansion and contraction, which are modeled using a conventional CTMC. In the simplest case, the rates of dispersal and local extinction are the same for all areas and time intervals, but both spatial and temporal variation in the rates may be accommodated (Buerki et al. 2011, Ree & Smith 2008, Ree et al. 2005, Salvo et al. 2010).

With respect to cladogenetic events, both methods allow diversification within a single area and disallow simultaneous speciation across multiple areas (**Figure 4a**). When the ancestral range includes two or more areas, however, the methods differ. DEC requires that between-area vicariance events separate a single area from the remainder of the ancestral range (alloperipatric speciation), whereas DIVA also permits classical vicariance events in which each daughter occupies more than one area. Furthermore, unlike DIVA, DEC allows speciation within one area of a wide ancestral range, which results in one descendant occupying the area of speciation whereas the other inherits the entire range (peripatric speciation).

Ree et al. (2005, p. 2302) say that allowing vicariance with multiple areas on each side, as in DIVA, “invokes a historical event without considering the spatial and temporal context, and we wish to avoid making this kind of ad hoc hypothesis.” However, spatial and temporal contexts can be considered—or ignored—within both DIVA and DEC. The question that begs an answer is how stochastic DIVA and DEC models (and other reticulate models) would stack up against one another. Which model describes empirical data best? Comparing parsimony results from DIVA with ML results from DEC (Buerki et al. 2011, Clark et al. 2008) can produce only limited insights into this question.

The DEC model can be extended by incorporating fossil and geological information into the rate matrix, which is allowed to vary over time (Ree & Smith 2008, Ree et al. 2005). Dating uncertainty can be accommodated by integrating DEC reconstructions over a BI posterior sample of dated trees (Smedmark et al. 2010, Smith 2009). Recent studies show that dating uncertainty can

have a profound influence if a stratified model is used to pose temporal constraints on area connectivity and dispersal routes (Buerki et al. 2011, Smith 2009). Complex stratified paleogeographic models also can cause various other types of problems in DEC analyses (Buerki et al. 2011).

Goldberg et al. (2011) combine DEC with an explicit model of speciation and extinction of lineages. The result is GeoSSE (geological state speciation and extinction), a model inspired by the binary state speciation extinction model (FitzJohn et al. 2009, Maddison et al. 2007). Lineage extinction is not considered a unique type of event in GeoSSE; it is simply the consequence of local extinction in a species restricted to a single area. In a model of two areas, A and B, speciation is treated using three different rates,  $s_A$ ,  $s_B$ , and  $s_{AB}$ ; the first two correspond to speciation within areas and the third to vicariance between areas. Other parts of the model correspond to DEC (Figure 4b).

GeoSSE cannot be tackled using standard computational machinery because of the intimate connection between state transitions on one hand and speciation-extinction on the other; slower numerical methods must be used instead. Nevertheless, ML or BI results can be obtained given that the state space is sufficiently small and the phylogenies of moderate size. Goldberg et al. (2011) apply GeoSSE to the evolution of habitat occupancy in Californian plant communities, contrasting chaparral and forest. They find higher rates of speciation in chaparral and evidence for expanding habitat tolerances, nicely demonstrating both the utility of GeoSSE and that biogeographic areas need not be circumscribed solely on the basis of geographic coordinates or geological features.

The DEC and GeoSSE models are arguably the most realistic of the stochastic models considered in historical biogeography thus far, but their fundamental properties are still poorly understood. For instance, it is impossible to derive the equilibrium state of the DEC model because it is incomplete: it does not explicitly model the rate of speciation even though speciation affects distributions (Figure 4a). The GeoSSE model provides a complete probabilistic description of the biogeographic process, but it is rather complicated to derive its equilibrium state, and this has not yet been attempted, not even for the two-area case. This causes computational problems, which have been addressed thus far using unorthodox ad hoc procedures (Goldberg et al. 2011; see also FitzJohn et al. 2009).

Because of the complexity of the DEC and GeoSSE models, statistical power is a concern. Simulations performed by Ree & Smith (2008) indicate that DEC analyses estimate extinction rates as close to zero regardless of their true value and that estimated dispersal rates are biased downward. The GeoSSE model fared better in similar studies (Goldberg et al. 2011), but part of the reason may be the restricted state space considered in those analyses (two areas).

The state space of DEC and GeoSSE (as well as DIVA) is cumbersome, which has been discussed repeatedly in the literature. For  $n$  unit areas, these models consider  $2^n - 1$  states (all possible area combinations except complete absence), a number that grows quite quickly with increasing  $n$ . In practice, it is difficult to perform ML inference under DEC with more than seven or eight areas (corresponding to approximately 250 states) unless one restricts attention to only the most relevant area combinations (Salvo et al. 2010). Even with such constraints, there will be an inordinate number of parameters if each area is to have its own dispersal and extinction rates.

**GeoSSE:** geological state speciation extinction model

## SUMMARY POINTS

1. Four classes of models have been considered in phylogenetic biogeography: (a) diffusion models, (b) island models, (c) HVM models, and (d) reticulate models.

2. Both parsimony and statistical methods (BI and ML) can be used for inference under these models. The full power of parsimony methods has not yet been explored; nevertheless, statistical approaches are rapidly gaining in popularity. Advantages include flexibility in expanding biogeographic models to incorporate related processes and the availability of well-developed computational machinery for BI or ML inference under many types of models.
3. The diffusion model is primarily suited for phylogeographic analysis but can also offer insights into geographic discontinuities of significance on longer time scales.
4. Statistical approaches (ML and BI) to the island model are well developed and typically offer more detailed insights than parsimony analysis for both organism-focused (taxon biogeography) and process-focused (area biogeography) studies.
5. Parsimony may be the best option for HVM analysis, because existing statistical approaches do not accommodate dispersal in a realistic fashion.
6. Statistical approaches to reticulate models are still in their infancy. Only a couple of models have been explored, and they are still poorly understood. Even so, these models probably do the best job currently of integrating diversification processes into biogeographic analysis.

## FUTURE ISSUES

1. Statistical analysis of island models should be extended to allow dispersal rates and carrying capacities to change over time.
2. There is a need to develop statistical approaches to the HVM that incorporate realistic models of dispersal.
3. The equilibrium dynamics and statistical power of the DEC and GeoSSE models should be studied in more detail.
4. The reticulate model space needs to be better explored. Does DIVA, DEC, or another variant offer the best fit to empirical data?
5. Across the board, there is a lack of understanding of the power of different statistical approaches to biogeographic analysis. Both theoreticians and empiricists need to address this.
6. The flexibility of stochastic modeling should be used to further expand biogeographic models by incorporating potentially relevant processes. Examples include speciation, extinction, competition, population-genetic processes, geological events, and climate change.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We would like to thank Seraina Klopstein, Gonzalo Nieto, and Javier Fuertes for valuable comments on the manuscript. Our apologies to all colleagues whose work we did not mention because of our ignorance or because of space constraints and our subjective and possibly misinformed choice of topics to discuss. This work was supported by the Swedish Research Council (VR 2008-5629, to F.R.), the Consejo Superior de Investigaciones Científicas (CSIC visiting fellowship PA2002030, to F.R.), the Spanish Ministry of Education and Science (CGL2009-13322-C03-01, to I.S.), and the National Evolutionary Synthesis Center (NESCent) (US NSF #EF 0423641-) Biogeography Working Group (to F.R. and I.S.).

## LITERATURE CITED

- Ahn KJ, Jeon MJ, Branham MA. 2010. Phylogeny, biogeography and the stepwise evolutionary colonization of intertidal habitat in the Liparocephalini based on morphological and molecular characters (Coleoptera: Staphylinidae: Aleocharinae). *Cladistics* 26:344–58
- Akerborg O, Sennblad B, Arvestad L, Lagergren J. 2009. Simultaneous Bayesian gene tree reconstruction and reconciliation analysis. *Proc. Natl. Acad. Sci. USA* 106:5714–19
- Arvestad L, Berglund AC, Lagergren J, Sennblad B. 2003. Bayesian gene/species tree reconciliation and orthology analysis using MCMC. *Bioinformatics* 19(Suppl. 1):i7–15
- Arvestad L, Berglund AC, Lagergren J, Sennblad B. 2004. Gene tree reconstruction and orthology analysis based on an integrated model for duplications and sequence evolution. *Proc. Annu. Int. Conf. Res. Comput. Mol. Biol., 8th, San Diego*, March 27–31, pp. 326–35. New York: ACM
- Barber PH, Bellwood DR. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and new world tropics. *Mol. Phylogenetics Evol.* 35:235–53
- Biswas S, Pawar SS. 2006. Phylogenetic tests of distribution patterns in South Asia: towards an integrative approach. *J. Biosci.* 31:95–113
- Bloomquist EW, Lemey P, Suchard MA. 2010. Three roads diverged? Routes to phylogeographic inference. *Trends Ecol. Evol.* 25(11):626–32
- Bremer K. 1995. Ancestral areas: Optimization and probability. *Syst. Biol.* 44:255–59
- Bremer K, Janssen T. 2006. Gondwanan origin of major monocot groups inferred from dispersal-vicariance analysis. *Aliso* 22:22–27
- Brooks DR. 1985. Historical ecology: a new approach to studying the evolution of ecological associations. *Ann. Mo. Bot. Garden* 72:660–80
- Brooks DR. 2005. Historical biogeography in the age of complexity: expansion and integration. *Rev. Mex. Biodivers.* 76:79–94
- Brundin L. 1965. On the real nature of transantarctic relationships. *Evolution* 19:496–505
- Brundin L. 1966. Transantarctic relationships and their significance, as evidenced by chironomid midges: with a monograph of the subfamilies Podonominae and Aphroteniinae and the Austral Heptagyiæ. *Kungl. Svens. Vetenskapakad. Handl.* 11:1–472
- Buerki S, Forest F, Álvarez N, Nylander JAA, Arrigo N, Sanmartín I. 2011. An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *J. Biogeogr.* 38:531–50
- Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* 61:1439–54
- Christin P-A, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, et al. 2008. Oligocene CO<sub>2</sub> decline promoted C<sub>4</sub> photosynthesis in grasses. *Curr. Biol.* 18:37–43
- Clark JR, Ree RH, Alfaro ME, King MG, Wagner WL, Roalson EH. 2008. A comparative study in ancestral range reconstruction methods: retracing the uncertain histories of insular lineages. *Syst. Biol.* 57:693–707

- Cook LG, Crisp MD. 2005. Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *J. Biogeogr.* 32:741–54
- Crisci JV, Katinas L, Posadas P. 2003. *Historical Biogeography: An Introduction*. Cambridge, MA: Harvard Univ. Press. 264 pp.
- Darwin C. 1859. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray
- de Queiroz A. 2005. The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* 20:68–73
- Donoghue MJ, Moore BR. 2003. Toward an integrative historical biogeography. *Integr. Comp. Biol.* 43:261–70
- Donoghue MJ, Smith SA. 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philos. Trans. R. Soc. Lond. B* 359:1633–44
- Ebach MC, Humphries CJ, Newman RA, Williams DM, Walsh SA. 2005a. Assumption 2: opaque to intuition? *J. Biogeogr.* 32:781–87
- Ebach MC, Newman RA, Humphries CJ, Williams DM. 2005b. 3item version 2.0: Three-Item Analysis for Cladistics and Area Cladistics. Software package.
- Ebach MC, Tangney RS, eds. 2007. *Biogeography in a Changing World*. Boca Raton, FL: CRC Press. 212 pp.
- Edwards SV, Liu L, Pearl DK. 2007. High-resolution species trees without concatenation. *Proc. Natl. Acad. Sci. USA* 104:5936–41
- Evans MEK, Smith SA, Flynn RS, Donoghue MJ. 2009. Climate, niche evolution, and diversification of the “bird-cage” evening primroses (*Oenothera*, sections *Anogra* and *Kleinia*). *Am. Nat.* 173:225–40
- Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansion. *Annu. Rev. Ecol. Evol. Syst.* 40:481–501
- Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17:368–76
- Felsenstein J. 2003. *Inferring Phylogenies*. Sunderland, MA: Sinauer
- FitzJohn RG, Maddison WP, Otto SP. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* 58:595–611
- Funk VA, Wagner WL. 1995. Biogeographical patterns in the Hawaiian islands. In *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*, ed. WL Wagner, VA Funk, 17:379–419. Washington, DC/London: Smithsonian. Inst. Press. 467 pp.
- Giribet G, Boyer SL. 2010. “Moa’s Ark” or “Goodbye Gondwana”: is the origin of New Zealand’s terrestrial invertebrate fauna ancient, recent, or both? *Invertebr. Syst.* 24(1):1–8
- Goldberg EE, Lancaster LT, Ree RH. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst. Biol.* 60:451–65
- Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58(8):1781–93
- Green PJ. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82:711–32
- Hall R. 2009. Southeast Asia’s changing paleogeography. *Blumea* 54:148–61
- Harbaugh DT, Wagner WL, Allan GJ, Zimmer EA. 2009. The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *J. Biogeogr.* 36:230–41
- Hausdorf B, Hennig C. 2003. Biotic element analysis in biogeography. *Syst. Biol.* 52:717–23
- Havran JC, Sytsma KJ, Harvey EB Jr. 2009. Evolutionary relationships, interisland biogeography, and molecular evolution in the Hawaiian violets (*Viola*: Violaceae). *Am. J. Bot.* 96:2087–99
- Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27:570–80
- Hennig W. 1966. *Phylogenetic Systematics*. Urbana: Univ. Ill. Press
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965–78
- Hileman LC, Vasey MC, Parker VT. 2001. Phylogeny and biogeography of the Arbutioideae (Ericaceae): implications for the Madrean-Tethyan hypothesis. *Syst. Bot.* 26:131–43
- Holder M, Lewis PO. 2003. Phylogeny estimation: traditional and Bayesian approaches. *Nat. Rev. Genet.* 4:275–84

---

Presents the GeoSSE model, the first complete probabilistic description of DEC.

---

- Hosaka K, Castellano MA, Spatafora JW. 2008. Biogeography of Hysterangiales (Phallomycetidae, Basidiomycota). *Mycol. Res.* 112:448–62
- Hovmöller R, Alexandrov B, Hardman J. 2010. Tracking the geographical spread of avian influenza (H5N1) with multiple phylogenetic trees. *Cladistics* 26:1–13
- Huelsenbeck JP, Imennov NS. 2002. Geographic origin of human mitochondrial DNA: accommodating phylogenetic uncertainty and model comparison. *Syst. Biol.* 51:155–65
- Huelsenbeck JP, Rannala B, Larget B. 2000. A Bayesian framework for the analysis of cospeciation. *Evolution* 54:352–64
- Humphries CJ, Parenti L. 1999. *Cladistic Biogeography: Interpreting Patterns of Plant and Animal Distribution*. New York: Oxford Univ. Press
- Huttunen S, Hedenäs L, Ignatov MS, Devos N, Vanderpoorten A. 2008. Origin and evolution of the northern hemisphere disjunction in the moss genus *Homalothecium* (Brachytheciaceae). *Am. J. Bot.* 95:720–30
- Jolivet P, Verma KK. 2010. Good morning Gondwana. *Ann. Soc. Entomol. Fr.* 46(1–2):53–61
- Knowles LL. 2009. Statistical phylogeography. *Annu. Rev. Ecol. Syst.* 40:593–612
- Knowles LL, Alvarado-Serrano DF. 2010. Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Mol. Ecol.* 19:3727–45
- Kodandaramaiah U. 2010. Use of dispersal-vicariance analysis in biogeography—a critique. *J. Biogeogr.* 37:3–11
- Kozak KH, Graham CH, Wiens JJ. 2008. Integrating GIS-based environmental data into evolutionary biology. *Trends Ecol. Evol.* 23:141–48
- Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5(9):e1000520
- Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* 27:1877–85
- Lemmon AR, Lemmon EM. 2008. A likelihood framework for estimating phylogeographic history on a continuous landscape. *Syst. Biol.* 57:544–61
- Lieberman BS. 2000. Paleobiogeography: using fossils to study global changes, plate tectonics, and evolution. *Topics Geobiol.* 16:1–208
- Lieberman BS. 2004. Range expansion, extinction, and biogeographic congruence: a deep time perspective. In *Frontiers in Biogeography: New Directions in the Geography of Nature*, ed. MV Lomolino, LR Heaney, pp. 111–24. Sunderland, MA: Sinauer
- Maddison DR, Ruvolo M, Swofford DL. 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Syst. Biol.* 41:111–24
- Maddison WP, Maddison DR. 2007. Mesquite: a modular system for evolutionary analysis. Version 2.01. <http://mesquiteproject.org>
- Maddison WP, Midford PE, Otto SP. 2007. Estimating a binary character’s effect on speciation and extinction. *Syst. Biol.* 56:701–10
- Mansion G, Rosenbaum G, Schoenenberger N, Bacchetta G, Rossello JA, Conti E. 2008. Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean Basin by the angiosperm family Araceae. *Syst. Biol.* 57:269–85
- McGuire JA, Witt CC, Altshuler DL, Remsen JV. 2007. Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. *Syst. Biol.* 56:837–56
- Micó E, Sanmartín I, Galante E. 2009. Mediterranean diversification of the grass-feeding Anisopliina beetles (Scarabaeidae, Rutelinae, Anomalini) as inferred by bootstrap-averaged dispersal–vicariance analysis. *J. Biogeogr.* 36:546–60
- Moore MJ, Francisco-Ortega J, Santos-Guerra A, Jansen RK. 2002. Chloroplast DNA evidence for the roles of island colonization and extinction in *Tolpis* (Asteraceae: Lactuceae). *Am. J. Bot.* 89:518–26
- Morrone JJ. 2009. *Evolutionary Biogeography. An Integrative Approach with Case Studies*. New York: Columbia Univ. Press
- Morrone JJ, Crisci JV. 1995. Historical biogeography: Introduction to methods. *Annu. Rev. Ecol. Syst.* 26:373–401

---

Adds stochastic variable selection to Bayesian island biogeography to help find the best dispersal model.

---



---

Develops BI for diffusion models, accommodating variation in the dispersal rate across the tree.

---



---

Develops ML inference for diffusion models and discusses the possibility of incorporating environmental layers.

---



- Nelson G, Ladiges PY. 1996. Paralogy in cladistic biogeography and analysis of paralogy-free subtrees. *Am. Mus. Novit.* 3167:1–58
- Nelson G, Platnick NI. 1981. *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia Univ. Press
- Nepokroeff MK, Sytsma J, Wagner WL, Zimmer EA. 2003. Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understory tree genus *Psychotria* (Rubiaceae): a comparison of parsimony and likelihood approaches. *Syst. Biol.* 820–38
- Nordlander G, Zhiwei L, Ronquist F. 1996. Phylogeny and historical biogeography of the cynipoid wasp family Ibaliidae (Hymenoptera). *Syst. Entomol.* 21:151–66
- Nylander JAA, Olsson U, Alstrom P, Sanmartín I. 2008. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal-vicariance analysis of the thrushes (Aves: *Turdus*). *Syst. Biol.* 57:257–68
- Oberprieler C. 2005. Temporal and spatial diversification of circum-Mediterranean Compositae-Anthemideae. *Taxon* 54:951–66
- Olsson U, Alström P, Gelang M, Ericsson PG, Sundberg P. 2006. Phylogeography of Indonesian and Sino-Himalayan bush warblers (*Cettia*, Aves). *Mol. Phylogenetics Evol.* 41:556–61
- Page RDM. 1995. Parallel phylogenies: reconstructing the history of host-parasite assemblages. *Cladistics* 10:155–73
- Page RDM, Charleston MA. 1998. Trees within trees: phylogeny and historical associations. *Trends Ecol. Evol.* 13:356–59
- Pagel MA. 1999. The maximum likelihood approach to reconstructing ancestral character states on phylogenies. *Syst. Biol.* 48:612–22
- Pagel MA, Meade A, Barker D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* 53:673–84
- Parent L. 2007. Common cause and historical biogeography. See Ebach & Tangney 2007, pp. 61–81
- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.* 34:102–17
- Pereira SL, Baker AJ. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Mol. Biol. Evol.* 23:1731–40
- Peterson AT, Soberón J, Sanchez-Cordero V. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285:1265–67
- Pirie MD, Lloyd KM, Lee WG, Linder HP. 2010. Diversification of *Chionochloa* (Poaceae) and biogeography of the New Zealand Southern Alps. *J. Biogeogr.* 37(2):379–92
- Prieto-Márquez A. 2010. Global historical biogeography of hadrosaurid dinosaurs. *Zool. J. Linn. Soc.* 159:503–25
- Ree RH, Moore BR, Webb CO, Donoghue MJ. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–311
- Ree RH, Sanmartín I. 2009. Prospects and challenges for parametric methods in historical biogeography. *J. Biogeogr.* 36:1211–20
- Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14
- Renner SS, Srijik JS, Strasberg D. 2010. Biogeography of the Monimiaceae (Laurales): a role for East Gondwana and long-distance dispersal, but not West Gondwana. *J. Biogeogr.* 37:1227–38
- Richards CL, Carstens BC, Knowles LL. 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *J. Biogeogr.* 34:1833–45
- Ronquist F. 1994. Ancestral areas and parsimony. *Syst. Biol.* 43:267–74
- Ronquist F. 1995a. Ancestral areas revisited. *Syst. Biol.* 44:572–75
- Ronquist F. 1995b. Reconstructing the history of host-parasite associations using generalised parsimony. *Cladistics* 11:73–89
- Ronquist F. 1996. Manual DIVA v.1.1. Computer program for MacOS and Win32. <http://diva.sourceforge.net/>

---

The first statistical approach to a reticulate model, DEC.

---



---

Put ML analysis of DEC on a firm computational foundation.

---



- Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46:195–203
- Ronquist F. 2003. Parsimony analysis of coevolving associations. In *Tangled Trees: Phylogeny, Cospeciation, and Coevolution*, ed. RDM Page, pp. 22–64. Chicago: Univ. Chicago Press
- Ronquist F, Deans A. 2010. Bayesian phylogenetics and its influence on insect systematics. *Annu. Rev. Entomol.* 55:189–206
- Ronquist F, Liljeblad J. 2001. Evolution of the gall wasp–host plant association. *Evolution* 55:2503–22
- Rosenbaum G, Lister GS, Duboz C. 2002. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. *J. Virtual Explor.* 8:107–26
- Salvo G, Ho SYH, Rosenbaum G, Ree R, Conti E. 2010. Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae). *Syst. Biol.* 59:705–22
- Sanmartín I. 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachyderminae (Coleoptera: Scarabaeoidea). *J. Biogeogr.* 30:1883–97
- Sanmartín I. 2007. Event-based biogeography: integrating patterns, processes, and time. See Ebach & Tangney 2007, pp. 135–59
- Sanmartín I, Anderson CL, Alarcón-Cavero ML, Ronquist F, Aldasoro JJ. 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. *Biol. Lett.* 6:703–7
- Sanmartín I, Engloff H, Ronquist F. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linn. Soc.* 73:345–90
- Sanmartín I, Ronquist F. 2002. New solutions to old problems: widespread taxa, redundant distributions and missing areas in event-based biogeography. *Anim. Biodivers. Conserv.* 25:75–93
- Sanmartín I, Ronquist F. 2004. Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* 53:216–43
- Sanmartín I, van der Mark P, Ronquist F. 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *J. Biogeogr.* 35:428–49
- Sanmartín I, Wanntorp L, Winkworth R. 2007. West Wind Drift revisited: testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *J. Biogeogr.* 34:398–416
- Smedmark JEE, Eriksson T, Bremer B. 2010. Divergence time uncertainty and historical biogeography reconstruction—an example from Urophylleae (Rubiaceae). *J. Biogeogr.* 37:2217–396
- Smith SA. 2009. Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographic analysis of the Northern Hemisphere plant clade Caprifoliaceae. *J. Biogeogr.* 36:2324–37
- Smith SA, Donoghue MJ. 2010. Combining historical biogeography with niche modeling in the *Caprifolium* clade of *Lonicera* (Caprifoliaceae, Dipsacales). *Syst. Biol.* 59:322–41
- Szumik CA, Cuezco F, Goloboff PA, Chalup. 2002. An optimality criterion to determine areas of endemism. *Syst. Biol.* 51:806–16
- Szumik CA, Goloboff PA. 2004. Areas of endemism: an improved optimality criterion. *Syst. Biol.* 53:968–77
- Vanderpoorten A, Gradstein SR, Carine MA, Devos N. 2010. The ghosts of Gondwana and Laurasia in modern liverwort distributions. *Biol. Rev.* 85:471–87
- Wallis GP, Trewick SA. 2009. New Zealand phylogeography: evolution on a small continent. *Mol. Ecol.* 18:3548–80
- Waltari E, Hijmans RJ, Peterson AT, Nyári ÁS, Perkins SL, et al. 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* 2(7):e563
- Waters JM. 2008. Driven by the West Wind Drift? A synthesis of southern temperate marine biogeography, with new directions for dispersalism. *J. Biogeogr.* 35:417–27
- Waters JM, Craw D. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Syst. Biol.* 55(2):351–56
- Weaver KF, Anderson T, Guralnick R. 2006. Combining phylogenetic and ecological niche modeling approaches to determine distribution and historical biogeography of Black Hills mountain snails (Oreohellicidae). *Divers. Distrib.* 12(6):756–66
- Wiley EO. 1988. Parsimony analysis and vicariance biogeography. *Syst. Zool.* 37:271–90
- Wojciki M, Brooks DR. 2005. PACT: an efficient and powerful algorithm for generating area cladograms. *J. Biogeogr.* 32:755–74

---

Proposes DIVA, a parsimony approach to a reticulate model accommodating dispersal, nonhierarchical vicariance, and extinction.

---



---

A comprehensive review of parsimony-based tree fitting methods.

---



---

Applies Bayesian island biogeography to a continental problem, the Rand Flora pattern.

---



---

Develops Bayesian island biogeography, a statistical analysis of dispersal patterns across many organism groups.

---

- Yang Z. 2006. *Computational Molecular Evolution*. Oxford, UK: Oxford Univ. Press. 376 pp.
- Yesson C, Culham A. 2006. Phyloclimatic modeling: combining phylogenetic and bioclimatic modeling. *Syst. Biol.* 55:785–802
- 

## RELATED RESOURCES

- Cox CB, Moore PD. 2010. *Biogeography: An Ecological and Evolutionary Approach*. Hoboken, NJ: Wiley. 498 pp. 8th ed.
- Lomolino MV, Riddle BR, Brown JH. 2005. *Biogeography*. Sunderland, MA: Sinauer Associates Inc. 845 pp. 3rd ed.
- Myers AA, Giller PS. 1995. *Analytical Biogeography: An Integrated Approach to the Study of Animal and Plant Distributions*. London: Chapman and Hall. 578 pp.



# Contents

Native Pollinators in Anthropogenic Habitats <i>Rachael Winfree, Ignasi Bartomeus, and Daniel P. Cariveau</i>	1
Microbially Mediated Plant Functional Traits <i>Maren L. Friesen, Stephanie S. Porter, Scott C. Stark, Eric J. von Wettberg, Joel L. Sachs, and Esperanza Martinez-Romero</i>	23
Evolution in the Genus <i>Homo</i> <i>Bernard Wood and Jennifer Baker</i>	47
Ehrlich and Raven Revisited: Mechanisms Underlying Codiversification of Plants and Enemies <i>Niklas Janz</i>	71
An Evolutionary Perspective on Self-Organized Division of Labor in Social Insects <i>Ana Duarte, Franz J. Weissing, Ido Pen, and Laurent Keller</i>	91
Evolution of <i>Anopheles gambiae</i> in Relation to Humans and Malaria <i>Bradley J. White, Frank H. Collins, and Nora J. Besansky</i>	111
Mechanisms of Plant Invasions of North America and European Grasslands <i>T.R. Seastedt and Petr Pyšek</i>	133
Physiological Correlates of Geographic Range in Animals <i>Francisco Bozinovic, Piero Calosi, and John I. Spicer</i>	155
Ecological Lessons from Free-Air CO <sub>2</sub> Enrichment (FACE) Experiments <i>Richard J. Norby and Donald R. Zak</i>	181
Biogeography of the Indo-Australian Archipelago <i>David J. Lobman, Mark de Bruyn, Timothy Page, Kristina von Rintelen, Robert Hall, Peter K.L. Ng, Hsi-Te Shib, Gary R. Carvalho, and Thomas von Rintelen</i>	205
Phylogenetic Insights on Evolutionary Novelty in Lizards and Snakes: Sex, Birth, Bodies, Niches, and Venom <i>Jack W. Sites Jr, Tod W. Reeder, and John J. Wiens</i>	227

The Patterns and Causes of Variation in Plant Nucleotide Substitution Rates <i>Brandon Gaut, Liang Yang, Shohei Takuno, and Luis E. Eguiarte</i> .....	245
Long-Term Ecological Records and Their Relevance to Climate Change Predictions for a Warmer World <i>K.J. Willis and G.M. MacDonald</i> .....	267
The Behavioral Ecology of Nutrient Foraging by Plants <i>James F. Cahill Jr and Gordon G. McNickle</i> .....	289
Climate Relicts: Past, Present, Future <i>Arndt Hampe and Alistair S. Jump</i> .....	313
Rapid Evolutionary Change and the Coexistence of Species <i>Richard A. Lankau</i> .....	335
Developmental Patterns in Mesozoic Evolution of Mammal Ears <i>Zhe-Xi Luo</i> .....	355
Integrated Land-Sea Conservation Planning: The Missing Links <i>Jorge G. Álvarez-Romero, Robert L. Pressey, Natalie C. Ban, Ken Vance-Borland, Chuck Willer, Carissa Joy Klein, and Steven D. Gaines</i> .....	381
On the Use of Stable Isotopes in Trophic Ecology <i>William J. Boecklen, Christopher T. Yarnes, Bethany A. Cook, and Avis C. James</i> ....	411
Phylogenetic Methods in Biogeography <i>Fredrik Ronquist and Isabel Sanmartín</i> .....	441
Toward an Era of Restoration in Ecology: Successes, Failures, and Opportunities Ahead <i>Katharine N. Suding</i> .....	465
Functional Ecology of Free-Living Nitrogen Fixation: A Contemporary Perspective <i>Sasha C. Reed, Cory C. Cleveland, and Alan R. Townsend</i> .....	489

## Indexes

Cumulative Index of Contributing Authors, Volumes 38–42 .....	513
Cumulative Index of Chapter Titles, Volumes 38–42 .....	517

## Errata

An online log of corrections to *Annual Review of Ecology, Evolution, and Systematics* articles may be found at <http://ecolsys.annualreviews.org/errata.shtml>