

# Historical Biogeography

Fredrik Ronquist

November 16, 2005

## 1 Introduction

Historical biogeographers and phylogeographers study very similar problems even though the time scales and spatial scales are typically larger in historical biogeography. Phylogeographers decided early on to ignore much of the existing literature in historical biogeography and the two disciplines still develop largely independently today.

Historical biogeography can be divided into two subdisciplines: *taxon biogeography* and *area biogeography*. In the former we are interested in reconstructing the geographic distribution history of a particular group of organisms. Questions we may be interested in answering include: Where did the organisms originate? How did they colonize the areas in which they live today? When did they first come to a particular area? How prone to dispersal is the lineage? In area biogeography, the focus is on general phenomena that affect many groups of organisms. Typical research questions include: How was a particular community assembled? What is the explanation of commonly observed disjunct distributions? What are the most important geographic dispersal barriers that affect or affected the distributions of organisms in a particular region of the world? Are differences in species diversity between two regions best explained by differential speciation or differential extinction?

In terms of quantitative methods, historical biogeography is still dominated by parsimony or optimization techniques. There is an overwhelming number of approaches that have been discussed in the literature. They can be classified into *pattern-based* and *event-based*; the former approaches are simple numerical recipes whereas the latter identify some biologically interesting events, associate these events with costs, and then find the minimum-cost solutions. In recent years, there has also been a few papers discussing likelihood approaches to problems in historical biogeography, but the development of these methods is still in its infancy.

## 2 Area models

Before discussing the techniques used in taxon and area biogeography, it is useful to describe the different models or scenarios that are used to describe how organism distributions evolve over time.

The simplest model is the *island model*. Geographic areas are represented as islands that remain constant through time (Fig. ??). Important processes that shape the distribution of lineages include dispersal between islands and speciation within islands. The use of island models is certainly not restricted to true islands. Many types of geographic situations can be described in terms of islands. For instance, patches of particular habitat like mountaintops, freshwater lakes, or rivers can all be thought of as islands.

Figure 1: Different area models used in historical biogeography



The *vicariance model* starts with a large geographic area, which is split into smaller and smaller parts by the appearance of dispersal barriers (Fig. ??). Organism lineages that respond to the appearance of a dispersal barrier by speciating produce a vicariance pattern in which sister groups have complementary distributions, together corresponding to a larger ancestral distribution.

The most complex scenario is the *reticulation model*. Here, distribution areas can both be split through vicariance or united through fusion, producing very complex patterns of events that potentially affect many groups of organisms (Fig. ??).

### 3 Taxon biogeography

In the simplest of all event-based approaches, we use the island model and Fitch optimization to reconstruct ancestral areas. Strictly speaking, this approach implies that a single evolutionary lineage always inhabits a single island and that dispersal of a lineage to a new island results in extinction on the old island (Fig. ??). We can then find the optimal ancestral areas by simply using Fitch optimization, minimizing the number of inter-island dispersals (and associated extinctions on the pre-colonization islands). An alternative formulation, which also leads to Fitch optimization under some mild conditions, models dispersal as being coupled with speciation (Fig. ??). In the former interpretation, there is no extinction without associated dispersal, and in the latter there is no extinction at all. Alternatively, we could say that if we included such extinction events in the model, and if they had a positive cost, they would still never be included in optimal solutions.

What if the islands form a geographic chain? In such a case, one might expect dispersal between neighboring islands to be more frequent than dispersal between more distant islands. This could be modeled using Wagner parsimony, in which islands are ordered in terms of their geographic position. This can of course be extended to more complex situations, in which Sankoff cost matrices could be used to describe the expected frequency of dispersal between different islands. The more frequent a particular dispersal event is expected to be, the lower the associated cost should be. Of course, the exact costs need to be determined somehow, and there are no standard methods for estimating them from distribution data.

As soon as we abandon Fitch optimization in favor of Wagner parsimony or more complex Sankoff cost optimization, there is a potential that optimal solutions will involve extinction in “intermediate” areas. However, as long as costs are metric, there will always be alternative interpretations that do not involve such extinction. To prove this it is sufficient to refer to the triangle inequality. This property of metric distances means that going from any point A to any other point C through an intermediate point B must be at least as costly as going directly from A to C.

It would be relatively straight forward to implement likelihood techniques in the context of island models but the development of this approach is still in its infancy and basic concepts, such as the correct interpretation of stationary state frequencies and transition rates, have not been adequately

Figure 2: Different models of inter-island dispersal leading to Fitch optimization techniques.



explored yet. An important reason for this may be the lack of statistical power, particularly in taxon biogeographic problems, which makes standard empirical estimation of parameters difficult, particularly in a maximum likelihood framework.

What if we want to do taxon biogeography while taking the possibility of vicariance into account? A simple technique, which has been used occasionally and that does not have a name as far as we know but could be referred to as *areasum optimization*, assumes that ancestral distributions are the sum of all the descendant distributions (Fig. ??). In the case of a pure vicariance history, this technique would accurately reconstruct the ancestral areas. It should also work well when vicariance is occasionally followed by secondary back colonization of previously occupied areas. However, colonization of new areas or extinction would cause mismatches between the reconstructed and the true ancestral areas.

A simple technique that combines the island and vicariance perspectives is *dispersal-vicariance optimization*. It is based on modeling distributions in terms of a set of unit areas, or islands if you wish. Speciation is assumed to result in division of an ancestral distribution into two disjoint sets of

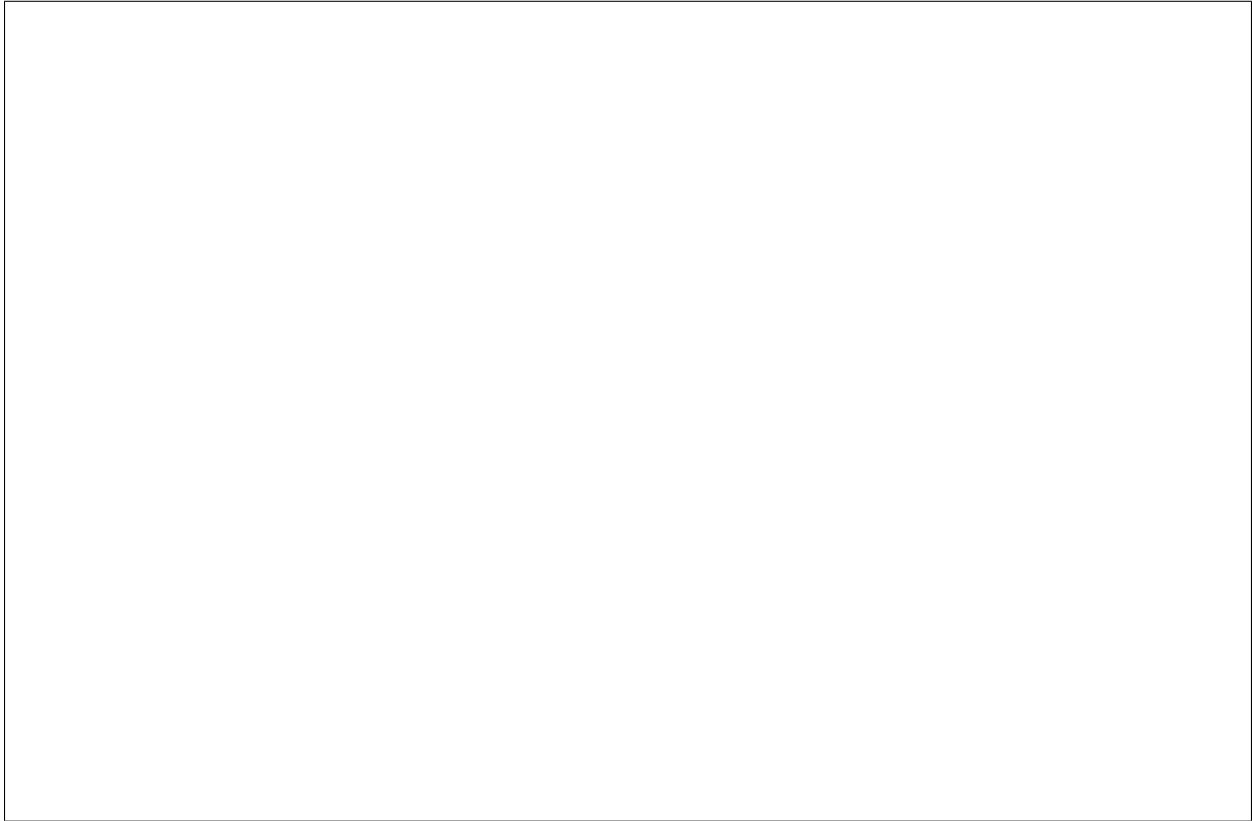
Figure 3: The areasum optimization technique.



unit areas. However, if speciation occurs in an ancestor occupying a single unit area, both daughter lineages are assumed to occur in the same area. Both of these events have zero cost. Lineages can then colonize new unit areas or go extinct in previously occupied unit areas; a unit cost is used for each addition or deletion of a unit area from a distribution. The resulting technique will in some cases favor an island colonization history and in some cases a vicariance history (Fig. ??). Note that the vicariance events need not conform to a general pattern; one lineage can be reconstructed as having a vicariance history involving a completely different sequence of geographical splitting than its sister lineage. This also means that, in principle, dispersal-vicariance optimization can reconstruct reticulate distribution histories.

Before ending this section, it is worth pointing out that results from area biogeography approaches, such as general estimates of dispersal rates, could be used in taxon biogeography to increase the precision of the reconstructions of past distributions for single organism groups. However, this is still rarely done.

Figure 4: The dispersal-vicariance optimization technique.



## 4 Area biogeography

The simplest approach to generalities in historical biogeography is to use one of the taxon biogeography techniques and then simply summarize the results across organism groups. For instance, we could apply Fitch optimization over the same set of islands for a large set of organism groups and then summarize the number of dispersals between particular islands to get an estimate of the relative frequency of different types of dispersal events. We might, for instance, discover asymmetric dispersal between some islands, perhaps showing the effects of prevailing winds or ocean currents on long-distance dispersal. For an example involving dispersal-vicariance optimization and the historical biogeography of terrestrial animals in the northern Hemisphere, see the paper by Sanmartin, Ronquist, and Enghoff (2001). They used dated phylogenies to be able to separate biogeographic events temporally and thus address a reticulate geological history.

A disadvantage with the above approach is that we lose some information by ignoring the presence of generalities when reconstructing the distribution history of each individual group. For instance, if we have a choice between explaining a particular pattern in one organism group as the result of

event A or event B, it might be helpful to know that across a larger set of organisms, event A is common and event B rare. To take such information into account, we need techniques that try to identify general events and put a premium on using these in explaining the distribution history of individual groups of organisms.

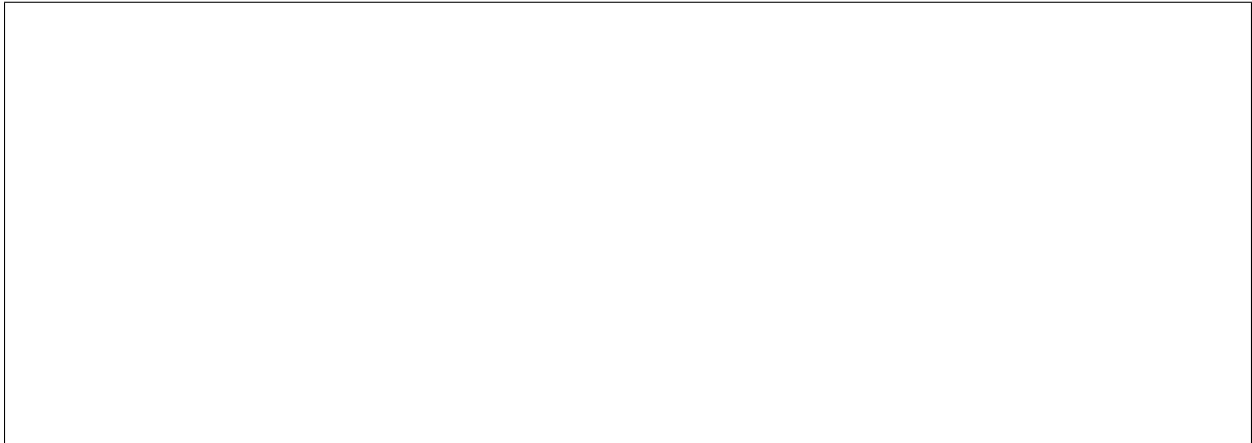
Most of the work on such methods has focused on parsimony techniques based on the vicariance model, and it exemplifies the more general problem of tree fitting, which we will examine in more detail in the next lecture.

If the vicariance model is true, then we can describe the history of areas in terms of branching diagrams, which are called *area cladograms* or sometimes general area cladograms. Each split in the area cladogram corresponds to the appearance of a particular geographic area splitting one ancestral area into two smaller daughter areas. If all organisms had a perfect vicariant history, then the area cladogram could be found by simply replacing the organisms with their distributions in any of the phylogenies of organisms inhabiting the areas. However, if processes such as random dispersal and extinction occur on top of a generally vicariant history, we need more sophisticated techniques for measuring the fit between a particular area cladogram and one or more organism phylogenies.

The first step in describing such a method is to identify the events. Typically, there are four types of events that are considered in parsimony-based approaches to finding general area cladograms (Fig. ??). These are vicariance, duplication (speciation within an area), sorting (partial extinction), dispersal (speciation-associated), extinction, and non-speciation dispersal. In *vicariance*, a lineage of organisms responds to the appearance of a general dispersal barrier by speciating. *Duplication* within one area can result either from sympatric speciation or from allopatric speciation in response to a temporary local dispersal barrier. A *sorting* event results when an organism lineage, in response to the appearance of a general dispersal barrier, either goes extinct in one of the daughter areas or fails to make it to one of the areas (missing the boat). Finally, a *dispersal* event is typically modeled as being associated with speciation. True extinction and dispersal associated with extinction in the original area are difficult to consider in parsimony methods because they completely erase the information about past distributions. Therefore, they are typically ignored.

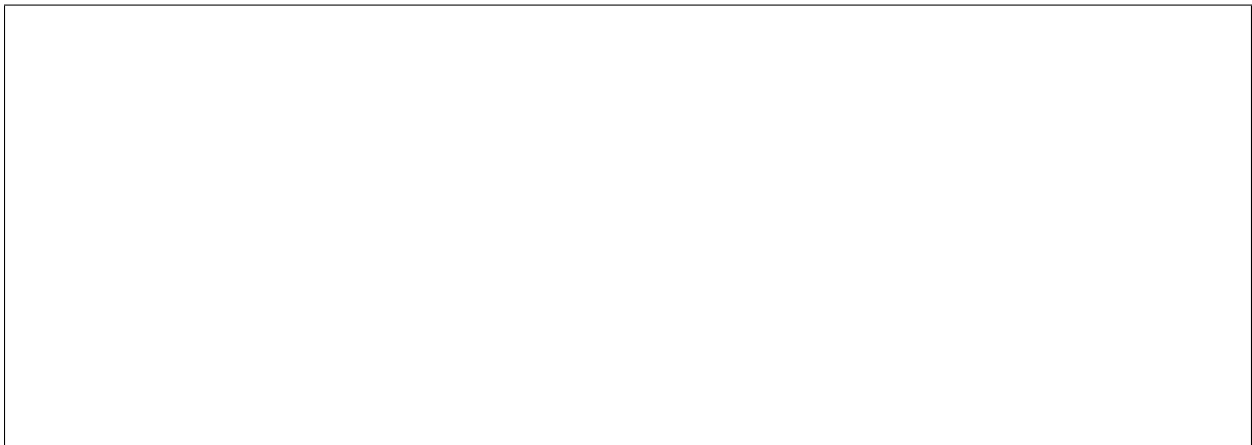
Assuming a model with these four events, the next step is to identify a cost for each. In other contexts where parsimony methods are used, it is possible to use the same cost for all events. A simple example is the standard four-by-four model of DNA substitution, where Fitch parsimony uses the same cost for all transitions. However, in the biogeographic context this is not possible, at least not if we wish to reconstruct vicariance scenarios. Consider a simple example in which

Figure 5: Biogeographic vents considered in parsimony-based methods for finding general area cladograms.



two sister species occur in two areas (Fig. ??). The pattern can be explained by a dispersal event, a vicariance event, or a duplication event followed by two sorting events. Although there are no patterns that will fit a vicariance scenario better, the vicariance explanation will not be the single most parsimonious reconstruction unless a vicariance event costs *less* than a dispersal event; thus, these two events cannot have the same cost. Similarly, a vicariance event must cost less than the sum of a duplication and two sorting events.

Figure 6: Alternative explanations for a perfectly vicariant pattern.



A simple solution is to maximize the number of vicariance events with no consideration of the number of other events that must be postulated. This is equivalent to using a negative cost for vicariance events and a zero cost for the three remaining events. The reason that this approach works is that the postulated events are all constrained by the requirement that the descendant lineages show traces of the events in their distributions. Thus, it is not possible to introduce extra



events even if their cost is non-positive.

Even so, the maximization has several disadvantages. One of the most serious is that maximization is inefficient in distinguishing between random and non-random patterns, in particular if duplication is common. A more powerful approach is to associate all events with a positive cost. Consideration of a range of hypothetical patterns suggests that optimal event costs are similar in most cases: duplication and vicariance events should have a low cost and sorting and dispersal events a high cost. For instance, one can associate vicariance and duplication events with a very low cost, sorting events with a unit cost, and dispersal events with a cost of 2. This cost assignment scenario would order the three explanations of the vicariant sister lineage pattern (Fig. ??) such that the dispersal and sorting explanations would be roughly equivalent and much more costly than the vicariance explanation.

In the next lecture, we will explore various parsimony approaches to tree fitting in more detail, and we will also look at some likelihood-based approaches as well.

## 5 Study Questions

1. What is the difference between historical biogeography and phylogeography?
2. What is the difference between taxon biogeography and area biogeography?
3. Describe different area models used in historical biogeography.
4. How is Fitch parsimony used in taxon biogeography? How can it be used in area biogeography?
5. Can extinction events be postulated in optimal dispersal-vicariance reconstructions?
6. Define the four events commonly used in parsimony-based techniques of finding general area cladograms.
7. What is maximum vicariance analysis?