Integrative Biology 200A "PRINCIPLES OF PHYLOGENETICS"
University of California, Berkeley   Kipling Will- In revision 2012

Biogeography—from Area Homology and patterns to why organisms are where they are--
The field that attempts to document and understand the spatial patterns of organisms in the past and present and develop process explanations and historical scenarios for variation in distribution and/or seeks a natural classification of areas.

I. Ecological Biogeography
      A. Typically deals with essentially atemporal, relatively recent patterns and interactions with an ecological, physiological and phenological emphasis.
            1. From this view questions addressed might be like "What allows a species to occur in one area and prevents it from expanding into other areas?"
      B. Community based.
      C. Can involve paleontological data, so is not restricted to currently extant populations.
      D. Studies often involve the impact of human activities; succession theory; the dynamics of communities and populations; fire ecology; restoration ecology; invasive species; species pulses or waves; island biogeography.

e.g., Elias, S.A., Berrnan, D., Alfimov, A. 2000. Late Pleistocene Beetle Faunas of Beringia: Where East meets West. Journal of Biogeography. 27:1349-1~63.

II. Historical Biogeography
      A. Usually involves historical (temporal), typically older patterns inferred by looking at clades (often as species and higher taxa that are monophyletic)
            1. How species accumulate (or are lost)
            2. How biotas come into being
      B. Focus on why lineages are represented in certain areas and not others and why is a pattern of distributions frequently repeated in different lineages.
      C. Typically area relationships and general patterns of diversity are emphasized over single taxon distributions.

e.g., Marshall, C. J., Liebherr, J.K. 2000. Cladistic biogeography of the Mexican transition zone. Journal of Biogeography. 27(1):203-216.

III. Brief Historical overview of "periods" and discoveries in historical biogeography
      A. At the time of a limited view of dynamics and diversity
            1. Creation myths, dispersal from Noah's ark, etc.
            2. Little understanding beyond local flora/fauna
      B. Age of exploration
            1. 17th century led to discovery of too many species for the Ark
            2. Realization that environmentally similar but distantly isolated regions have distinct assemblages of organisms (Buffon's Law)
            3. Islands have lower lineage diversity
            4. Similar floristic zonation (Humbolt)
      C. 19th century, advances in geology and evolutionary theory
            1. Lyell, Darwin, Wallace, Sclater, Hooker, etc.
            2. Abandonment (by most) of the idea of static distribution and immutability of species.
            3. Land bridges/megacontinents (e.g. Hooker, Wallace) vs. dispersal (e.g. Darwin)

D. 20th century

    1. Continental drift (Wegener 1912)- but not really accepted until 1960s- (Dietz & Holder 1966) Persistent dispersalists like Simpson and Darlington maintained that even if continents moved dispersal was still more important.

    2. Panbiogeography of Croizat, phylogenetics of Hennig and implementation Brundin (1966) for a trans-Gondwanan group of flies

        a. Start of phylogenetic biogeography

        b. similarity of phylogenetic and distributional patterns implies a shared biogeographical history.

    3. Vicariance biogeography fully developed by Platnick & Nelson (1978)

    4. Event based estimations of biogeographic history

IV. Generally used methods of Historical (Vicariance) Biogeography

    A. Similar systems like parasites and their hosts can use the same or similar methods.

Host - Associate
Organism - genes
Host organism - parasite organism
Geographic area - organism

IV. Historical explanation for disjunct patterns can be either dispersal or fragmentation of a continuous range (vicariant).

    A. Vicariance biogeography considers dispersal to be an ad hoc explanation that could be used to account for any pattern. Dispersal is not falsifiable.

    B. Implied process assumptions in pattern methods, which are more explicit in event-based methods

        1. allopatric speciation is relative more common

        2. dispersal is relatively less common than stasis

        3. current distributions generally reflect ancient events

    C. The basics

        1. You need monophyletic groups that occur in three or more areas

            a. Requires that you assume the phylogeny is correct

        2. Replace the OTU names with area names to make taxon-area cladograms

            a. Requires you assume the known ranges are adequate

        3. Summarize the set of taxon-area cladograms to make a general area cladogram that shows the congruent vicariant pattern

            a. Assumption of simplicity or at least that the majority of taxa respond the same most of the time

        4. "Missing", widespread taxa and redundant areas are problems.

            a. Areas may be absent in groups because

                i. no member ever occur in the area

                ii. sampling error

                iii. extinction

            b. Taxa may occur in more than one area because

                i. non-response to a vicariant event

                ii. dispersal (secondary sympatry)

            c. Redundant distributions

                i. dispersal in one or more taxa

                ii. sympatric speciation

V. *A priori methods*- Derive general biogeographic patterns from shared cladistic patterns and distributional data. No secondary or "total evidence" analysis.

A. Component Analysis. TAC, FAC, RAC, GAC: Use taxon area cladograms (TAC) to derive reduced area cladograms (sometimes called fundamental area cladograms (FAC)) from phylogenetic hypotheses and distributional data under Assumptions 0, 1, 2 (Zandee & Roos 1987; Platnick & Nelson 1978). General Area cladograms (GAC) are the intersection of reduced area cladograms (RAC) from two or more sets of trees.

Assumption 0: assumes that widespread taxa are the result of non-response to only the most recent vicariant event, so that areas that are occupied by widespread taxa are interpreted as being most closely related. Clearly, the existence of more than one widespread taxon, each occupying overlapping but not identical sets of areas are conflicting evidence for area relationships. The conflict is resolved for these redundant areas for the most apical cladistic position.

Assumption 1: assumes that widespread taxa are the result of non-response to the most recent OR earlier vicariant event. Redundant area representations are all considered valid. The assumption will often produce more trees (the convex set) than the more restrictive assumption 0, but will always include Assumption 0 trees as a subset.

Assumption 2: allows for the possibility that dispersal, as well as non-response to a vicariant event, may be the explanation for widespread taxa. (fix one move the other to all possible positions).

B. Reconciled Trees (Page 1990, 1994): Mapping one tree onto another by the inclusion of unobserved events.
    1. User tree (host or area relationships) mapped to observed or associate tree (taxa or parasites)
    2. Items of Error (IOE): Mis-fit of the trees. Number of nodes in reconciled tree – number of nodes in associate tree (in Componant 2.0 "leaves added" = ½ IOE)
    3. Various area cladograms are reconciled to find an optimal fit.
    4. Heuristic searches may focus on trees that minimizes leaves added, however this may overestimate the number of events.
C. We have been using "areas" but what are Areas of Endemism?
    1. "if they are not exactly congruent, but are largely congruent, then they may be considered equivalent for purposes of biogeographical analysis." (Wiley 1981)
    2. What the terminals are in a biogeographic study is fundamental to doing the analysis in the first place. However, there has been much less attention given to area definition/discovery than to tree conversion/combination in to a general area cladogram. An Area of Endemism is an area of nonrandom distributional congruence. (Morrone 1994).
    3. Most frequently AofE are determined intuitively through inspection of distributional information and often endemic areas are equivalent to species distributions. However, various grid methods have been used (Morrone 1994; Linder 2001; Szumik et a1. 2002, all in Sys. Bio).
    4. Linder (2001) published these criteria:
        i. The ranges of the species endemic to the area of endemism should be maximally congruent. Ideally all taxa would be evenly distributed in the area. [generally not the case]
        ii. The areas of endemism should be narrower than the whole study area, so that several areas are located.

iii. Areas of endemism must be mutually exclusive, and grid squares (or any other units) cannot belong to two areas of endemism.

iv. If areas of endemism are to be determined before cladistic biogeographical analysis, then that definition should not be open to retrospective modification, which would weaken the power of incongruent distributions in falsifying any patterns.

5. A simple example using PAE (parsimony analysis of endemisity, Rosen 1988; Morrone 1994)

i. Draw quadrats (grid) on a map, consider quadrats only where at least one locality record exists for a study taxon.

ii. Construct a data matrix where columns are species and rows are quadrants. Use 0-1 = absence - presence coding.

iii. Add an all zero quadrant (the empty or what we might call the Stanford quadrat) as a hypothetical root (where nothing exists).

iv. Analyze for the parsimony solution.

v. Delimit groups of quadrats with at least two species

vi. Superimpose the delimited groups on the map to delineate the AofE boundries.

vii. This is a non-historical scale dependent method.

viii. How is the host-parasite system different than organism-areas system?

VI. *A Posteriori* methods- taxon distributional data is compiled into a single matrix and subsequently analyzed.

A. Brooks parsimony analysis -BPA (Brooks 1985,1990; Wiley 1987)

See the recent debate in

Siddall, M.E., Perkins, S.L., 2003. Brooks Parsimony Analysis: a valiant failure. Cladistics, 19:554-564.

Brooks, Daniel R., Dowling, Ashley P. G., van Yeller, Marco G. P. & Hoberg, Eric P. 2004. Ending a decade of deception: a valiant failure, a not-so-valiant failure, and a success story. Cladistics 20 (1):32-46.

1. Convert taxon cladogram an area x taxon/node matrix (MRP)

2. Analyze to find the parsimony solution (primary BPA)

3. Widespread, redundant and missing taxa may cause the solution to include impossible ancestor-descendent relationships. If so, re-code offending areas (duplicate) and re-run analysis or force characters (taxa and their ancestor-descendent relationships) to be polarized and irreversible (secondary BPA).

4. Combine many such matrices and read the resulting relationships from the parsimony tree(s).

VII. Comparative Biogeography (sensu Parenti & Ebach 2009)

1. Focus is on homology of natural areas not on speciation or other diversity generation processes.

2. Defined as "the comparative study of biotic areas and their relationships through time"

3. Uses "paralogy-free subtree analysis"

VIII. Event based optimality methods for estimation of biogeographic history (you will do this in 200B and be sure to read Ronquist & Sanmartin, 2011 and Crisp et al., 2011)

1. Biogeographic history is another step of abstraction from phylogeny

2. Event-based methods is rapidly expanding due to advances in ML and Bayesian methods, parametric statistics, DNA data and fossil data.

3. Calibrated phylogenies

    4. DIVA, DEC, HVM…

IX. Use of Phylobetadiversity methods to look at fauna/floral assembly
    1. Various processes, e.g. speciation, extinction, vicariance, dispersal, environmental and ecological filtering, etc. have contribute to the current composition
    2. Phylobetadiversity metrics show changes over space in the phylogenetic relationships of members of a community (=sample) (Graham & Fine 2008)
    3. Should be applicable to all scales and ages.
    4. Example from New Caledonia: Fragment islands and continental areas should exhibit higher phylogenetic dispersion than oceanic islands, which should be relatively 'clustered' with respect to phylogenetic relationships.