11. Ancestral reconstruction
EEOB563

Spring 2021

"Character state reconstructions can provide a powerful mechanism for studying many facets of the evolutionary process. However, the zeal with which these techniques are sometimes advocated belies the complexity of the problem". Swofford and Maddison, 1992.

1 Preliminary considerations

The pathways and patterns shown by character evolution are important to our understanding of evolutionary processes and predicting the future. Could we see into the past, we would like to know the history of changes both in time and in relationship to each other. This would allow us to learn how many times traits evolved, the conditions that favored their evolution, and the lineages in which changes occurred. Of course, in most cases we don't know the past so we have to be satisfied with inferring the history of character evolution indirectly ("statistical paleontology"). A number of methods have been developed to do this.

1.1 Some examples of ancestral character state reconstruction

- Michael Ryan and Stanley Rand (1995) reconstructed ancestral advertisement calls in Tungara Frogs and observed reactions of female frogs to them.

- Gaucher et al. (2003) explored the palaeoenvironment of ancient bacteria on the basis of resurrected proteins.

- Bridgham et al. (2009) investigated the importance of epistatis in protein evolution.

1.2 Do early branching lineages signify ancestral traits? (after Crisp and Cook, 2005)

Failure to distinguish between present-day descendants and long-dead ancestors has led to incorrect interpretation of phylogenetic trees. Misinterpretation becomes evident when authors use the terms "basal" or "early diverging" to refer to extant taxa.

For example are monotremes "ancient/ancestral mammals"? Need to remember that whereas some character states can be relatively ancestral (plesiomorphic) or derived (apomorphic), these concepts should not be applied to whole organisms. Only nodes on a tree can be referred to as "basal" or "derived" relative to each other.

A common misinterpretation occurs when a species-poor/less diverse sister group is labeled as "basal" or "early-diverging" with respect to its species-rich sister. This misreading of phylogenies

is encouraged by asymmetrical trees. Simplistic interpretation of phylogenies is also common in biogeography, where, for example, researchers may attempt to trace the history of a lineage back to its ancestral area.

2 Methods and models for inferring ancestral states

In reconstructing character evolution, our goal is to reconstruct the character states of hypothetical ancestors throughout the phylogenetic tree. The usual practice is to estimate the states of the hypothetical ancestors at the nodes or branch points of the tree, because these provide convenient reference points. Many methods are available (e.g., table 1 from Crisp and Cook below)

Table 1. Methods and models for inferring ancestral areas

Method or model	Type	Properties	Deficiencies	Examples of appropriate use	Refs
Parsimony	P[a]	Minimizes discrete-state changes over the tree; state changes can be weighted equally (Fitch) or differentially using step matrices	Cannot indicate probability of estimates; does not use branch lengths, thus underestimates change when it is frequent relative to speciation	Double fertilization, long thought to be a unique defining character of angiosperms, might have originated independently in gymnosperms, or earlier, in the first seed plants	[14,19, 41]
Dispersal vicariance analysis (DIVA)	P	Uses cost matrices to estimate ancestral areas; differentially costs vicariance and dispersal	Bias towards sympatric speciation and against early dispersal	Ancestral area reconstructions for oaks were more consistent with fossil record using DIVA than using Fitch parsimony or strict vicariance	[16,24]
Markov continuous time transition model	ML[b]	Estimates rate of discrete-state changes, allowing asymmetry	Evolutionary assumptions, e.g. that rate of change is constant throughout the tree, can be unrealistic (see punctuated models)	Overturned the parsimony estimate of origin of ruminant digestion in artiodactyls by taking rate of change into account	[21][c]
Punctuated evolution model	ML	Assumes evolutionary change occurs only at speciation	Evolutionary assumption is equivalent to having equal branch lengths and likely to be unrealistic, as in parsimony	For *Psychotria* plants in Hawaii, ancestral area inferences differ from both parsimony and ML models using differential branch lengths	[3,29]
Generalized least squares (=Squared change 'parsimony')	ML	Minimizes sum of squared changes over the tree; Brownian motion model of evolution; uses branch lengths	Does not indicate probability of estimates; non-directional version cannot reconstruct ancestral values outside range of descendant values	Size (area) of the first lower molar in mammals was reconstructed accurately, with verification from the fossil record	[18,28, 54]
Maximum likelihood with General time reversible model	ML	Based on Markov model, adapted from nucleotide modelling; allows differential rates and symmetry	Can fail if model unrealistic	Estimated ancestral areas, and rates and direction of dispersal of plants among Hawaiian islands	[3,55]
Stochastic (Bayesian inference), e.g. using Markov continuous time transition model	B[d]	Models multiple parameters including tree topology; posterior probabilities given for estimates	Can fail if model unrealistic	In contrast to parsimony, Bayesian analysis found multiple credible histories of gains and losses of horned soldiers in aphids	[26]

[a]Parsimony.
[b]Maximum likelihood.
[c]Implemented in Mesquite http://mesquiteproject.org.
[d]Bayesian inference.

2.1 Finding the most parsimonious ancestral states

Ancestral state reconstruction using parsimony is done in two or three passes up and down the tree. The details of the process may differ with the different assumptions used; the algorithm for unordered characters is presented below. It creates three sets of character states. The first set is called the downpass set of the node, because this is the set of states preferred by that part of the tree above the node. The second set is called uppass set and is created by uppass optimization. The

uppass state set of a node is calculated from the uppass set of the node below it (its ancestor) with the downpass set of the node beside it (its sister node), using the same calculation for combining state sets as for the downpass. For the final state of each node, consider the uppass set of that node, and the downpass sets of its two descendant nodes. Choose the state that has the greatest number in all three sets. If none is in a majority, it remains ambiguous.
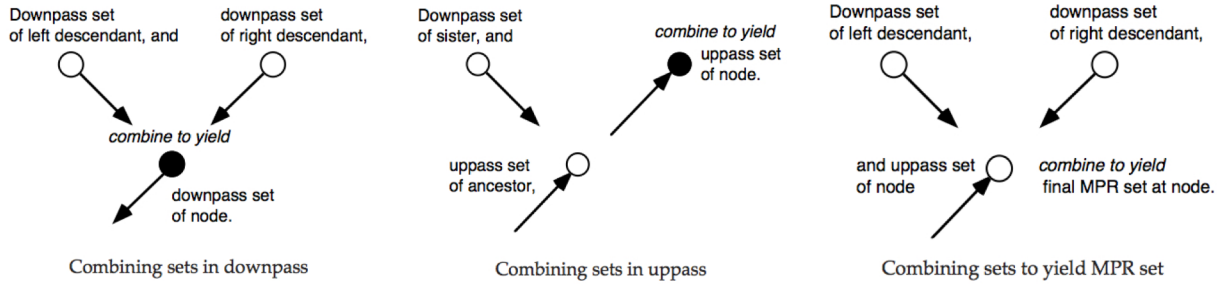

Figure 1: Parsimony reconstruction of ancestral states using Fitch algorithm

2.1.1 Reconstruction uncertainty

The first expression of uncertainty is ambiguity in the reconstruction or the existence of multiple equally parsimonious reconstructions. Ideally, when faced with multiple reconstructions, we should examine all of them. In reality, the number of possible reconstruction is often prohibitory large and only a few of them are examined/presented.

Two popular methods of choice are ACCTRAN and DELTRAN, which yield extremes of reversals versus parallelisms in the reconstruction.
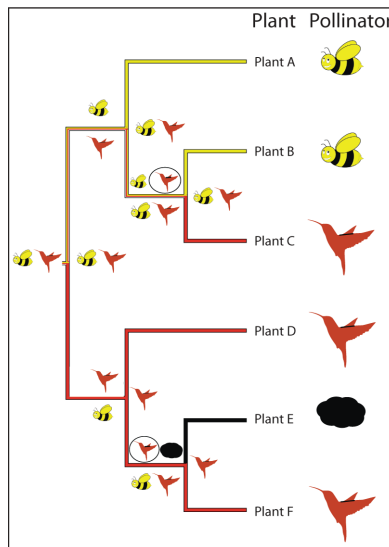

Figure 2: Parsimony reconstruction of the class example using Fitch algorithm without prior info on the root. Pictures on the left of each node show dowpass (above) and uppass(below) reconstructions. Pictures on the right of each node show final reconstruction.

2.1.2 ACCTRAN and DELTRAN

For characters of unordered and ordered type, ambiguities in character tracings can be resolved so as to choose the assignments that delay or accelerate transformations (look at the figure above and think what will happen to our reconstruction if we assign white/black color to node G). The DELTRAN option prefers most parsimonious assignments that delay changes away from the root; this maximizes parallel changes. The ACCTRAN tracing shows those assignments that accelerate changes toward the root; this procedure maximizes early gains and thus forces subsequent reversals. ACCTRAN and DELTRAN are but two of various methods to select from among the most parsimonious reconstructions.

Uncertainty does not end with ambiguity in the reconstruction. Even unequivocal assignment may be incorrect. Reconstructions of ancestral states are subject to error, as are all estimates of history. Studies of the reliability of ancestral state reconstructions have yielded mixed results, although it is clear that when rates of evolution are high over the time scale of the tree, error rates can be high. Also, the rate of change from one character state to another character state can be different (e.g., losses can be much easier than gains).

2.2 Probabilistic Methods (ML and Bayesian)

The use of probabilistic methods (ML) provides several advantages: 1) Use an explicit model of character evolution 2) Consider branch length 3) Can estimate the relative probability of each character state at every node.

Several models are used to estimate ancestral states of discrete and continuous characters based on phylogenetic data and character states of extant descendants. For discrete-valued traits a continuous-time Markov chain is usually used; for continuous-valued traits a Brownian motion or an Ornstein-Uhlenbeck process is used.

2.2.1 Discrete-state models

A continuous-time Markov chain process is usually used to model the evolution of discrete traits. In this model, each state has associated rates of transition to all of the other states. ML methods can be used to find the ML values for the parameters (rates of transitions) and then the likelihood of each possible state for a node given these rates. One may also use this substitution model as the basis for a Bayesian inference procedure, which would consider the posterior belief in the state of an ancestral node given some user-chosen prior.

2.2.2 Binary state speciation and extinction model

The binary state speciation and extinction model (BiSSE) is a discrete-space model that allows estimation of ancestral binary character states jointly with diversification rates associated with different character states. In its most basic form, this model involves 6 parameters: 2 speciation
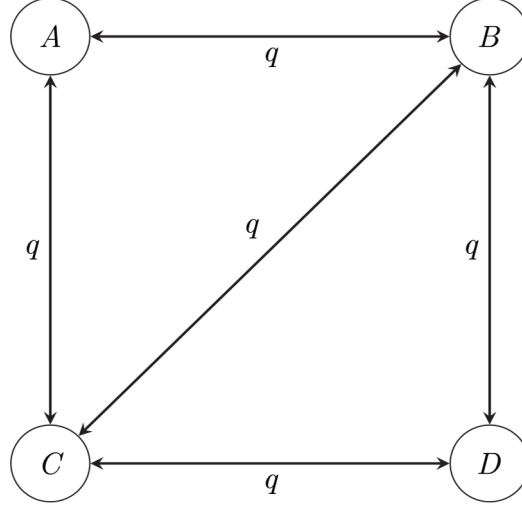
Figure 3: Example of a four-state 1 parameter Markov chain model. (From: Joy et al. 2016)

rates (one each for lineages in states 0 and 1); similarly, 2 extinction rates; and 2 rates of character change. This model allows for hypothesis testing a binary character's effect on speciation and extinction.

2.2.3 Continuous-state models

Although continuous traits can be split into discrete categories, it is more appropriate to model their evolution as some continuous process. In this case the likelihoods of transitions in state between adjacent nodes will be given by a continuous probability distribution.

2.2.4 Discrete-state models

Brownian motion: If nodes U and V are adjacent in the phylogeny (say U is the ancestor of V) and separated by a branch of length t, the likelihood of a transition from U being in state x to V being in state y is given by a Gaussian density with mean 0 and variance $2\sigma t$. In this case, there is only one parameter (σ^2), and the model assumes that the trait evolves freely without a bias toward increase or decrease, and that the rate of change is constant throughout the branches of the phylogenetic tree.

Ornstein-Uhlenbeck process: a continuous stochastic process that behaves like a Brownian motion, but attracted toward some central value, where the strength of the attraction increases with the distance from that value. Applicable when a trait is subject to stabilizing selection around a certain value (say 0). Under this model the above-described transition of U being in state x to V being in state y would have likelihood defined by the transition density of an Ornstein-Uhlenbeck process with two parameters: σ^2, which describes the variance of the driving Brownian motion, and α, which describes the strength of its attraction to 0.

Stable models of continuous character evolution: though Brownian motion is appealing
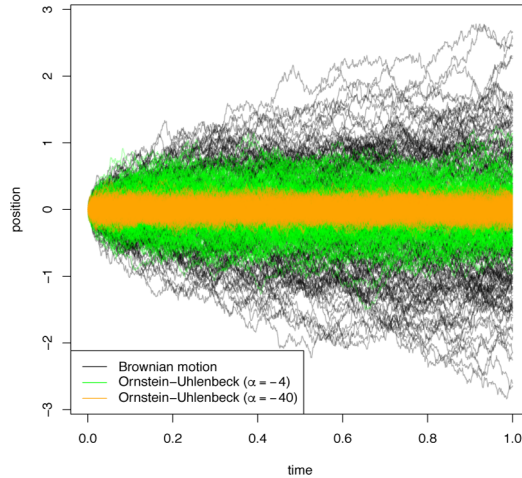
Figure 4: Plots of 200 trajectories of each of: Brownian motion with drift 0 and $\sigma^2 = 1$ (black); Ornstein-Uhlenbeck with $\sigma^2 = 1$ and $\alpha = -4$ (green) and Ornstein-Uhlenbeck with $\sigma^2 = 1$ and $\alpha = -40$ (orange). (From: Joy et al. 2016)

and tractable as a model of continuous evolution, it does not permit non-neutrality in its basic form, nor does it provide for any variation in the rate of evolution over time. Instead, one may use a stable process, one whose values at fixed times are distributed as stable distributions, to model the evolution of traits. Stable processes behave as Brownian motions but also incorporate discontinuous jumps. This allows one to appropriately model scenarios in which short bursts of fast trait evolution are expected. In this setting, maximum likelihood methods are poorly suited due to a rugged likelihood surface and because the likelihood may be made arbitrarily large, so Bayesian methods are more appropriate.

Bayesian methods provide additional advantages 1) avoid potential errors in fixed parameter estimates 2) address phylogenetic uncertainty

2.3   Software

We will be using Bayes Traits for the lab exercise. Bayes Traits is a computer package which performs analyses on discrete or continuous characters in ML and Bayesian frameworks and allows testing of hypotheses about models of evolution, ancestral states, and correlations among pairs of traits.

There are many other software packages available; a selection is described below:

Within the R statistical language: APE implements a variety of ancestral state reconstruction methods for both discrete and continuous characters. Diversitree implements ancestral state reconstruction under Mk2 and BiSSE models.

HyPhy is a modular software package for hypothesis testing using phylogenies in a maximum likelihood framework. HyPhy implements a fast joint likelihood method of ancestral sequence reconstruction that can be readily adapted to reconstructing discrete ancestral character states or

geographic ranges.

Mesquite implements parsimony, maximum likelihood, and Bayesian methods of ancestral state reconstruction for both discrete and continuous characters and has several display methods for resulting reconstructions.

Bayes Traits is a computer package which performs analyses on discrete or continuous characters in ML and Bayesian frameworks and allows testing of hypotheses about models of evolution, ancestral states, and correlations among pairs of traits.

BEAST Also performs ancestral state and ancestral sequence reconstruction analyses.

Lagrange is an application which allows analyses reconstruction of geographic range evolution on phylogenetic trees available from http://www.reelab.net/home/software/.

Phylomapper implements a likelihood-based statistical framework for estimating historical patterns of gene flow and ancestral geographic locations. Available from: http://www.evotutor.org/LemmonLab/PhyloMapp