1. Can we use the CFN model for morphological traits?

2. Can we use something like the GTR model for morphological traits?

3. Stochastic Dollo.

4. Continuous characters.

Mk models

k-state variants of the Jukes-Cantor model – all rates equal.

$$\Pr(i \to i|\nu) = \frac{1}{k} + \left(\frac{k-1}{k}\right) e^{-\left(\frac{k}{k-1}\right)\nu}$$

$$\Pr(i \to j|\nu) = \frac{1}{k} - \left(\frac{1}{k}\right) e^{-\left(\frac{k}{k-1}\right)\nu}$$

Sampling morphological characters

Using our models assumes that our characters can be thought of as having been a random sample from a universe of iid characters.

1. We never have constant morphological characters.
   (a) There are plenty of attributes that do not vary.
   (b) The "rules" of coding morphological characters are well-defined.
   (c) How many constant characters "belong" in our matrix?

Solutions to the lack of constant characters

1. Score our taxa for a random selection of characters – *not* a selection of characters that are chosen because they are appropriate for our group. (Is this possible or desirable?)
2. Account for the fact that our data is filtered.

Mk_v model

Introduced by Lewis (2001) using a trick Felsenstein used for restriction site data.

We condition our inference on the fact that we know that (by design) our characters are variable.

If $\mathcal{V}$ is the set of variable data patterns, then we do inference on:

$$\Pr(x_i|T, \nu, x_i \in \mathcal{V})$$

rather than:

$$\Pr(x_i|T, \nu)$$

Conditional likelihood

If $x_i \in \mathcal{V}$, then:

$$\Pr(x_i|T, \nu, x_i \in \mathcal{V}) \Pr(x_i \in \mathcal{V}|T, \nu) = \Pr(x_i|T, \nu)$$

So:

$$\Pr(x_i|T, \nu, x_i \in \mathcal{V}) = \frac{\Pr(x_i|T, \nu)}{\Pr(x_i \in \mathcal{V}|T, \nu)}$$

Note that:

$$\Pr(x_i \in \mathcal{V}|T, \nu) = 1 - \Pr(x_i \notin \mathcal{V}|T, \nu)$$

If $\mathcal{C}$ is the set of constant data patterns:

$$x_i \notin \mathcal{V} \equiv x_i \in \mathcal{C}$$

So:

$$\Pr(x_i \in \mathcal{V}|T, \nu) = 1 - \Pr(x_i \in \mathcal{C}|T, \nu)$$

There are not that many constant patterns, so we can just calculate the likelihood for each one of them.

Inference under M2_v

1. Calculate $\Pr(x_i|T,\nu)$ for each site i
2. Calculate

$$\Pr(x \in \mathcal{C}|T,\nu) = \Pr(000\ldots0|T,\nu)+\Pr(111\ldots1|T,\nu)$$

3. For each site, calculate:

$$\Pr(x_i|T,\nu,x_i \in \mathcal{V}) = \frac{\Pr(x_i|T,\nu)}{1-\Pr(x \in \mathcal{C}|T,\nu)}$$

4. Take the product of $\Pr(x_i|T,\nu,x_i \in \mathcal{V})$ over all characters.

Mk_v and M$k_{pars-inf}$

The following were proved by Allman et al. (2010)

1. Mk_v is a consistent estimator of the tree and branch lengths,
2. If you filter your data to only contain parsimony-informative charecters:
   (a) A four-leaf tree cannot be identified!
   (b) Trees of eight or more leaves can be identified using inference under M$k_{pars-inf}$

Can we estimate biases in state-transitions and state frequencies from morphological data?

Can we estimate biases in state-transitions and state frequencies from morphological data?

Of course! (remember Pagel's model, which we have already encountered).

But we have to bear in mind that 0 in one character has nothing to do with 0 in another.

This means that we have to use character-specific parameters or mixtures models (to reduce the number of parameters). Typically this is done in a Bayesian setting.

Other tidbits about likelihood modeling of non-molecular data

1. We can use the No-common-mechanism model (Tuffley and Steel, 1997) to generate a likelihood score from a parsimony score (for combined analyses).
2. By setting some rates to 0 we can test transformation assumptions about irreversibility.
3. Modification to the pruning algorithm lead to models of Dollo's law (no independent gain of a character state). For further details, see Alekseyenko et al. (2008).
4. The use of ontologies to describe characters may revolutionize modeling of morphological data and the prospects for constructing "morphological super-matrices"

References

Alekseyenko, A., Lee, C., and Suchard, M. (2008). Wagner and Dollo: a stochastic duet by composing two parsimonious solos. *Systematic Biology*, 57(5):772–784.

Allman, E. S., Holder, M. T., and Rhodes, J. A. (2010). Estimating trees from filtered data: Identifiability of models for morphological phylogenetics. *Journal of Theoretical Biology*, 263(1):108–119.

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, 50(6):913–925.

Tuffley, C. and Steel, M. (1997). Links between maximum

likelihood and maximum parsimony under a simple model of site substitution. *Bulletin of Mathematical Biology*, 59(3):581–607.