



modelling trait
evolution on
phylogenies:
discrete traits

There are no better words than those of **Luke Harmon** to kickstart this:

Comparative methods combine biology, mathematics, and computer science to learn about a wide variety of topics in evolution using phylogenetic trees and other associated data.

For example, we can find out which processes must have been common, and which rare, across clades in the tree of life; whether evolution has proceeded differently in some lineages compared to others; and whether the evolutionary potential that we see playing out in real time is sufficient to explain the diversity of life on earth, or whether we might need additional processes that may come into play only very rarely or over very long timescales.

What can you do with phylogenetic comparative methods?

- model fitting - testing evolutionary hypotheses on how trait evolve.
- ancestral state(s) reconstructions - **will do in the practical!**
- diversification - speciation and extinction rates.
- biogeography .
- ...

Character Types:

- **Discrete Characters:** traits with a finite number of distinct states.
Examples: presence or absence of wings, flower color categories.
- **Continuous Characters:** traits that vary along a measurable scale.
Examples: body length, enzyme activity levels, beak depth.

Analytical Approaches:

- **Discrete Characters:** analyzed using models like the Mk model, which describes the evolution of discrete traits as a Markov process.
- **Continuous Characters:** often examined with models as Brownian Motion (BM) and Ornstein-Uhlenbeck (OU) processes.

Discrete characters

The most commonly used model for discrete character evolution on trees is a model called the **Mk model**.

- **M** for the process modelled follows the Markov property
- **k** as the model is generalized to include an arbitrary number (**k**) states

The Q matrix is the core of the model and gives the instantaneous transition rates between states... **does it remember you anything ?!**

Depending on the assumptions we make about the transition rates, the Mk model can be parameterised in several different ways:

- **ER** - Equal Rates: all transitions have the same rate.
- **SYM** - Symmetrical: different rates but forward and reverse trans. equal.
- **ARD** - All Rates Different: each transition has its own rate.

Joint *versus* marginal ancestral reconstruction

- Joint reconstruction is finding the set of character states at all nodes that (jointly) maximize the joint probability (likelihood) of the entire tree.
- marginal reconstruction is finding the state at the current node that maximizes the likelihood integrating over all other states at all nodes.

Stochastic character mapping for **ancestral reconstruction**

In stochastic character mapping we sample character histories in direct proportion to their posterior probability, under a model.

This is accomplished by first sampling a transition matrix Q (from its posterior probability distribution), then sampling a set of ancestral states at the nodes of the tree from their joint conditional probability distribution given Q . Finally, we simulate character histories along all the edges of the tree conditioned on Q and our sampled node states.

... what about parsimony for ancestral reconstruction?

We get a parsimony reconstruction of our character on the tree from stochastic mapping if we put a very strong prior on Q to be small.

This suggests that parsimony implicitly assumes that Q is very small – even if contrary evidence exist in our data suggesting Q is large.

This means that the parsimony reconstruction will only accurately reflect the evolutionary process for our character when Q is very small.

State-dependent Speciation and Extinction (SSE) models

- Explore how specific traits influence speciation and extinction rates.
- Integrate trait evolution with diversification processes to understand the relationship between character states and lineage diversification.

Key models:

- BiSSE (Binary-State Speciation and Extinction)
- MuSSE (Multi-State Speciation and Extinction)
- QuaSSE (Quantitative State Speciation and Extinction)
- HiSSE (Hidden-State Speciation and Extinction)
- GeoSSE (Geographic State Speciation and Extinction)

FINISH