Discrete trait models

Introduction to phylogenetic comparative methods

BIO508- Rosana Zenil-Ferguson

Pollination

- Plants depend on pollination to start the breeding process

- Pollen can be transferred from plant to plant by insects, wind, vertebrates, or even water

- There are multiple types of pollination, and frequently within a flowering plant/genus or family we can find different strategies for pollination.

- Often flower shape tells us something about the type of pollination we can expect.
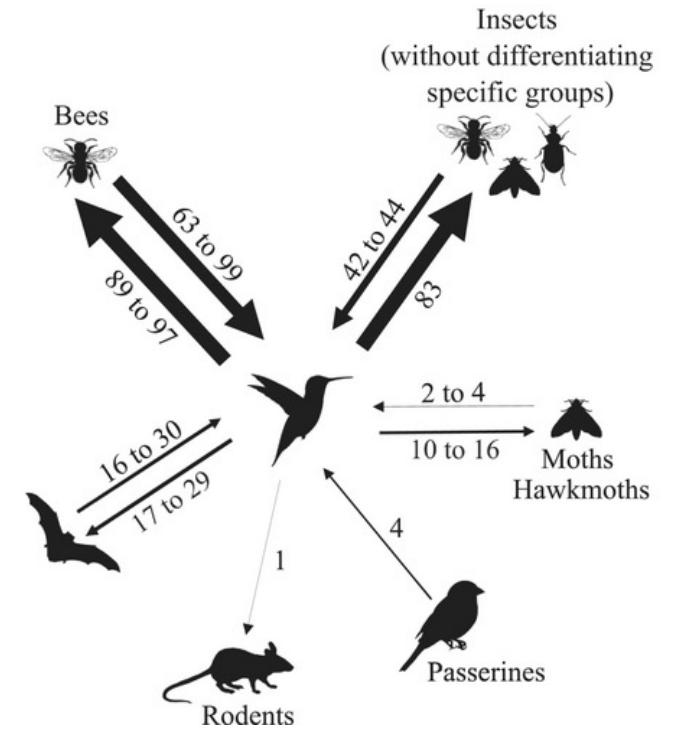
Barreto et al. 2024
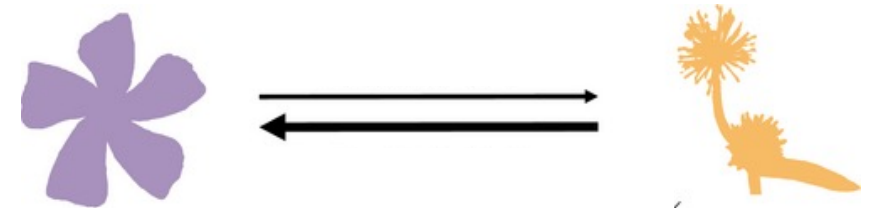
How do we estimate the number of transitions and the tempo?

Cumberland Falls State Park, KY

Thalictrum pubescences

Thalictrum thalictroides

Thalictrum
multiple changes between
insect and wind pollination

Verónica di Stilio
University of Washington

Fotos: Tennessee-Kentucky Plant Atlas, iNaturalist, and RZF

Pollination modes

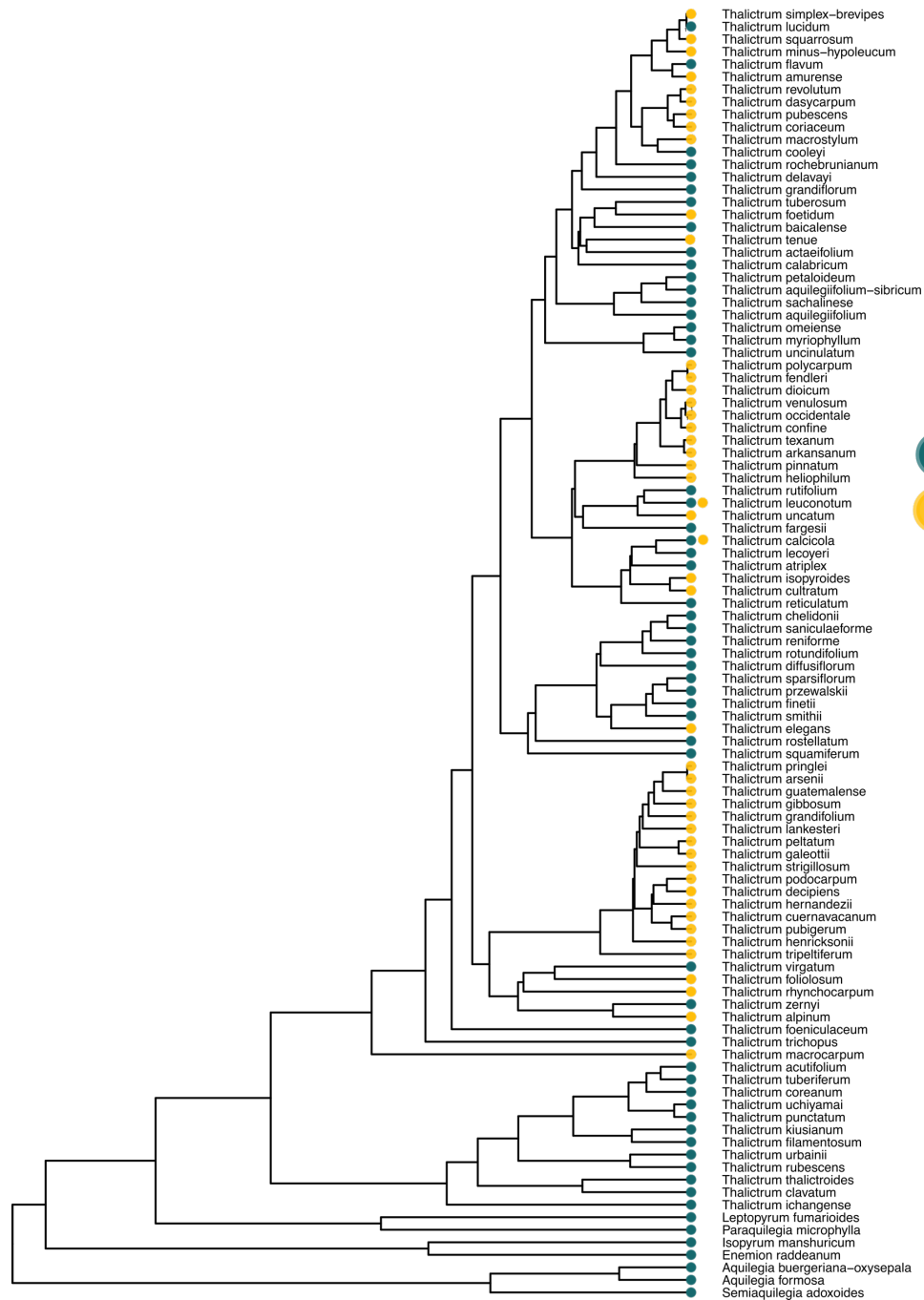Insect= 0

many pollination mode changes

Wind =1

- Always plot your data!

- Check for nestedness of trait!

- Transitions are the effective sample size for these analyses

Felsenstein's worst case scenario

Random variable definition

Continuous-Time Markov Chains (CTMC)

$$\{X(t), t \geq 0\}$$

Stochastic models that follow change in time with an associated **probability**

$X(t)$= phenotype (trait) value at time t

t= millions of years or expected number of substitutions

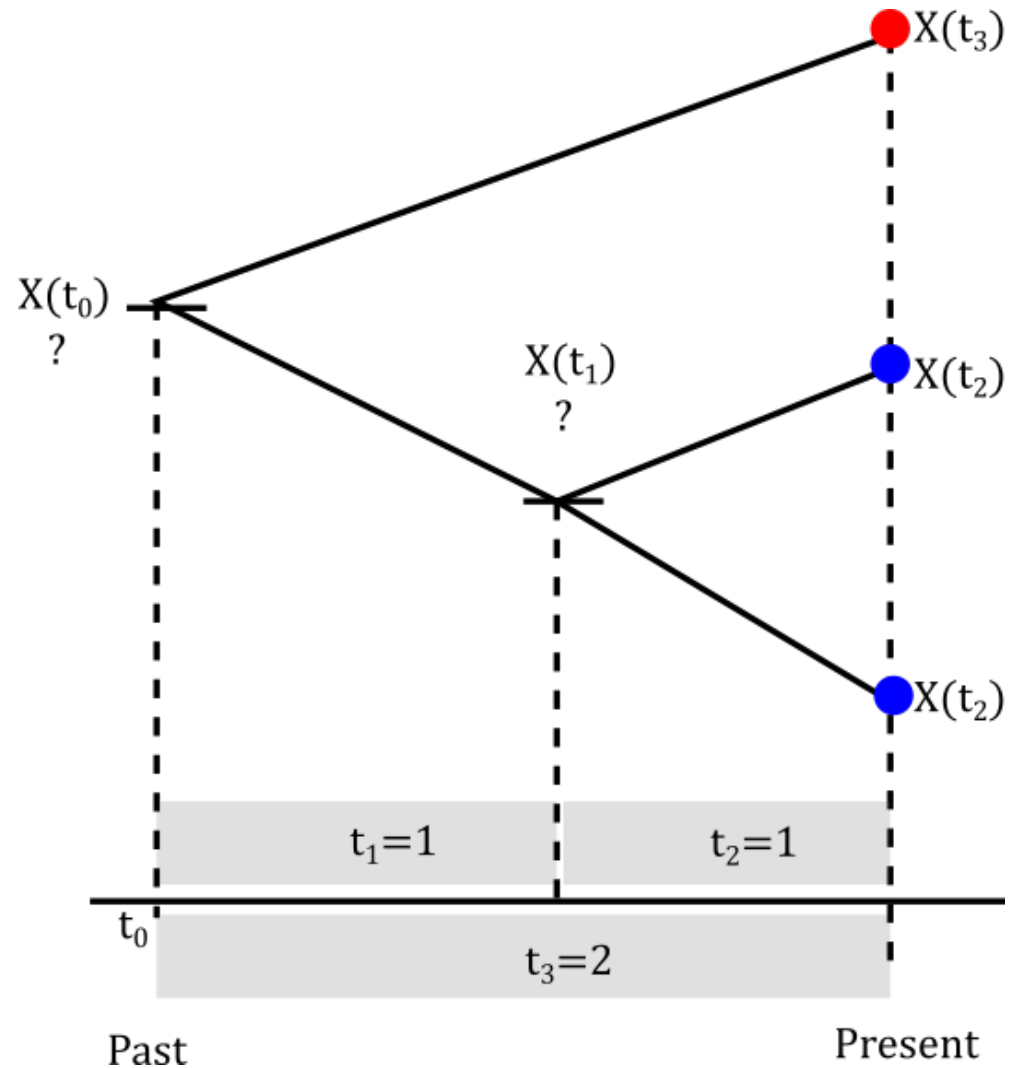# Discrete models of evolution

$X(t)$= phenotype (trait) value at time t

$X(t)$= Insect (0), Wind (1)

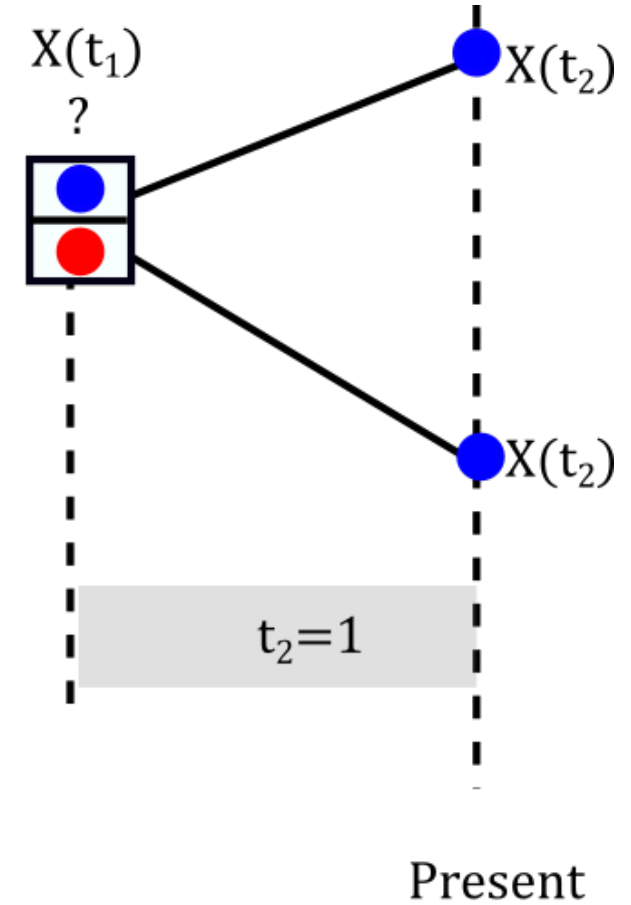We use conditional probabilities. For example:

$$P(X(t) = 1 | X(0) = 0)$$

Drawing the model

Data: Phylogenetic tree and traits

Calculating transition probabilities

$$P(X(t) = 1|X(0) = 0) = \frac{1}{q_{01} + q_{10}}\left(q_{10} - q_{01}e^{-(q_{01}+q_{01})t}\right)$$

X(t_1)

?

X(t_2)

X(t_2)

$t_2=1$

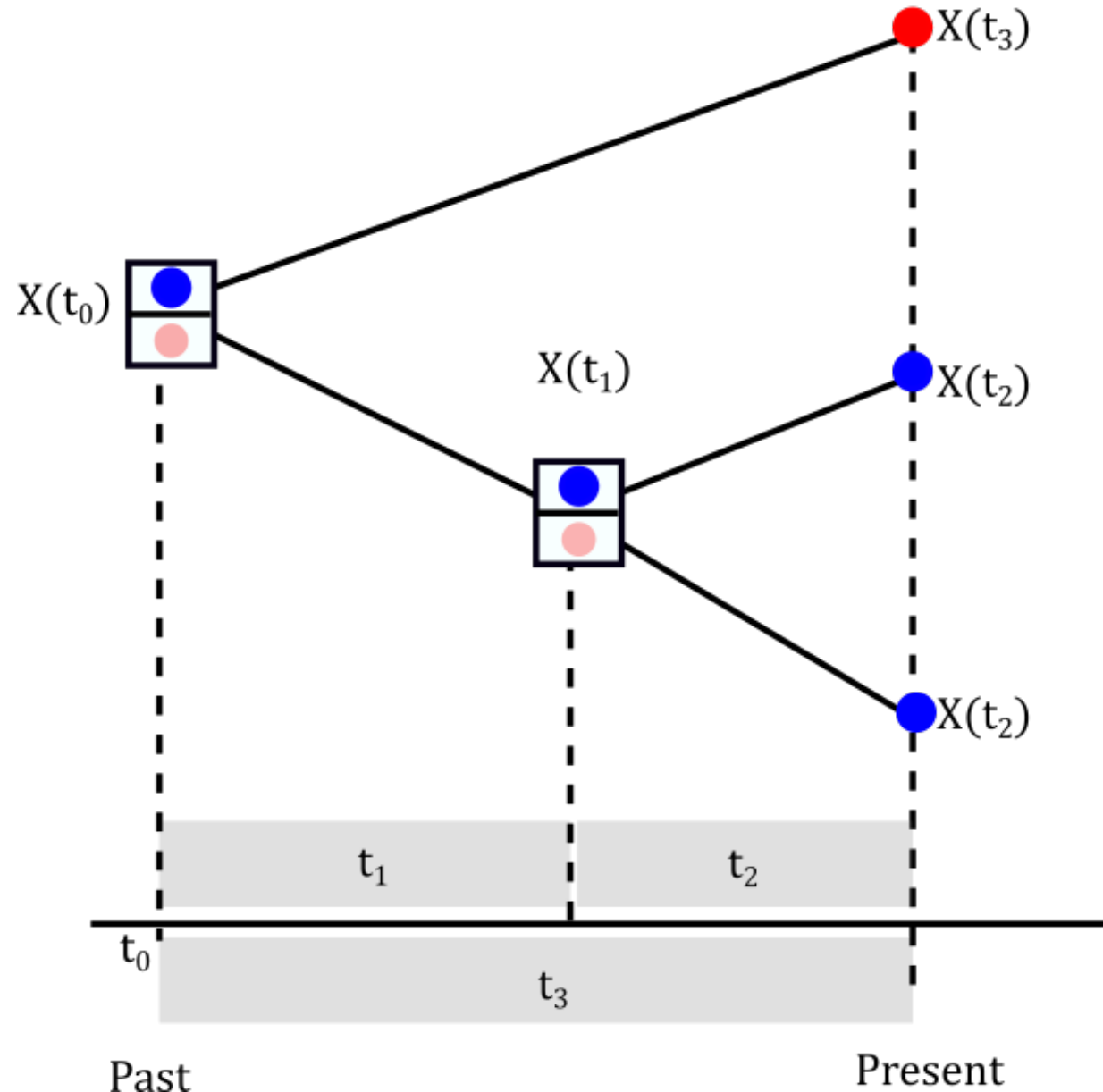Past   Present
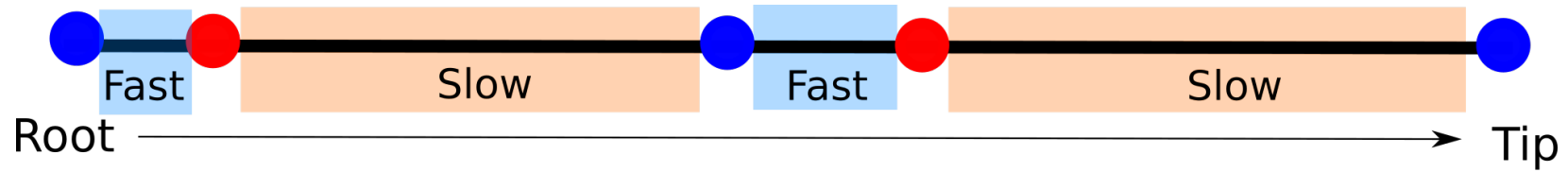
Probabilities and the Q-matrix

$$P(t) = \frac{1}{q_{01} + q_{10}} \begin{pmatrix} q_{10} + q_{01}e^{-(q_{01}+q_{10})t} & q_{01} - q_{01}e^{-(q_{01}+q_{10})t} \\ q_{10} - q_{10}e^{-(q_{01}+q_{10})t} & q_{01} + q_{10}e^{-(q_{01}+q_{10})t} \end{pmatrix}$$

$$P(t) = e^{Qt}$$

What about the rest of the tree?
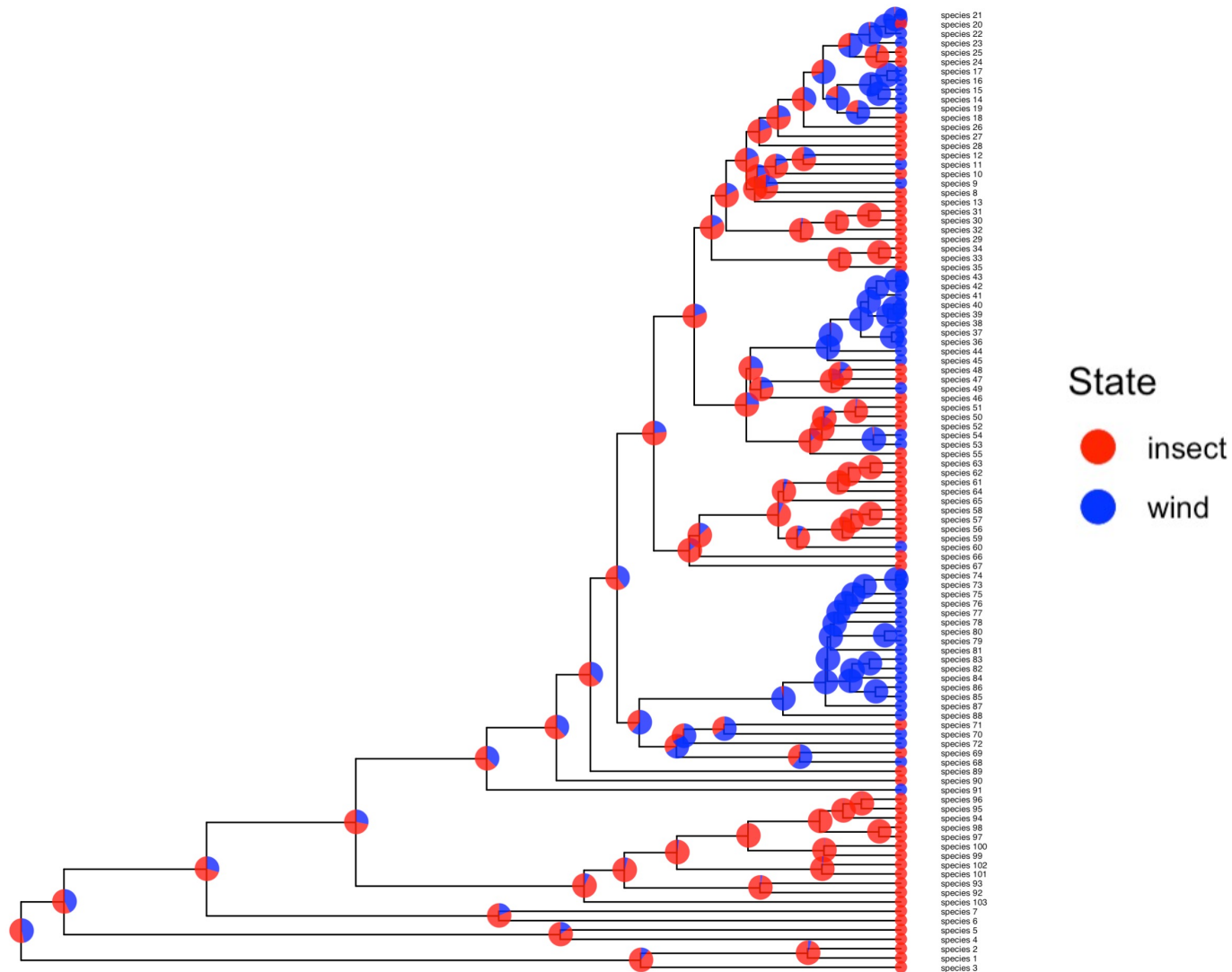
Rates: Change per time

Likelihood Estimation

Hypothesis testing- Likelihood

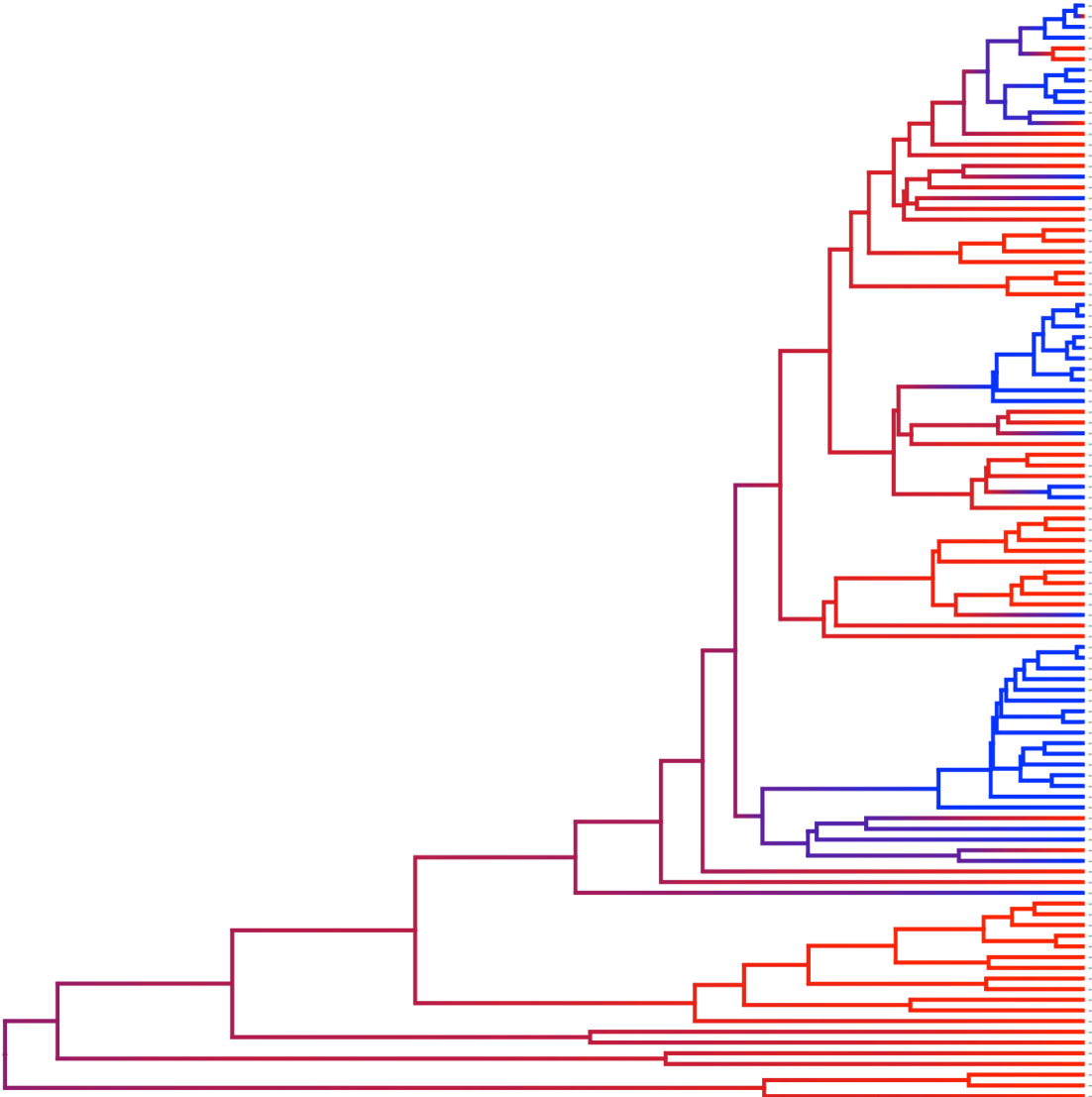Null hypothesis $H_0: q_{01} = q_{10}$

Ancestral state reconstruction (marginal)

- **Marginal likelihood function**

This is a posterior distribution pie

Stochastic mapping

- Many simulations= stochastic maps

- Posterior distribution in small time slices along branches (summarized as the MAP in each of the chunks)

- *Can be summarized with other statistics

Can you talk about the number of transitions with an stochastic map?