Q- the Matrix
Introduction to Markov Chains

Modeling discrete traits

NSF-DEB 2323170

Rosana Zenil-Ferguson
Assistant Professor. University of Kentucky
SSB Breakthrough Meeting 2026.

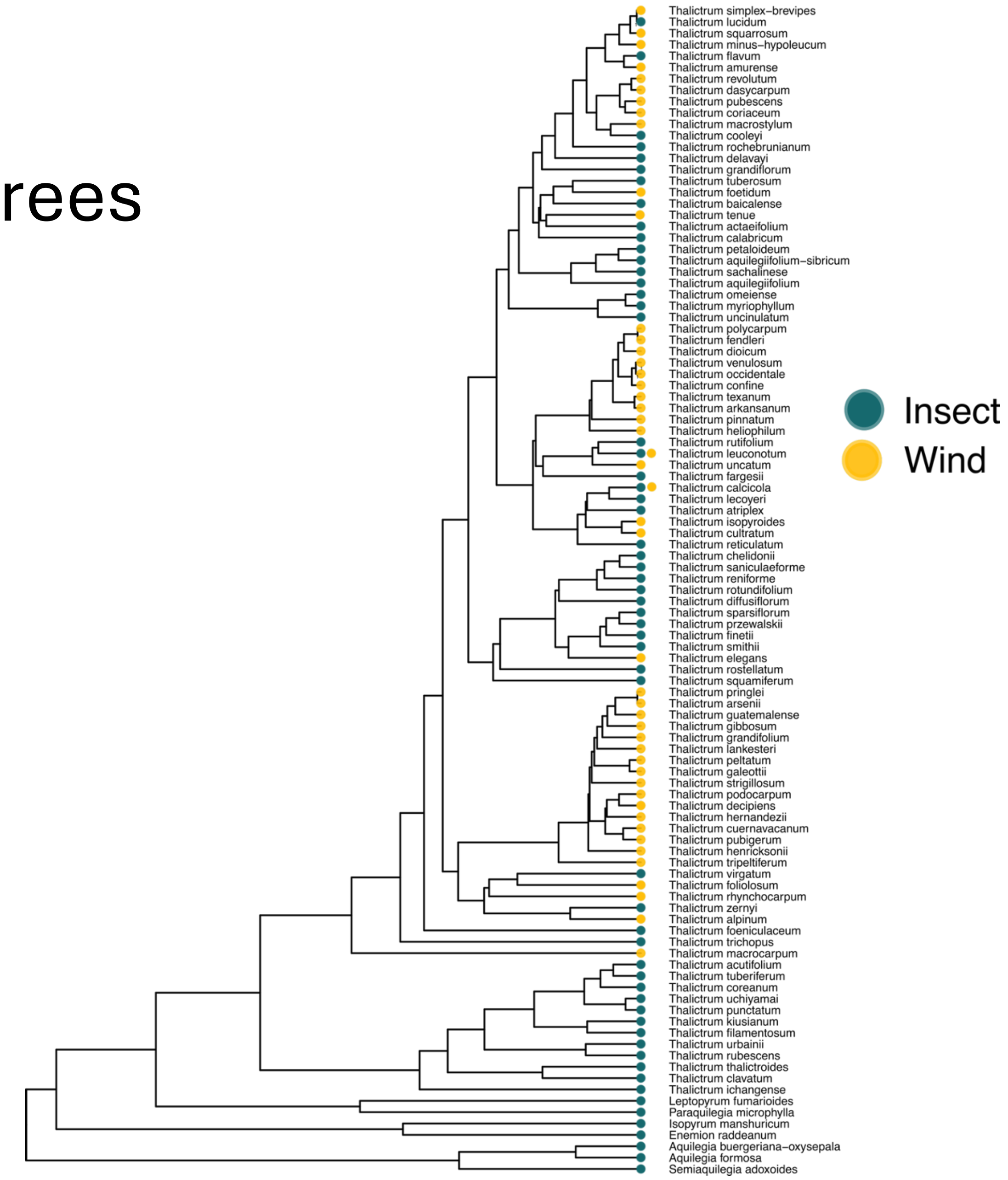
Download slides and files

https://roszenil.github.io/mytutorials

SCAN ME

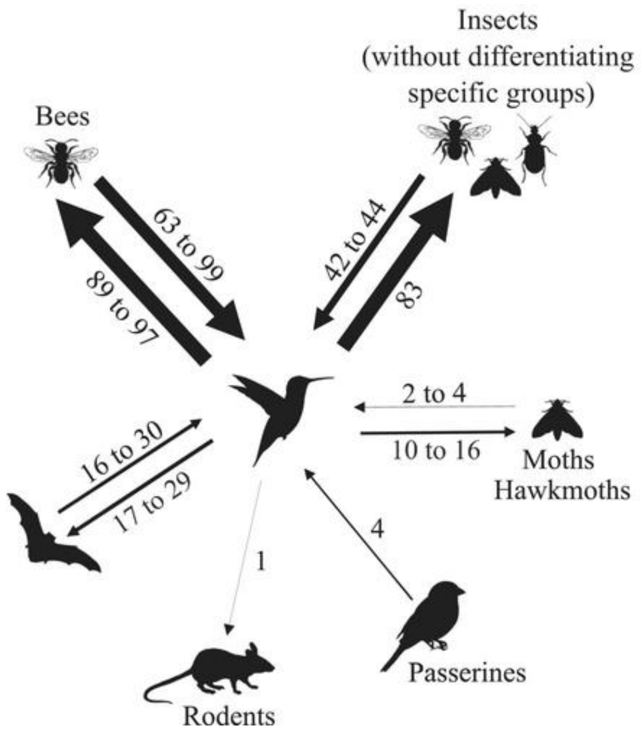
Discrete traits and phylogenetic trees

- We assume we "know the phylogenetic tree"
- We are interested in the number of times a state evolved (and how quickly



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.

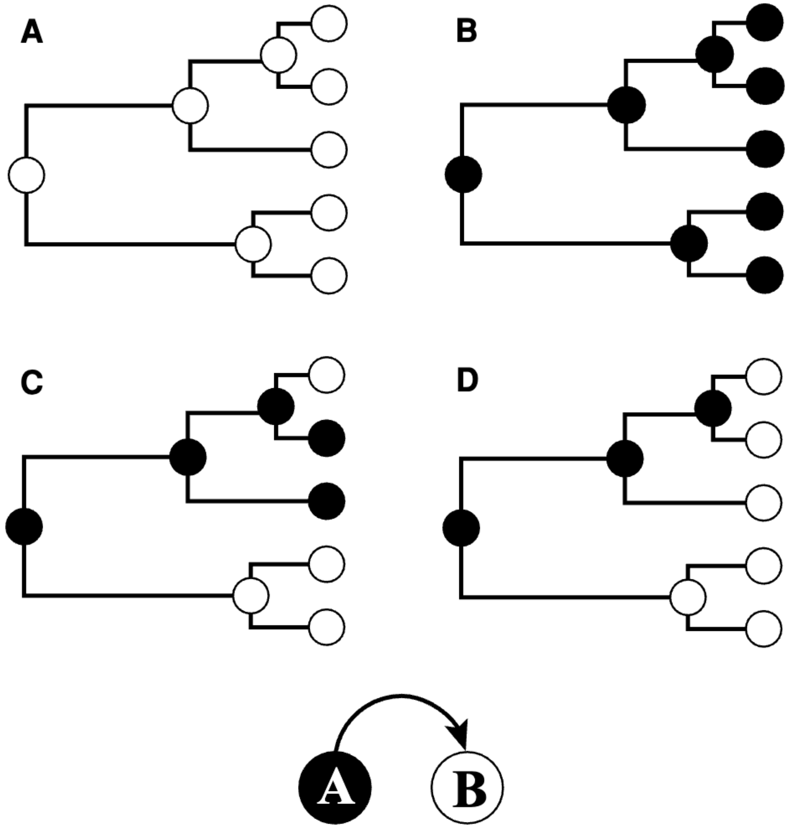


Barreto et al. 2024

How do we estimate the number of transitions and the tempo?

Biologically what are we testing?

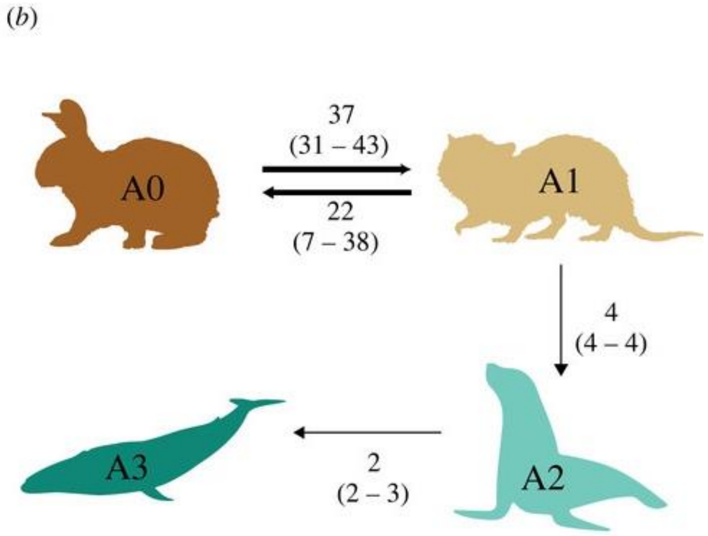
- Dollo’s law of irreversibility: An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those which it has previously lived (Louis Dollo (1893)



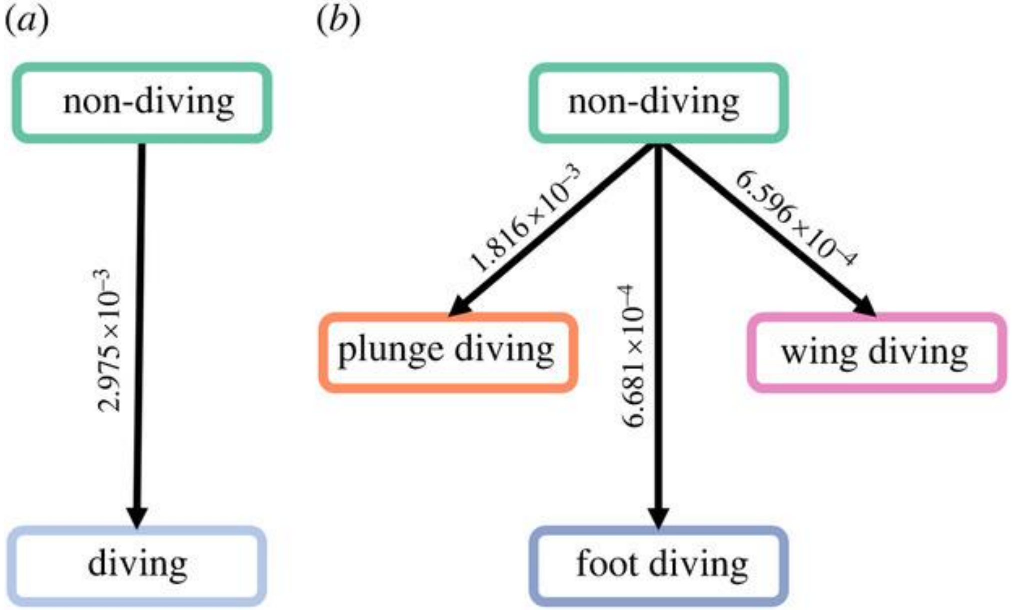
Goldberg and Igić. 2008

Biologically what are we testing?

Dollo’s law of irreversibility: An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those which it has previously lived (Louis Dollo (1893)

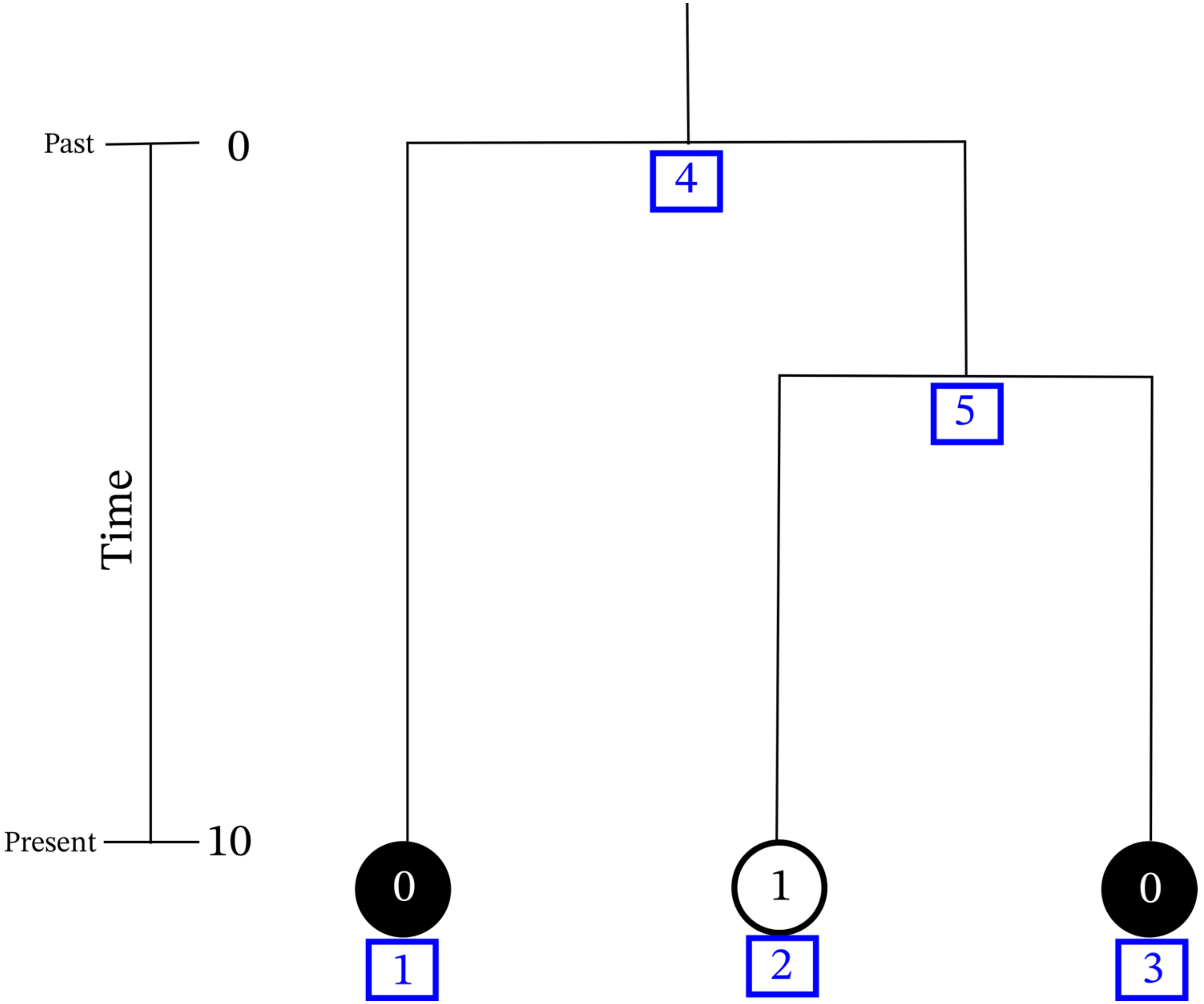


Secondary adaptation to aquatic environments (Farina and Silvestro, 2023)



Waterbirds diving
Tyler and Younger, 2022

Enumerating nodes in the phylogenetic tree



Random variable definition

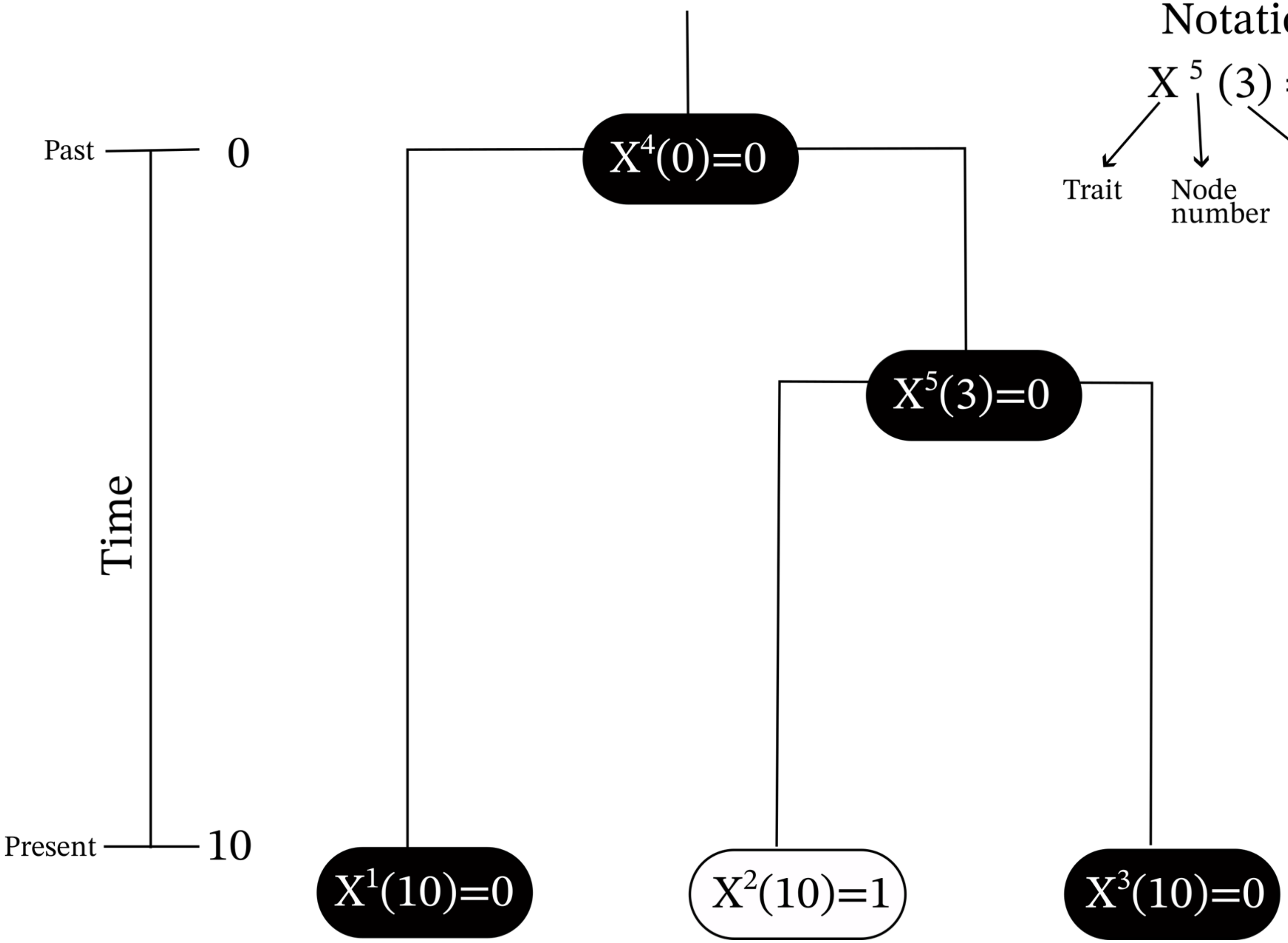
Continuous-Time Markov Chains (CTMC)

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

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

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

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



Notation

$X^5(3) = 0$ or equivalent $X_0^5(3)$

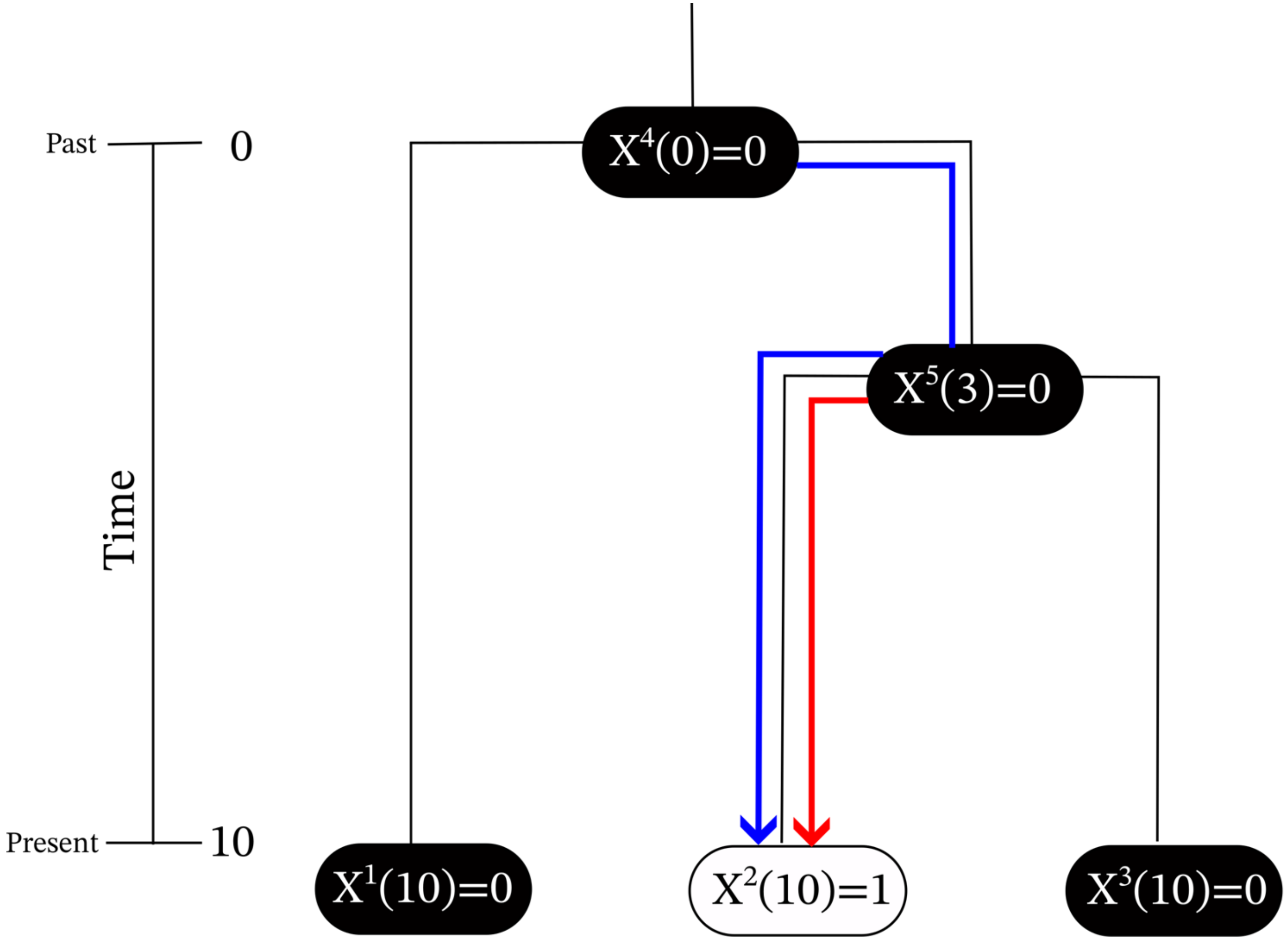
Trait

Node number

Time (mya)

State value

The Markovian in Markovian chains



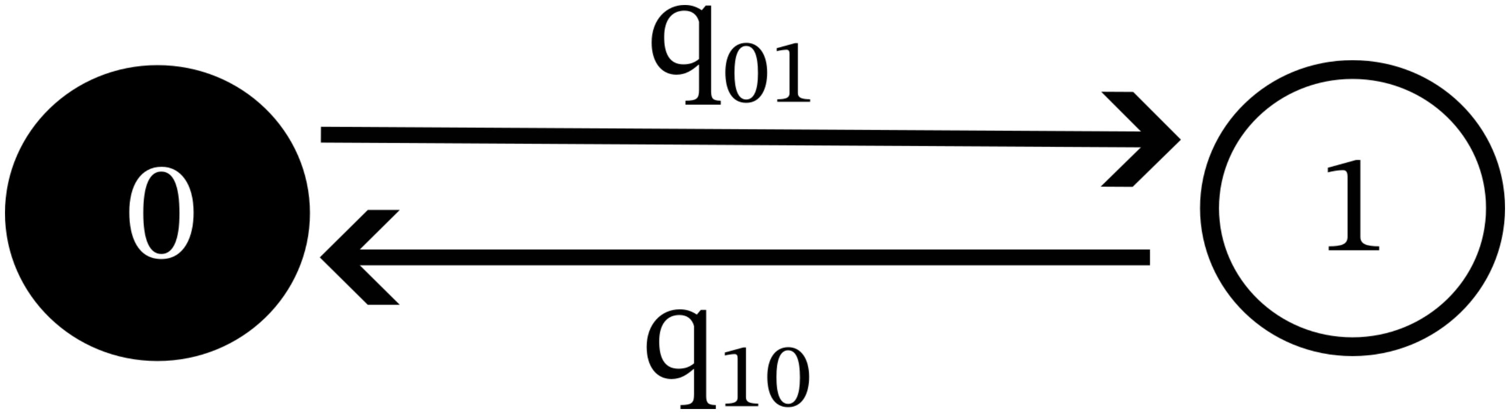
Conditional probabilities

$$P(X^{2}(10)=1|X^{5}(3)=0)=p_{01}(10-3)=p_{01}(7)$$

Transition probabilities- The P matrix

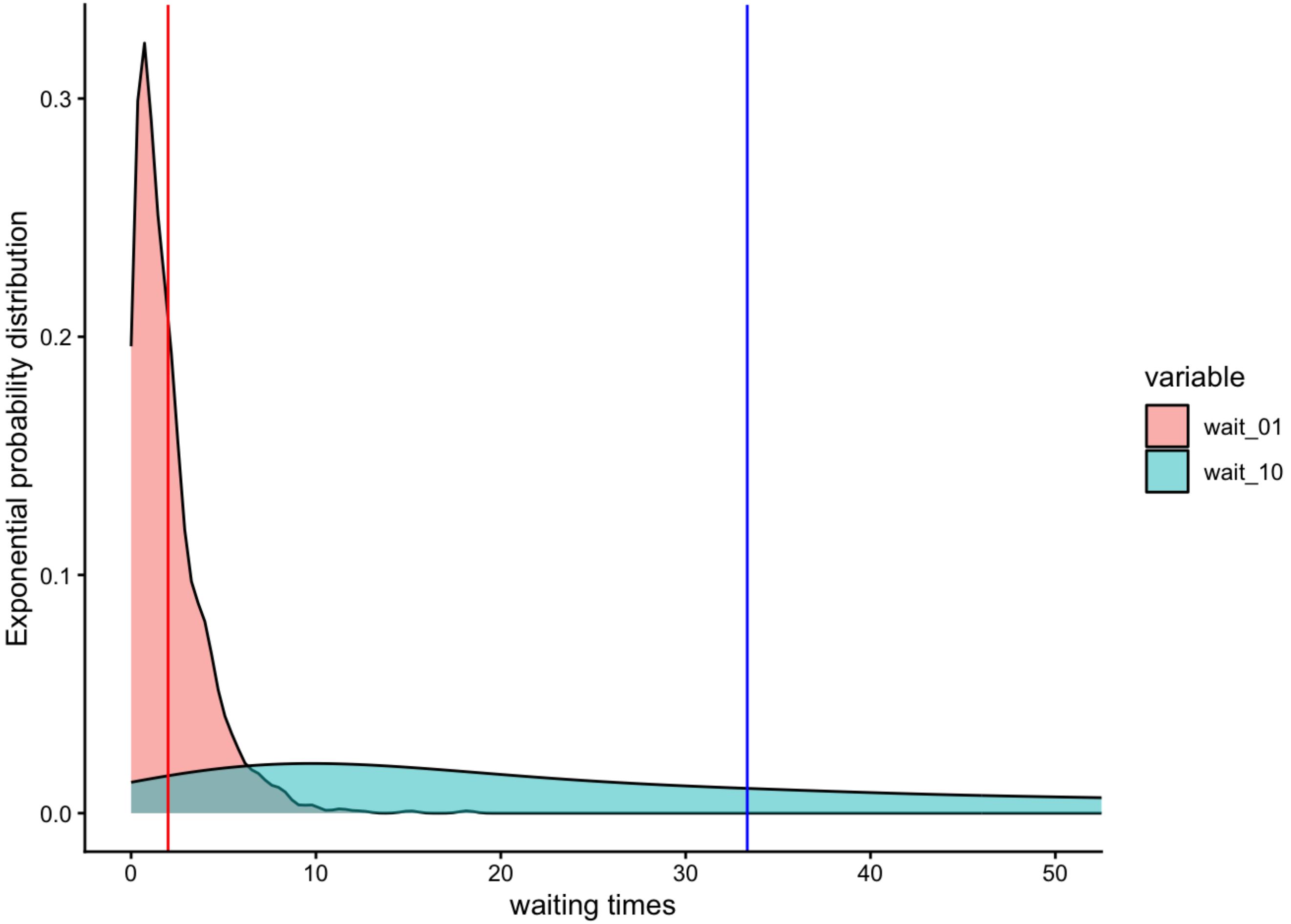
$$P(t) = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix}$$

Transition rates define CTMC

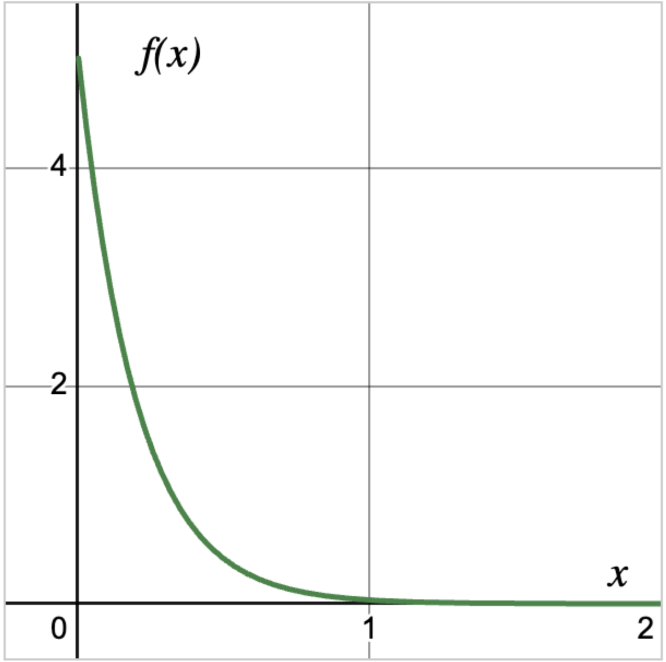


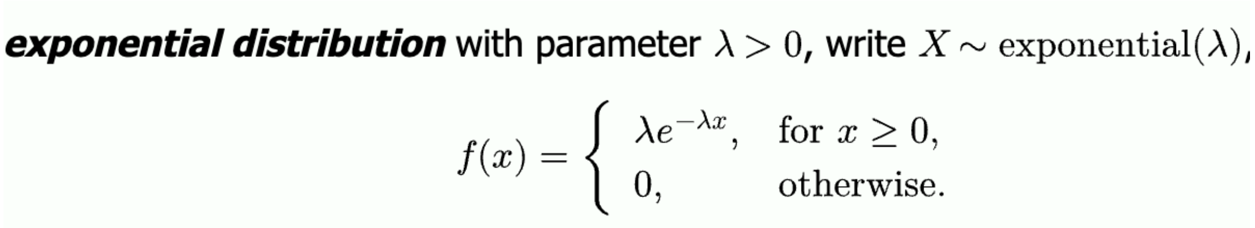
$$\lim_{h \to 0} P(X(t+h) = j | X(t) = j) = \lim_{h \to 0} p_{ij}(h) = q_{ij}$$

1/(transition rates)= average waiting time



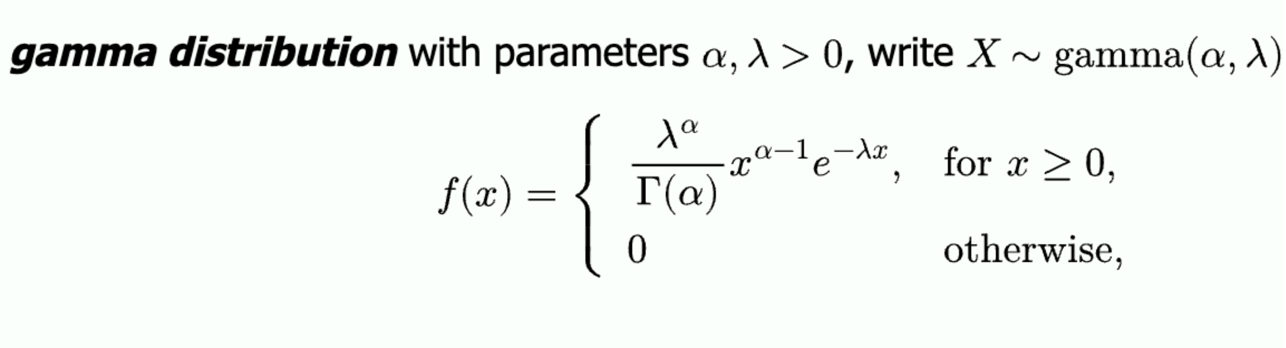
Exponential is a Gamma distribution





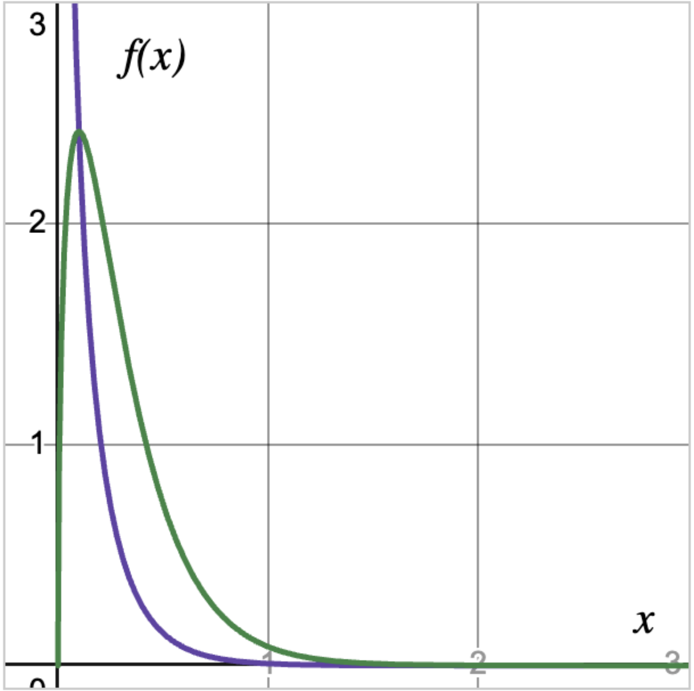
exponential distribution with parameter $\lambda > 0$, write $X \sim \mathrm{exponential}(\lambda)$,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{for } x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$



gamma distribution with parameters $\alpha, \lambda > 0$, write $X \sim \mathrm{gamma}(\alpha, \lambda)$

$$f(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{for } x \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$



The Q to the Matrix

$$Q = \begin{pmatrix} -q_{01} & q_{01} \\ q_{10} & -q_{10} \end{pmatrix}$$

P-matrix and the Q-matrix

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

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

Rev to the Bayes
Introduction to RevBayes

Very brief Bayesian Statistics through probabilistic modeling

Rosana Zenil-Ferguson
Assistant Professor. University of Kentucky
SSB Breakthrough Meeting 2026.

Rev to the Bayes

BASICS OF BAYESIAN STATISTICS

DRAWING PROBABILISTIC MODELS

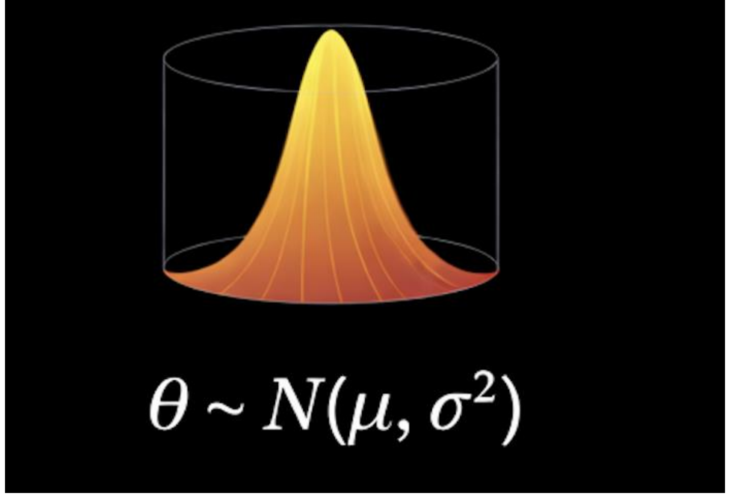
Bayesian Inference: Two important concepts

1. Probabilities in Bayesian statistics are based on previous knowledge or beliefs about uncertainty.

E.g. What is the probability that an asteroid hit the Earth today?

2. Every parameter that we want to infer is unknown and uncertain.

Parameters are random variables with probability distributions



$$\theta \sim N(\mu, \sigma^2)$$

Cumberland Falls State Park, KY

Thalictrum pubescences

Thalictrum thalictroides

Thalictrum
multiple changes between
insect and wind pollination

Verónica di Stilio
University of Washington

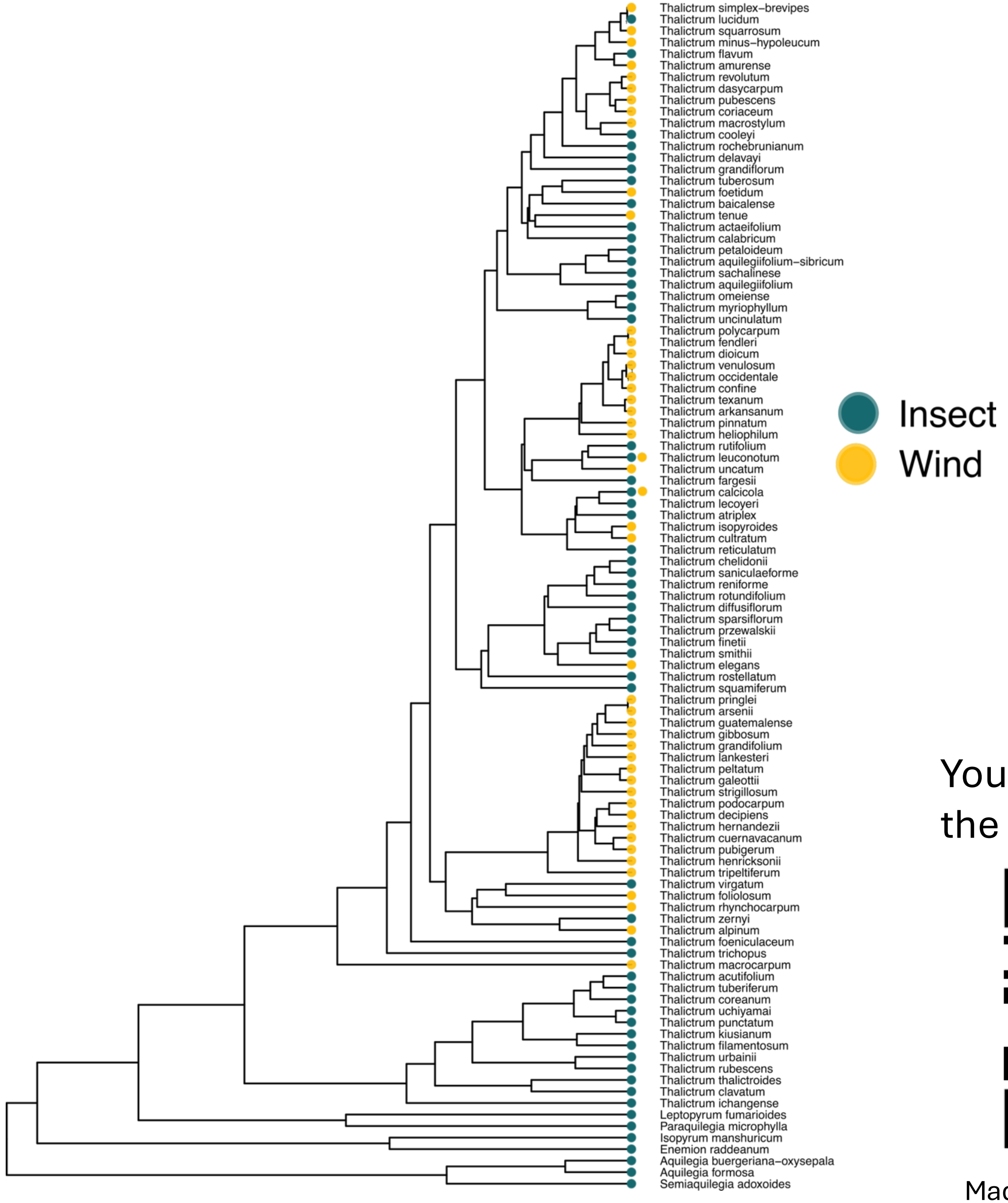
Fotos: Tennessee-Kentucky Plant Atlas, iNaturalist, and RZF

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

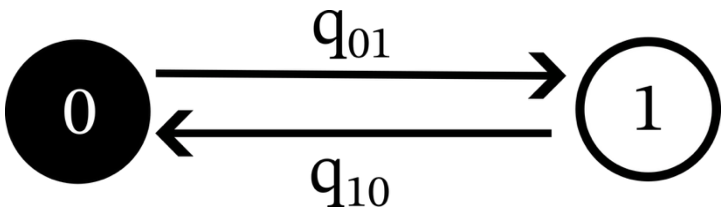
You cannot applied discrete models if you are in the ”Felsenstein worst case scenario”

Maddison and Fitzjohn, 2014

Uyeda, Zenil-Ferguson, and Pennell, 2018

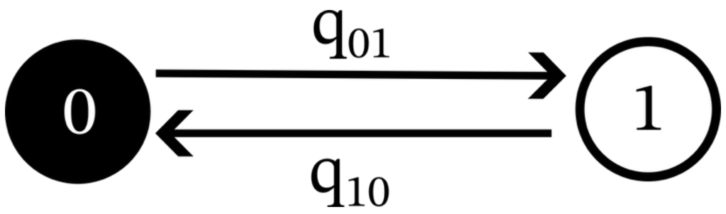
Unknown and uncertain parameters in our Mkn models

1. Transition rates

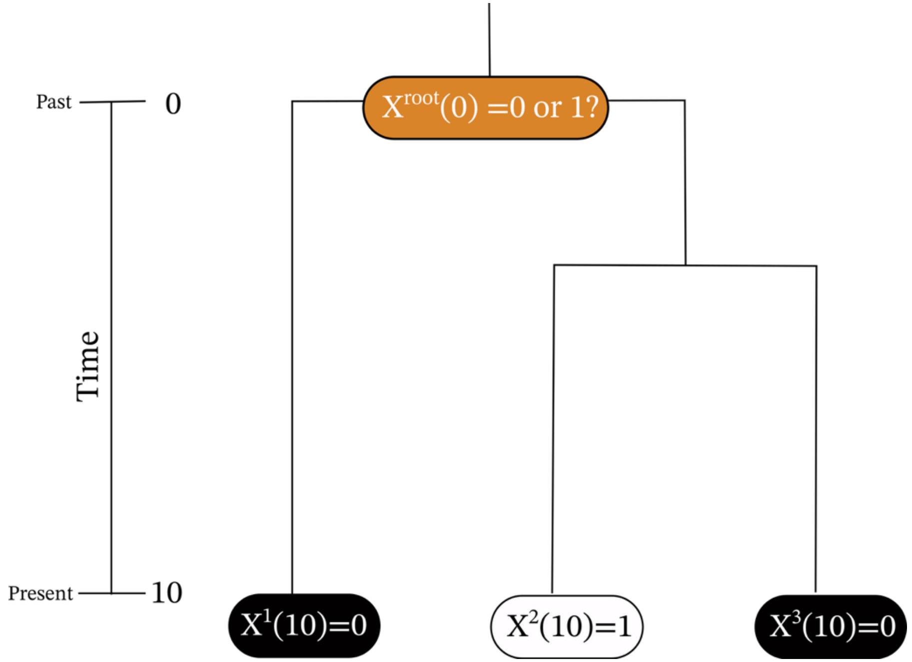


Unknown and uncertain parameters in our Mkn models

1. Transition rates



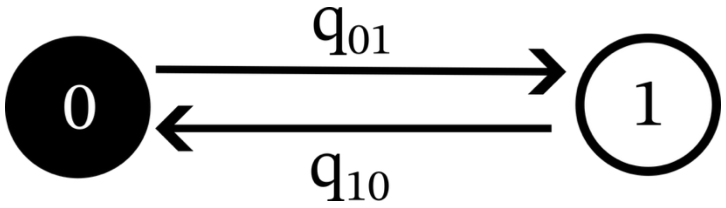
2. Frequencies of states in the root

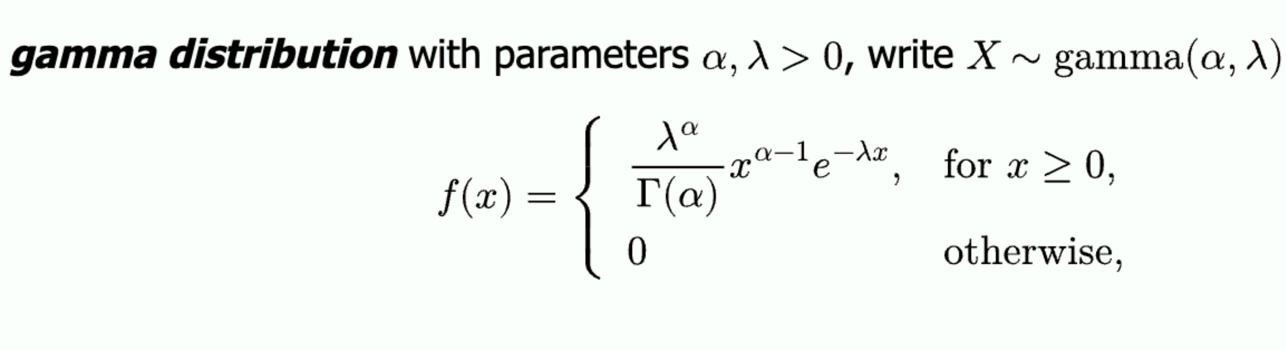


Graphical probabilistic models in RevBayes and their computer notation

□	a) Constant node	**<-**	constant variable
○	b) Stochastic node	**~**	stochastic variable
◌	c) Deterministic node	**:=**	deterministic variable
●	d) Clamped node (observed)	**node.clamp(data)**	observed value of stochastic variable
⬚	e) Plate	**for (i in 1:N) { ...}**	plates (multiple connected nodes)
		=	for fixed operations (not part of graphical model)

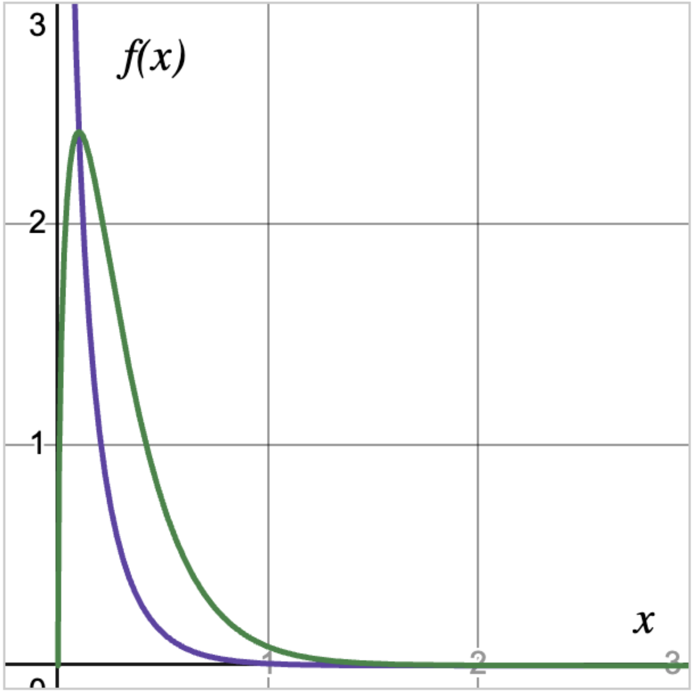
Transition rate prior distributions- Gamma



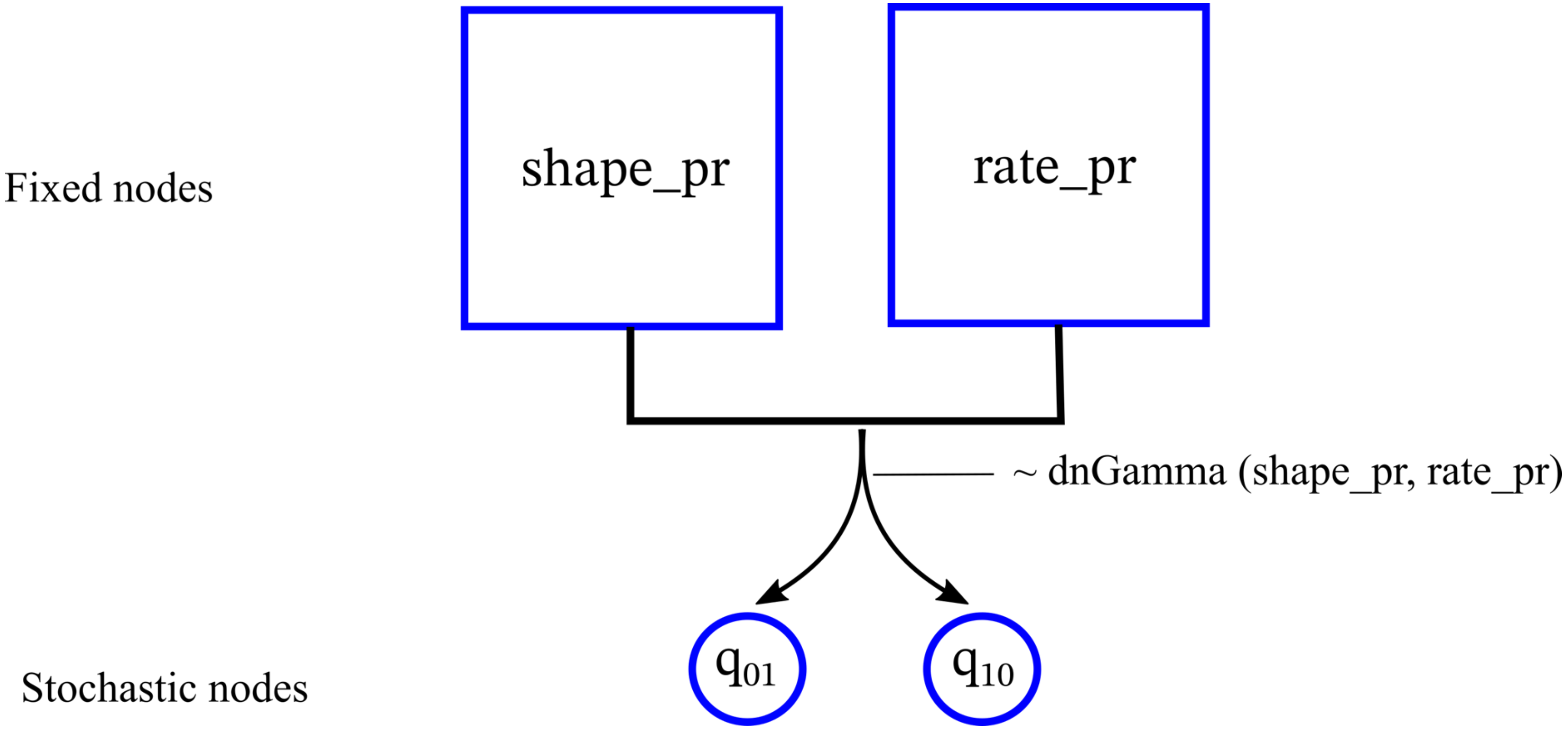


gamma distribution with parameters $\alpha, \lambda > 0$, write $X \sim \mathrm{gamma}(\alpha, \lambda)$

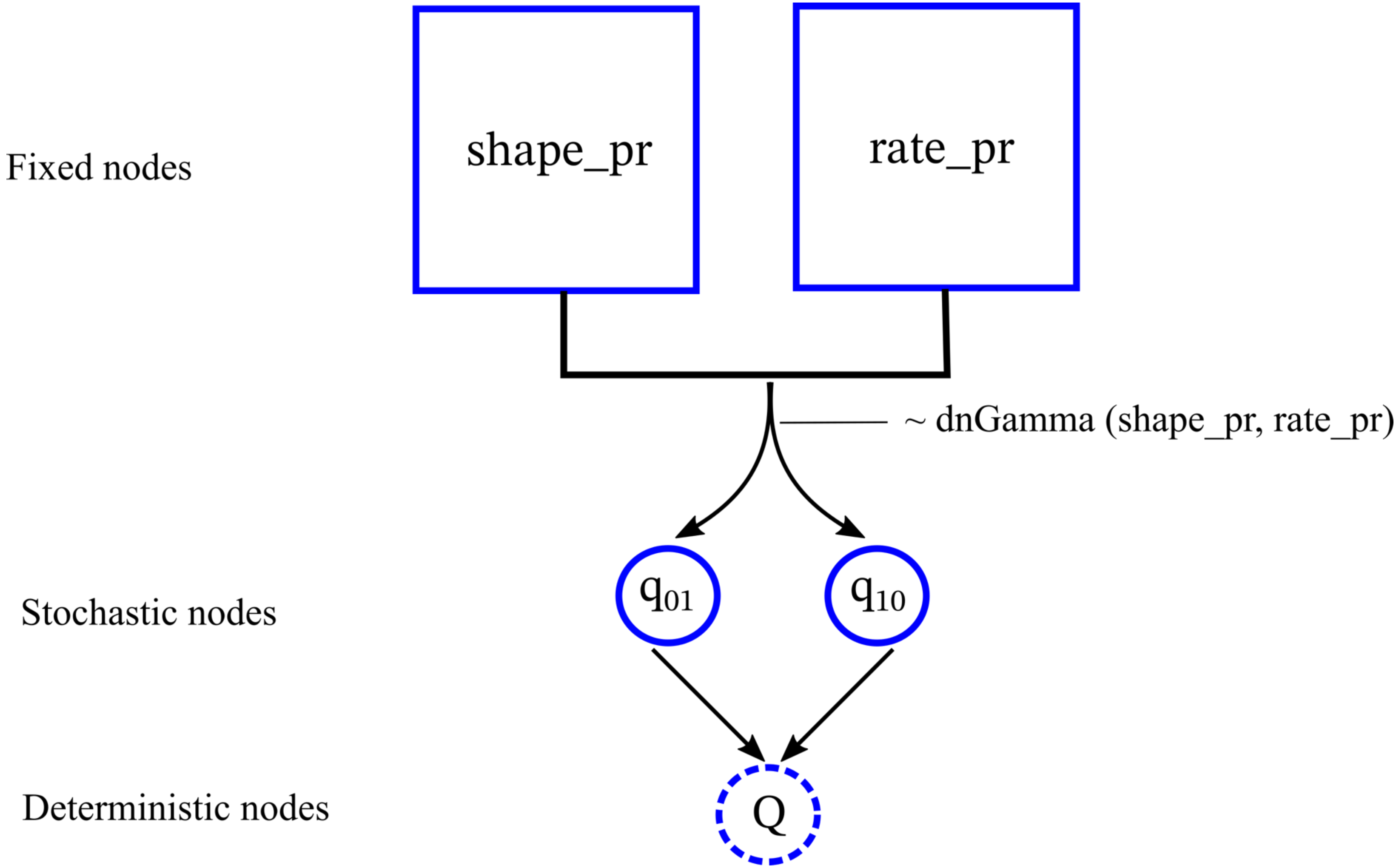
$$f(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{for } x \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$



Transition rate prior distributions



Transition rate prior distributions

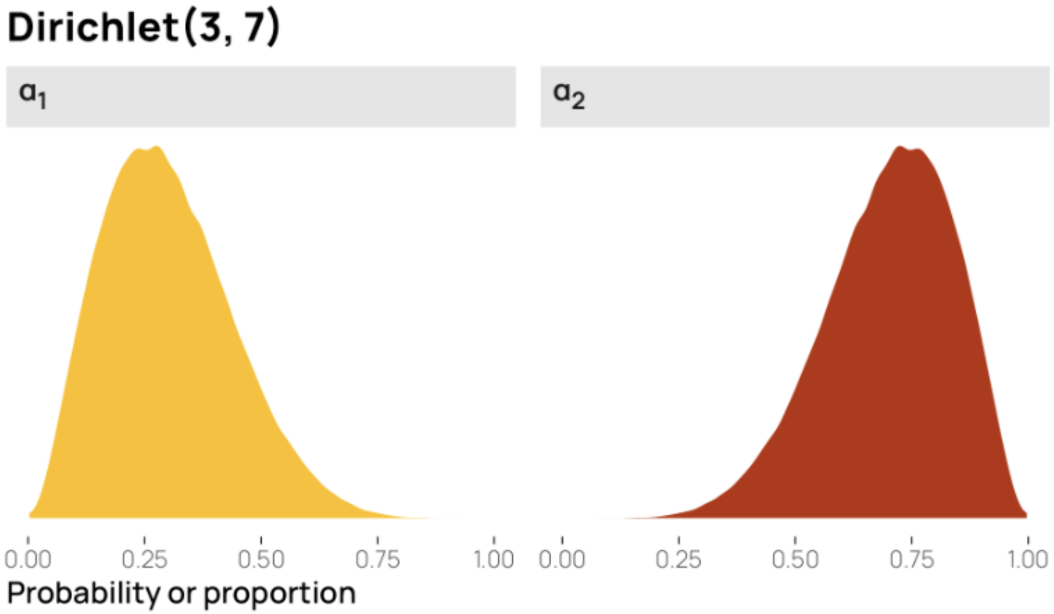


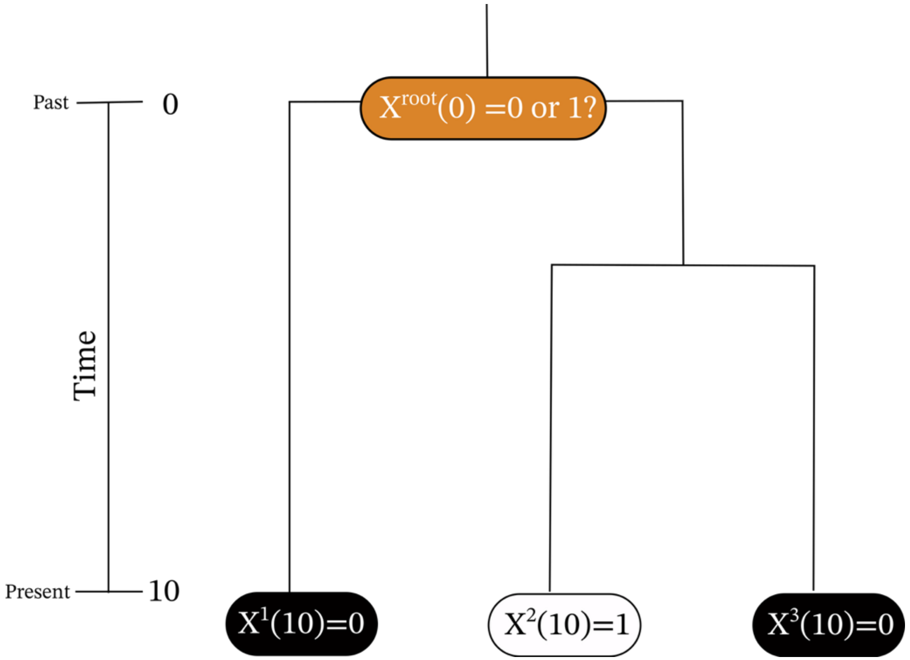
Fixed nodes

Stochastic nodes

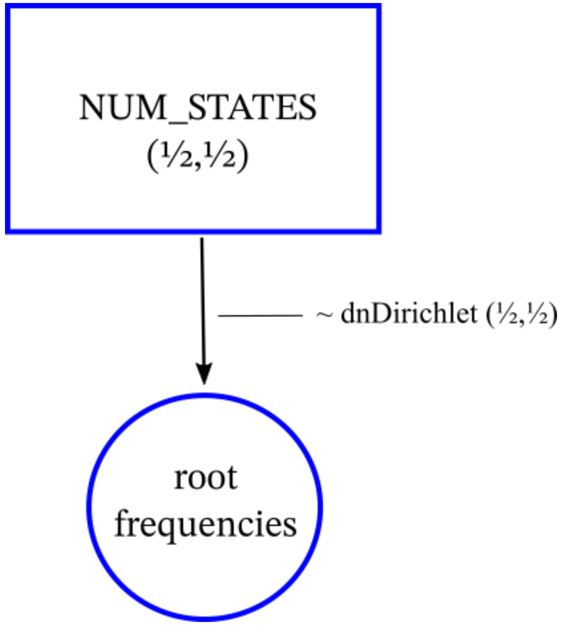
Deterministic nodes

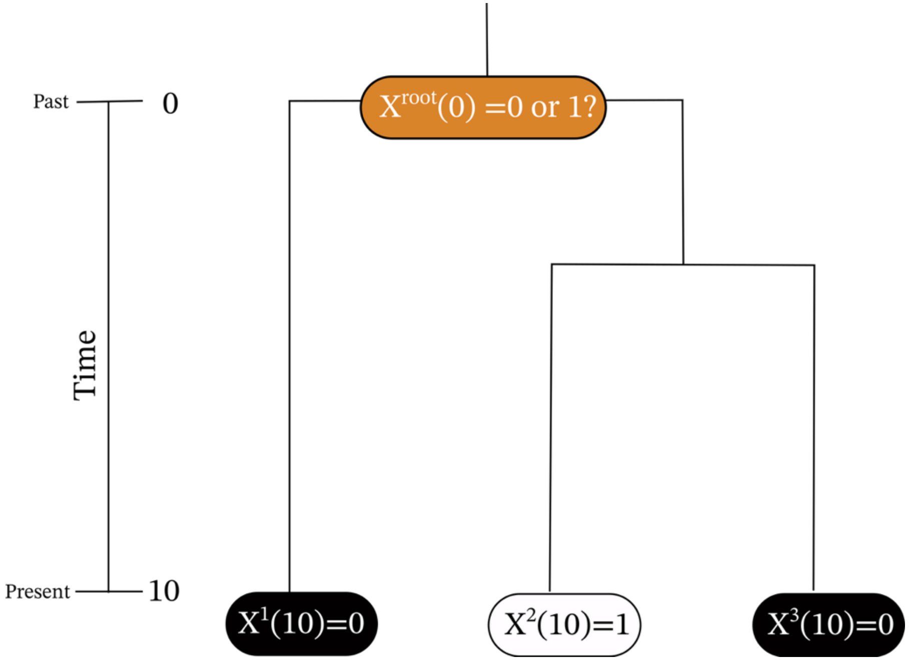
Frequencies of the root



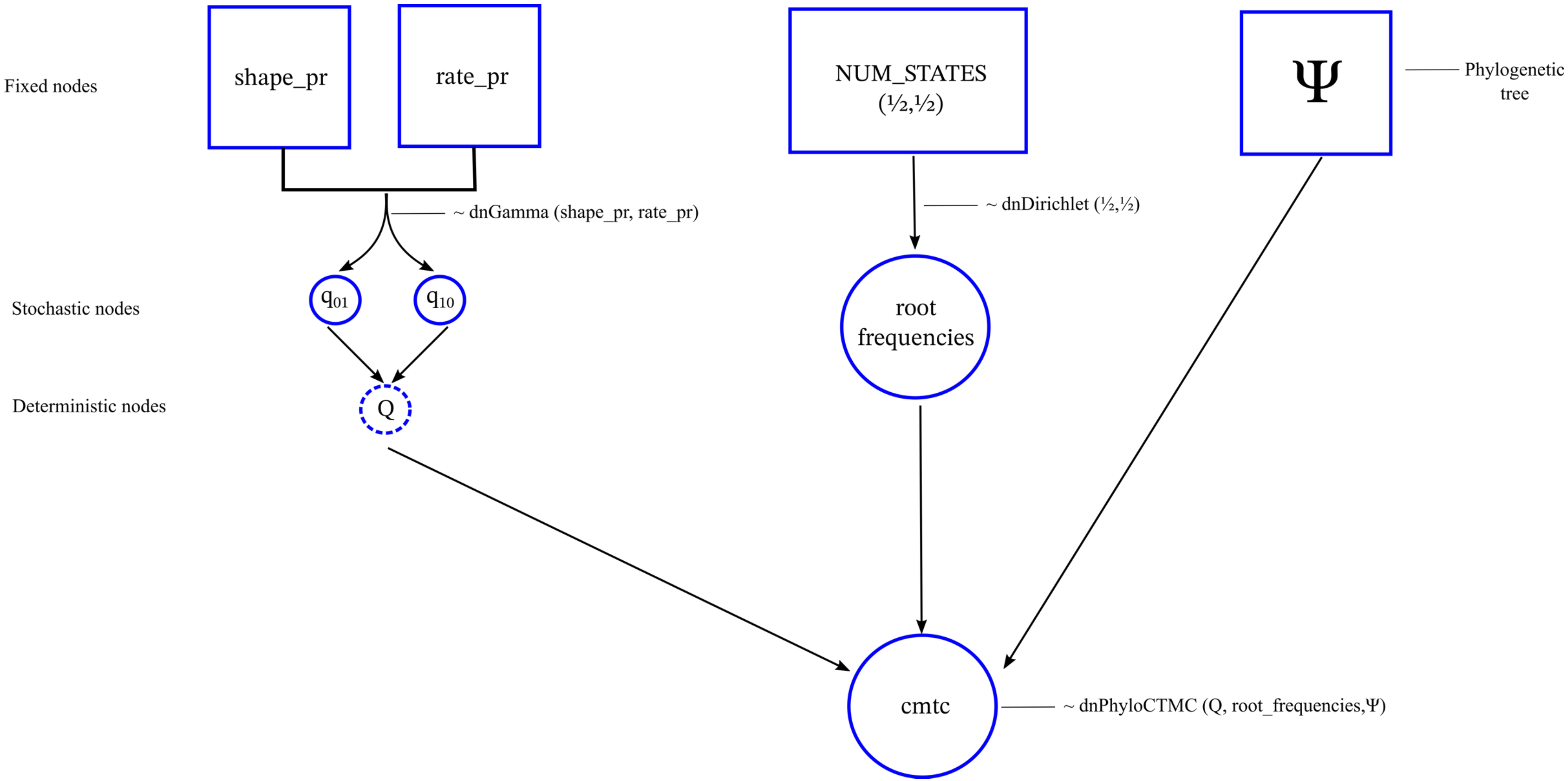


Frequencies of the root





Mk2- Continuous-Time Markov Chain with two states in RevBayes



Fixed nodes

shape_pr

rate_pr

NUM_STATES (½,½)

Ψ

Phylogenetic tree

~ dnGamma (shape_pr, rate_pr)

~ dnDirichlet (½,½)

Stochastic nodes

q_{01}

q_{10}

root frequencies

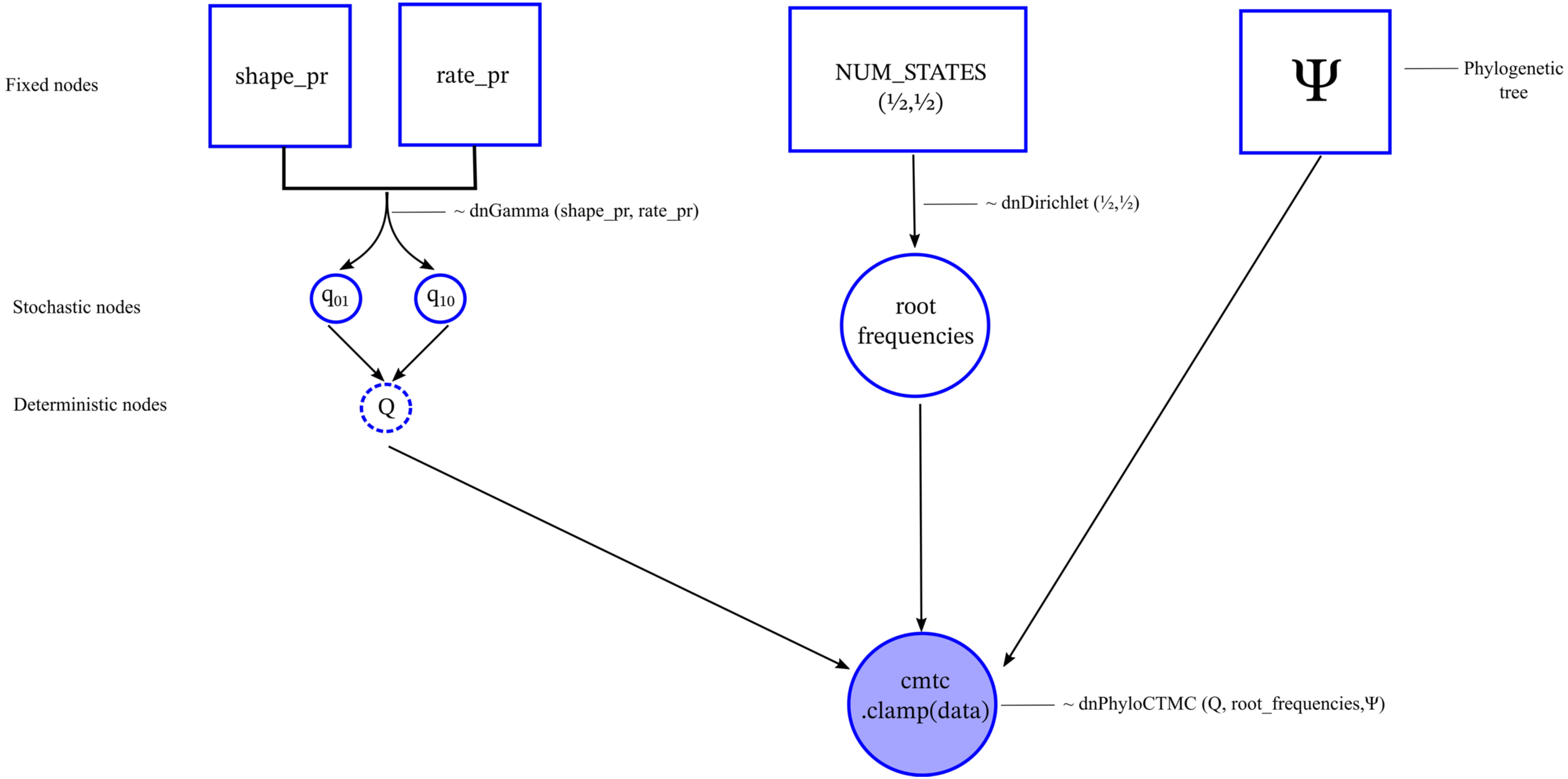
Deterministic nodes

Q

cmtc

~ dnPhyloCTMC (Q, root_frequencies,Ψ)

One more step- Add data (a.k.a. Clamping)



Fixed nodes

shape_pr

rate_pr

NUM_STATES (½,½)

Ψ

Phylogenetic tree

~ dnGamma (shape_pr, rate_pr)

~ dnDirichlet (½,½)

Stochastic nodes

q_{01}

q_{10}

root frequencies

Deterministic nodes

Q

cmtc .clamp(data)

~ dnPhyloCTMC (Q, root_frequencies,Ψ)

What are moves?

moves.append( )

Reflect proposals for searching randomly values of the parameters for my graphical model.

A parameter value gets proposed its probability evaluated if it is very good with high chances it gets accepted. Otherwise is rejected and we keep the previous value.

Watch Dr. Paul Lewis

Bayesian inference and MCMC

Lecture

What are monitors?

monitors.append( ) Store the full inference.

In Bayesian Statistics, the object of inference is the **posterior distribution**. We care about **credible intervals** (95% probability).

In likelihood inference (frequentist) often we care about **maximum likelihood estimates** and hopefully **confidence-likelihood intervals**.

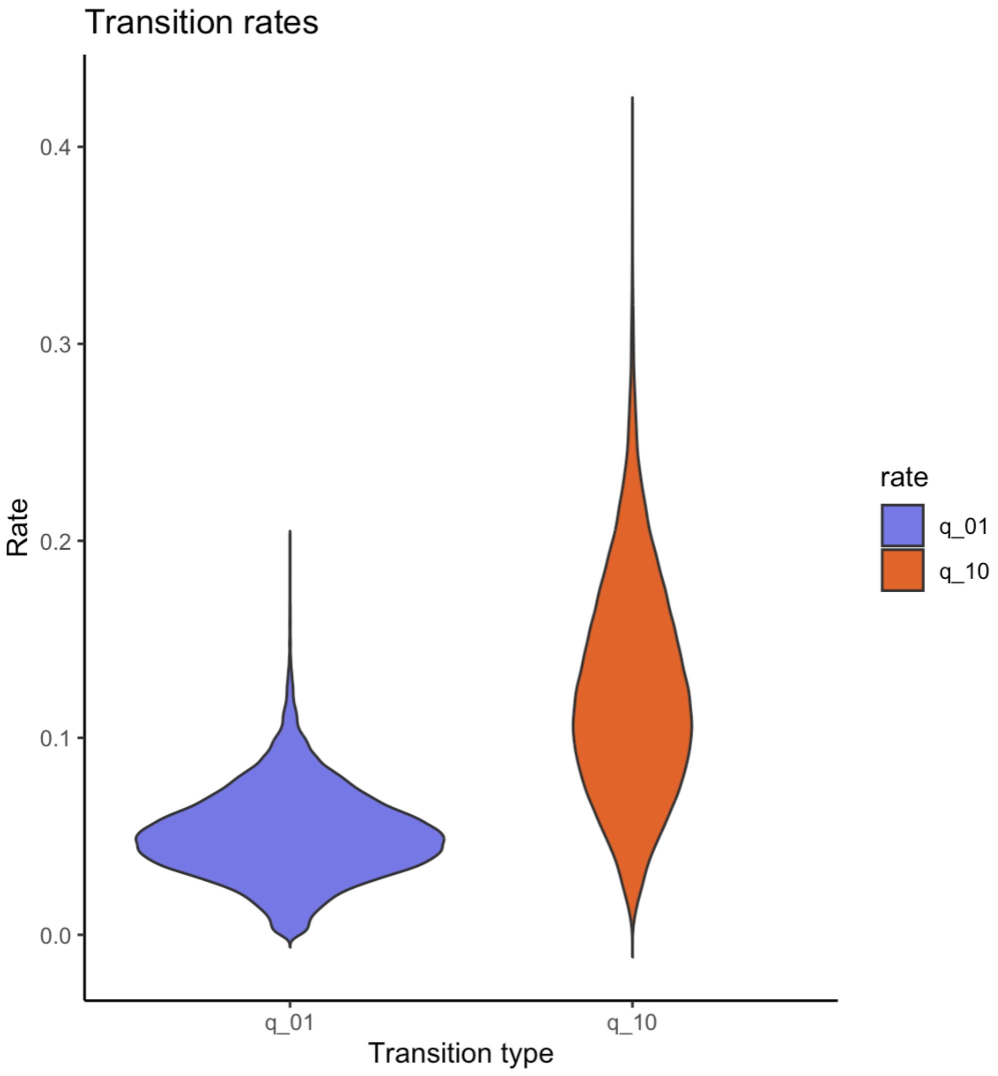
Credible Interval
(Bayesian)

Given our data and prior, there's a 95% probability the true value lies here

Confidence Interval
(Frequentist)

In 95 of 100 repeated experiments, the interval will contain the true value

Posterior distribution: Transition rates

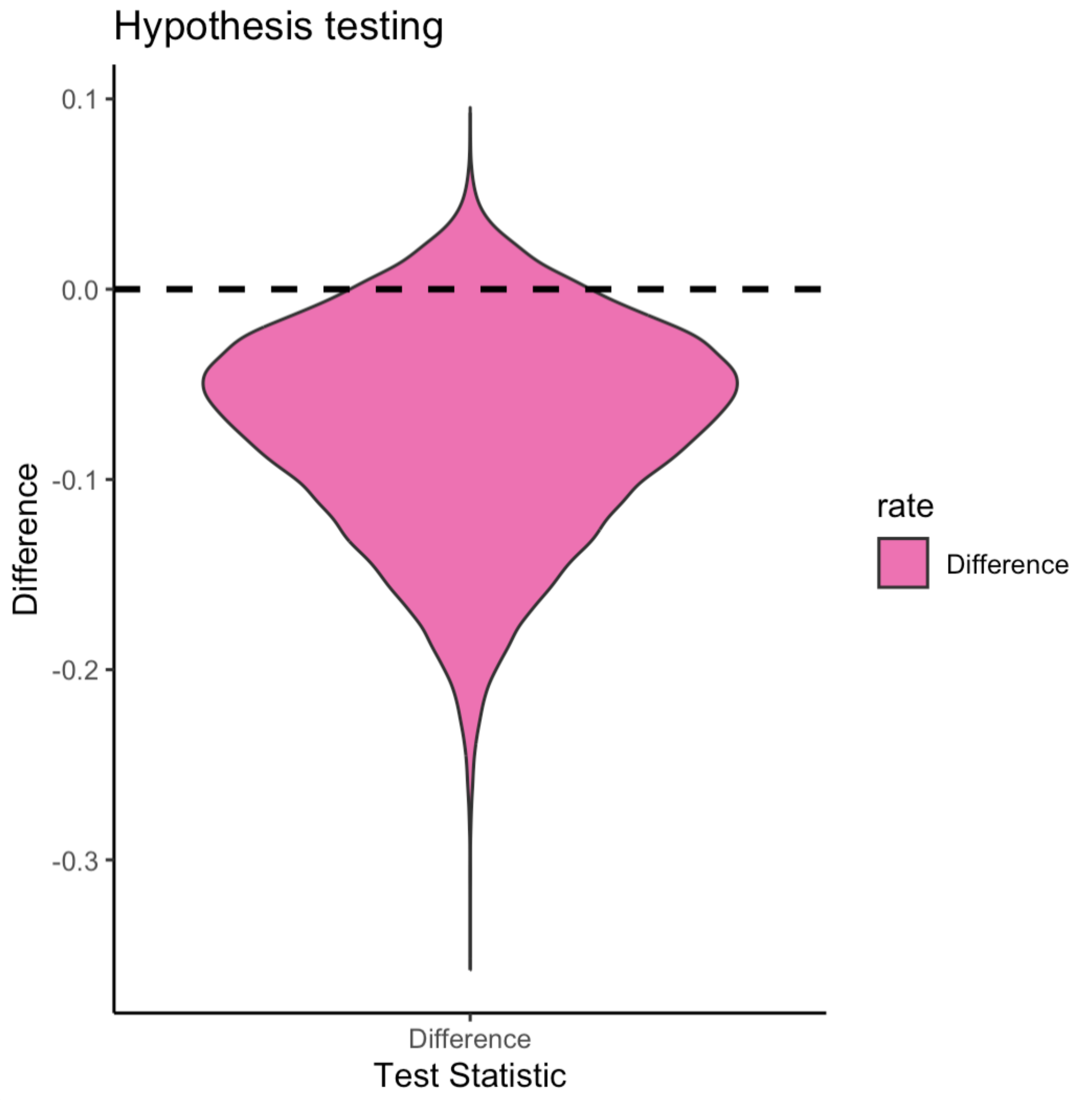


Hypothesis testing- Bayesian framework

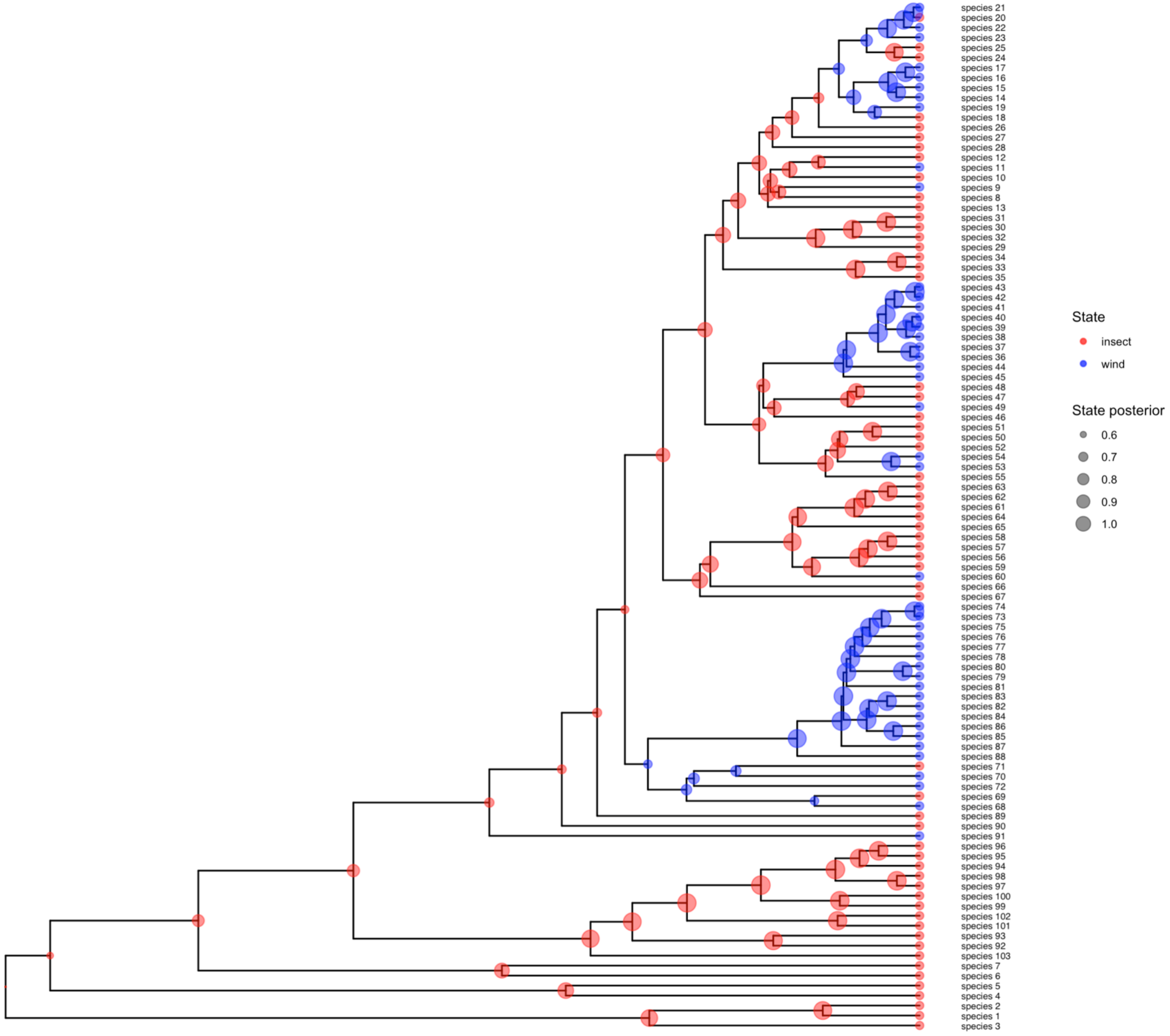
Null hypothesis $H_0: q_{01} = q_{10}$

Test statistic $D = q_{01} - q_{10}$

HPD (D)=



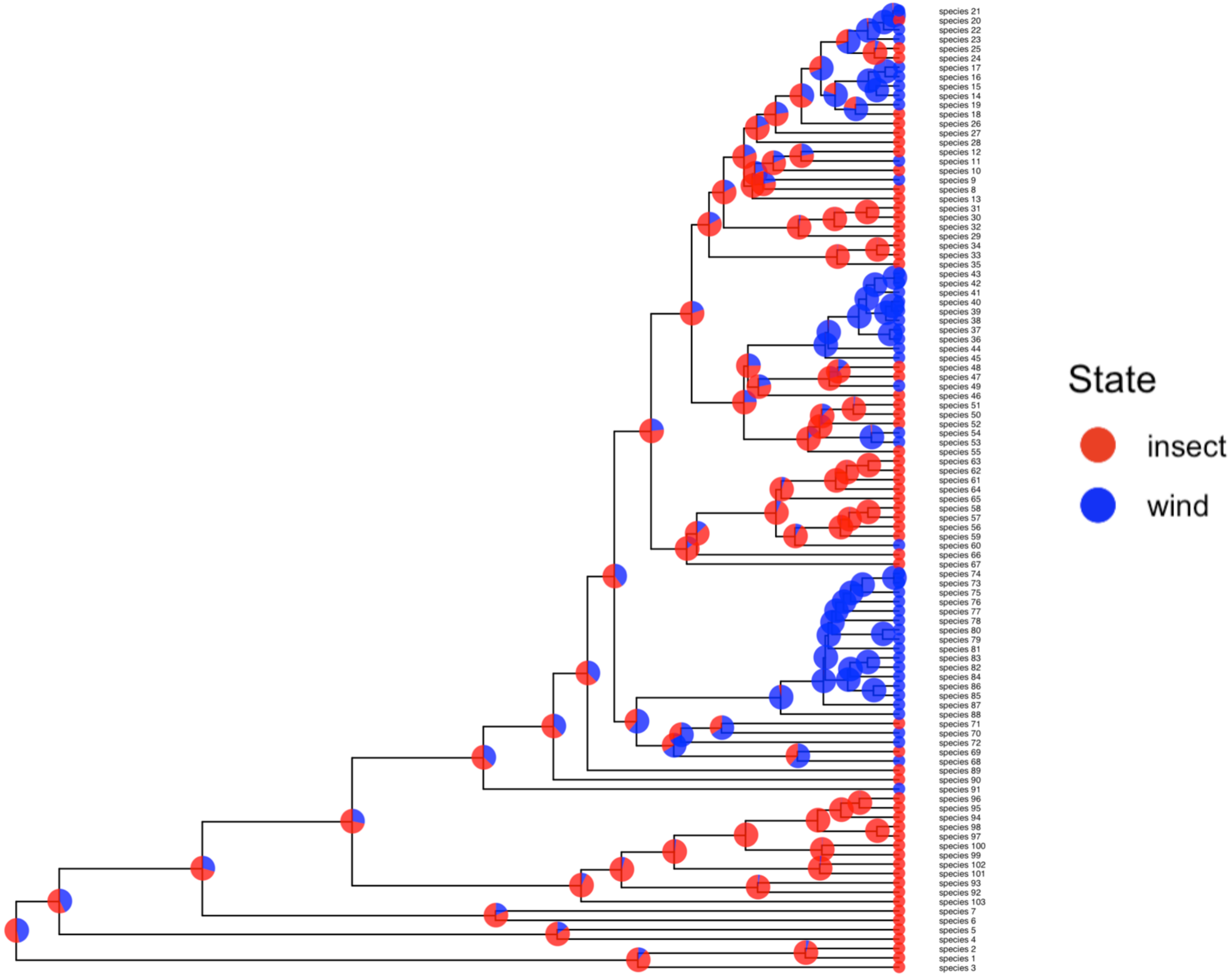
Ancestral state reconstructions



- Bayesian: Posterior distribution at each node.
- Marginal: Focus on one node at a time summing (integrating) over the rest of the nodes and maximizing the probability of the given node
- Global: Find the combination of node values that maximizes the total probability

Can you talk about the number of transitions with an ancestral reconstruction?

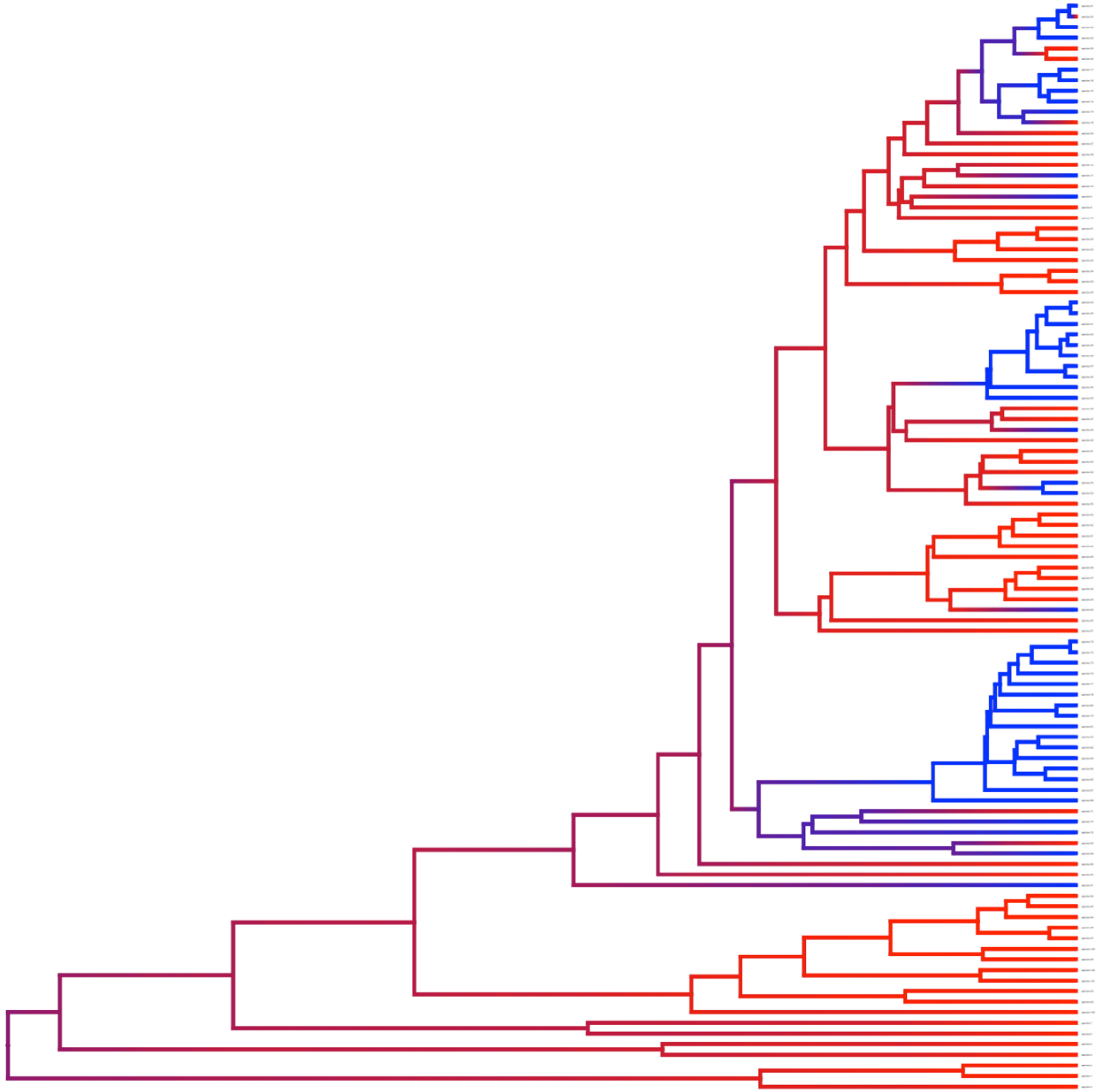
Ancestral state reconstruction (marginal)



This is a posterior distribution pie

- **Warning**: Pies are often used by other non Bayesian software to do marginal ancestral state reconstructions. Those packages use the **marginal likelihood** function

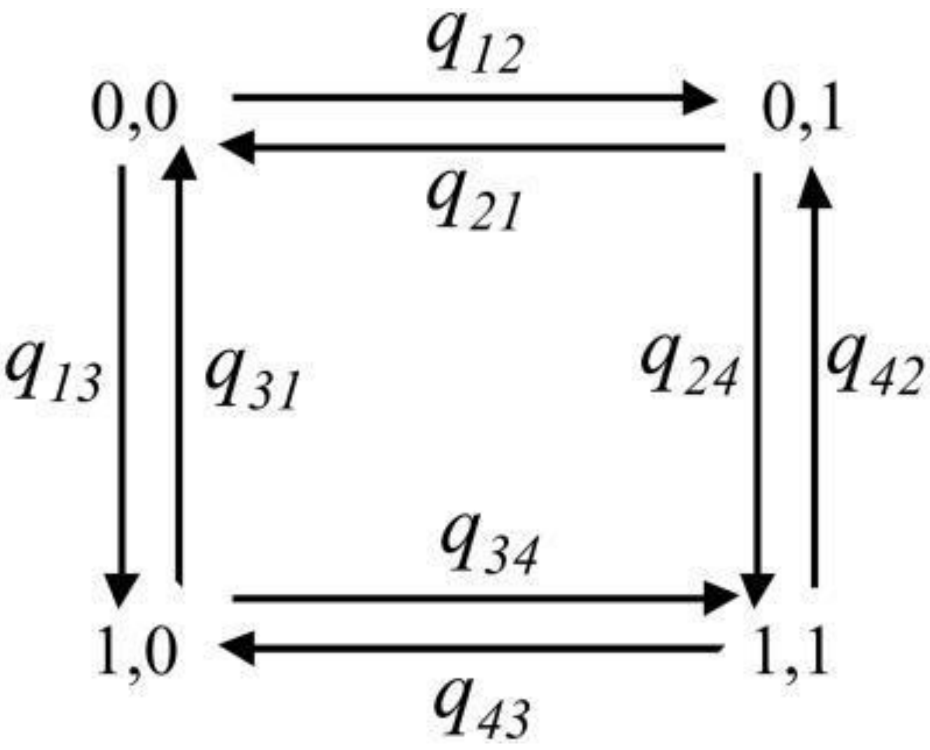
Stochastic mapping



- 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?

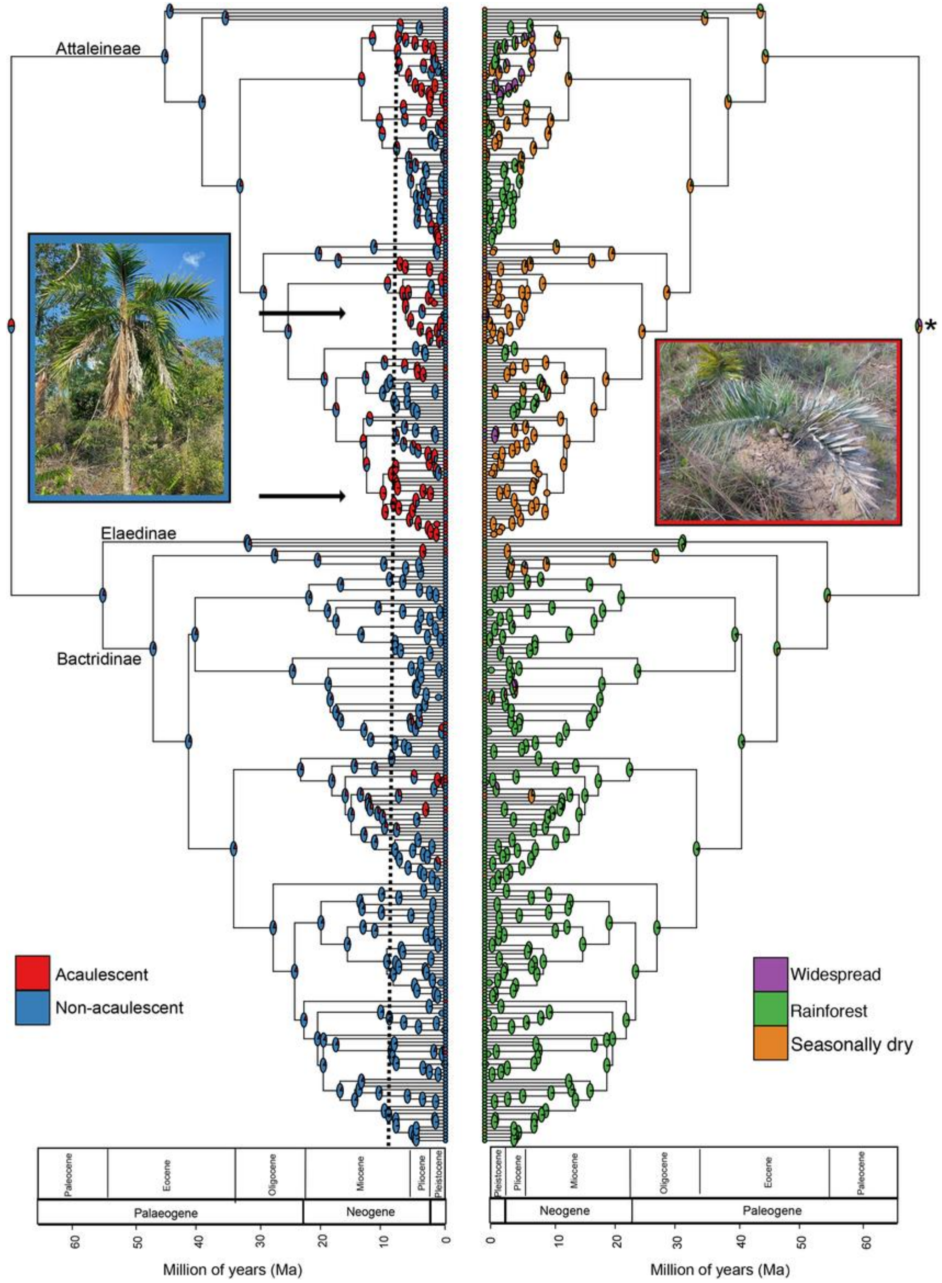
The Q-matrix as the center of the universe (for discrete traits)



Pagel's correlation test

Pagel 1994

Acaulescence and dry environments
Cassia Silva et al. Ecography. 2022

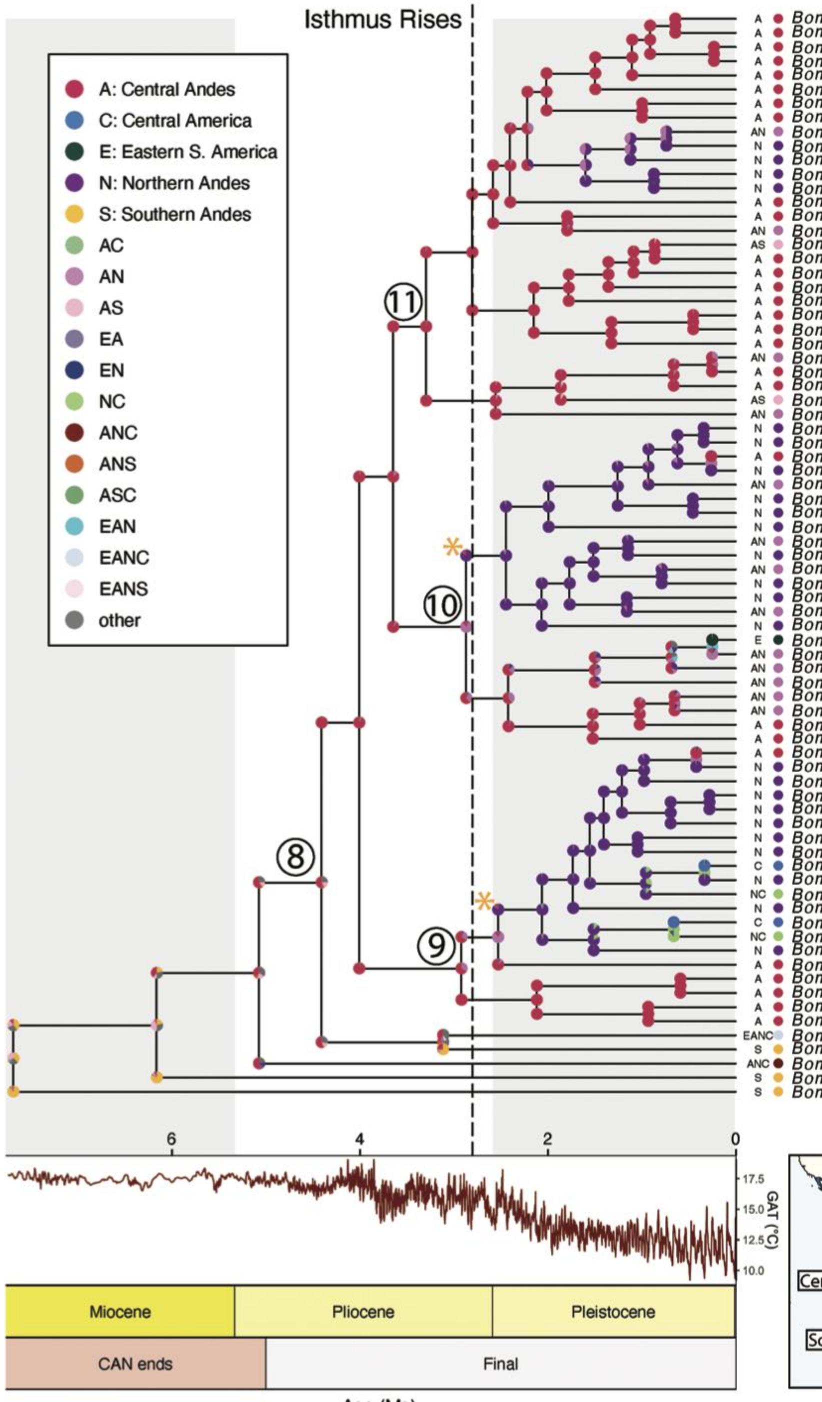


Bomarea Biogeography
Tribble et al. 2023

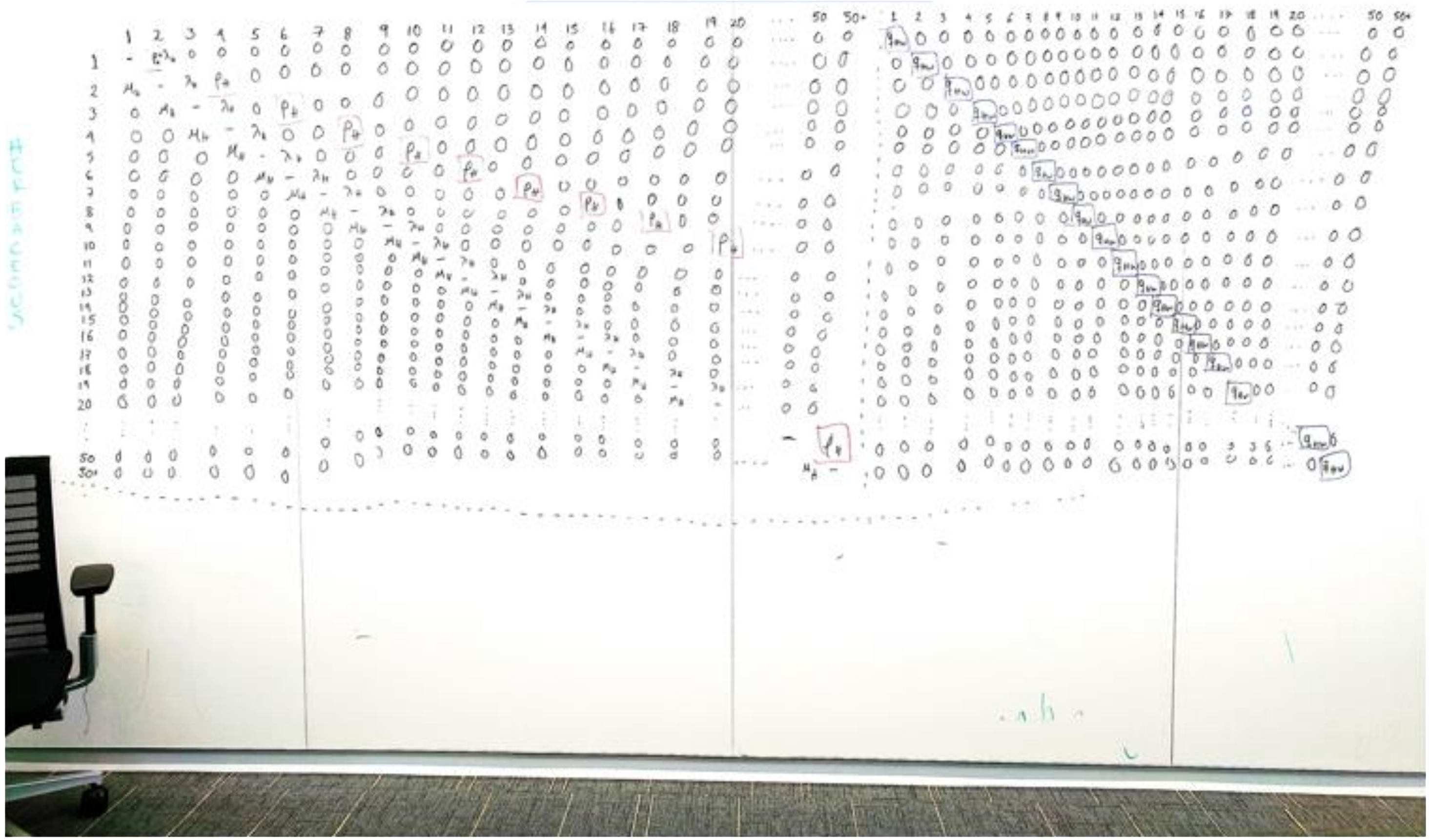
$$
\mathbf{Q} = \begin{array}{c|cccccccc}
 & \emptyset & A & B & C & AB & AC & BC & ABC \\
\hline
\emptyset & - & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
A & e_A & - & 0 & 0 & d_{AB} & d_{AC} & 0 & 0 \\
B & e_B & 0 & - & 0 & d_{BA} & 0 & d_{BC} & 0 \\
C & e_C & 0 & 0 & - & 0 & d_{CA} & d_{CB} & 0 \\
AB & 0 & e_B & e_A & 0 & - & 0 & 0 & d_{AC} + d_{BC} \\
AC & 0 & e_C & 0 & e_A & 0 & - & 0 & d_{AB} + d_{CB} \\
BC & 0 & 0 & e_C & e_B & 0 & 0 & - & d_{BA} + d_{CA} \\
ABC & 0 & 0 & 0 & 0 & e_C & e_B & e_A & -
\end{array}
$$

Dispersal-Extinction Cladogenesis model (DEC)

Ree et al. 2005; Ree and Smith (2008)



$$Q_{ij} = \begin{cases} \gamma_a & j = i+1, \\ \delta_a & j = i-1, \\ \rho_a & j = 2i, \\ \eta_a & j = 1.5i, \\ 0 & \text{otherwise,} \end{cases}$$



ChromEvol (Mayrose et al. 2010)
BiChrom (Zenil-Ferguson et al. 2017)