
Lecture 2.2

Bayesian Phylogenetics 1

The Bayesian framework

Bayesian phylogenetic analysis

- Bayesian phylogenetic analysis was developed in the mid 1990s
- Now one of the most widely used methods

MrBayes



BEAST 1

RevBayes



BEAST 2

Bayesian phylogenetic analysis

- Bayes's theorem (1763)
- Reverend Thomas Bayes



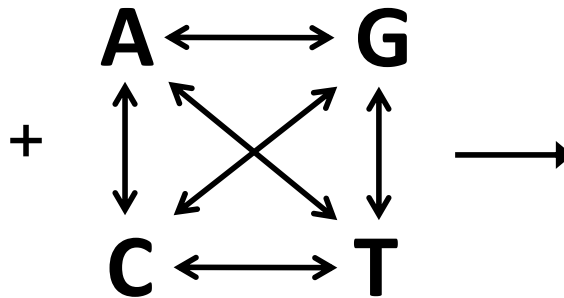
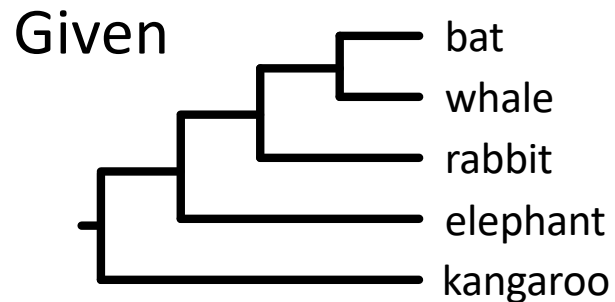
*Image probably
not of Thomas Bayes*

In phylogenetics

$$\Pr(\theta \mid D) \propto \Pr(\theta) \Pr(D \mid \theta)$$

Bayesian phylogenetic analysis

Maximum likelihood

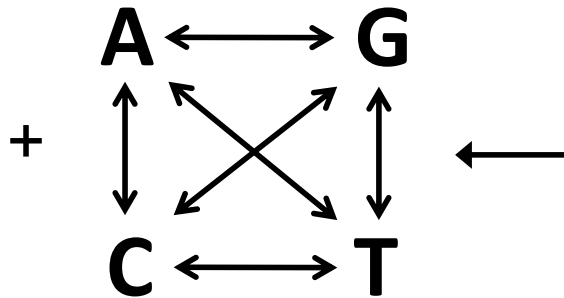
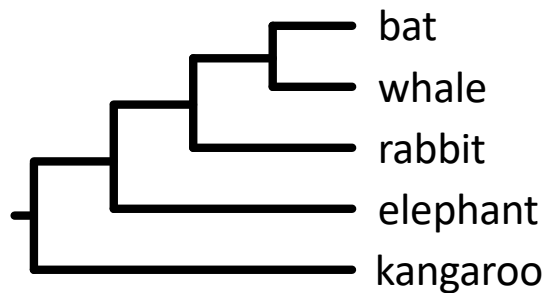


Probability of?

bat	CCGTTAGTAACT
whale	CCGTTAGTAACT
rabbit	CCGATAGTTACT
elephant	TCGTTAGTTACC
kangaroo	TCATTGGTTACT

Bayesian inference

Probability of?



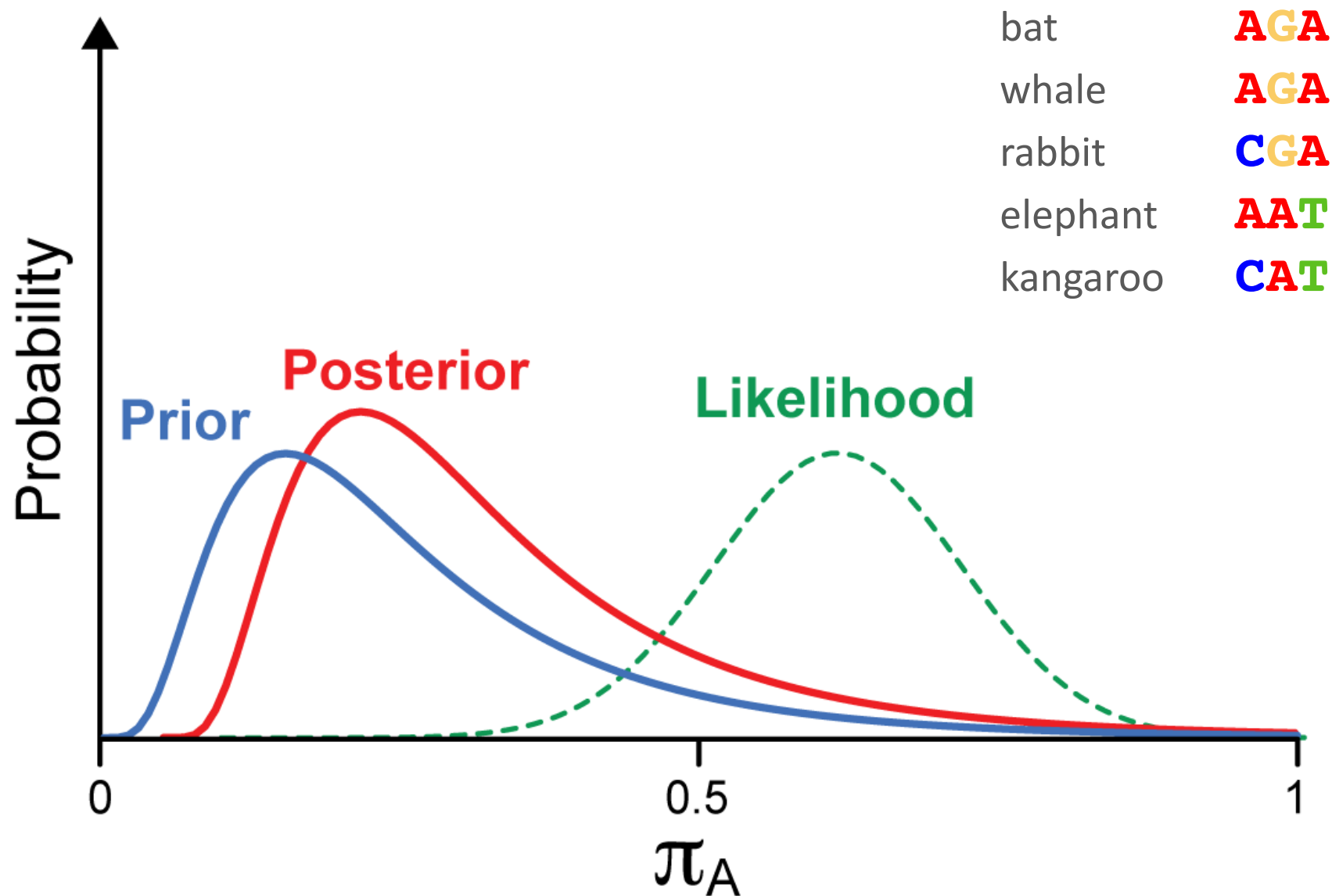
Given

bat	CCGTTAGTAACT
whale	CCGTTAGTAACT
rabbit	CCGATAGTTACT
elephant	TCGTTAGTTACC
kangaroo	TCATTGGTTACT

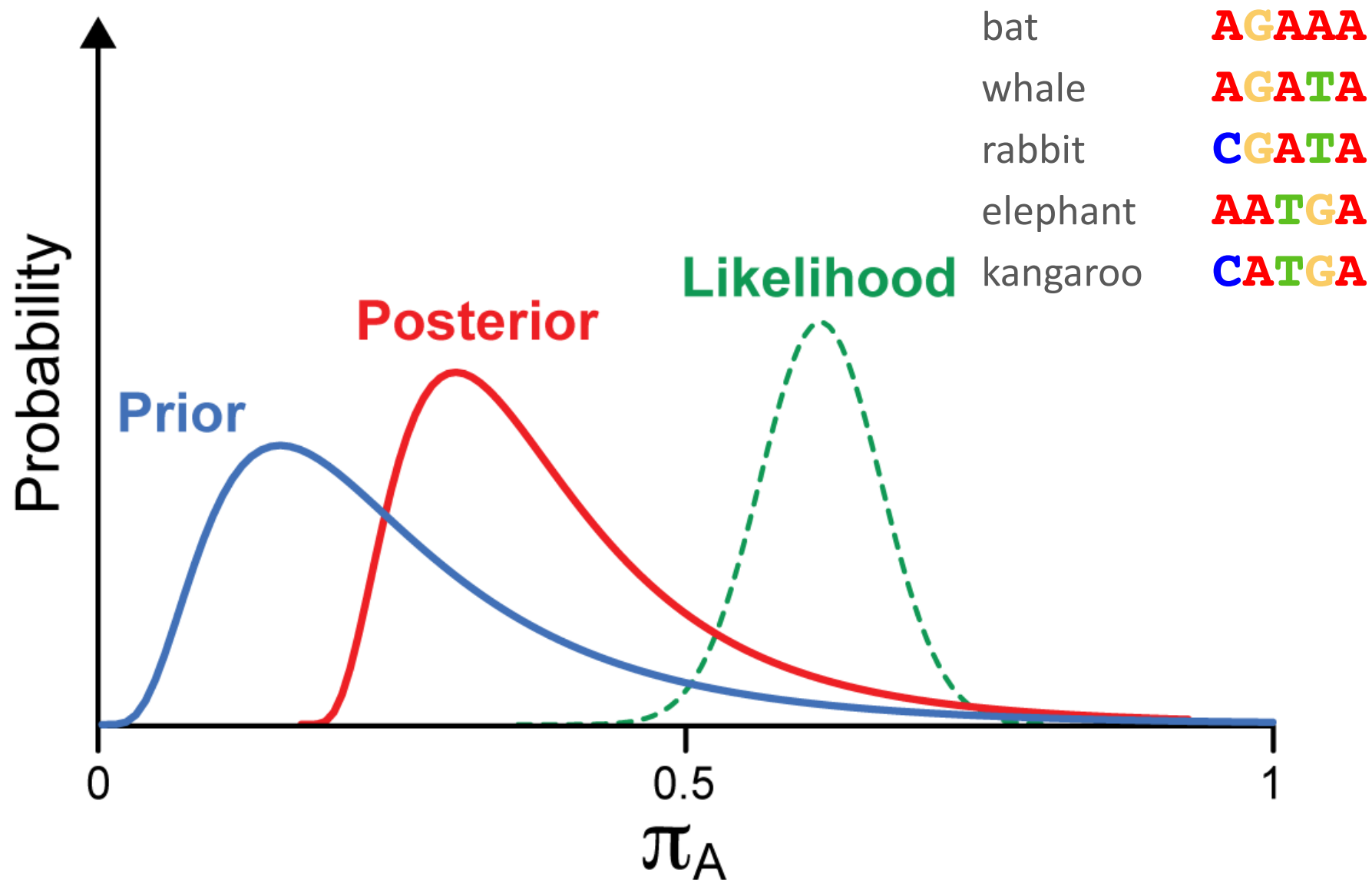
The Bayesian paradigm

- Contrast with frequentist statistics (likelihood)
- Parameters have **distributions**
- Before the data are observed, each parameter has a **prior distribution**
- The **likelihood** of the data is computed
- The prior distribution is combined (updated) with the likelihood to yield the **posterior distribution**

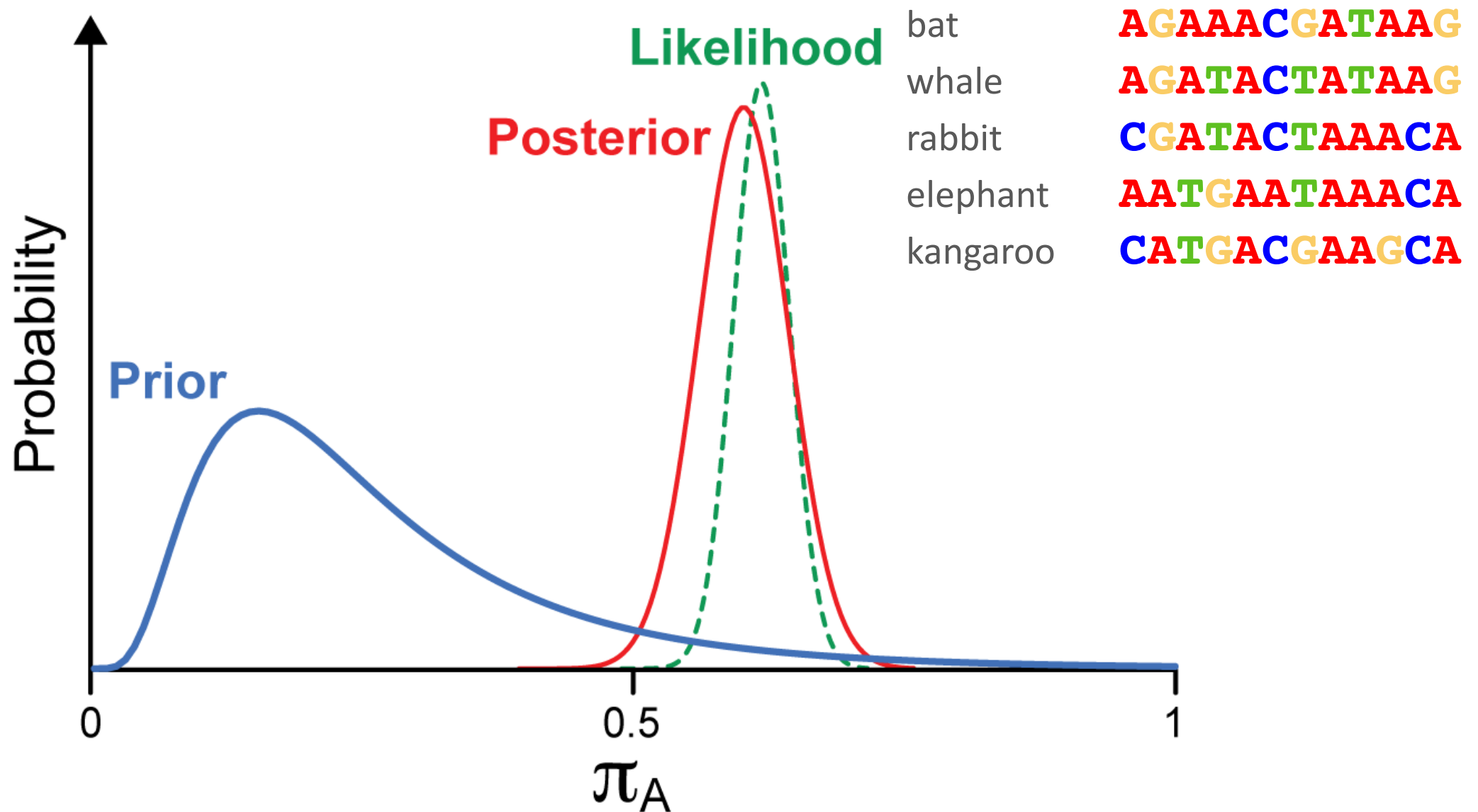
Simple example



Simple example



Simple example



Bayesian inference

Prior

Specified by user,
independent of data

Likelihood

Calculated from data

$$\Pr(\theta \mid D) = \frac{\Pr(\theta) \Pr(D \mid \theta)}{\Pr(D)}$$

The diagram shows the equation for Bayesian inference. The numerator consists of two terms, $\Pr(\theta)$ and $\Pr(D \mid \theta)$, each enclosed in a light blue rounded rectangle. A leader line from the 'Prior' label points to $\Pr(\theta)$, and another leader line from the 'Likelihood' label points to $\Pr(D \mid \theta)$. The denominator is $\Pr(D)$, also enclosed in a light blue rounded rectangle. A leader line from the text 'normalising constant' points to $\Pr(D)$. The entire equation is centered on the slide.

Posterior

This is what we
want to estimate

normalising constant
marginal likelihood of the data
model likelihood

Bayesian inference

Prior prob of tree

Topology

Branch lengths

**Prior prob of substitution
model parameters**

Rate parameters

Base frequencies

$$\boxed{\Pr(\tau, M \mid D)} = \frac{\boxed{\Pr(\tau)} \boxed{\Pr(M)} \boxed{\Pr(D \mid \tau, M)}}{\Pr(D)}$$

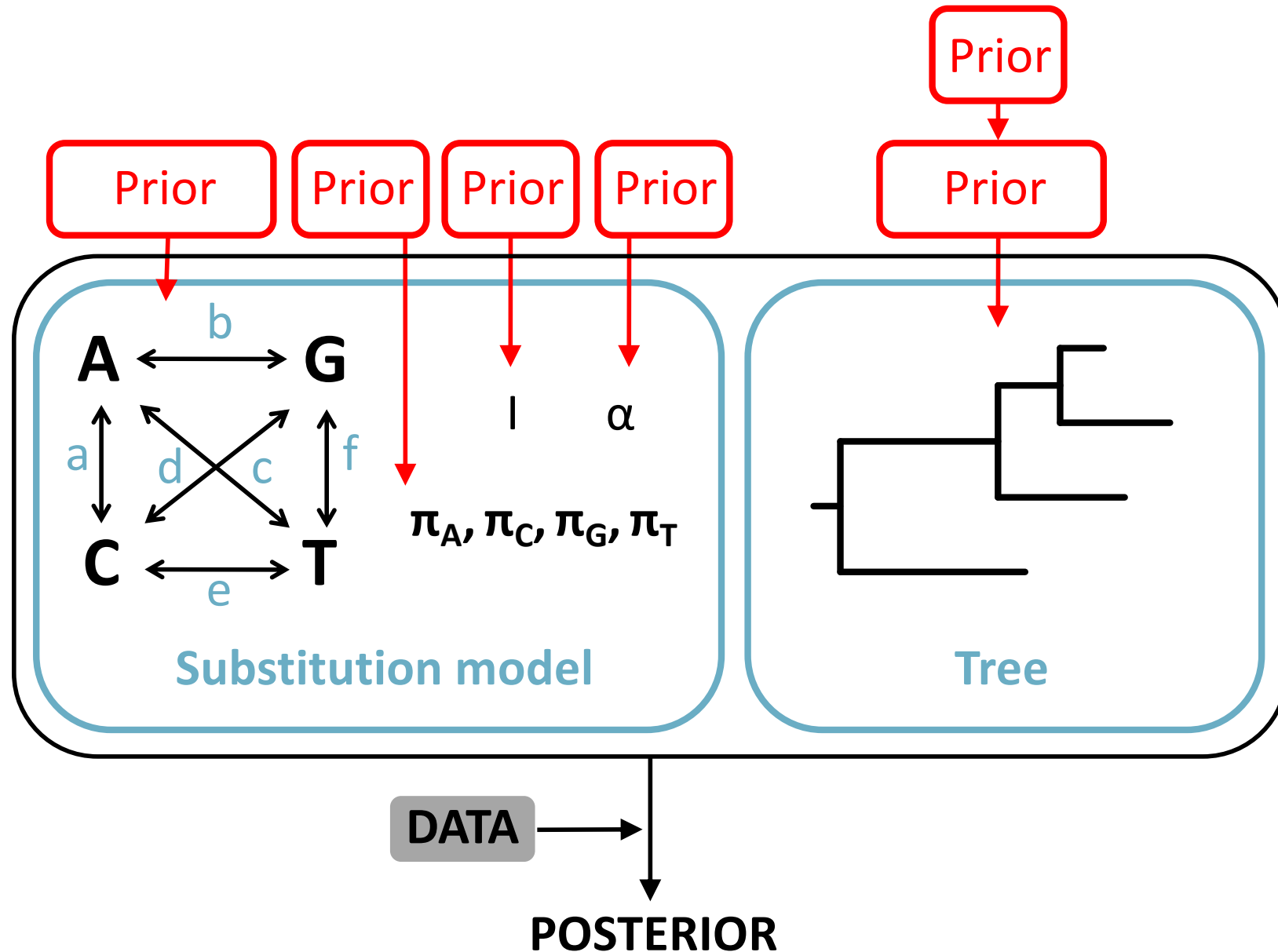
Posterior

This is what we
want to estimate

Likelihood

Calculated from data

Bayesian hierarchical model



Priors

Priors

- Priors are chosen in the form of probability distributions
- Reflect our prior expectations (and uncertainty) about values of parameters (without knowledge of the data)
 - Past observations
 - Personal beliefs
 - Use of a biological model

Continuous distributions

- Uniform
- Normal

Used to specify prior distributions of various continuous parameters

- Exponential
- Lognormal
- Gamma

Used to specify prior distributions of continuous parameters that cannot take negative values

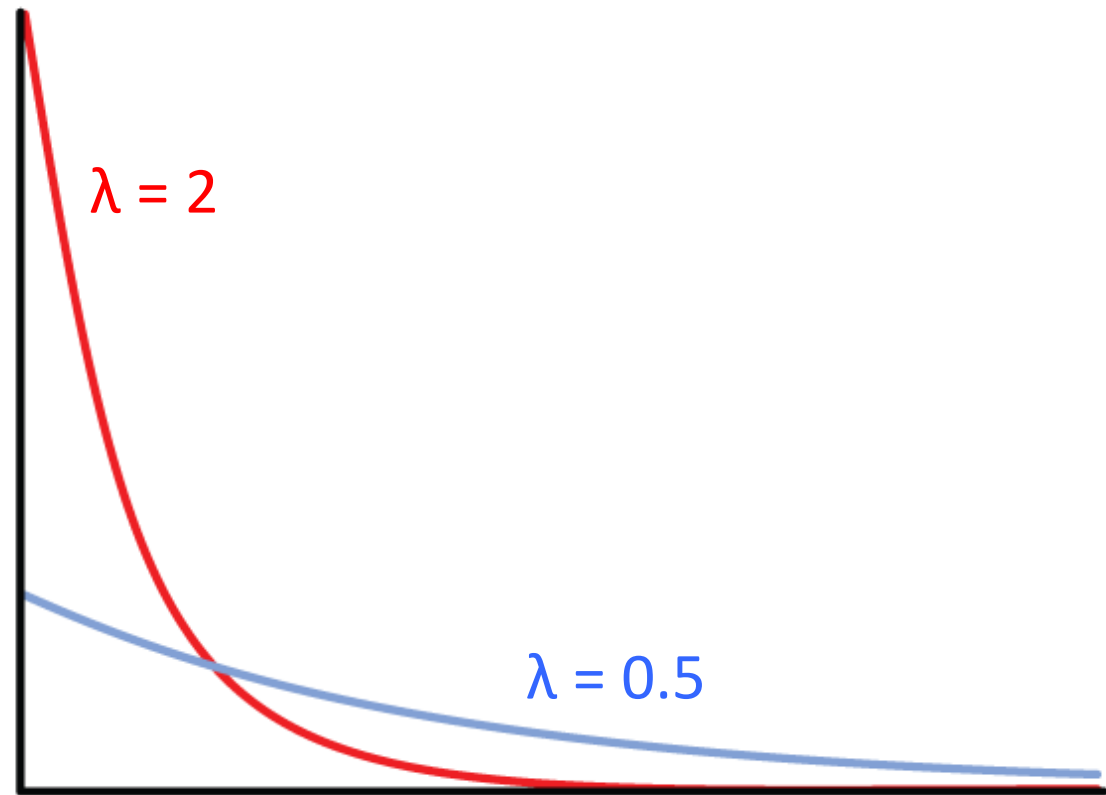
- Beta
- Dirichlet

Continuous distributions

- Uniform
- Normal
- **Exponential**
- Lognormal
- Gamma
- Beta
- Dirichlet

Parameters

- λ = rate of decay



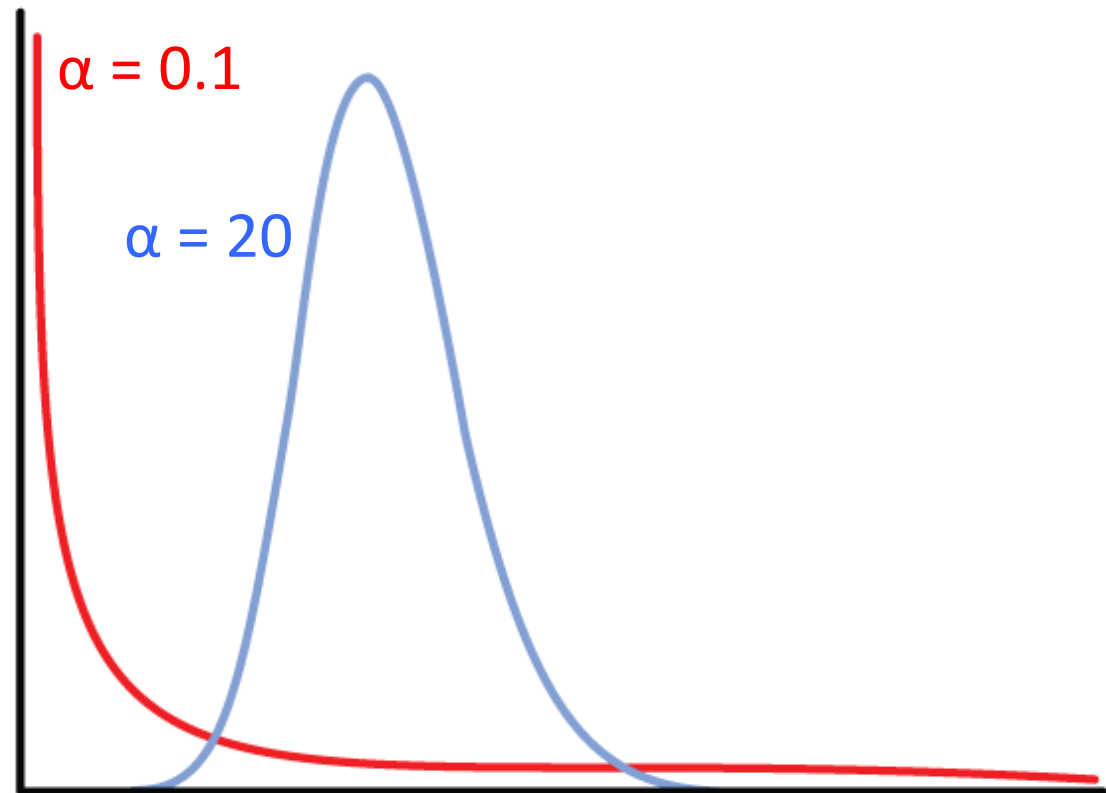
Continuous distributions

- Uniform
- Normal
- Exponential
- Lognormal
- **Gamma**
- Beta
- Dirichlet

Parameters

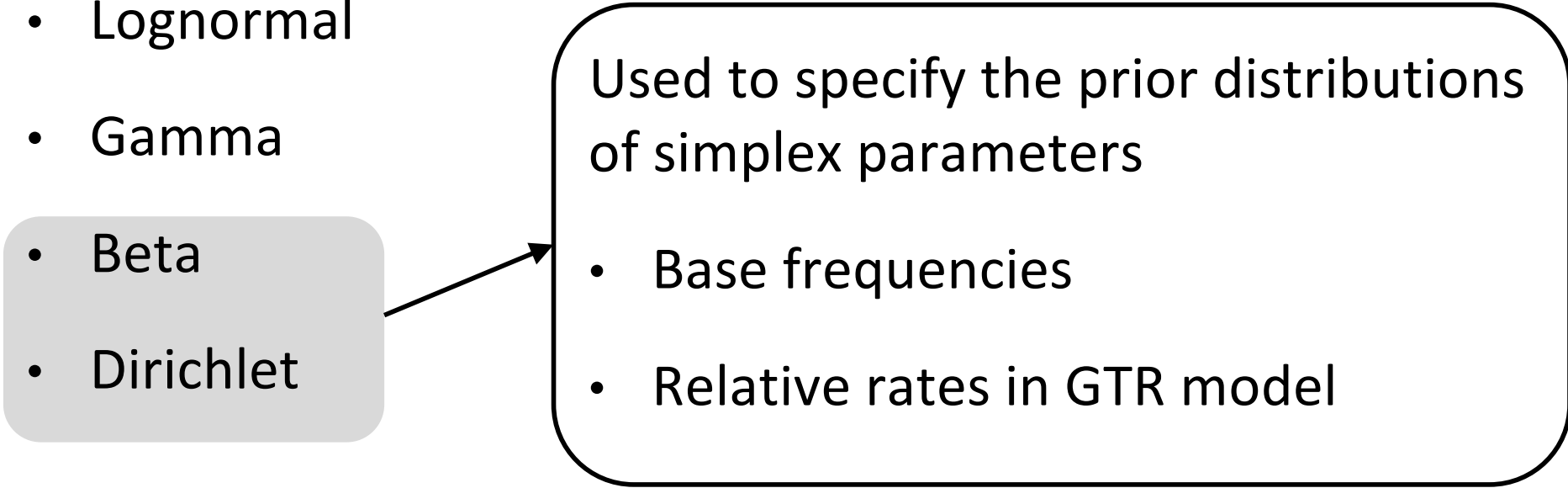
- α = shape
- β = inverse scale

Scaled gamma:
 $\alpha=\beta$



Continuous distributions

- Uniform
- Normal
- Exponential
- Lognormal
- Gamma
- Beta
- Dirichlet



Used to specify the prior distributions of simplex parameters

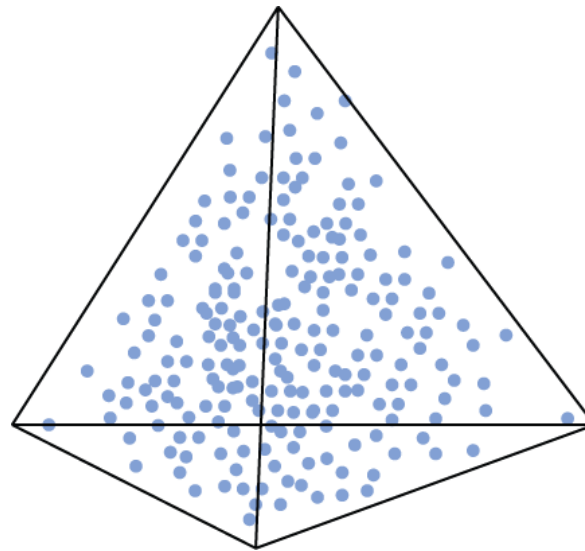
- Base frequencies
- Relative rates in GTR model

Continuous distributions

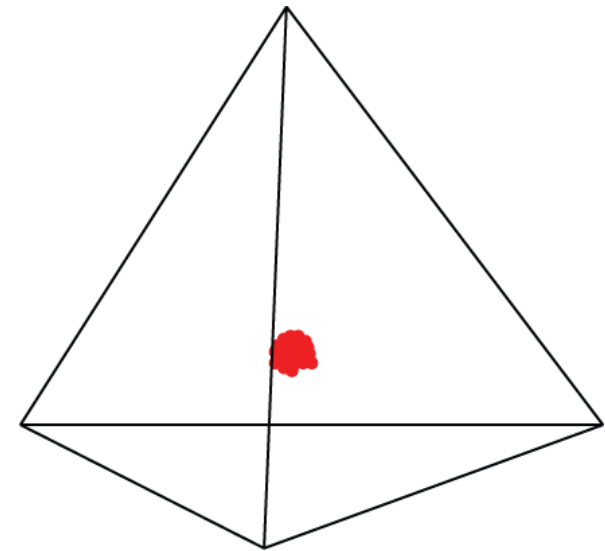
- Uniform
- Normal
- Exponential
- Lognormal
- Gamma
- Beta
- **Dirichlet**

Parameters

- $\alpha_1, \alpha_2, \dots$ = shape parameters



$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$$



$$\alpha_1 = \alpha_2 = \alpha_3 = 300, \alpha_4 = 1$$

Discrete distributions

- Bernoulli distribution
- Binomial
- Multinomial
- Poisson

Default priors

	<i>BEAST2</i>	<i>MrBayes</i>
Rate matrix parameters	Gamma(0.05,10)	Dirichlet(1,1,1,1,1,1)
Base frequencies	Uniform(0,1)	Dirichlet(1,1,1,1)
Shape parameter (α)	Exponential(1)	Exponential(2)
Proportion invariable	Uniform(0,1)	Uniform(0,1)

Uninformative priors

- **Flat or uniform prior**
- **Jeffreys prior**
 - Invariant to reparameterisation
 - Only works well for models with a single parameter
- **Reference priors**
 - Maximise expected distance between prior and posterior
 - Allows data to have maximum effect on posterior

Tree Prior

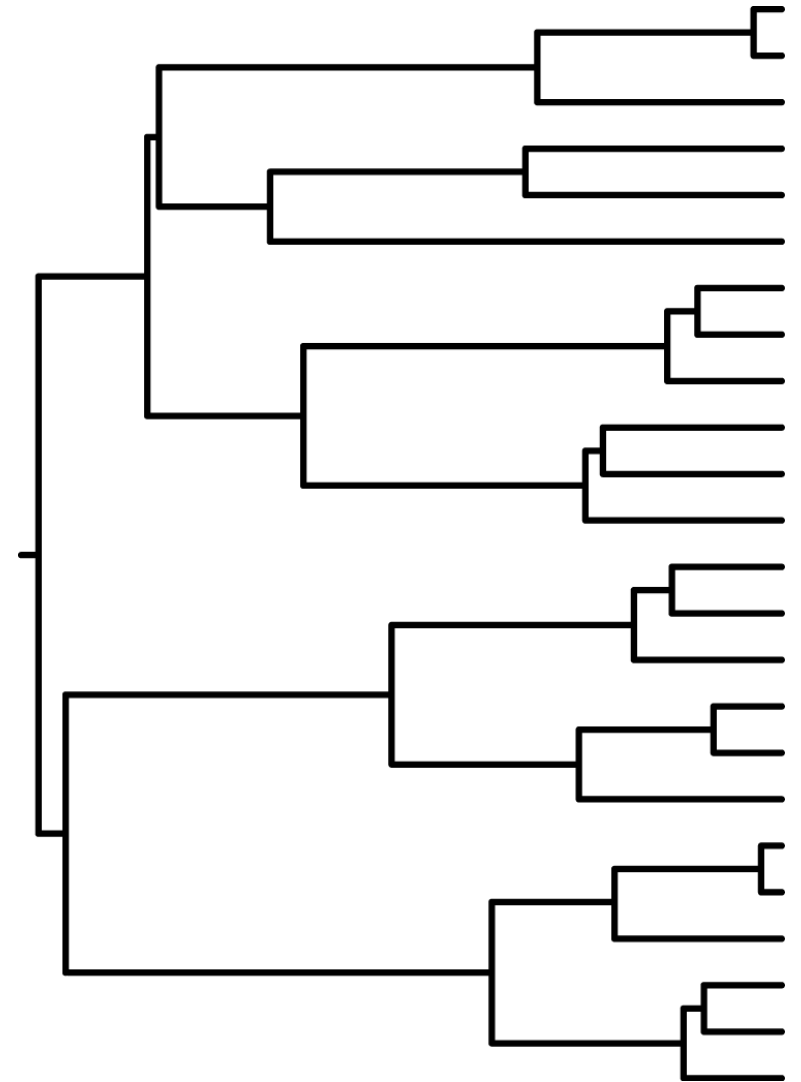
Tree prior

1. Use a **flat prior** (*MrBayes*)
 - All trees have equal probability
 - Also need a prior for branch lengths or node times

 2. Use a **biological model** (*BEAST* and *MrBayes*)
 - Among species: speciation model
 - Within species: coalescent model
- } Priors on rooted trees

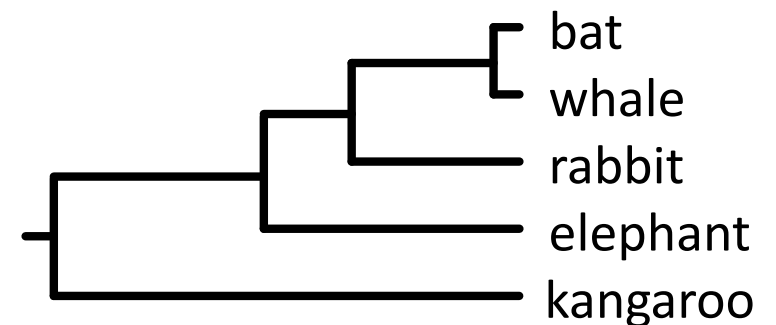
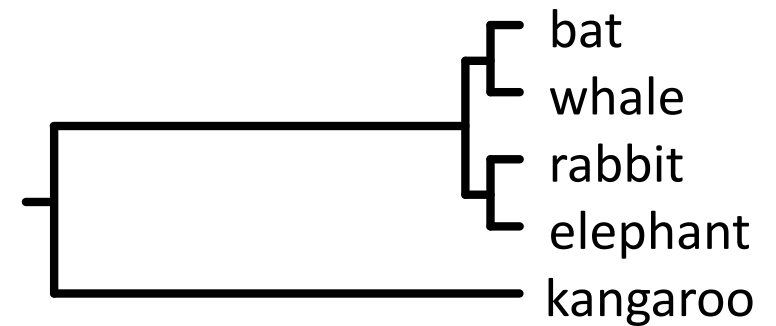
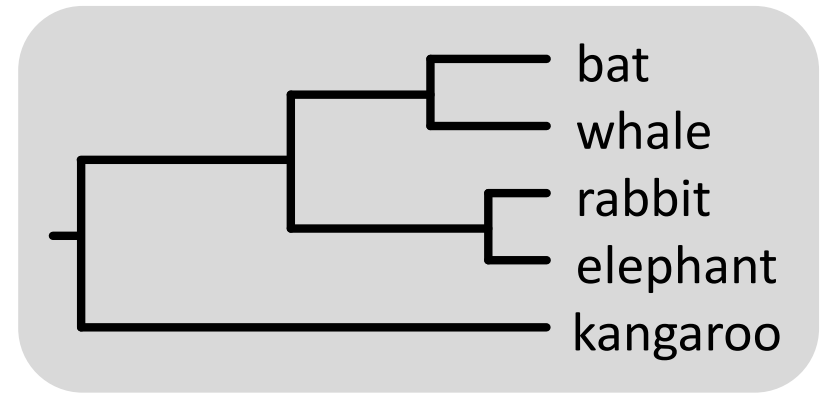
Speciation model

- Tree shape described by a stochastic branching process
- **Yule process**
 - The root lineage splits into two
 - Lineages split at a constant rate
 - Simulates speciation process
- **Birth-death process**
 - Allow lineages to go extinct



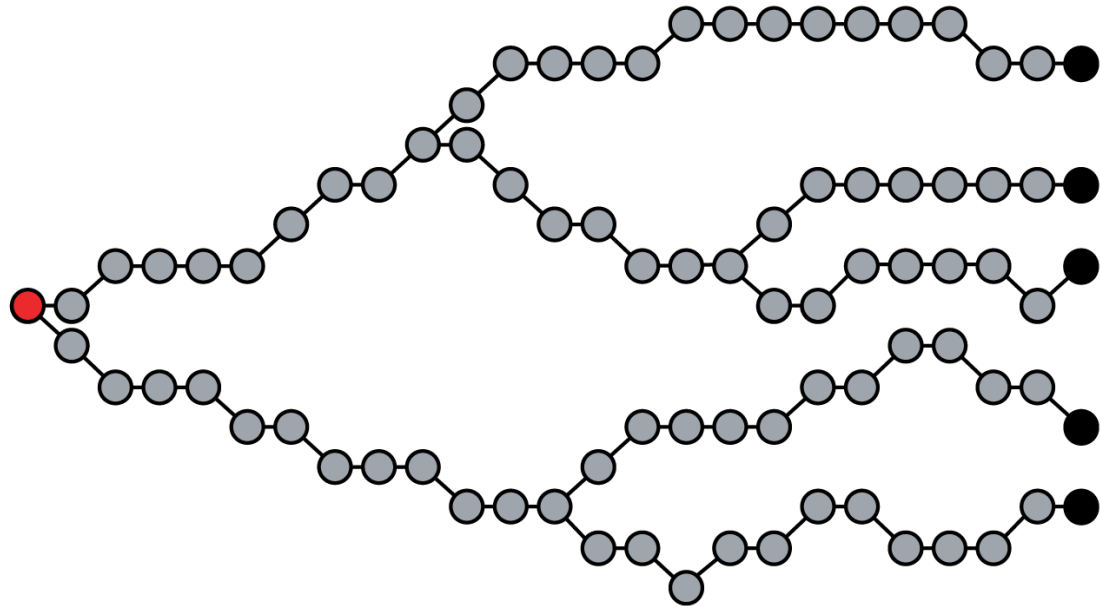
Speciation model

- Tree shape described by a stochastic branching process
- **Yule process**
 - The root lineage splits into two
 - Lineages split at a constant rate
 - Simulates speciation process
- **Birth-death process**
 - Allow lineages to go extinct

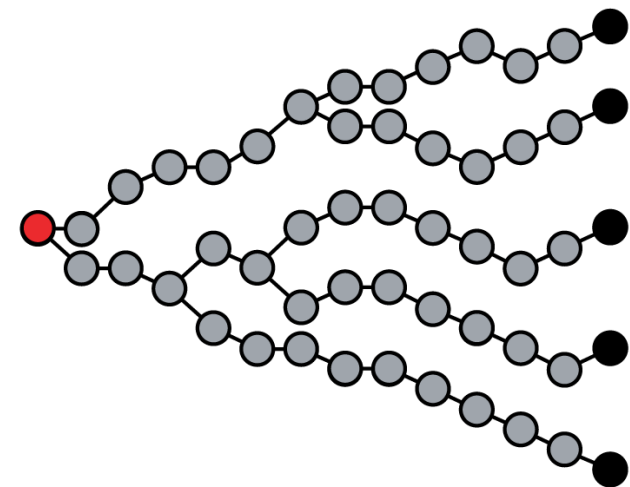


Coalescent model

Constant size



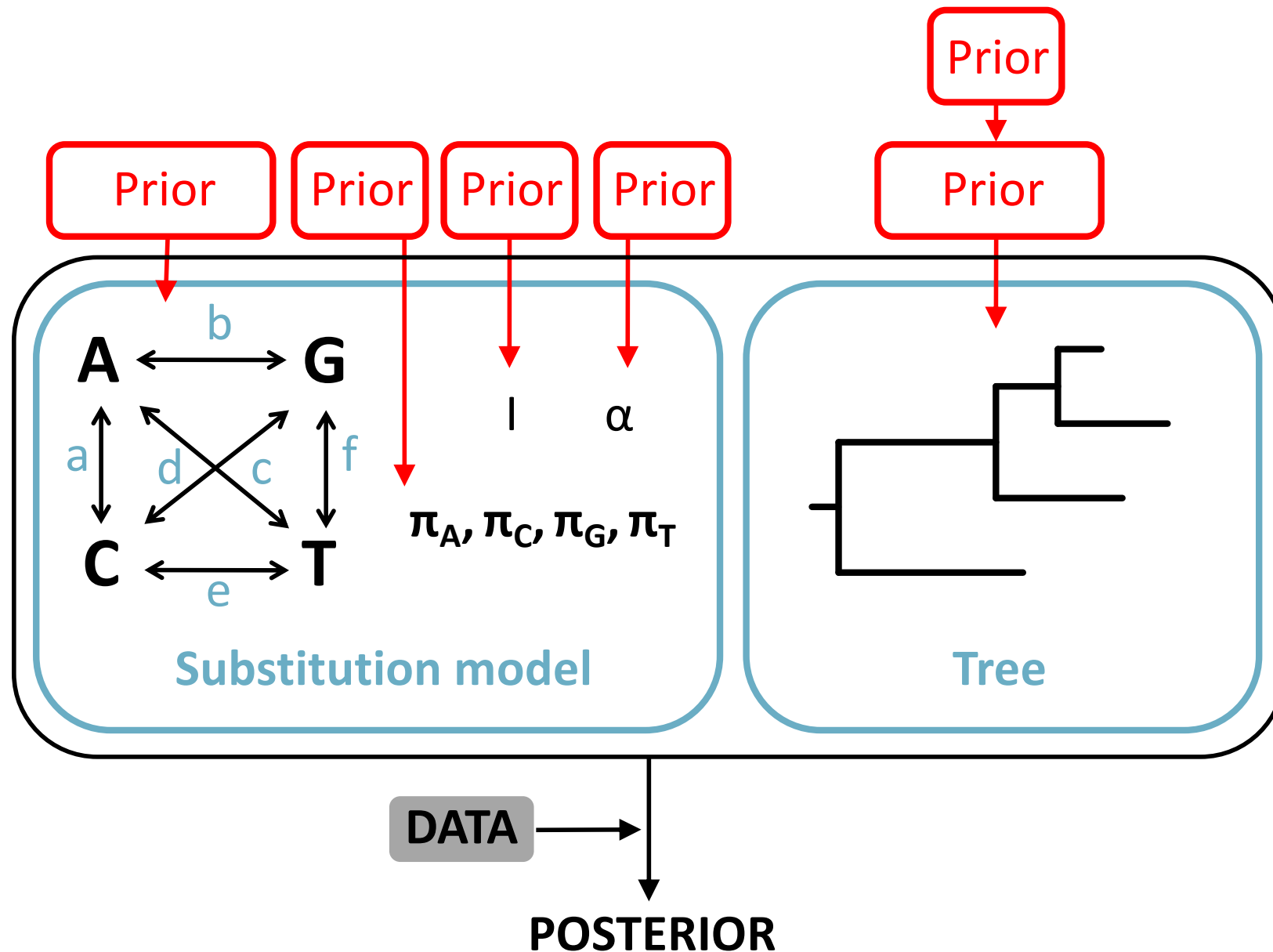
Exponential growth



Choosing a tree prior

- Test whether inferences are robust to the choice of tree prior
- Mixed data sets: multiple sequences from each species
 - Birth-death prior generally works well
- Compare tree priors using Bayesian model selection

Bayesian hierarchical model



Useful references

