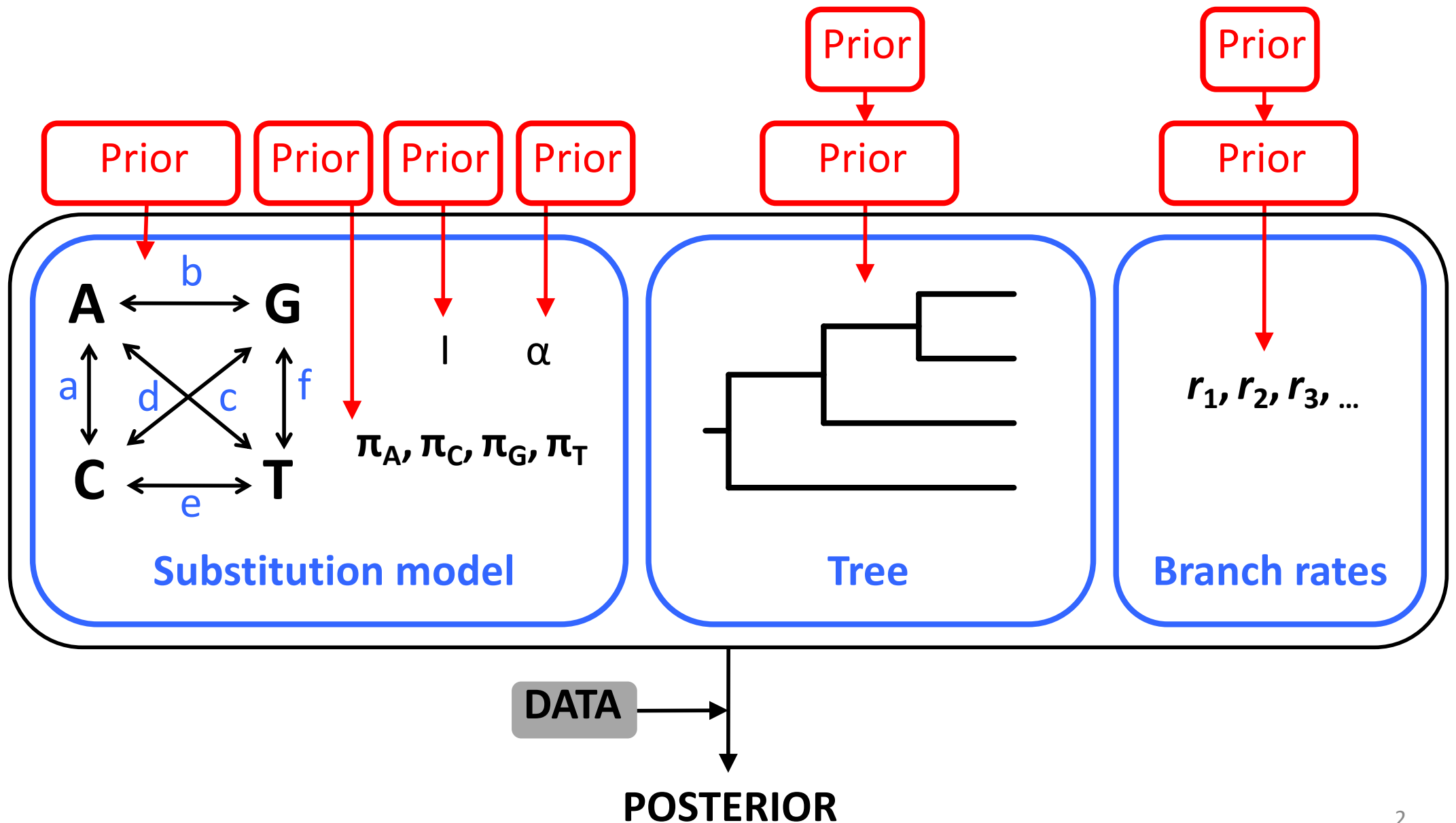

Lecture 2.2

Models and Priors

Bayesian hierarchical model



Prior Distributions

Priors

- 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

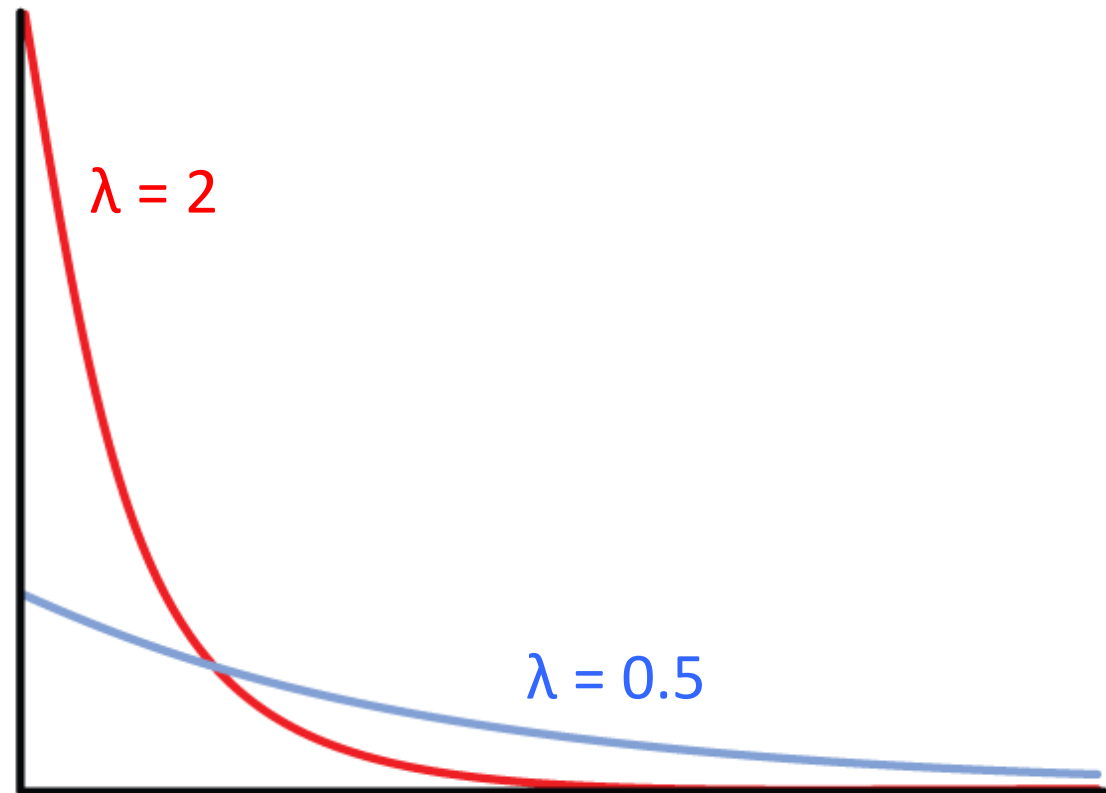
- Beta
- Dirichlet

Continuous distributions

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

Parameters

- λ = rate of decay



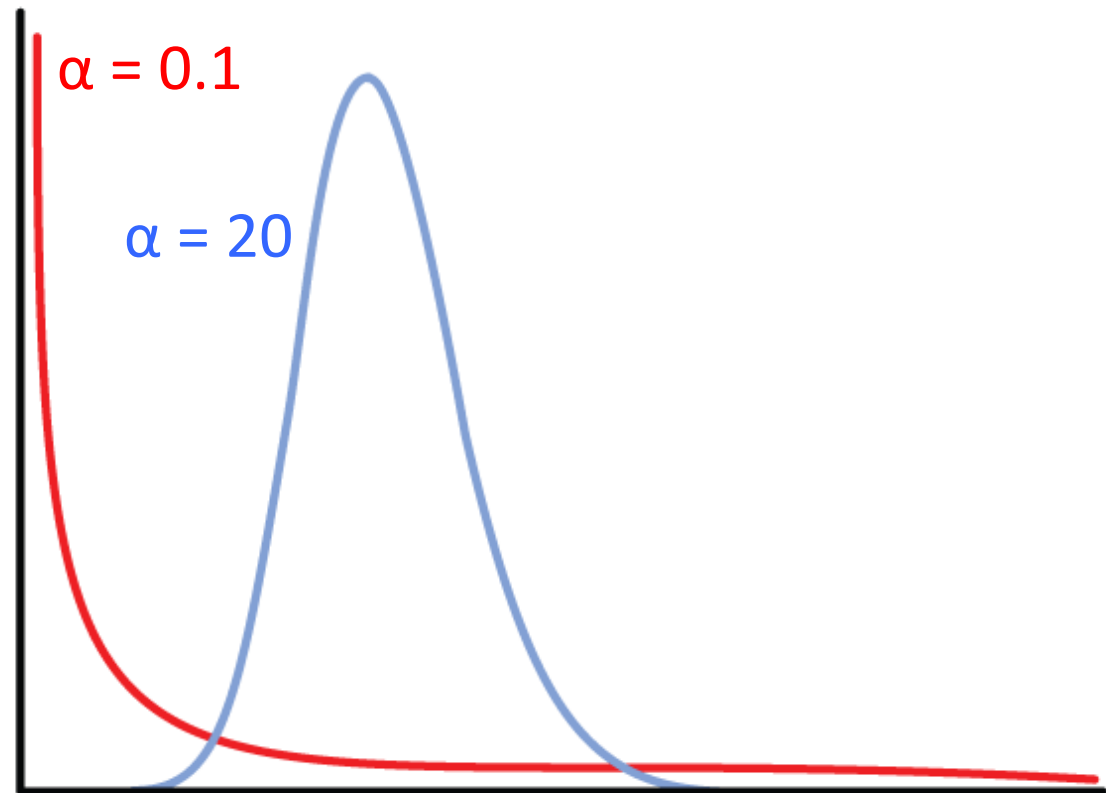
Continuous distributions

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

Parameters

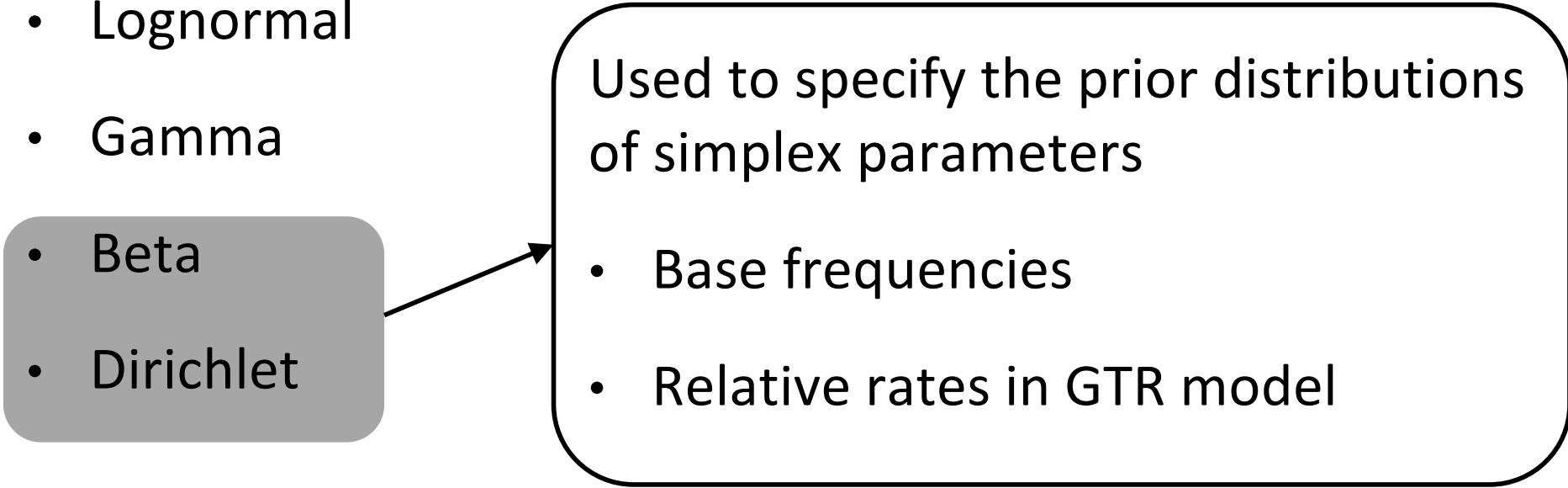
- α = shape
- β = inverse scale

Scaled gamma:
 $\alpha=\beta$



Continuous distributions

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



Used to specify the prior distributions of simplex parameters

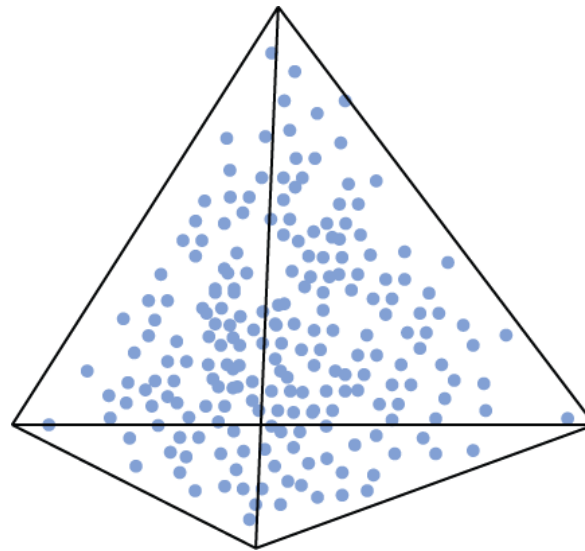
- Base frequencies
- Relative rates in GTR model

Continuous distributions

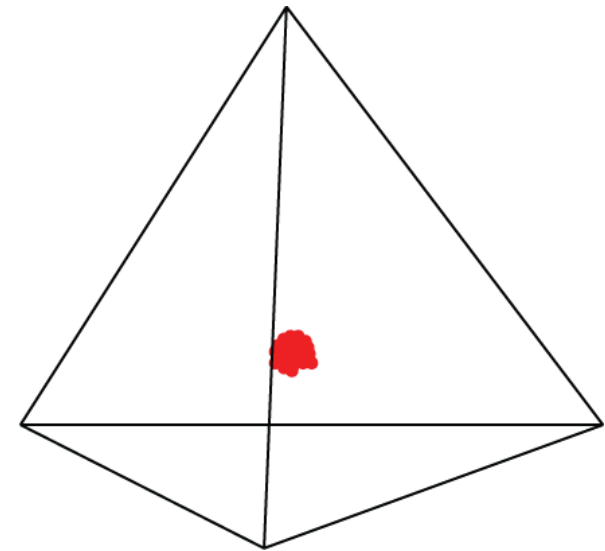
- Uniform
- Normal
- Exponential
- Lognormal
- Gamma
- Beta
- **Dirichlet**

Parameters

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



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



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

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

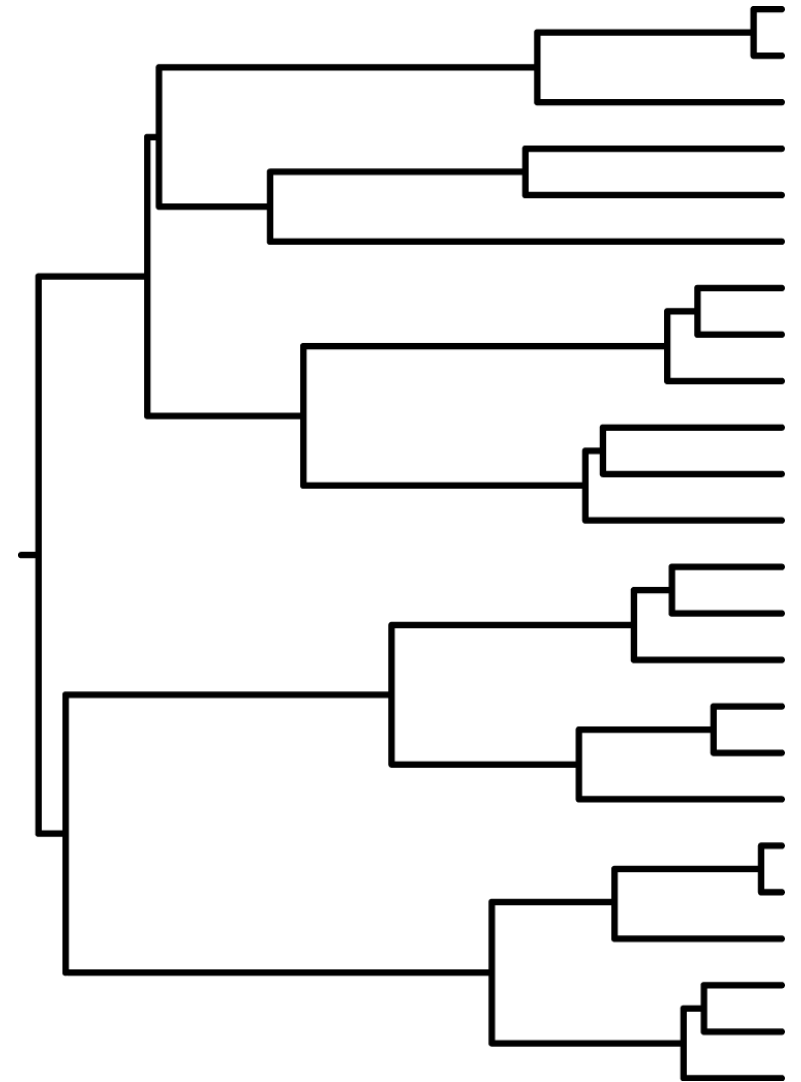
Tree prior

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**

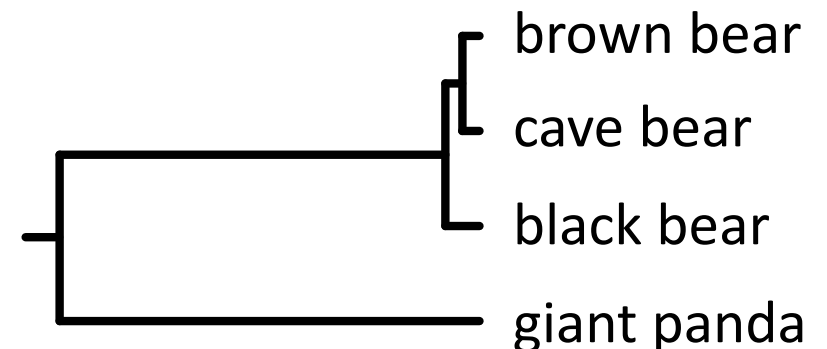
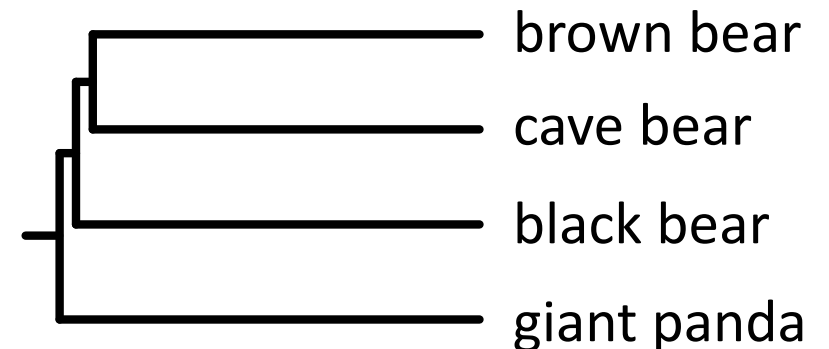
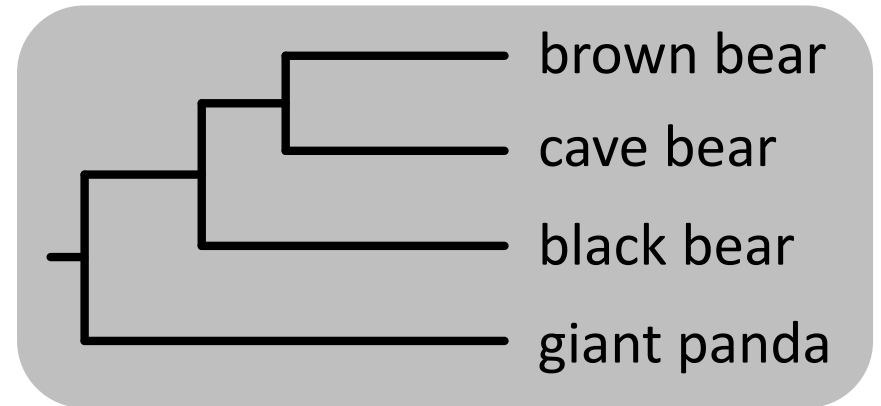
Speciation model

- 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



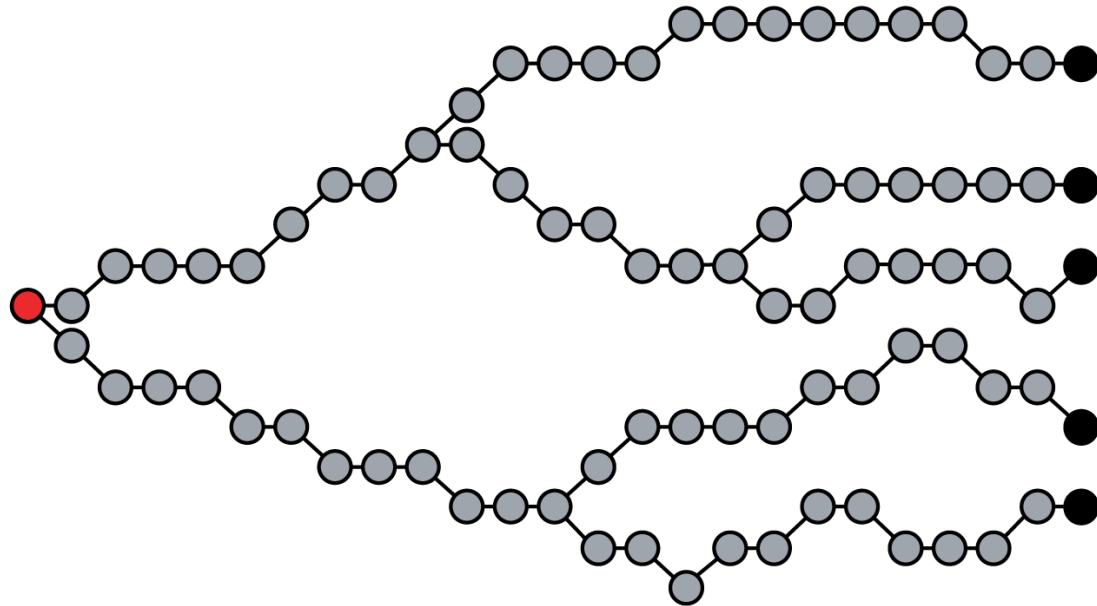
Speciation model

- 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

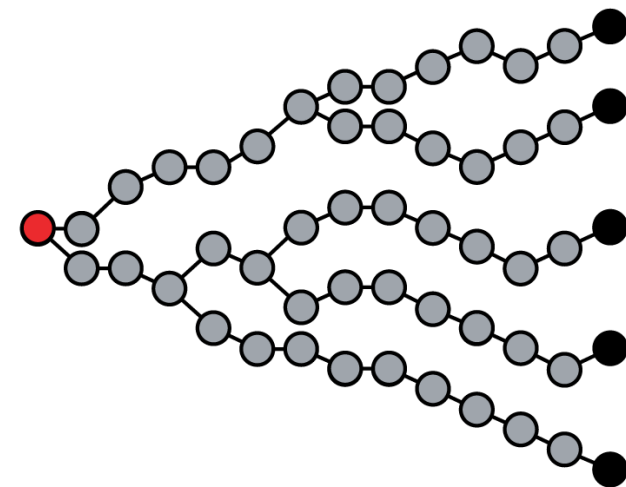


Coalescent model

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 Model Selection

Bayesian model selection

- Bayesian model selection is usually based on the marginal probability of the data, conditioned on the model:

$$\Pr(D | M)$$

- This is a weighted average of the likelihood
- Weights are given by the prior distribution

Marginal likelihood of the model

Bayesian model selection

- Compare marginal likelihoods of competing models
- Ratio of marginal likelihoods is the **Bayes factor**

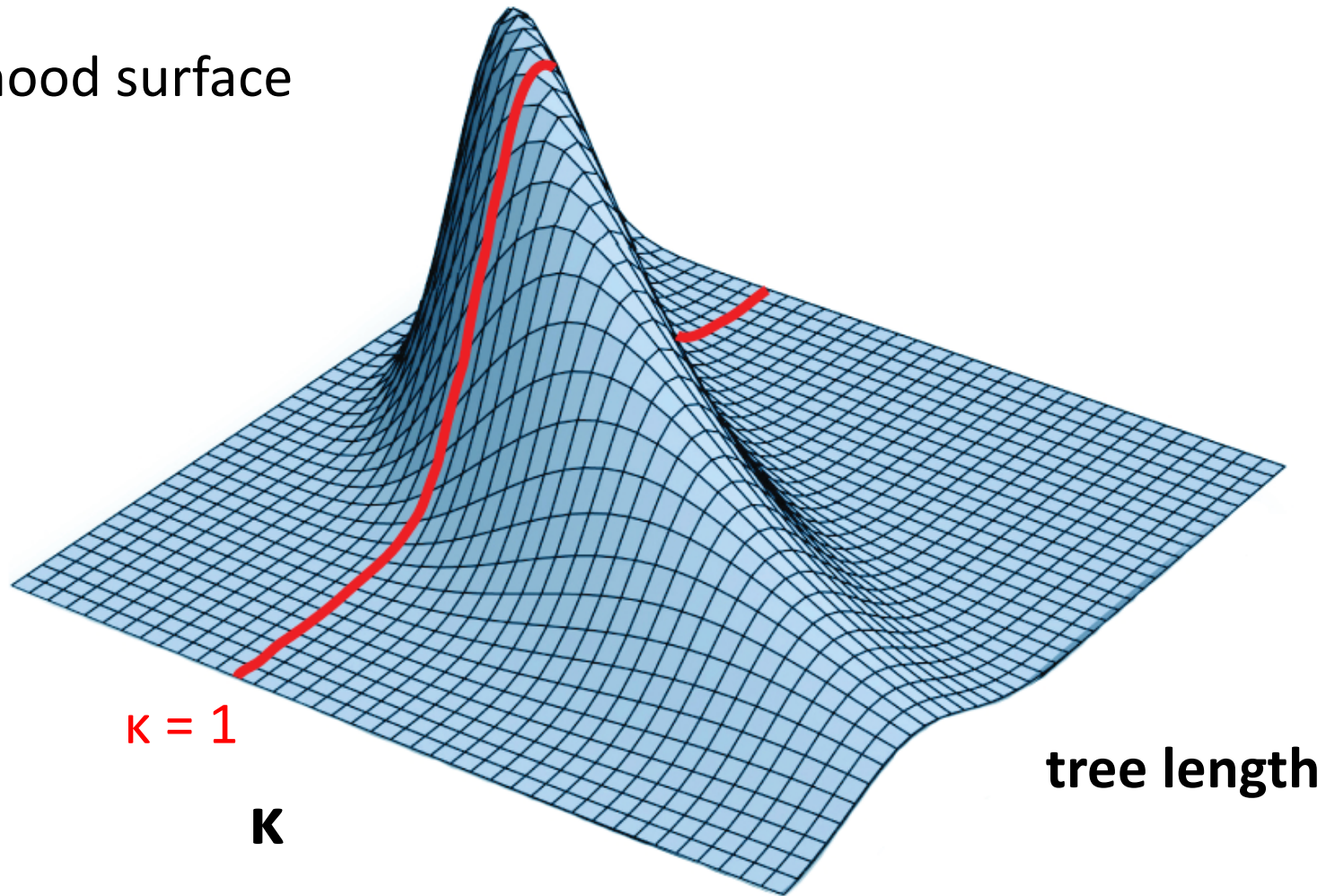
$$\mathbf{BF} = \frac{\mathbf{Pr(D | M_1)}}{\mathbf{Pr(D | M_2)}}$$

$$\mathbf{\log BF = \log Pr(D | M_1) - \log Pr(D | M_2)}$$

- Models do not need to be nested
- Do not need to correct for number of parameters

Bayesian model selection

Likelihood surface



Bayesian model selection

- Interpreting the Bayes factor

BF	$\log BF$	Evidence against M_2
1 – 3	0 – 1	Not worth mentioning
3 – 20	1 – 3	Positive
20 – 150	3 – 5	Strong
> 150	> 5	Very strong

Useful references

- **The impact of the tree prior on molecular dating of data sets containing a mixture of inter- and intraspecies sampling**
Ritchie, Lo, & Ho (2017) *Syst Biol*, 66: 413–425.

