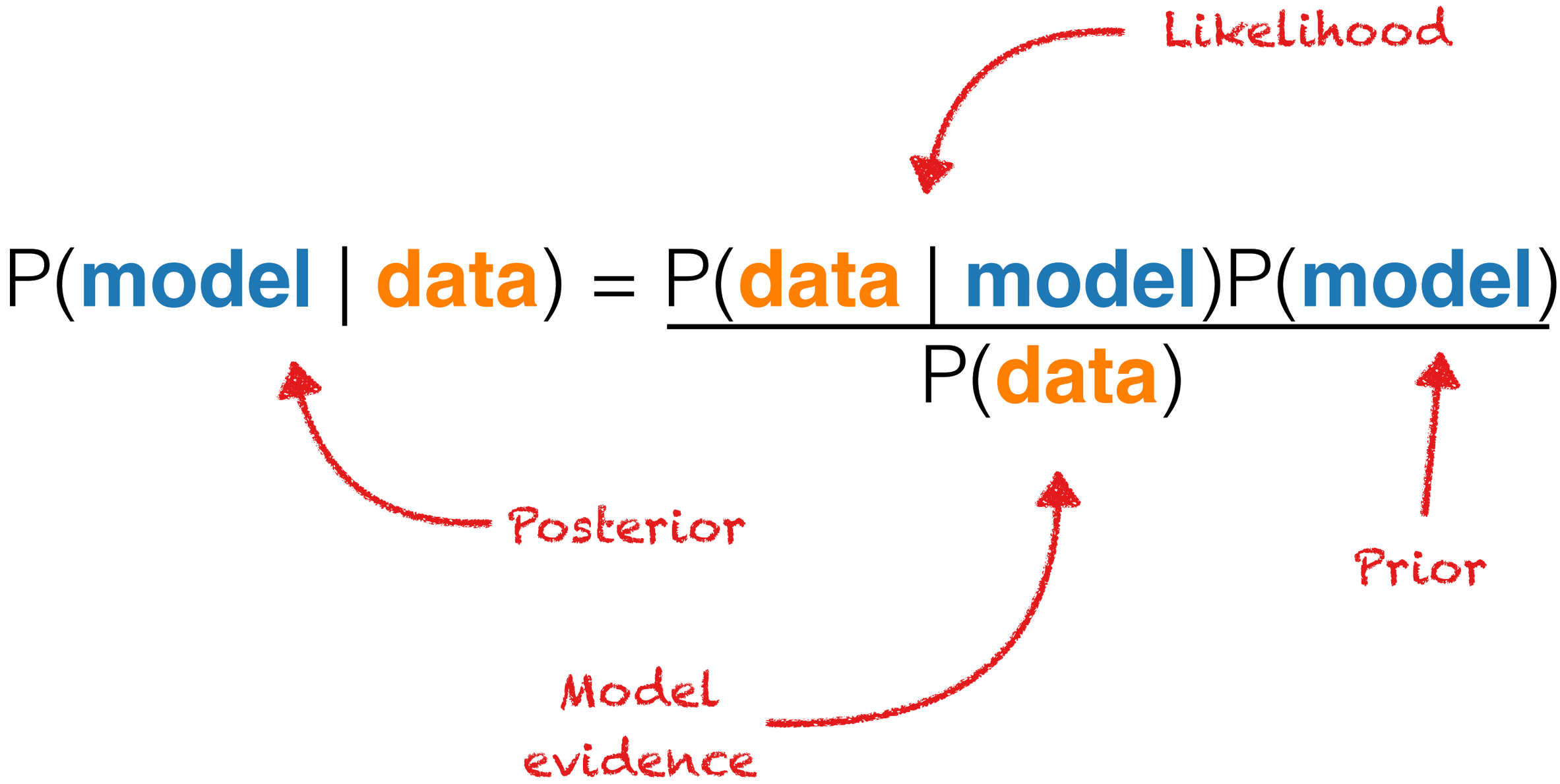
Taming the BEAST

Bayesian Evolutionary Analyses by Sampling Trees

Louis du Plessis

Setting priors

Some slides inspired by slides from Veronika Bošková, Jana Huisman, Vladimir Minin and Chi Zhang



Prior → P(model)

- Original probability for the model parameters/components (before data collected or analysed)
- Belief in our hypothesis
- All parameters have priors, whether you specify them or not!

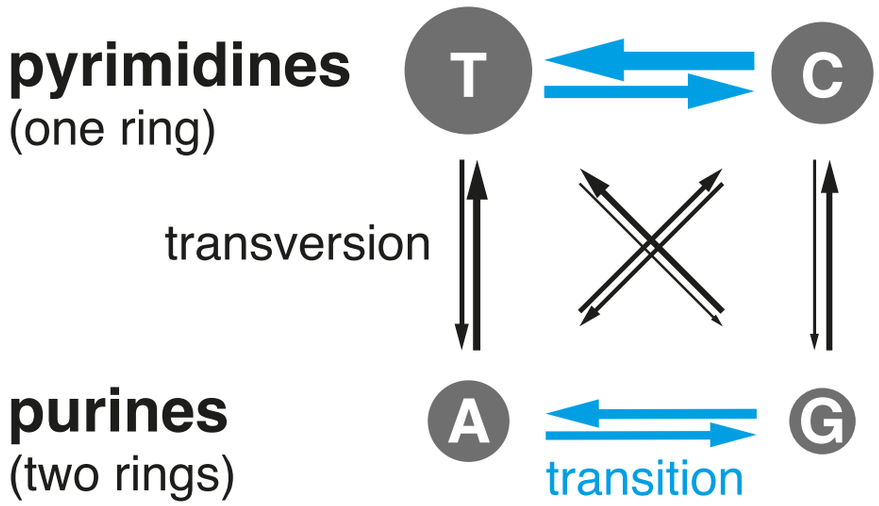
What is a **prior**?

- Usually a **distribution** for some model **parameter**
  e.g. clock rate, tree, …
- Often a **parametric distribution**
  e.g. uniform, normal, gamma, beta, lognormal, Laplace, …
- Sometimes a **prior** on a model component
  e.g. substitution model (HKY, GTR, JC, …)
- Priors can have priors which can have priors *ad infinitum* (hyperpriors)
- Parameter bounds are part of the prior
  e.g. normal distribution with lower bound 0.

How to pick a **prior**?

- Should be chosen based on your beliefs about model parameters (from independent evidence/experiments)
- Should be chosen with a particular analysis in mind (no priors are universal!)
- Think carefully about plausible ranges of parameters
- Be conservative if you are unsure about the parameter (use diffuse priors)
- Conjugate priors? (no — those are just for mathematical convenience)
- Reference priors? (yes — if you are an objective Bayesian)
- Do not use improper priors if you can help it (priors that integrate to infinity, e.g. $1/x$)

Example: HKY-model (HKY85)



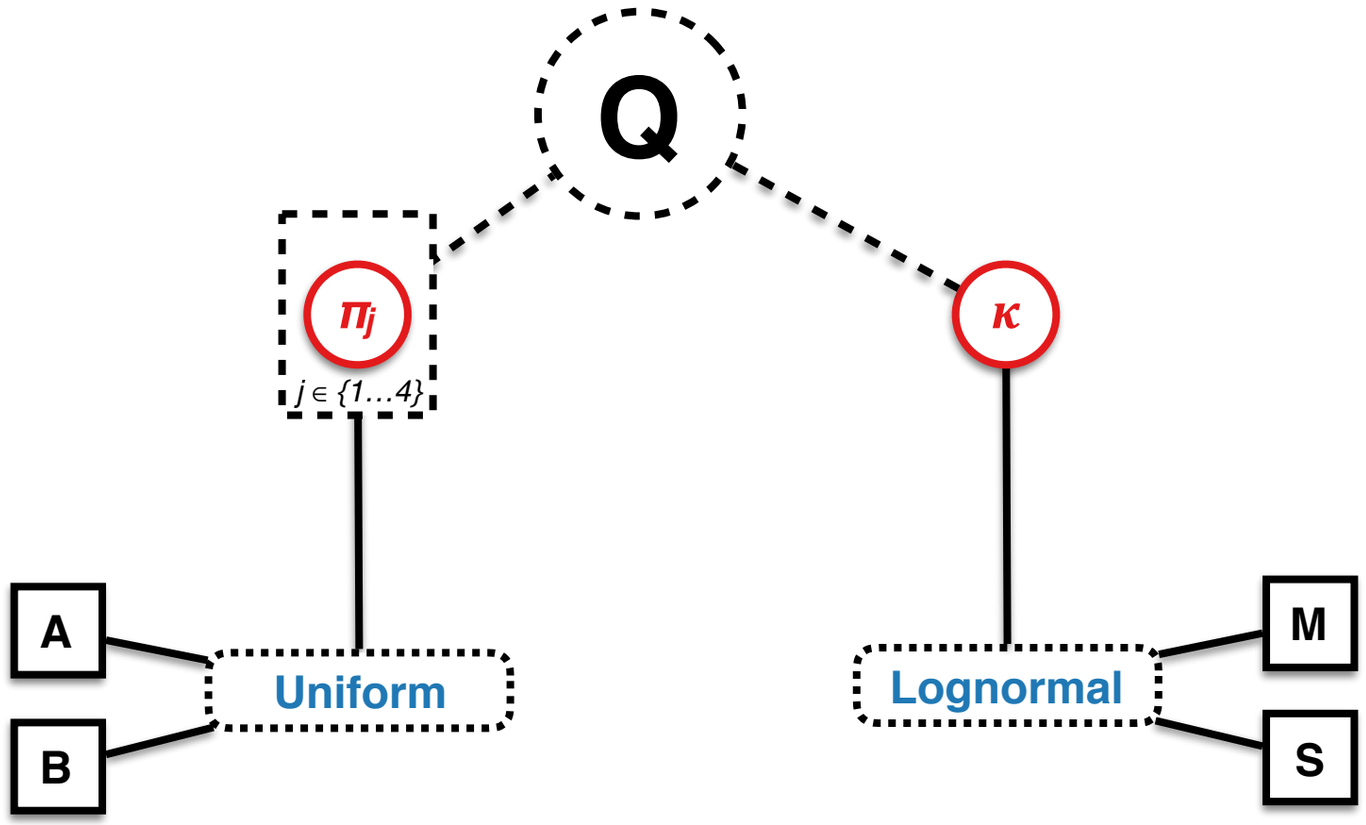
$(\pi_T, \pi_C, \pi_A, \pi_G)$

$$
\begin{array}{c} \\ T \\ C \\ A \\ G \end{array}
\begin{array}{c}
\begin{array}{cccc} T & C & A & G \end{array} \\
\begin{pmatrix} \cdot & \alpha\pi_C & \beta\pi_A & \beta\pi_G \\ \alpha\pi_T & \cdot & \beta\pi_A & \beta\pi_G \\ \beta\pi_T & \beta\pi_C & \cdot & \alpha\pi_G \\ \beta\pi_T & \beta\pi_C & \alpha\pi_A & \cdot \end{pmatrix}
\end{array}
$$

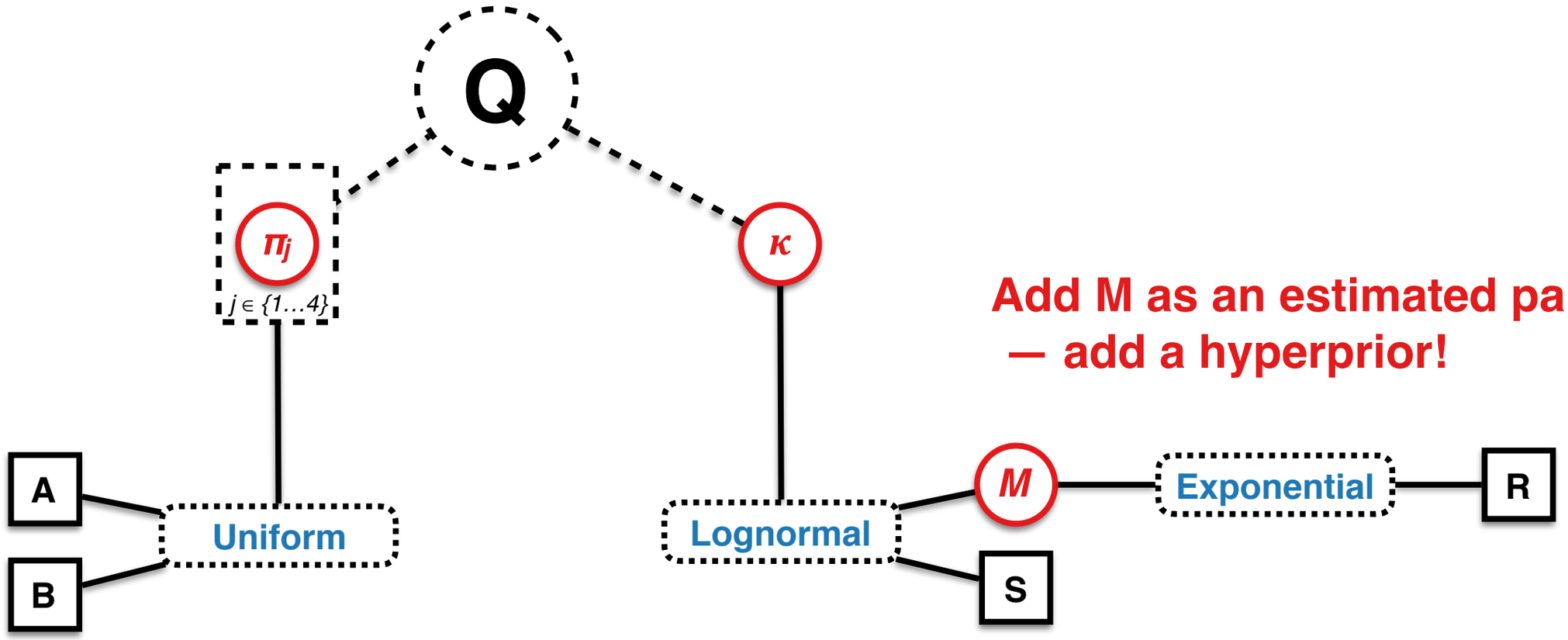
$$
= \begin{pmatrix} \cdot & \alpha & \beta & \beta \\ \alpha & \cdot & \beta & \beta \\ \beta & \beta & \cdot & \alpha \\ \beta & \beta & \alpha & \cdot \end{pmatrix} \cdot \begin{pmatrix} \pi_T & 0 & 0 & 0 \\ 0 & \pi_C & 0 & 0 \\ 0 & 0 & \pi_A & 0 \\ 0 & 0 & 0 & \pi_G \end{pmatrix}
$$

- **Q** matrix gives relative rates of substitution between nucleotides
- 5 parameters:
  - $\boldsymbol{\kappa} = \alpha/\beta$
  - $\boldsymbol{\pi_T, \pi_C, \pi_A, \pi_G}$ — assume these all have the same prior probability

Example: HKY-model (HKY85)

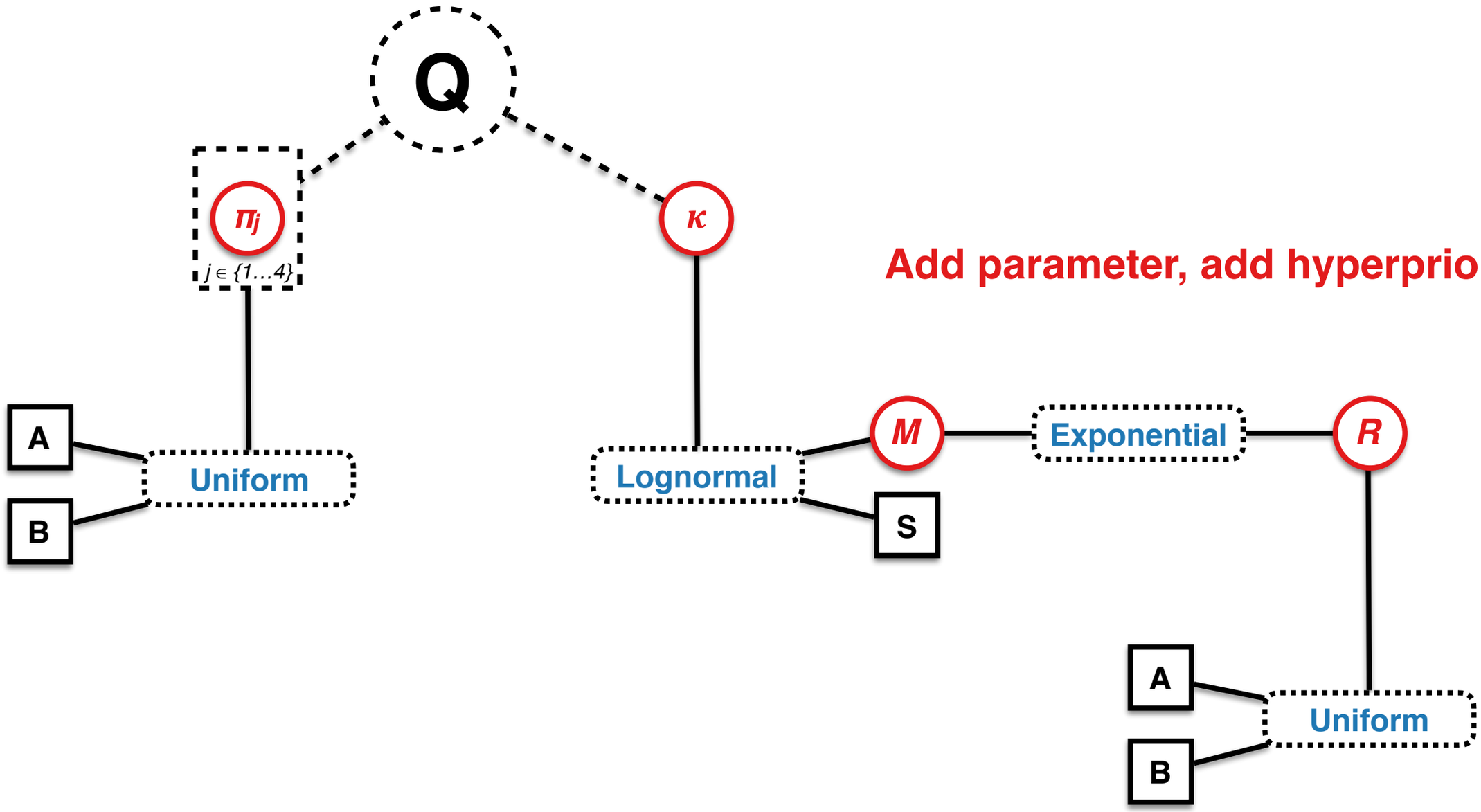


Example: HKY-model (HKY85)



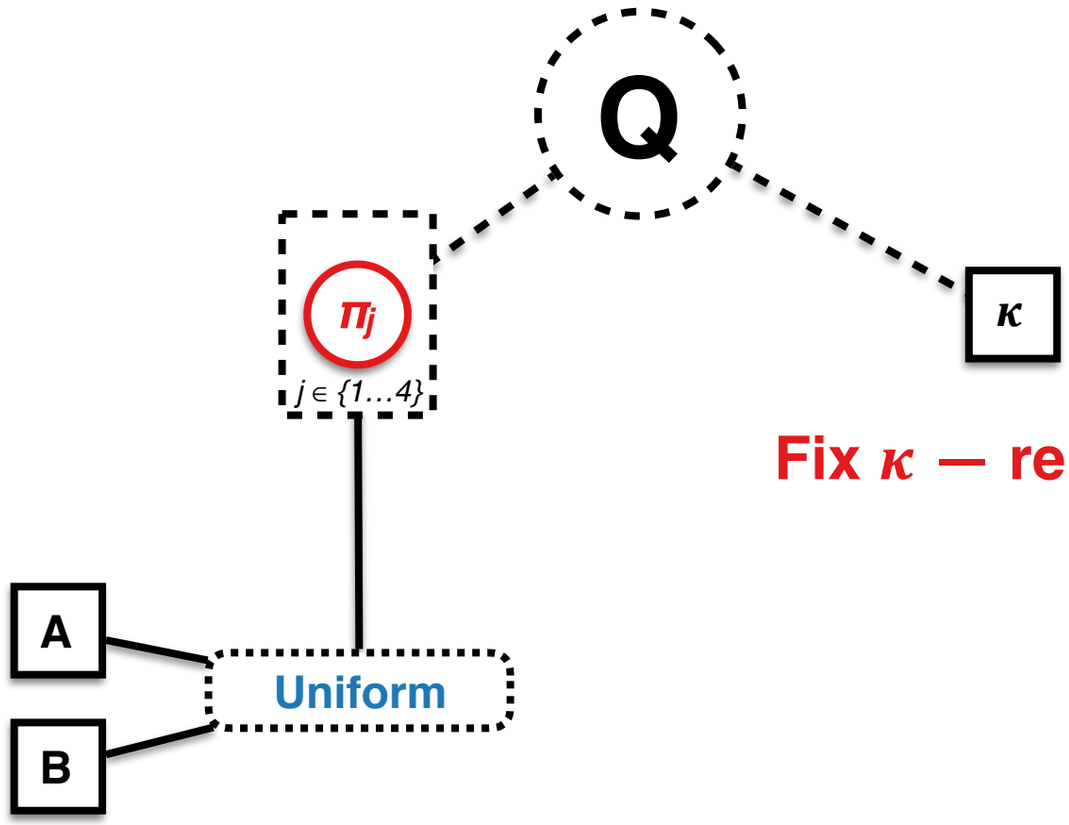
Add M as an estimated parameter — add a hyperprior!

Example: HKY-model (HKY85)



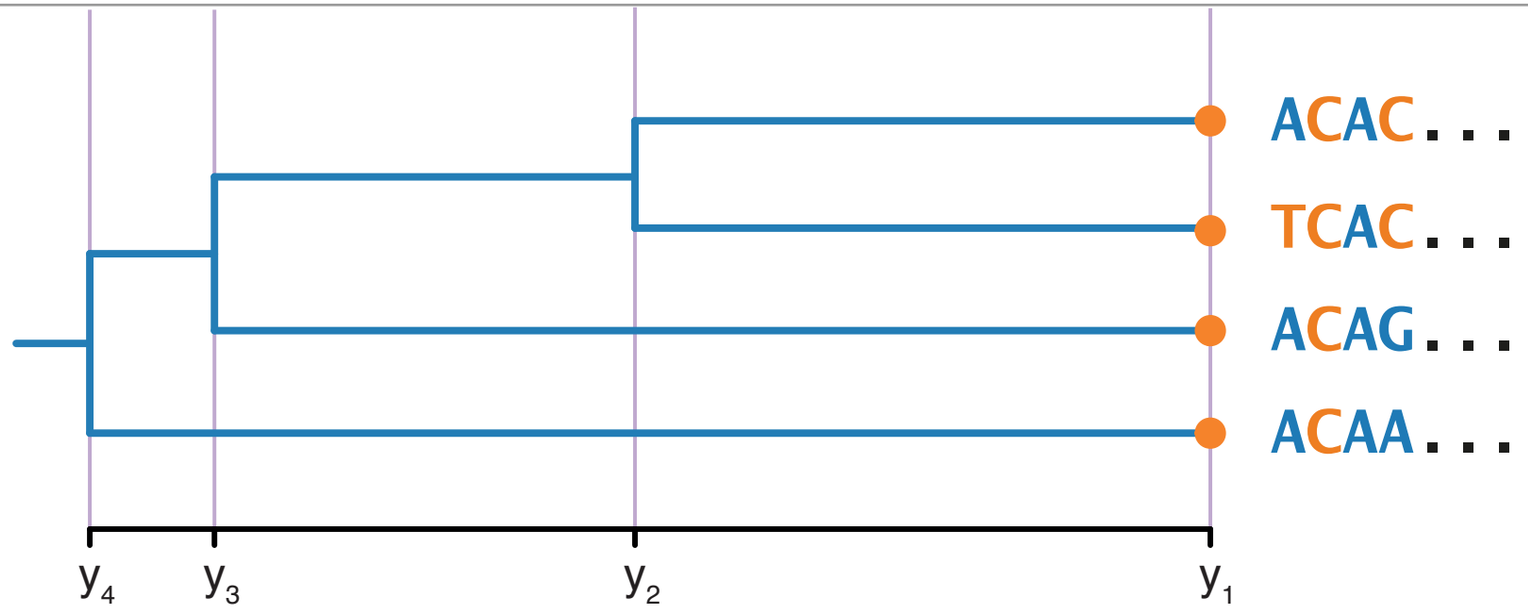
Add parameter, add hyperprior…

Example: HKY-model (HKY85)



Fix κ — remove prior!

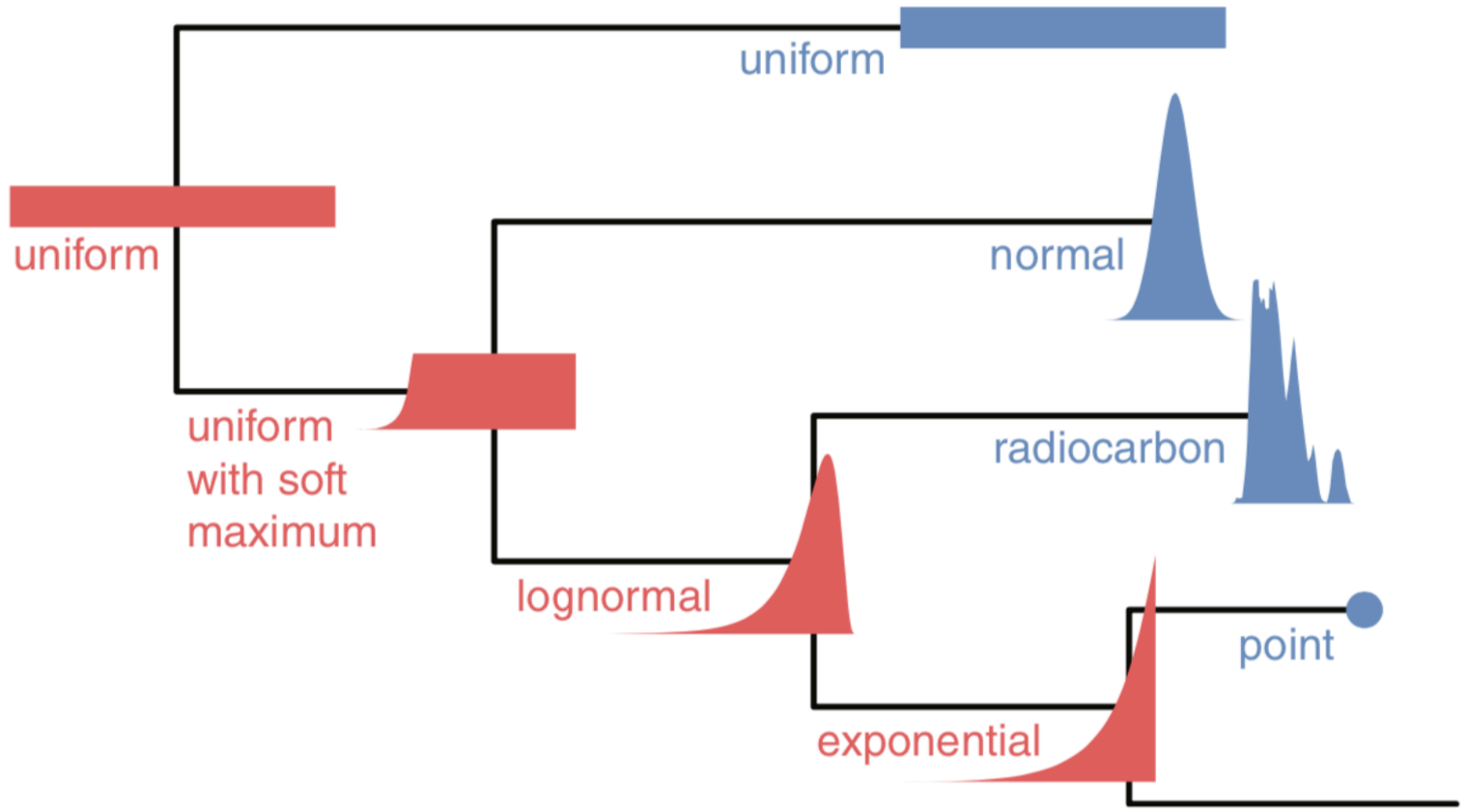
Calibration nodes



Homochronous trees need external **(prior)** information to date:

- Fix the clock rate
- Use a **calibration node** (or nodes)

Calibration nodes

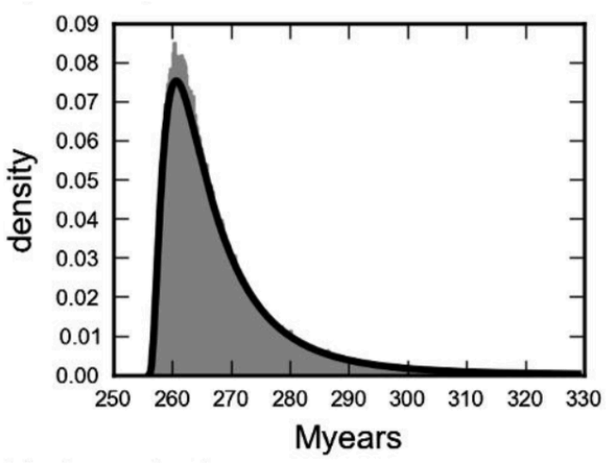


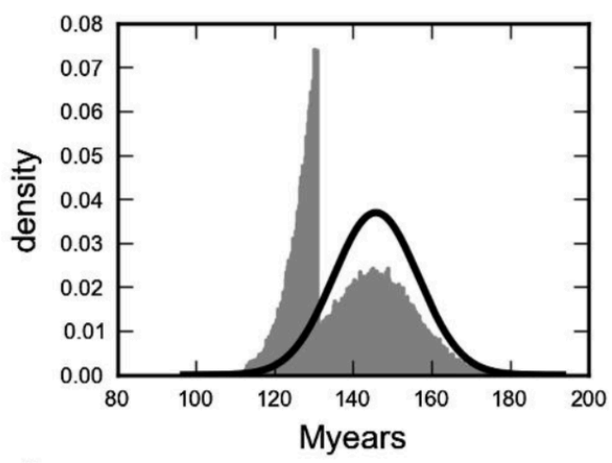
Bromham *et al.* **2018**

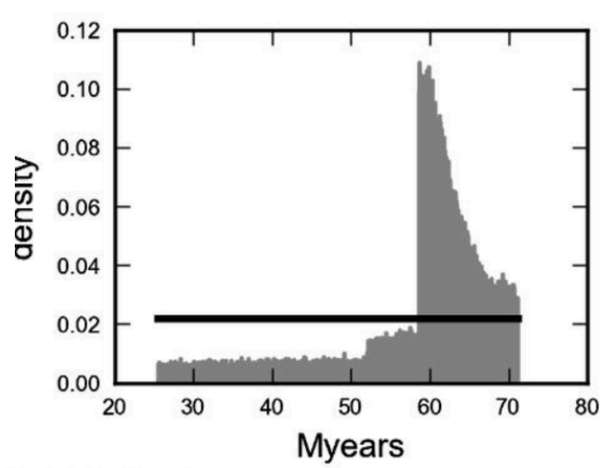
Calibration nodes

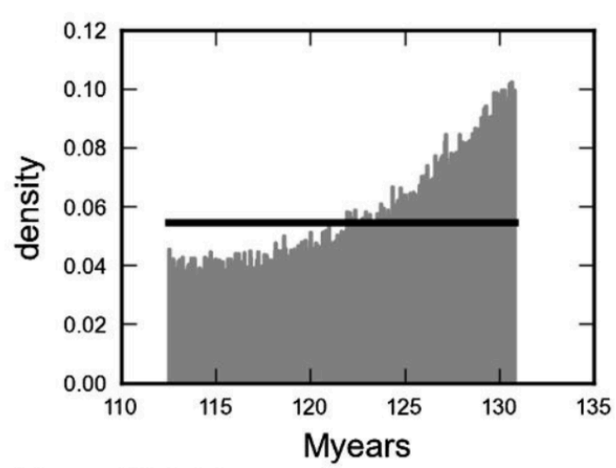
More calibration nodes mean more information to calibrate the clock

but is more better?





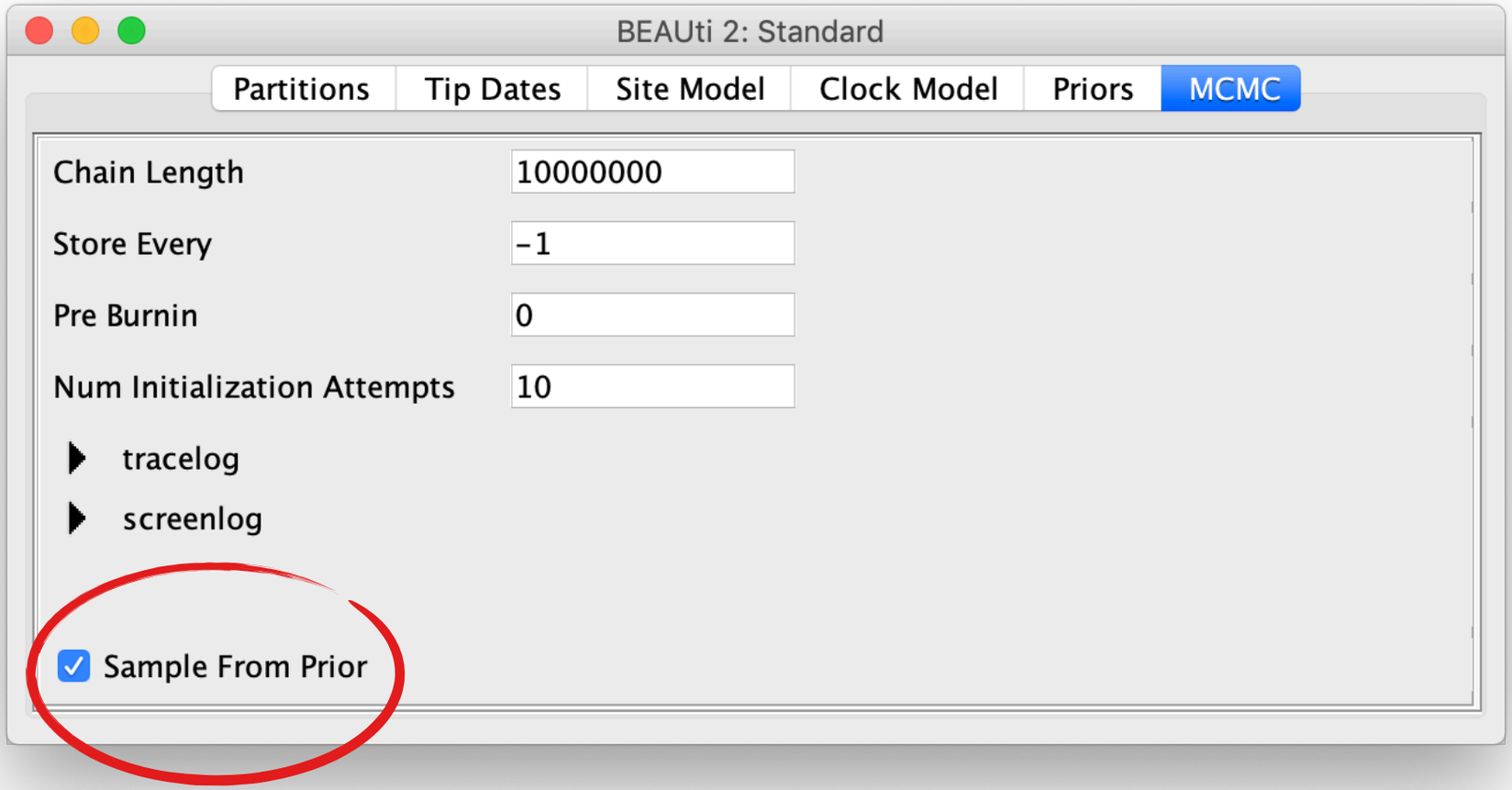




Induced prior may be different to the prior you set!

Heled and Drummond **2012**

Setting priors best practice



- Sample or simulate from prior
- For key distributions plot prior and posterior together
- Mess with priors and see how sensitive/robust posterior is to prior perturbations