

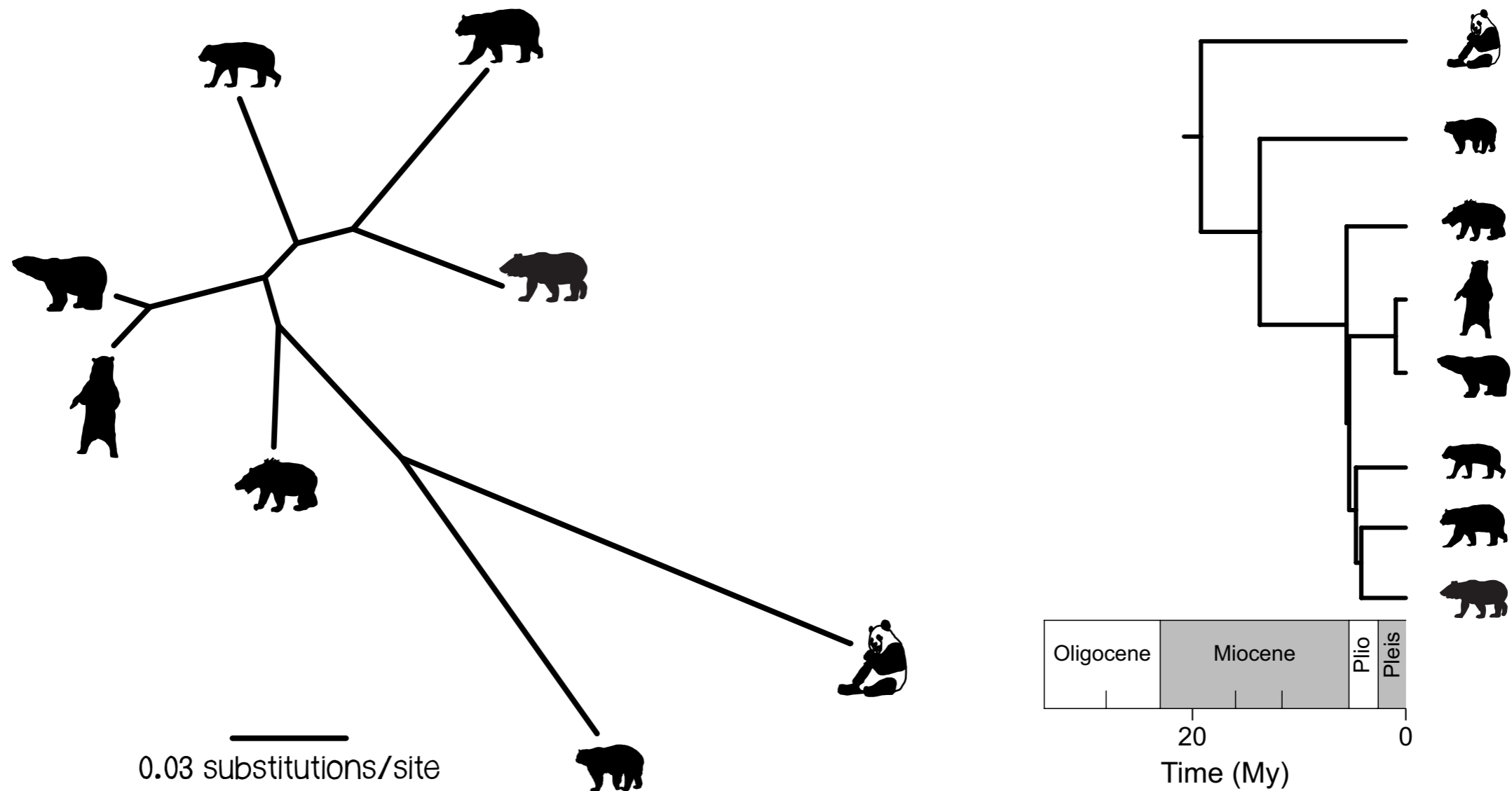


# Bayesian Divergence-Time Estimation



# A Time-Scale for Evolution

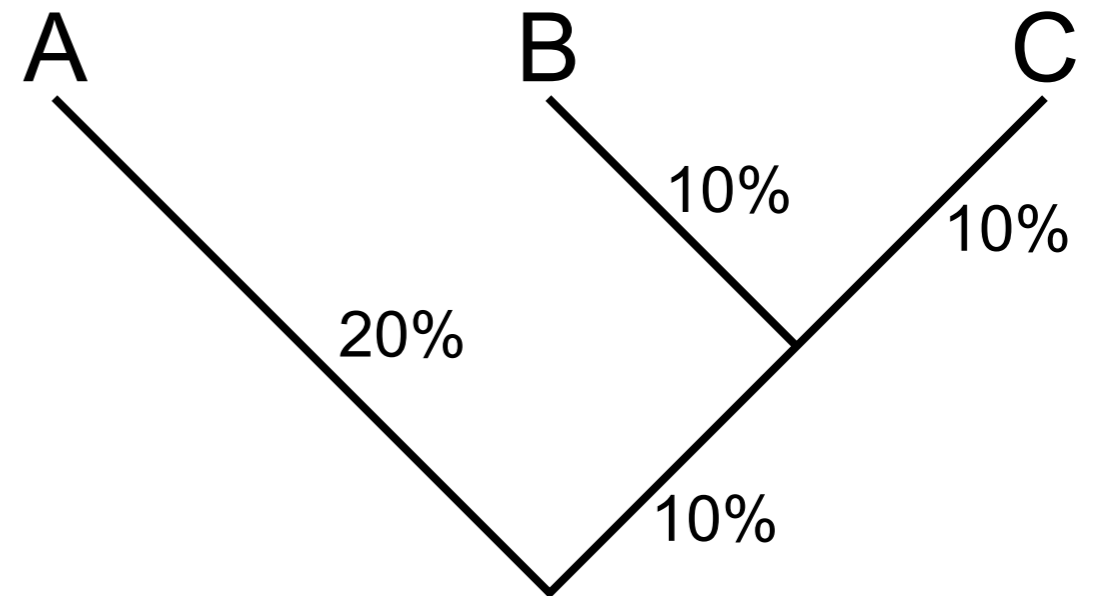
Phylogenies with branch lengths proportional to time provide more information about evolutionary history than unrooted trees with branch lengths in units of substitutions/site.



# The Global Molecular Clock

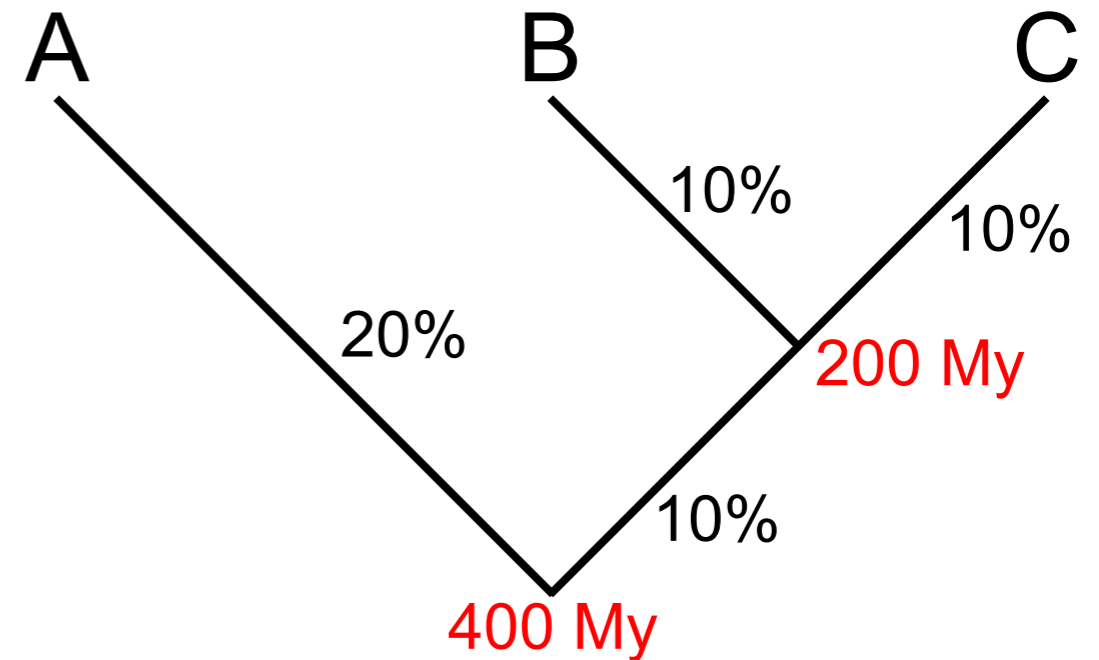
Assume that the rate of evolutionary change is constant over time

(branch lengths equal percent sequence divergence)



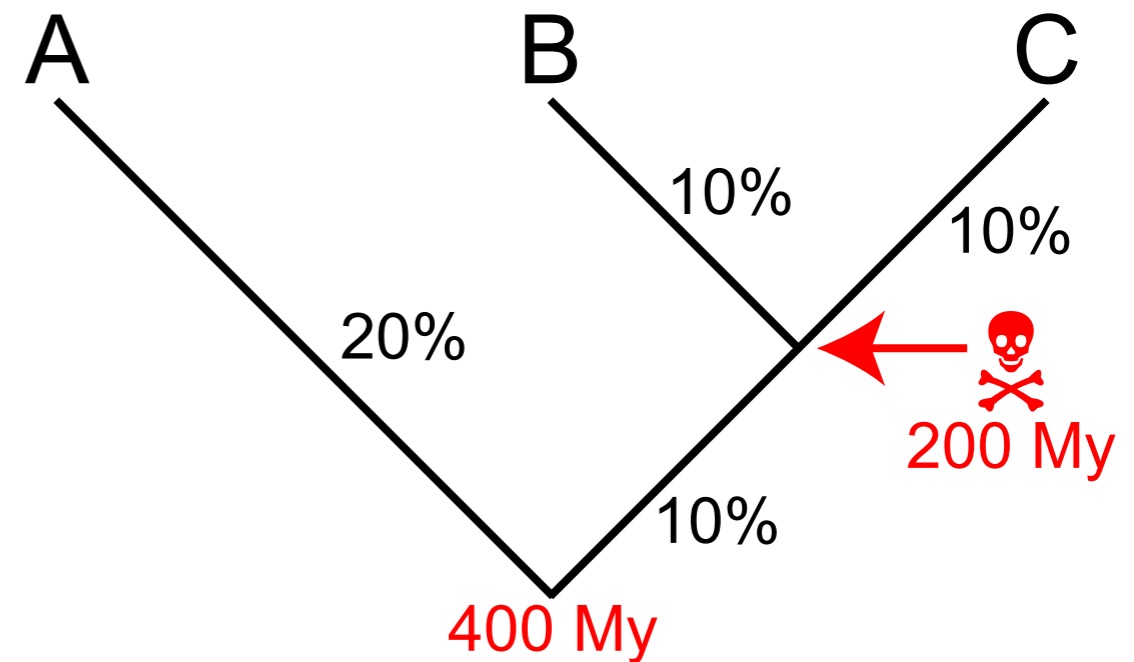
# The Global Molecular Clock

We can date the tree if we know the rate of change is 1% divergence per 10 My



# The Global Molecular Clock

If we found a fossil of the MRCA of **B** and **C**, we can use it to calculate the rate of change & date the root of the tree



# Rejecting the Global Molecular Clock

Rates of evolution vary across lineages and over time

## **Mutation rate**

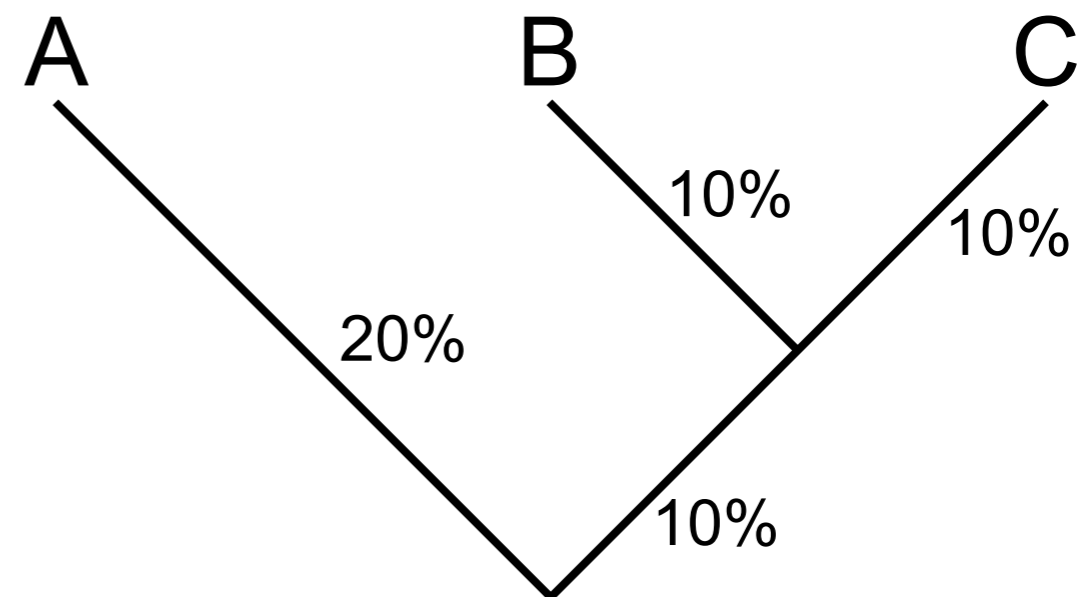
Variation in

- metabolic rate
- generation time
- DNA repair

## **Fixation rate**

Variation in

- strengths/targets of selection
- population sizes

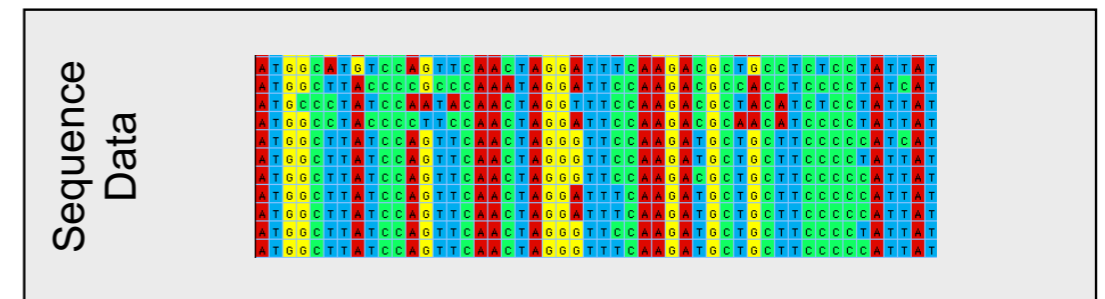


# Unconstrained Analysis

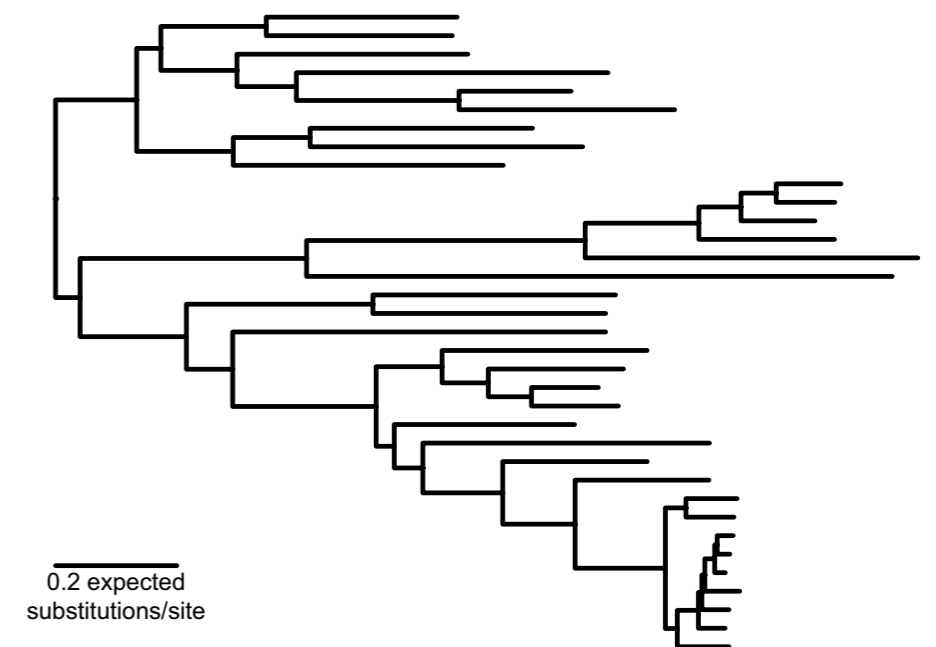
Sequence data provide information about **branch lengths**

In units of the **expected # of substitutions per site**

branch length = rate  $\times$  time



Phylogenetic Relationships

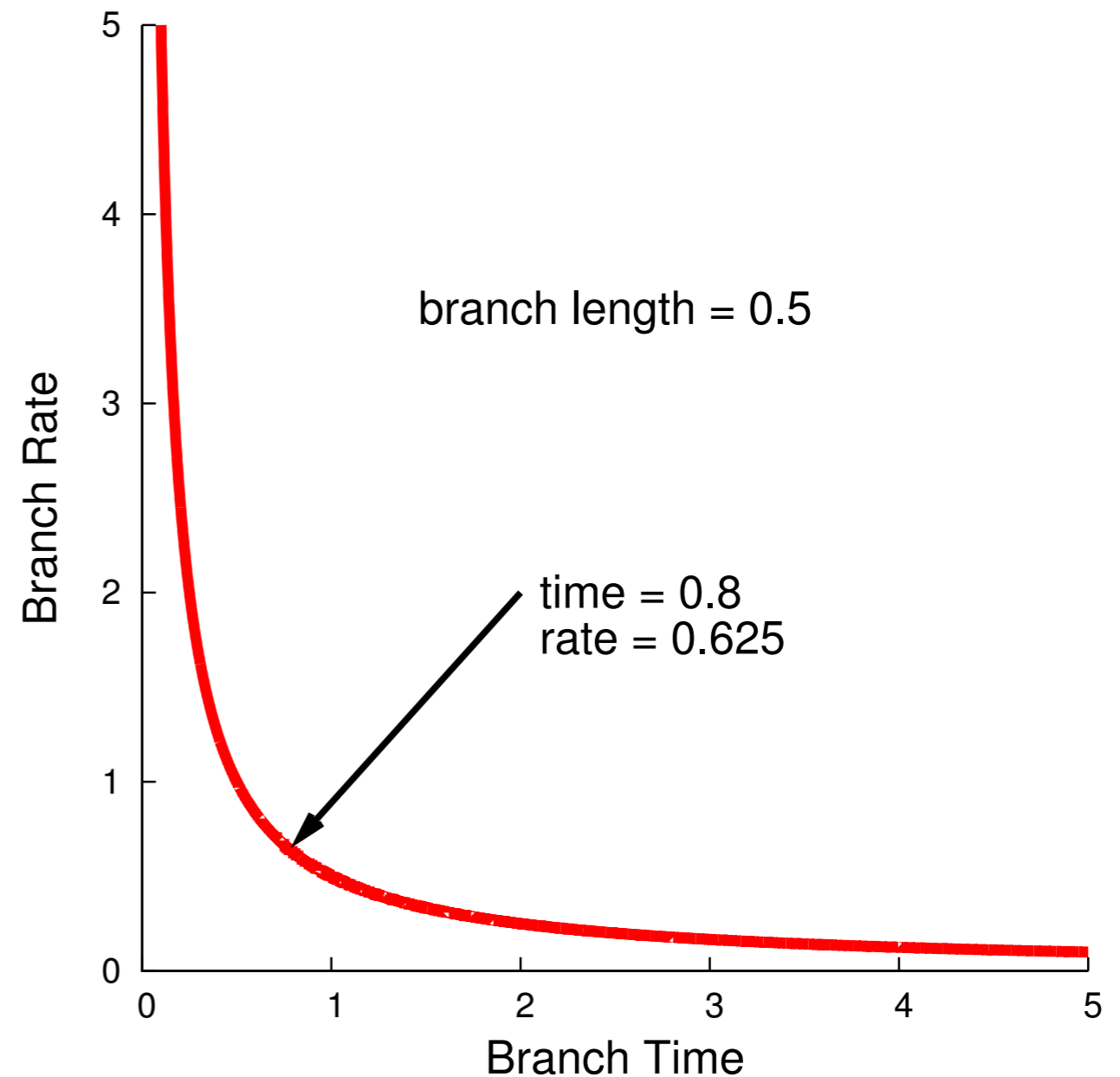




# Estimating Rate & Time

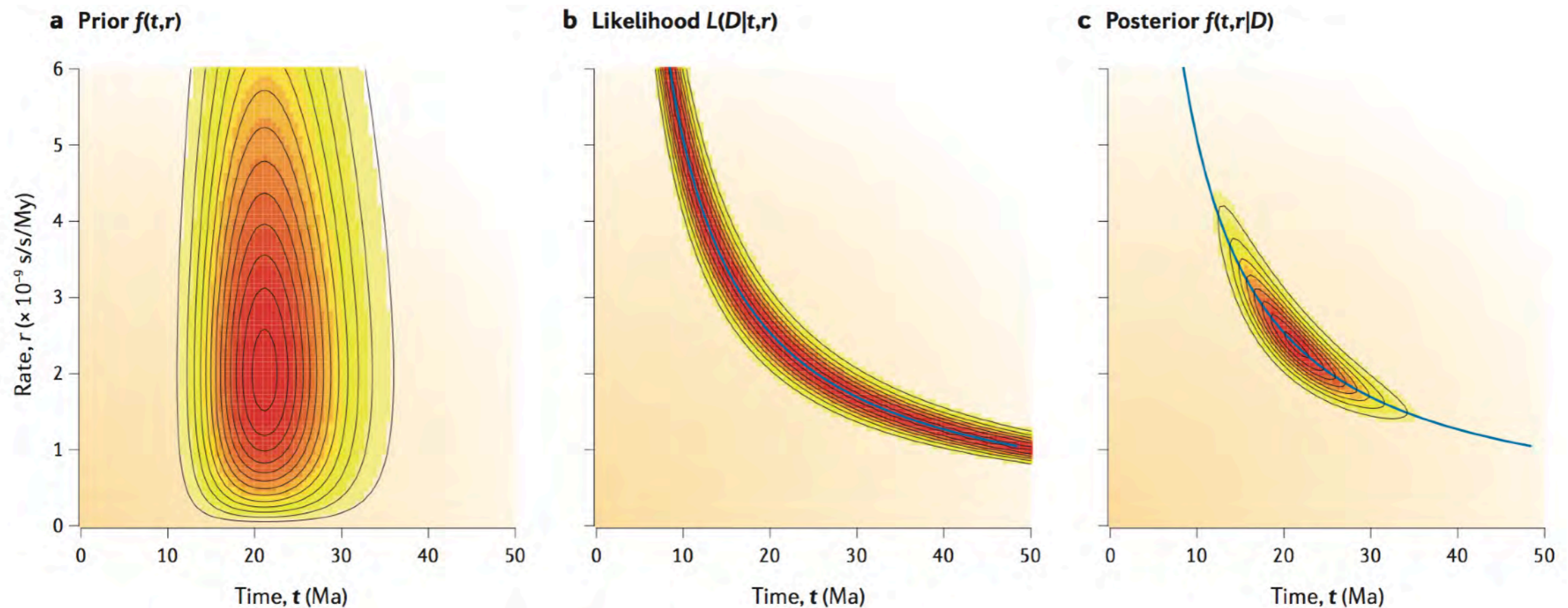
The sequence data provide information about branch length

for any possible rate, there's a time that fits the branch length perfectly



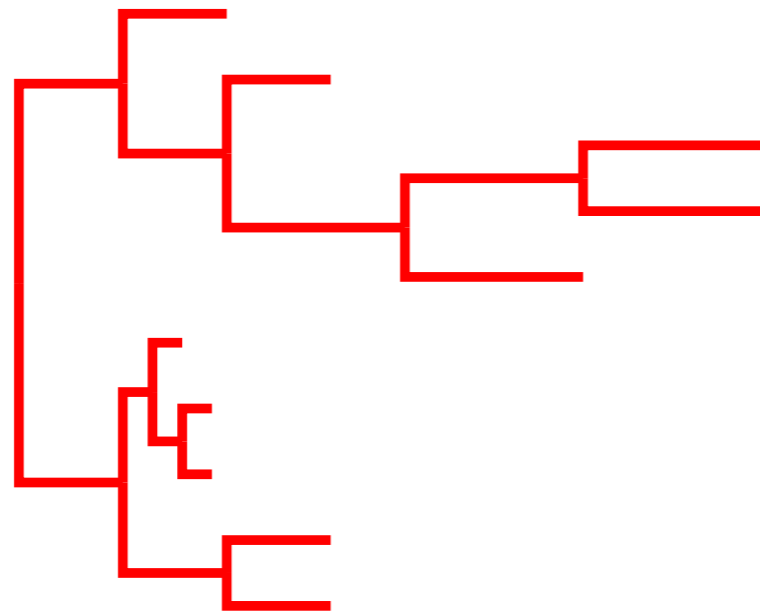
# Estimating Rate & Time

Methods for dating species divergences estimate the **substitution rate** and **time** separately

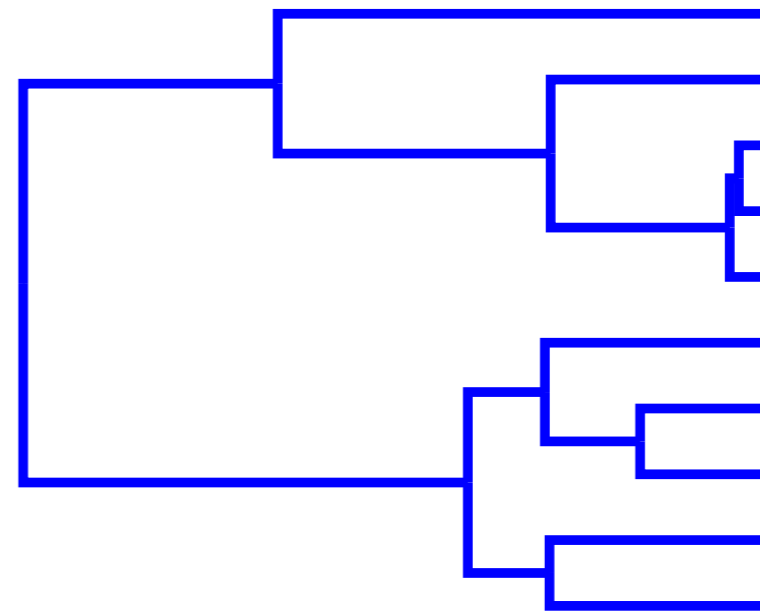


Tree-time priors for molecular phylogenies are only informative on a **relative** time scale

# Bayesian Divergence Time Estimation



length = rate

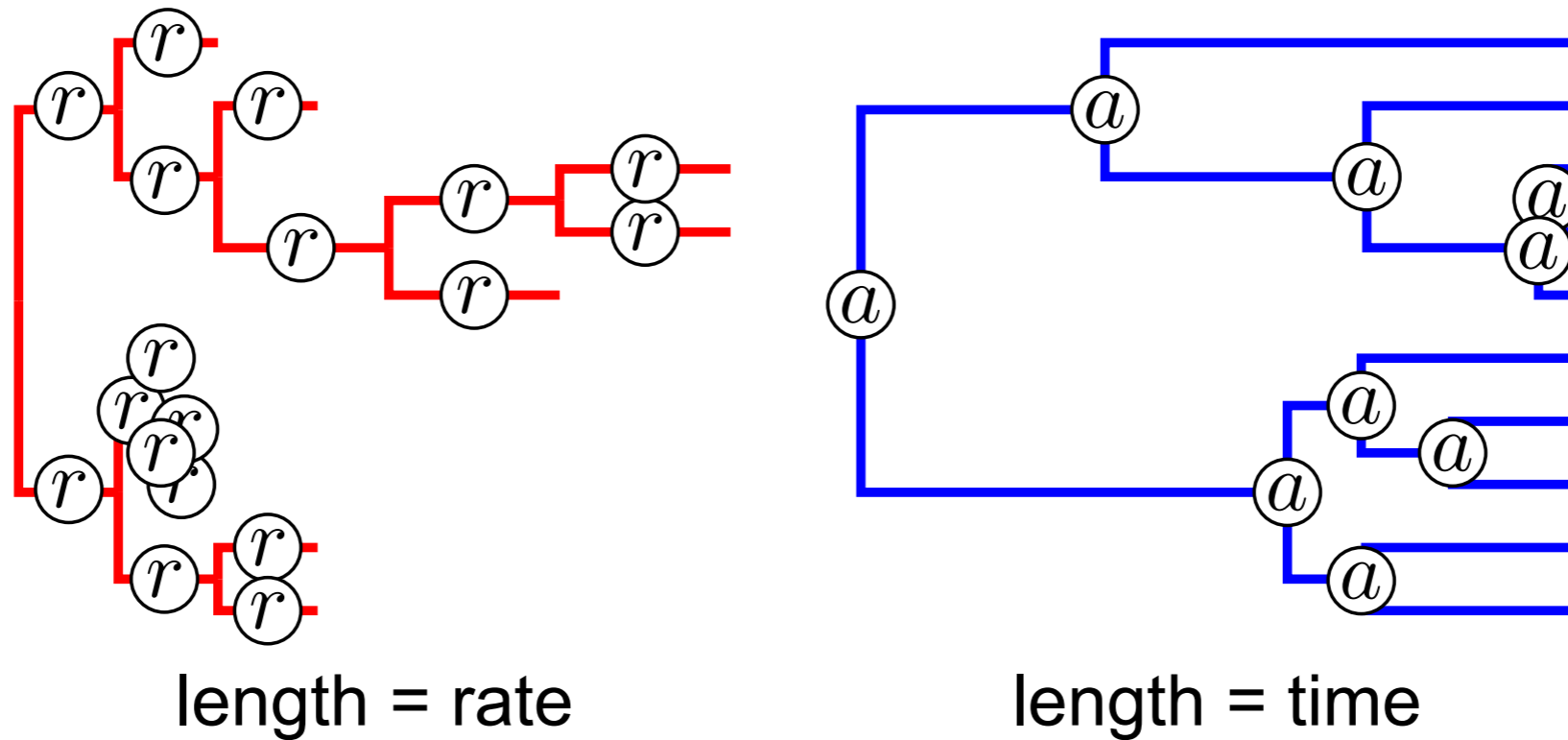


length = time

$$\mathbf{R} = (r_1, r_2, r_3, \dots, r_{2N-2})$$

$$\mathbf{A} = (a_1, a_2, a_3, \dots, a_{N-1})$$

# Bayesian Divergence Time Estimation



$$\mathbf{R} = (r_1, r_2, r_3, \dots, r_{2N-2})$$

$$\mathbf{A} = (a_1, a_2, a_3, \dots, a_{N-1})$$

# Bayesian Divergence Time Estimation

Posterior probability

$$f(\mathbf{R}, \mathbf{A}, \Psi, \theta_R, \theta_A, \theta_S | D)$$

$\mathbf{R}$	Vector of rates on branches
$\mathbf{A}$	Vector of internal node ages
$\theta_R, \theta_A, \theta_S$	Model Parameters
$D$	Character data
$\Psi$	Tree topology

# Bayesian Divergence Time Estimation

$$f(\mathbf{R}, \mathbf{A}, \Psi, \theta_R, \theta_A, \theta_s | D) =$$

$$\frac{f(D | \mathbf{R}, \mathbf{A}, \theta_s) f(\mathbf{R} | \theta_R) f(\mathbf{A}, \Psi | \theta_A) f(\theta_s)}{f(D)}$$

$\mathbf{R}$  Vector of rates on branches

$\mathbf{A}$  Vector of internal node ages

$\theta_R, \theta_A, \theta_s$  Model Parameters

$D$  Character data

$\Psi$  Tree topology

# Modeling Rate Variation

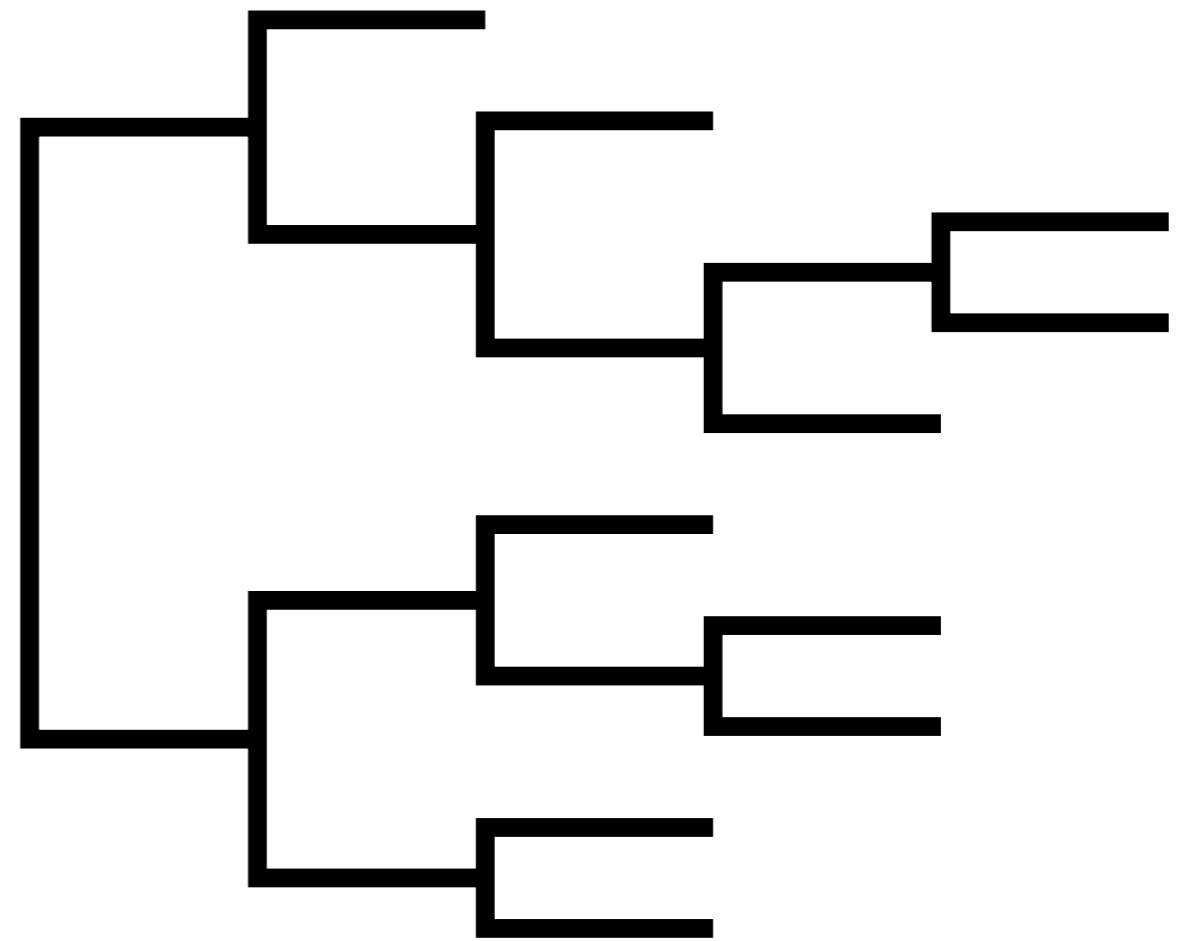
Models describing lineage-specific substitution rate variation:

- **Global/strict clock** ([Zuckermandl & Pauling, 1962](#))
- **Local molecular clocks** ([Hasegawa, Kishino & Yano 1989](#); [Kishino & Hasegawa 1990](#); [Yoder & Yang 2000](#); [Yang & Yoder 2003](#), [Drummond and Suchard 2010](#))
- **Punctuated rate change model** ([Huelsenbeck, Larget and Swofford 2000](#))
- **Autocorrelated rates** ([Thorne, Kishino & Painter 1998](#); [Kishino, Thorne & Bruno 2001](#); [Thorne & Kishino 2002](#); [Lepage et al. 2007](#))
- **Mixture models on branch rates** ([Heath, Holder, Huelsenbeck 2012](#))
- **Uncorrelated/independent rates models** ([Drummond et al. 2006](#); [Rannala & Yang 2007](#); [Lepage et al. 2007](#))

# Global Molecular Clock

The substitution rate is constant over time

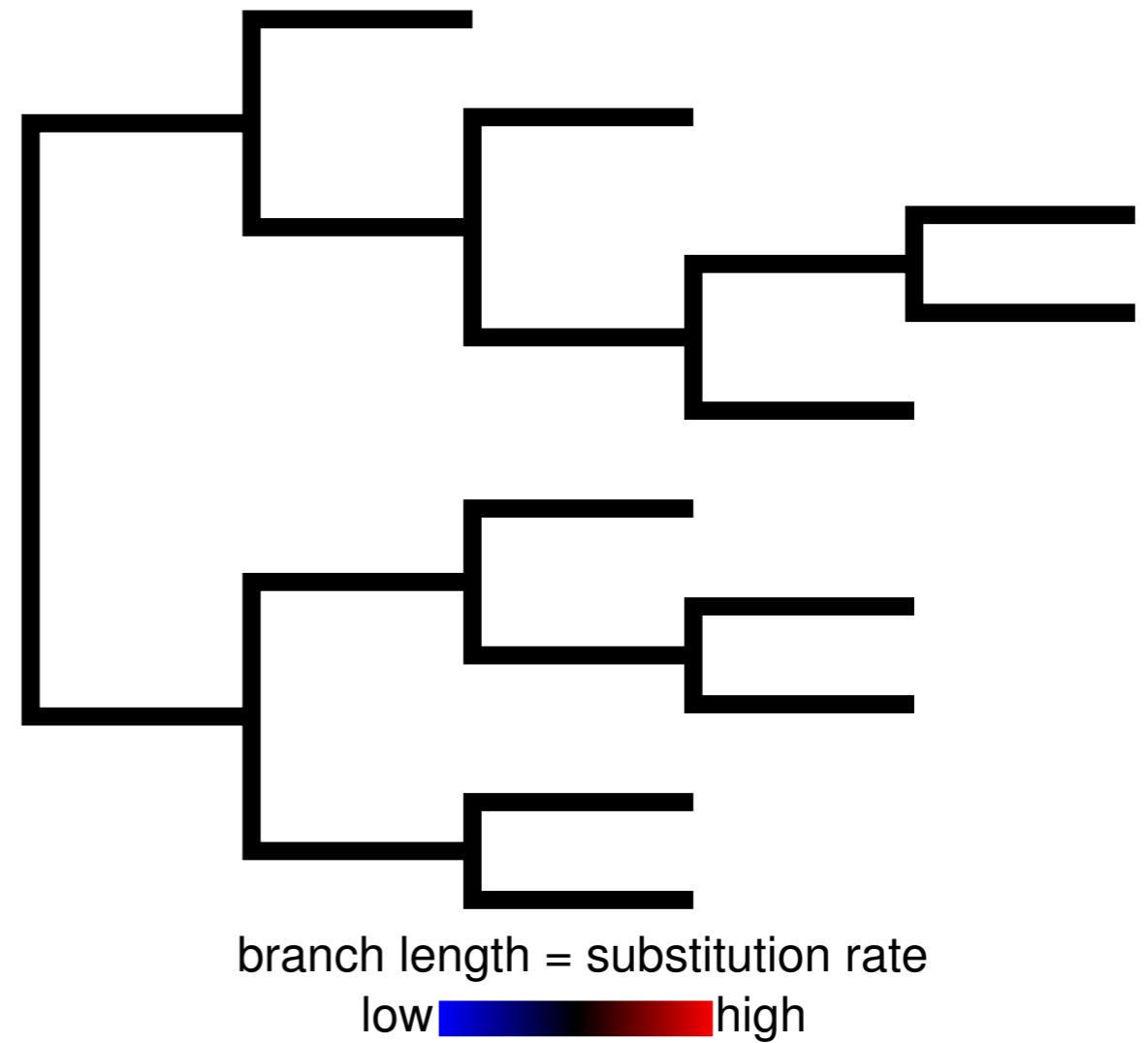
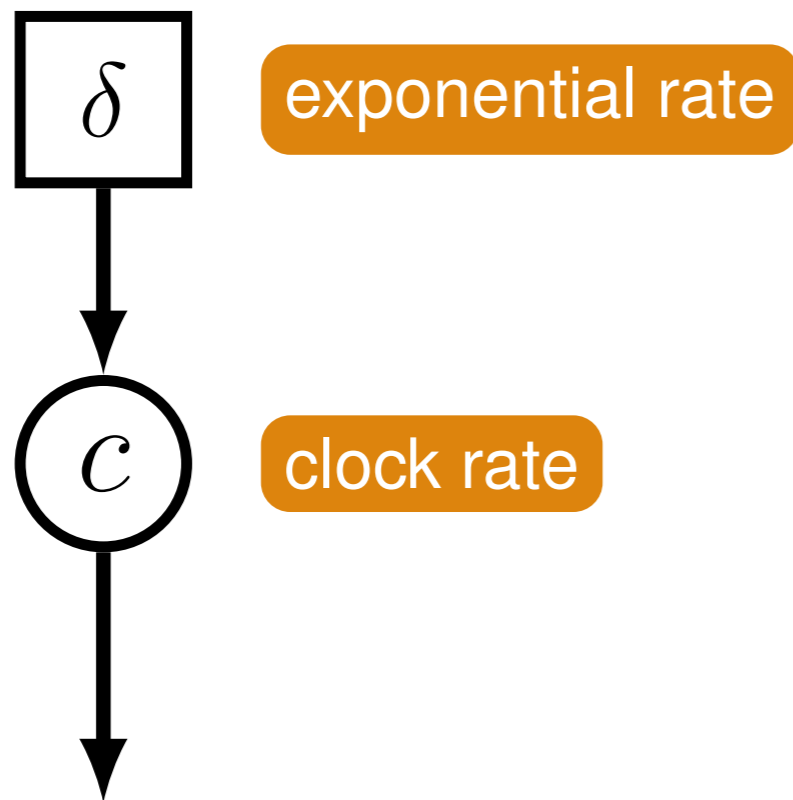
All lineages share the same rate





# Global Molecular Clock

$$c \sim \text{Exponential}(\delta)$$



# Global Molecular Clock

$$c \sim \text{Gamma}(\alpha, \beta)$$

$\alpha$

gamma shape

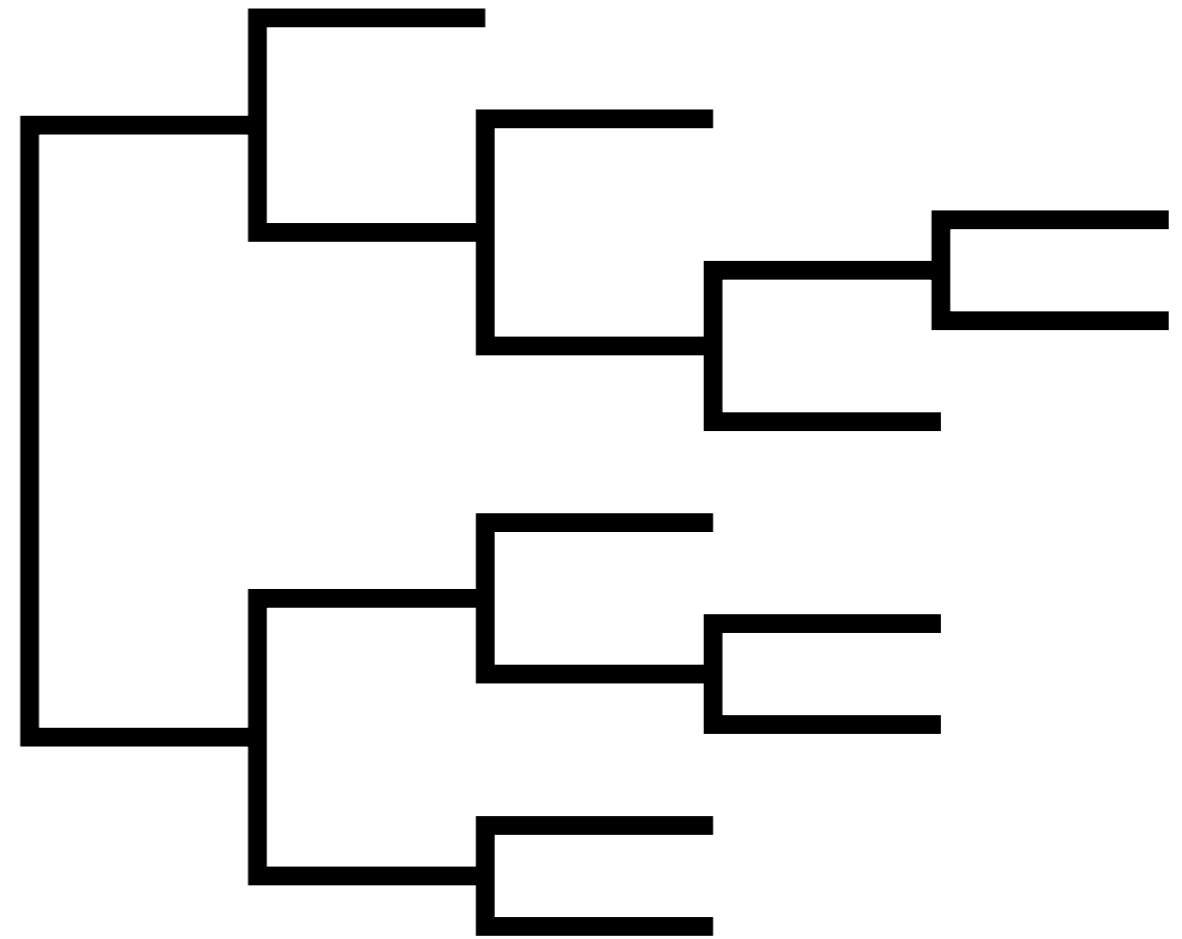
$\beta$


gamma scale

$c$

clock rate

Draw the graphical model!



branch length = substitution rate  
low  high

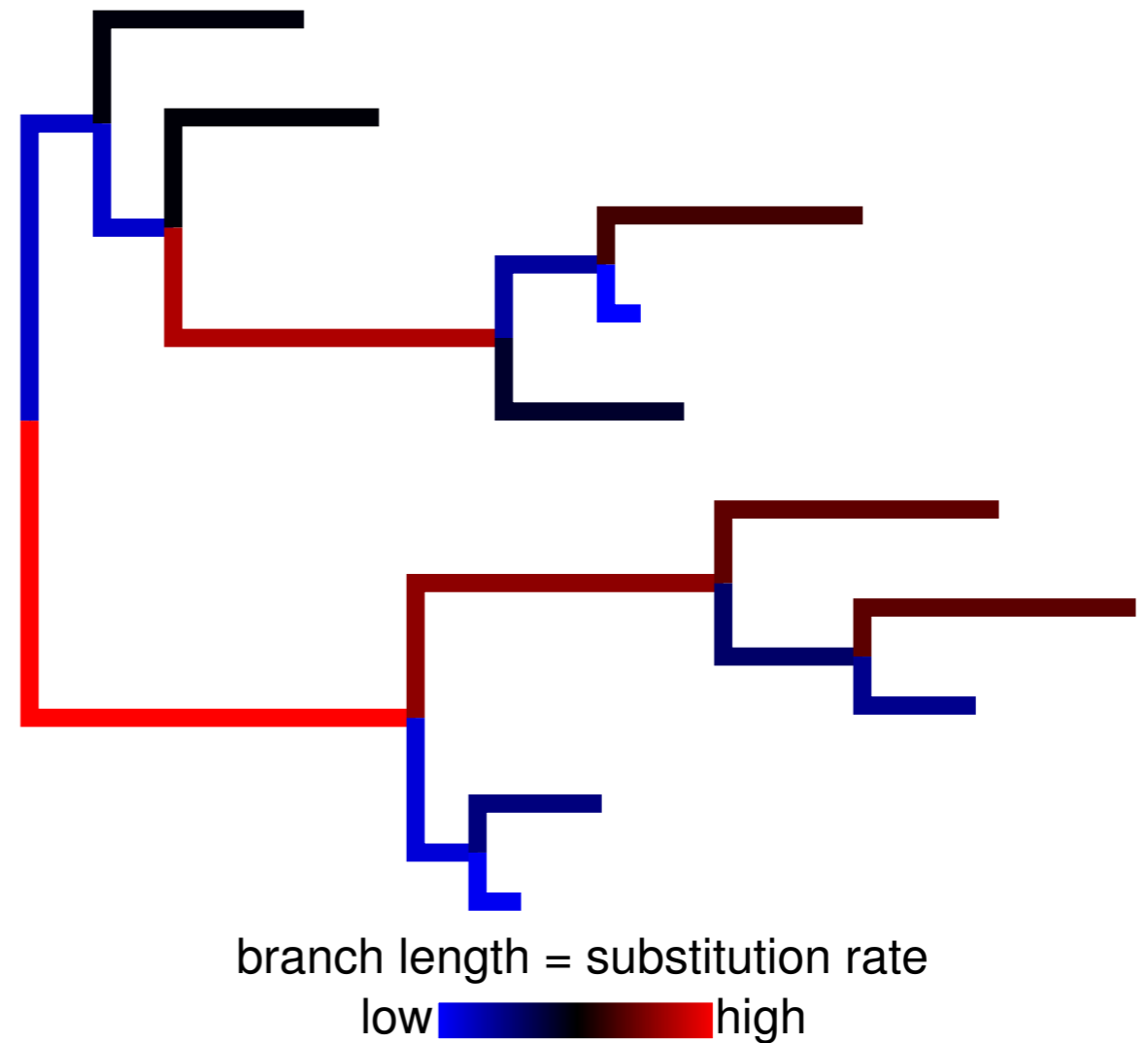
# Relaxed-Clock Models

To accommodate variation in substitution rates 'relaxed-clock' models estimate lineage-specific substitution rates

- **Global/strict clock**
- **Local clocks**
- **Punctuated rate change model**
- **Autocorrelated rates**
- **Mixture models on branch rates**
- **Uncorrelated/independent rates models**

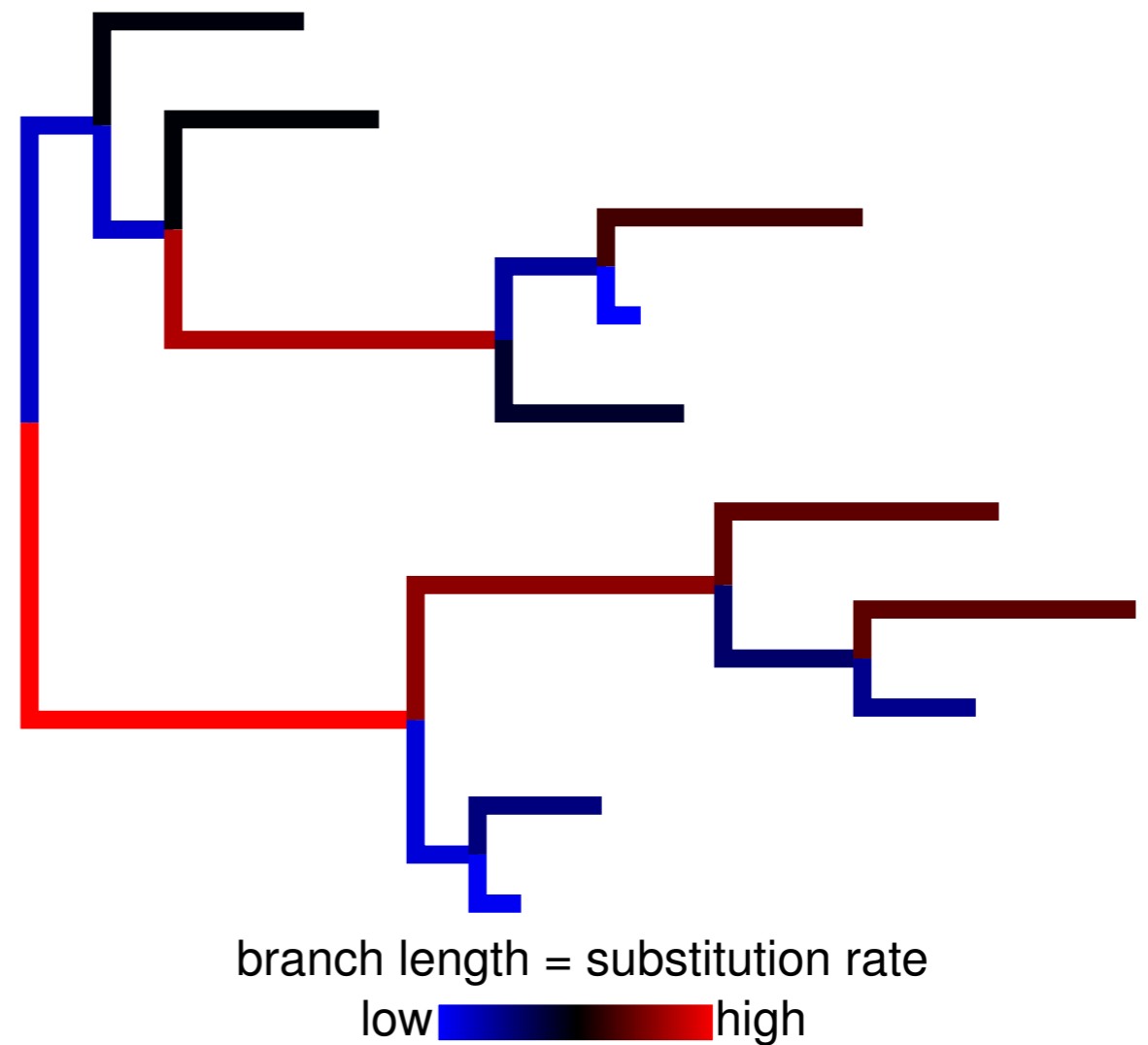
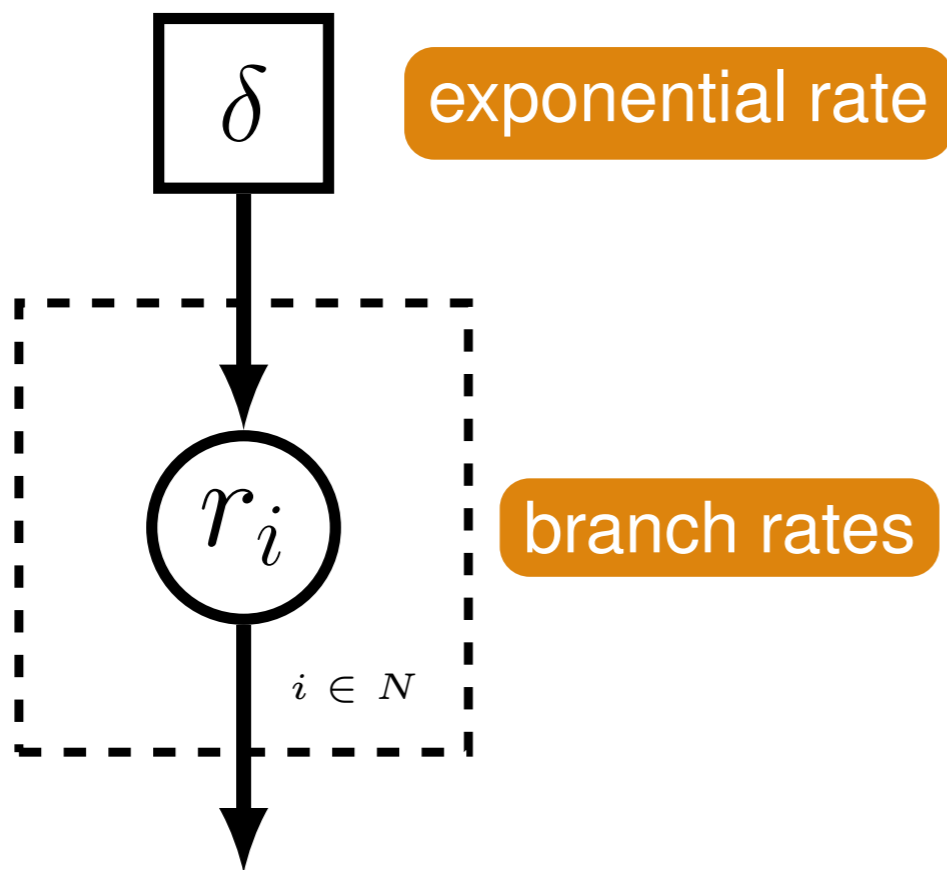
# Independent/Uncorrelated Rates

Lineage-specific rates are uncorrelated when the rate assigned to each branch is independently drawn from an underlying distribution



# Independent/Uncorrelated Rates

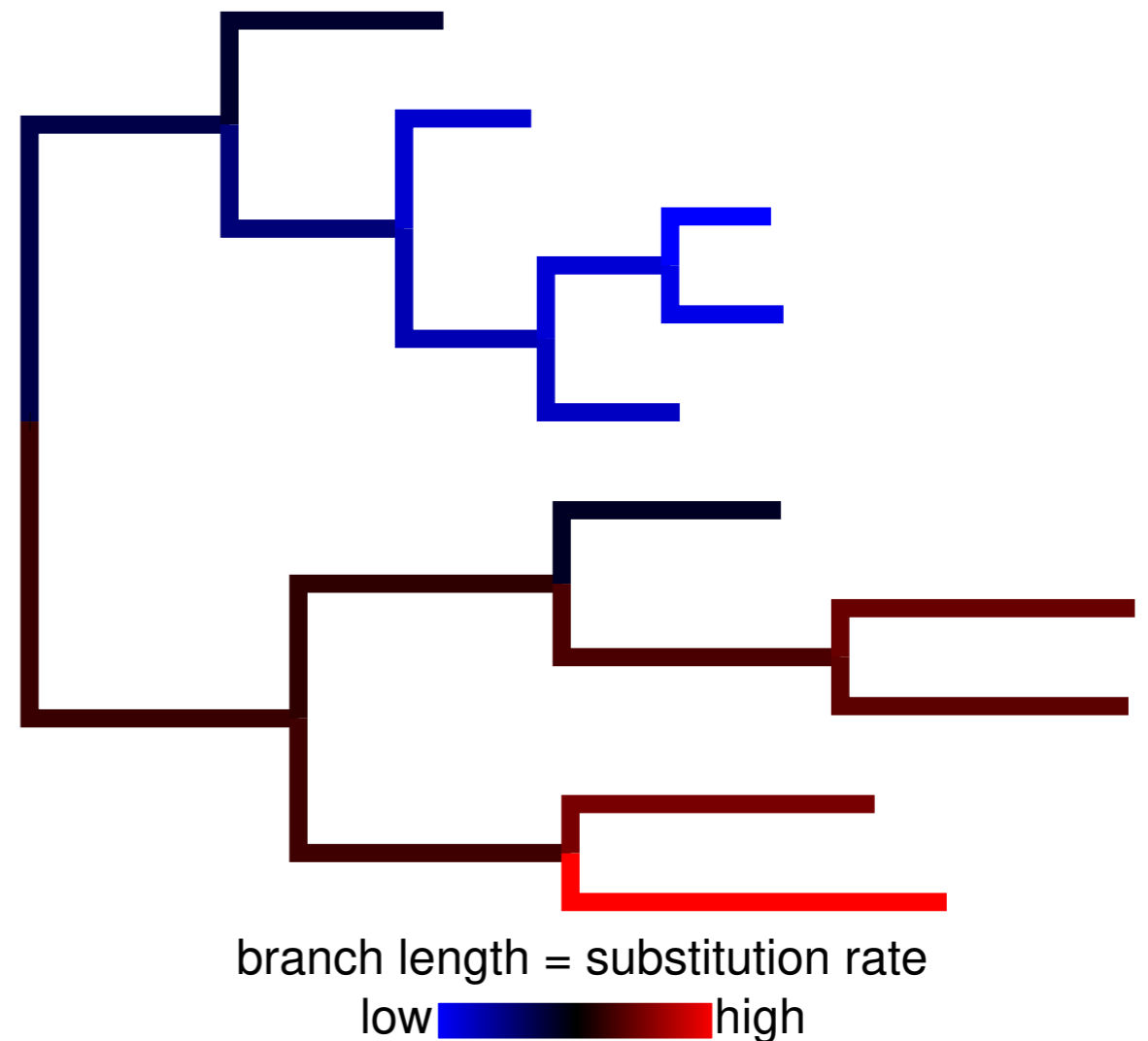
$$r_i \sim \text{Exponential}(\delta)$$



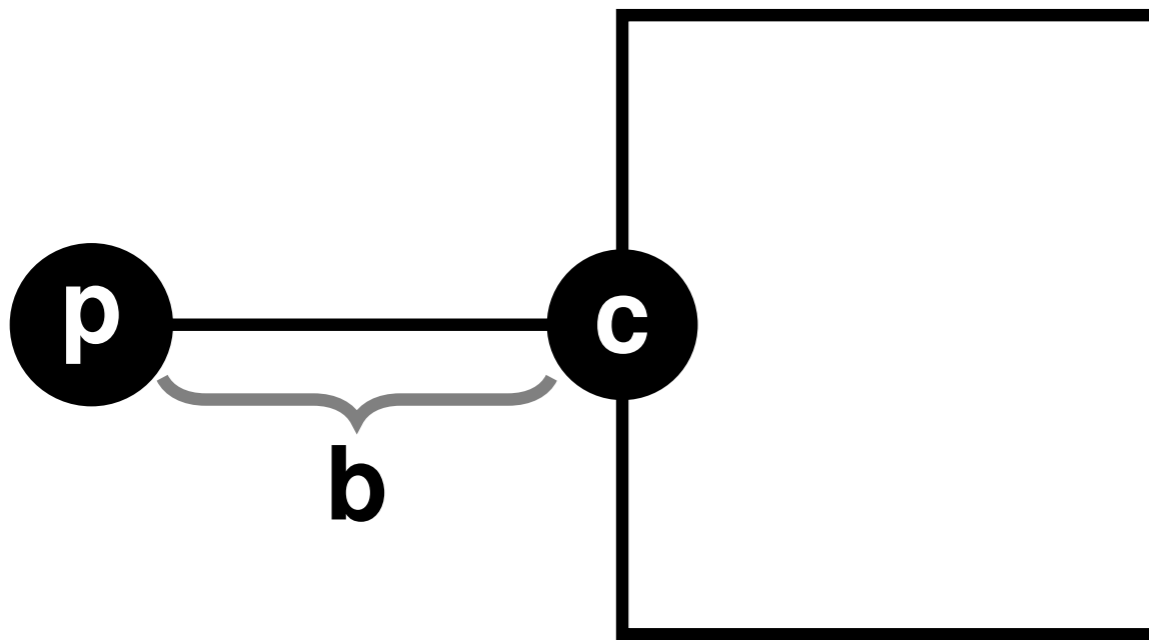
# Autocorrelated Rates

Substitution rates evolve gradually over time – closely related lineages have similar rates

The rate at a node is drawn from a distribution with a mean equal to the parent rate



# Autocorrelated Rates



p = parent node

c = child node

b = branch

$$r_c \sim \text{Lognormal}(\mu_c, \sigma_c)$$

$$\sigma_c := \nu t_b$$

$$\mu_c := \ln(r_p) - \frac{\sigma_c^2}{2}$$

$$r_b := \frac{r_p + r_c}{2}$$

$\nu$  = variance parameter

$t_b$  = time duration of branch

$$r_c$$

$$\mu_c$$

$$r_b$$

$$r_p$$

$$\sigma_c$$

$$t_b$$

$$\nu$$

# Autocorrelated Rates

Draw the graphical model!

$$r_c \sim \text{Lognormal}(\mu_c, \sigma_c)$$

$$\sigma_c := \nu t_b$$

$$\mu_c := \ln(r_p) - \frac{\sigma_c^2}{2}$$

$$r_b := \frac{r_p + r_c}{2}$$

$\nu$  = variance parameter

$t_b$  = time duration of branch

$$r_c$$

$$\mu_c$$

$$r_b$$

$$r_p$$

$$\sigma_c$$

$$t_b$$

$$\nu$$



# Relaxed-Clock Models

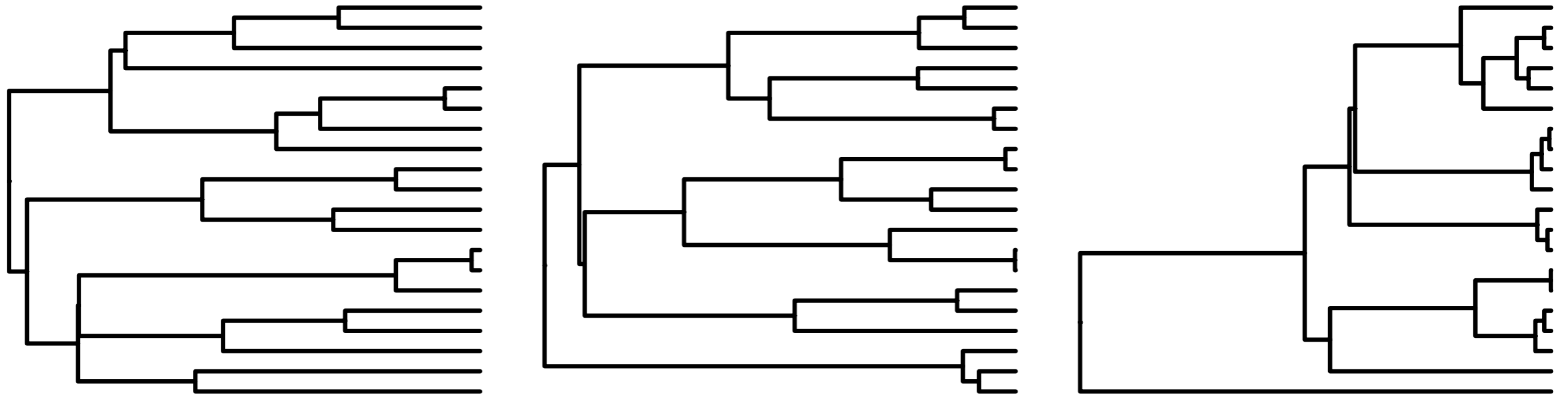
These are only a subset of the available models for branch-rate variation

- **Global/strict clock**
- **Local clocks**
- **Punctuated rate change model**
- **Autocorrelated rates**
- **Mixture models on branch rates**
- **Uncorrelated/independent rates models**

Considering model selection, uncertainty, & plausibility is **very** important for Bayesian divergence time analysis

# Priors on the Tree & Node Ages

Relaxed clock Bayesian analyses require a prior distribution on time trees



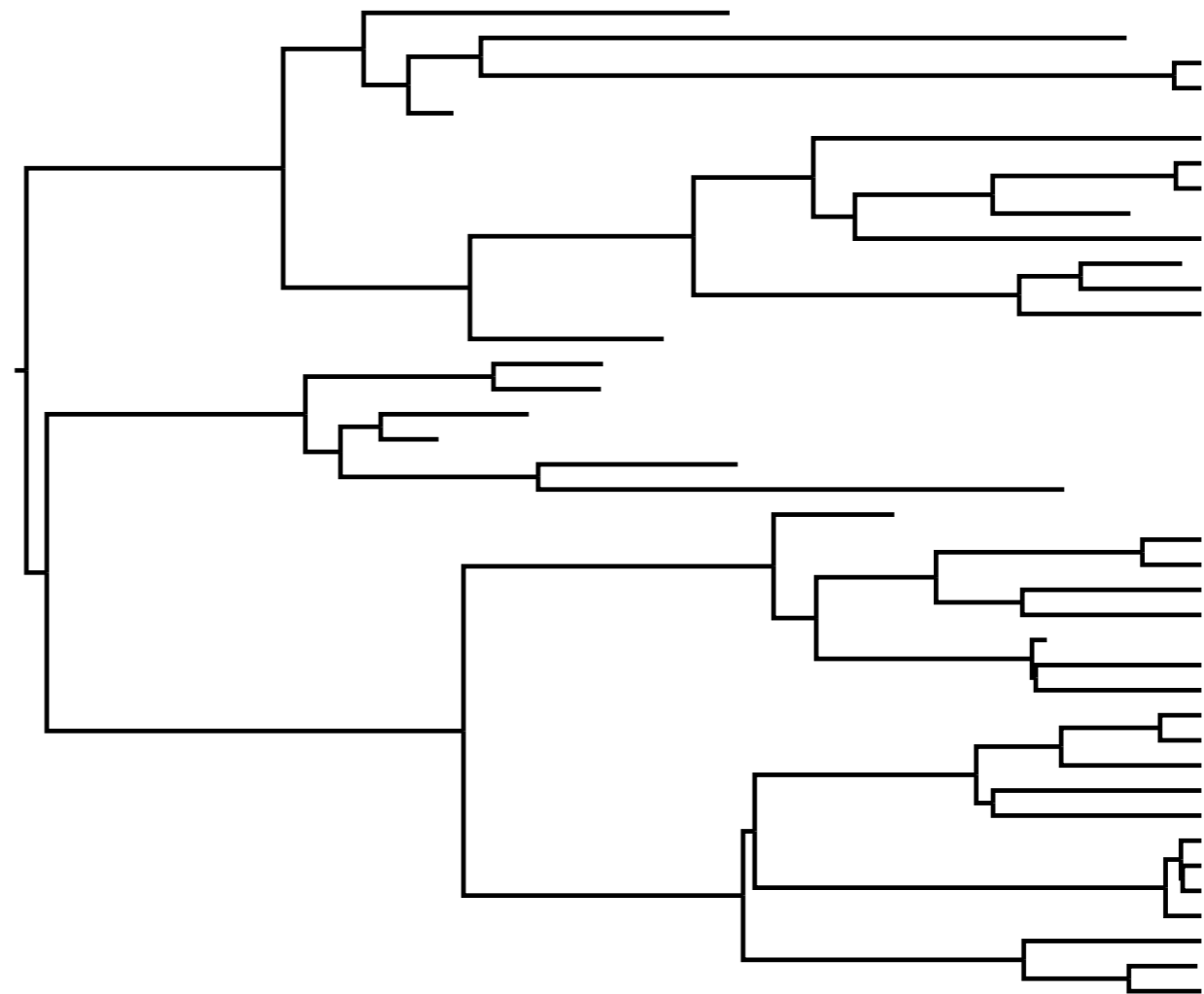
Different tree priors make different assumptions about the timing of divergence events and shape of the tree topology

# The Birth-Death Process

A stochastic branching process where new lineages arise at rate  $\lambda$  = births and go extinct at rate  $\mu$  = deaths ([Kendall 1948](#))

$\lambda$  = speciation

$\mu$  = extinction



Most commonly used tree priors are based on the birth-death process

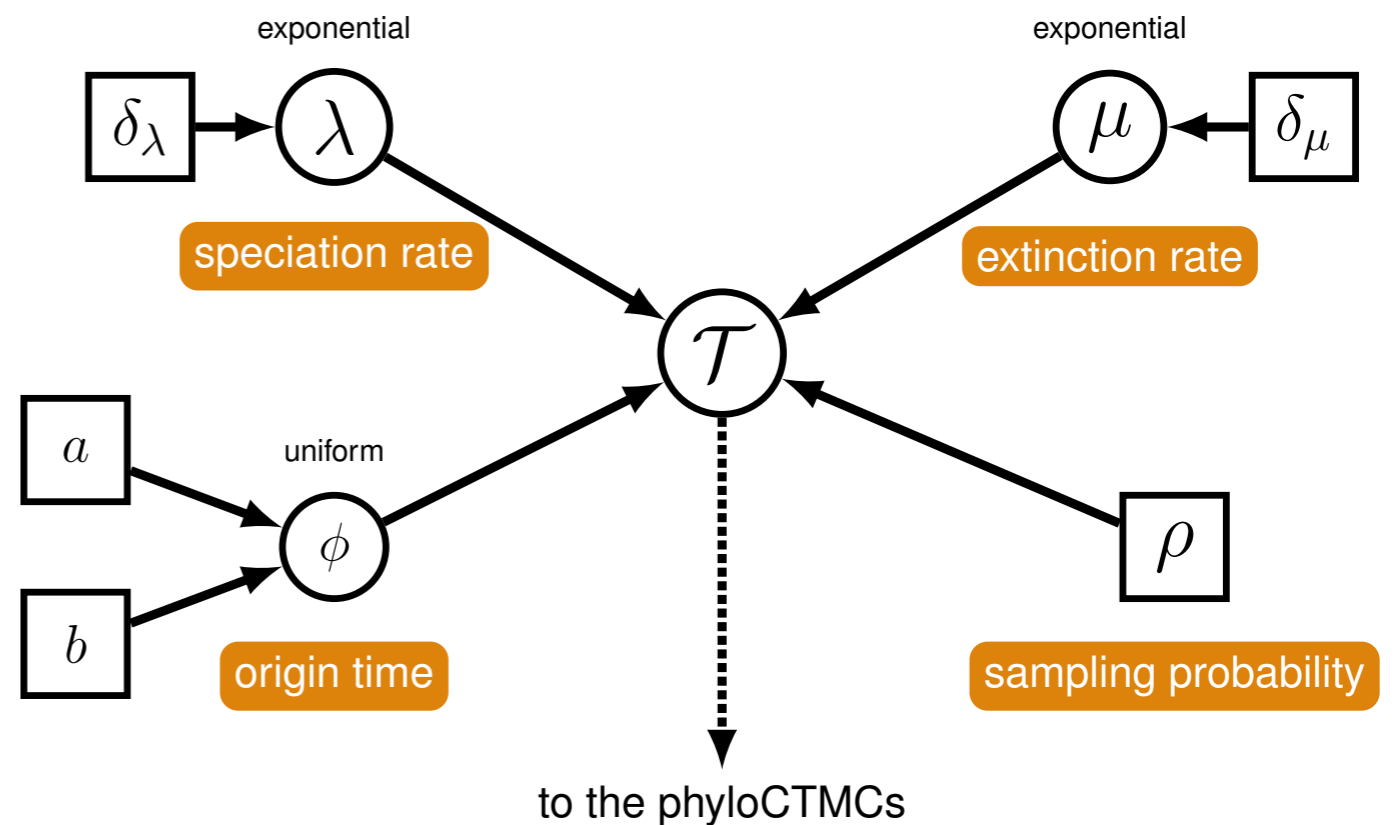
# The Birth-Death Process

Tree priors based on stochastic models of lineage diversification

## Birth-death-sampling

**process:** at any point in time a lineage can speciate at rate  $\lambda$  or go extinct with a rate of  $\mu$

Conditions on a probability of sampling a tip  $\rho$ , and the origin time of the process  $\phi$

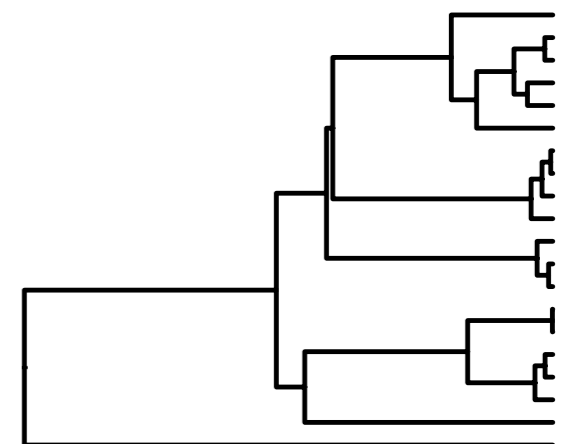
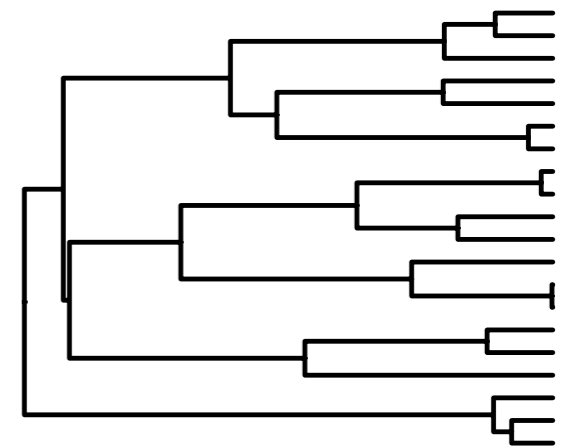
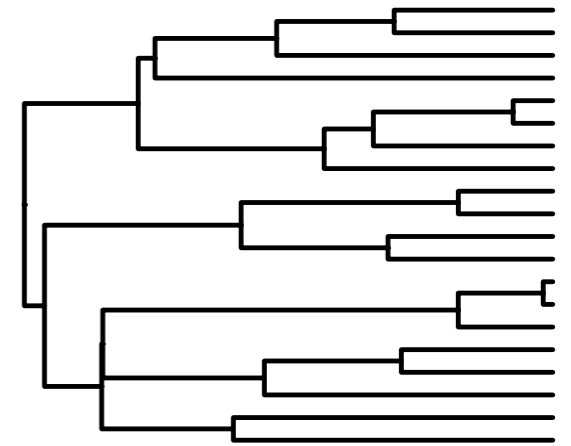


# The Birth-Death Process

Different values of  $\lambda$  and  $\mu$  lead to different trees

Bayesian inference under these models can be very sensitive to the values of these parameters

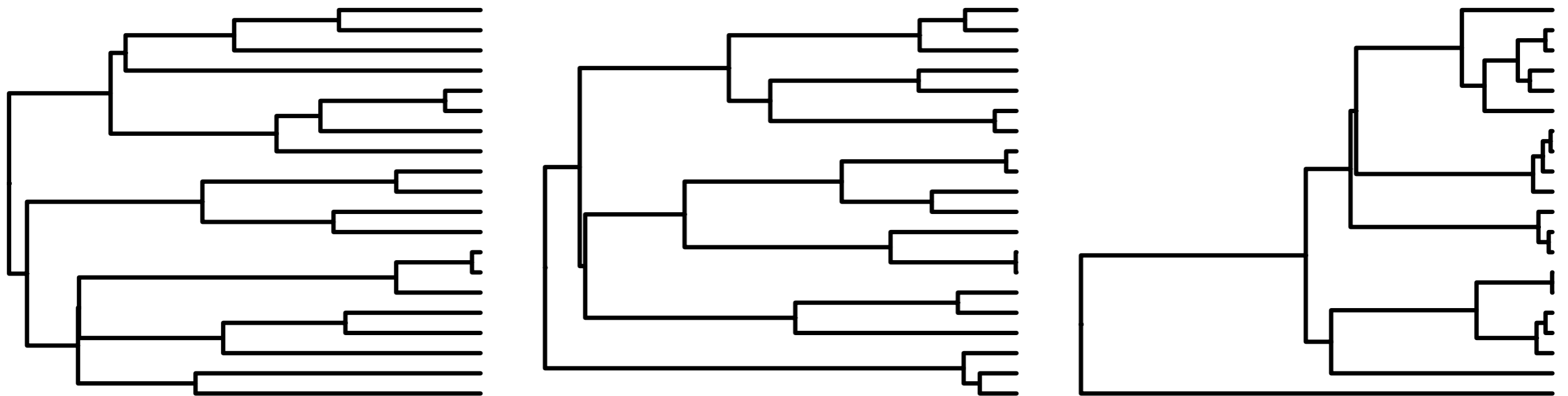
Using hyperpriors on  $\lambda$  and  $\mu$  accounts for uncertainty in these hyperparameters



# Priors on the Tree & Node Ages

Sequence data are only informative on *relative* rates & times

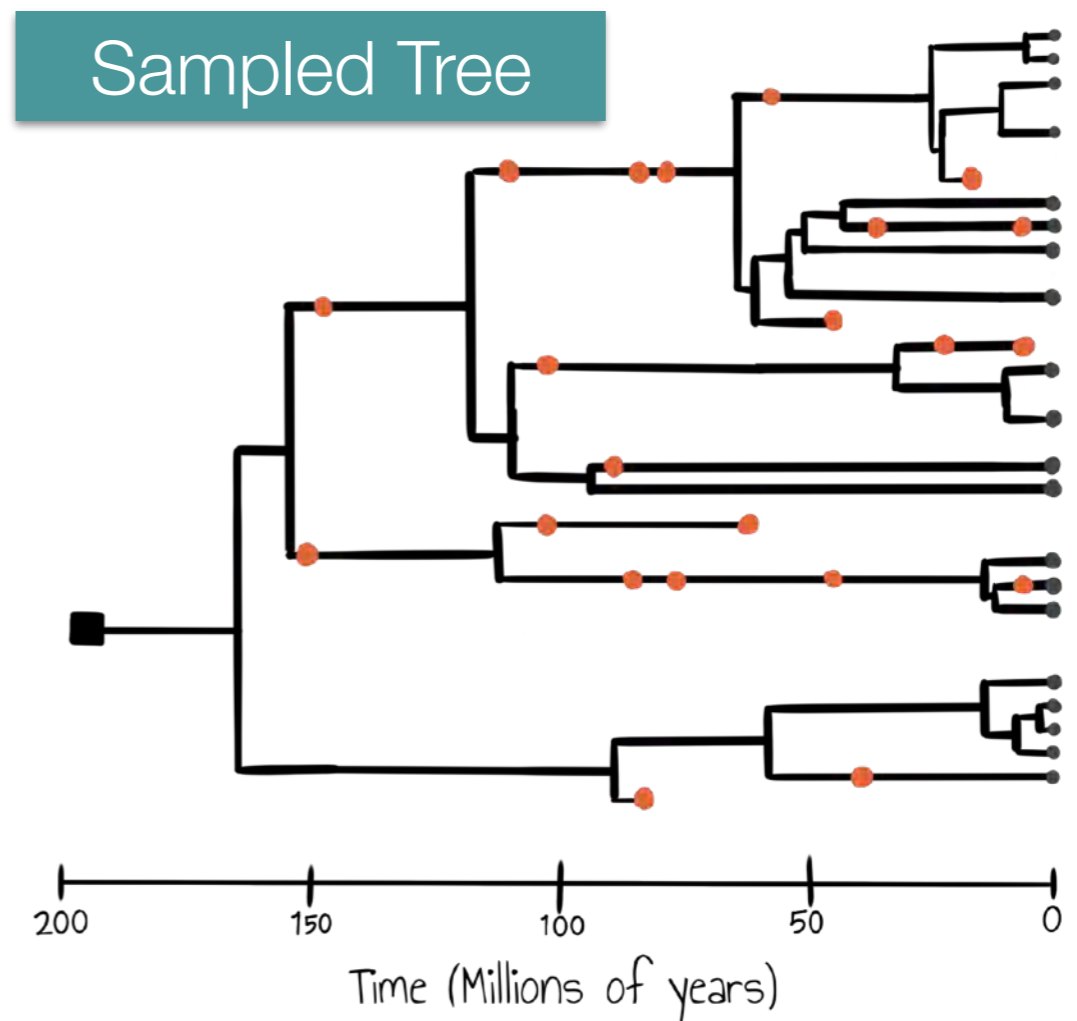
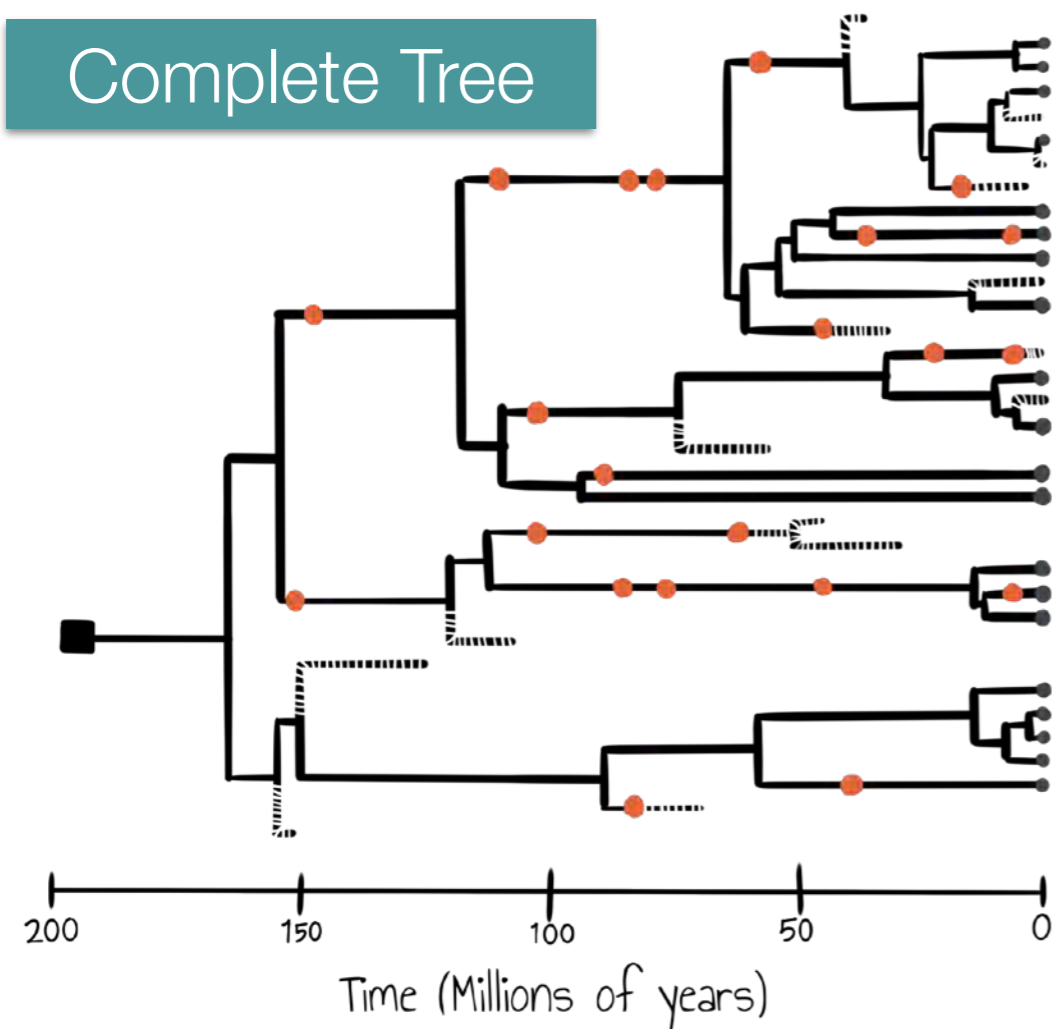
Most tree priors cannot give precise estimates of *absolute* node ages



We need additional data (like fossils) to provide an absolute time scale

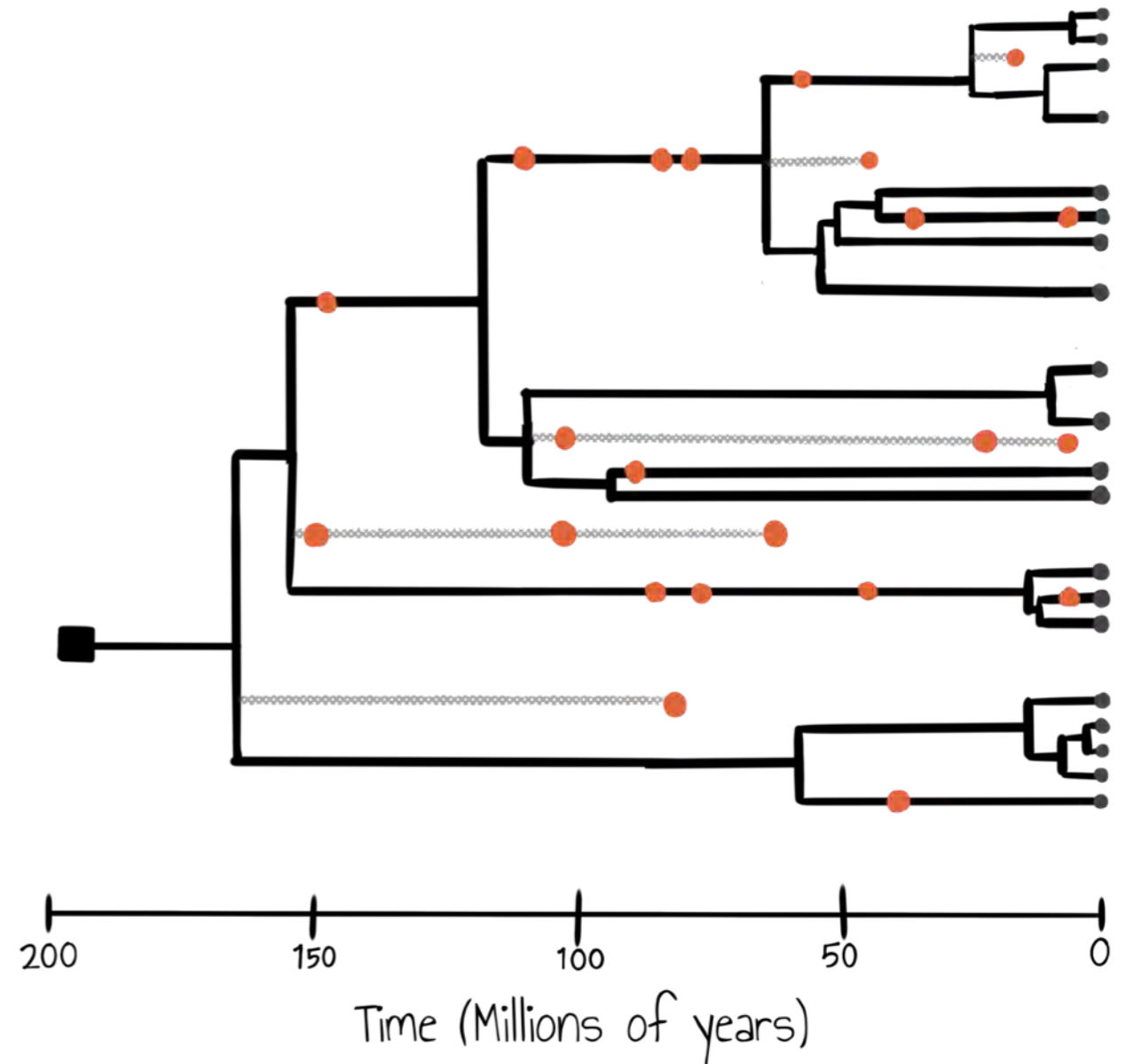
# Modeling the Tree & Occurrence Times

Stadler (2010) introduced a generating model for a serially sampled time tree – this is the *fossilized birth-death process*



# The Fossilized Birth-Death Process (FBD)

Recovered fossil specimens provide historical observations of the diversification process that generated the tree of extant species





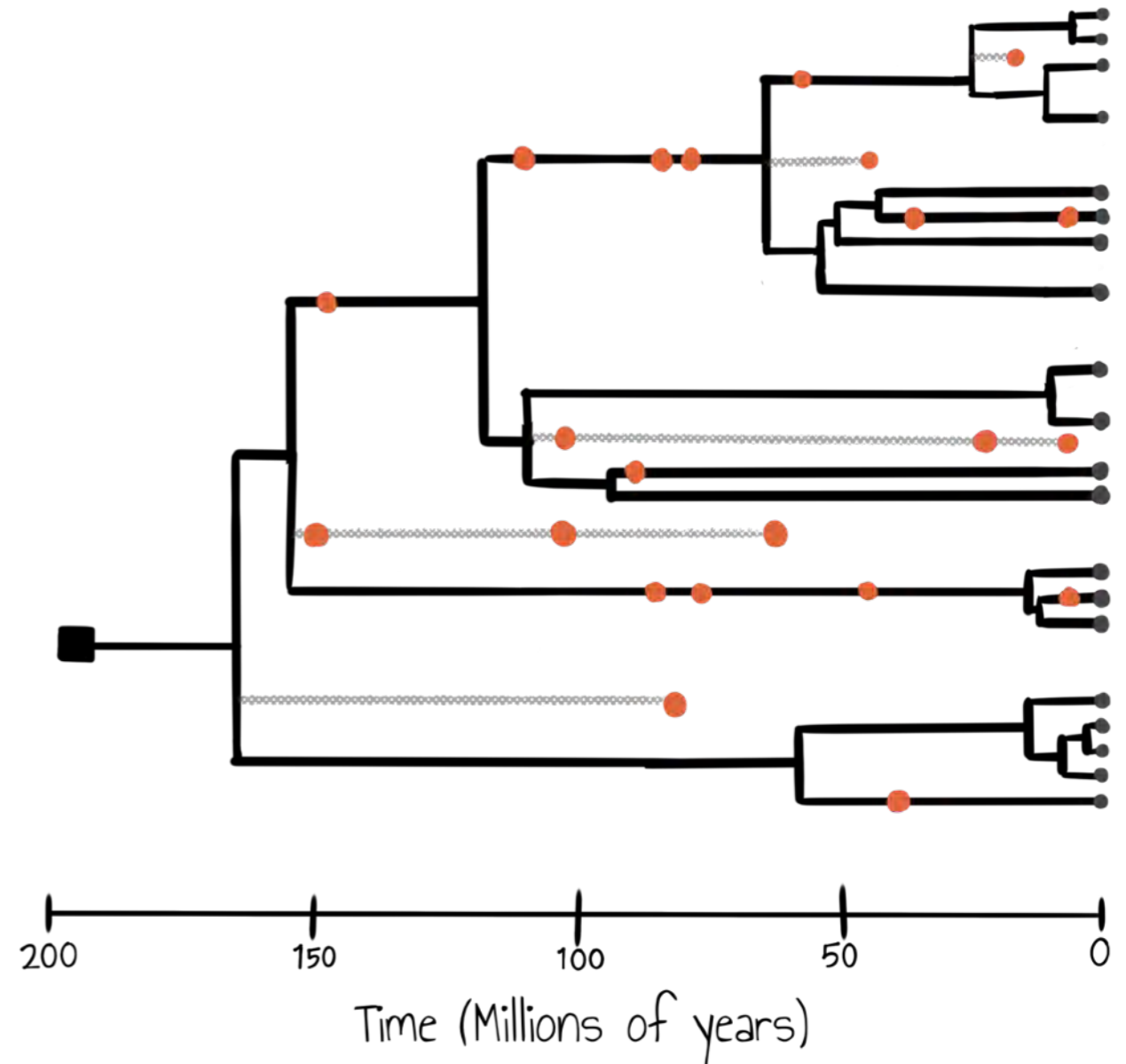
# The Fossilized Birth-Death Process (FBD)

We can compute the probability of the tree and fossil observations under a birth-death model with rate parameters:

$\lambda$  = speciation

$\mu$  = extinction

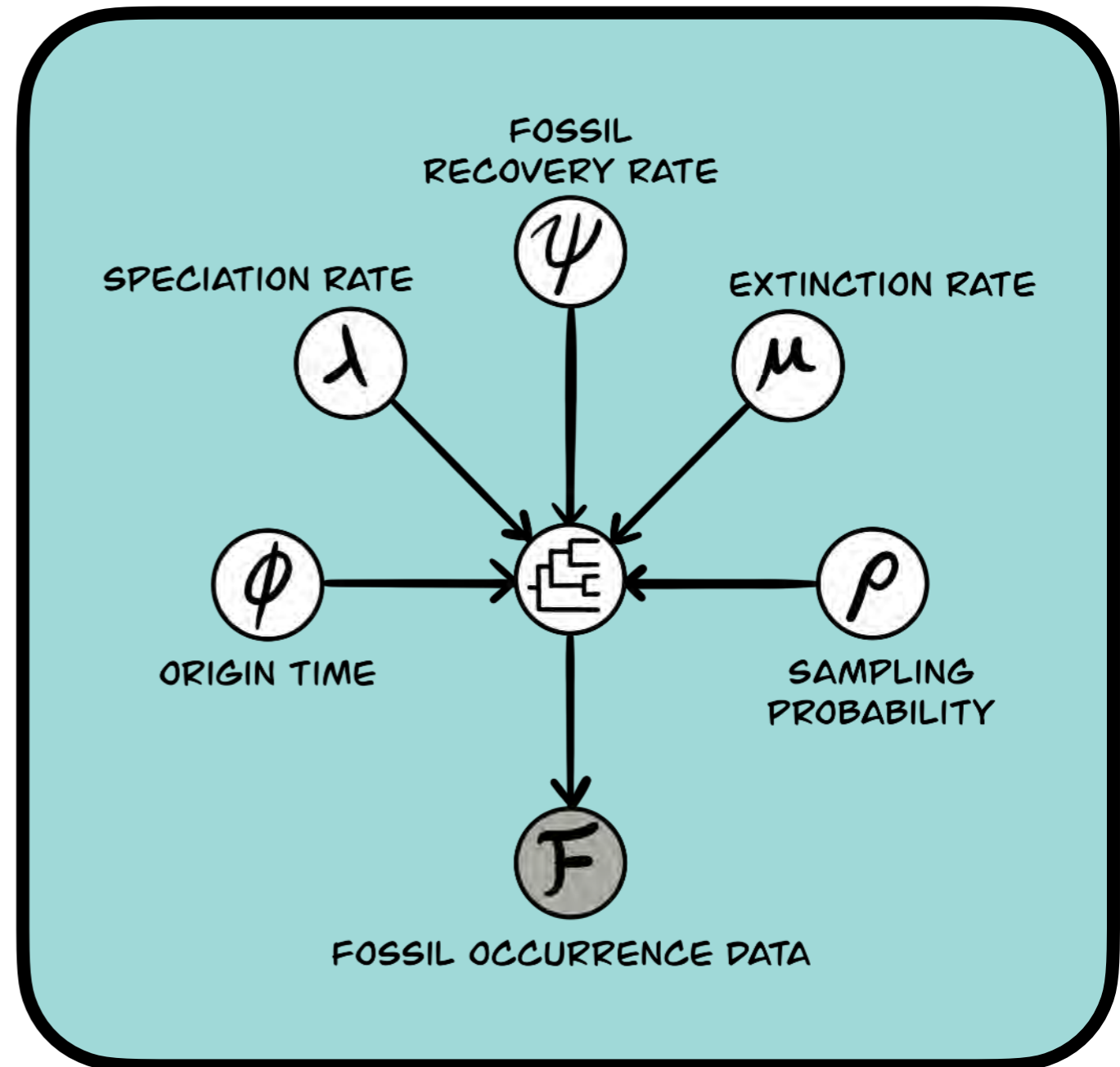
$\psi$  = fossilization/recovery



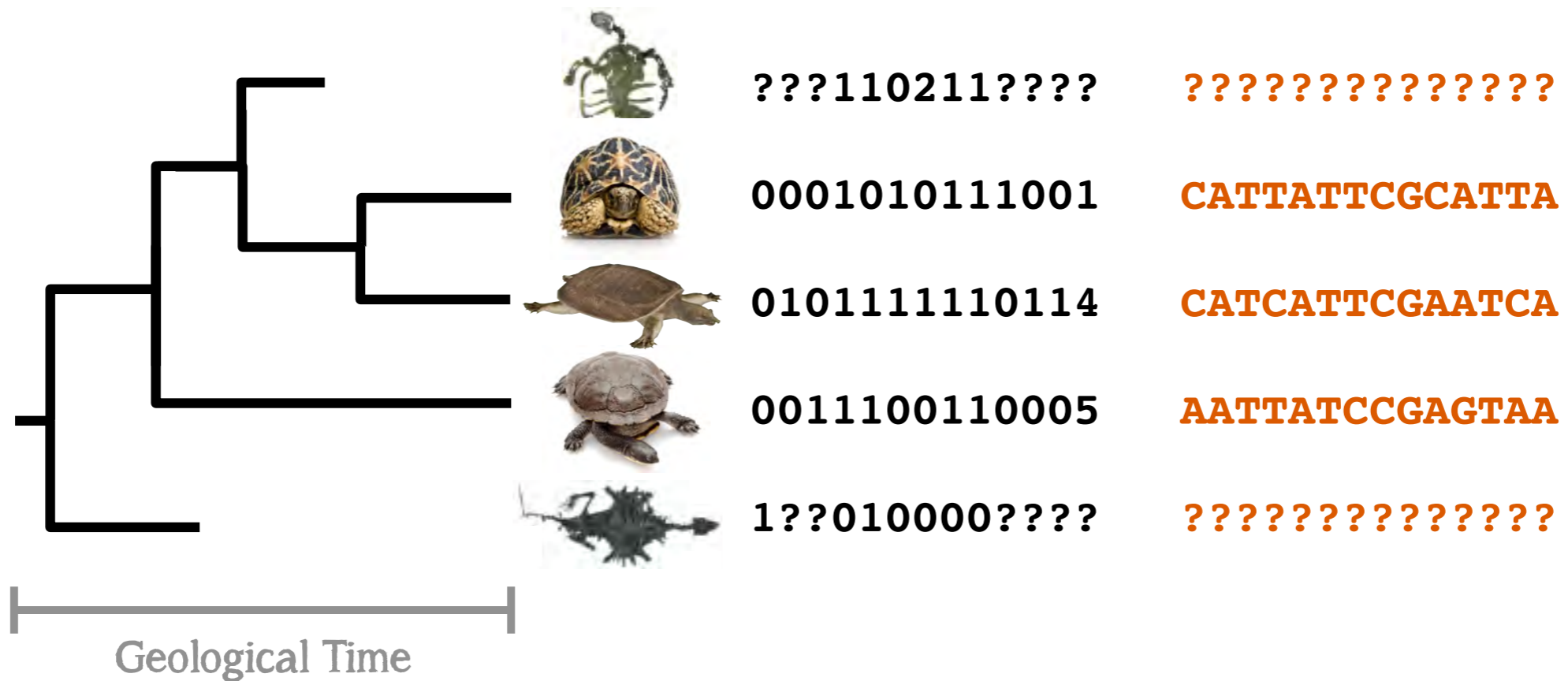
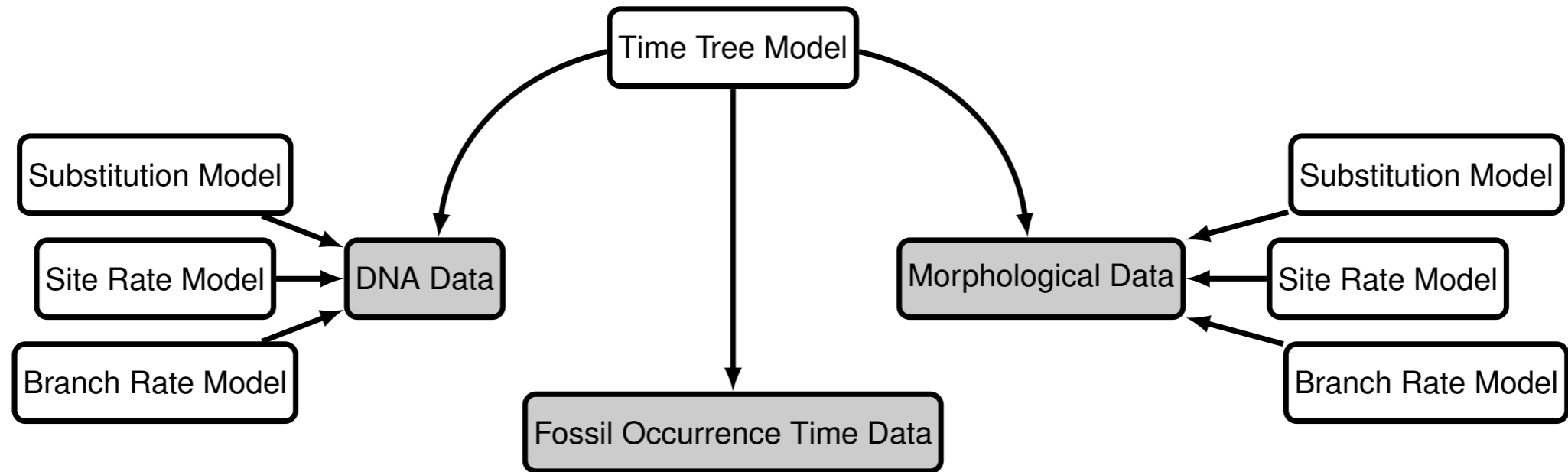
# The Fossilized Birth-Death Process (FBD)

The tree topology, divergence times, and fossil occurrences are conditioned on

- the origin time
- speciation rate
- extinction rate
- fossil recovery rate
- probability of sampling an extant species



# Combining Fossil & Extant Data



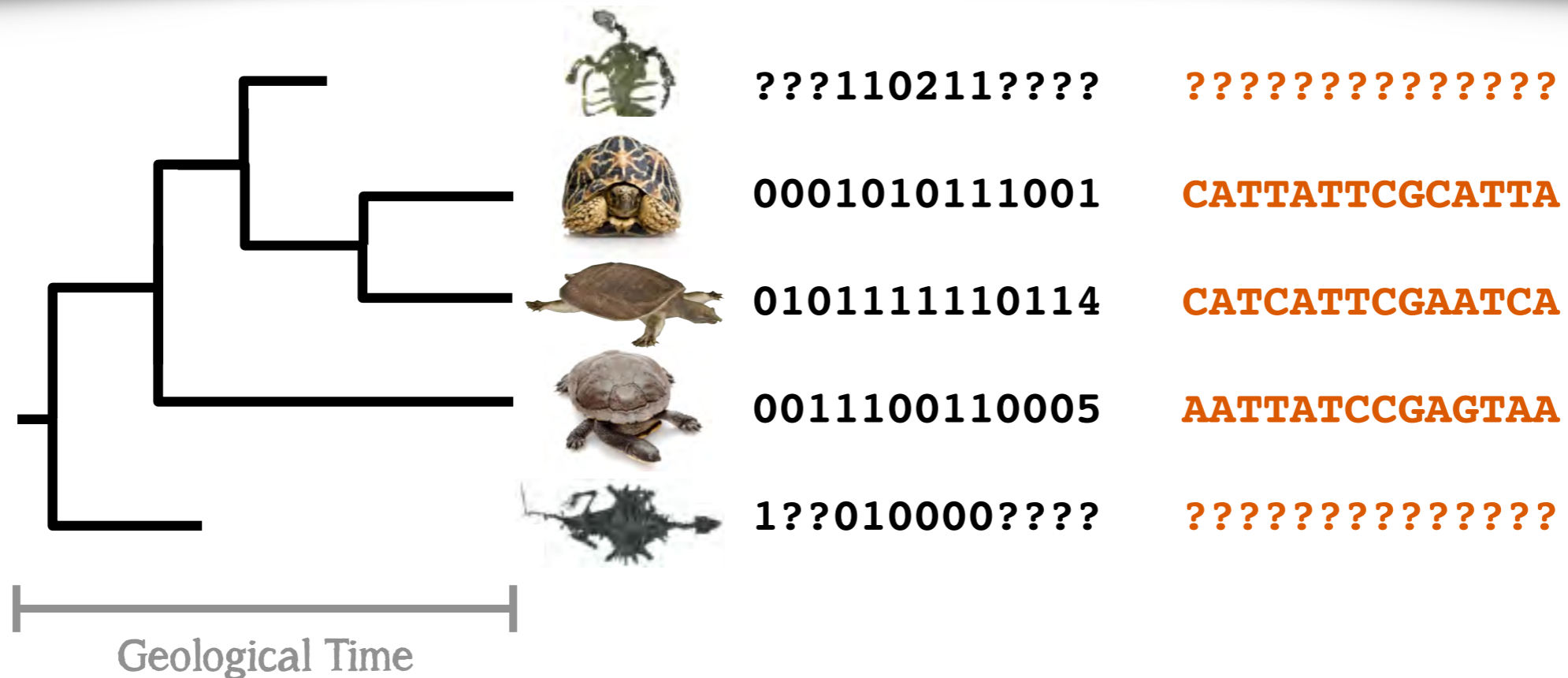
# Modeling Morphological Character Change

*Syst. Biol.* 50(6):913–925, 2001

## A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data

PAUL O. LEWIS

*Department of Ecology and Evolutionary Biology, The University of Connecticut, Storrs, Connecticut 06269-3043, USA;  
E-mail: paul.lewis@uconn.edu*



# Modeling Morphological Character Change

## The Lewis Mk model

Assumes a character can take  $k$  states

Transition rates between states are equal (symmetric)

$$Q = \alpha \begin{bmatrix} 1 - k & 1 & \dots & 1 \\ \vdots & 1 - k & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 - k \end{bmatrix}$$

**T1**    **0**

**T2**    **0**

**T3**    **1**

**T4**    **2**

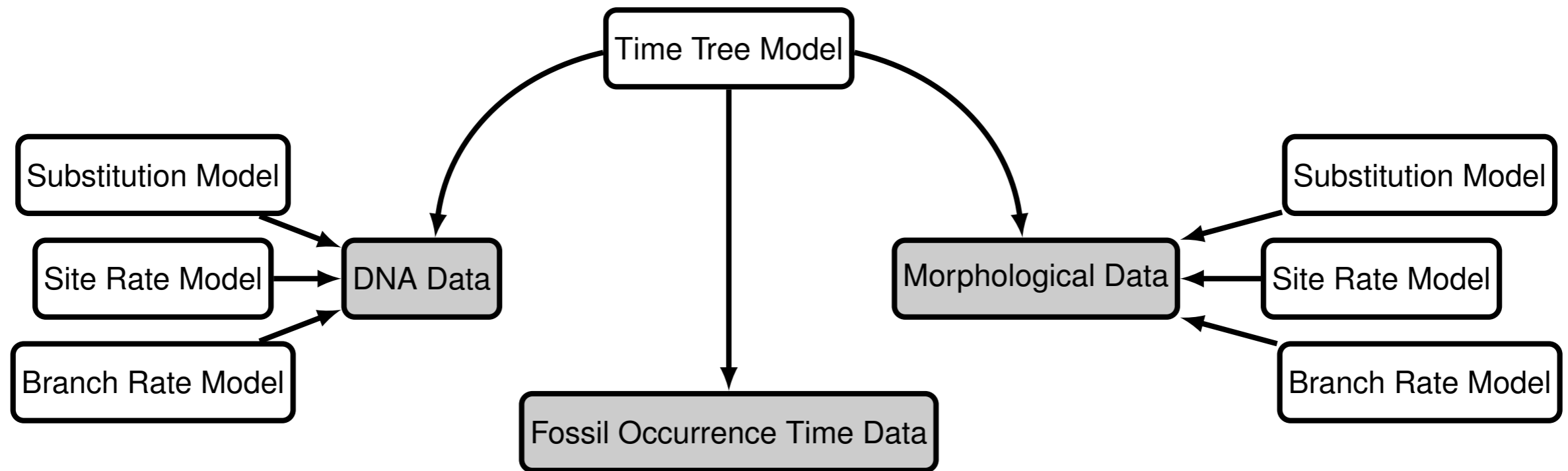
**T5**    **2**

**T6**    **1**

**T7**    **1**

# Combined-Evidence Analysis

Integrating models of molecular and morphological evolution with improved tree priors enables joint inference of the tree topology (extant & extinct) and divergence times



# Penguin Diversity in Deep Time

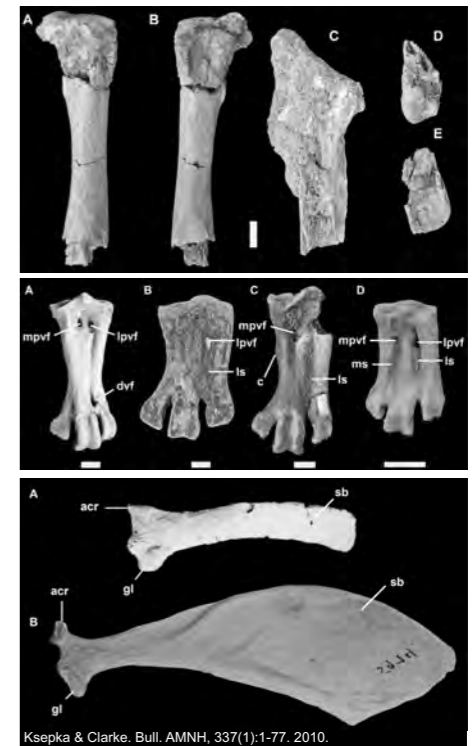
Does our understanding of penguin evolution improve when we consider both extant and fossil taxa?



"Penguin Party" by Kate Dzikiewicz



Photo of Daniel Ksepka and Karuku specimen by R.E. Fordyce (<http://blogs.scientificamerican.com/observations/giant-prehistoric-penguin-was-bigger-than-an-emperor/>)

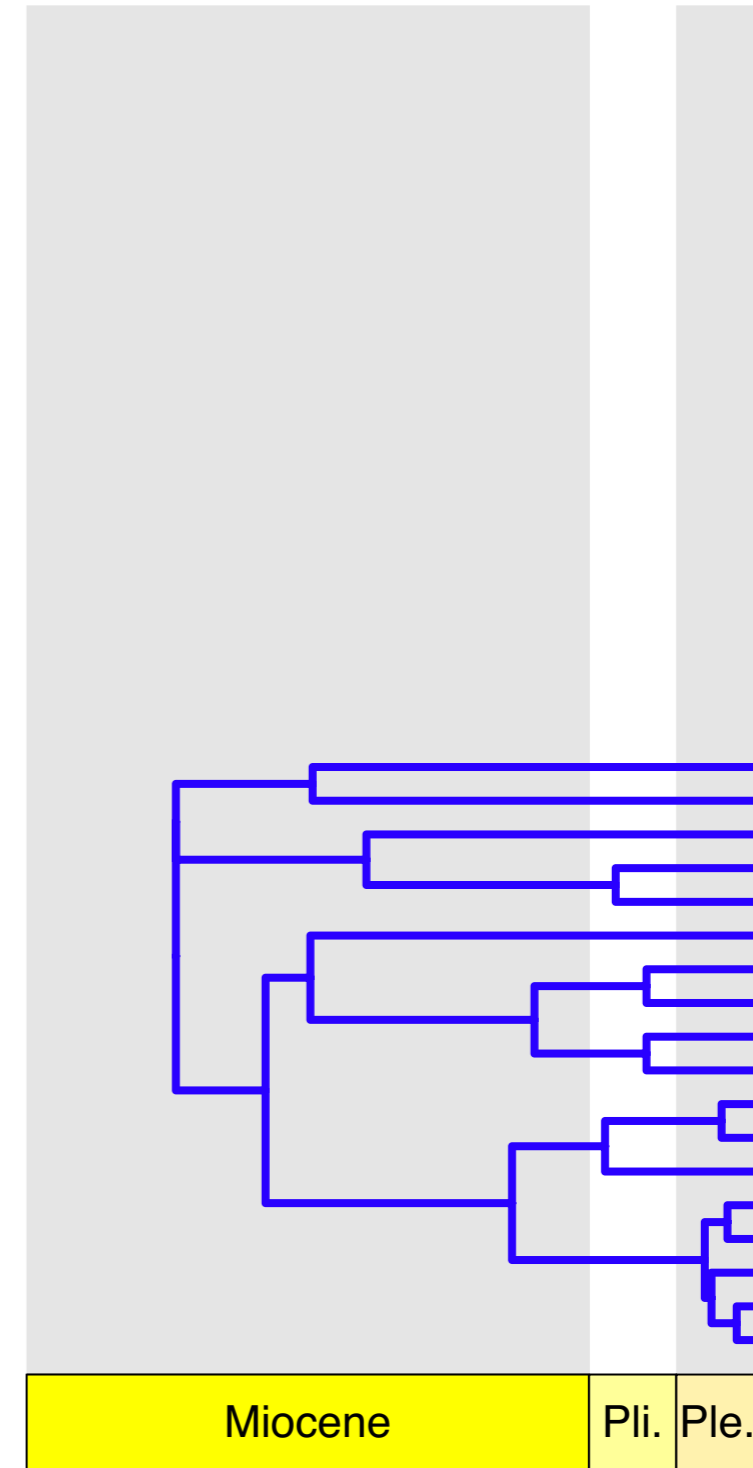


Ksepka & Clarke. Bull. AMNH, 337(1):1-77, 2010.



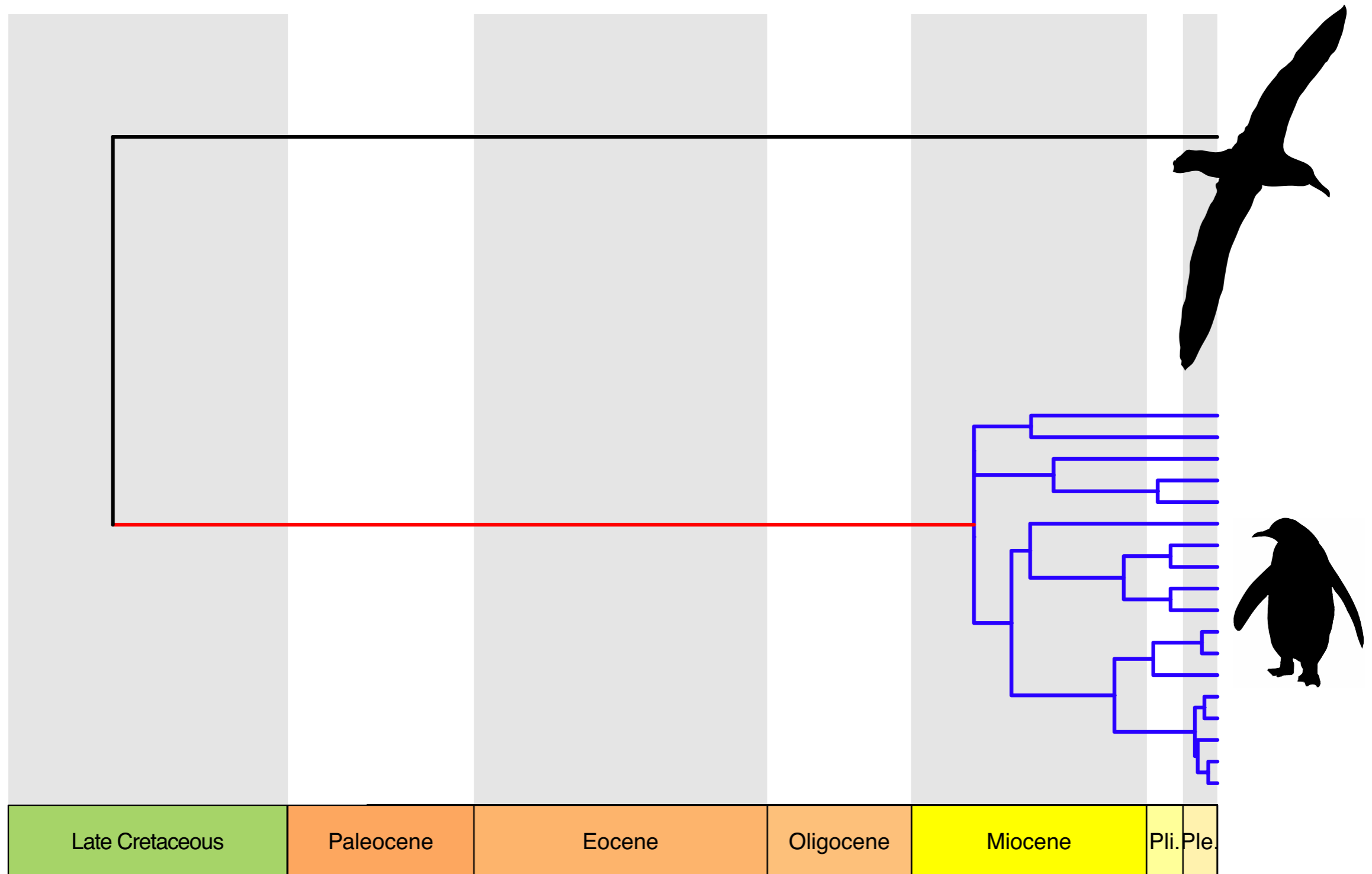
Artistic reconstructions by: Stephanie Abramowicz for Scientific American  
Fordyce, R.E. and D.T. Ksepka. The Strangest Bird Scientific American 307, 56 – 61 (2012)

# Penguin Diversity

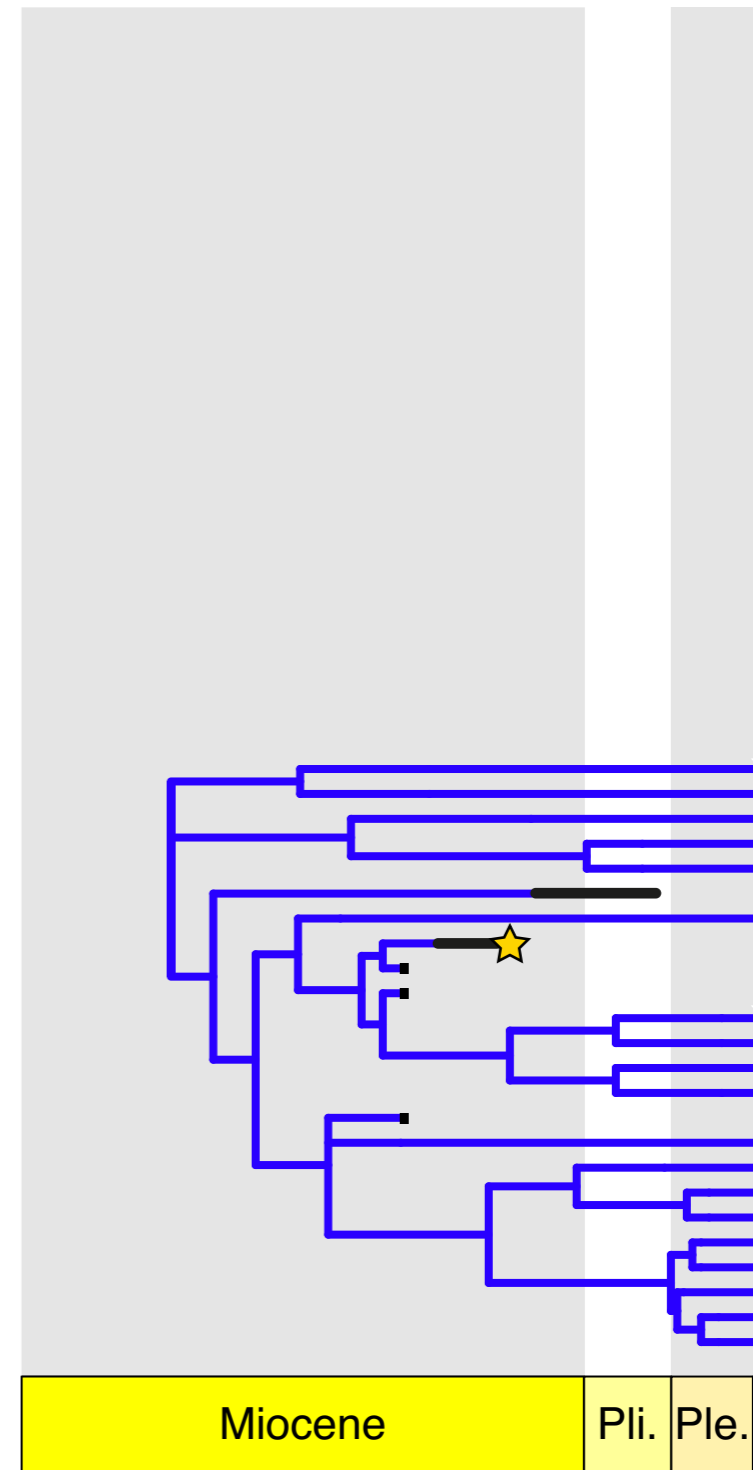




# Penguin Diversity



# Fossil Penguin Diversity



# Penguins in the Oligocene

## *Kairuku*

- ~1.5 m tall
- slender, with narrow bill
- scapula and pygostyle (tail bones) are more similar to non-penguin birds
- ~27 Mya



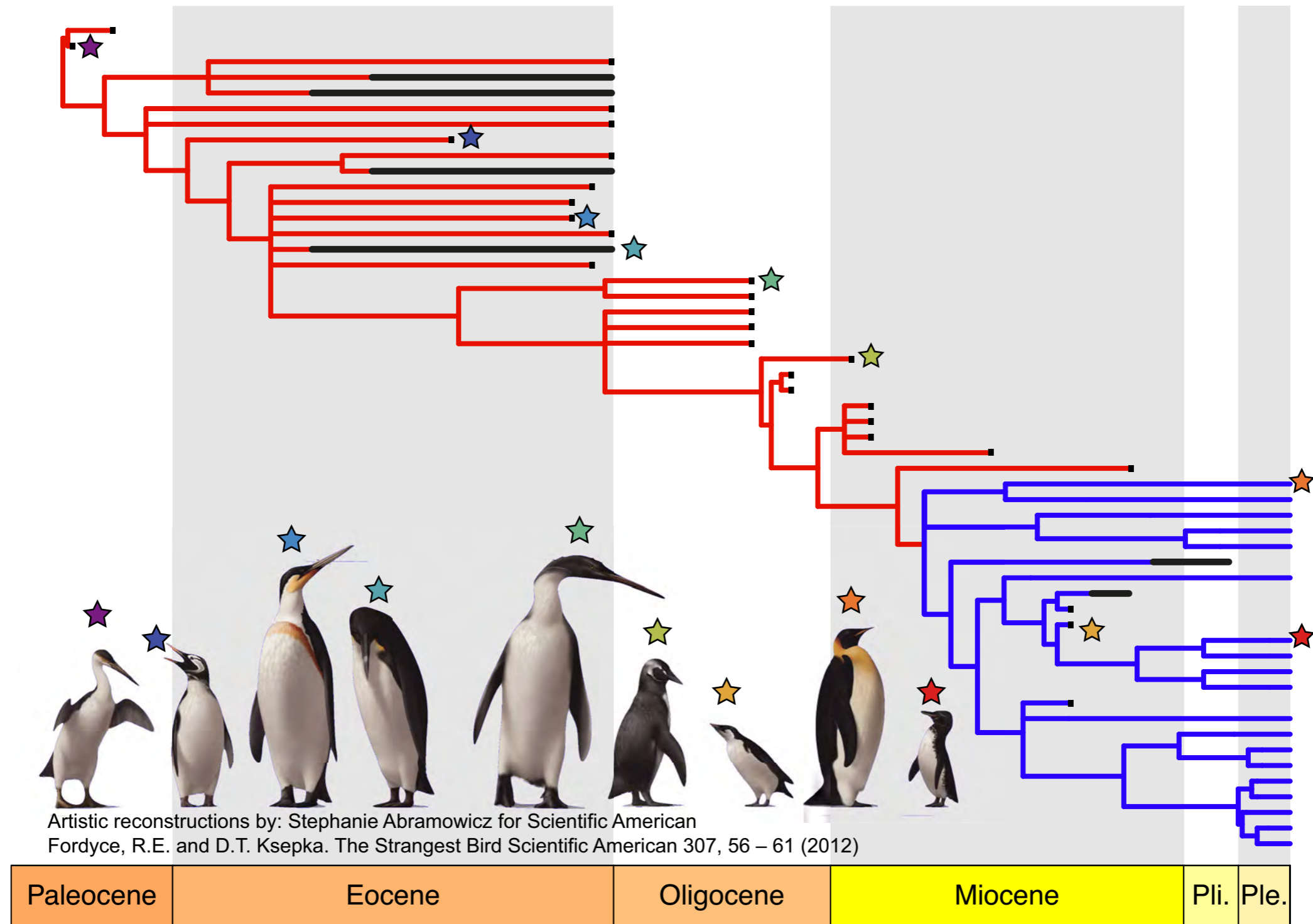
# Penguins in the Paleocene

## *Waimanu*

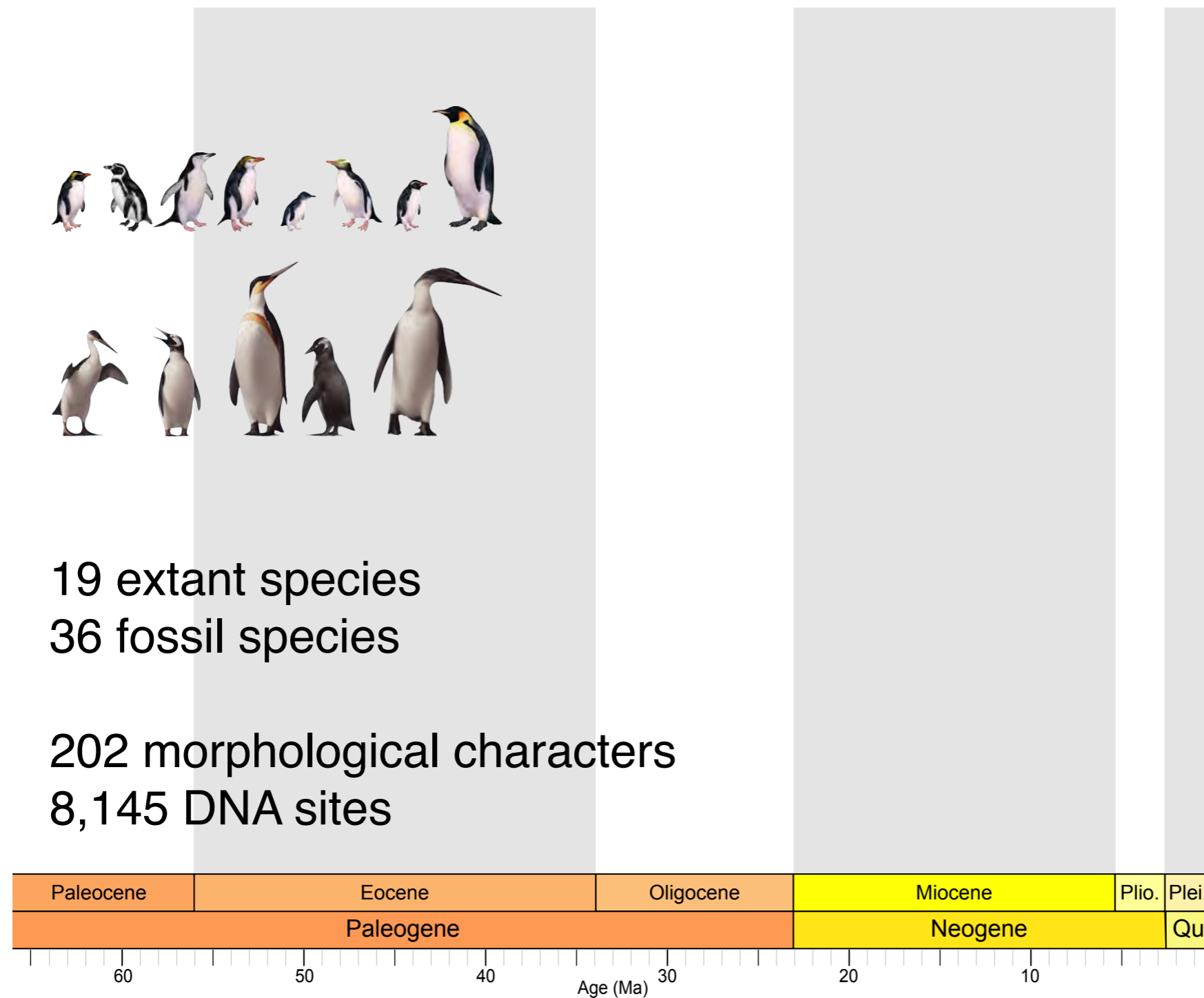
- oldest known penguin species
- intermediate wing morphology
- ~60 Mya



# Penguin Diversity in Deep Time

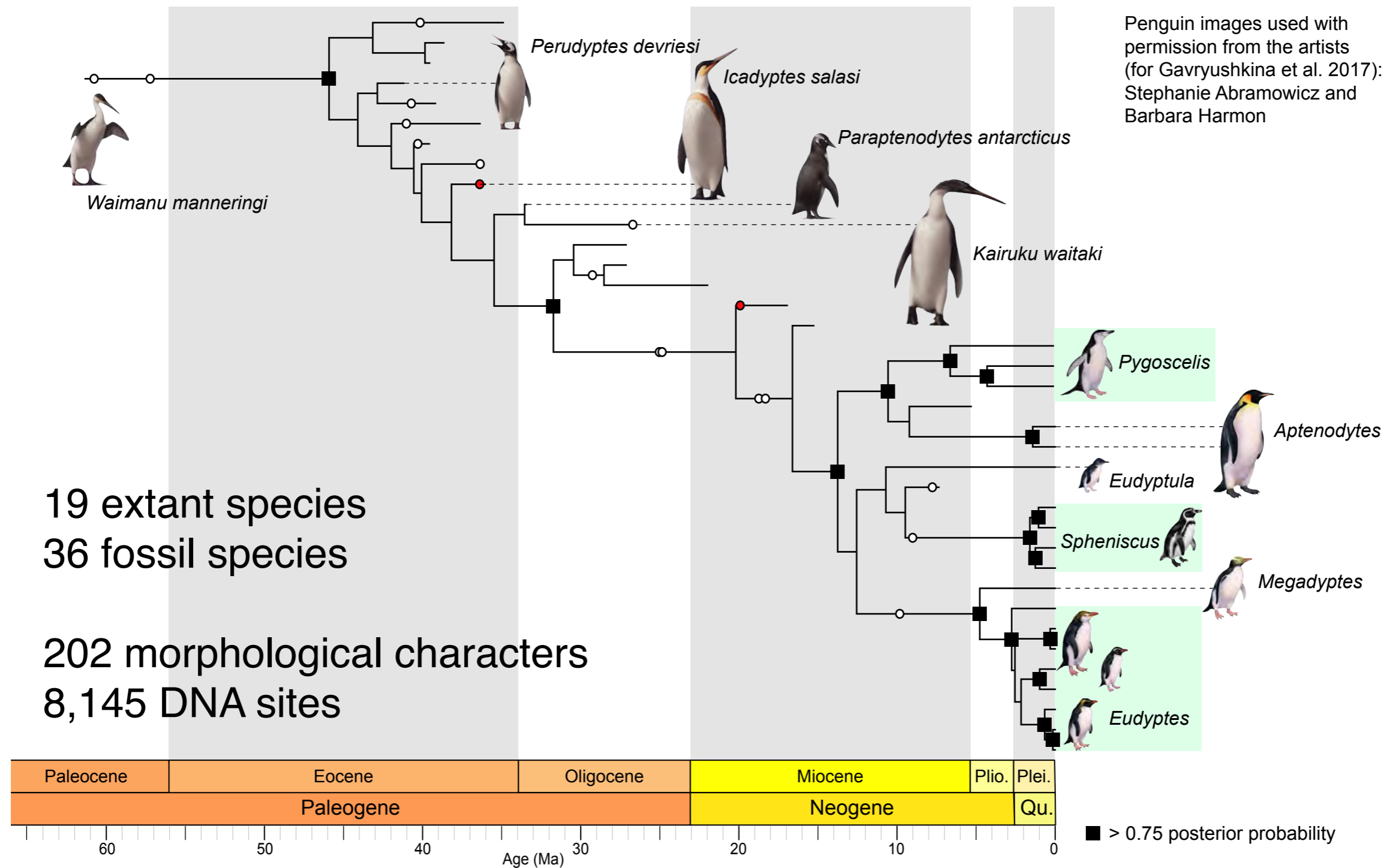


# Penguin Diversity in Deep Time

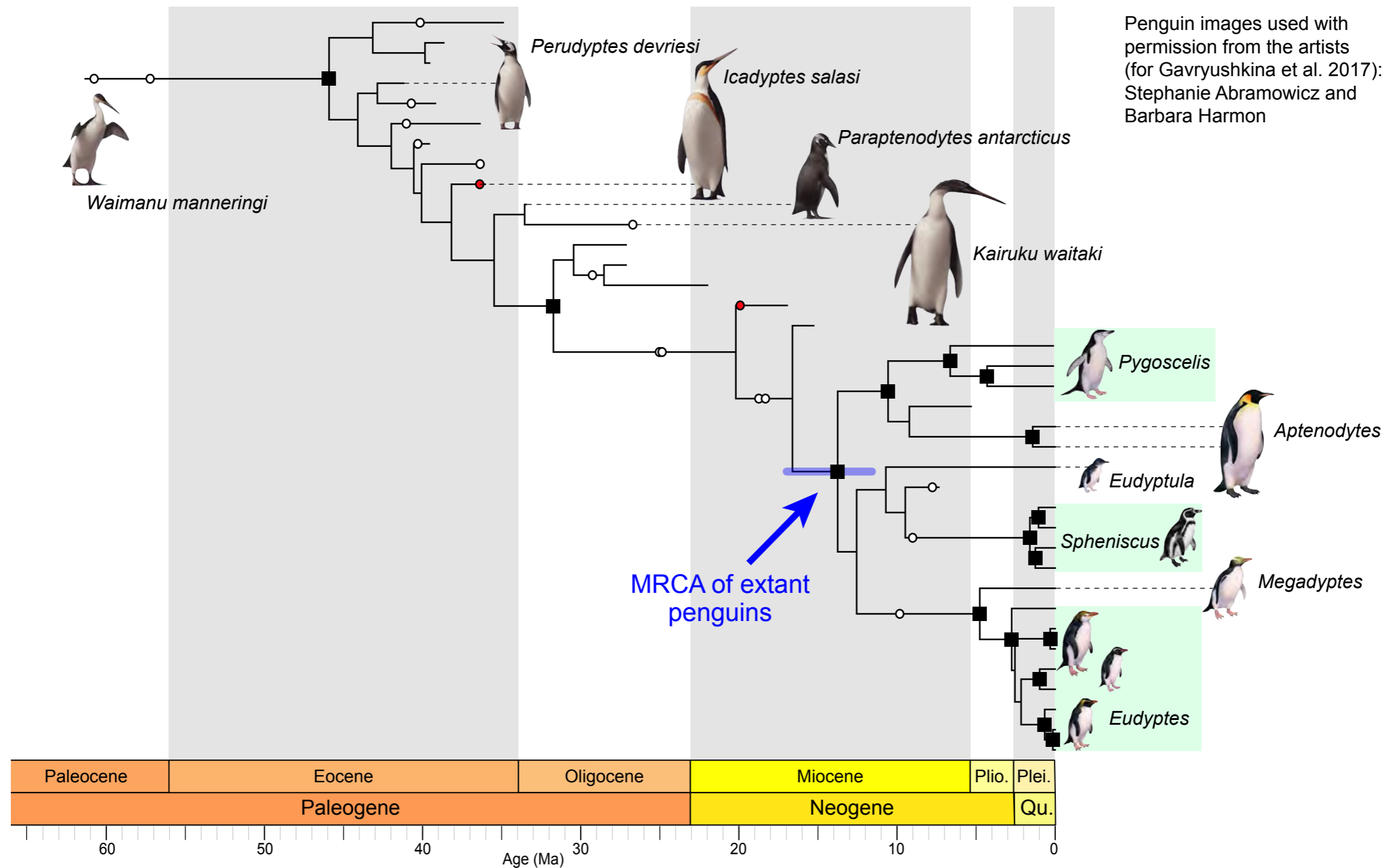


Penguin images used with permission from the artists (for Gavryushkina et al. 2017): Stephanie Abramowicz and Barbara Harmon

# Penguin Diversity in Deep Time

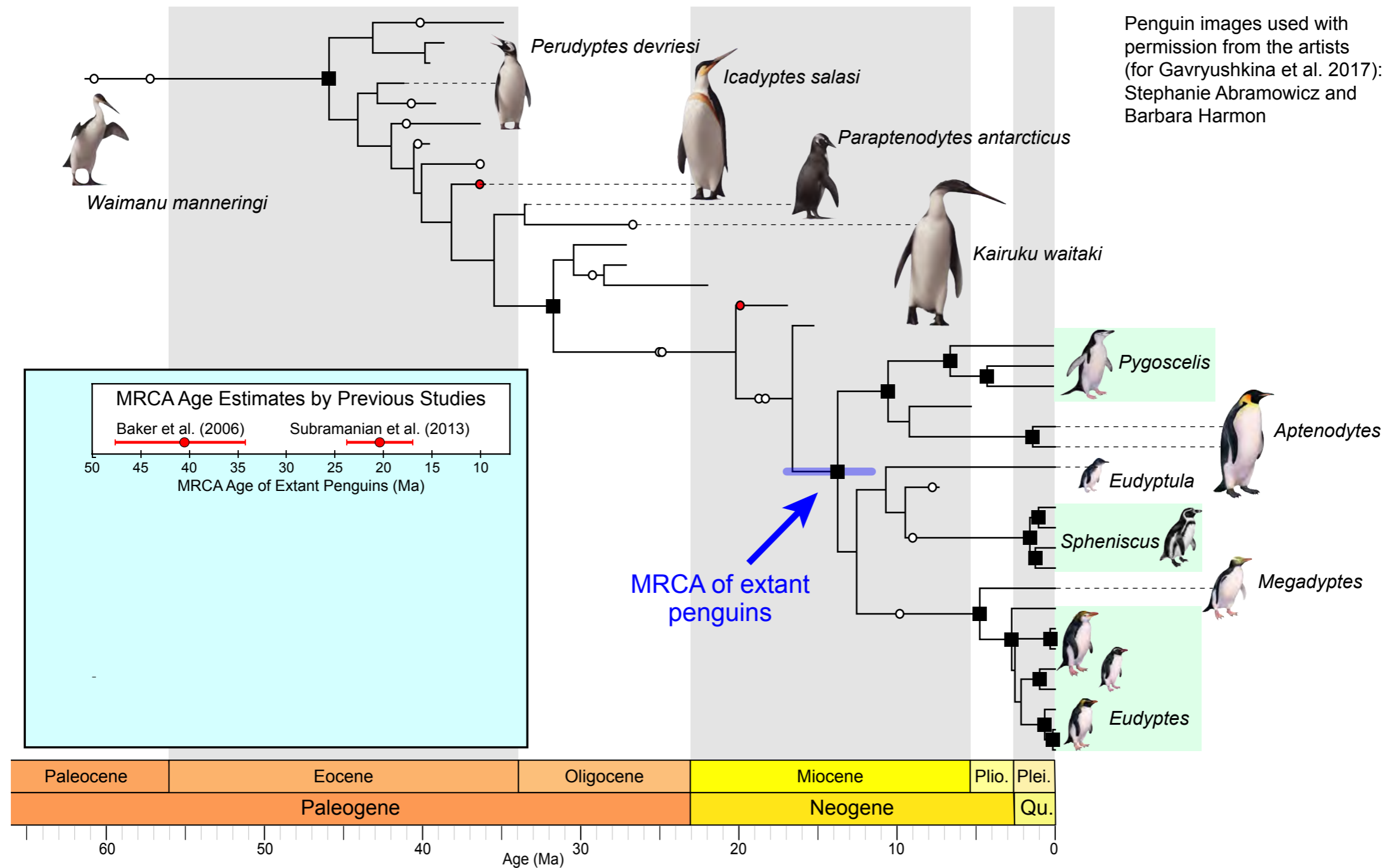


# Penguin Diversity in Deep Time



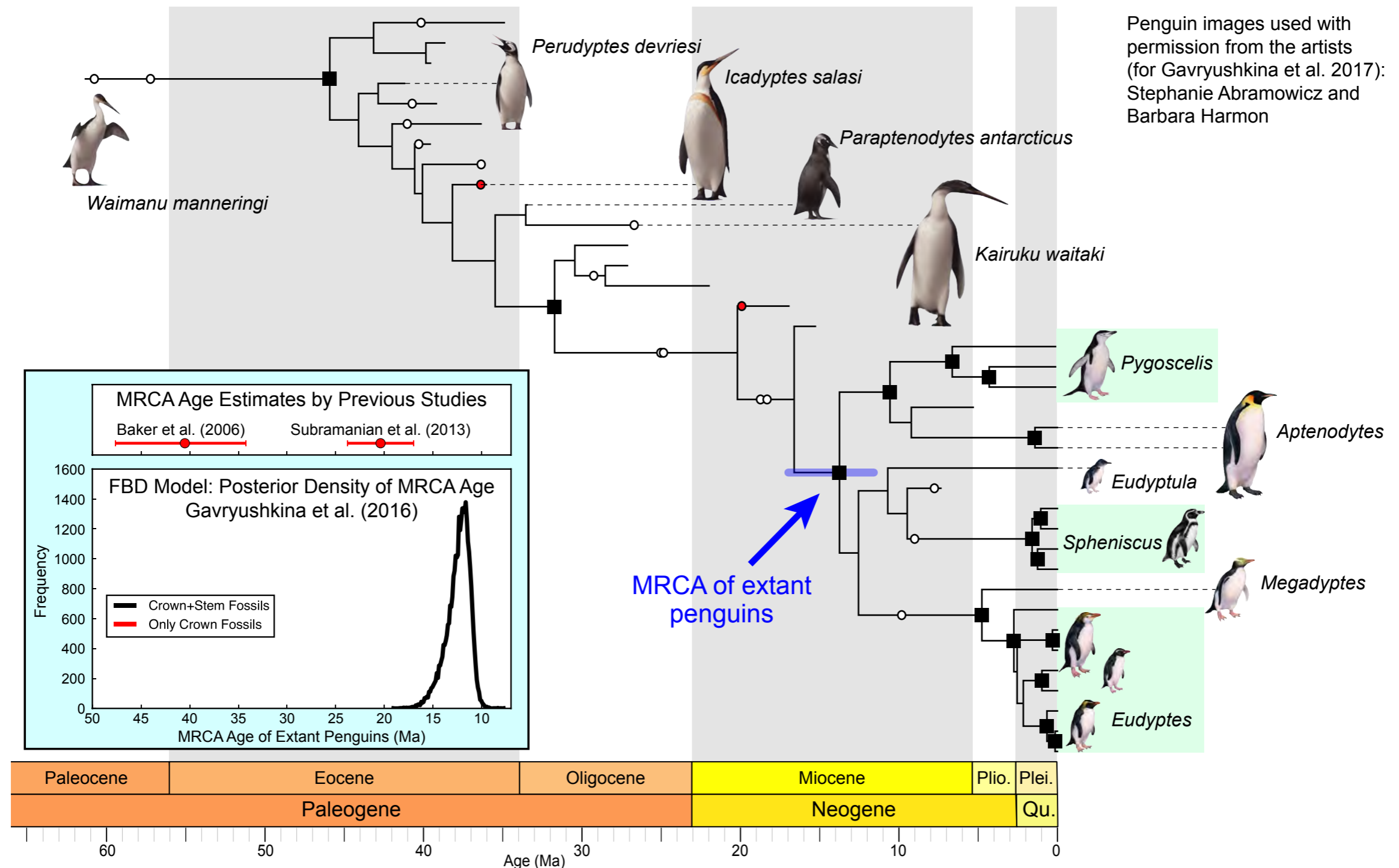


# Penguin Diversity in Deep Time



Penguin images used with permission from the artists (for Gavryushkina et al. 2017): Stephanie Abramowicz and Barbara Harmon

# Penguin Diversity in Deep Time



# Penguin Diversity in Deep Time

