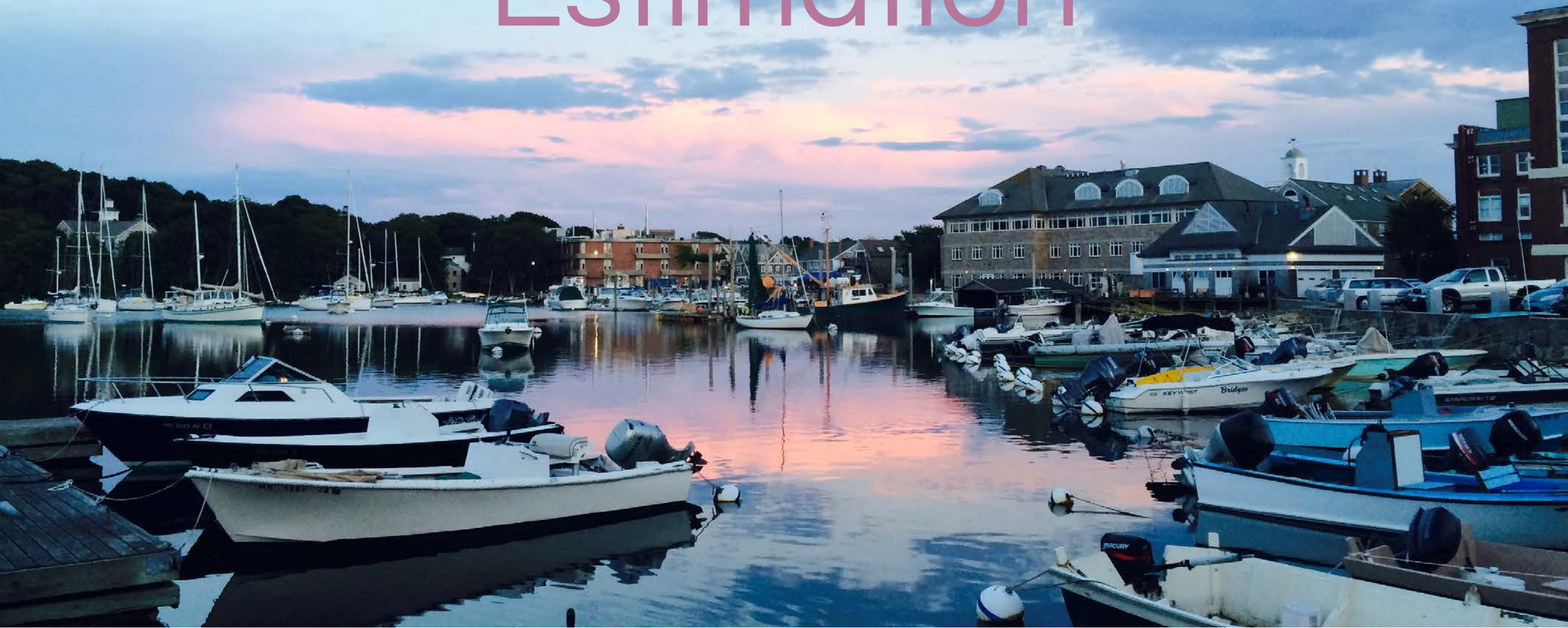


Bayesian Divergence-Time Estimation



@trayc7



phylo@iastate.edu



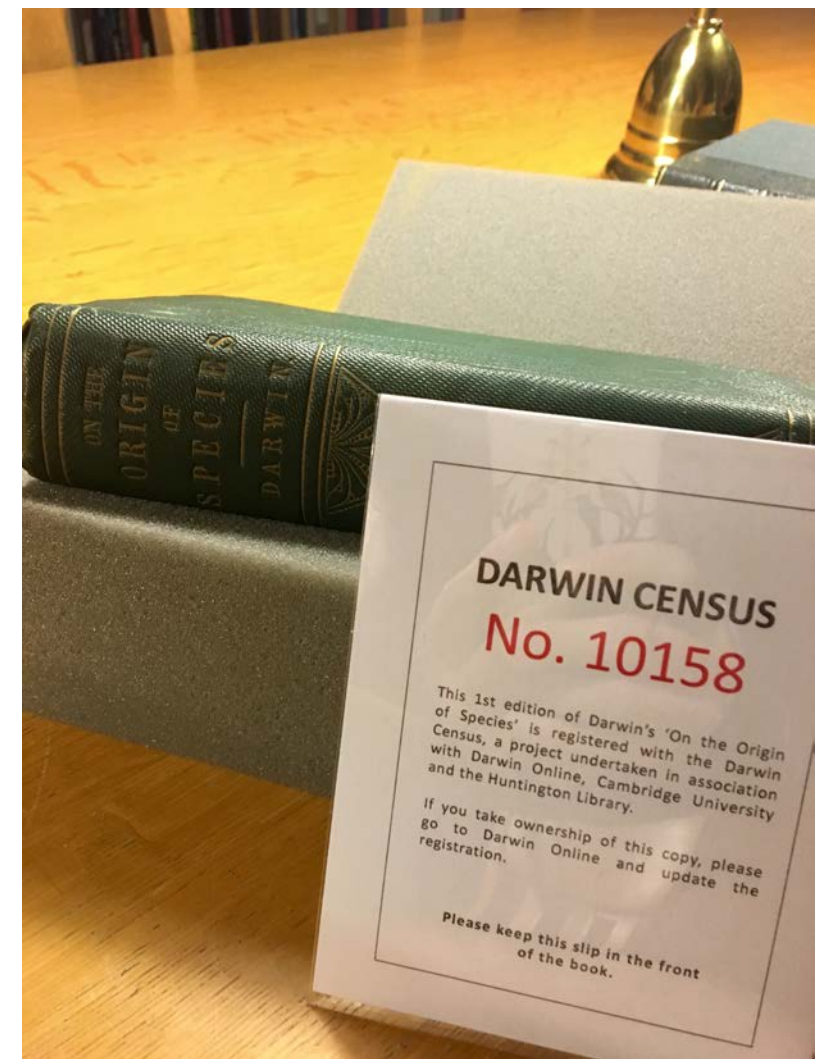
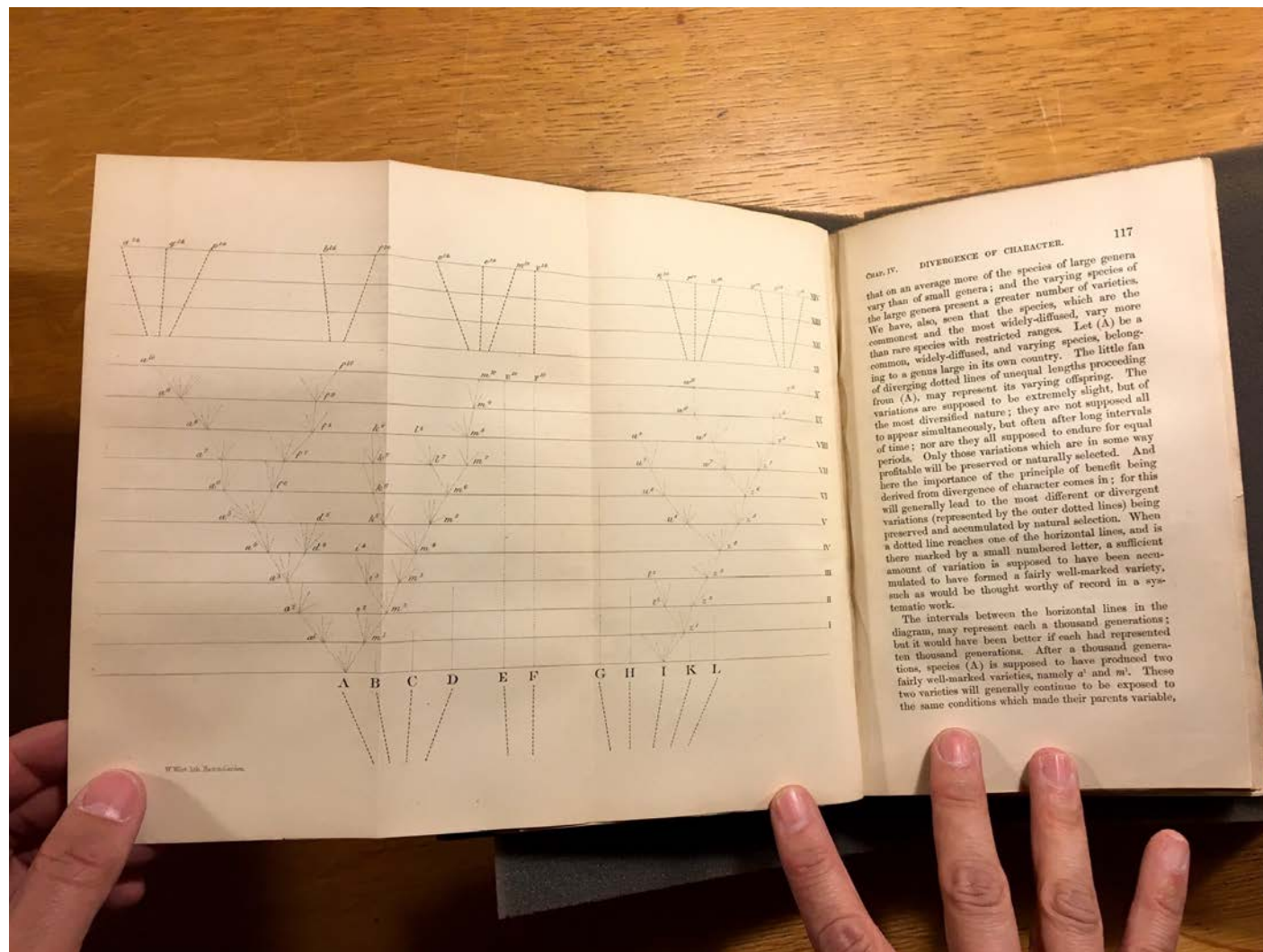
phyloworks.org

Tracy Heath

Ecology, Evolution, & Organismal Biology
Iowa State University

A Time-Scale for Evolution

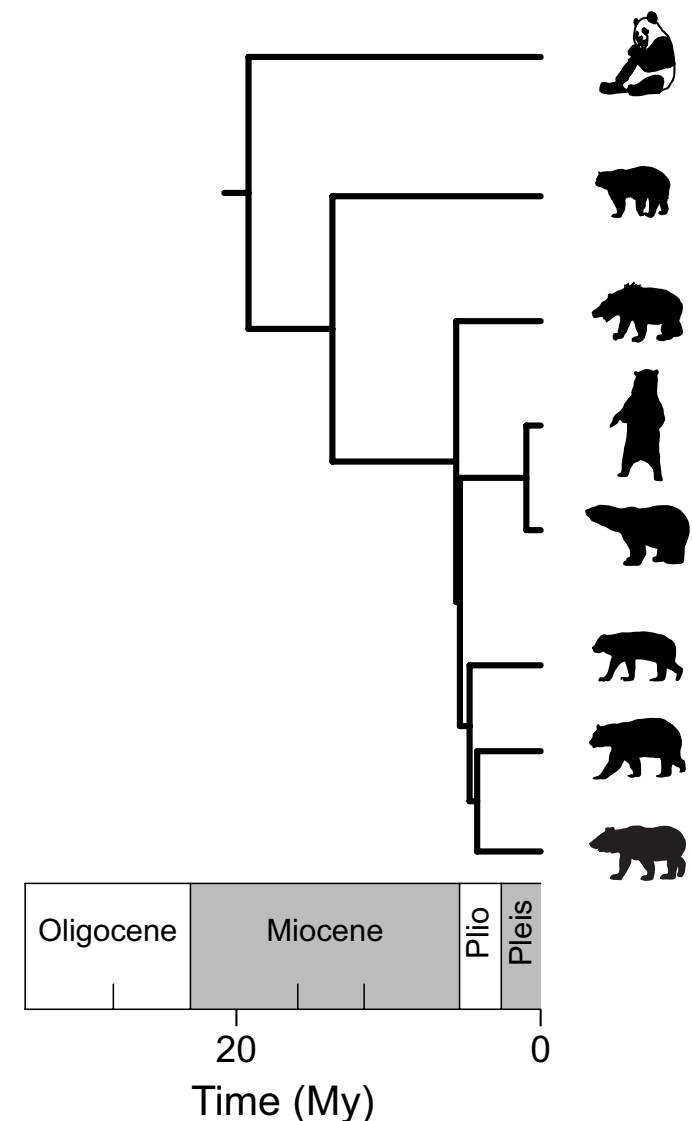
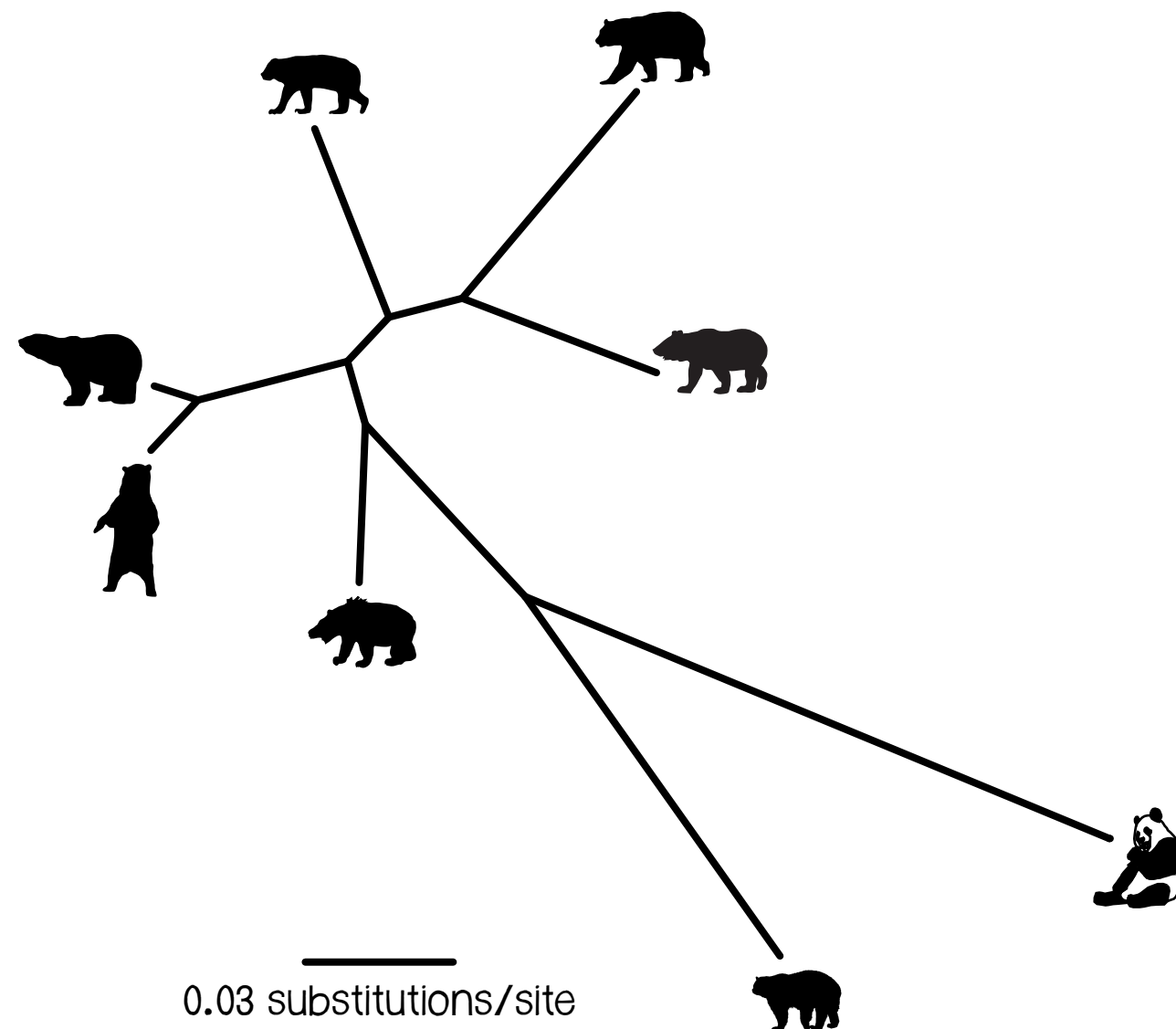
Phylogenies with branch lengths proportional to time provide valuable information about evolutionary history.



1st edition of Darwin's On the Origin of Species at Grinnell College (Grinnell, Iowa USA)

A Time-Scale for Evolution

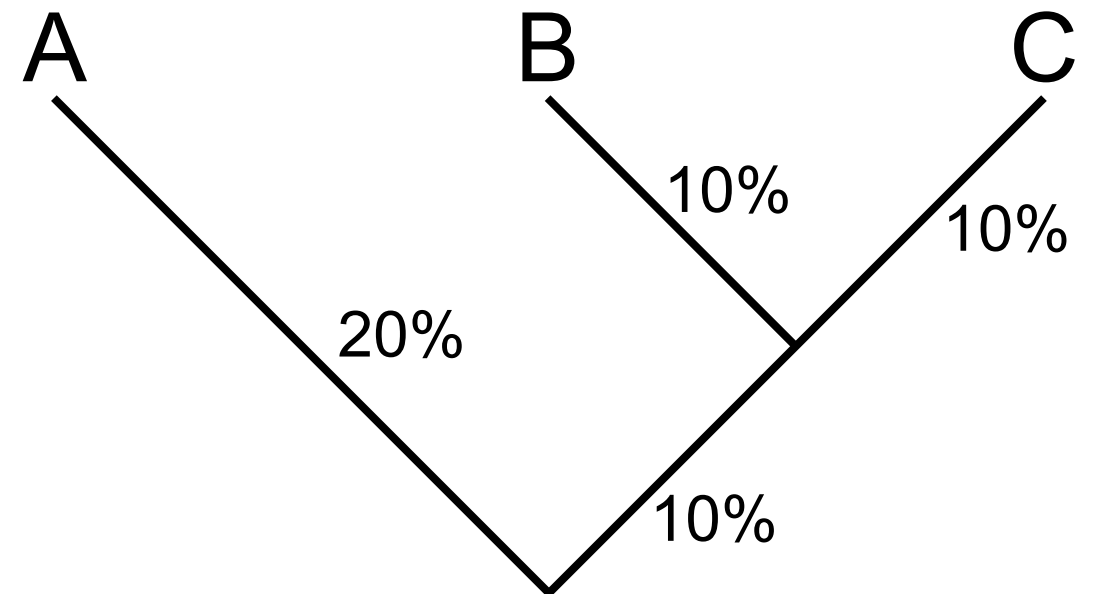
Phylogenies with branch lengths proportional to time provide more information about evolutionary history than unrooted trees with branch lengths in units of substitutions/site.



The Global Molecular Clock

Assume that the rate of evolutionary change is constant over time

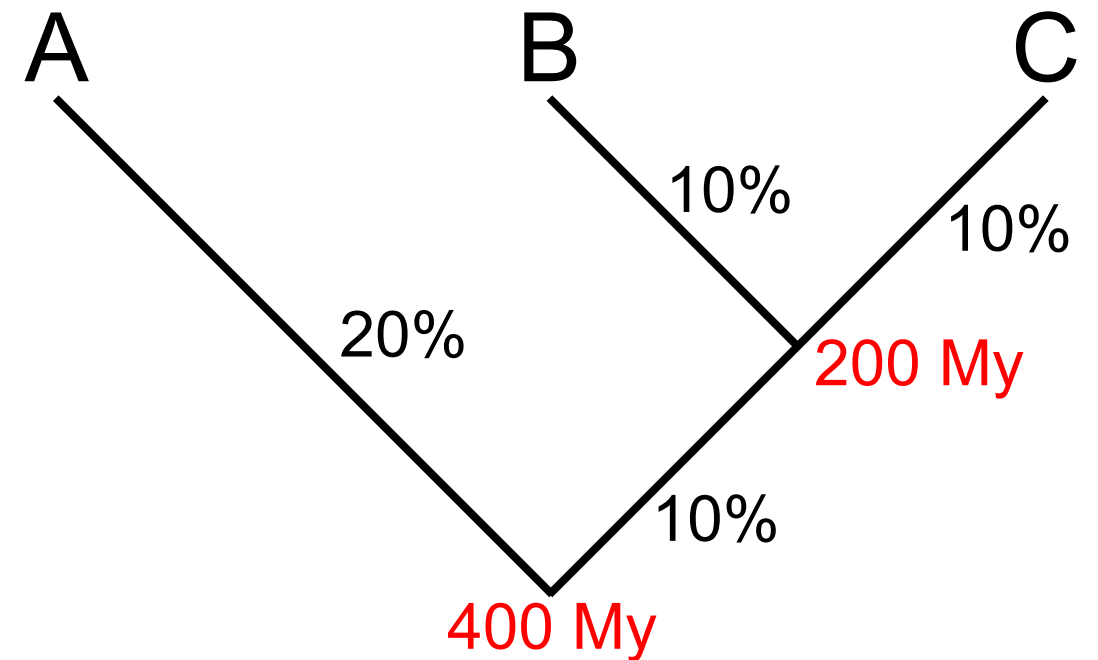
(branch lengths equal percent sequence divergence)



(Based on slides by Jeff Thorne; <http://statgen.ncsu.edu/thorne/compmolevo.html>)

The Global Molecular Clock

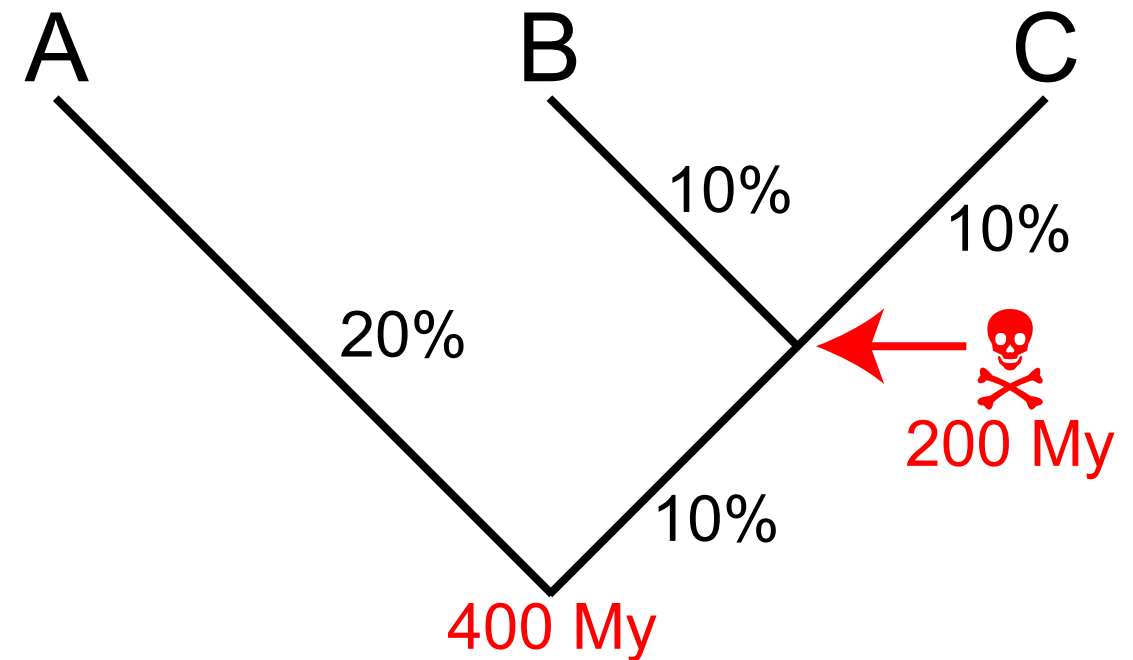
We can date the tree if we know the rate of change is 1% divergence per 10 My



(Based on slides by Jeff Thorne; <http://statgen.ncsu.edu/thorne/compmolevo.html>)

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



(Based on slides by Jeff Thorne; <http://statgen.ncsu.edu/thorne/compmolevo.html>)

Rejecting the Global Molecular Clock

Rates of evolution vary across lineages and over time

Mutation rate

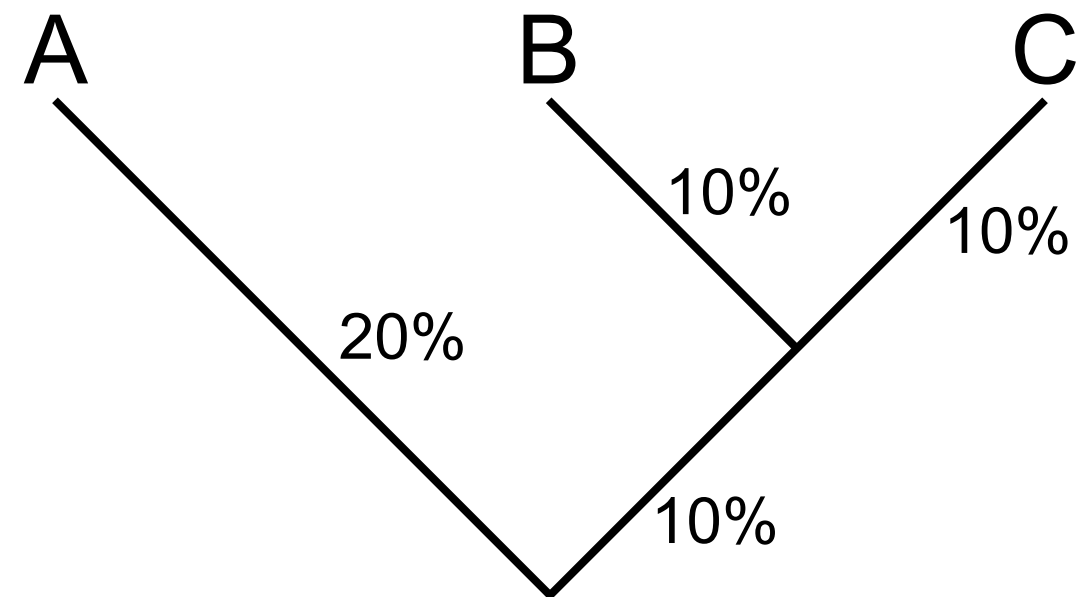
Variation in

- metabolic rate
- generation time
- DNA repair

Fixation rate

Variation in

- strengths/targets of selection
- population sizes

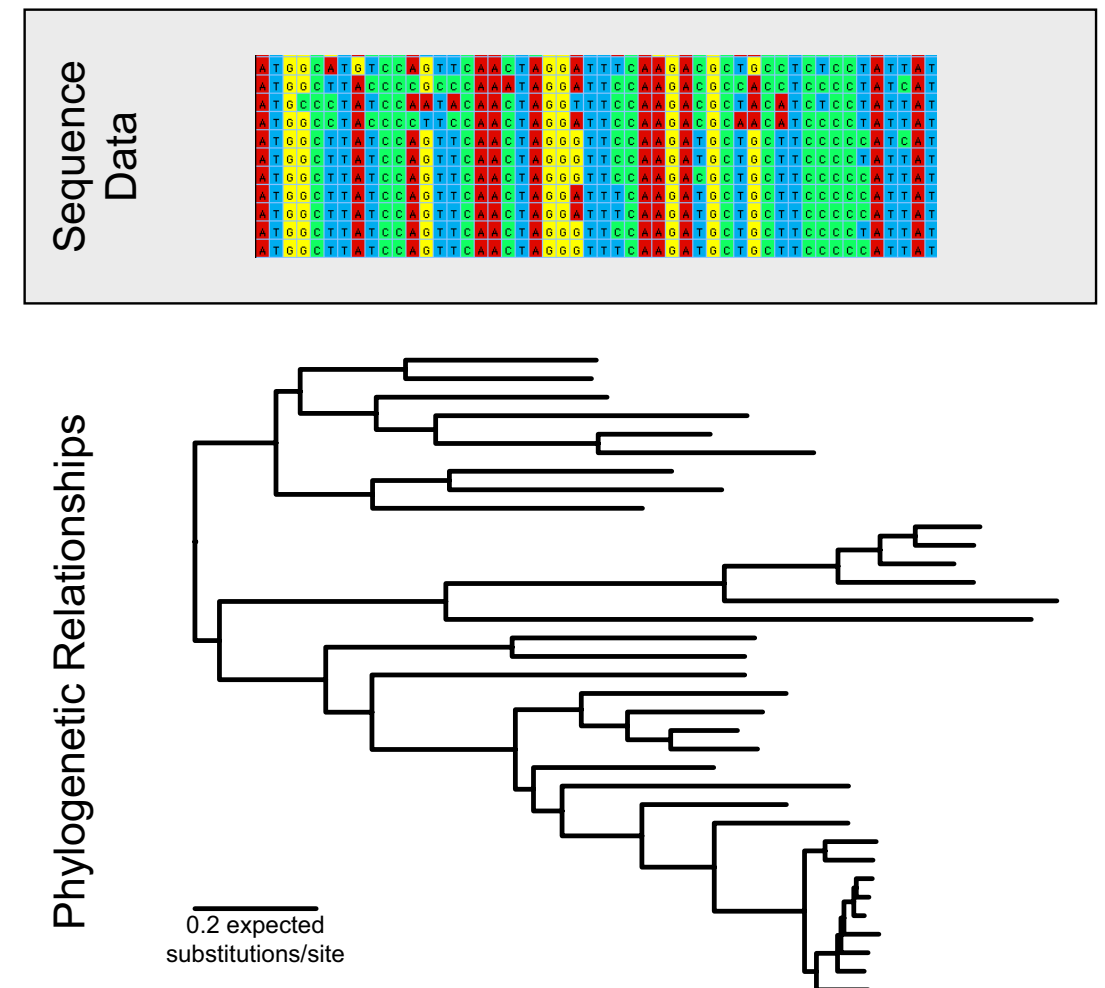


Unconstrained Analysis

Sequence data provide information about **branch lengths**

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

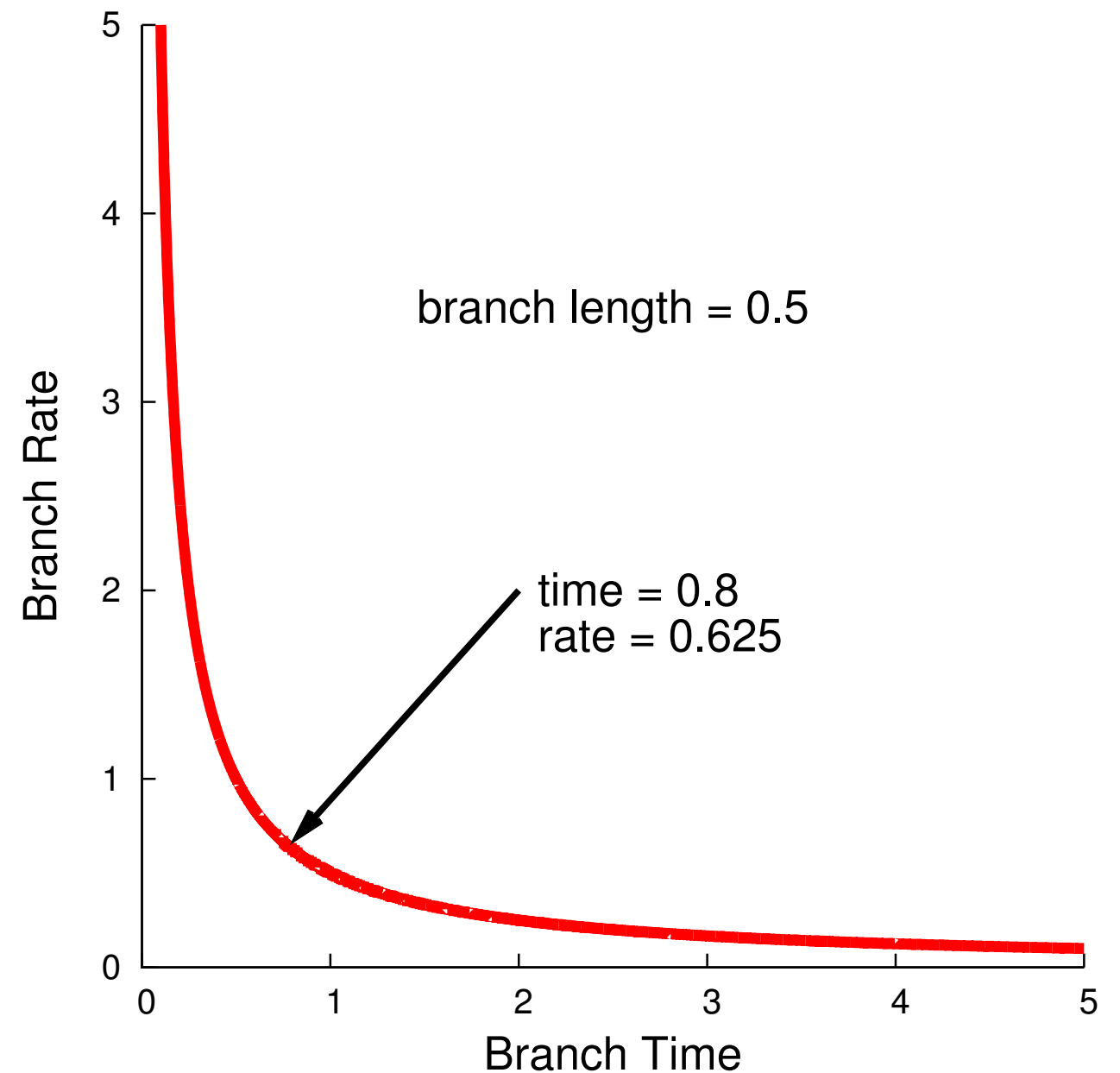
branch length = rate \times time



Estimating Rate & Time

The sequence data provide information about branch length

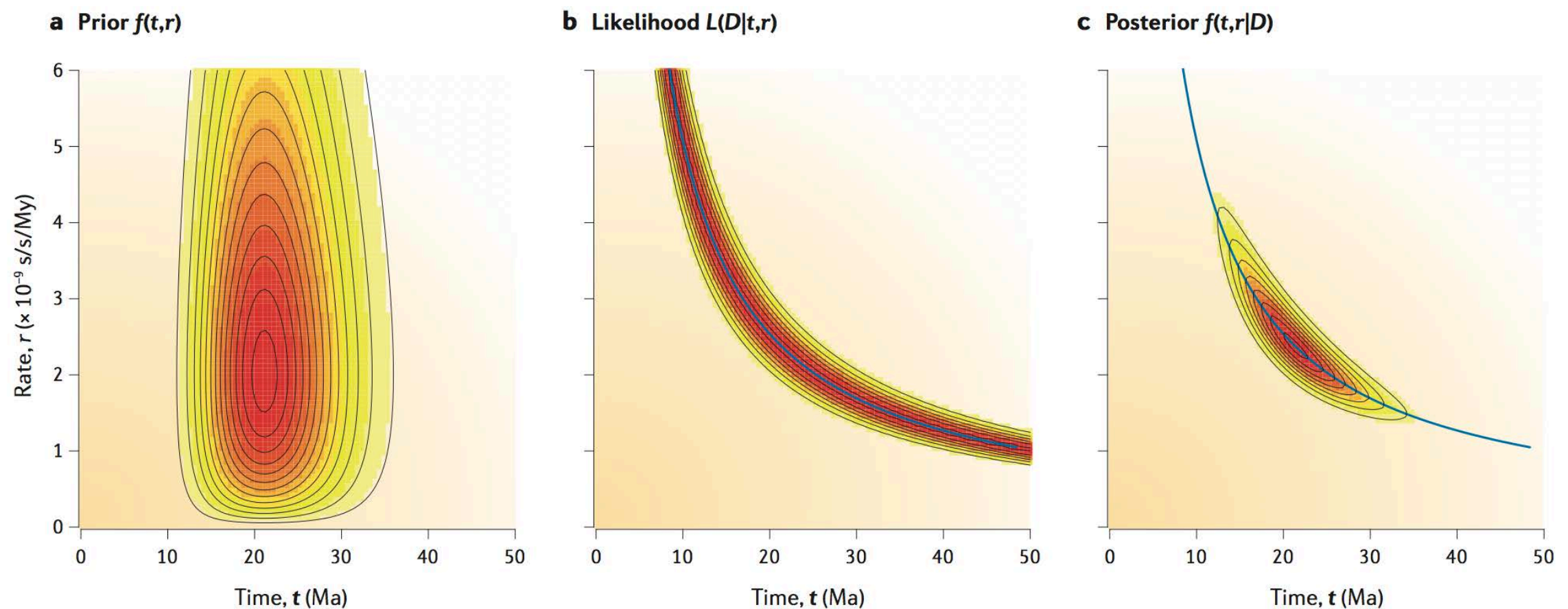
for any possible rate, there's a time that fits the branch length perfectly



(figure based on Thorne & Kishino, 2005)

Estimating Rate & Time

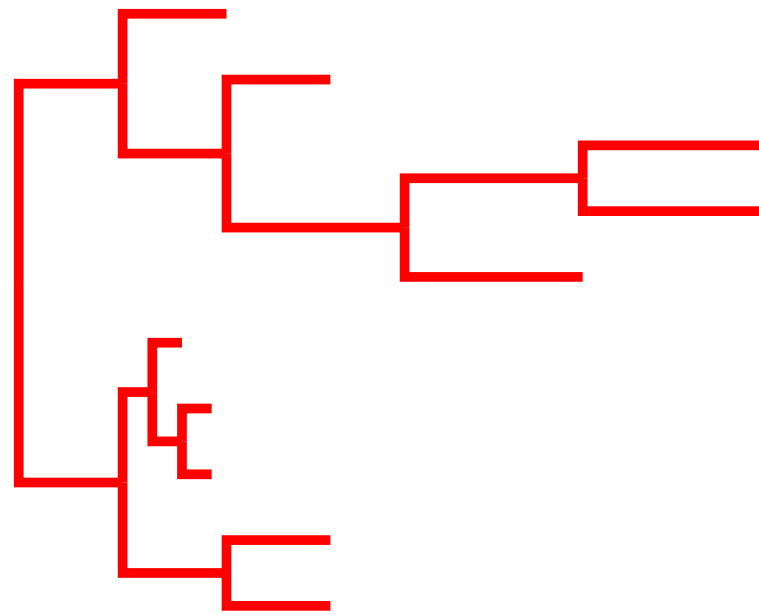
Methods for dating species divergences estimate the **substitution rate** and **time** separately



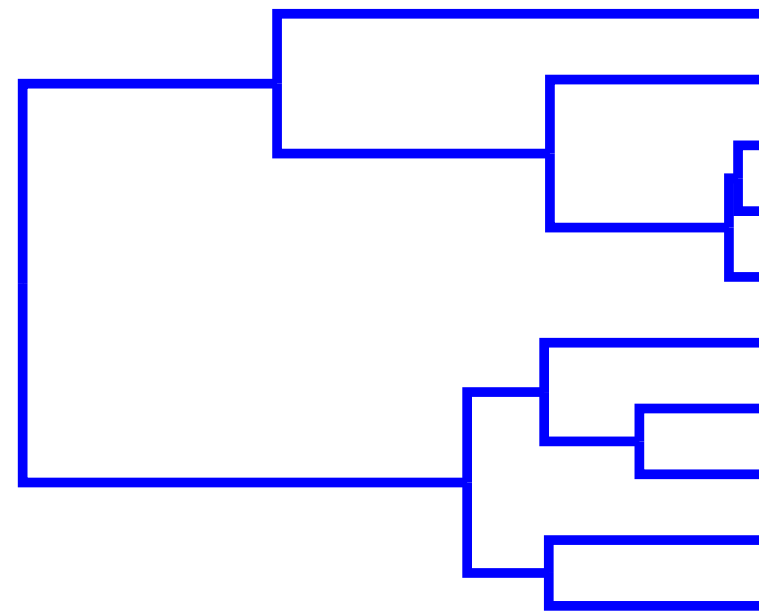
(dos Reis et al. Nature Reviews Genetics, 2016)

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

Bayesian Divergence Time Estimation



length = rate

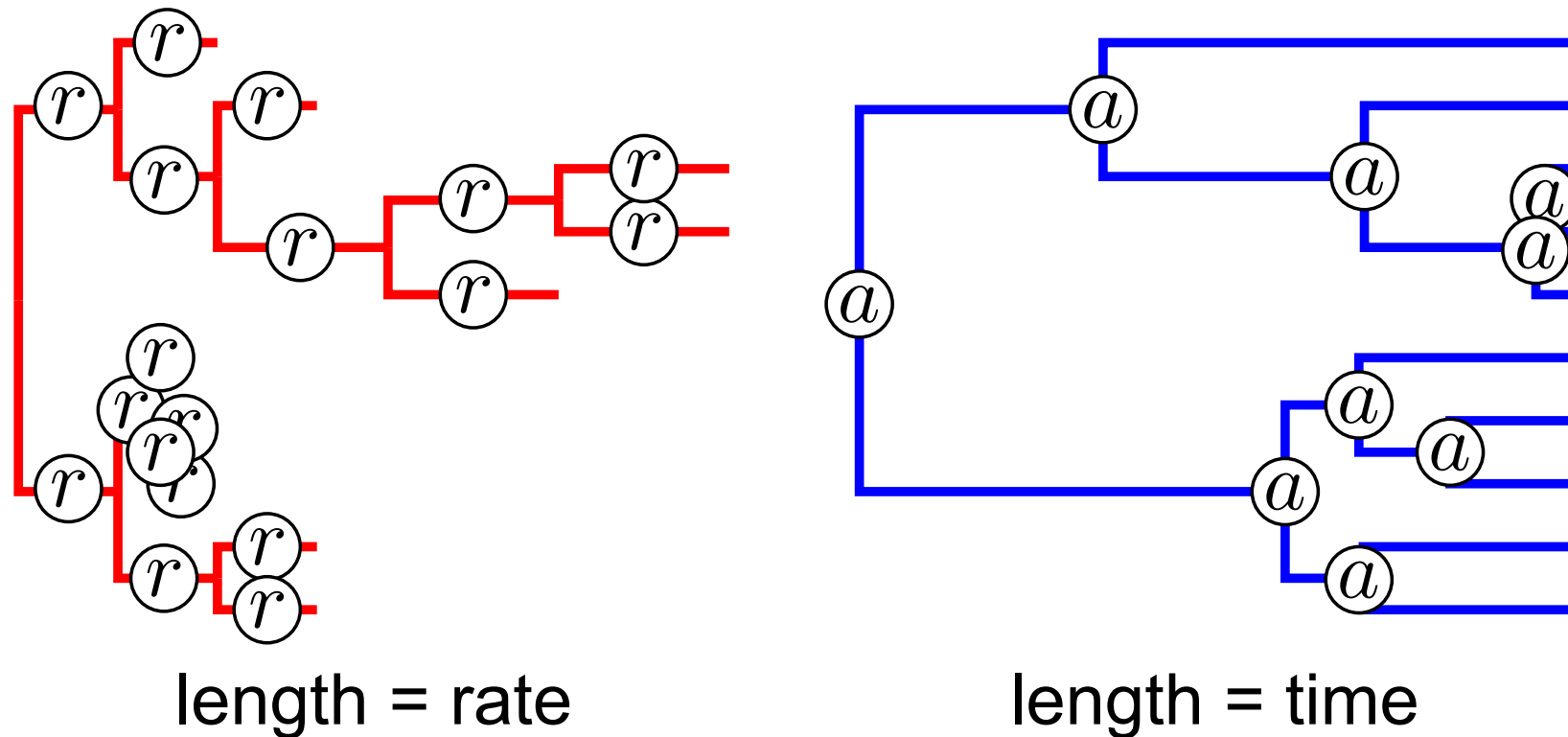


length = time

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

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

Bayesian Divergence Time Estimation



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

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

Bayesian Divergence Time Estimation

Posterior probability

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

\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
Ψ	Tree topology

Bayesian Divergence Time Estimation

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

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

\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

Ψ Tree topology

Modeling Rate Variation

Models describing lineage-specific substitution rate variation:

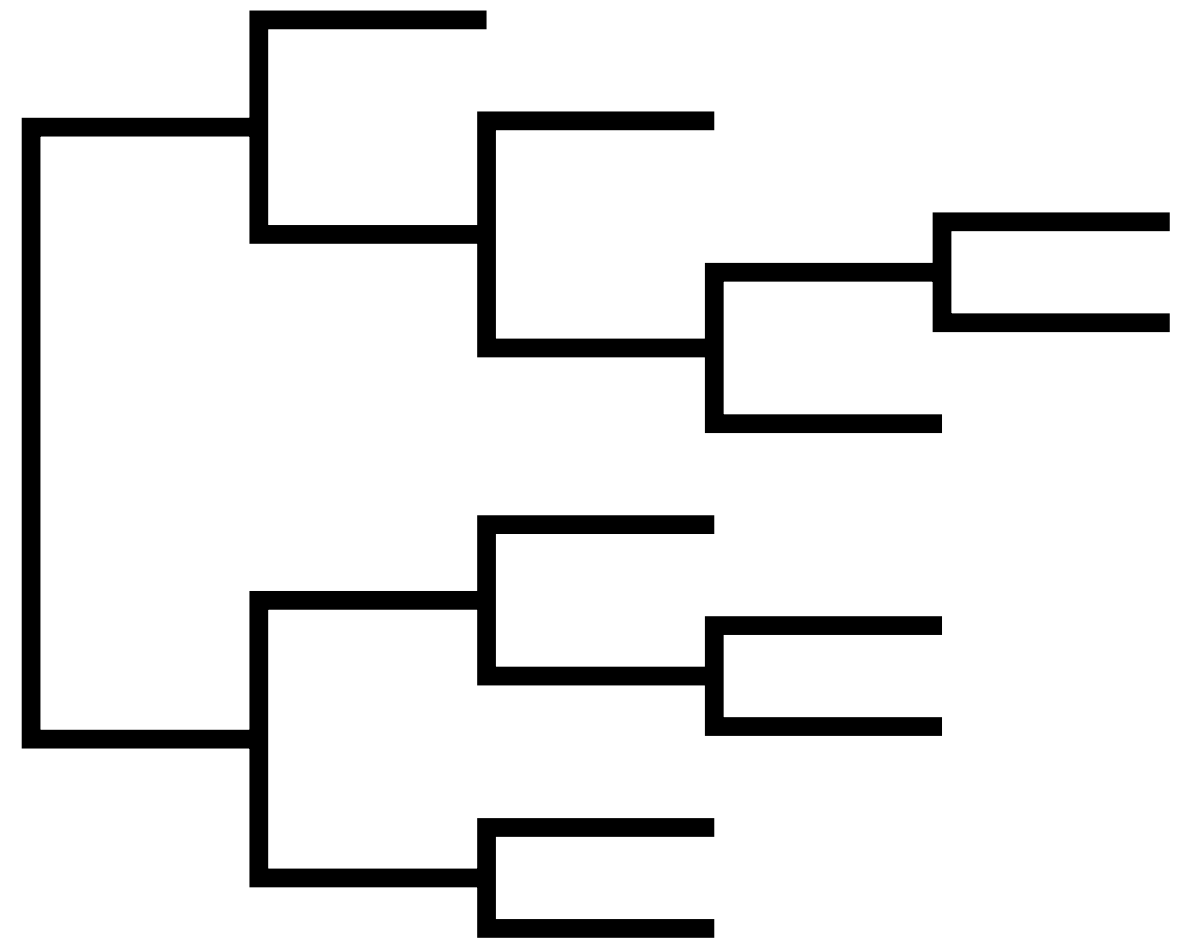
- **Global/strict clock** (Zuckerkandl & Pauling, 1962)
- **Local clocks** (Hasegawa, Kishino & Yano 1989; Kishino & Hasegawa 1990; Yoder & Yang 2000; Yang & Yoder 2003, Drummond and Suchard 2010)
- **Punctuated rate change model** (Huelsenbeck, Larget and Swofford 2000)
- **Autocorrelated rates** (Thorne, 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

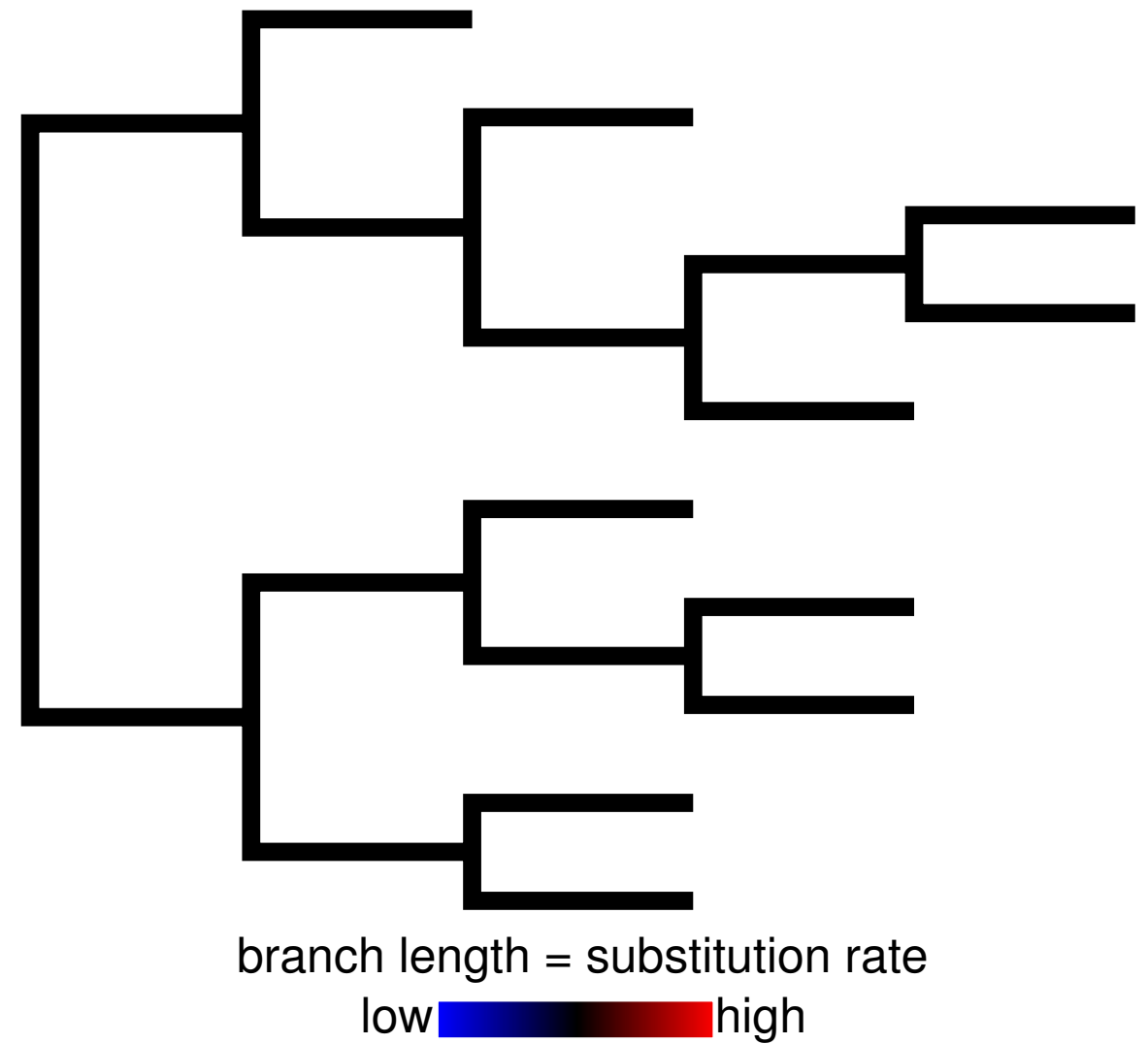
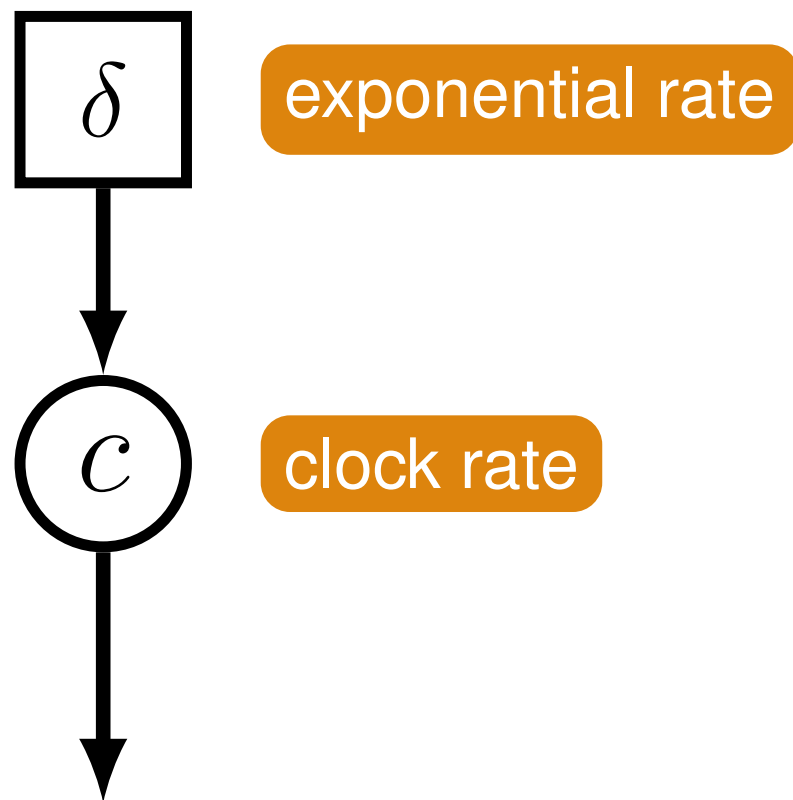
(Zuckerkandl & Pauling, 1962)



branch length = substitution rate
low  high

Global Molecular Clock

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



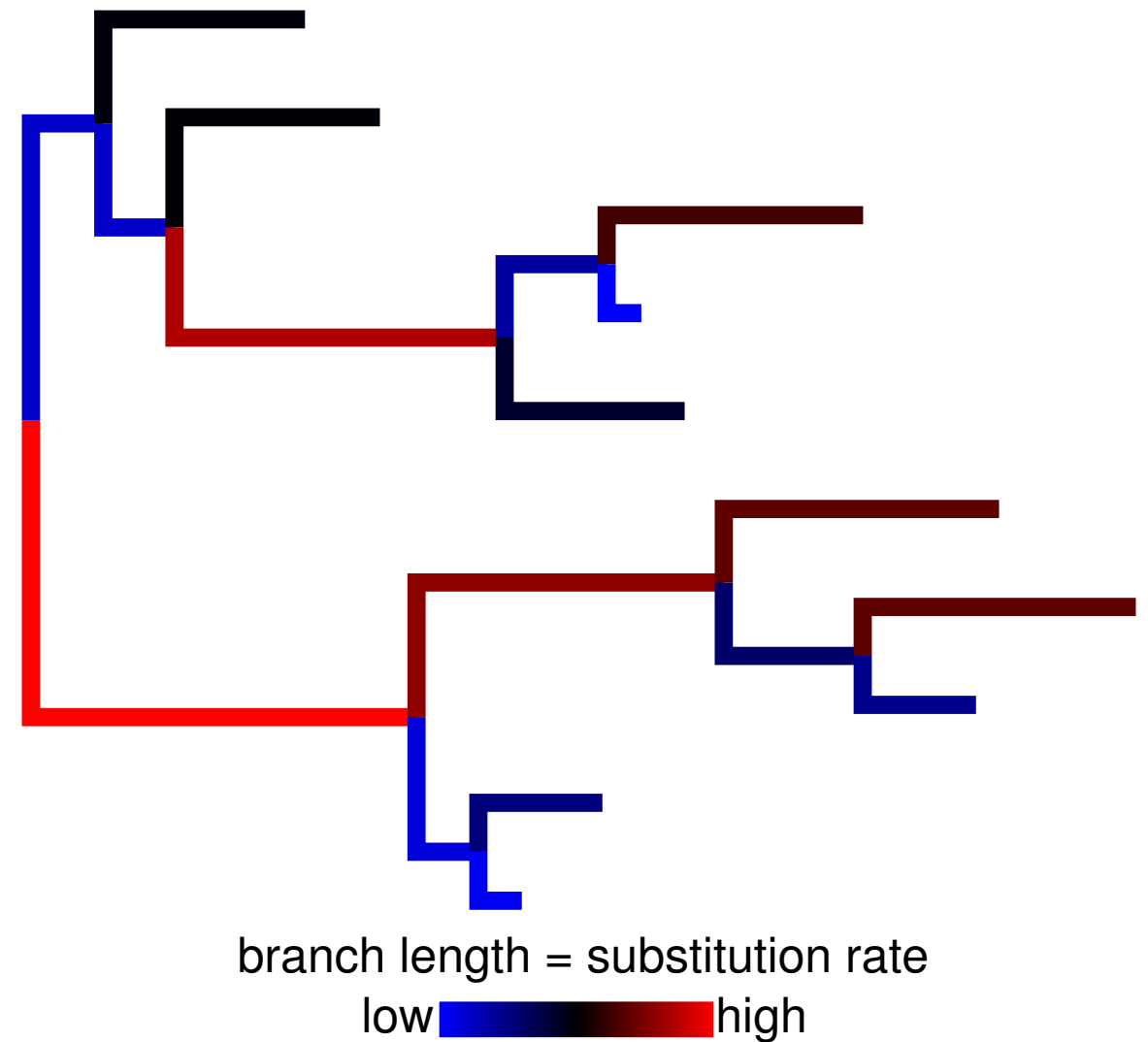
Relaxed-Clock Models

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

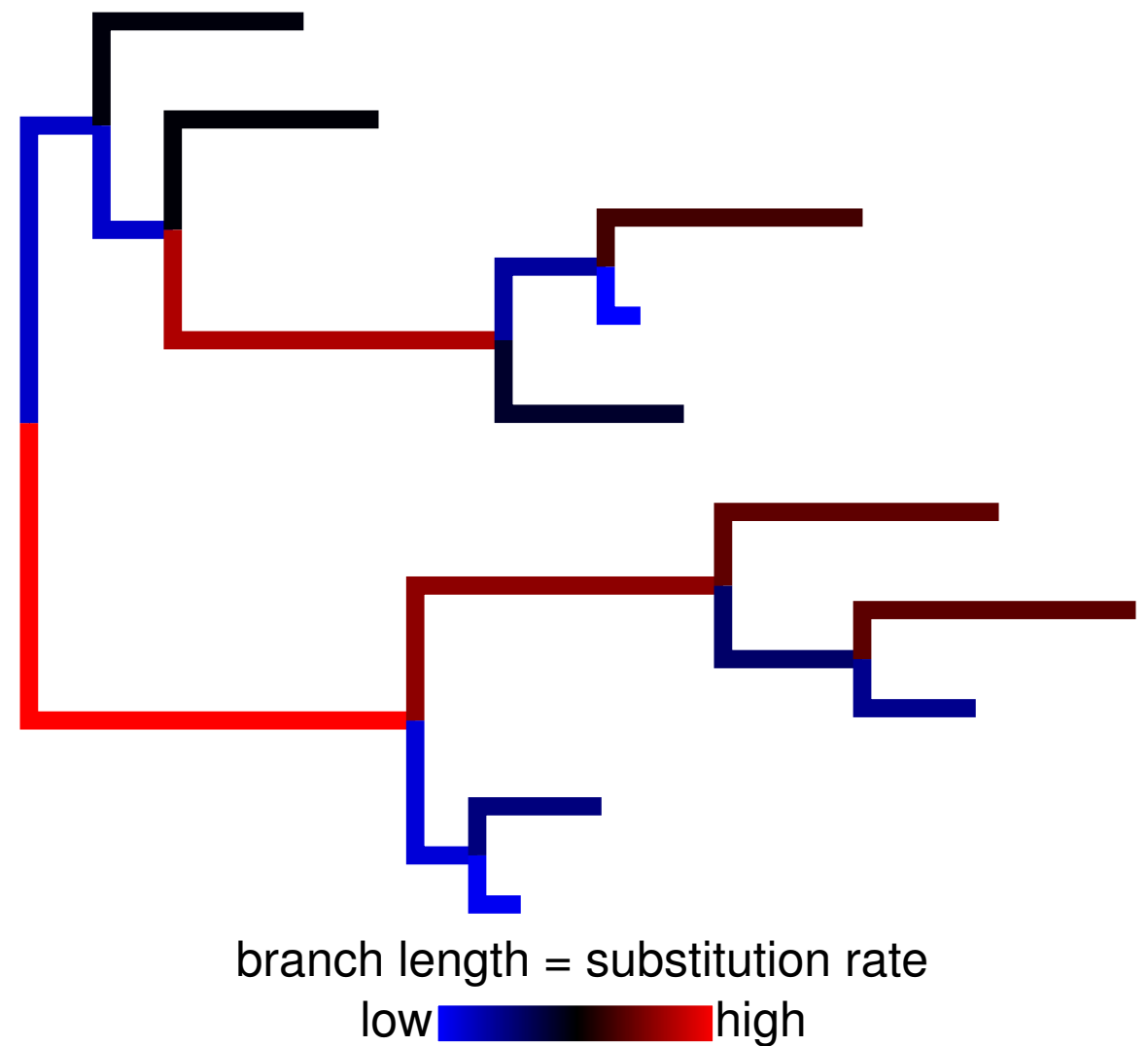
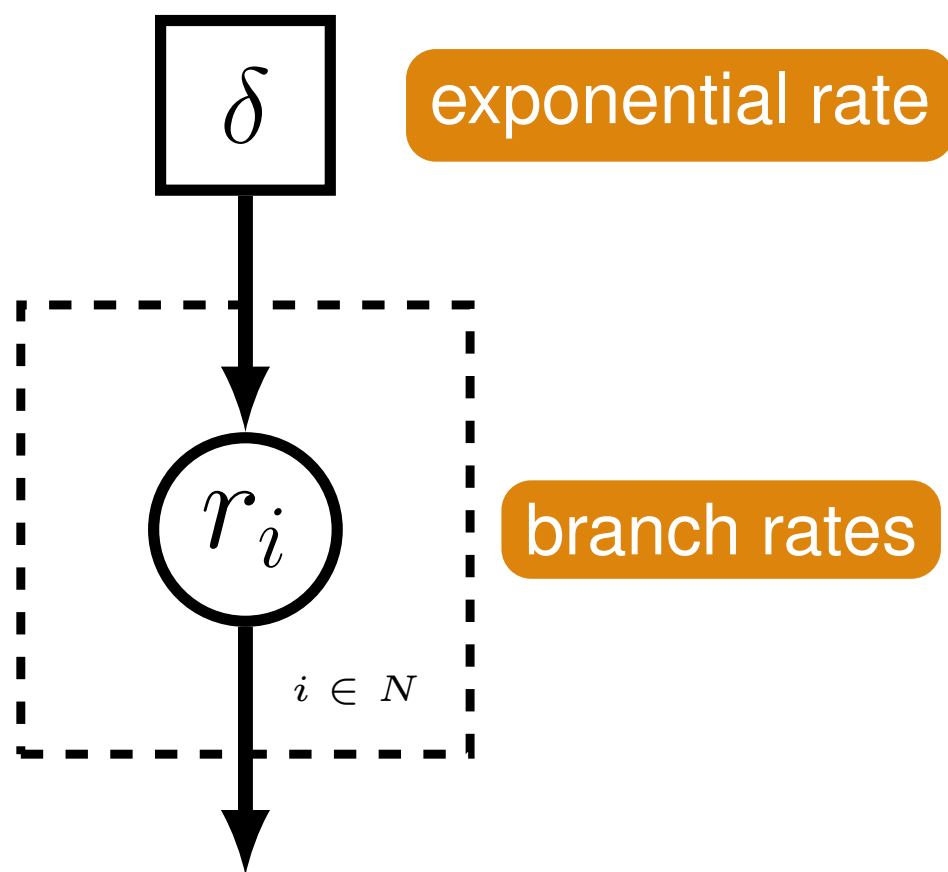
Lineage-specific rates are uncorrelated when the rate assigned to each branch is independently drawn from an underlying distribution



(Drummond et al. 2006; Rannala & Yang 2007; Lepage et al. 2007)

Independent/Uncorrelated Rates

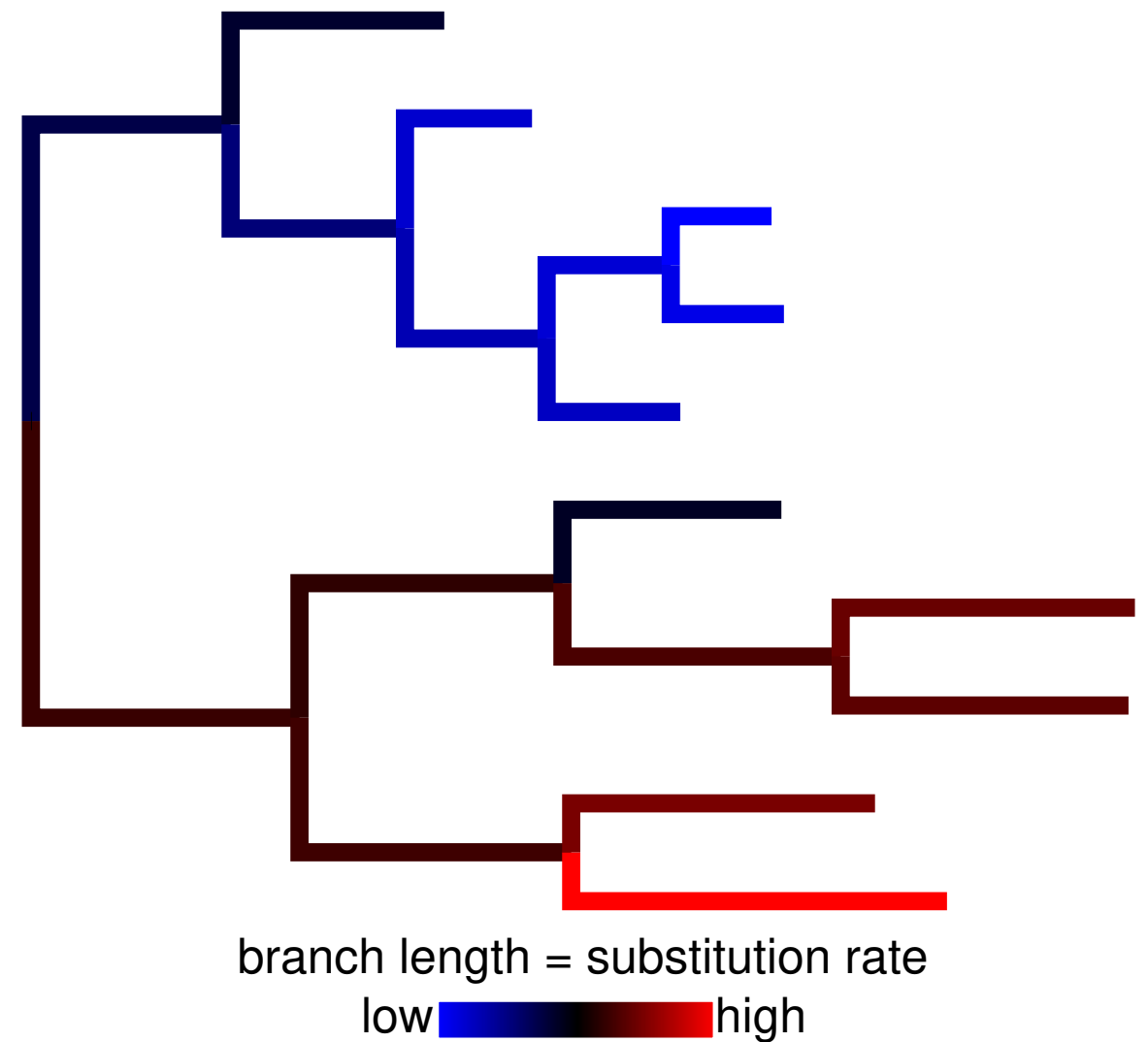
$$r_i \sim \text{Exponential}(\delta)$$



Autocorrelated Rates

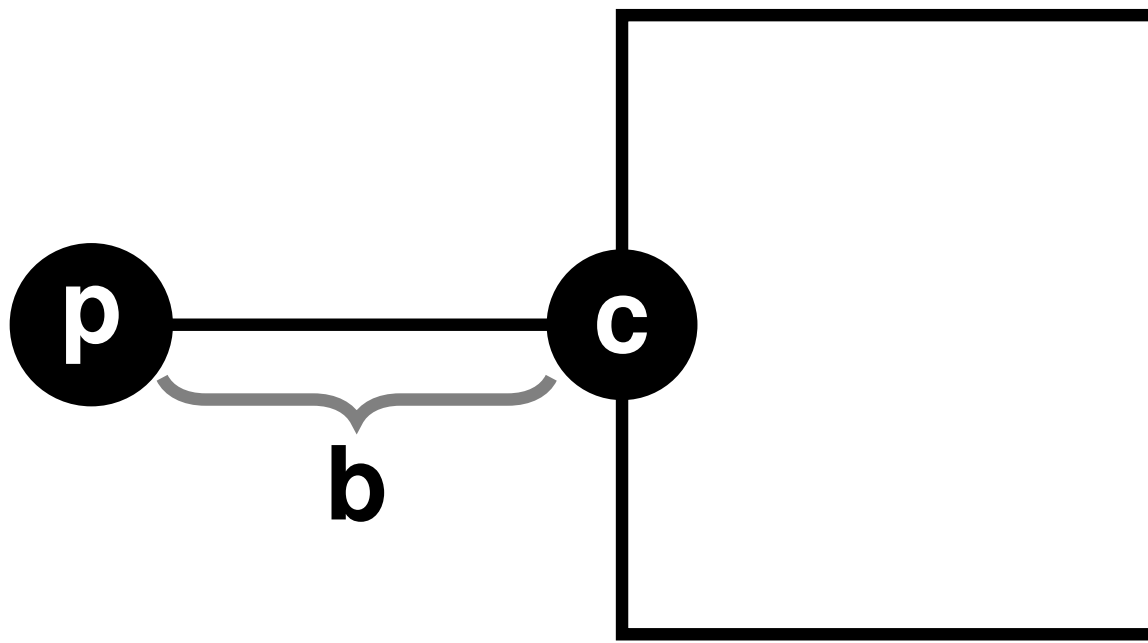
Substitution rates evolve gradually over time – closely related lineages have similar rates

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



(Thorne, Kishino, Painter, 1998; Kishino, Thorne, Bruno, 2001)

Autocorrelated Rates



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

$$\sigma_c := \nu t_b$$

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

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

p = parent node

c = child node

b = branch

ν = variance parameter

t_b = time duration of branch

r_c

μ_c

r_b

r_p

σ_c

t_b

ν

Autocorrelated Rates

Draw the graphical model!

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

$$\sigma_c := \nu t_b$$

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

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

ν = variance parameter

t_b = time duration of branch

$$r_c$$

$$\mu_c$$

$$r_b$$

$$r_p$$

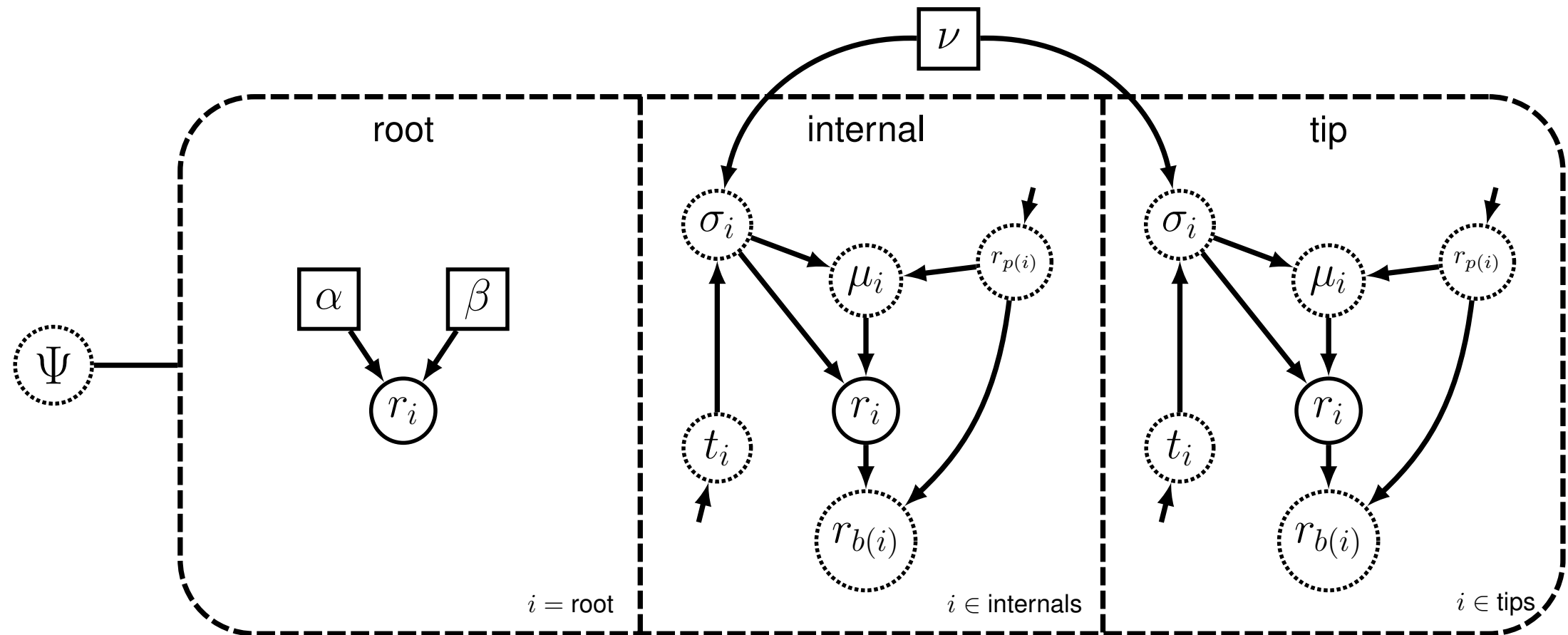
$$\sigma_c$$

$$t_b$$

$$\nu$$

Autocorrelated Rates

The rate associated with each node is a stochastic node, drawn from a distribution centered on its parent node



There is a gamma prior distribution on the rate at the root node

Relaxed-Clock Models

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

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

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

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

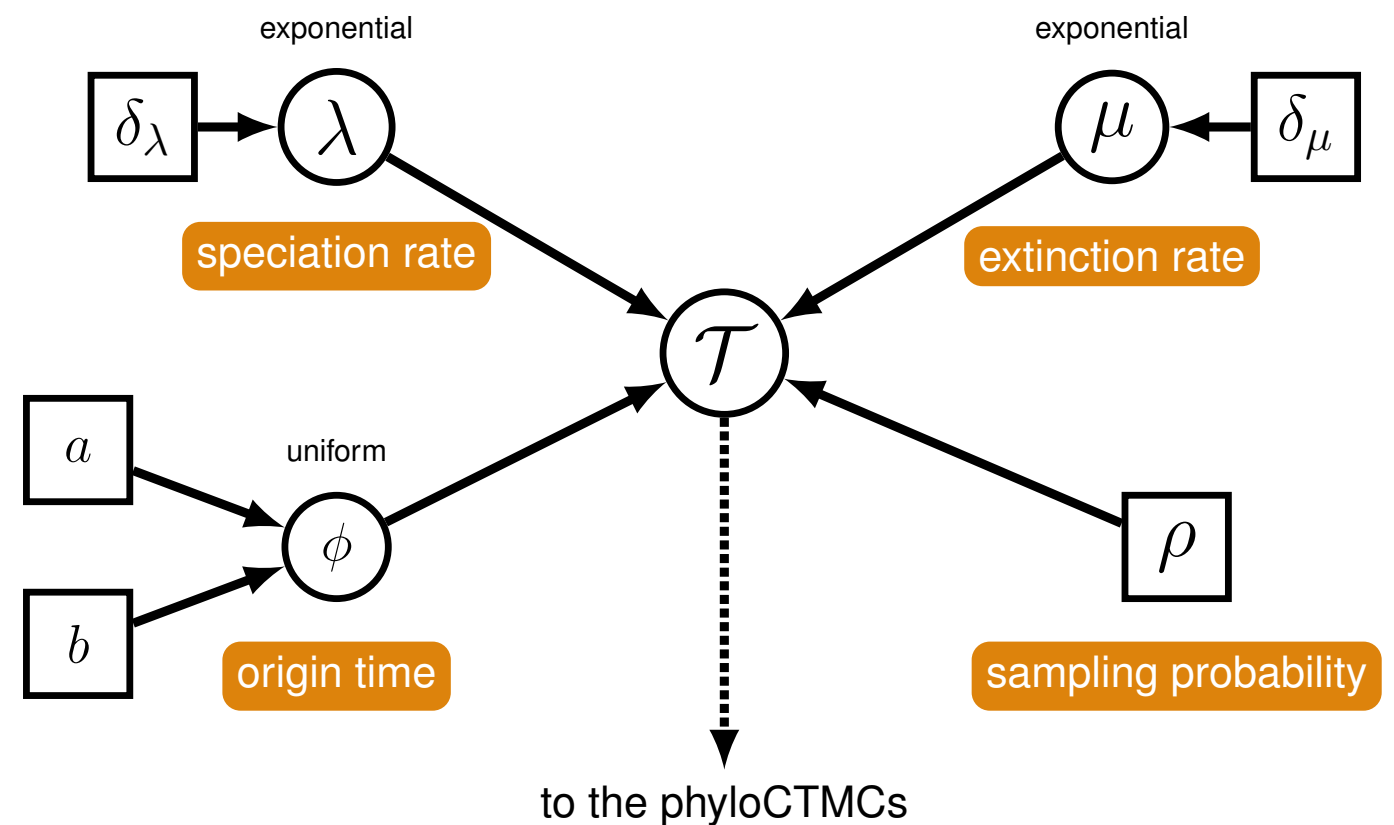
Stochastic Branching Processes

Tree priors based on stochastic models of lineage diversification

Birth-death-sampling

process: at any point in time a lineage can speciate at rate λ or go extinct with a rate of μ

Conditions on a probability of sampling a tip ρ , and the origin time of the process ϕ

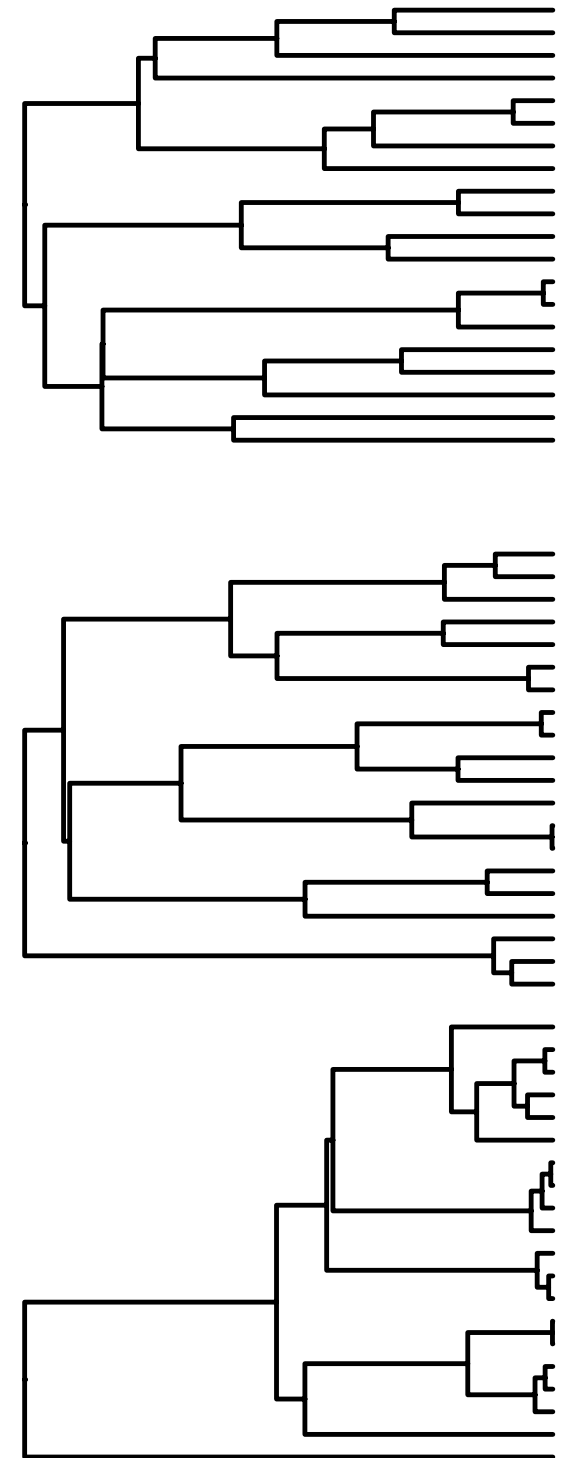


Stochastic Branching Processes

Different values of λ and μ lead to different trees

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

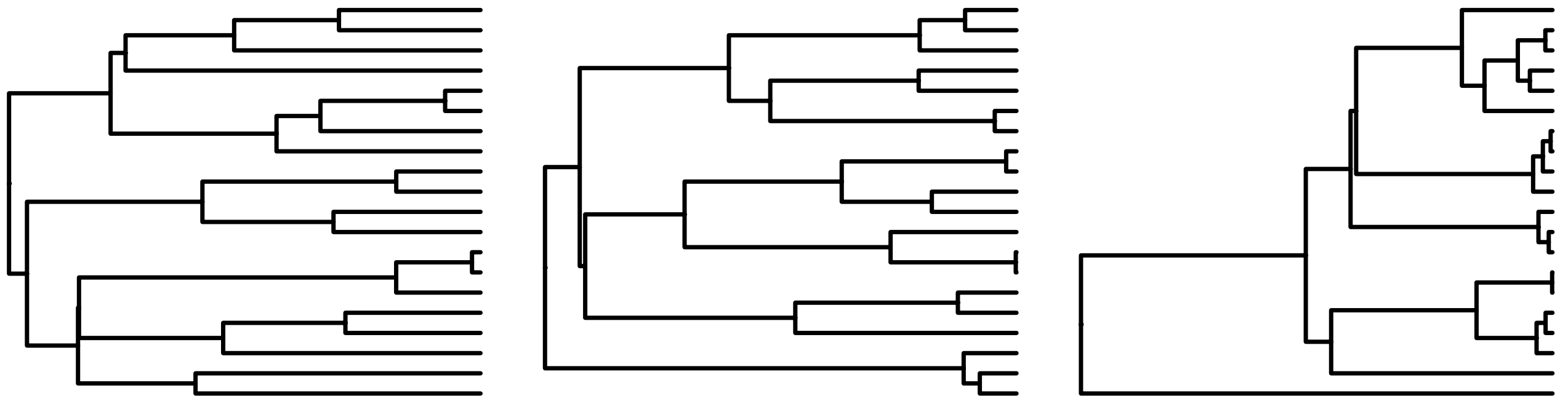
Using hyperpriors on λ and μ accounts for uncertainty in these hyperparameters



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



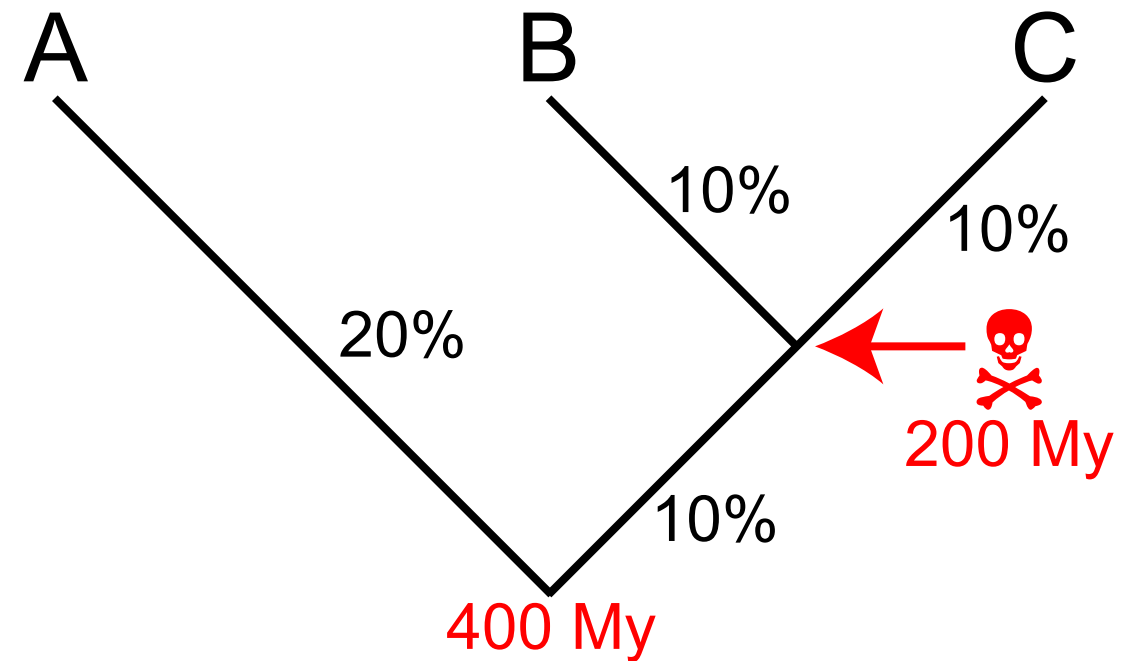
We need additional data (like fossils) to provide an absolute time scale

Calibrating Divergence Times

Fossils (or other data) are necessary to estimate *absolute* node ages

There is **no information** in the sequence data for absolute time

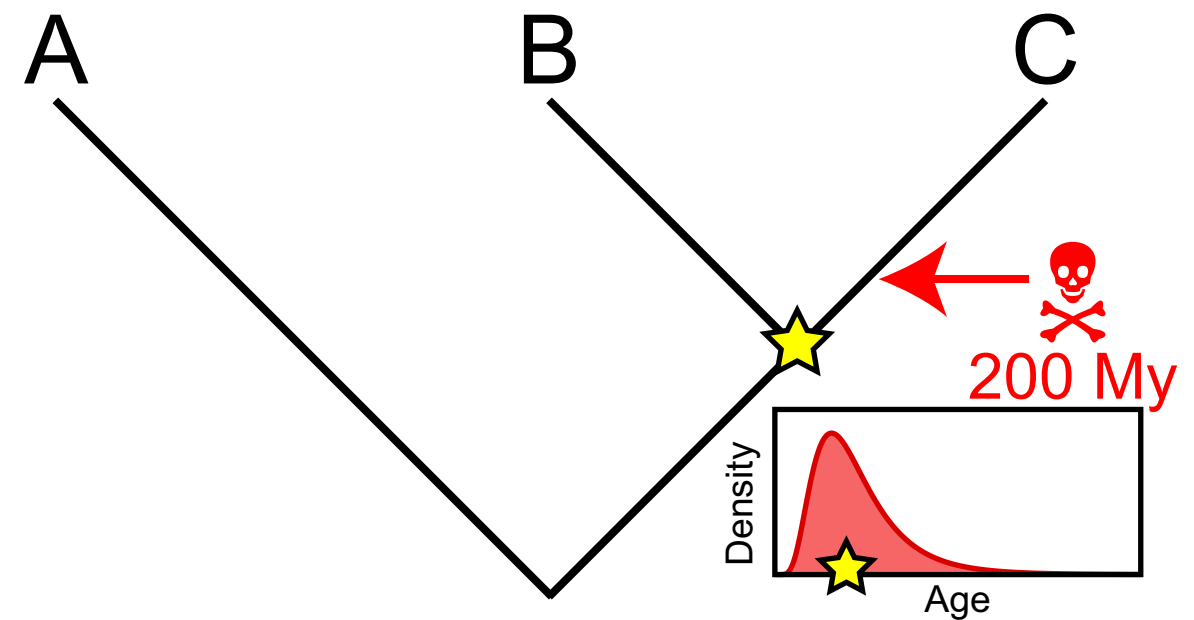
The placement and age of fossil taxa are always uncertain



Calibrating Divergence Times

Bayesian inference is well suited to accommodating uncertainty in the age of the calibration node

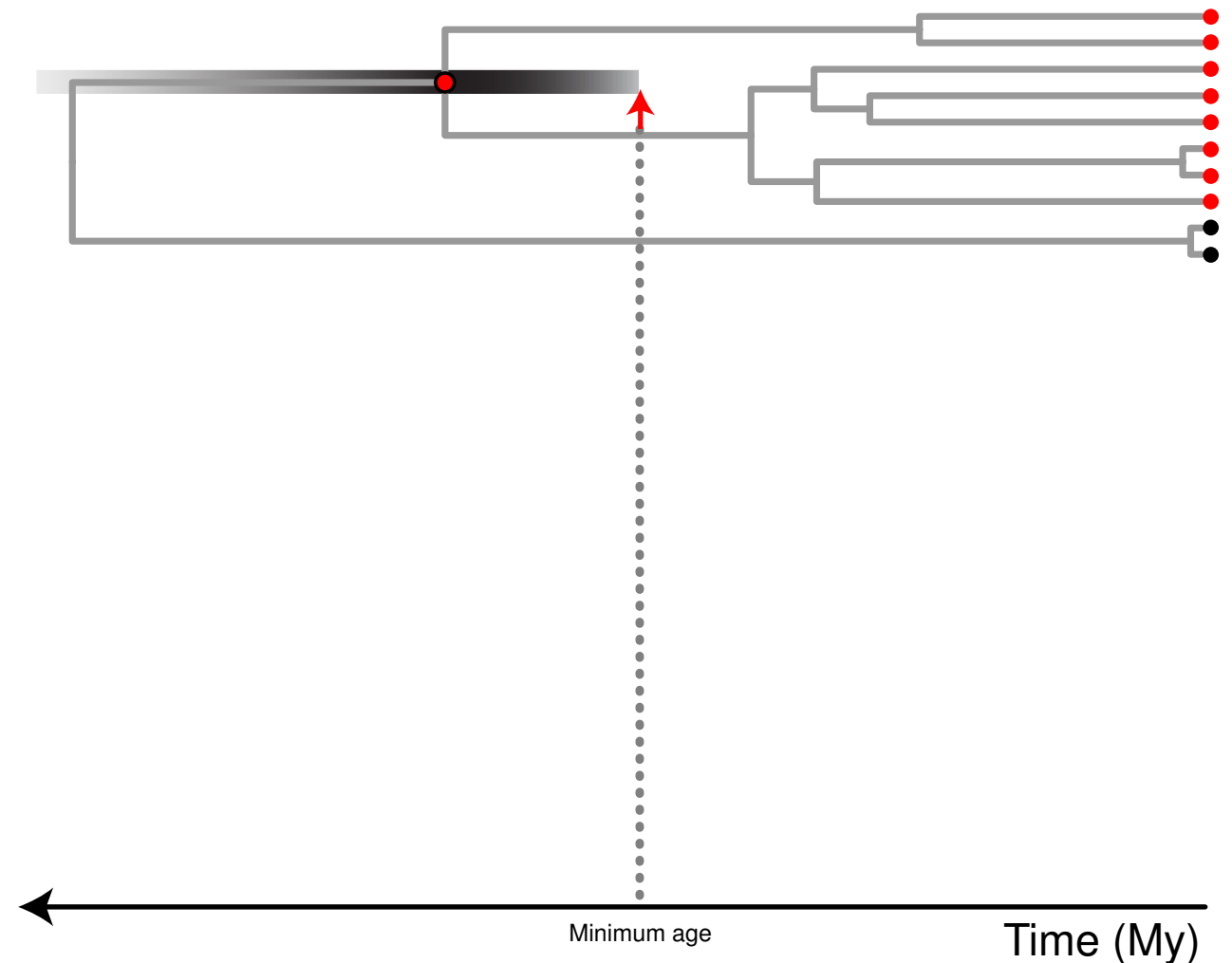
Divergence times are calibrated by placing parametric densities on internal nodes offset by age estimates from the fossil record



Fossil Calibration

Age estimates from fossils can provide **minimum** time constraints for internal nodes

Reliable **maximum** bounds are typically unavailable

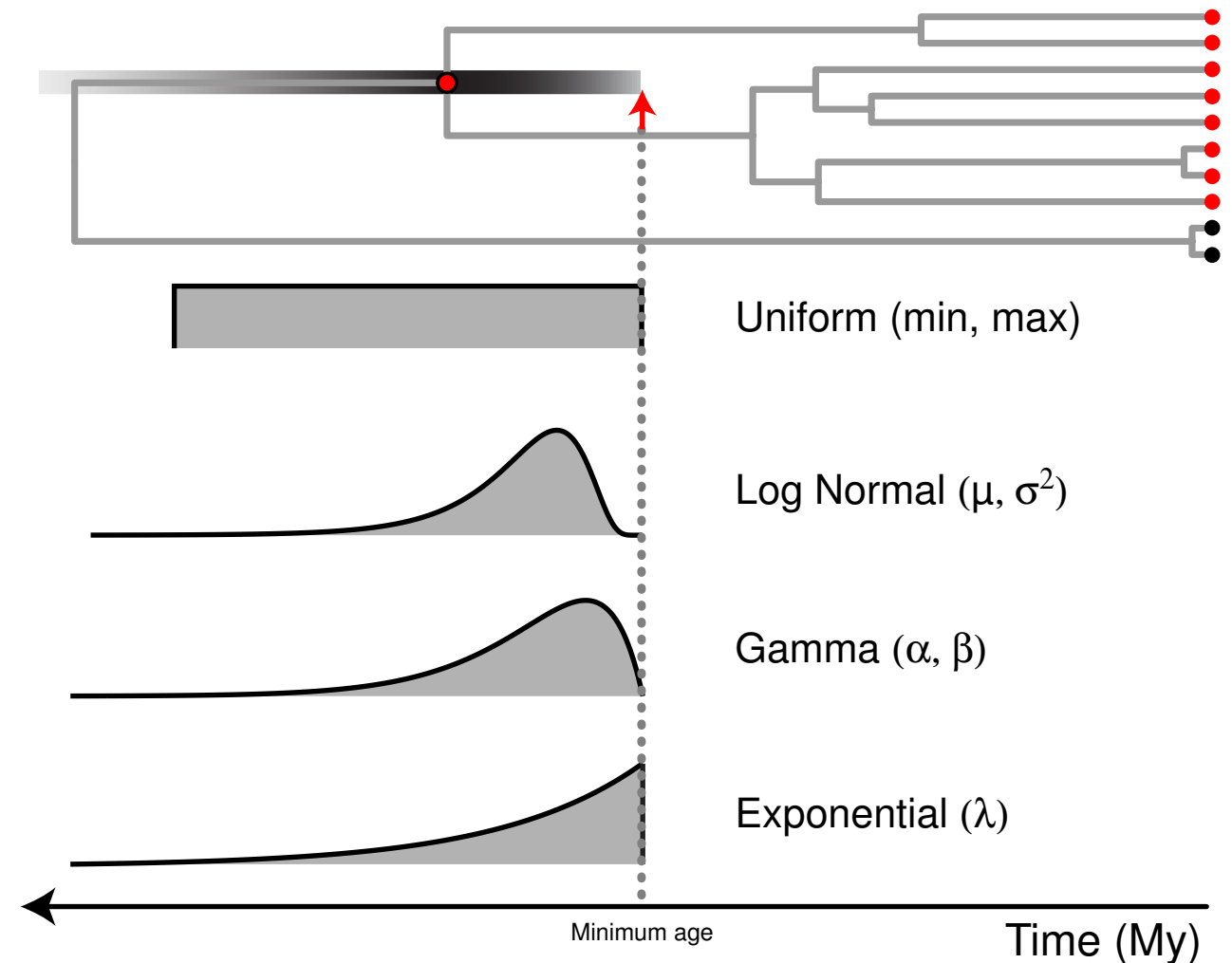


Prior Densities on Calibrated Nodes

Common practice in Bayesian divergence-time estimation

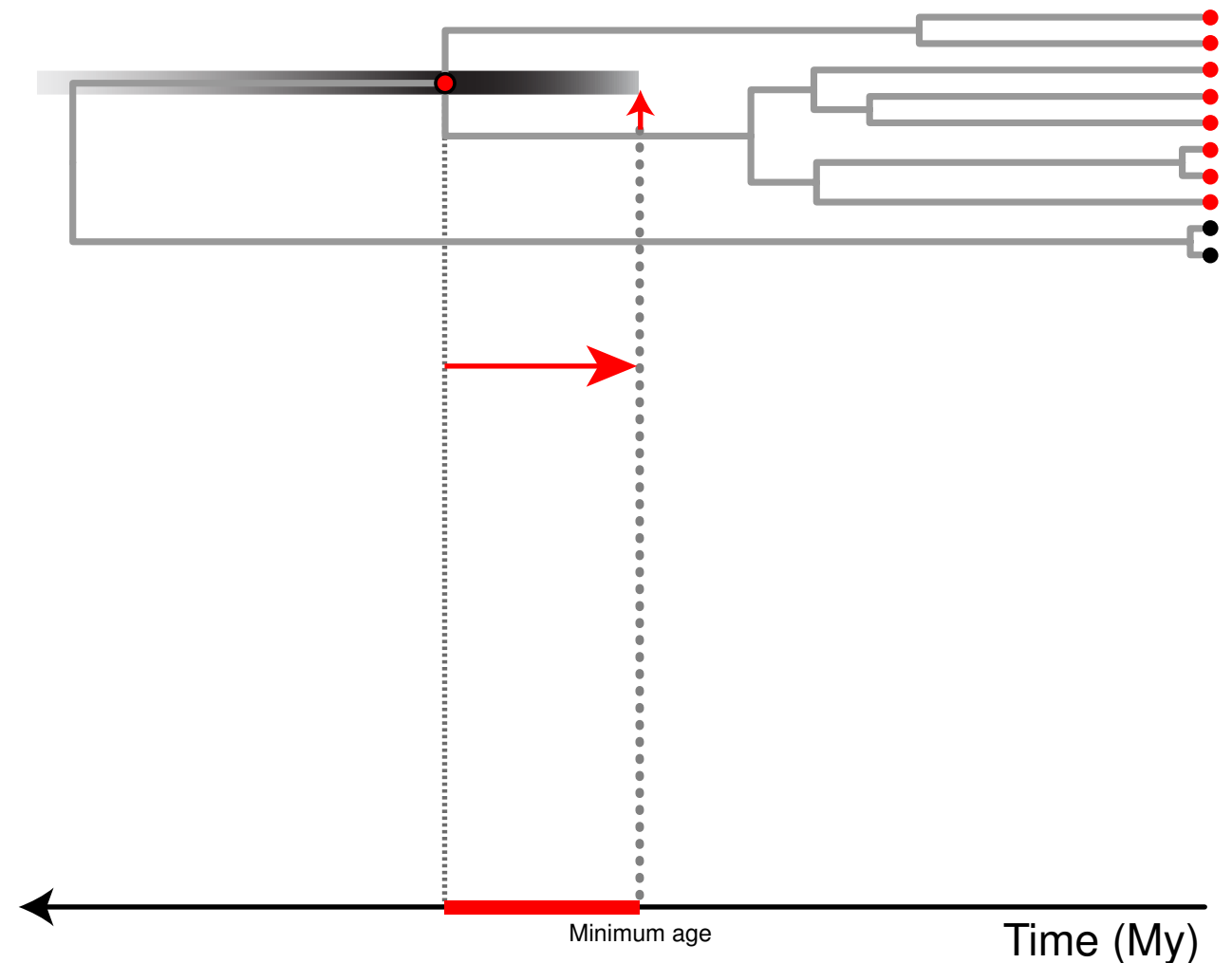
Parametric distributions are typically off-set by the age of the oldest fossil assigned to a clade

These prior densities do not (necessarily) require specification of maximum bounds



Prior Densities on Calibrated Nodes

Describe the waiting time between the divergence event and the age of the oldest fossil

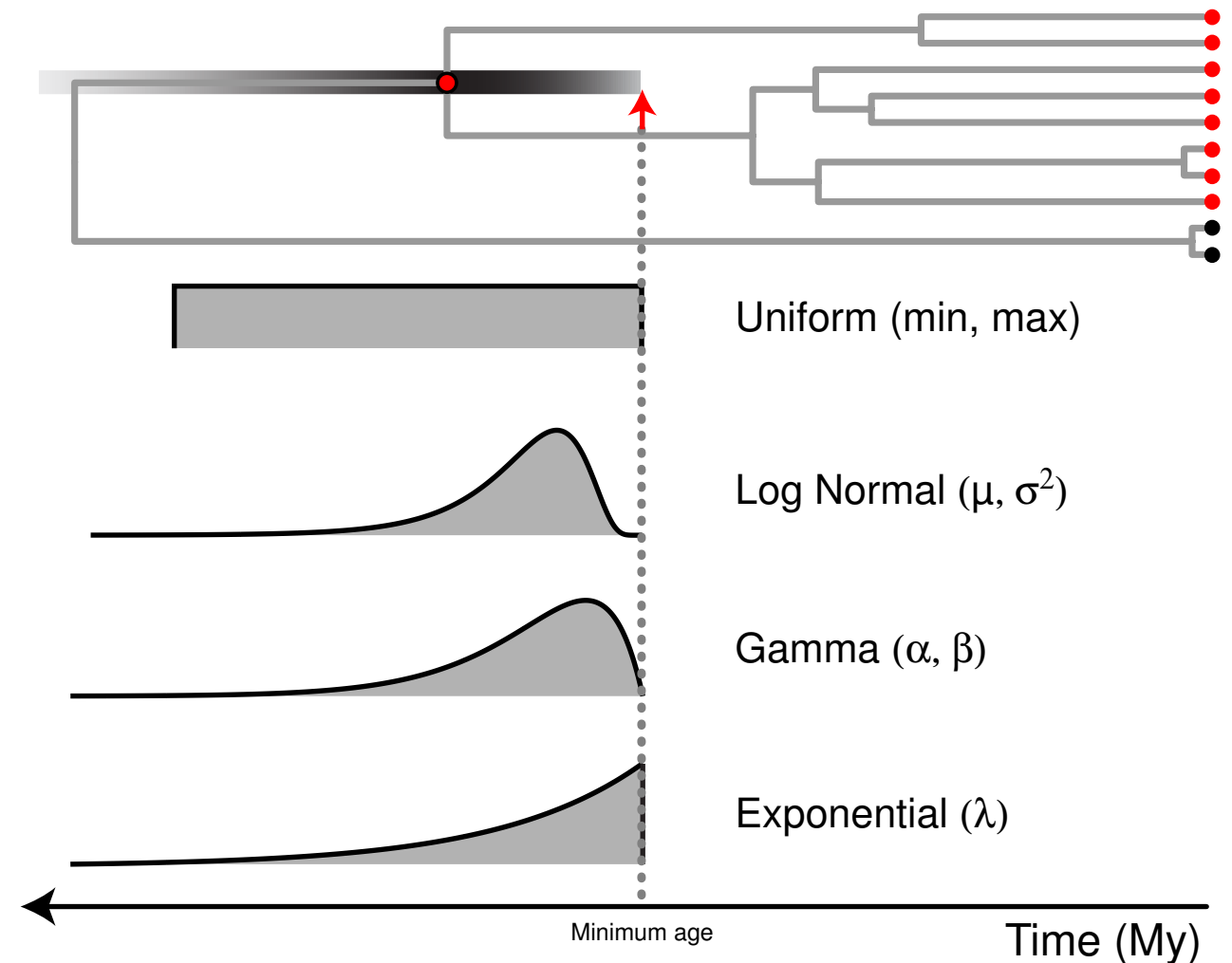


Prior Densities on Calibrated Nodes

Common practice in Bayesian divergence-time estimation

Estimates of absolute node ages are driven primarily by the calibration density

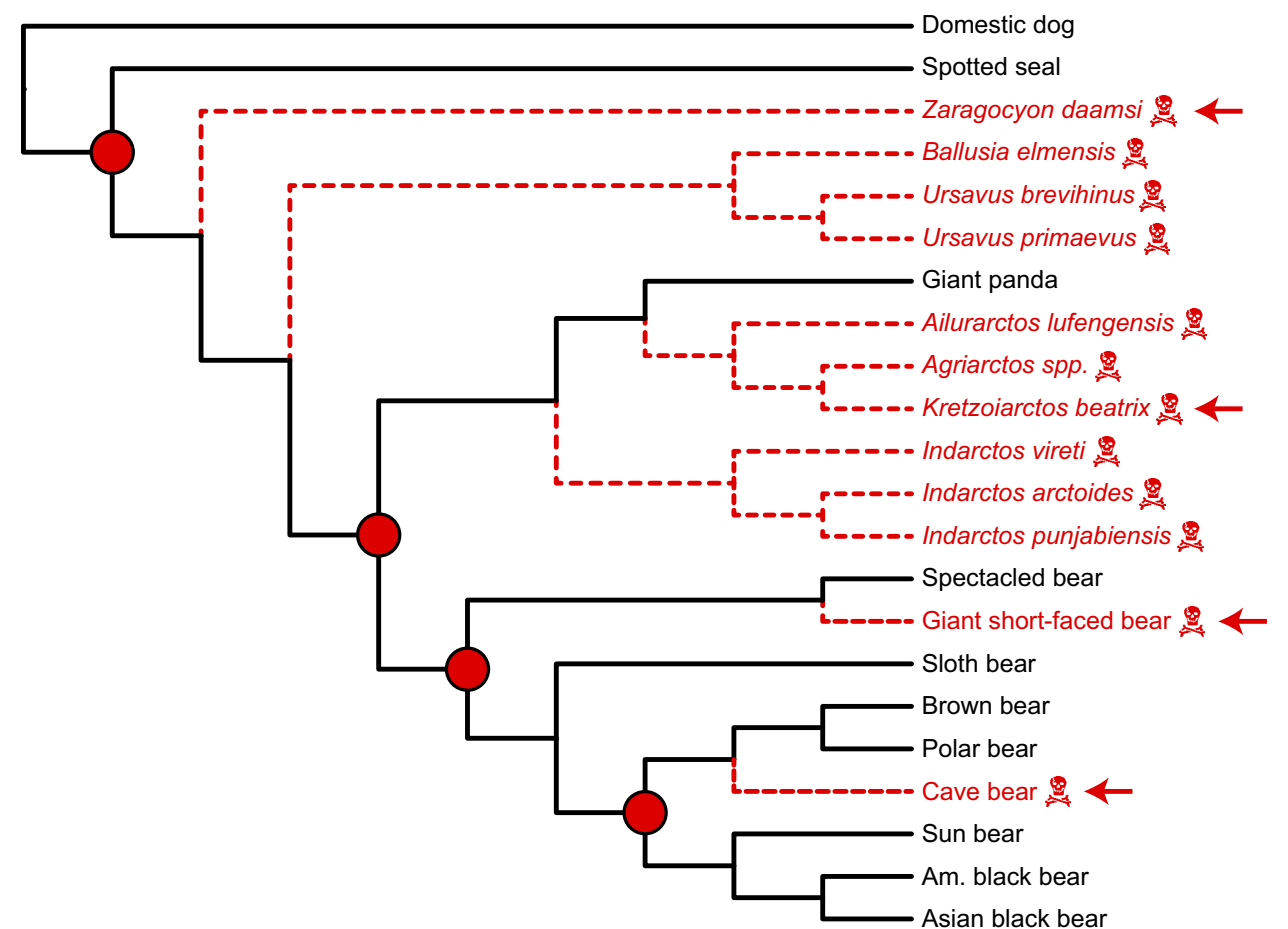
Specifying appropriate densities is a challenge for most molecular biologists



Improving Fossil Calibration

We would prefer to eliminate the need for *ad hoc* calibration densities

Calibration densities do not account for diversification of fossils

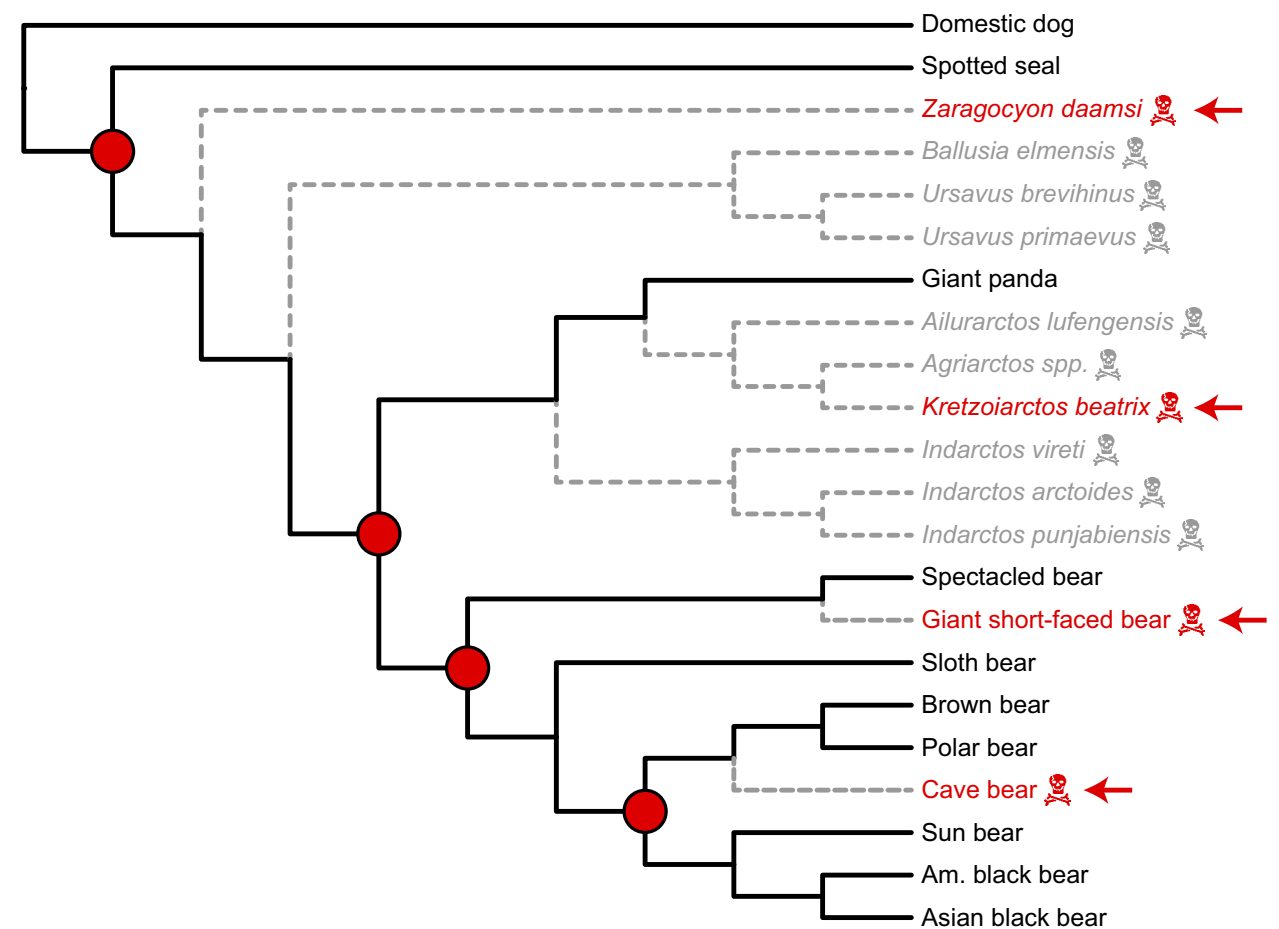


(Krause et al. BMC Evol. Biol. 2008; Abella et al. PLoS ONE 2012)

Improving Fossil Calibration

Example: Bears

12 fossils are reduced to 4 calibration ages if using node densities

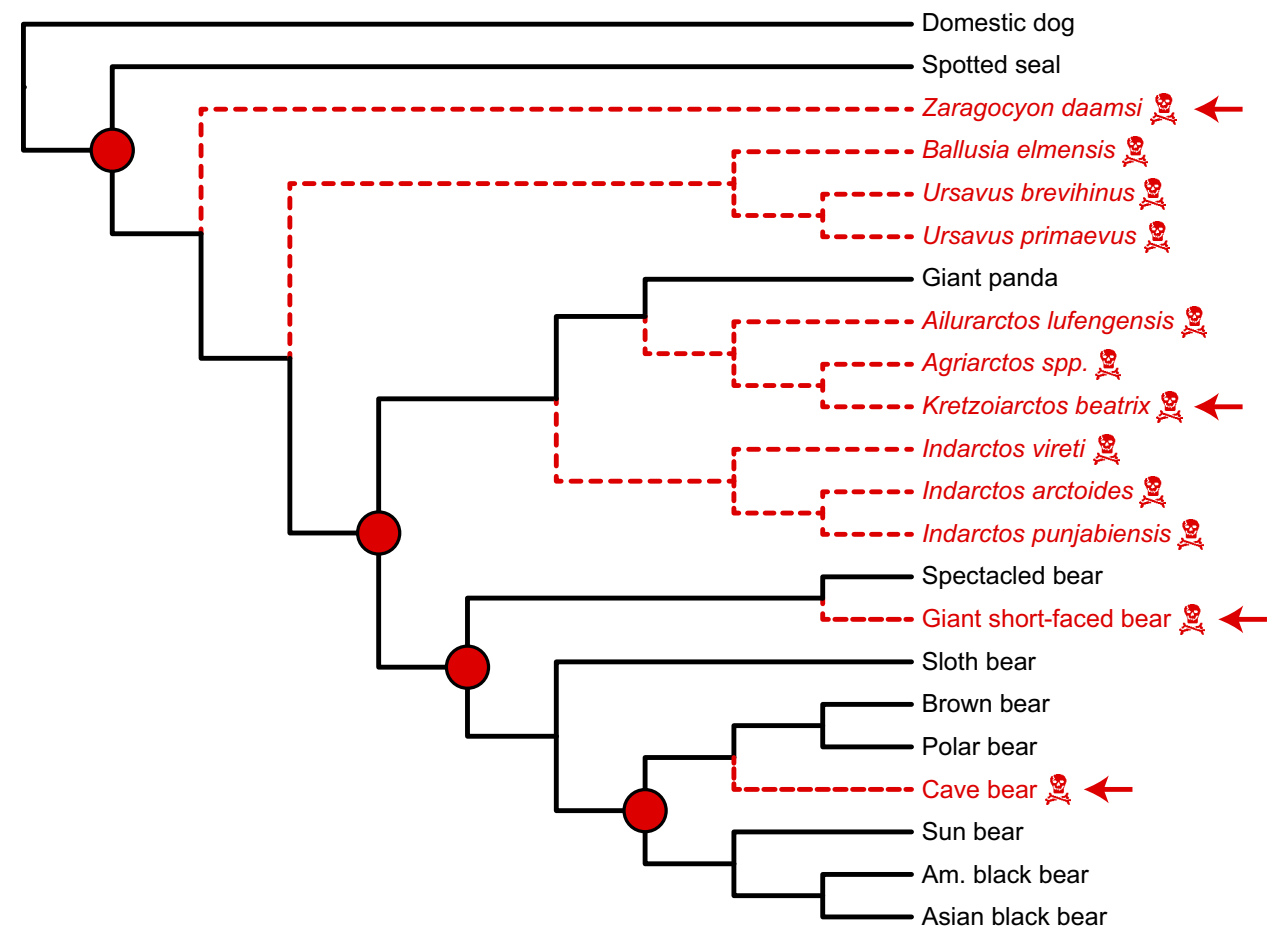


(Krause et al. BMC Evol. Biol. 2008; Abella et al. PLoS ONE 2012)

Improving Fossil Calibration

We want to use *all* of the available fossils

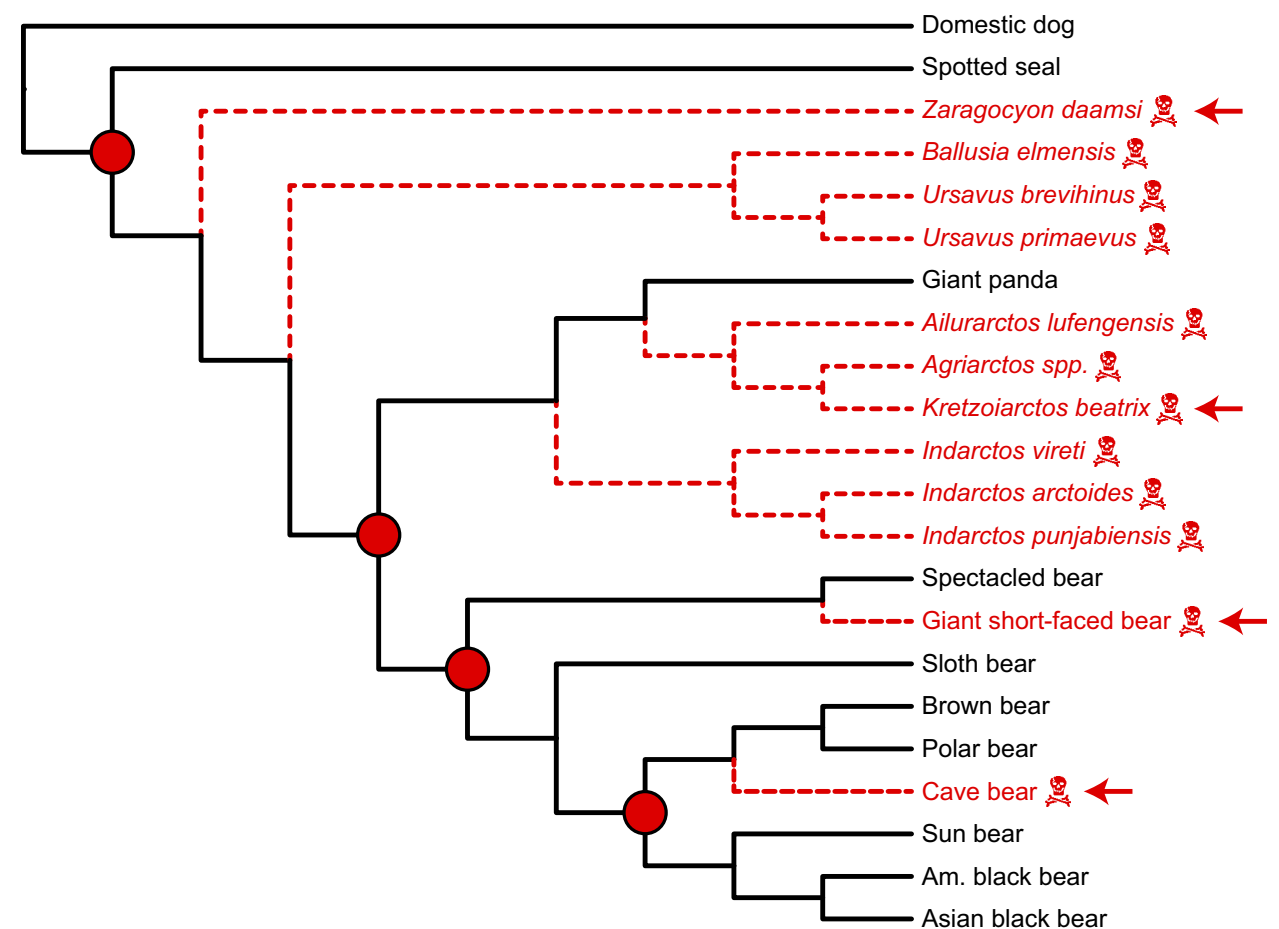
Because fossils are part of the same diversification process that gave rise to extant taxa, we can include them in the birth-death model



(Krause et al. BMC Evol. Biol. 2008; Abella et al. PLoS ONE 2012)

Improving Fossil Calibration

This relies on a branching model that accounts for **speciation, extinction** and rates of **fossilization, preservation, and recovery**



(Krause et al. BMC Evol. Biol. 2008; Abella et al. PLoS ONE 2012)

Paleontology & Neontology

"Except during the interlude of the [Modern] Synthesis, there has been limited communication historically among the disciplines of evolutionary biology, particularly between students of evolutionary history (paleontologists and systematists) and those of molecular, population, and organismal biology. **There has been increasing realization that barriers between these subfields must be overcome if a complete theory of evolution and systematics is to be forged.**"

Reaka-Kudla, M.L. & Colwell, R.: in E.C. Dudley (ed.), *The Unity of Evolutionary Biology: Proceedings of the Fourth International Congress of Systematic & Evolutionary Biology*, Discorides Press, Portland, OR, p. 16.

Paleontology & Neontology



Biology and Philosophy **19**: 687–720, 2004.

© 2004 Kluwer Academic Publishers. Printed in the Netherlands.

**The role of fossils in phylogeny reconstruction:
Why is it so difficult to integrate paleobiological and
neontological evolutionary biology?**

TODD GRANTHAM

Department of Philosophy, College of Charleston, Charleston, SC 29424, USA

(e-mail: granthamt@cofc.edu)

Combining Fossil & Extant Data

Statistical methods provide a way to integrate paleontological and neontological data

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*

Syst. Biol. 61(6):973–999, 2012

© The Author(s) 2012. Published by Oxford University Press

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI:10.1093/sysbio/sys058

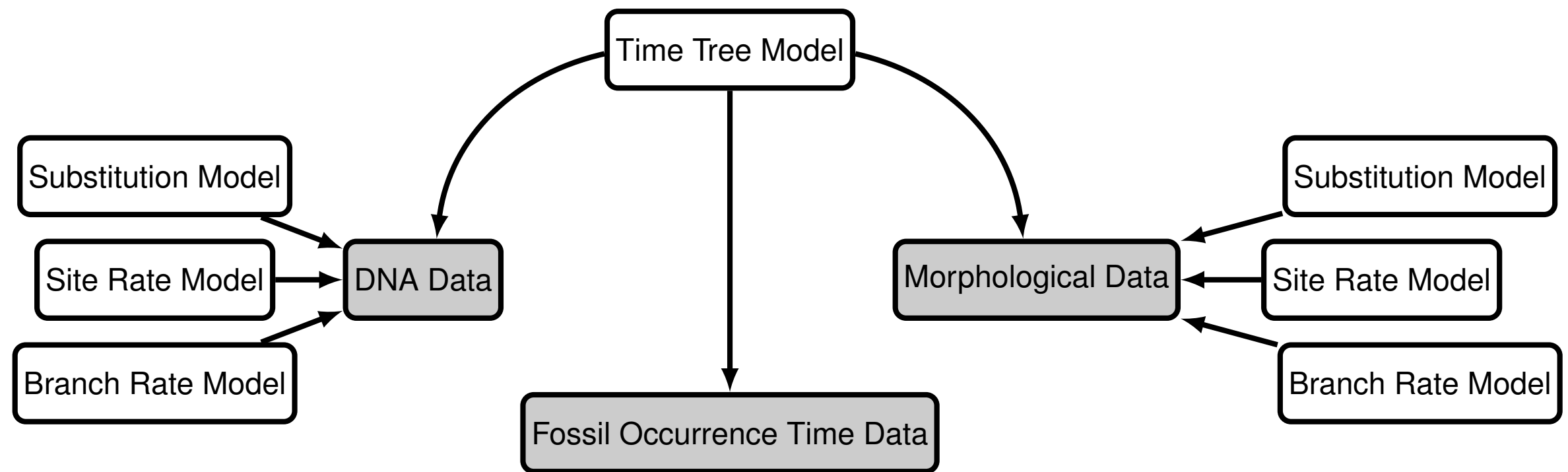
Advance Access publication on June 20, 2012

A Total-Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera

FREDRIK RONQUIST^{1,*}, SERAINA KLOPFSTEIN¹, LARS VILHELMSSEN², SUSANNE SCHULMEISTER³, DEBRA L. MURRAY⁴, AND
ALEXANDR P. RASNITSYN⁵

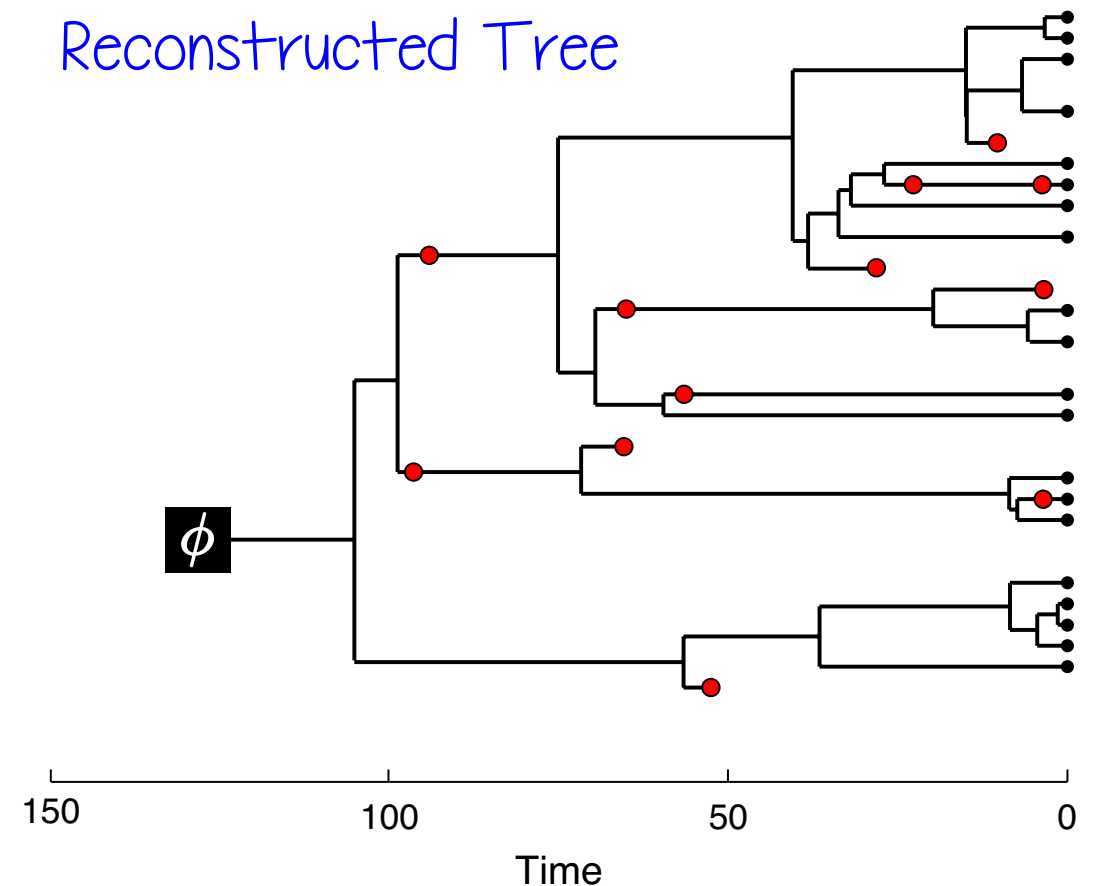
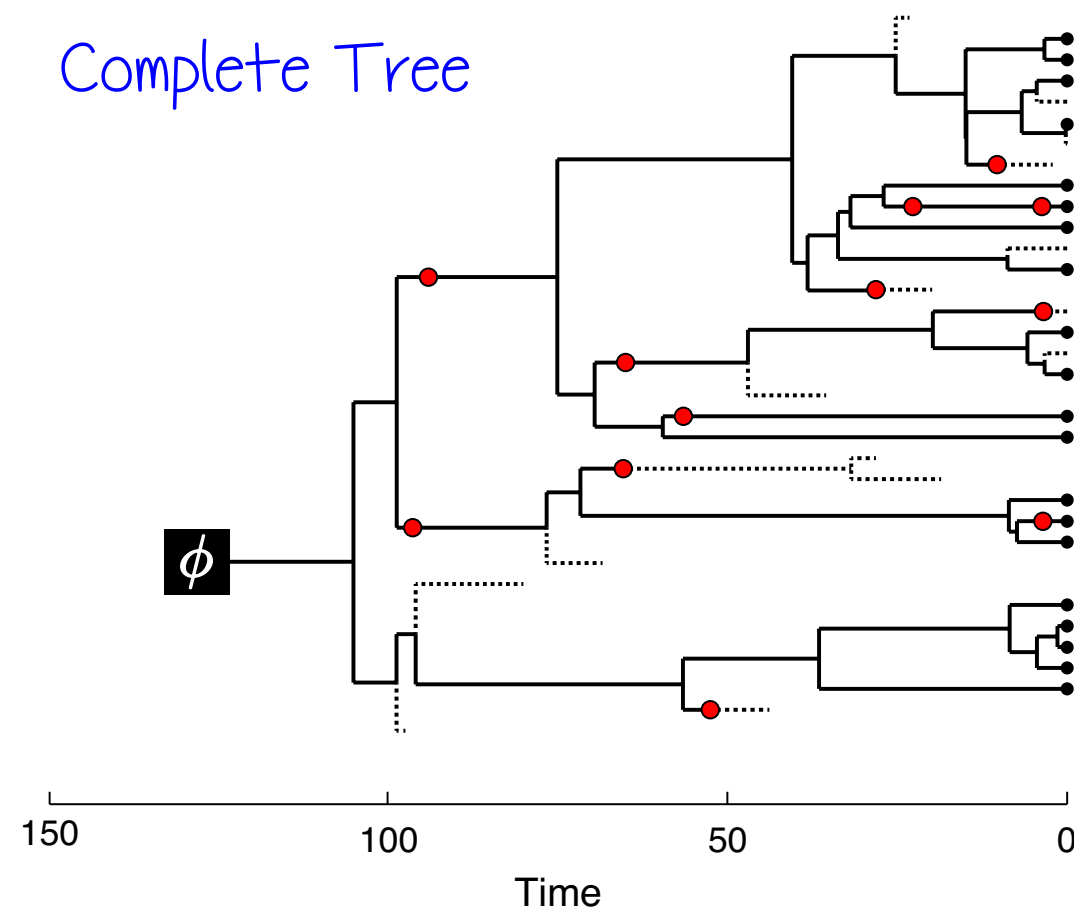
Combining Fossil & Extant Data

Combine models for sequence evolution, morphological change, & fossil recovery to jointly estimate the tree topology, divergence times, & lineage diversification rates



Modeling the Tree & Occurrence Times

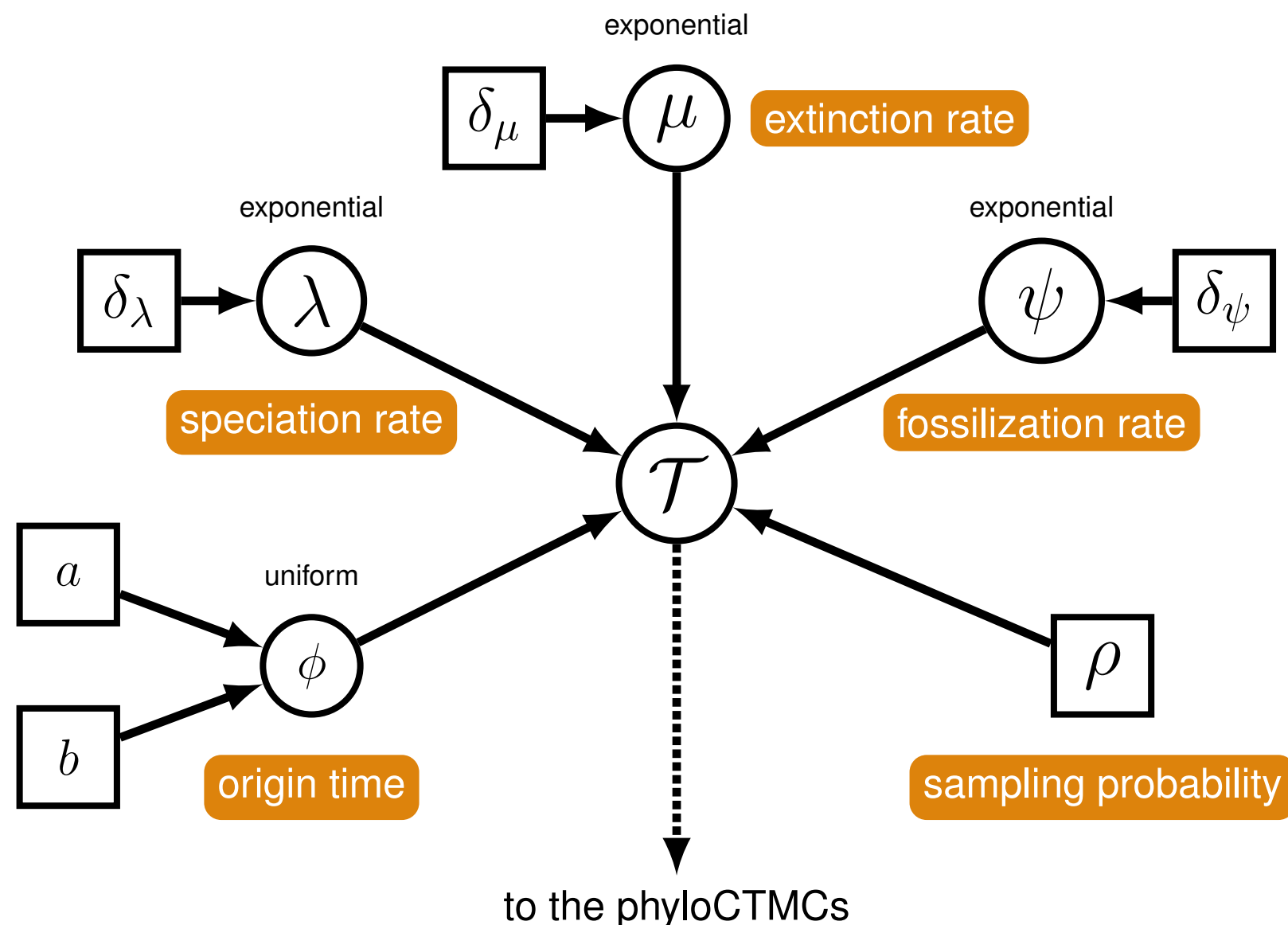
Stadler (2010) introduced a generating model for a serially sampled time tree – this is the *fossilized birth-death process*



(Stadler. *Journal of Theoretical Biology*. 2010)

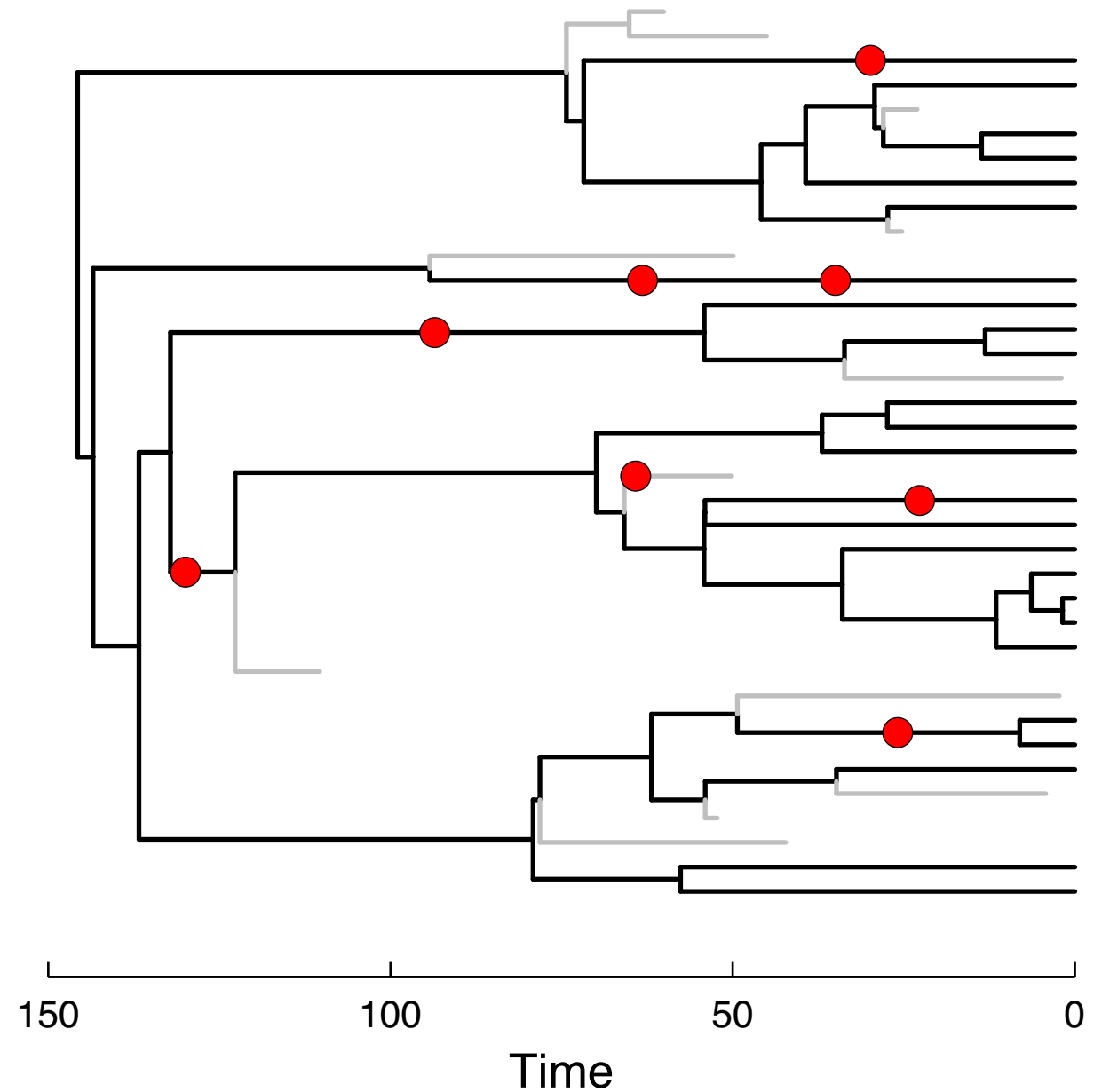
Parameters of the FBD

This graph shows the conditional dependence structure of the FBD model, which is a generating process for a sampled, dated time tree and fossil occurrences



The Fossilized Birth-Death Process (FBD)

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



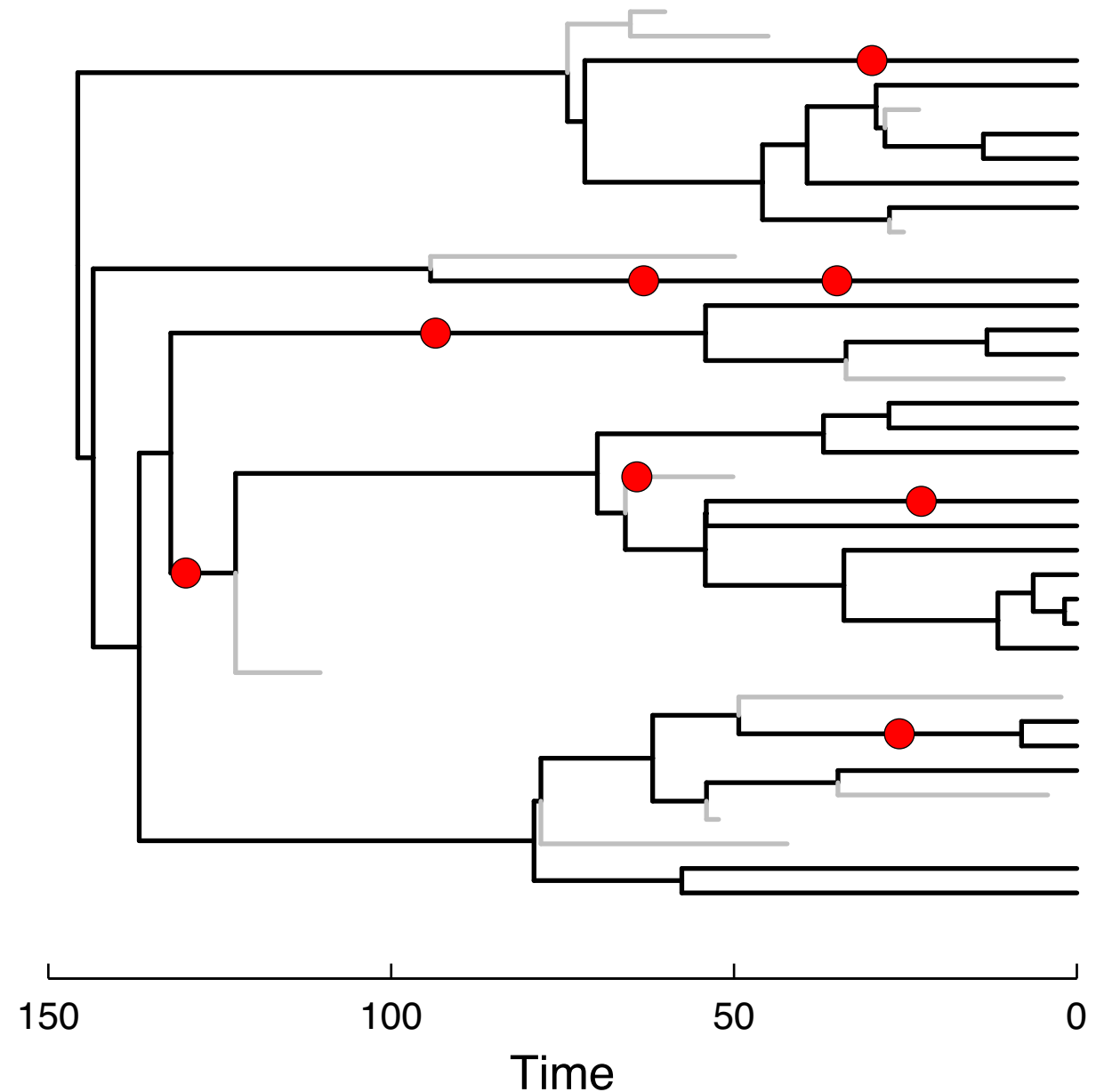
The Fossilized Birth-Death Process (FBD)

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

λ = speciation

μ = extinction

ψ = fossilization/recovery



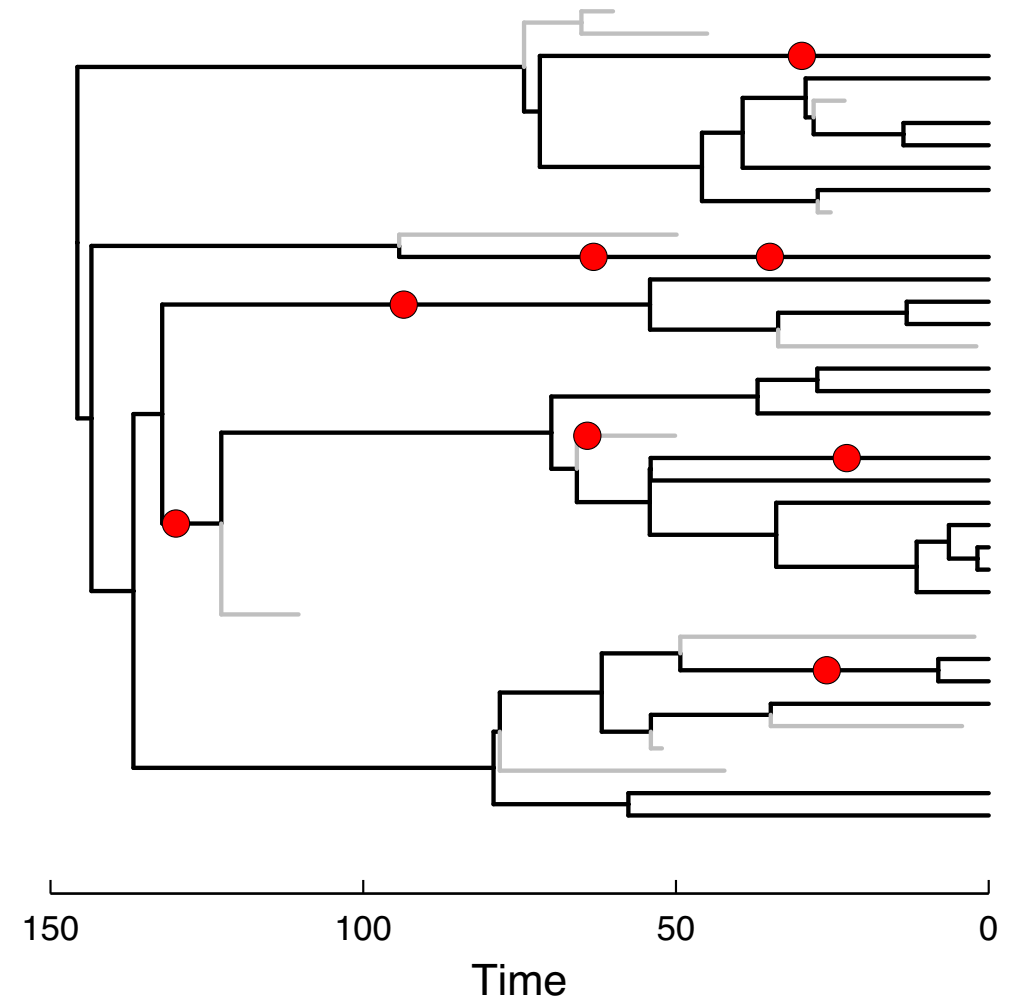
Sampled Ancestors

Sampled lineages with sampled decedents

Paleobiology, 22(2), 1996, pp. 141–151

On the probability of ancestors in the fossil record

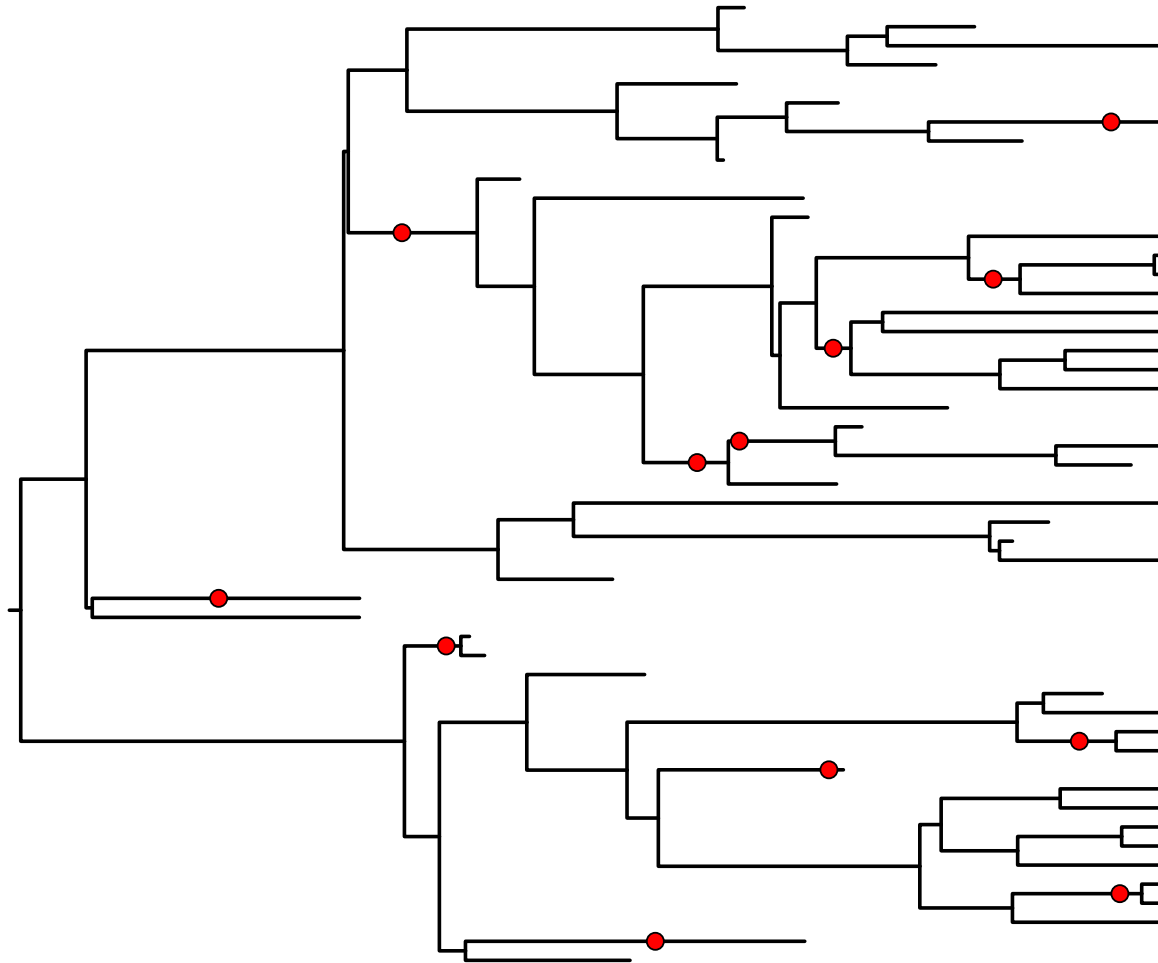
Mike Foote



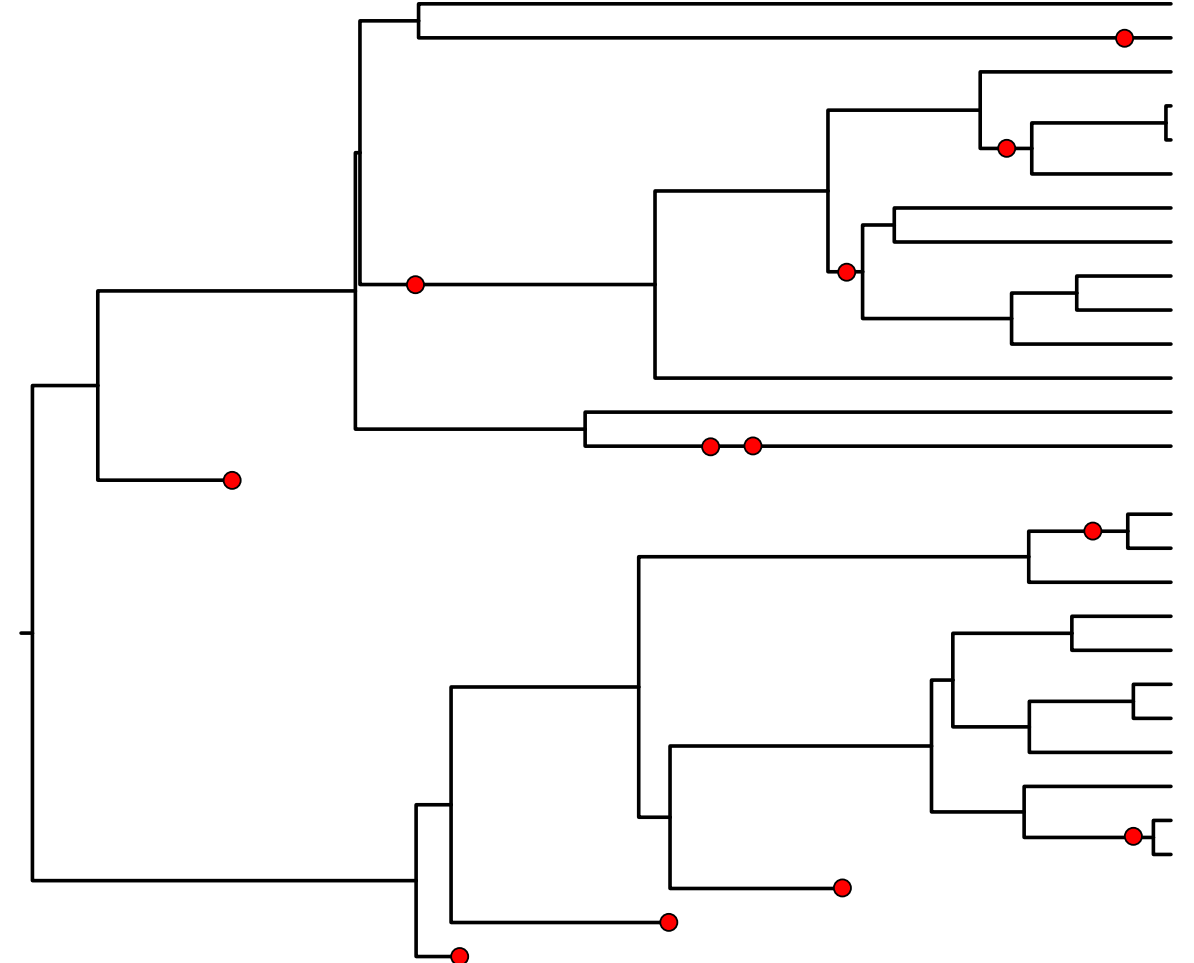
There is a non-zero probability of sampling ancestor-defendant relationships from the fossil record

Sampled Ancestors

COMPLETE FBD TREE



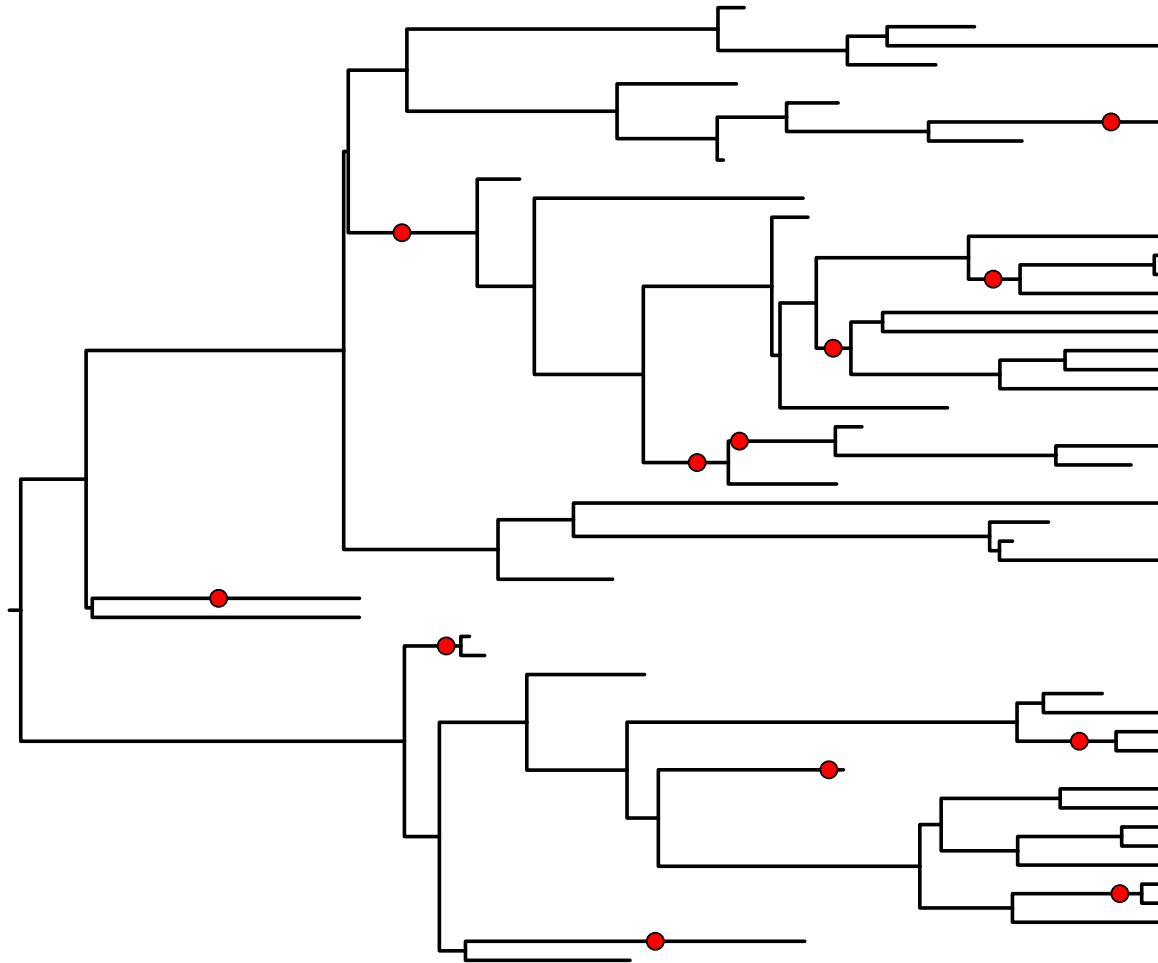
RECONSTRUCTED FBD TREE



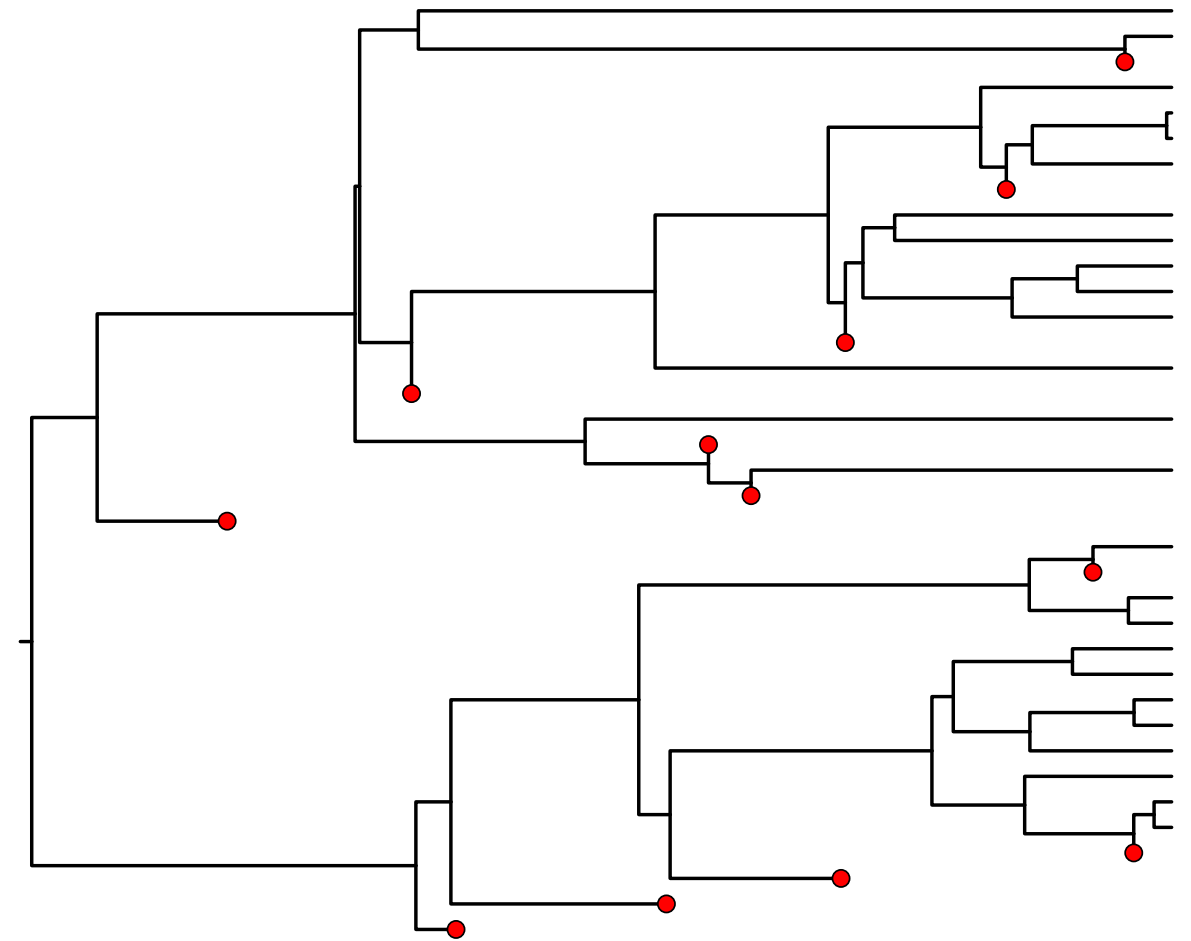
Because fossil & living taxa are assumed to come from a single diversification process, there is a non-zero probability of sampled ancestors

Sampled Ancestors

COMPLETE FBD TREE

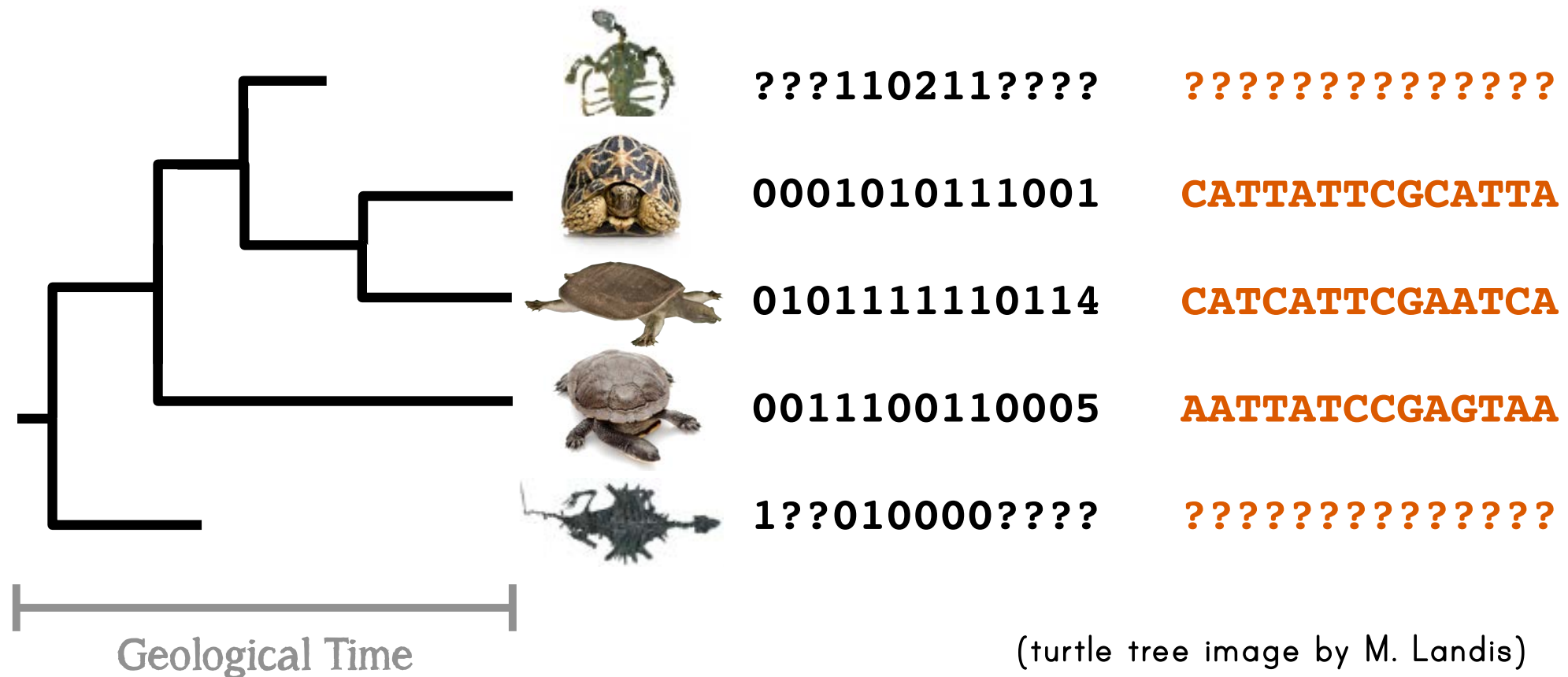
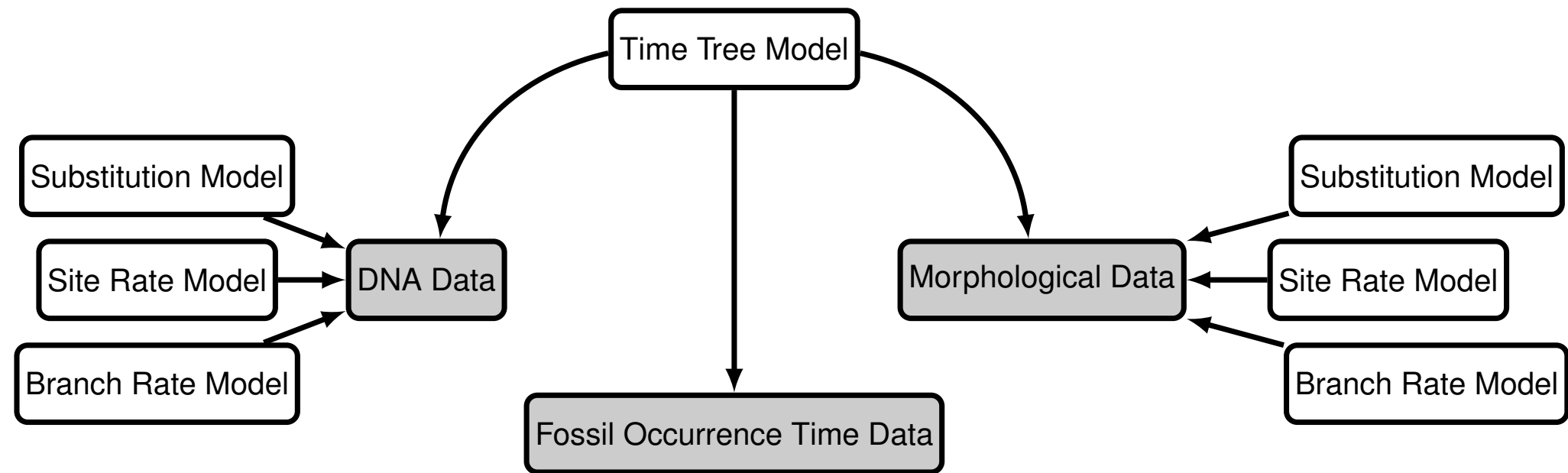


NO SAMPLED-ANCESTOR TREE



If all fossils are forced to be on separate lineages, this induces additional speciation events and will, in turn, influence rate & node-age estimates

Combining Fossil & Extant Data



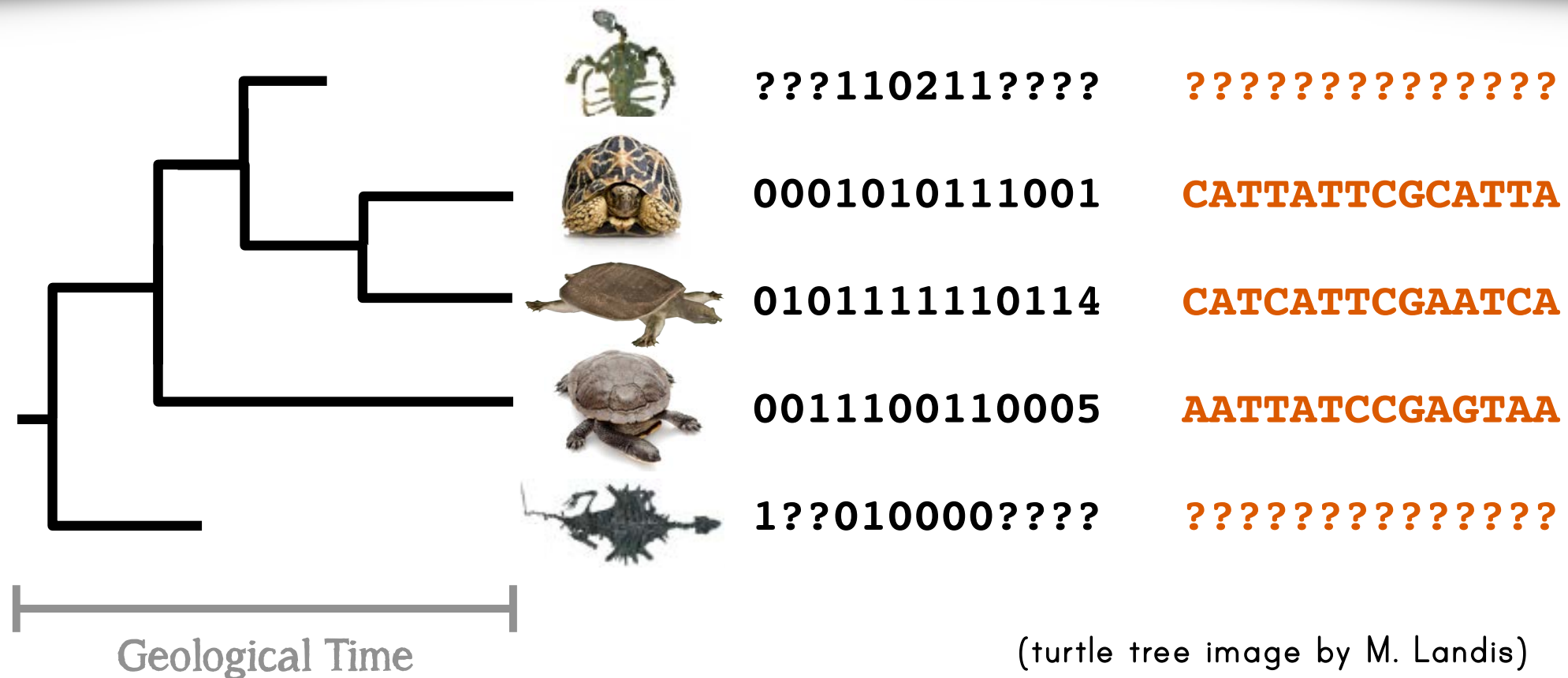
Modeling Morphological Character Change

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**

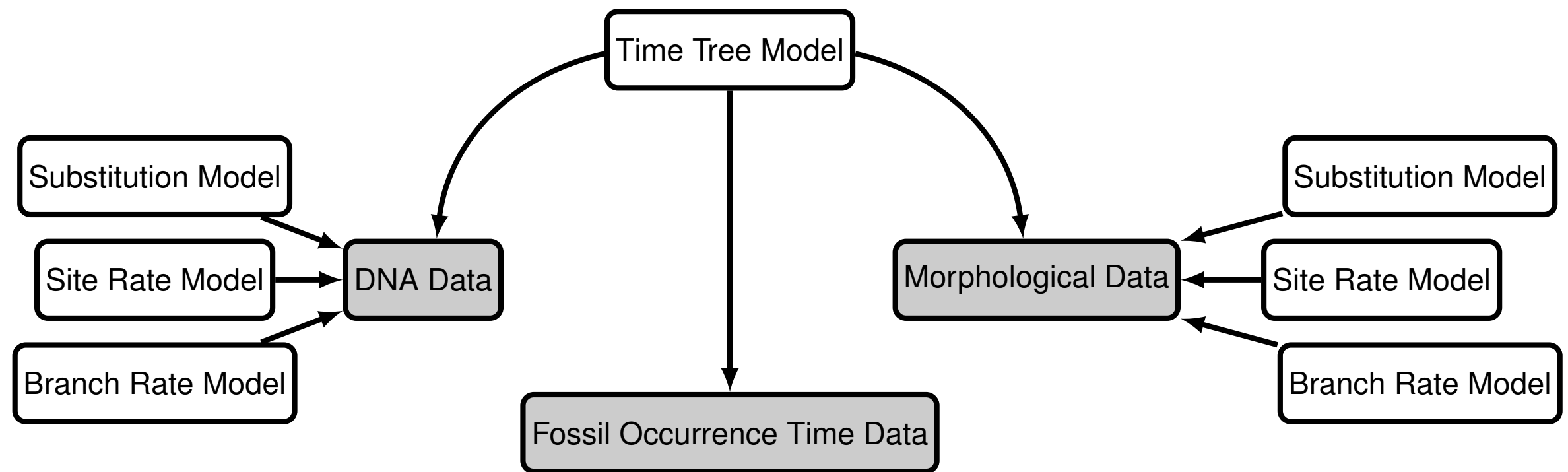
T6 **1**

T7 **1**

(Lewis. *Systematic Biology*. 2001)

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

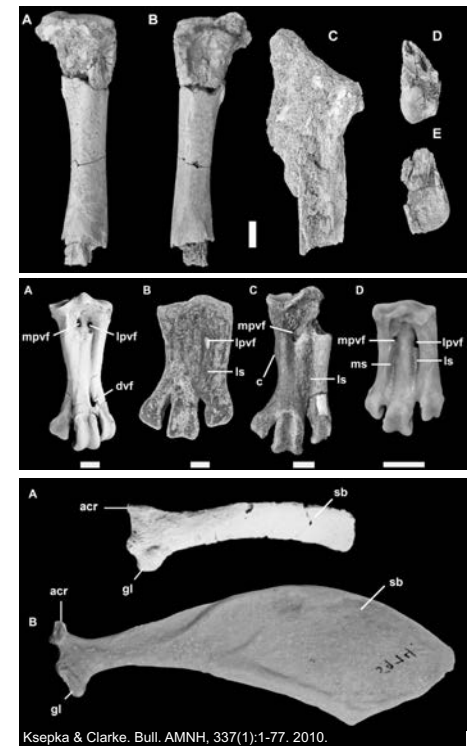
Does our understanding of penguin evolution improve when we consider both extant and fossil taxa?



"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/>)

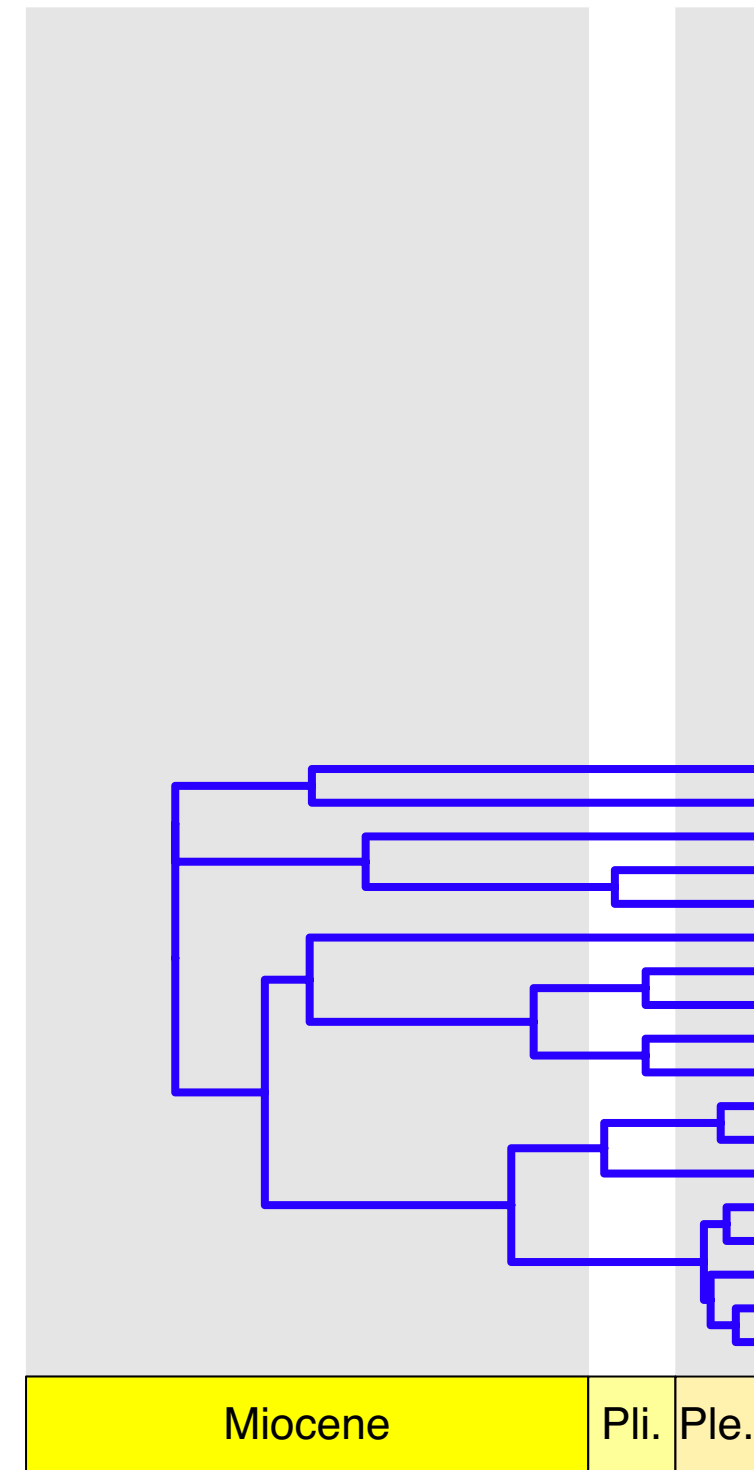
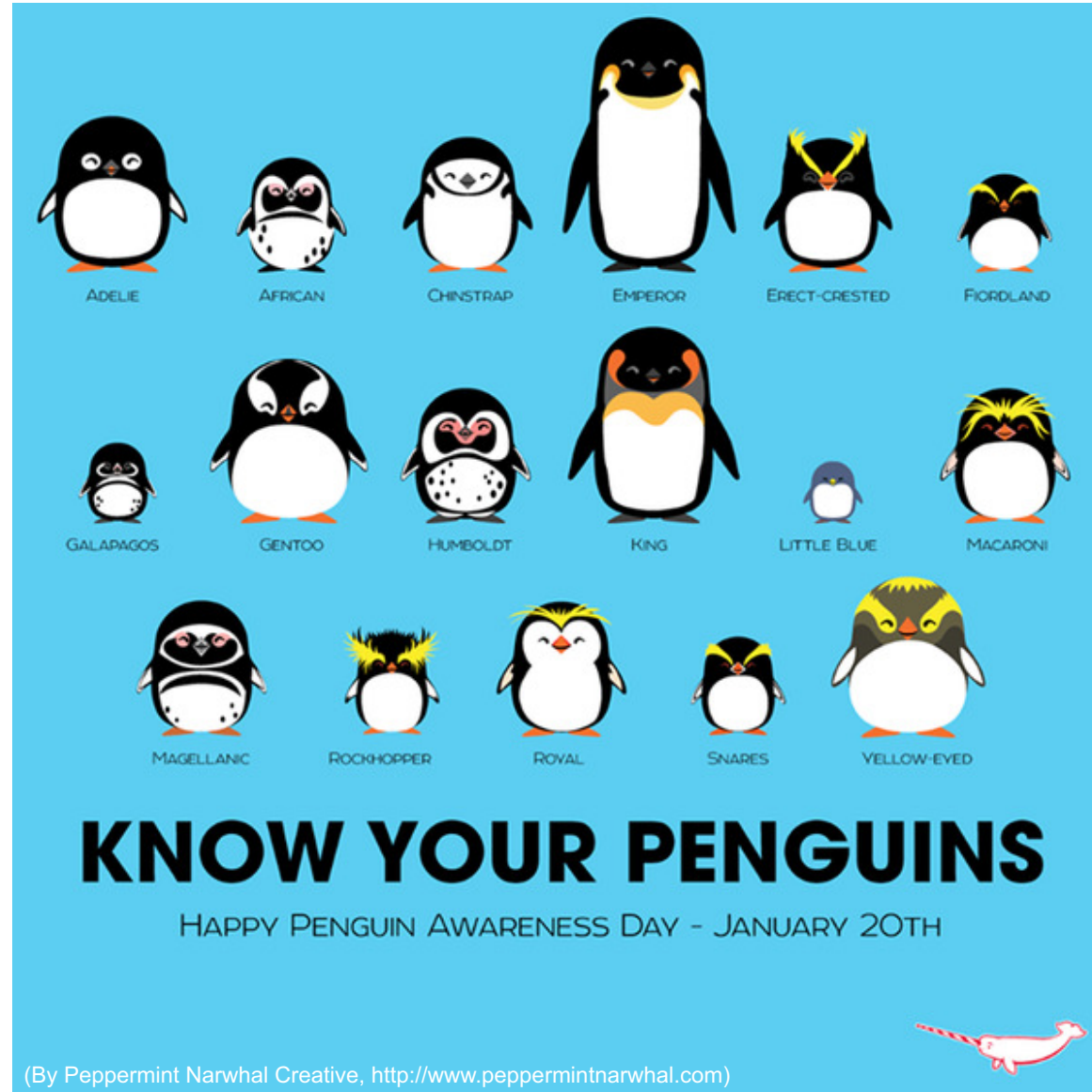


Ksepka & Clarke. Bull. AMNH, 337(1):1-77. 2010.

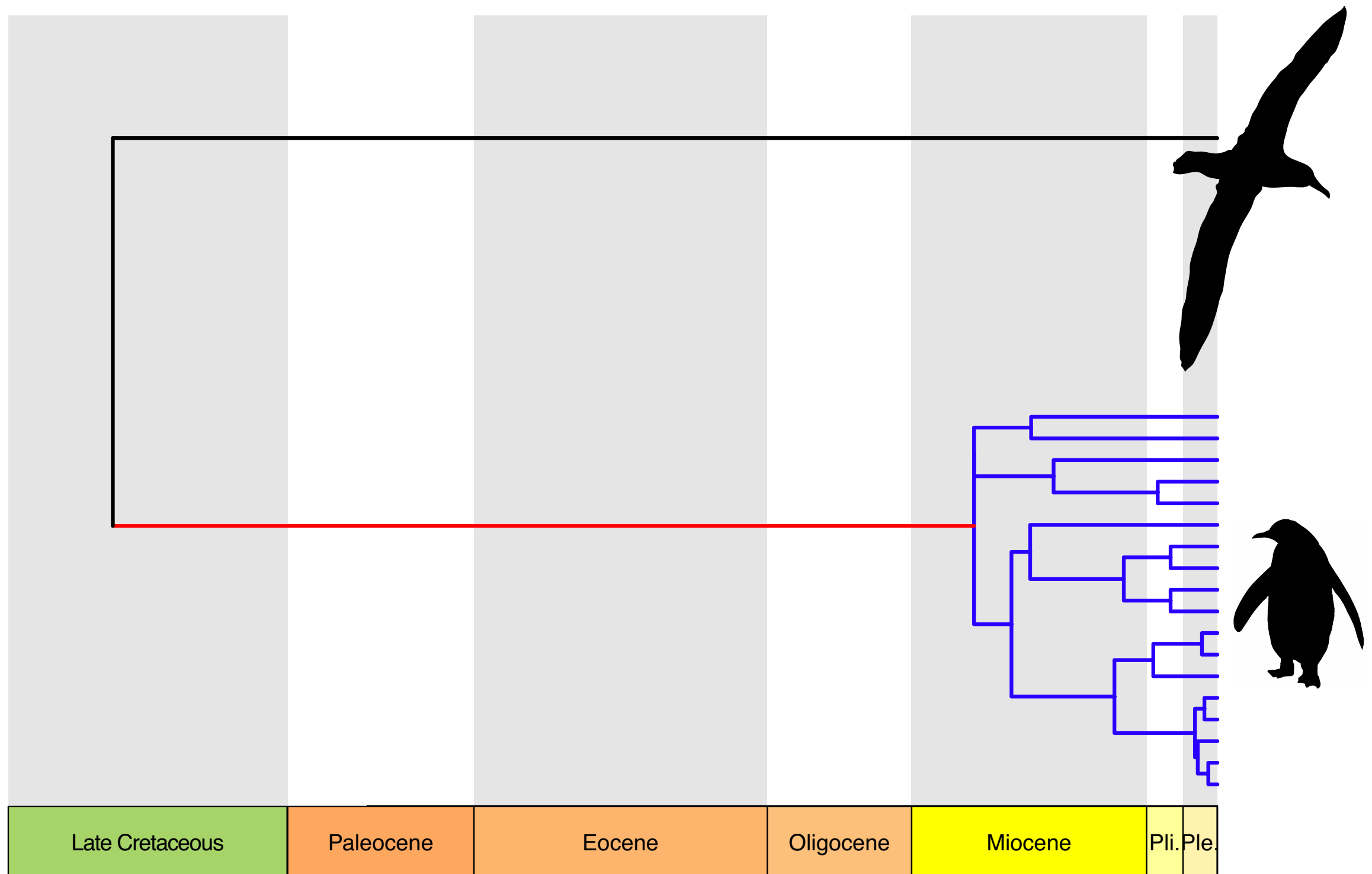


Artistic reconstructions by: Stephanie Abramowicz for Scientific American
Fordyce, R.E. and D.T. Ksepka. The Strangest Bird Scientific American 307, 56 – 61 (2012)

Penguin Diversity

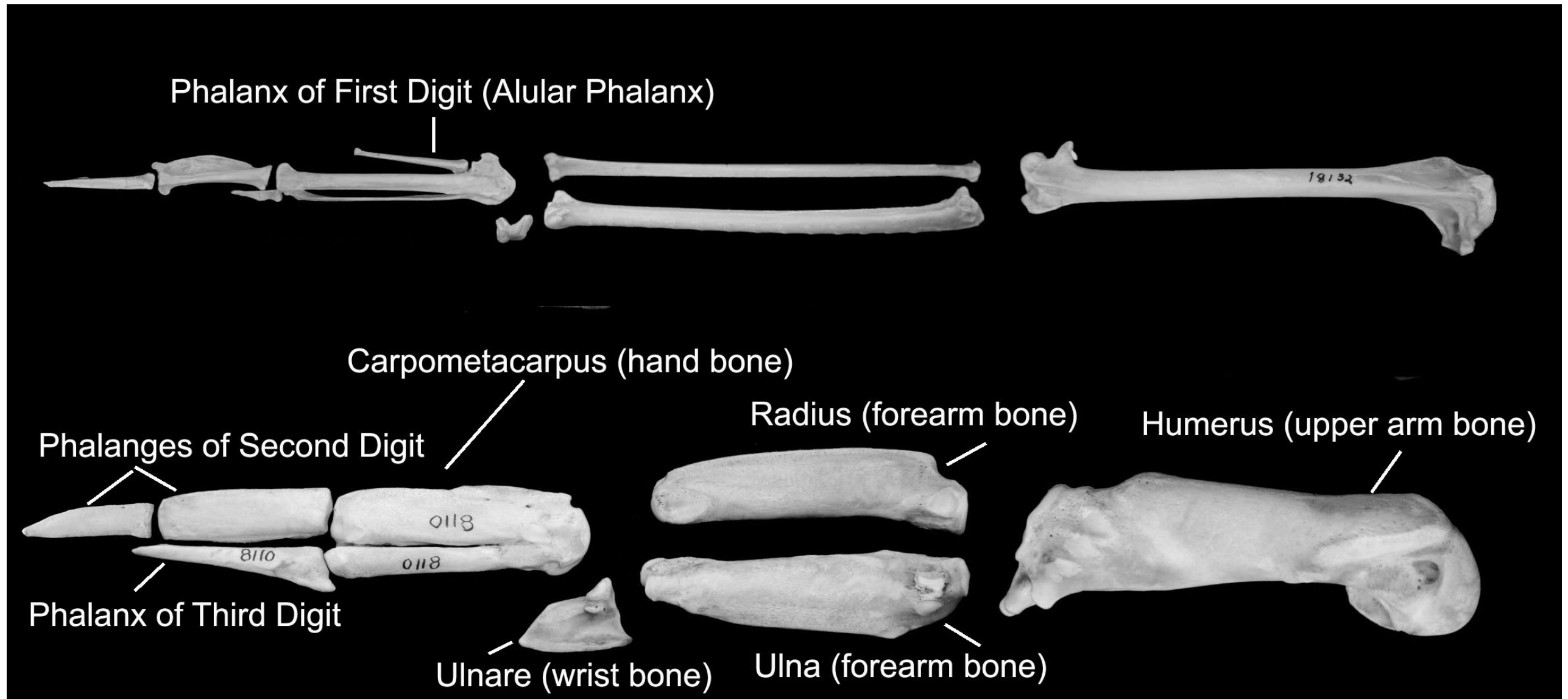


Penguin Diversity



What Makes a Penguin a Penguin?

Flattened, solid wing bones

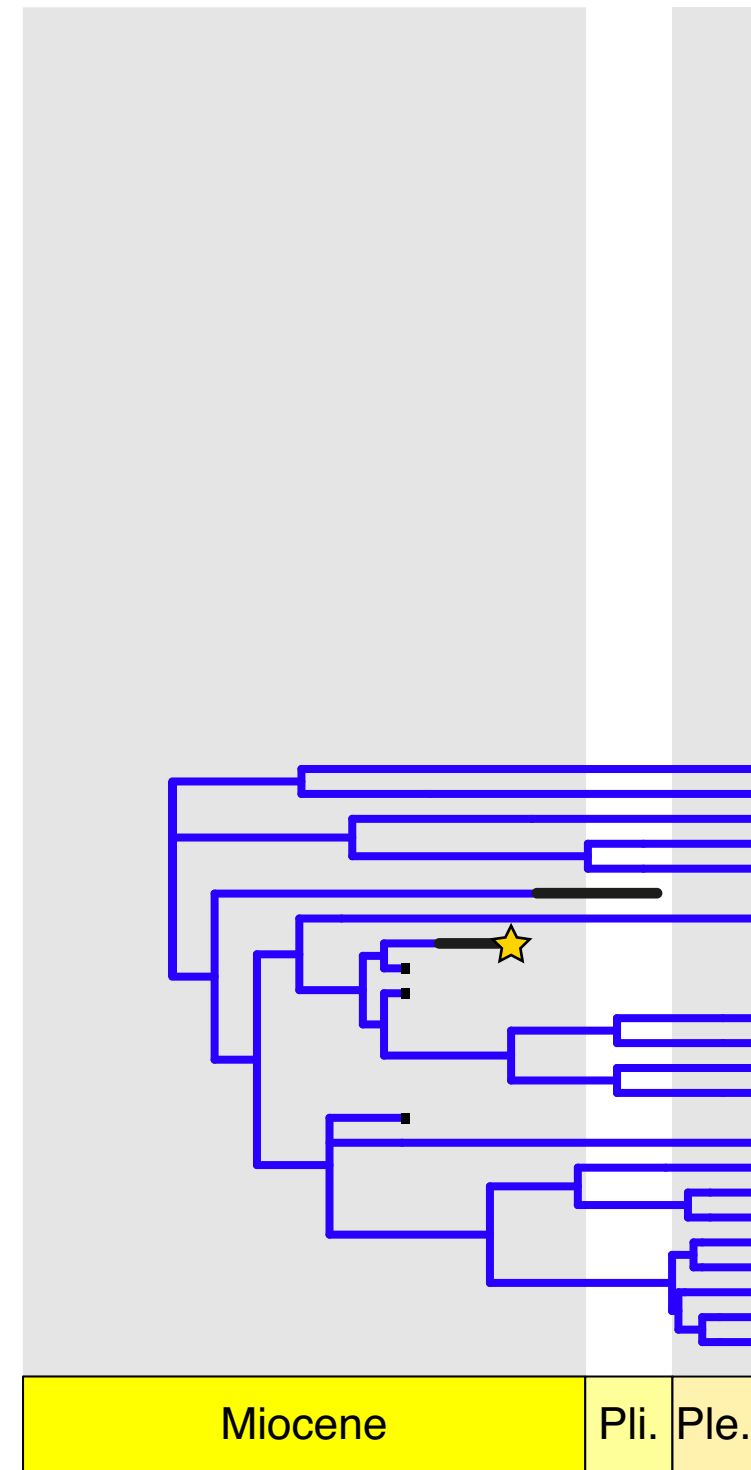


(image source: <https://fossilpenguins.wordpress.com>)

Fossil Penguin Diversity



(*S. urbinai* holotype fossil, 5-7 MYA, image by Martin Chávez)



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

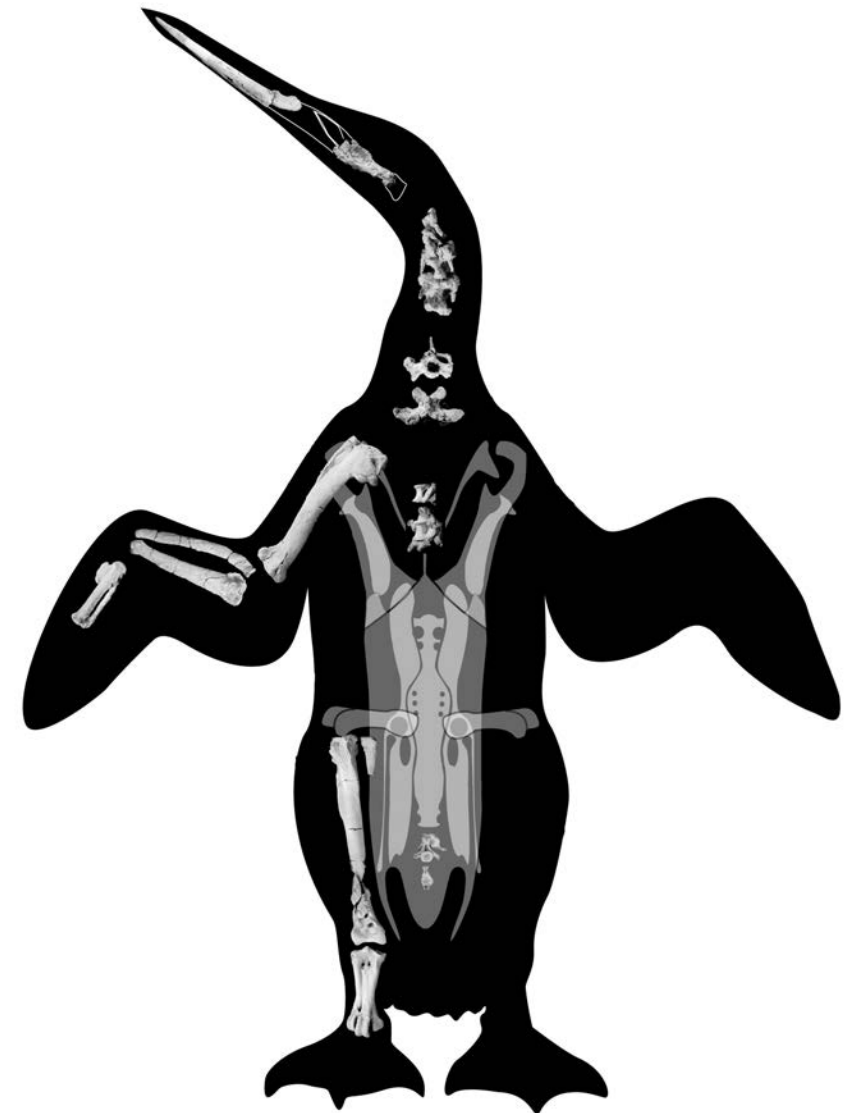


(Ksepka, Fordyce, Ando, & Jones, *J. Vert. Paleo.* 2012)

Penguins in the Paleocene

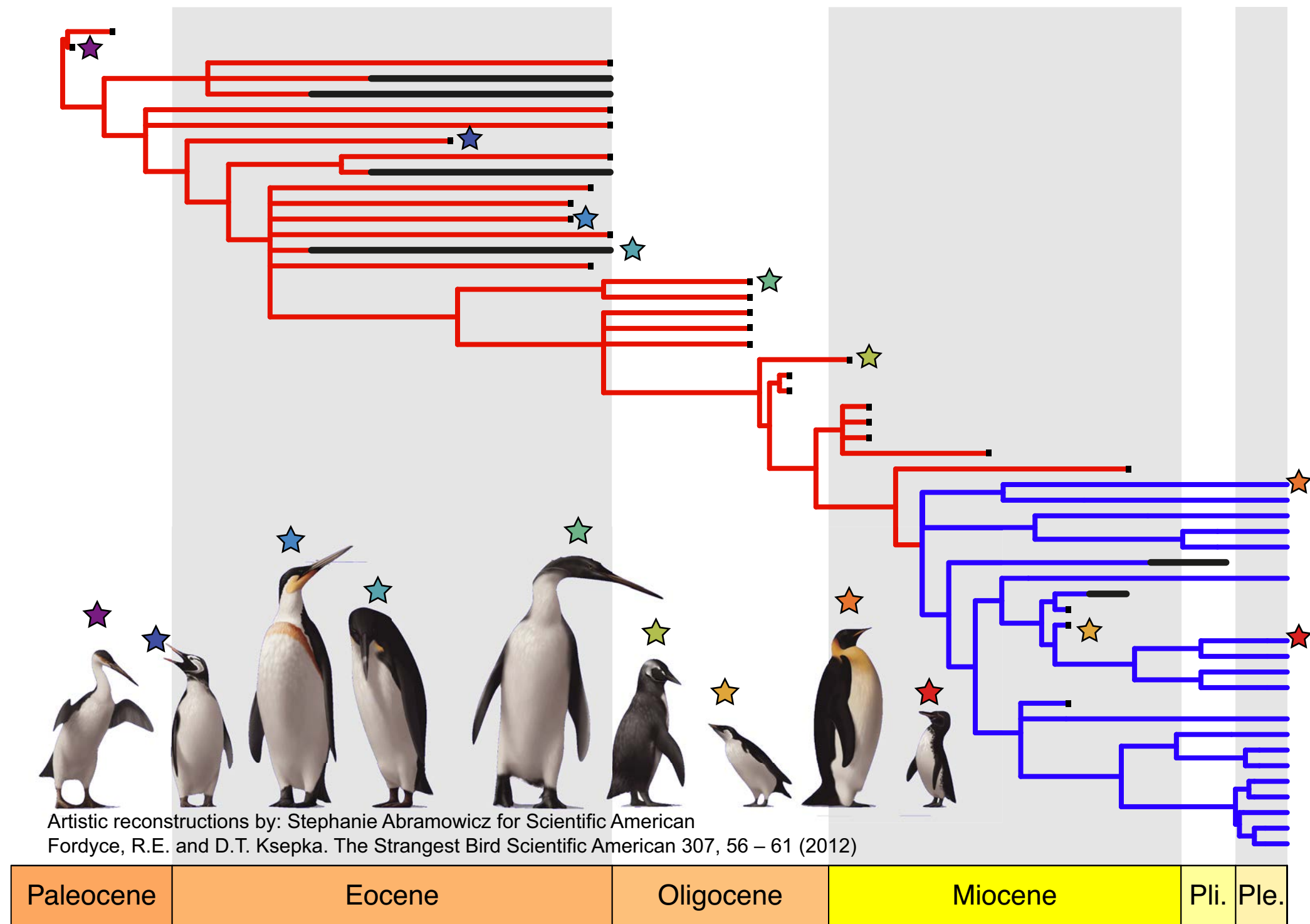
Waimanu

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

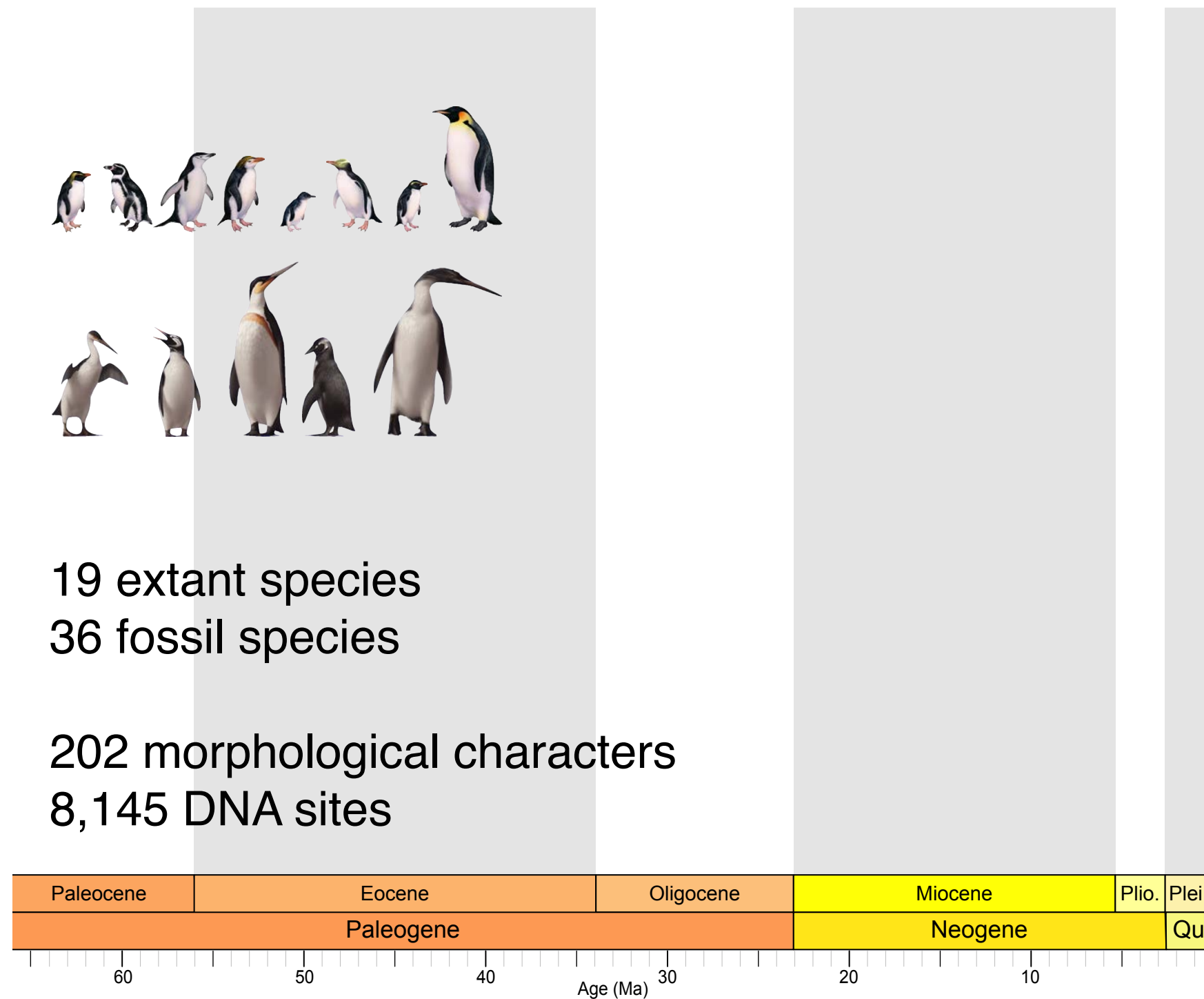


(Slack et al. *Mol. Biol. Evol.* 2006)

Penguin Diversity in Deep Time



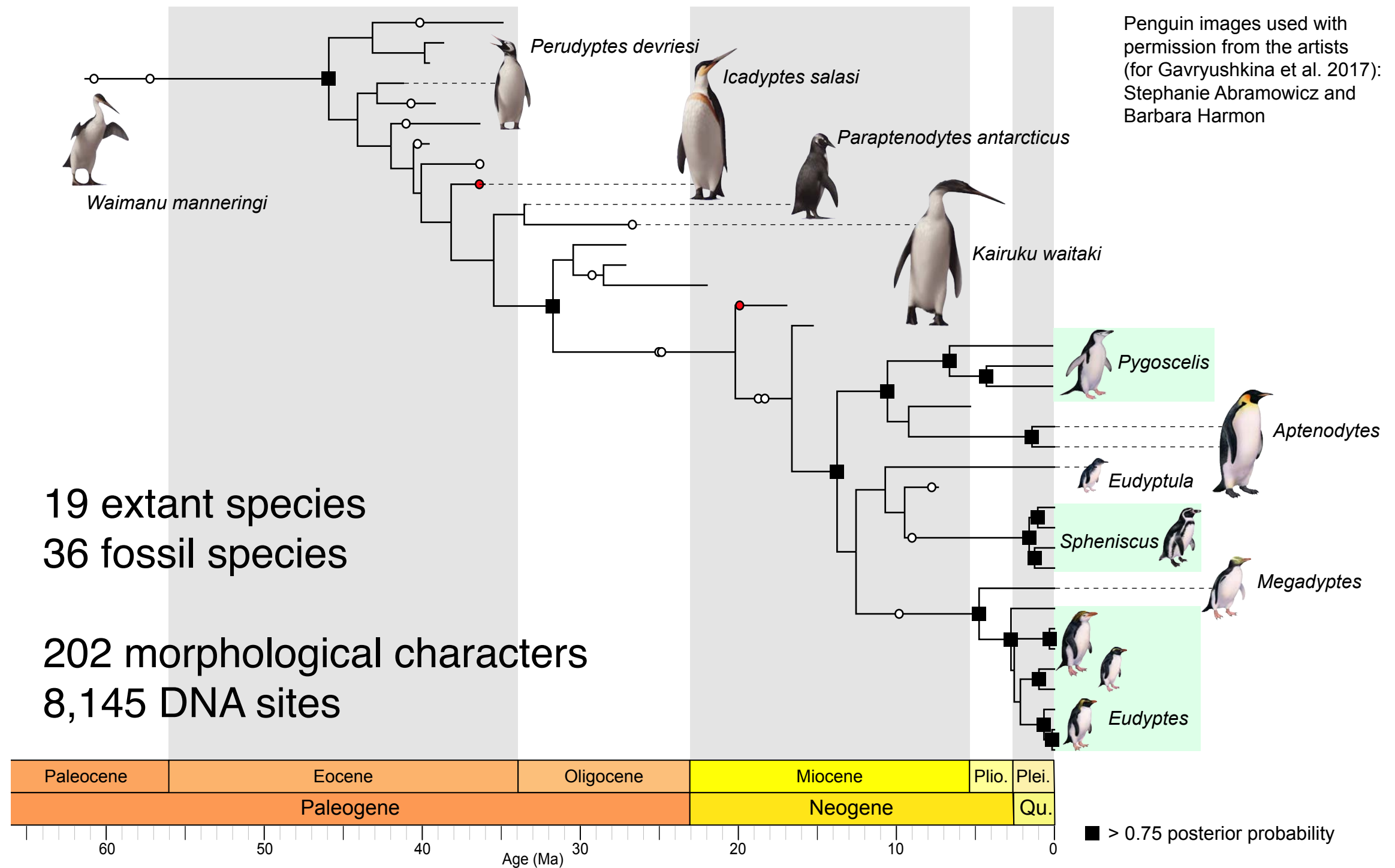
Penguin Diversity in Deep Time



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

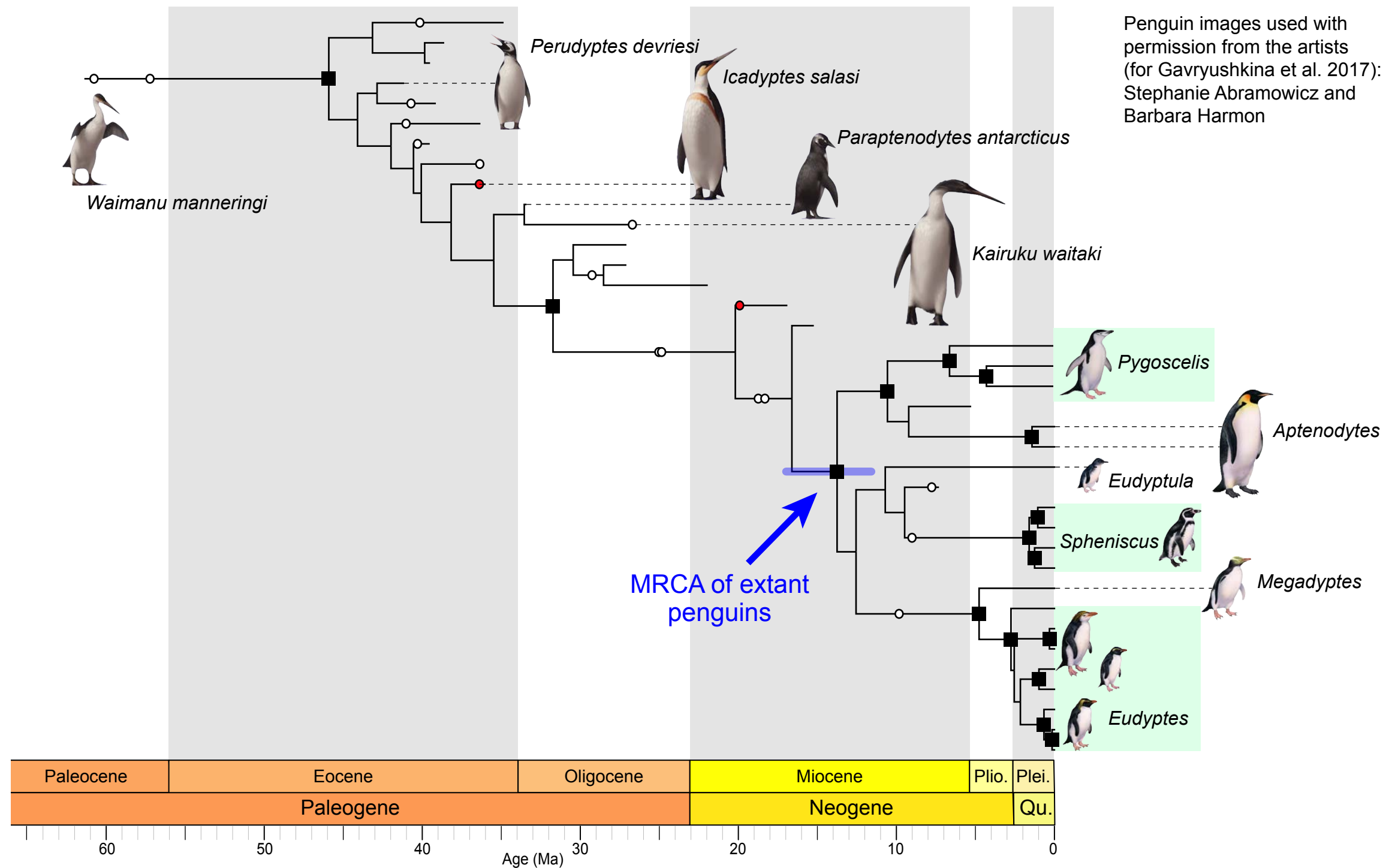
(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Penguin Diversity in Deep Time



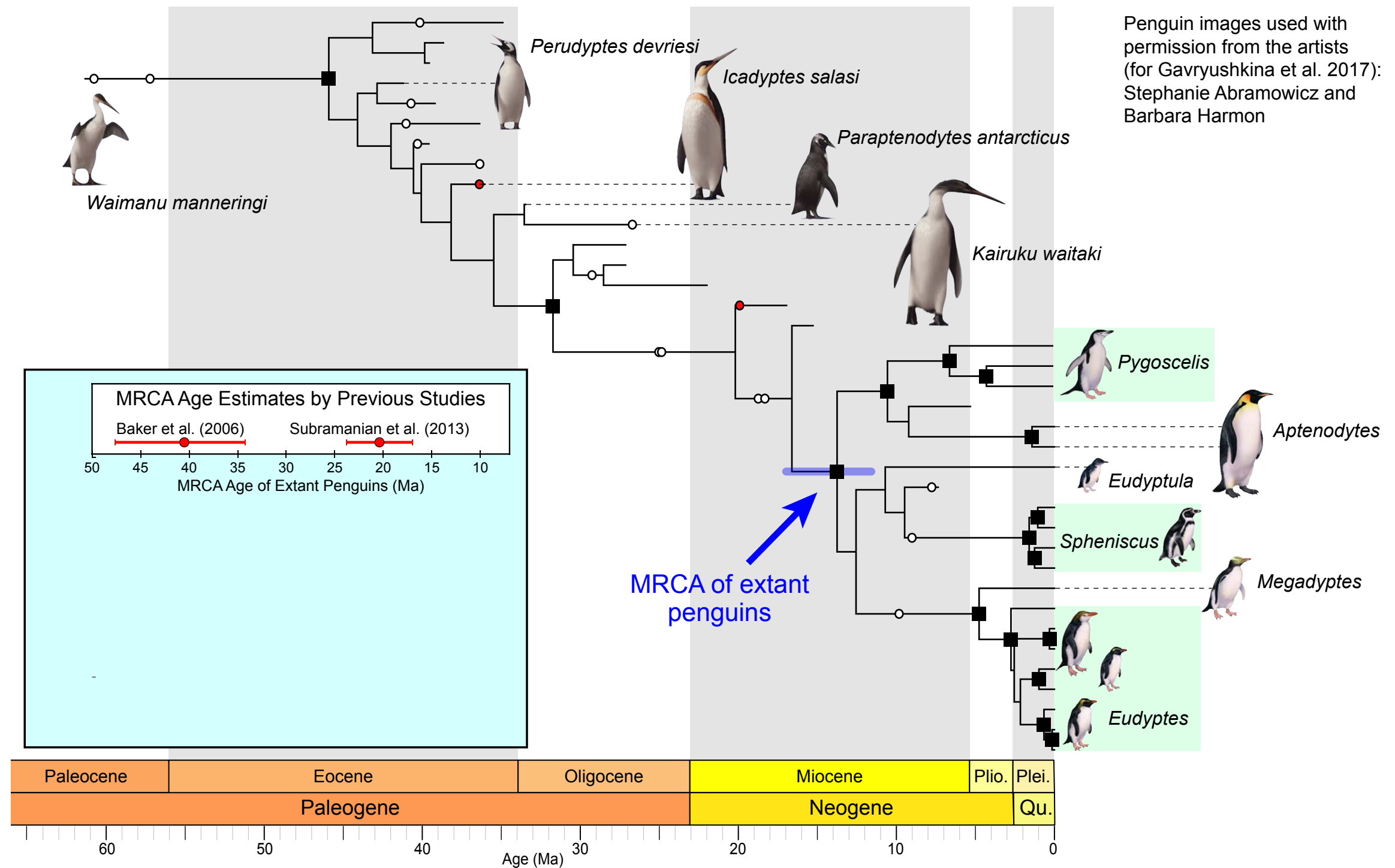
(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Penguin Diversity in Deep Time



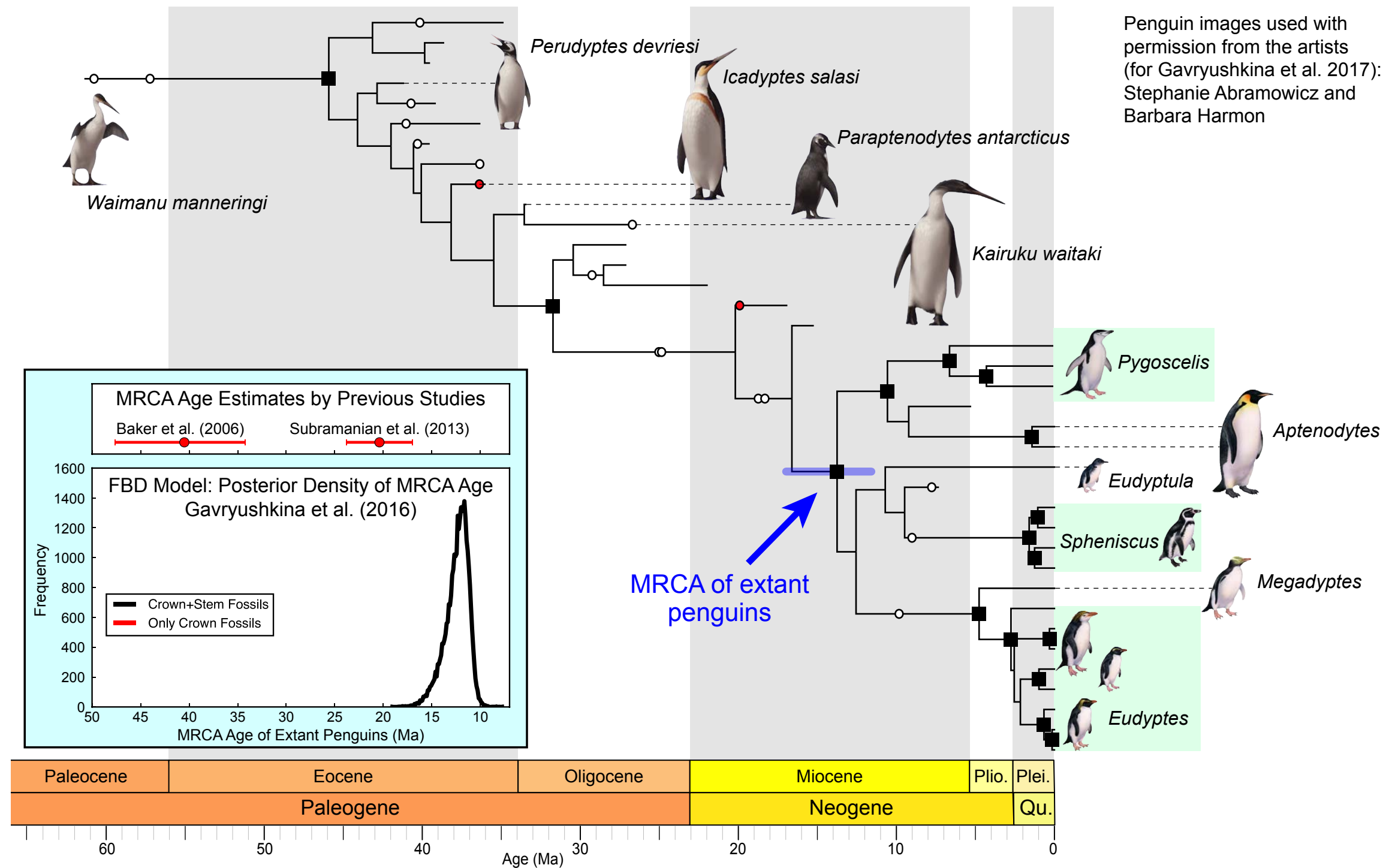
(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Penguin Diversity in Deep Time



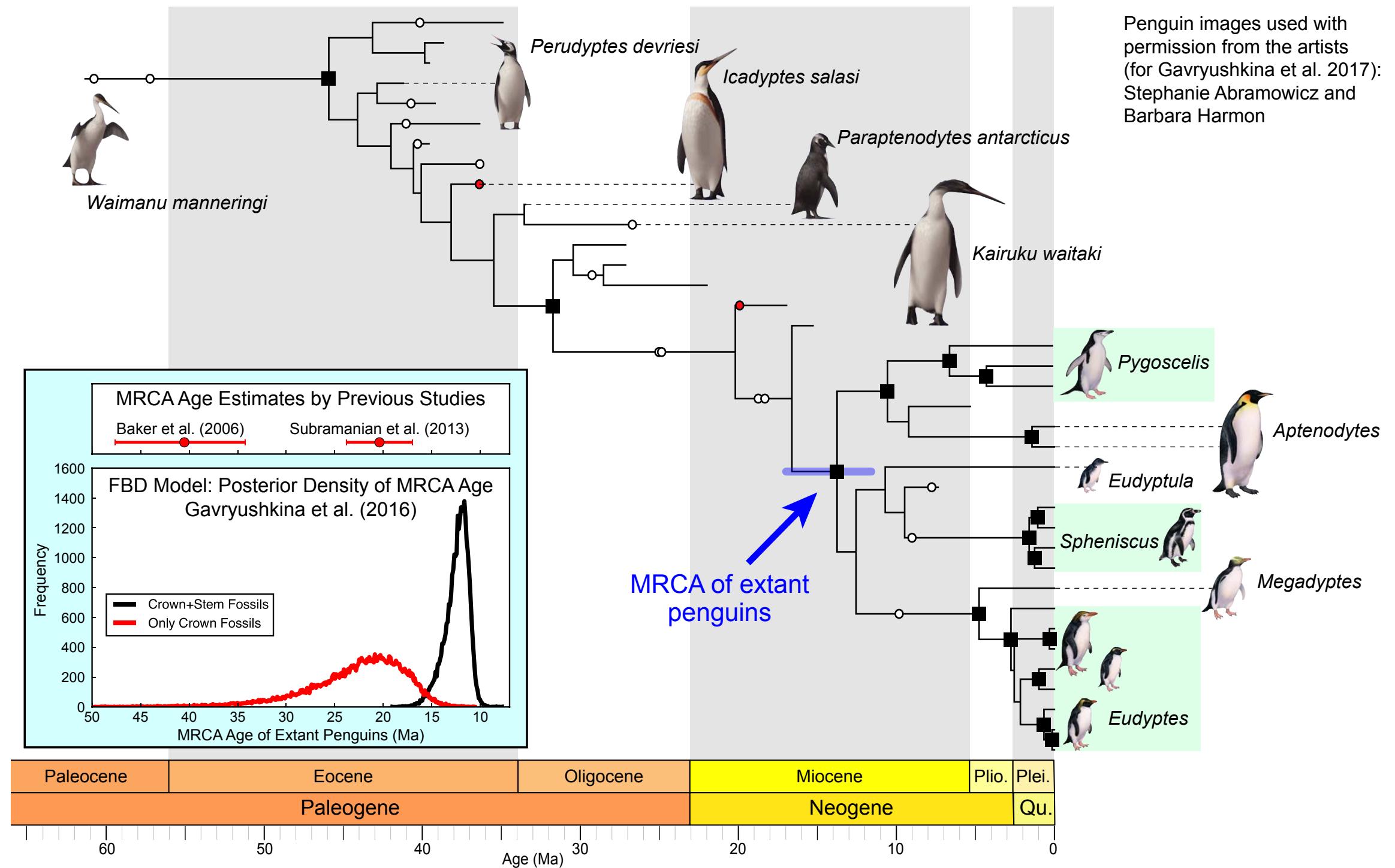
(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Penguin Diversity in Deep Time



(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Penguin Diversity in Deep Time



(Gavryushkina, Heath, Ksepka, Welch, Stadler, Drummond. 2017. *Syst. Biol.*)

Tutorial

[Download](#)[Tutorials](#)[Documentation](#)[RevScripter](#)[Workshops](#)[Jobs](#)[Developer](#)

Combined-Evidence Analysis and the Fossilized Birth-Death Process for Analysis of Fossil and Extant Specimens

Joint inference of divergence times and phylogenetic relationships of fossil and extant taxa

Tracy A. Heath, April M. Wright, and Walker Pett

Last modified on August 3, 2019

https://revbayes.github.io/tutorials/fbd/fbd_specimen.html