

BAYESIAN DIVERGENCE TIME ESTIMATION

Tracy Heath

Ecology, Evolution, & Organismal Biology
Iowa State University

@tracy7
<http://phyloworks.org>

2016 Workshop on Molecular Evolution
Woods Hole, MA USA

OUTLINE

Overview of divergence time estimation

- Relaxed clock models – accounting for variation in substitution rates among lineages
- Tree models – lineage diversification and sampling

break

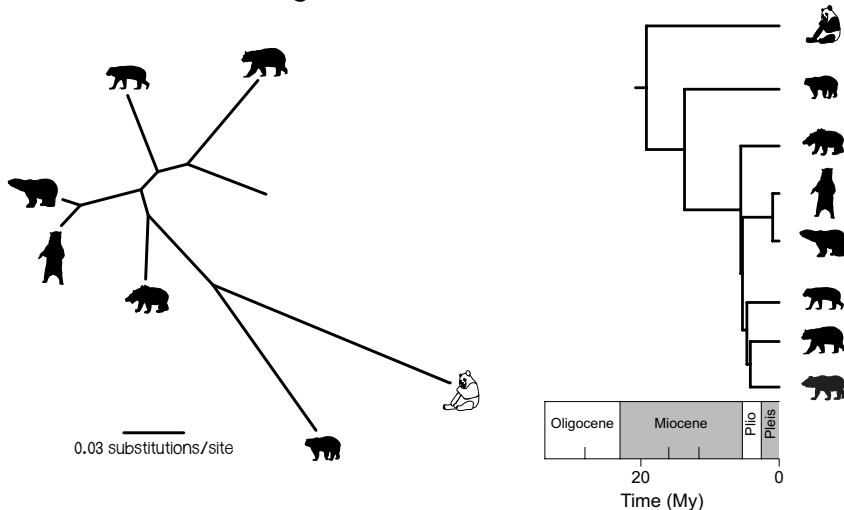
BEAST v2 Tutorial — Divergence-time estimation under birth-death processes

- Dating Bear Divergence Times with the Fossilized Birth-Death Process

dinner

A TIME-SCALE FOR EVOLUTION

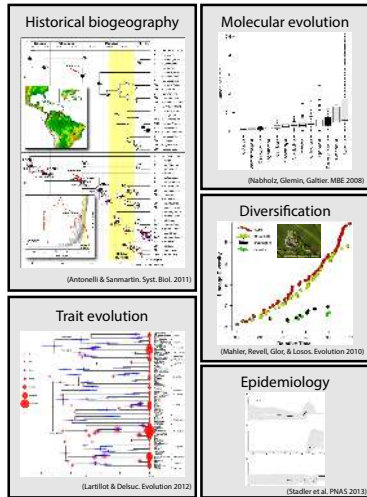
Phylogenies with branch lengths proportional to time provide more information about evolutionary history than unrooted trees with branch lengths in units of substitutions/site.



A TIME-SCALE FOR EVOLUTION

Phylogenetic divergence-time estimation

- What was the spacial and climatic environment of ancient angiosperms?
- How has mammalian body-size changed over time?
- How has the infection rate of HCV in Egypt changed over time?
- Is diversification in Caribbean anoles correlated with ecological opportunity?
- How has the rate of molecular evolution changed across the Tree of Life?



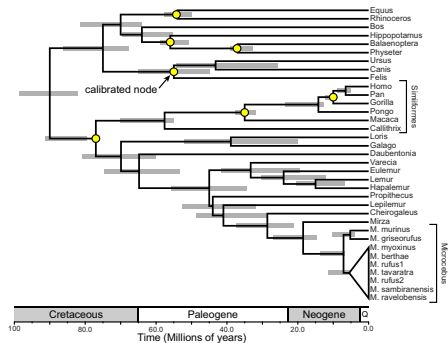
DIVERGENCE TIME ESTIMATION

Goal: Estimate the ages of interior nodes to understand the timing and rates of evolutionary processes

Model how rates are distributed across the tree

Describe the distribution of speciation events over time

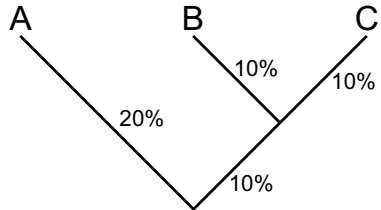
External calibration information for estimates of absolute node times



THE GLOBAL MOLECULAR CLOCK

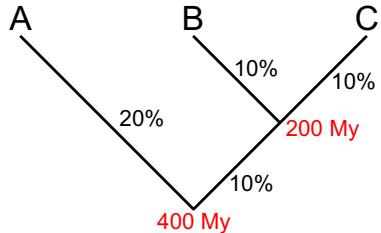
Assume that the rate of evolutionary change is constant over time

(branch lengths equal percent sequence divergence)



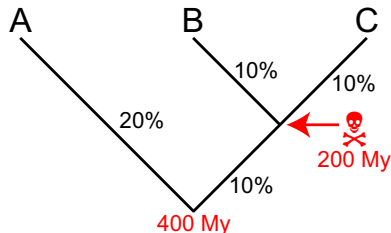
THE GLOBAL MOLECULAR CLOCK

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:

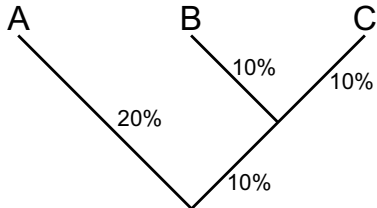
Variation in

- metabolic rate
- generation time
- DNA repair

Fixation rate:

Variation in

- strength and targets of selection
- population sizes

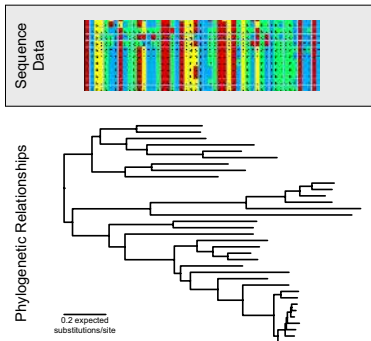


UNCONSTRAINED ANALYSIS

Sequence data provide information about **branch lengths**

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

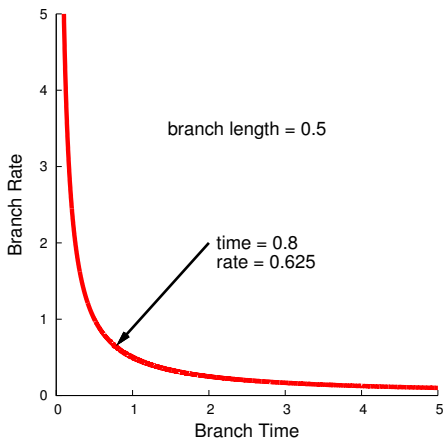
branch length = rate $\times$ time



RATE AND TIME

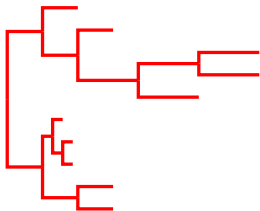
The sequence data provide information about branch length

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

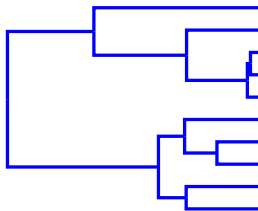


Methods for dating species divergences estimate the **substitution rate** and **time** separately

BAYESIAN DIVERGENCE TIME ESTIMATION



length = rate



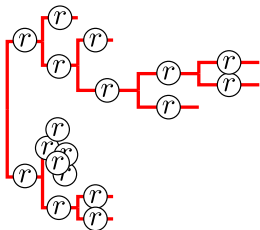
length = time

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

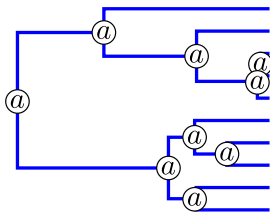
$$\mathcal{A} = (a_1, a_2, a_3, \dots, a_{N-1})$$

$$N = \text{number of tips}$$

BAYESIAN DIVERGENCE TIME ESTIMATION



length = rate



length = time

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

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

$$N = \text{number of tips}$$

BAYESIAN DIVERGENCE TIME ESTIMATION

Posterior probability

$$f(\mathcal{R}, \mathcal{A}, \theta_{\mathcal{R}}, \theta_{\mathcal{A}}, \theta_s \mid D, \Psi)$$

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

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

$\theta_{\mathcal{R}}, \theta_{\mathcal{A}}, \theta_s$ Model parameters

D Sequence data

Ψ Tree topology

BAYESIAN DIVERGENCE TIME ESTIMATION

$$f(\mathcal{R}, \mathcal{A}, \theta_{\mathcal{R}}, \theta_{\mathcal{A}}, \theta_s | D) =$$

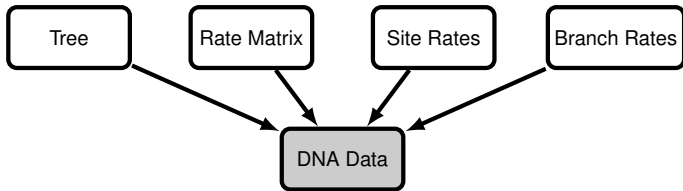
$$\frac{f(D | \mathcal{R}, \mathcal{A}, \theta_s) f(\mathcal{R} | \theta_{\mathcal{R}}) f(\mathcal{A} | \theta_{\mathcal{A}}) f(\theta_s)}{f(D)}$$

$f(D \mathcal{R}, \mathcal{A}, \theta_{\mathcal{R}}, \theta_{\mathcal{A}}, \theta_s)$	Likelihood
$f(\mathcal{R} \theta_{\mathcal{R}})$	Prior on rates
$f(\mathcal{A} \theta_{\mathcal{A}})$	Prior on node ages
$f(\theta_s)$	Prior on substitution parameters
$f(D)$	Marginal probability of the data

BAYESIAN DIVERGENCE TIME ESTIMATION

Estimating divergence times relies on 2 main elements:

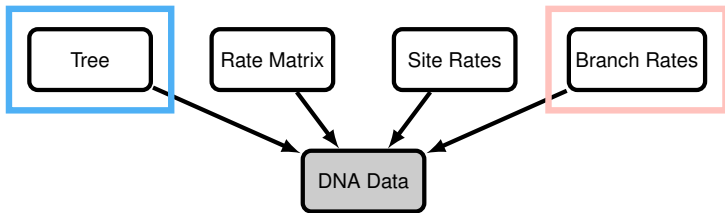
- Branch-specific rates: $f(\mathcal{R} | \theta_{\mathcal{R}})$
- Node ages & Topology: $f(\mathcal{A} | \theta_{\mathcal{A}})$



BAYESIAN DIVERGENCE TIME ESTIMATION

Estimating divergence times relies on 2 main elements:

- Branch-specific rates: $f(\mathcal{R} | \theta_{\mathcal{R}})$
- Node ages & Topology: $f(\mathcal{A} | \theta_{\mathcal{A}})$



MODELING RATE VARIATION

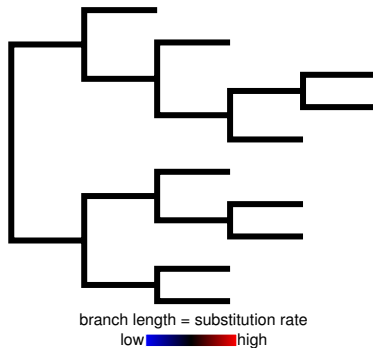
Some models describing lineage-specific substitution rate variation:

- **Global molecular clock** (Zuckerkandl & 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)
- **Log-normally distributed autocorrelated rates** (Thorne, Kishino & Painter 1998; Kishino, Thorne & Bruno 2001; Thorne & Kishino 2002)
- **Uncorrelated/independent rates models** (Drummond et al. 2006; Rannala & Yang 2007; Lepage et al. 2007)
- **Mixture models on branch rates** (Heath, Holder, Huelsenbeck 2012)

GLOBAL MOLECULAR CLOCK

The substitution rate is constant over time

All lineages share the same rate



RELAXED-CLOCK MODELS

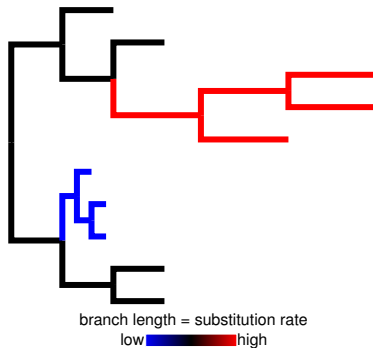
To accommodate variation in substitution rates
'relaxed-clock' models estimate lineage-specific substitution rates

- **Local molecular clocks**
- **Punctuated rate change model**
- **Log-normally distributed autocorrelated rates**
- **Uncorrelated/independent rates models**
- **Mixture models on branch rates**

LOCAL MOLECULAR CLOCKS

Rate shifts occur
infrequently over the tree

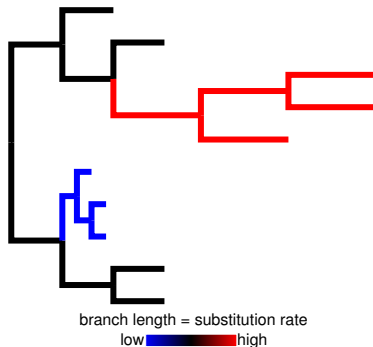
Closely related lineages
have equivalent rates
(clustered by sub-clades)



LOCAL MOLECULAR CLOCKS

Most methods for estimating local clocks required specifying the number and locations of rate changes *a priori*

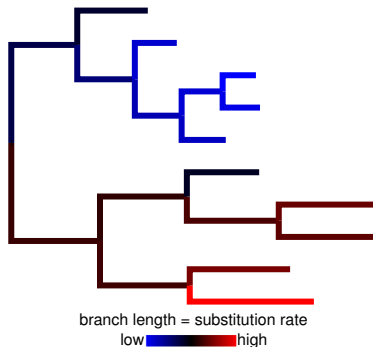
Drummond and Suchard (2010) introduced a Bayesian method that samples over a broad range of possible *random local clocks*



AUTOCORRELATED RATES

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

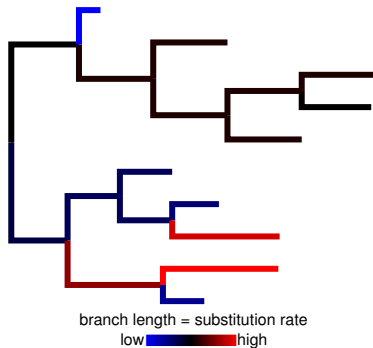
The rate at a node is drawn from a lognormal distribution with a mean equal to the parent rate (geometric brownian motion)



PUNCTUATED RATE CHANGE

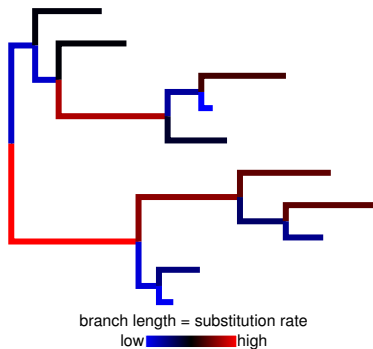
Rate changes occur along lineages according to a point process

At rate-change events, the new rate is a product of the parent's rate and a Γ -distributed multiplier



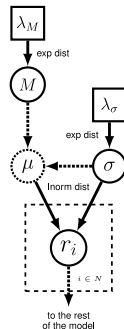
INDEPENDENT/UNCORRELATED RATES

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



INDEPENDENT/UNCORRELATED RATES

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

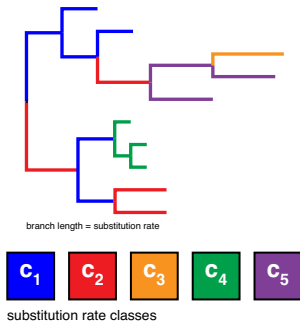


INFINITE MIXTURE MODEL

Dirichlet process prior:

Branches are partitioned into distinct rate categories

The number of rate categories and assignment of branches to categories are random variables under the model



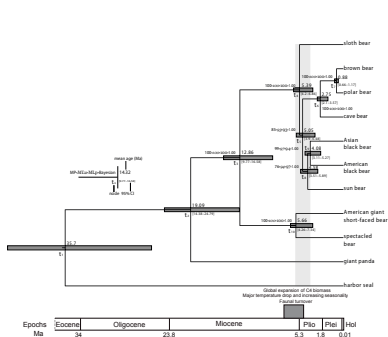
MODELING RATE VARIATION

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

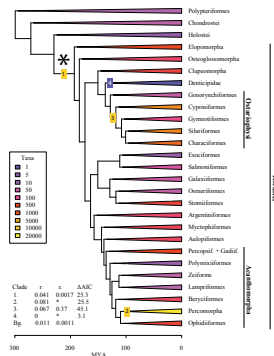
- **Global molecular clock**
- **Local molecular clocks**
- **Punctuated rate change model**
- **Log-normally distributed autocorrelated rates**
- **Uncorrelated/independent rates models**
- **Dirchlet process prior**

MODELING RATE VARIATION

Are our models appropriate across all data sets?



Krause et al., 2008. Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. BMC Evol. Biol. 8.



Santini et al., 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol. Biol. 9.

MODELING RATE VARIATION

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

- **Global molecular clock**
- **Local molecular clocks**
- **Punctuated rate change model**
- **Log-normally distributed autocorrelated rates**
- **Uncorrelated/independent rates models**
- **Dirichlet process prior**

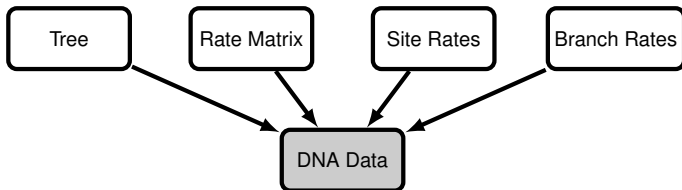
Considering model selection, uncertainty, & plausibility is **very** important for Bayesian divergence time analysis



BAYESIAN DIVERGENCE TIME ESTIMATION

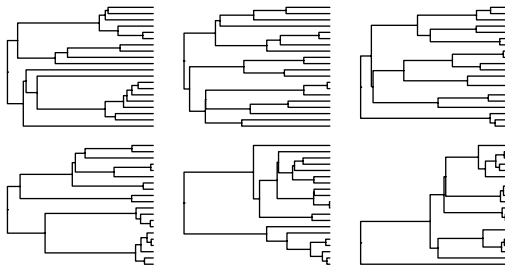
Estimating divergence times relies on 2 main elements:

- Branch-specific rates: $f(\mathcal{R} | \theta_{\mathcal{R}})$
- Node ages: $f(\mathcal{A} | \theta_{\mathcal{A}})$



PRIORS ON THE TREE AND NODE AGES

Relaxed clock Bayesian analyses require a prior distribution on time trees



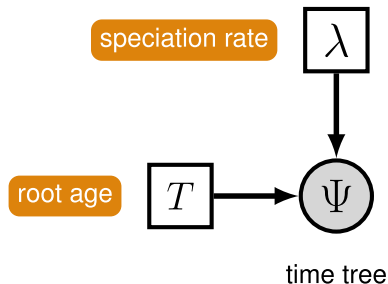
Different node-age priors make different assumptions about the timing of divergence events

STOCHASTIC BRANCHING PROCESSES

Node-age priors based on stochastic models of lineage diversification

Yule process: assumes a constant rate of speciation, across lineages

A pure birth process—every node leaves extant descendants (no extinction)

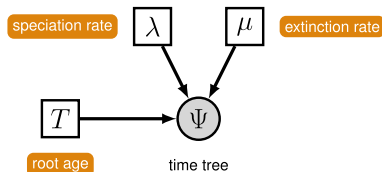


STOCHASTIC BRANCHING PROCESSES

Node-age priors based on stochastic models of lineage diversification

Constant-rate birth-death

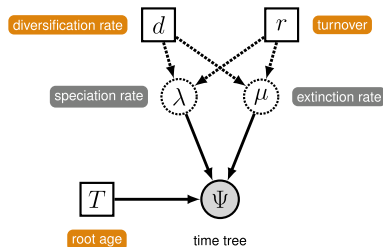
process: at any point in time a lineage can speciate at rate λ or go extinct with a rate of μ



STOCHASTIC BRANCHING PROCESSES

Node-age priors based on stochastic models of lineage diversification

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

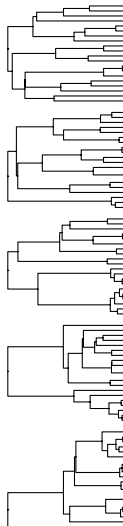


STOCHASTIC BRANCHING PROCESSES

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 (or d and r) accounts for uncertainty in these hyperparameters



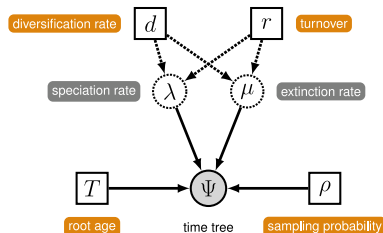
STOCHASTIC BRANCHING PROCESSES

Node-age priors based on stochastic models of lineage diversification

Birth-death-sampling

process: an extension of the constant-rate birth-death model that accounts for random sampling of tips

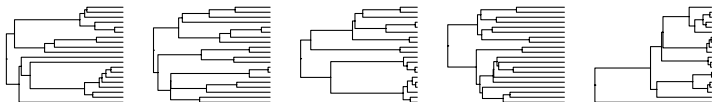
Conditions on a probability of sampling a tip, ρ



PRIORS ON NODE TIMES

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

Node-time priors cannot give precise estimates of *absolute* node ages



We need external information (like fossils) to provide absolute time scale

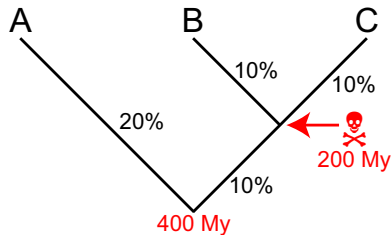


CALIBRATING DIVERGENCE TIMES

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

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

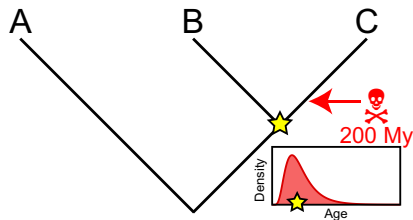
Uncertainty in the placement of fossils



CALIBRATION DENSITIES

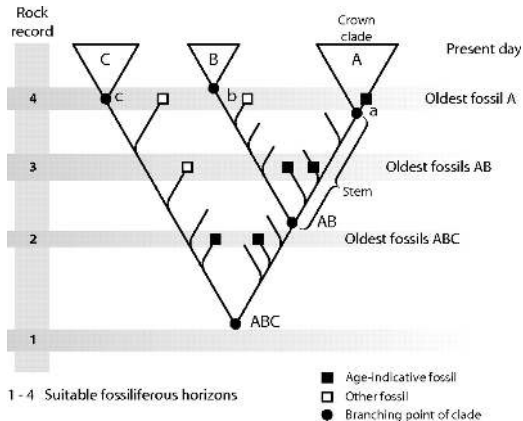
Bayesian inference is well suited to accommodating uncertainty in the age of the calibration node

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



ASSIGNING FOSSILS TO CLADES

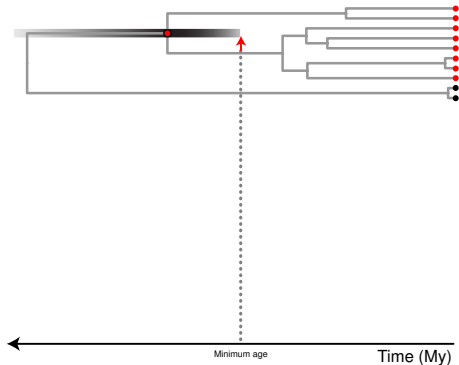
Misplaced fossils can affect node age estimates throughout the tree – if the fossil is older than its presumed MRCA



FOSSIL CALIBRATION

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

Reliable **maximum** bounds are typically unavailable

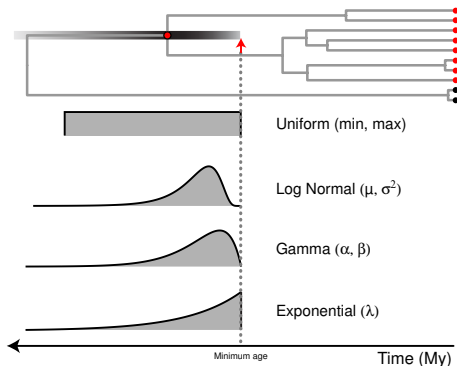


PRIOR DENSITIES ON CALIBRATED NODES

Common practice in Bayesian divergence-time estimation:

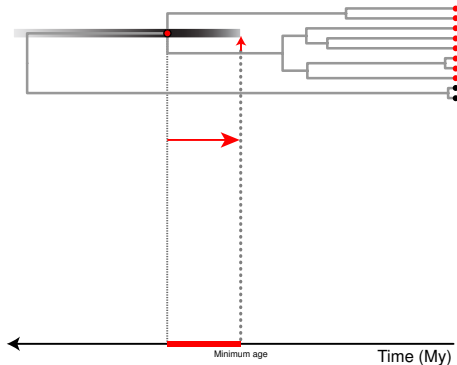
Parametric distributions are typically off-set by the age of the oldest fossil assigned to a clade

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



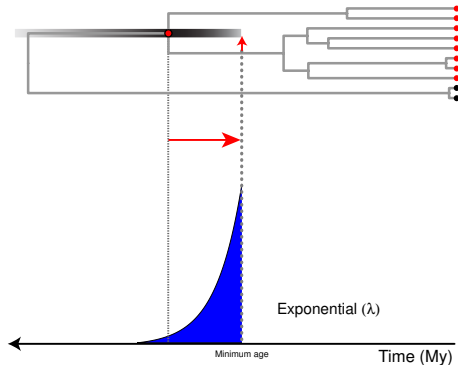
PRIOR DENSITIES ON CALIBRATED NODES

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



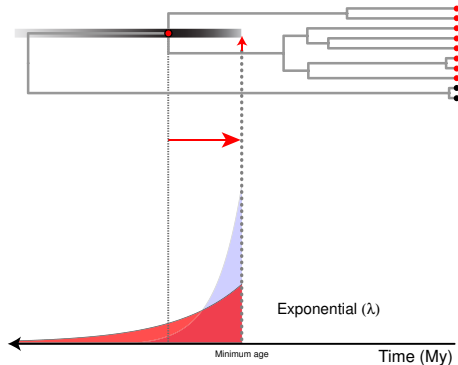
PRIOR DENSITIES ON CALIBRATED NODES

Overly **informative** priors can bias node age estimates to be too young



PRIOR DENSITIES ON CALIBRATED NODES

Uncertainty in the age of the MRCA of the clade relative to the age of the fossil may be better captured by **vague** prior densities

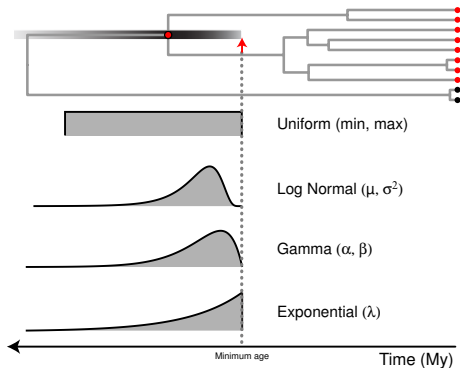


PRIOR DENSITIES ON CALIBRATED NODES

Common practice in Bayesian divergence-time estimation:

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

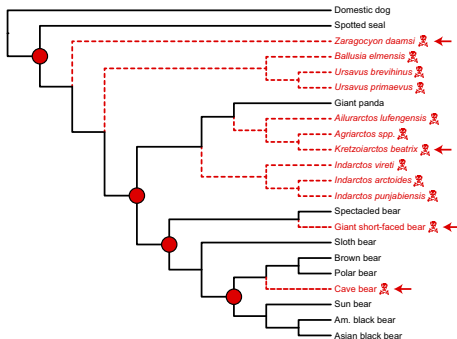
Specifying appropriate densities is a challenge for most molecular biologists



IMPROVING FOSSIL CALIBRATION

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

Calibration densities do not account for diversification of fossils

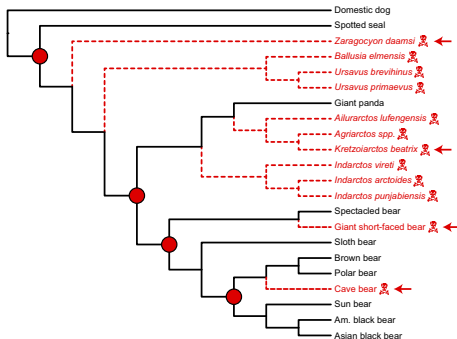


IMPROVING FOSSIL CALIBRATION

We want to use all
of the available fossils

Example: Bears

12 fossils are reduced
to 4 calibration ages
with calibration density
methods

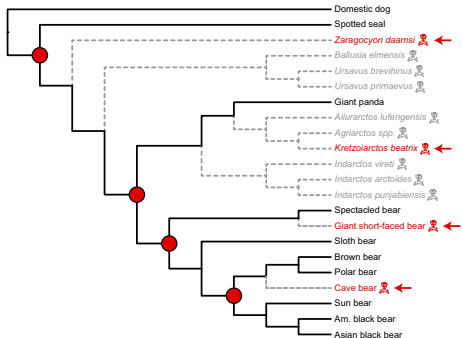


IMPROVING FOSSIL CALIBRATION

We want to use all
of the available fossils

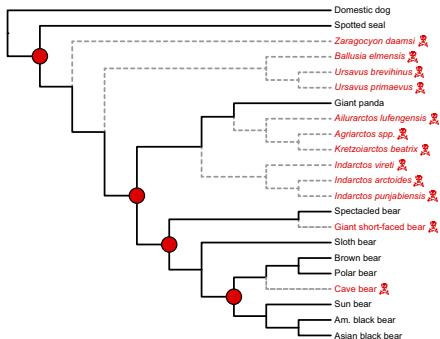
Example: Bears

12 fossils are reduced
to 4 calibration ages
with calibration density
methods



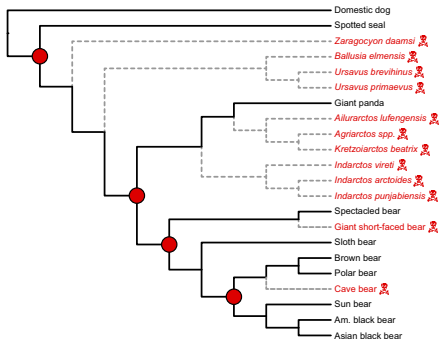
IMPROVING FOSSIL CALIBRATION

Because fossils are part of the diversification process, we can combine fossil calibration with birth-death models



IMPROVING FOSSIL CALIBRATION

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



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 & Colwell: in Dudley (ed.), *The Unity of Evolutionary Biology: Proceedings of the Fourth International Congress of Systematic & Evolutionary Biology*, Discorides Press, Portland, OR, p. 16. (1994)

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 & 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

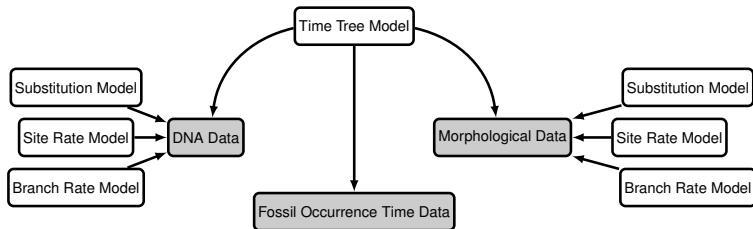
Advance Access publication on June 20, 2012

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

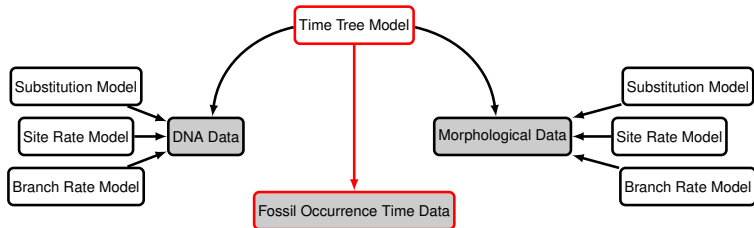
COMBINING FOSSIL & EXTANT DATA

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



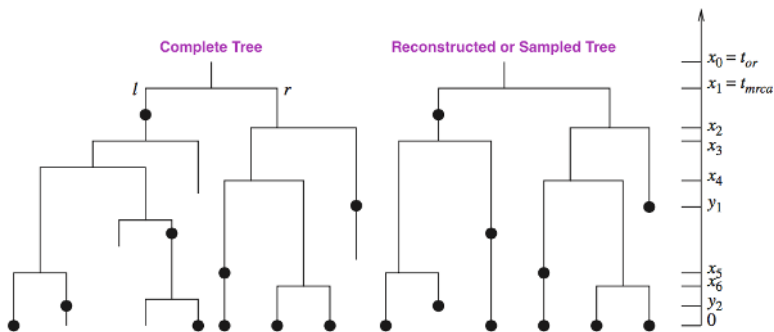
COMBINING FOSSIL & EXTANT DATA

Until recently, analyses combining fossil & extant taxa used simple or inappropriate models to describe the tree and fossil ages



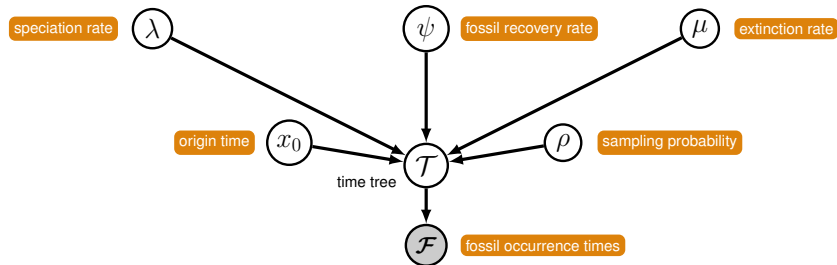
MODELING THE TREE & OCCURRENCE TIMES

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



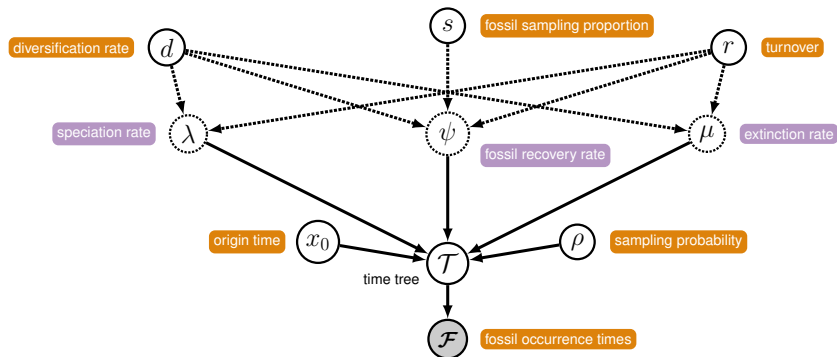
PARAMETERS OF THE FBD

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



PARAMETERS OF THE FBD

We re-parameterize the model so that we are directly estimating the diversification rate, turnover and fossil sampling proportion



$$\lambda = \frac{d}{1-r} \quad \mu = \frac{rd}{1-r} \quad \psi = \frac{s}{1-s} \frac{rd}{1-r}$$

THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

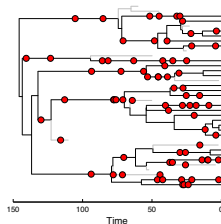
Improving statistical inference of absolute node ages

Eliminates the need to specify arbitrary calibration densities

Useful for 'total-evidence' analyses

Better capture our statistical uncertainty in species divergence dates

All reliable fossils associated with a clade are used

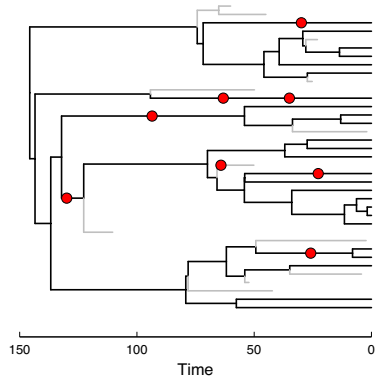


The fossilized birth–death process for coherent calibration of divergence-time estimates

Tracy A. Heath^{a,b}, John P. Huelsenbeck^{a,c}, and Tanja Stadler^{d,e,1}

THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



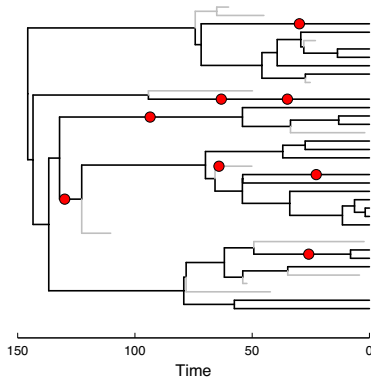
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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

λ = speciation

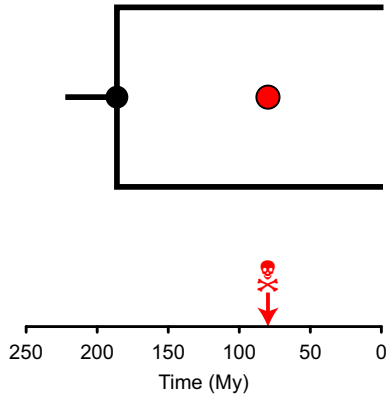
μ = extinction

ψ = fossilization/recovery



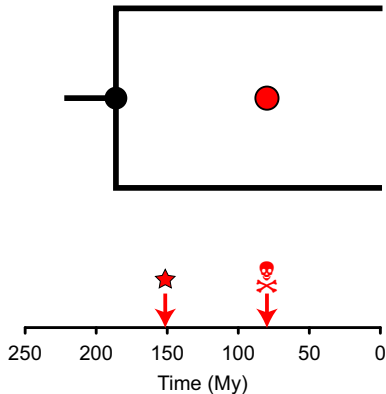
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

The occurrence time of the fossil 🦴 indicates an observation of the birth-death process before the present



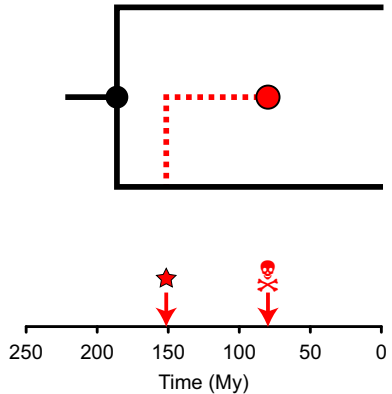
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

The fossil must attach to the tree at some time and to some branch: ★



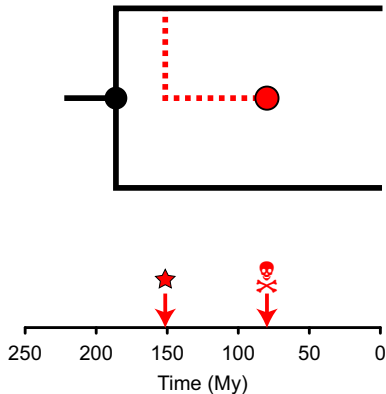
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

If it is the descendant of an unobserved lineage, then there is a speciation event at time ★



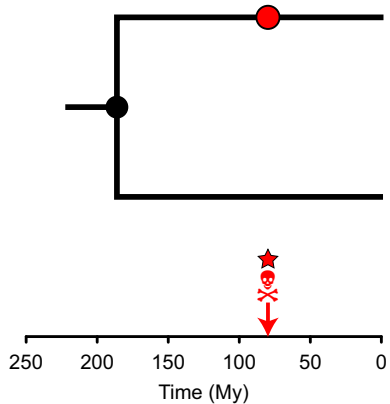
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

MCMC is used to propose new topological placements for the fossil



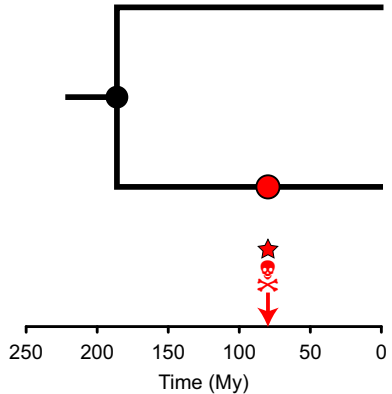
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Using rjMCMC, we can propose ★ = 💀, which means that the fossil is a “sampled ancestor”



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

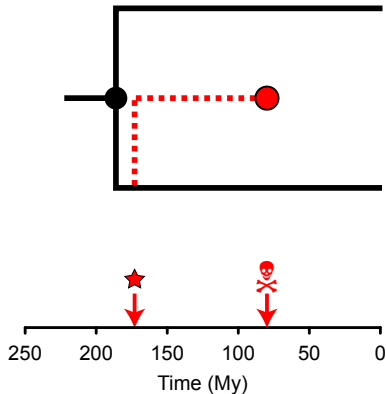
Using rjMCMC, we can propose ★ = 💀, which means that the fossil is a “sampled ancestor”



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

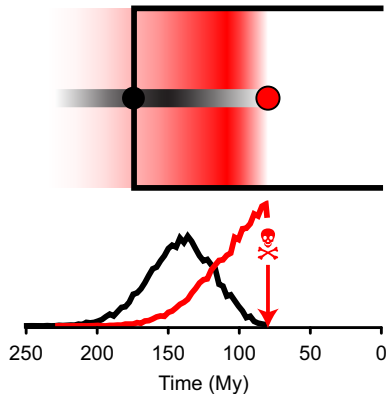
The probability of any realization of the diversification process is conditional on:

λ , μ , and ψ



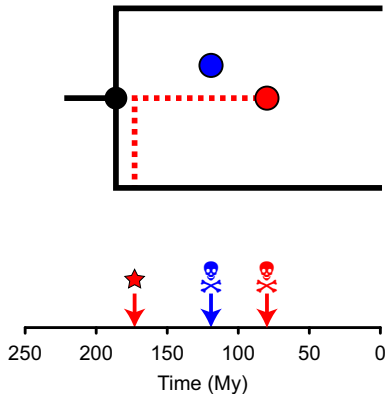
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Using MCMC, we can sample the age of the MRCA ● and the placement and time of the fossil lineage



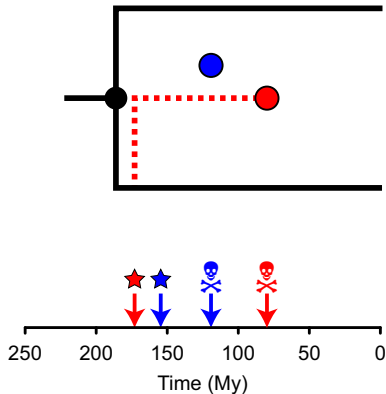
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Under the **FBD**, multiple fossils are considered, even if they are descended from the same MRCA node in the extant tree



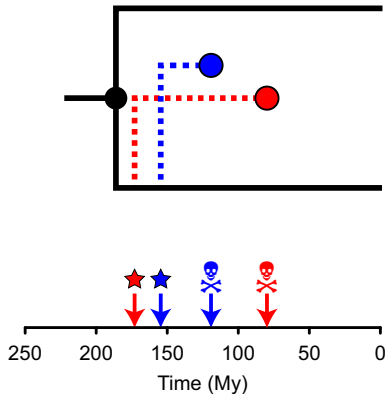
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Given ★ and ☠, the new fossil can attach to the tree via speciation along either branch in the extant tree at time ★



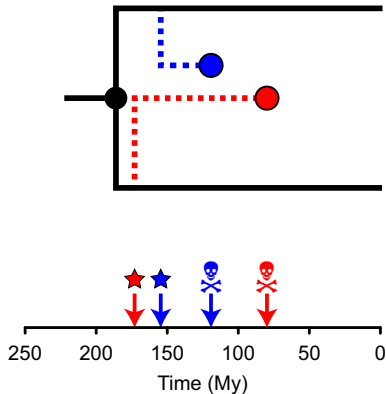
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Given ★ and ☠, the new fossil can attach to the tree via speciation along either branch in the extant tree at time ★



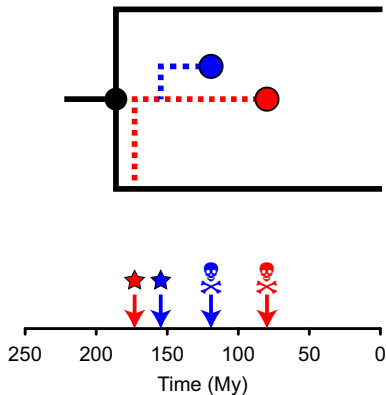
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Given ★ and ☠, the new fossil can attach to the tree via speciation along either branch in the extant tree at time ★



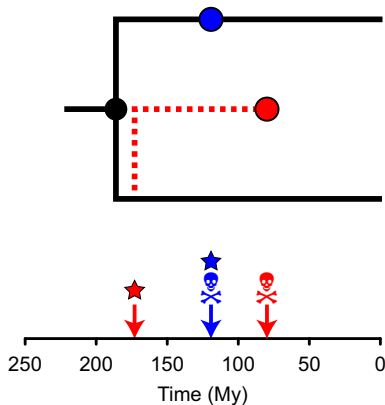
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Or the unobserved branch leading to the other fossil



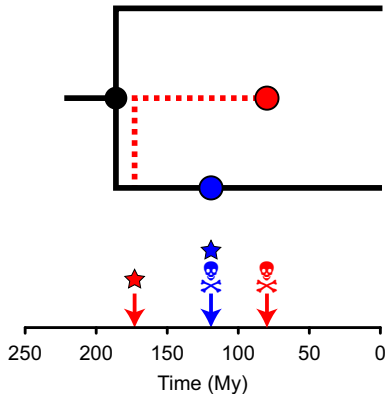
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

If ★ = ☠, then the new fossil lies directly on a branch in the extant tree



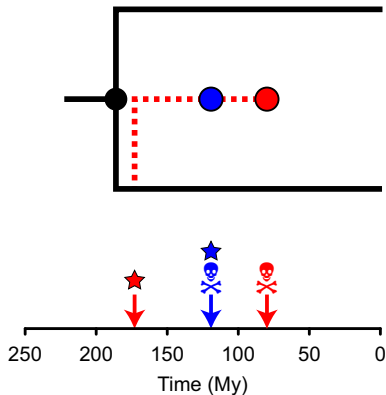
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

If ★ = ☠, then the new fossil lies directly on a branch in the extant tree



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

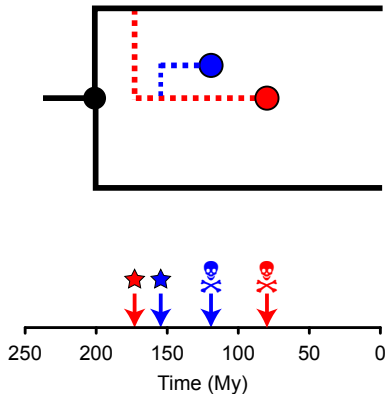
Or it is an ancestor of the other sampled fossil



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

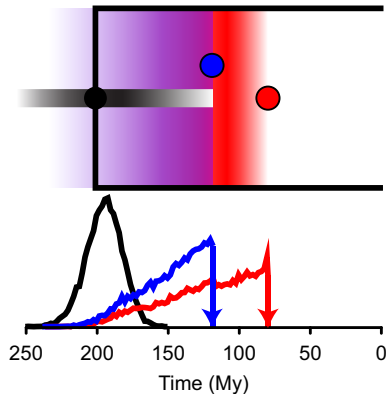
The probability of this realization of the diversification process is conditional on:

λ , μ , and ψ



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Using MCMC, we can sample the age of the MRCA ● and the placement and time of all fossil lineages



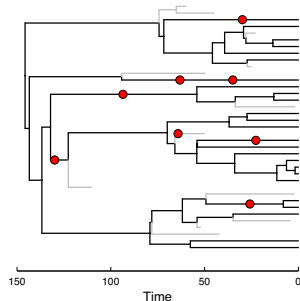
SAMPLED ANCESTORS

Sampled lineages with sampled descendants

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

On the probability of ancestors in the fossil record

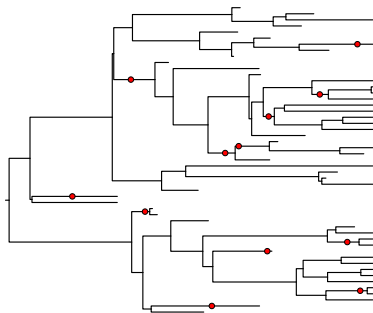
Mike Foote



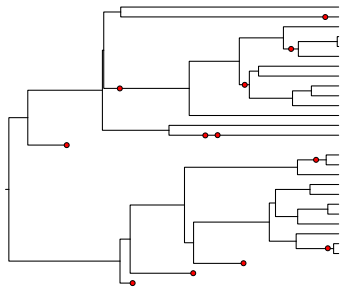
There is a non-zero probability of sampling ancestor-descendant relationships from the fossil record

SAMPLED ANCESTORS

Complete FBD Tree



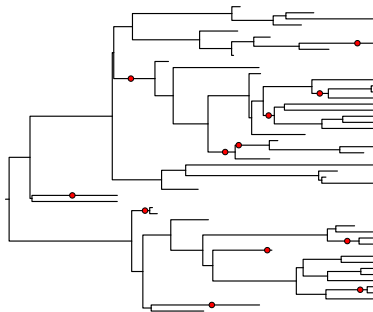
Reconstructed FBD Tree



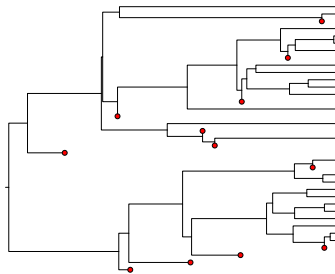
Because fossils & living taxa are assumed to come from a single diversification process, there is a non-zero probability of sampled ancestors

SAMPLED ANCESTORS

Complete FBD Tree

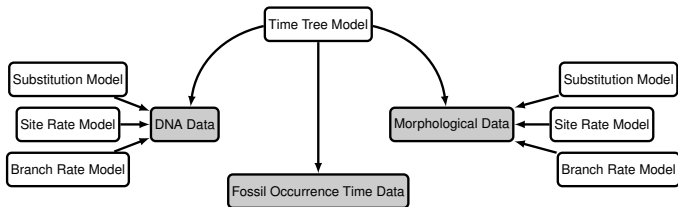


No Sampled Ancestor Tree



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

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

T1 **0**

T2 **0**

T3 **1**

T4 **2**

T5 **2**

T6 **1**

T7 **1**

MODELING MORPHOLOGICAL CHARACTER CHANGE

The Lewis Mk_v model

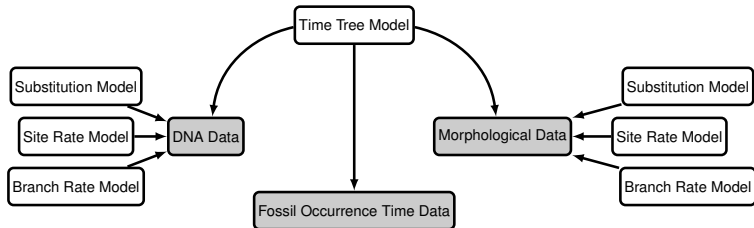
Accounts for the ascertainment bias in morphological datasets by conditioning the likelihood on the fact that invariant characters are not sampled

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

T1	0	0	1	2
T2	0	0	1	2
T3	1	0	1	2
T4	2	0	1	2
T5	2	0	1	2
T6	1	0	1	2
T7	1	0	1	2

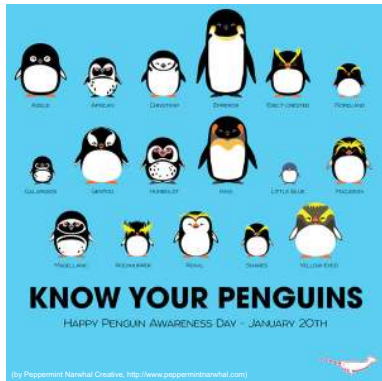
"TOTAL-EVIDENCE" ANALYSIS

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

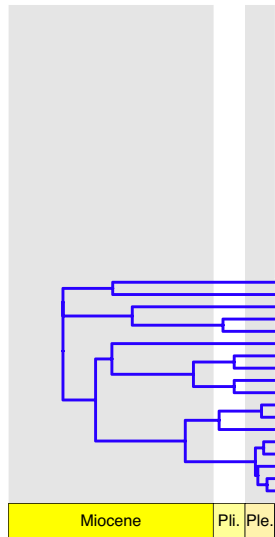
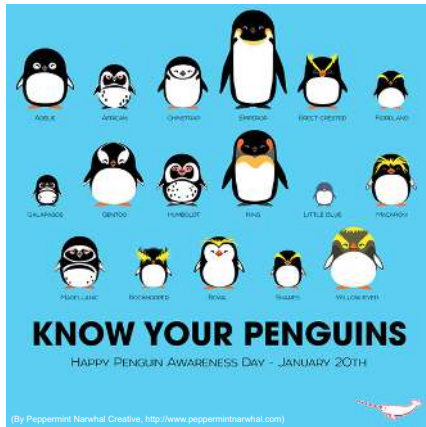


PENGUIN DIVERSITY IN DEEP TIME

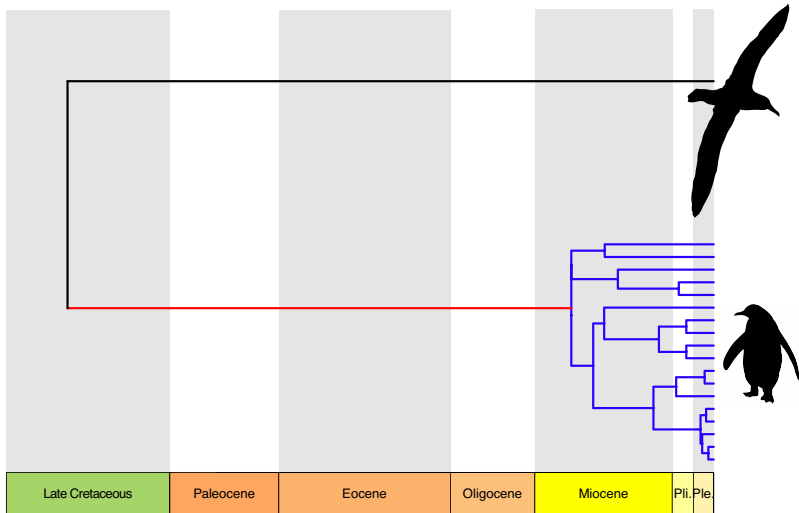
How does our understanding of penguin evolution improve when we consider both extant and fossil taxa?



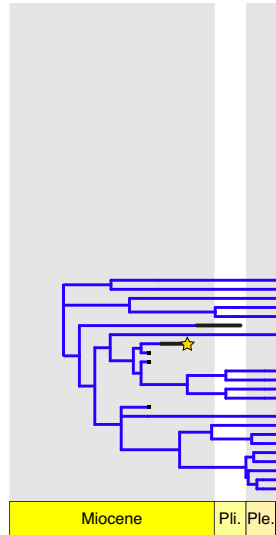
PENGUIN DIVERSITY



PENGUIN DIVERSITY



FOSSIL PENGUIN DIVERSITY

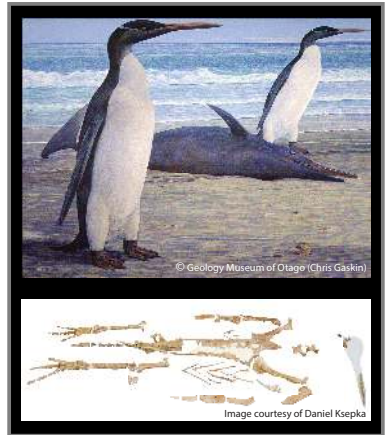


(*S. urbinai* holotype fossil, 5-7 MYA, image by Martin Chávez)

PENGUINS IN THE OLIGOCENE

Kairuku

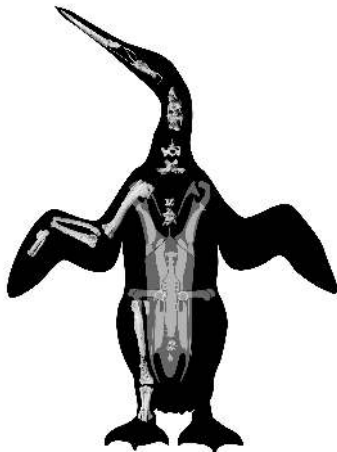
- ~1.5 m tall
- slender, with narrow bill
- scapula & pygostyle are more similar to non-penguins
- ~27 Mya



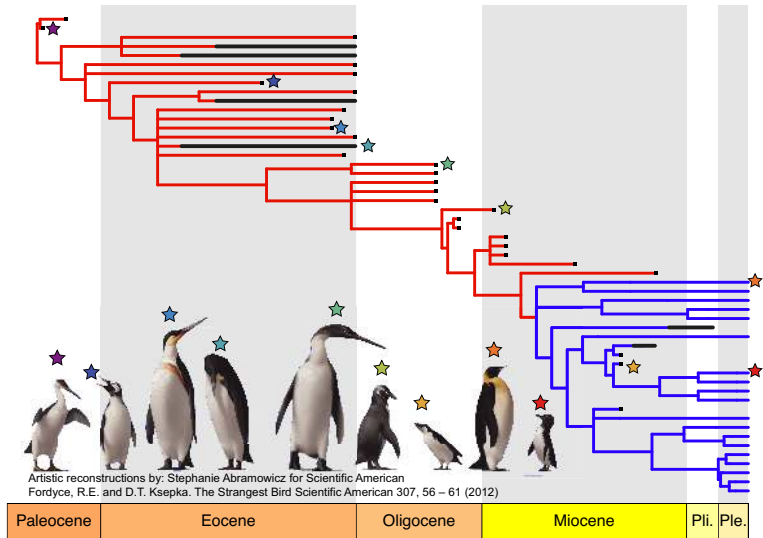
PENGUINS IN THE PALEOCENE

Waimanu

- oldest known penguin species
- intermediate wing morphology
- ~58–61.6 Mya



PENGUIN DIVERSITY IN DEEP TIME

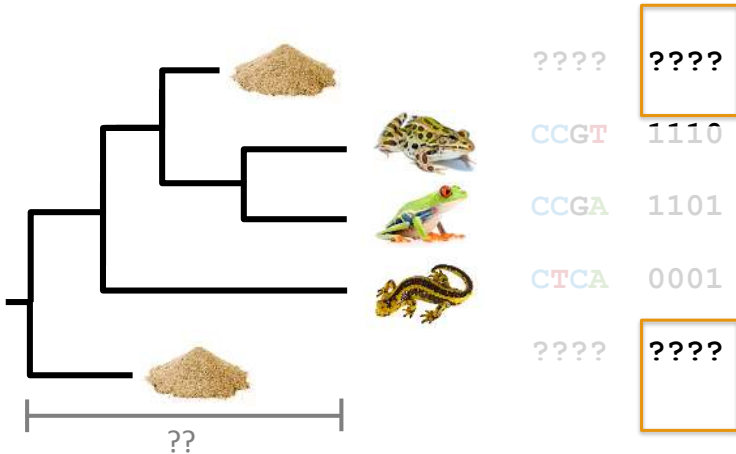


PENGUIN DIVERSITY IN DEEP TIME

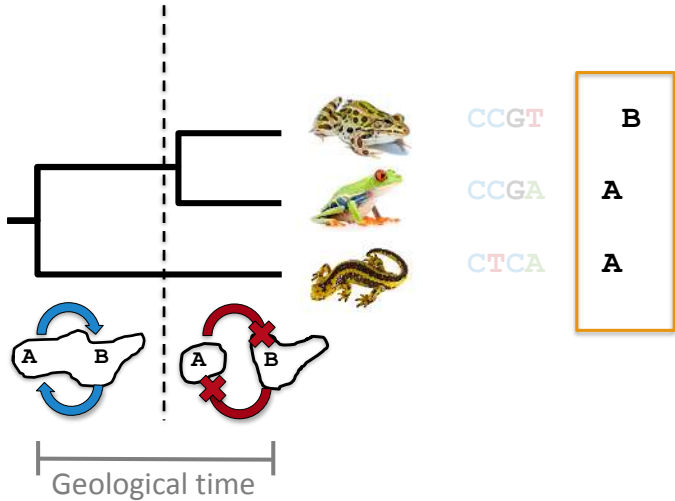
PENGUIN DIVERSITY IN DEEP TIME

MOLECULES + MORPHOLOGY + FOSSILS

...but I study amphibians...



Molecules + biogeography + paleogeography

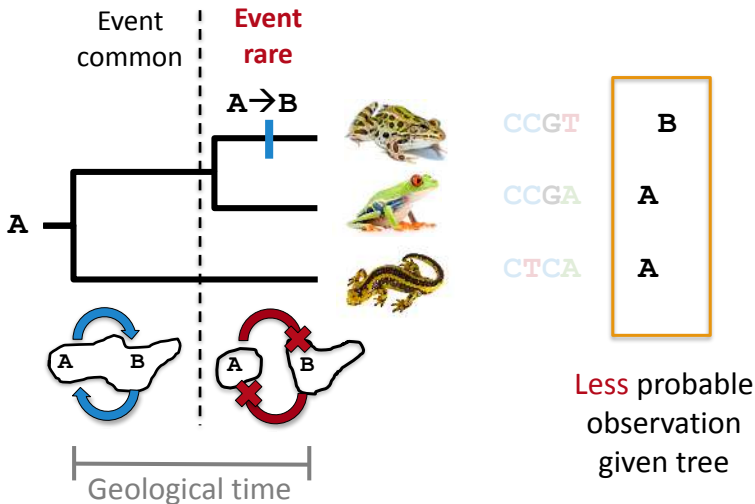


+ Paleogeography

Landis, 2016

(slides courtesy of M. Landis; <http://bit.ly/2aIHqB4>)

Events should occur *before* areas split

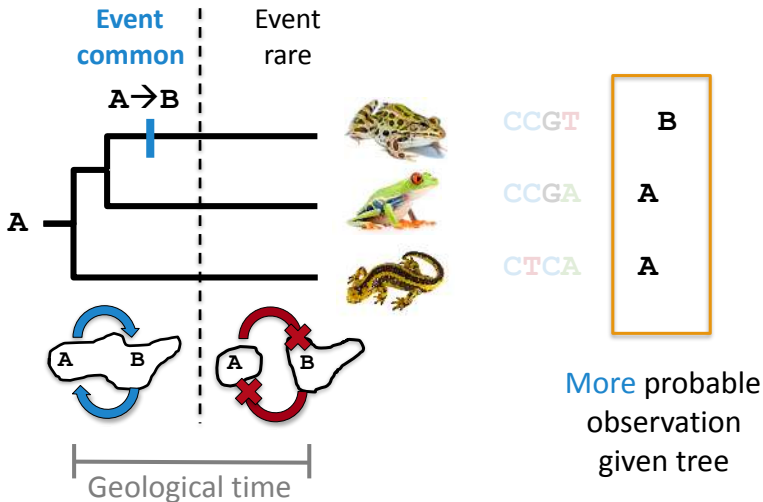


+ Paleogeography

Landis, 2016

(slides courtesy of M. Landis; <http://bit.ly/2aIHqB4>)

Events should occur *before areas split*

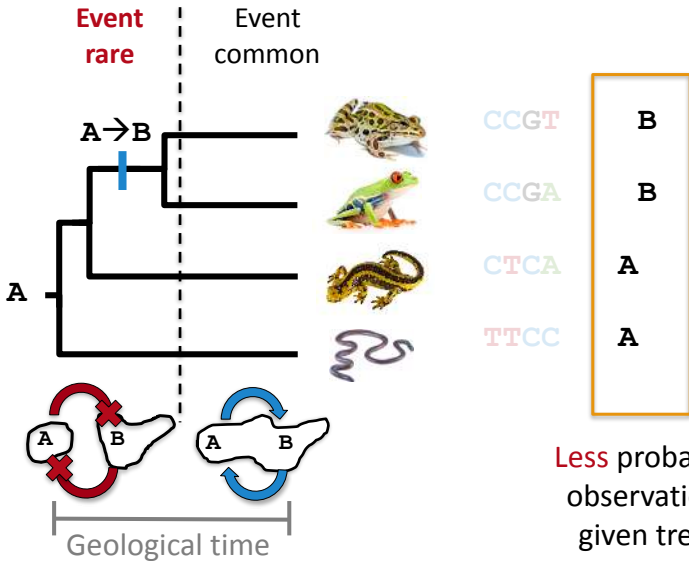


+ Paleogeography

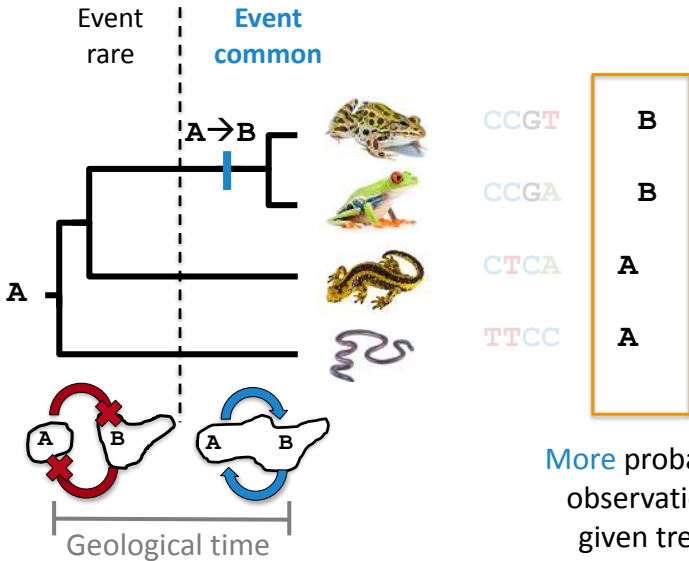
Landis, 2016

(slides courtesy of M. Landis; <http://bit.ly/2aIHqB4>)

Events should occur *after* areas merge



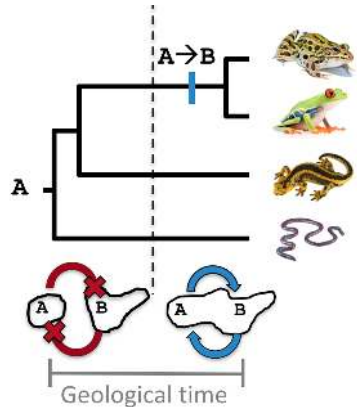
Events should occur *after* areas merge



BIOGEOGRAPHIC DATING

Fossil-free calibration

- data: molecular sequences
- data: biogeographic ranges
- empirical paleogeographic model that alters the rates of biogeographic change over time

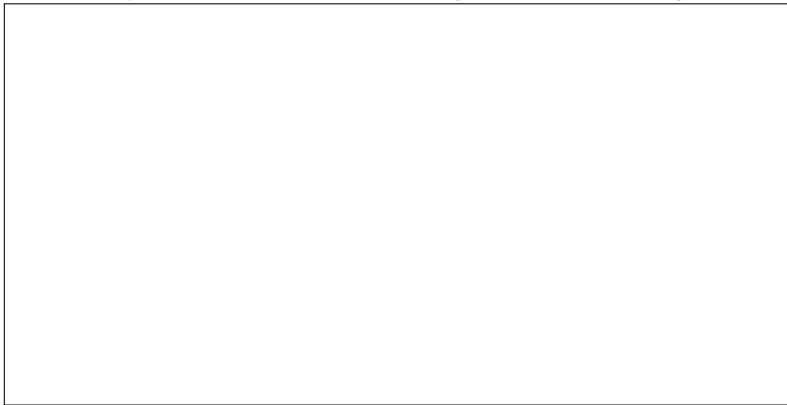


Landis. In Press. "Biogeographic Dating of Speciation Times Using Paleogeographically Informed Processes". *Systematic Biology*, doi: 10.1093/sysbio/syw040.

BIOGEOGRAPHIC DATING

25 areas, 26 time-slices, 540–0Ma

strong=share land, **weak**=nearby land, **none**=all pairs

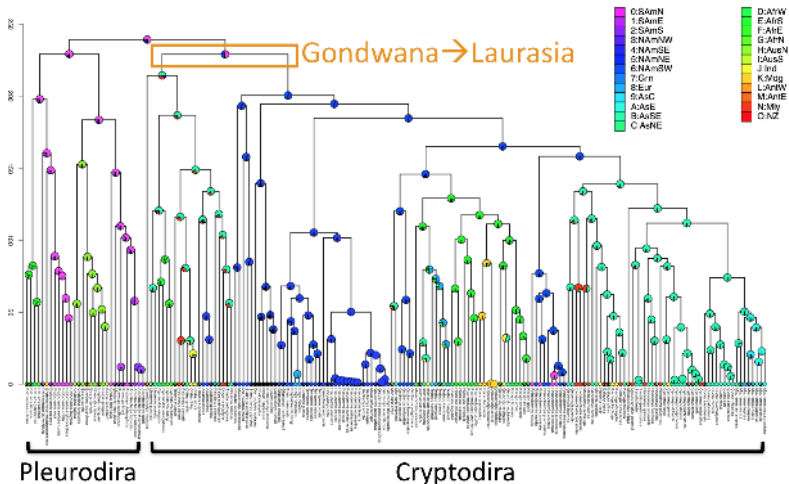


Connectivity model constructed using literature review and GPlates (<http://www.gplates.org/>)

(model & animation courtesy of M. Landis; <http://bit.ly/2alHqB4>)

DATING + ANCESTRAL AREA RECONSTRUCTION

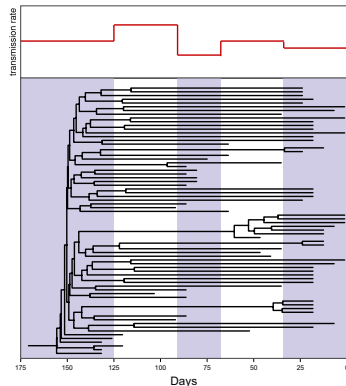
Ancestral area estimates (+G)



SKYLINE BIRTH-DEATH PROCESS

A piecewise shifting model where parameters change over time

Used to estimate epidemiological parameters of an outbreak



Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV)

Tanja Stadler^{a,1,2}, Denise Kühnert^{b,c,1}, Sebastian Bonhoeffer^a, and Alexei J. Drummond^{b,c}

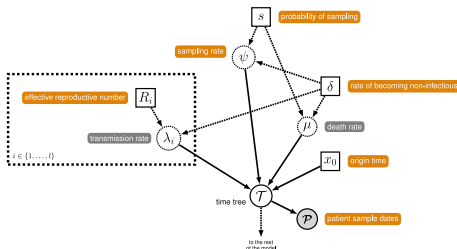
SKYLINE BIRTH-DEATH PROCESS

l is the number of parameter intervals

R_i is the effective reproductive number for interval $i \in l$

δ is the rate of becoming non-infectious

s is the probability of sampling an individual after becoming non-infectious



$$R_i = \frac{\lambda_i}{\mu + \psi}, \quad \delta = \mu + \psi, \quad s = \frac{\psi}{\mu + \psi}$$

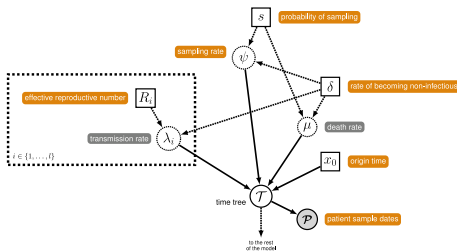
SKYLINE BIRTH-DEATH PROCESS

l is the number of parameter intervals

λ_i is the transmission rate for interval $i \in l$

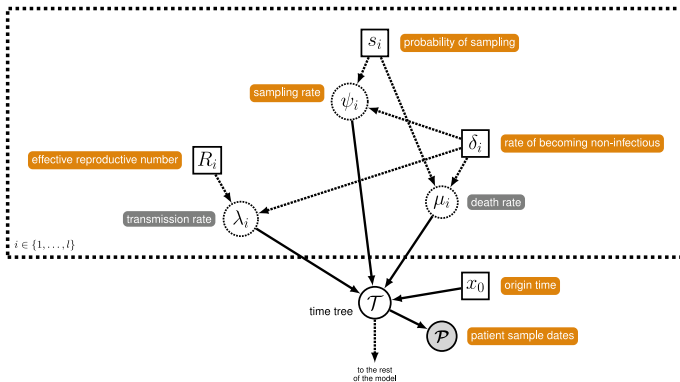
μ is the viral lineage death rate

ψ is the rate each individual is sampled



$$\lambda_i = R_i \delta, \quad \mu = \delta - s\delta, \quad \psi = s\delta$$

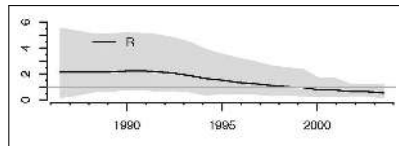
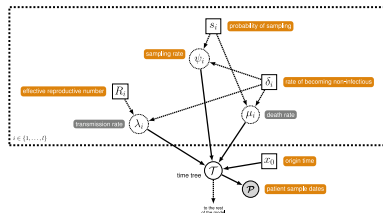
SKYLINE BIRTH-DEATH PROCESS



SKYLINE BIRTH-DEATH PROCESS

A decline in R over the history of HIV-1 in the UK is consistent with the introduction of effective drug therapies

After 1998 R decreased below 1, indicating a declining epidemic



EXERCISE IN BEAST2

This tutorial uses sequence data and fossil occurrence times to date species divergences using a relaxed-clock model and the fossilized birth-death process

