

BAYESIAN DIVERGENCE-TIME ESTIMATION

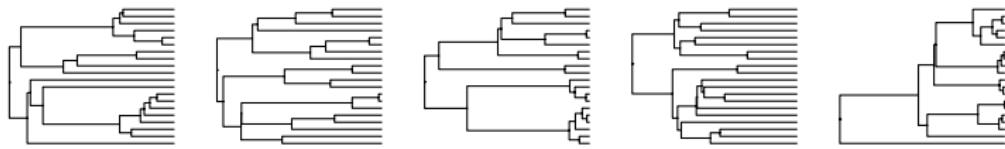
Slides by Tracy Heath
Voice acting by Joëlle Barido-Sottani

2017 Taming the BEAST Workshop, London

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



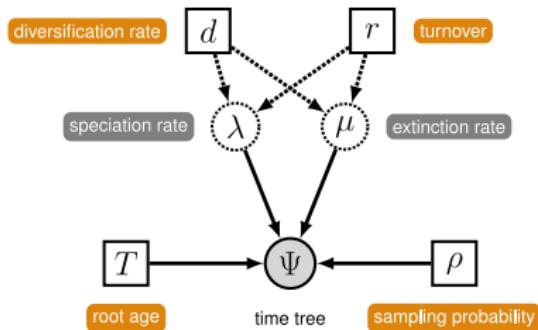
STOCHASTIC BRANCHING PROCESSES

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

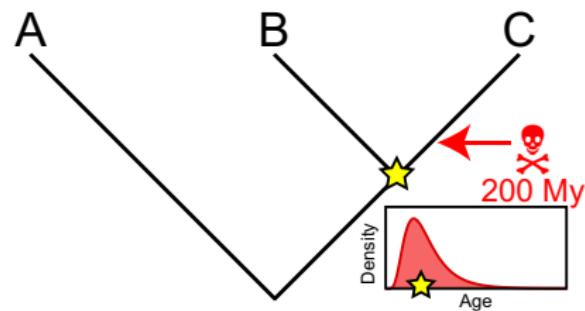
Conditions on a probability of sampling a tip, ρ



CALIBRATION DENSITIES

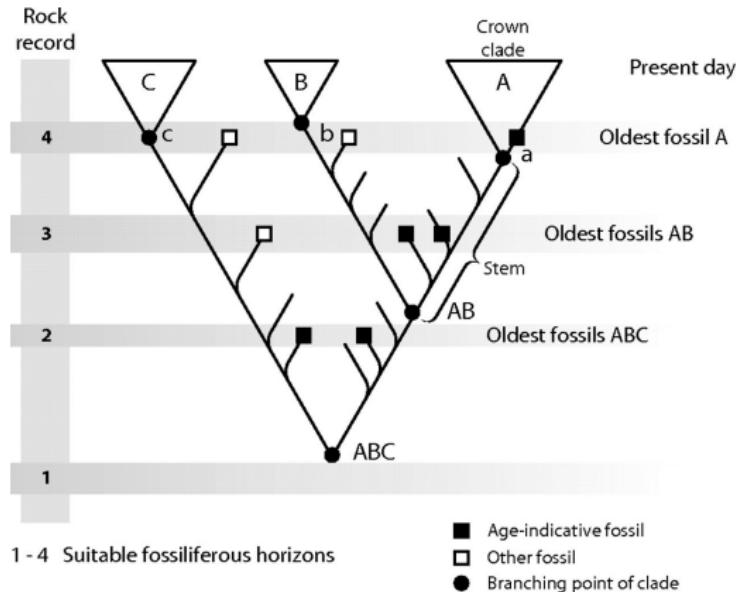
Bayesian inference is well suited to accommodating uncertainty in the age of the calibration node

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



ASSIGNING FOSSILS TO CLADES

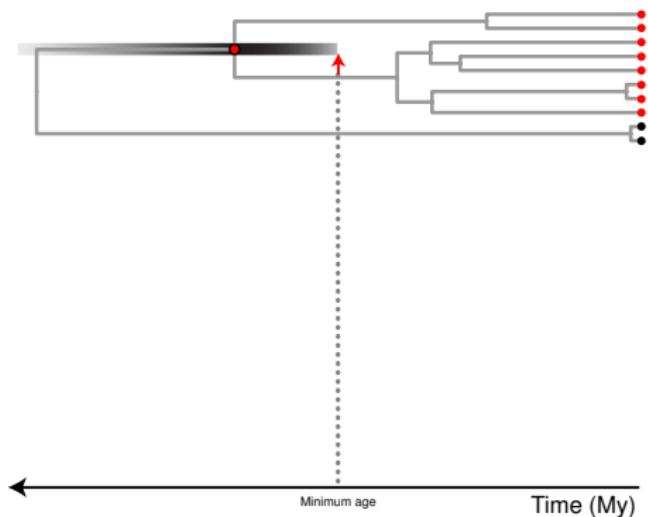
Misplaced fossils can affect node age estimates throughout the tree – if the fossil is older than its presumed MRCA



Fossil Calibration

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

Reliable **maximum** bounds are typically unavailable

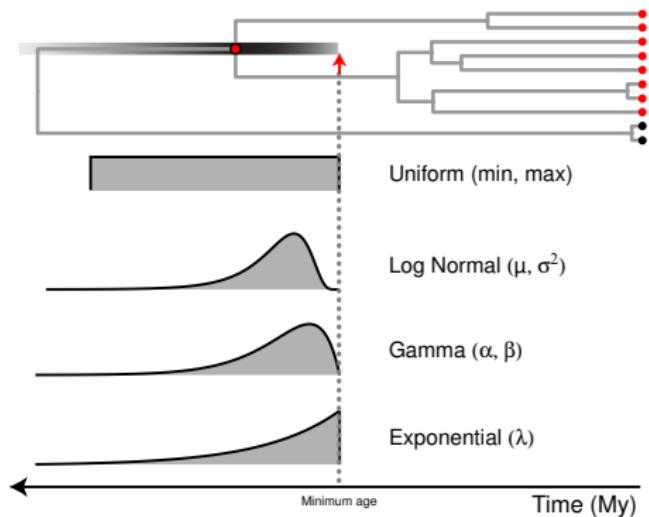


PRIOR DENSITIES ON CALIBRATED NODES

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

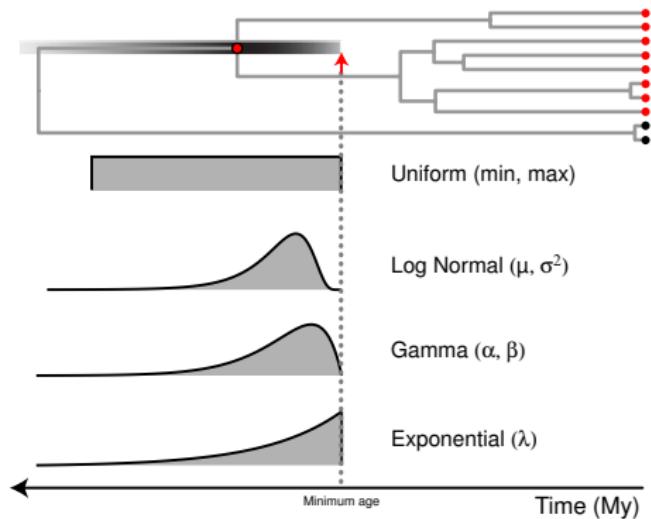


PRIOR DENSITIES ON CALIBRATED NODES

Common practice in Bayesian divergence-time estimation:

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

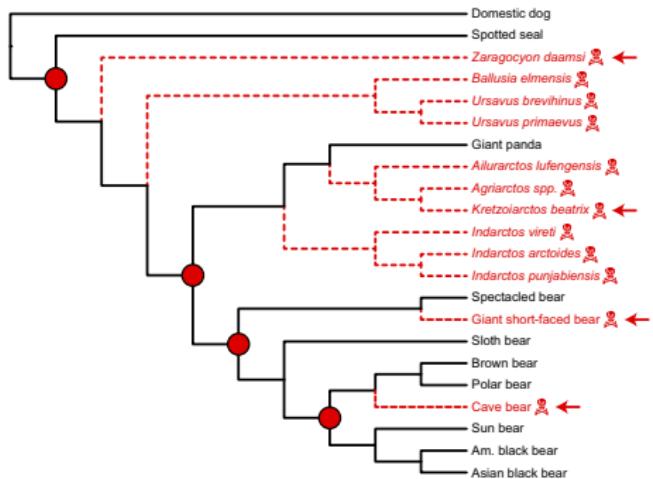
Specifying appropriate densities is a challenge for most molecular biologists



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

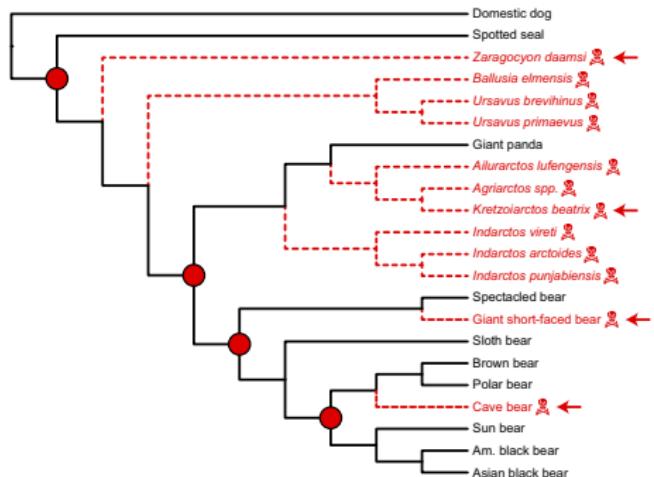


IMPROVING FOSSIL CALIBRATION

We want to use all of the available fossils

Example: Bears

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

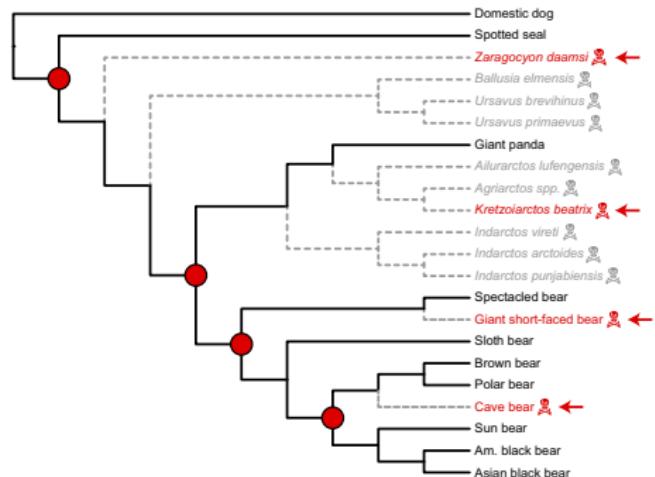


IMPROVING FOSSIL CALIBRATION

We want to use all of the available fossils

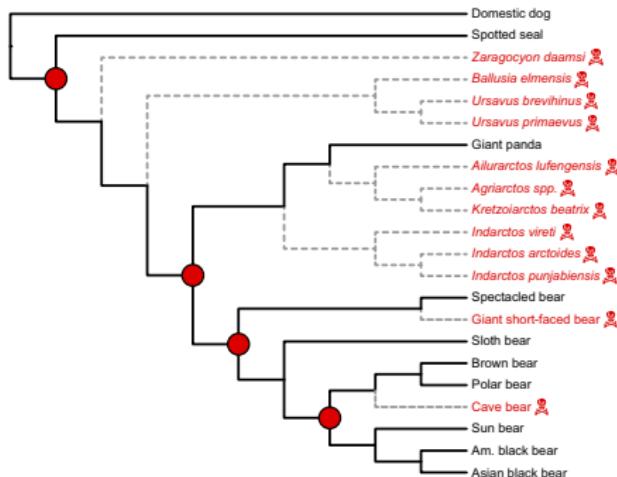
Example: Bears

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



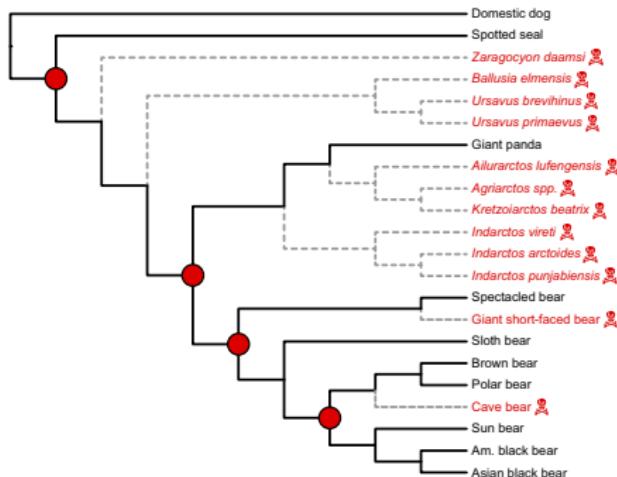
IMPROVING FOSSIL CALIBRATION

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



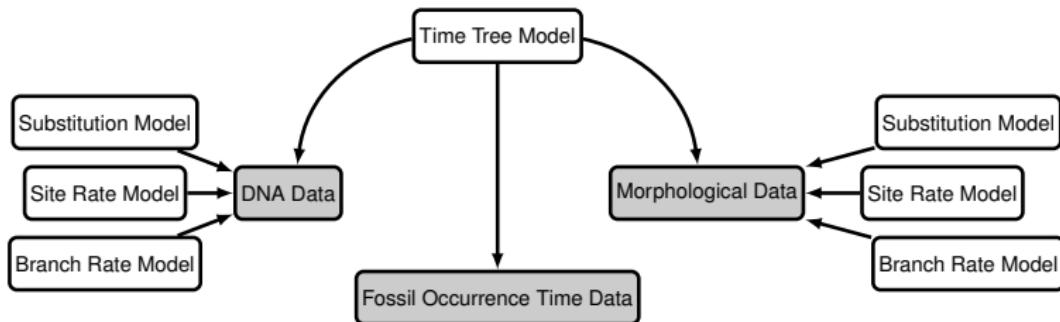
IMPROVING FOSSIL CALIBRATION

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



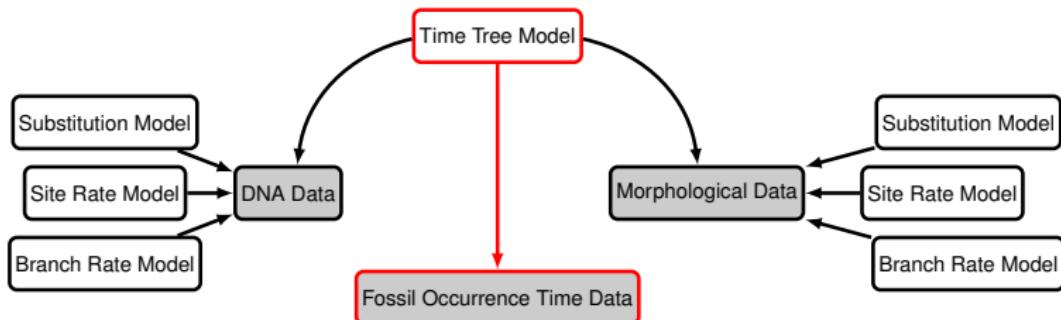
COMBINING FOSSIL & EXTANT DATA

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



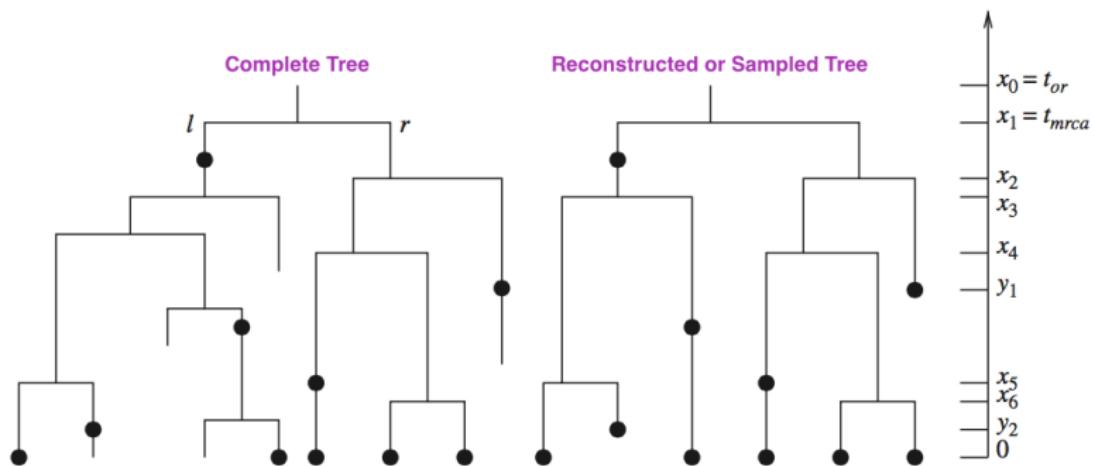
COMBINING FOSSIL & EXTANT DATA

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



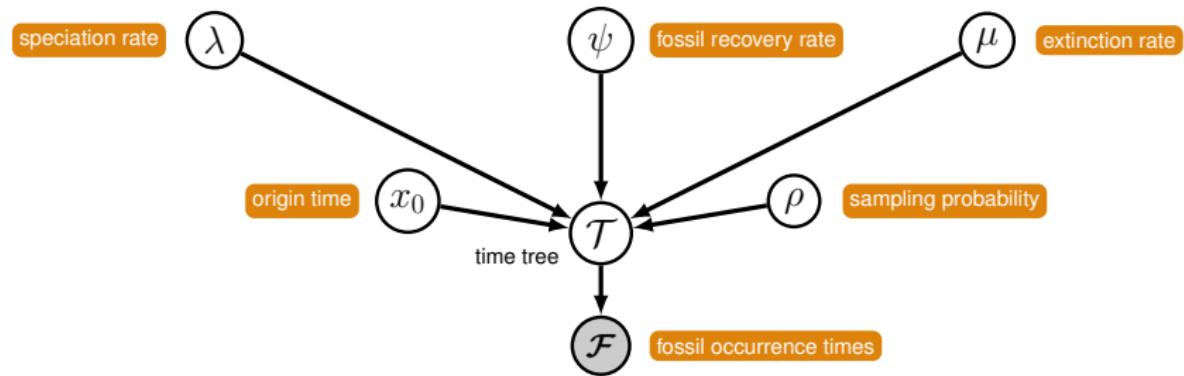
MODELING THE TREE & OCCURRENCE TIMES

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



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

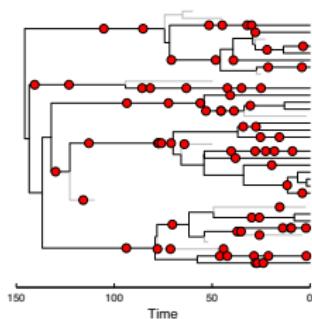
Improving statistical inference of absolute node ages

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

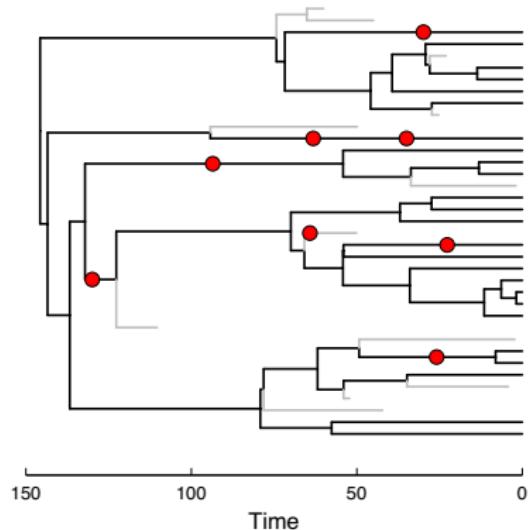


The fossilized birth–death process for coherent calibration of divergence-time estimates

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



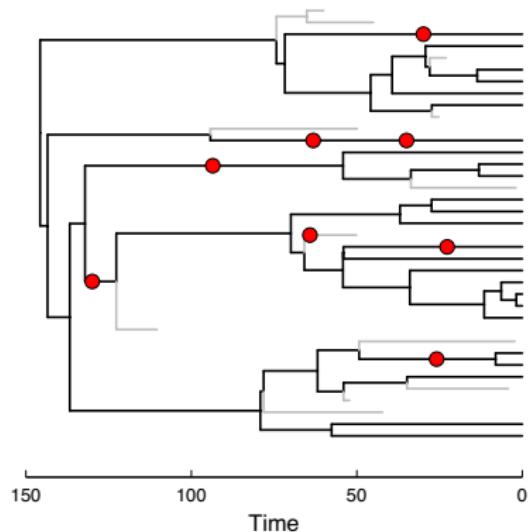
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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

λ = speciation

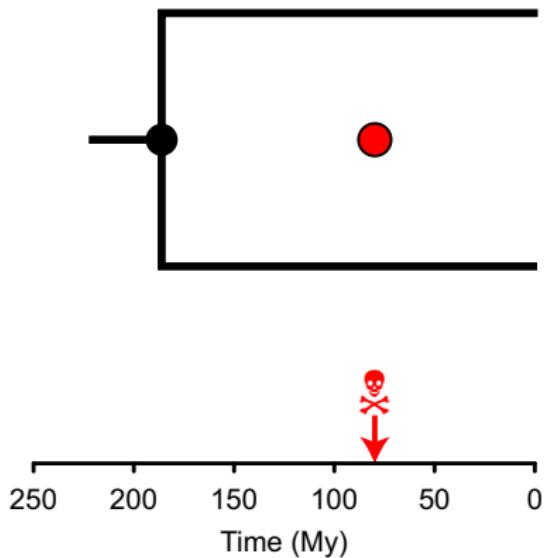
μ = extinction

ψ = fossilization/recovery



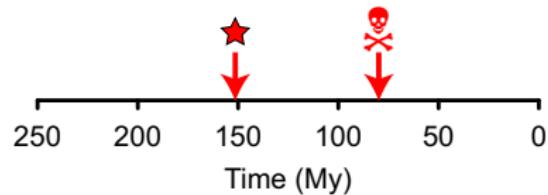
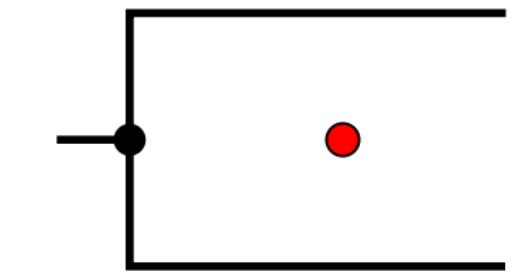
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



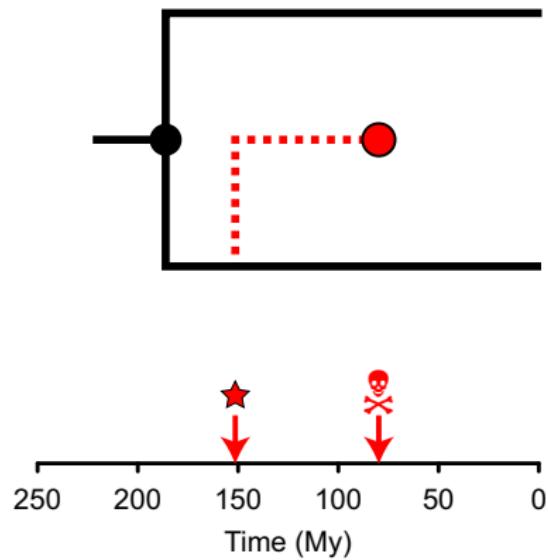
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



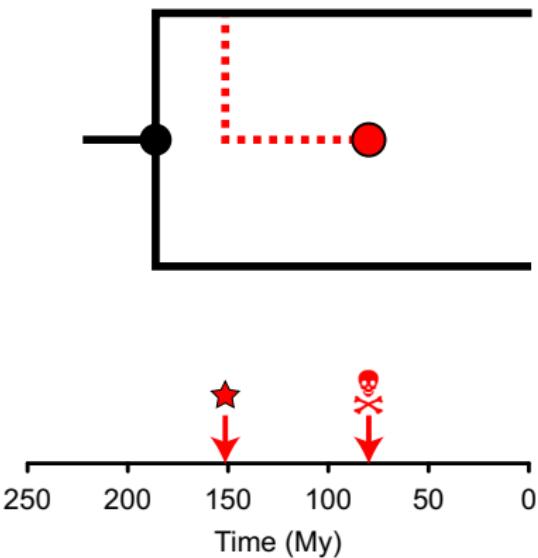
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



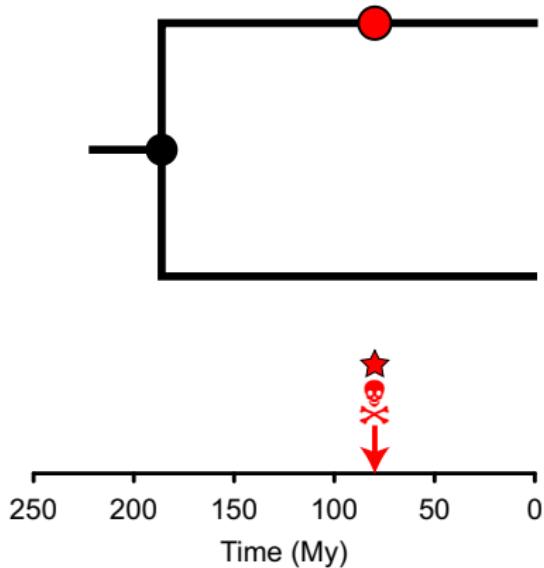
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

MCMC is used to propose new topological placements for the fossil



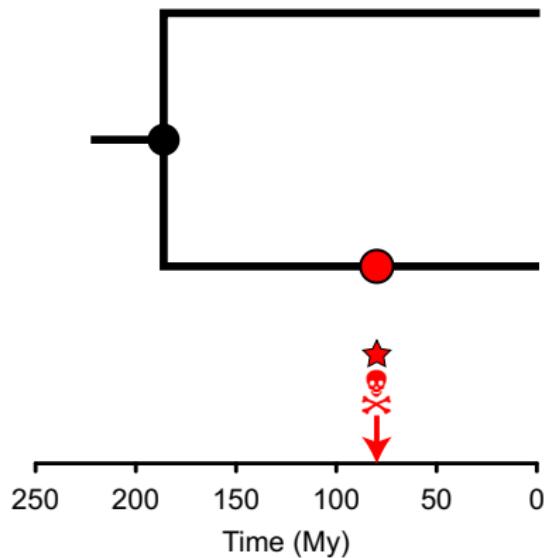
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Using rjMCMC, we can propose $\star = \text{💀}$, which means that the fossil is a "sampled ancestor"



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

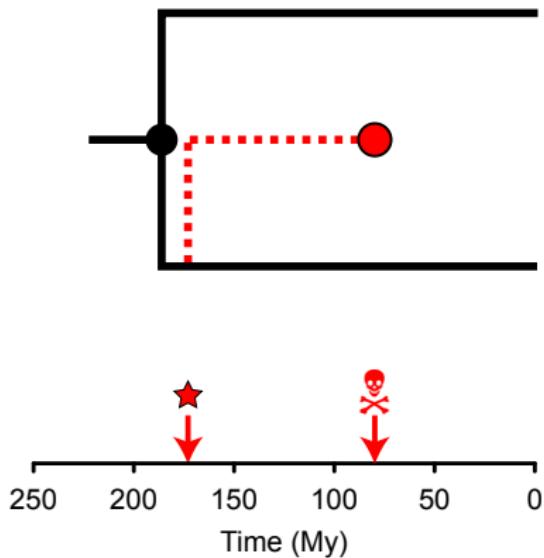
Using rjMCMC, we can propose $\star = \text{💀}$, which means that the fossil is a "sampled ancestor"



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

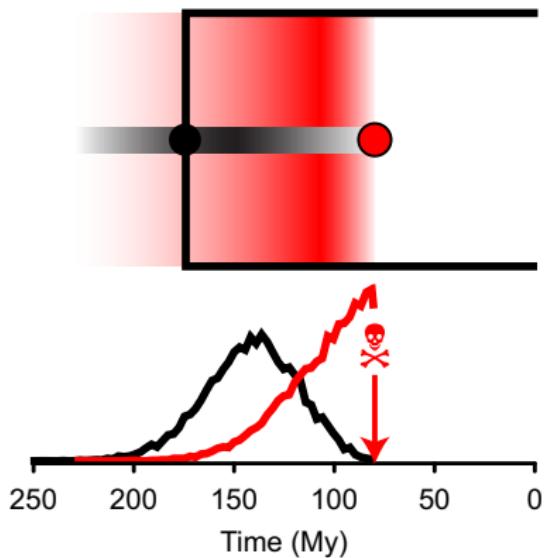
The probability of any realization of the diversification process is conditional on:

λ , μ , and ψ



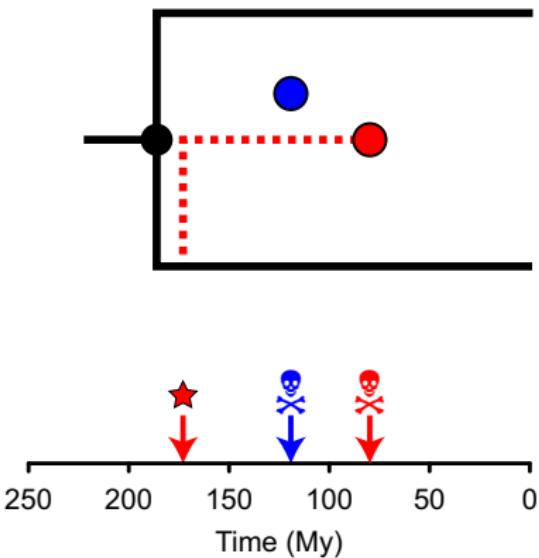
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



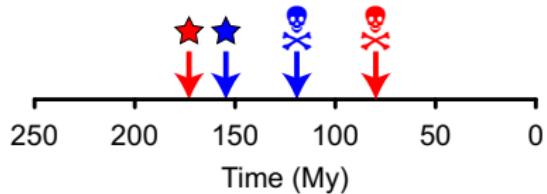
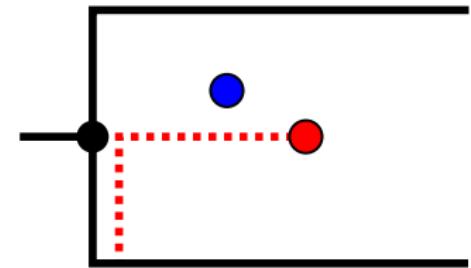
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



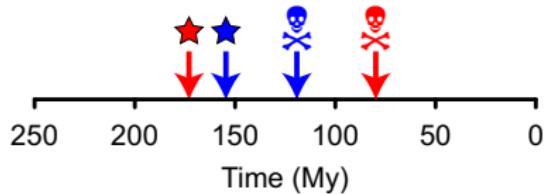
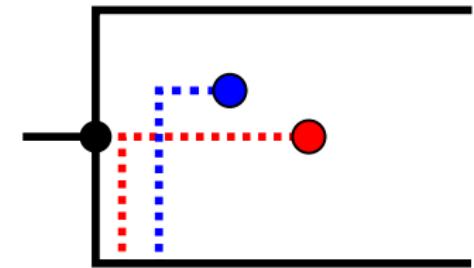
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



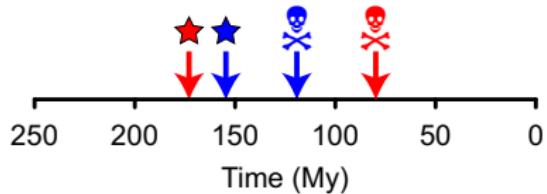
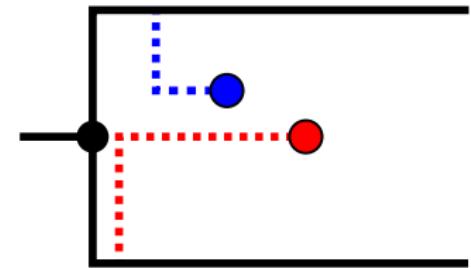
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



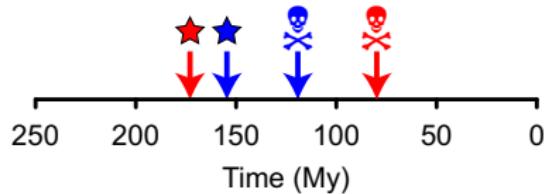
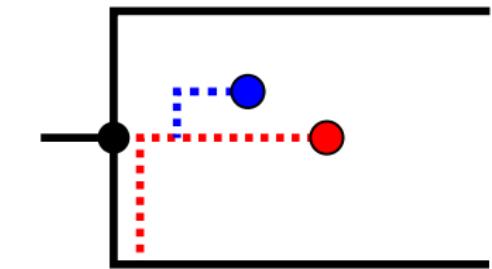
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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



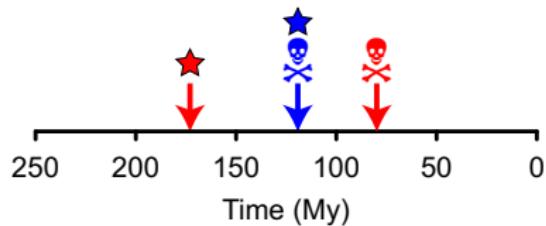
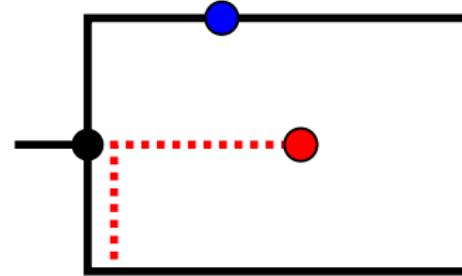
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

Or the unobserved branch leading to the other fossil



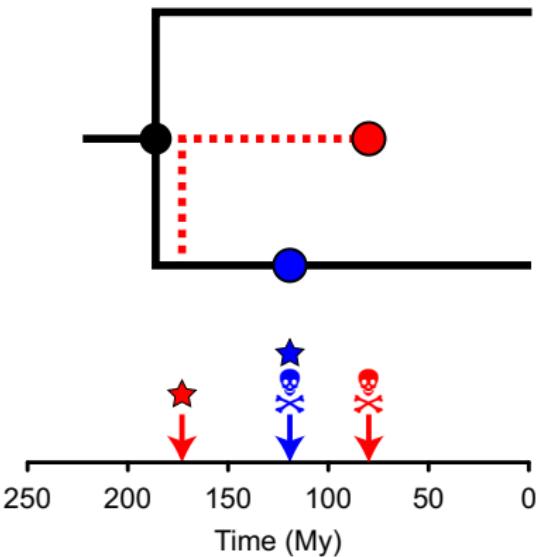
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

If $\star = \text{💀}$, then the new fossil lies directly on a branch in the extant tree



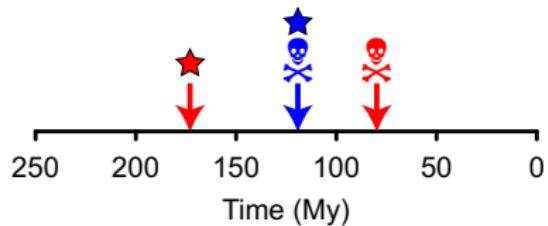
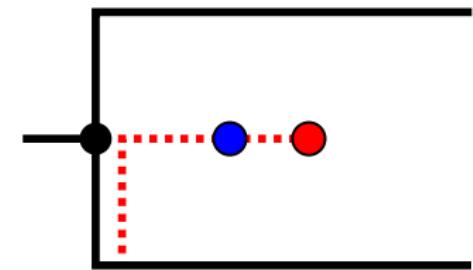
THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

If $\star = \text{💀}$, then the new fossil lies directly on a branch in the extant tree



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

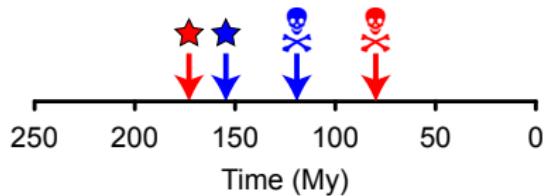
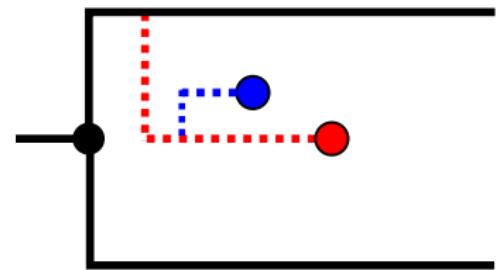
Or it is an ancestor of the other sampled fossil



THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

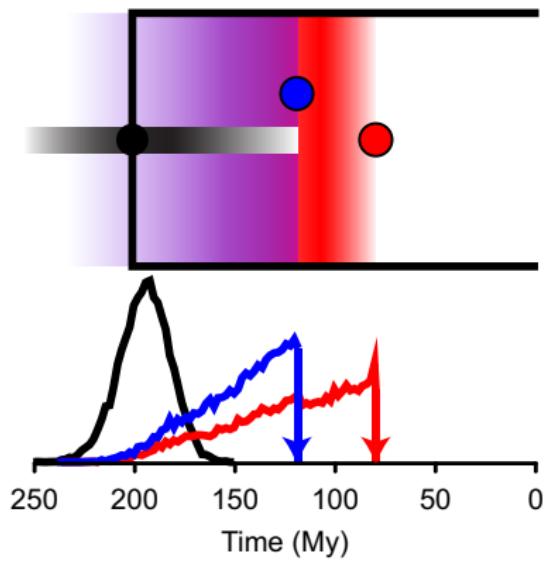
The probability of this realization of the diversification process is conditional on:

λ , μ , and ψ

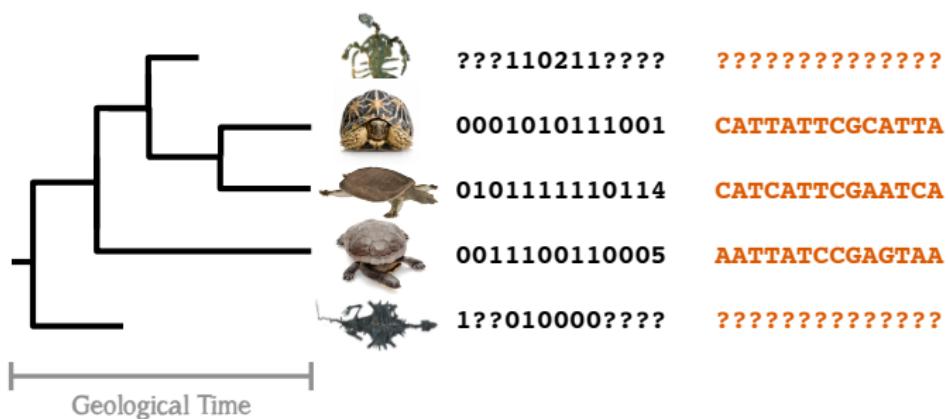
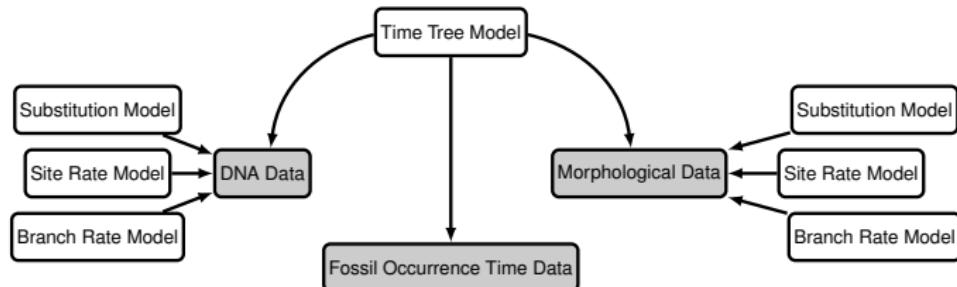


THE FOSSILIZED BIRTH-DEATH PROCESS (FBD)

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

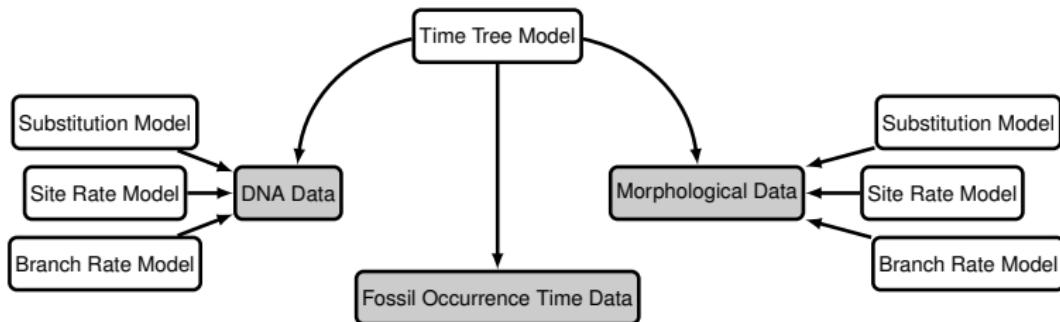


COMBINING FOSSIL & EXTANT DATA



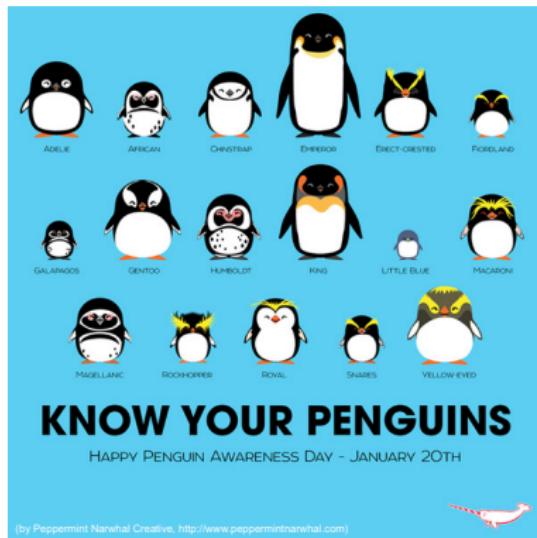
"TOTAL-EVIDENCE" ANALYSIS

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



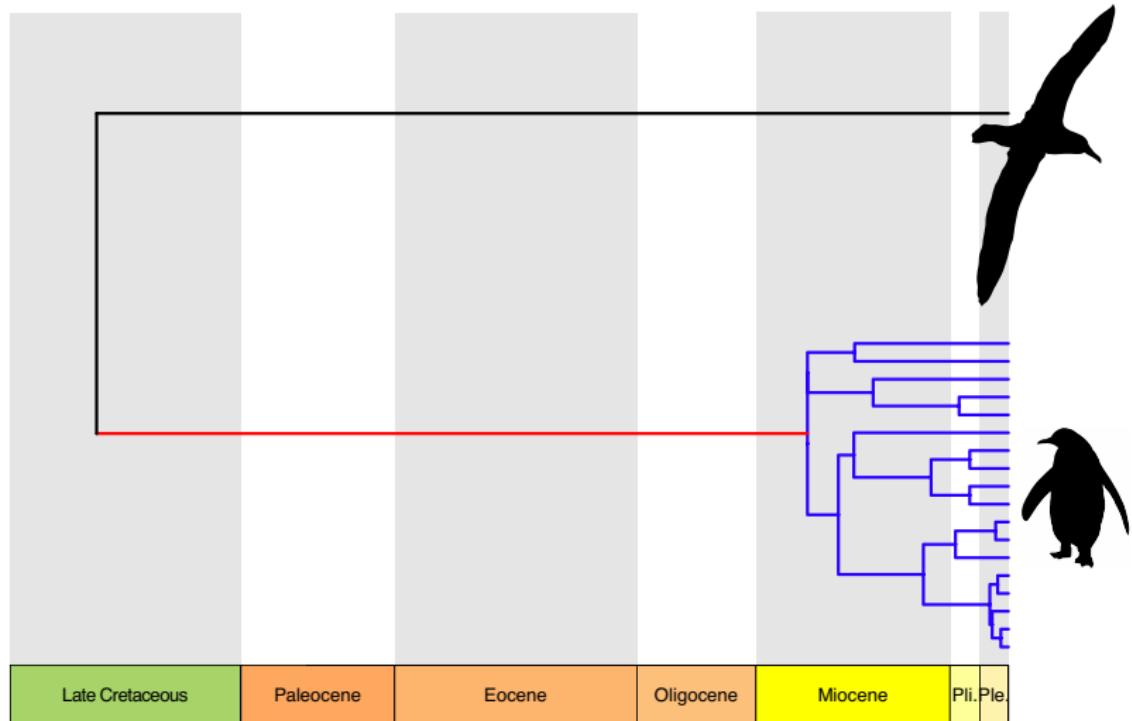
PENGUIN DIVERSITY IN DEEP TIME

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

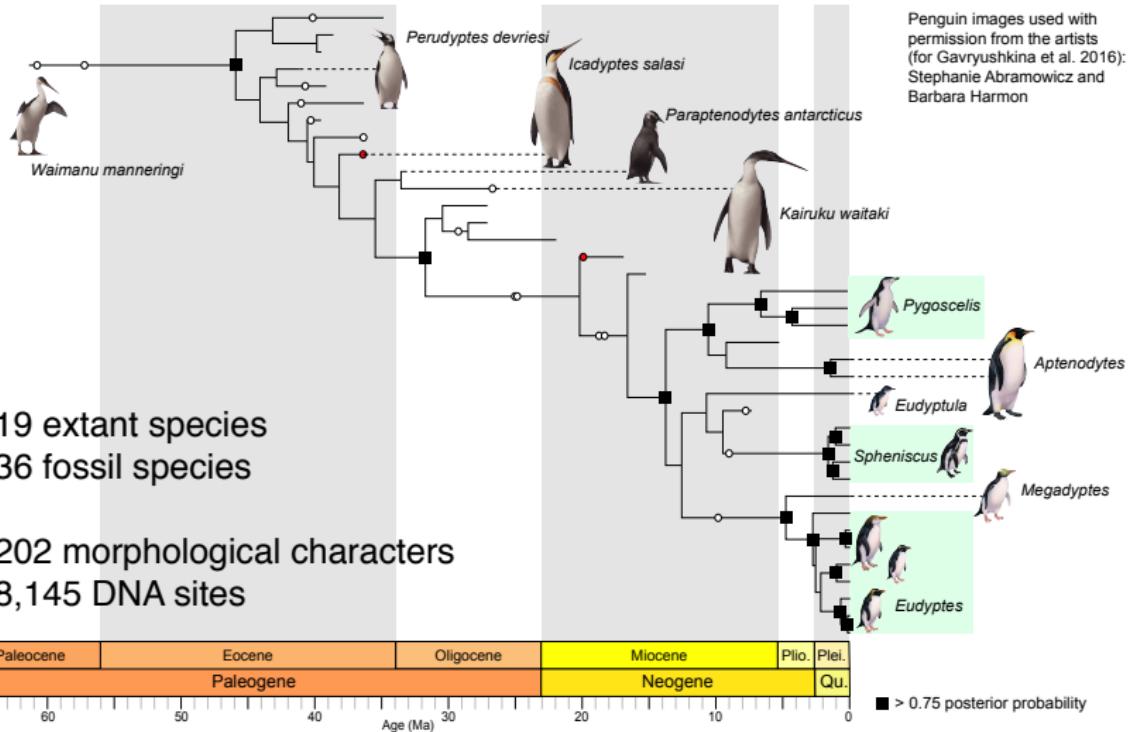


Artistic reconstructions by: Stephanie Abramowicz for Scientific American
Fordyce, R.E. and D.T. Keayka. The Strangest Bird. Scientific American 307, 56 – 61 (2012)

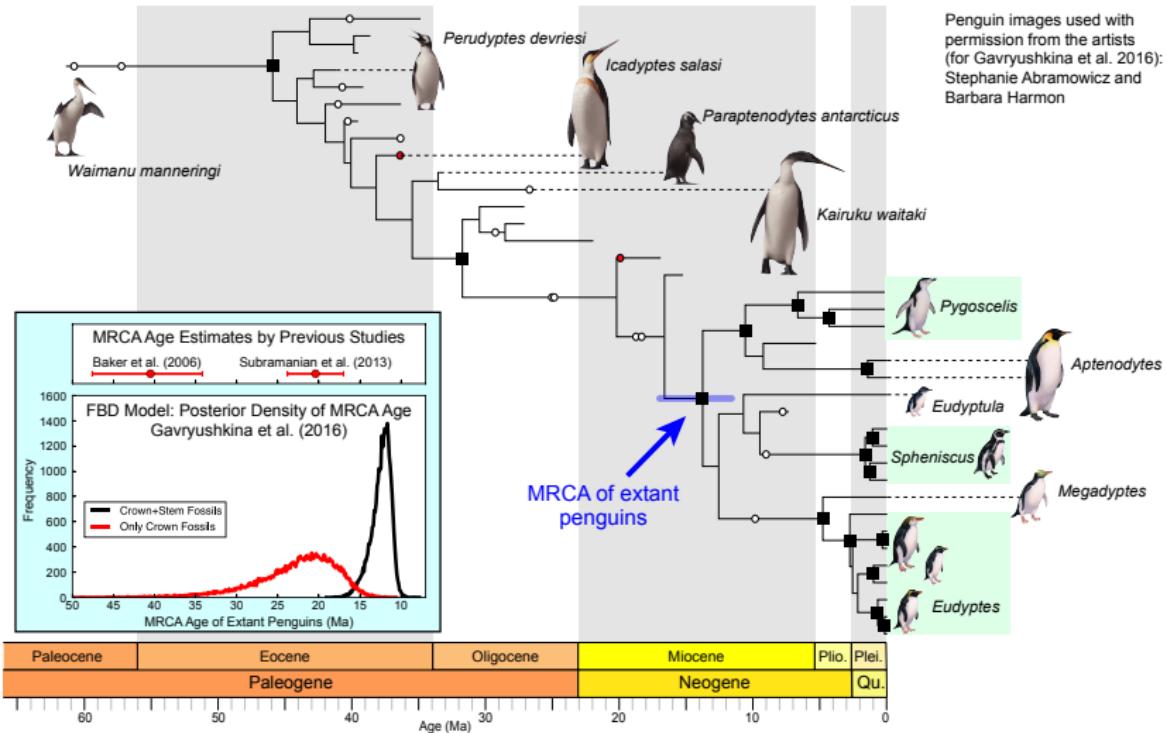
PENGUIN DIVERSITY



PENGUIN DIVERSITY IN DEEP TIME



PENGUIN DIVERSITY IN DEEP TIME



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

