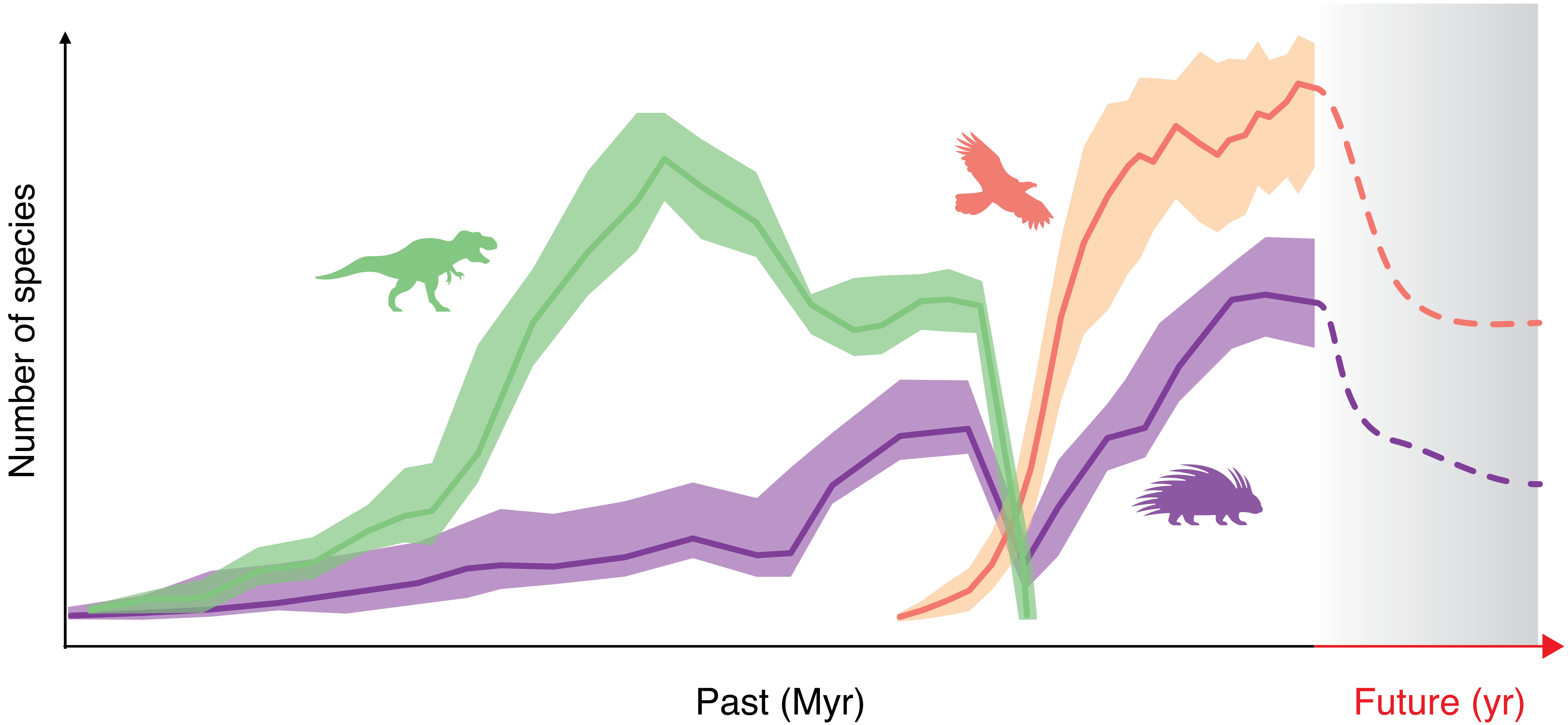
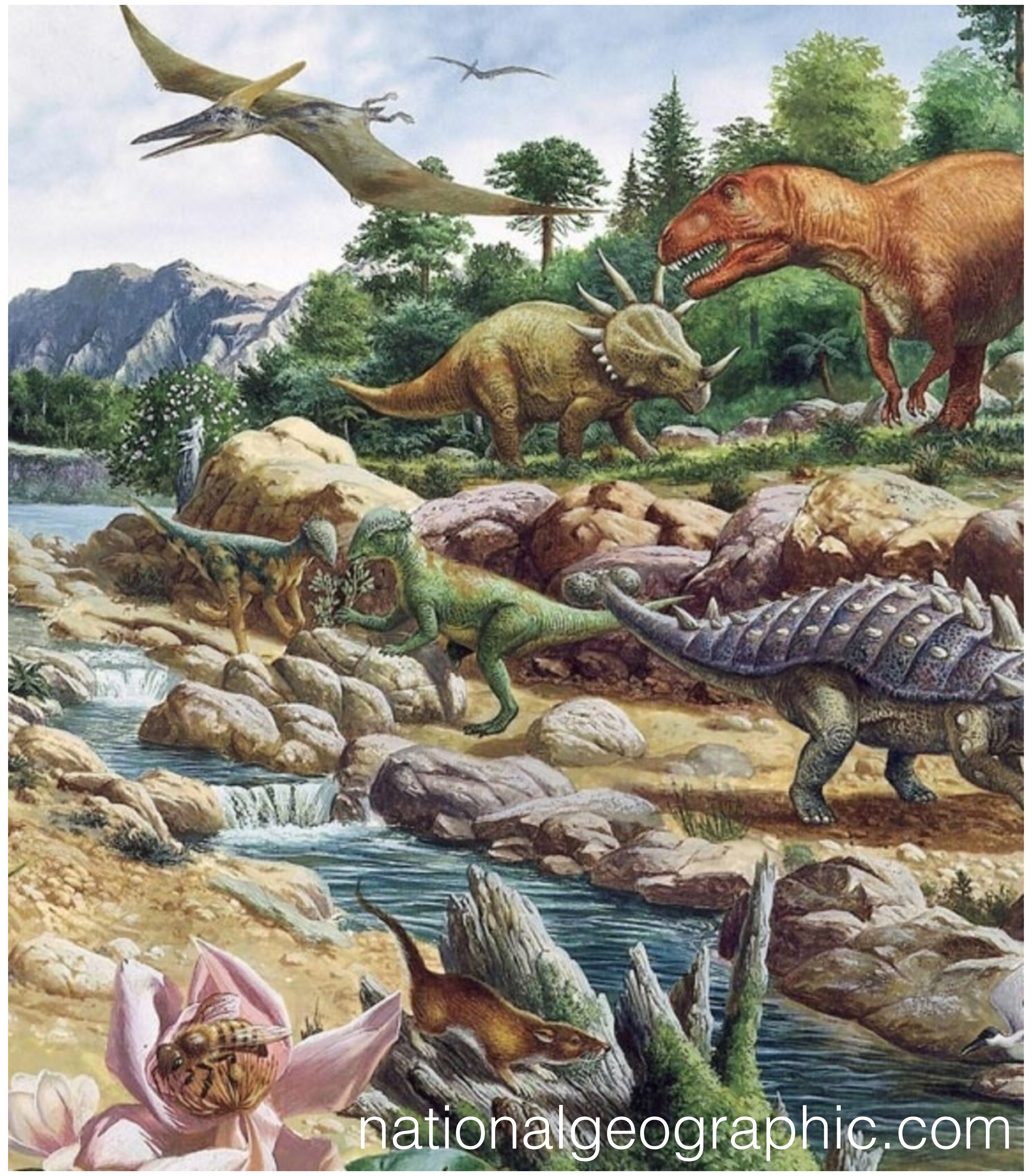
Modeling the evolutionary dynamics of biodiversity



Origin of biodiversity – How old are flowering plants (and other clades)?

> 300,000 species
of flowering plants
exist today

Flowering plants have
been around for a
long time



nationalgeographic.com

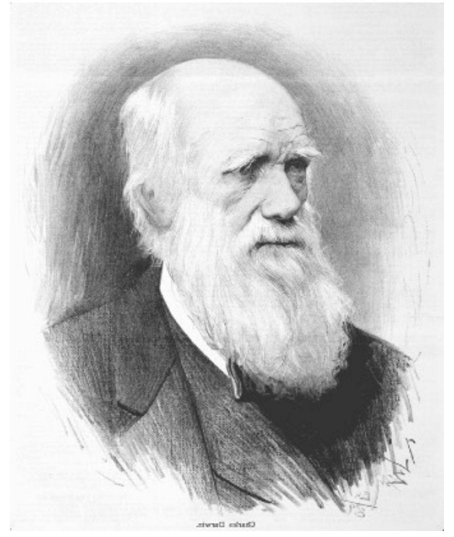
Mismatch between fossil data and molecular clocks

Origin of angiosperms

125 Ma vs 200 Ma

Origin of arthropods

540 Ma vs 800 Ma

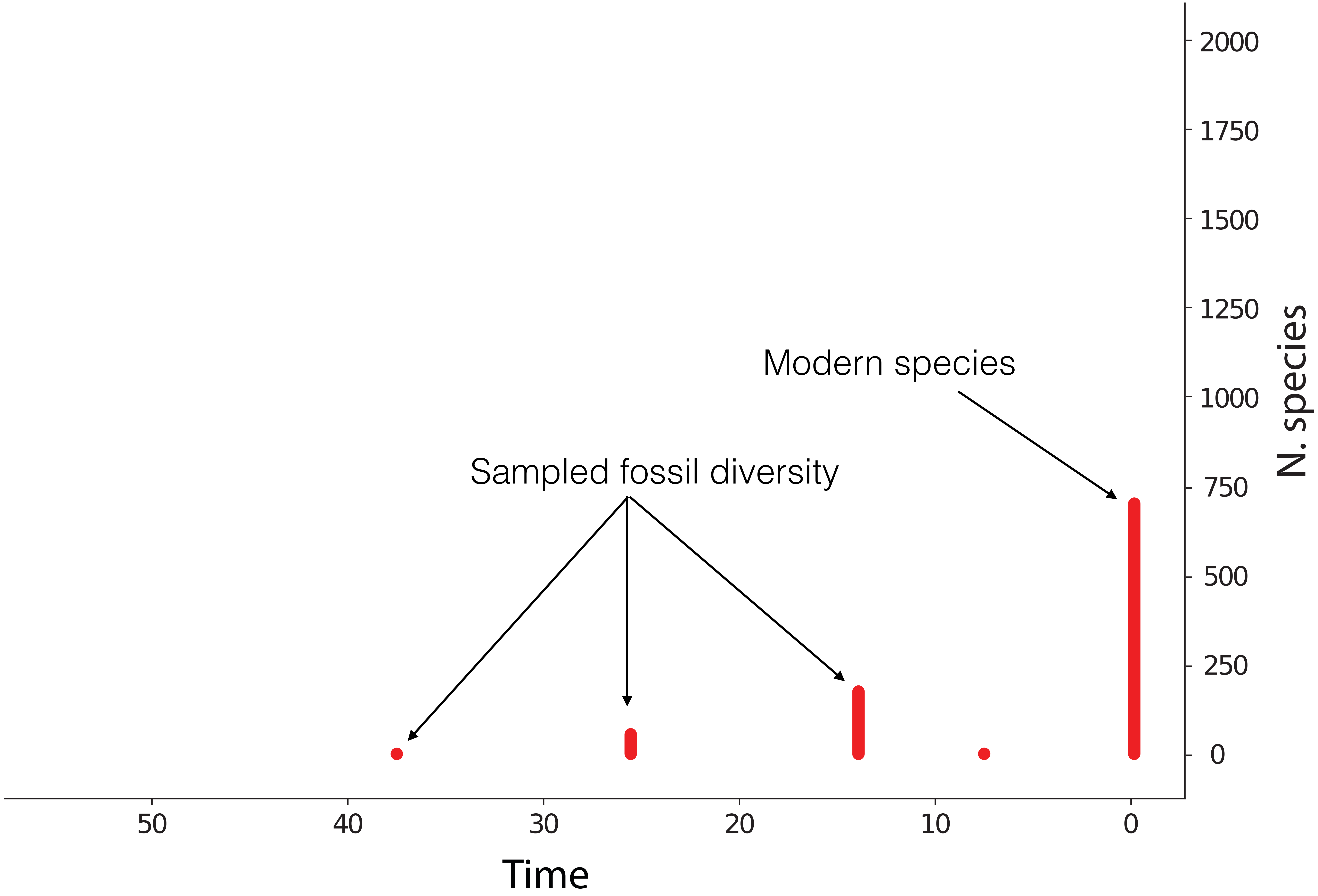


The sudden appearance of
angiosperms in the fossil record
has beed debated ever since
Darwin described it as an
“abominable mystery”

Estimating clade age using a Bayesian Brownian Bridge

Data

- Number of living species
- Number of fossil species in time bins



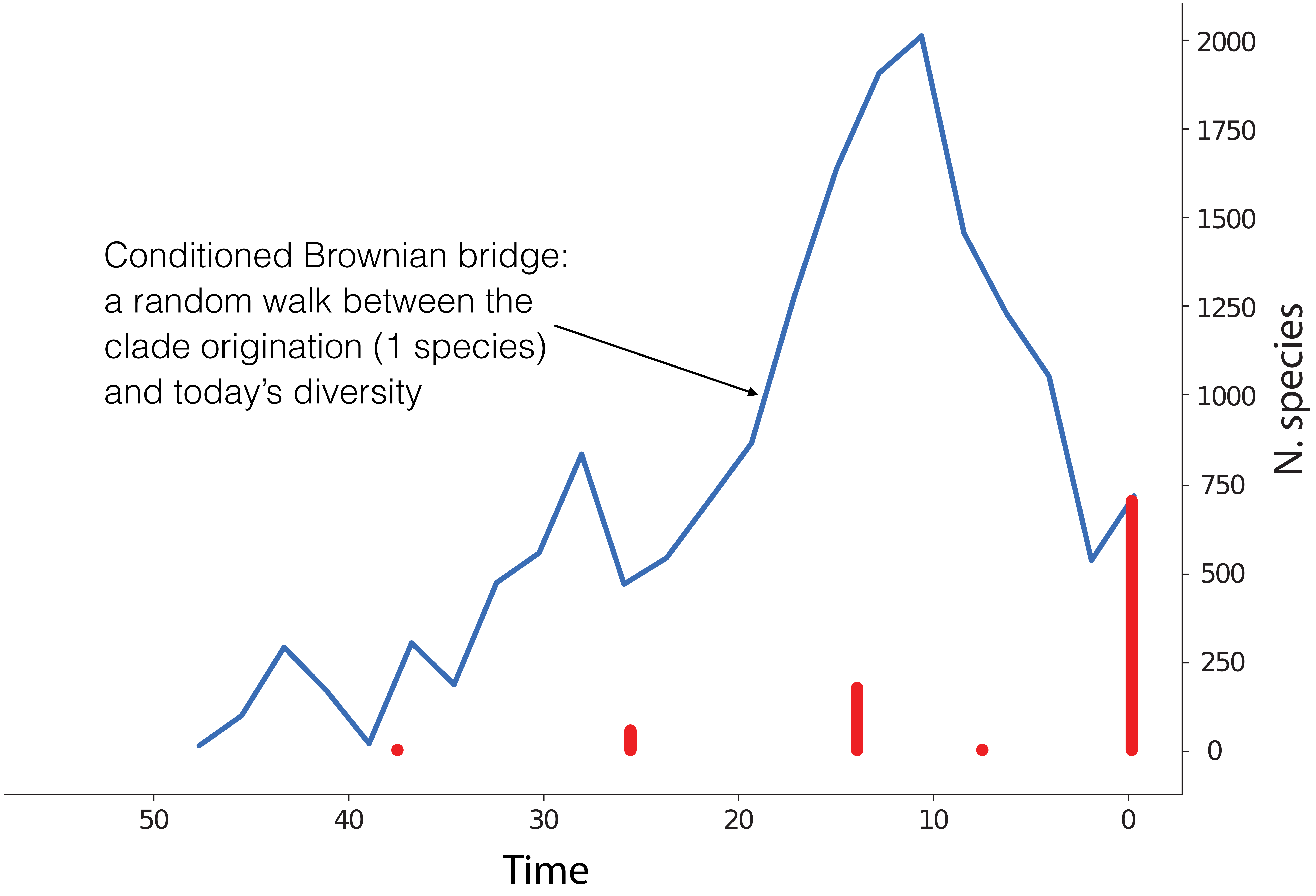
github.com/dsilvestro/rootBBB

Silvestro et al. 2021 Nature Ecol Evol

Diversity trajectories as conditioned Brownian bridges

Data

- Number of living species
- Number of fossil species in time bins


Conditioned Brownian bridge: a random walk between the clade origination (1 species) and today's diversity

github.com/dsilvestro/rootBBB

Silvestro et al. 2021 Nature Ecol Evol

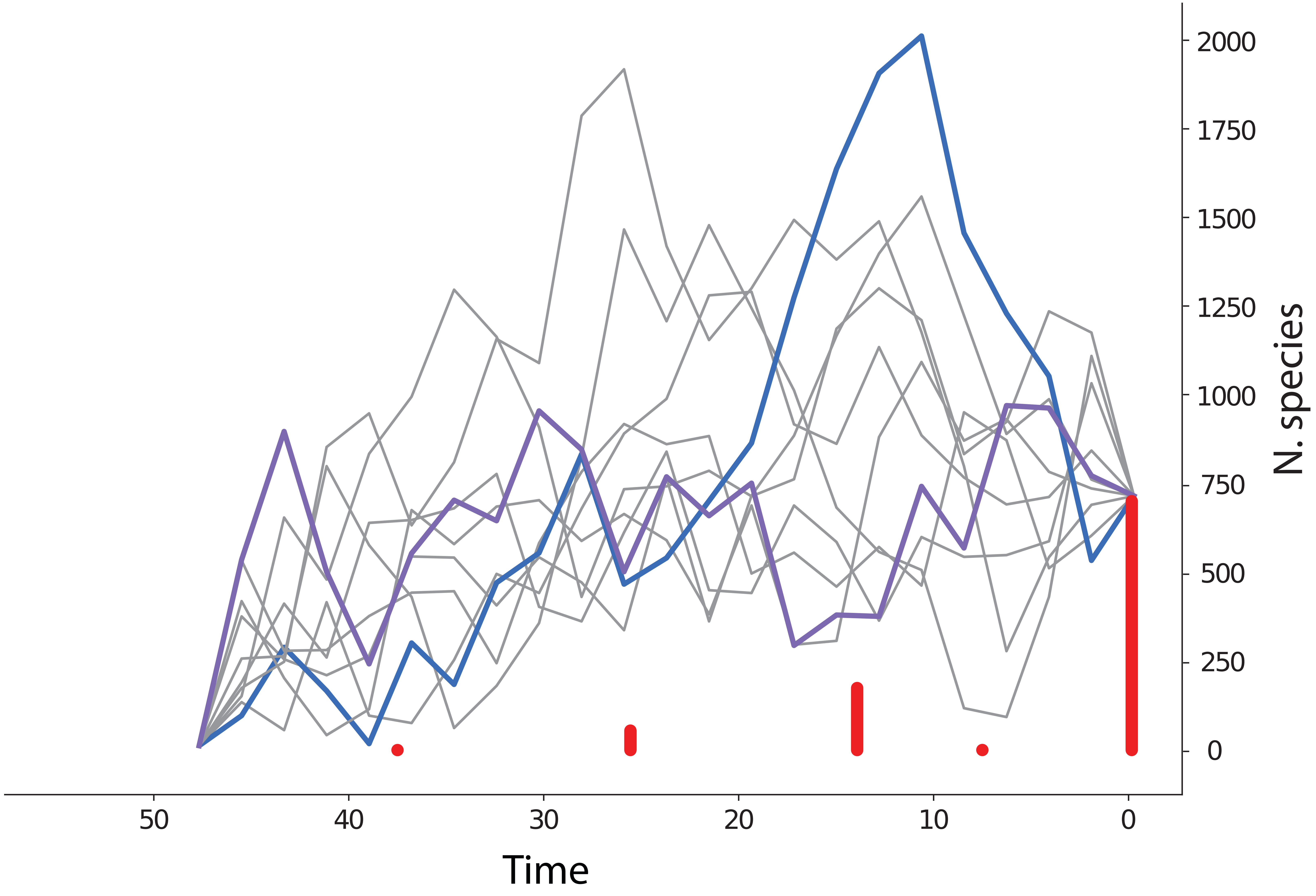
Diversity trajectories as conditioned Brownian bridges

Data

- Number of living species
- Number of fossil species in time bins

Data augmentation

- Integrating across trajectories from a conditioned Brownian bridge



github.com/dsilvestro/rootBBB

Silvestro et al. 2021 Nature Ecol Evol

Diversity trajectories as conditioned Brownian bridges

Data

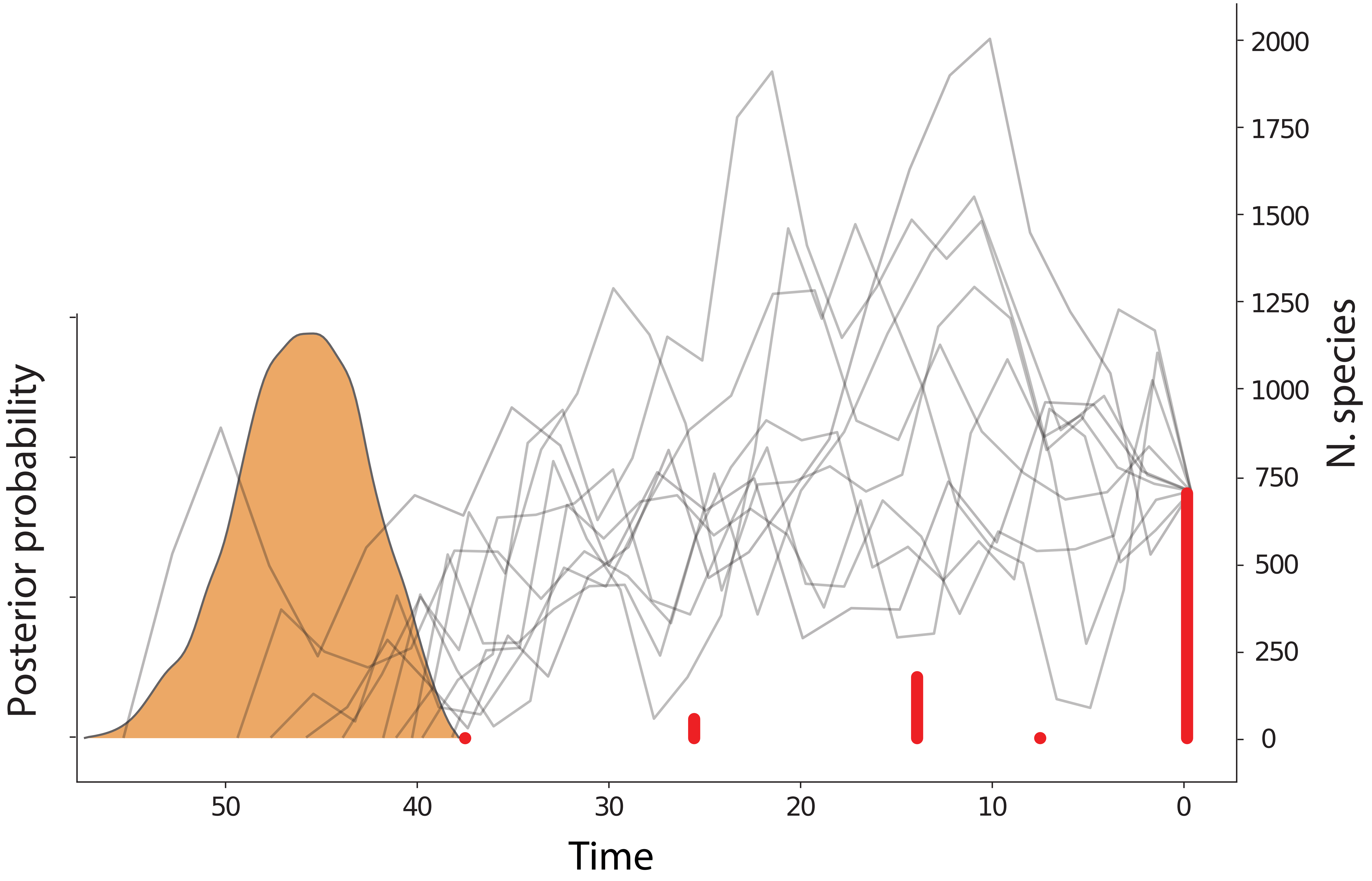
- Number of living species
- Number of fossil species in time bins

Data augmentation

- Integrating across trajectories from a conditioned Brownian bridge

Parameter estimation via MCMC

- Clade age
- Variance of the BB
- Average sampling probability



github.com/dsilvestro/rootBBB

Silvestro et al. 2021 Nature Ecol Evol

Diversity trajectories as conditioned Brownian bridges

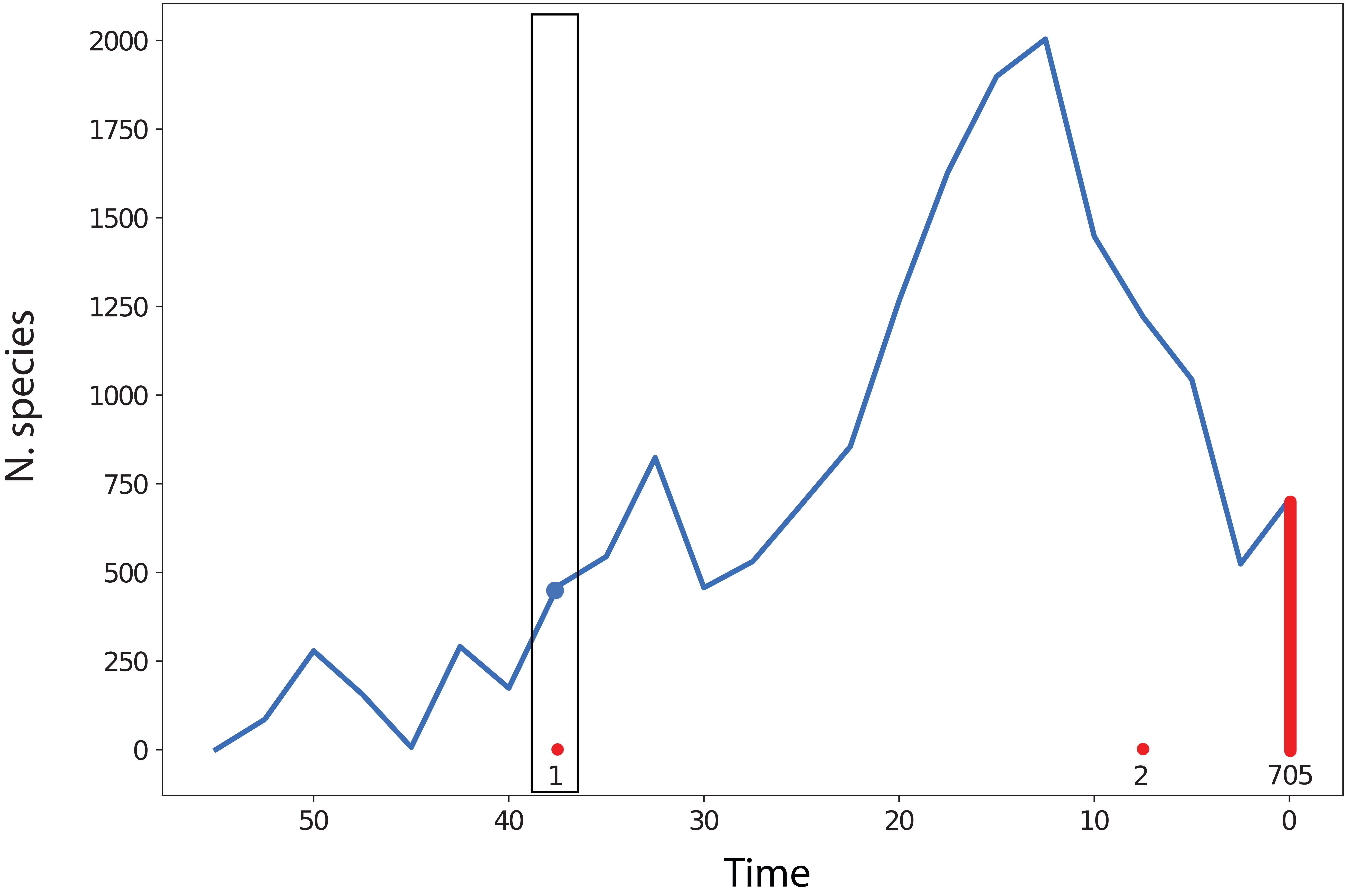
Binomial likelihood

$$P(x_i, d_i|q) = \binom{d_i}{x_i} q^{x_i}(1-q)^{d_i - x_i}$$

The BB is conditioned on:

- being in the positive range in all time bins (i.e. a clade cannot go extinct and reappear again)
- Being greater than (or equal to) the number of fossil species in time bins with fossils

$$P(x_t = 1, d_t = 450|q_{avg})$$



github.com/dsilvestro/rootBBB

Diversity trajectories as conditioned Brownian bridges

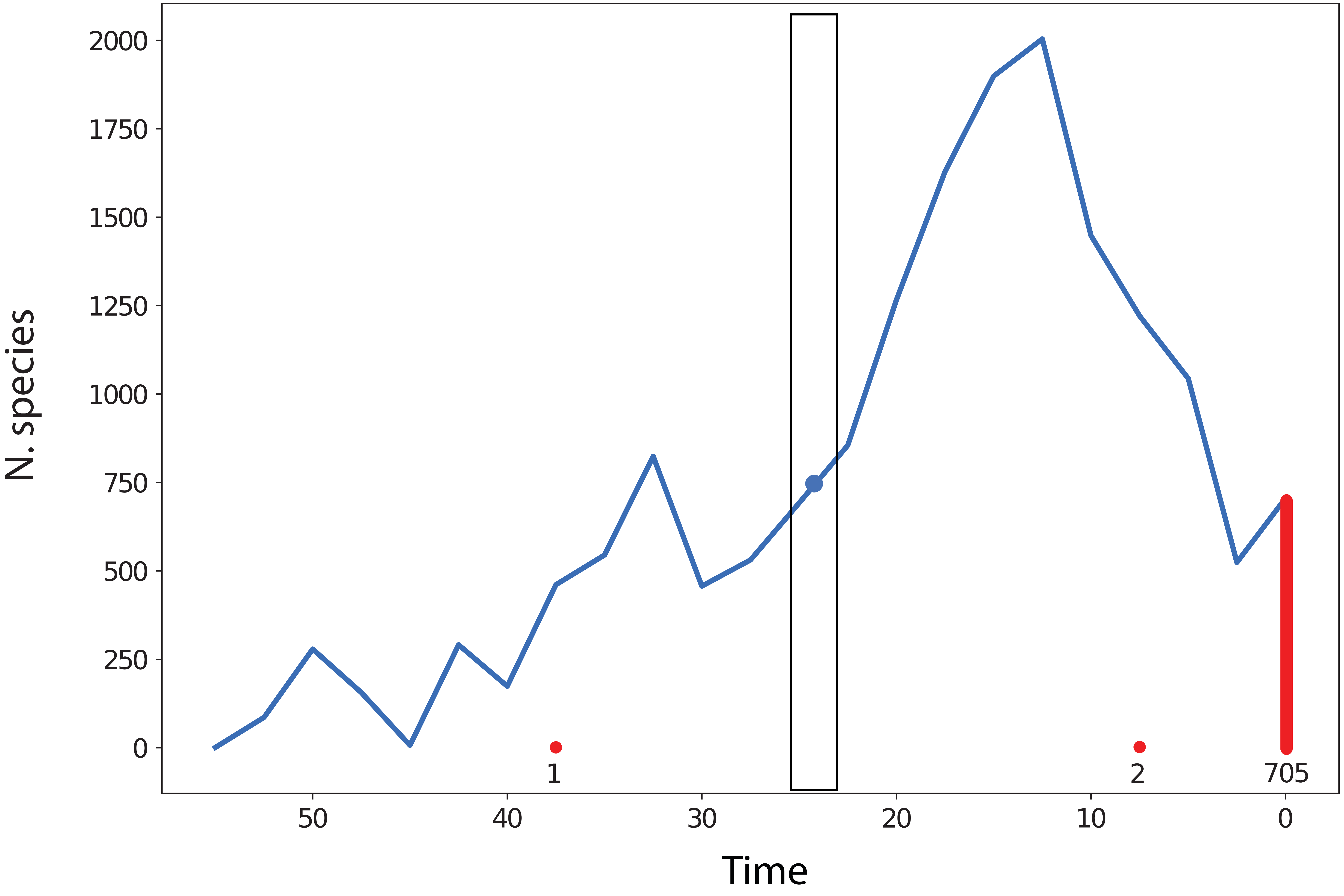
Binomial likelihood

$$P(x_i, d_i | q) = \binom{d_i}{x_i} q^{x_i} (1-q)^{d_i - x_i}$$

The BB is conditioned on:

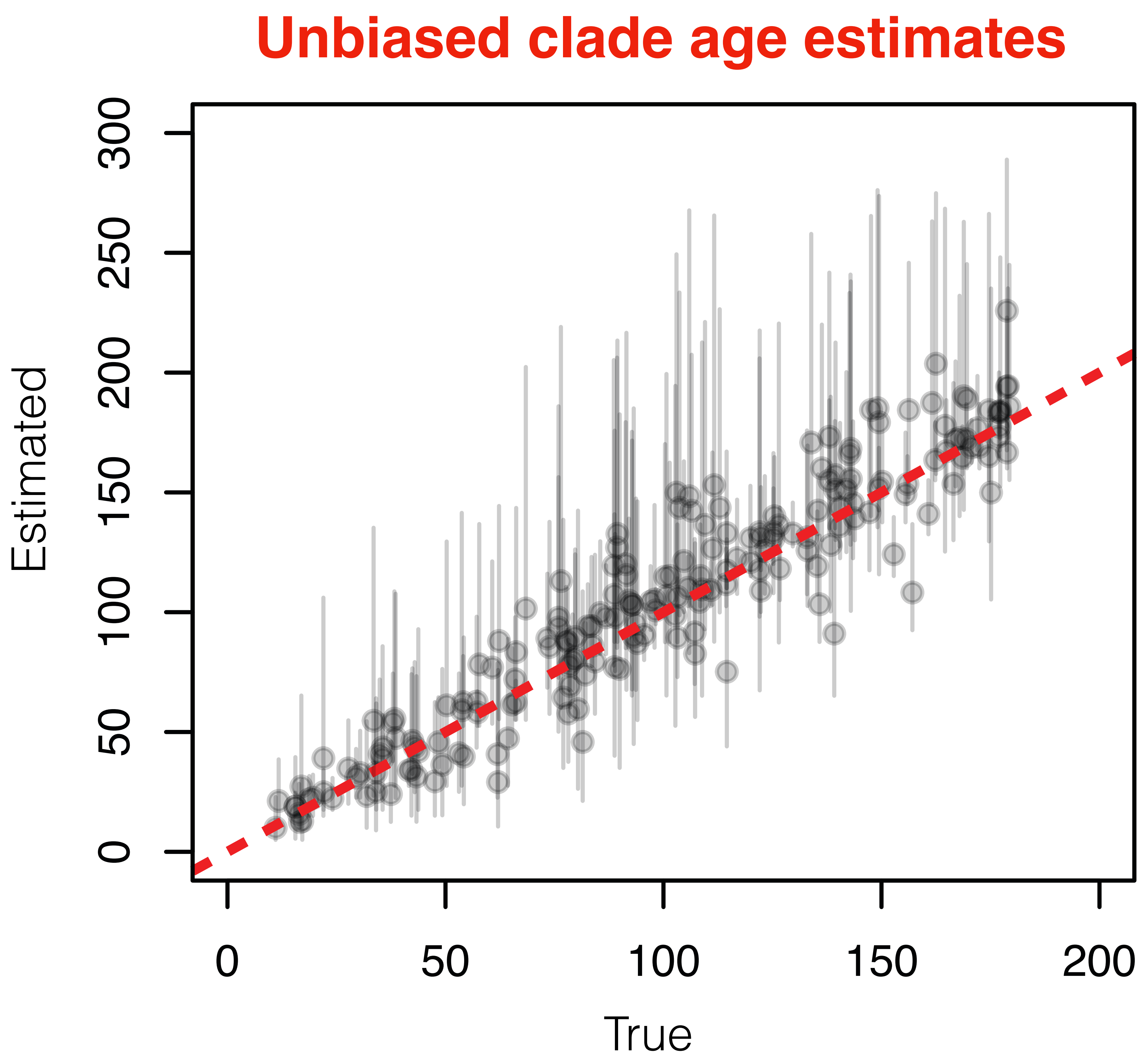
- being in the positive range in all time bins (i.e. a clade cannot go extinct and reappear again)
- Being greater than (or equal to) the number of fossil species in time bins with fossils

$$P(x_t = 0, d_t = 720 | q_{avg})$$

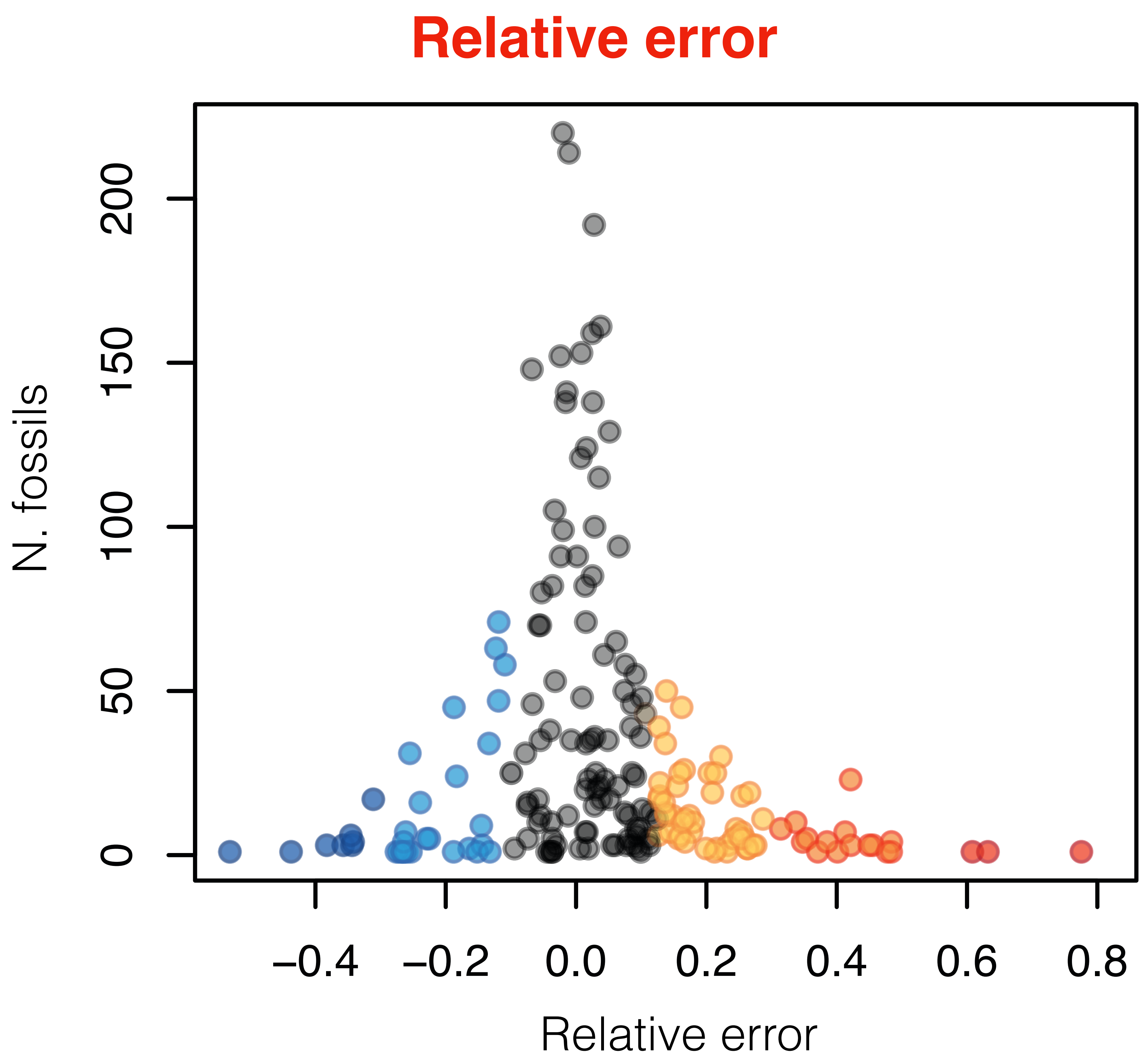


github.com/dsilvestro/rootBBB

Validation through simulations



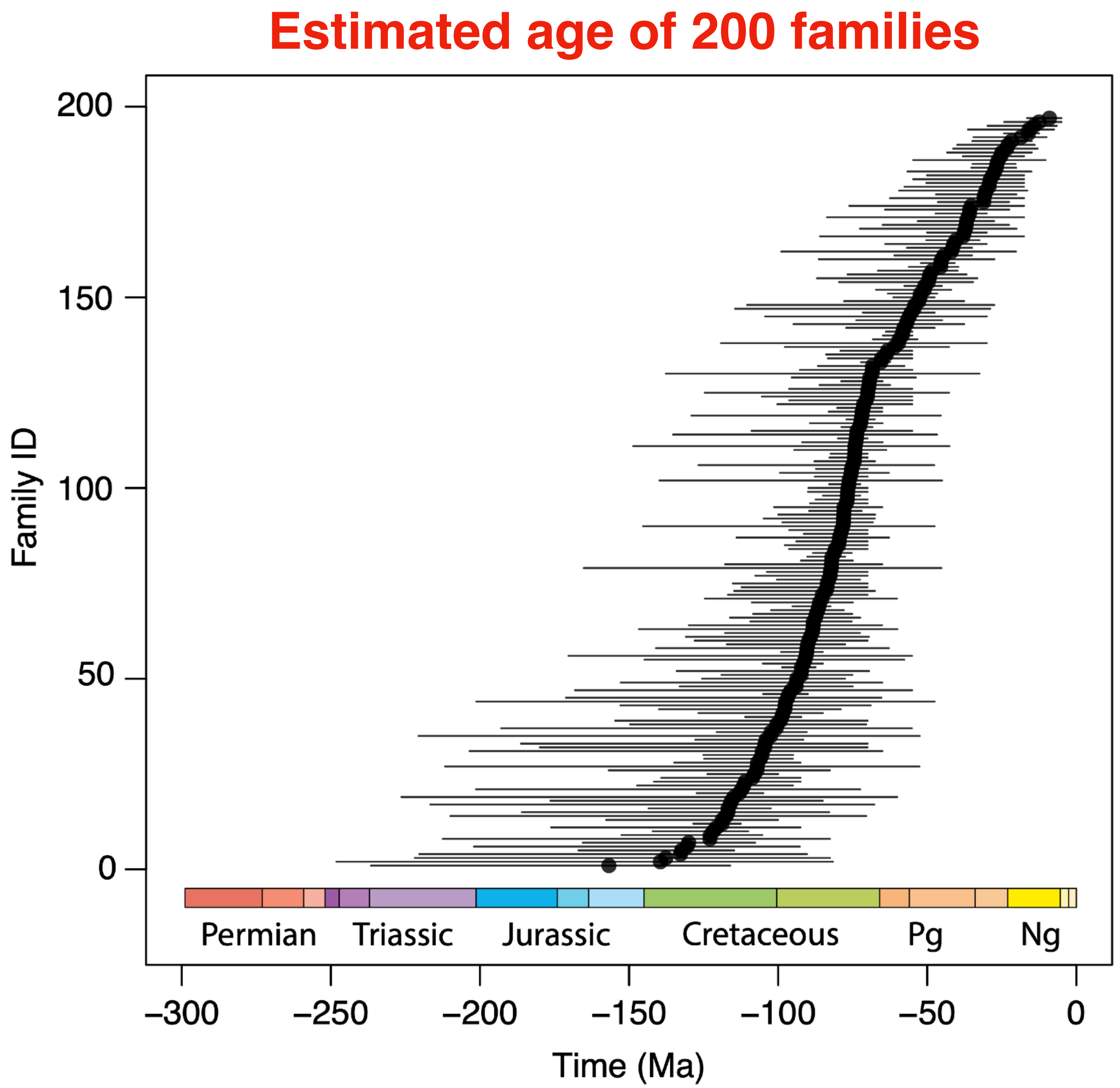
Coverage: 97%

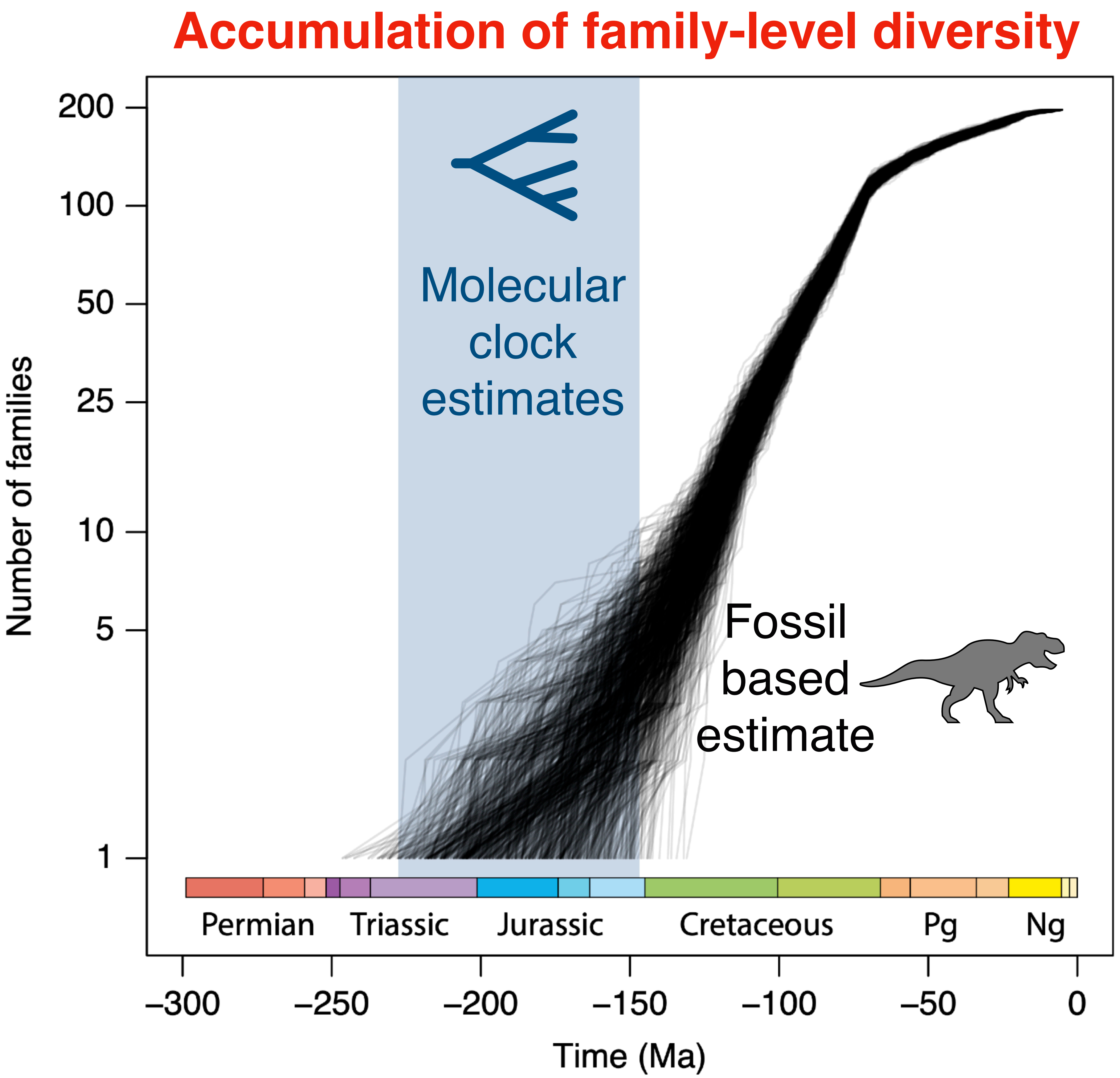


Mean Absolute Percentage Error: 16%

Revisiting the origin of angiosperms using the BBB model

15,408 fossils of Cretaceous and Cenozoic flowering plants

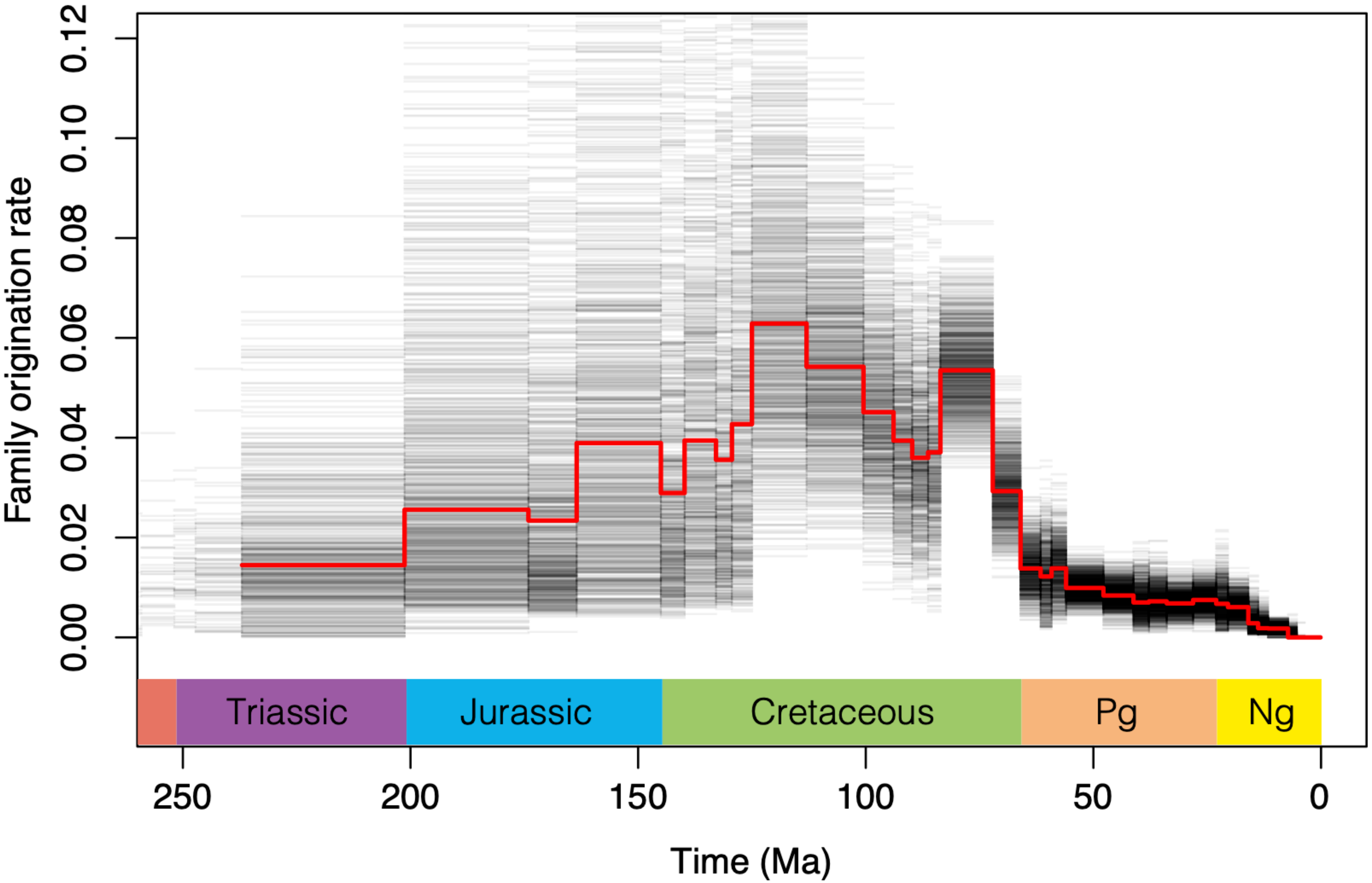




Q Zhang

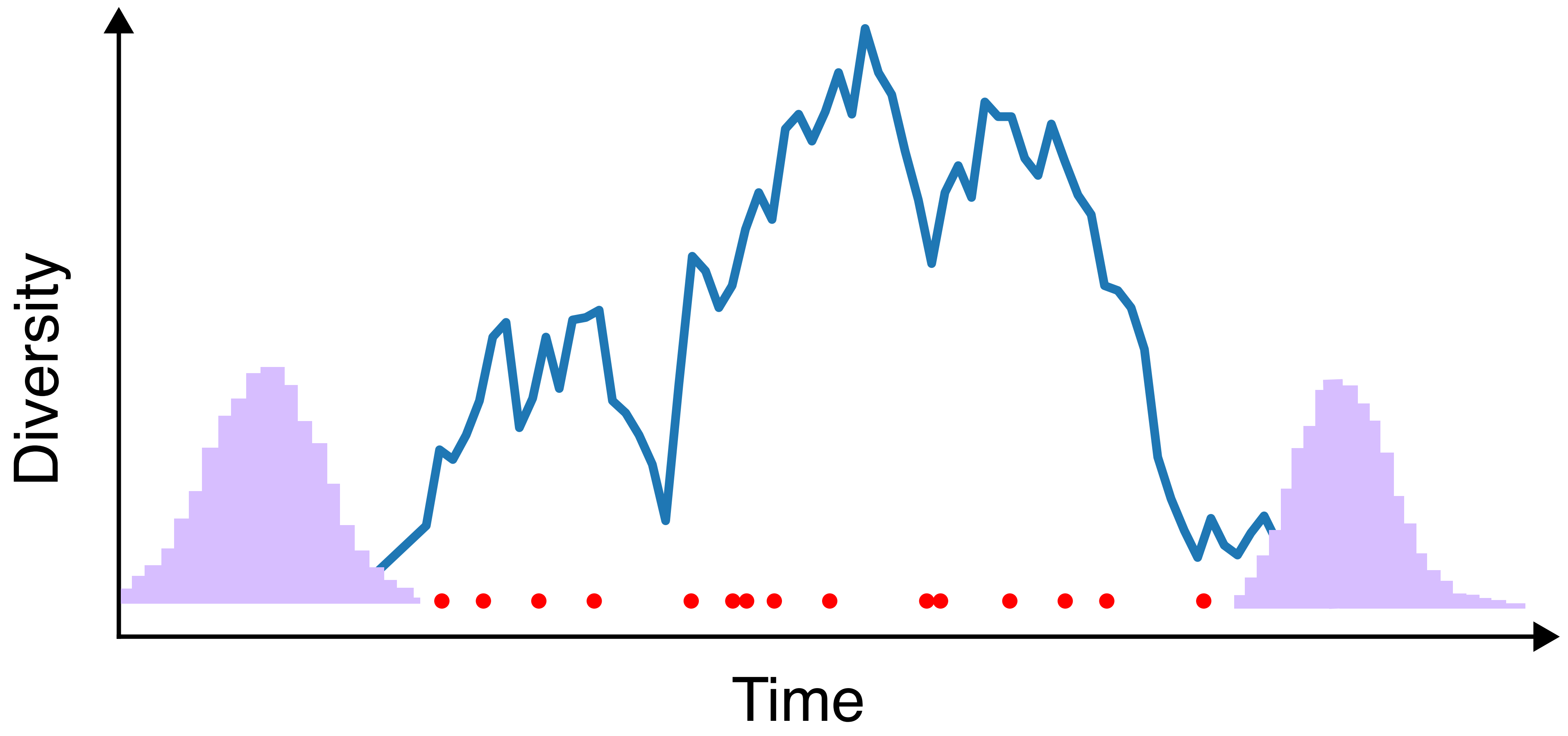
Silvestro et al. 2021 Nature Ecol Evol

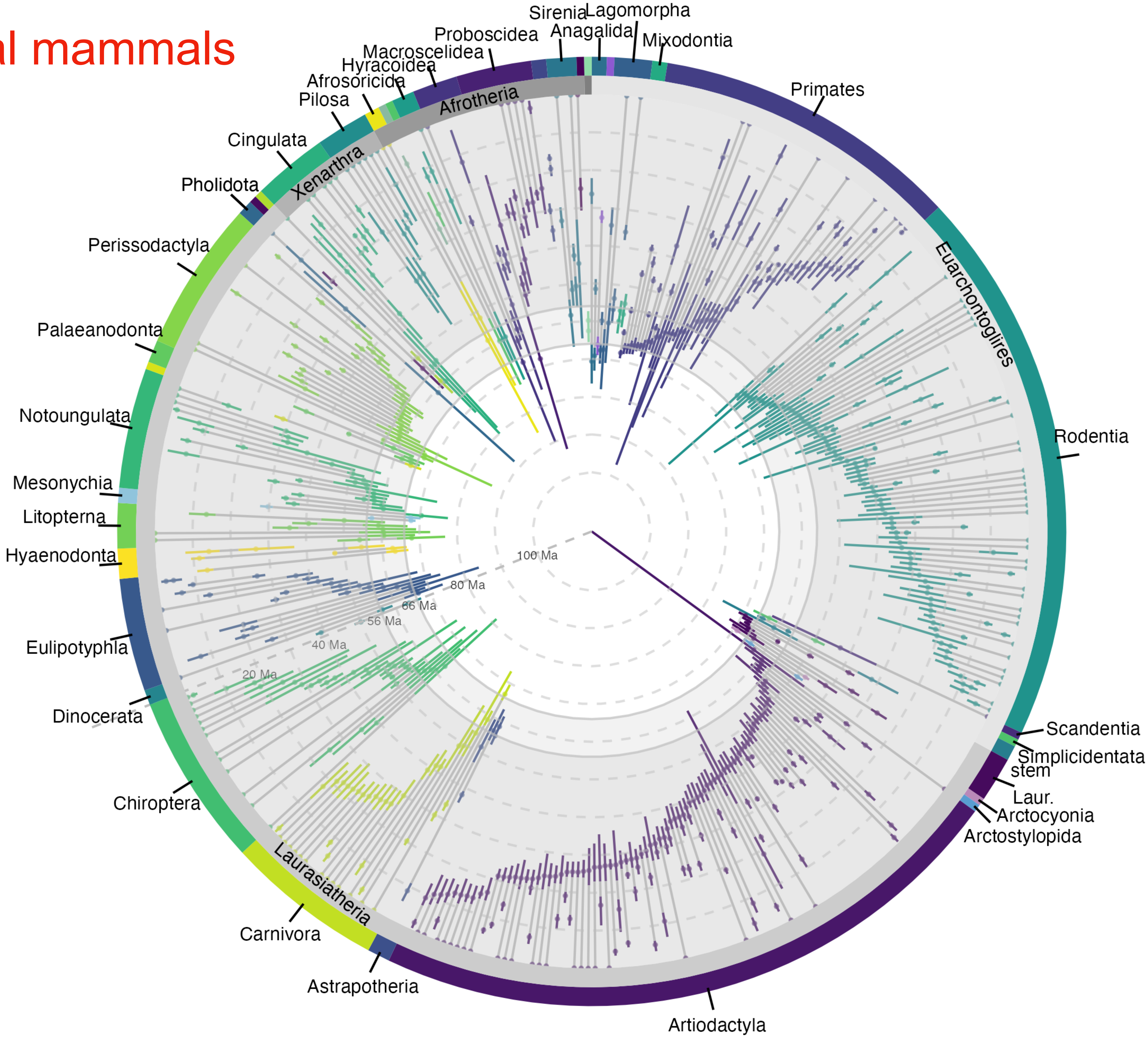
Family-level “empirical” origination rates

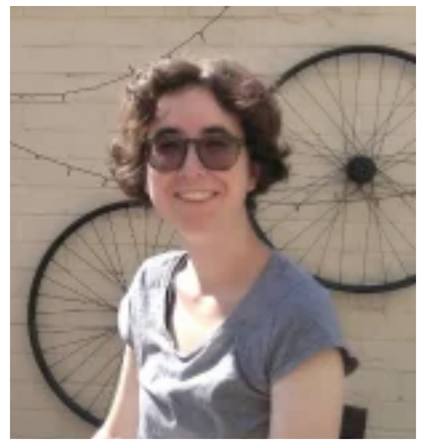


A timescale for the origin of placental mammals

A BBB model for extinct and extant clades







E Carlisle

Carlisle et al. 2023 Current Biology