

Diversification & SSE models

A word about Ancestral State Reconstruction...

Why care about diversification?

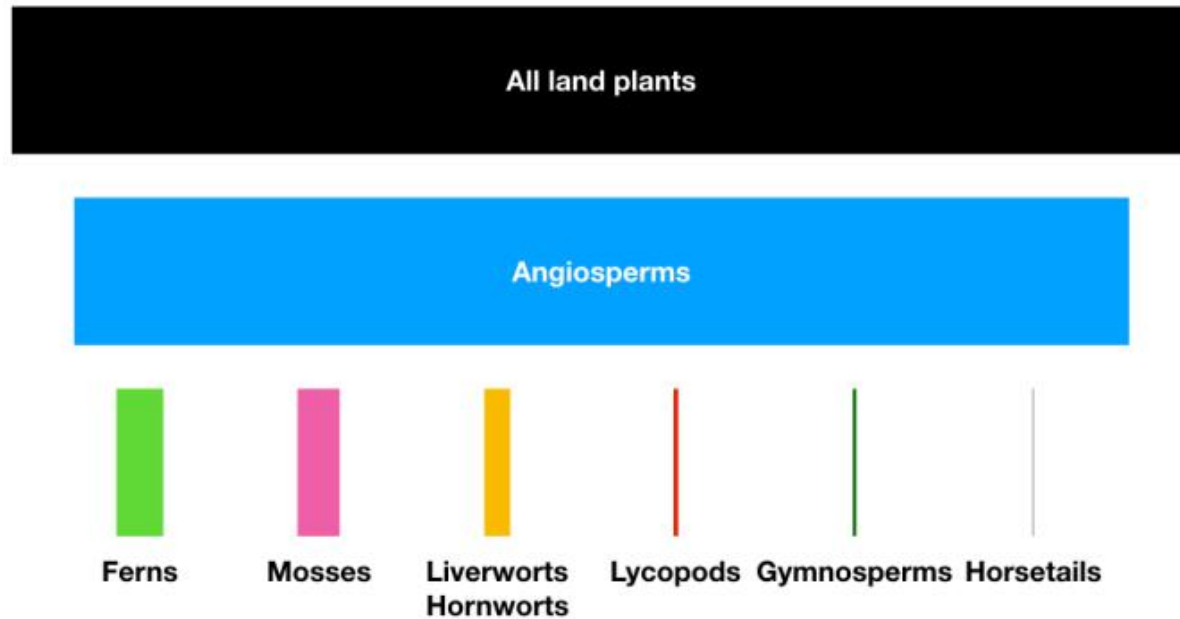


Figure 10.1. Diversity of major groups of embryophytes (land plants); bar areas are proportional to species diversity of each clade. Angiosperms, including some 250,000 species, comprise more than 90% of species of land plants. Figure modified from Crepet and Niklas (2009)

Birth-Death model

Poisson process (like CTMC models) - but events
= Speciation/Extinction rather than substitution

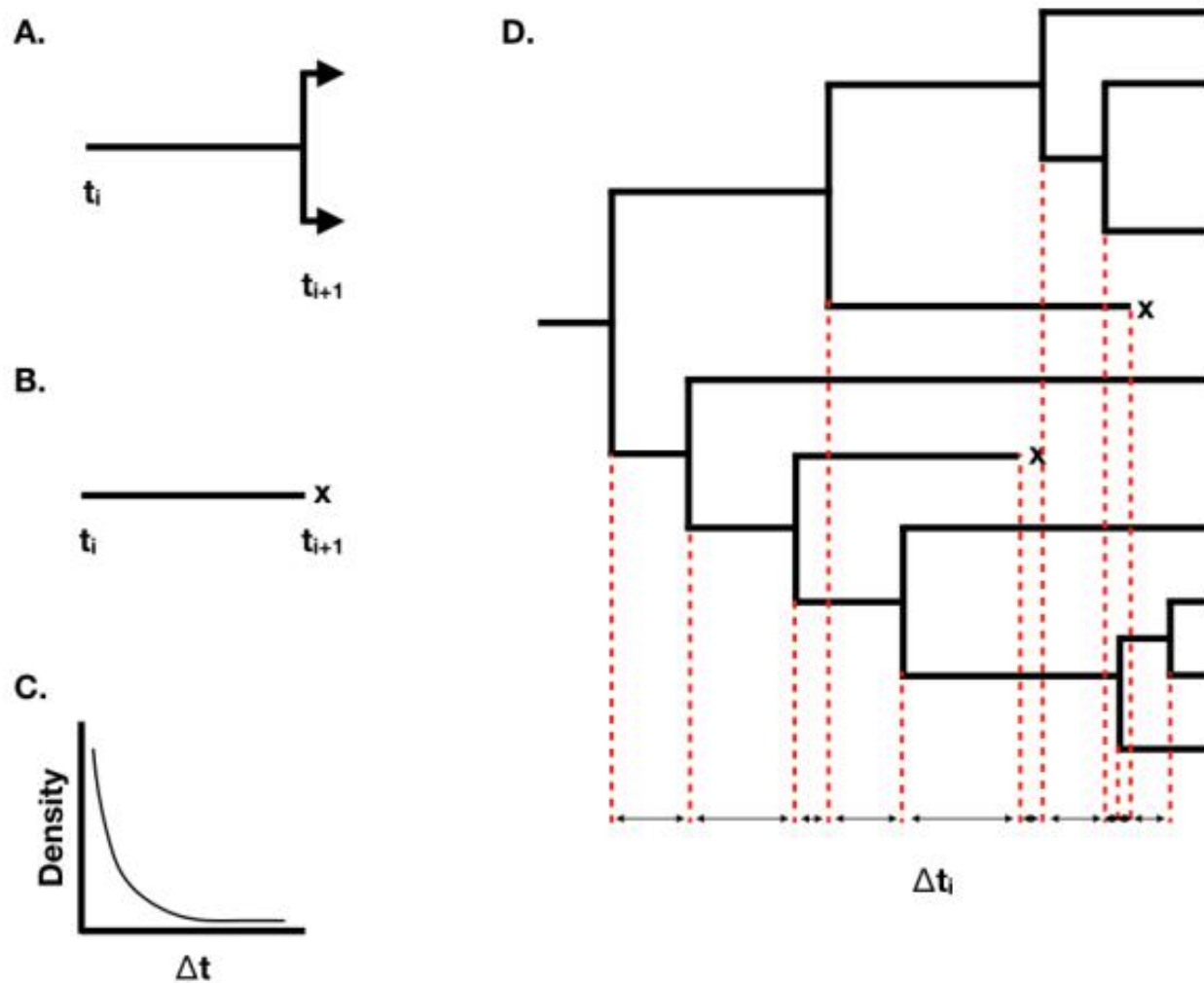
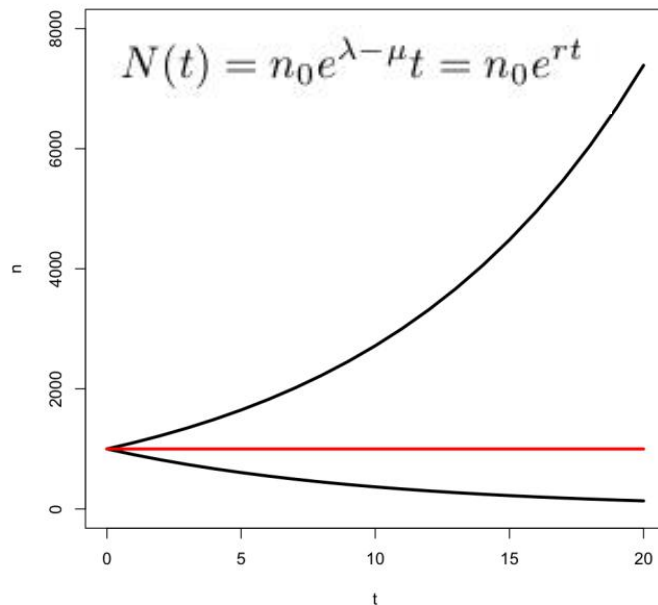
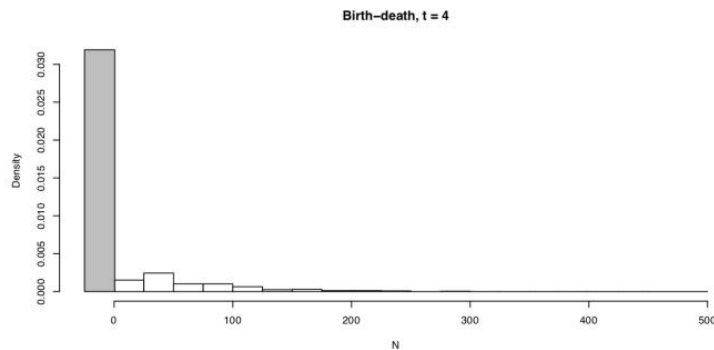
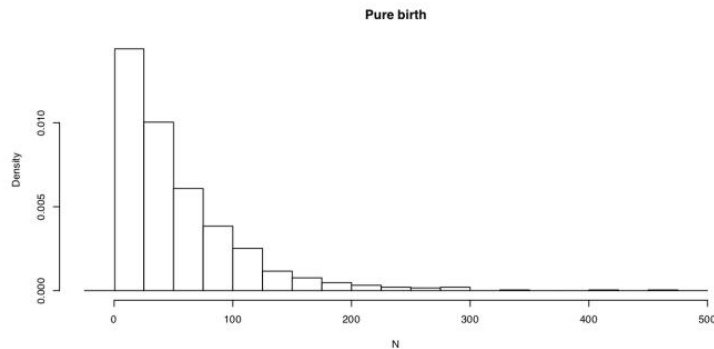


Figure 10.2. Illustration of the basic properties of birth-death models. A. Waiting time to a speciation event; B. Waiting time to an extinction event; C. Exponential distribution of waiting times until the next event; D. A birth-death tree with waiting times, with x denoting extinct taxa.



The Dub's view @DaBauz · Feb 13

You're such a #LIAR4Atheism! Show me where I - "...new species can't arise once every 3 years..."! You, and your #IDIOTS4Atheism friends, @curvemudgen, @ExMissionary & @Wearetheleavers... 1 / 3

cc - @pseudacris & @OregonState



The Dub's view

@DaBauz

Follow

Replying to @DaBauz @pseudacris and 45 others

The #ToE says - "...evolutionary change takes about one million years" - bit.ly/2BstLUz - and Mr. (possibly Professor) Josef Uyeda seems to agree. Using 1m years, per species, we'd need...

WAIT FOR IT...

WAIT FOR IT...

$5 \cdot 10^{15}$ - Five Quadrillion years... HA!

2 / 3



Lasting evolutionary change takes about one million years

CORVALLIS, Ore. - In research that will help address a long-running debate and apparent contradiction between short- and long-term evolutionary change, scientists have discovered that...

today.oregonstate.edu

11:01 AM - 13 Feb 2018

1 Like



Diversification vs. Turnover

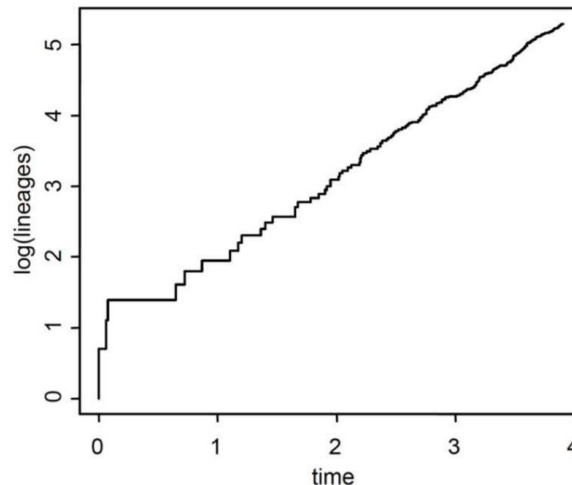
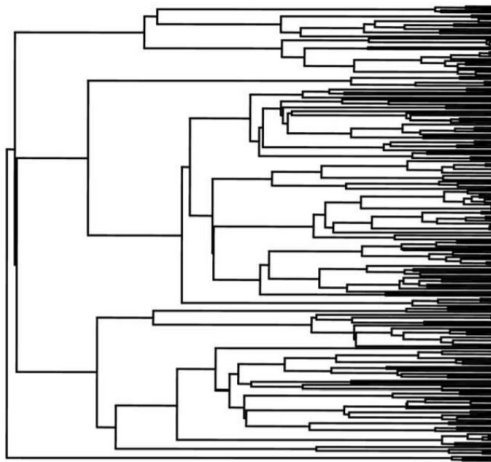
λ = Speciation rate

μ = Extinction rate

$\lambda - \mu$ = Diversification rate

μ/λ = Turnover rate

Yule process/Pure Birth model $\rightarrow \mu = 0$



Under pure birth,
lineages accrue
log-linearly with time

A few notes...

Speciation and extinction correlated in nature,
and in estimation

Estimating diversification and turnover often
easier

Can we even estimate extinction from extant-
only data?

Lineage through time plots

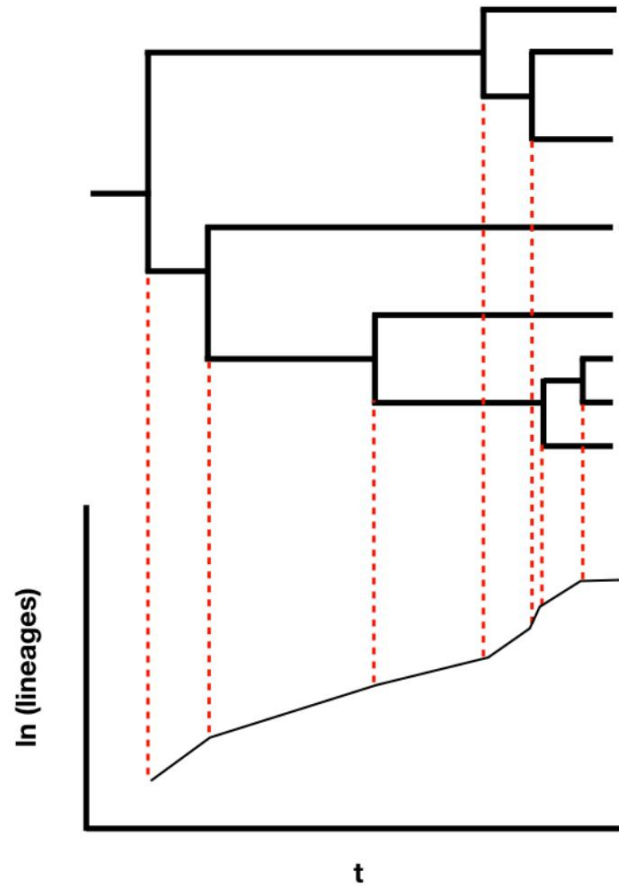
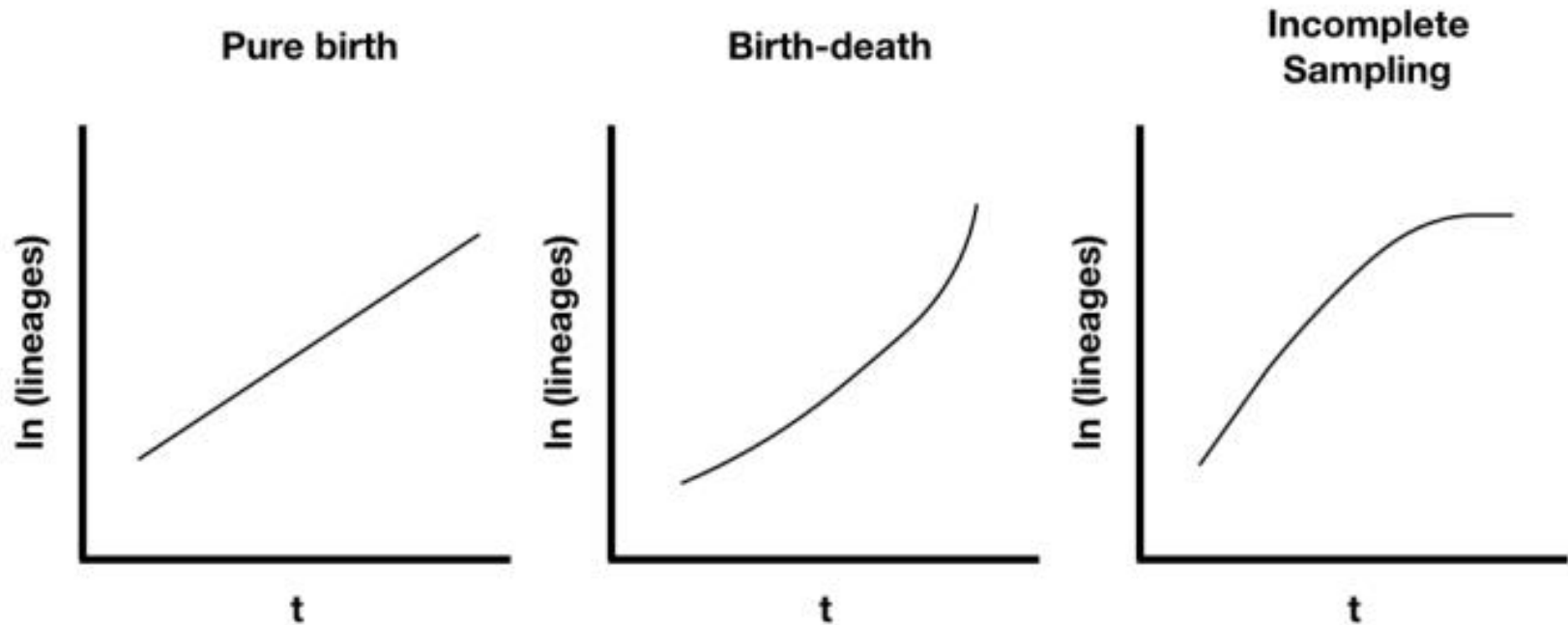


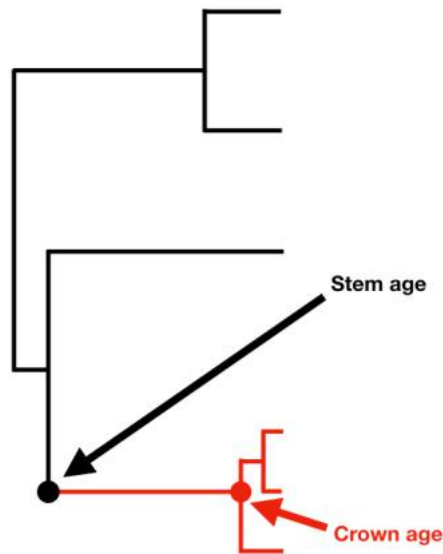
Figure 10.9. Lineage-through-time plot.

Lineage through time plots



Estimating diversification

Net diversification methods using clade age & diversity (can be used on unresolved trees)



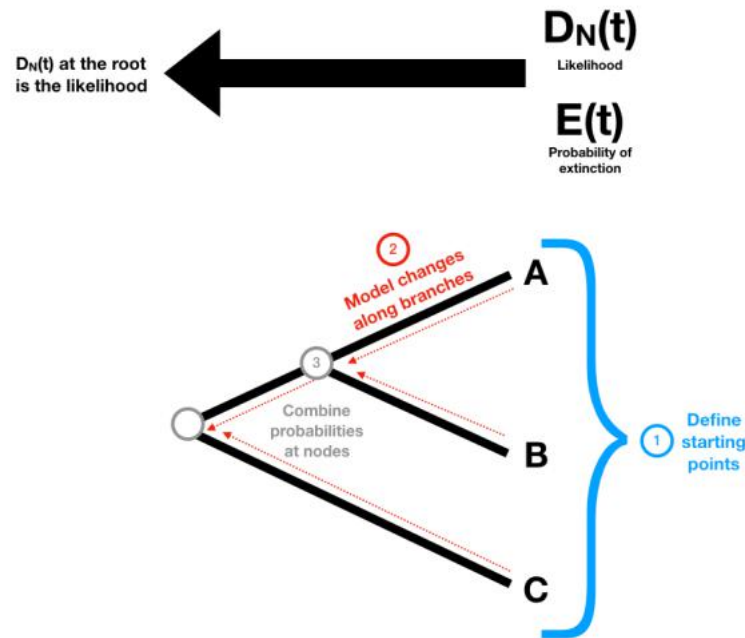
$$\hat{r} = \frac{\ln(n)}{t_{stem}}$$

$$\hat{r} = \frac{\ln(n) - \ln(2)}{t_{crown}}$$

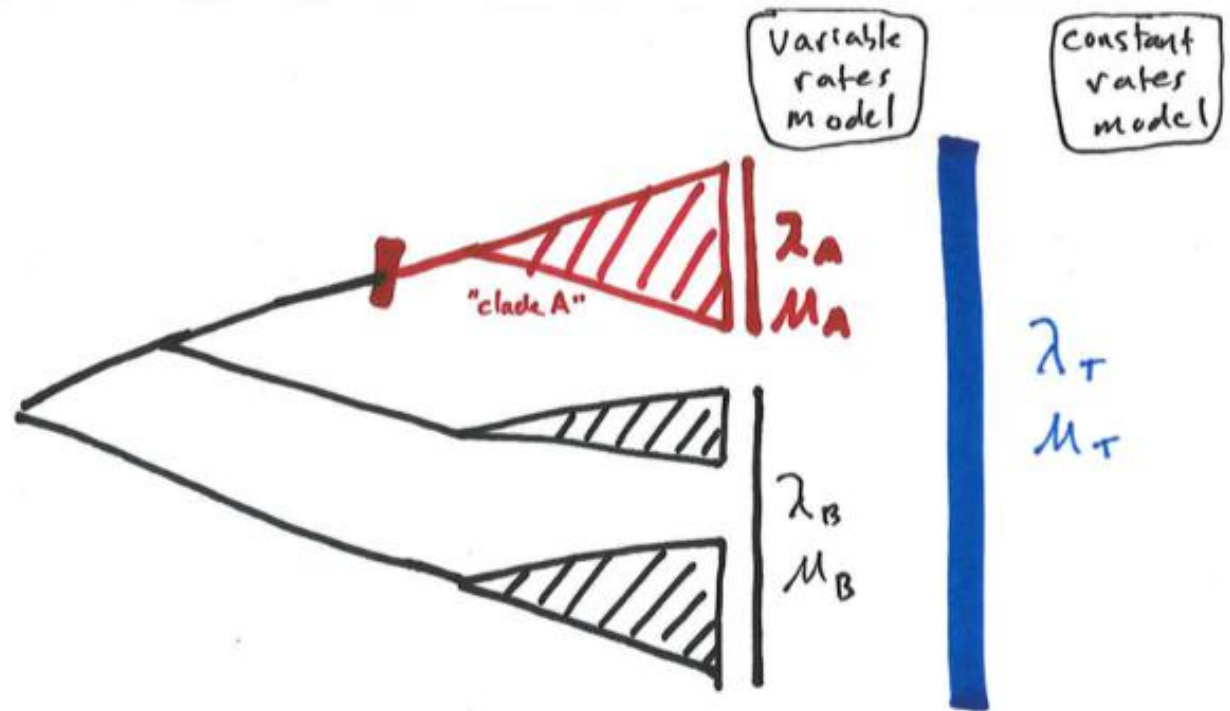
Magallon & Sanderson, 2001

Estimating diversification

Using the distribution of branch lengths/waiting times (requires whole tree or partial tree and assumptions about sampling)



Modeling heterogeneity in diversification



Hypothesis
testing

Sampling

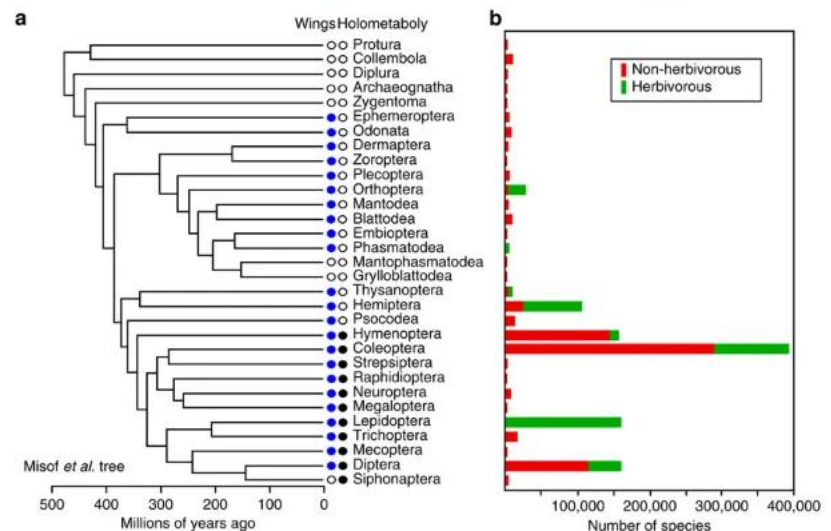
Usually we don't have random sampling of taxa from a phylogeny

Supply each tip with # of taxa it represents

Or

Assume all sampling is random, & provide sampling fraction

Figure 1: Phylogenies of insect orders and distributions of species richness and herbivory and other traits among clades.



Changing diversification over time

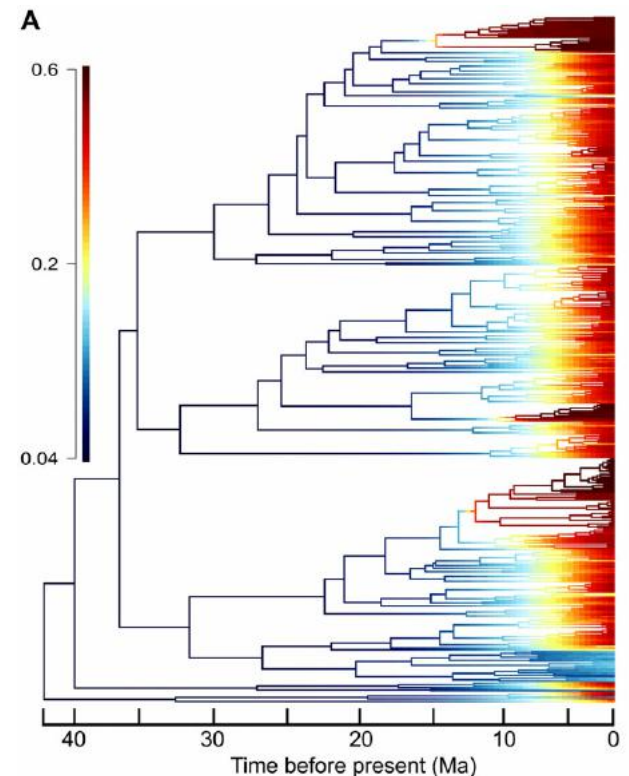
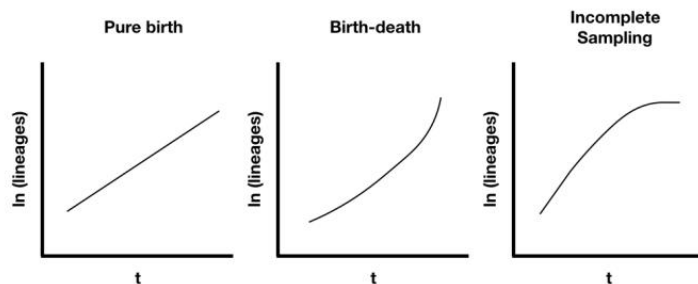
Does diversification increase or decrease toward the present?

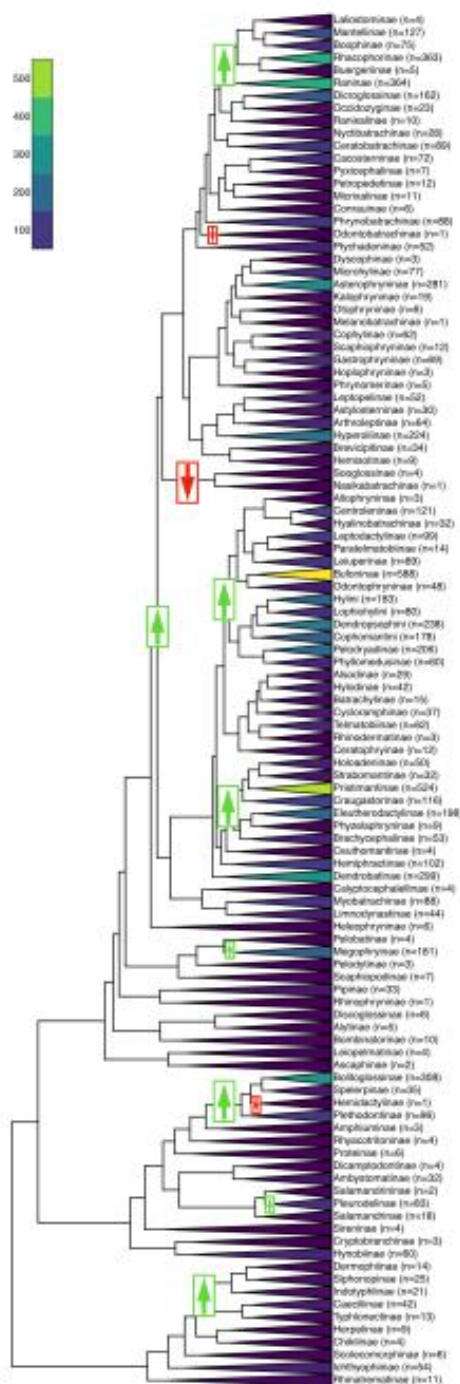
Biases...

“Pull of the present”

“Push of the past”

Protracted speciation





Finding evolutionary shifts in diversification

Bayesian - BAMM (But see controversies)

Likelihood (Stepwise AIC) - MEDUSA (but see problems with stepwise AIC)

Easy to do with a phylogeny, but
controversies abound

Can we ACTUALLY estimate extinction?

ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2009.00926.x

Frequently estimated to be 0, diversification rates often viewed as more reliable, but....

EXTINCTION RATES SHOULD NOT BE ESTIMATED FROM MOLECULAR PHYLOGENIES

Daniel L. Rabosky^{1,2,3,4}

¹Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853

²Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, New York 14850

³Department of Integrative Biology, University of California, Berkeley, California 94720

⁴E-mail: drabosky@berkeley.edu

BRIEF COMMUNICATION

doi:10.1111/evo.12614



Extinction can be estimated from moderately sized molecular phylogenies

Jeremy M. Beaulieu^{1,2} and Brian C. O'Meara³

¹National Institute for Biological and Mathematical Synthesis, University of Tennessee, Knoxville, Tennessee 37996

²E-mail: jbeaulieu@nimbios.org

³Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996

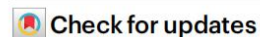
Extant timetrees are consistent with a myriad of diversification histories

<https://doi.org/10.1038/s41586-020-2176-1>

Received: 14 September 2019

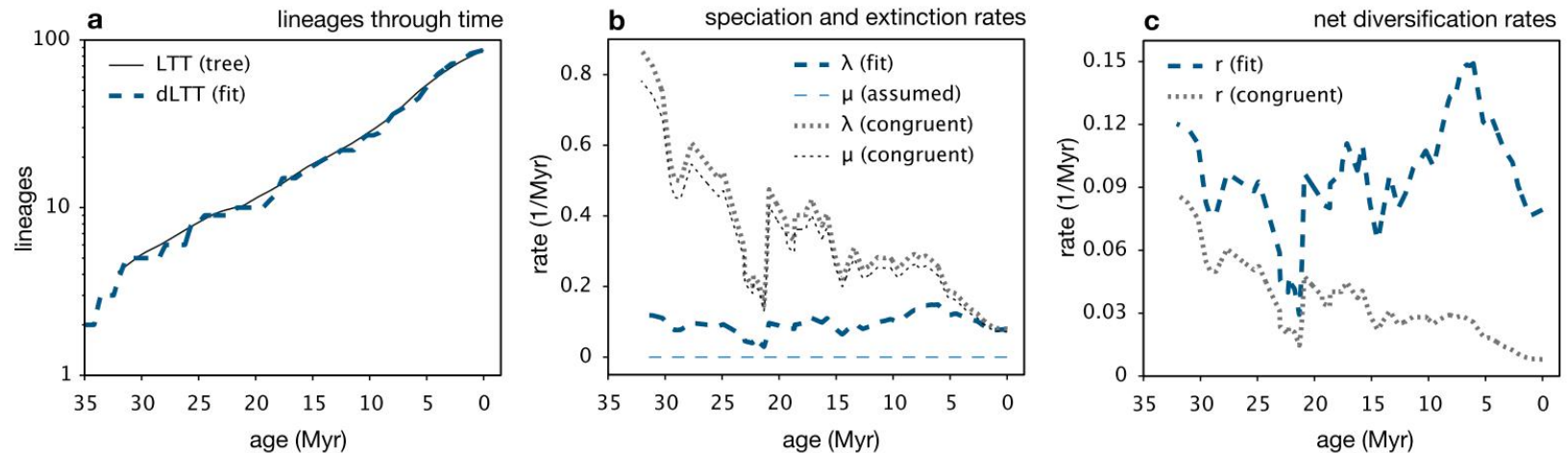
Accepted: 10 March 2020

Published online: 15 April 2020



Stilianos Louca^{1,2} & Matthew W. Pennell^{3,4}

Time-calibrated phylogenies of extant species (referred to here as ‘extant timetrees’) are widely used for estimating diversification dynamics¹. However, there has been considerable debate surrounding the reliability of these inferences^{2–5} and, to date, this critical question remains unresolved. Here we clarify the precise information that can



Extended Data Fig. 3 | Previous studies are likely to have over-interpreted phylogenetic data. Time-dependent birth–death model fitted to a nearly complete extant timetree of the Cetacea, under the assumption of extinction rates of zero ($\mu = 0$), compared to a congruent model in which the extinction

rate is close to the speciation rate ($\mu = 0.9\lambda$). **a**, LTT of the tree, compared to the dLTT predicted by the two models. **b**, Speciation rates (λ) and extinction rates (μ) of the two models. **c**, Net diversification rates ($r = \lambda - \mu$) of the two models. Further details are provided in Supplementary Information section S.4.

What's the deal with the BAMM fights?

BAMM's likelihood is wrong - model tries to evaluate diversification shifts as a Poisson Point Process

Such a process however, produces shifts on branches that then increase in extinction rate; biasing the sample of what is observed

Accounting for those “unobserved shifts” is not done correctly, and we don't have a method that does it correctly - hard problem.

But....results probably correlated with the truth/good heuristic

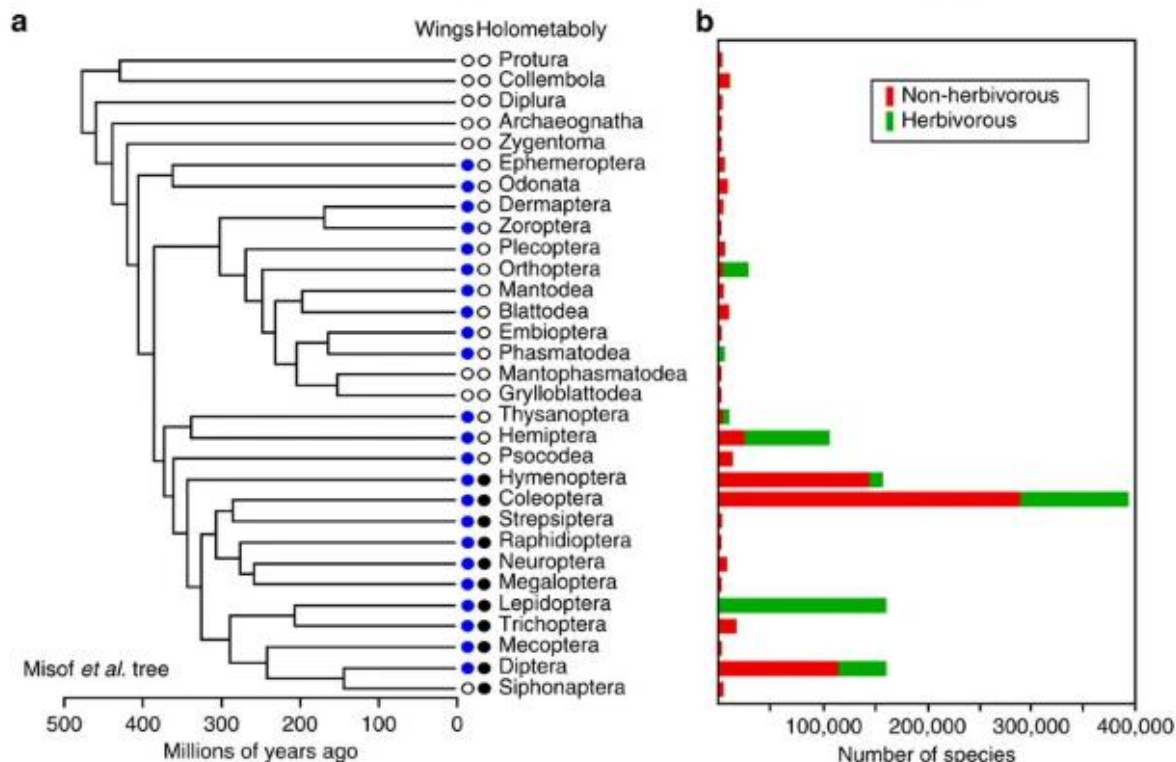
Other active areas of research:

- Incorporating fossils & sampling
- Diversity-dependent diversification
- Time-varying diversification
- Protracted Speciation
- Neutral theory (Hubbell) from ecology
- Trait-dependent diversification

Trait-dependent diversification

One approach: Estimate net div in a bunch of clades, estimate trait (prop/freq) in same clades, see if correlated (w/PGLS)

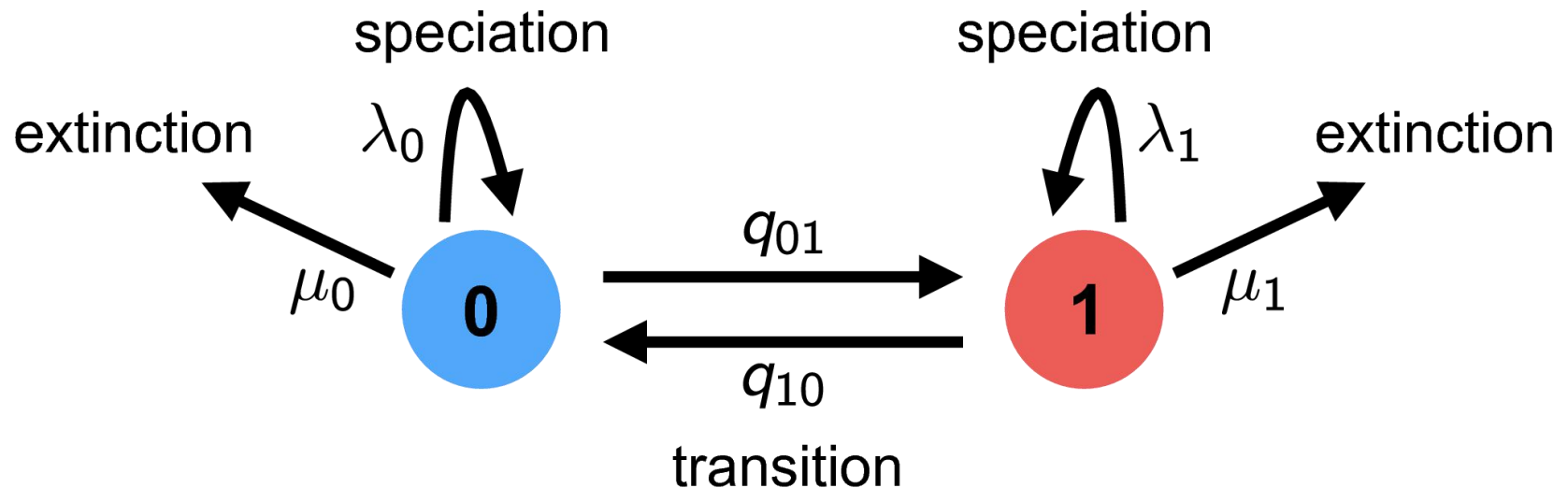
Figure 1: Phylogenies of insect orders and distributions of species richness and herbivory and other traits among clades.



Trait-dependent Diversification

Species Selection (sometimes species sorting)

-- Controversial?!? Not anymore!



Why important?

Ancestral state reconstructions *biased* with trait-dependent diversification!
(Maddison 2007)

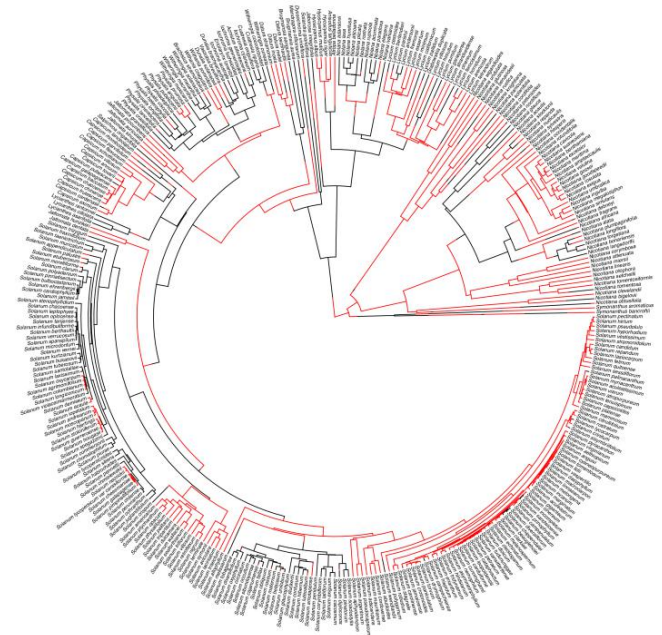
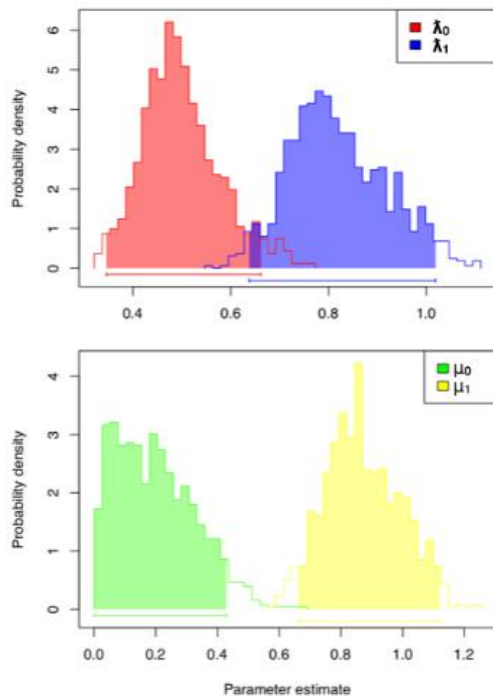


Figure 13.4. Data from Goldberg and Igic (2012) showing presence (red) and absence (black) of self-incompatibility among Solanaceae. Branches colored using stochastic character mapping under a model with distinct forwards and backwards rates; these reconstructions are biased if characters affect diversification rates.

Example:

Self-incompatibility in plants: Prevents selfing
Equal distribution but...

Selfing results in higher speciation rates AND
higher extinction rates

Ancestor more likely SI than SC

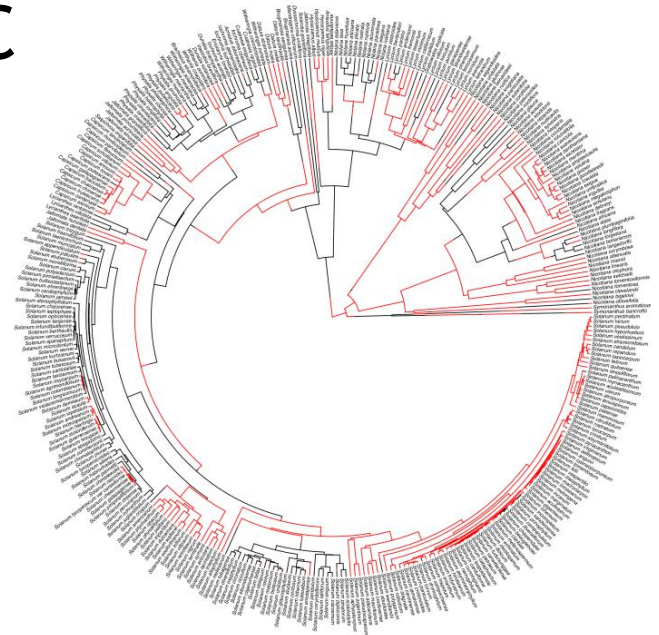
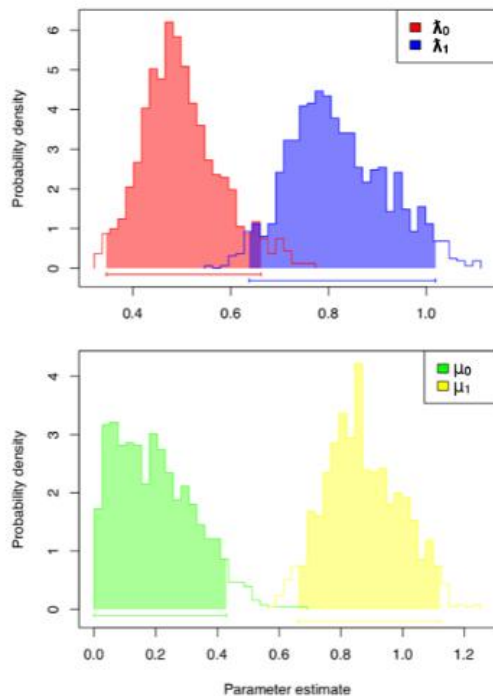


Figure 13.4. Data from Goldberg and Igic (2012) showing presence (red) and absence (black) of self-incompatibility among Solanaceae. Branches colored using stochastic character mapping under a model with distinct forwards and backwards rates; these reconstructions are biased if characters affect diversification rates.

The SSE family of models

Trait models:

BiSSE: Binary State-dependent Speciation and Extinction (extensions for incompletely resolved trees)

MuSSE: Multiple SSE

QuaSSE: Quantitative SSE (but slow)

Models where something special happens at nodes:

GeoSSE: Geographic SSE (DEC + SSE)

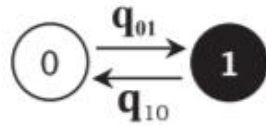
BiSSE-ness/ClaSSE: Cladogenetic SSE

Hidden-state models that account for background heterogeneity:

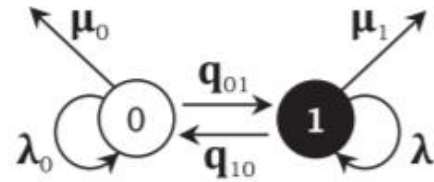
HiSSE: Hidden-State SSE (Solution to problem pointed out by Rabosky and Goldberg)

GeoHiSSE: Geographic Hidden-State SSE

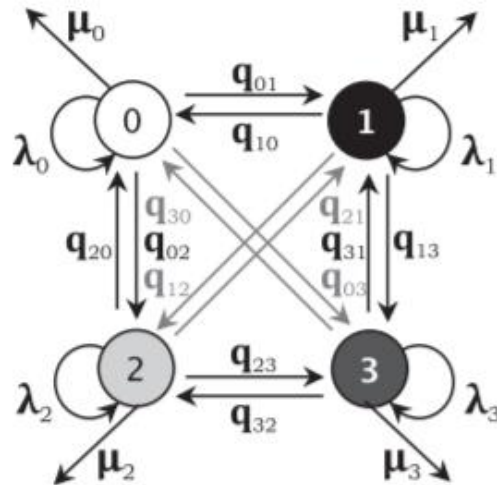
(a) Pagel / Mk2



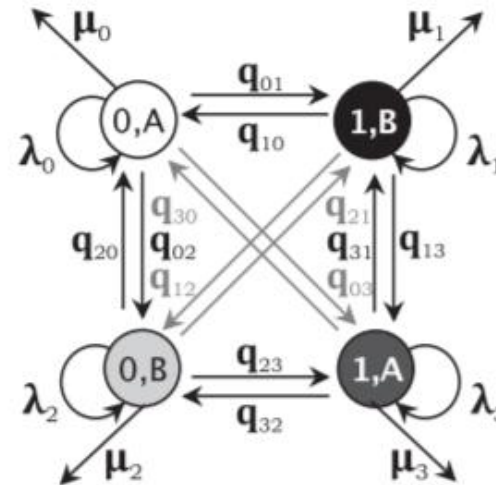
(b) BiSSE



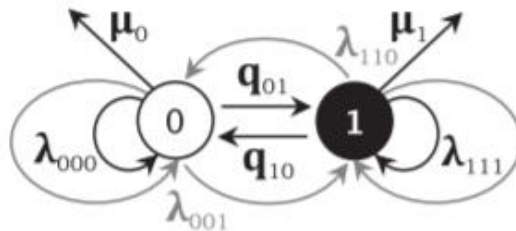
(c) MuSSE (multiple states)



(d) MuSSE (binary trait combinations)



(e) ClaSSE



(f) GeoSSE

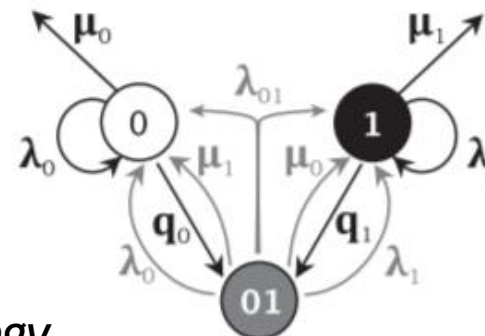


Table 1 Alternative hypotheses that can be tested using BiSSE to explain disparities in clade size associated with binary traits. BiSSE simultaneously estimates rate parameters for the ancestral and derived state (0 and 1, respectively, represented as circles), each of which can be constrained for hypothesis testing: speciation (λ_0 , λ_1), extinction (μ_0 , μ_1) and transition rates (q_{01} , q_{10}). Diversification rates (r_0 or r_1) can be calculated as $r_0 = \lambda_0 - \mu_0$ and $r_1 = \lambda_1 - \mu_1$. Different hypotheses can be tested for statistical significance using either a maximum-likelihood framework, with likelihood ratio tests for nested models or Akaike information criterion, or a Bayesian framework by comparing credibility intervals from a Markov chain Monte Carlo sample of parameter values. Larger circles (for 0 or 1) indicate the state that is found in a higher proportion of extant taxa. Thicker arrows indicate higher parameter values (rates) whereas dotted arrows represent lower values. Note that these processes are not necessarily mutually exclusive.

Process	Expectation	Schematic of expectation
Higher proportion of taxa with ancestral character state		
(A1) <i>Evolutionary dead end</i> : Increased extinction rates associated with state 1 and irreversible character evolution	$\lambda_0 = \lambda_1$ or $\lambda_0 > \lambda_1$ $\mu_0 < \mu_1$ $q_{10} = 0$	
(A2) <i>Asymmetrical diversification</i> : Higher diversification rates in the ancestral state	$r_0 > r_1$ $q_{01} = q_{10}$	
(A3) <i>Asymmetrical transitions (directional evolution)</i> : Higher rate of character loss than gain	$r_0 = r_1$ $q_{01} < q_{10}$	
(A4) <i>Nonequilibrium dynamics</i> : Low transition from the ancestral state to the derived state	$r_0 = r_1$ $q_{01} = q_{10} \sim 0$	
Higher proportion of taxa with derived character state		
(B1) <i>Key innovation</i> : Increased diversification of species with state 1	$\lambda_0 < \lambda_1$ $\mu_0 = \mu_1$ or $\mu_0 > \mu_1$ $q_{01} = q_{10}$	
(B2) <i>Asymmetrical transitions (directional evolution)</i> : Higher rate of character gain than loss	$r_0 = r_1$ $q_{01} > q_{10}$	

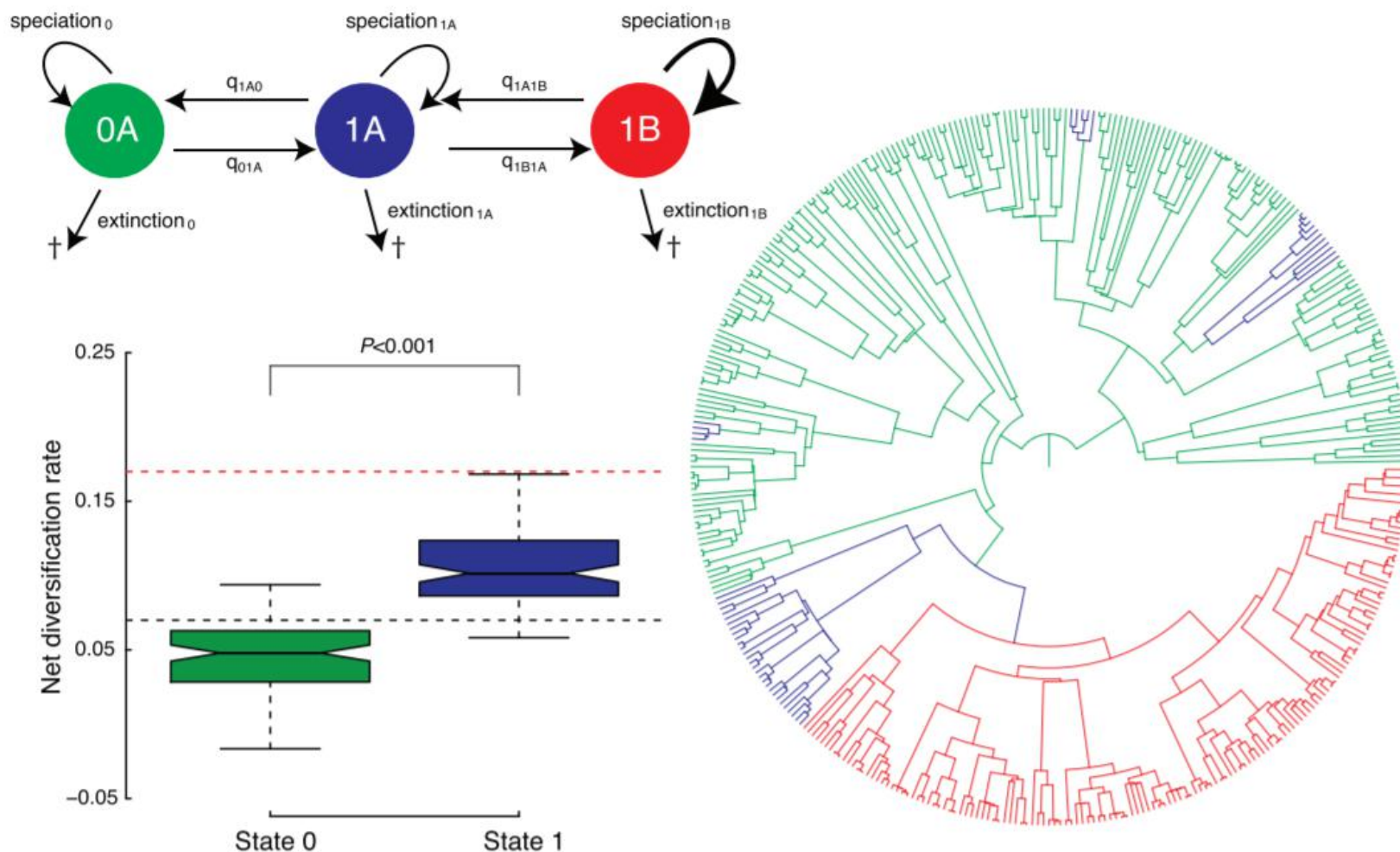


FIGURE 1. The conceptual problem with the presence of hidden states in the application of state-dependent models of speciation and extinction. Here related to state 0 and 1 is an unmeasured third variable with states A and B, and state B has twice the diversification rate of A. This trait is “hidden” because it is not observed in the tip data. If state 1 happens to be a prerequisite for evolving state B, all the state 0 branches will have state A, but some branches in state 1 will have state A and some will have state B. Thus, state 1 actually takes on two states, 1A when the hidden state with higher diversification rate is absent, and 1B when the hidden state with higher diversification rate is present. As shown in the example tree from a simulated tree and trait data, transitions to this unmeasured variable naturally produces nested shifts toward higher rates of diversification within clades comprised of species observed in state 1. When we run 100 simulations of this particular model and fit the resulting data sets in BiSSE, the model infers state 1 as being associated with a significantly higher diversification rates.

Next time: Diversitree lab

