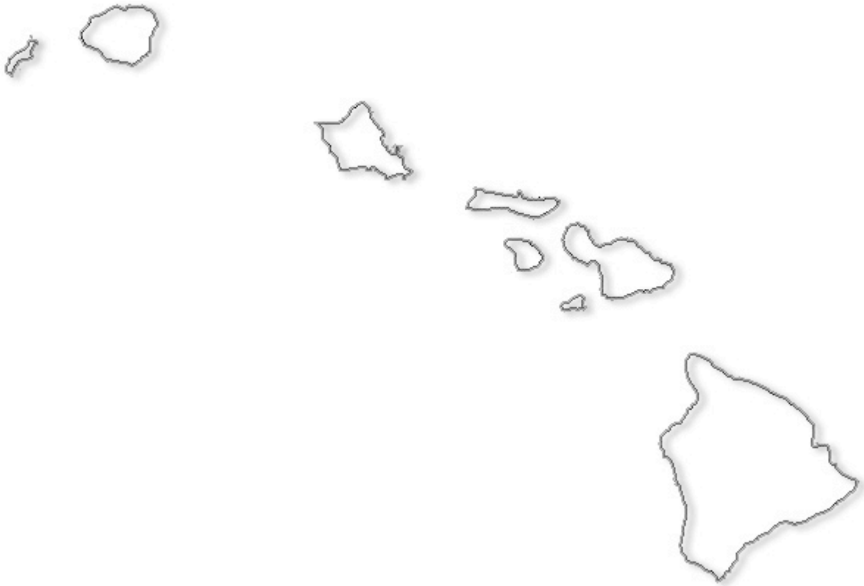
Phylogenetic Approaches for Exploring Differential Rates of Diversification*

*lineage diversification = (speciation – extinction)

Brian R. Moore & Mike May
Department of Evolution & Ecology
UC, Davis

Differential Rates of Diversification

adaptive radiation



©Tom Dempsey Photoseek.com

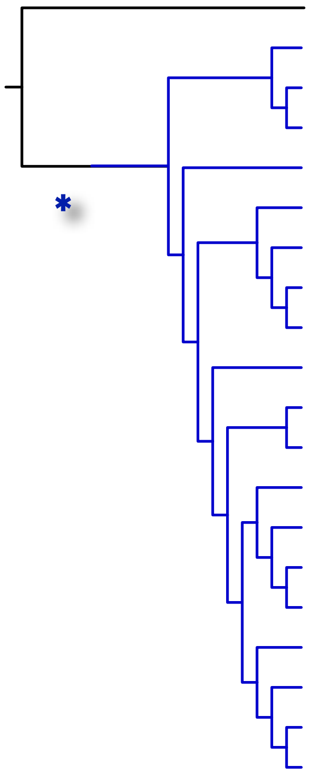
Argyroxiphium sandwicense ssp. *macrocephalum* © G. D. Carr

Hawaiian silverswords

Differential Rates of Diversification

adaptive radiation

key innovations



nectar spurs in columbines

Differential Rates of Diversification

adaptive radiation

key innovations

rapid cladogenesis

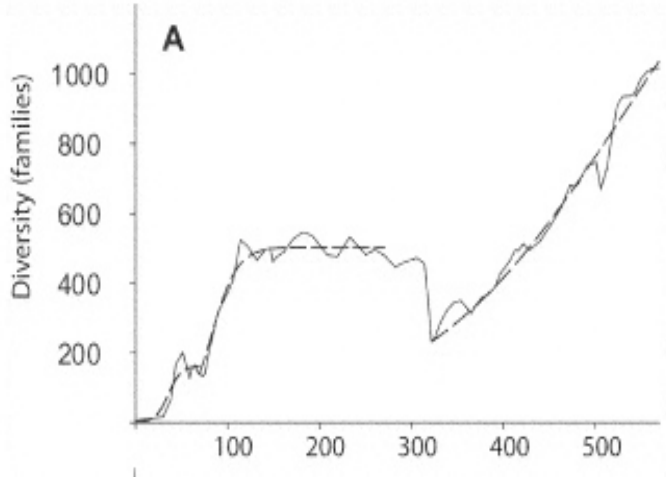
cichlids

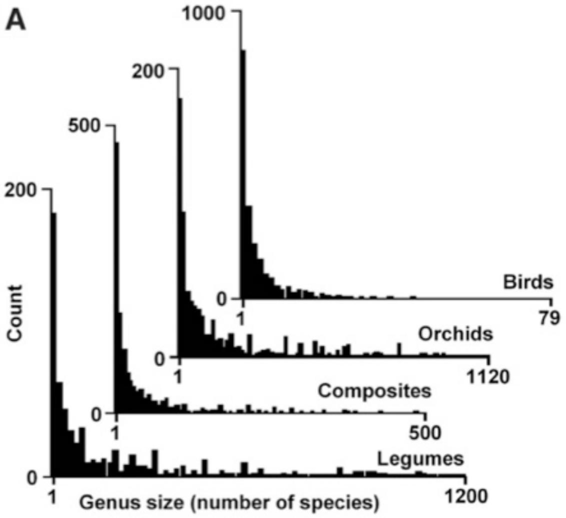
Approaches for Exploring Diversification Rates

Paleontological

Taxonomic

Phylogenetic





Statistical Phylogenetic Approaches for the Study of Diversification Rates

Statistical paradigm	Statistical phylogenetic paradigm
pose a substantive question	What is the net rate of diversification in my study group?
develop a stochastic model with parameters that, if known, would answer the question	develop a phylogenetic model that includes a parameter for net diversification-rate
collect observations that are informative about model parameters	collect a sample of sequence data from the study organisms
find the best estimate of the model parameters using some method conditioned on the data at hand	find the best estimate of the phylogenetic model parameters using ML or Bayesian inference

from the Huelsenbook

Outline: Phylogenetic Approaches For Exploring Rates of Lineage Diversification

I. What are the fundamental questions specific to this research area?

A beginners guide to the types of methods available

II. What phylogenetic information is relevant to inferring diversification rates?

Tree-based observations come in two flavors

III. What models are typically used to infer rates of lineage diversification?

Choosing the right tool for the job at hand

IV. How do we estimate under these lineage diversification models?

Different approaches and different statistical frameworks

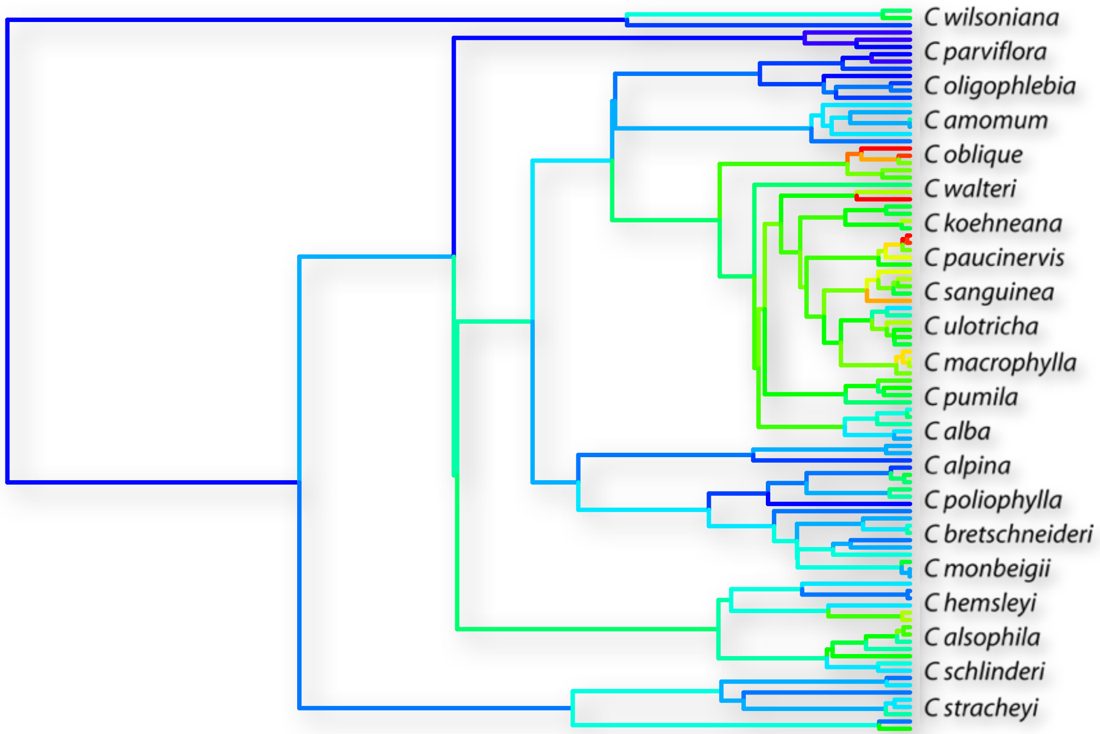
V. Some considerations regarding the application of these methods

Hidden assumptions, if violated, can spoil a nice study

Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

Has the tree diversified under a stochastically constant rate?

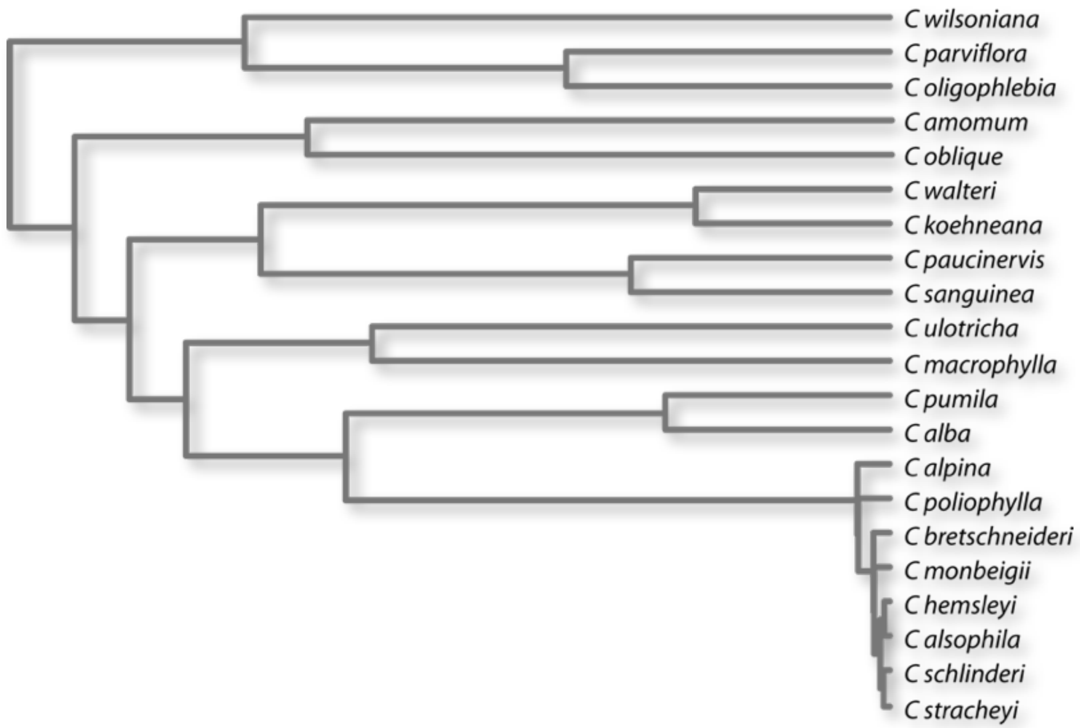


Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

Along which branches have significant rate shifts occurred?

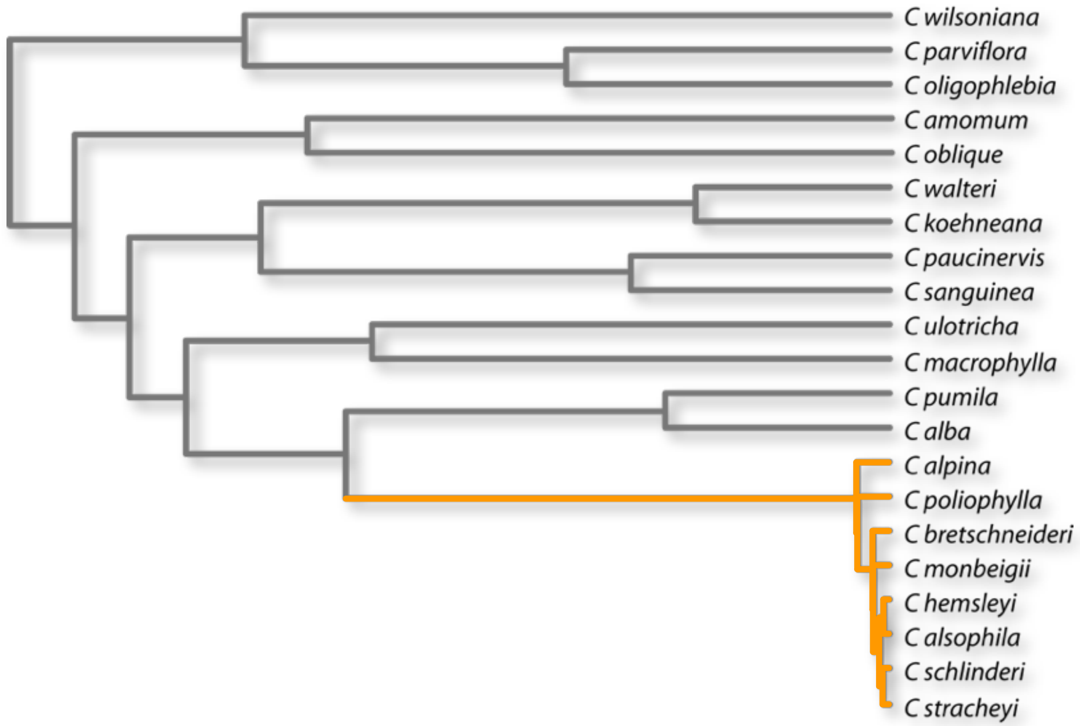


Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

Along which branches have significant rate shifts occurred?



Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

Have tree-wide rates of diversification changed through time?

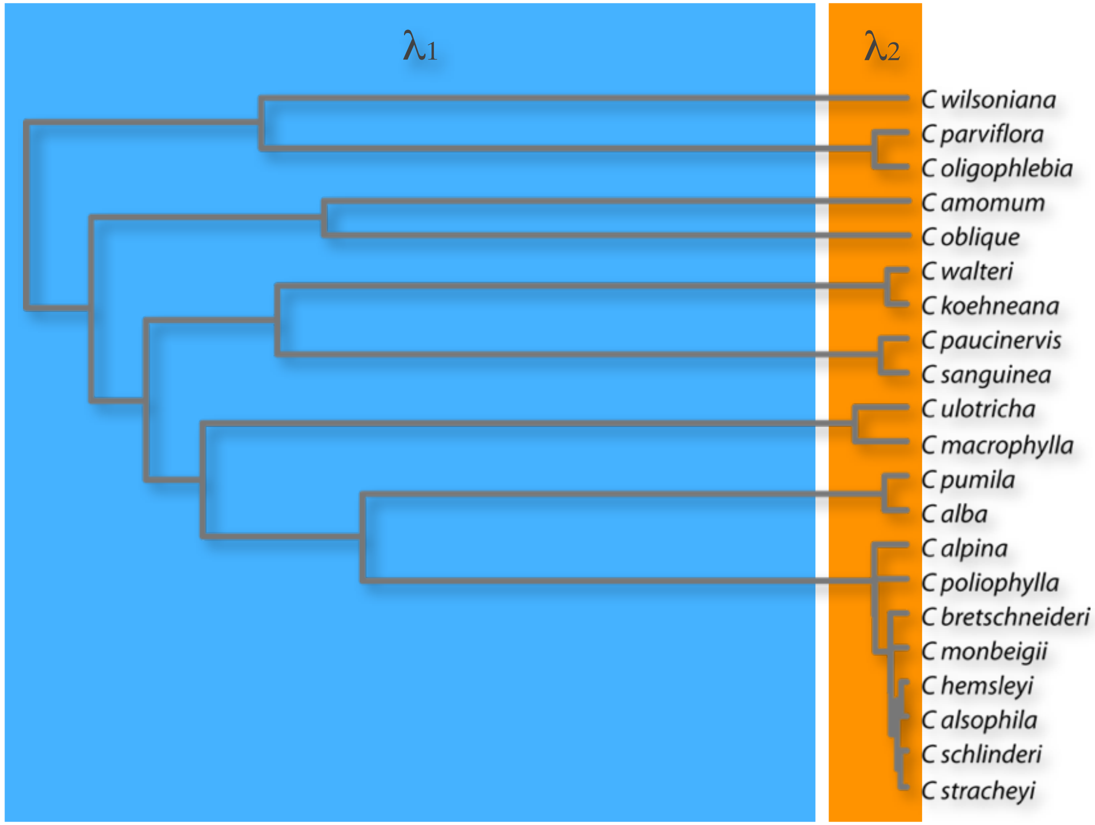
Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

Have tree-wide rates of diversification changed through time?



Five Fundamental Questions in the Study of Diversification Rates

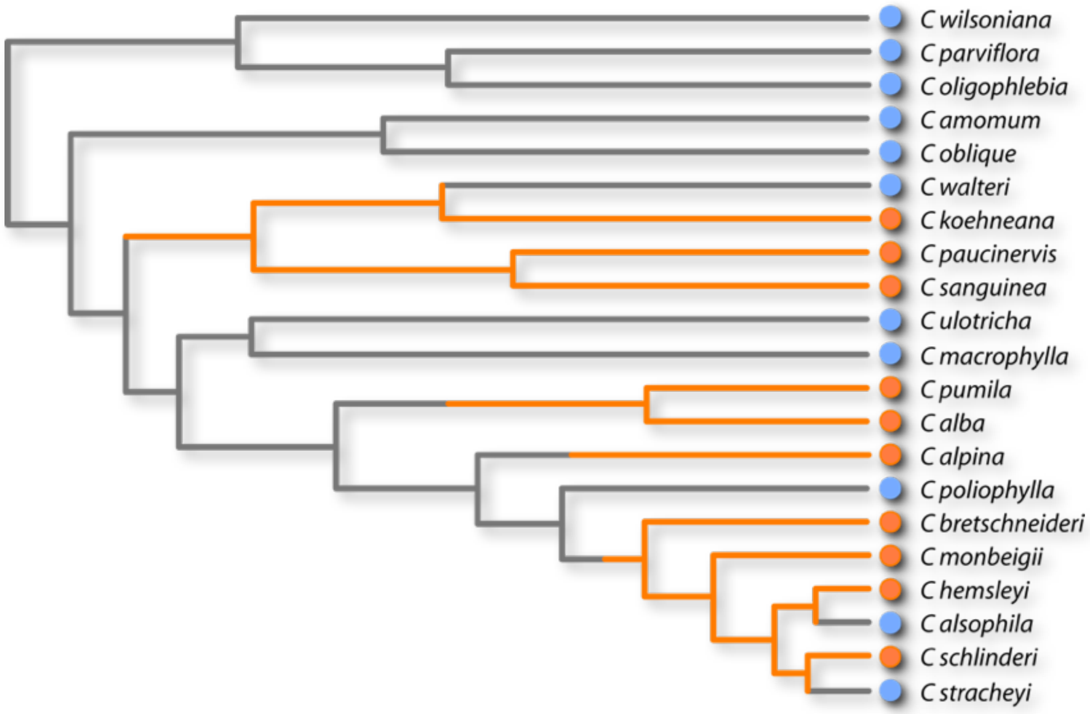
I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

Are diversification rates correlated with other evolutionary variables?



Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

What are the absolute rates of the branching process?

- speciation rate, λ
- extinction rate, μ
- diversification rate, $\partial = (\lambda - \mu)$
- relative extinction rate, $\varepsilon = (\mu / \lambda)$

Outline: Phylogenetic Approaches For Exploring Rates of Lineage Diversification

I. What are the fundamental questions specific to this research area?

A beginners guide to the types of methods available

II. What phylogenetic information is relevant to inferring diversification rates?

Tree-based observations come in two flavors

III. What models are typically used to infer rates of lineage diversification?

Choosing the right tool for the job at hand

IV. How do we estimate under these lineage diversification models?

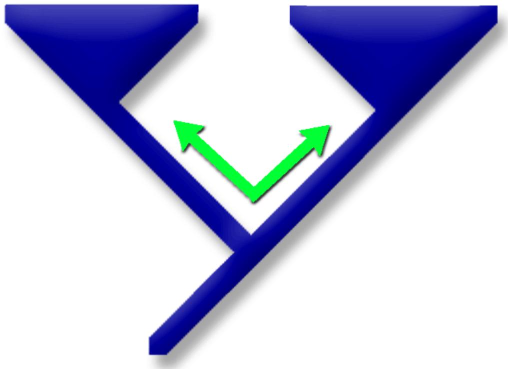
Different approaches and different statistical frameworks

V. Some considerations regarding the application of these methods

More cautionary tales to alarm the uninitiated

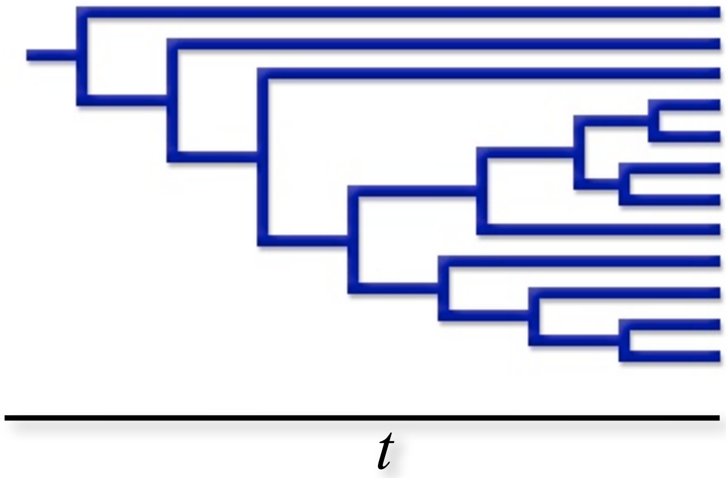
What Are the Basic Phylogenetic Data for Inferring Diversification Rates?

Topological Information



Distribution of species diversity across the tree

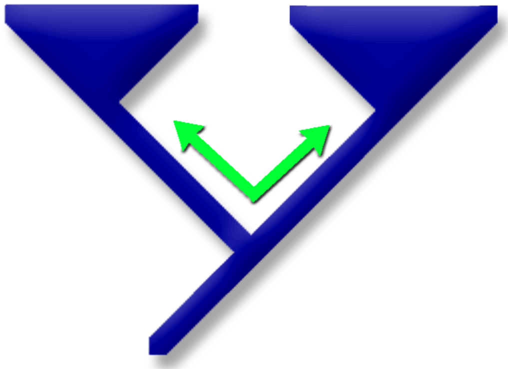
Temporal Information



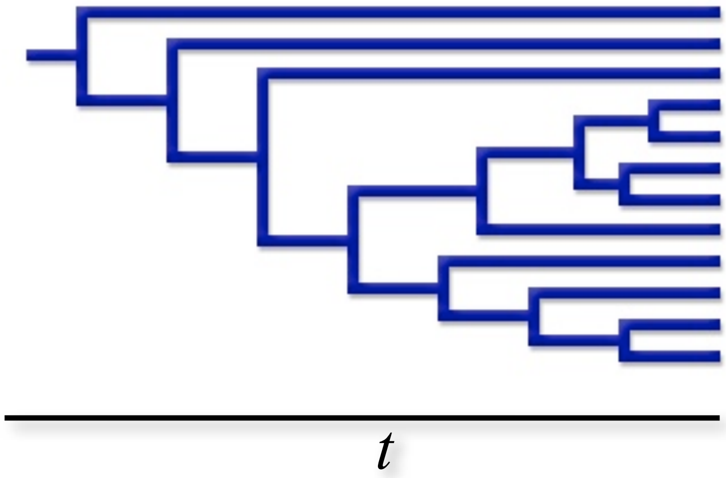
Distribution of speciation events through time

These Pseudo-data Are the Realization of a Stochastic Process

Topological Information



Temporal Information



The primary phylogenetic 'observations'—both the topological distribution of species diversity and the temporal distribution of waiting times—arise via stochastic processes

These 'observations' are estimates from data (and therefore associated with uncertainty)

The general approach is to compare our observations to expectations generated under an appropriate stochastic branching processes

Outline: Phylogenetic Approaches For Exploring Rates of Lineage Diversification

I. What are the fundamental questions specific to this research area?

A beginners guide to the types of methods available

II. What phylogenetic information is relevant to inferring diversification rates?

Tree-based observations come in two flavors

➡ III. What models are typically used to infer rates of lineage diversification?

Choosing the right tool for the job at hand

IV. How do we estimate under these lineage diversification models?

Different approaches and different statistical frameworks

V. Some considerations regarding the application of these methods

More cautionary tales to alarm the uninitiated

Stochastic-Branching Process Models

"All models are wrong, but some are useful" Box (1976)

Common Stochastic-Branching Process (SBP) Models

Generalized Birth-Death (GBD: Kendall, 1948)

- $\lambda(t)$ and $\mu(t)$, $\lambda > \mu$

Constant Birth-Death (CBD: Kendall, 1948)

- λ and μ, $\lambda > \mu$

Generalized Pure Birth (GPB: Harris, 1964)

- $\lambda(t)$, $\mu = 0$

Constant Pure Birth (CPB: Yule, 1924)

- λ, $\mu = 0$

Stochastic-Branching Process Models

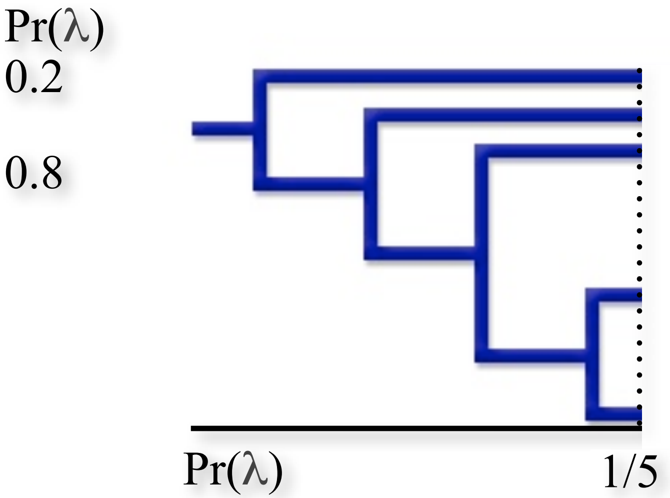
"All models are wrong, but some are useful" Box (1976)

General properties of SBP Markov models

Assume $\lambda(t)$ and/or $\mu(t)$ are:

- constant across all lineages at and instant
- independent across lineages at and instant
- occur instantaneously

A simple Yule simulation



Stochastic-Branching Process Models

"All models are wrong, but some are useful" Box (1976)

Some more exotic SBP models!

- non-Markov SBP models (Chan and Moore, 1999)

  relax assumption that events are instantaneous

  explicitly model the geographic context of speciation

- Multi-type SBP models

  relax assumption that events are equiprobable across all lineages at any instant

Outline: Phylogenetic Approaches For Exploring Rates of Lineage Diversification

I. What are the fundamental questions specific to this research area?

A useful guide to the types of methods available

II. What phylogenetic information is relevant to inferring diversification rates?

Tree-based observations come in two flavors

III. What models are typically used to infer rates of lineage diversification?

Choosing the right tool for the job at hand

➡ IV. How do we estimate under these lineage diversification models?

Different approaches and different statistical frameworks

V. Some considerations regarding the application of these methods

More cautionary tales to alarm the uninitiated

Statistical Inference Under Stochastic-Branching Process Models

SBP models are employed in two qualitatively different ways

- null modeling approaches generate expectations against which we can compare our phylogenetic ‘observations’—i.e., the topological distribution of species diversity and the temporal distribution of waiting times
- model fitting approaches estimate parameters of the SBP models from the phylogenetic ‘observations’—i.e., the topological distribution of species diversity and the temporal distribution of waiting times

Statistical Inference Under Stochastic-Branching Process Models

Inference under SBP models may exploit different statistical approaches

- Maximum Likelihood Estimation approaches involve numerical optimization algorithms to identify the joint parameter estimates that collectively maximize the profile likelihood of the phylogenetic ‘observations’ under the SBP model

- Bayesian approaches involve numerical algorithms to approximate the joint posterior probability density of the SBP model parameters given the phylogenetic ‘observations’

- Quasi-Bayesian approaches involve making ML estimates of the SBP model parameters over a marginal posterior probability density of some aspect of the phylogenetic ‘observations’

Why Pursue These Questions in a Bayesian Statistical Framework?

Recent developments have largely been developed in Quasi-Bayesian ML framework

The study of diversification entails several sources of uncertainty

- phylogeny/topology
- branch lengths/durations
- rate parameters
- event histories

Bayesian framework provides a means for accommodating uncertainty

Exploit advantages of associated numerical and analytical techniques

Outline: Phylogenetic Approaches For Exploring Rates of Lineage Diversification

I. What are the fundamental questions specific to this research area?

A useful guide to the types of methods available

II. What phylogenetic information is relevant to inferring diversification rates?

Tree-based observations come in two flavors

III. What models are typically used to infer rates of lineage diversification?

Choosing the right tool for the job at hand

IV. How do we estimate under these lineage diversification models?

Different approaches and different statistical frameworks

➡ V. Some considerations regarding the application of these methods

More cautionary tales to alarm the uninitiated

Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

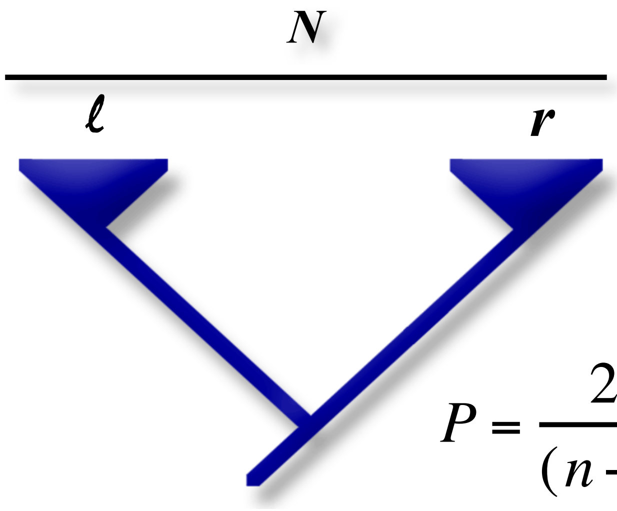
IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

Detecting Diversification Rate Variation Across Branches

Whole-Tree Model-Based Methods

(Chan & Moore, 2002; Moore, Chan & Donoghue, 2004)



$$P = \frac{2r}{(n-1)}$$

Detecting Diversification Rate Variation Across Branches

Whole-Tree Model-Based Methods

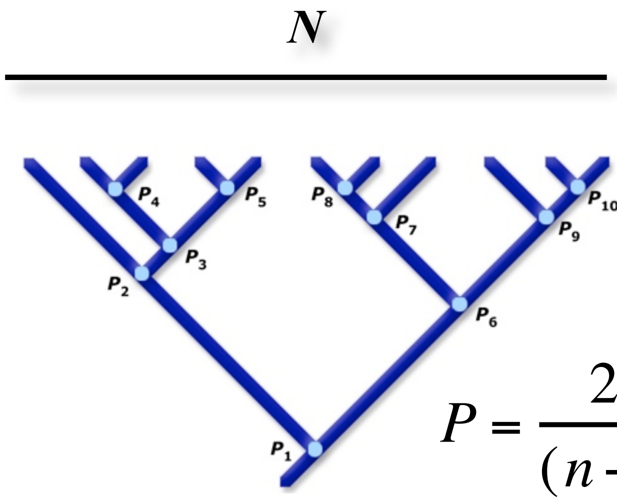
(Chan & Moore, 2002; Moore, Chan & Donoghue, 2004)

Topological approach

Combines individual ERM nodal probabilities over internal nodes as the product, $M\prod$, sum, $M\sum$, which may be weighted by the sample size of each node, $M^*\prod$, sum, $M^*\sum$

Significance estimated by Monte Carlo simulation of the null distribution of the test statistic under an ERM stochastic branching process

Power: HIGH
+ statistically robust (accommodates phylogenetic uncertainty, polytomies)
+ can accommodate incomplete taxon sampling
+ incorporates more of the tree
- does not identify anomalously large/small groups



$$P = \frac{2r}{(n-1)}$$

$$\Pi^* = \frac{\sum_{i=1}^{n-1} \ln(n_i)\ln(P_i)}{\sum_{i=1}^{n-1} \ln(n_i)} \qquad \Sigma^* = \frac{\sum_{i=1}^{n-1} \ln(n_i)P_i}{\sum_{i=1}^{n-1} \ln(n_i)}$$

Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

Locating Shifts in Diversification Rate Along Branches

Relative Cladogenesis Statistic

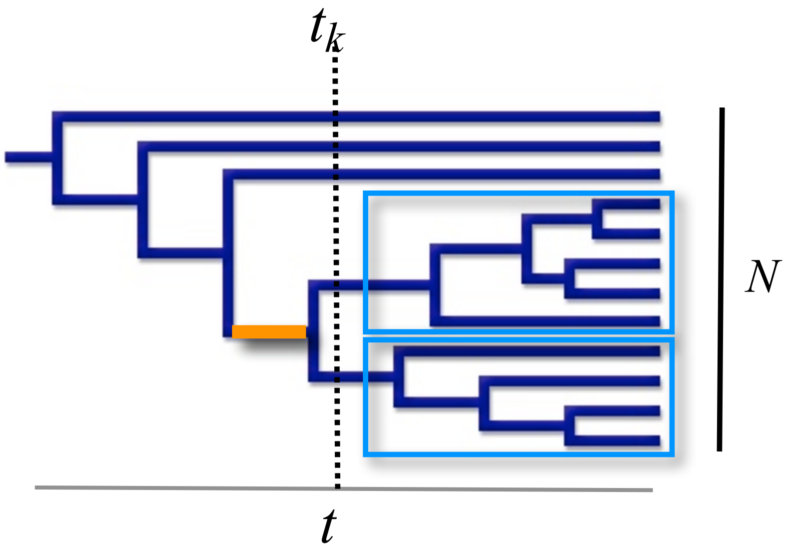
(Purvis et al., 1995)

Incorporates information on the topological distribution of species diversity of contemporaneous lineages

Compares observed topological distribution of descendant lineages to expectations under Yule stochastic branching process model

Uses implicit parsimony to assign rate shifts to internal branches

-susceptible to rate trickle-down problem



Locating Shifts in Diversification Rate Along Branches

AIC Model-Fitting

(Alfaro et al., 2009)

Incorporates information on the topological distribution of (unsampled) species diversity of dated lineages

Iteratively assess the fit of these data to a set of k candidate rate-parameter models using ML

Selects among the set of k rate-parameter models using AIC

+allows inclusion of unsampled species

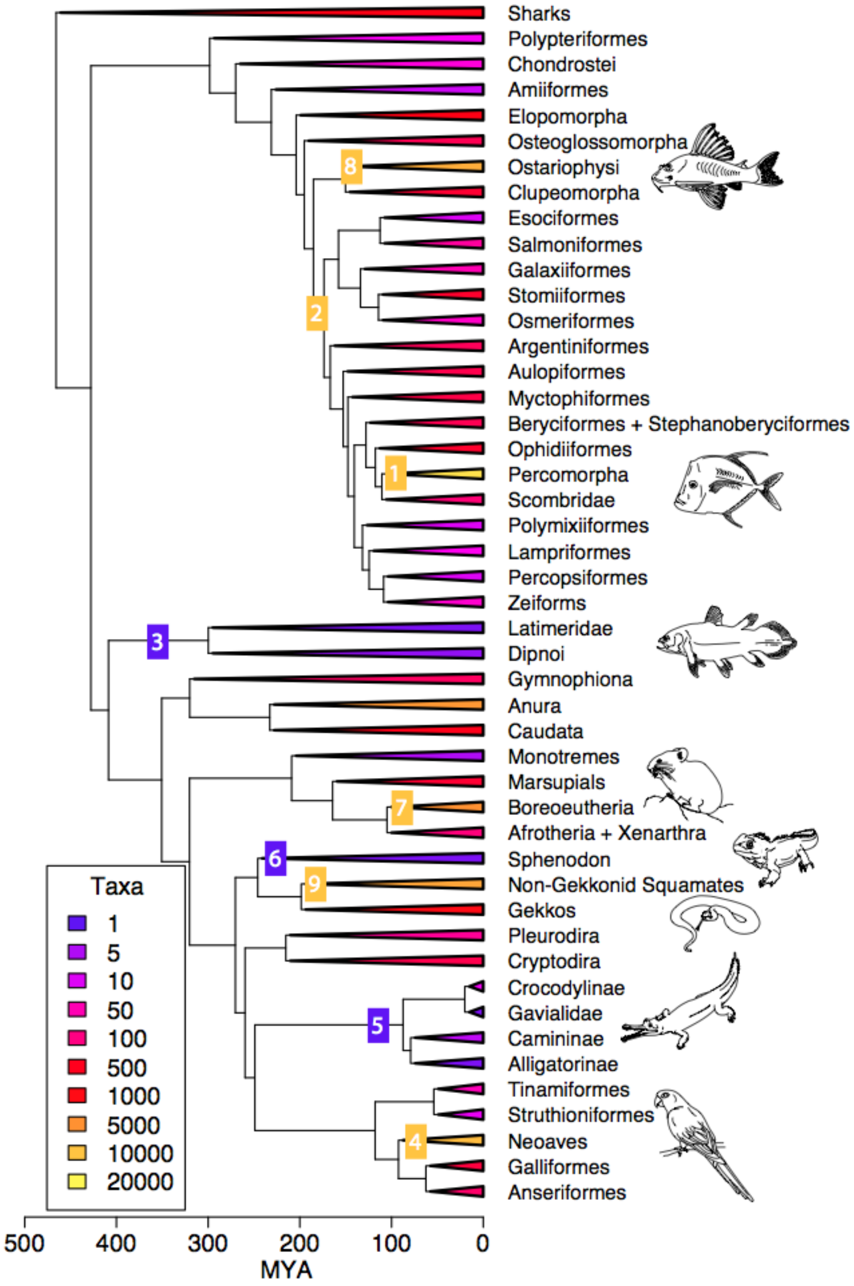
-small subset of models evaluated

-AIC model choice may be biased

-simple rate estimators

-does not accommodate phylogenetic uncertainty

-susceptible to rate trickle-down problem



Locating Shifts in Diversification Rate Along Branches

ML Model-Fitting

(Moore et al., 2004)

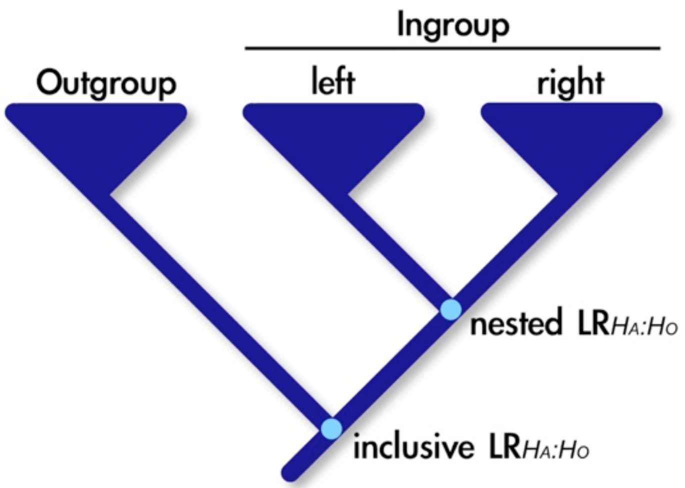
Iterative maximum likelihood model-fitting approach (Sanderson & Donoghue, 1994)

Three-taxon tree framework

Equal-rates Markov (ERM) branching process

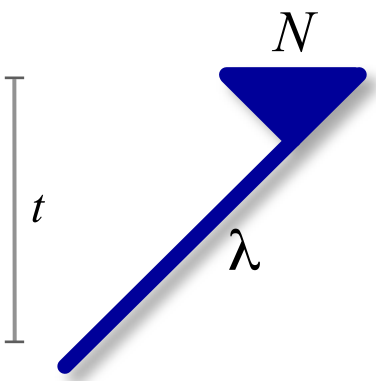
Simplified: one- and two-rate parameter models only, do not integrate through time

Generalized: iterated over all internal branches



Locating Shifts in Diversification Rate Along Branches

Likelihood Calculations



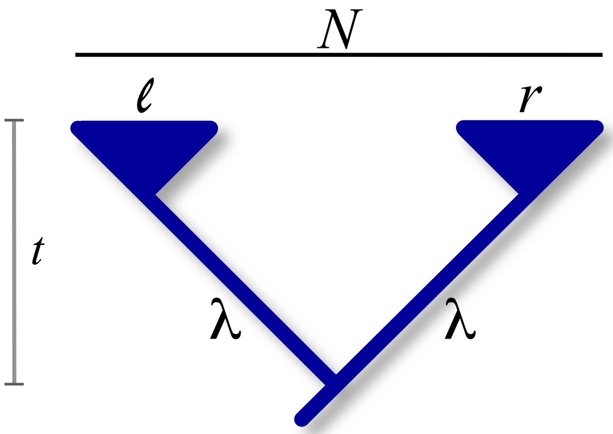
$$P(N \mid \lambda, t) = e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{N-1}$$

(Kendall, 1948; Harris, 1964)

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

Likelihood Calculations: One-Rate Parameter Model (H_O)



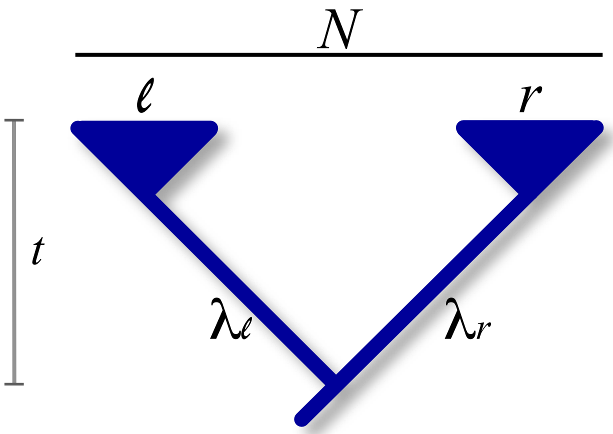
$$P(\ell,r\mid H_O)=\frac{P(\ell\mid\lambda,t)P(r\mid\lambda,t)}{\sum_{i=1}^{N-1}P(i\mid\lambda,t)P(N-i\mid\lambda,t)}$$

$$P(\ell,r\mid H_O)=\frac{\left(e^{-\lambda}\left(1-e^{-\lambda}\right)^{\ell-1}\right)\left(e^{-\lambda}\left(1-e^{-\lambda}\right)^{r-1}\right)}{\sum_{i=1}^{N-1}\left(e^{-\lambda}\left(1-e^{-\lambda}\right)^{i-1}\right)\left(e^{-\lambda}\left(1-e^{-\lambda}\right)^{N-i-1}\right)}$$

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

Likelihood Calculations: Two-Rate Parameter Model (HA)



$$P(\ell, r \mid H_A) = \frac{P(\ell \mid \lambda_\ell, t)\,P(r \mid \lambda_r, t)}{\sum_{i=1}^{N-1} P(i \mid \lambda_\ell, t)\,P(N-i \mid \lambda_r, t)}$$

$$P(\ell, r \mid H_A) = \frac{\left(e^{-\lambda_\ell}\left(1-e^{-\lambda_\ell}\right)^{\ell-1}\right)\left(e^{-\lambda_r}\left(1-e^{-\lambda_r}\right)^{r-1}\right)}{\sum_{i=1}^{N-1}\left(e^{-\lambda_\ell}\left(1-e^{-\lambda_\ell}\right)^{i-1}\right)\left(e^{-\lambda_r}\left(1-e^{-\lambda_r}\right)^{N-i-1}\right)}$$

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

Likelihood Ratio of One- & Two Rate Parameter Models

$$LR_{H_A:H_O} = \ln\left(\frac{P(\ell \mid \lambda_\ell, t)P(r \mid \lambda_r, t)}{\sum_{i=1}^{N-1} P(n_i \mid \lambda_\ell, t)P(N - n_i \mid \lambda_r, t)}\right) - \ln\left(\frac{P(\ell, r \mid \lambda, t)}{\sum_{i=1}^{N-1} P(n_i \mid \lambda, t)P(N - n_i \mid \lambda, t)}\right)$$

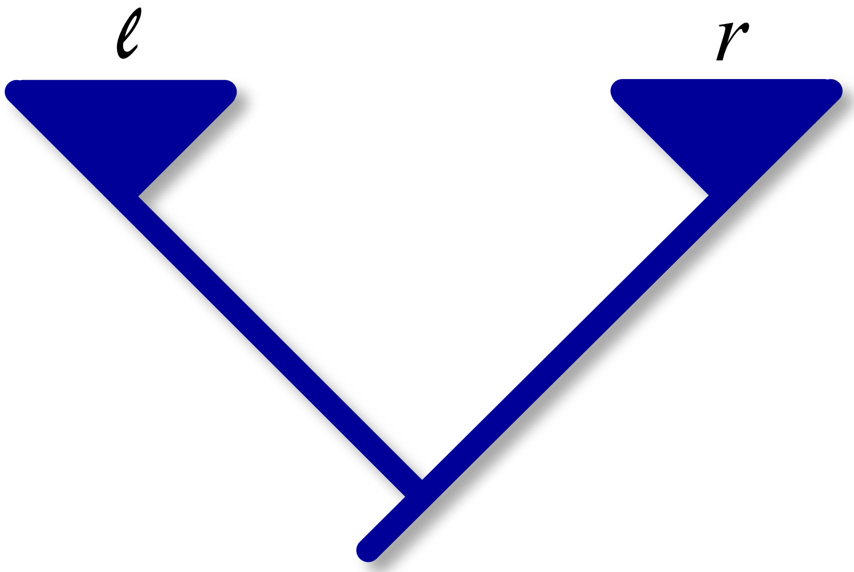
(two-rate model) (one-rate model)

- assess significance by Monte Carlo simulation

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

The 'Trickle-Down' Problem

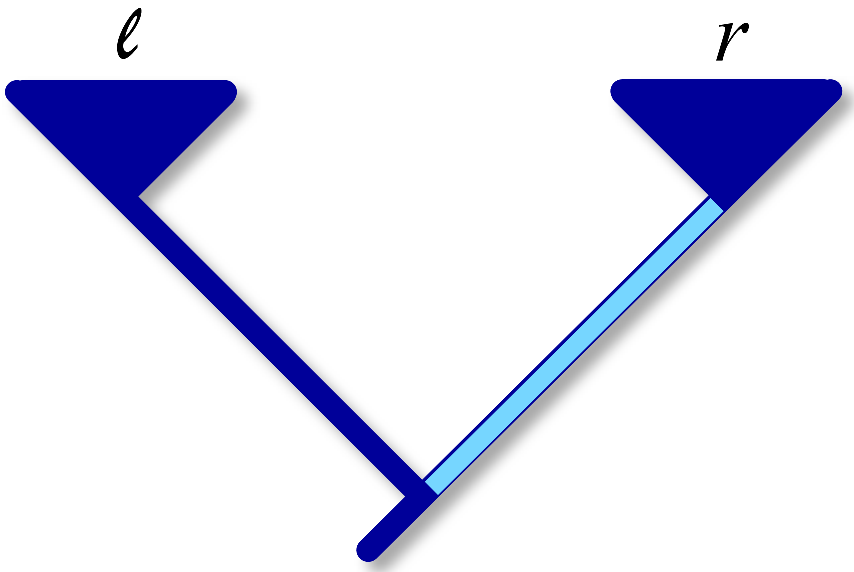


$\lambda_\ell \neq \lambda_r$

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

The 'Trickle-Down' Problem

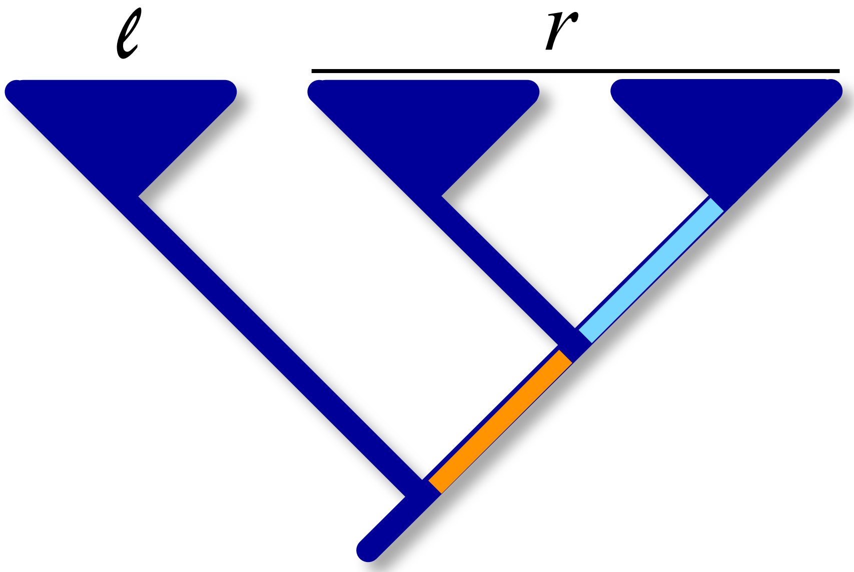


$\lambda_\ell \neq \lambda_r$

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

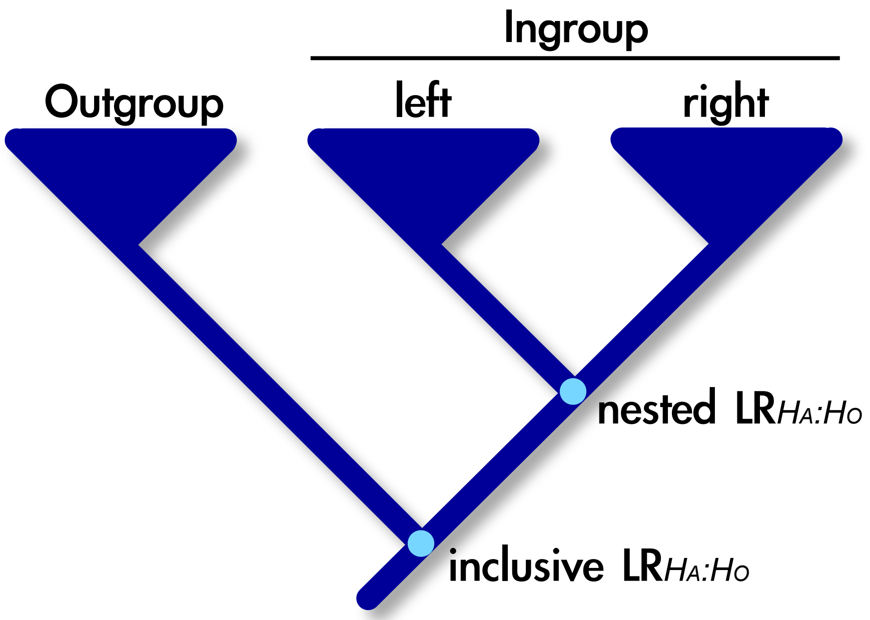
The 'Trickle-Down' Problem



Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

The Δ_1 Shift Statistic



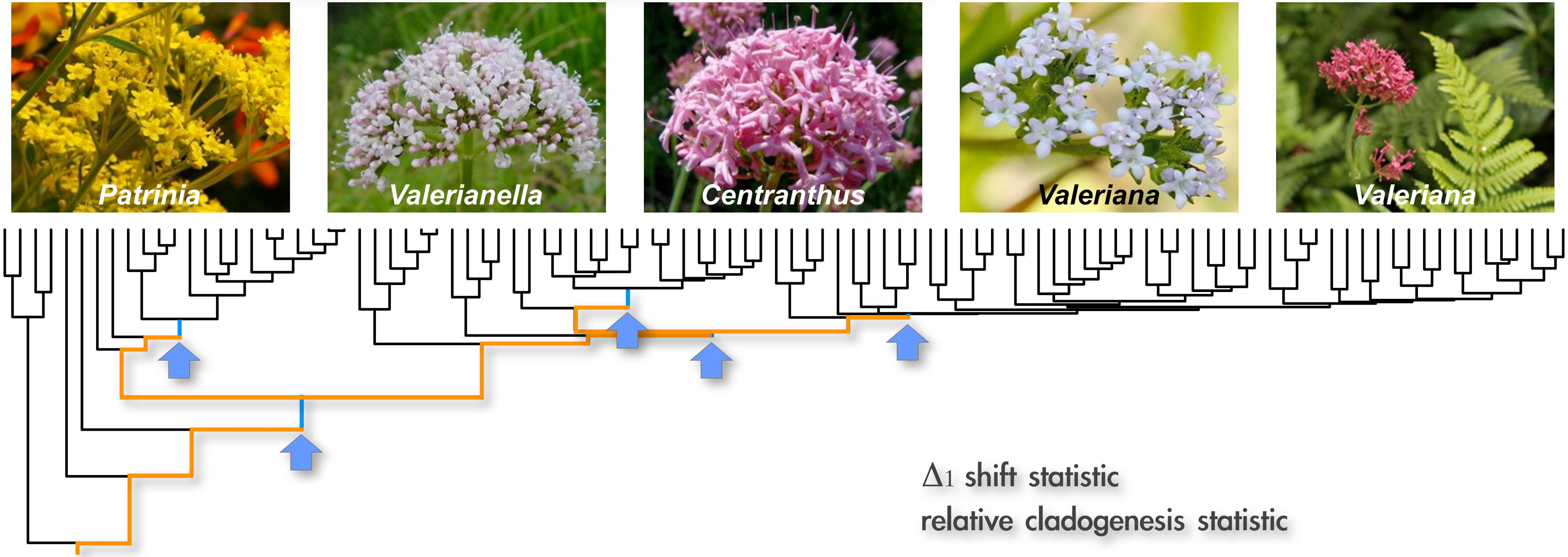
$$\Delta_1 = \left(LR_{H_A:H_O} n_{OG} : n_{IG}\right) - \left(LR_{H_A:H_O} n_{IG_L} : n_{IG_R}\right)$$

(inclusive node) (nested node)

Moore, Chan & Donoghue (2004)

Locating Shifts in Diversification Rate Along Branches

The 'Trickle-Down' problem in action: Valerianaceae



Δ_1 shift statistic

relative cladogenesis statistic

Moore & Donoghue (*American Naturalist,* 2007)

Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

Locating Shifts in Diversification Rate Through Time

Lineage-Thru-Time Plots

(e.g., Nee et al., 1992, 1994a,b)

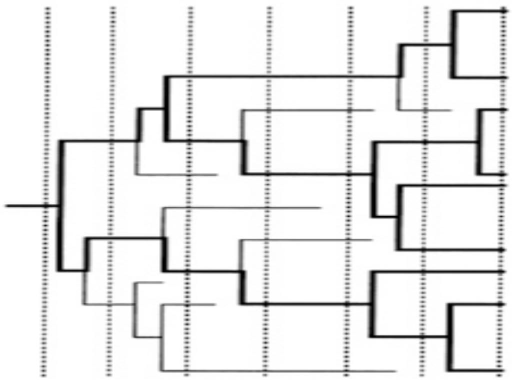
Semi-logarithmic accumulation of lineages

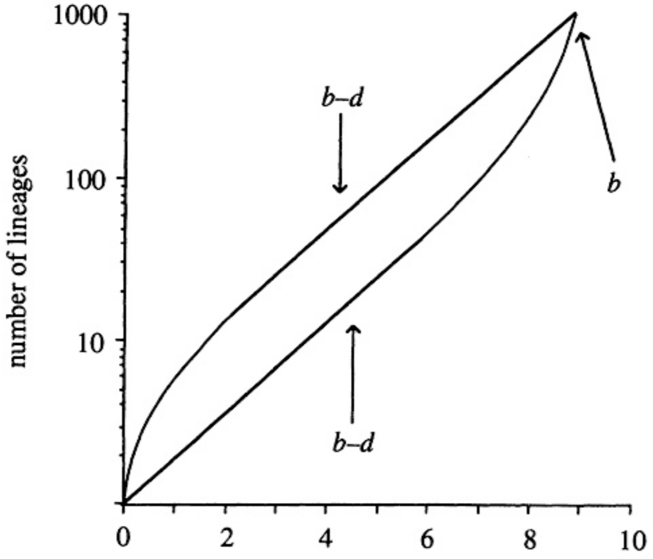
Primarily descriptive, characterized by:

- Push of the Past
- slope = diversification rate, $\partial = (\lambda - \mu)$

If extinction rate, $\mu = 0$, also characterized by:

- Pull of the Present
- increased offset slope
- asymptotic behavior under random sub sampling





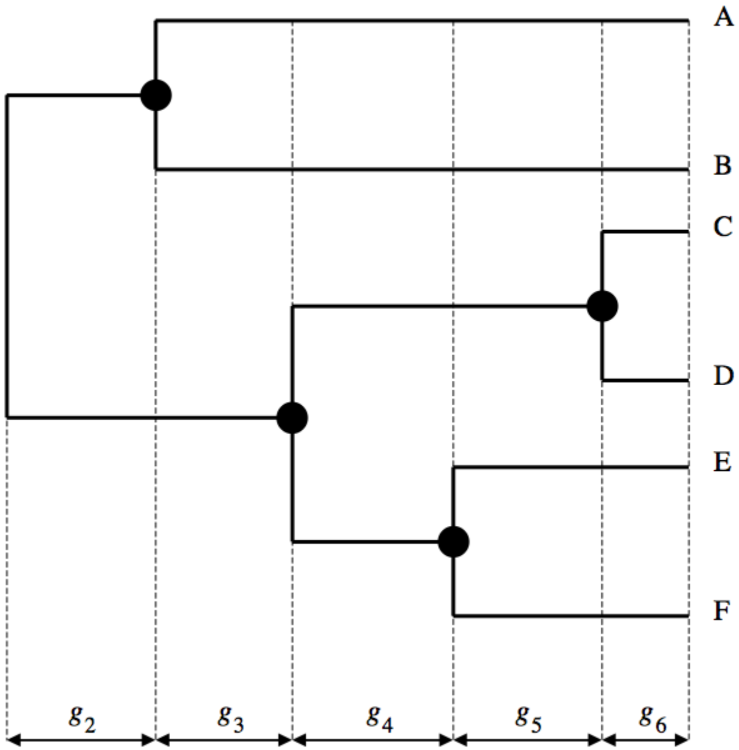
Locating Shifts in Diversification Rate Through Time

Gamma Statistic

(Pybus & Harvey, 2001)

Based on the average distance of branching events relative to the root

- assumes constant rates across lineages
- can accommodate incomplete and RANDOM species sampling
- can ONLY detect diversification rate decreases but NOT diversification rate increases
- Power: FAIR (if N is large and $\sim \mu = 0$)



$$\gamma = \frac{\left(\frac{1}{n-2}\sum_{i=2}^{n-1}\left(\sum_{k=2}^{i} k g_k\right)\right) - \left(\frac{T}{2}\right)}{T\sqrt{\frac{1}{12(n-2)}}}, \quad T = \left(\sum_{j=2}^{n} j g_j\right).$$

Locating Shifts in Diversification Rate Through Time

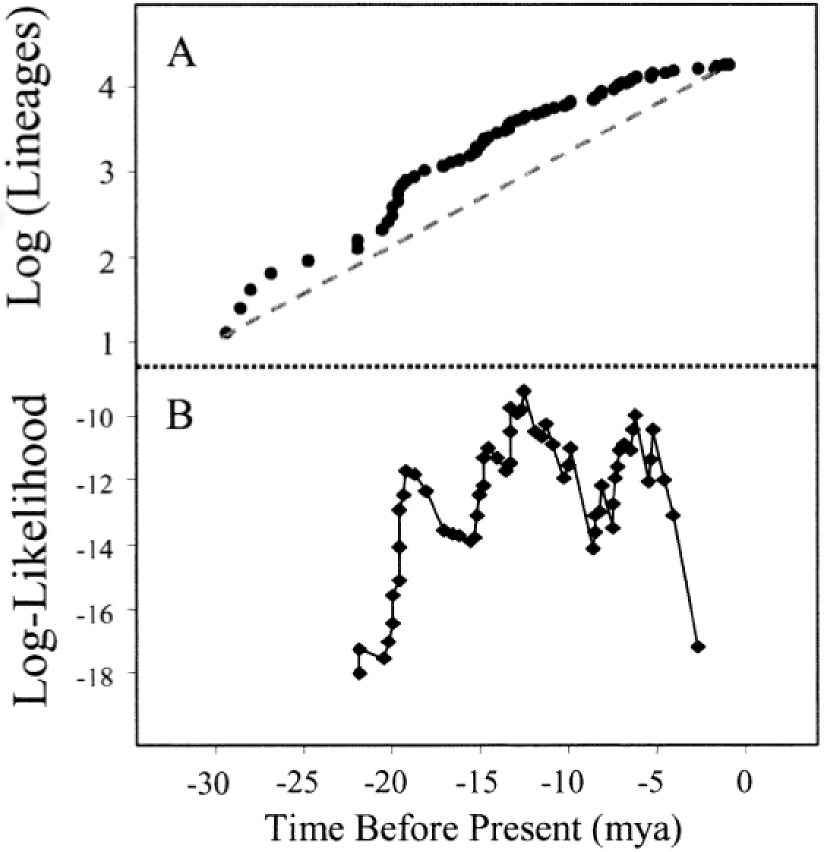
BDL Method

(Rabosky, 2006)

Based on the relative MLEs of the estimated waiting times under RC and RV models

- calculated as DeltaAICrc
- compare test statistic to distribution generated by Monte Carlo simulation under a Yule model
- can detect rate decreases and rate increases (sort of, if $\mu = 0$)
- assumes constant rates across lineages
- assumes complete/random sampling
- Power: Good (if N is large and $\mu = 0$, trees are large sampling is complete, and no diversification rate variation across lineages)

$$\Delta AIC_{RC} = AIC_{RC} - AIC_{RV}$$



Locating Shifts in Diversification Rate Through Time

'Density-dependent' diversification

LTT plots estimated from empirical data exhibit pervasive rate decreases through time

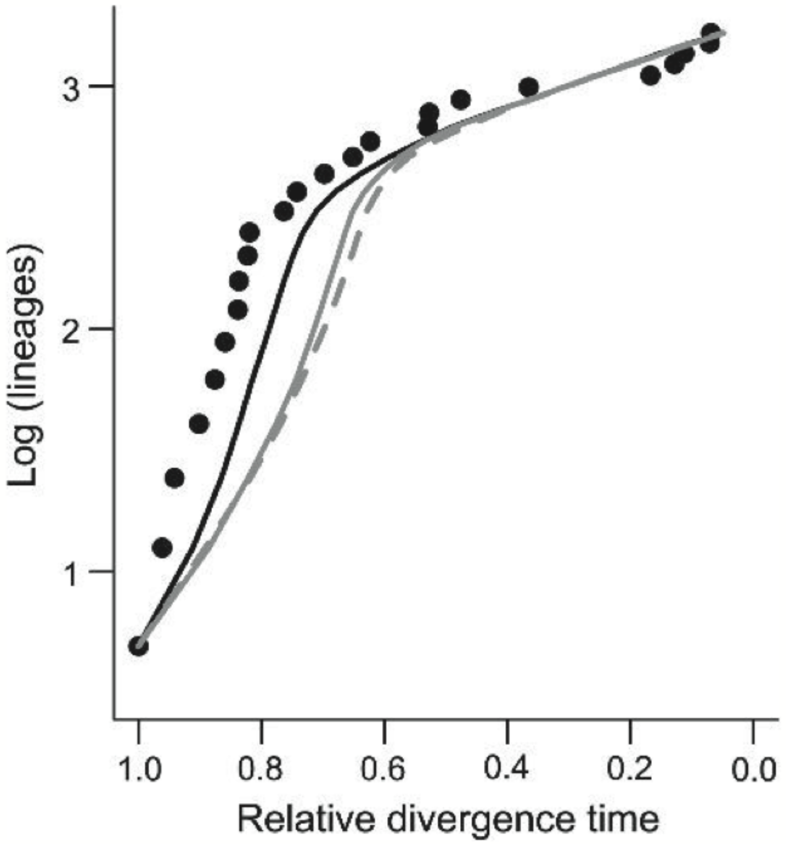
- rate-constant and more complex multi-rate models are fit to the 'observations'
- standard statistical methods are then used to select among the models

The better fit of rate-variable models to phylogenetic 'observations' has been used to argue for a deterministic causal mechanism

- diversification is regulated by ecological/ geographic interactions among species

It is critical to assess the adequacy of these inferences, as it is known that violation of the assumptions leads to identifiability problems

- an infinite number of models can generate an identical LTT plot if assumptions are not met (Kubo & Iwasa, 1995)



(e.g., Rabosky, 2009)

Locating Shifts in Diversification Rate Through Time

Methodological Concerns

Effects of branch-length estimation biases

- substitution-rate profile may lead to systematic estimation bias of deeper/shallower nodes (e.g., Revell et al., 2005)

Departure from random species sampling

- more realistic taxon sampling may produce spurious decrease in rate (e.g., Cusimano & Renner, 2010)

Effects of among-lineage rate variation

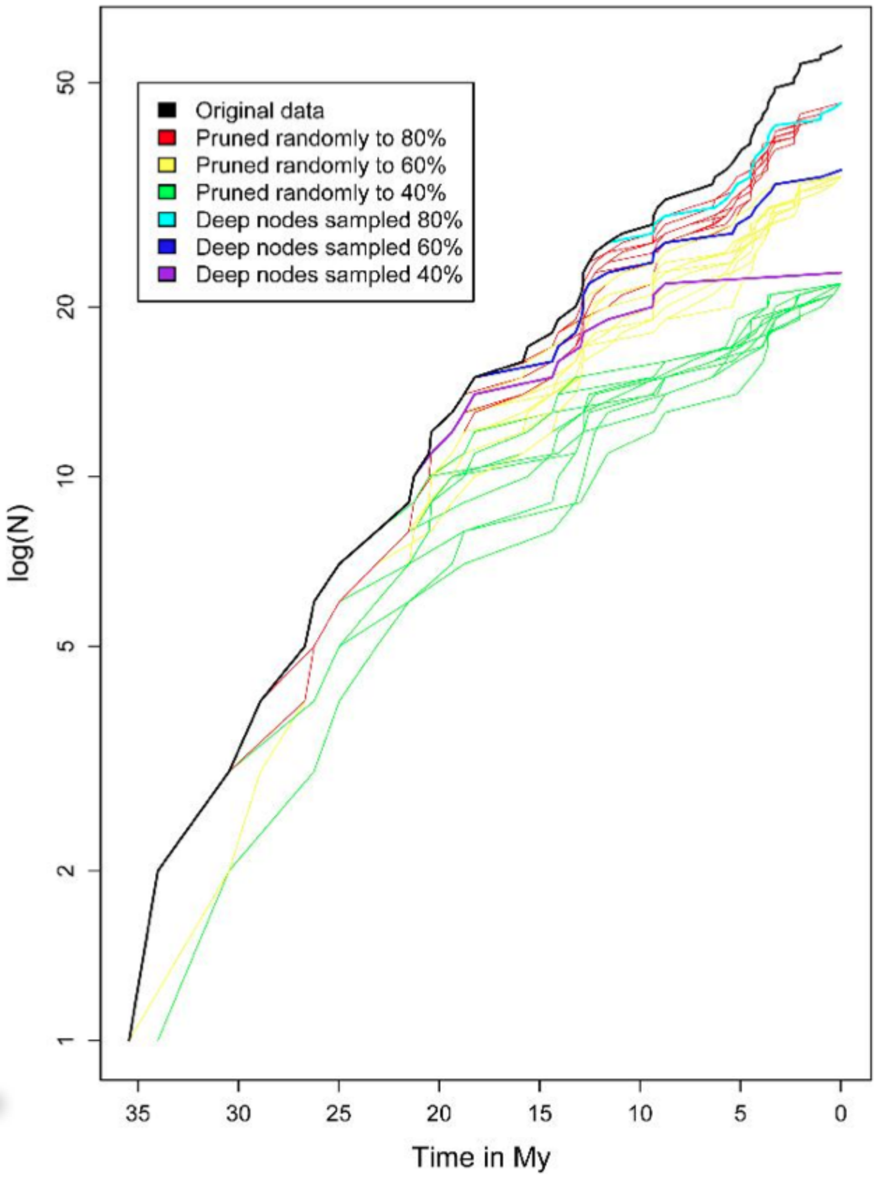
- violation of stochastically constant rates across lineages may produce spurious rate decreases

Effects of model-selection bias

- reliance on AIC may lead to inflated rejection of simpler rate constant models

Effects of phylogenetic uncertainty

- reliance on point estimates of phylogeny and divergence times may render inferences unreliable



Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

Evaluating Correlates of Differential Diversification Rates

Two fundamentally different types of questions:

Is there a general correlation between trait state and diversification rate?

e.g., Are nectar spurs correlated with increase rates of diversification?

Is there a correlation between a specific event and and diversification rate?

e.g., Was a particular origin of nectar spurs in columbines correlated with increase rates of diversification?

Evaluating Correlates of Differential Diversification Rates

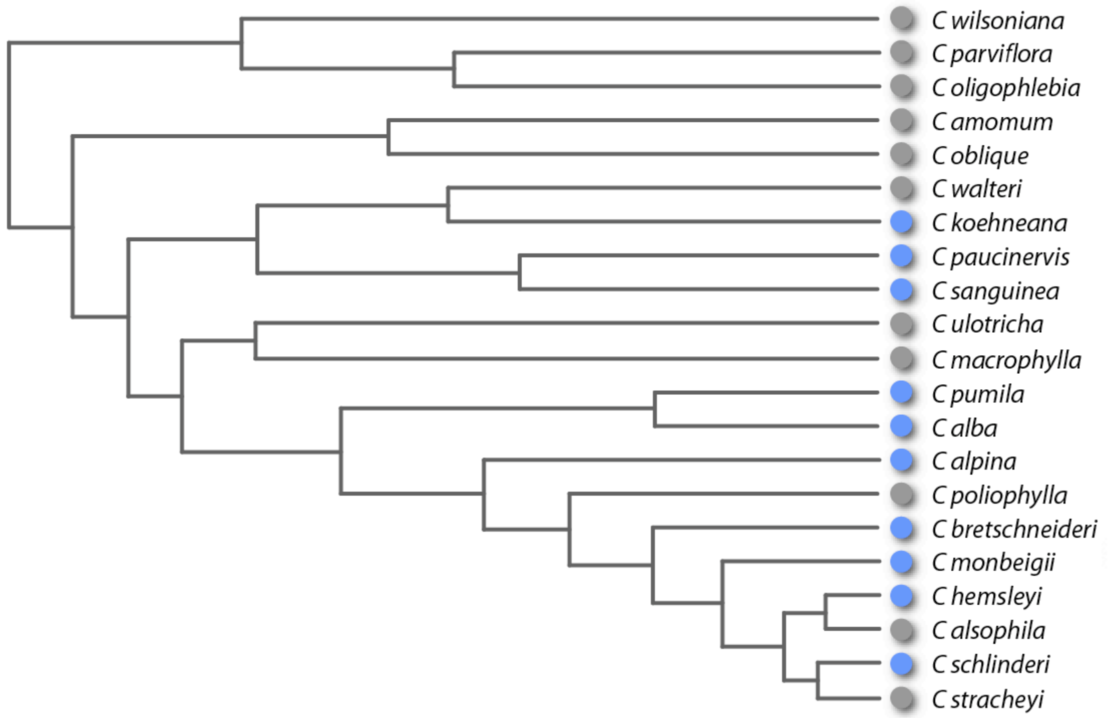
Recent advances for evaluating diversification-rate correlates

Stochastic character mapping; keymap (Ree, 2005)

CVPPD; tRate (Moore & Donoghue, 2009)

*SSE model; DiversiTree (Maddison et al., 2007; Fitzjohn et al., 2009; Fitzjohn, 2010...)

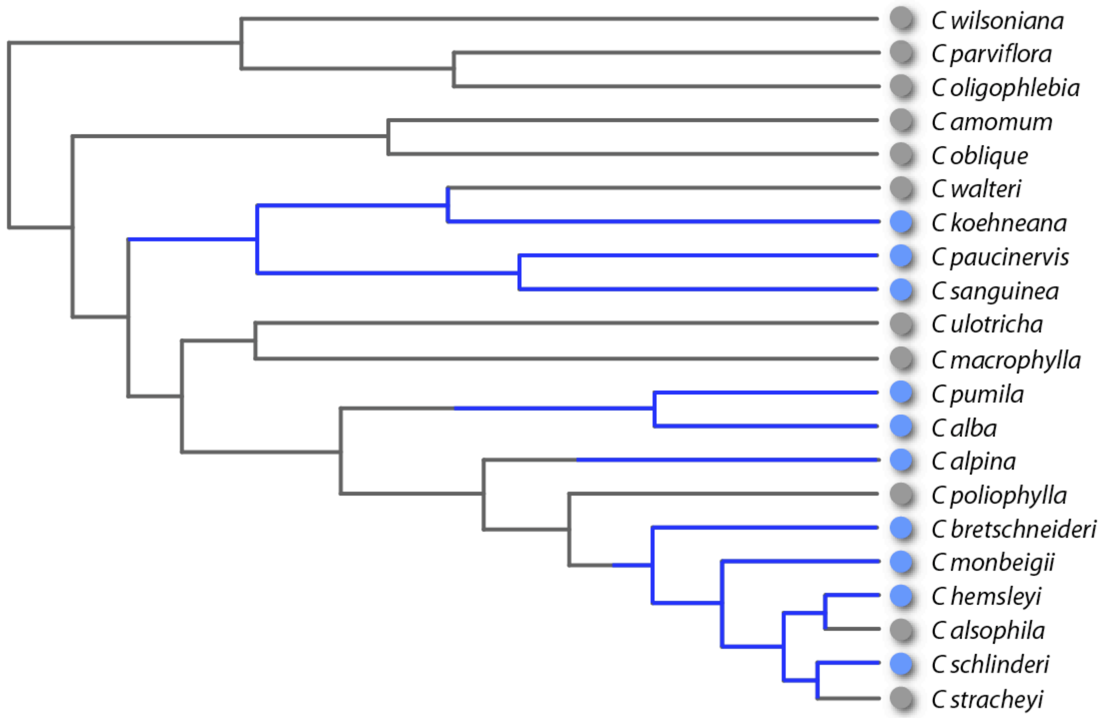
Evaluating Correlates of Differential Diversification Rates



Relies on stochastic character mapping (e.g, Neilsen, 2002; Huelsenbeck et al., 2003)

Ree (***Evolution***, 2005)

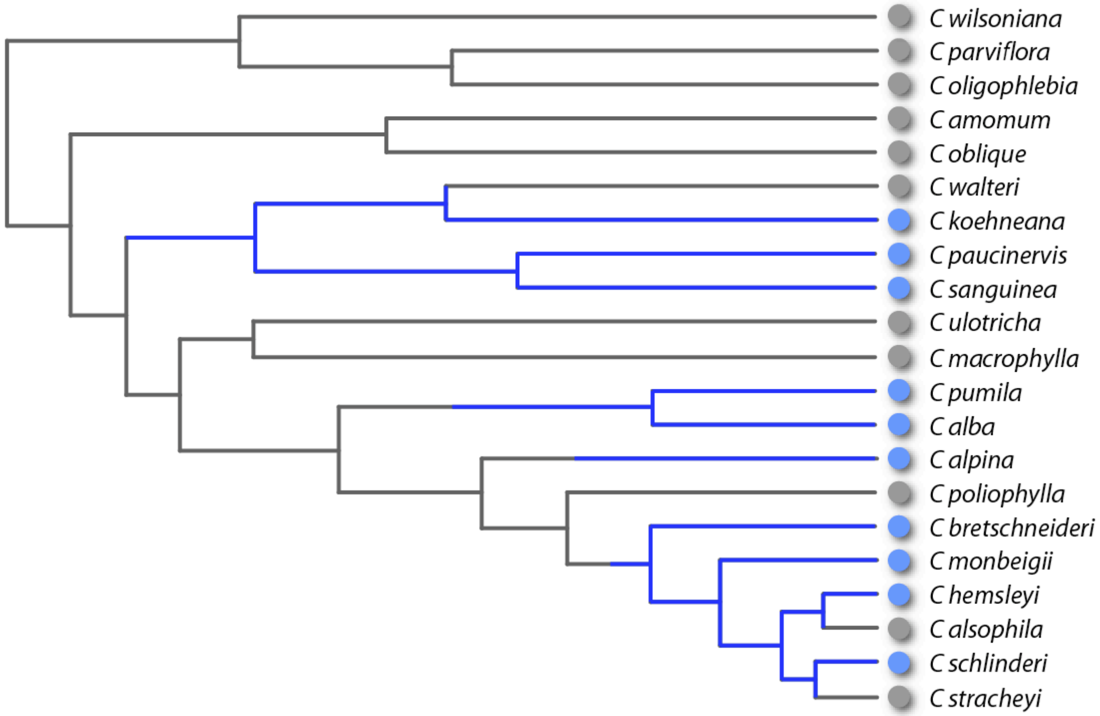
Evaluating Correlates of Differential Diversification Rates



$$\lambda_i = N_i/t_i$$

Ree (*Evolution*, 2005)

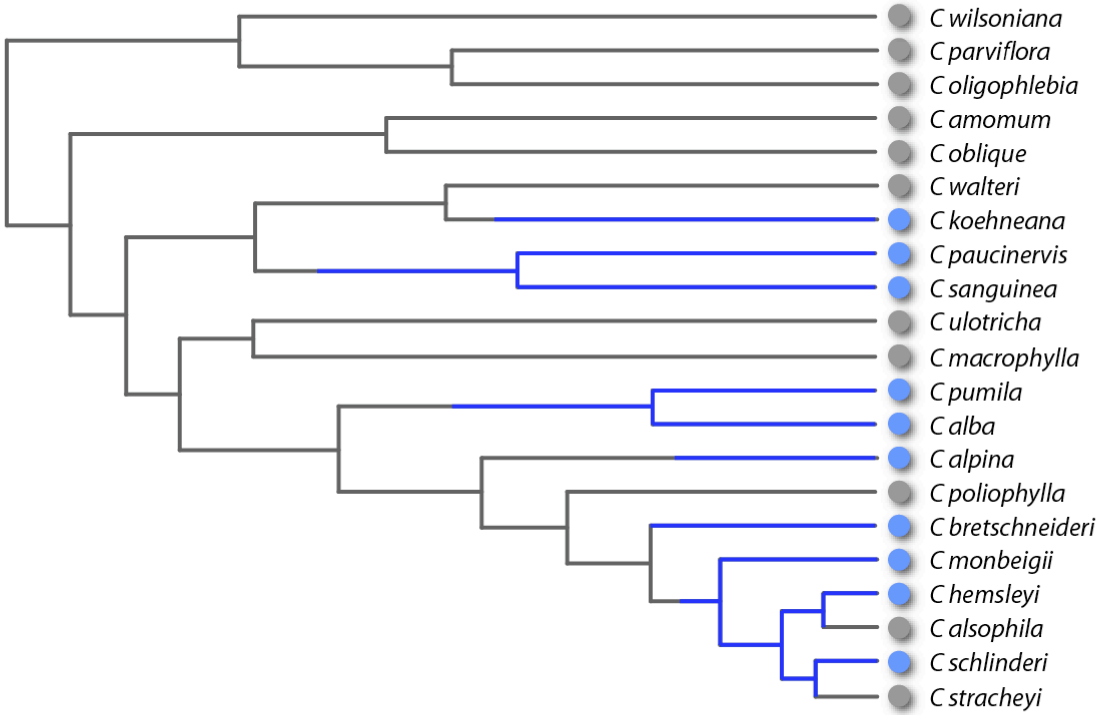
Evaluating Correlates of Differential Diversification Rates



$$\delta_1 = \lambda_{●} - \lambda_{●}$$

Ree (***Evolution***, 2005)

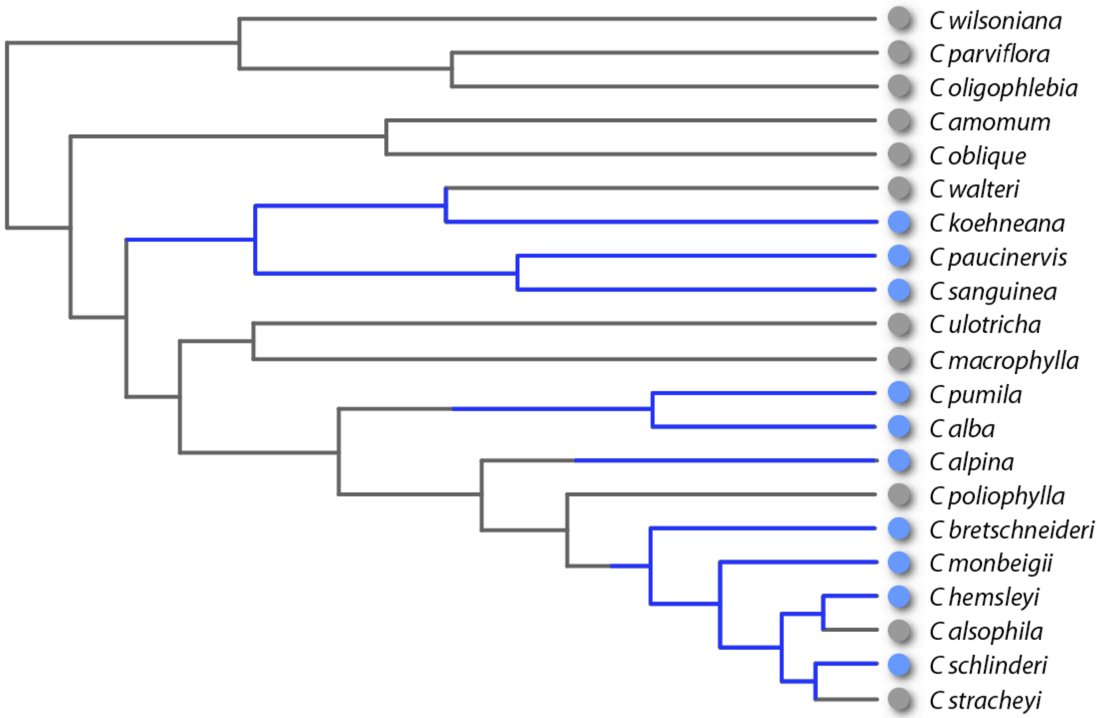
Evaluating Correlates of Differential Diversification Rates



$$\delta_2 = \lambda_{\bullet} - \lambda_{\bullet}$$

Ree (*Evolution*, 2005)

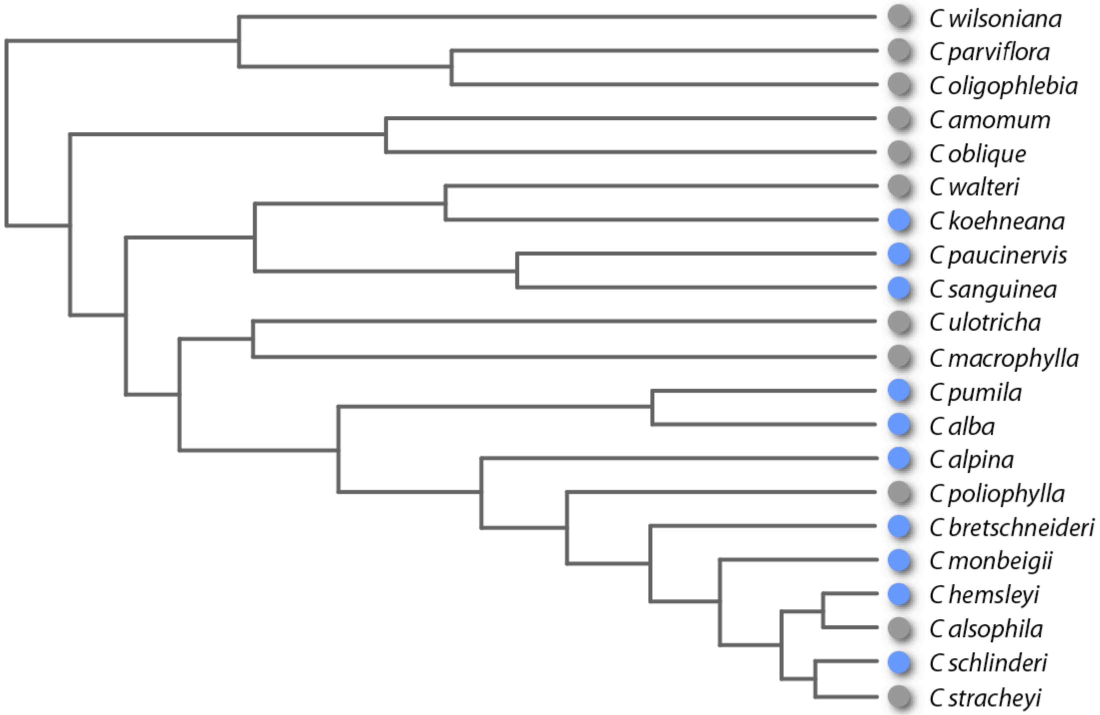
Evaluating Correlates of Differential Diversification Rates



$$\delta_n = \lambda_{\bullet} - \lambda_{\bullet}$$

Ree (*Evolution*, 2005)

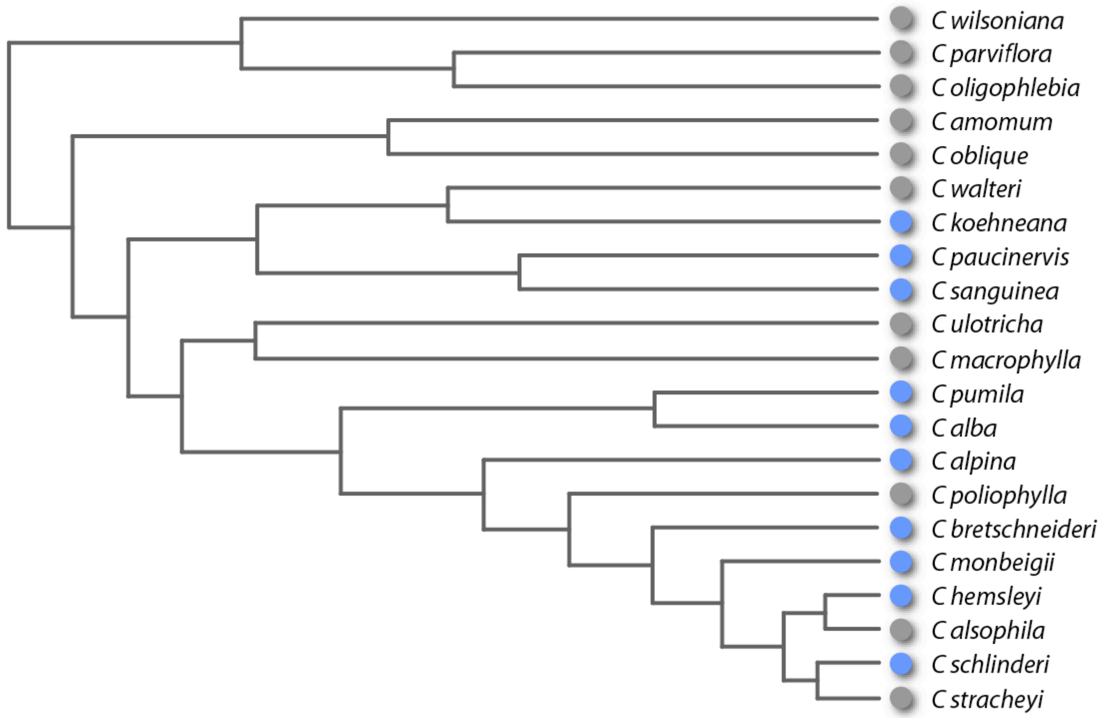
Evaluating Correlates of Differential Diversification Rates



$$\bar{\delta}_{obs} = \frac{1}{n}\sum_{i=1}^{n}\delta_i$$

Ree (*Evolution*, 2005)

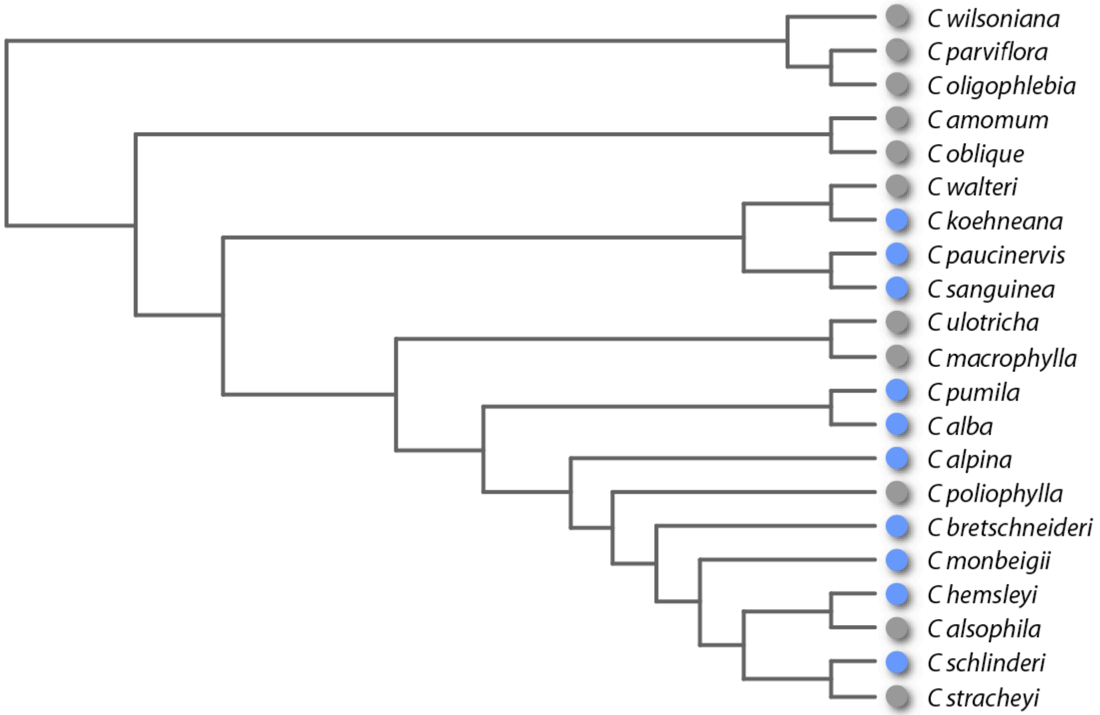
Evaluating Correlates of Differential Diversification Rates



$\bar{\delta}_{null\,1}$

Ree (***Evolution***, 2005)

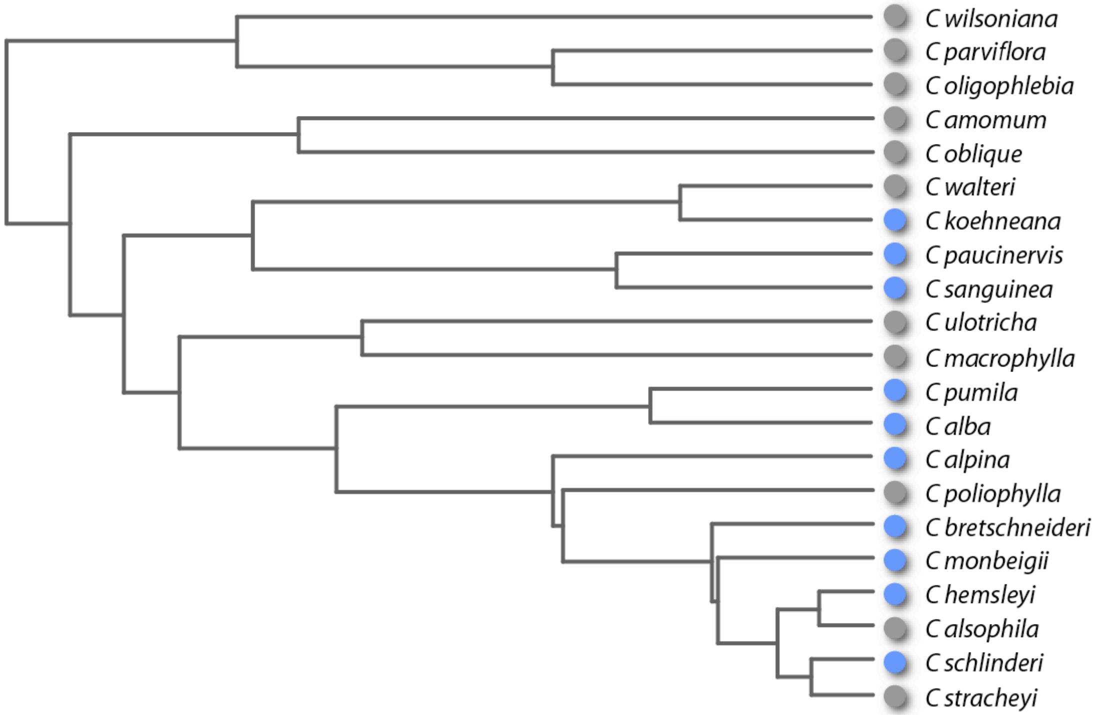
Evaluating Correlates of Differential Diversification Rates



$\bar{\delta}_{null\,2}$

Ree (***Evolution***, 2005)

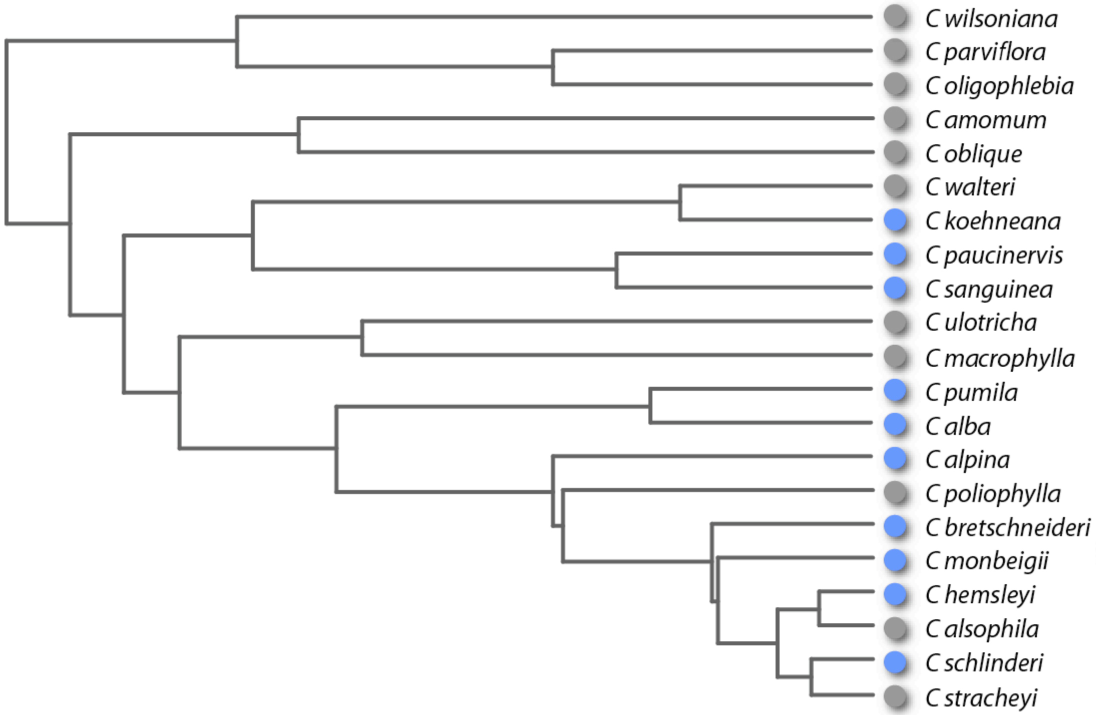
Evaluating Correlates of Differential Diversification Rates



$\bar{\delta}_{null\,r}$

Ree (***Evolution***, 2005)

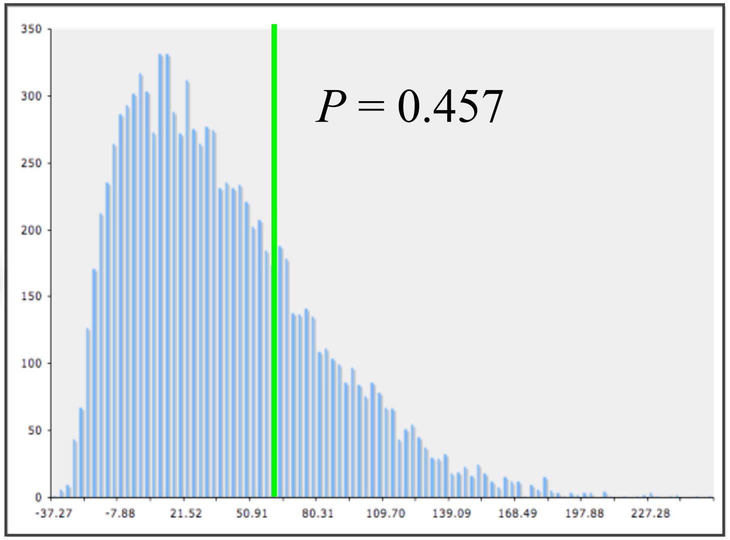
Evaluating Correlates of Differential Diversification Rates



$\bar{\delta}_{obs}$

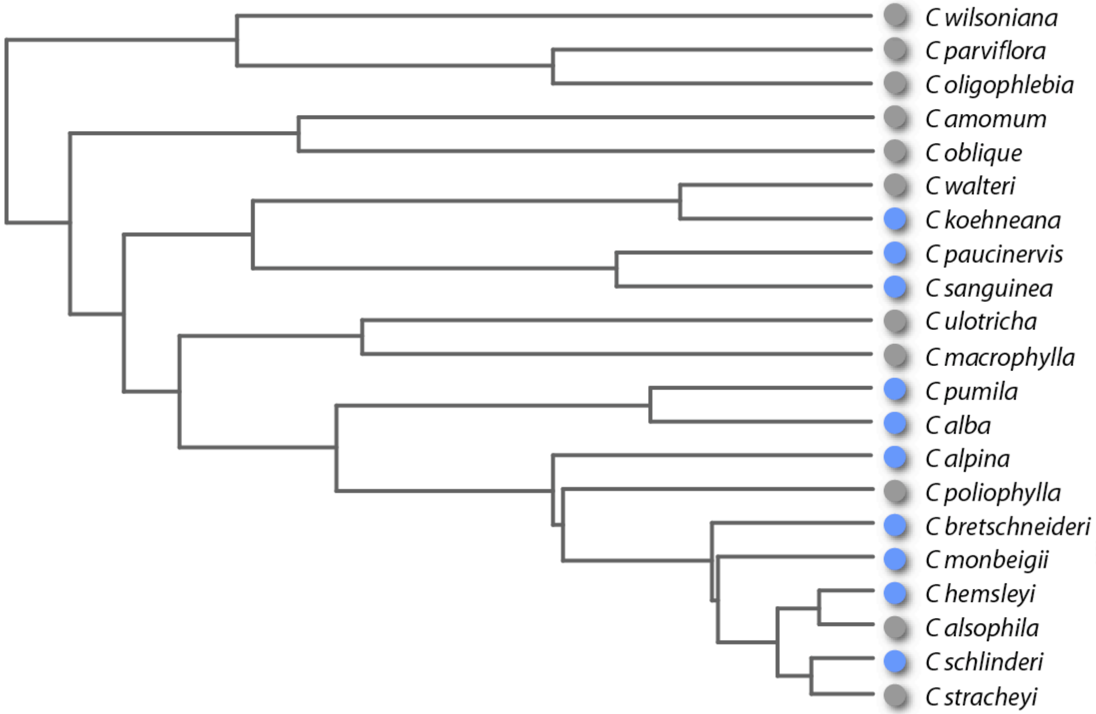
Trait A

$f\left(\bar{\delta}_{null}\right)$



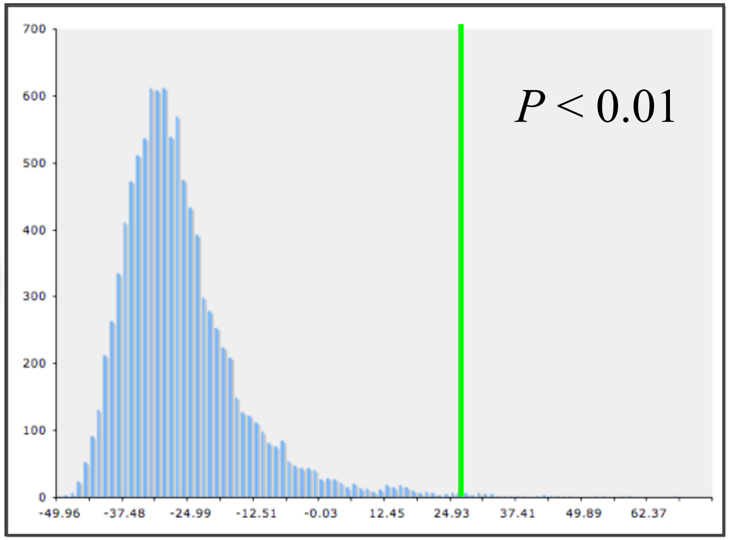
Ree (*Evolution*, 2005)

Evaluating Correlates of Differential Diversification Rates



Trait B

$\bar{\delta}_{obs}$

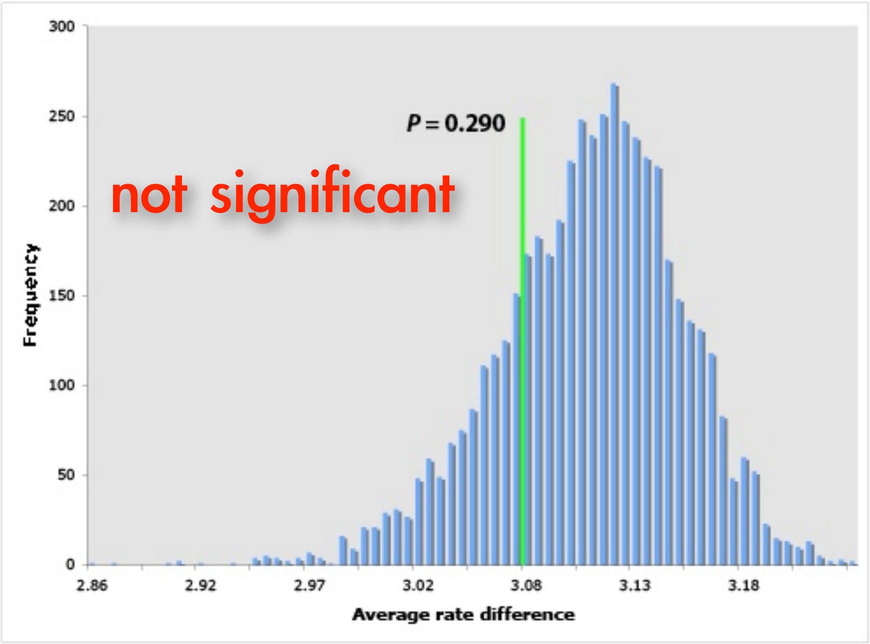


$f\left(\bar{\delta}_{null}\right)$

Ree (*Evolution*, 2005)

Fruit Type

● 1 seed
● >1 seed



not significant

Evaluating Correlates of Differential Diversification Rates

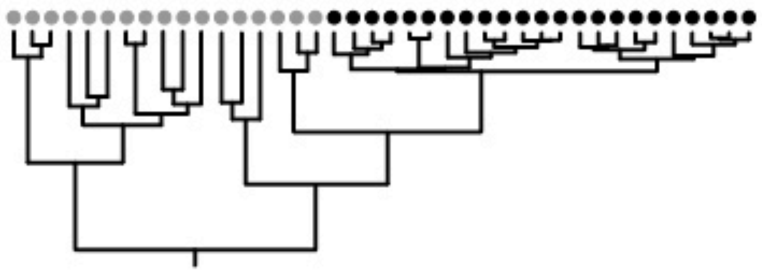
Recent advances for evaluating diversification-rate correlates

Stochastic character mapping; keymap (Ree, 2005)

CVPPD; tRate (Moore & Donoghue, 2009)

*SSE model; DiversiTree (Maddison et al., 2007; Fitzjohn et al., 2009; Fitzjohn, 2010...)

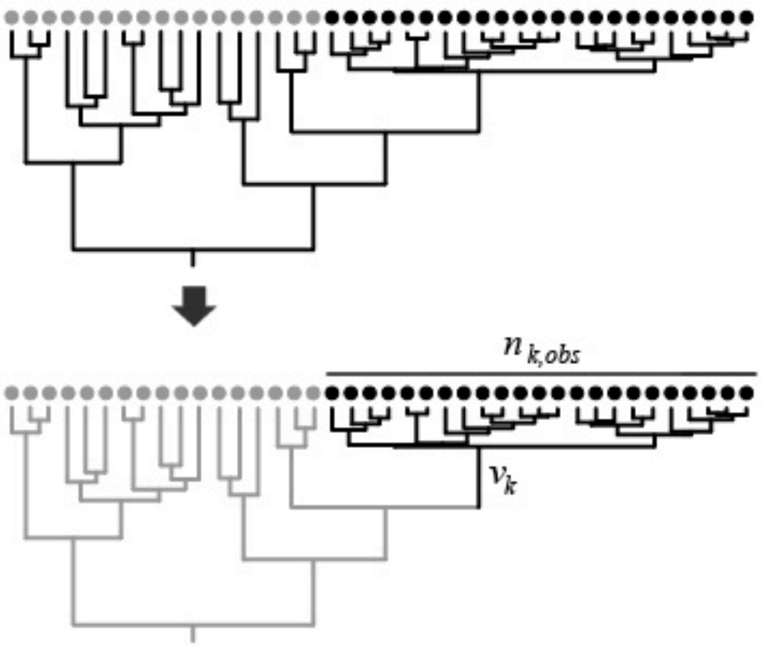
Cross-Validation Predictive Densities



Step 1. Jointly estimate phylogeny and divergence times

Moore & Donoghue (*PNAS*, 2009)

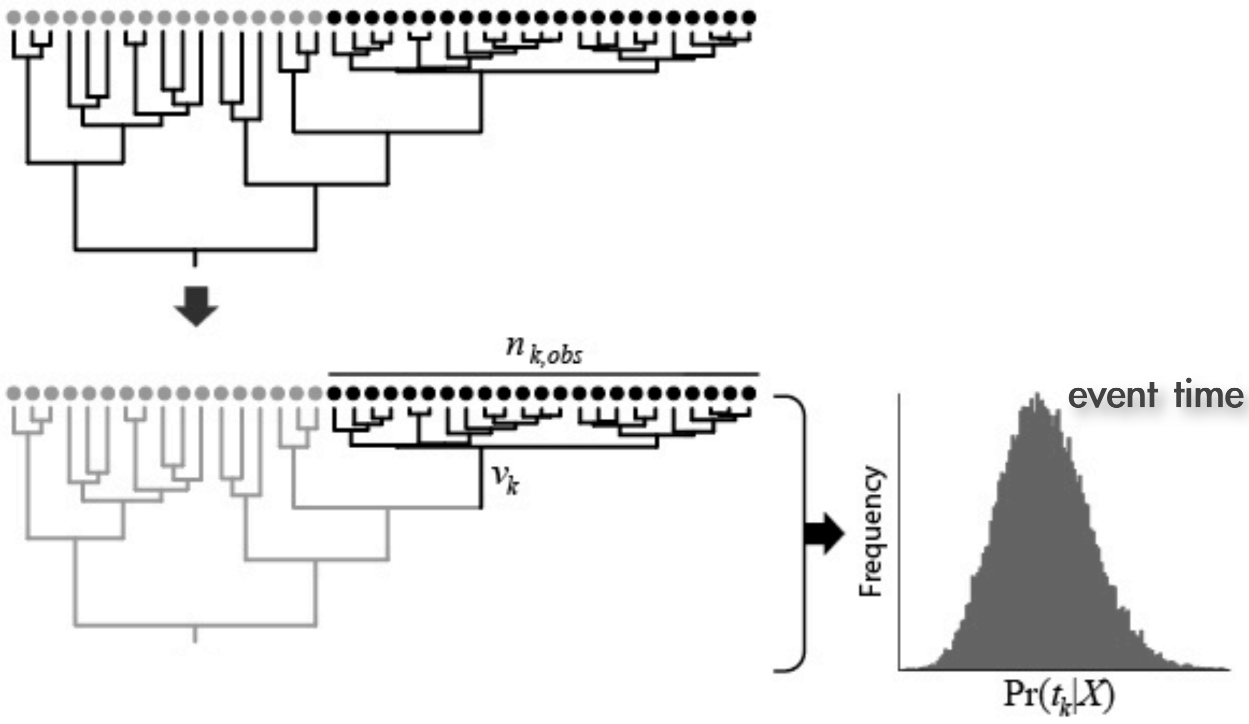
Cross-Validation Predictive Densities



Step 2. Estimate event history

Moore & Donoghue (*PNAS*, 2009)

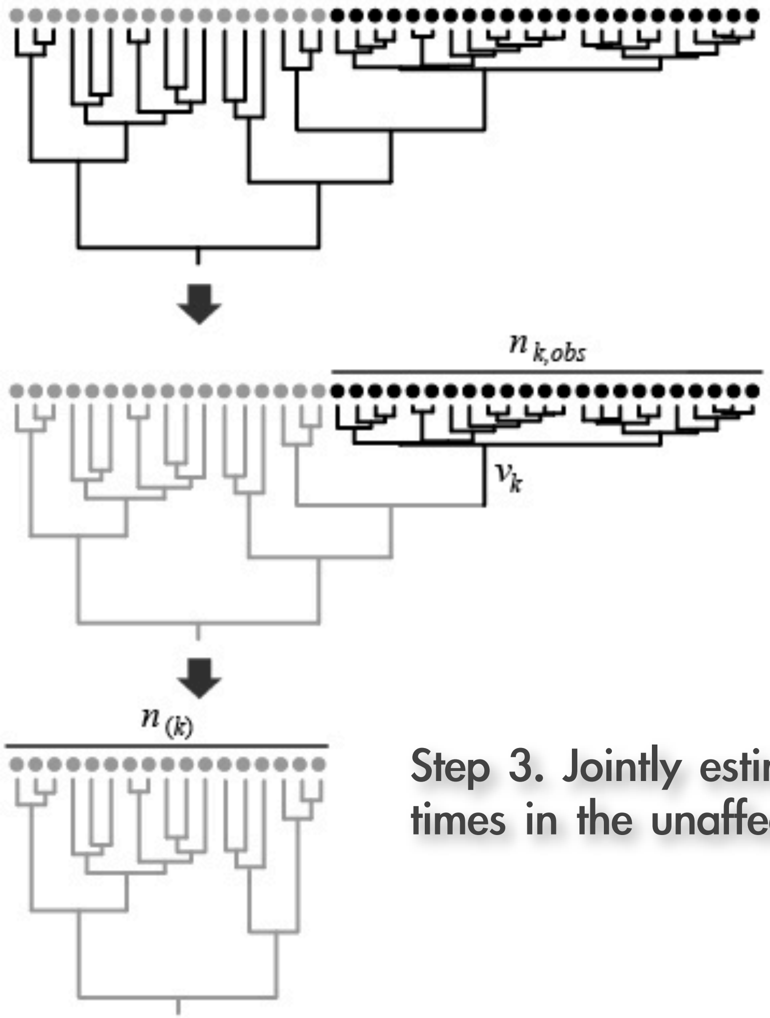
Cross-Validation Predictive Densities

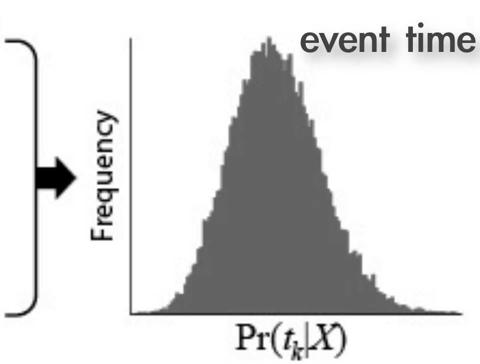


event time

Moore & Donoghue (*PNAS*, 2009)

Cross-Validation Predictive Densities



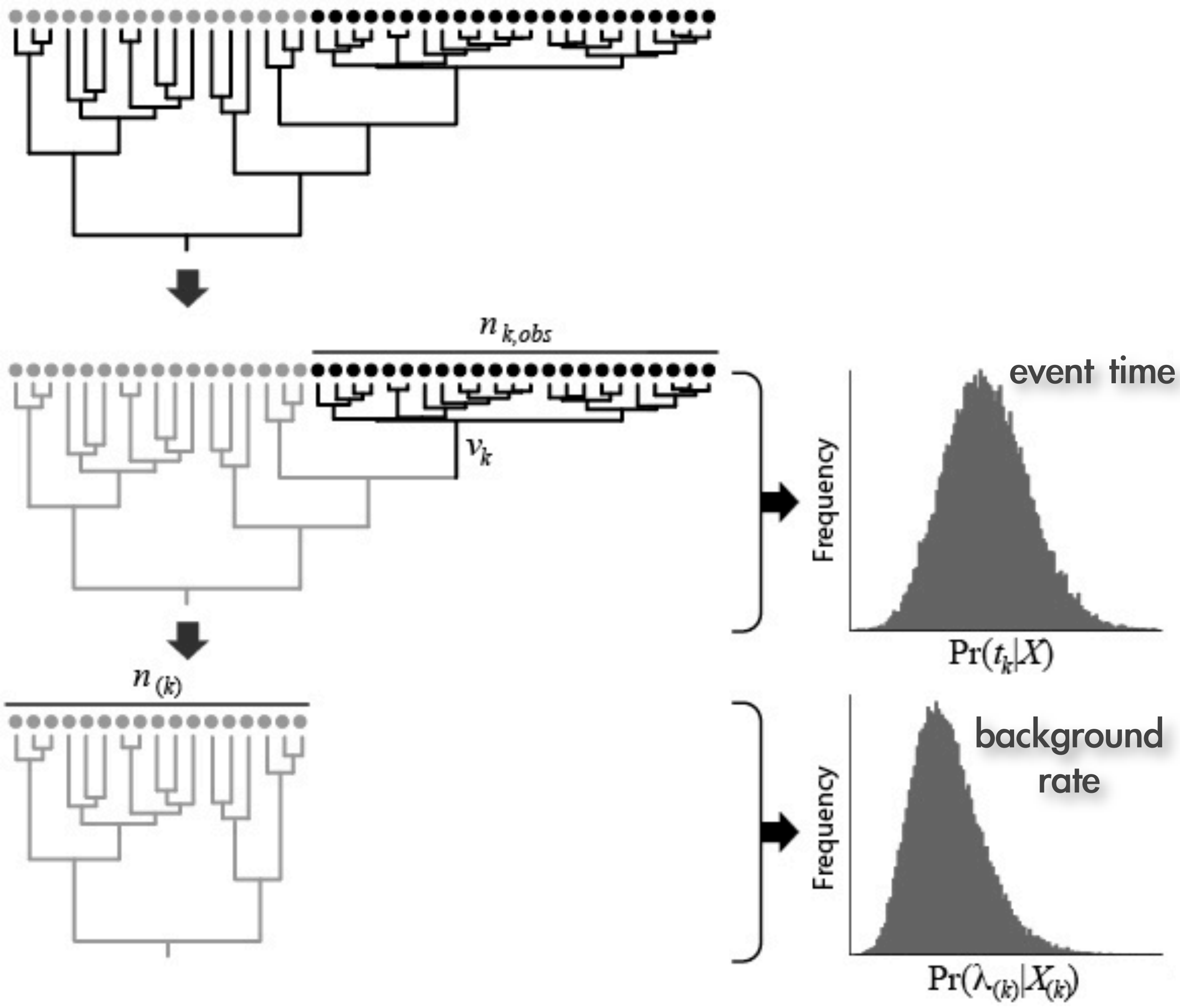


event time

Step 3. Jointly estimate phylogeny and divergence times in the unaffected part of the tree

Moore & Donoghue (*PNAS*, 2009)

Cross-Validation Predictive Densities

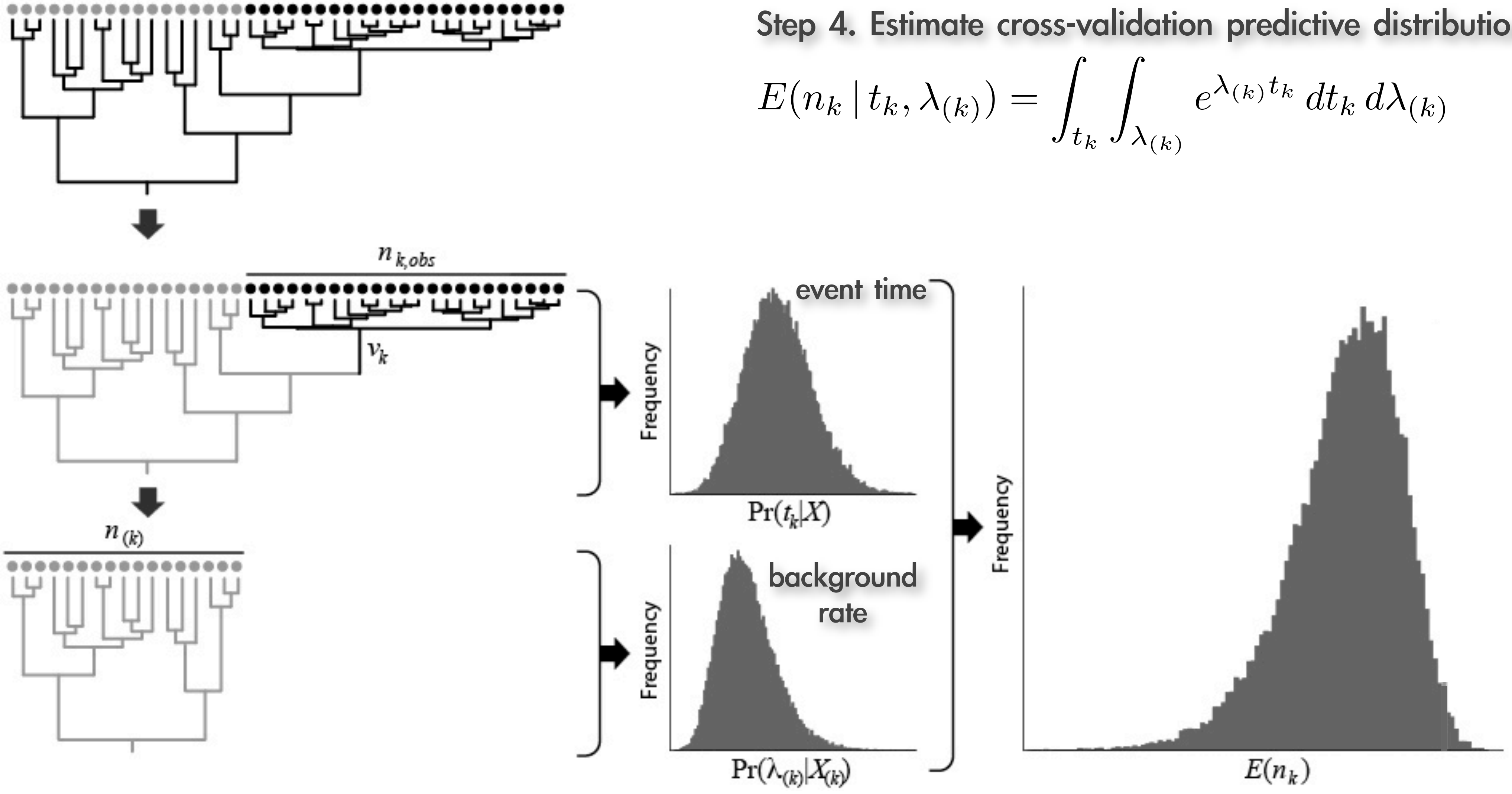


Moore & Donoghue (*PNAS*, 2009)

Cross-Validation Predictive Densities

Step 4. Estimate cross-validation predictive distribution

$$E(n_k \,|\, t_k, \lambda_{(k)}) = \int_{t_k} \int_{\lambda_{(k)}} e^{\lambda_{(k)} t_k} \, dt_k \, d\lambda_{(k)}$$



event time

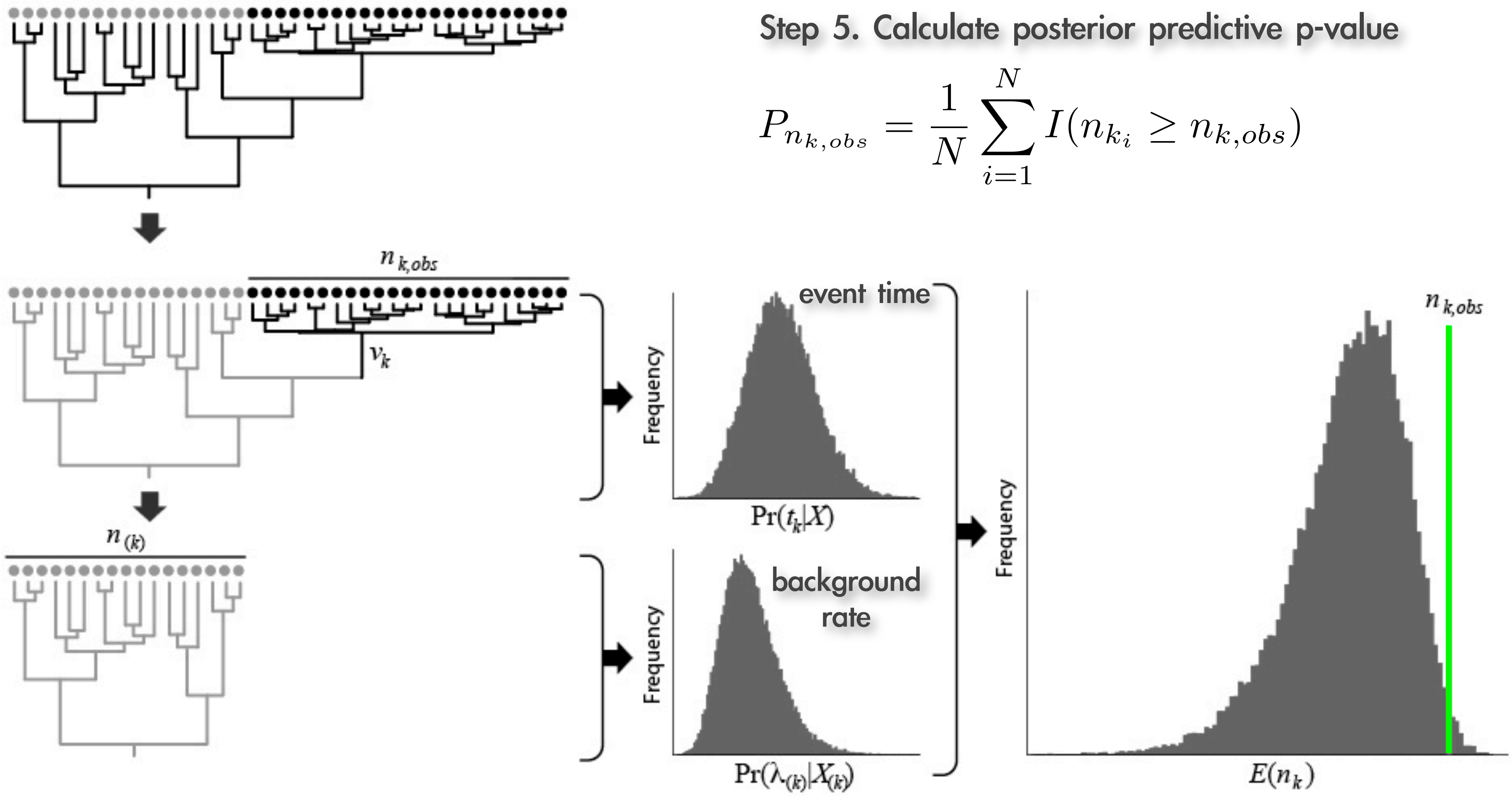
background rate

Moore & Donoghue (*PNAS*, 2009)

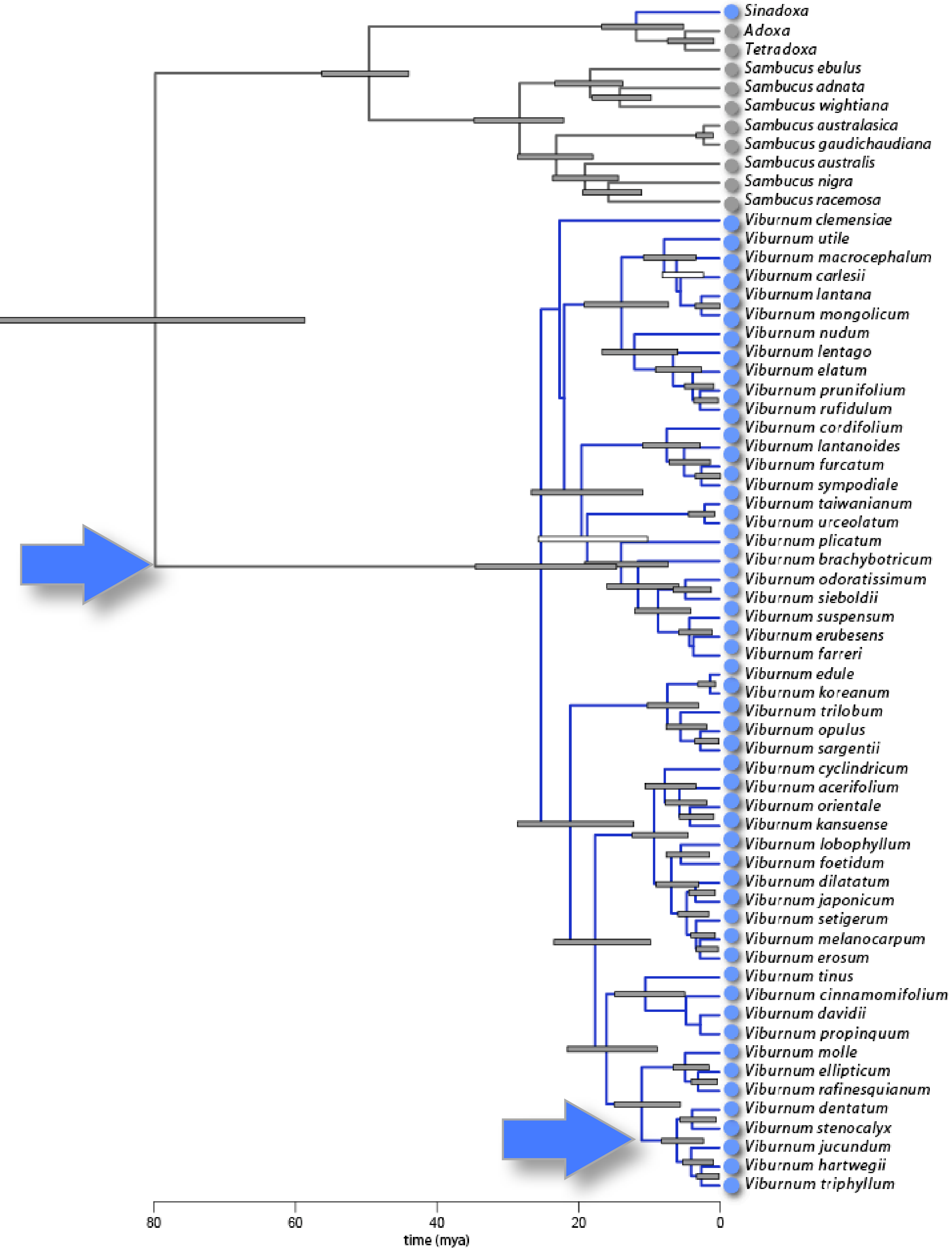
Cross-Validation Predictive Densities

Step 5. Calculate posterior predictive p-value

$$P_{n_{k,obs}} = \frac{1}{N}\sum_{i=1}^{N} I(n_{k_i} \geq n_{k,obs})$$



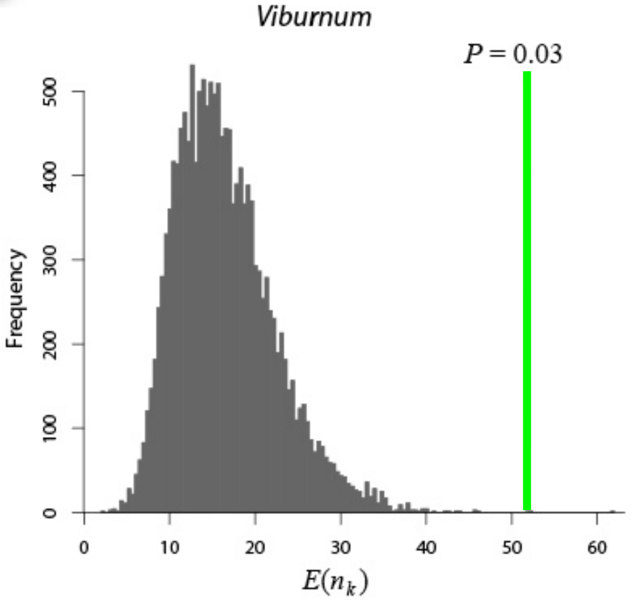
Moore & Donoghue (*PNAS*, 2009)

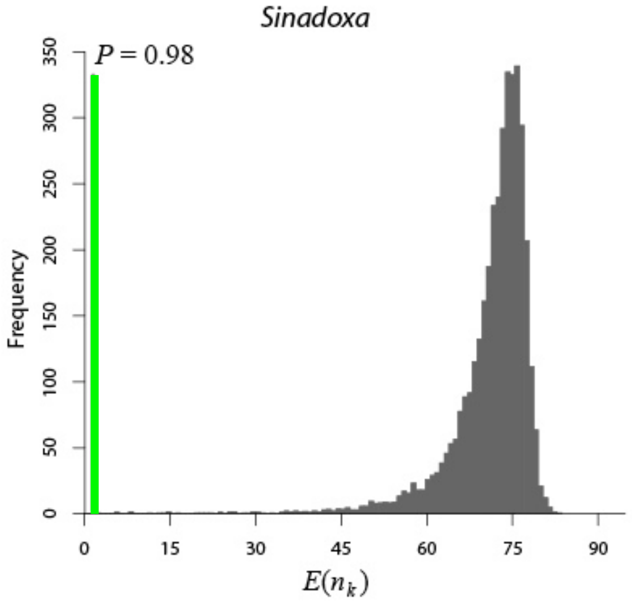


Fruit Type

● 1 seed

● >1 seed





Evaluating Correlates of Differential Diversification Rates

Recent advances for evaluating diversification-rate correlates

Stochastic character mapping; keymap (Ree, 2005)

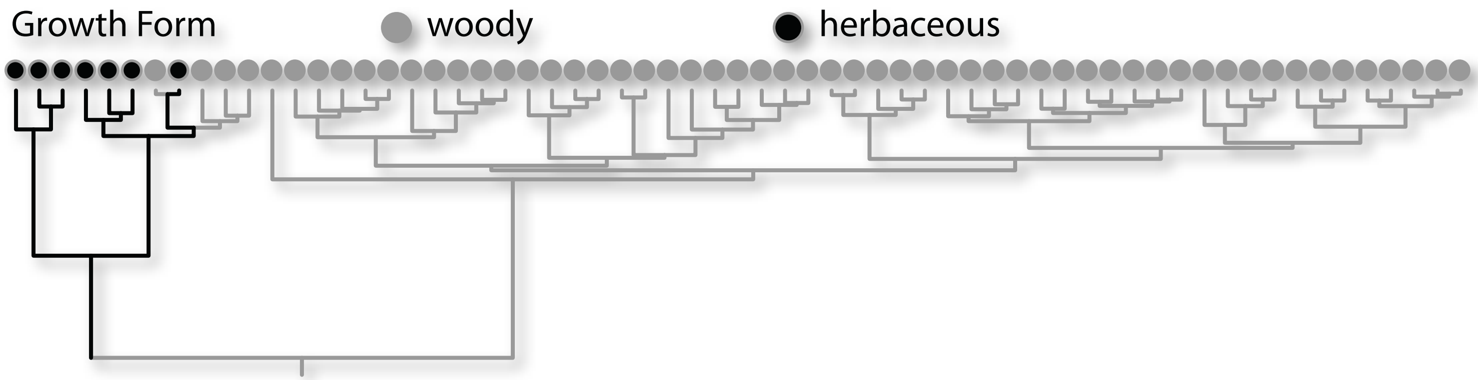
CVPPD; tRate (Moore & Donoghue, 2009)

➡ *SSE model; DiversiTree (Maddison et al., 2007; Fitzjohn et al., 2009; Fitzjohn, 2010...)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

Motivation: Imagine that we see a preponderance of species with a given state



Explanations include:

increased woody speciation rate $\lambda_{\circ}$

decreased woody extinction rate $\mu_{\circ}$

decreased herb speciation rate $\lambda_{\bullet}$

increased herb extinction rate $\mu_{\bullet}$

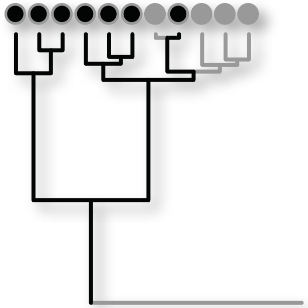
biassed exchangeability rate $q_{\bullet\circ}$

Maddison (***Evolution***, 2006)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

Solution: Specify a model that includes the parameters to tease these scenarios apart



λ_0	speciation rate for state 0
λ_1	speciation rate for state 1
μ_0	extinction rate for state 0
μ_1	extinction rate for state 1
q_{01}	rate of change to state 1
q_{10}	rate of change to state 0

Solve numerically with PDEs:

$$\begin{aligned} D_{N0}(t+\delta t) = {} & (1-\mu_0\delta t)[(1-q_{01}\delta t)(1-\lambda_0\delta t)D_{N0}(t) + (q_{01}\delta t)(1-\lambda_0\delta t)D_{N1}(t) \\ & +(1-q_{01}\delta t)(\lambda_0\delta t)E_0(t)D_{N0}(t) + (1-q_{01}\delta t)(\lambda_0\delta t)E_0(t)D_{N0}(t)] + (\mu_0\delta t)0 \end{aligned}$$

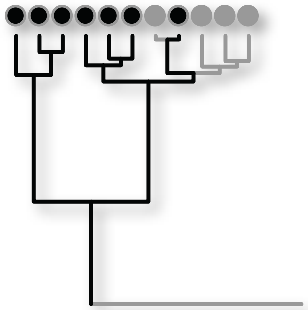
$$\begin{aligned} D_{N1}(t+\delta t) = {} & (1-\mu_1\delta t)[(1-q_{10}\delta t)(1-\lambda_1\delta t)D_{N1}(t) + (q_{10}\delta t)(1-\lambda_1\delta t)D_{N0}(t) \\ & +(1-q_{10}\delta t)(\lambda_1\delta t)E_1(t)D_{N1}(t) + (1-q_{10}\delta t)(\lambda_1\delta t)E_1(t)D_{N1}(t)] + (\mu_1\delta t)0 \end{aligned}$$

Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

Solution: Specify a model that includes the parameters to tease these scenarios apart



λ_0	speciation rate for state 0
λ_1	speciation rate for state 1
μ_0	extinction rate for state 0
μ_1	extinction rate for state 1
q_{01}	rate of change to state 1
q_{10}	rate of change to state 0

Solve numerically with PDEs:

$$E_0(t+\delta t) = \mu_0\delta t + (1-\mu_0\delta t)(1-q_{01}\delta t)(1-\lambda_0\delta t)E_0(t) + (1-\mu_0\delta t)$$
$$(q_{01}\delta t)(1-\lambda_0\delta t)E_1(t) + (1-\mu_0\delta t)(1-q_{01}\delta t)(\lambda_0\delta t)E_0(t)^2$$

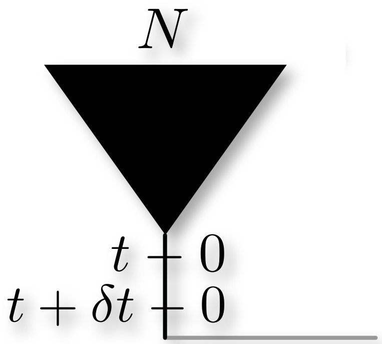
$$E_1(t+\delta t) = \mu_1\delta t + (1-\mu_1\delta t)(1-q_{10}\delta t)(1-\lambda_1\delta t)E_1(t) + (1-\mu_1\delta t)$$
$$(q_{10}\delta t)(1-\lambda_1\delta t)E_0(t) + (1-\mu_1\delta t)(1-q_{10}\delta t)(\lambda_1\delta t)E_1(t)^2$$

Maddison et al. (***Syst Biol***, 2007)

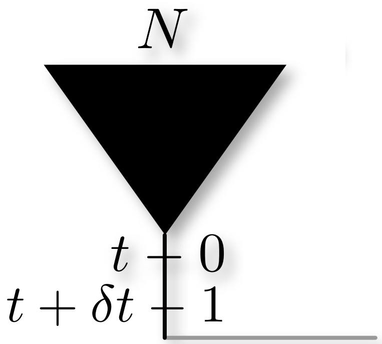
Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

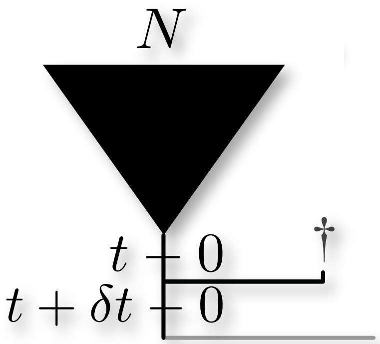
no state change,
no speciation



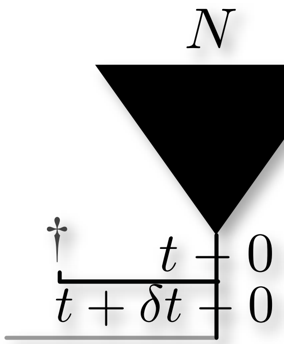
state change,
no speciation



no state change,
speciation & extinction



no state change,
speciation & extinction



Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

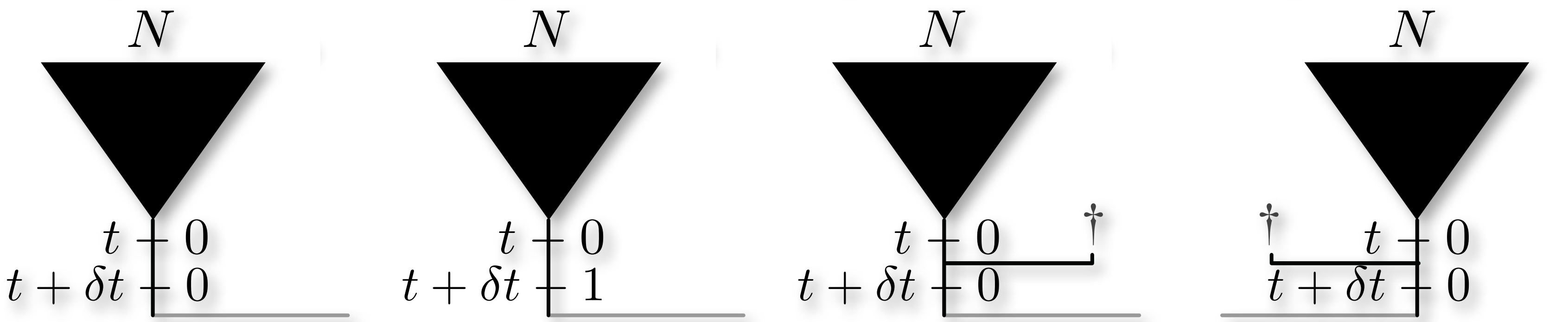
Binary-State Speciation Extinction (BiSSE) Model

no state change,
no speciation

state change,
no speciation

no state change,
speciation & extinction

no state change,
speciation & extinction



PDEs for branch probabilities:

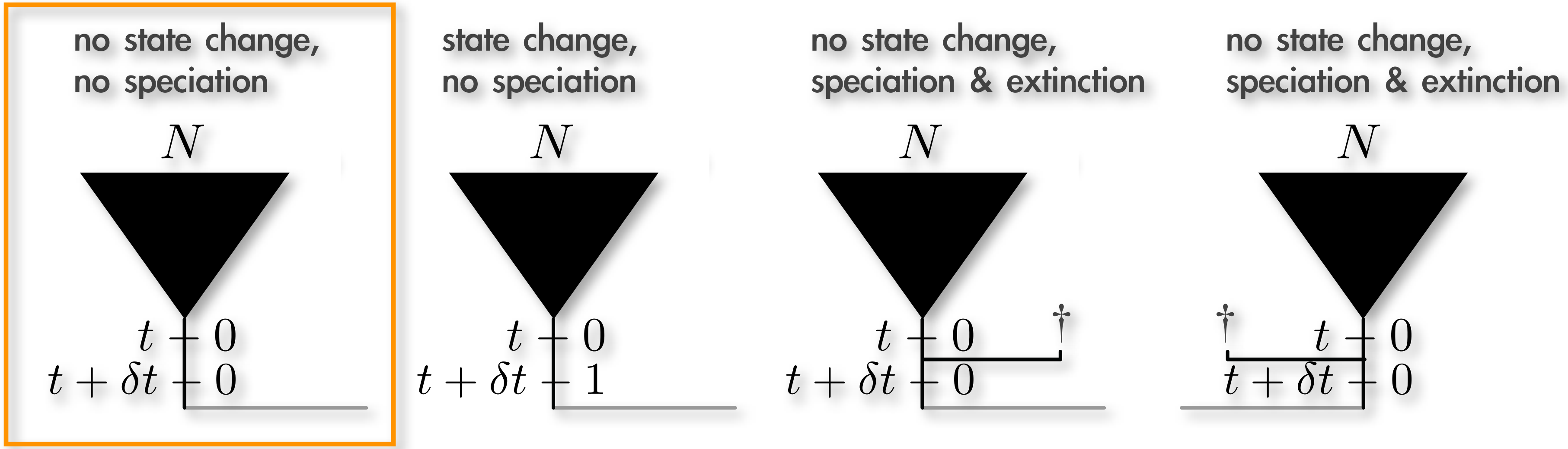
$$D_{N0}(t+\delta t) =$$

$$(1-\mu_0\delta t)$$ no extinction in the interval

Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model



no state change,
no speciation

state change,
no speciation

no state change,
speciation & extinction

no state change,
speciation & extinction

PDEs for branch probabilities: $D_{N0}(t + \delta t) =$

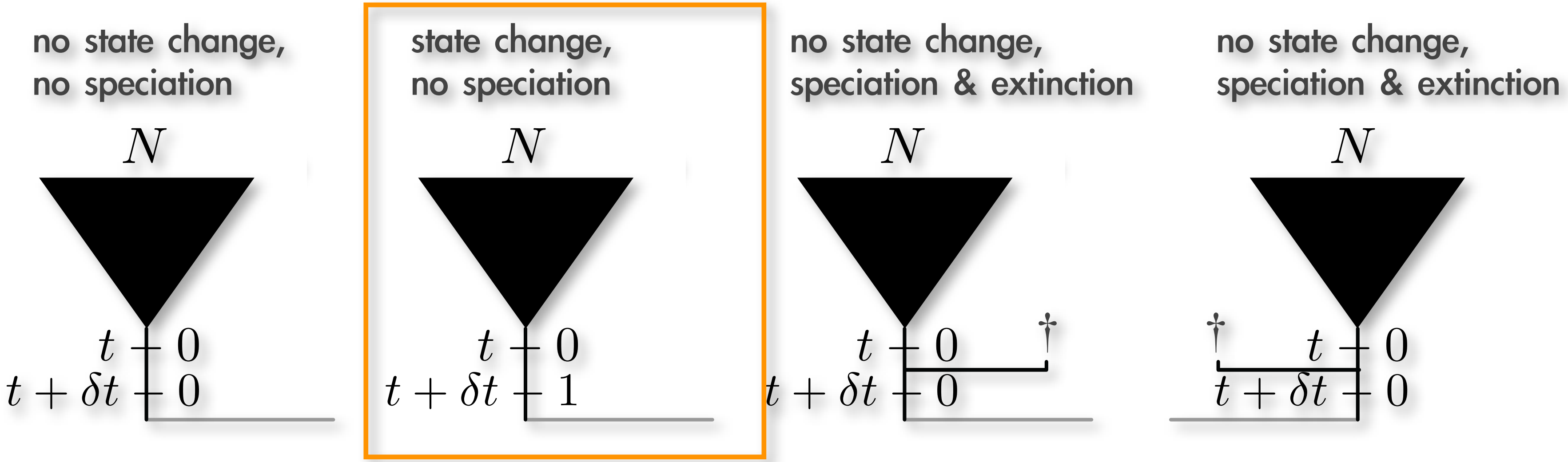
$(1 - \mu_0 \delta t)$ no extinction in the interval

$[(1 - q_{01} \delta t)(1 - \lambda_0 \delta t) D_{N0}(t)$ no state change, no speciation

Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model


no state change, no speciation

state change, no speciation

no state change, speciation & extinction

no state change, speciation & extinction

PDEs for branch probabilities: $D_{N0}(t+\delta t) =$

$(1-\mu_0\delta t)$ no extinction in the interval

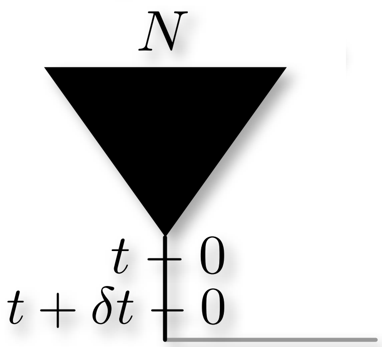
$[(1-q_{01}\delta t)(1-\lambda_0\delta t)D_{N0}(t)$ no state change, no speciation

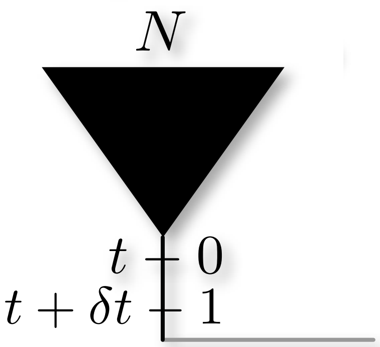
$+(q_{01}\delta t)(1-\lambda_0\delta t)D_{N1}(t)$ state change, no speciation

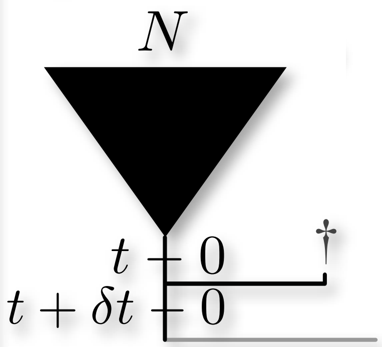
Maddison et al. (***Syst Biol***, 2007)

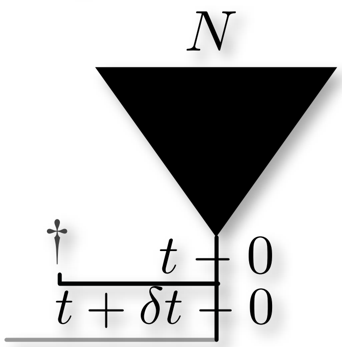
Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model


no state change,
no speciation


state change,
no speciation


no state change,
speciation & extinction


no state change,
speciation & extinction

PDEs for branch probabilities:

$$D_{N0}(t + \delta t) =$$

$$(1 - \mu_0 \delta t)$$ no extinction in the interval

$$[(1 - q_{01}\delta t)(1 - \lambda_0 \delta t) D_{N0}(t)$$ no state change, no speciation

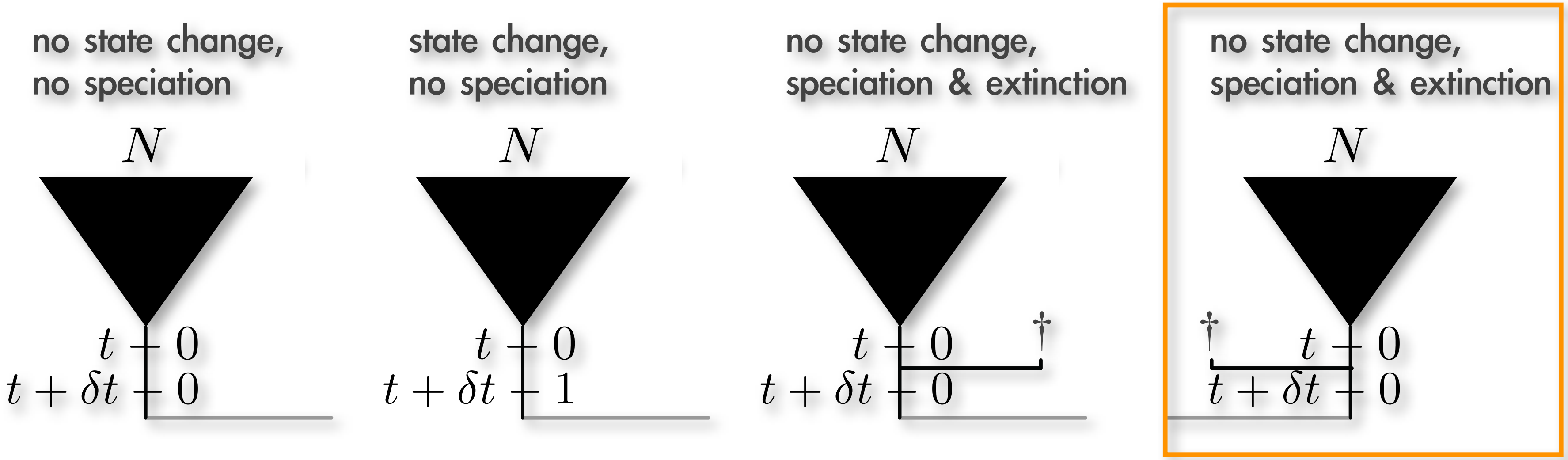
$$+(q_{01}\delta t)(1 - \lambda_0 \delta t) D_{N1}(t)$$ state change, no speciation

$$+(1 - q_{01}\delta t)(\lambda_0 \delta t) E_0(t) D_{N0}(t)$$ no state change, speciation & extinction

Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model



no state change,
no speciation

state change,
no speciation

no state change,
speciation & extinction

no state change,
speciation & extinction

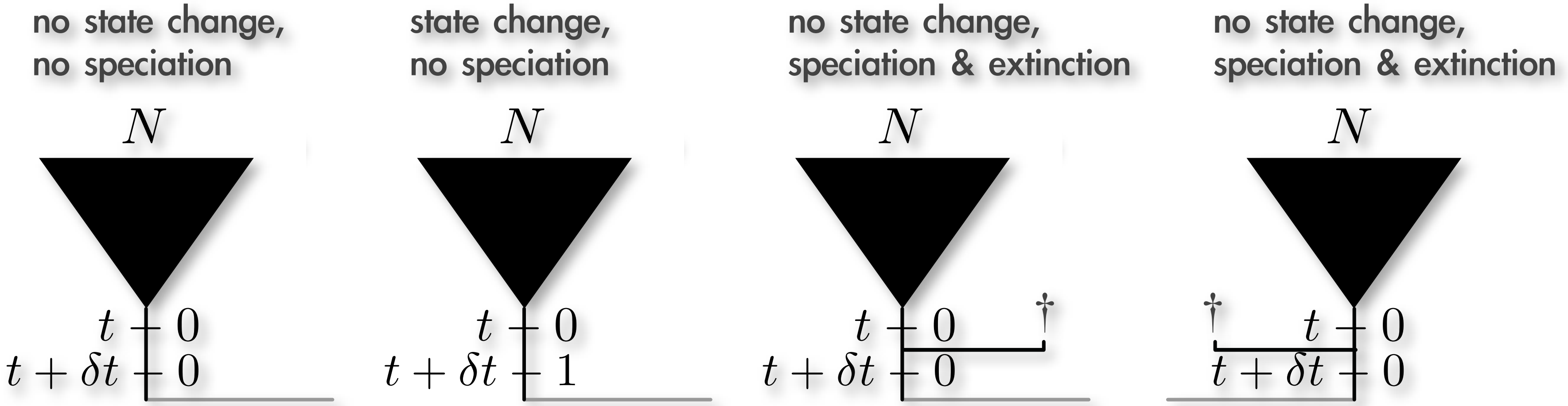
PDEs for branch probabilities: $D_{N0}(t + \delta t) =$

$(1 - \mu_0 \delta t)$	no extinction in the interval
$[(1 - q_{01}\delta t)(1 - \lambda_0 \delta t) D_{N0}(t)$	no state change, no speciation
$+(q_{01}\delta t)(1 - \lambda_0 \delta t) D_{N1}(t)$	state change, no speciation
$+(1 - q_{01}\delta t)(\lambda_0 \delta t) E_0(t) D_{N0}(t)$	no state change, speciation & extinction
$+(1 - q_{01}\delta t)(\lambda_0 \delta t) E_0(t) D_{N0}(t)]$	no state change, speciation & extinction

Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model


no state change, no speciation

state change, no speciation

no state change, speciation & extinction

no state change, speciation & extinction

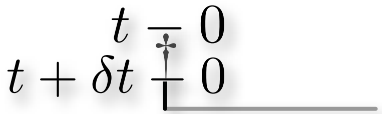
PDEs for branch probabilities: $D_{N0}(t+\delta t) =$

$(1-\mu_0\delta t)$	no extinction in the interval
$[(1-q_{01}\delta t)(1-\lambda_0\delta t)D_{N0}(t)$	no state change, no speciation
$+(q_{01}\delta t)(1-\lambda_0\delta t)D_{N1}(t)$	state change, no speciation
$+(1-q_{01}\delta t)(\lambda_0\delta t)E_0(t)D_{N0}(t)$	no state change, speciation & extinction
$+(1-q_{01}\delta t)(\lambda_0\delta t)E_0(t)D_{N0}(t)]$	no state change, speciation & extinction
$+(\mu_0\delta t)0$	if extinct, zero probability of being observed

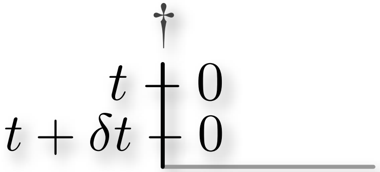
Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

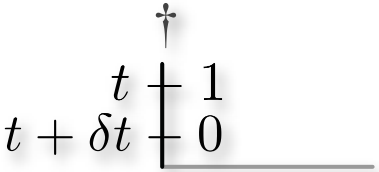
Binary-State Speciation Extinction (BiSSE) Model



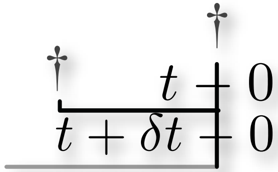
extinction in interval



no state change, no speciation, extinction since t



state change, no speciation, extinction since t



no state change, speciation, extinction of both since t

Extinction scenarios

Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

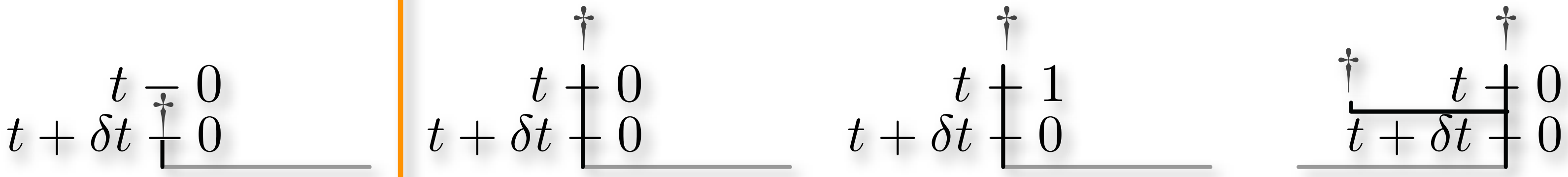
Binary-State Speciation Extinction (BiSSE) Model

extinction in interval

no state change, no speciation, extinction since t

state change, no speciation, extinction since t

no state change, speciation, extinction of both since t



PDEs for branch probabilities:

$$E_0(t + \delta t) =$$

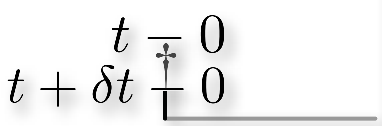
$\mu_0 \delta t$ extinction in the interval

Maddison et al. (*Syst Biol*, 2007)

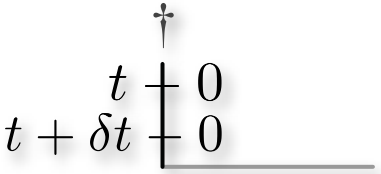
Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

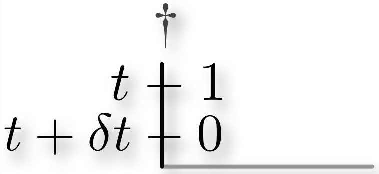
extinction in interval



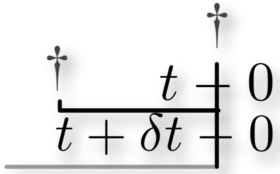
no state change, no speciation, extinction since t



state change, no speciation, extinction since t



no state change, speciation, extinction of both since t



PDEs for branch probabilities: $E_0(t+\delta t) =$

$\mu_0 \delta t$ extinction in the interval

$+(1-\mu_0\delta t)(1-q_{01}\delta t)(1-\lambda_0\delta t)E_0(t)$ no state change, no speciation

Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

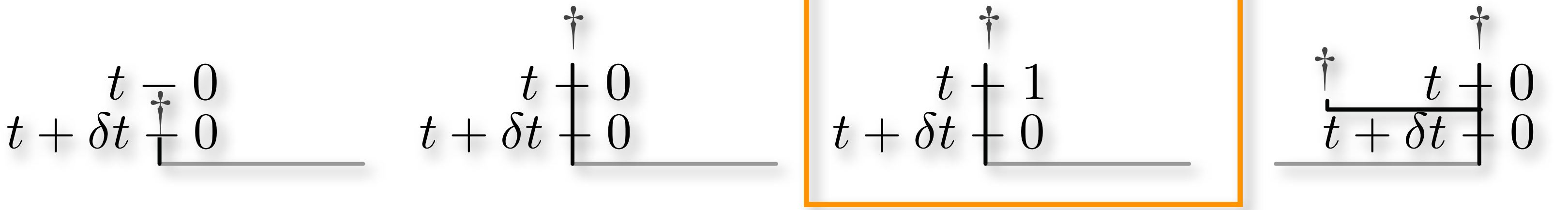
Binary-State Speciation Extinction (BiSSE) Model

extinction in interval

no state change, no speciation, extinction since t

state change, no speciation, extinction since t

no state change, speciation, extinction of both since t



PDEs for branch probabilities:

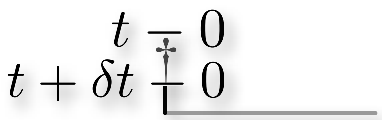
$$
\begin{aligned}
E_0(t + \delta t) &= \\
&\mu_0 \delta t && \text{extinction in the interval} \\
&+(1 - \mu_0 \delta t)(1 - q_{01} \delta t)(1 - \lambda_0 \delta t) E_0(t) && \text{no state change, no speciation} \\
&+(1 - \mu_0 \delta t)(q_{01} \delta t)(1 - \lambda_0 \delta t) E_1(t) && \text{state change, no speciation}
\end{aligned}
$$

Maddison et al. (*Syst Biol*, 2007)

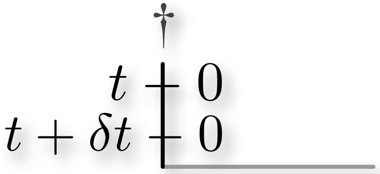
Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

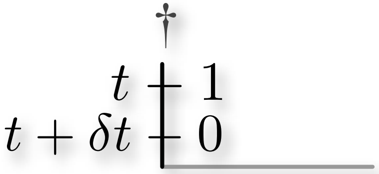
extinction in
interval



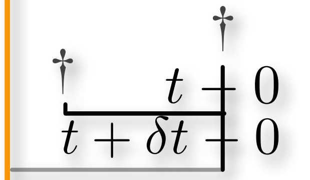
no state change,
no speciation,
extinction since t



state change,
no speciation,
extinction since t



no state change,
speciation, extinction
of both since t



PDEs for branch probabilities: $E_0(t + \delta t) =$

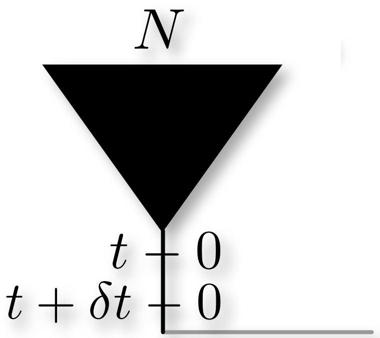
$\mu_0 \delta t$	extinction in the interval
$+(1 - \mu_0 \delta t)(1 - q_{01} \delta t)(1 - \lambda_0 \delta t) E_0(t)$	no state change, no speciation
$+(1 - \mu_0 \delta t)(q_{01} \delta t)(1 - \lambda_0 \delta t) E_1(t)$	state change, no speciation
$+(1 - \mu_0 \delta t)(1 - q_{01} \delta t)(\lambda_0 \delta t) E_0(t)^2$	no state change, speciation

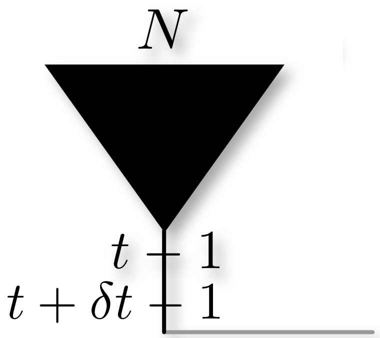
Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

We start at the tips of the tree and move root-ward in small increments...





If N is a tip species with state 1:

$$D_{N0}(t_0) = 0$$

$$D_{N1}(t_0) = 1$$

If N is a tip species with state 0:

$$D_{N0}(t_0) = 1$$

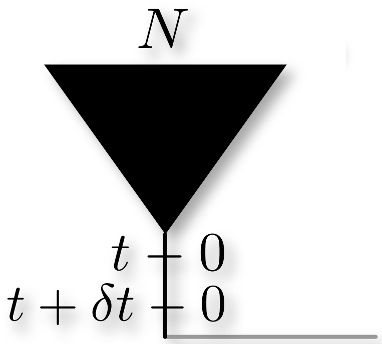
$$D_{N1}(t_0) = 0$$

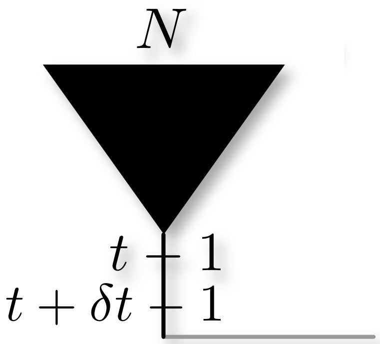
Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

solving coupled differential equations to obtain conditional branch probabilities...





We take the derivative of the PDEs to shrink δt:

$$\frac{dD_{N0}}{dt} = -(\lambda_0 + \mu_0 + q_{01})D_{N0}(t) + q_{01}D_{N1}(t) + 2\lambda_0 E_0(t) D_{N0}(t)$$

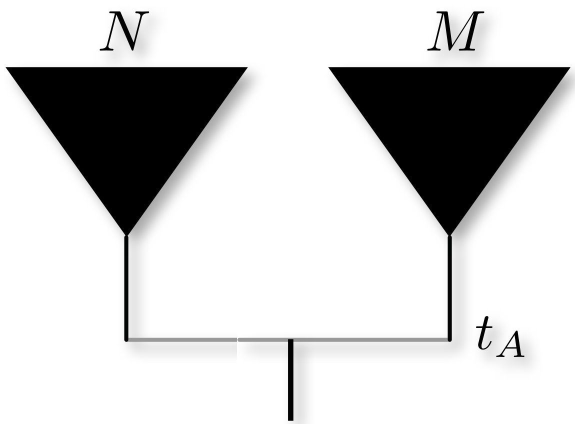
$$\frac{dD_{N1}}{dt} = -(\lambda_1 + \mu_1 + q_{10})D_{N1}(t) + q_{10}D_{N0}(t) + 2\lambda_1 E_1(t) D_1(t)$$

Maddison et al. (*Syst Biol*, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

and then combine branch probabilities at internal nodes...



$$D_{A0}(t_A) = D_{N0}(t_A)D_{M0}(t_A)\lambda_0$$

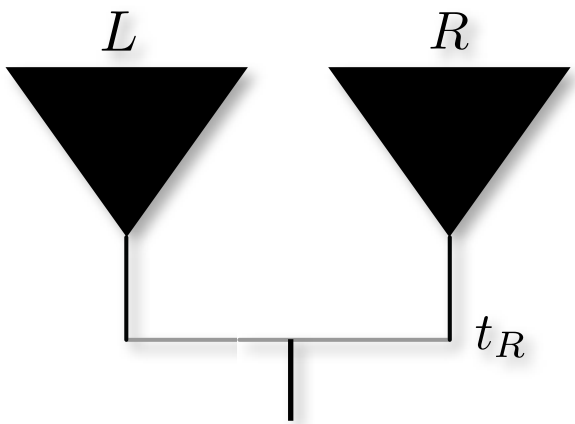
$$D_{A1}(t_A) = D_{N1}(t_A)D_{M1}(t_A)\lambda_1$$

Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

Binary-State Speciation Extinction (BiSSE) Model

and at the root scale the conditional probabilities by the stationary frequencies



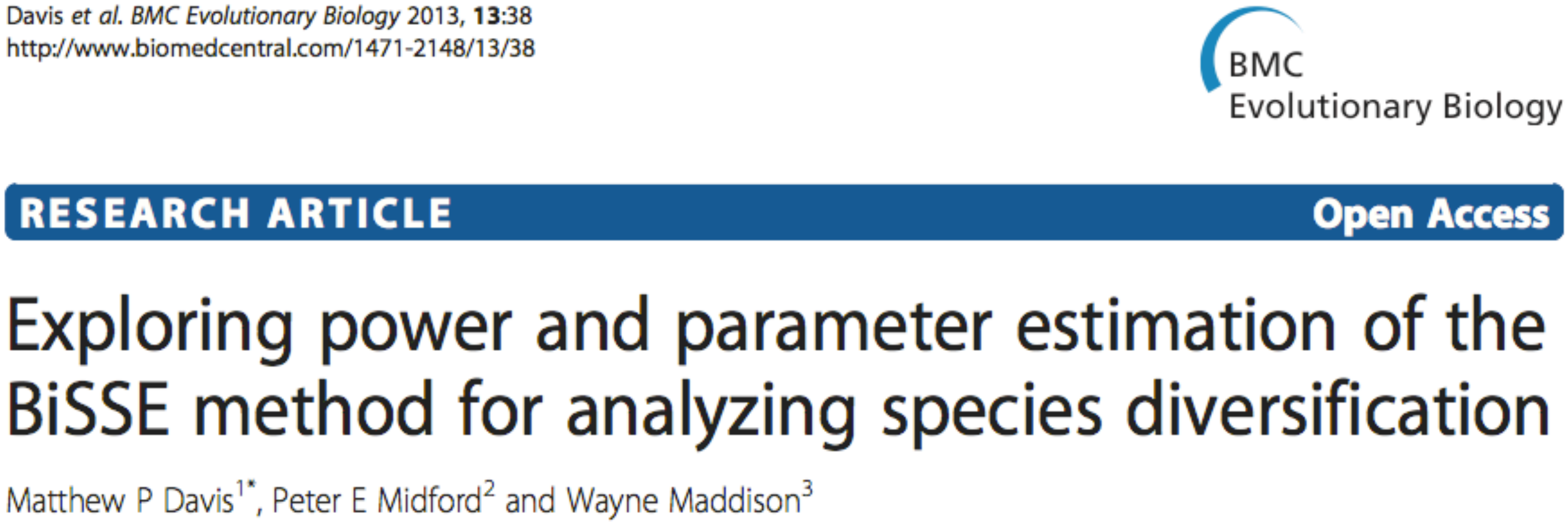
$$D_0 = D_{R0}(t_R)\pi_0$$

$$D_1 = D_{R1}(t_R)\pi_1$$

Maddison et al. (***Syst Biol***, 2007)

Evaluating Correlates of Differential Diversification Rates

How well can we do inference under the BiSSE model?



Davis *et al. BMC Evolutionary Biology* 2013, **13**:38
http://www.biomedcentral.com/1471-2148/13/38

BMC
Evolutionary Biology

RESEARCH ARTICLE **Open Access**

Exploring power and parameter estimation of the BiSSE method for analyzing species diversification

Matthew P Davis[1*], Peter E Midford[2] and Wayne Maddison[3]

Davis et al. (***BMC Biology***, 2013)

Evaluating Correlates of Differential Diversification Rates

How well can we do inference under the BiSSE model?

Set 1:

$$\lambda_{\color{gray}\bullet} \neq \lambda_{\bullet}$$
$$\mu_{\color{gray}\bullet} = \mu_{\bullet}$$
$$q_{\bullet\color{gray}\bullet} = q_{\color{gray}\bullet\color{black}\bullet}$$

Set 2:

$$\lambda_{\color{gray}\bullet} = \lambda_{\bullet}$$
$$\mu_{\color{gray}\bullet} = \mu_{\bullet}$$
$$q_{\bullet\color{gray}\bullet} \neq q_{\color{gray}\bullet\color{black}\bullet}$$

Set 3:

$$\lambda_{\color{gray}\bullet} = \lambda_{\bullet}$$
$$\mu_{\color{gray}\bullet} \neq \mu_{\bullet}$$
$$q_{\bullet\color{gray}\bullet} = q_{\color{gray}\bullet\color{black}\bullet}$$

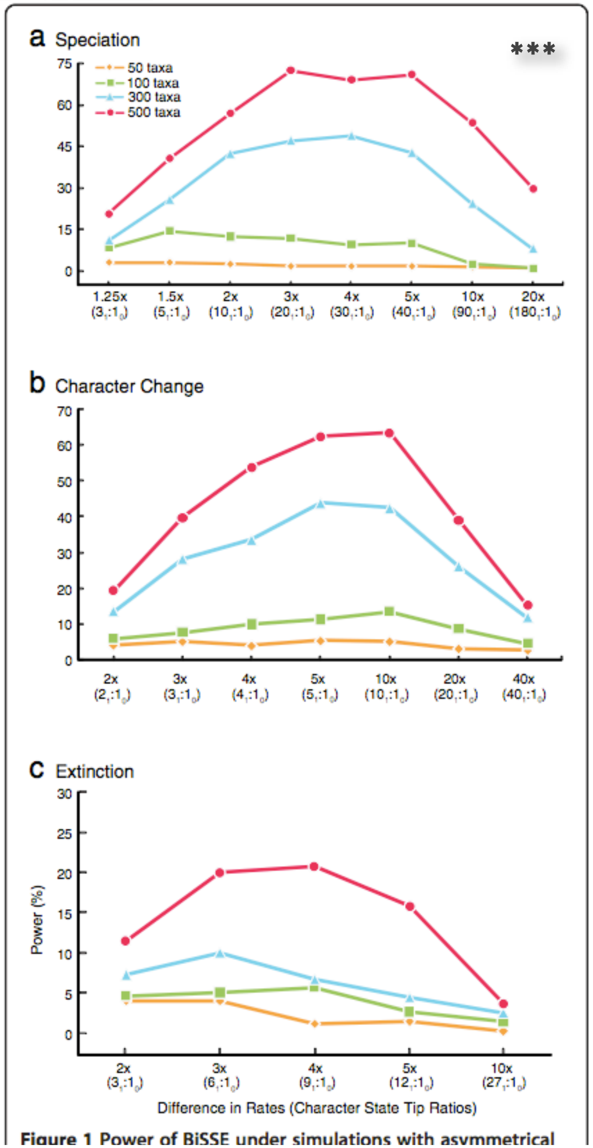


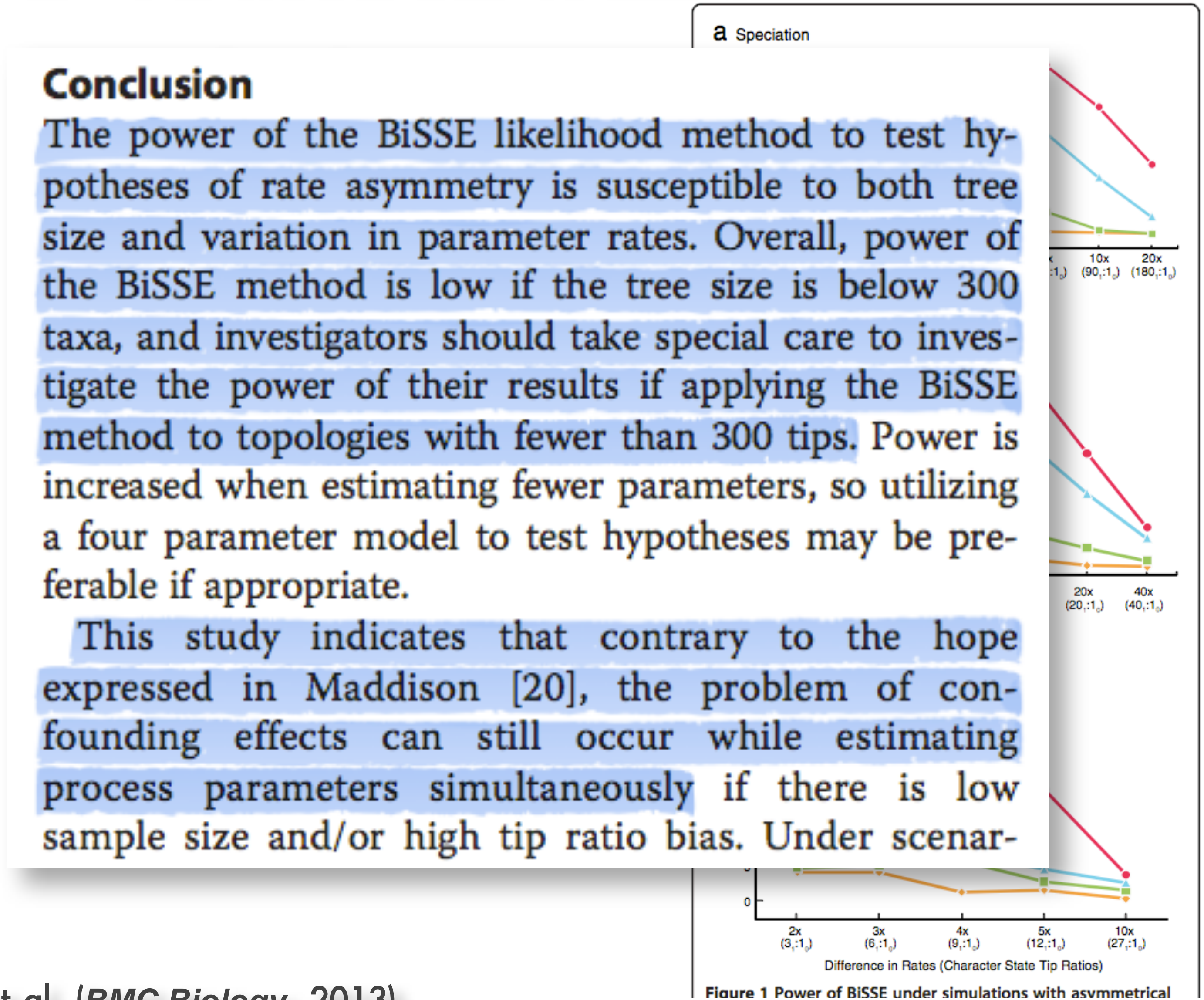
Figure 1 Power of BiSSE under simulations with asymmetrical

***Assuming no phylogenetic uncertainty

Davis et al. (*BMC Biology*, 2013)

Evaluating Correlates of Differential Diversification Rates

How well can we do inference under the BiSSE model?



Conclusion

The power of the BiSSE likelihood method to test hypotheses of rate asymmetry is susceptible to both tree size and variation in parameter rates. Overall, power of the BiSSE method is low if the tree size is below 300 taxa, and investigators should take special care to investigate the power of their results if applying the BiSSE method to topologies with fewer than 300 tips. Power is increased when estimating fewer parameters, so utilizing a four parameter model to test hypotheses may be preferable if appropriate.

This study indicates that contrary to the hope expressed in Maddison [20], the problem of confounding effects can still occur while estimating process parameters simultaneously if there is low sample size and/or high tip ratio bias. Under scenar-

Figure 1 Power of BiSSE under simulations with asymmetrical

Davis et al. (*BMC Biology*, 2013)

Evaluating Correlates of Differential Diversification Rates

How well can we do inference under the BiSSE model?

Set 1: $\lambda_{\bullet} = \lambda_{\bullet}$

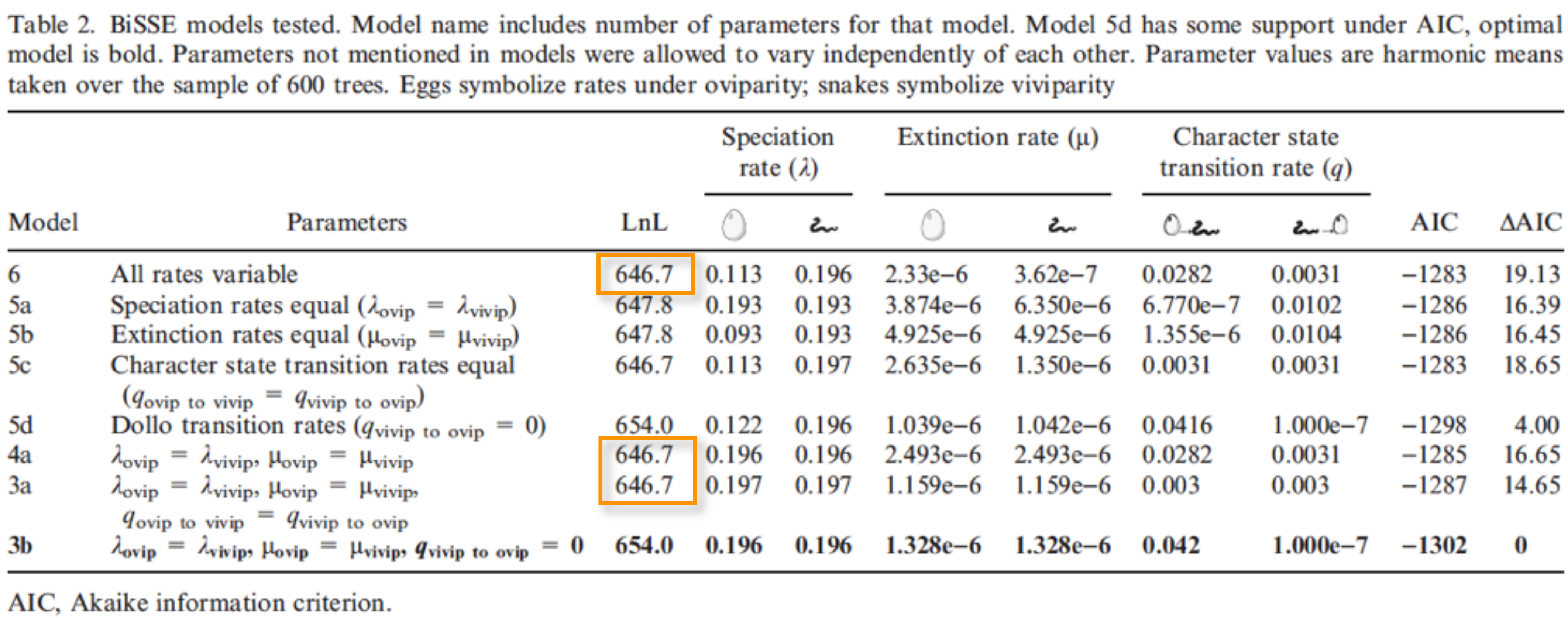


Table 2. BiSSE models tested. Model name includes number of parameters for that model. Model 5d has some support under AIC, optimal model is bold. Parameters not mentioned in models were allowed to vary independently of each other. Parameter values are harmonic means taken over the sample of 600 trees. Eggs symbolize rates under oviparity; snakes symbolize viviparity

			Speciation rate (λ)		Extinction rate (μ)		Character state transition rate (q)			
Model	Parameters	LnL	🥚	🐍	🥚	🐍	🥚→🐍	🐍→🥚	AIC	ΔAIC
6	All rates variable	646.7	0.113	0.196	2.33e−6	3.62e−7	0.0282	0.0031	−1283	19.13
5a	Speciation rates equal ($\lambda_{ovip} = \lambda_{vivip}$)	647.8	0.193	0.193	3.874e−6	6.350e−6	6.770e−7	0.0102	−1286	16.39
5b	Extinction rates equal ($\mu_{ovip} = \mu_{vivip}$)	647.8	0.093	0.193	4.925e−6	4.925e−6	1.355e−6	0.0104	−1286	16.45
5c	Character state transition rates equal ($q_{ovip\ to\ vivip} = q_{vivip\ to\ ovip}$)	646.7	0.113	0.197	2.635e−6	1.350e−6	0.0031	0.0031	−1283	18.65
5d	Dollo transition rates ($q_{vivip\ to\ ovip} = 0$)	654.0	0.122	0.196	1.039e−6	1.042e−6	0.0416	1.000e−7	−1298	4.00
4a	$\lambda_{ovip} = \lambda_{vivip}, \mu_{ovip} = \mu_{vivip}$	646.7	0.196	0.196	2.493e−6	2.493e−6	0.0282	0.0031	−1285	16.65
3a	$\lambda_{ovip} = \lambda_{vivip}, \mu_{ovip} = \mu_{vivip}, q_{ovip\ to\ vivip} = q_{vivip\ to\ ovip}$	646.7	0.197	0.197	1.159e−6	1.159e−6	0.003	0.003	−1287	14.65
3b	$\boldsymbol{\lambda_{ovip} = \lambda_{vivip}, \mu_{ovip} = \mu_{vivip}, q_{vivip\ to\ ovip} = 0}$	**654.0**	**0.196**	**0.196**	**1.328e−6**	**1.328e−6**	**0.042**	**1.000e−7**	**−1302**	**0**

AIC, Akaike information criterion.

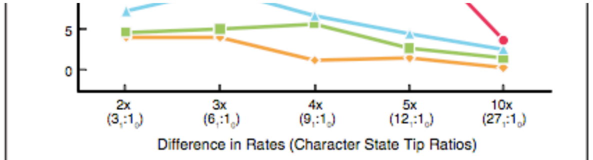


Figure 1 Power of BiSSE under simulations with asymmetrical

Davis et al. (*BMC Biology*, 2013)

Evaluating Correlates of Differential Diversification Rates

And now for some complicated models...meet the *SSE family!

BiSSE (Binary-State Speciation Extinction): Maddison et al. (Syst Biol, 2007)

- Two states, 6 parameters

MuSSE (Multi-State Speciation Extinction): Fitzjohn et al. (Syst Biol, 2009)

- $(2^N - 2) + 2N$ parameters

QuaSSE (Quantitative-State Speciation Extinction): Fitzjohn (Syst Biol, 2010)

- lots and lots of parameters

GeoSSE (Geographic-State Speciation Extinction): Goldberg et al. (Syst Biol, 2011)

- even more parameters!

BiSSEness (BiSSE-Node Enhanced State Shift): Magnuson-Ford & Otto (Am Nat, 2012)

- two states, 10 parameters

Five Fundamental Questions in the Study of Diversification Rates

I. Detecting significant diversification rate variation across the tree

II. Locating significant diversification rate shifts along branches

III. Identifying significant diversification rate shifts through time

IV. Evaluating correlates of differential diversification rates

V. Estimating parameters of the branching process

Diversification-Rate Parameter Estimation

ML Estimation from Lineage-Thru-Time Plots

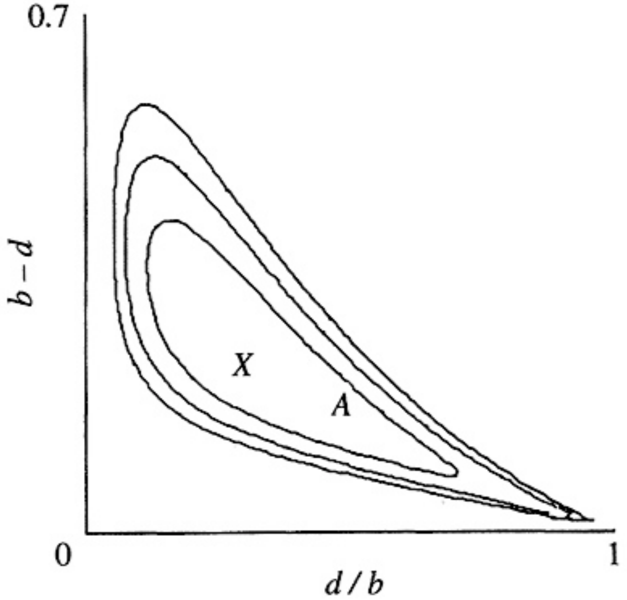
(e.g., Nee , 2001, Magallon & Sanderson, 2001)

Maximum likelihood is used to estimate composite rate parameters:

- diversification rate, $\partial = (\lambda - \mu)$
- relative extinction rate, $\varepsilon = (\mu/\lambda)$

It is not possible to estimate individual rate parameters (e.g., Kubo & Iwasa, 1995; Paradis, 2004)

Comparing absolute diversification rates across clades of different ages is largely invalid if $\mu > 0$ (e.g., Kubo & Iwasa)



Reliable (valid) parameter estimation REQUIRES

- accommodating variance in divergence-time estimates
- no bias in divergence time estimates
- large phylogenies
- complete or phylogenetically unbiased species sampling
- demonstration that diversification has been constant
  - across lineages
  - through time

Summary: Some General Advice for Exploring Diversification Rates

1. Expectations under SBP models are diffuse to a degree that defies intuition

   this makes it difficult to detect departures from stochastic expectations

2. When using methods that entail SBP models for estimation it's critical to:

   carefully assess model fit/adequacy

   carefully assess our ability to reliably estimate under the assumed model

3. It's important to accommodate various sources of phylogenetic uncertainty

   inferences based on point estimates are unlikely to be reliable

4. The statistical behavior of many recent methods is poorly characterized

   power analyses may be useful for assessing the ability of methods to make the desired inferences from your data

5. Make an effort to understand–and assess–the (implicit) assumptions

   the assumptions are often cryptic, critical and frequently violated in real data