

Phylogeny & Biogeography

Michael Landis

Phylogenetics Using RevBayes

NESCent workshop

28 August 2014

Biogeography

How life is distributed in space and time



Watercolor, Joanna Barnum

“Every species has come into existence coincident both in space and time with a pre-existing closely allied species.”

AR Wallace, 1855

Outline

Definitions and examples

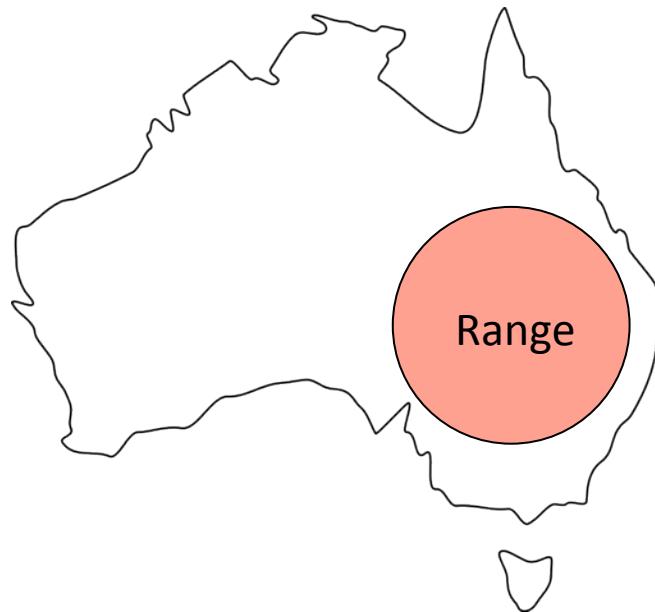
Phylogenetic inference

Discrete models

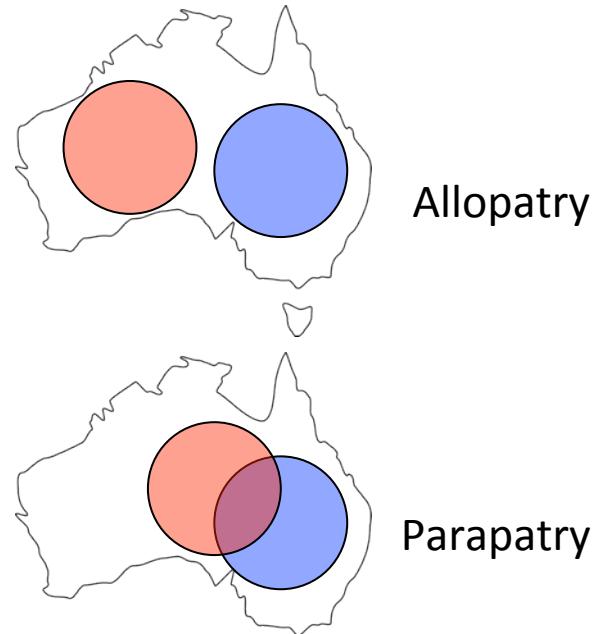
Continuous models

Biogeography lab

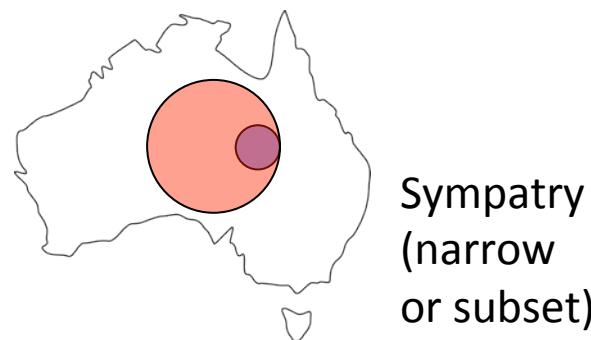
Biogeographic patterns



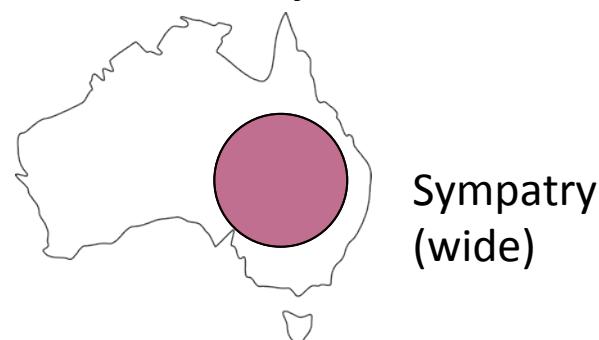
Range



Allopatry

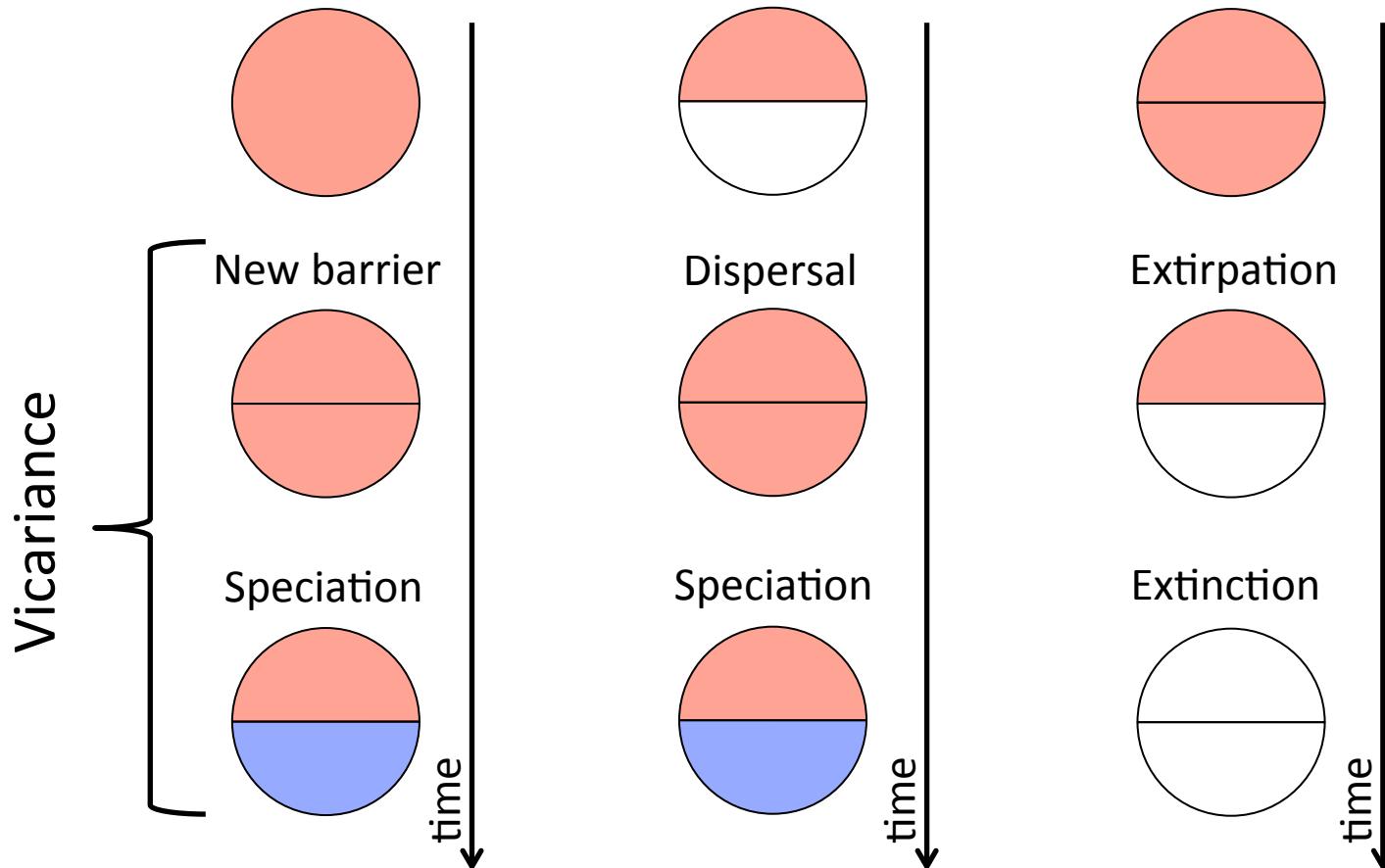


Sympatry
(narrow
or subset)



Sympatry
(wide)

Biogeographic processes



Some of the big puzzles

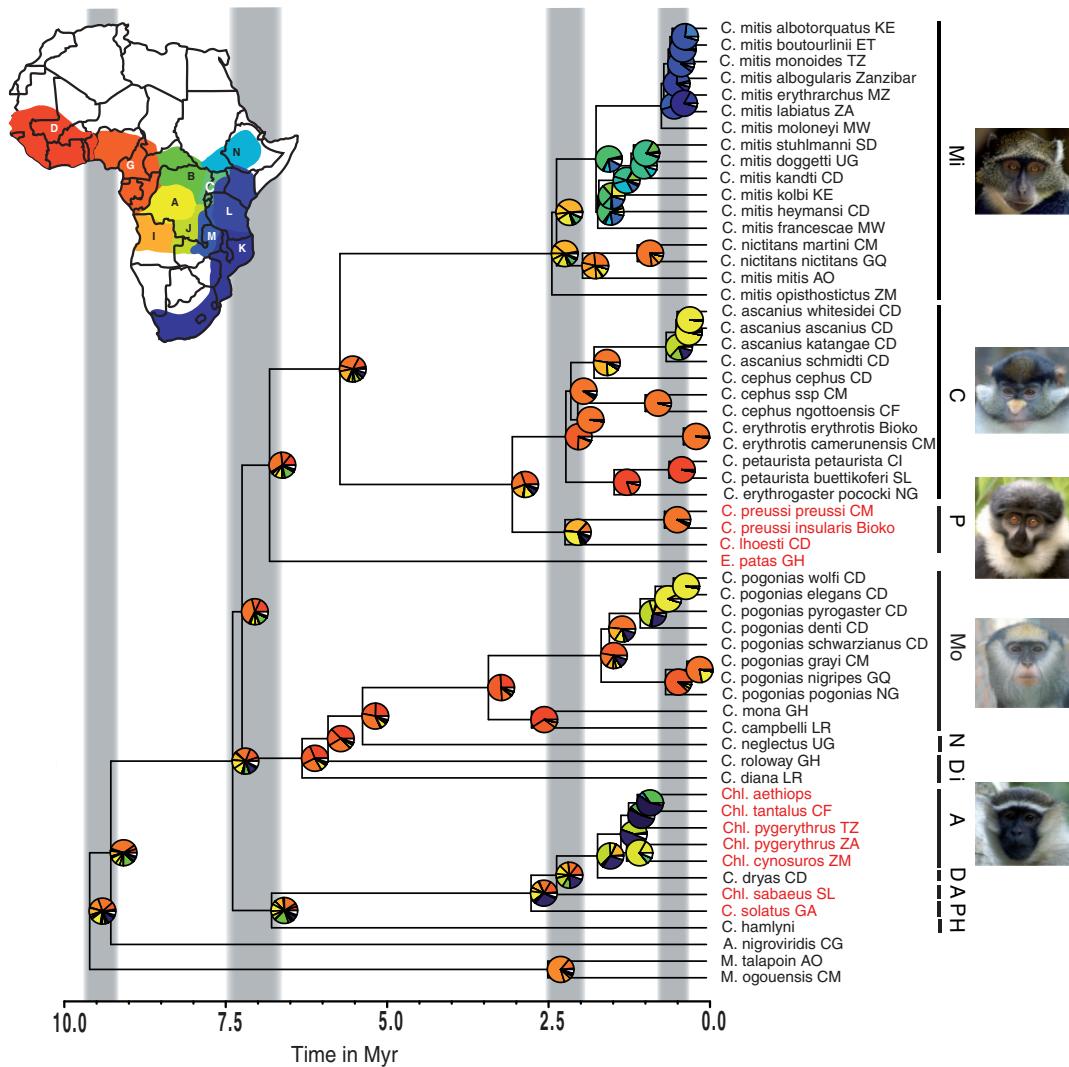
Where did ancestral species live?

How does range size affect speciation/extinction?

What species traits help/prevent colonization?

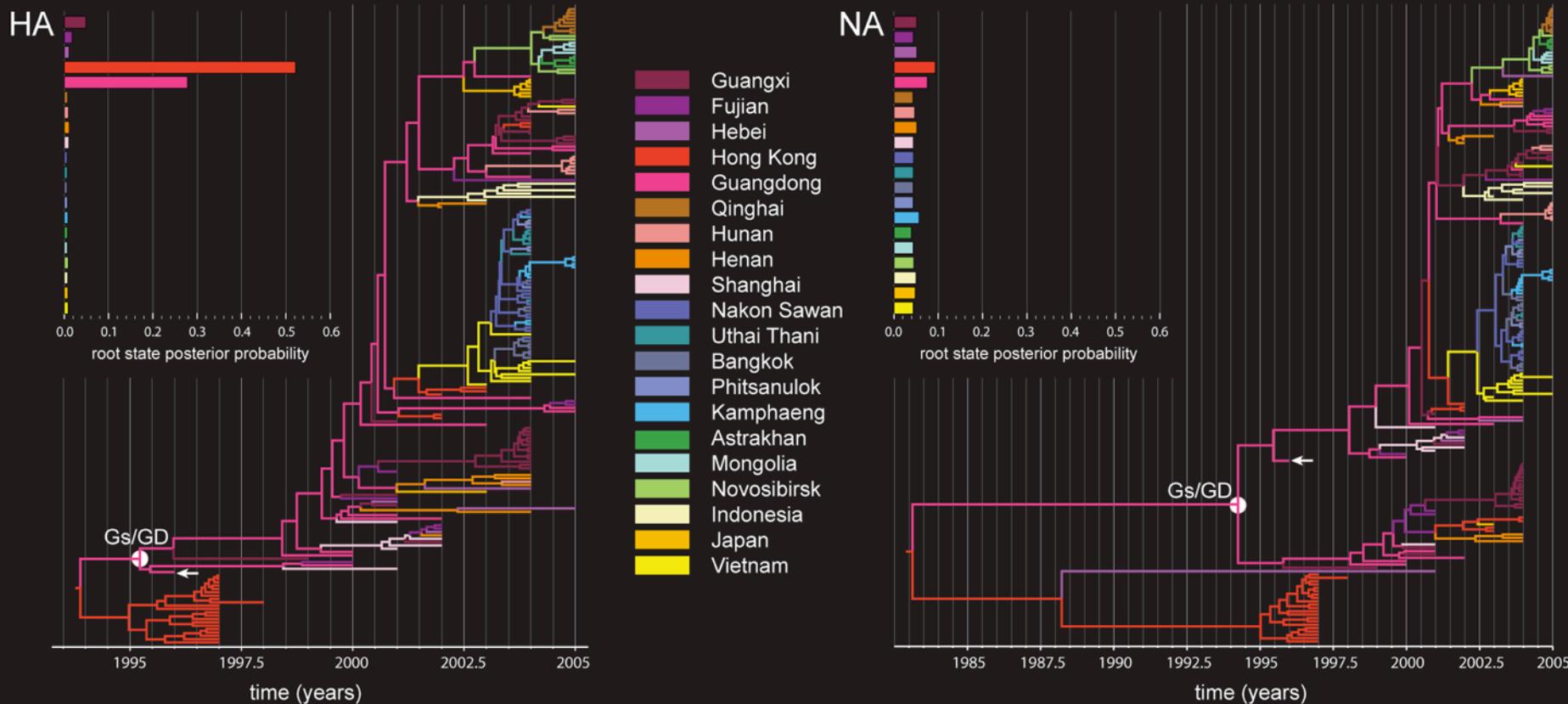
How does geography affect range?

Conservation



Cercopithecidae
(Primates)

Epidemiology



Hemagglutin (HA)

H5N1
(Avian Flu)

Neuraminidase (NA)

Lemey *et al.*, 2008 (PLoS Comp Biol)

Statistical phylogenetics

Familiar strategy:

Data matrix (homology)

Time-calibrated phylogeny

Transition probability of change along branches

Integrate over ancestral characters

Gives us $\mathcal{L}(X; \theta, T, M)$



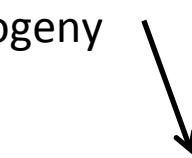
Octodon degus

Character states



Ancestral state
reconstructions

Time-calibrated
phylogeny



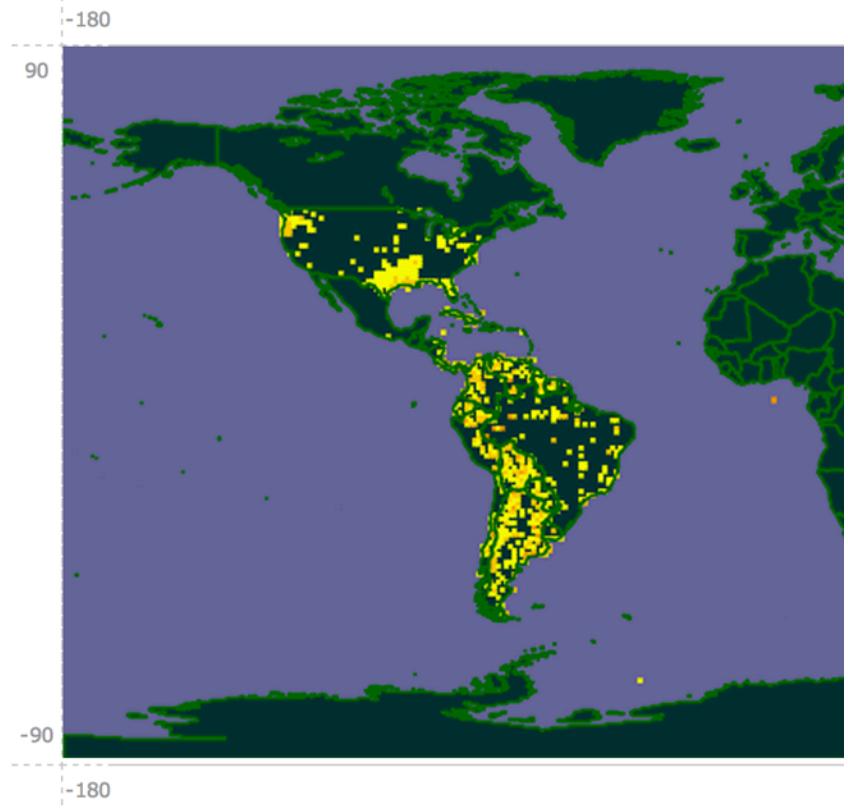
(b)

Data matrix



Species occurrence data (gbif.org, 2013)

Map of results



► Your search returned **13,264** occurrences with coordinates.

Discrete presence-absence (Upham & Patterson, 2012)

A	B	C	D	E	F	G	H	I
Abrocoma cinerea	Abrocoma bennettii	Octodon gliroides	Octodon degus	Octodon lunatus	Octodon degus	Acnoaemys porteri	Acnoaemys sagei	Acnoaemys fuscus
Spalacopus cyaneus	Octomys mimax	Pipanacotomys aureus	Tympanocotomys barrerae	Ctenomys haigi	Ctenomys boliviensis	Ctenomys steinbachi	Isothrix sinnamariensis	Isothrix bistrata
Isothrix barbarabrownae	Makalata macrura	Phyllomys brasiliensis	Makalata didelphoides	Phyllomys blainvillii	Toromys grandis	Echimys chrysurus	Mesomys ocellatus	Mesomys hispidus
Lonchothrix emiliae	Dactylomys dactylinus	Dactylomys boliviensis	Kannabateomys albimonyx	Proechimys roberti	Proechimys longicaudatus	Proechimys simonsi	Proechimys quadruplicatus	Hoplomys gymnurus
Capromys pilorides	Trinomys dimidiatus	Trinomys iheringi	Trinomys yonenagae	Thrichomys apereoides	Thrichomys apereoides	Myocastor coypus	Capromys pilorides	Trinomys eliasi
Trinomys setosus	Clomys laticeps	Trinomys paratus	Trinomys setosus	Myoprocta acouchy	Trinomys setosus	Lagostomus maximus	Chinchilla lanigera	Erethizon dorsatum
Clomys laticeps	Euryzygomatomys spinosus	Coendou bicolor	Clomys laticeps	Dasyprocta leporina	Euryzygomatomys spinosus	Lagidium viscacia	Sphiggurus melanurus	Coendou bicolor
Euryzygomatomys spinosus	Dinomys branickii	Myoprocta acouchy	Euryzygomatomys spinosus	Hydrochoerus hydrochaeris	Dinomys branickii	Chinchilla lanigera	Dolichotis patagonum	Microcavia australis
Dinomys branickii	Lagostomus maximus	Dasyprocta leporina	Hydrochoerus hydrochaeris	Kerodon rupestris	Lagostomus maximus	Sphiggurus melanurus	Cavia tschudii	Microcavia australis
Lagostomus maximus	Chinchilla lanigera	Hydrochoerus hydrochaeris	Kerodon rupestris	Dolichotis patagonum	Chinchilla lanigera	Dolichotis patagonum	Cavia tschudii	Microcavia australis
Chinchilla lanigera	Sphiggurus melanurus	Dolichotis patagonum	Dolichotis patagonum	Microcavia australis	Sphiggurus melanurus	Microcavia australis	Cavia porcellus	Cavia porcellus
Sphiggurus melanurus	Erethizon dorsatum	Microcavia australis	Microcavia australis	Cavia tschudii	Erethizon dorsatum	Cavia tschudii	Cavia aperea	Cavia aperea
Erethizon dorsatum	Coendou bicolor	Cavia tschudii	Cavia tschudii	Cavia aperea	Coendou bicolor	Cavia aperea	Galea musteloides	Galea musteloides
Coendou bicolor	Myoprocta acouchy	Cavia aperea	Cavia aperea	Galea musteloides	Myoprocta acouchy	Galea musteloides	Cuniculus pacas	Cuniculus pacas
Myoprocta acouchy	Dasyprocta leporina	Dasyprocta leporina	Dasyprocta leporina	Cuniculus pacas	Dasyprocta leporina	Dasyprocta leporina	Cuniculus tacjanowskii	Cuniculus tacjanowskii
Dasyprocta leporina	Hydrochoerus hydrochaeris	Hydrochoerus hydrochaeris	Hydrochoerus hydrochaeris	Cuniculus pacas	Hydrochoerus hydrochaeris	Hydrochoerus hydrochaeris		
Hydrochoerus hydrochaeris	Kerodon rupestris	Kerodon rupestris	Kerodon rupestris	Cuniculus pacas	Kerodon rupestris	Kerodon rupestris		
Kerodon rupestris	Dolichotis patagonum	Dolichotis patagonum	Dolichotis patagonum	Cuniculus pacas	Dolichotis patagonum	Dolichotis patagonum		
Dolichotis patagonum	Microcavia australis	Microcavia australis	Microcavia australis	Cuniculus pacas	Microcavia australis	Microcavia australis		
Microcavia australis	Cavia tschudii	Cavia tschudii	Cavia tschudii	Cuniculus pacas	Cavia tschudii	Cavia tschudii		
Cavia tschudii	Cavia porcellus	Cavia porcellus	Cavia porcellus	Cuniculus pacas	Cavia porcellus	Cavia porcellus		
Cavia porcellus	Cavia aperea	Cavia aperea	Cavia aperea	Cuniculus pacas	Cavia aperea	Cavia aperea		
Cavia aperea	Galea musteloides	Galea musteloides	Galea musteloides	Cuniculus pacas	Galea musteloides	Galea musteloides		
Galea musteloides	Cuniculus pacas							
Cuniculus pacas	Cuniculus tacjanowskii	Cuniculus tacjanowskii	Cuniculus tacjanowskii	Cuniculus pacas	Cuniculus tacjanowskii	Cuniculus tacjanowskii		



Individual or range

The individuals in a taxon share a range.

	Discrete	Continuous
Individual or Endemic	single occupied area	geographical point
Range	set of occupied areas	set of geographical points

Data matrix

X_{ij} taxon i , character j

Continuous

e.g. latitude-longitude

$$X_i = (\phi, \lambda) = (38.54^\circ\text{N}, 121.75^\circ\text{W})$$

Discrete

e.g. single area presence-absence (range)

$$X_i = \text{Africa} \quad X_i = (0, 0, 1, 0, 0, 1, 1, 1)$$

Models

Continuous

e.g. Brownian motion (Gaussian)

$$X \sim N(\mu, \Sigma)$$

Discrete

e.g. continuous-time Markov chain

$$P(x \rightarrow y; t) = [e^{Qt}]_{x,y}$$

Continuous models

Brownian motion

Each taxon is an individual sample

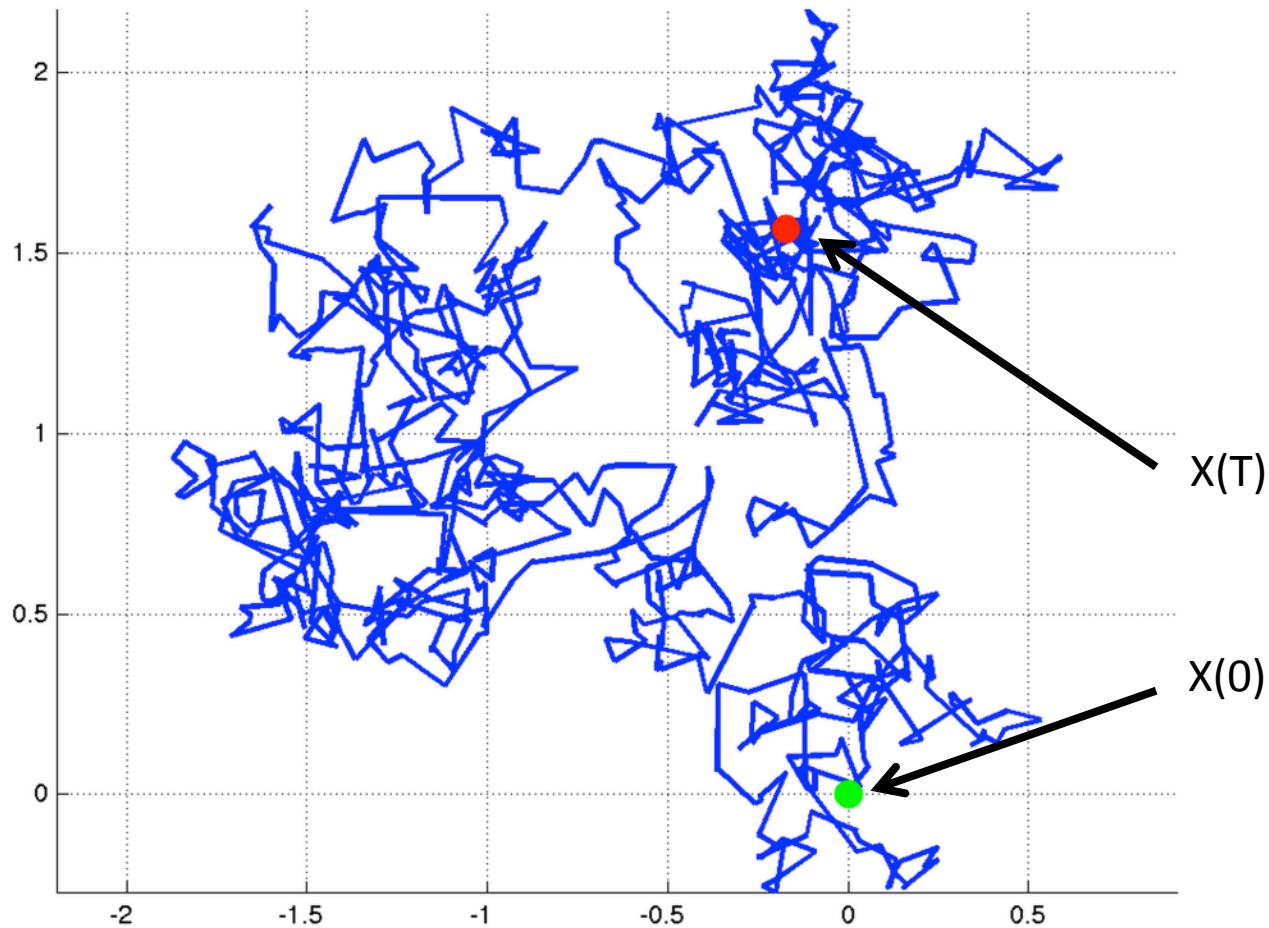
Epidemiology models

Work by:

Lemmon & Lemmon, 2008 (Syst Biol)

Lemey *et al.*, 2010 (Mol Biol Evol)

2D Brownian motion



Relaxed random walk

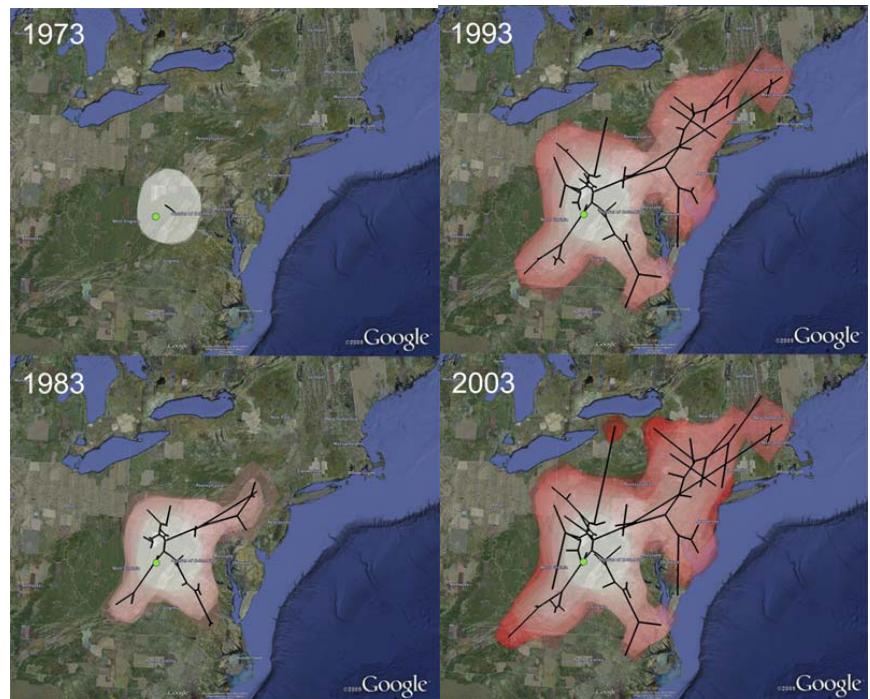
Joint inference of gene tree
using relaxed molecular clock

Latitude, longitude diffuse by
Brownian motion for each branch:

$$X_b \sim N(X_{\text{pa}(b)}, t_b \phi_b \Sigma)$$

Branch rate rescaled (“relaxed”):

$$\phi_b \sim \text{Gamma}(\nu/2, \nu/2)$$



Continuous models for
ranges or multiple individuals

Diffusion of set of individual coordinates

???

Diffusion of range as polygon

???

Hard problem, underexplored

Island Model

Dispersal-only model

One area per taxon (endemic/individual)

Learn favored dispersal routes

Work by:

Sanmartín *et al.*, 2008 (J Biogeog)

Lemey *et al.*, 2009 (PLoS Comp Biol)

Island Model

I. Sanmartín, P. van der Mark and F. Ronquist

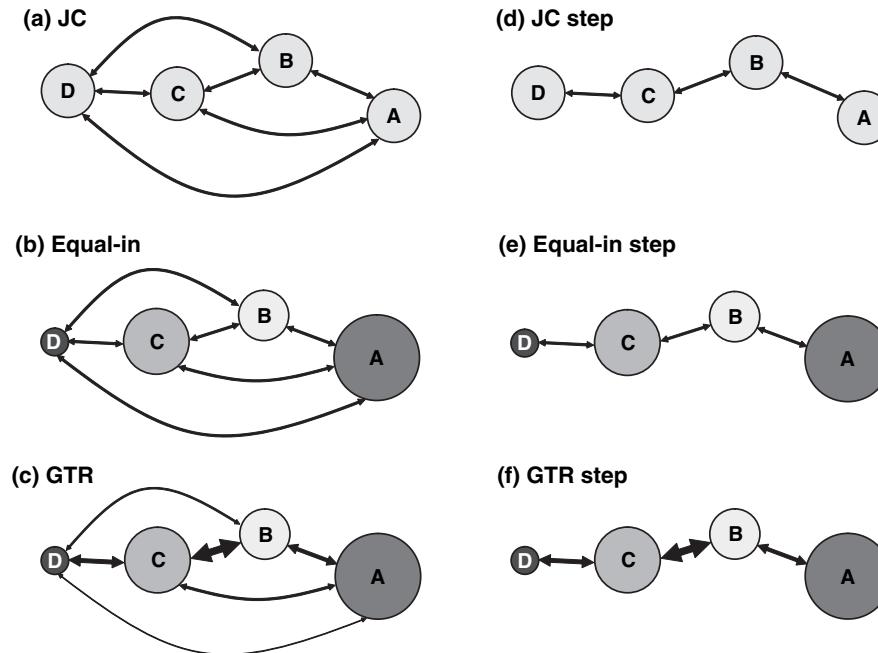


Figure 2 Bayesian Island Models: Each circle represents an island; circle size represents the relative carrying capacity of the island (expected number of lineages at equilibrium); arrow width represents the relative dispersal rate between two single islands. (a) Jukes–Cantor (JC) model: all carrying capacities equal, all dispersal rates equal. (b) Equal-in model: unequal carrying capacities, equal dispersal rates. (c) General Time Reversible (GTR) model: unequal carrying capacities, unequal dispersal rates. (d–f) Stepping-stone variant of each model. (d) JC step: all carrying capacities equal, dispersal rates equal between adjacent islands, zero between non-adjacent islands. (e) Equal-in step: unequal carrying capacities, all dispersal rates equal between adjacent islands, zero between non-adjacent islands. (f) GTR step: all carrying capacities unequal, all dispersal rates unequal between adjacent islands, zero between non-adjacent islands.

Embedding the graph in Q

General Time
Reversible
(a – c)

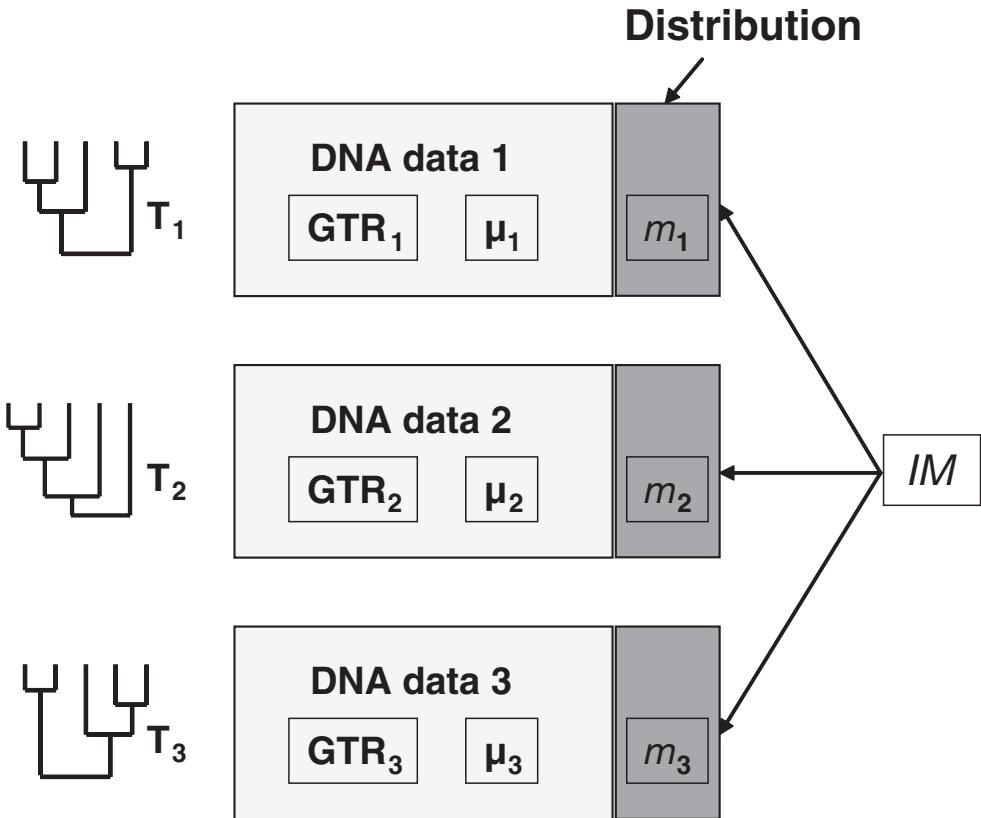
$$Q = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \left(\begin{matrix} - & \pi_B r_{AB} & \pi_C r_{AC} & \pi_D r_{AD} \\ \pi_A r_{AB} & - & \pi_C r_{BC} & \pi_D r_{BD} \\ \pi_A r_{AC} & \pi_B r_{BC} & - & \pi_D r_{CD} \\ \pi_A r_{AD} & \pi_B r_{BD} & \pi_C r_{CD} & - \end{matrix} \right) \end{matrix}$$

Stepping
Stone
(b – f)

$$Q = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \left(\begin{matrix} - & \pi_B r_{AB} & 0 & 0 \\ \pi_A r_{AB} & - & \pi_C r_{BC} & 0 \\ 0 & \pi_B r_{BC} & - & \pi_D r_{CD} \\ 0 & 0 & \pi_C r_{CD} & - \end{matrix} \right) \end{matrix}$$

Shared:
Dispersal process

Independent:
molecular process,
molecular speed,
dispersal speed,
clock tree



Bayes Factors (harmonic mean)

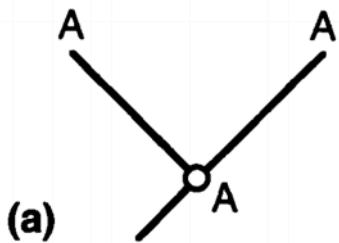
13 groups, 393 species, 954 taxa

Island model	Ln model likelihood
JC step	-101704.09
Equal-in	-101667.87
JC	-101649.92
Equal-in step	-101628.31
GTR	-101624.19
GTR step	-101618.94 ←
	*(-101642.9)

*Model likelihood for the ‘long analysis’ (30 million generations, four runs); see text.

DIVA Parsimony

Dispersal



Vicariance

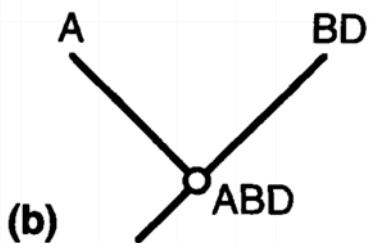


FIGURE 1. Expected biogeographic outcomes of speciation. (a) When the ancestral species occurs in a single area, the daughter species will be expected to occur in the same area. (b) When the ancestral species occurs in several areas, the daughter species will be expected to occur in mutually exclusive sets of areas.

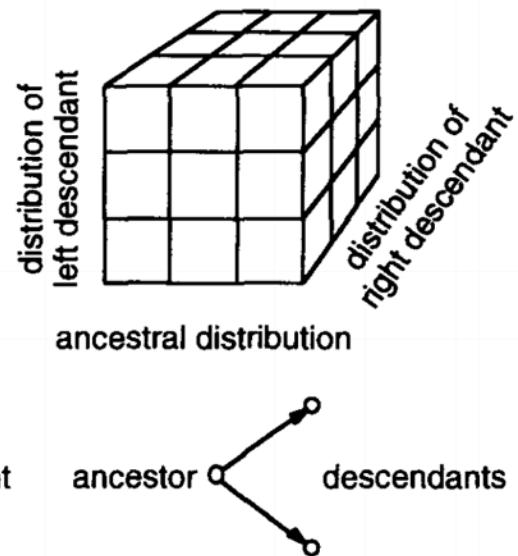
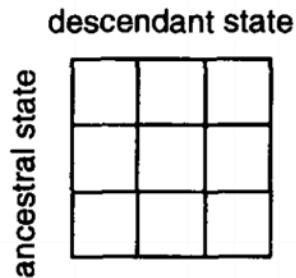


FIGURE 2. The difference between an ordinary step matrix and the cost matrix needed for the reconstruction of ancestral areas. An ordinary step matrix is two dimensional and specifies the cost of moving between states along an internode. The cost matrix used in dispersal–vicariance analysis is three dimensional and specifies the cost of combinations of ancestral, left descendant, and right descendant distributions.

LAGRANGE

Dispersal-(Local) Extinction-Cladogenesis (DEC)

Many areas per taxon (range)

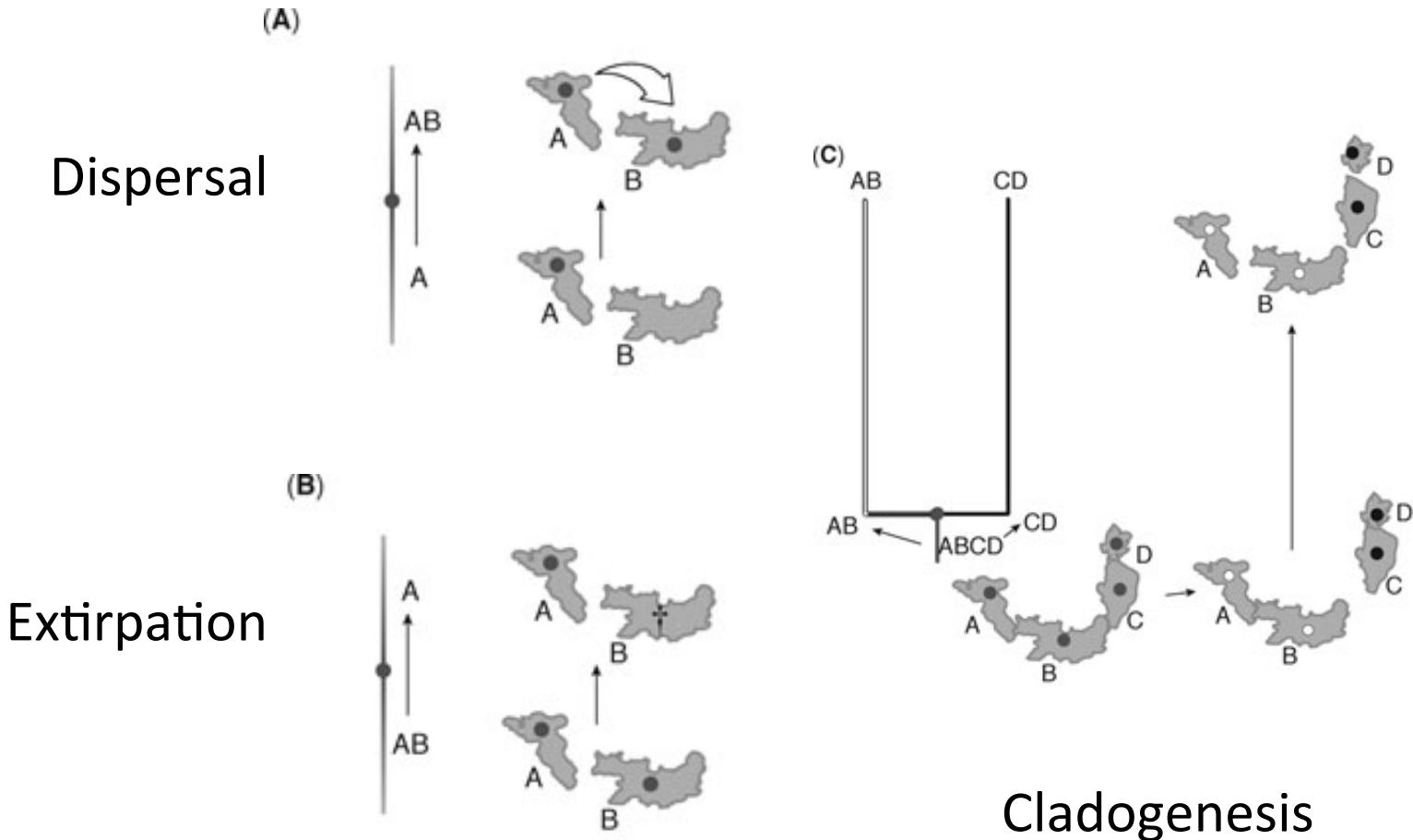
D,E as parameterized event classes

Work by:

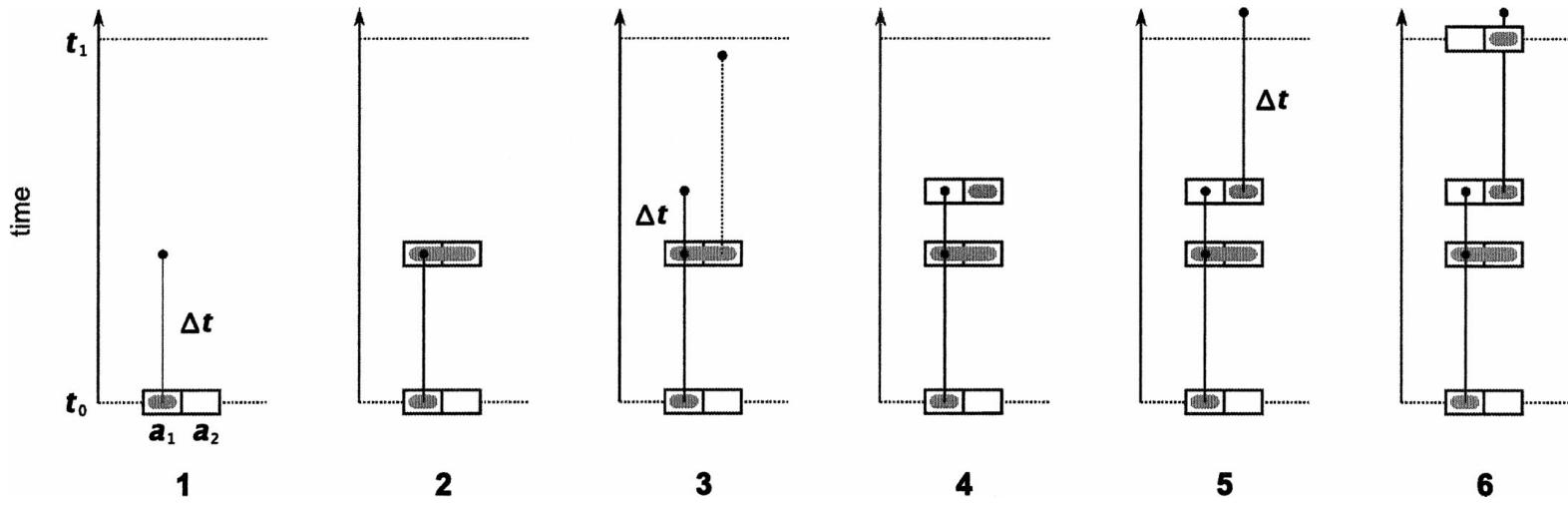
Ree *et al.*, 2005 (Evolution)

Ree & Smith, 2008 (Syst Biol)

DEC event types



Dispersal & Extinction



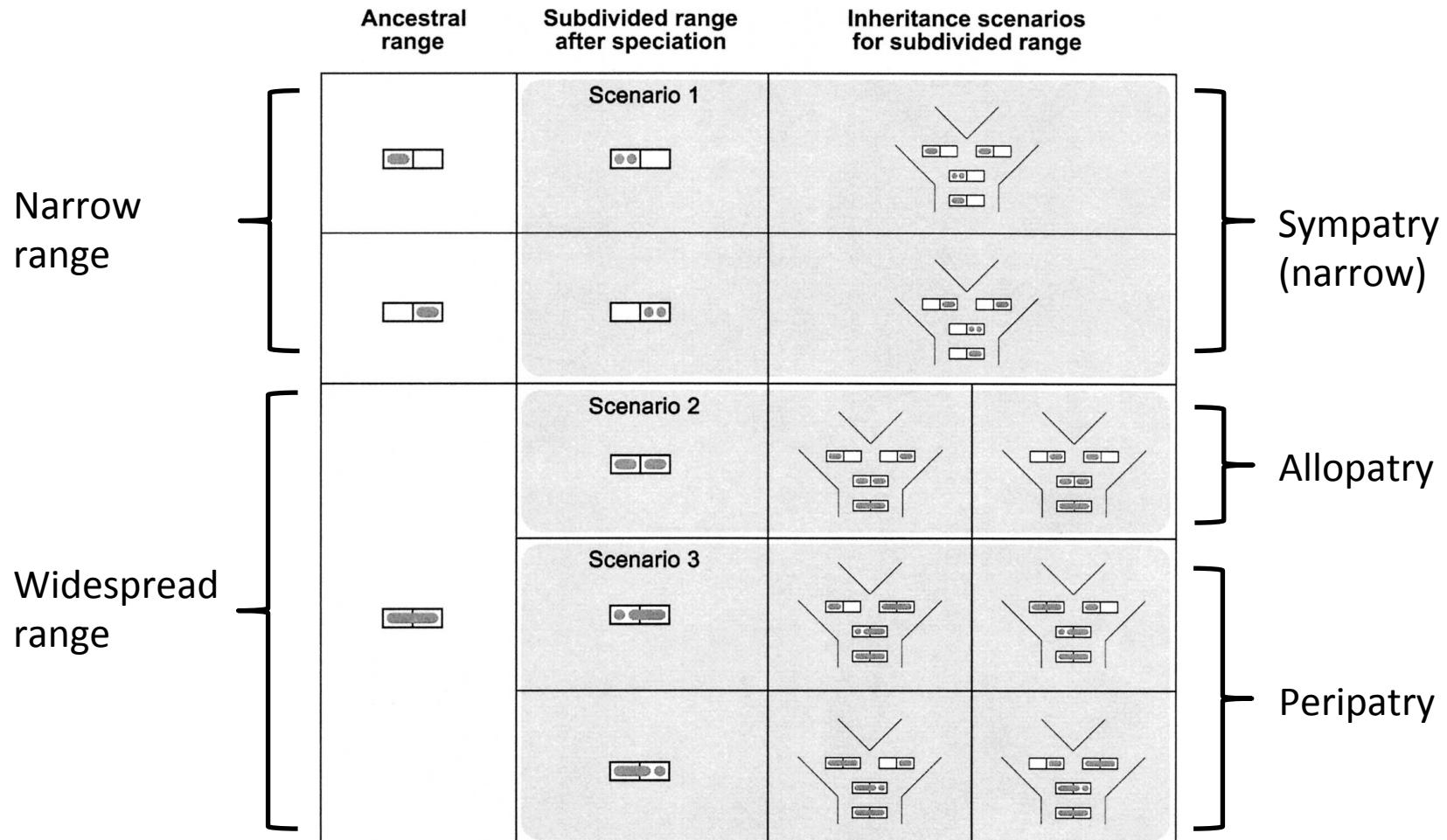
Exponentially-distributed times between events

Rate matrix for anagenesis

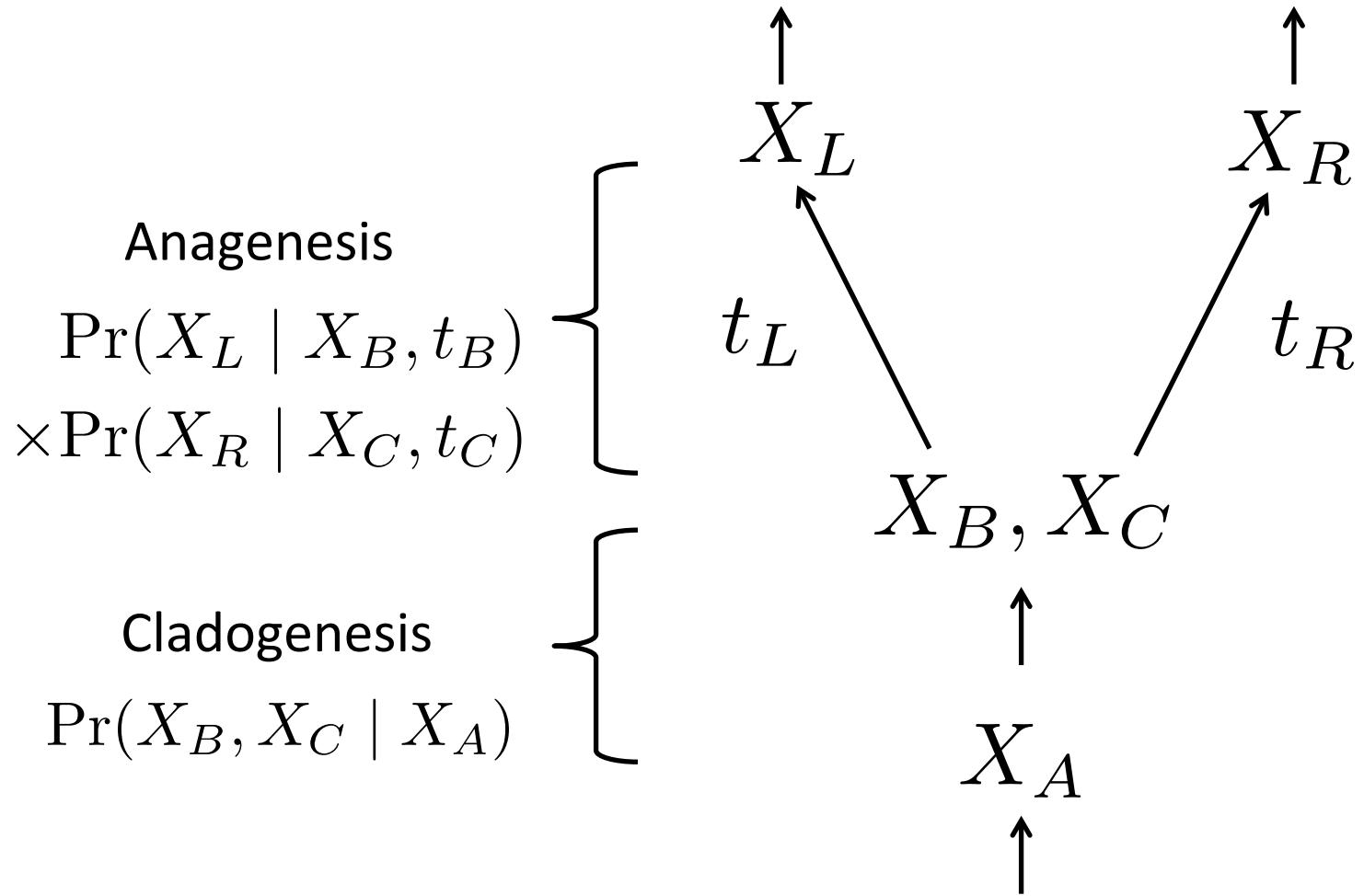
$$Q = \left[\begin{array}{c|ccccccccc} & \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & 123 \\ \hline \emptyset & — & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & E_1 & — & 0 & 0 & D_{12} & D_{13} & 0 & 0 \\ 2 & E_2 & 0 & — & 0 & D_{21} & 0 & D_{23} & 0 \\ 3 & E_3 & 0 & 0 & — & 0 & D_{31} & D_{32} & 0 \\ 12 & 0 & E_2 & E_1 & 0 & — & 0 & 0 & D_{13} + D_{23} \\ 13 & 0 & E_3 & 0 & E_1 & 0 & — & 0 & D_{12} + D_{32} \\ 23 & 0 & 0 & E_3 & E_2 & 0 & 0 & — & D_{21} + D_{31} \\ 123 & 0 & 0 & 0 & 0 & E_3 & E_2 & E_1 & — \end{array} \right]$$

$$\mathbf{P}_{ij}(t) = [\exp \{\mathbf{Qt}\}]_{ij}$$

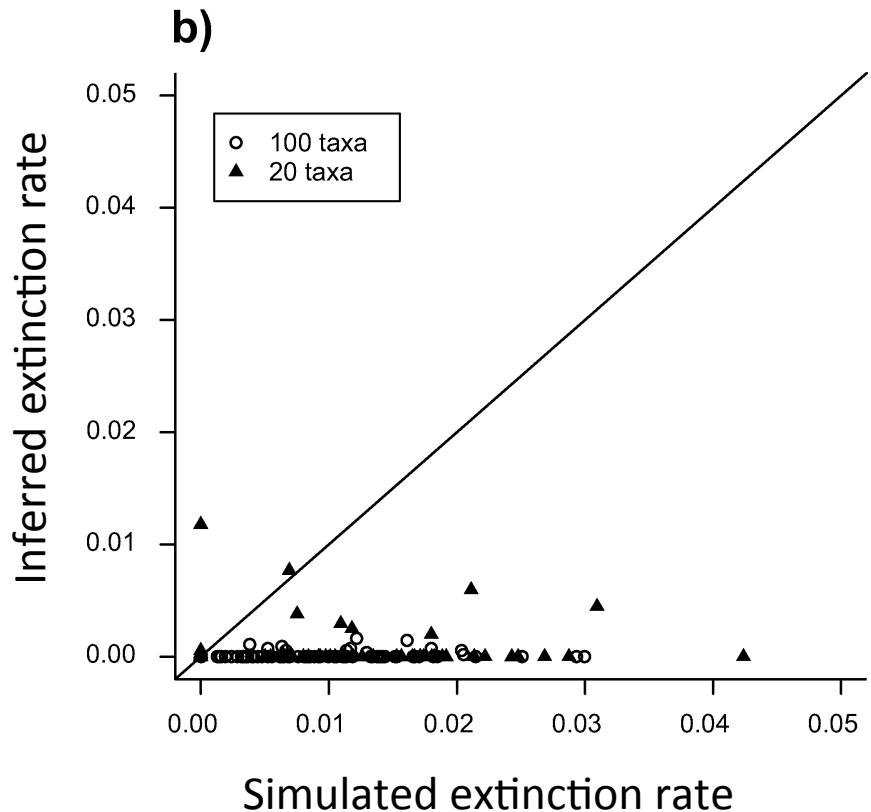
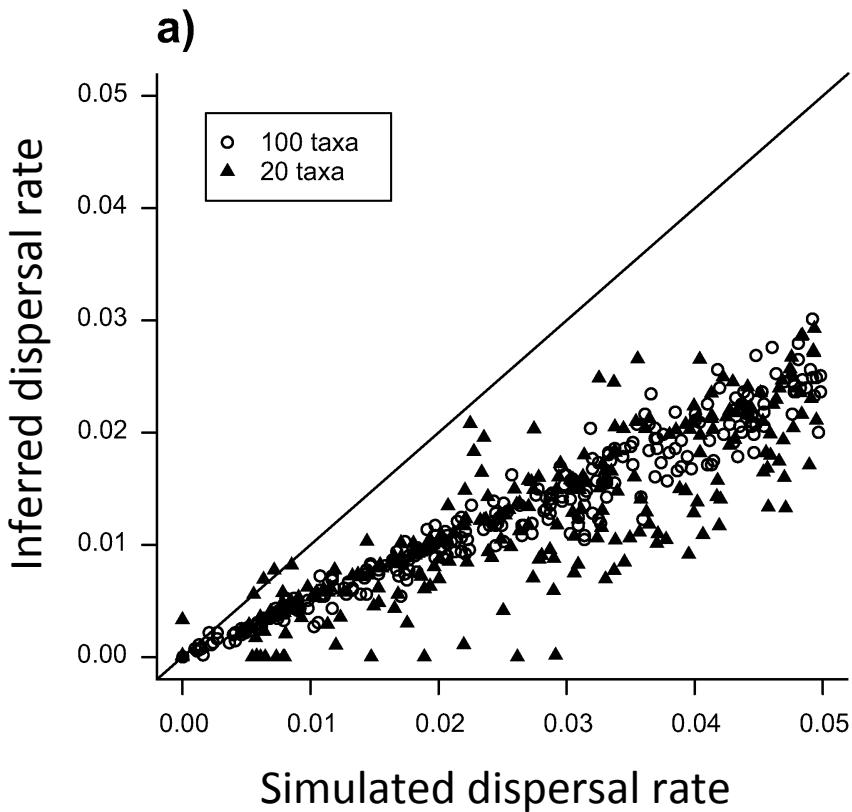
Cladogenesis



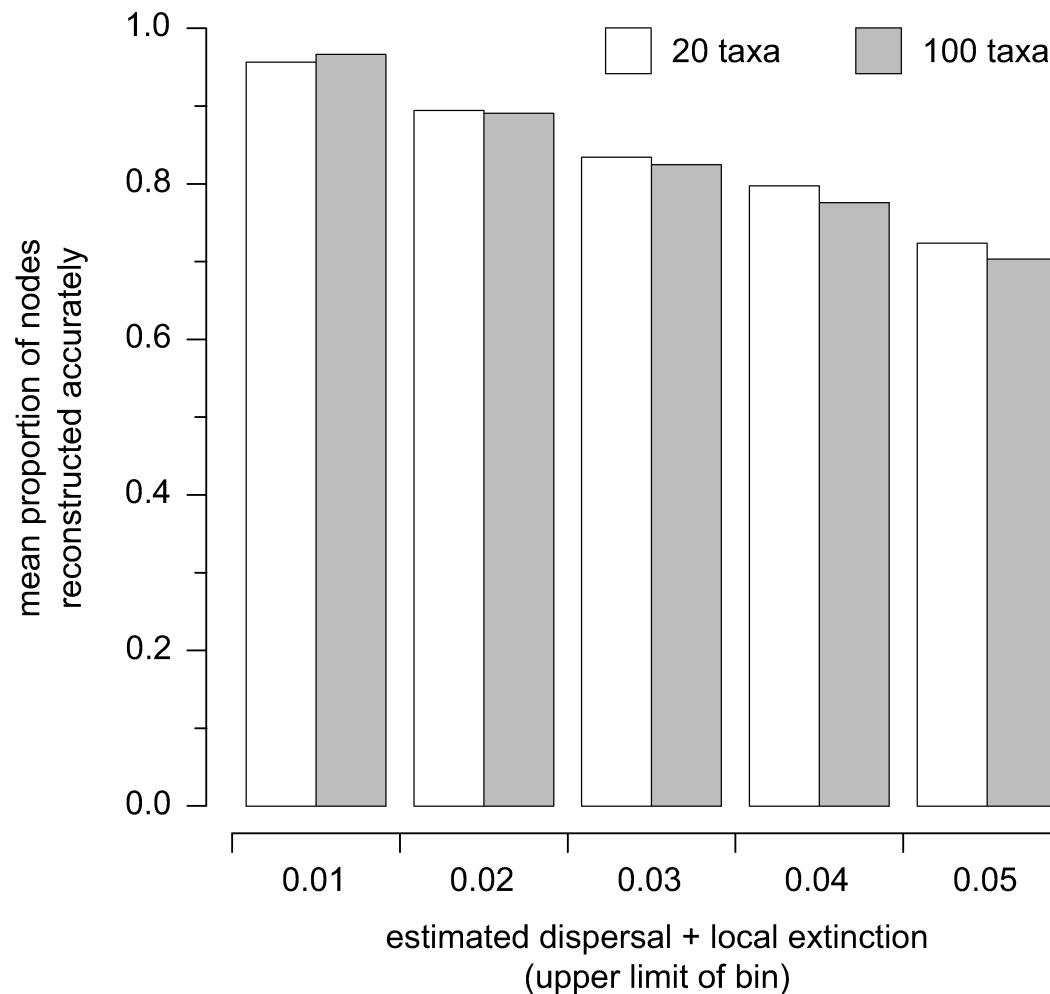
DEC likelihood



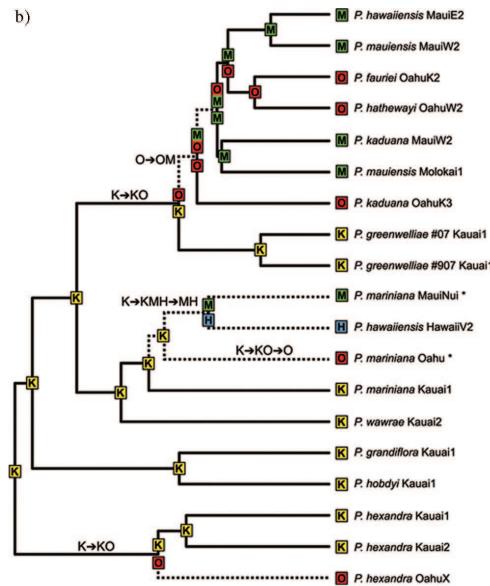
MLE inference vs truth



MLE range reconstructions

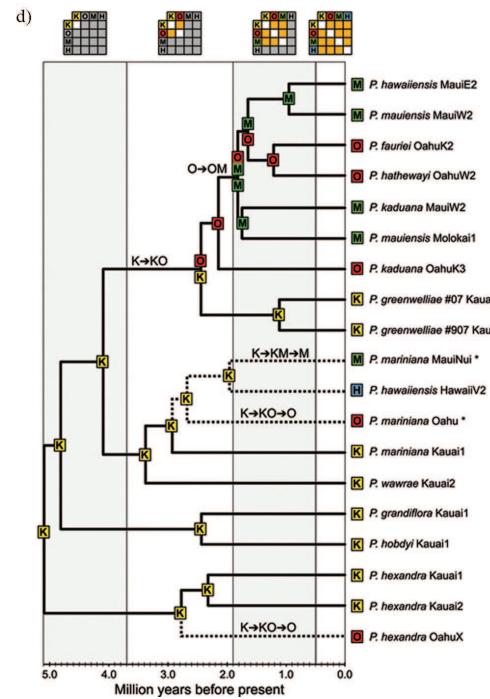
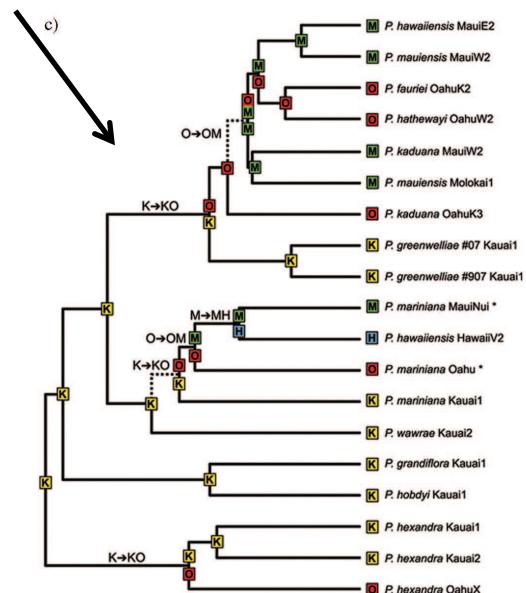


Unconstrained
(GTR)



Psychotria mariniana

Stepping stone
(Small adjacent ranges:
K, O, M, H, KO, OM, MH)



Stratified
(time-dependent GTR)

BioGeoBEARS

Generalized DEC model

Several areas per taxon (range)

Seven parameterized event classes

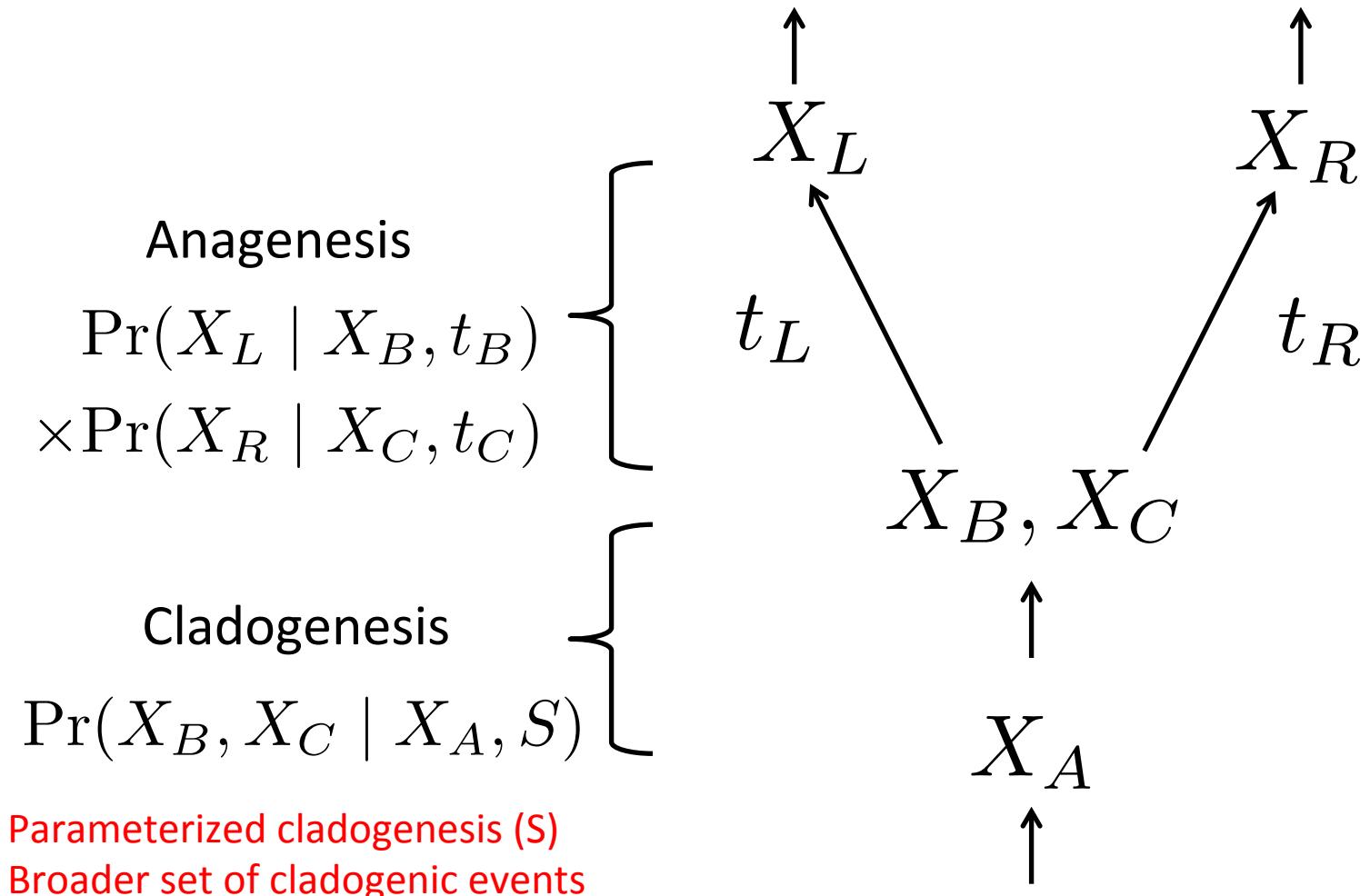
Work by:

Matzke, 2014 (Syst Biol)

Generalized DEC model

	Process	Ranges Before	Ranges After	Character mapping	DIVA	DEC (GeoSSE, LAGRANGE)	BayArea, BBM (RASP)	Parameter of BioGeoBEARS Supermodel
Anagenic	Dispersal				✓	✓	✓	d (& x, b)
	Extinction				✓	✓	✓	e (& u, b)
	Range-switching			✓				a (& x, b)
Cladogenic	Sympatry (narrow)		 		✓	✓	✓	y (& $mx0 \mid y$)
	Sympatry (widespread)		 			✓		y (& $mx0 \mid y$)
	Sympatry (subset)		 			✓		S (& $mx0 \mid s$)
	Vicariance (narrow)		 		✓	✓		V (& $mx0 \mid v$)
	Vicariance (widespread)		 		✓			V (& $mx0 \mid v$)
	Founder	 	 				j (& $x, mx0 \mid j$)	

DEC+J likelihood



13 island clades
55 scenarios

DEC

DEC+J

Model

d: rate of dispersal/
range expansion

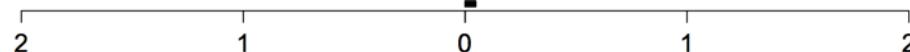
e: rate of extinction/
range contraction

j: weight of jump dispersal/
founder-event speciation

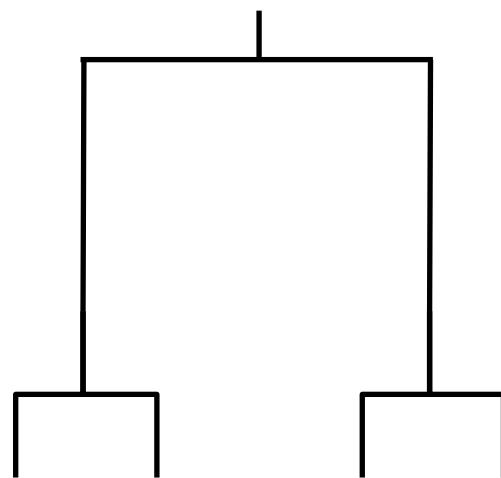
Dispersal

Extinction

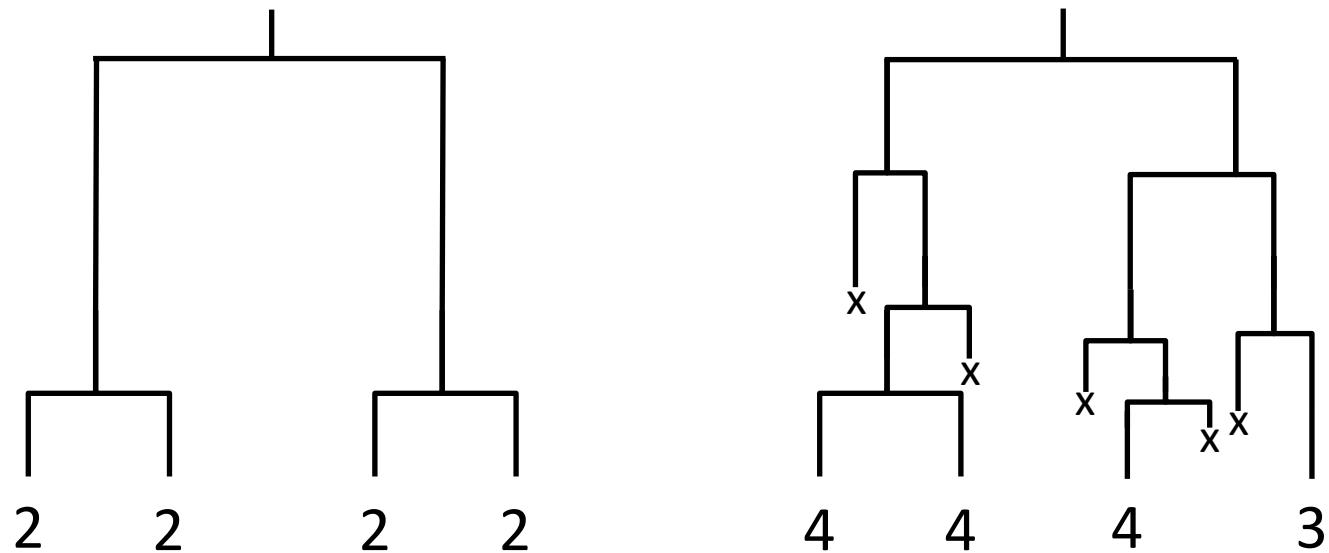
Jump-dispersal



Difference in parameter value



Speciation hidden by extinction



Geographic State Speciation Extinction (GeoSSE)

DEC model

Joint birth-death process & range evolution

Accounts for “hidden” speciation

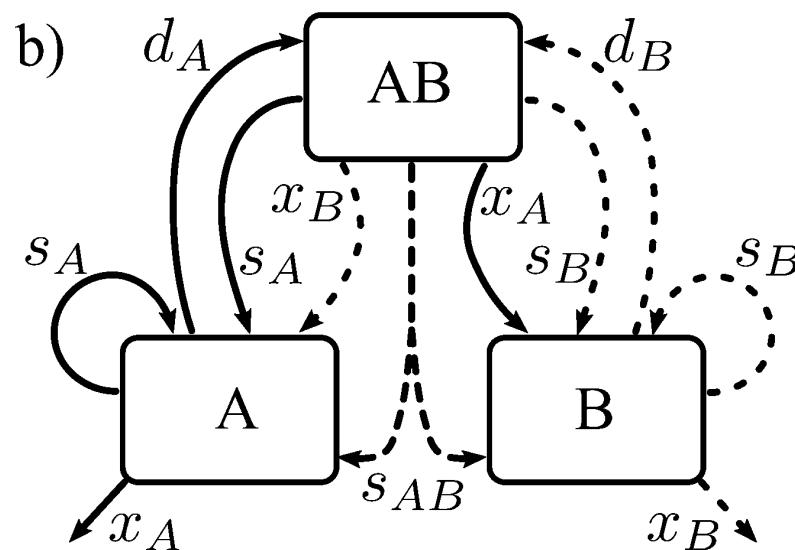
Range evolution, speciation, extinction intertwined

Work by:

Goldberg *et al.*, 2011 (Syst. Biol.)

Goldberg & Igić, 2012 (Evolution)

	Parameter	Areas	Event
Speciation	s_A	A	New lineage in area A
	s_B	B	New lineage in area B
	s_{AB}	AB	New lineage in area A or B
Dispersal	d_A, d_B	A or B	This lineage gains an area
Extinction	x_A, x_B	AB	This lineage loses an area
		A or B	This lineage goes extinct



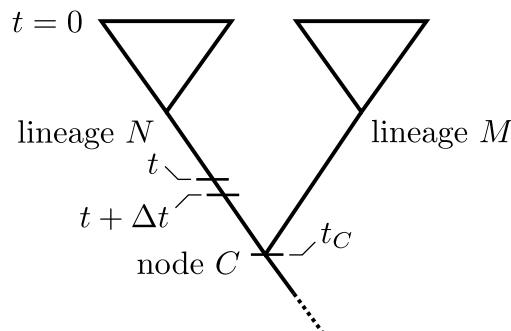
GeoSSE likelihood

Likelihood of tree and character states

$$\frac{dD_{NA}}{dt} = -(s_A + d_A + x_A)D_{NA}(t) + d_A D_{NAB}(t) + 2s_A D_{NA}(t)E_A(t),$$

$$\frac{dD_{NB}}{dt} = -(s_B + d_B + x_B)D_{NB}(t) + d_B D_{NAB}(t) + 2s_B D_{NB}(t)E_B(t),$$

$$\begin{aligned} \frac{dD_{NAB}}{dt} = & -(s_A + s_B + s_{AB} + x_A + x_B)D_{NAB}(t) \\ & + x_A D_{NB}(t) + x_B D_{NA}(t) \\ & + s_A [E_A(t)D_{NAB}(t) + E_{AB}(t)D_{NA}(t)] \\ & + s_B [E_B(t)D_{NAB}(t) + E_{AB}(t)D_{NB}(t)] \\ & + s_{AB}[E_A(t)D_{NB}(t) + E_B(t)D_{NA}(t)], \end{aligned}$$



Likelihood of extinction

$$\frac{dE_A}{dt} = -(s_A + d_A + x_A)E_A(t) + x_A + d_A E_{AB}(t) + s_A E_A(t)^2, \quad (3a)$$

$$\frac{dE_B}{dt} = -(s_B + d_B + x_B)E_B(t) + x_B + d_B E_{AB}(t) + s_B E_B(t)^2, \quad (3b)$$

$$\begin{aligned} \frac{dE_{AB}}{dt} = & -(s_A + s_B + s_{AB} + x_A + x_B)E_{AB}(t) + x_A E_B(t) \\ & + x_B E_A(t) + s_A E_{AB}(t)E_A(t) + s_B E_{AB}(t)E_B(t) \\ & + s_{AB} E_A(t)E_B(t). \end{aligned} \quad (3c)$$

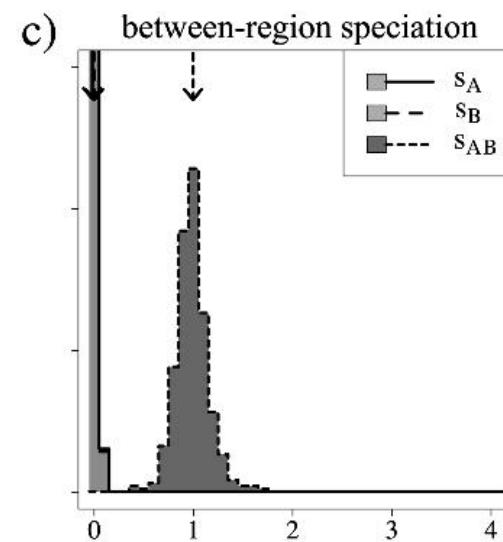
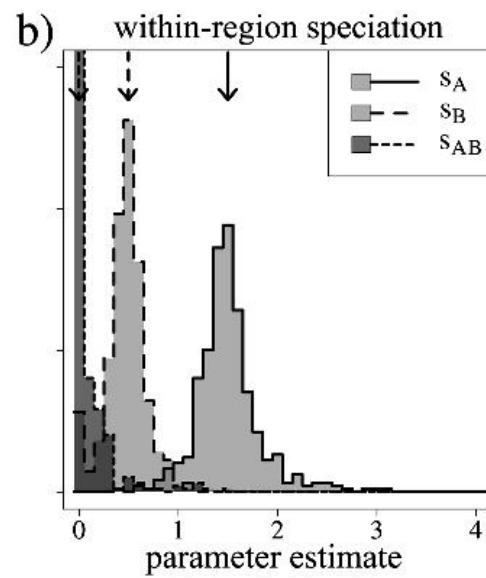
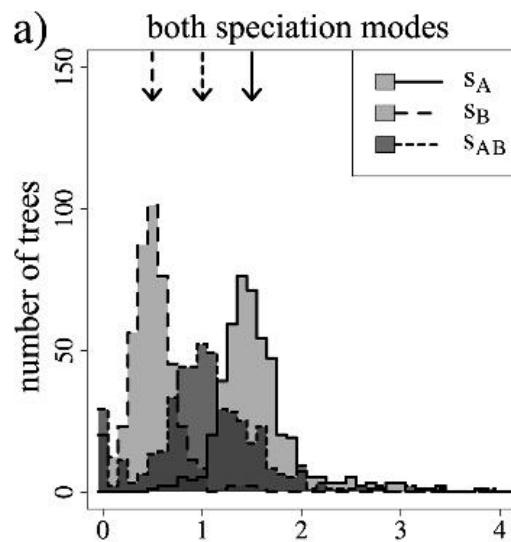
Likelihood of “observed” speciation

$$D_{CA}(t_C) = D_{NA}(t_C)D_{MA}(t_C)s_A, \quad (2a)$$

$$D_{CB}(t_C) = D_{NB}(t_C)D_{MB}(t_C)s_B, \quad (2b)$$

$$\begin{aligned} D_{CAB}(t_C) = & \frac{1}{2}[D_{NAB}(t_C)D_{MA}(t_C) + D_{NA}(t_C)D_{MAB}(t_C)]s_A \\ & + \frac{1}{2}[D_{NAB}(t_C)D_{MB}(t_C) + D_{NB}(t_C)D_{MAB}(t_C)]s_B \\ & + \frac{1}{2}[D_{NA}(t_C)D_{MB}(t_C) + D_{NB}(t_C)D_{MA}(t_C)]s_{AB}. \end{aligned} \quad (2c)$$

MLE inference vs truth



$$s_{AB} > 0$$

BayArea

Dispersal-Extirpation model

Distance effects as a free parameter

Scales for many areas

Work by:

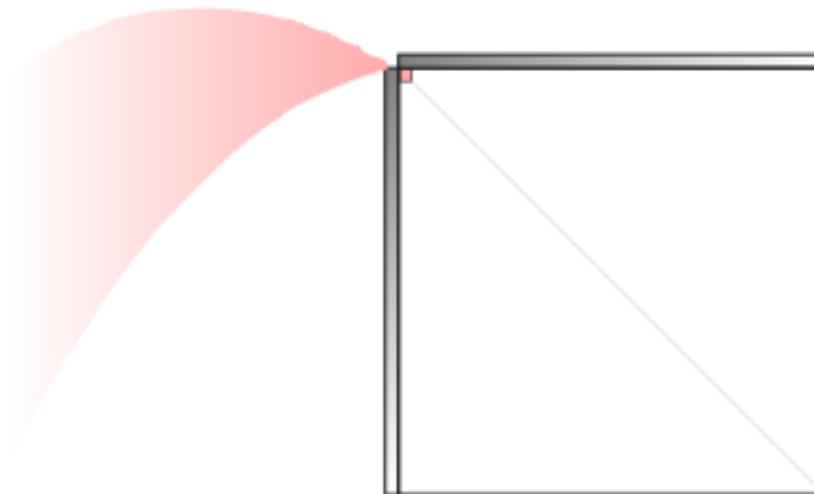
Landis *et al.*, 2013 (Syst Biol)

For more areas, Q explodes

3 areas

0	0	0	0	0	1	1	1	1
0	0	1	1	1	0	0	1	1
0	1	0	1	0	1	0	1	1
0	0	0	0	0	0	0	0	0
0	0	1	1	1	0	0	0	0
0	1	0	1	1	0	0	0	0
1	0	0	0	0	1	1	1	1
1	0	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1

10 areas



$$2^3 \times 2^3 = 8 \times 8$$

$$2^{10} \times 2^{10} = 1024 \times 1024$$

Intractable for more than ten areas

How to infer large Q?

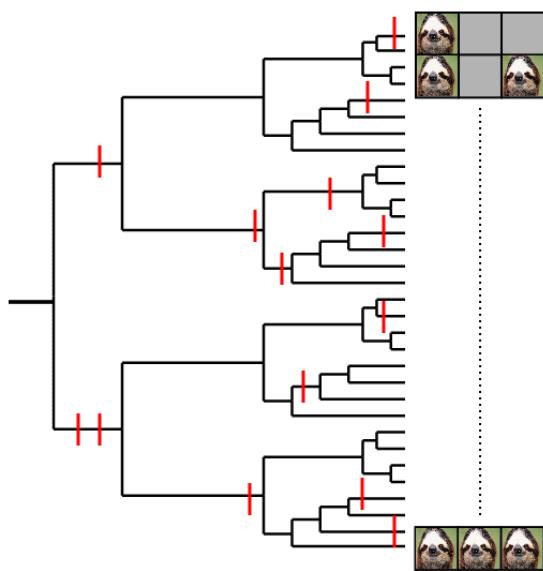
Inspired by Robinson *et al.*, 2003 (*Mol Biol Evol*)

Key concepts

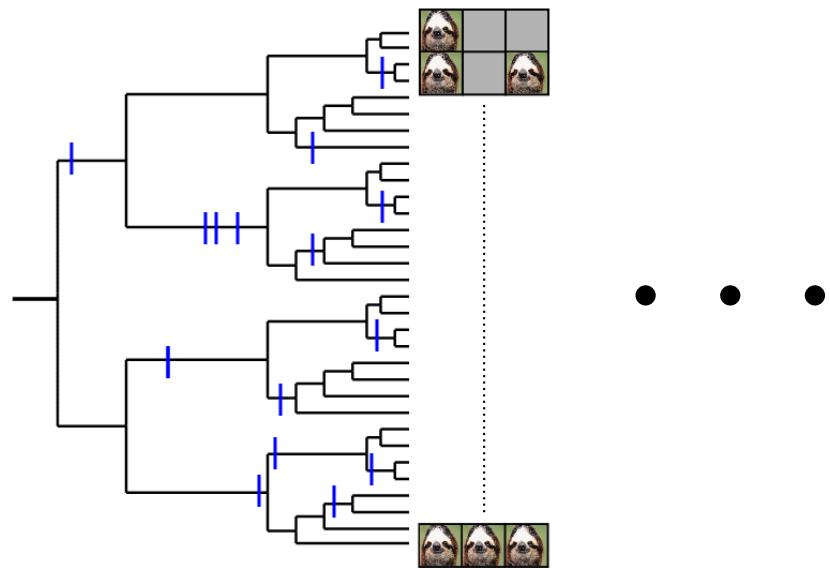
1. Sample over biogeographic histories, H
2. Compute likelihood, $\mathcal{L}_{\theta, H}$
3. Approximate $P(\theta, H \mid D)$ using
Markov chain Monte Carlo (MCMC)

1. Propose biogeographic histories, H

MCMC iteration n



MCMC iteration n+1



Stochastic mapping

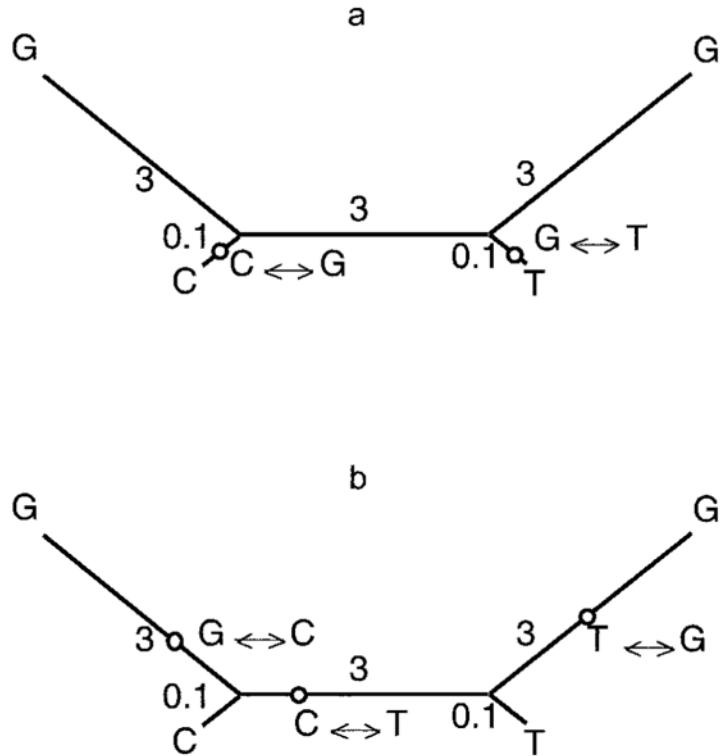


FIGURE 1. Two alternative mutational mappings. The numbers along the lineages indicate the length of the lineages measured in expected number of substitutions. The circles indicate mutations. (a) Parsimony mapping, which requires only two mutations. (b) The alternative mapping is almost 1,000 times as likely to be the right mapping.

e.g. 50 mappings

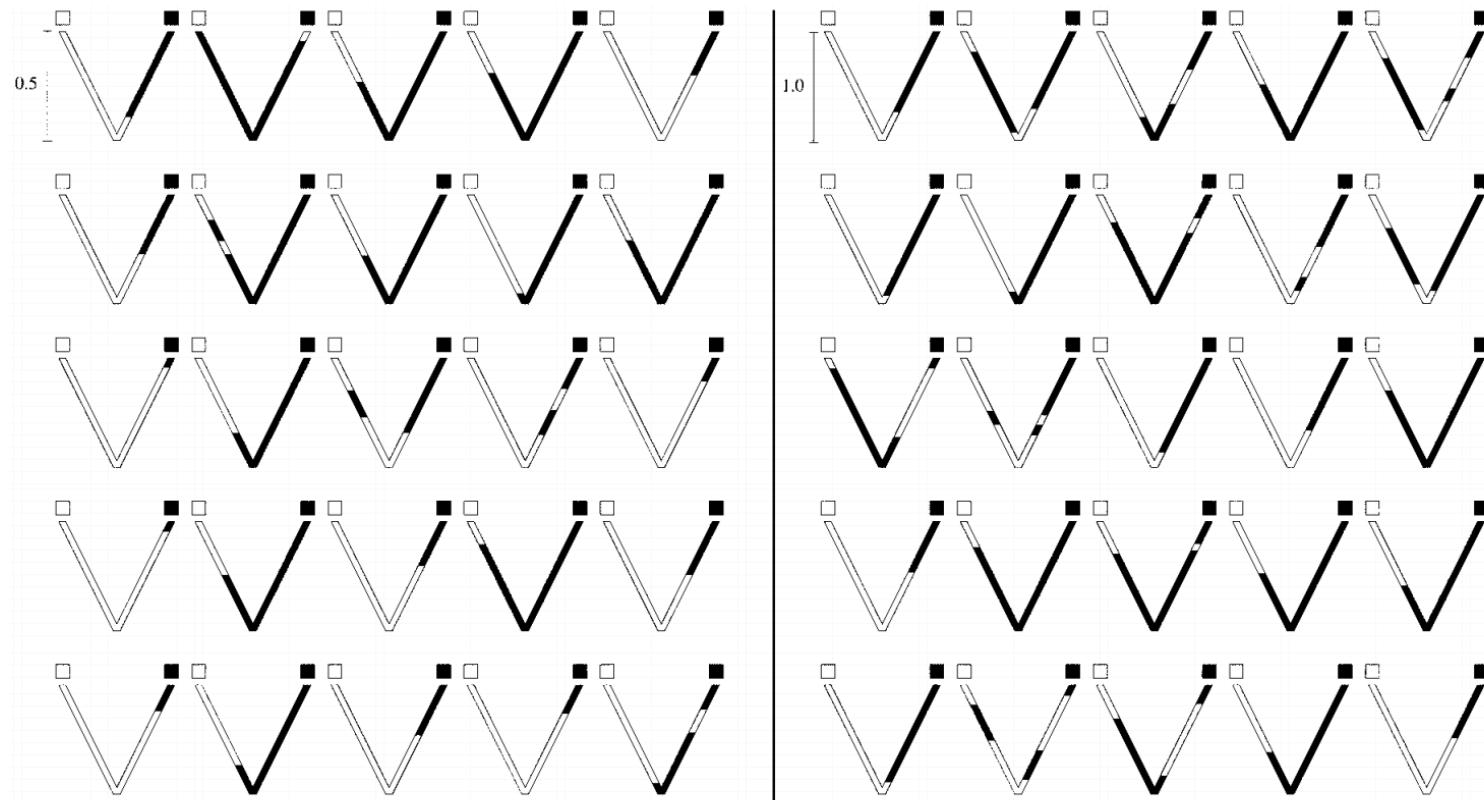


FIGURE 2. Fifty random realizations of character histories consistent with the observations at the tips of the tree when there are two character states with equal rates of change from $\blacksquare \rightarrow$ and $\square \rightarrow \blacksquare$. The tree length for the 25 realizations to the left of the line is one substitution, and the tree length for the 25 realizations to the right of the line is two substitutions.

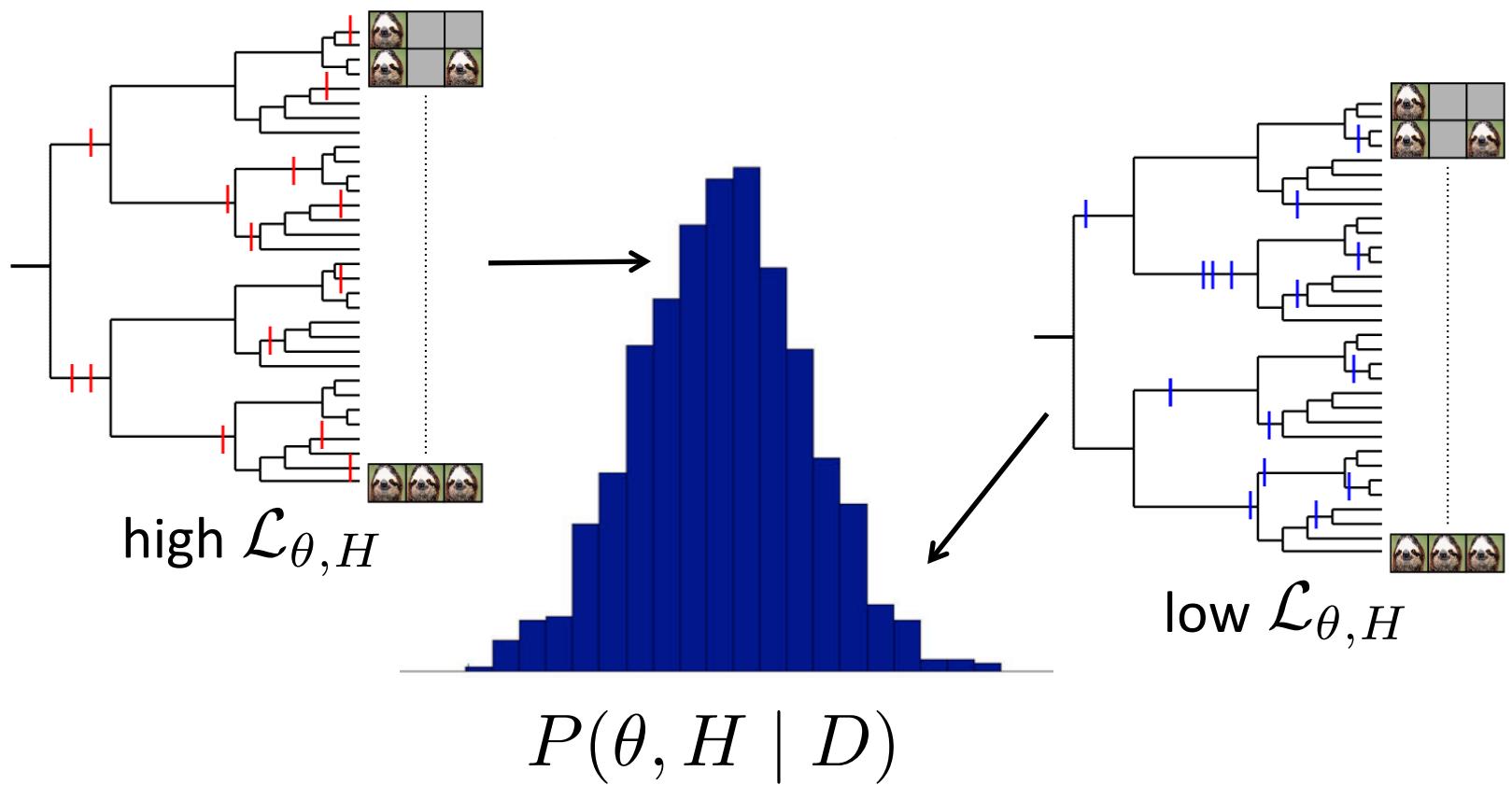
2. Compute likelihood, $\mathcal{L}_{\theta,H}$

Range evolution events from range $j \rightarrow i$:

$$\begin{aligned} r &= \sum r_j && \text{sum of rates leaving } j \\ re^{-rt} & && \text{prob any event at time } t \\ r_i/r & && \text{prob next event is } j \rightarrow i \end{aligned}$$

$\mathcal{L}_{\theta,H}$ = product of event types & times over tree

3. Approximate $P(\theta, H \mid D)$ using MCMC



3. Approximate $P(\theta, H \mid D)$ using MCMC

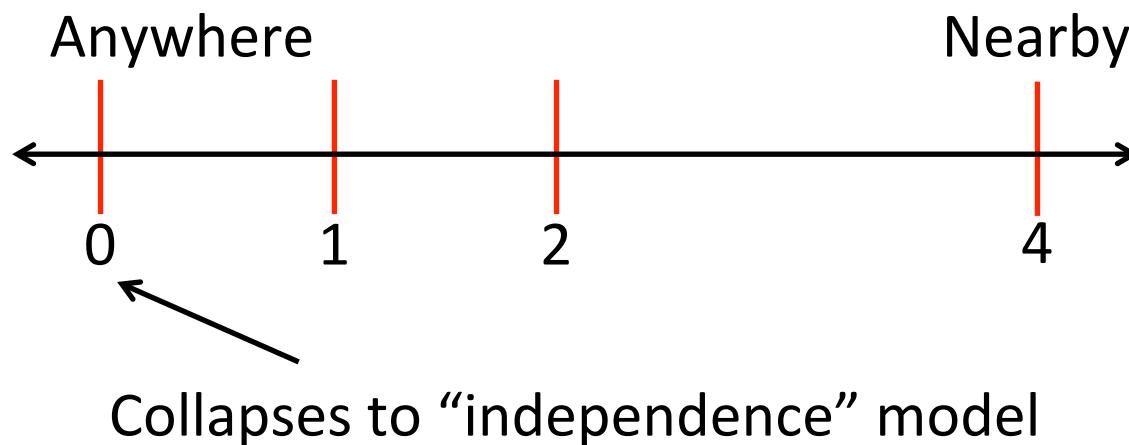
MCMC proposals

1. Parameter updates
2. Character history updates either
 1. Resamples single path
 2. Resamples single node and incident paths

Distance-dependent dispersal model

Infer distance effect parameter

Where is next dispersal event given current range?



Distance dependent rate matrix

$$R_{Y_i, Y_j}^{(a)} = \begin{cases} \lambda_0 & \text{if } Y_{j,a} = 0 \\ \lambda_1 \eta(Y_i, Y_j, a, \beta) & \text{if } Y_{j,a} = 1 \\ 0 & \text{if } Y_i \text{ and } Y_j \text{ differ at more than one area} \\ 0 & \text{if } Y_j = (0, 0, \dots, 0) \end{cases}$$

Extrapolation
Dispersal
Extinction

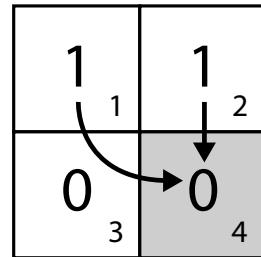
Uniform at random
Modified by distance
Forbidden

Distance-dependent rate modifier

$$\eta(Y_i = (1, 1, 0, 0) \rightarrow Y_j = (1, 1, 0, 1), a = 4, \beta) =$$

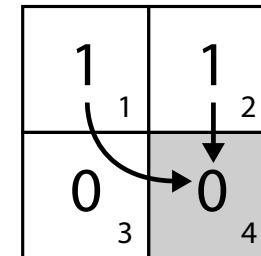
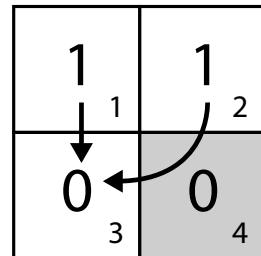
$$\underbrace{d(G_1, G_4)^{-\beta} + d(G_2, G_4)^{-\beta}}$$

Rate-modifier



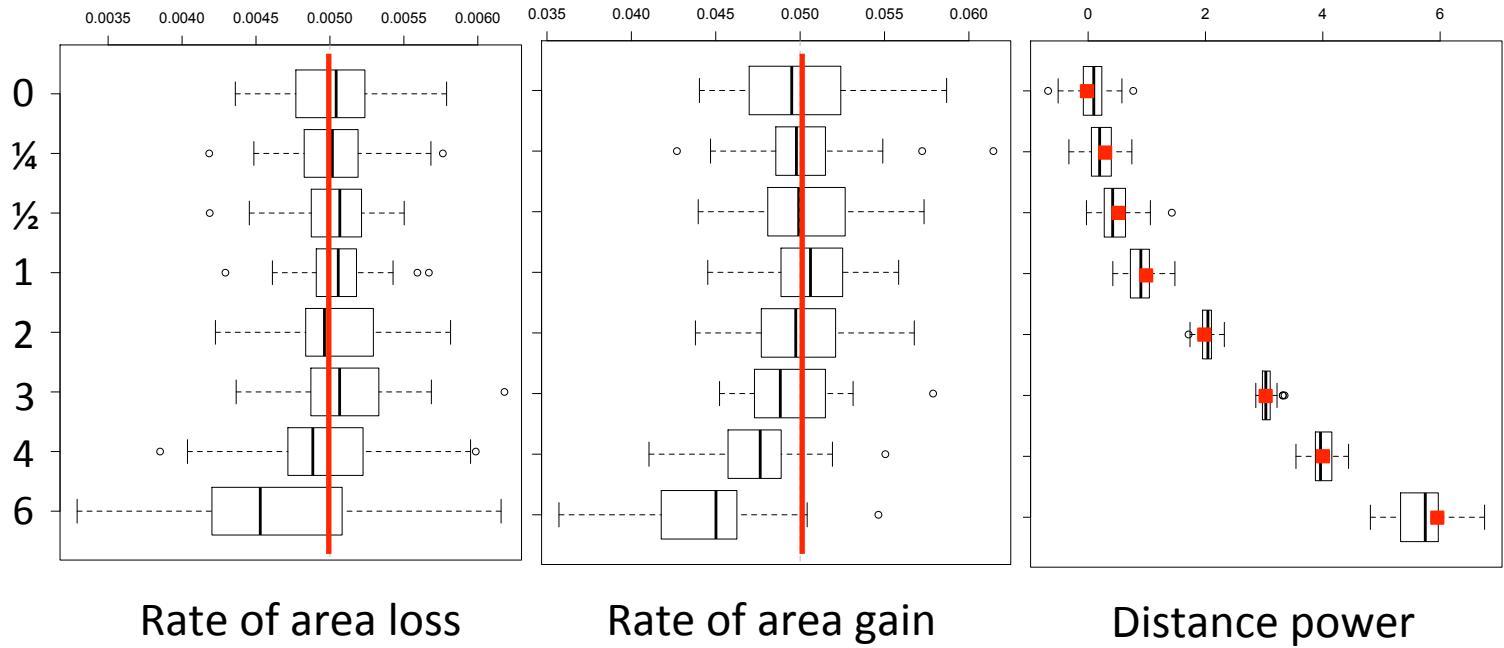
$$\times \underbrace{\frac{2}{d(G_1, G_3)^{-\beta} + d(G_2, G_3)^{-\beta}}}_{+} \underbrace{\frac{d(G_1, G_4)^{-\beta} + d(G_2, G_4)^{-\beta}}{d(G_1, G_4)^{-\beta} + d(G_2, G_4)^{-\beta}}}$$

Normalization

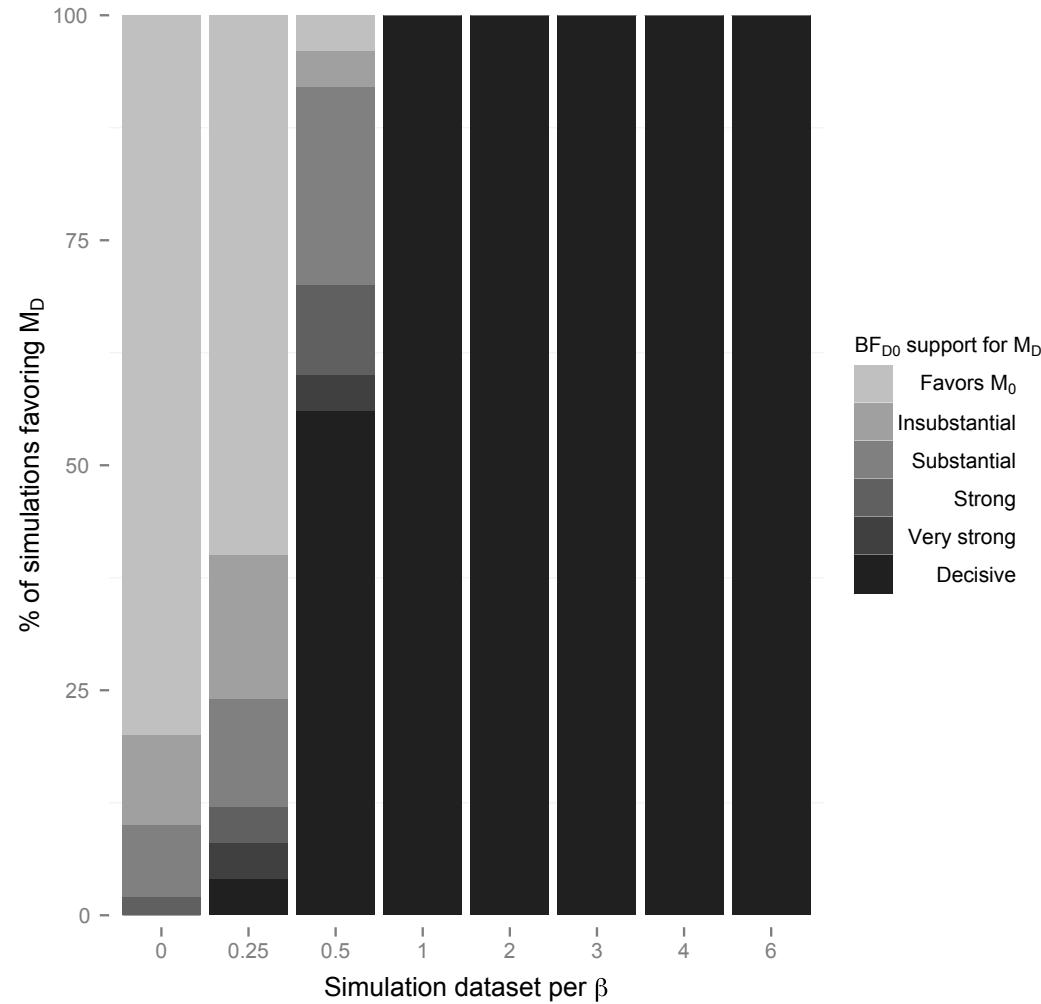


Bayesian inference vs truth

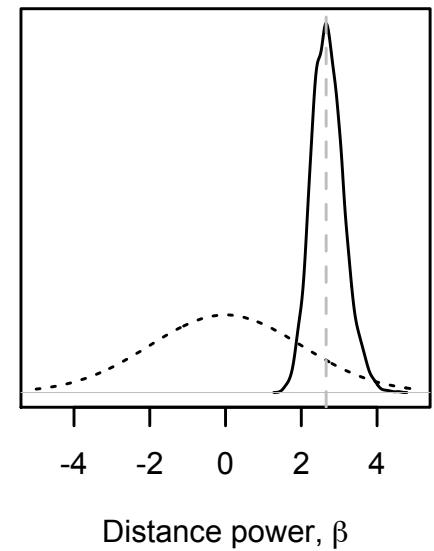
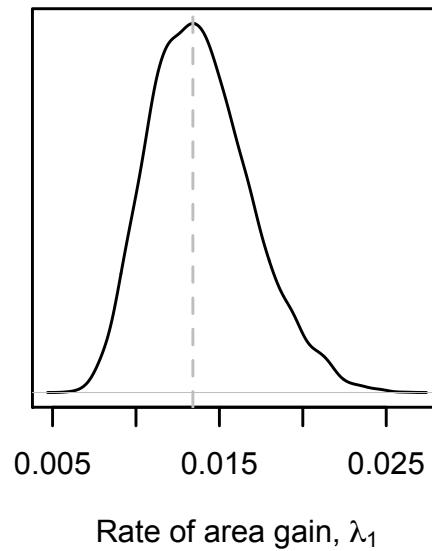
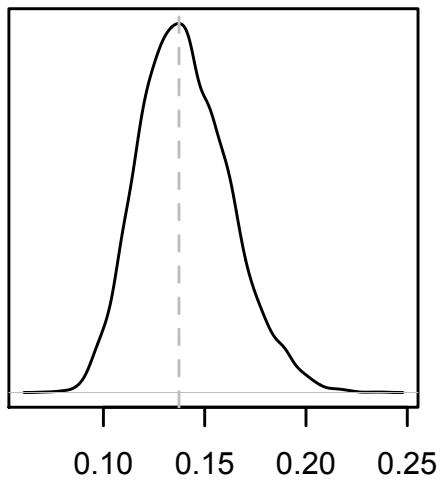
Distance power parameter



Bayes factors



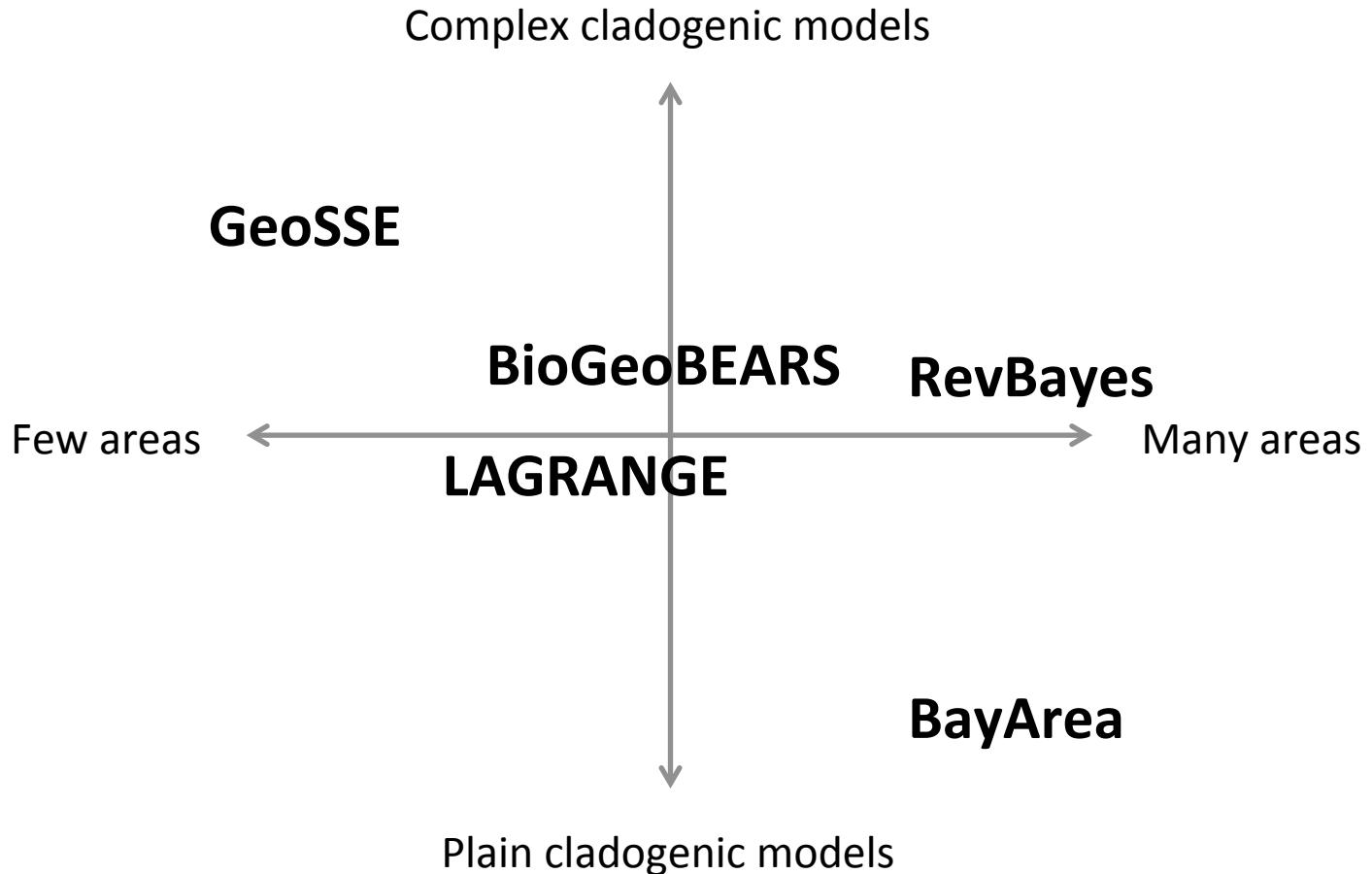
Vireya parameter inference



Discrete vs. continuous models

	Discrete	Continuous
Data	Transformed	As is
Model	CTMC Asymmetry easy	Diffusion (BM) Asymmetry hard
Individual/Endemic	Yes	Yes
Range	Yes (scales poorly)	No (currently)
Dispersal/Extirpation	Yes	NA
Cladogenesis	Yes	NA
Speciation/Extinction	Yes (for 2-3 areas)	Yes (for individuals)
Spatial heterogeneity	Easy	Hard
Temporal heterogeneity	Easy	Easy

Tradeoffs (my view)



Lab

RevBayes

Input

Model

Posterior parameter

Posterior range

Phylowood

Visualization

Papers

- Bielejec, F., A. Rambaut, M. A. Suchard, and P. Lemey. 2011. Spread: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* 27:2910–2912.
- Goldberg, E. E., L. T. Lancaster, and R. H. Ree. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* 60:451–465.
- Lamm, K. S. and B. D. Redelings. 2009. Reconstructing ancestral ranges in historical biogeography: properties and prospects. *Journal of Systematics and Evolution* 47:369–382.
- Landis, M. J. and T. Bedford. 2014. Phylowood: interactive web-based animations of biogeographic and phylogeographic histories. *Bioinformatics* 30:123–124.
- Landis, M. J., N. J. Matzke, B. R. Moore, and J. P. Huelsenbeck. 2013. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* 62:789–804.
- Lemey, P., A. Rambaut, A. J. Drummond, and M. A. Suchard. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* 5:e1000520.
- Lemey, P., A. Rambaut, J. J. Welch, and M. A. Suchard. 2010. Phylogeography takes a relaxed random walk incontinuous space and time. *Molecular Biology and Evolution* 27:1877–1885.
- Lemmon, A. A. and E. M. Lemmon. 2008. A likelihood framework for estimating phylogeographic history on a continuous landscape. *Systematic Biology* 57:544–561.
- Matzke, N. J. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* 5.
- Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.
- Ree, R. H. and I. Sanmartin. 2009. Prospects and challenges for parametric models in historical biogeographical inference. *Journal of Biogeography* 36:1211–1220.
- Ree, R. H. and S. A. Smith. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57:4–14.
- Ronquist, F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46:195–203.
- Ronquist, F. and I. Sanmartin. 2011. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics* 42:441–464.
- Sanmartin, I. and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* 53:216–243.
- Webb, C. O. and R. H. Ree. 2012. Historical biogeography inference in Malesia. Pages 191–215 in *Biotic evolution and environmental change in Southeast Asia* (D. Gower, K. Johnson, J. Richardson, B. Rosen, L. Ruber, and S. Williams, eds.) Cambridge University Press.

Software

BayArea

<https://code.google.com/p/bayarea/>

BEAST

http://beast.bio.ed.ac.uk/Main_Page

BioGeoBEARS

<http://cran.r-project.org/web/packages/BioGeoBEARS/index.html>

GeoSSE

<http://www.zoology.ubc.ca/prog/diversitree/>

LAGRANGE

<https://code.google.com/p/lagrange/>

RevBayes

<http://github.com/revbayes/revbayes>

SHIBA

<http://phylodiversity.net/shiba/>