

# Probabilistic Models of Geographic Range Evolution

Will Freyman

IB200, Spring 2016

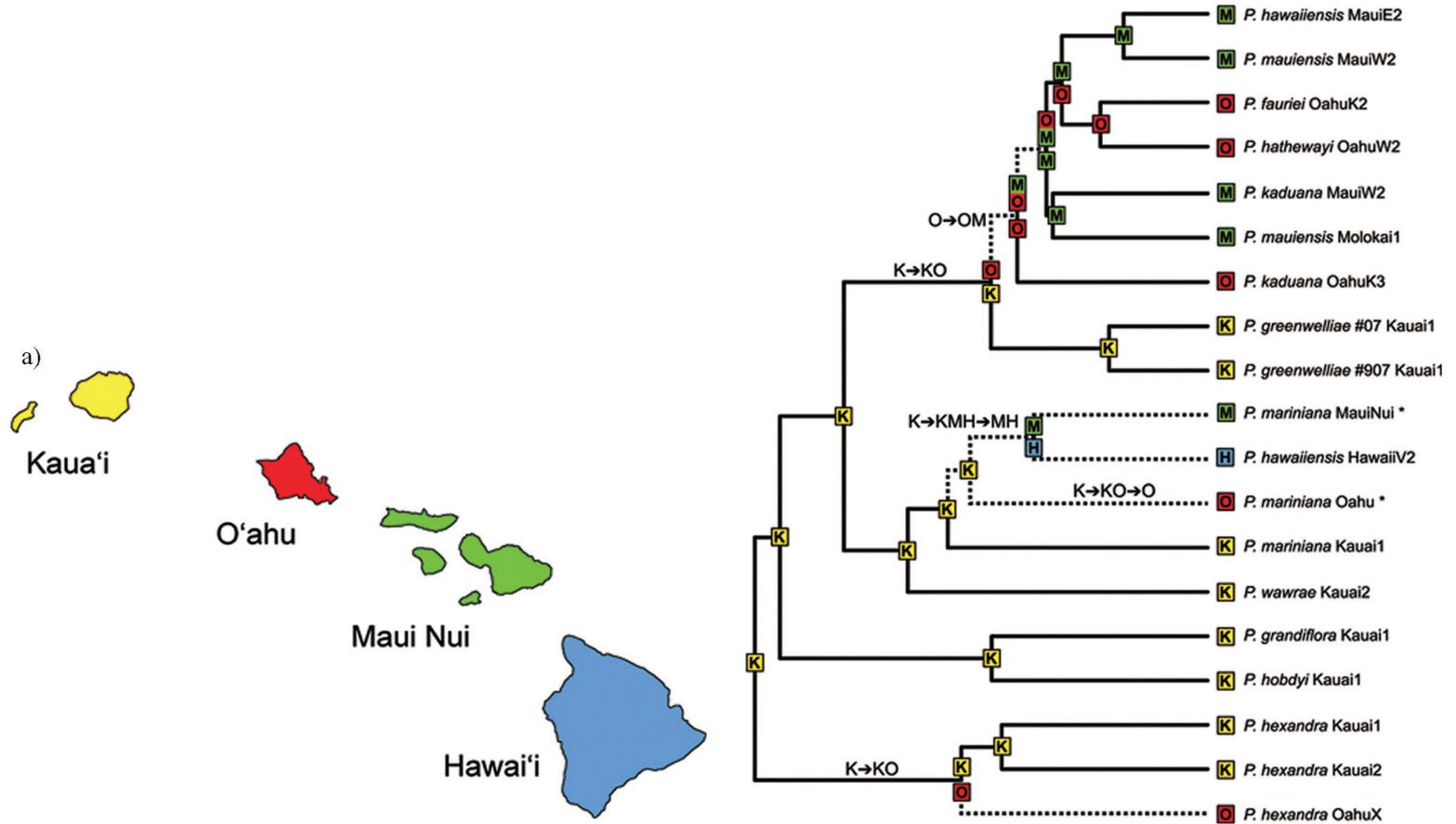


Image: Richard H Ree and Stephen A Smith. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1):4–14, 2008.

# Biogeographic histories on a phylogeny:

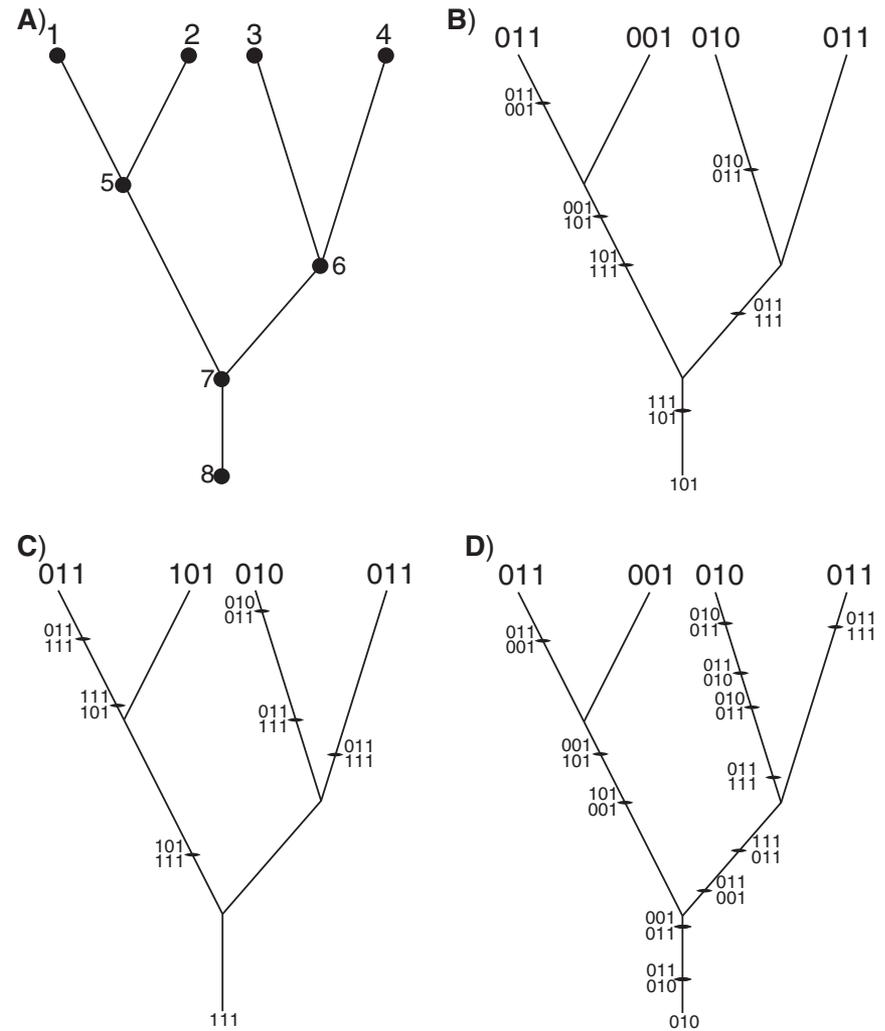


FIGURE 1. An example of a tree with  $M=4$  species. A) Nodes on the tree are labeled such that the tips of the tree have the labels  $1, 2, \dots, M$  whereas the interior nodes of the tree are labeled  $M+1, M+2, \dots, 2M$ . Note that in this article we also consider the “stem” branch of the tree, which connects the root node (node 7) and its immediate common ancestor (node 8). B–D) Several possible biogeographic histories—comprising 6, 6, and 12 events, respectively—that can explain the observed species ranges.

# The Dispersal-Extinction-Cladogenesis (DEC) model:

## Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis

RICHARD H. REE<sup>1</sup> AND STEPHEN A. SMITH<sup>2</sup>

<sup>1</sup>*Department of Botany, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605, USA; E-mail: rree@fieldmuseum.org*

<sup>2</sup>*Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520, USA*

*Abstract.*—In historical biogeography, model-based inference methods for reconstructing the evolution of geographic ranges on phylogenetic trees are poorly developed relative to the diversity of analogous methods available for inferring character evolution. We attempt to rectify this deficiency by constructing a dispersal-extinction-cladogenesis (DEC) model for geographic range evolution that specifies instantaneous transition rates between discrete states (ranges) along phylogenetic branches and apply it to estimating likelihoods of ancestral states (range inheritance scenarios) at cladogenesis events. Unlike an earlier version of this approach, the present model allows for an analytical solution to probabilities of range transitions as a function of time, enabling free parameters in the model, rates of dispersal, and local extinction to be estimated by maximum likelihood. Simulation results indicate that accurate parameter estimates may be difficult to obtain in practice but also show that ancestral range inheritance scenarios nevertheless can be correctly recovered with high success if rates of range evolution are low relative to the rate of cladogenesis. We apply the DEC model to a previously published, exemplary case study of island biogeography involving Hawaiian endemic angiosperms in *Psychotria* (Rubiaceae), showing how the DEC model can be iteratively refined from inspecting inferences of range evolution and also how geological constraints involving times of island origin may be imposed on the likelihood function. The DEC model is sufficiently similar to character models that it might serve as a gateway through which many existing comparative methods for characters could be imported into the realm of historical biogeography; moreover, it might also inspire the conceptual expansion of character models toward inclusion of evolutionary change as directly coincident, either as cause or consequence, with cladogenesis events. The DEC model is thus an incremental advance that highlights considerable potential in the nascent field of model-based historical biogeographic inference. [Ancestral state reconstruction; dispersal; extinction; Hawai'i; historical biogeography; *Psychotria*; speciation; vicariance.]

Instantaneous transition rate matrix of anagenetic DE events:

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

$$p(y \rightarrow z; t, \mathbf{Q}) = \left[ e^{-\mathbf{Q}t} \right]_{yz}$$

# Transition probabilities for cladogenetic events:

Types of speciation, and example descendant ranges:

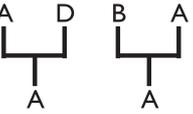
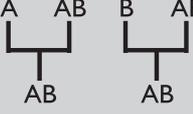
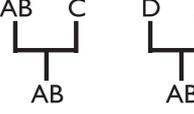
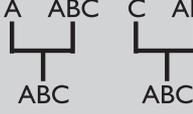
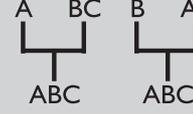
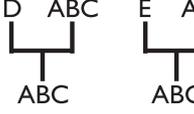
Ancestral ranges:	Sympatric (range copying)	Sympatric (subset)	Vicariance	Founder Event
A		--	--	
AB				
ABC				
ABCD				

FIGURE 1. Example range-inheritance events at cladogenesis. The events allowed by LAGRANGE's DEC model are highlighted in gray. Each allowed event is fixed to have equal probability in the DEC model. DEC requires that, at cladogenesis, at least one daughter species must have a range size of 1, which excludes some events. Founder-event speciation (column 4) is also excluded from DEC, but included in DEC+J.

# Ancestral state reconstructions using DEC:

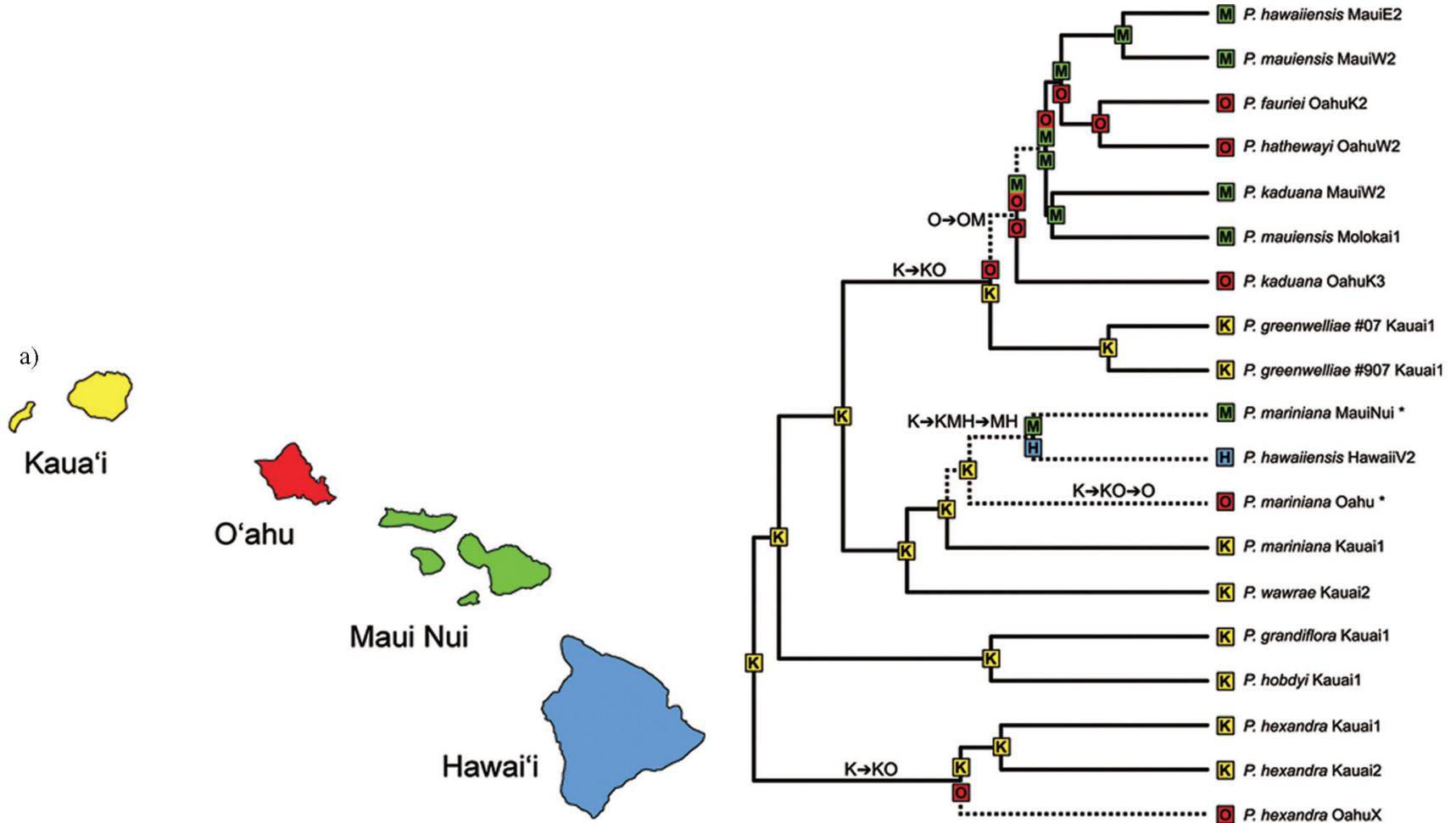


Image: Richard H Ree and Stephen A Smith. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1):4–14, 2008.

# Time stratified models:

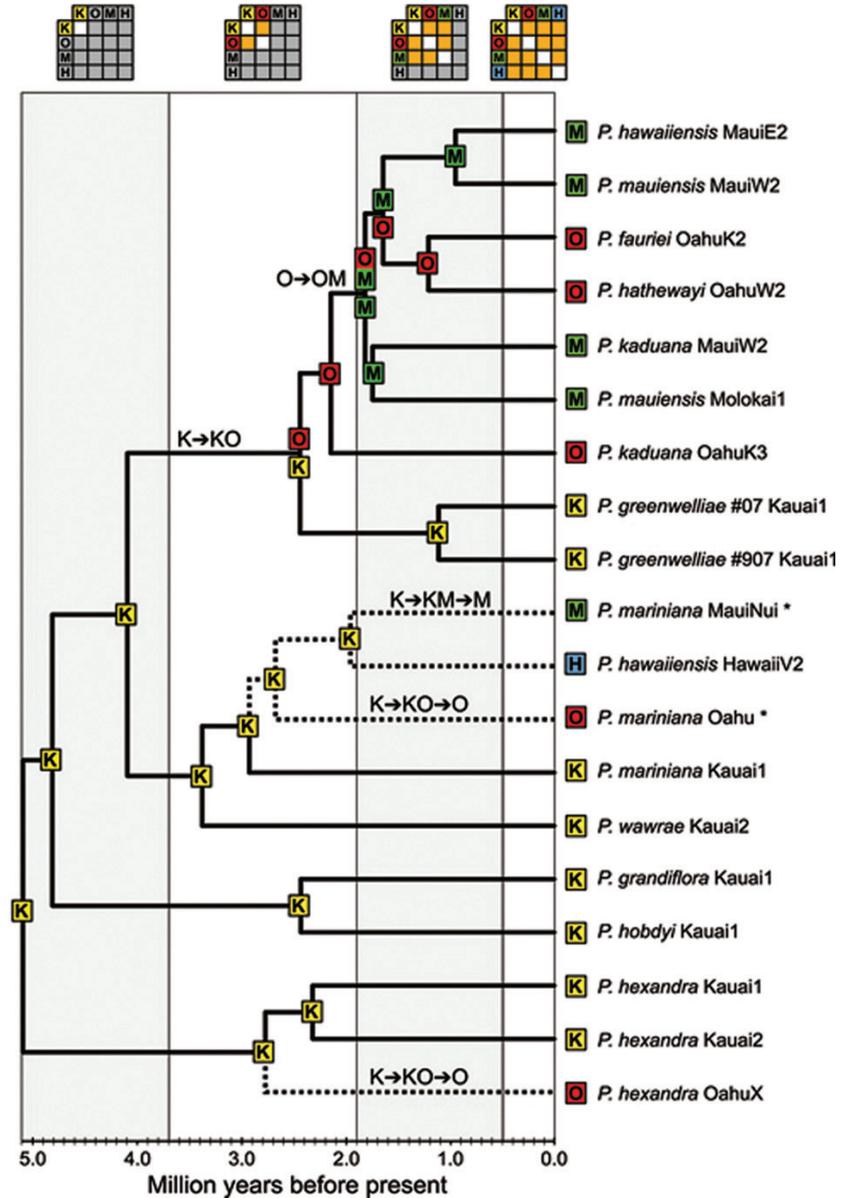
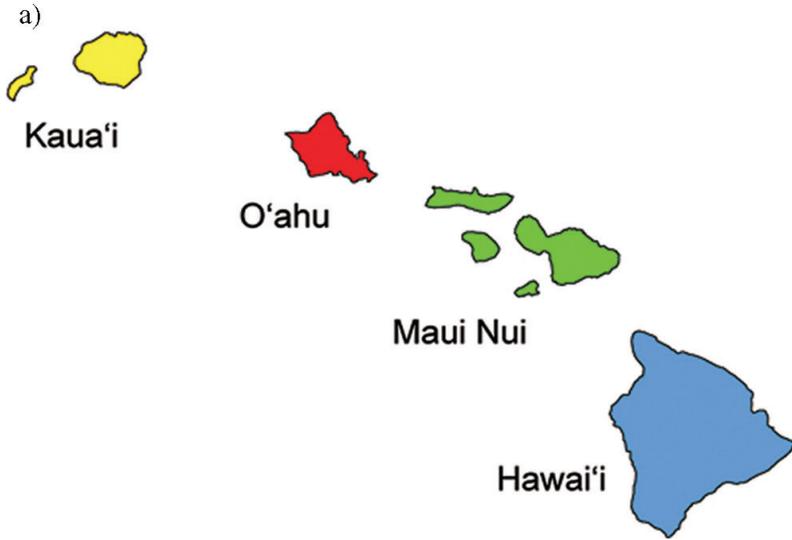


Image: Richard H Ree and Stephen A Smith. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1):4–14, 2008.

# Extending the DEC:

	Process	Ranges		Character mapping	DIVA	DEC (GeoSSE, LAGRANGE)	BayArea, BBM (RASP)	Parameter of BioGeoBEARS Supermodel
		Before	After					
Anagenetic	Dispersal				✓	✓	✓	$d$ (& $x, b$ )
	Extinction				✓	✓	✓	$e$ (& $u, b$ )
	Range-switching			✓				$a$ (& $x, b$ )
Cladogenetic	Sympatry (narrow)			✓	✓	✓	✓	$y$ (& $mx0ly$ )
	Sympatry (widespread)					✓	$y$ (& $mx0ly$ )	
	Sympatry (subset)					✓	$s$ (& $mx0ls$ )	
	Vicariance (narrow)				✓	✓	$v$ (& $mx0lv$ )	
	Vicariance (widespread)				✓		$v$ (& $mx0lv$ )	
	Founder							$j$ (& $x, mx0lj$ )

# Future directions:

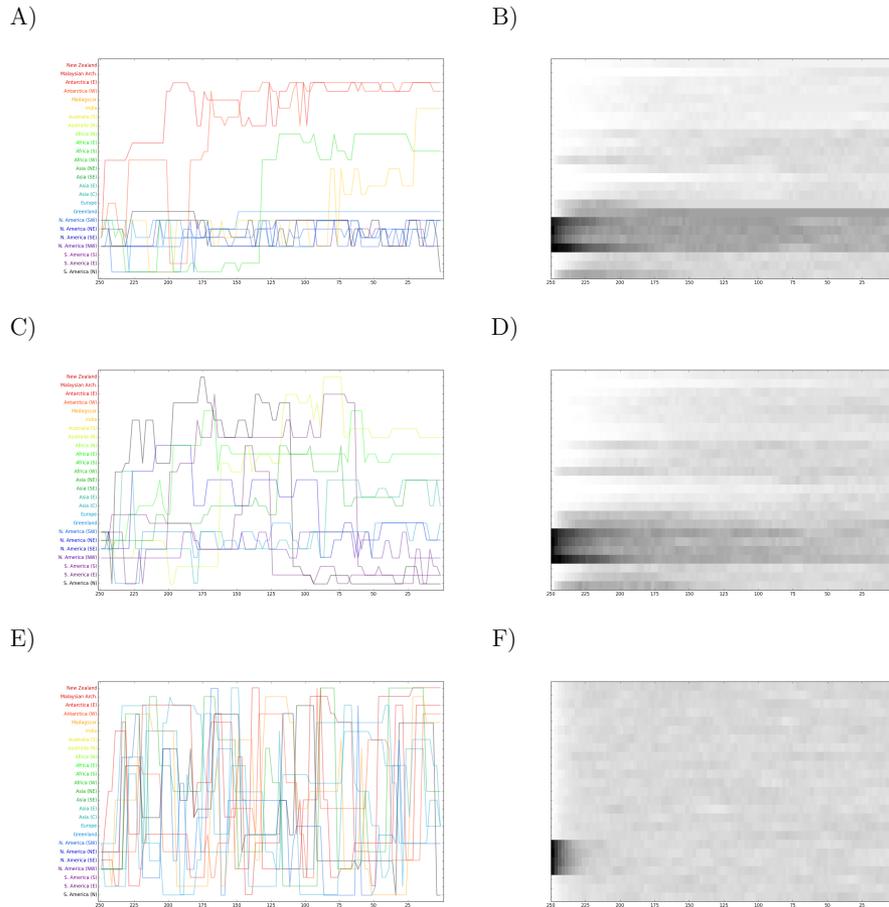


Figure 6: Sample paths for paleogeographically informed biogeographic process. The top, middle, and bottom panels show dispersal histories simulated by the pure short (A,B), medium (C,D), and long (E,F) distance process components. All processes originate in one of the four North American areas 250 Ma. The left column shows 10 of 2000 sample paths. Color indicates the area the lineage is found in the present (A,C,E). Colors for areas match those in Figure 8. The right column heatmap reports the sample frequencies for any of the 2000 dispersal process being in that state at that time (B,D,F).

# Biogeographic dating of speciation times using paleogeographically informed processes

MICHAEL J. LANDIS<sup>1</sup>

<sup>1</sup>*Department of Integrative Biology  
University of California, Berkeley, CA 94720-3140, U.S.A.*

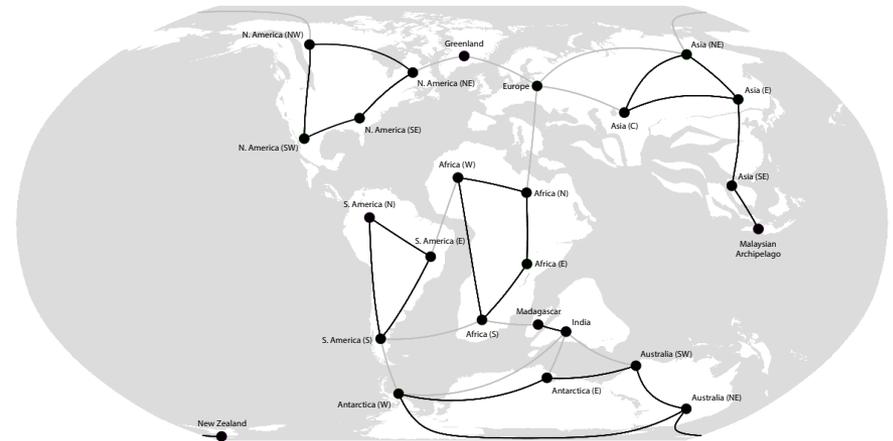


Figure 5: Dispersal graph for Epoch 14, 110-100Ma: India and Madagascar separate from Australia and Antarctica. A gplates (Gurnis et al. 2012) screenshot of Epoch 14 of 26 is displayed. Areas are marked by black vertices. Black edges indicate both short- and medium distance dispersal routes. Gray edges indicate exclusively medium distance dispersal routes. Long distance dispersal routes are not shown, but are implied to exist between all area-pairs. The short, medium, and long dispersal graphs have 8, 1, and 1 communicating classes, respectively. India and Madagascar each have only one short distance dispersal route, which they share. Both areas maintain medium distance dispersal routes with various Gondwanan continents during this epoch. The expansion of the Tethys Sea impedes dispersal into and out of Europe.