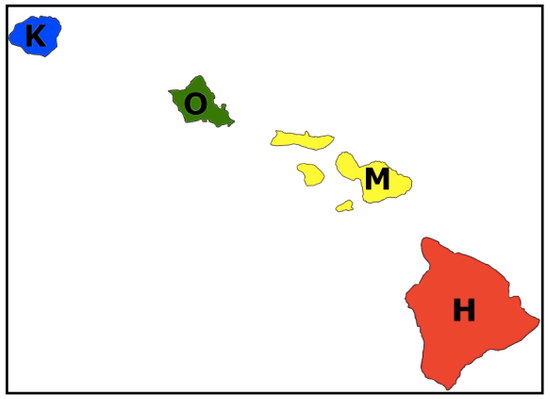
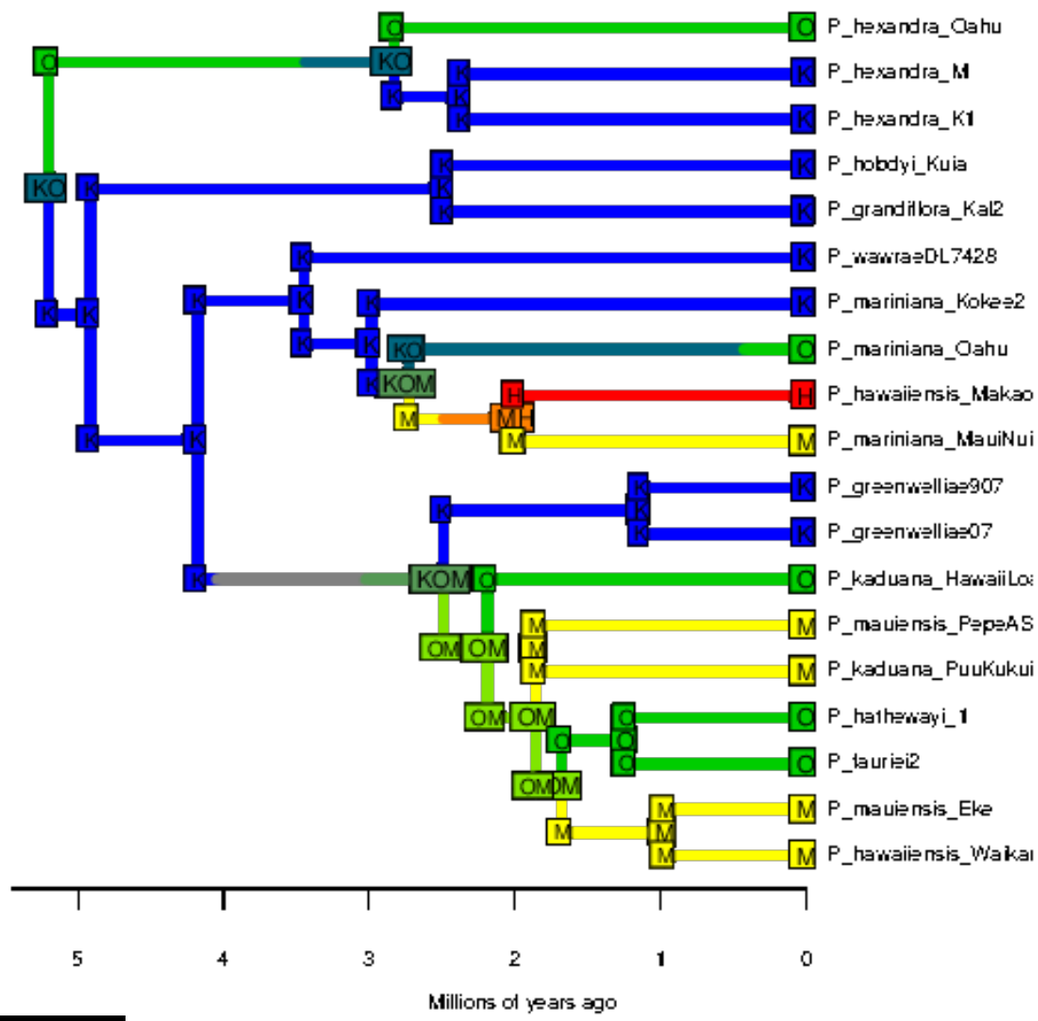


rexpokit and cladoRcpp: R packages integrating FORTRAN and C++ for faster matrix exponentiation and likelihood calculations in historical biogeography

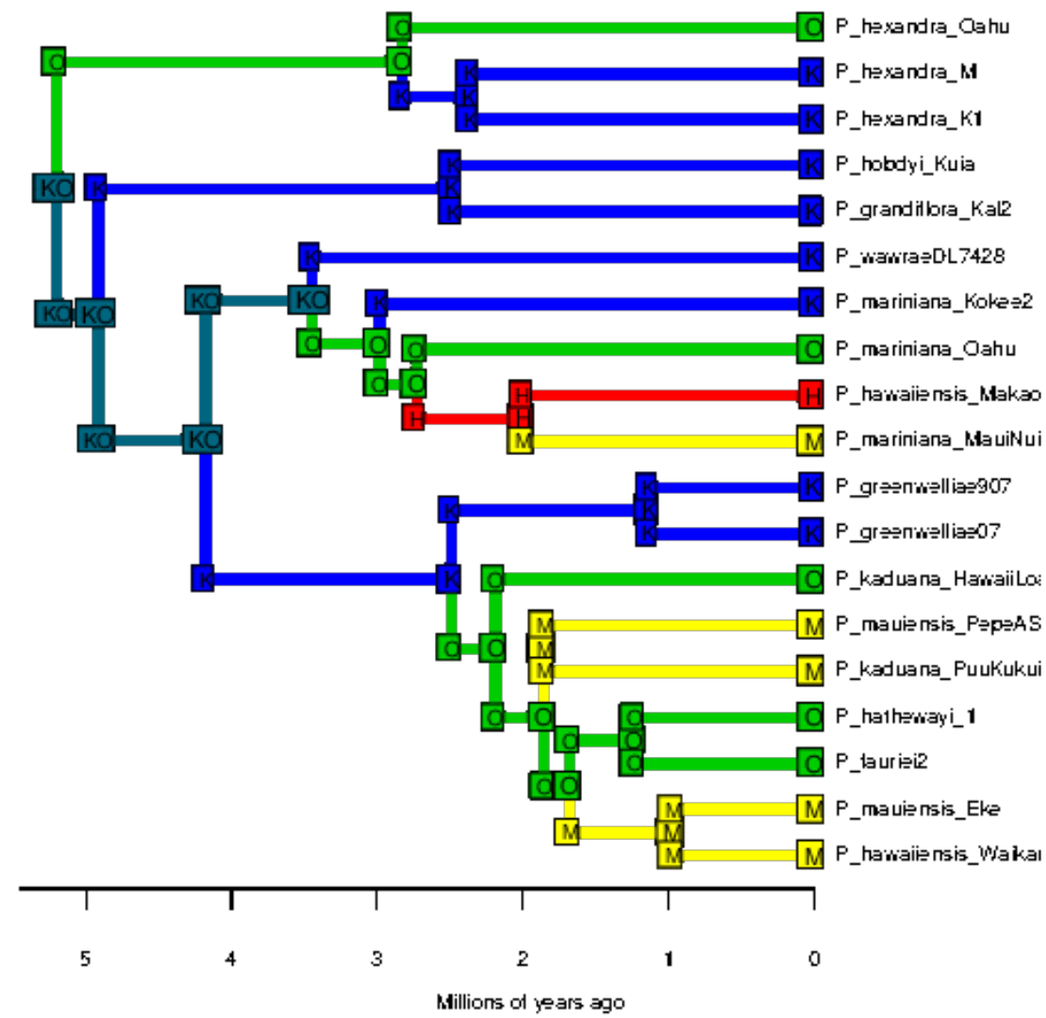
Nicholas J. Matzke, Postdoctoral Fellow, NIMBioS
(National Institute of Mathematical, www.nimbios.org)
iEvoBio 2014, Raleigh, NC, Tuesday, June 24, 2014
Software Bazaar, Room 402, 3:15-5:00 pm
Coauthor: Drew Schmidt, University of Tennessee

Hawaii *Psychotria* BioGeoBEARS stochastic mapping:

DEC (LnL=-34.5)



DEC+J (LnL=-20.9)



	Process	Ranges Before	Ranges After	Character mapping	DIVA	DEC (GeoSSE, LAGRANGE)	BayArea, BBM (RASP)	Parameter of BioGeoBEARS Supermodel
Anagenetic	Dispersal				✓	✓	✓	*d* (& *x,b*)
	Extinction				✓	✓	✓	*e* (& *u,b*)
	Range-switching			✓				*a* (& *x,b*)
Cladogenetic	Sympatry (narrow)			✓	✓	✓	✓	*y* (& *mx01y*)
	Sympatry (widespread)						✓	*y* (& *mx01y*)
	Sympatry (subset)					✓		*s* (& *mx01s*)
	Vicariance (narrow)				✓	✓		*v* (& *mx01v*)
	Vicariance (widespread)				✓			*v* (& *mx01v*)
	Founder							*j* (& *x, mx01j*)

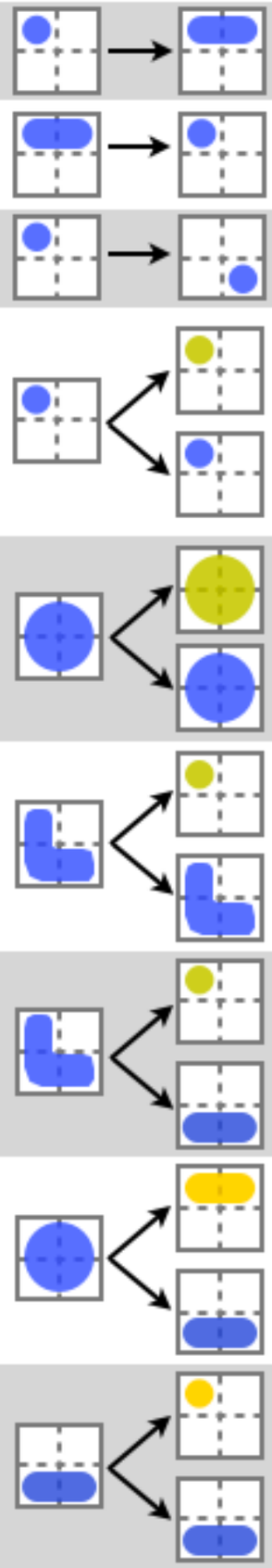
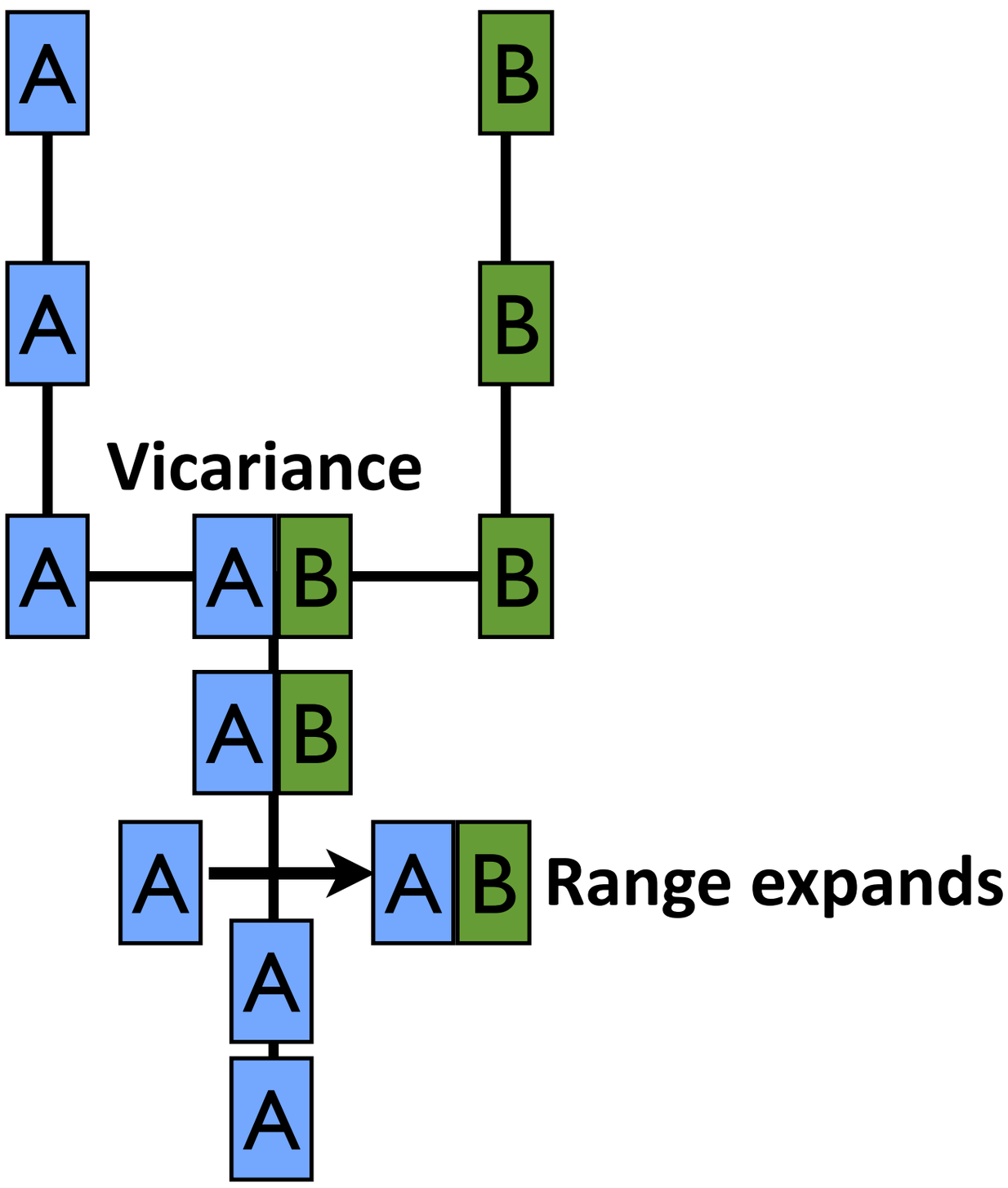


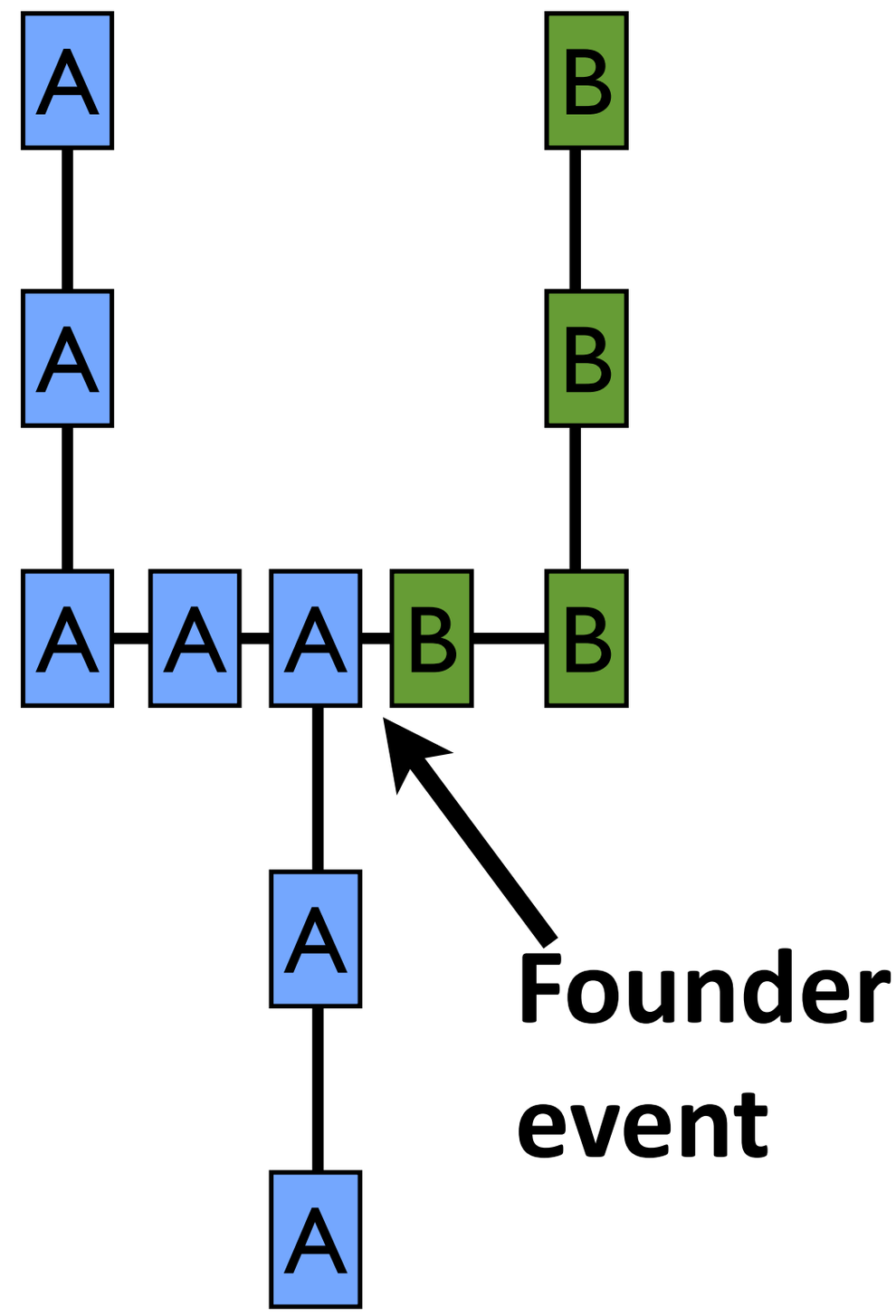
Figure 1, Matzke 2013, *Frontiers of Biogeography*

BioGeoBEARS: testing biogeographical models

LAGRANGE or DIVA

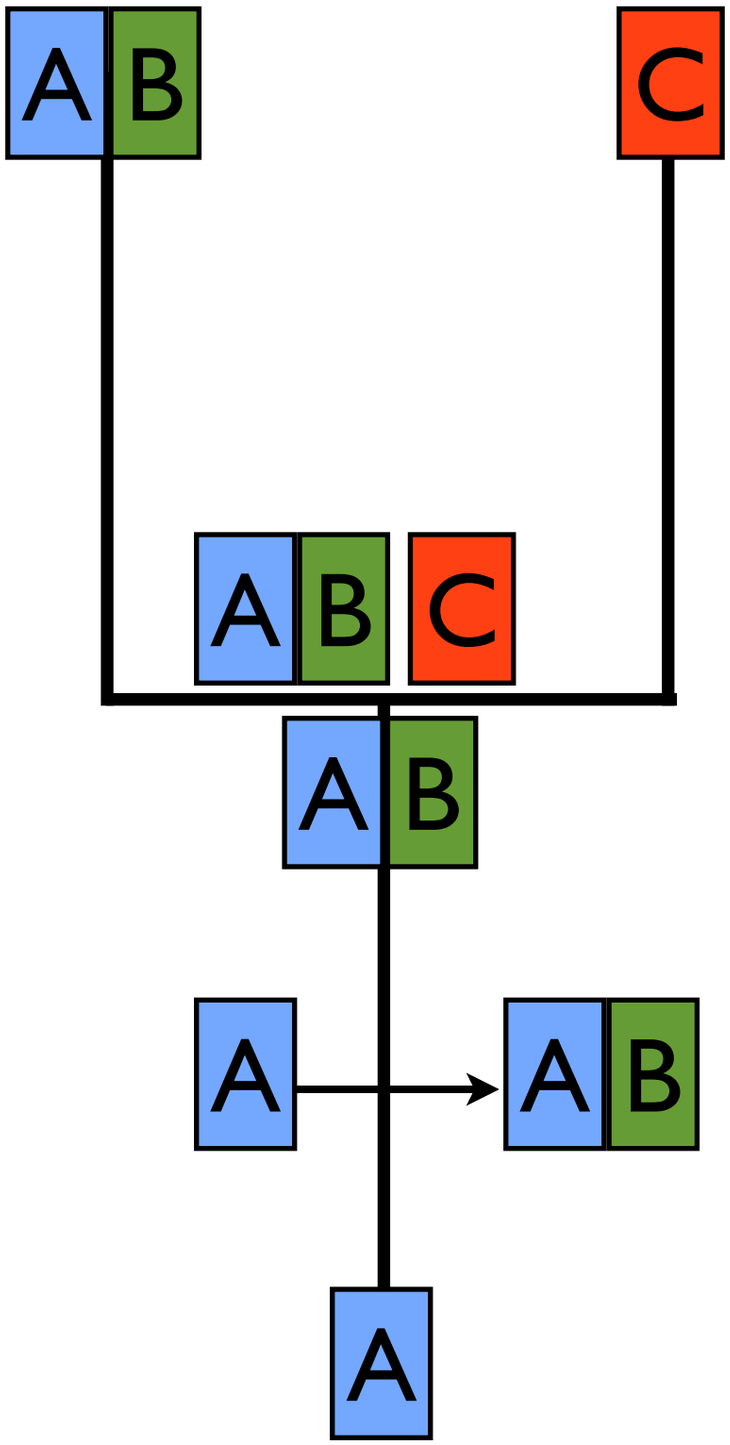


Founder-event



Test biogeographic processes in BioGeoBEARS

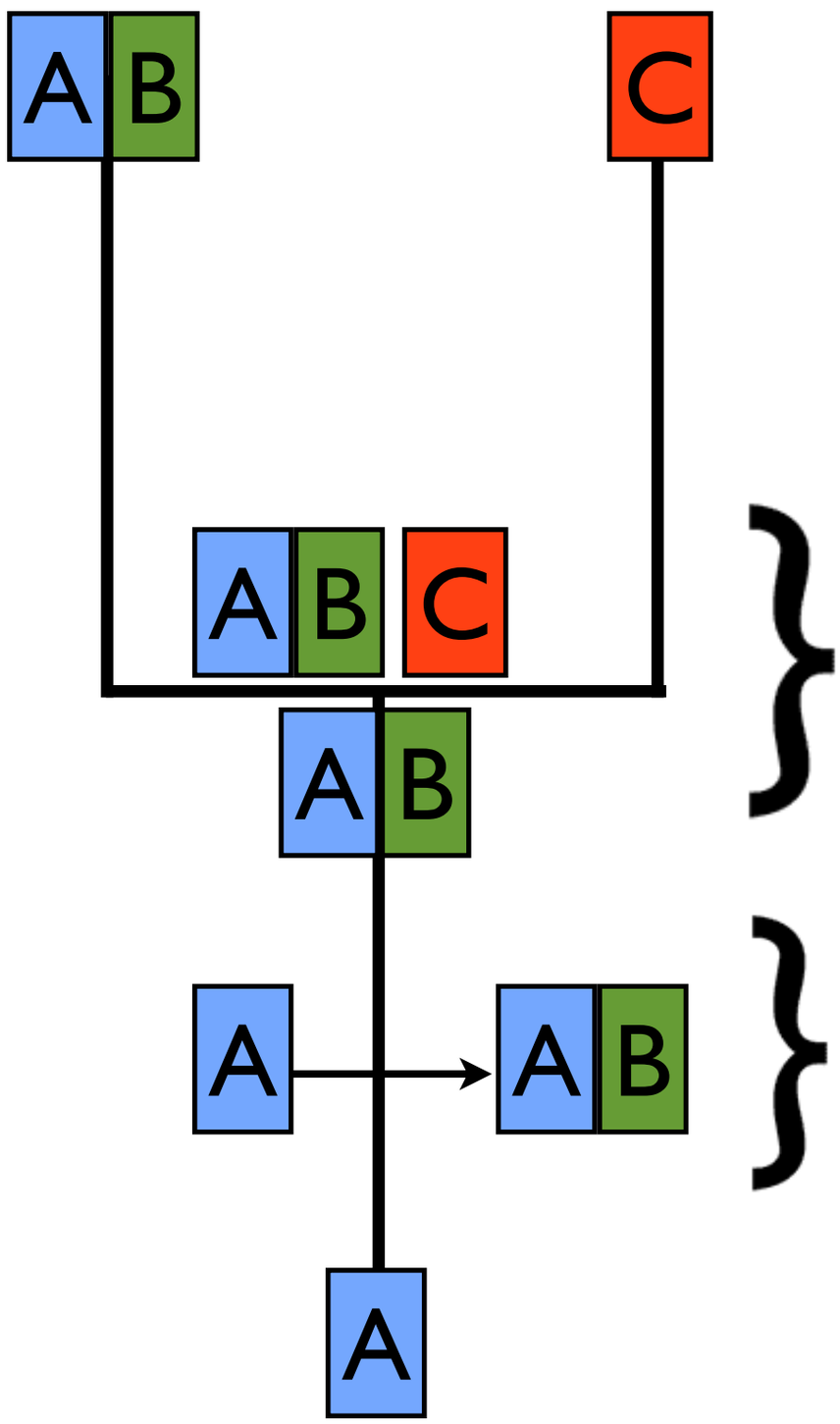
I implemented this method in *three* R packages:



Package #1, **rexpokit**, handles anagenic events requiring matrix exponentiation (FORTRAN EXPOKIT library, C++, Rcpp)

Test biogeographic processes in BioGeoBEARS

I implemented this method in *three* R packages:

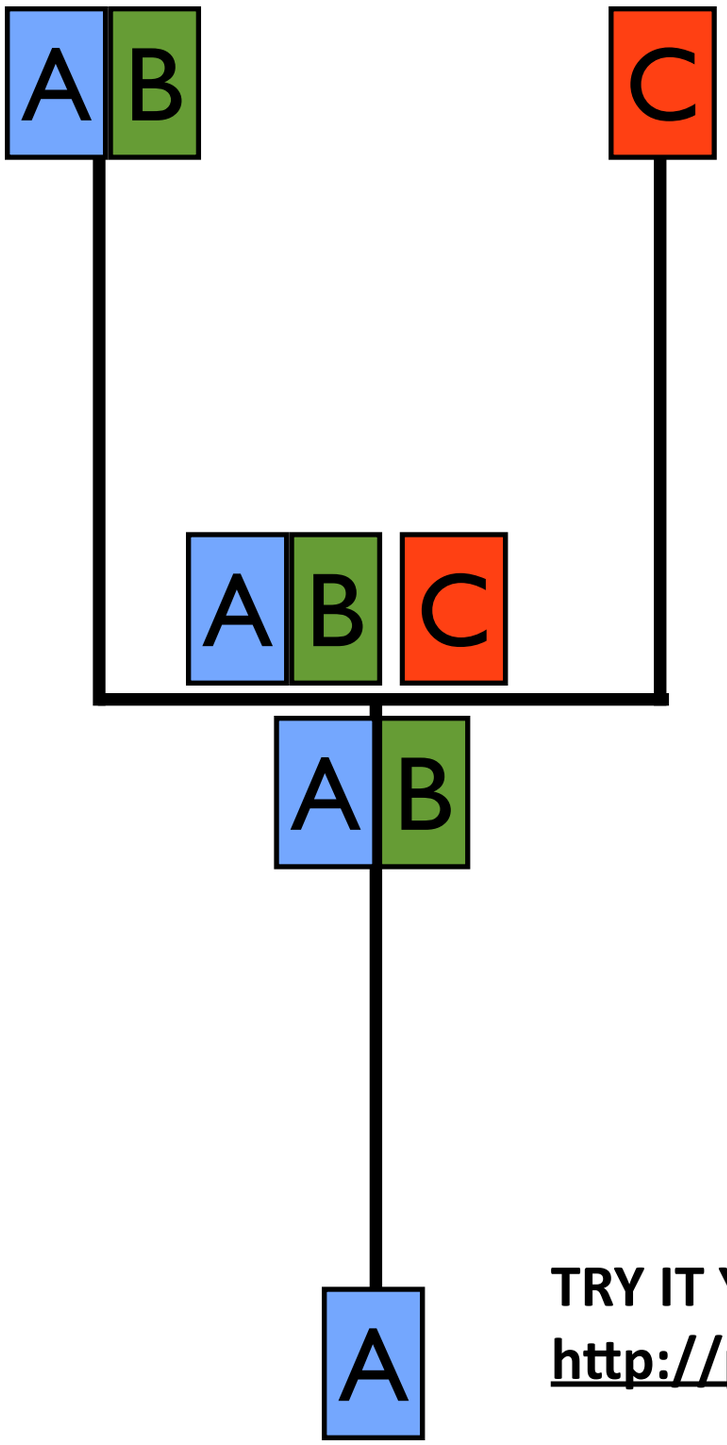


Package #2, **cladoRcpp**, handles cladogenesis events (C++, Rcpp)

Package #1, **rexpokit**, handles anagenic events requiring matrix exponentiation (FORTRAN EXPOKIT library, C++, Rcpp)

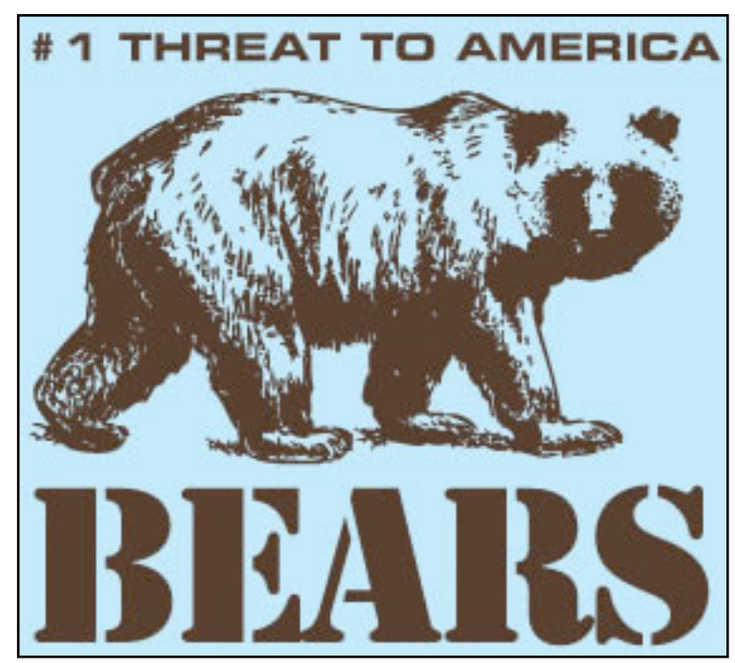
Test biogeographic processes in BioGeoBEARS

I implemented this method in *three* R packages:



Package #3, **BioGeoBEARS**, does the ML/Bayesian searches, model testing, etc.

BioGeographic Bayesian Evolutionary Analysis with R Scripts



TRY IT YOURSELF!
http://phylo.wikidot.com/biogeobears

Dealing with slowness #1: Matrix exponentiation

$$e^{\boldsymbol{Q}\nu} = \boldsymbol{P}(\nu)$$

v = branch length
Q = Q matrix
P = Probability transition matrix

Dealing with slowness #1: Matrix exponentiation

$$e^{\boldsymbol{Q}\nu} = \boldsymbol{P}(\nu)$$

DNA: 4x4 Q matrix
Amino acids: 20x20 matrix
Biogeography: # ranges x # ranges matrix

Number of states in biogeography

$$\# \text{ states} = 2^{\text{number of areas}}$$

2 areas = 2^2 = 4 states

4 areas = 2^4 = 16 states

8 areas = 2^8 = 256 states

10 areas = 2^{10} = 1024 states

Exponentiating a 1024x1024 matrix is slow

Dealing with slowness #1: Matrix exponentiation

SIAM REVIEW
Vol. 45, No. 1, pp. 3–000

Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later*

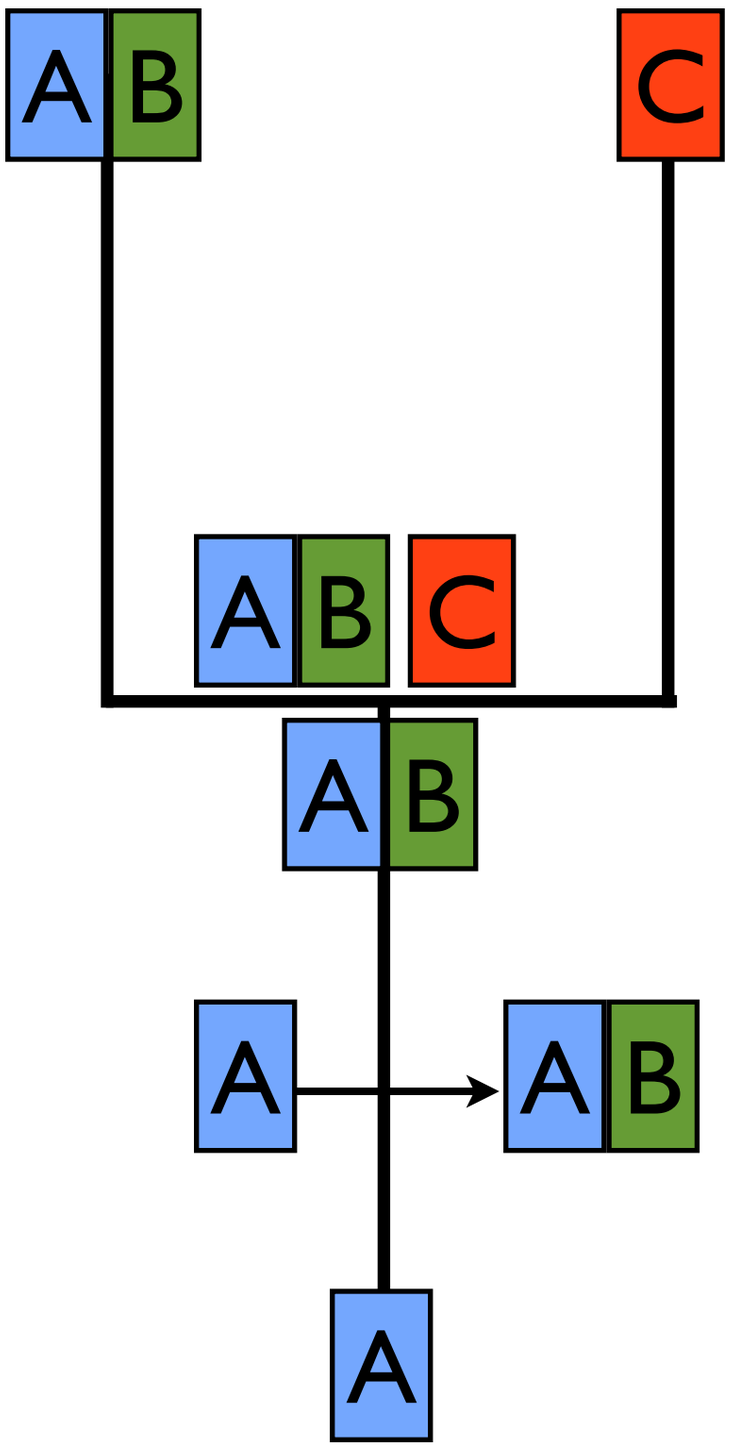
Cleve Moler†
Charles Van Loan‡

Abstract. In principle, the exponential of a matrix could be computed in many ways. Methods involving approximation theory, differential equations, the matrix eigenvalues, and the matrix characteristic polynomial have been proposed. In practice, consideration of computational stability and efficiency indicates that some of the methods are preferable to others, but that none are completely satisfactory.
Most of this paper was originally published in 1978. An update, with a separate bibliography, describes a few recent developments.

Key words. matrix, exponential, roundoff error, truncation error, condition

AMS subject classifications. 15A15, 65F15, 65F30, 65L99

PII. S0036144502418010



Package #1, **rexpokit**, handles anagenic events requiring matrix exponentiation (FORTRAN EXPOKIT library, C++, Rcpp)

R Package: rexpokit

rexpokit: R wrappers for EXPOKIT; other matrix functions

This package wraps some of the matrix exponentiation utilities from EXPOKIT (http://www.maths.uq.edu.au/expokit/), a FORTRAN library that is widely recommended for matrix exponentiation (Sidje RB, 1998. "Expokit: A Software Package for Computing Matrix Exponentials." ACM Trans. Math. Softw. 24(1): 130-156). EXPOKIT includes functions for exponentiating both small, dense matrices, and large, sparse matrices (in sparse matrices, most of the cells have value 0). Rapid matrix exponentiation is useful in phylogenetics when we have a large number of states (as we do when we are inferring the history of transitions between the possible geographic ranges of a species), but is probably useful in other ways as well.

Version:	0.24.1
Depends:	methods, SparseM, Rcpp (≥ 0.9.10)
LinkingTo:	Rcpp
Published:	2013-07-15
Author:	Nicholas J. Matzke [aut, cre, cph], Roger B. Sidje [aut, cph]
Maintainer:	Nicholas J. Matzke <matzke at berkeley.edu>
License:	GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL:	http://phylo.wikidot.com/rexpokit
NeedsCompilation:	yes
Citation:	rexpokit citation info
Materials:	NEWS
CRAN checks:	rexpokit results

Downloads:

Reference manual:	rexpokit.pdf
Package source:	rexpokit_0.24.1.tar.gz
Windows binaries:	r-devel: rexpokit_0.24.1.zip, r-release: rexpokit_0.24.1.zip, r-oldrel: rexpokit_0.24.1.zip
OS X Snow Leopard binaries:	r-release: rexpokit_0.24.1.tgz, r-oldrel: rexpokit_0.24.1.tgz

R Package: rexpokit

Download | Support

Expokit

- Download
- copyright
- Changes
- Paper (PDF)

Frequently Asked Questions

Expokit is a software package that provides matrix exponential routines for small dense or *very* large sparse matrices, real or complex. Here you will find the source code in **Fortran** and **Matlab**. The native Fortran version is embeddable in C/C++.

Understanding Expokit

1. w(t) = exp(tA)v is the analytic solution of the homogeneous ODE problem: w'(t) = Aw(t), w(0) = v.
2. w(t) = exp(tA)v + t*phi(tA)u, where phi(x) = (exp(x)-1)/x, is the analytic solution of the nonhomogeneous ODE problem: w'(t) = Aw(t) + u, w(0) = v.

Expokit handles both cases.

From Computing Reviews: *"This amazingly complete paper gives the theoretical background of the methods, presents applications, provides computational results for a standard set of pathological matrices, and closes with a mini-user's manual (...) It is entirely self-contained (...) Here we are given several sure ways to compute (...) exponentials, and these with rather sharp error bounds."*
[Full Review]

Moler and Van Loan: *"The most extensive software for computing the matrix exponential that we are aware of is Expokit."*
[Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. SIAM Review, Vol 45, No 1, March 2003.]

R Package: rexpokit

Dense matrix exponentiation: EXPOKIT's DGPADM —

(faster than the other R options, esp. for large matrices)

Sparse matrix exponentiation: EXPOKIT's DMEXPV

(DMEXPV's imprecision affected ML analysis, needs more exploration)

R Package: rexpokit

Use of multicore mapply() to farm out the matrix exponentiations on the branches

Dealing with slowness #2: conditional probabilities at cladogenesis

Number of possible cladogenetic range-inheritance events

states = $2^{\text{number of areas}}$

2 areas = 4 states = 4-1 = 3 non-null states

At every node,

3 possible ancestor states,

3 possible left descendants,

3 possible right descendants

= 3x3x3 = 27 combinations to assign probabilities to

Number of possible cladogenetic range-inheritance events

states = $2^{\text{number of areas}}$

9 areas = 512 states = 512-1 = 511 non-null states

At every node,

511 possible ancestor states,

511 possible left descendants,

511 possible right descendants

= 511x511x511 = 133,432,831 combinations to assign probabilities to

Problem can be simplified, according to your cladogenesis model

E.g., range AB giving rise to:

left = CDEFG, right = DEFGH

…should get probability 0.

But, searching intelligently involves lots of nested for-loops

R Package: cladoRcpp

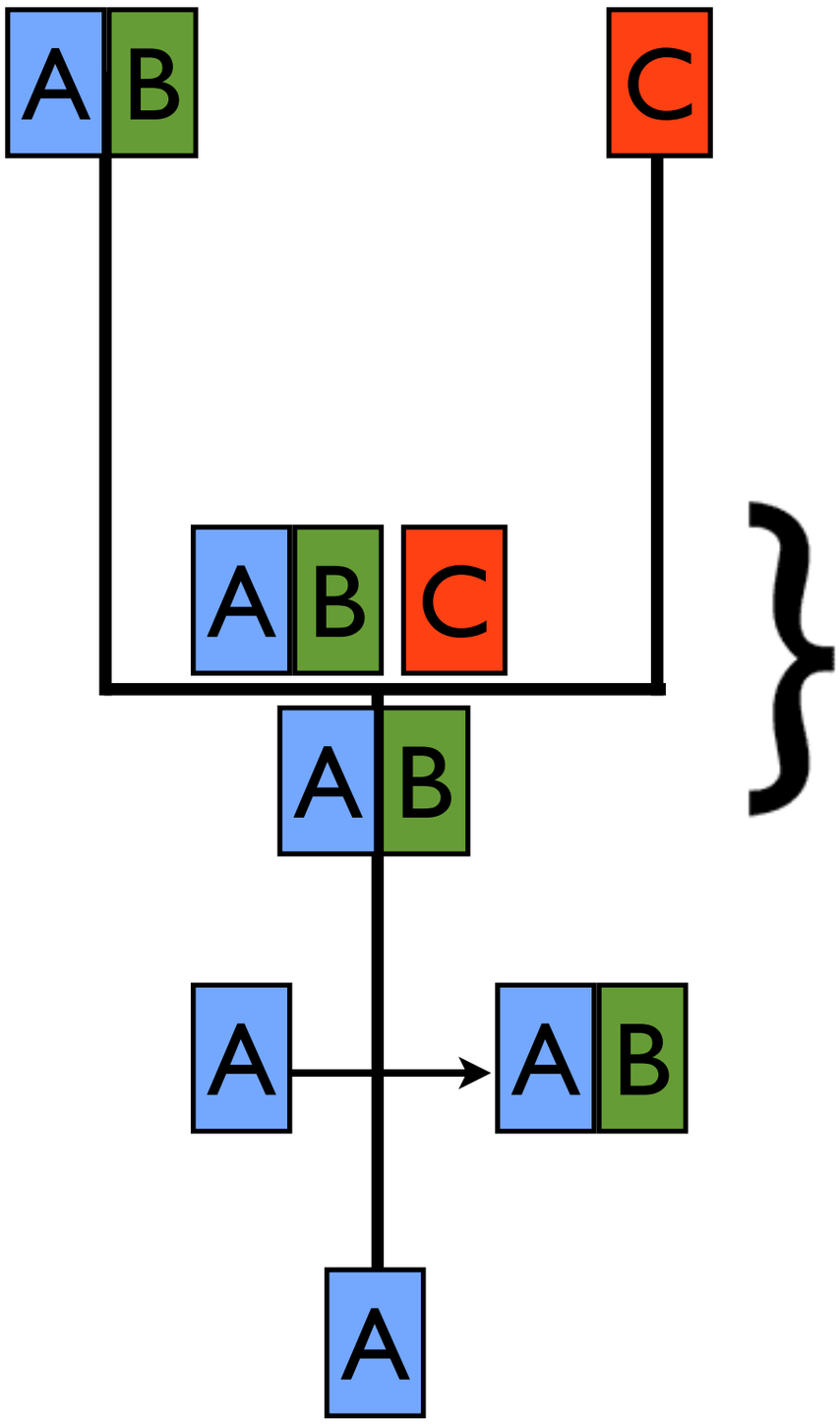
cladoRcpp: C++ implementations of phylogenetic cladogenesis calculations

This package implements in C++/Rcpp various cladogenesis-related calculations that are slow in pure R. These include the calculation of the probability of various scenarios for the inheritance of geographic range at the divergence events on a phylogenetic tree, and other calculations necessary for models which are not continuous-time markov chains (CTMC), but where change instead occurs instantaneously at speciation events. Typically these models must assess the probability of every possible combination of (ancestor state, left descendent state, right descendent state). This means that there are up to (# of states)^3 combinations to investigate, and in biogeographical models, there can easily be hundreds of states, so calculation time becomes an issue. C++ implementation plus clever tricks (many combinations can be eliminated a priori) can greatly speed the computation time over naive R implementations. CITATION INFO: This package is the result of my Ph.D. research, please cite the package if you use it! Type: citation(package="cladoRcpp") to get the citation information.

Version:	0.14.4
Imports:	Rcpp
LinkingTo:	Rcpp, RcppArmadillo
Published:	2014-05-18
Author:	Nicholas J. Matzke [aut, cre, cph]
Maintainer:	Nicholas J. Matzke <matzke at berkeley.edu>
License:	GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL:	http://phylo.wikidot.com/biogeobears
NeedsCompilation:	yes
Materials:	NEWS
CRAN checks:	cladoRcpp results

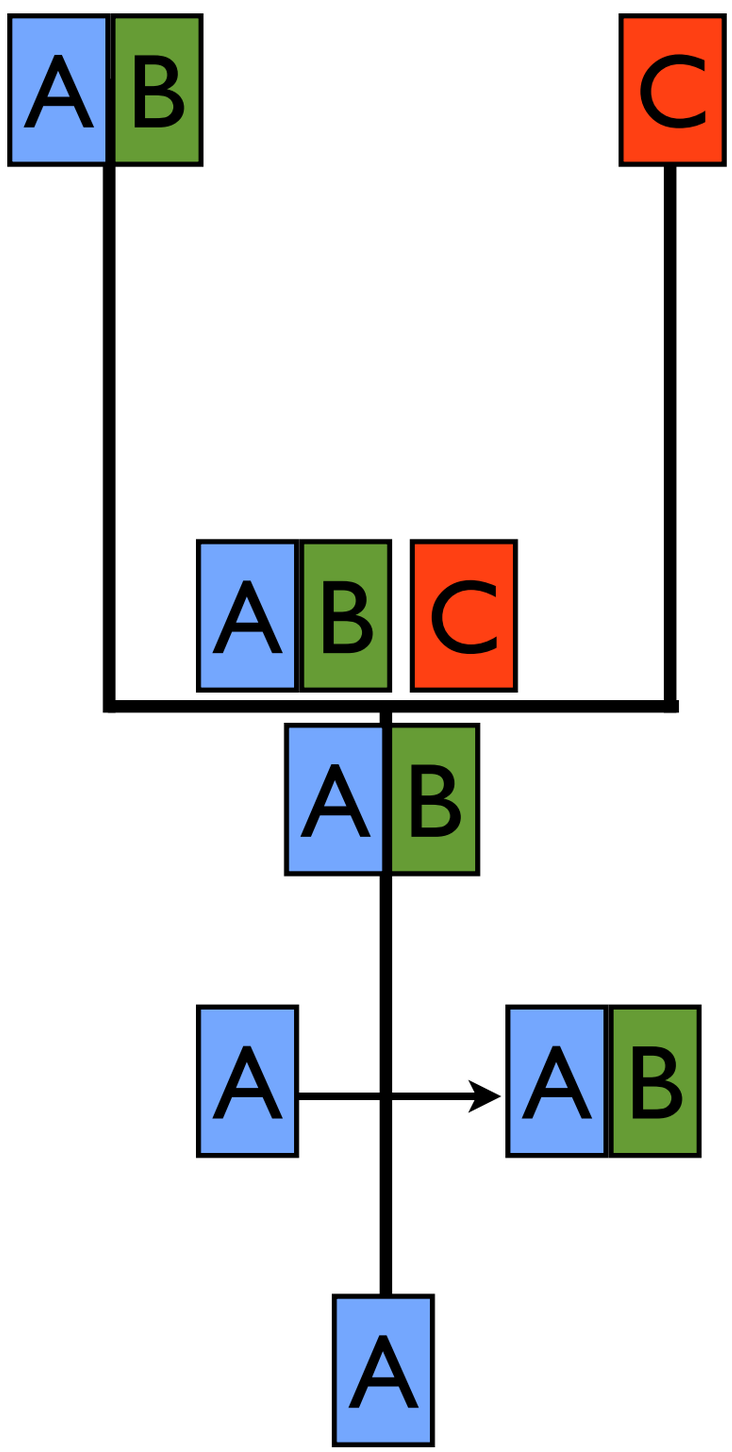
Downloads:

Reference manual:	cladoRcpp.pdf
Package source:	cladoRcpp_0.14.4.tar.gz
Windows binaries:	r-devel: cladoRcpp_0.14.4.zip, r-release: cladoRcpp_0.14.4.zip, r-oldrel:



Package #2, **cladoRcpp**, handles cladogenesis events (C++, Rcpp)

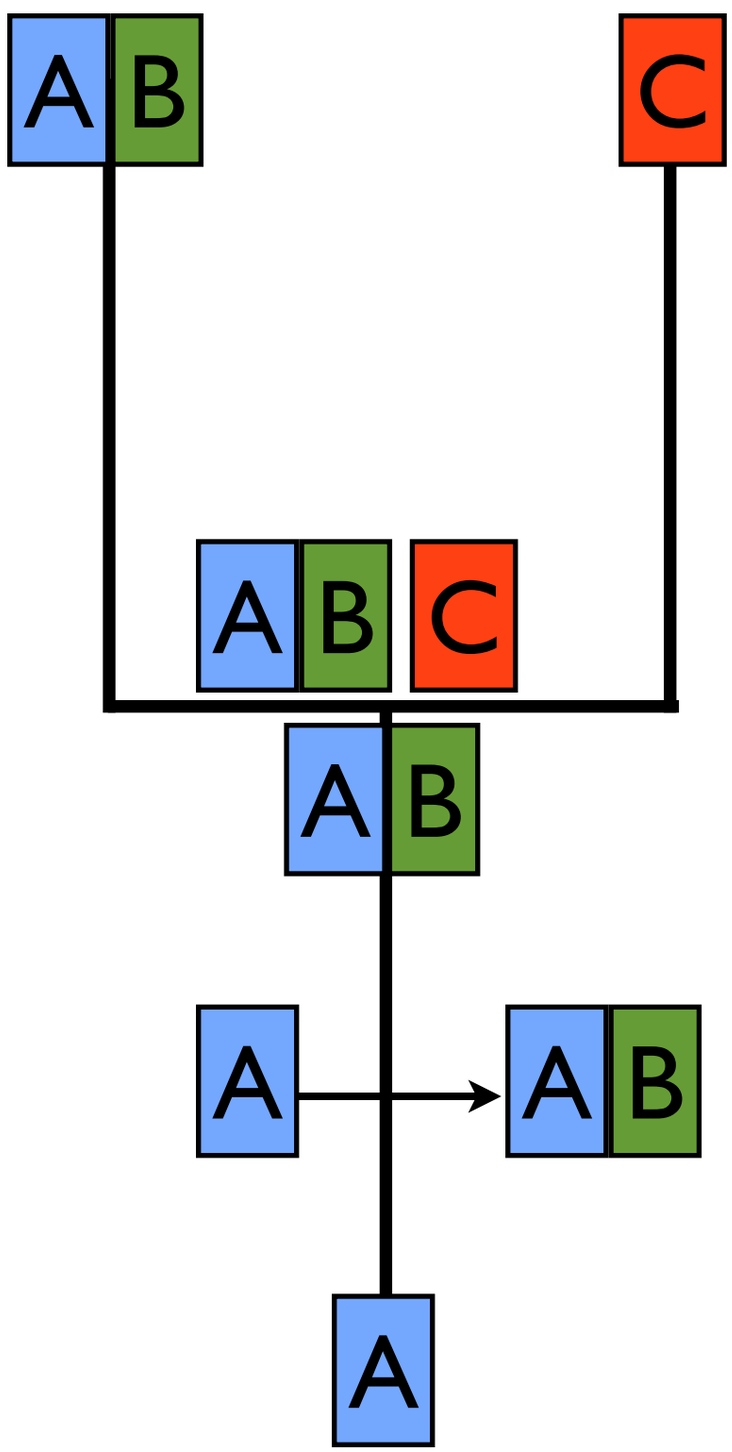
BioGeoBEARS speed



Small analysis (4 areas, 16 states)

C++ Lagrange: < 1 second
Python Lagrange: a few sec.
BioGeoBEARS: ~10 seconds

BioGeoBEARS speed



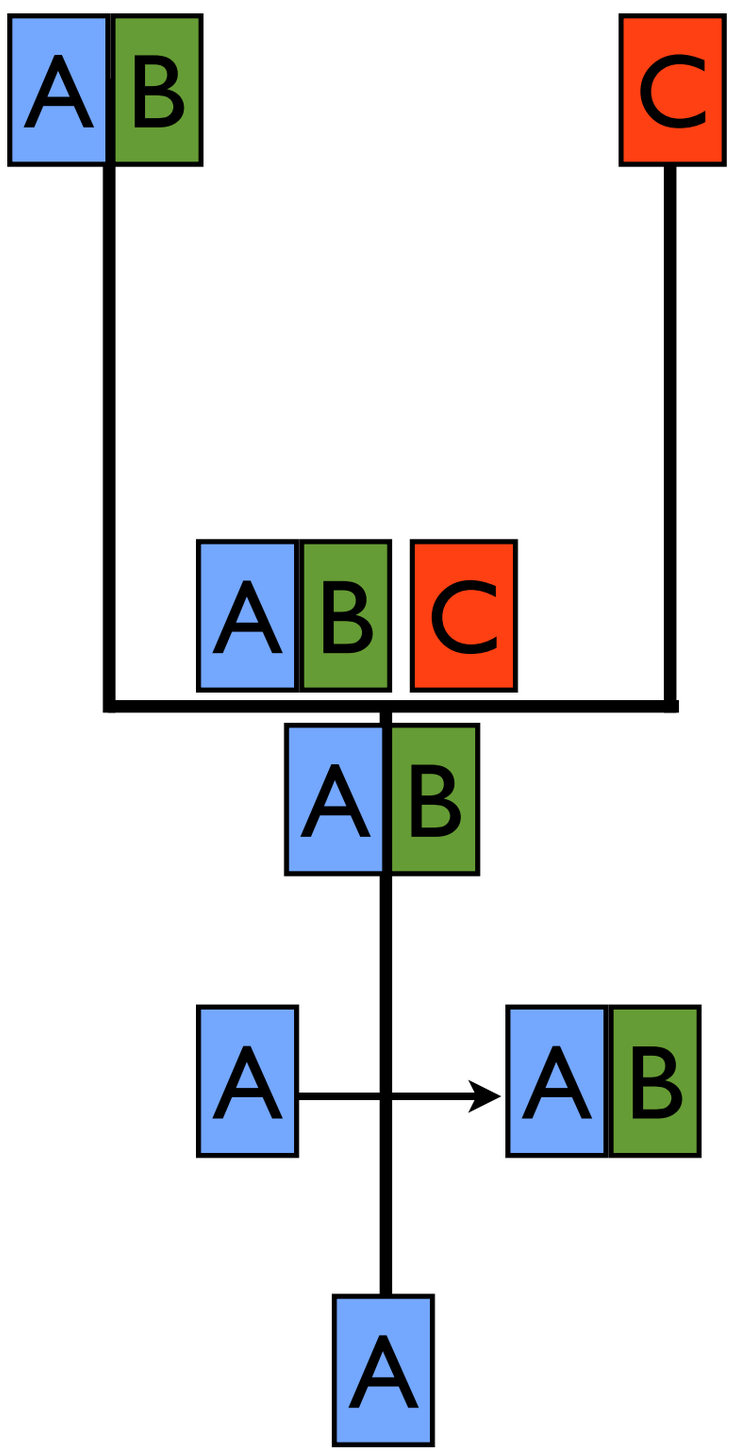
Large analysis (10 areas, 6 max, 848 states)

C++ Lagrange: 208 min.
Python Lagrange: won't run

BioGeoBEARS single-core: 187 minutes

BioGeoBEARS twelve cores: 22 minutes

rexpokit improvements



Coauthor Drew Schmidt, University of Tennessee

Gets a 40% speed improvement via various fixes in the R/FORTRAN interaction

New version of rexpokit going on CRAN this week

See his talk at UseR conference, June 30-July 3, UCLA

R Package: at PhyloWiki

NickMatzke | My account ▼ Search this site Search

PhyloWiki

Assisting research and education in phylogenetics and evolution

example menu | contact

Welcome page
What is a Wiki Site?
How to edit pages?
How to join this site?
Site members
Recent changes
List all pages
Page Tags
Site Manager

Page tags

start
free icons

Add a new page

new page

Cheap soma generic

edit this panel

BioGeoBears

UPDATE 6/23/2013: I have just uploaded my Evolution 2013 talk, which was Sunday, 11:45 am at the Ernst Mayr symposium. The PDF can be found at the bottom of the file. I have also uploaded a zipped source package for BioGeoBEARS. This is more up-to-date than github. It is the same as what I just submitted to CRAN as 0.1.1.

See also by iEvoBio Lightning Talk, 1:30 pm Tuesday I think.

The goal of my R package "BioGeoBEARS" — short for "BioGeography with Bayesian Evolutionary Analysis in R Scripts", if you were wondering — is to infer the history of species' geographic ranges on phylogenetic trees. It implements the LAGRANGE model (2 free parameters) as well as models with fewer or more free parameters. Standard model-testing procedures may then be applied.

The default implementation is actually ML (maximum likelihood), for comparison with LAGRANGE, but there is a Bayesian option making use of LaplacesDemon — contact me at matzke@berkeley.edu, if you can't figure it out from the functions.

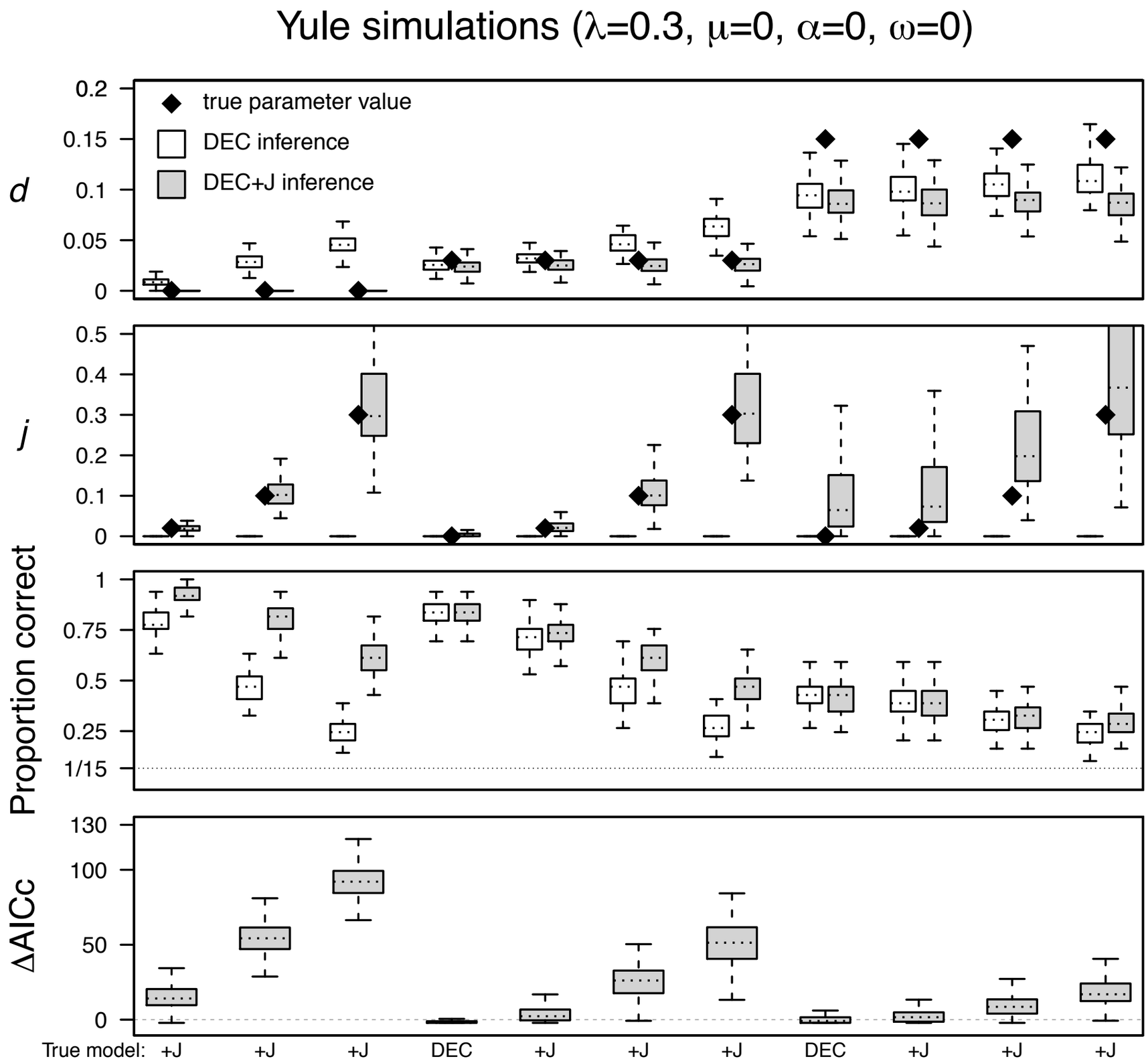
See also my IBS poster.

Author: Nicholas J. Matzke

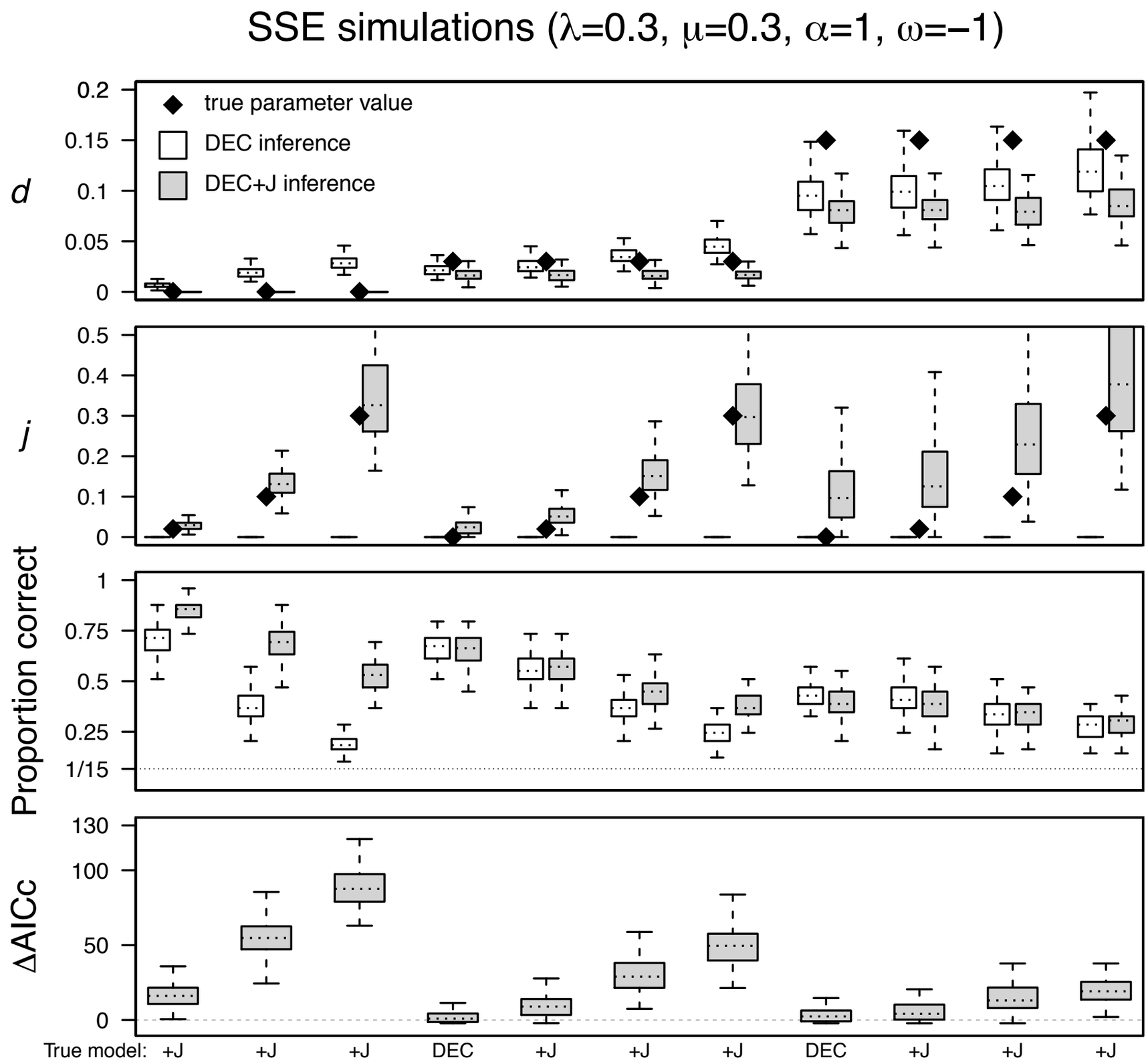
Download BioGeoBEARS

A draft version of BioGeoBEARS is available for download on github at: https://github.com/nmatzke/BioGeoBEARS. It is easier to download it from within R (see the install_github command, below).

Accurate when simulated under Yule process



Accurate even with simulation under SSE



Papers using BioGeoBEARS

Journal of Biogeography (J. Biogeogr.) (2014)

ORIGINAL ARTICLE

Incubator birds: biogeographical origins and evolution of underground nesting in megapodes (Galliformes: Megapodiidae)

Rebecca B. Harris[1,2]*, Sharon M. Birks[2] and Adam D. Leaché[1,2]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA, [2]Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA

ABSTRACT

Aim Unique amongst birds, megapodes (family Megapodiidae) have exchanged the strategy of incubating eggs with the warmth of their bodies for incubation behaviours that rely entirely on environmental heat sources. Typically, mound-builders capture heat released from the decomposition of organic materials, while burrow-nesters lay their eggs in geothermal or solar-heated soils. The evolutionary path towards novel incubation behaviours has led to ecological and physiological adaptations unique to megapodes. Here, we present a species tree for all extant megapodes that settles long-standing debates about megapode evolution: namely, their biogeographical origins and ancestral nesting behaviour.

Location Australasia.

Methods A time-calibrated multilocus species tree for all extant megapodes was constructed using *BEAST. We estimated and compared divergence dates for megapodes obtained from molecular rates, fossils, and a combination of fossils and rates. Using this tree, Bayesian estimation of ancestral nesting behaviour was conducted in BayesTraits and ancestral ranges were estimated in BioGeoBEARS.

Zoologica Scripta

KUNGL. VETENSKAPS-AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Molecular phylogenetics of mouse opossums: new findings on the phylogeny of *Thylamys* (Didelphimorphia, Didelphidae)

R. Eduardo Palma, Dusan Boric-Bargetto, J. Pablo Jayat, David A. Flores, Horacio Zeballos, Víctor Pacheco, Ricardo A. Cancino, Fernando D. Alfaro, Enrique Rodríguez-Serrano & Ulyses F. J. Pardiñas

Submitted: 7 June 2013
Accepted: 5 February 2014
doi:10.1111/zsc.12051

Palma, R.E., Boric-Bargetto, D., Jayat, J.P., Flores, D.A., Zeballos, H., Pacheco, V., Cancino, R.A., Alfaro, F.D., Rodríguez-Serrano, E. & Pardiñas, U.F.J. (2014). Molecular phylogenetics of mouse opossums: new findings on the phylogeny of *Thylamys* (Didelphimorphia, Didelphidae) — *Zoologica Scripta*, *43*, 217–234.

The mouse opossums of the genus *Thylamys* constitute a group of species mainly adapted to open xeric-like habitats and restricted to the southern portion of South America. We used molecular data (mitochondrial and nuclear sequences) to evaluate the phylogenetic and biogeographical relationships of all currently known living species of the genus, recognizing a new taxon from the middle and high elevations of the Peruvian Andes and evaluating the phylogenetic structuring within *T. pallidior* and *T. elegans*, as well as the validity of *T. sponsorius*, *T. cinderella* and *T. tatei*, and the haplogroups recognized within *T. pusillus*. Our results confirm the monophyly of the genus and that the Caatinga and the Cerrado inhabitants *Thylamys karimii* and *T. velutinus* are the most basal species in the radiation of *Thylamys*. We also calibrated a molecular clock which hypothesized a time of origin of the genus of about 24 My, with most species differentiating in middle and late Miocene and Plio-Pleistocene times of South America.

Molecular Phylogenetics and Evolution 73 (2014) 97–105

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

ELSEVIER

Diversification in an Afro-Asian songbird clade (*Erythropygia–Copsychus*) reveals founder-event speciation via trans-oceanic dispersals and a southern to northern colonization pattern in Africa

Gary Voelker[a,*], Joshua V. Peñalba[b], Jerry W. Huntley[a], Rauri C.K. Bowie[b,c]

[a] Department of Wildlife and Fisheries Sciences and Biodiversity Research and Teaching Collections, Texas A&M University, College Station, TX 77843, USA
[b] Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA
[c] Percy FitzPatrick Institute, DST/NRF Centre of Excellence, University of Cape Town, Rondebosch 7701, South Africa

ARTICLE INFO

Article history:
Received 2 December 2013
Revised 22 January 2014
Accepted 25 January 2014
Available online 4 February 2014

Keywords:
Africa
Asia
Erythropygia
Historical biogeography
LaGrange
Scrub-robins
Systematics

ABSTRACT

Erythropygia scrub-robins and their allies are distributed throughout Africa, Europe, Southeast Asia, India, Madagascar and the Seychelles. This broad distribution, as well as the distribution of *Erythropygia* taxa across Africa, presents an interesting opportunity to explore the mechanisms by which this biogeographic distribution was achieved. Multilocus sequence data (3310 base pairs from two mitochondrial and two nuclear genes) were generated for all species of *Erythropygia* and *Cercotrichas* scrub-robins, as well as from genera previously shown to render *Erythropygia* paraphyletic. Using model-based phylogenetic methods and molecular clock dating, we constructed a time-calibrated molecular phylogenetic hypothesis for the lineage. Ancestral area reconstructions were performed on the phylogeny using probabilistic approaches implemented in LaGrange and BioGeoBEARS. Our results confirm that *Erythropygia* is not monophyletic, and that one of the two *Erythropygia* clades is more closely related to a clade of Asian and Indian Ocean islands distributed species. Overall, the *Erythropygia* and allies clade originated in Africa in the late Miocene c. 6.9 Ma. Subsequently, a number of overwater dispersals occurred to include an initial colonization of Southeast Asia, and an ensuing progression of colonizations from Southeast Asia to the Seychelles, from there to Madagascar, and from these Indian Ocean islands back to Southeast Asia. Within the two clades of *Erythropygia*, ancestral area reconstructions within Africa indicate a Southern Africa origin, with subsequent lineage divergence in each clade indicating northward colonization. Overall, this clade of non-migratory songbirds shows a remarkable number of trans-oceanic colonization events, that were possibly facilitated by wind-driven dispersal; repeated Africa to Asia colonizations, two of which occur in this clade, are exceptionally rare in birds. Also rare is our finding that colonization patterns in Africa indicate a southern to northern progression.

ARTICLE IN PRESS

Molecular Phylogenetics and Evolution xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

ELSEVIER

Biogeography of the marmosets and tamarins (Callitrichidae)

Janet C. Buckner[a,*], Jessica Lynch Alfaro[b,c], Anthony B. Rylands[d], Michael E. Alfaro[a]

[a] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA
[b] Institute for Society and Genetics, University of California, Los Angeles, CA 90095, USA
[c] Department of Anthropology, University of California, Los Angeles, CA 90095, USA
[d] Conservation International, Arlington, VA 22202, USA

ARTICLE INFO

Article history:
Received 1 October 2013
Revised 9 April 2014
Accepted 15 April 2014
Available online xxxx

Keywords:
Biogeography
Phylogenetics
Neotropics
Platyrrhini
Leontocebus
Saguinus

ABSTRACT

The marmosets and tamarins, Family Callitrichidae, are Neotropical primates with over 60 species and subspecies that inhabit much of South America. Although callitrichids exhibit a remarkable widespread distribution, attempts to unravel their biogeographic history have been limited by taxonomic confusion and the lack of an appropriate statistical biogeographic framework. Here, we construct a time-calibrated multi-locus phylogeny from GenBank data and the callitrichid literature for 38 taxa. We use this framework to conduct statistical biogeographic analyses of callitrichids using BioGeoBEARS. The DIVAj model is the best supported reconstruction of biogeographic history among our analyses and suggests that the most recent common ancestor to the callitrichids was widespread across forested regions c. 14 Ma. There is also support for multiple colonizations of the Atlantic forest region from the Amazon basin, first by *Leontopithecus* c. 11 Ma and later by *Callithrix* c. 5 Ma. Our results show support for a 9 million year old split between a small-bodied group and large-bodied group of tamarins. These phylogenetic data, in concert with the consistent difference in body size between the two groups and geographical patterns (small-bodied tamarins and large-bodied tamarins have an unusually high degree of geographic overlap for congeners) lend support to our suggestion to split *Saguinus* into two genera, and we propose the use of distinct generic names; *Leontocebus* and *Saguinus*, respectively.

See also:

Model Selection in Historical Biogeography Reveals that Founder-event Speciation is a Crucial Process in Island Clades

Journal:	*Systematic Biology*
Manuscript ID:	Draft
Manuscript Type:	Regular Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Matzke, Nicholas; University of California, Berkeley, Integrative Biology
Keywords:	historical biogeography, cladogenesis, founder-event speciation, jump dispersal, LAGRANGE, BioGeoBEARS, extinction, GeoSSE

SCHOLARONE™
Manuscripts

In review: email me at matzke@nimbios.org for this

frontiers of biogeography
the scientific magazine of the International Biogeography Society

Download PDF | Show Abstract | Email | Share | Cite | + Save | 1

resources ISSN 1948-6596

thesis abstract

Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing

Nicholas Joseph Matzke

PhD Thesis, Department of Integrative Biology, University of California, Berkeley. 3060 Valley Life Sciences Building, Berkeley, CA, 94720; matzke@berkeley.edu.
Current address: National Institute for Mathematical and Biological Synthesis (NIMBioS, www.nimbios.org). 1122 Volunteer Blvd., Suite 106, University of Tennessee, Knoxville, TN 37996-3410; matzke@nimbios.org.

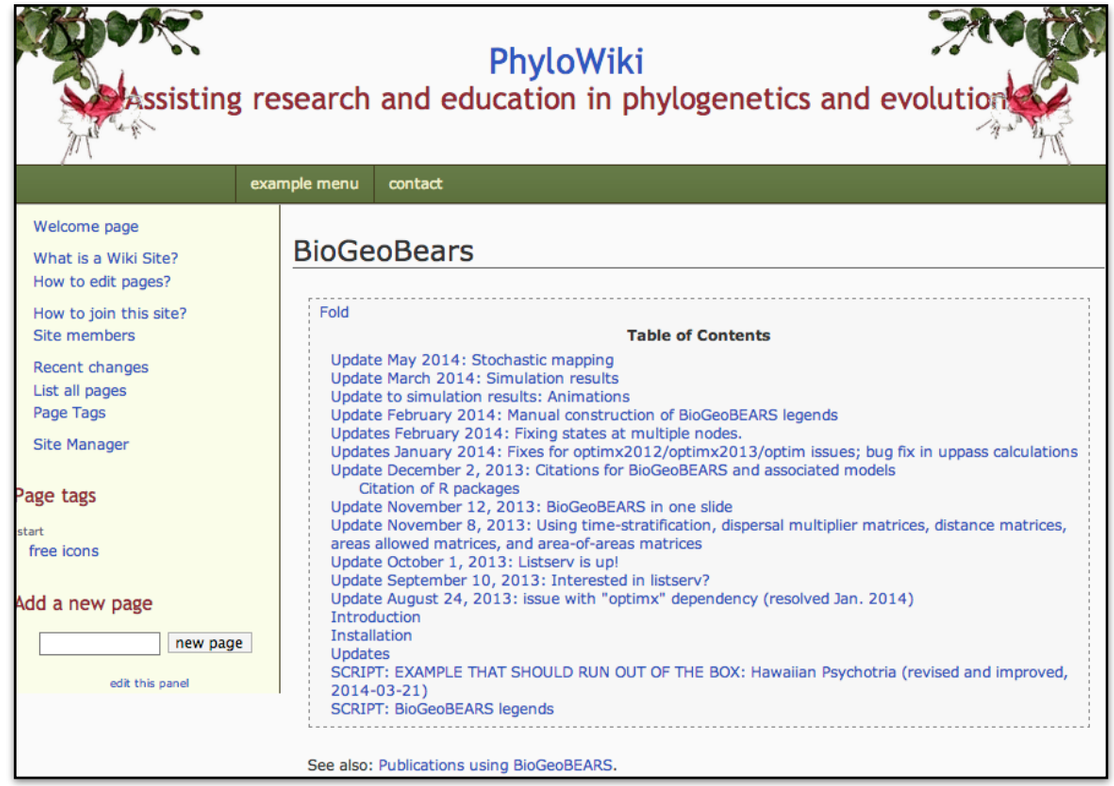
Abstract. Historical biogeography has been characterized by a large diversity of methods and unresolved debates about which processes, such as dispersal or vicariance, are most important for explaining distributions. A new R package, BioGeoBEARS, implements many models in a common likelihood framework, so that standard statistical model selection procedures can be applied to let the data choose the best model. Available models include a likelihood version of DIVA ("DIVALIKE"), LAGRANGE's DEC model, and BAYAREA, as well as "+J" versions of these models which include founder-event speciation, an important process left out of most inference methods. I use BioGeoBEARS on a large sample of island and non-island clades (including two fossil clades) to show that founder-event speciation is a crucial process in almost every clade, and that most published datasets reject the non-J models currently in widespread use. BioGeoBEARS is open-source and freely available for installation at the Comprehensive R Archive Network at http://CRAN.R-project.org/package=BioGeoBEARS. A step-by-step tutorial is available at http://phylo.wikidot.com/biogeobears.

Keywords. cladogenesis, DIVA, founder-event speciation, historical biogeography, jump dispersal, LAGRANGE, phylogenetics

Thanks! I welcome questions/comments/ collaborations at:

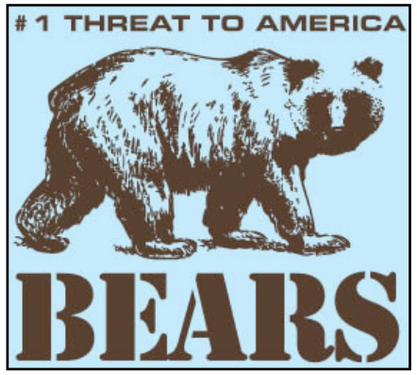
matzke@nimbios.org

(also: seeking a job!)



Thanks especially to:

NIMBioS
Brian O'Meara
Jeremy Beaulieu
Katie Massana
Michael Landis

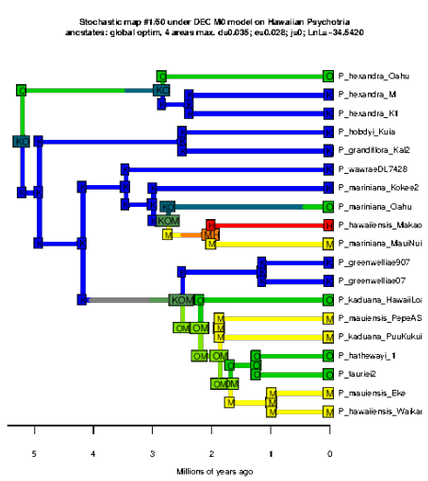


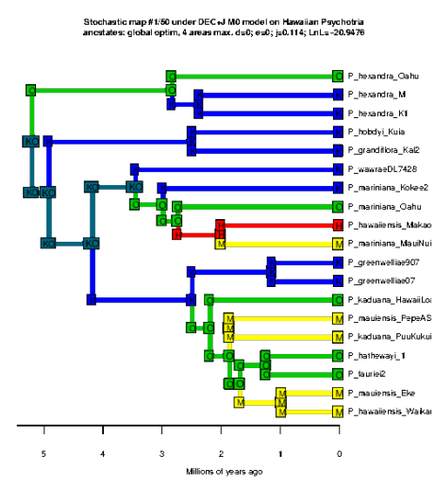
TRY IT YOURSELF AT:

http://phylo.wikidot.com/biogeobears

Ph.D. committee
John Huelsenbeck
Tony Barnosky
David Jablonski
Roger Byrne

Systematic Biology editors & reviewers





Funding: NIMBioS
NSF "Bivalves in Time and Space"
UC Berkeley Wang Fellowship
UC Berkeley Tien Fellowship
Google Summer of Code
NIMBioS

NIMBioS National Institute for Mathematical and Biological Synthesis

EE Ecology & Evolutionary Biology The University of Tennessee, Knoxville

The University of Tennessee Knoxville

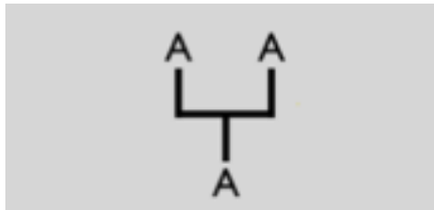
Bivalves in Time & Space — Bivalves in Time and Space: Testing the accuracy of methods to reconstruct ancestral morphology, dates, geography, and diversification patterns

Center for Theoretical Evolutionary Genomics — Bachtrog • Eisen • Hallatschek • Huelsenbeck • Nielsen • Slatkin • Song

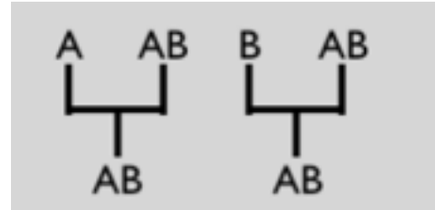
DEC: Dispersal-extinction cladogenesis

Cladogenesis -- range evolution at speciation events **(LAGRANGE model)**

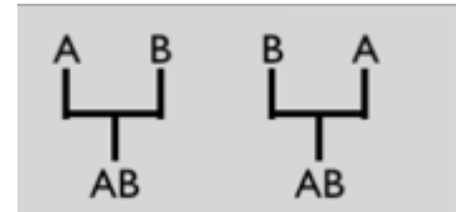
Probabilites per speciation event:



sympatric



subset

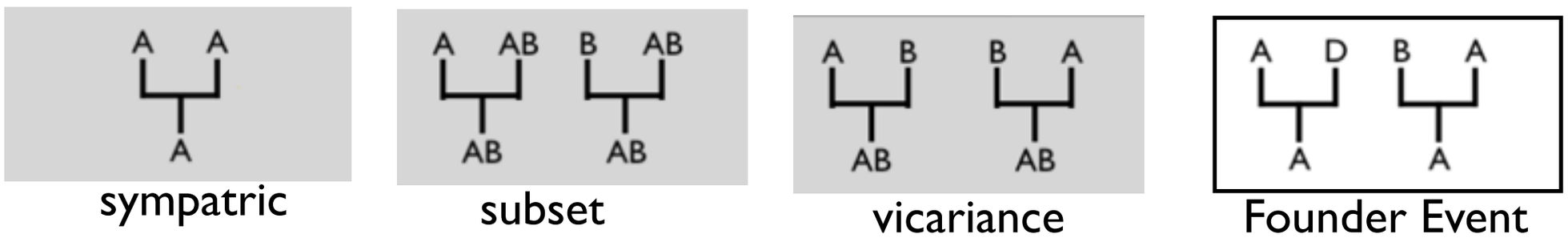


vicariance

		ending range																																																			
	left:	A	A	A	A	A	A	A	B	B	B	B	B	B	B	C	C	C	C	C	C	C	AB	AB	AB	AB	AB	AB	AB	BC	BC	BC	BC	BC	BC	BC	AC	AC	AC	AC	AC	AC	AC	ABC	ABC	ABC	ABC	ABC	ABC	ABC	
	right:	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	count
	A	sym																																																			1
	B									sym																																											1
starting	C																	sym																																			1
range	AB		vic		sub				vic			sub											sub	sub																													6
	BC										vic		sub				vic			sub												sub	sub																				6
	AC			vic			sub									vic					sub																sub		sub														6
	ABC					vic		sub						vic	sub				vic			sub			vic					vic								vic						sub	sub	sub					12

DEC+J: Dispersal-extinction cladogenesis PLUS founder-event speciation

Probabilites per speciation event:

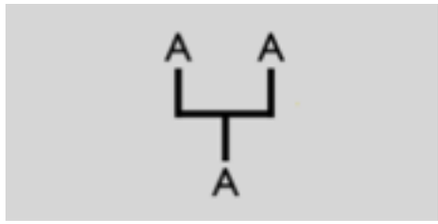


		ending range																																																					
	left:	A	A	A	A	A	A	A	B	B	B	B	B	B	B	C	C	C	C	C	C	C	AB	AB	AB	AB	AB	AB	AB	BC	BC	BC	BC	BC	BC	BC	AC	AC	AC	AC	AC	AC	AC	ABC	ABC	ABC	ABC	ABC	ABC	ABC	
	right:	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	count
	A	sym	**j**	**j**					**j**							**j**																																					1
	B		**j**						**j**	sym	**j**						**j**																																			1
starting	**C**			**j**							**j**					**j**	**j**	sym																																		1
range	**AB**		vic		sub				vic			sub							**j**				sub	sub	**j**																										6
	BC					**j**					vic		sub				vic			sub									**j**	sub	sub																				6
	AC			vic			sub							**j**		vic					sub															sub	**j**	sub													6
	ABC					vic		sub						vic	sub				vic			sub			vic					vic							vic						sub	sub	sub						12

BayArea cladogenesis model

Cladogenesis -- "no event" during speciation (i.e., pure continuous time)

Probabilites per speciation event:



Sympatric speciation (range duplication)

		ending range																																																
	left:	A	A	A	A	A	A	A	B	B	B	B	B	B	B	C	C	C	C	C	C	C	AB	AB	AB	AB	AB	AB	AB	BC	BC	BC	BC	BC	BC	BC	AC	AC	AC	AC	AC	AC	AC	ABC	ABC	ABC	ABC	ABC	ABC	ABC	
	right:	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	A	B	C	AB	BC	AC	ABC	count
	A	sym																																																1
	B									sym																																								1
starting	C																	sym																																	1
range	AB																									sym																									1
	BC																																	sym																	1
	AC																																									sym									1
	ABC																																																	sym	1