

SSB 2015 Model-Based Biogeography

Practicum: “BioGeoBEARS”

Jeet Sukumaran

Ecology and Evolutionary Biology
University of Michigan
Ann Arbor, MI

jeetsukumaran@gmail.com

May 20, 2015

Preliminaries and General Best Practices

Background Housekeeping Concepts: A *GOOD* Text Editor

- A *good* general-purpose text-editor is an important tool to know and to know well.
- *Not* general-purpose text editors:
 - Microsoft Office/Word, LibreOffice, OpenOffice, Pages, etc.
 - TextEdit (Macs)
 - NotePad (Windows)
 - RStudio
 - PAUP*
- Good text editors:
 - Sublime Text
 - TextWrangler
 - NotePad++
 - Vim

Background Housekeeping Concepts: Know Your Data File Formats!

- Learn to recognize the various file formats data comes in (Newick, NEXUS, FASTA, etc.) by visual inspection.
- Understand the file formats required by various programs.
- Understand that many of these formats do not have a single standard.
- Different programs may or may not strictly follow any one convention for each format.
- Many programs add or support their own idiosyncratic variation of formats.
- Worse, many files found “in the wild” may or may not conform to any standard convention.

Background Housekeeping Concepts: Know Your Filename Extensions!

- Filename extensions are how your operating system recognizes types of files.
- Not always handled correctly by the operating system.
- As with file formats, *you* need to inform yourself to be the final arbitrator.
- Set your operating system *NOT* to hide file extensions.
- Set your operating system *NOT* to hide any files.

Background Housekeeping Concepts: Best Practices for Naming Files and Directories

- Stick to the 26 character standard Roman (unaccented) alphabet ("A-Za-z"), numbers ("0-9"), underscores ("_"), periods (".") and/or dashes (" - ").
- Use casing as you wish, but do not count on case differences to distinguish names.
- Do *NOT* use spaces ANYWHERE in the file/directory name.
- Do *NOT* use ANY other punctuation apart from the underscore, period or dash.
- Do *NOT* start a file or directory name with an underscore, period, or dash.
- Meaningful file/directory names are *good* (even if long).
- Meaningful filename extensions are *good*.
- Primary criterion: clear indication of contents *and* easy to type.

Background Housekeeping Concepts: Best Practices for Naming Taxa

- Stick to the 26 character standard Roman (unaccented) alphabet (A-Z, a-z) or periods (".").
- If you use the underscore ("_"), be aware that the standard Nexus/Newick format treats these as equivalent to spaces.
- Unfortunately, not all programs respect this convention, and, of course, not all formats (such as FASTA) follow this as well.
- I highly recommend avoiding spaces AND underscores in your taxon labels for this reason.
- Do not count on case difference to disambiguate your names.

Background Housekeeping Concepts: File System Navigation Concepts

- Relative vs. absolute directory paths
- Current working directory
- Navigating directories

Background Housekeeping Concepts: File System Navigation Concepts

- For the following exercises, we assume that your current working directory is the “top” BioGeoBears lab directory.
- All paths specified will be specified *relative* to the current working directory, so make sure you have this set correctly.

Getting Started

Download the Lab Files

First, go to the parent directory under which you wish to work, creating it first if needed. E.g.:

```
$ mkdir ssb2015-workshop  
$ cd ssb2015-workshop
```

Using Git:

```
$ git clone http://github.com/jeetsukumaran/2015-SSB-AnnArbor-Biogeography
```

Alternatively, download and unpack from:

```
https://github.com/jeetsukumaran/2015-SSB-AnnArbor-Biogeography/releases
```

Navigate to the Project Directory

- Go to the project directory:

```
$ cd 2015-SSB-AnnArbor-Biogeography/biogeobears
```

- Confirm that you are in the correct directory

```
$ pwd
```

```
/Users/jeet/Documents/Courses/Biogeography/2015-SSB-AnnArbor-Biogeography/
```

```
$ ls -l
```

```
total 0
```

```
drwxr-xr-x  2 jeet  staff   68B May 15 09:59 data
```

```
drwxr-xr-x 27 jeet  staff  918B May 15 11:25 libexec
```

```
drwxr-xr-x  4 jeet  staff  136B May 15 17:40 scripts
```

Data Required for an Analysis: Phylogeny

- The phylogeny should be **ultrametric** (with branch lengths proportional to time).
- The phylogeny is taken as a the *true* phylogeny:
 - Uncertainty in the phylogeny is not automatically treated
 - Multiple analyses, with an independent analysis carried out on each MCMC sample, and results averaged over all analyses can be used to treat uncertainty in phylogenies.
- The phylogeny is specified in Newick format.
- NO spaces allowed in the taxon names!
- Taxon names must match the names in the phylogeny exactly.
- Example file: "data/Psychotria_5.2.newick".

Data Required for an Analysis: Geography

A specification of the localities associated with each tip (in “pseudo-PHYLIP” format):

- First row: number of taxa and number of areas (with optional specification of area names in parenthesis)
- Subsequent rows: taxon name followed by sequence of “0’s” (for absence in an area) or “1’s” (for presence in area).
- NO spaces allowed in the taxon names!
- Taxon names must match the names in the phylogeny exactly.
- Example file: “data/Psychotria_geog.data”.

NOTE: while area names are optional, I highly recommend that (a) you use them, and (b) you stick to single-character labels.

Start R

```
$ R
```

```
R version 3.0.3 (2014-03-06) -- "Warm Puppy"  
Copyright (C) 2014 The R Foundation for Statistical Computing  
Platform: x86_64-apple-darwin10.8.0 (64-bit)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.
```

```
  Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.
```

```
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.
```

```
>
```

Install “BioGeoBEARS”

“BioGeoBEARS” can be installed using the “install.packages()” command inside R:

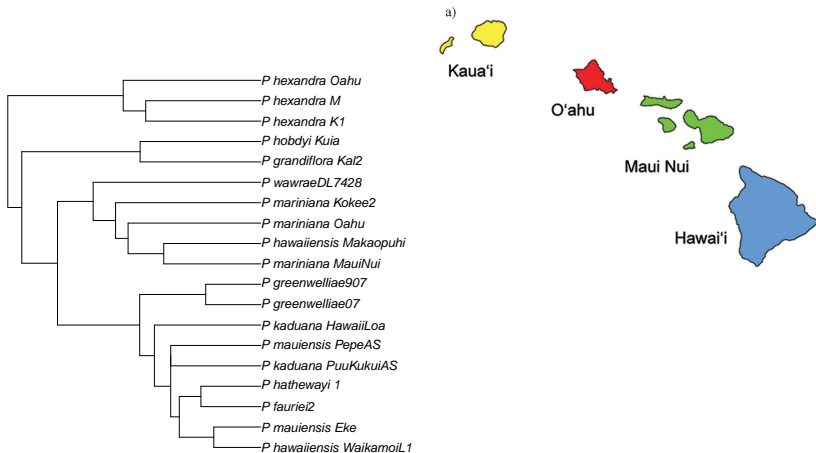
```
> install.packages("optimx", dependencies=TRUE)  
> install.packages("BioGeoBEARS", dependencies=TRUE)
```

This will be executed if you source the file “scripts/00-01-install.R” from inside R:

```
> source("scripts/00-01-install.R")
```


The Dataset

The Dataset



Exercise I: A Basic “BioGeoBEARS” Run

“BioGeoBEARS”: BioGeography with Bayesian (and likelihood) Evolutionary Analysis in R Scripts.

- “BioGeoBEARS” is an R package by Nick Matzke that implements the DEC model with various elaborations and extensions.
- Full information (with lots of examples and exercises) can be found at: <http://phylo.wikidot.com/biogeobears>.
- All software used in a project *must* be cited, not only to give academic credit, but also as part of the full documentation of your workflow. You would be very disappointed with a study that left out the markers/genes sampled or collecting localities; you should be just as disappointed with any work that does not properly credit the software used. To get the citation to use, you can type:

```
> citation("BioGeoBEARS")
```
- **NOTE:** Code presented in these exercises or bundled with this lab are based on scripts by Nick Matzke, and used as permitted under the GPL3 license.

Loading the “BioGeoBEARS” Library Package

In R, you normally load a package (that has already been installed) using the “`library()`” function. However, “BioGeoBEARS” requires some special start-up patches and modifications. These have been packaged for you in the function “`load.biogeobears()`”, in the script “`scripts/biogeobears-utilities.R`”:

```
> source("scripts/biogeobears-utilities.R")  
> load.biogeobears()
```

Recipe for a “BioGeoBEARS” Analysis

- 1 Create a BioGeoBEARS run object.
- 2 Configure the object: data, model, parameters, run settings etc.
- 3 Check the object configuration and settings.
- 4 Run the analysis.
- 5 Save and inspect the results.

Create the “BioGeoBEARS” Run Object

We create a “BioGeoBEARS” run object called “dec.run”:

```
> dec.run = define_BioGeoBEARS_run()
```

Inspect the “BioGeoBEARS” Run Object

We create a “BioGeoBEARS” run object called “`dec.run`”:

```
> dec.run
```

- Lots and lots and *lots* of settings!
- You do not need (or want!) to the meanings of *all* of these at this stage.
- We will focus on just a few of immediate interest for now and work the others in later as needed.

Check the Phylogeny Data

Let us quickly check the phylogeny and make sure that it is in order. We could do this by opening it up in our text editor (which is what I would normally do) and visualizing it in, e.g., FigTree, but here we are going to do it from inside R:

```
> moref("data/Psychotria_5.2.newick")  
> psychotria.tree = read.tree("data/Psychotria_5.2.newick")  
> plot(psychotria.tree)  
> prt(psychotria.tree)
```

Setting the Phylogeny Data on the “BioGeoBEARS” Run Object

Specify the path to the tree file by setting the “trfn” property of the “BioGeoBEARS” run object:

```
> dec.run$trfn = "data/Psychotria_5.2.newick"
```

Check that it has been set:

```
> dec.run$trfn  
[1] "data/Psychotria_5.2.newick"
```

Check the Geography Data

Let us also inspect the geography data:

```
> moref("data/Psychotria_geog.data")  
> getranges_from_LagrangePHYLIP(lgdata_fn="data/Psychotria_geog.data")
```

Setting the Geography Data on the “BioGeoBEARS” Run Object

Specify the path to the geography file by setting the “geogfn” property of the “BioGeoBEARS” run object:

```
> dec.run$geogfn = "data/Psychotria_geog.data"
```

Check that it has been set:

```
> dec.run$geogfn  
[1] "data/Psychotria_geog.data"
```

Defining the Model: Fixing or Freeing Parameters

- We specify the model to use by freeing or fixing parameters.
- Each parameter has the following three attributes:
 - **type** (which can be “fixed”, “free”, or some other special value)
 - **init** (the initial value for the ML optimization)
 - **min** (the minimum value for the ML optimization)
 - **max** (the maximum value for the ML optimization)
 - **est** (the current value of the parameter used when calculating the likelihood)
- We add a parameter to the model by setting its “**type**” to “free”, and providing it sensible initial, minimum, and maximum values.
- We remove a parameter from the model by settings it “**type**” to “fixed” and assigning its “**est**” property to an appropriate value.

Defining the Model: Example of Parameter Setting

Adding parameter “foo” (**note:** dummy parameter for example purposes):

```
dec.run$BioGeoBEARS_model_object@params_table["foo","type"] ="free"  
dec.run$BioGeoBEARS_model_object@params_table["foo","init"] = 0.050  
dec.run$BioGeoBEARS_model_object@params_table["foo","min"] = 0.001  
dec.run$BioGeoBEARS_model_object@params_table["foo","max"] = 1.000
```

Removing parameter “foo” (**note:** dummy parameter for example purposes):

```
dec.run$BioGeoBEARS_model_object@params_table["foo","type"] ="fixed"  
dec.run$BioGeoBEARS_model_object@params_table["foo","init"] = 0.0  
dec.run$BioGeoBEARS_model_object@params_table["foo","est"] = 0.0
```

Defining the Model: Example of Parameter Setting

Some special parameters can be selectively constrained with reference to other parameters:

```
dec.run$BioGeoBEARS_model_object@params_table["ysv","type"] = "2-j"  
dec.run$BioGeoBEARS_model_object@params_table["ys","type"] = "ysv*1/2"  
dec.run$BioGeoBEARS_model_object@params_table["y","type"] = "ysv*1/2"  
dec.run$BioGeoBEARS_model_object@params_table["v","type"] = "ysv*1/2"
```

(For more details, see: <http://phylo.wikidot.com/biogeobears>.)

Defining the Model: Available Parameters

		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 (& $m \times 0 y$)
	Sympatry (widespread)						✓	y (& $m \times 0 y$)
	Sympatry (subset)					✓		s (& $m \times 0 s$)
	Vicariance (narrow)				✓	✓		v (& $m \times 0 v$)
	Vicariance (widespread)				✓			v (& $m \times 0 v$)
	Founder							j (& $x, m \times 0 j$)

Defining the Model: Available Parameters

- Full information on parameters, with examples on how to set various parameters to define different biogeographic models (e.g., DIVA, "DEC-like", "BayArea-like", can be found at:
<http://phylo.wikidot.com/biogeobears>.
- By selectively fixing/constraining or freeing parameters, a broad range of models can be achieved, that vary by:
 - What sorts of anagenetic range evolution processes are allowed
 - What sorts of cladogenetic range evolution processes are allowed
- Interesting to note that, while estimation/optimization paradigms are fundamentally different in, e.g., DIVA and DEC, from a modeling perspective they are very similar in principal, differing only in the details of supported speciation modes.

Default Model: DEC

Inspect the current parameter configuration of the model:

```
> dec.run$BioGeoBEARS_model_object@params_table
```

- Anagenetic range evolution:
 - “**d**” : “dispersal”, i.e. rate of anagenetic range expansion
 - “**e**” : “extinction”, i.e., rate of anagenetic range contraction
- Cladogenetic range evolution:
 - “**y**” : single-area sympatric speciation
 - “**s**” : sympatric subset speciation
 - “**v**” : vicariance, narrow/peripheral allopatry
- This (the default) corresponds to the original DEC model.

Setting the Run Configuration

There are numerous “meta-settings” that allow for the configuration of the ML estimation and calculation (e.g., number of cores to use, speed-ups, optimization engines, etc. etc.) We will not go through them in detail, but sensible defaults are available to you by calling the “`configure.standard.biogeobears.run()`” function, passing in the the BioGeoBEARS run object as an argument:

```
> dec.run = configure.standard.biogeobears.run(dec.run)
```

One important thing we set here are options to make sure we recover the estimated ancestral states:

```
> dec.run$return_condlikes_table = TRUE  
> dec.run$calc_TTL_loglike_from_condlikes_table = TRUE  
> dec.run$calc_ancprobs = TRUE
```

This gets the ancestral states from the optimization run for us to inspect later.

Practical Considerations: Computational Effort and Numbers of Areas or Range Size

- We have to integrate over all possible ancestral states for each node.
- Each state in our system is not just an area, but a *range*, i.e. a distinct combination of areas.
- So, for e.g., for three areas, we have eight states:
 $\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$.
- This number goes up exponentially (i.e., 2^n , where n is the number of areas):

Number of Areas	Number of States
3	8
5	32
8	256
12	4096
20	104576

Practical Considerations: Computational Effort and Numbers of Areas or Range Size

- How many areas is too many? It depends, but the largest analysis I have done successfully is with 8 areas (256 states).
- What can we do if the number of areas are too many?
 - Change the model by aggregating areas.
 - Restrict the number of ranges by setting limits on the maximum numbers of areas in each possible range.

```
dec.run$max_range_size = 4
```
- Above are not always suitable or possible or desired: *use another approach!*.

Confirm and Check Configuration

```
> check_BioGeoBEARS_run(dec.run)
> dec.run
```

OK! Time to Execute the Analysis!

```
> dec.results = bears_optim_run(dec.run)
> save(dec.results, file="results/Psychotria.DEC.Rdata")
```

And Look at the Results!

Ancestral range estimates:

```
> plot.biogeobears.results.ranges(  
+   results.object=dec.results,  
+   analysis.title="DEC"  
+   )
```


Inspect the Results in Detail

A table listing the likelihoods of each range state at each node is given in the results object by:

```
> dec.results$ML_marginal_prob_each_state_at_branch_top_AT_node
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	0	0.000000e+00	0.000000e+00	1.000000e+00	0.000000e+00	0.000000e+00
...
[20,]	0	2.786140e-01	1.941047e-03	2.064356e-05	6.847621e-06	5.165582e-01
...
...

	[,7]	[,8]	[,9]	[,10]	[,11]
[1,]	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
...
[20,]	3.061937e-02	1.723071e-02	9.717510e-04	4.245526e-04	1.233218e-0
...

- The rows are the indexes of the *nodes*.
- The columns are the indexes of the *range states*.
- So we have to identify the estimation of the ancestral states of an MRCA of a group, we have to:
 - 1 Look up the index of the node of interest
 - 2 Find that index in the above table
 - 3 Read off the likelihoods of each state index
 - 4 Look up the list of areas that correspond to each state index

Inspect the Results in Detail

To make things more convenient, we can aggregate the tree node information and the ancestral states estimation results into a single table:

```
> dec.range.table = get.biogeobears.results.by.range.table(dec.results)
```

and save the this table for a file to be viewed in a spreadsheet or text-editor:

```
> write.table(dec.range.table,  
+ "results/Psychotria.DEC.ranges.tsv",  
+ sep="\t",  
+ row.names=F)
```

Inspect the Results in Detail

We can also inspect the results with the probabilities marginalized by area instead of ranges:

```
> plot.biogeobears.results.areas(  
+   results.object=dec.results,  
+   analysis.title="DEC"  
+ )  
> dec.area.table = get.biogeobears.results.by.range.table(dec.results)  
> write.table(dec.area.table,  
+   "results/Psychotria.DEC.areas.tsv",  
+   sep="\t",  
+   row.names=F)
```

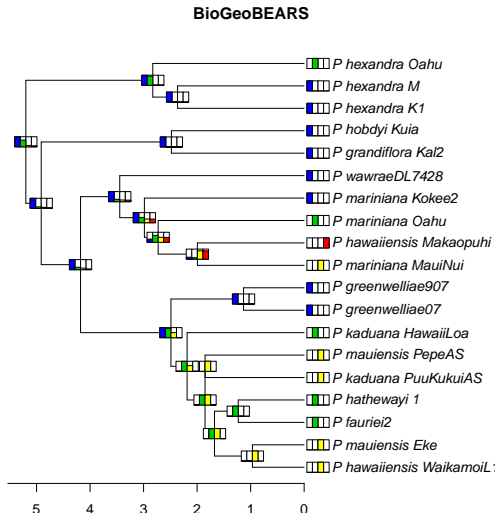
Inspect the Results in Detail

If we want to look up the range of an MRCA of a particular set of taxa, we first need to identify the index of the node, and then look it up:

```
> ndi = getMRCA(psychotria.tree, c("P_hawaiiensis_Makaopuhi", "P_wawraeDL7428"))  
> dec.range.table[ndi,]
```

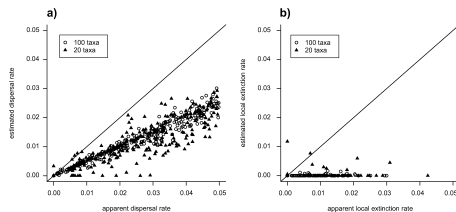
What Have We Learned?

- An estimate of the ancestral areas of the various clades (the primary information).
- The global dispersal rate and extinction rate (is this really useful?).



What are the Limitations of this Analysis?

- Number of areas
 - May not make sense to aggregate areas or limit range sizes.
 - Only other option is to use other methods.
- Extinction rate
 - A known issue with the DEC model family is that it underestimates the extinction rate, sometimes pretty considerably.



- Is this important? As a nuisance parameter, how can this affect the analyses?

Exercise II: Adding Founder-Event Speciation

Founder-Event Speciation

- Founder-effect speciation (dispersal of a lineage to a new area, followed by speciation) is not only a well-established biogeographical process, it is arguably a fairly important one in island systems.
- The “classic” DEC model *can* actually replicate the effect of this process by a sequence of two independent range evolution events in one of the daughter lineages following cladogenesis:
 - Anagenetic range gain through dispersal (new area added to range)
 - Anagenetic range loss of *all* old areas through extinction
- While possible, the joint probability of this consecutive sequence of events on any particular lineage, especially of the multiple extinctions required before, e.g. another cladogenesis event, is extremely low.
- This is exacerbated by the tendency of the DEC model to underestimate the extinction rate.
- Thus the founder-event speciation pattern is difficult to highly-unlikely to be recovered under the classic DEC model.

Effects of Adding Founder-Event Speciation

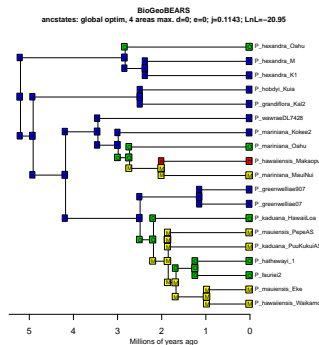
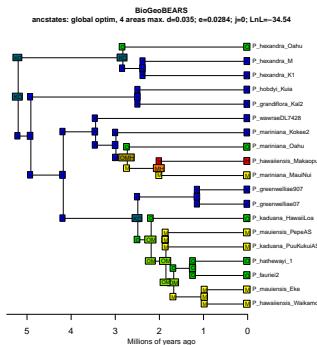
- Matzke (2014) extended the DEC model by (among other things) adding founder-event as a distinct cladogenetic range evolution process.
- This adds one additional parameter, j , to the classic DEC model.
- Both AIC as well as LRT model selection approaches dramatically preferred the more complex model.
- When including founder-event dispersal in island systems, in some cases estimated anagenetic dispersal and extinction rates dropped to nearly 0.
- This indicates that in many island systems, cladogenetic dispersal is the dominant mode structuring biogeographical relationships rather than vicariance.
- Furthermore, ancestral range estimates also changed dramatically: ancestral nodes were usually estimated to occupy only a single area instead of being widespread.

Adding the Founder-Event Speciation Process to the DEC Model

- Available in “BioGeoBEARS” by freeing the `j` parameter:

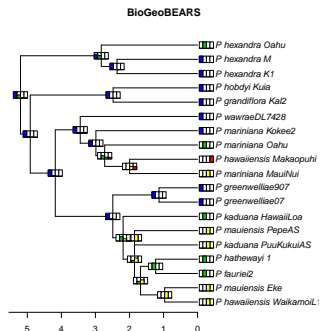
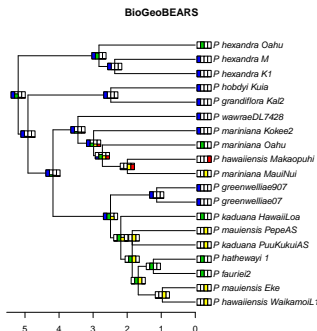
```
> run.object$BioGeoBEARS_model_object@params_table["j","type"] ="free"  
> run.object$BioGeoBEARS_model_object@params_table["j","init"] = 0.0001  
> run.object$BioGeoBEARS_model_object@params_table["j","est"] = 0.0001  
> run.object$BioGeoBEARS_model_object@params_table["j","min"] = 0.0000
```

DEC vs. DEC+J (Marginal Likelihood of Ranges)



	DEC	DEC+J
lnL	-34.54196	-20.94759
parameters	2	3
AIC	73.08	47.90
d	0.03504399	1.000000e-12
e	0.02835955	1.000000e-12
j	N/A	1.142538e-01

DEC vs. DEC+J (Marginal Likelihood of Areas)



	DEC	DEC+J
lnL	-34.54196	-20.94759
parameters	2	3
AIC	73.08	47.90
d	0.03504399	1.000000e-12
e	0.02835955	1.000000e-12
j	N/A	1.142538e-01

What Processes Do We Need to Include?

- As with any analyses, too many parameters can result in loss of power, while too little can lead to bias.
- How do we decide what processes are important enough to include?
 - By using our understanding of the system
 - Adopt a model selection approach, just like we do with models of nucleotide evolution: AIC, LRT, etc.
 - More sophisticated approaches in the future: model-averaging, transdimensional MCMC, etc.

Exercise III: Changing the Connectivity Between Areas

Constructing a Stepping-Stone System

- The previous analyses all assumed equal connectivity between all areas.
- We can construct a model that restricts the connectivity between areas to:
 - Incorporate information from geology etc.
 - Test different hypotheses
- We communicate this information to “BioGeoBEARS” through a *dispersal multiplier* matrix:

K	O	M	H
1	1	0	0
0	1	1	0
0	0	1	1
0	0	0	1

END

The Dispersal Multiplier Matrix File

- Multiple matrices can be listed in a file (with one matrix for each time period, most recent first, if doing a time-stratified analysis).
- List of matrices terminated with “END”.
- First row consists of area names, SEPARATED BY TAB-CHARACTERS (NOT SPACES).
- Subsequent rows consist of connectivity from area (given by row index) to another area (given by column index); AGAIN, COLUMNS ARE SEPARATED BY TAB-CHARACTERS!

```
K O M H
1 1 0 0
1 1 0 0
0 0 1 1
0 0 1 1
```

```
K O M H
1 1 0 0
0 1 1 0
0 0 1 1
0 0 0 1
```

```
K O M H
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
```

END

The Dispersal Multiplier Matrix File

- A “1” indicates dispersal is possible, while a “0” indicates that dispersal is not.
- Other values are possible, e.g. “0.1” to indicate a reduce rate of dispersal.
- Set “`dispersal_multipliers_fn`” on run object to path to dispersal matrix file, followed by calling “`readfiles_BioGeoBEARS_run`” on run object.
- Different hypothesis of connectivities can be compared using model selection approaches.

```
run$dispersal_multipliers_fn = "data/dispersal.txt"  
run = readfiles_BioGeoBEARS_run(run)
```

```
K O M H  
1 1 0 0  
1 1 0 0  
0 0 1 1  
0 0 1 1
```

```
K O M H  
1 1 0 0  
0 1 1 0  
0 0 1 1  
0 0 0 1
```

```
K O M H  
1 1 1 1  
1 1 1 1  
1 1 1 1  
1 1 1 1
```

END

Exercise IV: Time-Stratified Analysis

Changing the Model Over Time

- *Time-stratification* is the term used to describe a DEC analysis where we impose a schedule of changes to the geography over time.
- This can reflect information we have that we want to use to structure the analysis or maybe just a hypothesis for which we want to calculate a likelihood for testing.
- We specify the time periods using a simple file, with the beginning of each time period listed on its own row, with the most recent time first.
- Units are, of course, the same as given in the phylogeny.
- The final row specifies the start of the oldest time period, and this *must* be older than the root of the phylogeny.
- All other information (e.g., dispersal multipliers) must be provided for each time period.

Specifying Time-Stratified Dispersal Multipliers

Time Periods:

0.5

3.7

10

Dispersal Multipliers:

K O M H

1 1 1 1

1 1 1 1

1 1 1 1

1 1 1 1

K O M H

1 1 0 0

0 1 1 0

0 0 1 1

0 0 0 1

K O M H

1 1 0 0

1 1 0 0

0 0 0 0

0 0 0 0

END

Exercise V: Areas-Allowed

Dealing with Emergent/Disappearing Islands/Habitats/Areas

- Initially, areas that were not available during some time periods were modeled by prohibiting dispersal to and from them.
- This was not ideal, as it allows for lineages to persist in areas that should not exist.
- “BioGeoBEARS” allows for you specify a “schedule” of area availability, just like you would specify the dispersal multipliers.
- File and matrix format is exactly the same, with “1” specifying that area is available for that time period, and “0” specifying that it is not.

Time Periods:

0.5
1.9
3.7
10

Available Areas:

K O M H
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1

K O M H
1 1 1 0
1 1 1 0
1 1 1 0
0 0 0 0

K O M H
1 1 0 0
1 1 0 0
0 0 0 0
0 0 0 0

K O M H
1 0 0 0
0 0 0 0
0 0 0 0
0 0 0 0

END