

# Phylogenetic Inference using RevBayes

*Total-evidence Dating under the FBD Model*

Tracy A. Heath, April Wright, and Walker Pett

## 1 Introduction

?

### 1.1 Models

#### 1.1.1 Sequence Evolution

Point to other tutorials (e.g., GTR stuff)

#### 1.1.2 Morphological Character Change

Mk models and ascertainment bias

#### 1.1.3 Lineage-Specific Substitution Rates

Clocks (?) and relaxing them

#### 1.1.4 Lineage Diversification and Sampling

Birth-death processes and FBD

## 2 Prerequisites

What do you need to know before doing this?

### 2.1 Requirements

We assume that you have read and hopefully completed the following tutorials:

- `RB_Getting_Started`
- `RB_Basics_Tutorial`

Note that the `RB_Basics_Tutorial` introduces the basic syntax of `Rev` but does not cover any phylogenetic models. You may skip the `RB_Basics_Tutorial` if you have some familiarity with `R`. We tried to keep this tutorial very basic and introduce all the language concepts on the way. You may only need the `RB_Basics_Tutorial` for a more in-depth discussion of concepts in `Rev`.

### 3 Data and files

We provide the data file(s) which we will use in this tutorial. You may want to use your own data instead. In the **data** folder, you will find the following files

- **bears\_extant\_cytb.nex**: description of file/data (we also need more descriptive file names).

## 4 Exercise: Title

### 4.1 Getting Started

We will complete this analysis in **RevBayes** by entering the **Rev** code interactively.

### 4.2 Loading the Data

→ Download data and output files (if you don't have them already) from:  
<http://revbayes.github.io/tutorials.html>

Start up **RevBayes** at the command line. You should do this from within the **RB\_TotalEvidenceDating\_FBD\_Tutorial** directory.

```
./rb
```

If you've done this correctly, you should see the prompt shown below

```
>
```

If you do not see this, please ask for help.

First load in the sequences using the **readDiscreteCharacterData()** function. Note that copying and pasting instructions can lead to extraneous characters being copied. This can cause commands to fail. For learning, it's often better to 'live code' and type the commands in manually, rather than copying and pasting. **Ctrl+A** allows you to jump to the start of the command, if you need to delete extra characters from the front of a line. **Ctrl+E** allows you to jump to the end of a command. Pressing the up key will pull up previous commands, and allow you to edit them. If you do choose to copy and paste in commands, doing that from the tutorial script file will cause fewer errors.

```
cytb <- readDiscreteCharacterData("data/bears_cytb.nex")
```

Executing these lines initializes the data matrix as the respective **Rev** variables. To report the current value of any variable, simply type the variable name and press enter. For the **data** matrix, this provides information about the alignment:

```
cytb
  DNA character matrix with 10 taxa and 1000 characters
=====
Origination:                bears_cytb.nex
Number of taxa:              10
Number of included taxa:     10
Number of characters:        1000
Number of included characters: 1000
Datatype:                    DNA
```

Next, we import the morphological character matrix.

```
morpho <- readDiscreteCharacterData("data/bears_morphology.nex")
```

The information about the morphological character matrix should look like this:

```
morpho
  Standard character matrix with 18 taxa and 81 characters
=====
Origination:                bears_morphology.nex
Number of taxa:              18
Number of included taxa:     18
Number of characters:        81
Number of included characters: 81
Datatype:                    Standard
```

Now we read in the full list of taxa and create a workspace object with the total number of taxa.

```
taxa <- readTaxonData("data/bears_taxa.tsv", delimiter=TAB)
n_taxa <- taxa.size() # the number of taxa
```

If you open the bears taxa file, you'll notice that this is a tab-separated file of all of the taxon names, with the age in millions of years ago (mya) in the second column. An age of 0.0 indicates an extant bear. We will use this information to allow fossils to be incorporated as tips in the analysis.

The last bit of data preparation we will do is to add any taxa that are not found in the molecular partition (i.e. are only found in the fossil data) to the molecular partition as missing data, and vice versa.

```
cytb.addMissingTaxa( taxa )  
morpho.addMissingTaxa( taxa )
```

In order for all the taxa to appear on the same tree, they all need to be part of the same dataset, as opposed to present in separate datasets. This ensures that there is a unified taxon set that contains all of our tips.

### 4.3 Specify monophyletic clade constraints

Version dated: December 16, 2016