

# Setup Instructions

David Levy-Booth

April 7 - 9, 2020

## Setup Instructions

Please come to the workshop with your laptop setup with the required software and data files as described in our setup instructions.

## Making an RStudio Project

Projects allow you to divide your work into self-contained contexts.

Let's create a project to work in.

In the top-right corner of your RStudio window, click the "Project: (None)" button to show the projects drop-down menu. Select "New Project." > "New Directory" > "New Project." Under directory name, input "metagenomics" and choose a parent directory to contain this project on your computer.

## Installing Packages

Many R packages can be downloaded and installed from The Comprehensive R Archive Network (CRAN).

You can install each package with the following code:

```
install.packages("BiocManager",  
                 "devtools",  
                 "RColorBrewer",  
                 "ape",  
                 "tidyverse",  
                 "ggplot2",  
                 "phytools",  
                 "stringr",  
                 "reshape2",  
                 "cowplot",  
                 "grid",  
                 "circlize",  
                 "dplyr",  
                 "aplot")
```

To make things a bit more difficult, there are other source of R packages, such as Bioconductor, which provides tools for the analysis and comprehension of high-throughput genomic data.

```
BiocManager::install("pathview")}
```

## Install ggtree

And finally, we have to install the developmental version of *ggtree* from GitHub. **This is very important to ensure colored ranges are applied correctly** (the Bioconductor version does not apply colors to trees correctly)

```
library(devtools)
install_github("YuLab-SMU/ggtree")
```

## Loading Packages

Next, we should check if all packages can be loaded properly. If a package does not load because of a missing package, please install the package (as shown above) and try again.

```
# Suite of packages for data manipulation and visualization
library(tidyverse)
# Working with trees
library(ape)
library(phytools)
# Plotting Grammar of Graphics
library(ggplot2)
# Tree visualization
library(ggtree)
# String manipulation
library(stringr)
# Data structure manipulation
library(reshape2)
# Multi-panel figure layouts
library(cowplot)
library(grid)
# Color manipulation
library(RColorBrewer)
# Pathway analysis with KEGG Orthologs
library(pathview)
#plot circular graphs and contig maps
library(circlize)
```

## Downloading Workshop Material

The material for this workshop is hosted on GitHub.

Follow the above link and click the green “Clone or download” button to download all files as a .zip archive. Unzip the archive to access files.

## Loading custom functions

I have written some helpful functions to replace some time-consuming workflows. Please load them from the source R script like this:

```
#Custom functions for metagenomics and metatranscriptomics  
source("scripts/metagenome.utilities.R")
```

Please ensure all packages are installed and up to date, and that you have all data downloaded prior to the start of the workshop. If you require any assistance, please contact David Levy-Booth (dlevyboo@mail.ubc.ca).

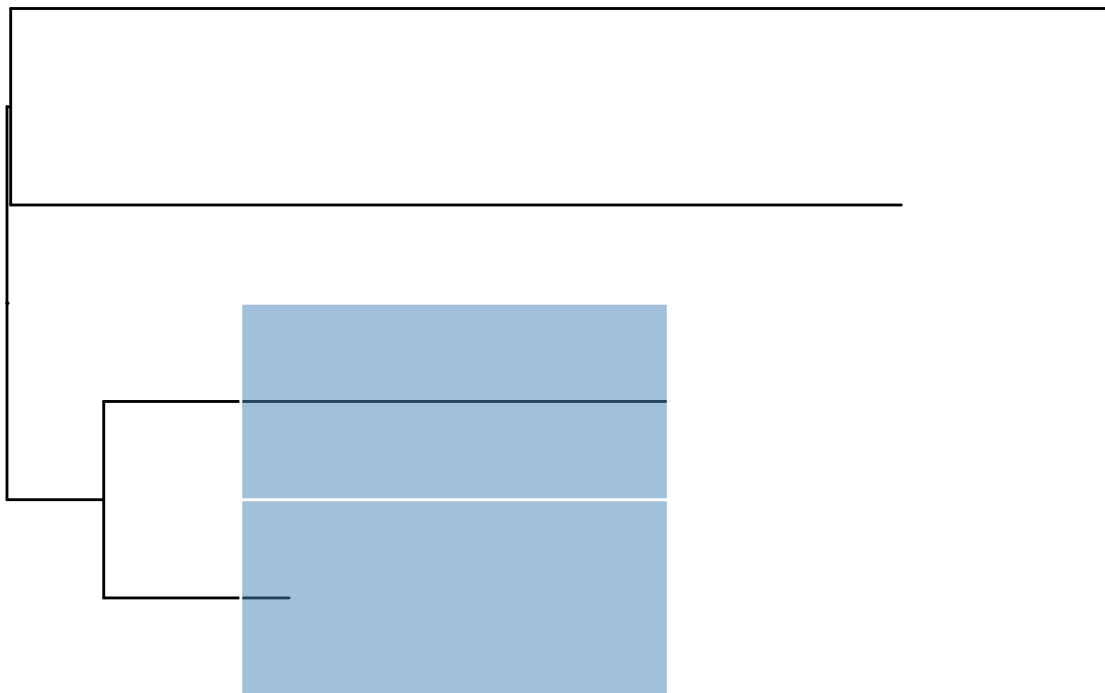
We will also spend a few minutes prior to the workshop ensuring that all attendees have completed setup and installation.

## Sanity Check

Because *ggtree* is the main package that we will work with throughout the workshop, please ensure that you can reproduce a simple example plot using this package.

Please enter the following code in RStudio and ensure your output matches the plot provided:

```
library(ape) #in case you did not load the package earlier.  
library(ggtree)  
  
set.seed(2020)  
example.tree <- rtree(4)  
ggtree(example.tree) + geom_balance(node=6)
```



If your plot **does not** look like this, you may have to re-install the *ggtree* package from GitHub and any dependent packages. If it does, congratulations, we are ready to begin!