

# Setup Instructions

David Levy-Booth

March 17 - 19, 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). First we must choose a CRAN mirror, or location from which we will download packages.

One way we can do this is to set our repository manually to automatically download packages from nearby mirrors:

- Simon Fraser University <https://mirror.its.sfu.ca/mirror/CRAN/>
- University of Toronto <http://cran.utstat.utoronto.ca/>
- Oregon State University <https://ftp.osuosl.org/pub/cran/>

```
local({r <- getOption("repos")
  r["CRAN"] <- "https://mirror.its.sfu.ca/mirror/CRAN/"
  options(repos=r)})
```

Next we will download the R packages required for this workshop.

You could simply install each package with the following code:

```
install.packages(<package name>)
```

Or we can create a function that checks if you're missing the packages and download them if needed. Please ensure that packages are the latest versions.

```
# List required packages that are found on CRAN
cran_packages <- c("BiocManager",
                  "devtools",
                  "RColorBrewer",
                  "ape",
                  "tidyverse",
                  "ggplot2",
                  "phytools",
                  "stringr",
                  "reshape2",
                  "cowplot",
                  "grid",
                  "circlize",
                  "dplyr")

for(package in cran_packages) {
  if(package %in% rownames(installed.packages()) == FALSE) {
    install.packages(package, repos='http://cran.utstat.utoronto.ca/')
    library(package, character.only = TRUE)
  }
}
```

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. Here's how we load the

```
#List required packages that are found on Bioconductor
bioc_packages <- c("pathview")

for(package in bioc_packages) {
  if(package %in% rownames(installed.packages()) == FALSE) {
    source("https://bioconductor.org/biocLite.R")
    BiocManager::install(package)
    library(package, character.only = TRUE)
  }
}
```

## 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

```
# 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. There are two easy ways to download the GitHub repository:

1. 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.
2. Clone the repository to your computer with git (must have git installed locally):

```
git clone https://github.com/davidlevybooth/Metagenomics_Workshop.git
```

## Downloading data

The data should have been downloaded for you using the above commands. If you would like to download the data manually, enter the following command downloads the data from the workshop GitHub. Since you’re working in a Project, saves it in the Project directory under the ‘data’ directory on your computer.

```
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/checkm.qa.tsv",
  "data/checkm.qa.tsv")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/gtdb.classification",
  "data/gtdb.classification.tree")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/gtdb.classification",
  "data/gtdb.classification.tsv")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/gtdb.taxonomy.tsv",
  "data/gtdb.taxonomy.tsv")
```

```
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/metagenome.counts",
  "data/metagenome.counts.tsv")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/K0.genomes.tsv",
  "data/K0.genomes.tsv")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/K0.pathways.tsv",
  "data/K0.pathways.tsv")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/LE.CH5.EKL.cov",
  "data/LE.CH5.EKL.cov")
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/data/LE.CH5.gff",
  "data/LE.CH5.gff")
```

## Downloading Custom Functions

Throughout this workshop we will develop and apply functions that we wish to recall later. The R script “metagenome.utilities.R” contains a series of helper scripts to carry out advanced metagenomic and metatranscriptomic analysis. Advanced users are encouraged to read through this file and alter the functions therein.

```
download.file(
  "https://raw.githubusercontent.com/davidlevybooth/Metagenomics_Workshop/master/scripts/metagenome.utilities.R",
  "scripts/metagenome.utilities.R")
```

### Loading custom functions

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

That’s it!

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.