**Bioassay data plotting in R**

by Peter van Dam, nov 2016

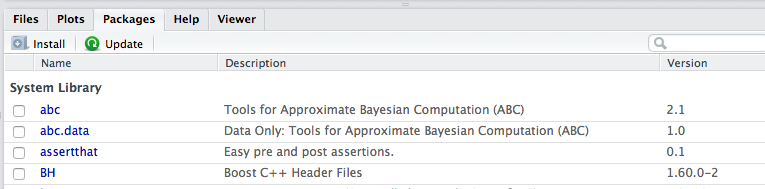
This script executes a number of tasks:

1. Import the data as a csv file for each csv file in a folder.
2. Reorganize the data for R to be used in a plot
3. Test for significance with ANOVA and Tukey Honest Significant Difference (Tukey-HSD)
4. Plot the data, with the labels from the statistical test above each bar. Freshweight measurements inside the /fw/ folder will be plotted as bar graphs ± SEM. Disease Indexes will be plotted as stacked bar graphs.
5. The calculated values and significance categories will be outputted to a csv file in the data folder.

Software needed:

1. R (<https://cran.r-project.org/bin/macosx/>); this is the programming language in which the script has been written.
2. Rstudio (<https://www.rstudio.com/products/rstudio/download/>); this is an easy way to write and test scripts. Importantly, it also has an easy way to install packages that are needed for the script to run (e.g. for reorganizing the data and plotting the actual graphs)

Once both are installed, launch Rstudio, and find the packages panel on the right side of the screen and click the ‘Install’ button (alternatively, go to Tools > Install Packages in the top bar).



Go ahead and install these packages one by one:

* ggplot2
* reshape2
* multcompView
* plyr
* tools
* gdata
* dplyr

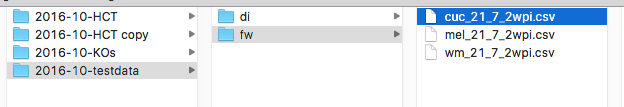
Preparing your data:

The script takes in a folder (please specify in the first lines of the script) that contains two folders: ‘/fw/’ and ‘/di/’. Please prepare these two folders manually.

These folders should contain comma-separated-value (CSV) files with the measurements. If no csv files are found, nothing will be executed. The filenames should be formatted like this: “<plantname>\_<temperature>\_<10^.. spores/ml>\_<time\_after\_inoculation>.csv”.

e.g.: “cuc\_21\_7\_2wpi.csv”

Each plant species tested is plotted with a different color (these are supported: cuc, mel, wm, tom).



The data should be formatted like this (so separated by comma’s):

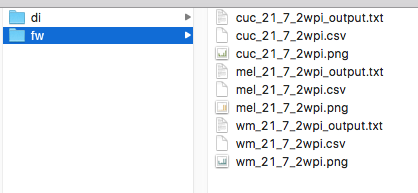
You can export your data from excel (file > save as > comma separated values).

|  |  |
| --- | --- |
| FW | DI |
| A,5.2,10.5,9.2,7,9.3,6.6,2.4,1.2  B,3.5,1.6,0.8,0.4,1.7,1.9,9.2,0.2  C,1.5,0.4,0.5,0.2,0.2,0.8,3.4,0.3  D,5.3,2.7,1.4,6.8,2.8,1.4,0.3,3  E,0.2,0.2,0.9,0.2,8.1,1.2,0.4,1.5  F,12.9,6.1,8.3,9.8,4.6,4.1,9,4.7  G,0.2,2,0.4,0.3,2.5,4.6,0.4,0.3  mock,15.8,9.5,10.4,12.2,8.6,18,12.1,11.5 | A,1,0,0,0,0,0,3,3  B,3,3,4,4,3,3,1,4  C,3,4,4,4,4,4,3,4  D,2,4,4,1,3,3,4,3  E,4,4,4,4,0,4,4,4  F,0,0,0,0,0,0,0,0  G,4,3,4,4,3,3,4,4  mock,0,0,0,0,0,0,0,0 |

Running the script:

1. Edit the first lines to point to the main directory (in this case: “/Users/Peter/Programming/R/FW\_plot\_and\_anova/data/2016-10-testdata/”).
2. To run all the code, press CMD+ALT+R
3. If everything goes well, you will see some output in the bottom left panel (the ‘console’) and in your ‘/fw/’ and ‘/di/’ folders you will see some output files appear.

The output:



For each of the csv files you should see a png file with the graph and an output (txt) file containing the data that was used for plotting, including the mean, stdev, standard error of the mean, as well as the statistical categories (a / b / ab / etc). This can be imported to excel or another graph plotting program in case you want to edit some of the settings.

These are two example plots:

