# Technical Bulletin – BIN06 RBioplot: Automated statistical analysis and data visualization Hanane Hadj-Moussa and Jing Zhang PhD, 2016.

#### **Overall:**

- I. Installation
  - A. R installation
  - B. RStudio installation
- II. Setting up RBioplot
- III. Statistical analysis (rbiostats)
- IV. Plotting
  - A. Histogram (rbioplot)
  - B. Heatmap (rbioplot heatmap)
  - C. Joint-point curve (rbioplot\_curve)
- V. RBioplot Resources

RBioplot is an R package designed for highly accessible yet comprehensive statistical analysis and data visualization. It represents a fully automated and versatile data processing solution for molecular biology and biochemistry. We need to break free from the confinements of SigmaPlot 12 and Jing Zhang PhD has developed RBioplot to help us do just that. RBioplot is a black box that will do all our graphing, plotting, and statistical analyses for us with minimal data reorganization, time, and energy. This techbull is the RBioplot for dummies user's manual.

\*Note: This techbull will refer you to various websites and to the RBioplot webpage that is hosted on our lab website (<a href="http://kenstoreylab.com/?page\_id=2448">http://kenstoreylab.com/?page\_id=2448</a>), where you can find all the sample input and output files used to make this techbull.

Before using RBioplot you should read the following paper that explains the core principles, functions, commands, and selected options. Also, the package is free of charge.

\*\*If using this program for data analysis and generating figures when publishing, remember to cite the paper below in your methods (or other appropriate) section\*\*:

Zhang J, Storey KB. (2016) RBioplot: an easy-to-use R pipeline for automated statistical analysis and data visualization in molecular biology and biochemistry. PeerJ 4:e2436.

#### I. Installing R and RStudio

To run RBioplot you must first install R and RStudio on your computer. RStudio is an IDE (interactive development environment) that allows you to interact with R in a more user friendly way. RStudio will be the main interface for R – even though you could launch R console alone, we almost never do so.

Links to both of these programs are on the lab website (<a href="www.kensotreylab.com">www.kensotreylab.com</a>) → Research → Research Tools → RBioplot (<a href="http://kenstoreylab.com/?page\_id=2448">http://kenstoreylab.com/?page\_id=2448</a>). Both R and RStudio are open-source software, meaning they are completely free. Additionally, both programs are cross platform, ie they will run on Windows, Mac and Linux computers.

\*Note: You will find that the "\*Notes" in this techbull will be useful for all the applications, not just the one it is listed in.

#### A. Installing "R"

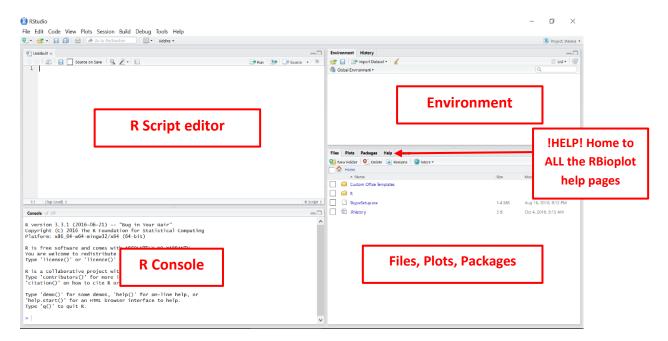
- 1) Install 'R' by visiting (<a href="https://www.r-project.org">https://www.r-project.org</a>) and then selecting your preferred CRAN mirror to download the installation file. We will use the University of Toronto's (<a href="http://cran.utstat.utoronto.ca/">http://cran.utstat.utoronto.ca/</a>).
- 2) Download the version of R that corresponds with your operating system. For example, if you are working on a Windows computer click Download R for Windows → install R for the first time → Download R 3.3.1 (or the latest version number if not 3.3.1) for Windows.
- 3) Save the installation .exe file → open the file → follow the R for Windows 3.3.1 Setup Wizard's simple installation instructions.
  - \*Note: You should install all the installation components and make sure to accept the default startup options.

#### B. Installing RStudio

- 1) To install **RStudio** visit (<a href="https://www.rstudio.com/">https://www.rstudio.com/</a>). Click **RStudio** → **Desktop** → **Download RStudio** Desktop
- 2) Download the RStudio version that corresponds with your operating system. For example, if you are working on a Windows computer click RStudio 0.99.903 (or the latest version number if not 0.99.903) Windows Vista/7/8/10
- 3) Save the installation .exe file → open the file → follow the RStudio Setup Wizard's simple installation instructions.
  - \*Note: Once the program has installed you should locate your RStudio shortcut and place it on your desktop.

#### **II. Setting up RBioplot**

Open the RStudio application, this will be the platform you will use to run RBioplot. The picture below is a breakdown of the different RStudio quadrants and panels. \*\*EVERYTHING IN RSTUDIO IS CASE-SENSITIVE\*\*



- 1) To setup the R Script text editor go to File → New File → R Script (make sure to select R Script, otherwise you will not have the option to run your commands). The R Script editor is the main window in which you will interact with R. You should prepare all your R commands before running your commands in the console. Also, the R script editor allows you to run multiple commands at the same time (as long as you highlight all the commands you want to run).
- 2) To use the latest version of RBioplot we need to first install the package **devtools.** This package is the main interface by which RBioplot can be installed. Both devtools and RBioplot are a one-time installation, unless there are recommended updates that are released. To install 'devtools' type the command below in the R Script editor.

#### install.packages("devtools")

\*Note: \*\*You must be connected to the internet\*\*. You should type all the commands written in this techbull and not just copy them into the text editor. Microsoft word has altered the formatting and syntax of the code written in this techbull. So if you copy/paste this into RStudio, your command will not run, because R Script quotation marks are uniquely annotated. If you copy/paste them then make sure you rewrite your quotation marks.

- 3) To run your R commands, you have two options:
  - a. **R Script editor**, (preferred): On Windows PC, highlight the command you want to run and press Ctrl+Enter. Or place your text cursor on the command you want to run and press Ctrl+Enter. If you separate your command into different lines then make sure all the lines are highlighted before pressing Ctrl+Enter, otherwise your full command will not run.
    - \*Note: on Mac, replace Ctrl key with Command key.

- b. **Console**: Copy your command from your text editor to the Console and press Enter.
- 4) Once your command has been successfully installed you should receive a message that reads something like this:

```
The downloaded binary packages are in C:\Users\Hanane\AppData\Local\Temp\Rtmp4gUWDW\downloaded_packages
```

- 5) Now you are ready to install the RBioplot package. Run the following command: devtools::install\_github("jzhangc/git\_R\_STATS\_KBS/package/RBioplot")
- 6) (optional) You may run into a few errors as the current version of **devtools** (1.12.0) has a bug that prevents from installing dependencies that RBioplot needs. In such case, you will need to manually install the missing packages.
  - \*Note: the bug will likely be fixed in the next release of devtools. To update devtools, just run the command again: install.packages("devtools")

Here are a few examples of errors and solutions, in this first one I was missing the scales package:

```
Error in loadNamespace(i, c(lib.loc, .libPaths()), versionCheck = vI[[i]])
: there is no package called 'scales'
```

so I installed it manually:

```
install.packages("scales")
```

After I installed the missing package (scales) I re-ran the RBioplot installation command in step 5.

```
Error in loadNamespace(j <- i[[1L]], c(lib.loc, .libPaths()), versionCheck = vI[[j]]): there is no package called 'sandwich'
```

Now, I am missing the package (**sandwich**), so I installed it with the command below and then re-ran the RBioplot installation command from step 5. Repeat this process until you have installed all the missing packages.

```
install.packages("sandwich")
```

Once RBioplot has successfully been installed you will get this message: \* DONE (RBioplot)

7) The next step is to setup your working directory, this is where you will place all your input data files and where all your plots, stats, and graphs will be exported to. In this example my working directory file is on my desktop and it's called "RBioplot stuffs". To setup your working directory run the following command and replace "C:\Users\Hanane\Desktop" with your folder address:

```
<u>For Windows</u>: setwd("C:\\Users\\Hanane\\Desktop\\RBioplot stuffs")

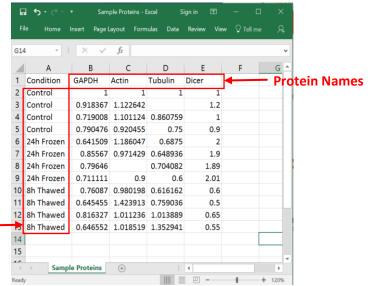
<u>For Mac and Linux</u>: setwd("C:/Users/Hanane/Desktop/RBioplot stuffs")
```

- \*Note: You will be required to reset your working domain once you exit and re-enter RStudio.
- 8) In the 'Files, Plots, and Packages' quadrant of the graph you should select 'Packages' and then find RBioplot in the list. Once you click on RBioplot you will be redirected to a page with the RBioplot documentation and all the help files for the different functions. If you get stuck using one of the commands use these help files to trouble shoot.
- 9) If your RBioplot stops working or there is an error that you have spent the last 3 weeks troubleshooting, then you should email the developer <a href="mailto:jzhangcad@gmail.com">jzhangcad@gmail.com</a> for help.

#### III. Statistical analysis

The rbiostats function is used to conduct statistical analysis. Before starting you should visit the **rbiostats** help page (or the **rbiostats mini-HELP! box** on page 6), this will outline all the available functions and how they work. All sample files used and generated can be downloaded from the website. **rbiostats** is a simple to use function for comprehensive statistical analyses that generates a statistical report for various statistical test including: t-tests, ANOVA and post-hoc tests (Tukey or Dunnett's). Make sure to read up on the various tests and to ensure you are using the one most appropriate for your data set and experimental design.

- Your data must be formatted in a very specific manner to ensure the values are analyzed properly by RBioplot. All excel files MUST be saved as a Microsoft Excel Comma Separated Values File (.csv).
- 2) The table on the right is my Sample Proteins.csv file, standardized western blot relative protein abundances. The values are standardized with Coomassie blue stained bands. Experimental conditions are listed in the first column. The A1 cell reads 'Condition' or something similar. Protein names are listed in the first row. The values I report are an average of the relative protein abundance for all my technical replicates that correspond to each of my biological replicates.



\*Note: You can analyze and graph your data even with outliers taken out in your table – there is no need to rearrange the table.

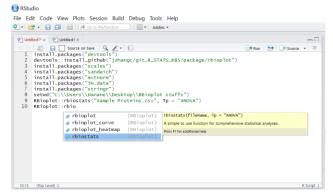
\*Note: Even though there are 4 different biological replicates for each experimental condition (ie. Control 1, 2, 3, and 4) in the example, make sure to keep the exact same identifier for each condition or the program will separate them into different conditions.

3) To run stats on your data using **rbiostats**, run the following command and replace "**Sample Proteins.csv**" with the file name for your .csv data file. Make sure to include the full file name (with .csv) in the quotations, and select the stats test you need. (ANOVA is the default, meaning that if you don't type the "Tp = " part, i.e. RBioplot::rbiostats("Sample Proteins.csv"), the function will run ANOVA)

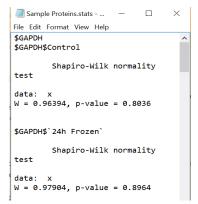
RBioplot::rbiostats("Sample Proteins.csv", Tp = "ANOVA")

\*Note: everything is case sensitive and has to be exact, meaning only "ANOVA" will work for ANOVA test. When you are not sure what to type, run command ?rbiostats to view the help page for this function, or see the section below.

4) As you are typing your command you will notice that RStudio has an auto-fill setting that you will find quite useful. If you hover your mouse over one of the options that appear, you will get a description of the function and the various arguments and customizations available.



7) Outputs a **stats.txt** file with Shapiro-Wilk normality test results, Bartlett test of homogeneity (equal variance) results, and the results of the statistical analysis of interest, this is automatically exported to your working directory. In this case, my file (on the left) was called **Sample Proteins.stats.txt**, the complete report can be downloaded from the website. Briefly, you need \*\*p values greater than 0.05\*\* for both Shapiro-Wilk normality test and Bartlett test to proceed to the follow up statistical tests (e.g. t-test, ANOVA). For a detailed break-down of the statistical report refer to Zhang and Storey (2016).



## rbiostats mini-HELP!

#### **Usage**

rbiostats(fileName, Tp = "ANOVA")

#### **Arguments**

Type of the intended statistical test. Case sensitive and type with quotation marks. Options are: "t-test", "ANOVA", "Tukey" and "Dunnett". Default is "ANOVA".

#### **IV. Plotting**

You can plot various types of graphs with your data using the corresponding functions in each of the RBioplot packages. They are excellent for applications such as western blotting, protein abundance levels, transcription factor-DNA binding activity, miRNA expression analyses, mRNA abundance, enzyme studies and many others. All sample files used and generated can be downloaded from the website.

\*Note: The following case studies don't explore all of the customizable arguments and settings available for each function. There are many additional functions that you can use to tailor your dataset graph requirements such as; plot width, plot height, using standard deviation for your error bars, etc.

#### A. Histogram

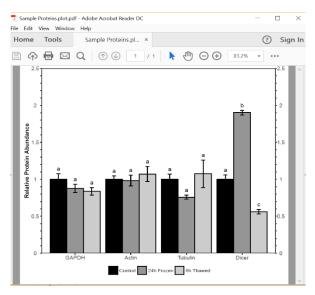
**rbioplot** is a simple to use function for plotting basing on the statistical analysis of choice. Before starting visit the **rbioplot** help page, this will outline the different functions and arguments and how to properly use them. You can also use the **rbioplot mini-HELP! box** on page 10.

- 1) For this example I used the western blotting data set from **Sample Proteins.csv** (check out rbiostats, step1). The protein abundance, PCR quantification, activity levels, etc. should all be in this format for **rbioplot.**
- 2) To generate a histogram run the following rbioplot command:
   RBioplot::rbioplot("Sample Proteins.csv", Tp = "Tukey", Title = NULL,
   errorbar = "SEM", errorbarWidth = 0.2, fontType = "sans", xLabel = NULL,
   xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel =
   NULL, yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE, plotWidth
   = 170, plotHeight = 150, y\_custom\_tick\_range = FALSE, y\_lower\_limit = 0,
   y\_upper\_limit, y\_major\_tick\_range, y\_n\_minor\_ticks = 4)
  - \*Note: Since most of the arguments have a default value, you can run something like rbioplot("Sample Protein") with all the default values. However, it is recommended that you play with the settings to get the graph that you are want. This is true of all the functions in the package.
- 3) The above command is very versatile and should be tailored to your specific dataset. You can use the customizable arguments to design what your graph looks like. Some of the customizable options include: Type of statistical test, Title, error bars, font, axes labels, Y-axis ticks, and plot height and width. For more information about the individual arguments go to the help pages described above.

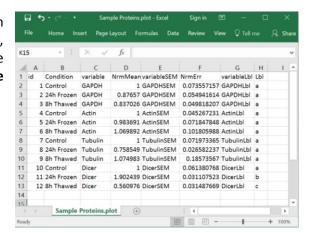
4) Once you run the command, a preview of your graph will appear in the lower right quadrant of RStudio. A pdf version of your graph (600 dpi resolution.) will also be automatically exported into your working directory. In this example the pdf generated was titled **Sample Proteins.plot.pdf**.

\*Note: If you re-run the above command using the same input file name then the new generated graph will overwrite your old graph.

\*Note: Make sure the .pdf and .plot.pdf file that were generated are closed before re-running. You cannot generate a new file with the exact same name if the old version is active.



5) The above command also outputs a .csv file with detailed metrics for the plot, including Mean, SEM and significance labels. In this sample the exported file was named Sample Proteins.plot.csv:



- 6) If you notice that the default axes on your newly generated graph are not optimal or do not match the criteria required for publication in your journal of interest, then you can use the autorange\_bar\_y function. It is a function that allows you to get custom lower/upper limit, major tick range, as well as minor tick options for y axis, based on a user-defined major tick number. The purpose of this function is to allow the user to determine the optimal number of minor ticks based on a user-defined major tick number.
- 7) For the autorange\_bar\_y function, data should be arranged as the same input file for rbioplots, in this example the file "Sample Proteins.csv" was used, remember to replace this with your "file name.csv". To find the perfect tick numbers for your y-axis run the following autorange\_bar\_y command: RBioplot::autorange\_bar\_y("Sample Proteins.csv", Nrm = TRUE, errorbar = "SEM", nMajorTicks = 5, DfltZero = TRUE)
  - \*Note: If you have negative values then you must select 'DfltZero = FALSE'.
  - \*Note: If you do not require a specific number of major ticks then the default setting of 5 will be selected.

8) The output is a list containing lower\_limit, upper\_limit, major\_tick\_range, and minor\_tick\_options. This list will appear directly in your R Console:

```
$y_axis_range
lower_limit upper_limit major_tick_range $minor_tick_options
0.0 2.5 0.5 [1] 0 4
```

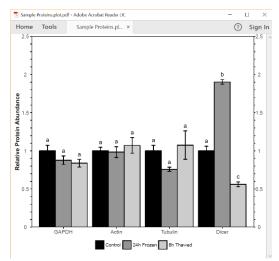
\*Note: The above values are just recommendations. The minor\_tick\_options sections shows all the available numbers of minor ticks (0 or 4) for this particular data set with the user-defined number of major ticks set to 5.

9) Once autorange\_bar\_y has detected the y-axis range that best represents your data you can input the above values and arguments into the main rbioplot command, see following (highlighted):

RBioplot::rbioplot("Sample Proteins.csv", Tp = "Tukey", Title = NULL, errorbar = "SEM", errorbarwidth = 0.2, fontType = "sans", xLabel = NULL, xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel = "Relative Protein Abundance", yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE, plotwidth = 170, plotHeight = 150, y\_custom\_tick\_range = TRUE, y\_lower\_limit = 0, y\_upper\_limit = 2.5, y\_major\_tick\_range = 0.5, y\_n\_minor\_ticks = 4)

\*Note: Generally, the default graphing settings will try to optimize the plot area to our dataset, which means you only need to use the **autorange\_bar\_y** function if you have specific major and minor tick values you would like to use.

- 10) The following **Sample Proteins.pdf** was generated. In this case, we can see that the new graph is identical to the old graph, that's because the predicted y\_axis\_range was the same as the default settings. This is not always the case and I recommend that you run the **autorange\_bar\_y** function for all your histograms to ensure that your graphs are visually appealing and easy to read.
- 11) When customizing the command arguments, you could be forcing the graphing functions to generate a non-optimal graph. Be sure to visually inspect your plots and verify that they make sense. You may need to manually change the settings accordingly.



12) For more information on what the individual arguments mean and how to properly use them read the **autorange bar y** help page on RStudio.

# rbioplot mini-HELP!

```
Usage
```

```
rbioplot(fileName, Tp = "Tukey", Title = NULL, errorbar = "SEM",
  errorbarWidth = 0.2, fontType = "sans", xLabel = NULL,
 xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5,
  yLabel = NULL, yTickLblSize = 10, yTickItalic = FALSE,
  legendTtl = FALSE, plotWidth = 170, plotHeight = 150,
  y custom tick range = FALSE, y lower limit = 0, y upper limit,
  y_major_tick_range, y n minor ticks = 4)
```

#### Arguments

fileName Input file name. Case sensitive and be sure to type with quotation marks. Currently only

takes .csv files.

Type of the intended statistical test. Case sensitive and be sure to type with quotation qΤ

marks. Options are: "t-test", "Tukey" and "Dunnett". Default is "Tukey".

Title The displayed title on top of the plot. Be sure to type with quotation marks. Default is NULL. errorbar

Set the type of errorbar. Options are standard error of mean ("SEM"), or standard deviation

("SD"). Default is "SEM".

errorbarWidth Set the width for errorbar. Default is 0.2.

The type of font in the figure. Default is "sans". For all options please refer to R font table, fontType

which is available on the website: http://kenstoreylab.com/?page\_id=2448.

xLabel x axis label. Type with quotation marks. Default is NULL.

xTickLblSize Font size of x axis ticks. Default is 10.

xTickItalic Set x axis tick font to italic. Default is FALSE.

xAngle The rotation angle (degrees) of the x axis marks. Default is 0 - horizontal.

xAlign The alignment type of the x axis marks. Options are 0, 0.5 and 1. The default value at 0 is

especially useful when xAngle = 90.

yLabel y axis label. Type with quotation marks. Default is NULL.

yTickLblSize Font size of y axis ticks. Default is 10.

yTickItalic Set y axis tick font to italic. Default is FALSE.

legendTtl Hide/Display legend title. If TRUE or T. the name of the first column of the raw data file will

display as the legend title. Default is FALSE.

plotWidth The width of the plot (unit: mm). Default is 170. Default will fit most of the cases. plotHeight The height of the plot (unit: mm). Default is 150. Default will fit most of the cases.

y custom tick range To initiate setting the

custom y upper limit, y lower limit, y major tick range, y n minor ticks.

Default is FALSE.

y lower limit Can only be set when y custom tick range = TRUE. Set custom lower limt for y axis.

Default is 0. Value can be obtained from autorange bar y.

y upper limit Can only be set when y custom tick range = TRUE. Set custom upper limt for y axis.

Value can be obtained from autorange bar y.

y major tick range Can only be set when y custom tick range = TRUE. Set custom major tick range for

y axis. Value can be obtained from autorange bar y.

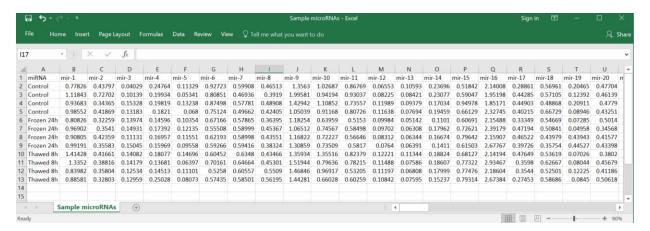
y n minor ticks Can only be set when y custom tick range = TRUE. Set custom numbers of minor

ticks. Default is 4. Value can be obtained from autorange bar v.

#### B. Heatmap

The **rbioplot\_heatmap** function can be used to analyze and graph largescale high-throughput studies such as large explorations of microRNA differential expression. **rbioplot\_heatmap** is a function for plotting simple heatmap based on the statistical analysis of your choice. Before starting visit the **rbioplot\_heatmap** help page, this will outline the different functions and arguments and how to properly use them. You can also use the **rbioplot\_heatmap mini-HELP! box** on page 13 for a quick reference.

1) The table below is my **Sample miRNAs.csv** file, the values are standardized miRNA relative abundance levels (obtained from qRT-PCR and normalized to the reference gene). The A1 cell should contain the word **microRNA** (or whatever you are measuring). The values I report are an average of the relative miRNA abundance for all my technical replicates that correspond to each of my biological replicates. This is just a small portion of the file- the full file contains 152 microRNAs.



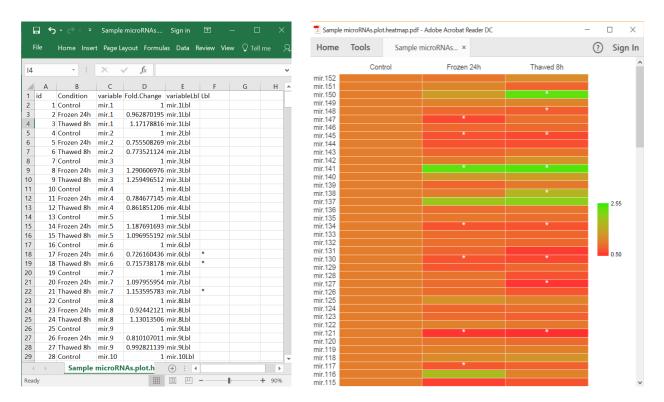
2) To generate a heatmap run the following **rbioplot** command:

RBioplot::rbioplot\_heatmap("Sample microRNAs.csv", Tp = "Dunnett", rmCntl = FALSE, Title = NULL, fontType = "sans", tileLow = "firebrick1", tileHigh = "green2", tileLbl = TRUE, tileLblSize = 5, tileTxtColour = "white", tileLblPos = 0.5, xLabel = NULL, xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel = NULL, yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE, legendPos = "right", plotWidth = 170, plotHeight = 600)

\*Note: In this case, the plotHeight argument was adjusted to 600 fit all make the individual heatmap cells bigger.

3) The above command is very versatile and should be tailored to your specific dataset. You can also easily change the colour scheme of your heatmap, refer to **V. RBioplot Resources** for the full colour palette available on with R. For more information on how to customize the arguments go to the help page or the mini-help box.

4) The above **rbioplot\_heatmap** command outputs a **.plot.csv** file with detailed metrics for the plot, including normalized mean and significance labels, as well as a plot image file (**.pdf**), with 600 dpi resolution.



```
rbioplot_heatmap mini-HELP!
Usage
rbioplot heatmap(fileName, Tp = "Dunnett", rmCntl = FALSE, Title = NULL,
  fontType = "sans", tileLow = "skyblue", tileHigh = "midnightblue",
  tileLbl = TRUE, tileLblSize = 10, tileTxtColour = "white",
  tileLblPos = 0.5, xLabel = NULL, xTickLblSize = 10,
  xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel = NULL,
  yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE,
  legendPos = "bottom", plotWidth = 170, plotHeight = 150)
Arguments
fileName
                         Input file name. Case sensitive and be sure to type with quotation marks. Currently only
                         takes .csv files.
Tp
                         Type of the intended statistical test. Case sensitive and be sure to type with quotation
                         marks. Options are: "t-test", "Tukey" and "Dunnett". Default is "Dunnett".
rmCntl
                         Remove the first column (i.e., control). Default is FALSE.
Title
                         The displayed title on top of the plot. Be sure to type with quotation marks. Default is NULL.
                         The type of font in the figure. Default is "sans". For all options please refer to R font table,
fontType
                         which is available on the website: http://kenstoreylab.com/?page_id=2448.
tileLow
                         Set the colour for the lower limit of the heatmap. Default is skyblue. For full colour options
                         and names, refer to the websitehttp://kenstoreylab.com/?page_id=2448.
                         Set the colour for the upper limit of the heatmap. Default ismidnightblue. For full colour
tileHigh
                         options and names, refer to the website http://kenstoreylab.com/?page_id=2448.
tileLbl
                         Enable or disable significant notation on the tiles. Default is TRUE.
tileLblSize
                         Set the font size of the tile label. Default is 10.
tileTxtColour
                         Set the colour of the on tile label. Default is "white". For full colour options and names,
                         refer to the website <a href="http://kenstoreylab.com/?page_id=2448">http://kenstoreylab.com/?page_id=2448</a>.
tileLblPos
                         Set the position of the tile lables. Options are 0, 0.5 and 1. Default is 0.5.
xLabel
                         x axis label. Type with quotation marks. Default is NULL.
xTickLblSize
                         Font size of x axis ticks. Default is 10.
xTickItalic
                         Set x axis tick font to italic. Default is FALSE.
xAngle
                         The rotation angle (degrees) of the x axis marks. Default is 0 - horizontal.
xAliqn
                         The alignment type of the x axis marks. Options are 0, 0.5 and 1. The default value at 0 is
                         especially useful when xAngle = 90.
yLabel
                         y axis label. Type with quotation marks. Default is NULL.
vTickLblSize
                         Font size of y axis ticks. Default is 10.
yTickItalic
                         Set y axis tick font to italic. Default is FALSE.
```

legendTtl Hide/Display legend title. Default is FALSE.

Set the legend position. Options are "top", "bottom", "left"and "right". Default legendPos

is "bottom".

plotWidth The width of the plot (unit: mm). Default is 170. Default will fit most of the cases. The height of the plot (unit: mm). Default is 150. Default will fit most of the cases. plotHeight

y custom tick range To initiate setting the

custom y upper limit, y lower limit, y major tick range, y n minor ticks.

Default is FALSE.

y upper limit Can only be set when y custom tick range = TRUE. Set custom upper limt for y axis.

Value can be obtained from autorange bar y.

y lower limit Can only be set when y custom tick range = TRUE. Set custom lower limt for y axis.

Default is 0. Value can be obtained from autorange bar y.

y major tick range Can only be set when y custom tick range = TRUE. Set custom major tick range for

y axis. Value can be obtained from autorange bar y.

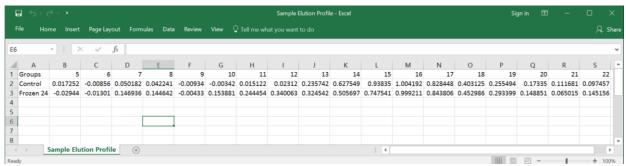
y n minor ticks Can only be set when y\_custom tick range = TRUE. Set custom numbers of minor

ticks. Default is 4. Value can be obtained from autorange bar y.

#### C. Joint-point curve

**rbioplot\_curve** is a simple to use function for plotting joining-point curve figures with continuous X-axis and Y-axis values. Before starting visit the **rbioplot\_curve** help page, this will outline the different functions and arguments and how to properly use them. You can also use the **rbioplot\_curve mini-HELP! box** on page 17.

1) In this example below I plotted my imaginary enzyme's activity for 18 eluted fractions from both control and 24 h frozen. I used the **Sample Elution Profile.csv** (below). The independent variables are listed in column A.

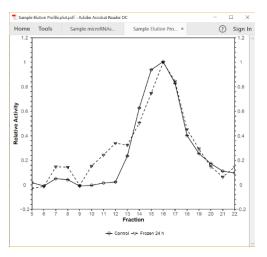


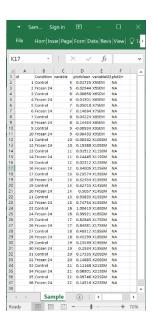
\*Note: biological replicates can also be accepted by this function. The format is the same as functions rbioplot and rbioplot\_heatmap. The rbioplot\_curve function will automatically detect the replicates and add appropriate error bars to the graph.

2) To generate a histogram run the following **rbioplot** command:

```
RBioplot::rbioplot_curve("Sample Elution Profile.csv", Title = NULL, errorbar = "SEM", errorbarWidth = 0.2, fontType = "sans", symbolSize = 2, xLabel = "Fraction", xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel = "Relative Activity", yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE, plotWidth = 170, plotHeight = 150, x_custom_tick_range = TRUE, x_lower_limit = 5, x_upper_limit = 23, x_major_tick_range = 1, x_n_minor_ticks = 0, y_custom_tick_range = TRUE, y_lower_limit = -0.2, y_upper_limit = 1.2, y_major_tick_range = 0.2, y_n_minor_ticks = 4)
```

3) The above **rbioplot\_curve** command outputs a **.plot.csv** file with detailed metrics for the plot, including Mean and SEM, as well as a plot image file (**.pdf**), with 600 dpi resolution.





- 4) With a joint-point curve both the y- and x- axes are continuous which means that both of their ranges can be tailored to your dataset of interest. To do this automatically you can use the **autorange\_curve** function. This function allows you to get custom lower/upper limit, major tick range, as well as minor tick options for both axes of a joint-point curve with continuous x AND y values, based on a user-defined major tick number. The purpose of this function is to allow the user to determine the optimal number of x and y minor ticks based on a user-defined major tick number.
  - \*Note: Generally, the default graphing settings will try to optimize the plot area to the dataset, which means you only need to use the **autorange\_curve** function if you have specific major and minor tick values you would like to use.
- 5) For the **autorange\_curve** function, data should be arranged as the same input file for **rbiocurve**, in this example the file "Sample Elution Profile.csv" was used, remember to replace this with your "file name.csv". To find the perfect axes for your dataset run the following **autorange\_curve** command:

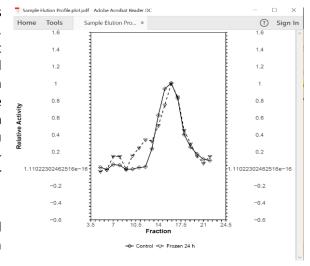
RBioplot::autorange\_curve("Sample Elution Profile.csv", errorbar = "SEM", x\_nMajorTicks = 5, x\_DfltZero = FALSE, y\_nMajorTicks = 10, y\_DfltZero = FALSE)

- \*Note: The DfltZero argument is set to "TRUE" by default which forces the axis to start from 0. You might need to set the x\_DfltZero and y\_DltZero arguments to 'FALSE' if the dataset contains negative values and/or large initial points into account.
- \*Note: If you do not require a specific number of major ticks then the default setting of 5 and 10 will be selected for x and y- axis, respectively.
- 6) The output is a list containing lower\_limit, upper\_limit, major\_tick\_range, and minor\_tick\_options. This list will appear directly in your R Console:
- \$x\_axis\_range
   x\_lower\_limit x\_upper\_limit x\_major\_tick\_range \$x\_minor\_tick\_options
   3.5 24.5 3.5 [1] 0 4 6 34
- - \*Note: The above values are just recommendations.
- 7) Once **autorange\_curve** has detected the x- and y-axes range that best represents your data you can input the above values and arguments into the main **rbioplot** command, see following (axis settings are highlighted):

RBioplot::rbioplot\_curve("Sample Elution Profile.csv", Title = NULL, errorbar = "SEM", errorbarwidth = 0.2, fontType = "sans", symbolSize = 2, xLabel = "Fraction", xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5, yLabel = "Relative Activity", yTickLblSize = 10, yTickItalic = FALSE, legendTtl = FALSE, plotwidth = 170, plotHeight = 150, x\_custom\_tick\_range = TRUE, x\_lower\_limit = 3.5, x\_upper\_limit = 24.5, x\_major\_tick\_range = 3.5, x\_n\_minor\_ticks = 4, y\_custom\_tick\_range = TRUE, y\_lower\_limit = -0.6, y\_upper\_limit = 1.6, y\_major\_tick\_range = 0.2, y\_n\_minor\_ticks = 4)

\*Note: When manually changing your x and y axis values you may get this error message, Error: `breaks` and `labels` must have the same length. This means that the upper and lower limit values you set do not match the major and minor tick values. For example, if your major tick value is 5 then you cannot set your lower limit to 3 as this is not a multiple of 5.

- 8) The following **Sample Elution Profile.pdf** was generated with the recommended x and y axes. As you can see, the graphing conditions were not optimal: the x axis is a fraction number and setting the lower limit to 3.5 does not make much sense since my first fraction was 5; whereas the default 10 major ticks setting produced an irregular y-axis tick setup. To change this, you may need to manually type in values for the x-axis, and set another major tick number other than 10 using the **autorange\_curve** function.
- 9) For more information on what the individual arguments mean and how to properly use them read the **autorange\_curve** help page on RStudio.



# rbioplot\_curve mini-HELP!

```
Usage
```

```
rbioplot_curve(fileName, Title = NULL, errorbar = "SEM",
  errorbarWidth = 0.2, fontType = "sans", symbolSize = 2, xLabel = NULL,
  xTickLblSize = 10, xTickItalic = FALSE, xAngle = 0, xAlign = 0.5,
  yLabel = NULL, yTickLblSize = 10, yTickItalic = FALSE,
  legendTtl = FALSE, plotWidth = 170, plotHeight = 150,
  x_custom_tick_range = FALSE, x_lower_limit = 0, x_upper_limit,
  x_major_tick_range, x_n_minor_ticks = 0, y_custom_tick_range = FALSE,
  y lower limit = 0, y upper limit, y major tick range, y n minor ticks = 4)
```

#### **Arguments**

fileName Input file name. Case sensitive and be sure to type with quotation marks. Currently only

takes .csv files. Note that the column names (excluding the first column) need to be numeric.

Title The displayed title on top of the plot. Be sure to type with quotation marks. Default is NULL.

errorbar Set the type of errorbar. Options are standard error of mean ("SEM"), or standard deviation ("SD").

Default is "SEM".

errorbarWidth Set the width for errorbar. Default is 0.2.

fontType The type of font in the figure. Default is "sans". For all options please refer to R font table, which is

avaiable on the website: http://kenstoreylab.com/?page\_id=2448.

symbolSize Set the size of symbols. Default is 2.

x Label x axis label. Type with quotation marks. Default is NULL.

xTickLblSize Font size of x axis ticks. Default is 10.
xTickItalic Set x axis tick font to italic. Default is FALSE.

xAngle The rotation angle (degrees) of the x axis marks. Default is 0 - horizontal.

xAlign The alignment type of the x axis marks. Options are 0, 0.5 and 1. The default value at 0 is especially

useful when xAngle = 90.

yLabel y axis label. Type with quotation marks. Default is NULL.

yTickLblSize Font size of y axis ticks. Default is 10.
yTickItalic Set y axis tick font to italic. Default is FALSE.

legendTtl Hide/Display legend title. If TRUE or T, the name of the first column of the raw data file will display as

the legend title. Default is FALSE.

plotWidth The width of the plot (unit: mm). Default is 170. Default will fit most of the cases.

plotHeight The height of the plot (unit: mm). Default is 150. Default will fit most of the cases.

x\_custom\_tick\_range To initiate setting the

custom x\_upper\_limit, x\_lower\_limit,x\_major\_tick\_range, x\_n\_minor\_ticks. Default

IS FALSE

x\_lower\_limit Can only be set when x custom tick range = TRUE. Set custom lower limt for x axis. Default

is 0. Value can be obtained from autorange curve.

x\_upper\_limit Can only be set when x custom tick range = TRUE. Set custom upper limt for x axis. Value can

be obtained from <u>autorange curve</u>.

Value can be obtained from autorange curve.

x\_n\_minor\_ticks Can only be set when x custom tick range = TRUE. Set custom numbers of minor ticks. Default

is 4. Value can be obtained from autorange curve.

y custom tick range To initiate setting the

custom y upper limit, y lower limit, y major tick range, y n minor ticks. Default

**is** FALSE.

y\_lower\_limit Can only be set when y custom tick range = TRUE. Set custom lower limt for y axis. Default

is 0. Value can be obtained from autorange\_curve.

y\_upper\_limit Can only be set when y custom tick range = TRUE. Set custom upper limt for y axis. Value can

be obtained from autorange curve.

Value can be obtained from autorange curve.

y\_n\_minor\_ticks Can only be set when y\_custom\_tick\_range = TRUE. Set custom numbers of minor ticks. Default

is 4. Value can be obtained from autorange curve.

### V. RBioplot Resources

To customize your graphs and make them more visually appealing (or usually you have to make them more boring for journals), you can refer to the R colour palette and R font table below.

## R colour palette

	darkorange4	gray	gray57	hotpink3	lightsalmon4	navajowhite1	plum3	slategray3	antiquewhite
brown3	darkorange3	goldenrod4	gray56	hotpink2	lightsalmon3	navajowhite	plum2	slategray2	aliceblue
brown2	darkorange2	goldenrod3	gray55	hotpink1	lightsalmon2	moccasin	plum1	slategray1	white
brown1	darkorange1	goldenrod2	gray54	hotpink	lightsalmon1	mistyrose4	plum	slategray	yellowgreen
brown	darkorange	goldenrod1	gray53	honeydew4	lightsalmon	mistyrose3	pink4	slateblue4	yellow4
blueviolet	darkolivegreen4	goldenrod	gray52	honeydew3	lightpink4	mistyrose2	pink3	slateblue3	yellow3
blue4	darkolivegreen3	gold4	gray51	honeydew2	lightpink3	mistyrose1	pink2	slateblue2	yellow2
blue3	darkolivegreen2	gold3	gray50	honeydew1	lightpink2	mistyrose	pink1	slateblue1	yellow1
	darkolivegreen1	gold2	gray49		lightpink1		pink	A CONTRACTOR OF THE PARTY OF TH	yellow
blue2				honeydew		mintcream	peru	slateblue	
blue1	darkolivegreen	gold1	gray48	greenyellow	lightpink	midnightblue		skyblue4	whitesmoke
blue	darkmagenta	gold	gray47	green4	lightgrey	mediumvioletred	peachpuff4	skyblue3	wheat4
olanchedalmond	darkkhaki	ghostwhite	gray46	green3	lightgreen	mediumturquoise	peachpuff3	skyblue2	wheat3
black	darkgrey	gainsboro	gray45	green2		ediumspringgree	n peachpuff2	skyblue1	wheat2
bisque4	darkgreen	forestgreen	gray44		htgoldenrodyello		peachpuff1	skyblue	wheat1
bisque3	darkgray	floralwhite	gray43	green	1000, 10	mediumseagreen	peachpuff	sienna4	wheat
bisque2	darkgoldenrod4	firebrick4	gray42	gray100	lightgoldenrod3	mediumpurple4	papayawhip	sienna3	violetred4
bisque1	darkgoldenrod3	firebrick3	gray41	gray99	lightgoldenrod2	mediumpurple3	palevioletred4	sienna2	violetred3
bisque	darkgoldenrod2	firebrick2	gray40	gray98	lightgoldenrod1	mediumpurple2	palevioletred3	sienna1	violetred2
beige	darkgoldenrod1	firebrick1	gray39	gray97	lightgoldenrod	mediumpurple1	palevioletred2	sienna	violetred1
azure4	darkgoldenrod	firebrick	gray38	gray96	lightcyan4	mediumpurple	palevioletred1	seashell4	violetred
azure3	darkcyan	dodgerblue4	gray37	gray95	lightcyan3	mediumorchid4	palevioletred	seashell3	violet
azure2	darkblue	dodgerblue3	gray36	gray94	lightcyan2	mediumorchid3	paleturquoise4	seashell2	turquoise4
azure1	cyan4	dodgerblue2	gray35	gray93	lightcyan1	mediumorchid2	paleturquoise3	seashell1	turquoise3
azure	cyan3	dodgerblue1	gray34	gray92	lightcyan	mediumorchid1	paleturquoise2	seashell	turquoise2
aquamarine4	cyan2	dodgerblue	gray33	gray91	lightcoral	mediumorchid	paleturquoise1	seagreen4	turquoise1
aquamarine3	cyan1	dimgrey	gray32	gray90	lightblue4	mediumblue	paleturquoise	seagreen3	turquoise
aquamarine2	cyan	dimgray	gray31	gray89	lightblue3 m	ediumaquamarin	e palegreen4	seagreen2	tomato4
aquamarine1	cornsilk4	deepskyblue4	gray30	gray88	lightblue2	maroon4	palegreen3	seagreen1	tomato3
aquamarine	cornsilk3	deepskyblue3	gray29	gray87	lightblue1	maroon3	palegreen2	seagreen	tomato2
antiquewhite4	cornsilk2	deepskyblue2	gray28	gray86	lightblue	maroon2	palegreen1	sandybrown	tomato1
antiquewhite3	cornsilk1	deepskyblue1	gray27	gray85	lemonchiffon4	maroon1	palegreen	salmon4	tomato
	cornsilk	deepskyblue	gray26	gray84	lemonchiffon3	maroon	palegoldenrod	salmon3	thistle4
antiquewhite1									
	cornflowerblue	deeppink4	gray25	gray83	lemonchiffon2	magenta4	orchid4	salmon2	thistle3
antiquewhite	coral4	deeppink3	gray24	gray82	lemonchiffon1	magenta3	orchid3	salmon1	thistle2
aliceblue	coral3	deeppink2	gray23	gray81	lemonchiffon	magenta2	orchid2	salmon	thistle1
white	coral2	deeppink1	gray22	gray80	lawngreen	magenta1	orchid1	saddlebrown	thistle
bisque3	coral1	deeppink	gray21	gray79	lavenderblush4	magenta	orchid	royalblue4	tan4
bisque2	coral	darkviolet	gray20	gray78	lavenderblush3	linen	orangered4	royalblue3	tan3
bisque1	chocolate4	darkturquoise	gray19	gray77	lavenderblush2	limegreen	orangered3	royalblue2	tan2
bisque	chocolate3	darkslategrey	gray18	gray76	lavenderblush1	lightyellow4	orangered2	royalblue1	tan1
beige	chocolate2	darkslategray4	gray17	gray75	lavenderblush	lightyellow3	orangered1	royalblue	tan
azure4	chocolate1	darkslategray3	gray16	gray74	lavender	lightyellow2	orangered	rosybrown4	steelblue4
azure3	chocolate	darkslategray2	gray15	gray73	khaki4	lightyellow1	orange4	rosybrown3	steelblue3
azure2	chartreuse4	darkslategray1	gray14	gray72	khaki3	lightyellow	orange3	rosybrown2	steelblue2
azure1	chartreuse3	darkslategray	gray13	gray71	khaki2	lightsteelblue4	orange2	rosybrown1	steelblue1
azure	chartreuse2	darkslateblue	gray12	gray70	khaki1	lightsteelblue3	orange1	rosybrown	steelblue
aquamarine4	chartreuse1	darkseagreen4	gray11	gray69	khaki	lightsteelblue2	orange	red4	springgreen4
aquamarine3	chartreuse	darkseagreen3	gray10	gray68	ivory4	lightsteelblue1	olivedrab4	red3	springgreen3
aquamarine2	cadetblue4	darkseagreen2	gray9	gray67	ivory3	lightsteelblue	olivedrab3	red2	springgreen2
aquamarine1	cadetblue3	darkseagreen1	gray8	gray66	ivory2	lightslategrey	olivedrab2	red1	springgreen1
aquamarine	cadetblue2	darkseagreen	gray7	gray65	ivory1	lightslategray	olivedrab1	red	springgreen
antiquewhite4	cadetblue1	darksalmon	gray6	gray64	ivory	lightslateblue	olivedrab	purple4	snow4
antiquewhite3	cadetblue	darkred	gray5	gray63	indianred4	lightskyblue4	oldlace	purple3	snow3
antiquewhite2	burlywood4	darkorchid4	gray4	gray62	indianred3	lightskyblue3	navyblue	purple2	snow2
antiquewhite1	burlywood3	darkorchid3	gray3	gray61		lightskyblue2	navy	purple1	snow1
antiquewhite	burlywood2	darkorchid3 darkorchid2			indianred2	lightskyblue1			
			gray2	gray60	indianred1		navajowhite4	purple	snow
aliceblue	burlywood1	darkorchid1	gray1	gray59	indianred	lightskyblue	navajowhite3	powderblue	slategrey
white	burlywood	darkorchid	gray0	gray58	hotpink4	lightseagreen	navajowhite2	plum4	slategray4

#### R font table

	plain		
URWTimes	NimbusRom		
	URWPalladio		
	CenturySch		
	NimbusSanCond		
URWHelvetica	NimbusSan		
	NimbusMon		
	URWBookman		
	URWGothic		
	Palatino		
	NewCenturySchoolbook		
	Helvetica-Narrow		
	Bookman		
	AvantGarde		
serif	Times		
sans	Helvetica		
mono	Courier		