Guide for R beginners

- We encourage running this script in R studio, which can be freely downloaded at https://rstudio.com/ (https://rstudio.com/).
- The first lines of the script download packages that contain specialized data analysis tools. The first time you run the script, you will need to install these packages to your computer. You can use the following code:

```
install.packages("tidyverse")
```

  This command will then install the tidyverse package (and let you know if any errors occured during installation). You'll need to repeat this command for each package listed in lines 1-15. After you install packages, you can use the library(tidyverse) commands to load each package.

- To run code in R studio, click on a line of code and push Ctrl+Enter (PC) or Command+Enter (Mac). The code will show up in the console, which will also indicate whether there were any errors (errors generally show up as red text).
- The first step is to load your data into R using the read.csv command. However, R may not know where to look for your .csv file. To tell R which file folder you've to look in, use the following command:

```
setwd("C:/User/File Path")
```

  This command sets your working directory (wd), so any output files will also be saved to this file folder.

- You've probably noticed two different kinds of line in the R script. The first are commands, which perform functions on the data:

```
data <- read.csv("histone_ratios_basic.csv", skip = 1)
```

  The second are comments, which give you instructions on how to run the script:

```
# Reading in all the data. You'll need to change the filename to whatever matches your
file
```

  However, you will not run all of the commands in this script on any single dataset. So, if there is a command that isn't applicable to your dataset, you can turn it into a comment so R knows not to run it:

```
# data <- data %>%
#   filter(!str_detect(Peptide, "H33"))
```

  You may also need to uncomment a line if it is applicable to your data but is commented in the orginal script. You can comment or uncomment lines using by selecting lines and clicking Code -> Comment/ Uncomment Lines, or using the hotkey Ctrl+Shift+C (PC) or Command+Shift+C (Mac)

- The first command should create a dataframe (called "data"), which will show up in your Environment.You can click on this dataframe to open it and inspect it in R. If you want to open it in Excel instead, you can always export a dataframe as a .csv file using the following command:

```
write.csv(name_of_dataframe, "filename.csv")
```

As you continue to run commands, the dataframe will change. You can always reopen a dataframe to see what happens after you run a command.

- The R script takes you through data quality checks, filtering, normalization, calculations, and statistical analysis. Comments should give you instructions on where you'll need to make decisions or add information about your samples/experimental design. Good luck!

P.S. If you're interested in learning more about R, our lab highly recommends the free online textbook "R for Data Science": https://r4ds.had.co.nz/ (https://r4ds.had.co.nz/)!