1. Open Difference\_in\_Two\_Columns.R script.
2. On line 11, change the path of the .csv file in which you wish to compare two columns for unique sequences. On line 14, change the name of the .csv file.

Macintosh HD:Users:legalllab:Desktop:Screen Shot 2018-07-27 at 9.34.37 AM.png

Macintosh HD:Users:legalllab:Desktop:Screen Shot 2018-07-27 at 9.34.44 AM.png

1. On line 23, change the number to the column that the sequences are located in (the column in which you are wondering contains unique sequences).

Macintosh HD:Users:legalllab:Desktop:Screen Shot 2018-07-27 at 9.35.33 AM.png

1. On line 29, change the number to the column that the surface peptides are located in (the column with the sequences you have already put in your “database”).

Macintosh HD:Users:legalllab:Desktop:Screen Shot 2018-07-27 at 9.35.38 AM.png

1. On line 39, change the path and the name of the .csv file in which you want the output of the unique sequences stored.

Macintosh HD:Users:legalllab:Desktop:Screen Shot 2018-07-27 at 9.35.49 AM.png

1. Run the program by clicking the “Run” button on the top right of the script in Rstudio. If you want to run each line individually on a Mac, type “cmd-return” after each line of code. On a Window computer, you need to type “Ctrl-Enter” after each line of code.