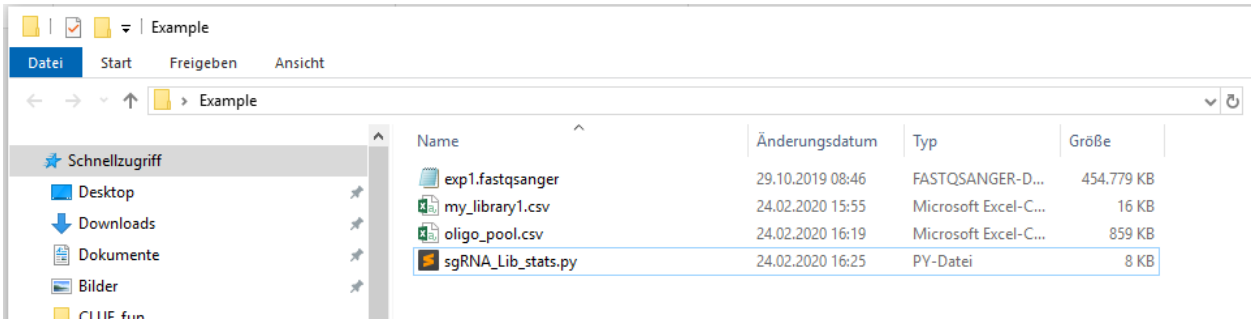
Analyzing NGS data from sequencing sgRNA libraries cloned with CRISPR Clue, using *sgRNA_Lib_stats.py*

This protocol provides a detailed description on how to analyze NGS data obtained from sequencing a sgRNA Library and wanting to know some quality statistics of your cloning. The aim of this script is to tell you how many reads map to your specific sgRNA library, to any other library of your oligo-pool or to the empty vector or could not be mapped at all. This description assumes you have installed Python and all dependencies necessary, as described in “Installing Python on your machine”.

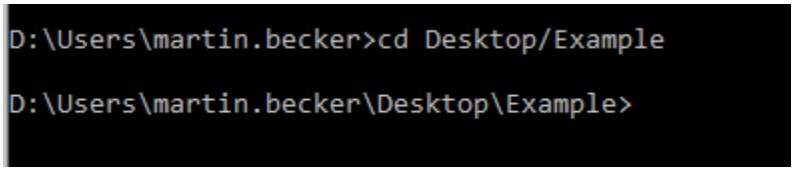
This Python script allows you to map NGS reads to your oligo-pool and to a specific sgRNA sub-library thereof. This will tell you about the specificity of your cloning and how well your PCRs and sequencing worked. To do so, the script requires a number of inputs:

Flag	Description	Example
-fq **--fastq**	File name of the fastq file containing your NGS data	exp1.fastq
-l **--library**	File name of your sgRNA library or oligo pool you want to map your NGS reads to	my_library1.csv
-p **--pool**	File name of your oligo-pool	ligo_pool.csv
-v **--vector**	Determines the sgRNA vector system you are using (either H1 or U6 promoter containing vectors from the Jeremias or Zhang labs, respectively)	H1 or U6
-o **--output**	Name of the output file for the statistics/read counts	stats_output

To get started, place all files you need together with the Python script in the same folder.



Now, open the command line by simply typing “cmd” into your search bar. Then navigate to your folder by typing “cd path/to/your/folder” and hitting enter.



So far, so good. Now it is time to run the script, which is as simple as typing the following:

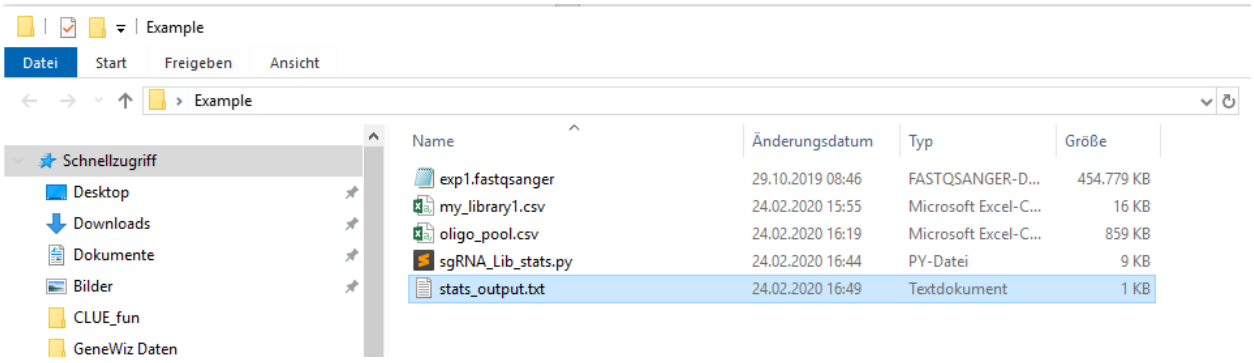
```
D:\Users\martin.becker\Desktop\Example>python sgRNA_Lib_stats.py -fq exp1.fastqsanger
-l my_library1.csv -p oligo_pool.csv -v H1 -o stats_output
```

Hit enter and wait for the script to finish.

```
D:\Users\martin.becker\Desktop\Example>python sgRNA_Lib_stats.py -fq exp1.fastqsanger
-l my_library1.csv -p oligo_pool.csv -v H1 -o stats_output
--------Script finished--------

D:\Users\martin.becker\Desktop\Example>
```

By now you will find the output file with the statistics and read counts in your folder.



Congratulations! You just analyzed your NGS data and mapped it to your sgRNA library and the oligo pool.