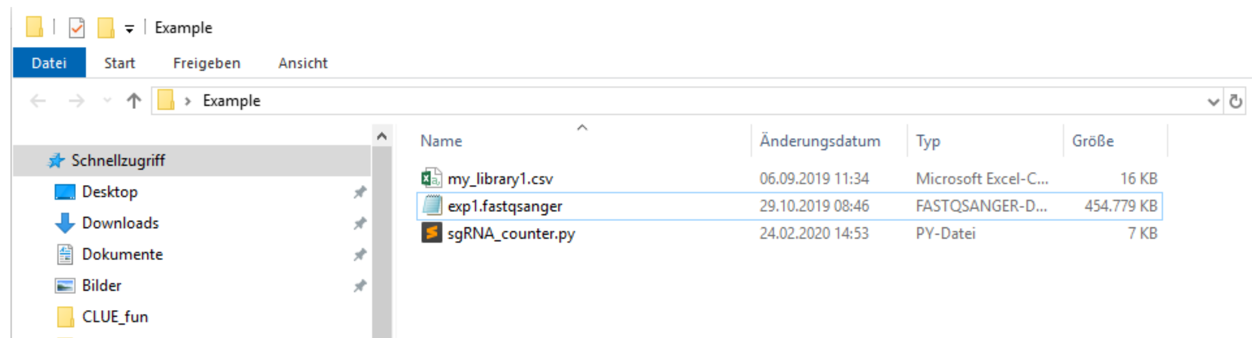
Analyzing NGS data from sequencing sgRNA libraries cloned with CRISPR Clue, using *sgRNA_counter.py*

This protocol provides a detailed description on how to analyze NGS data obtained from sequencing either sgRNA libraries or oligo pools cloned into the TOPO vector, using the CRISPR Clue pipeline. 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 a sgRNA library or Clue oligo pool of choice. It will generate a read count table as well as density and bar plots for you, showing you the distribution of sgRNAs. To do so, the script requires a number of inputs:

Flag	Description	Example
-fq **--fastq**	File name of the fastq file(s) containing your NGS data, you can enter several fastq file names separated by commas if they are to be mapped to the same library	exp1.fastq,exp2.fastq
-l **--library**	File name of your sgRNA library or oligo pool you want to map your NGS reads to	my_library1.csv
-t **--type**	Type of library you are mapping to (sgRNA sub-library or oligo pool)	sub-lib or oligo-pool
-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 read count table that will be generated. Note that all plots will be named as the fastq file	Exp1_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.

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

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

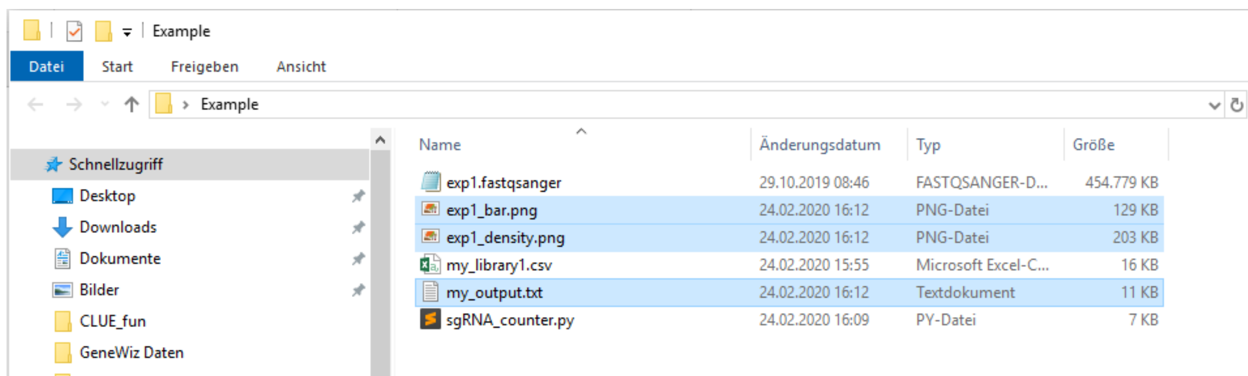
So far, so good. Now it is time to run the script, which is as simple as typing the following:

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

D:\Users\martin.becker\Desktop\Example>python sgRNA_counter.py -fq exp1.fastqsanger
-l my_library1.csv -t sub-lib -v H1 -o my_output
```

Hit enter and wait for the script to finish.

```
D:\Users\martin.becker\Desktop\Example>python sgRNA_counter.py -fq exp1.fastqsanger
-l my_library1.csv -t sub-lib -v H1 -o my_output
>>>sgRNA Dictionary generated<<<
>>>Bar Graph generated<<<
>>>Density Plot generated<<<
>>>Read count table generated<<<
--------Script finished--------

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

By now you will find several new files in your folder, which contain the results of your analysis.



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