

CRISPRclean[®] Single Cell RNA Boost Kit

This document describes how to run Cell Ranger to observe the benefits of CRISPRclean depletion on single cell sequencing. The user can download our example datasets from our website or use their own experimental data. 10X Genomics has an excellent set of [instructions](#) for novice users of Cell Ranger.

1. Downsample the fastq files to ensure that the same number of reads are being compared between the control and depleted samples. Below are the commands to execute, with the green highlighted portion being the read number for sample with the lowest number of reads. The yellow highlighted command is for zipped files (remove .gz for non-zipped)

- `mkdir ds`
- Check the lowest number of reads to downsample:
`for i in `ls *.fastq.gz` ; do echo $(zcat ${i} | wc -l)/4|bc; done`
- `ls *.fastq.gz | parallel -j 12 `seqtk sample -s1000 {} 1000000 > ds/{}.fastq``

2. Check to make sure the downsampling was successful

- `cd ds`
- `for i in `ls *.fastq.gz` ; do echo $(cat ${i} | wc -l)/4|bc; done`

3. Gzip the fastq files (if not already zipped)

- `ls *.fastq | parallel -j 25 `gzip {}``

4. Make individual folders for each condition and move the corresponding R1 and R2 files in that folder.

- `mkdir Depleted`
- `mkdir Normal`
- `mv *cleaned* ./Depleted/.`
- `mv *normal* ./normal/.`

5. Now run cell ranger count on each folder

- `cd Depleted`
- `cellranger count --id=output_folder --fastqs=/path/to/fastqs --transcriptome= path/to/masked_transcriptome_index`
- Where,
- `id=` the output directory
- `fastqs=` folder where fastqs are present
- `transcriptome=` masked transcriptomic index