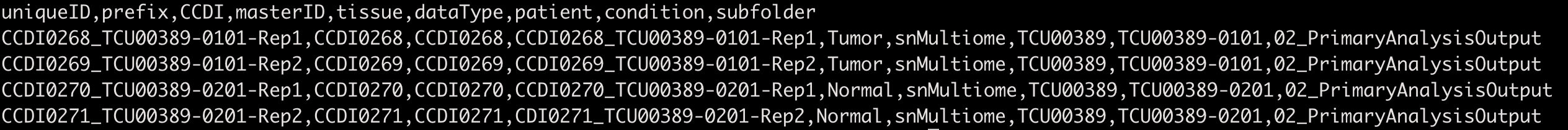
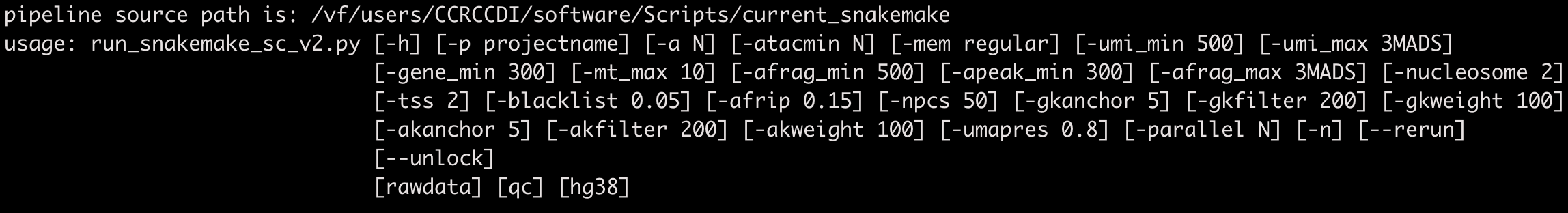
**snMultiome Pipeline Instruction Version 2**

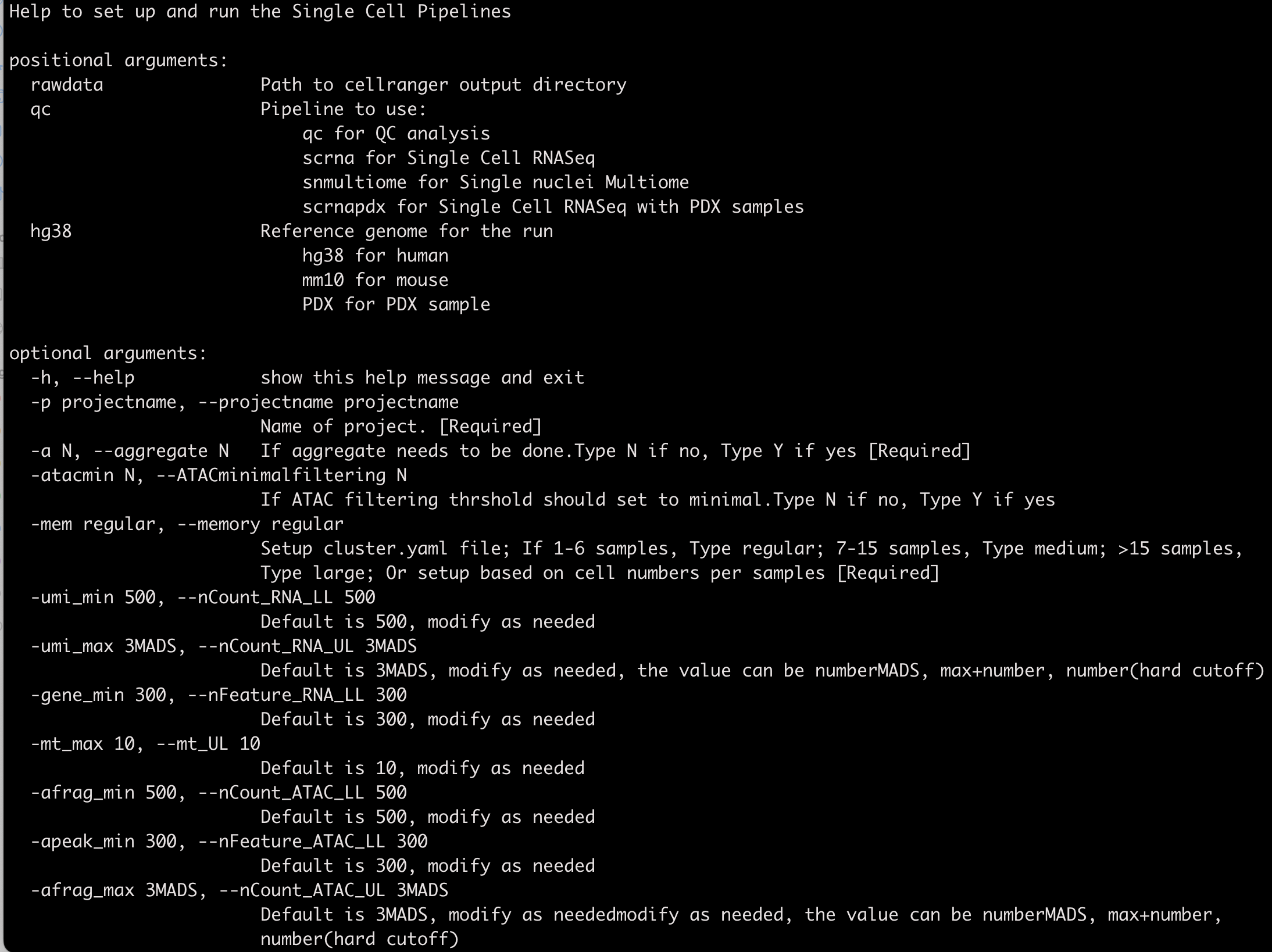
1. snMultiome Pipeline is saved in directory of “/data/CCRCCDI/software/Scripts/current\_snakemake/snMultiome\_split\_v2”
   1. config folder/cluster.yaml: the cluster.yaml file (other config files can go into here in the future)
   2. Snakefile: snakemake rules
   3. Submit.sh: for run the pipeline.
   4. scripts: R and Python scripts used in the pipeline.
2. Please create ANALYSIS\_DIR and put the manifest data in the folder ANALYSIS\_DIR/ assets/input\_manifest\_cellranger.csv before running the pipeline

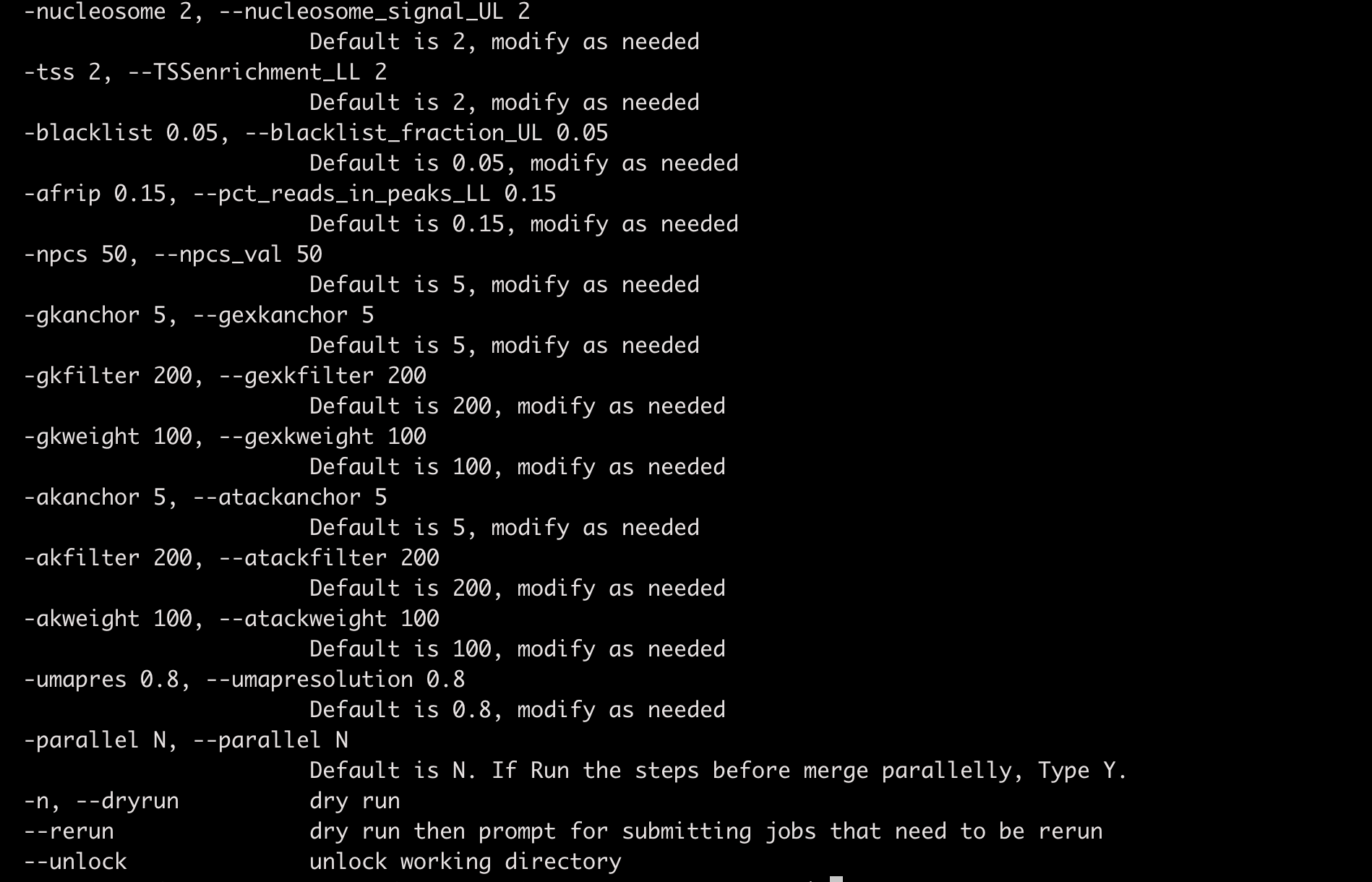


Folder structure: the wrapper will check the folder of rawdata/cellranger/folder/subfolder for cellranger.tar files and make softlink

1. Pipeline wrapper is in “/data/CCRCCDI/software/Scripts/current\_snakemake/run\_snakemake\_sc\_v2.py”
   1. usage: run\_snakemake\_sc.py -p projectname -a N -atacmin Y rawdata/cellranger/folder snmultiome hg38
   2. ./run\_snakemake\_sc.py --rerun to rerun the pipeline
   3. ./run\_snakemake\_sc.py --unlock to unlock the working directory







1. mkdir ANALYSIS\_DIR and cd ANALYSIS\_DIR
2. export PATH=$PATH:/data/CCRCCDI/software/Scripts/current\_snakemake
3. run the wrapper:
   1. example: run\_snakemake\_sc\_v2.py -p ccrccdi4 -a N -atacmin Y -mem regular -umapres 0.2 -parallel Y /data/CCRCCDI/rawdata/ccrccdi4 snmultiome hg38
   2. It will 1) create a folder (00\_FullCellrangerOutputs) with all the softlink to the cellranger output

2) create config.py file used in the Snakemake pipeline

3) copy config folder which includes cluster.yaml for cluster configuration

4) create Aggregate.csv file which will be used for cellranger aggregate step

5) create logs and stats folder to host the log file

6) copy the script folder for the pipeline

* 1. **The pipeline will not start right away**. Please check if the softlink is correct and consistent with the sample manifest in assets folder. Modify the config.py and config/cluster.yaml files if needed in terms of parameter setup.
  2. If everything set up correctly, using **./run\_snakemake\_sc.py --rerun** to start the pipeline

1. The results will be in the following folder
   1. cellranger\_output: untar file of the cellranger tar file
   2. cellranger\_finalreport: cellranger summary file (project.xlsx) combining the cellranger summary file for all the sample, the html report file of all the samples are also copied here
   3. Aggregate: results from cellranger-arc aggr
   4. result\_wSoupX: result from analysis using Seurat
      1. seurat: object.list.rds file from each step
      2. seurat/preprocessed: Seurat object for each sample at each step
      3. qc\_result\_plot: qc results for pre and post-filtering steps
      4. sample\_result\_plot: qc results for each sample
      5. integration: Seurat object from each integration steps
      6. integration\_plot: plots at the integration steps