# User Guide – SnapperRocks

## Requirements

SnapperRocks requires that Nextflow v20+ be installed to run, which is compatible with Linux and OSX.

Compute environments are provided in Singularity containers for local and cluster deployment. For AWS deployment, instances may be created from the containers.

## Running

SnapperRocks may be launched from the command line, using the Nextflow binary and the pipeline’s main.nf file. Any changes to the default values (see table below) may be specified with a `--` prefix and then the value. For Boolean parameters, a true value may be indicated with just the `--` prefix. Defaults may be permanently changed by editing the nextflow.config file in the path of SnapperRocks’ main.nf.

E.g.:

nextflow /path/to/SnapperRocks/main.nf --run\_id “run\_71” --fastq “/ftp/ingest/\*\_R{1,2}.fastq.gz" --trim --taxoprofile --assembly --typing --nesoni --cluster --executor “slurm”

At a minimum, the values that change between runs should be specified on the command line (e.g. `run\_id`, and depending on data ingest and handling `fastq` and/or `results`)

Parameters that will stay the same between runs may be updated in the nextflow.config or left as command line arguments and stored as a batch file for record keeping.

## Default Parameters

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Default** | **Description** |
| run\_id | " SR\_Run " | Run ID prefix tag |
| fastq | "./Raw/\*\_R{1,2}.fastq.gz" | Glob for input files. Accepts fastq and fastq.gz |
| store\_path | "/data/Compiled\_tables" | Filesystem location to store any required files for future processing |
| results | "./Results" | Output directory for results |
| executor | "local" | Executor to use. See NextFlow documentation for valid executor values |
| container | "singularity" | Container engine to use. Choice of ‘singularity’ and ‘docker’ |
| db | "/SnapperRocks/db" | Filesystem location for SnapperRocks database files. Default is correct value for provided DB containers. |
| scripts | "/SnapperRocks/scripts" | Filesystem location for SnapperRocks script files Default is correct value for provided DB containers |
| trim | false | Perform read trimming |
| fastqc | false | Perform FastQC analysis on samples |
| filterqc | false | Filters out samples that fail QC for further pipeline analysis |
| taxoprofile | false | Perform taxonomic profiling |
| metadata | “false” | Path to CSV containing metadata for samples. Use default value if not available |
| assembly | false | Perform de novo assembly |
| qc | false | Compile MultiQC reports for each sample and entire run |
| quast | false | Perform QUAST analysis on assembled samples. Requires `assembly` |
| amr | false | Perform AMRFinderPlus analysis. Requires `assembly` and `taxoprofile` |
| post | false | Perform post-assembly analysis. Requires `assembly` |
| typing | true | Perform sequence typing. Requires `assembly` |
| redcap | false | Perform Redcap metadata query |
| resprofile | false | Perform resistance profiling |
| intrinsic | false | Perform intrinsic analysis. Requires `assembly` |
| nesoni | false | Perform Nesoni SNP calling. Requires `taxoprofile` |
| pseudo | false | Perform pseudo-genome analysis. Requires `nesoni` |
| cluster | false | Perform clustering. Required for CATHAI. Requires `taxoprofile`, `typing` and `nesoni` |
| staph | false | Perform Staphylococcus analysis. Requires `taxoprofile` |
| njobs | 4 | Number of jobs to run concurrently |
| redcap\_key | "" | API key for Redcap access |
| ramdisk | false | If true, will emplace the Kraken2 database into a ramdisk and run it from there. Provides performance increase when used with krakenbulk otherwise decreases performance |
| ramdisk\_path | "/dev/shm" | Filesystem location for ramdisk |