Setup

Create environment and activate it

conda activate base

mamba create -c conda-forge -c bioconda -n name snakemake

conda activate name

Rule syntax

rule rule.name:

input:

“input.one”

“input.two”

output:

“output.one”

“output.two”

shell:

#python string containing shell commands to execute

#here you can refer to inputs or outputs put putting them in curly brackets

#example: “bwa mem {input} | samtools view -Sb - > {output}”

#when multiple inputs, they are concatenated with white space in between

script:

#path and filename to custom script

#python scripts: refer to inputs as: snakemake.input[num]

#remember that python starts with 0 as the first one

#r scripts: call with snakemake@input[[num]]

#starts with one here, because it’s in R

#or do snakemake@input[["filename"]]

Commands

conda activate snakemake #enter snakemake environment

Ctrl+C #exit running snakemake workflow (kills all runs)

conda deactivate #deactivate active snakemake environment

snakemake --help #see available tools

snakemake directory/file

#tells snakemake to generate the target files (this can be a rule output)

#options

-n #dry run (show execution plan without actually running)

-p #print the resulting shell command

--cores number #specify number of cores to use

--dag #visualize the DAG

#pipe into dot -Tsvg > dag.svg to save the dag visualization

expand(“directory/{something}.bam”, something=NAME) #helper funct for collecting input files

Best practice

#at top of file

rule all:

input:

"directory/finaltargetfile"