snakemake_TP

Introduction to workflow with Snakemake

Memo

Official documentation: https://snakemake.readthedocs.io/en/stable/

To date (2023/07/04, 170 standardized workflows are available: https://snakemake.github.io/snakemake-workflow-catalog/), in average over the last 6 months, 10 standardized workflows are added per month.

Have a look there to be FAIR !

Pre-requisites

1- Data: input data to run the workflow example are reduced RNASeq reads files from *Ostreococcus tauri* green algae ((with a focus on chr18) from runs SRR3099585-87 & SRR3105697-99, Bioproject PRJNA304086). To download data from zenodo here, the easiest way is to click on the download icon, but it would download data on our local computer. However, we wish to use these data on Cloud. So we just collect the link to data by doing a right-click. The link is "https://zenodo.org/record/3997237/files/FAIR_Bioinfo_data.tar.gz?download=1"

2- Snakemake: for the run of snakemake, deploy a BioPipes VM from Biosphere IFB Cloud (https://biosphere.france-bioinformatique.fr/cloud/)

```
$ ssh ubuntu@my.IP
The authenticity of host '193.49.167.73 (193.49.167.73)' can't be established.
ED25519 key fingerprint is SHA256:ek4Sqedg2Uhlict4wBE9WmsoqRBuxVTirRm2lEaMGrY.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])?

Type "yes"
```

Note : Replace "my.IP" by the IP address provided on your biosphere VM

2-1 Activate the environment to use Snakemake and check that the 'snakemake' program works by asking for the version used.

```
$ conda activate snakemake
(snakemake)$ snakemake --version
7.30.1
```

2-2 I. The architecture of your working directory

Create directories to get this architecture:

```
├── data -> /ifb/data
├── logs
├── results
└── workflow
```

Note: On Ubuntu, `mkdir` command does the job.

```
(snakemake)$ mkdir logs results workflow
```

At the end of the practice, the architecture should look like this:

```
├── data-> /ifb/data
       ├──  mydatalocal
             ├──  Data
                 ├──   SRR3099585_chr18.fastq.gz
                 ├──   SRR3099586_chr18.fastq.gz
                 ├──   ...
                 ├──   SRR3105699_chr18.fastq.gz
                 ├──   O.tauri_annotation.gff
                 ├──   O.tauri_genome.fna
        ├──  public
├── logs
├── results
├── workflow
       ├──  envs
            ├──  qc.yaml
       ├──  Snakefile.ex1.smk
       ├──  Snakefile.ex2.smk
       ├──  ...
```

Conda environment based on the qc.yml file:

Create a new file in workflow directory as follow,

```
name: qc # conda environment name
channels:
    - defaults
    - bioconda
    - conda-forge
dependencies:
    - fastqc=0.11.9 # quality check of fastq data (java)
    - multiqc=1.13 # reports aggregation (R package)
```

Data management

```
# Go to data/mydatalocal
(snakemake)$ cd ~/data/mydatalocal
(snakemake)$ wget "https://zenodo.org/record/3997237/files/FAIR_Bioinfo_data.tar.gz?
        download=1"
(snakemake)$ mv 'FAIR_Bioinfo_data.tar.gz?download=1' FAIR_Bioinfo_data.tar.gz
(snakemake)$ tar xvzf FAIR_Bioinfo_data.tar.gz
(snakemake)$ ls Data
O.tauri_annotation.gff  SRR3099585_chr18.fastq.gz  SRR3099587_chr18.fastq.gz
        SRR3105698_chr18.fastq.gz
O.tauri_genome.fna      SRR3099586_chr18.fastq.gz  SRR3105697_chr18.fastq.gz
        SRR3105699_chr18.fastq.gz
```

Workflow definition

1. The Snakefile example

The final objective is to create a `Snakefile` to manage a small workflow with 2 steps: i)

fastqc ii) multiqc

These two tools belonging to the bioinformatics domain allow to check the quality of high throughput sequence data. They are accessible via a Conda environment, `qc.yml`

note: If you have already run this notebook, you may need to run:

```
(snakemake)$ rm -Rf results/FastQC results/multiqc_data results/multiqc_report.html
```

2. Objective 1: Understanding the rule concept

Create a snakemake file named `Snakefile.ex1.smk` including the first step of the RNAseq workflow (the reads quality checking thank to the fastqc tool) on one of the RNAseq files

fastqc access: through a conda environment (see prerequisites on top)

```
fastqc command line: fastqc --outdir results/FastQC inputFileName
inputFileName: SRR3099585_chr18.fastq.gz in the ${PWD}/Data directory
outputfiles produced in outdir:
The 2 result files (*_fastqc.zip & *_fastqc.html) will be located in your outdirectory and
named based on the prefix of input file (eg. SRR3099585_chr18_fastqc.zip)
```

Create the Snakefile

```
# Go to workflow directory and create a new file called `Snakefile_ex1.smk`
(snakemake)$ cd ~/workflow
(snakemake)$ touch Snakefile.ex1.smk
# edit the file with `nano` editor for example:

rule fastqc:
    input:
        "data/mydatalocal/Data/SRR3099585_chr18.fastq.gz"
    output:
        "results/FastQC/SRR3099585_chr18_fastqc.zip",
        "results/FastQC/SRR3099585_chr18_fastqc.html"
    conda:
        envs/qc.yaml
    shell: "fastqc --outdir results/FastQC/ {input}"
```

Then execute snakemake:

```
(snakemake)$ pwd
/home/ubuntu
(snakemake)$ snakemake --snakefile workflow/Snakefile.ex1.smk --cores 1 --use-conda
Building DAG of jobs...
Creating conda environment workflow/envs/qc.yaml...
Downloading and installing remote packages.
Environment for /home/ubuntu/workflow/envs/qc.yaml created (location:
        .snakemake/conda/1cfd55d9dd88f128b691aad4280f97b3_)
Using shell: /usr/bin/bash
Provided cores: 1 (use --cores to define parallelism)
Rules claiming more threads will be scaled down.
Job stats:
job       count    min threads    max threads
------  -------  -------------  -------------
all           1              1              1
fastqc        1              1              1
total         2              1              1

Select jobs to execute...

rule fastqc:
    input: data/mydatalocal/Data/SRR3099585_chr18.fastq.gz
    output: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html
    jobid: 1
```

```
    reason: Missing output files: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html
    resources: tmpdir=/tmp

Activating conda environment: .snakemake/conda/1cfd55d9dd88f128b691aad4280f97b3_
Started analysis of SRR3099585_chr18.fastq.gz
Approx 5% complete for SRR3099585_chr18.fastq.gz
Approx 10% complete for SRR3099585_chr18.fastq.gz
...
Analysis complete for SRR3099585_chr18.fastq.gz

Finished job 1.
1 of 2 steps (50%) done
Select jobs to execute...

Finished job 1.
1 of 2 steps (50%) done
Select jobs to execute...

localrule all:
    input: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html
    jobid: 0
    reason: Input files updated by another job:
        results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html
    resources: tmpdir=/tmp

Finished job 0.
2 of 2 steps (100%) done
Complete log: .snakemake/log/2023-07-06T095157.460406.snakemake.log
```

3. Objective 2: One rule, 2 input files

Add a second input RNAseq file to the rule, for exemple Data/SRR3099586_chr18.fastq.gz

Don’t forget to add the cognate output files in the Snakefile !

Create a new Snakefile Snakefile.ex2.smk containing:

```
rule fastqc:
  output:
    "FastQC/SRR3099585_chr18_fastqc.zip",
    "FastQC/SRR3099585_chr18_fastqc.html",
    "FastQC/SRR3099586_chr18_fastqc.zip",
    "FastQC/SRR3099586_chr18_fastqc.html"
  input:
    "Data/SRR3099585_chr18.fastq.gz",
    "Data/SRR3099586_chr18.fastq.gz"
  shell: "fastqc --outdir FastQC/ {input}"
```

Then execute snakemake:

```
(snakemake)$ pwd
/home/ubuntu
(snakemake)$ snakemake --snakefile workflow/Snakefile.ex2.smk --cores 1 --use-conda

Building DAG of jobs...
Using shell: /usr/bin/bash
Provided cores: 1 (use --cores to define parallelism)
Rules claiming more threads will be scaled down.
Job stats:
job       count    min threads    max threads
------  -------  -------------  -------------
all           1              1              1
fastqc        1              1              1
```

```
total         2              1              1

Select jobs to execute...

rule fastqc:
    input: data/mydatalocal/Data/SRR3099585_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099586_chr18.fastq.gz
    output: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.html
    jobid: 1
    reason: Missing output files: results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.zip; Set of input files has changed since
        last execution
    resources: tmpdir=/tmp

Activating conda environment: .snakemake/conda/1cfd55d9dd88f128b691aad4280f97b3_

Finished job 1.
1 of 2 steps (50%) done
Select jobs to execute...

[Thu Jul  6 10:08:14 2023]
localrule all:
    input: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.html
    jobid: 0
    reason: Input files updated by another job:
        results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.html
    resources: tmpdir=/tmp

[Thu Jul  6 10:08:14 2023]
Finished job 0.
2 of 2 steps (100%) done
Complete log: .snakemake/log/2023-07-06T100804.103180.snakemake.log
```

4. Objective 3: Manage all the RNAseq files

Boring with writing all input and output file names ?

Hint:

Use the expand() function to manage all the input RNAseq files at once.

Create a Python list at the begining of the snakefile and containing all the basename of the input files (don’t include the .fastq.gz suffix).

```
Python list format: list_name = ["item1", "item2", ..., "itemN"]
Replace the filename lists of the input and output directives by the `expand()` function.
```

Create another Snakefile called Snakefile.ex3.smk containing:

```
SAMPLES=["SRR3099585_chr18","SRR3099586_chr18","SRR3099587_chr18", "SRR3105697_chr18",
"SRR3105698_chr18", "SRR3105699_chr18"]

rule fastqc:
  output:
        expand("FastQC/{sample}_fastqc.zip", sample=SAMPLES),
        expand("FastQC/{sample}_fastqc.html", sample=SAMPLES)
  input:
        expand("data/{sample}.fastq.gz", sample=SAMPLES)
  shell: "fastqc --outdir FastQC {input}"
```

Then execute snakemake:

```
(snakemake)$ pwd
/home/ubuntu
(snakemake)$ snakemake --snakefile workflow/Snakefile.ex3.smk --cores 1 --use-conda

Building DAG of jobs...
Using shell: /usr/bin/bash
Provided cores: 1 (use --cores to define parallelism)
Rules claiming more threads will be scaled down.
Job stats:
job       count    min threads    max threads
------  -------  -------------  -------------
all           1              1              1
fastqc        1              1              1
total         2              1              1

Select jobs to execute...

[Thu Jul  6 11:15:19 2023]
rule fastqc:
    input: data/mydatalocal/Data/SRR3099585_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099586_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099587_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105697_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105698_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105699_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099585_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099586_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099587_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105697_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105698_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105699_chr18.fastq.gz
    output: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3105697_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3105699_chr18_fastqc.html
    jobid: 1
    reason: Missing output files: results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3105697_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.html; Set of input files has changed since
        last execution
    resources: tmpdir=/tmp

Activating conda environment: .snakemake/conda/1cfd55d9dd88f128b691aad4280f97b3_
Started analysis of SRR3099585_chr18.fastq.gz
Approx 5% complete for SRR3099585_chr18.fastq.gz
...
Analysis complete for SRR3099585_chr18.fastq.gz
Started analysis of SRR3099586_chr18.fastq.gz
...
Finished job 1.
1 of 2 steps (50%) done
Select jobs to execute...

localrule all:
```

```
    input: results/FastQC/SRR3099585_chr18_fastqc.zip,
         results/FastQC/SRR3099586_chr18_fastqc.zip,
         results/FastQC/SRR3099587_chr18_fastqc.zip,
         results/FastQC/SRR3105697_chr18_fastqc.zip,
         results/FastQC/SRR3105698_chr18_fastqc.zip,
         results/FastQC/SRR3105699_chr18_fastqc.zip,
         results/FastQC/SRR3099585_chr18_fastqc.html,
         results/FastQC/SRR3099586_chr18_fastqc.html,
         results/FastQC/SRR3099587_chr18_fastqc.html,
         results/FastQC/SRR3105697_chr18_fastqc.html,
         results/FastQC/SRR3105698_chr18_fastqc.html,
         results/FastQC/SRR3105699_chr18_fastqc.html
    jobid: 0
    reason: Input files updated by another job:
         results/FastQC/SRR3105697_chr18_fastqc.zip,
         results/FastQC/SRR3099587_chr18_fastqc.zip,
         results/FastQC/SRR3105697_chr18_fastqc.html,
         results/FastQC/SRR3105699_chr18_fastqc.zip,
         results/FastQC/SRR3099585_chr18_fastqc.html,
         results/FastQC/SRR3105698_chr18_fastqc.html,
         results/FastQC/SRR3099587_chr18_fastqc.html,
         results/FastQC/SRR3099586_chr18_fastqc.html,
         results/FastQC/SRR3105698_chr18_fastqc.zip,
         results/FastQC/SRR3099585_chr18_fastqc.zip,
         results/FastQC/SRR3099586_chr18_fastqc.zip,
         results/FastQC/SRR3105699_chr18_fastqc.html
    resources: tmpdir=/tmp

Finished job 0.
2 of 2 steps (100%) done
```

5. Objective 4: Add a second step

With a second tool, it starts becoming an analysis workflow!

The second tool, multiqc will aggregate all the fastqc results.

```
multiqc command line : multiqc *_fastqc.zip
inputs: the fastqc zip files
2 outputs: a file multiqc_report.html & a repository multiqc_data (to manage with
directory("multiqc_data"))
```

Create another Snakefile called Snakefile.ex4.smk:

```
SAMPLES=["SRR3099585_chr18", "SRR3099586_chr18", "SRR3099587_chr18", "SRR3105697_chr18",
"SRR3105698_chr18", "SRR3105699_chr18"]

rule all:
    input:
       expand("results/FastQC/{sample}_fastqc.html", sample=SAMPLES),
       "results/multiqc/multiqc_report.html"

rule fastqc:
   input:
         expand("data/mydatalocal/Data/{sample}.fastq.gz", sample=SAMPLES)
   output:
         expand("results/FastQC/{sample}_fastqc.zip", sample=SAMPLES),
         expand("results/FastQC/{sample}_fastqc.html", sample=SAMPLES)
   conda:
        "envs/qc.yaml"
    shell: "fastqc --outdir results/FastQC/ {input}"

rule multiqc:
   input:
       expand("results/FastQC/{sample}_fastqc.zip", sample = SAMPLES)
   output:
        "results/multiqc/multiqc_report.html"
   conda:
```

```
        "envs/qc.yaml"
    shell:
        "multiqc --outdir results/multiqc {input}"
```

Then execute snakemake:

```
(snakemake)$ pwd
/home/ubuntu
(snakemake)$ snakemake --snakefile workflow/Snakefile.ex4.smk --cores 1 --use-conda

Building DAG of jobs...
Creating conda environment workflow/envs/qc.yaml...
Downloading and installing remote packages.
Environment for /home/ubuntu/workflow/envs/qc.yaml created (location:
        .snakemake/conda/efe5c6f7911ea7c09bf27b3f1015bb80_)
Using shell: /usr/bin/bash
Provided cores: 1 (use --cores to define parallelism)
Rules claiming more threads will be scaled down.
Job stats:
job        count    min threads    max threads
-------  -------  -------------  -------------
all            1              1              1
fastqc         1              1              1
multiqc        1              1              1
total          3              1              1

Select jobs to execute...

...

rule fastqc:
    input: data/mydatalocal/Data/SRR3099585_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099586_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3099587_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105697_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105698_chr18.fastq.gz,
        data/mydatalocal/Data/SRR3105699_chr18.fastq.gz
    output: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3105697_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3105699_chr18_fastqc.html
    jobid: 1
    reason: Missing output files: results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3105697_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099585_chr18_fastqc.zip
    resources: tmpdir=/tmp

Activating conda environment: .snakemake/conda/efe5c6f7911ea7c09bf27b3f1015bb80_

...
Finished job 1.
1 of 3 steps (33%) done
```

```
Select jobs to execute...

rule multiqc:
    input: results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip
    output: results/multiqc/multiqc_report.html
    jobid: 2
    reason: Missing output files: results/multiqc/multiqc_report.html; Input files updated
        by another job: results/FastQC/SRR3105697_chr18_fastqc.zip,
        results/FastQC/SRR3099585_chr18_fastqc.zip,
        results/FastQC/SRR3105699_chr18_fastqc.zip,
        results/FastQC/SRR3105698_chr18_fastqc.zip,
        results/FastQC/SRR3099587_chr18_fastqc.zip,
        results/FastQC/SRR3099586_chr18_fastqc.zip
    resources: tmpdir=/tmp

Activating conda environment: .snakemake/conda/efe5c6f7911ea7c09bf27b3f1015bb80_

  /// MultiQC 🔍 | v1.14

|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3099585_chr18_fastqc.zip
|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3099586_chr18_fastqc.zip
|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3099587_chr18_fastqc.zip
|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3105697_chr18_fastqc.zip
|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3105698_chr18_fastqc.zip
|           multiqc | Search path :
        /home/ubuntu/results/FastQC/SRR3105699_chr18_fastqc.zip
|         searching | ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 100% 6/6
|            fastqc | Found 6 reports
|           multiqc | Compressing plot data
|           multiqc | Report      : results/multiqc/multiqc_report.html
|           multiqc | Data        : results/multiqc/multiqc_data
|           multiqc | MultiQC complete

Finished job 2.
2 of 3 steps (67%) done
Select jobs to execute...

localrule all:
    input: results/FastQC/SRR3099585_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3105697_chr18_fastqc.html,
        results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3105699_chr18_fastqc.html, results/multiqc/multiqc_report.html
    jobid: 0
    reason: Input files updated by another job:
        results/FastQC/SRR3105698_chr18_fastqc.html,
        results/FastQC/SRR3105697_chr18_fastqc.html, results/multiqc/multiqc_report.html,
        results/FastQC/SRR3105699_chr18_fastqc.html,
        results/FastQC/SRR3099586_chr18_fastqc.html,
        results/FastQC/SRR3099587_chr18_fastqc.html,
        results/FastQC/SRR3099585_chr18_fastqc.html
    resources: tmpdir=/tmp

Finished job 0.
3 of 3 steps (100%) done
Complete log: .snakemake/log/2023-07-06T140543.494495.snakemake.log
```

6. Extra objective, the log file

Objective 6: In Unix systems, the output of a command is usually sent to 2 separate streams: the expected output to Standard Out (stdout, or ”>”), and the error messages to Standard Error (stderr, or ”2>”). To integrate stderr and stdout into the same log, use ”&>” (use it with care because output files are often printed to stdout).

Hint: Redirect the stdout and stderr streams of the fastqc and multiqc rules by adding a ”log:” directive with two variables, out and err to separately redirect each streams.

```
log:
  std="Logs/{sample}_fastqc.std",
  err="Logs/{sample}_fastqc.err"
shell: "fastqc --outdir FastQC/ {input} 1>{log.std} 2>{log.err}"

log:
  std="Logs/multiqc.std",
  err="Logs/multiqc.err"
shell: "multiqc {input} 1>{log.std} 2>{log.err}"
```