

## ## Comparative analysis of **Fastqc** using **Nextflow** vs **Bash** script

### 1. Bash script

{ CODE BLOCK

```
#!/bin/bash
```

```
# Define input and output directories
```

```
INPUT_DIR="/home/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files"
```

```
OUTPUT_DIR="/home/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files  
/output_bash_fastqc"
```

```
# Create the output directory if it doesn't exist
```

```
mkdir -p "$OUTPUT_DIR"
```

```
# Loop through all .fastq.gz files in the input directory
```

```
for file in "$INPUT_DIR"/*.fastq.gz
```

```
do
```

```
    echo "Running FastQC on $file..."
```

```
    fastqc "$file" --outdir "$OUTPUT_DIR"
```

```
done
```

```
echo "All FastQC analyses completed!"
```

```
}
```

```
Approx 55% complete for SRR33317478.fastq.gz  
Approx 60% complete for SRR33317478.fastq.gz  
Approx 65% complete for SRR33317478.fastq.gz  
Approx 70% complete for SRR33317478.fastq.gz  
Approx 75% complete for SRR33317478.fastq.gz  
Approx 80% complete for SRR33317478.fastq.gz  
Approx 85% complete for SRR33317478.fastq.gz  
Approx 90% complete for SRR33317478.fastq.gz  
Approx 95% complete for SRR33317478.fastq.gz  
Analysis complete for SRR33317478.fastq.gz  
All FastQC analyses completed!  
[ble: elapsed 42.864s (CPU 117.5%)] bash fastqc_bash.sh
```

**Bash takes 42 seconds to complete analysis**

## 2. Nextflow script

[CODE BLOCK]

```
#!/usr/bin/env nextflow

params.reads =
'/home/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files/*.fastq.gz'
params.outdir =
'/home/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files/output_fastqc'

process FastQC {
    publishDir params.outdir, mode: 'copy'

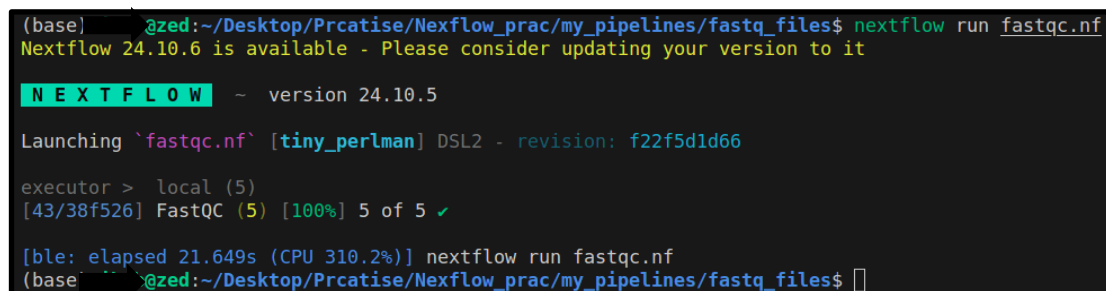
    input:
    path read_file

    output:
    path "*.html"
    path "*.zip"

    script:
    """
    fastqc $read_file
    """
}

workflow {
    Channel
    .fromPath(params.reads)
    .set { read_files }

    FastQC(read_files)
}
]
```



A terminal window showing the execution of a Nextflow script. The prompt is (base) @zed:~/Desktop/Prcatise/Nexflow\_prac/my\_pipelines/fastq\_files\$. The command nextflow run fastqc.nf is entered. The output shows Nextflow 24.10.6 is available - Please consider updating your version to it. Below that, a green bar displays NEXTFLOW ~ version 24.10.5. Then, it says Launching `fastqc.nf` [tiny\_perlman] DSL2 - revision: f22f5d1d66. The executor is local (5). A progress bar shows [43/38f526] FastQC (5) [100%] 5 of 5 ✓. Finally, it shows [ble: elapsed 21.649s (CPU 310.2%)] nextflow run fastqc.nf. The prompt returns to (base) @zed:~/Desktop/Prcatise/Nexflow\_prac/my\_pipelines/fastq\_files\$.

```
(base) @zed:~/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files$ nextflow run fastqc.nf
Nextflow 24.10.6 is available - Please consider updating your version to it

NEXTFLOW ~ version 24.10.5

Launching `fastqc.nf` [tiny_perlman] DSL2 - revision: f22f5d1d66

executor > local (5)
[43/38f526] FastQC (5) [100%] 5 of 5 ✓

[ble: elapsed 21.649s (CPU 310.2%)] nextflow run fastqc.nf
(base) @zed:~/Desktop/Prcatise/Nexflow_prac/my_pipelines/fastq_files$
```

**Nextflow take only 21 seconds to finish task**