

## Fastqc on HPC cluster:

### *What is Fastqc?*

FastQC aims to provide a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines. It provides a modular set of analyses which you can use to give a quick impression of whether your data has any problems of which you should be aware before doing any further analysis.

The documentation for the Fastqc is found in the following links:

[Documentation](#)

[GITHUB](#)

### Versions Available:

Fastqc– 0.11.5

### *How to load a version of Fastqc?*

To load a version of Fastqc on the HPC, use the following command:

```
module avail bio/fastqc
```

The version will be listed. To use a version of software, use following command:

```
module load bio/fastqc/0.11.5
```

Verify by using this command:

```
module list
```

The loaded software and dependencies will be shown.

## *How to use Fastqc on the cluster?*

There are three ways to run Fastqc on the cluster:

GUI method:

To run fastqc in graphical interface, users should jump into a compute node and run fastqc:

```
#Jump to a compute node
srun -p main --qos main -n 1 -c 6 --mem 6G --pty bash

# start fastqc graphical interface
fastqc
```

Note: User should connect to a server with X11 display forwarding to run a graphical interface. Use `-X` or `-Y` flag while connecting to the server through SSH.

```
ssh -X userid@uahpc.ua.edu
```

Interactive method:

To run the program interactively, follow the steps:

```
#Open a bash session on compute node
srun -p main --qos main -n 1 -c 12 --mem 10G --pty bash

#load the module
```

```
module load bio/fastqc/0.11.5

# start your commands here
fastqc -h
# This method is ideal for short runs which produces runtime output
and to debug the codes
```

## The Script:

To run a slurm job, the user must prepare input files. For this example, get input files with,

```
#download fastq files through this command
curl -o test2.fastq
https://zenodo.org/record/3736457/files/1_control_18S_2019_minq7.fastq
?download=1

curl -L test.fastq
https://zenodo.org/record/3736457/files/1_control_18S_2019_minq7%20-
%20Copy.fastq?download=1

# make a sbatch script
touch script.sbatch
```

Use the following template for the script,

```
#!/bin/bash
#SBATCH -J fastqc-test
#SBATCH -n 1
#SBATCH -mem=10G
#SBATCH -p main
#SBATCH --qos main # Quality of service
#SBATCH -o Fastqc_sim_out-%J.txt
#SBATCH -e Fastqc_sim_out-%J.txt
#SBATCH -t 30

module load bio/fastqc # load the module

cd $SLURM_SUBMIT_DIR # Moving to the test directory
```

```
#run fastqc
fastqc test.fastq test2.fastq
```

Schedule the job with the following sbatch command.

```
sbatch myscript.sh
```

All the processed files will be generated in the same directory as the sbatch script.

### ***Where to find help?***

If you are stuck on some part or need help at any point, please contact OIT at the following address.

<https://ua-app01.ua.edu/researchComputingPortal/public/oitHelp>