8/4/24, 10:10 AM OneNote

## Installation of nextflow on new Ubuntu with conda package manager

```
31 July 2024 12:58

    1. Make sure you install the latest version of Java.

              sudo apt install default-jre
        Use this command to confirm the Java Version - it should be above 11.
              java -version
    3. Follow nextflow.io to install NextFlow on Linux machine.
              curl -s https://get.nextflow.io | bash
              chmod +x nextflow
             sudo mv nextflow /usr/local/bin

    4. Use nextflow info to confirm the installation of NextFlow

              nextflow info
  5. Installing Conda - Ref and then add channels Bioconda and conda-forge
         a. <a href="https://www.anaconda.com/download/success">https://www.anaconda.com/download/success</a>
              wget <a href="https://repo.anaconda.com/archive/Anaconda3-2024.06-1-Linux-x86_64.sh">https://repo.anaconda.com/archive/Anaconda3-2024.06-1-Linux-x86_64.sh</a>
              Chmod +x /home/talktoamitdhir/Downloads/Anaconda3-2024.06-1-Linux-x8o 64.sh
              /home/talktoamitdhir/Downloads/Anaconda3-2024.06-1-Linux-x86 64.sh
              Yes
              Yes [when init is asked]
              /home/talktoamitdhir/anaconda3//bin/conda init
              conda config -set auto_activate_base false
  6. confirm if Conda is installed.
               (base) talktoamitdhir@talktoamitdhir-MS-7788:~$ conda
              conda -h
  7. Set up BiConda after installation of conda Ref
              conda config --add channels bioconda
              conda config --add channels conda-forge
             https://bioconda.github.io/
  8. Confirm Conda channels Ref
             Conda config -show channels
   9. Set up conda environment and create new conda Env.
              conda env list
              conda create -n bioinformatics
             conda activate bioinformatics
  10. Install samtools in new conda environment.
             conda install -c bioconda samtools
              conda install -c bioconda fastqc
              conda install -c bioconda nf-core
■ 11. Update Nextflow.
             nextflow self-update [updates and fixed the java issue]
  12. Use following command to instantiate the git repo
              eval $(ssh-agent -s)
              ssh-add ubuntu
   13. Create a new FastQSheet.csv file.
    14. Go to working directory and run following command
              sudo nextflow run nf-core/rnaseq --input FastQSheet.csv --outdir "/" --genome
 ■ 15. Follow these instructions for installation of the RNASeq.
             https://nf-co.re/docs/usage/installation/
              https://nf-co.re/rnaseq/3.14.0/
              https://nf-co.re/docs/usage/introduction#how-to-run-a-pipeline
             https://nf-co.re/docs/usage/configuration#custom-configuration-files
         Helpful Links
     1. Sample Data fastQ files - Link.
        Wget command sample to download the files on Linux.
        Use realpath command to get the path to the folder structure
               alktoamitdhir@talktoamitdhir-MS-7788:~/NextFlowProject$ realpath FastO_Data/
              home/talktoamitdhir/NextFlowProject/FastQ_Data
         IMPORTANT - Nfcore video link
```

```
EC2 Instance Connect Session Manager SSH client EC2 serial console

Instance ID

☐ i-04603ea46e2540408 (preludeNF02)

1. Open an SSH client.
```

8/4/24, 10:10 AM OneNote

- 2. Locate your private key file. The key used to launch this instance is preludeNF02.pem
- 3. Run this command, if necessary, to ensure your key is not publicly viewable.
  - chmod 400 "preludeNF02.pem"
- 4. Connect to your instance using its Public DNS:
  - d ec2-54-210-50-85.compute-1.amazonaws.com

## Example:

- ssh -i "preludeNF02.pem" ubuntu@ec2-54-210-50-85.compute-1.amazonaws.com
- 6. **Note:** In most cases, the guessed username is correct. However, read your AMI usage instructions to check if the AMI owner has changed the default AMI username.

```
https://d-90678d2a22.awsapps.com/start/#/console?account_id=
 896191365710&role_name=Sridhar_Raidi
 896191365710
 https://us-east-1.signin.aws/platform/login?workflowStateHandle=
 87b75226-fabd-4b90-8b64-7163e0c7da14
 Connection Type SSH using putty and ppk file
 EC2 Public IP Address: 54.210.50.85
 Port 22
 use ppk as authentication file.
 Connection Type terminal using pem file
 Go to /c/Users folder
 ssh -i hopde/OneDrive/Desktop/aws/preludeNF02.pem
 ubuntu@ec2-54-210-50-85.compute-1.amazonaws.com
Code to login to EC2
ssh -I hopde/OneDrive/Desktop/aws/preludeNF02.pem ubuntu@ec2-54-210-50-85.compute-
1.amazon.aws.com
Code to copy from local to EC2
scp -i ~/preludeNF02.pem local image file ubuntu@ec2-54-210-50-85.compute-
1.amazon.aws.com:/home/user/images/
App.pluto.bio
MultiQC
```

## Requirement

https://github.com/nf-core/rnaseg/tree/3.14.0

## Workflow: Fastg Files to Final Algorithm Processing

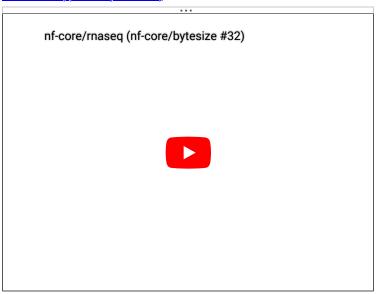
- 1. Start with Fastq Files:
  - Input: Raw <u>Fastq</u> files containing the RNA sequencing data.
- 2. NF-Core RNA-Seq Pipeline:

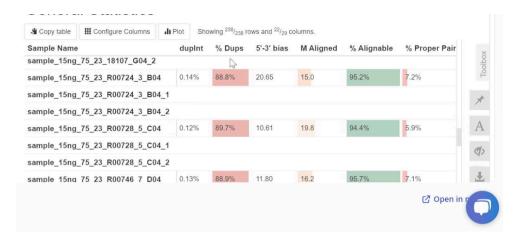
(Public Pipeline) <a href="https://nf-co.re/docs/nf-core-tools/installation">https://nf-co.re/docs/nf-core-tools/installation</a>

- QC and Alignment:
  - Use the NF-Core RNA-Seq pipeline to process the Fastq files. This pipeline will handle alignment of reads and will QC data.
- o This Output will be generated: Aligned

reads (BAM files) and QC reports.

https://nf-co.re/events/2022/bytesize-32-nf-core-rnaseq\_nf-core/rnaseq\_(nf-core/bytesize #32)





8/4/24, 10:10 AM OneNote