

# ~ NEXTFLOW ~

## Initiation Genologin

Devoir de l'étudiant : **Gatépé Cédoine KODJOVI**

---

### I - Connection à Genologin

A travers nos nom de fleurs : dans mon cas « **arome** » : avec la commande :  
« **ssh -XY arome@genologin.toulouse.inrae.fr** »  
mot de passe : « **\*\*\*\*\*** » ,  
et on se retrouve dans la base du compte associé à notre nom de fleur.

### II – Création de dossier et de l'environnement de travail

- Création en 1<sup>er</sup> d'un dossier **TP\_Nextflow** : dans lequel j'ai recrée un nouveau dossier **data\_TP** qui va contenir mes fichiers d'entrée (inputs).
- Création d'un script de lancement : « **job\_tp.sh** »

```
#!/bin/bash
#SBATCH --time=1-0
#SBATCH -J GatepeKODJOVI
#SBATCH -p workq
#SBATCH -e errorjob.out
#SBATCH --mem=6G
#SBATCH --cpus-per-task=4
#SBATCH --mail-type=BEGIN,END,FAIL

module purge

module load bioinfo/nfcore-Nextflow-v21.04.1

nextflow run nf-core/rnaseq -r 3.4 -profile genotoul \
--input ./data_TP/samplesheet.csv \
--fasta ./data_TP/ITAG2.3_genomic_Ch6.fasta \
--gtf ./data_TP/ITAG2.3_genomic_Ch6.gtf \
```

suivi de son lancement avec la commande : « **sbatch job\_tp.sh** »

# III – Réponses aux questions

## 1) Explication de la sortie « seff » et Intérêt du « resume »:

### - seff :

la fonction seff permet d'avoir : une sortie résumée de l'état d'un job qu'on a envoyé sur un serveur ou cluster de calcul, dans mon cas sur **Genotoul**.

Pour avoir l'information d'un job on fait « seff id\_du\_job », si on ne dispose pas l'id du job on a un fichier slurm qui est créé au lancement de chaque job qui contient les informations sur le déroulement du job . Le nom du slurm est : **slurm-id\_du\_job.out** , en copiant et collant l'id devant seff on peut avoir les informations.

Dans mon cas j'ai : « seff 37379873 » :

```
arome@genologin1 ~/work/arome/TP_nextflow $ seff 37379873
Job ID: 37379873
Cluster: genobull
User/Group: arome/formation
State: COMPLETED (exit code 0)
Nodes: 1
Cores per node: 4
CPU Utilized: 00:02:52
CPU Efficiency: 9.56% of 00:30:00 core-walltime
Job Wall-clock time: 00:07:30
Memory Utilized: 1.83 GB
Memory Efficiency: 30.50% of 6.00 GB
arome@genologin1 ~/work/arome/TP_nextflow $
```

- Job ID: 37379873	# l'identifiant du job
- Cluster: genobull	# le cluster sur lequel le calcul est lancé
- User/Group: arome/formation	# l'utilisateur et le groupe auquel il appartient le propriétaire du job dans mon cas user = arome et group = formation
- State: COMPLETED (exit code 0)	# l'état du job : en cours ou en fin
- Nodes: 1	# le nœud sur lequel on est connecté pour se connecter au cluster
- Cores per node: 4	# le nombre de cœurs par nœud
- CPU Utilized: 00:02:52	# le temps d'utilisation du CPU
- CPU Efficiency: 9.56% of 00:30:00 core-walltime	# temps d'occupation du CPU et son efficacité
- Job Wall-clock time: 00:07:30	# la durée total du job sur le cluster
- Memory Utilized: 1.83 GB	# la mémoire utilisé par le job
- Memory Efficiency: 30.50% of 6.00 GB	# taux d'occupation de la mémoire allouée

### - resume

la fonction ou le paramètre « resume » qu'on rajoute à la commande « sbatch nom-fichier.sh » permet de ne plus relancer le job depuis le début et de recréer tous les dossiers ou fichiers intermédiaires.

Le job prends moi de temps et on évite de créer des duplications de fichiers et/ou qui pourraient générés des conflits et donc arrêt du processus/job.

Dans mon cas j'ai lancé sans le resume , j'aurai pu le faire mais vu qu'au début j'avais des erreurs dans mes fichiers initiaux d'entrée j'ai préféré à chaque fois supprimé les sorties créés systématiquement pour chacune des tentatives inachevées.

**PS : Récupération des fichiers généré** sur son ordinateur en local :

le téléchargement des fichiers results ou n'importe lequel est possible grâce à la commande scp qui permet de copier depuis le serveur vers la machine local :

« **scp serveur+identifiant\_de\_connexion : chemin\_vers\_le\_fichier  
chemin\_vers\_ledossier\_qui\_va\_receptionner** »

Dans mon cas j'ai la commande

«**scp**

**arome@genologin.toulouse.inrae.fr:~/work/arome/TP\_nextflow/results/multiqc/star\_salmon/  
multiqc\_data/\* /home/guest/Bureau/multiqc/** » .

## 2) MultiQC

Dans le dossier Nextflow créer dans mon cas je l'ai appelé TP\_Nextflow , au sein de ce dossier la fin du job a généré 2 dossiers : **results** et **work** ; 2 fichiers: **slurm-37379873.out** et **errorjob.out**. Au sein du dossier results nous avons : **fastqc**, **genome**, **multiqc**, **pipeline\_info**, **star\_salmon**, **trimgalore** .

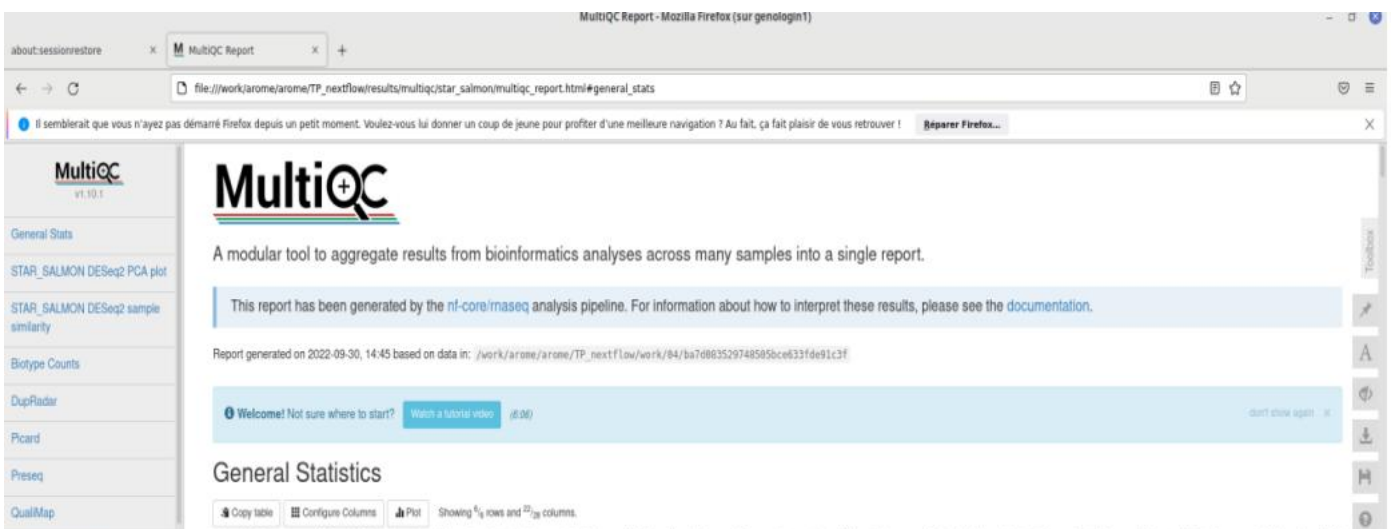
A cette étape nous parlerons de Multiqc.

Dans le répertoire multiqc j'ai un autre répertoire star\_salmon dans lequel se trouve un autre répertoire **multiqc\_data** et un fichier **multiqc\_report.html** qui ouvert nous envois direct sur une page web sur le quel on retrouve toutes les sorties des sous parties et logiciels qui ont été lancé quand on a mis en marche notre Nextflow.

**Remarque** : Multiqc est un outil de création de rapports qui analyse les statistiques récapitulatives des résultats et des fichiers journaux générés par d'autres outils bio-informatiques. Multiqc n'exécute pas d'autres outils, il est conçu pour être placé à la fin des pipelines d'analyse ou pour être exécuté manuellement lorsque vous avez terminé d'exécuter vos outils.

Lorsque nous avons lancé notre job rnaseq , il a eu lancement programmé de multiqc qui est ira récupérer de manière récursive dans tous les chemins de fichiers qui lui ont été fournis et trouvera les fichiers qu'il reconnaît. Il analysera les informations pertinentes à partir de celles-ci et générera un seul fichier de rapport HTML autonome c'est le : **multiqc\_report.html**.

La page se présente comme suit.



Multiqc enregistre également un répertoire de fichiers avec toutes les données analysées pour une utilisation ultérieure en aval : le répertoire **multiqc\_data**.

Et dans ce répertoire nous avons des fichiers .txt, .log, .json ci contre : au nombre de 16. les fichiers avec une extension .txt peuvent être convertis en .csv pour une visualisation en forme de table qui peut être plus facile à visualiser, comprendre et interpréter au besoin. Les fichiers .txt portent un nom qui peut renseigner sur le logiciel, l'outil et/ou le type de données, d'information qu'ils peuvent contenir.

```
total 1369
drwxr-xr-x 2 arome formation 4096 30 sept. 14:45 multiqc_data
-rw-r--r-- 1 arome formation 1398516 30 sept. 14:45 multiqc_report.html
arome@genologin1 ~/work/arome/TP_nextflow/results/multiqc/star_salmon $
arome@genologin1 ~/work/arome/TP_nextflow/results/multiqc/star_salmon/
total 991
-rw-r--r-- 1 arome formation 438 30 sept. 14:45 multiqc_cutadapt.txt
-rw-r--r-- 1 arome formation 974700 30 sept. 14:45 multiqc_data.json
-rw-r--r-- 1 arome formation 1240 30 sept. 14:45 multiqc_fastqc_1.txt
-rw-r--r-- 1 arome formation 1177 30 sept. 14:45 multiqc_fastqc.txt
-rw-r--r-- 1 arome formation 2670 30 sept. 14:45 multiqc_general_stat
-rw-r--r-- 1 arome formation 20658 30 sept. 14:45 multiqc.log
-rw-r--r-- 1 arome formation 417 30 sept. 14:45 multiqc_picard_dups.
-rw-r--r-- 1 arome formation 622 30 sept. 14:45 multiqc_rseqc_bam_st
-rw-r--r-- 1 arome formation 104 30 sept. 14:45 multiqc_rseqc_infer
-rw-r--r-- 1 arome formation 732 30 sept. 14:45 multiqc_rseqc_juncti
-rw-r--r-- 1 arome formation 1736 30 sept. 14:45 multiqc_rseqc_read d
-rw-r--r-- 1 arome formation 1088 30 sept. 14:45 multiqc_samtools fla
-rw-r--r-- 1 arome formation 100 30 sept. 14:45 multiqc_samtools_idx
-rw-r--r-- 1 arome formation 1796 30 sept. 14:45 multiqc_samtools sta
```

### 3) Interprétation des principaux résultats :

Comme précédemment décrits à la fin du job nous avons 2 dossiers créés : results et work ; il y a 2 fichiers qui sont aussi créés « slurm-37379873.out » et « errorjob.out » .


#### - errorjob.out :

Ce fichier est censé contenir les éventuels erreurs au cas où la tentative aurait échoué. J'ai pu remarquer qu'il fonctionne et recueille bien les erreurs lors de mes précédentes tentatives non abouties.

#### - slurm-37379873.out :

On peut retrouver dans ce fichier des informations concernant le job mais aussi les fichiers d'entrée et les paramètres qui accompagnent le job , la version du NEXTFLOW utilisé , le temps que prendra le job , la mémoire et le CPU qu'il pourrait occuper au max.

```
arome@genologin1:~/work
Fichier  Edition  Affichage  Rechercher  Terminal  Onglets  Aide
arome@genologin1:~/work/arome/TP_nextflow
N E X T F L O W  -  version 21.04.1
Launching 'nf-core/rnaseq' [festering_meitner] - revision: 964425e3fd [3.4]
WARN: Found unexpected parameters:
* --igenomesIgnore: true
- Ignore this warning: params.schema_ignore_params = "igenomesIgnore"

-----
NF-CORE 
nf-core/rnaseq v3.4
-----
Core Nextflow options
revision           : 3.4
runName            : festering_meitner
containerEngine    : singularity
launchDir          : /work/arome/arome/TP_nextflow
workDir            : /work/arome/arome/TP_nextflow/work
projectDir         : /home/arome/.nextflow/assets/nf-core/rnaseq
userName           : arome
profile            : genotoul
configFiles        : /home/arome/.nextflow/assets/nf-core/rnaseq/nextflow.conf

Input/output options
input              : ../data_TP/samplesheet.csv

Reference genome options
```



## - Répertoire work :

Ce répertoire contient d'autres sous répertoires : dans mon cas plus de 70, exactement 73. On se rend très vite compte que ces mêmes répertoires comportent d'autres répertoire jusqu'à ce qu'on tombe sur des fichier.yml, .txt ou .zip .

```
arome@genologin1 ~/work/arome/TP_nextflow/results $ cd ../work/
arome@genologin1 ~/work/arome/TP_nextflow/work $ ls
01 04 08 11 14 1c 24 3a 33 41 54 63 6e 7d 89 94 a0 a6 a9 ad b0 b2 b8 ba c0 c4 c7 ce d3 d9 df e7 e9 ed f5 fd tmp
02 09 0e 13 18 20 27 30 40 4c 56 6b 77 7f 8a 9c a4 a8 ac ae b1 b7 b9 be c1 c6 cd d1 d7 dd e1 e8 ea f0 fa ff
arome@genologin1 ~/work/arome/TP_nextflow/work $ cd 01
arome@genologin1 ~/work/arome/TP_nextflow/work/01 $ ls
7ac692ace55828397dfdf81c0ebcd7
arome@genologin1 ~/work/arome/TP_nextflow/work/01 $ cd 7ac692ace55828397dfdf81c0ebcd7/
arome@genologin1 ~/work/arome/TP_nextflow/work/01/7ac692ace55828397dfdf81c0ebcd7 $ ls
CONTROL_REP2.sorted.bam CONTROL_REP2.sorted.bam.bai CONTROL_REP2.sorted.bam.flagstat versions.yml
arome@genologin1 ~/work/arome/TP_nextflow/work/01/7ac692ace55828397dfdf81c0ebcd7 $
```

## - Répertoire results :

Contient les répertoires : fastqc, genome, multiqc, pipeline\_info, star\_salmon, trimgalore .  
Je ne parlerai plus de multiqc dans cette section.

- **pipeline\_info :**

Nous avons là un dossier qui contient des informations dans les fichiers au format html : des informations plus détaillées sur chaque sous partie ou analyse, leur durée et leur occupation mémoire et CPU.

```
arome@genologin1 ~/work/arome/TP_nextflow/results $ cd pipeline_info/
arome@genologin1 ~/work/arome/TP_nextflow/results/pipeline_info $ ls -l
total 3745
-rw-r--r-- 1 arome formation 3225761 30 sept. 14:45 execution_report_2022-09-30_14-38-43.html
-rw-r--r-- 1 arome formation 275579 30 sept. 14:45 execution_timeline_2022-09-30_14-38-43.html
-rw-r--r-- 1 arome formation 14109 30 sept. 14:45 execution_trace_2022-09-30_14-38-43.txt
-rw-r--r-- 1 arome formation 310057 30 sept. 14:45 pipeline_dag_2022-09-30_14-38-43.svg
-rw-r--r-- 1 arome formation 229 30 sept. 14:39 samplesheet.valid.csv
-rw-r--r-- 1 arome formation 1521 30 sept. 14:44 software_versions.yml
arome@genologin1 ~/work/arome/TP_nextflow/results/pipeline_info $ firefox execution_report_2022-09-30_14-38-43.html &
[1] 65028
```

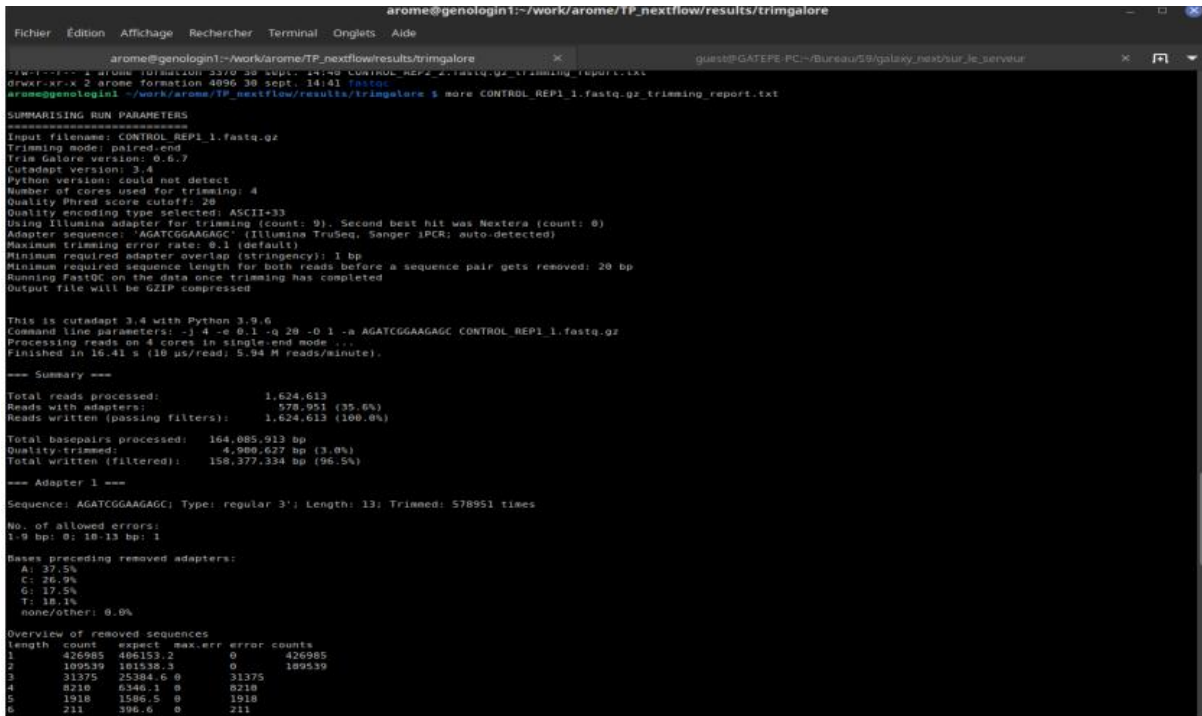
- **Genome :**

on retrouve des informations concernant le genome lui même, tel que sa longueur, les exons trouvés et a quel endroit .

```
arome@genologin1 ~/work/arome/TP_nextflow/results $ cd genome/
arome@genologin1 ~/work/arome/TP_nextflow/results/genome $ ls -l
total 2314
drwxr-xr-x 3 arome formation 4096 30 sept. 14:39 index
-rw-r--r-- 1 arome formation 320910 30 sept. 14:39 ITAG2.3_genomic_Ch6.bed
-rw-r--r-- 1 arome formation 29 30 sept. 14:39 ITAG2.3_genomic_Ch6.fasta.fai
-rw-r--r-- 1 arome formation 20 30 sept. 14:39 ITAG2.3_genomic_Ch6.fasta.sizes
-rw-r--r-- 1 arome formation 2034585 30 sept. 14:39 ITAG2.3_genomic_Ch6_genes.gtf
drwxr-xr-x 2 arome formation 4096 30 sept. 14:39 rsem
arome@genologin1 ~/work/arome/TP_nextflow/results/genome $ cd index/
arome@genologin1 ~/work/arome/TP_nextflow/results/genome/index $ ls -l
total 1
drwxr-xr-x 2 arome formation 4096 30 sept. 14:39 star
arome@genologin1 ~/work/arome/TP_nextflow/results/genome/index $ cd star/
arome@genologin1 ~/work/arome/TP_nextflow/results/genome/index/star $ ls -l
total 513035
-rw-r--r-- 1 arome formation 9 30 sept. 14:39 chrLength.txt
-rw-r--r-- 1 arome formation 20 30 sept. 14:39 chrNameLength.txt
-rw-r--r-- 1 arome formation 11 30 sept. 14:39 chrName.txt
-rw-r--r-- 1 arome formation 11 30 sept. 14:39 chrStart.txt
-rw-r--r-- 1 arome formation 396006 30 sept. 14:39 exonGeTrInfo.tab
-rw-r--r-- 1 arome formation 174928 30 sept. 14:39 exonInfo.tab
-rw-r--r-- 1 arome formation 53452 30 sept. 14:39 geneInfo.tab
-rw-r--r-- 1 arome formation 48320606 30 sept. 14:39 Genome
-rw-r--r-- 1 arome formation 627 30 sept. 14:39 genomeParameters.txt
-rw-r--r-- 1 arome formation 377398217 30 sept. 14:39 SA
-rw-r--r-- 1 arome formation 97867203 30 sept. 14:39 SAindex
-rw-r--r-- 1 arome formation 278238 30 sept. 14:39 sjdbInfo.txt
-rw-r--r-- 1 arome formation 332597 30 sept. 14:39 sjdbList.fromGTF.out.tab
-rw-r--r-- 1 arome formation 332535 30 sept. 14:39 sjdbList.out.tab
-rw-r--r-- 1 arome formation 153751 30 sept. 14:39 transcriptInfo.tab
arome@genologin1 ~/work/arome/TP_nextflow/results/genome/index/star $
```

- **Trimgalore :**

Contient les résultats de l'outil trimgalore : qui est une sorte d'enveloppe autour de Cutadapt et FastQC pour évaluer la qualité et effectuer le découpage de l'adaptateur sur les fichiers FastQ. On distingue des infos comme la séquence qui a été considéré comme adaptateur , le script de lancement, occurrence en pourcentage des bases précédant les adaptateurs retirés, etc.



```
arome@genologin1:~/work/arome/TP_nextflow/results/trimgalore
Fichier  Edition  Affichage  Rechercher  Terminal  Onglets  Aide
arome@genologin1:~/work/arome/TP_nextflow/results/trimgalore
guest@GATEPE-PC:~/Bureau/S9/galaxy_nextflow_le_serveur
arome@genologin1:~/work/arome/TP_nextflow/results/trimgalore $ more CONTROL_REP1_1.fastq.gz_trimming_report.txt
SUMMARISING RUN PARAMETERS
=====
Input filename: CONTROL_REP1_1.fastq.gz
Trimming mode: paired-end
Trim Galore version: 0.6.7
Cutadapt version: 3.4
Python version: could not detect
Number of cores used for trimming: 4
Quality Phred score cutoff: 20
Quality encoding type selected: ASCII-33
Using Illumina adapter for trimming (count: 0); Second best hit was Nextera (count: 0)
Adapter sequence: AGATCGGAAGAGC (Illumina TruSeq, Sanger iPCR; auto-detected)
Maximum trimming error rate: 0.1 (default)
Minimum required adapter overlap (stringency): 1 bp
Minimum required sequence length for both reads before a sequence pair gets removed: 20 bp
Running FastQC on the data once trimming has completed
Output file will be GZIP compressed

This is cutadapt 3.4 with Python 3.9.6
Command line parameters: -j 4 -e 0.1 -q 20 -0 1 -a AGATCGGAAGAGC CONTROL_REP1_1.fastq.gz
Processing reads on 4 cores in single-end mode ...
Finished in 16.41 s (10 us/read; 5.94 M reads/minute).

--- Summary ---
Total reads processed:      1,624,613
Reads with adapters:       578,951 (35.6%)
Reads written (passing filters): 1,624,613 (100.0%)

Total basepairs processed: 164,085,913 bp
Quality trimmed:          4,980,627 bp (3.0%)
Total written (filtered): 158,377,334 bp (96.5%)

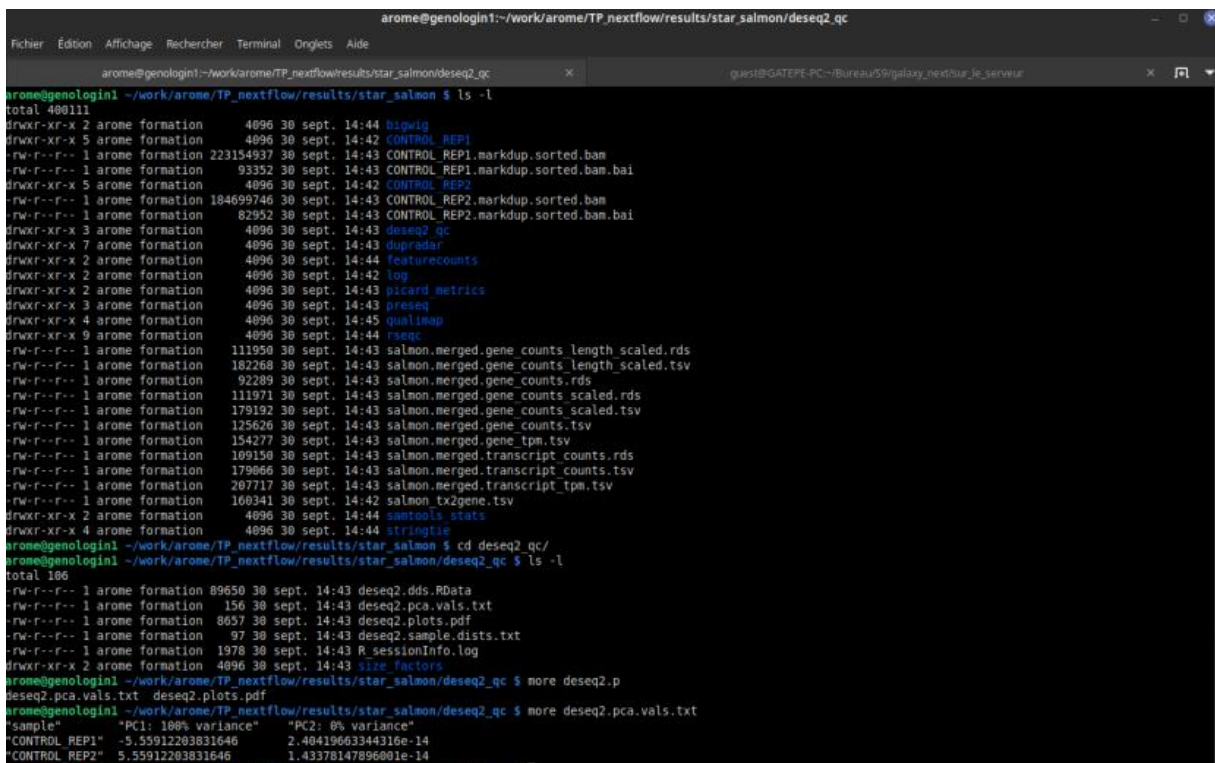
--- Adapter 1 ---
Sequence: AGATCGGAAGAGC; Type: regular 3'; Length: 13; Trimmed: 578951 times
No. of allowed errors:
1-9 bp: 0; 10-13 bp: 1

Bases preceding removed adapters:
A: 37.5%
C: 26.9%
G: 37.5%
T: 18.1%
none/other: 0.0%

Overview of removed sequences
length count expect max.err error counts
1 425985 460152.2 0 425985
2 109539 101538.3 0 109539
3 31375 25384.6 0 31375
4 8210 6346.1 0 8210
5 1918 1586.5 0 1918
6 211 396.6 0 211
```

- **Star\_salmon :**

On retrouve les résultats de salmon dans le répertoire mais aussi d'autres répertoires contenant les résultats d'autres outils (packages essentiellement) : deseq2, rseqc, dupradar etc .



```
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon/deseq2_qc
Fichier  Edition  Affichage  Rechercher  Terminal  Onglets  Aide
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon/deseq2_qc
guest@GATEPE-PC:~/Bureau/S9/galaxy_nextflow_le_serveur
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon $ ls -l
total 400111
drwxr-xr-x 2 arome formation 4096 30 sept. 14:44 bigwig
drwxr-xr-x 5 arome formation 4096 30 sept. 14:42 CONTROL_REP1
-rw-r--r-- 1 arome formation 223154937 30 sept. 14:43 CONTROL_REP1.markdup.sorted.bai
-rw-r--r-- 1 arome formation 93352 30 sept. 14:43 CONTROL_REP1.markdup.sorted.bai
drwxr-xr-x 5 arome formation 4096 30 sept. 14:42 CONTROL_REP2
-rw-r--r-- 1 arome formation 184699746 30 sept. 14:43 CONTROL_REP2.markdup.sorted.bai
-rw-r--r-- 1 arome formation 82932 30 sept. 14:43 CONTROL_REP2.markdup.sorted.bai
drwxr-xr-x 3 arome formation 4096 30 sept. 14:43 deseq2_qc
drwxr-xr-x 7 arome formation 4096 30 sept. 14:43 dupradar
drwxr-xr-x 2 arome formation 4096 30 sept. 14:44 featurecounts
drwxr-xr-x 2 arome formation 4096 30 sept. 14:42 log
drwxr-xr-x 2 arome formation 4096 30 sept. 14:43 picard_metrics
drwxr-xr-x 3 arome formation 4096 30 sept. 14:43 preseq
drwxr-xr-x 4 arome formation 4096 30 sept. 14:45 qualimap
drwxr-xr-x 9 arome formation 4096 30 sept. 14:44 rseqc
-rw-r--r-- 1 arome formation 111950 30 sept. 14:43 salmon.merged.gene_counts.length.scaled.rds
-rw-r--r-- 1 arome formation 182268 30 sept. 14:43 salmon.merged.gene_counts.length.scaled.tsv
-rw-r--r-- 1 arome formation 92289 30 sept. 14:43 salmon.merged.gene_counts.rds
-rw-r--r-- 1 arome formation 111971 30 sept. 14:43 salmon.merged.gene_counts.scaled.rds
-rw-r--r-- 1 arome formation 179192 30 sept. 14:43 salmon.merged.gene_counts.scaled.tsv
-rw-r--r-- 1 arome formation 125626 30 sept. 14:43 salmon.merged.gene_counts.tsv
-rw-r--r-- 1 arome formation 154277 30 sept. 14:43 salmon.merged.gene_tpm.tsv
-rw-r--r-- 1 arome formation 169150 30 sept. 14:43 salmon.merged.transcript_counts.rds
-rw-r--r-- 1 arome formation 179666 30 sept. 14:43 salmon.merged.transcript_counts.tsv
-rw-r--r-- 1 arome formation 207717 30 sept. 14:43 salmon.merged.transcript_tpm.tsv
-rw-r--r-- 1 arome formation 160341 30 sept. 14:42 salmon.tx2gene.tsv
drwxr-xr-x 2 arome formation 4096 30 sept. 14:44 smtools_stats
drwxr-xr-x 4 arome formation 4096 30 sept. 14:44 stringtie
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon $ cd deseq2_qc/
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon/deseq2_qc $ ls -l
total 106
-rw-r--r-- 1 arome formation 89650 30 sept. 14:43 deseq2.dds.RData
-rw-r--r-- 1 arome formation 156 30 sept. 14:43 deseq2.pca.vals.txt
-rw-r--r-- 1 arome formation 8657 30 sept. 14:43 deseq2.plots.pdf
-rw-r--r-- 1 arome formation 97 30 sept. 14:43 deseq2.sample.dists.txt
-rw-r--r-- 1 arome formation 1978 30 sept. 14:43 R_sessionInfo.log
drwxr-xr-x 2 arome formation 4096 30 sept. 14:43 size_factors
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon/deseq2_qc $ more deseq2.p
deseq2.pca.vals.txt deseq2.plots.pdf
arome@genologin1:~/work/arome/TP_nextflow/results/star_salmon/deseq2_qc $ more deseq2.pca.vals.txt
"sample" "FC1: 100% variance" "FC2: 0% variance"
"CONTROL_REP1" 5.55912203831646 2.40419663344316e-14
"CONTROL_REP2" 5.55912203831646 1.43378147896001e-14
```

- **Fastqc :**

Ce répertoire contient les résultats de fastqc qui est un outil de contrôle de qualité . La particularité de fastqc est qu'il peut repérer les problèmes qui proviennent soit du séquenceur, soit du matériel de départ de la bibliothèque.

Dans notre cas, j'ai 4 fichier « .html » et leur version « .zip » ; correspondant chacun à mes 4 reads de départ donc le dossier contient 8 fichiers au total. A l'ouverture des html on a des informations sur la séquence , la qualité le long du genome

