

Thesis, COLLÉGIALITÉ

Auteur : p211718

Promoteur(s) : Taminiau, Bernard

Faculté : Faculté de Médecine

Diplôme : Master en sciences biomédicales, à finalité approfondie

Année académique : 2022-2023

URI/URL : <http://hdl.handle.net/2268.2/17839>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

I. Lignes de commandes pour Unicycler et Trycycler :

A. Unicycler :

```
unicycler -1 short_reads_1.fastq.gz -2 short_reads_2.fastq.gz -l long_reads.fastq.gz -o output_dir
```

B. Trycycler :

```
trycycler subsample --reads reads.fastq --out_dir read_subsets
```

```
threads=16 # change as appropriate for your system → 28
```

```
mkdir assemblies
```

```
flye --nano-hq read_subsets/sample_01.fastq --threads "$threads" --out-dir assembly_01 && cp  
assembly_01/assembly.fasta assemblies/assembly_01.fasta && rm -r assembly_01
```

```
miniasm_and_minipolish.sh read_subsets/sample_02.fastq "$threads" > assembly_02.gfa &&
```

```
any2fasta assembly_02.gfa > assemblies/assembly_02.fasta && rm assembly_02.gfa
```

```
raven --threads "$threads" read_subsets/sample_03.fastq > assemblies/assembly_03.fasta && rm  
raven.cereal
```

```
flye --nano-hq read_subsets/sample_04.fastq --threads "$threads" --out-dir assembly_04 && cp  
assembly_04/assembly.fasta assemblies/assembly_04.fasta && rm -r assembly_04
```

```
miniasm_and_minipolish.sh read_subsets/sample_05.fastq "$threads" > assembly_05.gfa &&
```

```
any2fasta assembly_05.gfa > assemblies/assembly_05.fasta && rm assembly_05.gfa
```

```
raven --threads "$threads" read_subsets/sample_06.fastq > assemblies/assembly_06.fasta && rm  
raven.cereal
```

```
flye --nano-hq read_subsets/sample_07.fastq --threads "$threads" --out-dir assembly_07 && cp  
assembly_07/assembly.fasta assemblies/assembly_07.fasta && rm -r assembly_07
```

```
miniasm_and_minipolish.sh read_subsets/sample_08.fastq "$threads" > assembly_08.gfa &&
```

```
any2fasta assembly_08.gfa > assemblies/assembly_08.fasta && rm assembly_08.gfa
```

```
raven --threads "$threads" read_subsets/sample_09.fastq > assemblies/assembly_09.fasta && rm  
raven.cereal
```

```
flye --nano-raw read_subsets/sample_10.fastq --threads "$threads" --out-dir assembly_10 && cp  
assembly_10/assembly.fasta assemblies/assembly_10.fasta && rm -r assembly_10
```

```
miniasm_and_minipolish.sh read_subsets/sample_11.fastq "$threads" > assembly_11.gfa &&
```

```
any2fasta assembly_11.gfa > assemblies/assembly_11.fasta && rm assembly_11.gfa
```

```
raven --threads "$threads" read_subsets/sample_12.fastq > assemblies/assembly_12.fasta && rm  
raven.cereal
```

```
trycycler cluster --assemblies assemblies/*.fasta --reads reads.fastq --out_dir trycycler
```

```
trycycler msa --cluster_dir trycycler/cluster_001
```

```
trycycler msa --cluster_dir trycycler/cluster_002
```

```
trycyclcr msa --cluster_dir trycyclcr/cluster_003
trycyclcr consensus --cluster_dir trycyclcr/cluster_001
trycyclcr consensus --cluster_dir trycyclcr/cluster_002
trycyclcr consensus --cluster_dir trycyclcr/cluster_003
cat trycyclcr/cluster_*/7_final_consensus.fasta > trycyclcr/consensus.fasta
```

C. Lignes de code pour FeatureCounts :

```
Subread-buildindex -o trycyclcr_index fichier_annotation_gb.fasta
Subjunc -T 28 -i trycyclcr_index -r read_1 -R read_2 -o subjunc_results.bam
featureCounts -t transcript -T 28 -a fichier.gtf -o counts.ncbi.txt subjunc_results.bam
```

D. Ligne de code pour Deseq2:

```
Data.table <- read.delim("comptage_12h_230307.txt",row.names = 1)
Data.env <- read.delim("metadata_12H_230307.txt",row.names = 1)
Data.dds <- DESeqDataSetFromMatrix(countData = Data.table, colData=Data.env, design= ~group)
Data.dds<-DESeq(Data.dds)
resultsNames(Data.dds)
Data.res<-results(Data.dds, contrast=c("group", "F03_12H", "WOF_12H"), alpha=0.05)
Data.reOrd<-Data.res[order(Data.res$pvalue),]
```

E. Ligne de code pour FeedMaurice:

```
Perl FeedMaurice.pl --gff3 S0756_trycyclcr_Ref100000000_194893.current.gb.gff --kg
koala_ghost_genome_ko.txt --Dq_output Deseq2.noemie.analyse3.WOF_12H-WOF_24H.tsv --
kegg_class Kegg.genome.module.txt
```