Rorida Assembly Quality AMR Detection (FLAQ-AMR)

Advanced Molecular Detection Southeast Region Bioinformatics

Updates

Office Hours

- October 28, 2024 Sanibel
- November 18, 2024 FLAQ-SARS-CoV-2



Overview

Purpose:

Standard Bacterial Assembly

Used for:

- Taxonomic Identification
- Antimicrobial Resistance (AMR) Detection

<u>Dependencies</u>

- Python3
- Singularity
- Git



Workflow

R1 and R2 fastq.gz files

Run the pipeline on HPG

Analyzes
Illumina PairedEnd Sequencing

Outputs

AMR Detection

Quality Metrics Report Multi Locus
Sequencing
Typing (MLST)

De Novo Assembly

Annotated Assembly

Taxonomic Classification



Application

Objective:

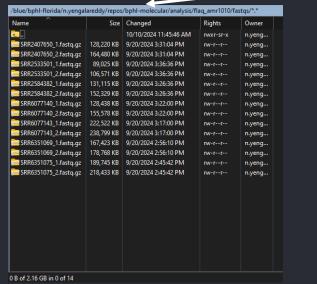
Use a training dataset (Salmonella enterica dataset) and analyze the dataset using the FLAQ-AMR pipeline.

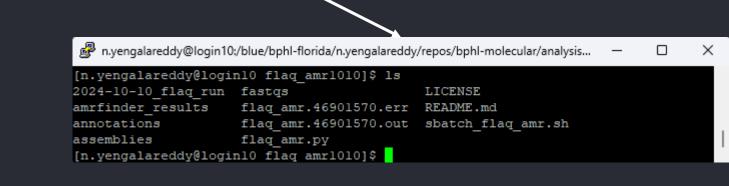


Application Cont.

```
cd /blue/bphl-<state>/<user>/repos/bphl-
molecular/
git clone https://github.com/BPHL-
Molecular/flaq_amr.git
mkdir analysis/
cd analysis/
cp /blue/bphl-<state>/<user>/repos/bphl-
molecular/flaq_amr/*
Mkdir fastqs/
cp /path/to/fastqs/*.fastq.gz fastqs/
```



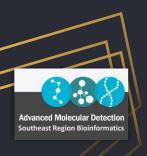


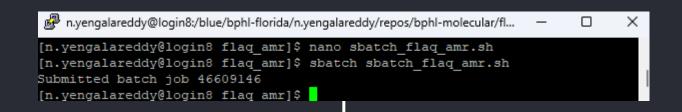


Application Cont.

nano sbatch_flaq_amr.sh

```
n.yengalareddy@login8:/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/fl...
  GNU nano 2.9.8
                                sbatch flag amr.sh
#! /bin/bash
#SBATCH --account=bphl-umbrella
#SBATCH --qos=bphl-umbrella
#SBATCH --job-name=flaq amr
#SBATCH --mail-type=END,FAII
#SBATCH --cpus-per-task=8
         --output=rraq amr.%j.out
#SBATCH --error=flaq_amr.%j.err
module load apptainer
module load conda
conda activate base
#Run script/command and use $SLURM CPUS ON NODE
python flaq_amr.py fastqs/
```



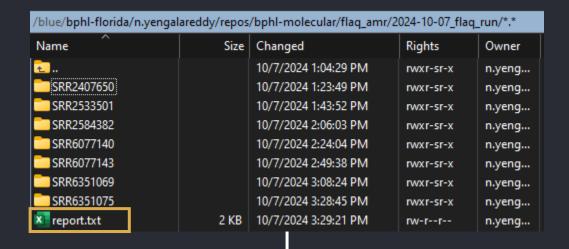


/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/flaq_amr/								
Name	Size	Changed	Rights	Owner				
<u>€</u>		9/23/2024 8:04:39 AM	rwxr-sr-x	n.yeng				
2024-10-07_flaq_run		10/7/2024 3:09:00 PM	rwxr-sr-x	n.yeng				
amrfinder_results		10/7/2024 3:29:21 PM	rwxr-sr-x	n.yeng				
annotations		10/7/2024 3:29:21 PM	rwxr-sr-x	n.yeng				
assemblies assemblies		10/7/2024 3:29:21 PM	rwxr-sr-x	n.yeng				
fastqs		9/20/2024 3:38:42 PM	rwxr-sr-x	n.yeng				
flaq_amr.46609146.err	0 KB	10/7/2024 1:04:26 PM	rw-rr	n.yeng				
flaq_amr.46609146.out	0 KB	10/7/2024 1:04:26 PM	rw-rr	n.yeng				
flaq_amr.py	14 KB	9/19/2024 2:01:32 PM	rwxr-xr-x	n.yeng				
LICENSE	2 KB	9/19/2024 2:01:32 PM	rw-rr	n.yeng				
README.md	5 KB	9/19/2024 2:01:32 PM	rw-rr	n.yeng				
sbatch_flaq_amr.sh	1 KB	9/19/2024 2:06:38 PM	rwxr-xr-x	n.yeng				

Application Cont.



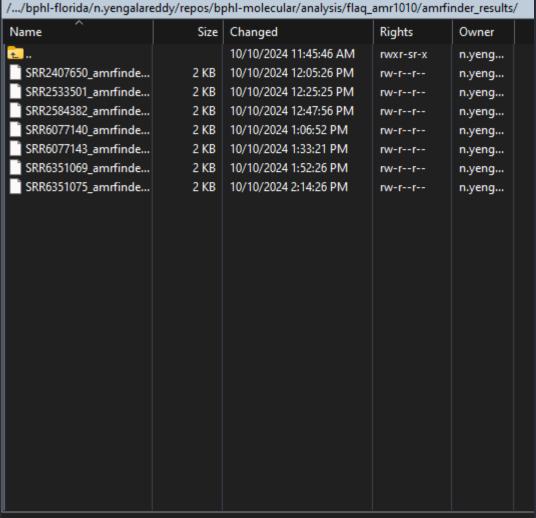
Output



sampleID	species ID_mash	nearest_neighbor_mash	mash_distance	speciesID_kraken	kraken_percent	mlst_scheme	mlst_st	num_clean_reads	avg_readlength	avg_read_qual	est_coverage	num_contigs	longest_contig	N50	L50 tot	al_length	gc_content	annotated_cds
SRR2407650	Salmonella_enterica	GCF_000623775.1	0.0011856	Salmonella enterica	97.05	senterica	11	1027736	225.85	36.03	49.14	32	1508692	400610	3	4722943	52	4433
SRR2533501	Salmonella_enterica	GCF_000623775.1	0.00060662	Salmonella enterica	97.74	senterica	11	937312	221.7	36.44	44.04	26	1508692	405853	3	4717664	52	4420
SRR2584382	Salmonella_enterica	GCF_000623775.1	0.000931334	Salmonella enterica	97.29	senterica	11	1460314	222.61	36.64	69.75	55	634220	197759	7	4660154	52	4356
SRR6077140	Salmonella_enterica	GCF_000623775.1	0.00213483	Salmonella enterica	96.82	senterica	11	1131246	210.83	36.38	50.66	36	1215647	405952	4	4707886	52	4411
SRR6077143	Salmonella_enterica	GCF_000623775.1	0.00343111	Salmonella enterica	96.83	senterica	11	2255536	211.83	36.39	101.3	27	1508692	478604	3	4716521	52	4421
SRR6351069	Salmonella_enterica	GCF_000623775.1	0.00213483	Salmonella enterica	96.27	senterica	11	1506344	200.02	36.15	63.89	32	1024859	440575	4	4715769	52	4422
SRR6351075	Salmonella_enterica	GCF_000623775.1	0.00157453	Salmonella enterica	96.24	senterica	11	1124130	183.18	35.64	44.13	29	1024859	406018	4	4665490	52	4365



Output Cont.



0 B of 10.5 KB in 0 of 7



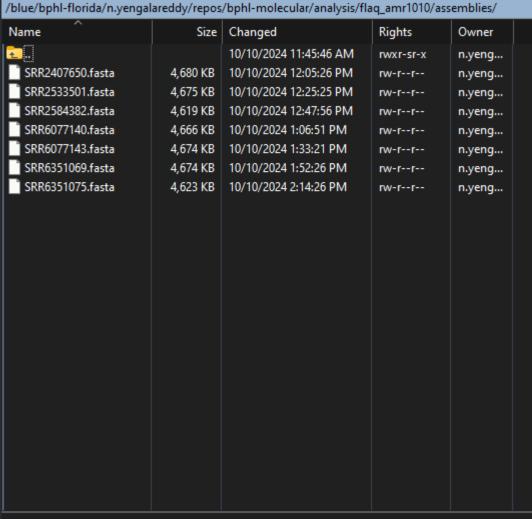
/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/flaq_amr1010/annotations/								
Name	Size	Changed	Rights	Owner				
₹		10/10/2024 11:45:46 AM	rwxr-sr-x	n.yeng				
SRR2407650.gff	6,682 KB	10/10/2024 12:05:26 PM	rw-rr	n.yeng				
SRR2533501.gff	6,673 KB	10/10/2024 12:25:25 PM	rw-rr	n.yeng				
SRR2584382.gff	6,586 KB	10/10/2024 12:47:56 PM	rw-rr	n.yeng				
SRR6077140.gff	6,657 KB	10/10/2024 1:06:52 PM	rw-rr	n.yeng				
SRR6077143.gff	6,671 KB	10/10/2024 1:33:21 PM	rw-rr	n.yeng				
SRR6351069.gff	6,670 KB	10/10/2024 1:52:26 PM	rw-rr	n.yeng				
SRR6351075.gff	6,595 KB	10/10/2024 2:14:26 PM	rw-rr	n.yeng				

0 B of 45.4 MB in 0 of 7



Output Cont.

Output Cont.



0 B of 31.8 MB in 0 of 7



Conclusion



Fundamentals of FLAQ-AMR



Installation and setup of FLAQ-AMR in HPG



Successfully executed job query for FLAQ-AMR



Generated Output Files





Advanced Molecular Detection Southeast Region Bioinformatics

Questions?

<u>bphl-sebioinformatics@flhealth.gov</u>

Molly Mitchell, PhD
Bioinformatics Supervisor
Molly.Mitchell@flhealth.gov

Nikhil Reddy, MS
Bioinformatician
Nikhil.Yengala@flhealth.gov

Sam Marcellus, MPHBioinformatician
Samantha.marcellus@flhealth.gov