

14 APR 2025

MIRA-NF

**Advanced Molecular Detection
Southeast Region Bioinformatics**

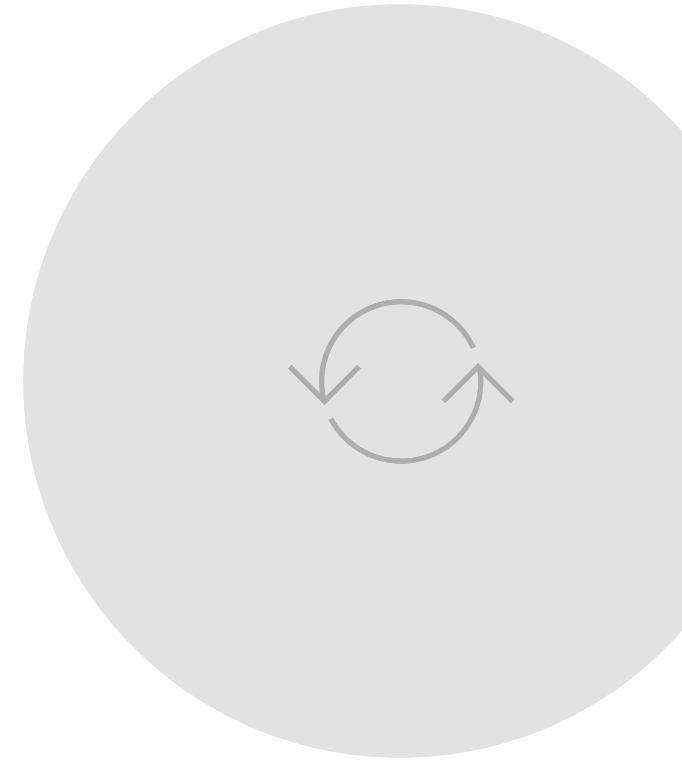
“

This resource was made possible through funding provided under the Epidemiology and Laboratory Capacity for Prevention and Control of Emerging Infectious Diseases (ELC) Cooperative Agreement (CK24-0002), Project D: Advanced Molecular Detection to the Florida Department of Health. The conclusions, findings, and opinions expressed by authors do not necessarily reflect the official position of the U.S. Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention.



Updates

- Still waiting on dates for our online and in-person trainings. Contractor changed from ABiL to UGA – delayed!



Overview

Purpose

- This pipeline is intended to provide a modular, automated and reproducible pipeline for QC, genome assembly, and variant analysis of viral pathogens focusing on SARS-CoV-2, Influenza, and RSV.

Usage

- It is used by public health labs and researchers to better understand outbreaks, track virus changes, and support studies on how the viruses spread and respond to the treatments.

Dependencies

- Nextflow
- Singularity
- Java



Experiment Types:

SARS-CoV-2
Genomic
Illumina Data

RSV
Illumina
Data

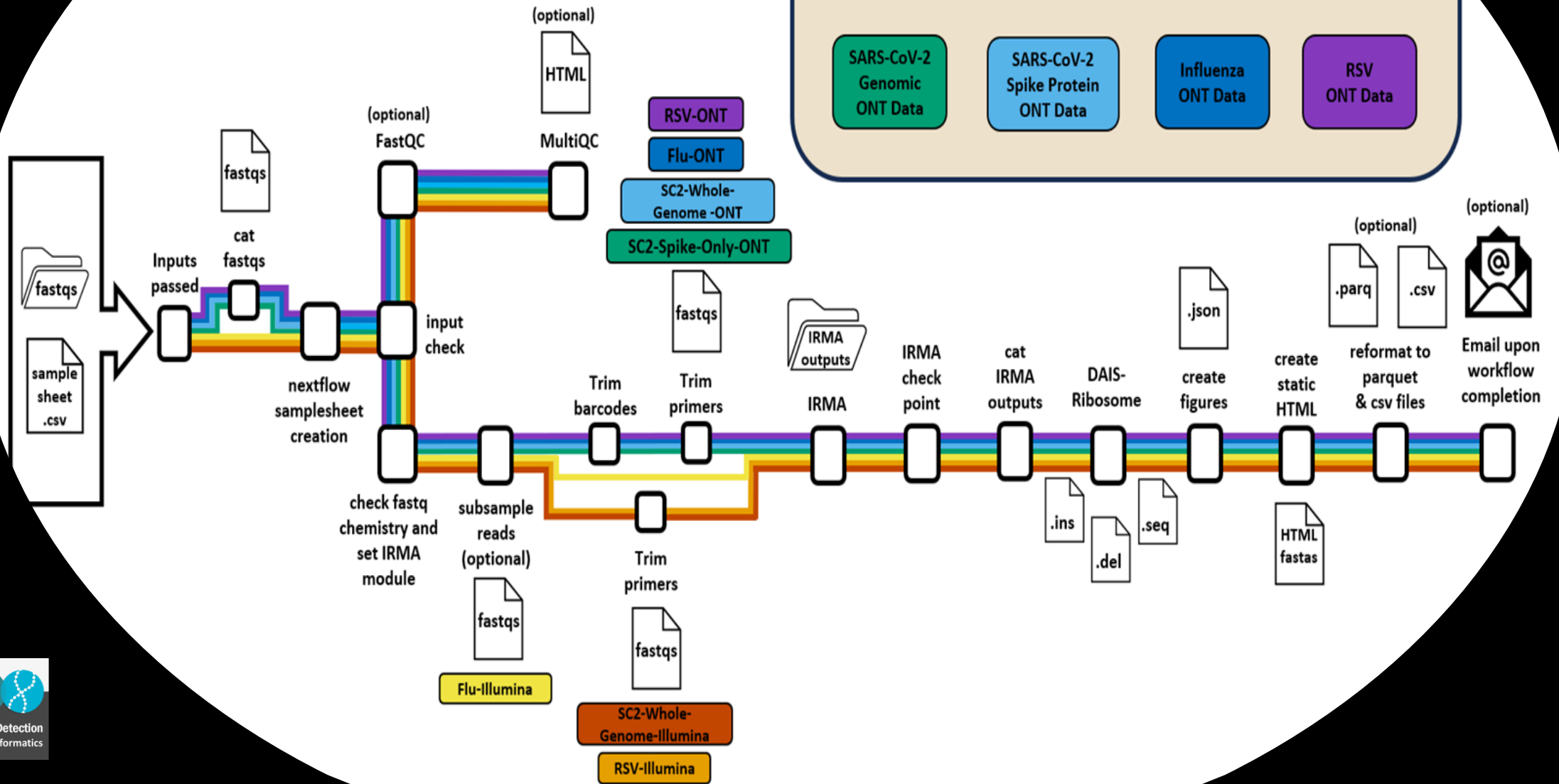
Influenza
Illumina
Data

SARS-CoV-2
Genomic
ONT Data

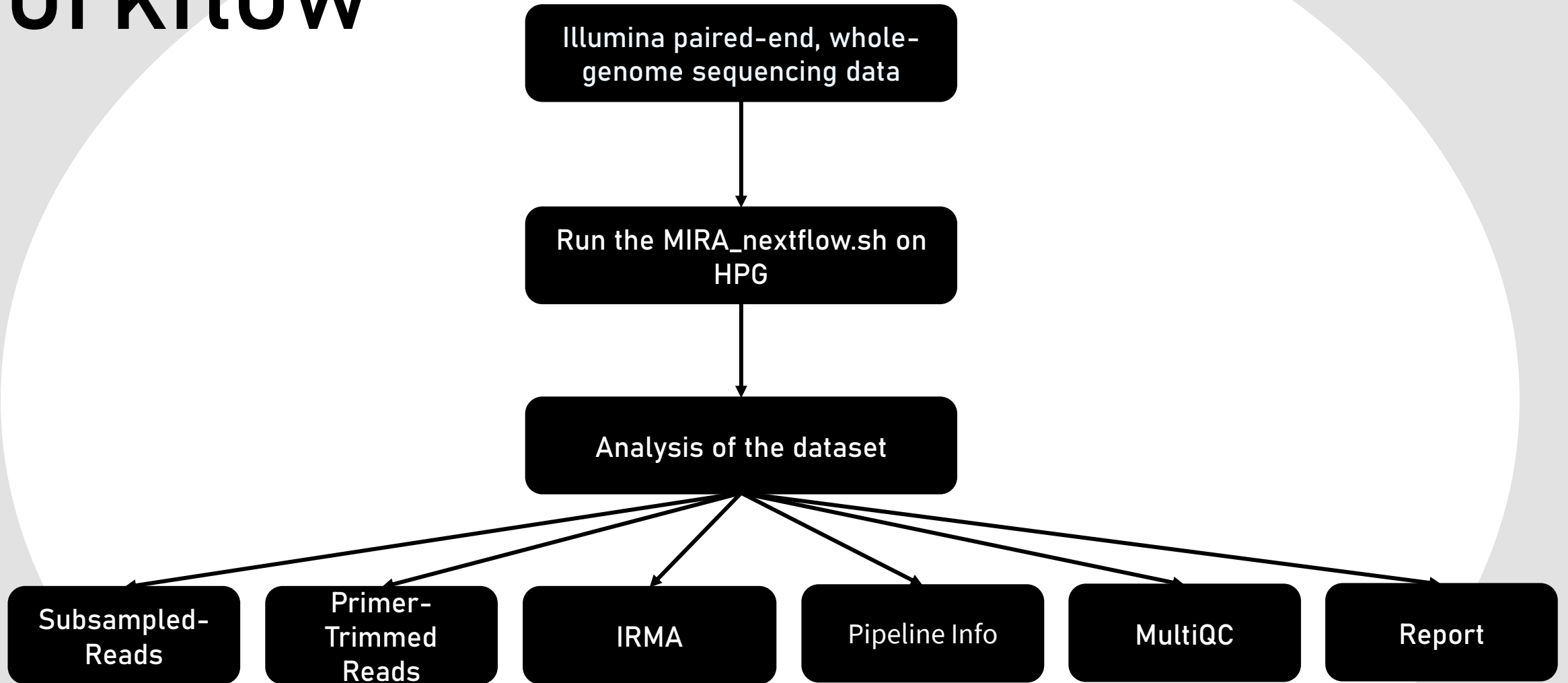
SARS-CoV-2
Spike Protein
ONT Data

Influenza
ONT Data

RSV
ONT Data



Workflow



Application

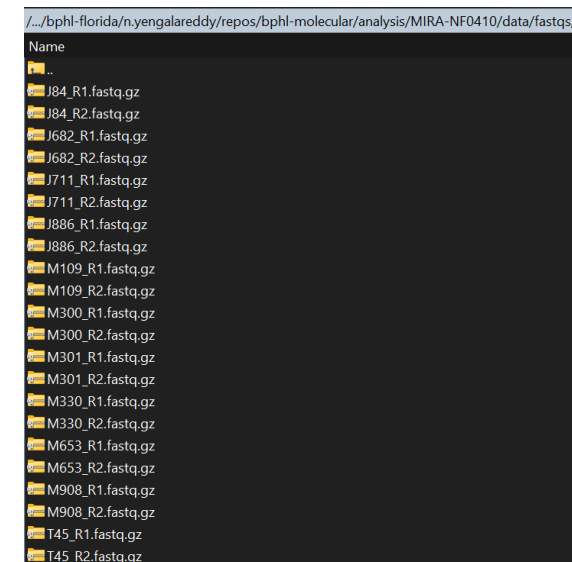
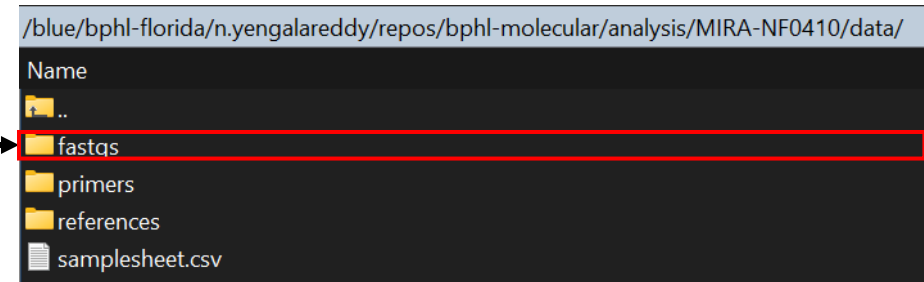
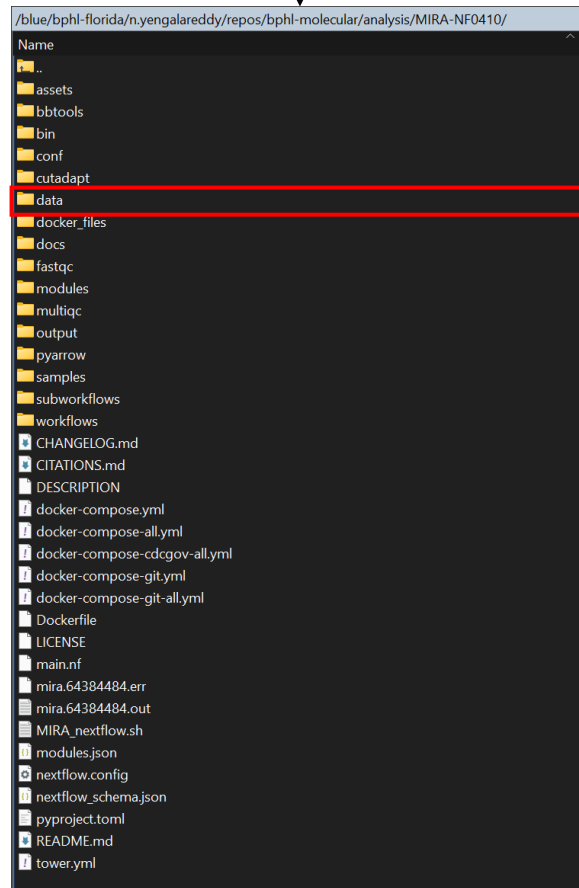
Objective

Use SARS-CoV-19 Illumina Paired-end dataset and analyze using MIRA-NF Pipeline.



Application Cont.

```
cd /blue/bphl-<state>/<user>/repos/bphl-molecular/  
git clone https://github.com/CDCgov/MIRA-NF  
mkdir analysis/  
cd analysis/  
cp /blue/bphl-<state>/<user>/repos/bphl-molecular/ MIRA-NF/  
copy .fastq to data/fastqs directory
```



Application Cont.

```
/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/  
Name  
..  
fastqs  
primers  
references  
samplesheet.csv
```

```
Sample ID,Sample Type,R1 Path,R2 Path  
J84,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J84_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J84_R2.fastq.gz  
J682,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J682_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J682_R2.fastq.gz  
J711,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J711_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J711_R2.fastq.gz  
J886,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J886_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/J886_R2.fastq.gz  
M109,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M109_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M109_R2.fastq.gz  
M300,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M300_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M300_R2.fastq.gz  
M301,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M301_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M301_R2.fastq.gz  
M330,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M330_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M330_R2.fastq.gz  
M653,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M653_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M653_R2.fastq.gz  
M908,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M908_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/M908_R2.fastq.gz  
T45,Test,/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/T45_R1.fastq.gz,/blue/bphl-  
florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/fastqs/T45_R2.fastq.gz
```



Application Cont.

```
GNU nano 2.9.8 MIRA_nextflow.sh

#!/bin/bash
#SBATCH --account=bphl-umbrella
#SBATCH --qos=bphl-umbrella
#SBATCH --job-name=mira-nf
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=20
#SBATCH --mem=128gb
#SBATCH --time=72:00:00
#SBATCH --output=mira.%j.out
#SBATCH --error=mira.%j.err
#SBATCH --mail-user=nikhil.yengala@flhealth.gov
#SBATCH --mail-type=FAIL,END

# === Set input and pipeline parameters ===
DIRNAME="/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410"
INPUT="$DIRNAME/data/samplesheet.csv"
OUTPATH="$DIRNAME/output"
RUNPATH="$DIRNAME"
EXPERIMENT_TYPE="SC2-Whole-Genome-Illumina"
APPLICATION="singularity"
PROCESSQ="long.q"

# === Primer schema fix ===
PRIMERS_DIR="$DIRNAME/data/primers"
cd "$PRIMERS_DIR"
rm -f trimmed_primers.fasta
cp -f artic_v5.3.2.fasta trimmed_primers.fasta

# === Optional arguments ===
REFORMAT="true"
READ_COUNTS="1000"
EMAIL="nikhil.yengala@flhealth.gov"
READS_QC="true"

# === Load modules ===
source /etc/profile
module load nextflow
module load singularity

# === Archive previous run outputs ===
TAR=True
if [ -d "$RUNPATH/dash-json/" ] && [ -n "${TAR}" ]; then
    tar --remove-files -czf $(RUNPATH)/previous_run_$(date +%Y%m%d-%H%M%S).tar.gz \
        $(RUNPATH)/html \
        $(RUNPATH)/fasta \
        $(RUNPATH)/txt \
        $(RUNPATH)/xlsx \
        $(RUNPATH)/IRMA \
        $(RUNPATH)/dash-json
fi

# === Clean up old primer directories that may conflict ===
find "$OUTPATH" -type d -name primers -exec rm -rf {} +

export TMPDIR="$HOME/tmp"
mkdir -p "$TMPDIR"

# === Run Nextflow pipeline ===
nextflow run "$DIRNAME"/main.nf \
    --input "$INPUT" \
    --outdir "$OUTPATH" \
    --runpath "$RUNPATH" \
    --o "$EXPERIMENT_TYPE" \
    --process_q "$PROCESSQ" \
    --custom_primers "$PRIMERS_DIR/trimmed_primers.fasta" \
    --reformat_tables "$REFORMAT" \
    --subsample_reads "$READ_COUNTS" \
    --email "$EMAIL" \
    --read_qc "$READS_QC" \
    --profile "$APPLICATION" \
    --resume \
    --with-trace "$OUTPATH/trace.txt" \
    --with-report "$OUTPATH/report.html"

rm -rf "$TMPDIR"
```

```
/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/
Name
├── ..
├── assets
├── bbttools
├── bin
├── conf
├── cutadapt
├── data
├── docker_files
├── docs
├── fastqc
├── modules
├── multiqc
├── output
├── pyarrow
├── samples
├── subworkflows
├── workflows
├── CHANGELOG.md
├── CITATIONS.md
├── DESCRIPTION
├── docker-compose.yml
├── docker-compose-all.yml
├── docker-compose-cdcgov-all.yml
├── docker-compose-git.yml
├── docker-compose-git-all.yml
├── Dockerfile
├── LICENSE
├── main.nf
├── mira.64384484.err
├── mira.64384484.out
├── MIRA_nextflow.sh
├── modules.json
├── nextflow.config
├── nextflow_schema.json
├── pyproject.toml
├── README.md
└── tower.yml
```



Application Cont.

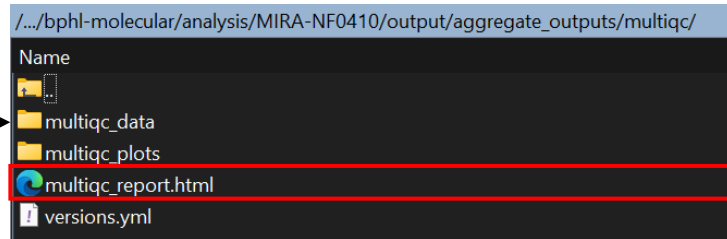
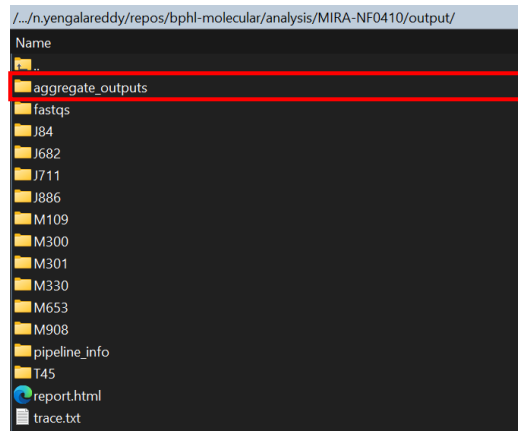
sbatch MIRA_nextflow.sh

```
/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/  
Name  
├── ..  
├── assets  
├── bbttools  
├── bin  
├── conf  
├── cutadapt  
├── data  
├── docker_files  
├── docs  
├── fastqc  
├── modules  
├── multiqc  
├── output  
├── pyarrow  
├── samples  
├── subworkflows  
├── workflows  
├── CHANGELOG.md  
├── CITATIONS.md  
├── DESCRIPTION  
├── docker-compose.yml  
├── docker-compose-all.yml  
├── docker-compose-cdcgov-all.yml  
├── docker-compose-git.yml  
├── docker-compose-git-all.yml  
├── Dockerfile  
├── LICENSE  
├── main.nf  
├── mira.64384484.err  
├── mira.64384484.out  
├── MIRA_nextflow.sh  
├── modules.json  
├── nextflow.config  
├── nextflow_schema.json  
├── pyproject.toml  
├── README.md  
└── tower.yml
```

```
../n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/output/  
Name  
├── ..  
├── aggregate_outputs  
├── fastqs  
├── J84  
├── J682  
├── J711  
├── J886  
├── M109  
├── M300  
├── M301  
├── M330  
├── M653  
├── M908  
├── pipeline_info  
├── T45  
├── report.html  
└── trace.txt
```



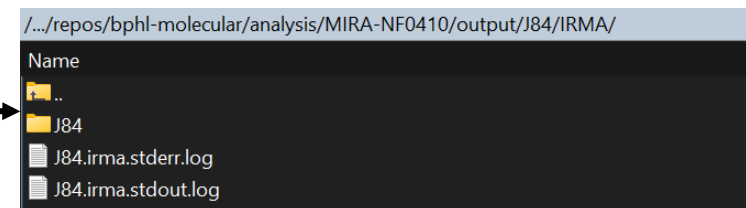
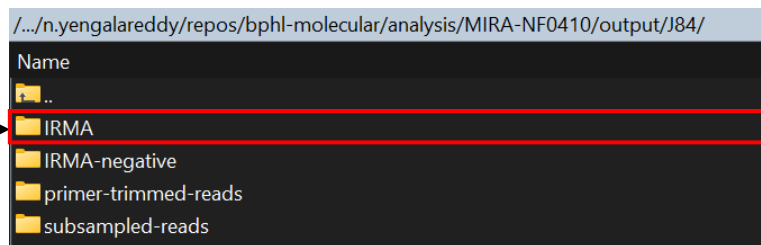
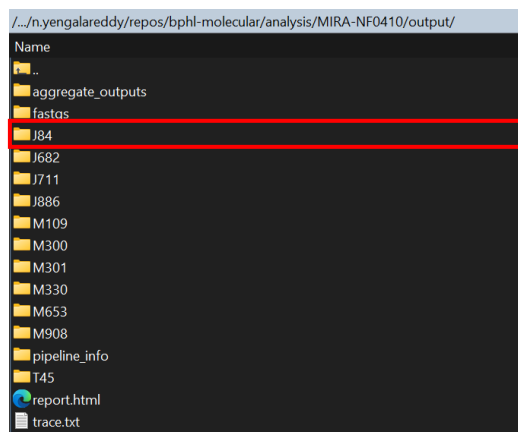
Application Cont.



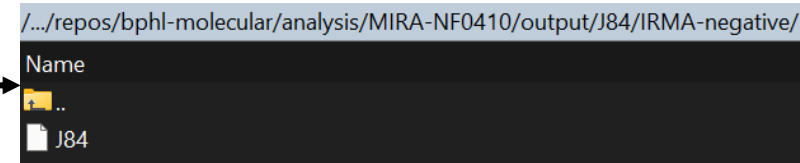
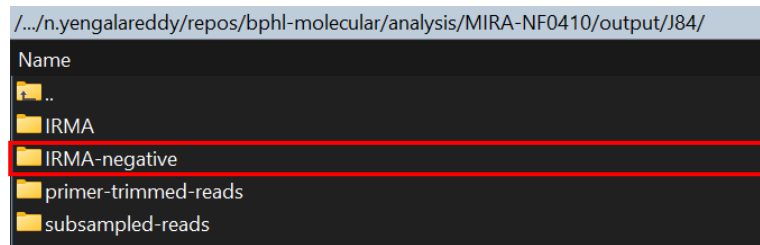
| General Statistics | | | | |
|--------------------|--------|------|--------------------|--------|
| Sample Name | % Dups | % GC | Median Read Length | M Seqs |
| J682_T1_1 | 73.3% | 40% | 242 bp | 0.1 |
| J682_T1_2 | 60.6% | 40% | 242 bp | 0.1 |
| J711_T1_1 | 85.9% | 38% | 151 bp | 0.2 |
| J711_T1_2 | 85.6% | 38% | 151 bp | 0.2 |
| J84_T1_1 | 87.6% | 38% | 251 bp | 0.4 |
| J84_T1_2 | 83.6% | 38% | 251 bp | 0.4 |
| J886_T1_1 | 84.7% | 38% | 151 bp | 0.2 |
| J886_T1_2 | 84.4% | 38% | 151 bp | 0.2 |
| M109_T1_1 | 90.0% | 39% | 251 bp | 0.5 |
| M109_T1_2 | 87.6% | 39% | 251 bp | 0.5 |
| M300_T1_1 | 85.0% | 38% | 251 bp | 0.1 |
| M300_T1_2 | 84.0% | 38% | 251 bp | 0.1 |
| M301_T1_1 | 57.6% | 45% | 251 bp | 0.2 |
| M301_T1_2 | 55.2% | 45% | 251 bp | 0.2 |
| M330_T1_1 | 88.4% | 38% | 251 bp | 0.2 |
| M330_T1_2 | 87.2% | 38% | 251 bp | 0.2 |
| M653_T1_1 | 89.6% | 38% | 251 bp | 0.4 |
| M653_T1_2 | 88.6% | 38% | 251 bp | 0.4 |
| M908_T1_1 | 87.6% | 38% | 251 bp | 0.3 |
| M908_T1_2 | 86.2% | 38% | 251 bp | 0.3 |
| T45_T1_1 | 63.8% | 41% | 251 bp | 0.2 |
| T45_T1_2 | 63.9% | 42% | 251 bp | 0.2 |



Application Cont.



Application Cont.



Application Cont.



Application Cont.

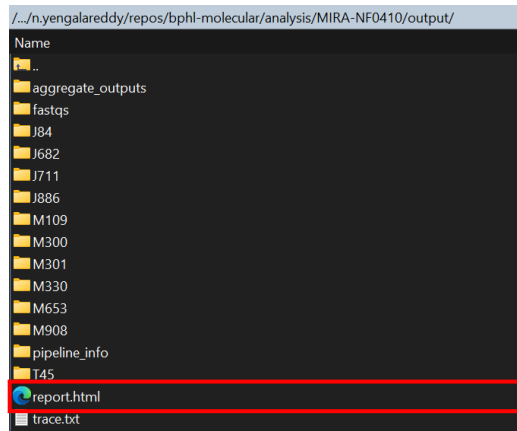
```
../n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/output/J84/  
Name  
..  
IRMA  
IRMA-negative  
primer-trimmed-reads  
subsampled-reads
```



```
../repos/bphl-molecular/analysis/MIRA-NF0410/output/J84/subsampled-reads/  
Name  
..  
J84.reformat.stderr.log  
J84.reformat.stdout.log  
J84_subsampled_R1.fastq  
J84_subsampled_R2.fastq  
trimmed_primers.fasta
```



Application Cont.



Nextflow workflow report

[scruffy_poitras]

Workflow execution completed successfully!

Run times

09-Apr-2025 10:01:28 - 09-Apr-2025 10:03:19 (duration: 1m 51s)

93 succeeded

Nextflow command

```
nextflow run /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/main.nf --input /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/samplesheet.csv --outdir /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/output --runpath /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410 --e SC2-whole-Genome-Illumina --process_q long.q --custom_primers /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/primers/trimmed_primers.fasta --reformat_tables true --subsample_reads 1000 --email nikhil.yengalareddy@flhealth.gov --read_qc true --profile singularity --resume --with-trace /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/output/trace.txt --with-report /blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/output/report.html
```

CPU-Hours

0.1

Launch directory

/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/primers

Work directory

/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410/data/primers/work

Project directory

/blue/bphl-florida/n.yengalareddy/repos/bphl-molecular/analysis/MIRA-NF0410

Script name

main.nf

Script ID

77418511fe436f978925b83be629067

Workflow session

273847e9-b873-45e6-8950-de63143ce246

Workflow profile

singularity

Nextflow version

version 24.10.5, build 5935 (04-03-2025 17:55 UTC)



Conclusion



Fundamentals of
MIRA-NF



Installation and setup
of MIRA-NF in HPG



Successfully executed
job query for MIRA-
NF



Generated output files





Advanced Molecular Detection

Southeast Region Bioinformatics

Questions?

bphl-sebioinformatics@flhealth.gov

Molly Mitchell, PhD

Bioinformatics Supervisor

Molly.Mitchell@flhealth.gov

Nikhil Reddy, MS

Bioinformatician

Nikhil.Yengala@flhealth.gov

Sam Bernhoft, MPH

Bioinformatician

Samantha.bernhoft@flhealth.gov