



# **Advanced Molecular Detection**

## **Southeast Region Bioinformatics**

**Office Hours – Session 1**  
03/06/2023

# Outline



Introductions



BRR Updates



Agenda



Singularity Demo



Open Discussion – Questions

# BRR Team



Sarah Schmedes, PhD  
Lead Bioinformatician  
Sarah.Schmedes@flhealth.gov



Molly Mitchell  
Bioinformatician  
Molly.Mitchell@flhealth.gov



Lakshmi Thsaliki, MS  
Bioinformatician  
Lakshmi.Thsaliki@flhealth.gov

For any BRR requests, please send an email to [bphl16bioinformatics@flhealth.gov](mailto:bphl16bioinformatics@flhealth.gov)  
Based on the requests one of us from the team will respond as soon as possible.

# ELC Updates/Reminders

**Point 1:** Program A, Activity 1f – AMD Regional Workforce Development Training Participant

- i. Please apply for this to send participants to an in-person training workshop (hosted by ABiL) in Atlanta, GA in Fall 2023
  - 1. \$3,000-\$10,000 available for each jurisdiction

**Point 2:** Program A, Activity 6a – AMD Platform Support Core Activities

- i. Please apply and participate in at least two Communities of Practice
  - 1. You do not need to be a subject matter expert in any of the CoPs. Participation in CoPs mainly allows supporting states to provide their needs and input for features for the new AMD Platform
  - 2. \$25,000 available for each jurisdiction
    - a. \*CDC said that most jurisdictions did not apply. There is an award available for all ELC jurisdictions.

# Agenda

**March 6** – Singularity

**March 20** - Conda

**April 3** – Nextflow

**April 17** - Phoenix #1 (Installation and Dependencies)

**May 1** - Phoenix #2 (Demo)

## Future Trainings

- StaPH-B Toolkit Programs/Pipelines
- GISAID flagged SARS-CoV-2
- Git (git clone, etc.)
- Generating R figures
- Long/short read sequencing and the best bioinformatic pipelines

# Why use Containers?

- Containers are an empty room which users can configure and customize
- Containers are a standardized unit of software that packages code and all dependencies
- Containers change the user space into a swappable component which means programs, custom configurations, and environment can be independent
- Consolidating a workflow into a Singularity container simplifies distribution and replication of scientific results

# Singularity

- Singularity is a container platform
- Software application that allows users to have ‘full control’ over their operating system without the need for any ‘super-user’ privileges using the notion of containers or images
- Allows you to create and run containers that package software in a way that is portable & reproducible
- **Note** – on HPG singularity is available as apptainer module (**module load apptainer**)
- Containers are built using Singularity which can run on HPC clusters
- A container is a single file

```
$ singularity --version
```

```
$ singularity help
```

# Importance of Singularity

- Used to run complex applications on HPC clusters in a simple, portable, and reproducible way
- Verifiable reproducibility and security which uses cryptographic signatures, an immutable container image format, and in-memory decryption
- Integration over isolation by default. Easily make use of GPUs, high speed networks, parallel filesystems on a cluster or server by default.
- Mobility of compute. The single file SIF container format is easy to transport and share.
- A simple, effective security model



# How to setup your directory for storing and using Singularity images on HiPerGator?

```
$ mkdir /blue/bphl-<state>/<user>/singularity/  
$ ln -s /blue/bphl-<state>/<user>/singularity/ ~/.singularity
```

- By default, Singularity will try to store your images in /home/.singularity/. However, these files can be quite large and may exceed your home directory storage quota.
- Use the above commands to make a symbolic link to a hidden directory on the /blue drive.

# How to get Singularity Container Images?

- A container image is a single executable file that defines the software environment and runs the container
- A single container image can be used to run multiple instances of the same container simultaneously for different jobs
- To get a container image for use on HiPerGator, you can either pull (i.e., download) pre-built container images into one of your directories or externally build a custom container image from a definition file and then transfer it to one of your directories

# Build Containers

- The build command accepts a target as input and produces a container as output
- The build command produces containers in two different formats:
  - A compressed read-only Singularity Image File (SIF) format suitable for production (default)
  - A writable root directory called a sandbox for interactive development (--sandbox option)
- The build command accepts an existing container as a target to create a container in either supported format or by converting existing containers from one format to another

```
$ singularity build
```

# Build Containers – StaPH-B Toolkit

- Path for prebuilt StaPH-B tools

```
/apps/staphb-toolkit/containers/
```

- For example, for shigatyper:

```
$ singularity build shigatyper_2.0.3.sif docker://staphb/shigatyper:2.0.3
```

- [GitHub - StaPH-B/docker-builds: Dockerfiles and documentation on tools for public health bioinformatics](#)
- Any tool from this link can be built and then used similarly when applying the command above

# Running Singularity Container Images

- Use the command **singularity exec** for executing commands within a container non-interactively

```
singularity exec
```

# Running Slurm Jobs with Singularity

Include the singularity commands to run your job scripts when running batch slurm jobs

Make sure to load singularity  
**module load apptainer**

```
thsalikilakshmi@login5/blue/bphl-florida/thsalikilakshmi/scripts
File Edit Options Buffers Tools Sh-Script Help
root@bphl:/bash
#SBATCH --account=bphl-umc00114
#SBATCH --qos=bphl-umc00114
#SBATCH --job-name=shigatyper
#SBATCH --mail-type=END,FAIL
#SBATCH --mail-user=ENTER Email
#SBATCH --nodes=1
#SBATCH --cpus-per-task=4
#SBATCH --mem=30gb
#SBATCH --time=04:00:00
#SBATCH --output=shigatyper_%j.out
#SBATCH --error=shigatyper_%j.err

module load apptainer

mkdir -p shigatyper_out

for f in $(cat files.txt)
do
    singularity exec -B $(pwd):/data /apps/staphb-toolkit/containers/shigatyper_2.0.1.sif shigatyper --R1 ${f}_1.fastq.gz --R2 ${f}_2.fastq.gz
    mv *.tsv shigatyper_out
done

--UU--:---F1 shigatyper.sh All L1 (Shell-script[bash]) -----
Indentation setup for shell type bash
```

# References

[Introduction to Singularity — Singularity User Guide 3.7 documentation \(sylabs.io\)](#)

# Time for Questions & Feedback

- Questions?
  - Do you need help with anything?
  - Requests for separate trainings?
- Feedback
  - What would you like to see?





# Advanced Molecular Detection Southeast Region Bioinformatics

## Questions?

[bphl16bioinformatics@flhealth.gov](mailto:bphl16bioinformatics@flhealth.gov)

**Lakshmi Thsaliki, MS**

Bioinformatician

[Lakshmi.Thsaliki@flhealth.gov](mailto:Lakshmi.Thsaliki@flhealth.gov)

**Molly Mitchell**

Bioinformatician

[Molly.Mitchell@flhealth.gov](mailto:Molly.Mitchell@flhealth.gov)

**Sarah Schmedes, PhD**

Lead Bioinformatician

[Sarah.Schmedes@flhealth.gov](mailto:Sarah.Schmedes@flhealth.gov)