Charles Pronghorn Homework Module

\*\*\* Next time remind them how to log in and where to go in this sheet then– then get into the homework.

All personal directories are located: /data/gpfs/assoc/biol\_bids-0/

Necessary files and directories are located:

1. /data/gpfs/assoc/biol\_bids-0/**[NETID]**
2. /data/gpfs/assoc/biol\_bids-0/atram\_files

In your personal directory you will need to create two directories:

mkdir temp

mkdir atram\_out

And create two files:

1. run\_atram.sh
2. class.slurm.sh

The ‘run\_atram.sh’ script is sent into the container to execute aTRAM and the path needs to be altered. Below is an example of the script and the bold sections within square brackets should be changed accordingly.

*run\_atram.sh*

#!/usr/bin/bash

python3 /aTRAM/atram.py -b /biol\_bids-0/atram\_files/atram\_lib/AlspHabad/AlspHabad -t /biol\_bids-0/**kristincharles/**temp -Q /biol\_bids-0/atram\_files/query\_files/100\_proteins.fasta -a trinity -o /biol\_bids-0/**kristincharles**/atram\_out/ --cpus 8

The ‘class.slurm.sh’ script needs to be altered as well. Add the name of the job (this can be anything you want), how many CPUs per task (you decide what you think is appropriate), and your email to get updates on the status of your job.

*class.slurm.sh*

#!/usr/bin/bash -l

#SBATCH --job-name=

#SBATCH --account=cpu-s5-biol\_bids-0

#SBATCH --partition=cpu-core-0

#SBATCH --nodes=1

#SBATCH --ntasks=1

#SBATCH --cpus-per-task=**8**

#SBATCH --hint=compute\_bound

#SBATCH --time=14-00:00

#SBATCH --mem-per-cpu=3500M

#SBATCH --output=hostname\_%j.out

#SBATCH --error=hostname\_%j.err

#SBATCH --mail-type=ALL

#SBATCH --mail-user=**kristincharles@nevada.unr.edu**

srun singularity exec -B /data/gpfs/assoc/biol\_bids-0:/biol\_bids-0 /apps/atram/aTRAM.sif run\_atram.sh

## this should be apps not app

Submit the job using:

sbatch class.slurm.sh

Answer the following questions:

1. What is the purpose of High Performance Computers (HPCs) and when might you use them?

High performance computers combine the computing power from multiple sources into a cluster to allow for rapid processing of data in such a way that a single computer would not be able to manage. These are extremely useful for large or complex datasets and tasks that would take an inordinate amount of time and space on one’s own personal device, such as genome assembly and annotation.

1. What is a slurm script? Paste the path your slurm script here.

A slurm script (.sl) is essentially a bash/shell script that includes parameters specifying which resources slurm should request when scheduling a job. /data/gpfs/assoc/biol\_bids-0/**kristincharles/charles.slurm.sh**

**## I am not sure why the homework specified that this should be a .sh file, but the example file provided was an .sl file… the files looked identical, does this make a difference?**

1. Paste the path to your gene assemblies on pronghorn.

If they had been successful, the gene assemblies should have been in /data/gpfs/assoc/biol\_bids-0/**kristincharles/**atram\_out

1. How many CPUs did you ask for?

8

1. How many genes assembled?

None, all of my runs failed almost immediately (see below) JobID JobName Partition Account AllocCPUS State ExitCode

------------ ---------- ---------- ---------- ---------- ---------- --------

2687521 charles\_j+ cpu-core-0 cpu-s5-bi+ 16 FAILED 127:0

2687521.bat+ batch cpu-s5-bi+ 16 FAILED 127:0

2687521.ext+ extern cpu-s5-bi+ 16 COMPLETED 0:0

2687521.0 singulari+ cpu-s5-bi+ 8 FAILED 127:0

2687522 charles\_j+ cpu-core-0 cpu-s5-bi+ 16 FAILED 127:0

2687522.bat+ batch cpu-s5-bi+ 16 FAILED 127:0

2687522.ext+ extern cpu-s5-bi+ 16 COMPLETED 0:0

2687522.0 singulari+ cpu-s5-bi+ 8 FAILED 127:0

2687523 charles\_j+ cpu-core-0 cpu-s5-bi+ 16 FAILED 127:0

2687523.bat+ batch cpu-s5-bi+ 16 FAILED 127:0

2687523.ext+ extern cpu-s5-bi+ 16 COMPLETED 0:0

1. Did all genes assemble? Why or why not?

The reservation Sebastian made for the class expired 2 weeks ago, and despite trying to submit the job on various queues or without a reservation the job failed shortly after starting, presumably because the requested resources were not available.

1. How long did it take for them to assemble? What processes might have slowed this down?

I believe that since there was no longer a reservation, other people’s jobs likely had priority in the queue and prevented mine from running because it exceeded some threshold resource that was allotted to others.