Edta on HPC

What is Edta?

This package is developed for automated whole-genome *de-novo* TE annotation and benchmarking the annotation performance of TE libraries.

The EDTA package was designed to filter out false discoveries in raw TE candidates and generate a high-quality non-redundant TE library for whole-genome TE annotations. Selection of initial search programs were based on benchmarking on the annotation performance using a manually curated TE library in the rice genome.

More info on E:

Edta - GitHub

Versions Available:

The following versions are available on the cluster:

- Edta –1.9.6

How to load Quast?

To load Edta on the HPC, use the following commands:

```
module load bio/edta
```

To verify if the module and dependencies are loaded correctly, use the following command.

```
module list
```

This should list all the software and dependencies that are loaded.

How to use Quast?

There are two ways to use most of the software on the cluster.

The Interactive Way:

To run the Edta in an interactive session, jump into a base terminal of a compute node using the following command:

```
srun -p main --qos main -n 1 -c 16  --mem 40G --pty bash
```

Now, load the Edta module and start your job interactively on the terminal.

```
#Use this to see the full syntax and options
EDTA.pl -h
```

Note: This is preferred usage in case of short runtime jobs and debugging. After you exit the terminal session, your job will be terminated unlike batch method.

The Script method:

To use this method, the user will need to write a batch script and submit it to a slrum.

To copy the example file, use the following command,

```
#Make a test dir on home
mkdir ~/ed-test  &&  cd ~/ed-test


#Copy the test files into the directory
cp -r /share/apps/EDTA/EDTA/test/  ~/ed-test

#Go to the test folder and make a script
cd * &&  touch script.sbatch
```

The sbatch script should be formatted like

```
#!/bin/bash
#SBATCH -J ap-test
#SBATCH -n 1
#SBATCH -c 12
#SBATCH --mem=8G
#SBATCH --time=00:15:00
#SBATCH -p main
#SBATCH --qos main
#SBATCH -e err.err
#SBATCH -o out.out

module load bio/edta

# Move to test data directory
cd $SLURM_SUBMIT_DIR

#Run the edta command here
EDTA.pl --genome genome.fa --cds genome.cds.fa --exclude
genome.exclude.bed --overwrite 1 --sensitive 1 --anno 1 --evaluate 1 -
-threads 10
```

All the output files will be under the test directory.

Where to find help?

If you are stuck on some part or need help at any point, please contact OIT at the following address.

https://ua-app01.ua.edu/researchComputingPortal/public/oitHelp