Price on HPC

What is Price?

Price is a computational tool used in genomics and bioinformatics to assemble Illumina paired-end reads into contiguous sequences called contigs. The tool is based on an iterative extension algorithm that uses the paired-end information to extend contigs in both directions. The iterative extension approach allows PRICE to assemble difficult regions such as repetitive sequences, homopolymers, and low-complexity regions that are challenging for other assemblers. The output of PRICE is a set of contigs and a set of paired-end reads that were not used in any contig. The tool is specially designed for assembling high-throughput sequencing data from complex genomes such as plants, fungi, and metagenomes.

Links:

PubMed

Manual

Versions Available:

The following versions are available on the cluster:

- PRICE Assembler v1.2

How to load Price?

To load Price, use the following commands:

```
#Load the Price module
module load bio/price/1.2
```

To verify if the module is loaded correctly, use the following command,

```
# List all the module loaded in the environment
module list
```

In a fresh environment, this should show gcc and price module listed.

How to use Price?

The general steps to run price in HPC environment are :

1. Prepare input data: The input for PRICE is a set of Illumina paired-end reads in FASTQ format.
2. Set up the environment: Make sure you have all the necessary dependencies installed on the cluster. PRICE is written in C++ and uses the Boost library. You may also need to load a module for gcc and other dependencies.
3. Download and compile PRICE: Download the latest version of PRICE from the official website, and then compile the source code on the cluster.
4. Submit a job: Once PRICE is compiled, you can submit a job to the cluster's job scheduler. The job should specify the input and output files, the number of threads to use, and any other parameters that you want to use.
5. Monitor the job: You can use the cluster's job scheduler to monitor the progress of your job and check for any errors.

It is good to note that, the command line usage and the parameters may vary depending on the specific version and implementation of PRICE you are using, it's important to check the manual for the specific version you are using.

Users can see a full detailed sample job in following link:

Sample Job Description

Here is a sample slurm script one can use to run price with 20 threads,

```
#!/bin/bash
#SBATCH --job-name=PRICE
#SBATCH --output=PRICE.out
#SBATCH --error=PRICE.err
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=20
#SBATCH --time=48:00:00
#SBATCH --mem=100GB
#SBATCH --p main
#SBATCH --qos main

# load dependencies
module load bio/price/1.2

# run PRICE
./PriceTI -fpp s_2_1_sequence.txt s_2_2_sequence.txt 300 95 -icf
sangerReads.fasta 1 1 5 -nc 30 -dbmax 72 -mol 30 -tol 20 -mpi 80 -
target 90 2 1 1 -o practice.fasta
```

-a $SLURM_CPUS_PER_TASK

Where to find help?

If you are confused or need help at any point, please contact OIT at the following address.

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

ice