Bismark on HPC cluster:

What is Bismark?

Bismark is a program to map bisulfite treated sequencing reads to a genome of interest and perform methylation calls in a single step. The output can be easily imported into a genome viewer, such as Seq Monk, and enables a researcher to analyze the methylation levels of their samples straight away.

Features:

- Bisulfite mapping and methylation calling in one single step
- Supports single-end and paired-end read alignments
- Supports ungapped, gapped or spliced alignments

The official documentation can be found on:

GitHub- Bismark

Documentation

Versions Available:

- Bismark = Version: v0.19.0

How to load a version of Bismark?

To load a version of Bismark on the HPC, use the following command:

```
module load bio/bismark
```

Verify by using this command:

```
module list
```

It should load bowtie as dependency.

Note: User are recommended to load Sam tools while working with bismark.

How to use Bismark on the cluster?

To see all the options available, use the following command available,

bismak --help

For this test run download the required file from following links:

```
# Download fastq file
wget
https://github.com/FelixKrueger/Bismark/raw/master/test_data.fastq

# Download fasta file
wget
https://ftp.ensembl.org/pub/release-
108/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.chromosome.1.fa.gz

# Unzip the file
gunzip *.fa.gz
```

To run the analysis on this dataset, use the following commands

Gene prepration:

```
# Run bismark gene preparation
# $(pwd) is my current data folder
bismark_genome_preparation $(pwd)
```

Run Bismark:

```
# Run bismark on the test data parallelly over 16 cores
bismark --parallel 16 --genome $(pwd) test_data.fastq
```

Dedublication:

```
deduplicate_bismark --bam test_data_bismark_bt2.bam
```

Running Methyl Extractor

```
bismark_methylation_extractor --parallel 16 --gzip --bedGraph
test_data_bismark_bt2.bam
```

To see the report, use the following command

```
bismark2report
```

These commands will process necessary data and write files in the current working directory. If user needs to run a batch job, user can wrap commands into a sbatch script and submit it to scheduler.

To learn more, please refer to the documentation,

Documentation

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