

Bam-readcount on HPC

What is Bam-readcount?

Bam-readcount is a utility that runs on a BAM or CRAM file and generates low-level information about sequencing data at specific nucleotide positions. Its outputs include observed bases, readcounts, summarized mapping and base qualities, strandedness information, mismatch counts, and position within the reads.

More info on bam -readcount:

[Bam-readcount - GitHub](#)

Versions Available:

The following versions are available on the cluster:

- Bam -readcount -v0.8.0

How to load Bam-readcount?

To load Bam-readcount on the HPC, use the following commands:

```
module load bio/bam-readcount
```

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 Bam-readcount?

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

The Interactive Way:

To run the Bam-readcount 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 12G --pty bash
```

Now, load the Bam-readcount module and start your job interactively on the terminal.

```
#Use this to see the full syntax and options available  
bam-readcount -h  
#You can run commands here!
```

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 slurm.

To copy the example file, use the following command,

```
#Make a test dir on home  
mkdir ~/bam-test && cd ~/bam-test  
  
#Get test files  
curl -o test.fa \  
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?db=nucleotide&id=NC_045512.2&rettype=fasta  
  
curl -o test.bam \  
http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeUwRepliSeq/wgEncodeUwRepliSeqBg02esG2AlnRep1.bam
```

```
#make a sbatch script
touch script.sbatch
```

The sbatch script should be formatted like

```
#!/bin/bash
#SBATCH -J bam-test
#SBATCH -n 1
#SBATCH -c 12
#SBATCH --mem=16G
#SBATCH -p main
#SBATCH --qos main
#SBATCH -e err.err
#SBATCH -o out.out

module load bio/bam-readcount

# Move to test data directory
cd $SLURM_SUBMIT_DIR

#Run the bam-readcount command here
bam-readcount -f test.fa test.bam > ans.out
```

To run the script, submit it using:

```
sbatch script.sbatch
```

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>

