

Bamtools on HPC cluster:

What is Bamtools?

BamTools provides both a programmer's API and an end-user's toolkit for handling BAM files. The command line toolkit allows manipulating and querying BAM files for data.

The documentation for the Bamtools is found in:

[Bamtools GitHub](#)

[Tutorial](#)

Versions Available:

Bamtools – 2.5.2

Bamtools – 2391b1a

How to load a version of Bamtools?

To list the versions of Bamtools on the HPC, use the following command:

```
module avail bio/bamtools
```

The version will be listed. To a version of software, use following command:

```
module load bio/bamtools/2.5.2
```

Verify by using this command:

```
module list
```

The loaded software and dependencies will be shown.

How to use Bamtools on the cluster?

Here is one example of how to use Bamtools on the cluster. To get a test file to process through Bamtools, use the following commands

```
cd ~ && mkdir Bamtools_test && cd Bamtools_test # Make a test directory on home #folder

#Download a test file

curl -L -o test.bam http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeUwRepliSeq/wgEncodeUwRepliSeqBg02esG1bAlnRep1.bam

# Now after you load your module and copied test files, we can make a script to run Bamtools
```

Interactive Mode:

To run Bamtools in interactive mode, start a bash session on compute node. **Do not use head node for this purpose.**

```
#Start a session on compute nodes

srun -p main --qos main -n 1 --mem 6G --pty bash

#Load the module if not loaded

module load bio/bamtools/

#Now you can use the bamtools

bamtools index -in ../../<path to bam file>/test.bam
```

The Script:

```
#!/bin/bash
$SBATCH -J Jobname # Jobname
#SBATCH -n 1      # Nodes per task
```

```
#SBATCH --mem=5G      # Must need to allocate memory to run Bamtools - at least 4G
#SBATCH -p main       # Partition
#SBATCH --qos main    # Quality of service
#SBATCH -o Bamtools_sim_out-%J.txt # STDOUT Out file
#SBATCH -e Bamtools_sim_out-%J.txt # Error file
#SBATCH -t 30         # time in minuets

module load bio/bamtools/2.5.2      # load the module

cd $SLURM_SUBMIT_DIR                # Moving to the test directory

#Now use the bamtools if run from diff Dir use ../../path/..test.bam instead of test.ba
m

bamtools stats -in test.bam
bamtools sort -in test.bam
```

Schedule the job with the following sbatch command.

```
sbatch myscript.sh
```

All the processed files will be generated in the submit directory as the script.

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>

Resources:

Tutorial