

Blat on HPC cluster:

What is Blat?

Blat is a precise and fast standalone sequence search command line tool. Blat produces two major classes of alignments:

- at the DNA level between two sequences that are of 95% or greater identity, but which may include large inserts.
- at the DNA level between two sequences that are of 95% or greater identity, but which may include large inserts.

The output of blat is flexible. By default, it is a simple tab-delimited file that describes the alignment, but which does not include the sequence of the alignment itself. Alternatively, blat can produce output compatible with BLAST or WU-BLAST, as well as several other formats.

The full documentation for the Blat is found in the following links:

[Documentation](#)

[Wiki](#)

[Research Paper](#)

Versions Available:

- BLAT v. 36x1

How to load a version of Blat?

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

```
module avail bio/blat
```

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

```
module load bio/blat/
```

Verify by using this command:

```
module list
```

Since this is a standalone software, it does not have dependencies. So, only Blat will be seen listed.

Note:

Two other binaries are provided with the blat module:

gfClient

gfServer

See docs to use these commands.

How to use Blat on the cluster?

There are two methods to run Blat on the cluster.

The Interactive Way:

To run the program interactively, follow the steps:

```
#Open a bash session on compute node
srun -p main --qos main -n 1 -c 12 --mem 10G --pty bash

#Load the module
module load bio/blat
# Start your commands here
blat
```

```
#Follow with commands to execute
```

This method is ideal for a short job run which produces runtime output and to debug the codes.

The Script:

To run a slurm job, the user must prepare input files. For this example, get input files with,

```
#Download all the required files to the test folder
wget
https://raw.githubusercontent.com/lifengtian/kent/master/blat/test/bad
Splice/chr16part.fa
wget
https://raw.githubusercontent.com/lifengtian/kent/master/blat/test/bad
Splice/small2.fa

# Make a script to run Blat
touch script.sbatch
```

Use the following template for the script,

```
#!/bin/bash
#SBATCH -J test
#SBATCH -n 1
#SBATCH -mem=16G
#SBATCH -p main
#SBATCH --qos main
#SBATCH -o blat-%J.txt
#SBATCH -e blat-%J.txt
#SBATCH -t 30

module load bio/blat      # load the module
```

```
cd $SLURM_SUBMIT_DIR           # Moving to the test directory

#Run Blat on the cluster
blat -q=rna -fine chr16part.fa small2.fa small2.psl > small2.log
```

Schedule the job with the following sbatch command.

```
sbatch script.sbatch
```

All the processed files will be generated in the same directory as the sbatch 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>