BioHPC: A High Performance Computing System for Life Sciences

Run/Install Software on BioHPC

Bioinformatics Facility
Biotechnology Resource Center
Cornell University
Software at BioHPC

- Many bioinformatics software packages pre-installed on BioHPC; new software can be installed on request.
- See the BioHPC software webpage for instructions on how to run individual software title;
- Email support@biohpc.cornell.edu if you would like software to be installed/updated;

Software Search Page:
https://biohpc.cornell.edu/lab/labsoftware.aspx

Instructions to Run Software

* We do not support module command to load software/pipeline.
To install software by yourself, we recommend to set up Conda environments in your home directory

- Anaconda for Python2
- Anaconda for Python3
- Miniconda for Python2
- Miniconda for Python3

- Light, no extra libraries;
- Python3 is more used now;
Where is software installed?

- On BioHPC servers, software installed in /programs;
- Software installed in /programs or your home directory are on network drive, and can be run from any BioHPC server except the login nodes.

Storage Locations:
Directories on local storage marked in green
BioHPC Network Storage in red
How to run software?

- Most software are not in the default path. You need to check the instructions page before running software.

```
export PATH=$HOME/miniconda2/bin:$PATH
```

User Installed Conda

Conda root:
```
export PATH=$HOME/miniconda2/bin:$PATH
bwa
```

Conda environment:
```
export PATH=$HOME/miniconda2/bin:$PATH
source activate bwa
bwa
```
Running software on BioHPC

- Use "screen" when running software;

- **Parallelization** of jobs:
  - perl_fork_univ.pl: perl_fork_univ.pl jobList 5
  - GNU parallel
  - Local scheduler (SLURM)
Supported containers:

Docker (a slightly restricted implementation for security)

https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=340#c

Singularity

https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=543#c
Other Software Related Information

(categorize by programming languages)
Some software require a specific version of library to compile.

Commonly used C libraries, including boost, HDF5 have multiple versions of libraries installed in /programs. Set LD_LIBRARY_PATH to use these libraries.

E.g.

- boost_1_46_1
- boost_1_55_0
- boost_1_57_0
- boost_1_58_0
- boost_1_62_0

Compilers:

Default: gcc v4.8.4

Alternative versions: 7.3.0, 5.5.0

To use alternative versions:

```bash
export LD_LIBRARY_PATH=/usr/local/gcc-7.3.0/lib:/usr/local/gcc-7.3.0/lib64
export PATH=/usr/local/gcc-7.3.0/bin:$PATH
```
**PYTHON**

**Python commands:**
- python : python v2.7.5
- python3 : python v3.6
- python2.7.15: python v2.7.15

(Corresponding pip commands: pip, pip3, pip2.7.15)

Updated python information can be found at:
[https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=556#c](https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=556#c)

**Jupyter Notebook:**
Implemented on each BioHPC server

To find more information about Jupyter Notebook on BioHPC, go to:
[https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=263#c](https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=263#c)
**R studio**

Both "R Studio Server" and "R Studio Desktop" are implemented on BioHPC.

To find more information about R studio on BioHPC, go to: [https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=199#c](https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=199#c)

---

**Commands:**

- **R**
  - Using parallel BLAS library
  - Location: /programs/R/bin/R
  - Using built-in BLAS library

---

For updated R information, see this page: [https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=37#c](https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=37#c)
Default Java version:

**Java 8 (jdk1.8.0_121)**

Update information can be found at:

https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=199#c

To use alternative versions

```bash
export JAVA_HOME=/usr/local/jdk1.7.0_51
export PATH=$JAVA_HOME/bin:$PATH
```
Use PERL5LIB to set library path

```
export PERL5LIB=/programs/PERL/lib/perl5
```

**$HOME**

- **$HOME/programs**
  - **/programs/PERL/lib**

- **/usr**
  - **usr/bin**
  - **usr/lib/perl5**

- **usr/local**
  - **bin**
  - **lib/perl5**

**v5.16**
- Distributed with system
- (perl in /usr/bin is alias to /usr/local/bin/perl)

**v5.22**
- Installed by BioHPC
- Default PERL

**BioHPC PERL modules**
Happy Computing!

The BioHPC Team

support@biohpc.cornell.edu