BioHPC: A High Performance Computing System for Life Sciences

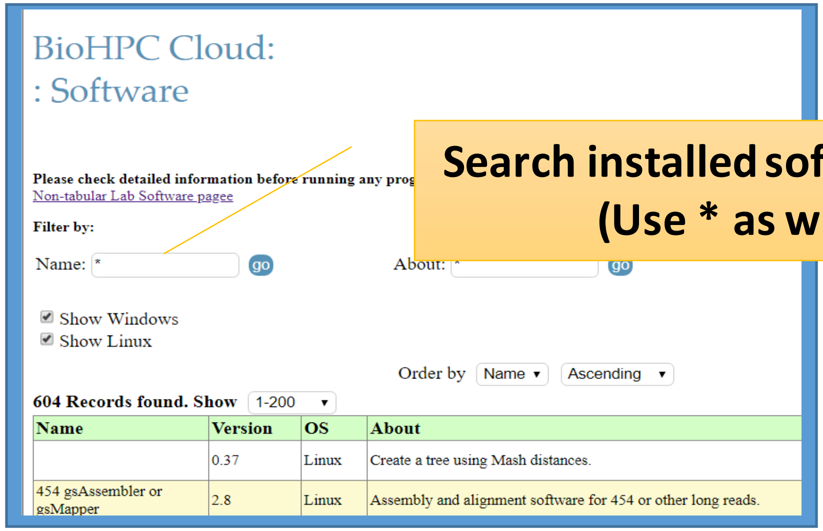
Run/Install Software on BioHPC

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



BioHPC Cloud:
: Software

Please check detailed information before running any prog
Non-tabular Lab Software pagee

Filter by:

Name: * go About: * go

☑ Show Windows
☑ Show Linux

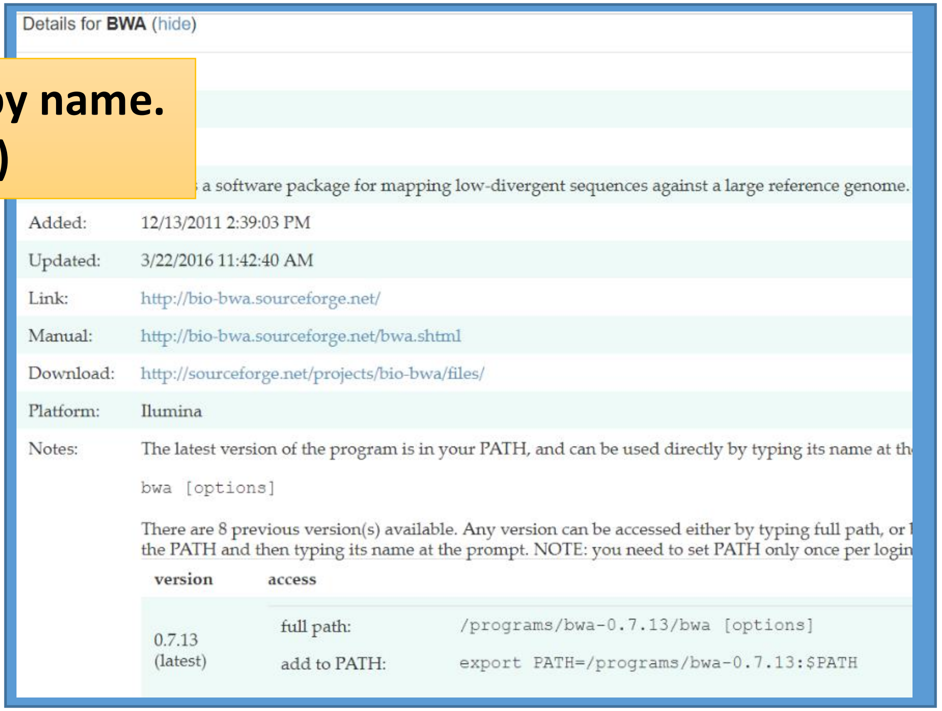
Order by Name ▾ Ascending ▾

604 Records found. Show 1-200 ▾

Name	Version	OS	About
	0.37	Linux	Create a tree using Mash distances.
454 gsAssembler or gsMapper	2.8	Linux	Assembly and alignment software for 454 or other long reads.

Search installed software by name.
(Use * as wild card)

Instructions to Run Software



Details for **BWA** (hide)

a software package for mapping low-divergent sequences against a large reference genome.

Added:	12/13/2011 2:39:03 PM
Updated:	3/22/2016 11:42:40 AM
Link:	http://bio-bwa.sourceforge.net/
Manual:	http://bio-bwa.sourceforge.net/bwa.shtml
Download:	http://sourceforge.net/projects/bio-bwa/files/
Platform:	Ilumina
Notes:	The latest version of the program is in your PATH, and can be used directly by typing its name at th

```
bwa [options]
```

There are 8 previous version(s) available. Any version can be accessed either by typing full path, or the PATH and then typing its name at the prompt. NOTE: you need to set PATH only once per login

version	access	
0.7.13 (latest)	full path:	`/programs/bwa-0.7.13/bwa [options]`
	add to PATH:	`export PATH=/programs/bwa-0.7.13:$PATH`

* 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

Miniconda for Python2

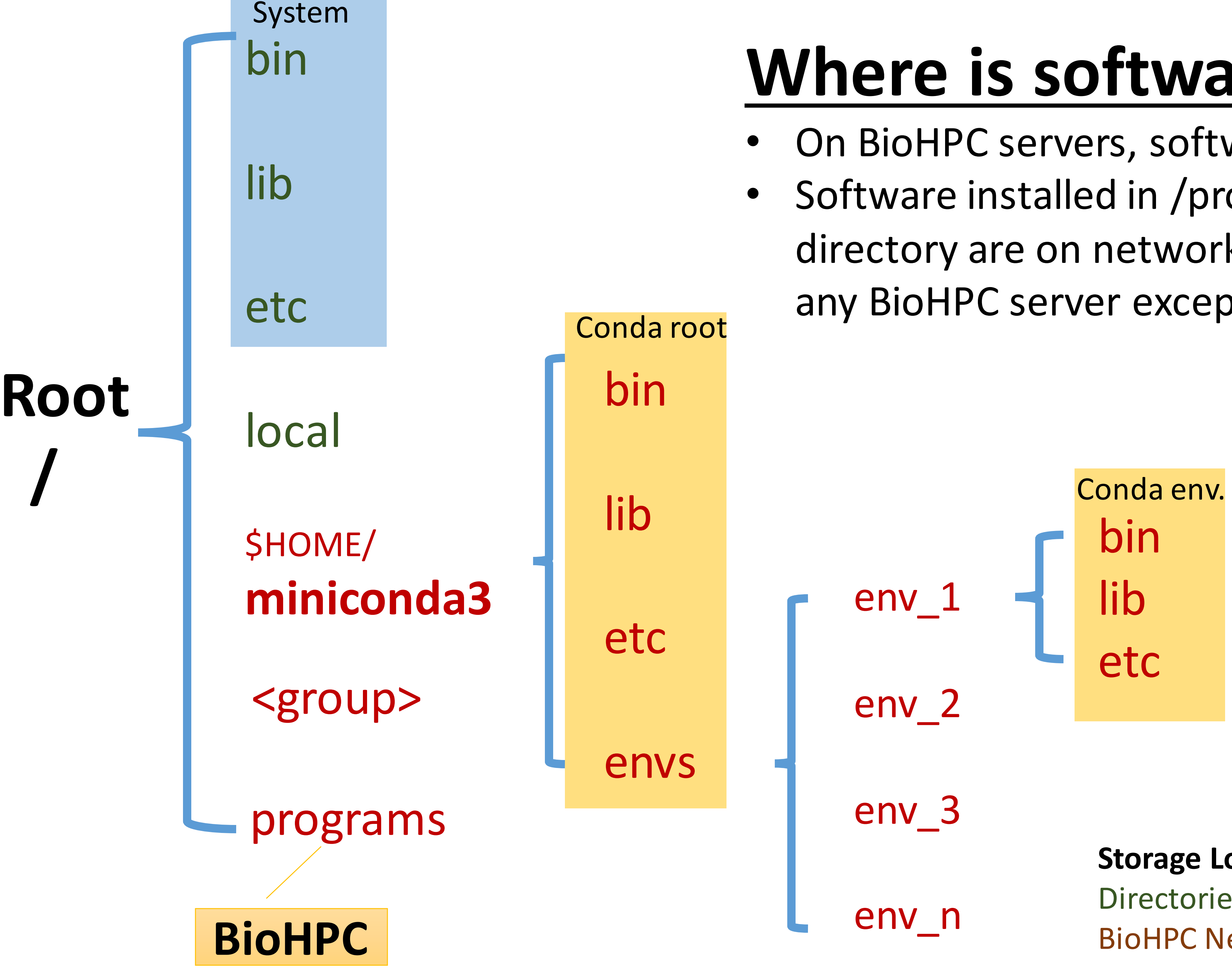
Anaconda for Python3

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.



Root

/

System

bin

lib

etc

local

$HOME/

miniconda3

<group>

programs

BioHPC

Conda root

bin

lib

etc

envs

env_1

env_2

env_3

env_n

Conda env.

bin

lib

etc

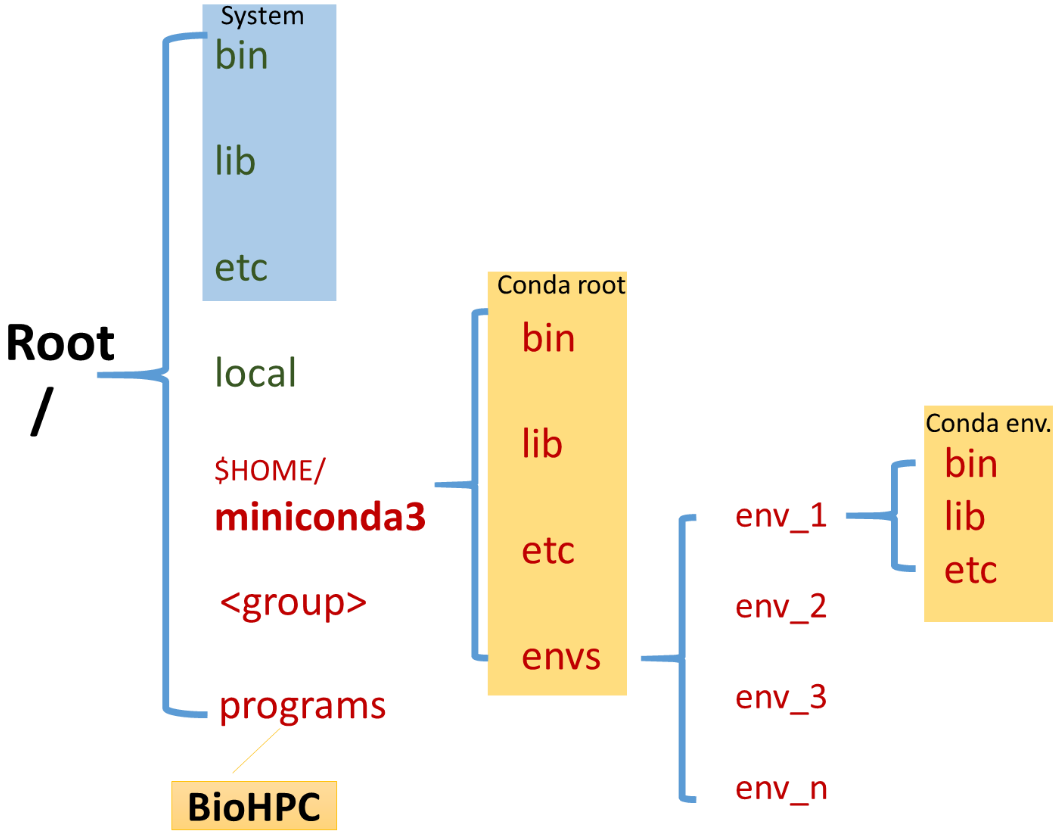
Storage Locations:

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



Root
/

System
bin
lib
etc

local

$HOME/
miniconda3

<group>

programs

BioHPC

Conda root
bin
lib
etc
envs

env_1
env_2
env_3
env_n

Conda env.
bin
lib
etc

BioHPC Software

```
bwa
```

or

```
/programs/bwa-0.7.8/bwa
```

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)

C

Compilers:

Default: gcc v4.8.4

Alternative versions:
7.3.0, 5.5.0

To use alternative versions:

```
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
```

Libraries:

- 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

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

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

- /
  - usr
    - bin
      - python2.7 (python); pip
      - Python3.6 (python3); pip3
    - lib
      - python2.7
      - Python3.6
    - lib64
      - python2.7
      - Python3.6
  - usr/local
    - bin
      - python2.7.15; pip2.7.15
    - Lib
      - python2.7
    - lib64
      - python2.7
  - $HOME
    - .local/lib
    - miniconda2
    - miniconda3
  - programs
    - pybedtools
    - Pysam
    - …
    - miniconda2
    - miniconda3
    - Anaconda2

R

Commands:

```
R
```

- Using parallel BLAS library
- Location: /programs/R/bin/R

```
/programs/R-3.5.0s/bin/R
```

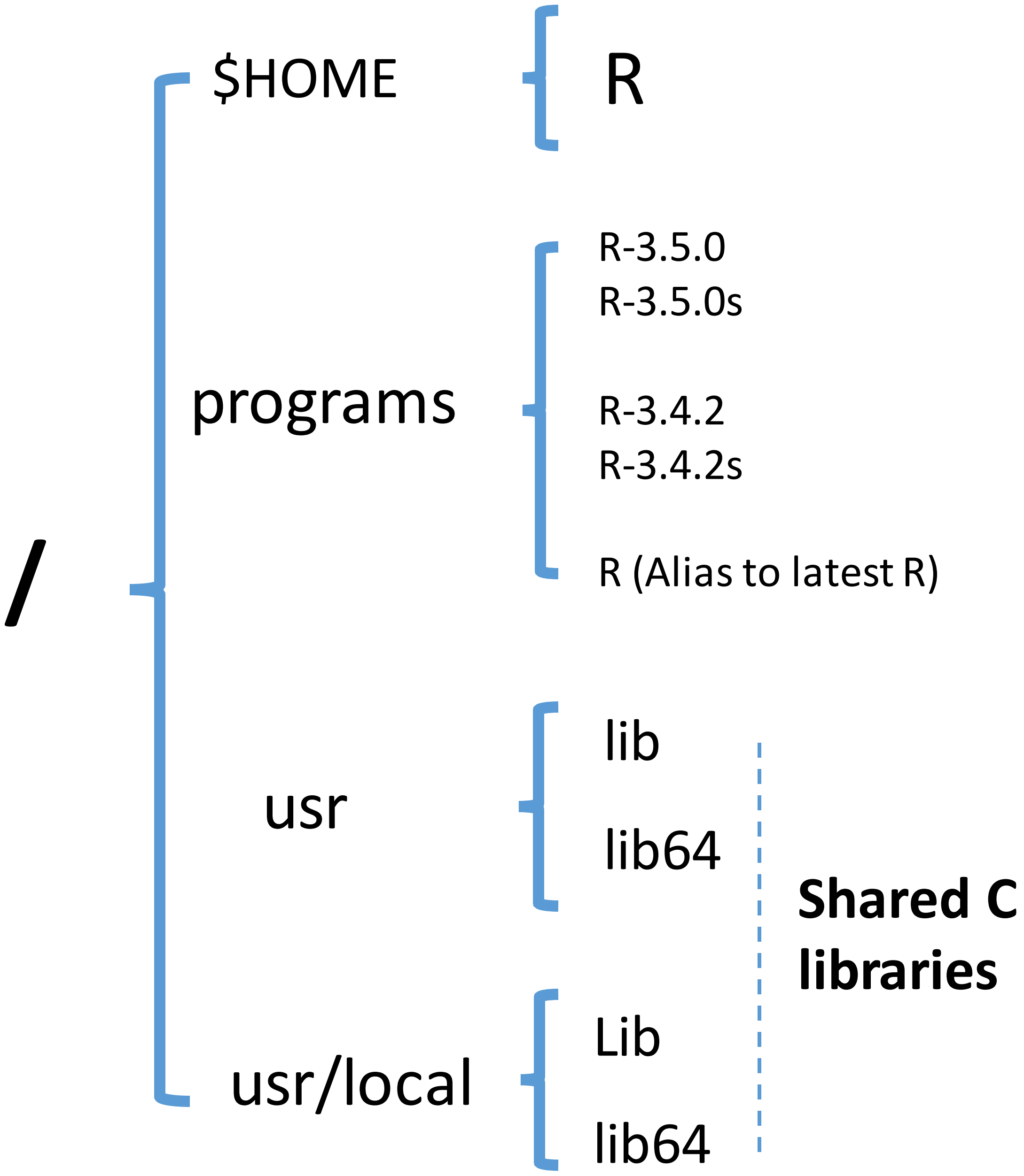
- Using built-in BLAS library

For updated R information, see this page:
https://biohpc.cornell.edu/lab/userguide.aspx?a=software&i=37#c

R studio

Both "RStudio Server" and "RStudio 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



/
- $HOME
  - R
- programs
  - R-3.5.0
  - R-3.5.0s
  - R-3.4.2
  - R-3.4.2s
  - R (Alias to latest R)
- usr
  - lib
  - lib64
- usr/local
  - Lib
  - lib64

Shared C libraries

JAVA

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

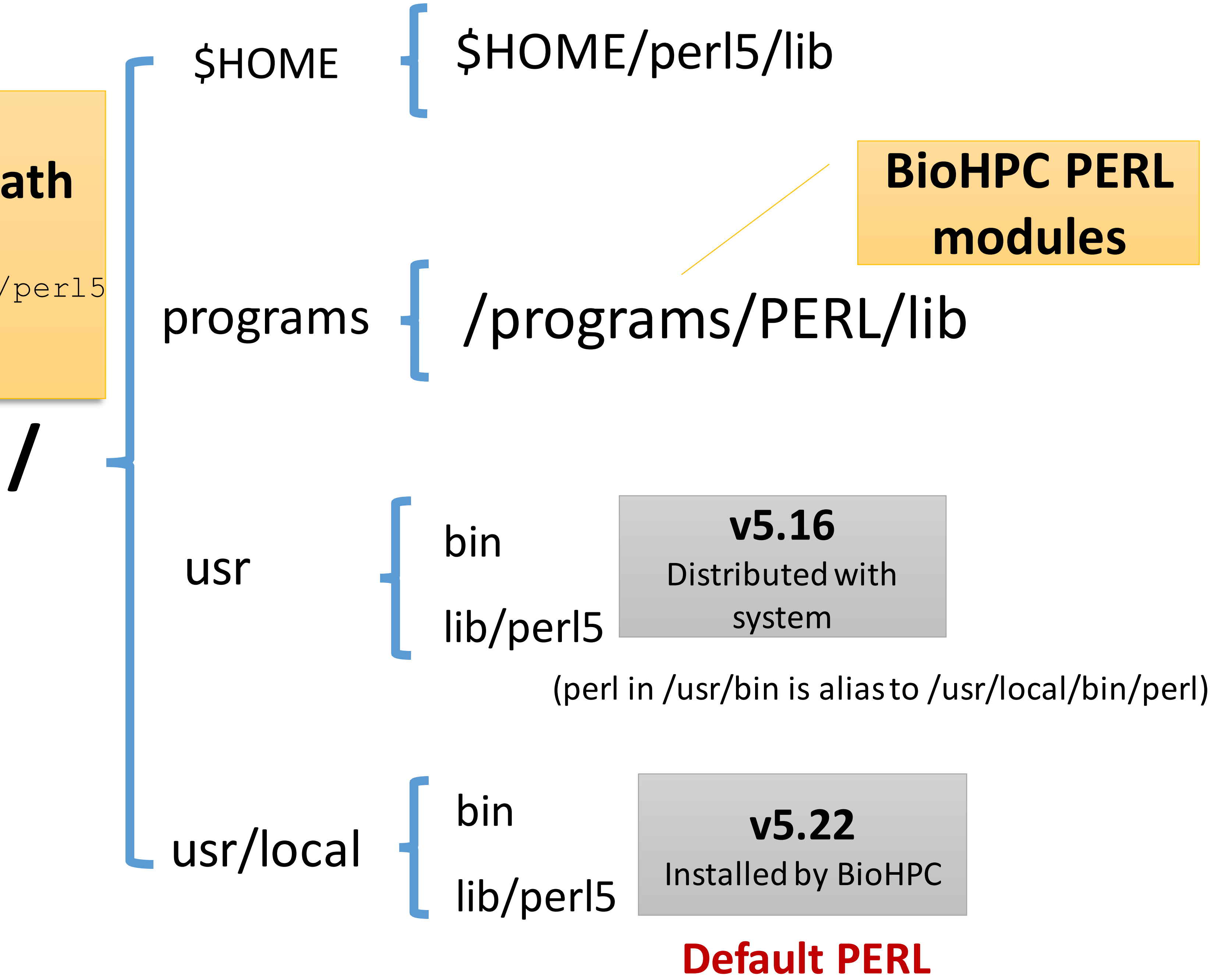
To use alternative versions

```
export JAVA_HOME=/usr/local/jdk1.7.0_51
export PATH=$JAVA_HOME/bin:$PATH
```

PERL

Use PERL5LIB to set library path

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



/

$HOME

$HOME/perl5/lib

programs

/programs/PERL/lib

BioHPC PERL modules

usr

bin

lib/perl5

v5.16
Distributed with system

(perl in /usr/bin is alias to /usr/local/bin/perl)

usr/local

bin

lib/perl5

v5.22
Installed by BioHPC

Default PERL

Happy Computing !

The BioHPC Team

support@biohpc.cornell.edu