

User Guide

I. System Requirements

1. UNIX or Linux operating system
2. jdk1.3.1 for unix/linux operating system
 - 1) Go to <http://java.sun.com>
 - 2) Select jdk1.3.1 for unix/linux operating system
 - 3) Click button for download
 - 4) Run setup program and follow instructions for installation. Normally put it in directory: `/usr/bin/java`
3. Exceed: If you have a windows machine, you need X window System to see the tree generated by java program. You can download it from <http://depot.berkeley.edu/software/index.html#XWindowSystem>

II. Installation

1. Create basic webpage and directories:
 - 1) Create a directory called `public_html`
 - 2) Under the `public_html` directory, create a file called `index.html`.
You will have a basic webpage.
 - 3) Under directory `public_html`, create two directories called `blast` and `cgi-bin`.
 - 4) Under directory `public_html`, create directories with name of organisms that you want to use (eg, *drosophila*, *celegans*, *mouse*, *zebrafish*, *xenopus*.)

For `blast` and those 5 organisms, you can either download by following the instructions below, or send e-mail to lyan@socrates.berkeley.edu for *.tgz files (totally 111.8M)

2. Download blast and organisms' protein database:
 - 1) Blast:
 - i. Under your blast directory: enter command:
`ftp ftp.ncbi.nih.gov`
 - ii. Login as anonymous and type following commands:
`cd blast`
`cd executables`
`get blast.linux.tar.Z`

The download will take one to two minutes.

iii. In your blast directory, type:

```
tar xzvf blast.linux.tar.Z
```

You will get all tools of blast, such as *blastall* and *fastacmd*.

2) Organisms:

i. In your linux ssh: type *netscape &*

ii. Go to <http://www.ncbi.nlm.nih.gov/Entrez/>

At option “*search*”: select “*protein*”. At “*for*”: input organism name you need. Then click “*limits*” and select “*organism*”. Click “*go*”, you will find a webpage showing records of this organism.

At option “*display*”, select “*FASTA*” and click “*save*”. You will be asked which directory you want to save it, save it under the organism you searched, with extension *.faa*

iii. After finish download, under that organism directory (eg, *drosophila*) type:

```
../blast/formatdb -p T -i drosophila.faa -o T
```

iv. Repeat above steps for each organism.

3) Change the permission of *cgi-bin*:

```
chmod 777 cgi-bin
```

3. Go to directory *cgi-bin*, download and unzip *bp.tgz*, which has 5 files: *.htaccess*, *blast.htm*, *blast.cgi*, *bp.pl9*, *tree*; and *javacode.zip*, which 6 files:

Metric.java, *Phylogenetics.java*, *intree.java*, *OrthoTable.java*, *dismat2.java* and *newtree.java*.

4. Put this file called *.htaccess* in your *public_html* directory:

```
Order allow, deny
```

```
allow from all
```

5. Now open your webpage and go to *cgi-bin/blast.htm*, you should find a webpage similar to this one: <http://sahara.lbl.gov/~lyan/cgi-bin/blast.htm>

6. Fill in form and press the “submit” button when done. For example inputs, check *sample_input.txt*

7. After pressing the “submit” button and allow the programs finish, you will find those files in your *cgi-bin* directory:

```
org.file.txt
```

```
Proteinsequences
```

```
Proteinnames
```

Pathwayname
threshold
gabe.txt
**.seq*
**_out*

8. Run exceed and enter the command: *perl tree*
A phylogenetic tree will show up.