Linux

Saba Ghaffari
Edited by: Negin Valizadegan
Linux commands

Using ClustalW to align two sequences
Step 0: Connect to VPN & Open MobaXterm
Step 1A: Accessing the IGB Biocluster

• Open **MobaXterm** on your desktop

• In a new session, select **SSH** and type the following host name:

  biologin.igb.illinois.edu

• Click **OK**
Step 1A: Accessing the IGB Biocluster

• Enter login credentials assigned to you.

• Example username: **class100**.

• You will not see any characters on screen when typing in password. Just type it.
• MobaXterm 10.5 •
  (SSH client, X-server and networking tools)

> SSH session to class101@biologin.igb.illinois.edu
  • SSH compression : ✓
  • SSH-browser : ✓
  • X11-forwarding : ✓ (remote display is forwarded through SSH)
  • DISPLAY : ✓ (automatically set on remote server)

> For more info, ctrl+click on help or visit our website

Last login: Fri Jun 18 10:43:45 2021 from phalanx.igb.illinois.edu

# Institute for Genomic Biology
# University of Illinois Urbana-Champaign
# http://biocluster.igb.illinois.edu

*Please follow the guide at http://help.igb.illinois.edu/Biocluster
*All data on this cluster is NOT backed up. It costs $8.75 per terabyte per month
*HIPAA data is not allowed on the biocluster
*Please email help@igb.illinois.edu with any questions

[class101@biologin-1 ~]$
Step 1B: Listing files and directories (ls)

$ ls
# listing files in your current directory. When you first login, your directory is your home directory.

$ = command prompt
# = notes
Step 1C: Making Directories (mkdir)

$ mkdir ~/01_Linux
# create a subdirectory in your home directory. The tilde ~ character refers to your home directory.
$ ls
# to see the directory you just created.
Step 1D: Changing directory (cd)

The lab is located in the following directory:

```
/home/classroom/mayo/2020/01_Linux
```

$ cd /home/classroom/mayo/2020/01_Linux

# tip: use "tab" for auto-completion for path

$ ls

# to see the contents. You should see seqs.fa

Step 1E: Print working directory (pwd)

```
$pwd
```

# to see the full pathname. You should see

"/home/classroom/mayo/2020/01_Linux"
Step 1F: Copying files (cp)

Copy seqs.fa from the data directory to your working directory.

```
$ cp /home/classroom/mayo/2020/01_Linux/seqs.fa ~/01_Linux/
# tip: use “tab” for autocompletion for path
$ cd ~/01_Linux/
```

Step 1G: Displaying the contents of a file on the screen (cat)

```
$ cat seqs.fa
# you should see two sequences on your screen
>seq1
GATCGAGCGATCGTGCAGC
GCAGAATGCGCGCTAG
>seq2
GGTAGGGTAAATTGCCTAC
CGTCGATCGAGTA
```
## Commands Summary

<table>
<thead>
<tr>
<th>Command</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ls</code></td>
<td>list files and directories</td>
</tr>
<tr>
<td><code>mkdir directory</code></td>
<td>make a directory</td>
</tr>
<tr>
<td><code>cd directory</code></td>
<td>change to named directory</td>
</tr>
<tr>
<td><code>cd ~</code></td>
<td>change to home directory</td>
</tr>
<tr>
<td><code>cd ..</code></td>
<td>change to parent directory</td>
</tr>
<tr>
<td><code>pwd</code></td>
<td>display the path of the current directory</td>
</tr>
<tr>
<td><code>cp file1 file2</code></td>
<td>cp file1 and call it file2</td>
</tr>
<tr>
<td><code>cat file</code></td>
<td>display the contents of a file</td>
</tr>
</tbody>
</table>
Useful tips

<table>
<thead>
<tr>
<th>Command</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>tab</td>
<td>auto-complete path</td>
</tr>
<tr>
<td>↑</td>
<td>retrieve previous commands</td>
</tr>
</tbody>
</table>
Step 1H: Run sequence alignment program

Accessing the IGB Biocluster
Step 1H: Run sequence alignment program

```
$ srun -p classroom -c 2 --mem 8000 --pty bash
# Open interactive session on biocluster with 2 cpus and 8G memory.

$ module load ClustalW2       # Load sequence aligner into the shell environment.

$ module list #See loaded tools

$ clustalw2 -INFILE=seqs.fa   # Run the clustalw sequence aligner.
```
Step 1H: Run sequence alignment program

You will see this on your screen, when the program is done.

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence format is Pearson
Sequence 1: seq1 35 bp
Sequence 2: seq2 32 bp
Start of Pairwise alignments
Aligning...

Sequences (1:2) Aligned. Score: 21
Guide tree file created: [seqs.dnd]

There are 1 groups
Start of Multiple Alignment

Aligning...
Group 1: Delayed
Alignment Score 47
Step 1H: Run sequence alignment program

The alignment result is in `seqs.aln`. Use `cat` command to see the result.

```
$ cat seqs.aln
# You should see the following on your screen.
CLUSTAL 2.1 multiple sequence alignment
seq1  GATCGAGCGA-TCGTGCAGCAGAATCGCGCTAG
seq2  GGTAGGGTAATTGCTACCGTGATCGAGTA----
       * * * *  * * *   * **  **  * *
```
Exit **MobaXterm** by either of the following:

- Close the window

- Type ‘exit’ in the command prompt twice and then press <return>.

  $ exit
  $ exit