

Linux

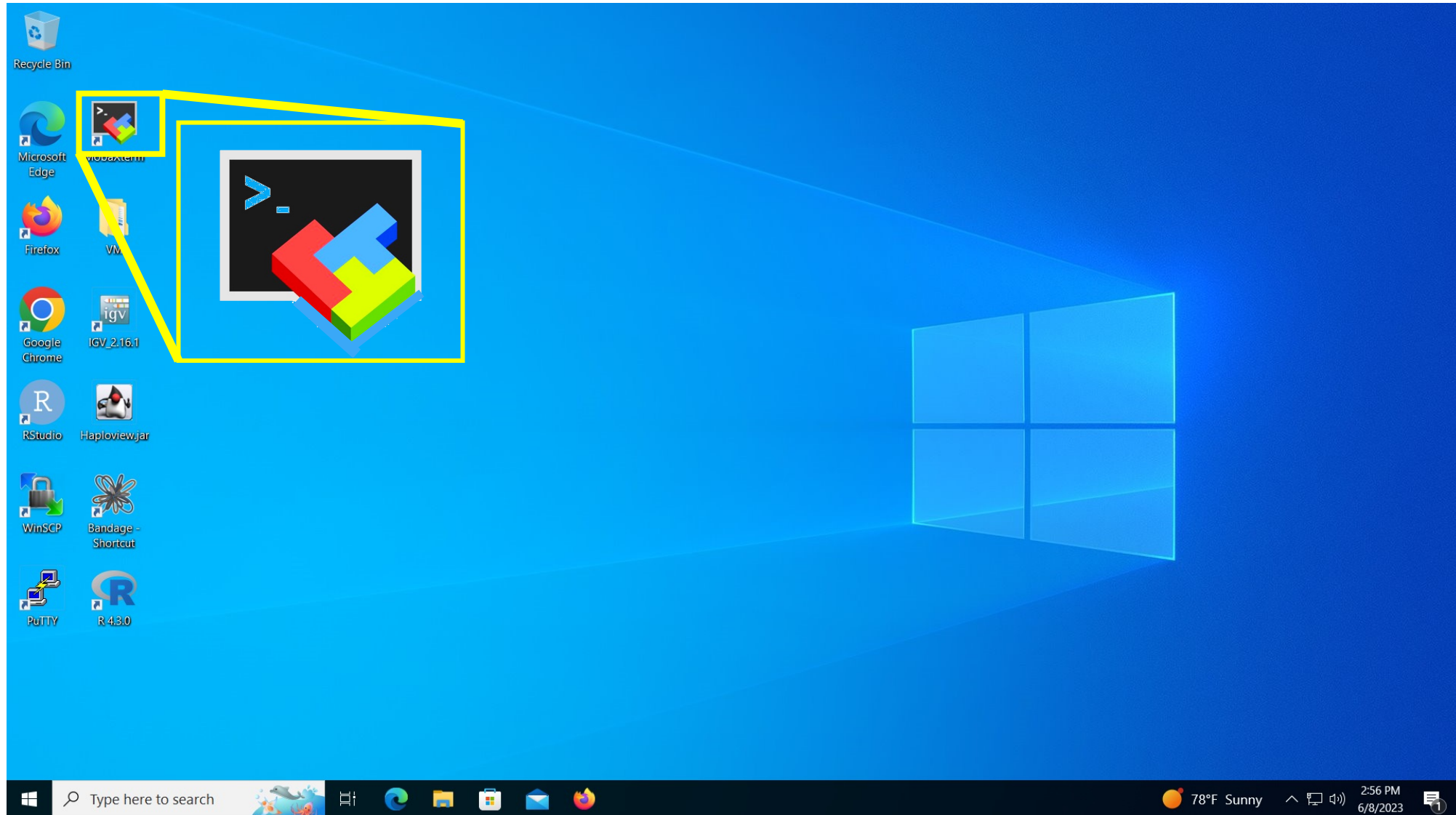
Saba Ghaffari

Edited by: Gio Madrigal & Roberto Cucalón Tamayo

Linux commands

Using ClustalW to align two sequences

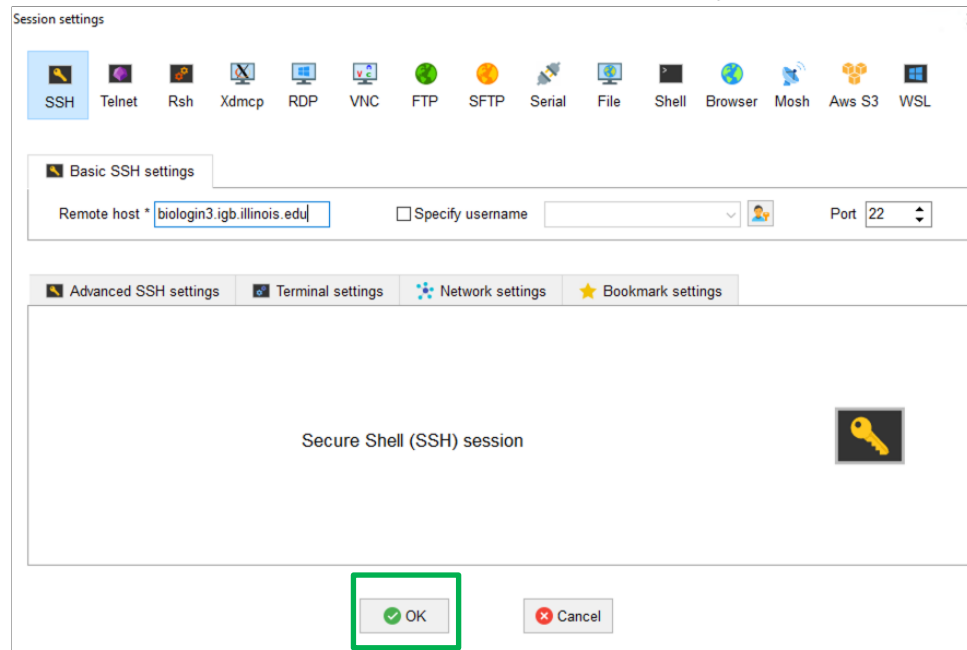
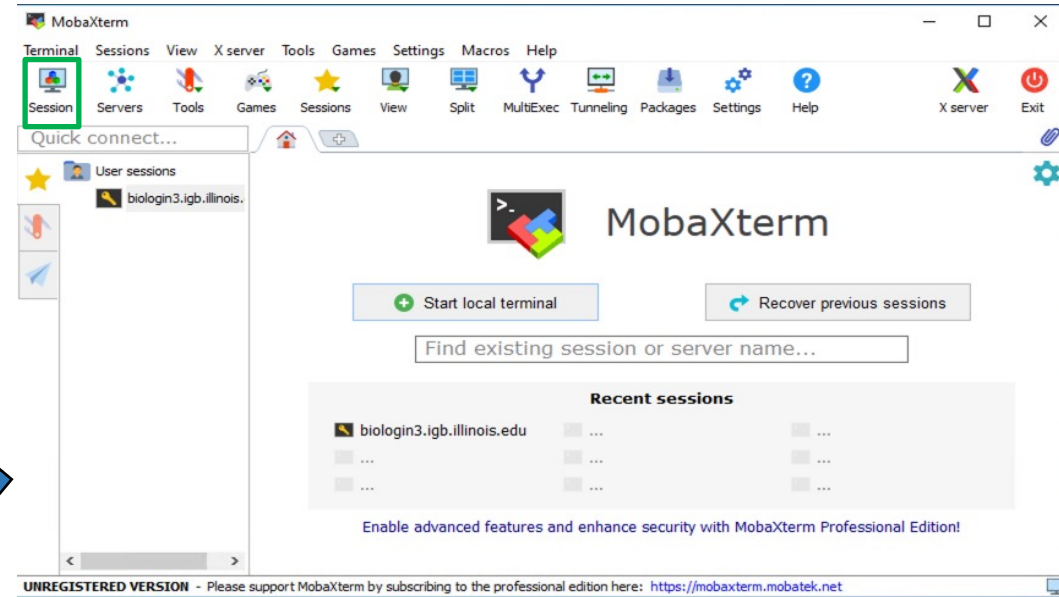
Step 0: Open MobaXterm from VM



Step 0A: Accessing the IGB Biocluster for First Time

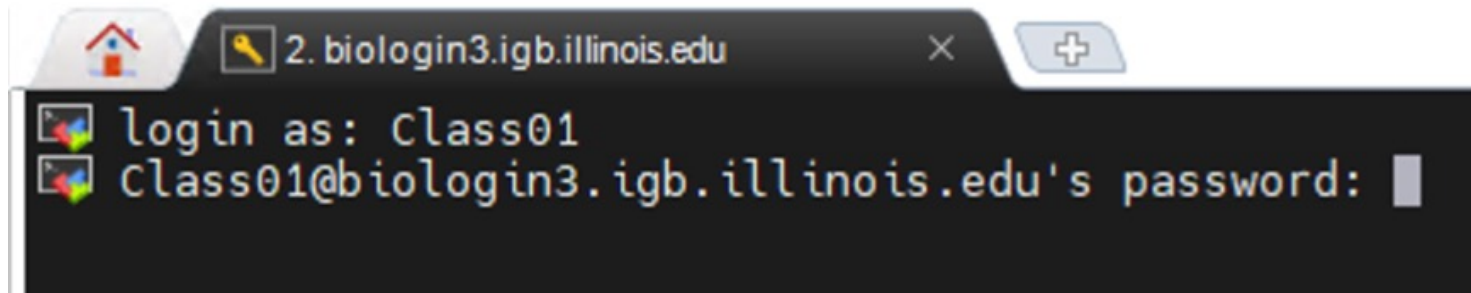
- Open **MobaXterm** from the VM
- In a new session, select **SSH** and type the following host name:

`biologin3.igb.illinois.edu`
- Click **OK**



Step 0A: Accessing the IGB Biocluster

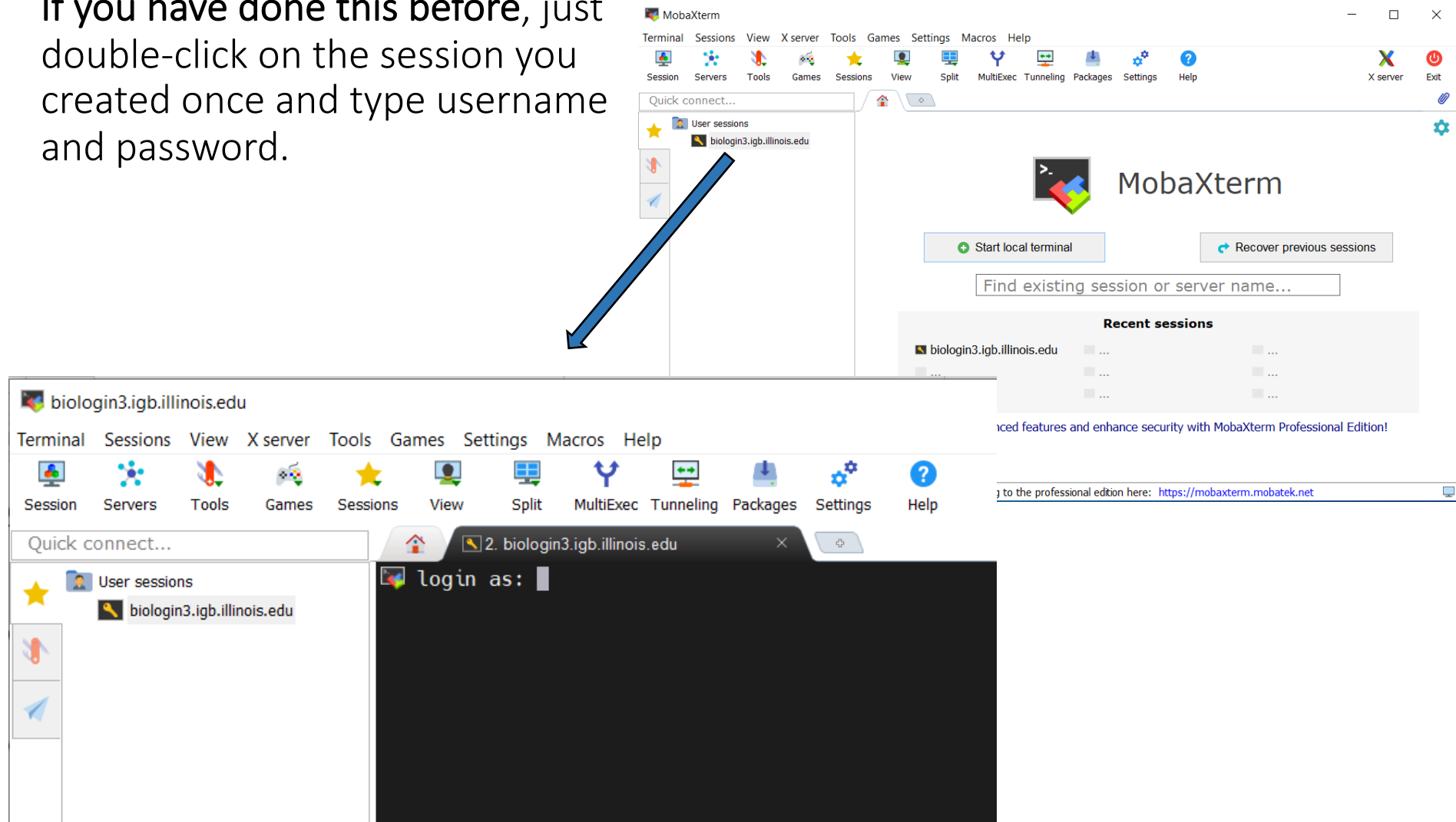
- Enter login credentials assigned to you.
- Example username: **Class01**
- You will not see any characters on screen when typing in password. Just type it.



```
login as: Class01
Class01@biologin3.igb.illinois.edu's password: █
```

Step 0A: Accessing the IGB Biocluster

If you have done this before, just double-click on the session you created once and type username and password.



• MobaXterm Personal Edition v23.1 •
(SSH client, X server and network tools)

- ▶ SSH session to [Class02@biologin3.igb.illinois.edu](https://class02@biologin3.igb.illinois.edu)
 - Direct SSH : ✓
 - SSH compression : ✓
 - SSH-browser : ✓
 - X11-forwarding : ✓ (remote display is forwarded through SSH)
- ▶ For more [info](#), ctrl+click on [help](#) or visit our [website](#).

```
#####  
#                                                                 #  
#           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  
  
Last login: Wed May 31 11:27:24 2023 from 128.174.127.200  
ln: failed to create symbolic link './dropbox': File exists  
Created dropbox symbolic link  
[Class02@biologin-2 ~]$ █
```

Step 1B: Listing files and directories (ls)

```
*Please email help@igb.illinois.edu with any questions
```

```
[class101@biologin-1 ~]$
```

```
$ 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 auto-completion 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
```

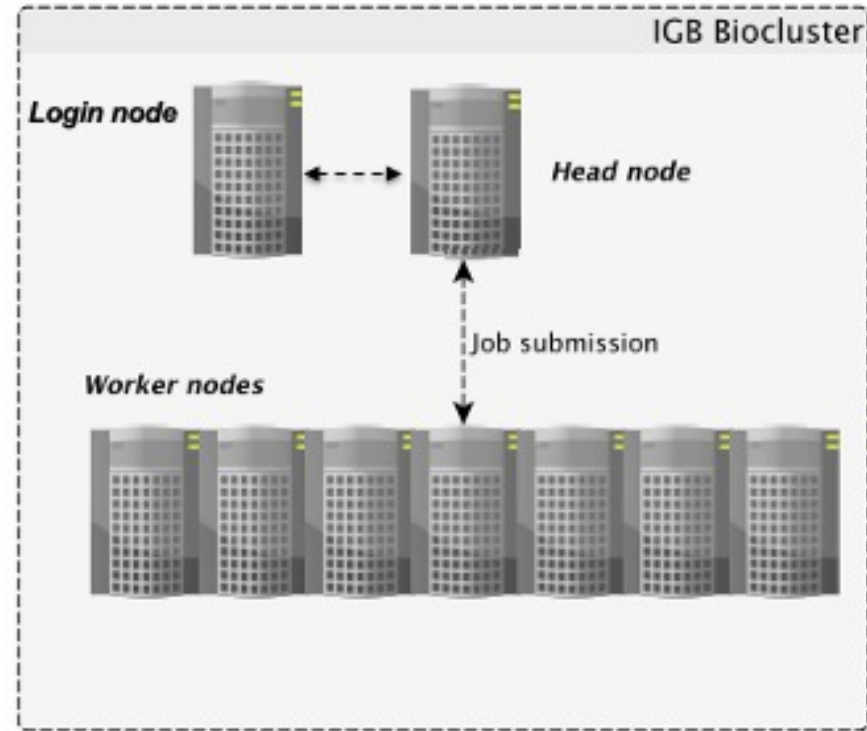
Commands Summary

Command	Meaning
ls	list files and directories
mkdir directory	make a directory
cd directory	change to named directory
cd ~	change to home directory
cd ..	change to parent directory
pwd	display the path of the current directory
cp file1 file2	cp file1 and call it file2
cat file	display the contents of a file

Useful tips

Command	Meaning
tab	auto-complete path
↑	retrieve previous commands

Step 1H: Run sequence alignment program



Chris Fields

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

CLUSTAL-Alignment file created [seqs.aln]
```


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-TCGTGCAGCGCAGAATGCGCGCTAG
seq2          GGTAGGGTAAATTGCCTACCGTCGATCGAGTA----
              * * * * * * * * * * * * * * * * * *
```

Exit **MobaXterm** by either of the following:

- Close the window
- Type 'exit' in the command prompt **twice** and then press <return>.

```
$ exit # first to exit interactive session  
$ exit # again to exit the login node  
<return>
```