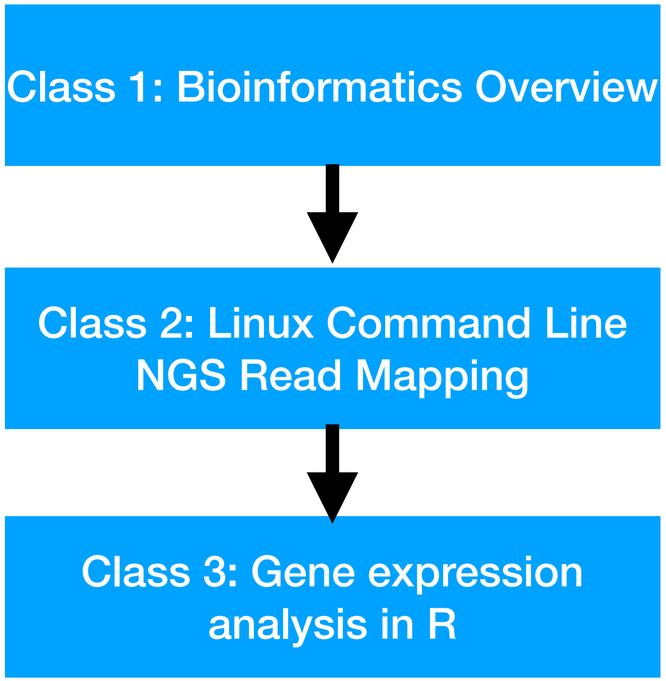
Class 2a: Linux Command Line

June 21, 2023

Course Overview

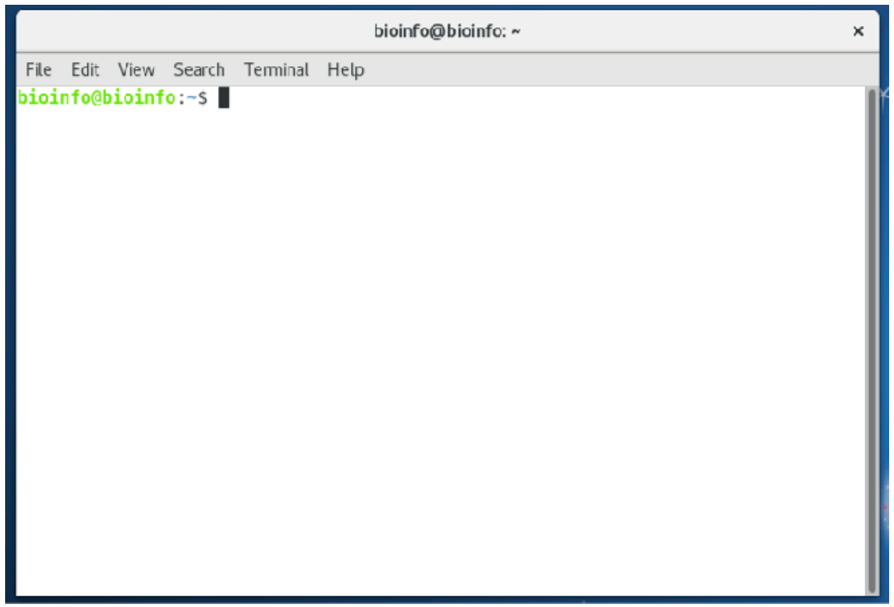


Linux Command Line
Class Content

- Introduction to the Virtual Machine
- Shells, Terminals, & file system navigation
- Anatomy of a UNIX command
- Wildcards, shortcuts and special characters
- File permissions
- Compression UNIX commands
- Networking UNIX commands

What is a Terminal?

- A terminal is a textual interface for interacting with a computer (a shell!)
- Using the terminal, one can issue powerful and concise command-line instructions for the computer to follow.

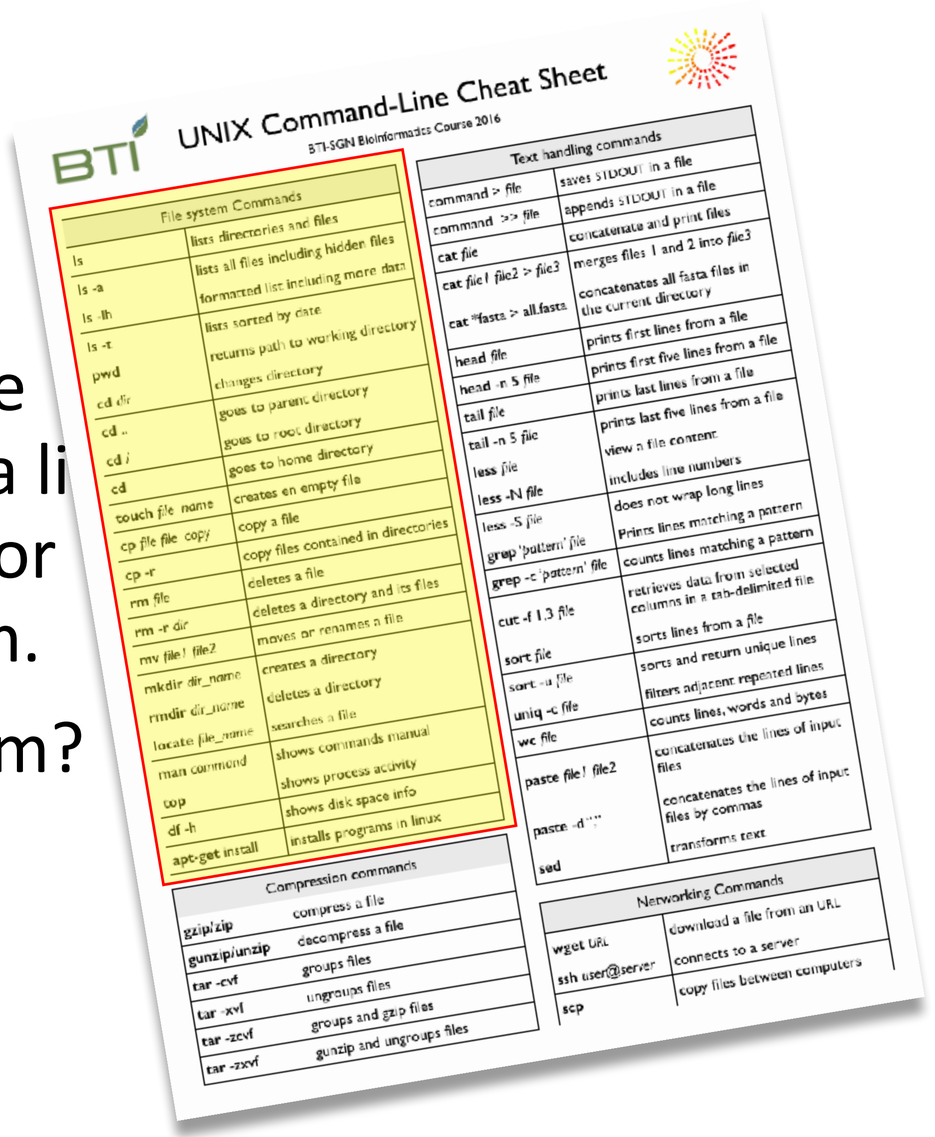


Why use the command-line (terminal)?

- Most software for biological big data analysis is used through UNIX command-line operations.
- Most of the servers for biological data analysis use Linux/Unix as their operating system.
- Data analysis on calculation servers are much faster since we can use more CPUs and RAM than in a PC or laptop (*e.g.* BTI's "Boyce" server has 64 cores and 1TB RAM).
- Large NGS data files can not be opened or loaded in most graphical software and web sites.
- App development is often harder and takes longer when using GUIs

Command-line File System Navigation

- The cheat sheet you have been provided contains a li of common commands for navigating the file system.
- But what *is* the file system?



BTI

UNIX Command-Line Cheat Sheet

BTI-SGN Bioinformatics Course 2016

File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd dir	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch file_name	creates an empty file
cp file file_copy	copy a file
cp -r	copy files contained in directories
rm file	deletes a file
rm -r dir	deletes a directory and its files
mv file1 file2	moves or renames a file
mkdir dir_name	creates a directory
rmdir dir_name	deletes a directory
locate file_name	searches a file
man command	shows commands manual
top	shows process activity
df -h	shows disk space info
apt-get install	installs programs in linux

Text handling commands	
command > file	saves STDOUT in a file
command >> file	appends STDOUT in a file
cat file	concatenate and print files
cat file1 file2 > file3	merges files 1 and 2 into file3
cat *fasta > all.fasta	concatenates all fasta files in the current directory
head file	prints first lines from a file
head -n 5 file	prints first five lines from a file
tail file	prints last lines from a file
tail -n 5 file	prints last five lines from a file
less file	view a file content
less -N file	includes line numbers
less -S file	does not wrap long lines
grep 'pattern' file	Prints lines matching a pattern
grep -c 'pattern' file	counts lines matching a pattern
cut -f 1,3 file	retrieves data from selected columns in a tab-delimited file
sort file	sorts lines from a file
sort -u file	sorts and return unique lines
uniq -c file	filters adjacent repeated lines
wc file	counts lines, words and bytes
paste file1 file2	concatenates the lines of input files
paste -d ";"	concatenates the lines of input files by commas
sed	transforms text

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

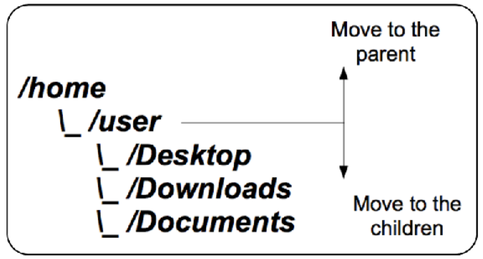
Networking Commands	
wget URL	download a file from an URL
ssh user@server	connects to a server
scp	copy files between computers

Connect to remote server ‘thompson’

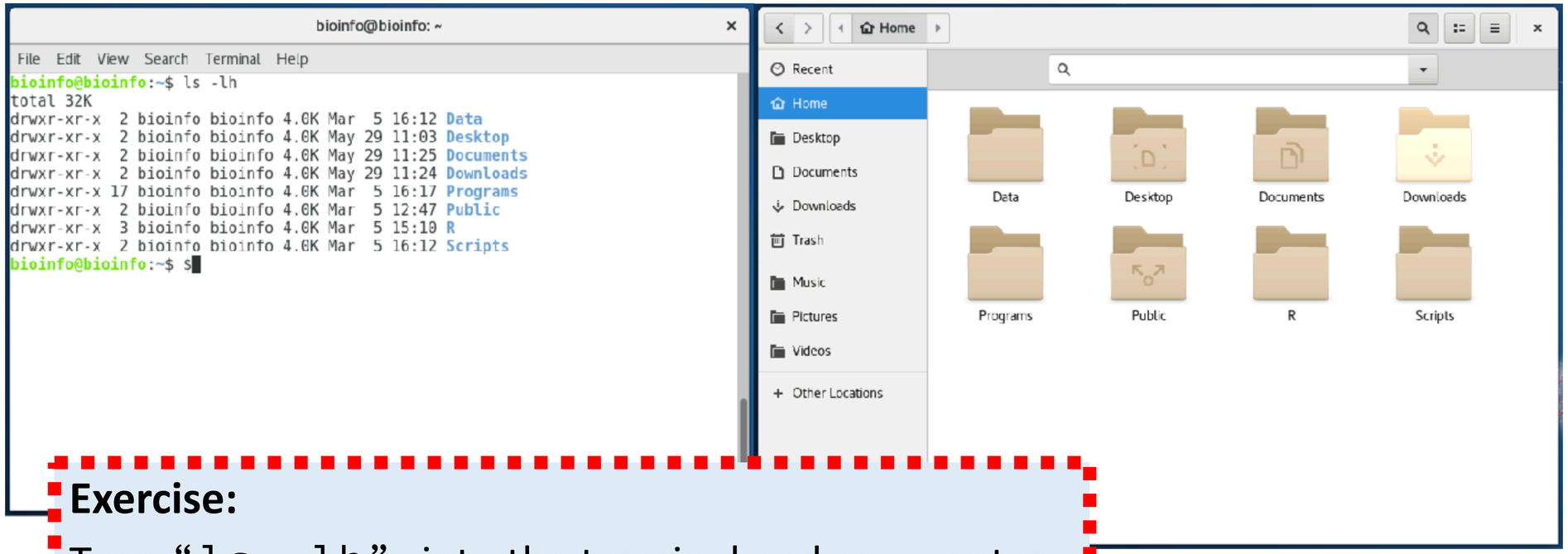
- Open terminal or Putty and type the following (using your username)

ssh bioinfo0@thompson.sgn.cornell.edu

- #The password is bioinfo00



The File System



Exercise:

Type “`ls -lh`” into the terminal and press enter.

Home and Root Directories

```
Noes-MacBook-Pro:~ Noe$ ls -lht
total 0
drwx------+ 29 Noe  staff   986B May 31 11:24 Desktop
drwx------@  8 Noe  staff   272B May 31 08:26 Dropbox
drwx------+ 54 Noe  staff   1.8K May 30 16:01 Downloads
drwx------+  8 Noe  staff   272B May 28 21:06 Pictures
drwxr-xr-x  18 Noe  staff   612B May 17 11:12 BTI
drwxr-xr-x   5 Noe  staff   170B May  8 11:44 programs
drwx------+ 15 Noe  staff   510B Apr 10 08:33 Documents
drwxr-xr-x   6 Noe  staff   204B Mar 18 09:22 VirtualBox VMs
drwxr-xr-x   8 Noe  staff   272B Mar 14 19:26 py_devel
drwx------@ 51 Noe  staff   1.7K Mar 11 15:08 Library
```

Home directory

/home/bioinfo
/home/noe
/home/noe/Desktop

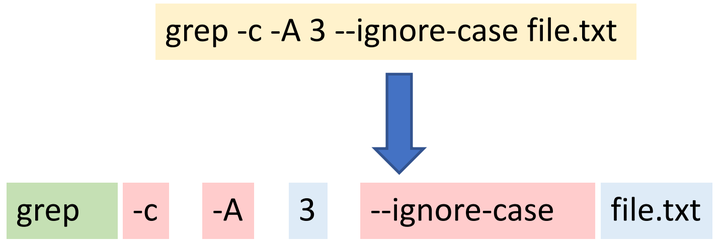
```
noe@debian-virtualbox:~$ ls -l /
total 108
drwxr-xr-x   2 root root  4096 Sep 26  2012 bin
drwxr-xr-x   3 root root  4096 Nov  9  2012 boot
drwxr-xr-x  15 root root  3140 May 31 12:46 dev
drwxr-xr-x 130 root root 12288 May 31 12:45 etc
drwxr-xr-x   5 root root  4096 Feb 28 13:54 export
drwxr-xr-x   4 root root  4096 Nov  7  2012 home
lrwxrwxrwx   1 root root    30 Sep 26  2012 initrd.img
drwxr-xr-x  12 root root 12288 Nov  9  2012 lib
drwxr-xr-x   2 root root 12288 Nov  9  2012 lib32
lrwxrwxrwx   1 root root     4 Sep 26  2012 lib64 -> /l
drwx------   2 root root 16384 Sep 26  2012 lost+found
drwxr-xr-x   3 root root  4096 Sep 26  2012 media
drwxr-xr-x   2 root root  4096 May  1  2012 mnt
drwxr-xr-x   2 root root  4096 Sep 26  2012 opt
dr-xr-xr-x 134 root root     0 May 31 12:45 proc
drwx------  10 root root  4096 Nov 15  2012 root
drwxr-xr-x   2 root root  4096 Nov  9  2012 sbin
drwxr-xr-x   2 root root  4096 Jul 21  2010 selinux
drwxr-xr-x   2 root root  4096 Sep 26  2012 srv
drwxr-xr-x  13 root root     0 May 31 12:45 sys
drwxrwxrwt  11 root root  4096 May 31 19:56 tmp
drwxr-xr-x  11 root root  4096 Sep 26  2012 usr
drwxr-xr-x  14 root root  4096 Sep 26  2012 var
```

Root directory

/bin, /lib, /usr	code and code libraries
/var	logs and other data
/home	user directories
/tmp	temporary files
/etc	configuration information
/proc	special file system in Linux

Anatomy of a UNIX Command

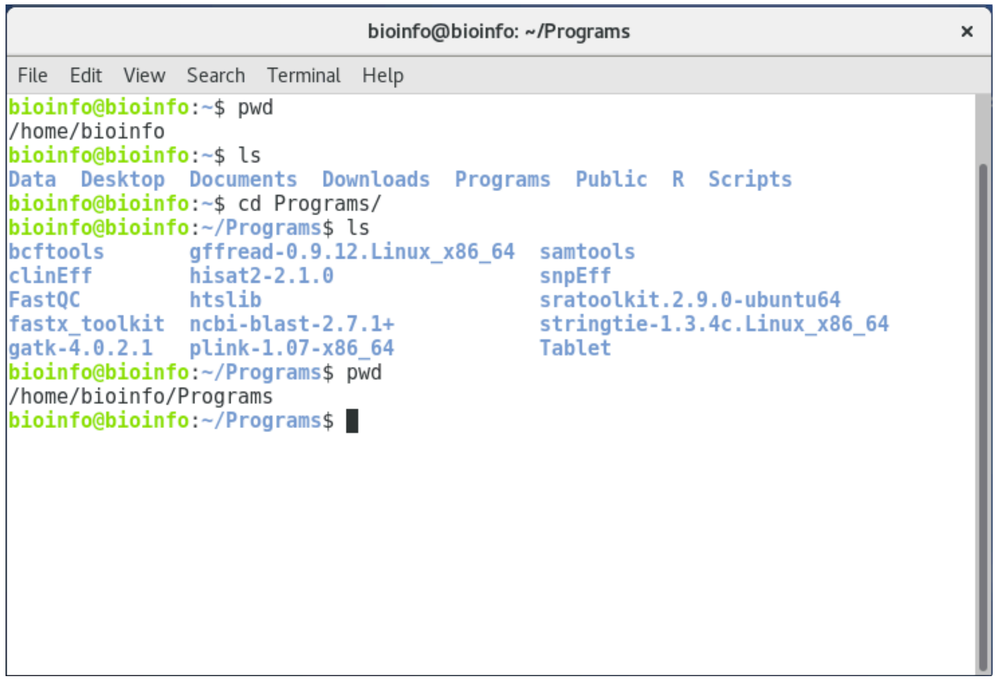
- Every UNIX command is made up of a series of space-separated strings.
- The first of these strings is always the command you will run.
- The user can also provide options (shown in red) and arguments (shown in blue).
- A command can require arguments, as can an option.
- In this example, the “-A” option is being given the argument “3” and the command is being given the argument “file.txt”.



- Options often have a short form (starting with “-”) and a long form (starting with “--”).
- If an option flag does not have any arguments, then can often be concatenated (*e.g.* “ls -h -l -a” is equivalent to “ls -hla”).

pwd, ls, and cd

- **pwd** — where am I?
- **ls** — what is inside the current directory?
- **cd** — change my current directory.

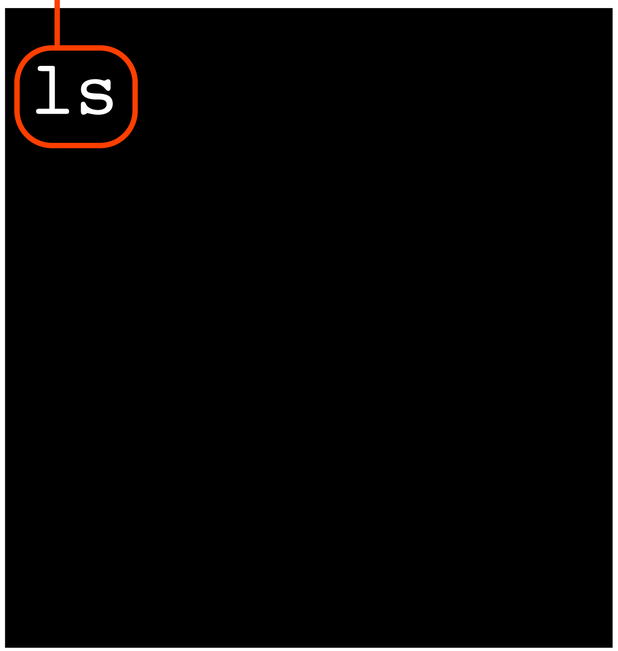


Exercise:

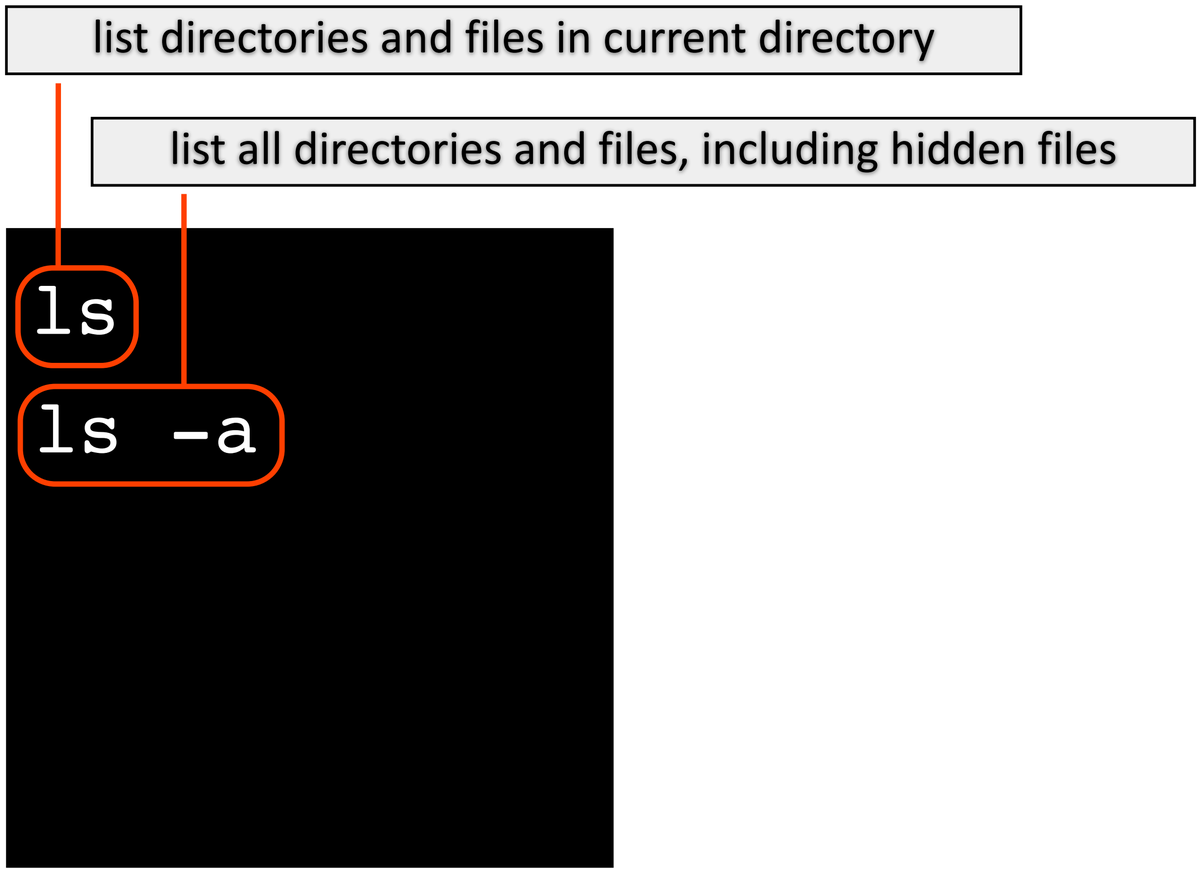
Change into the “Programs” directory and list the files found there.

ls lists files and directories

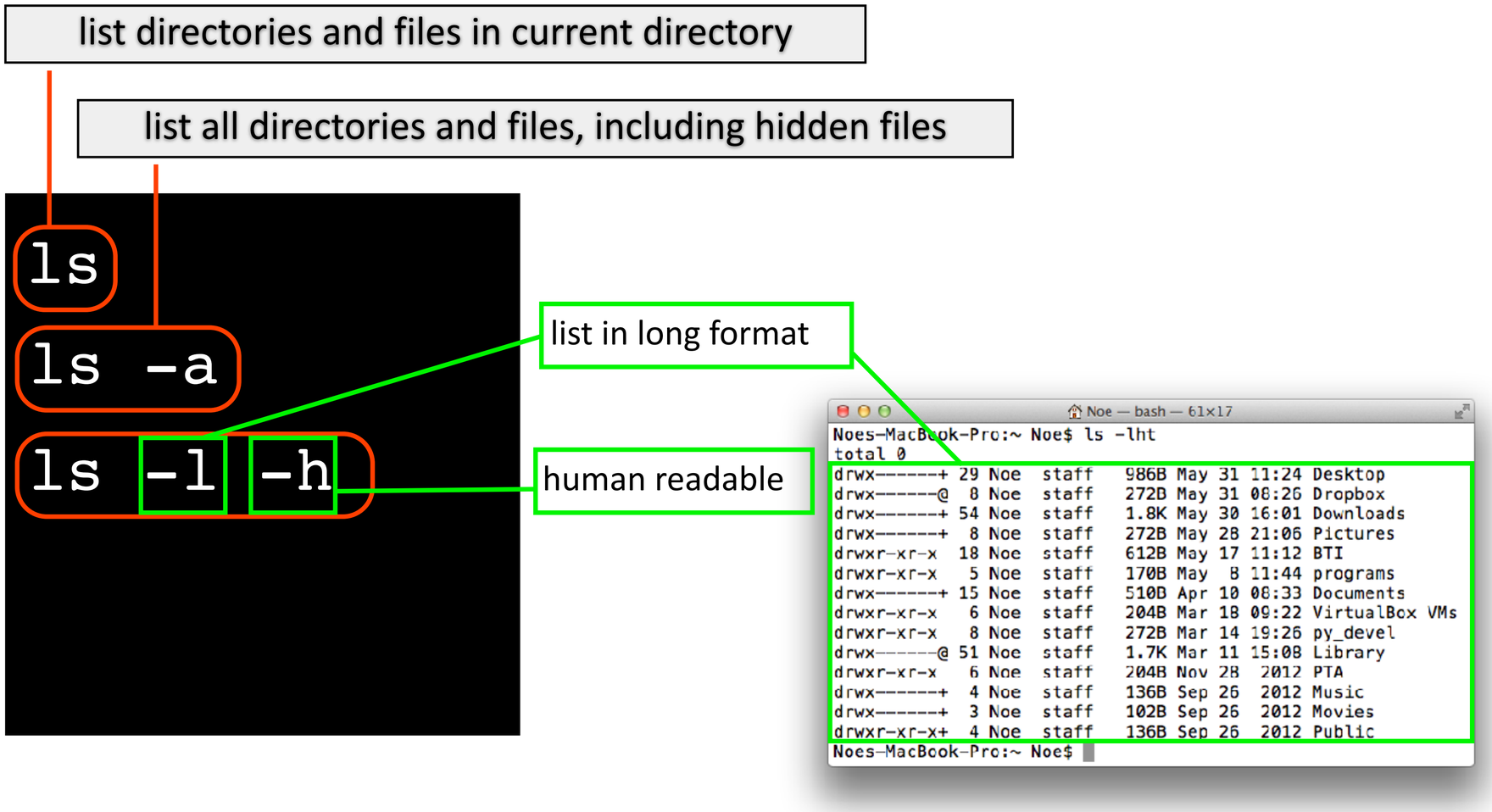
list directories and files in current directory



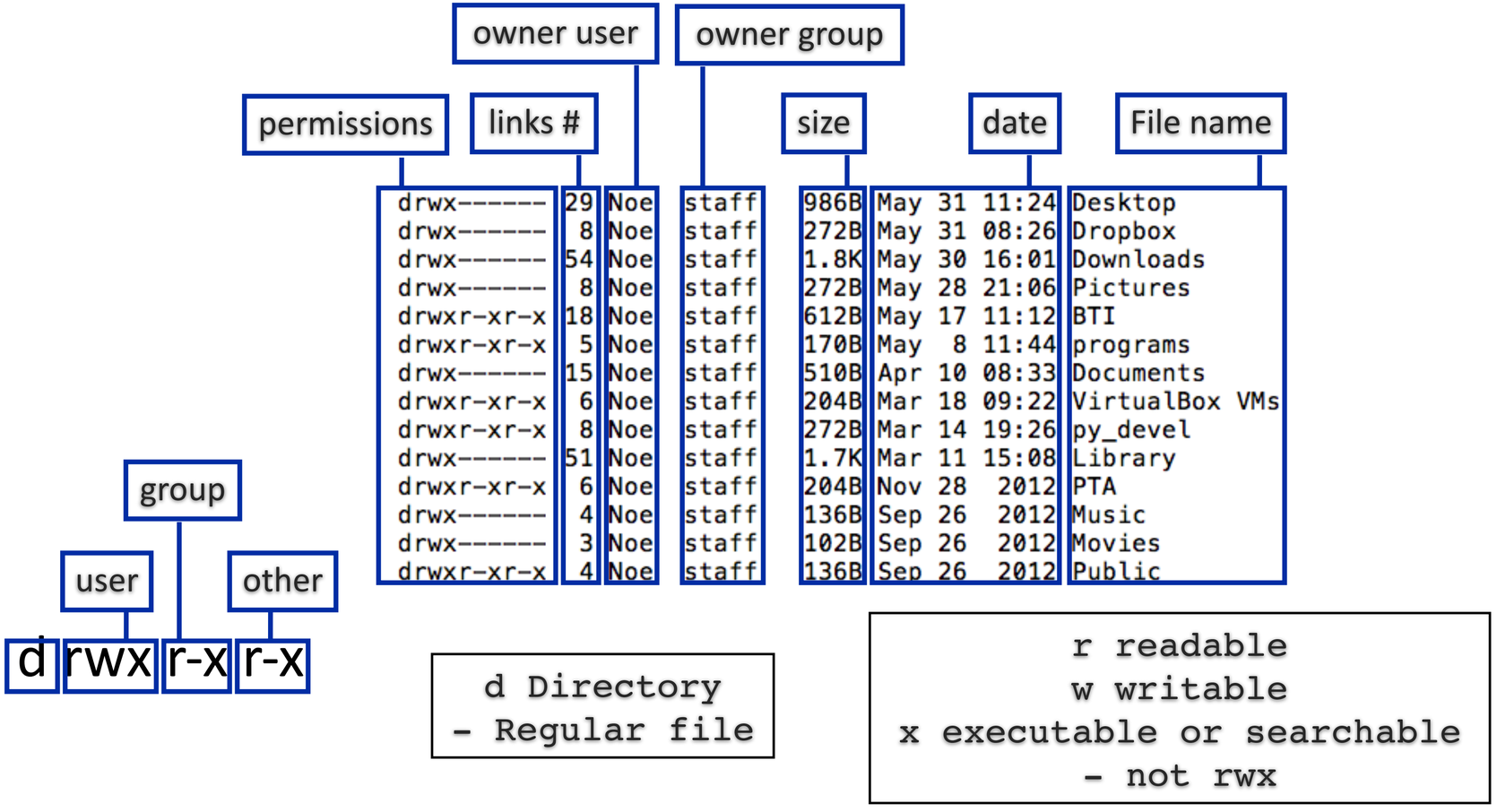
ls lists files and directories



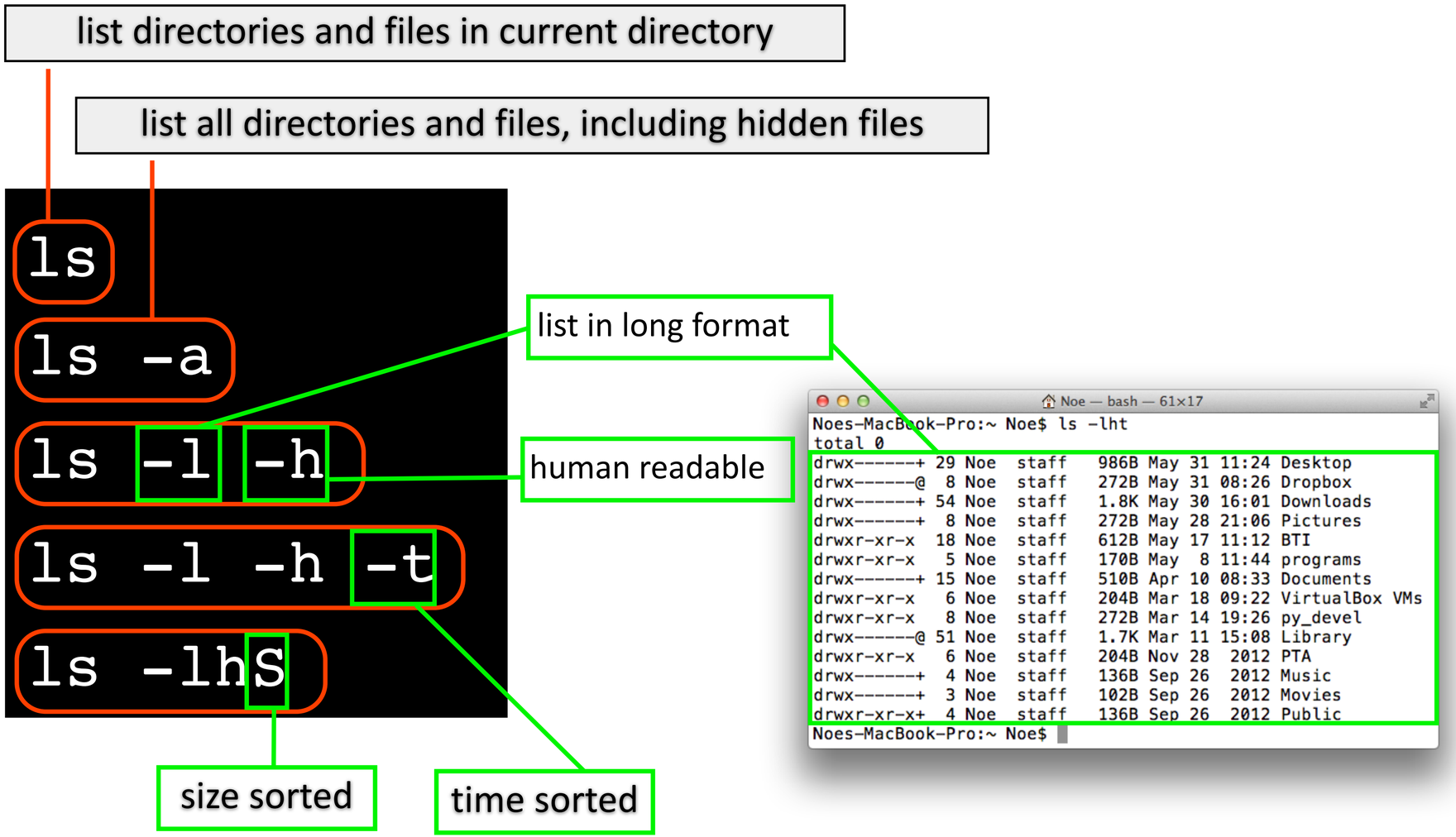
ls lists files and directories



The ls list output

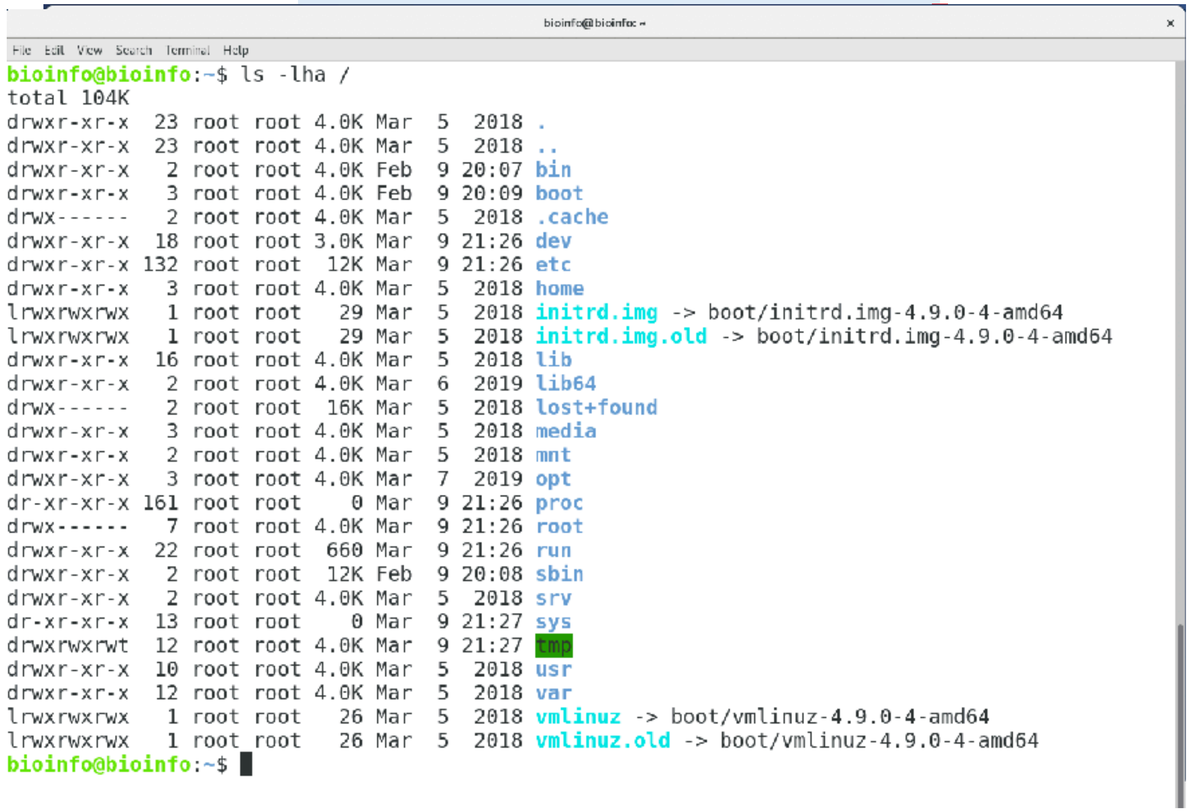


ls lists files and directories



Exercise:

List ALL files present in the root directory (including hidden files)



cd changes directory

changes directory to Desktop

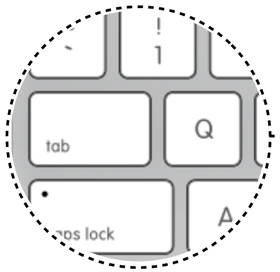
goes to parent directory

```
cd Desktop
cd ..
cd /
cd
cd -
```

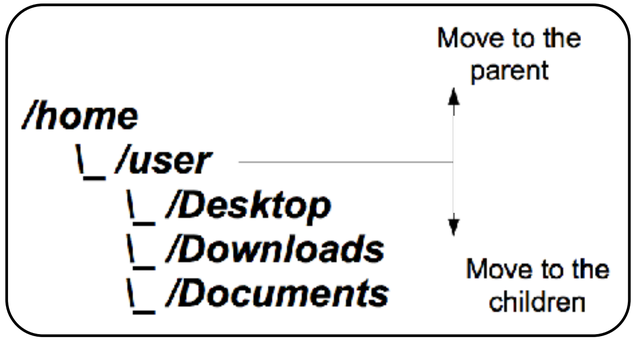
goes to root directory

goes to home directory

goes to previous directory



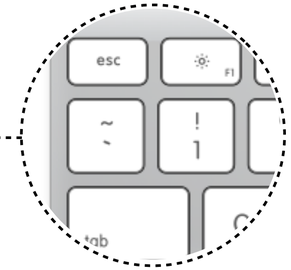
Use tab key to autocomplete names



Absolute and relative paths

list files in Desktop using an absolute path

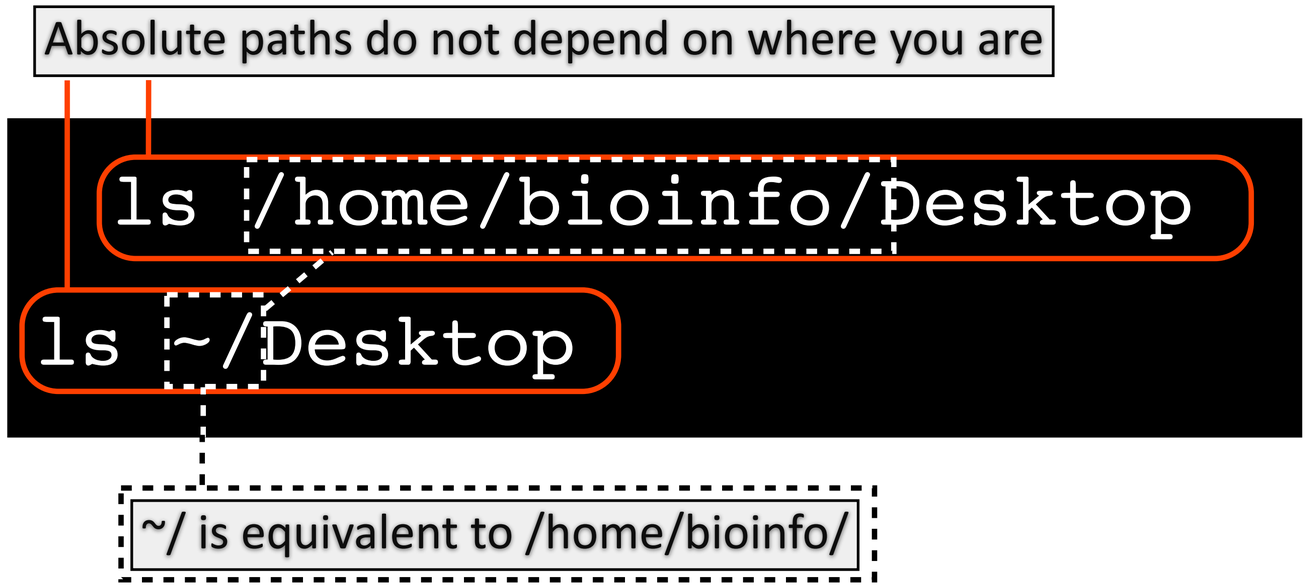
```
ls /home/user/Desktop
ls Desktop/
ls ~/Desktop
```



list files in Desktop using your home as a reference

list files in Desktop using a relative path (from your home: /home/bioinfo)

Absolute and relative paths

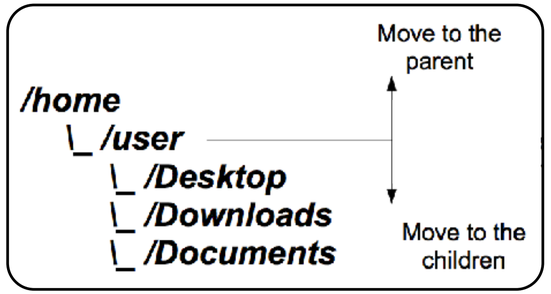


Absolute and relative paths

goes to *Desktop* from when you are in your home (/home/bioinfo)

```
cd Desktop/
ls ../Documents
```

list files from *Documents* when you are in *Desktop*



Wildcards, Shortcuts, and Command History

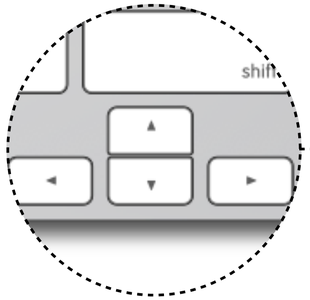
```
ls *txt
```

list all txt files in current directory

```
ls d*v
```

list files starting with d and ending with v

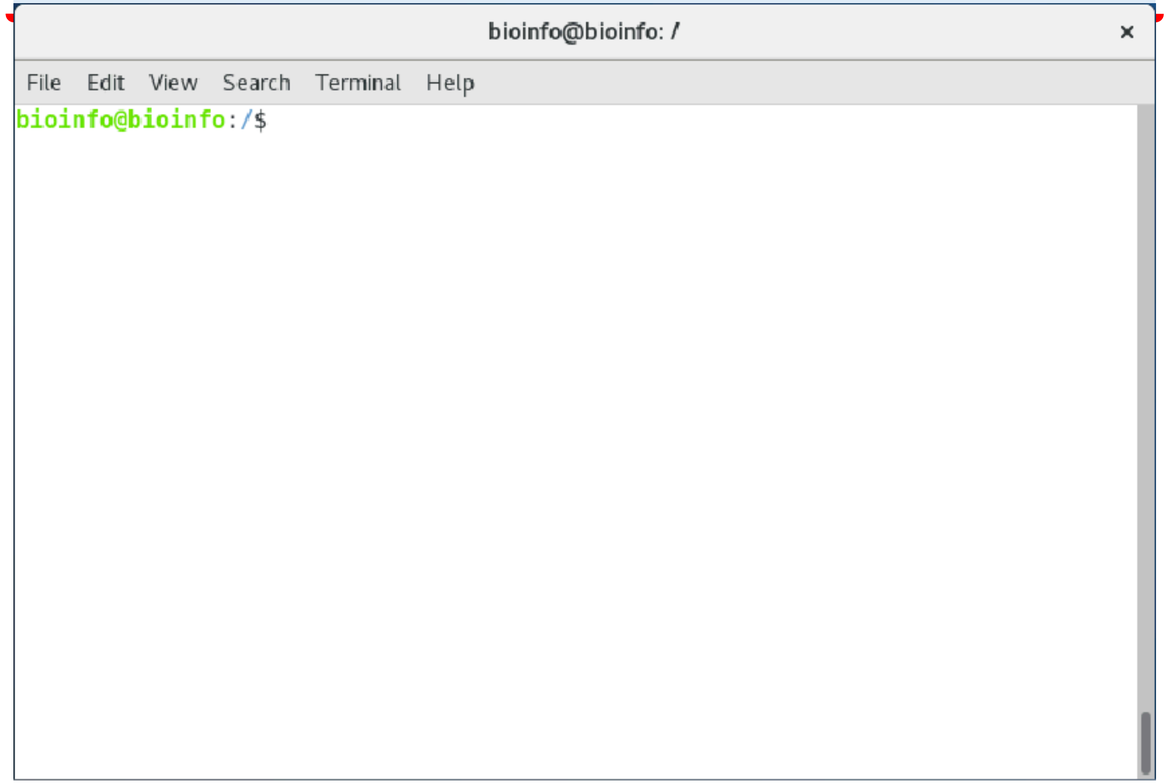
`ctrl-c`	stop process
`ctrl-a`	go to begin of line
`ctrl-e`	go to end of line
`ctrl-r`	search in command history



Use up and down arrows to navigate the command history

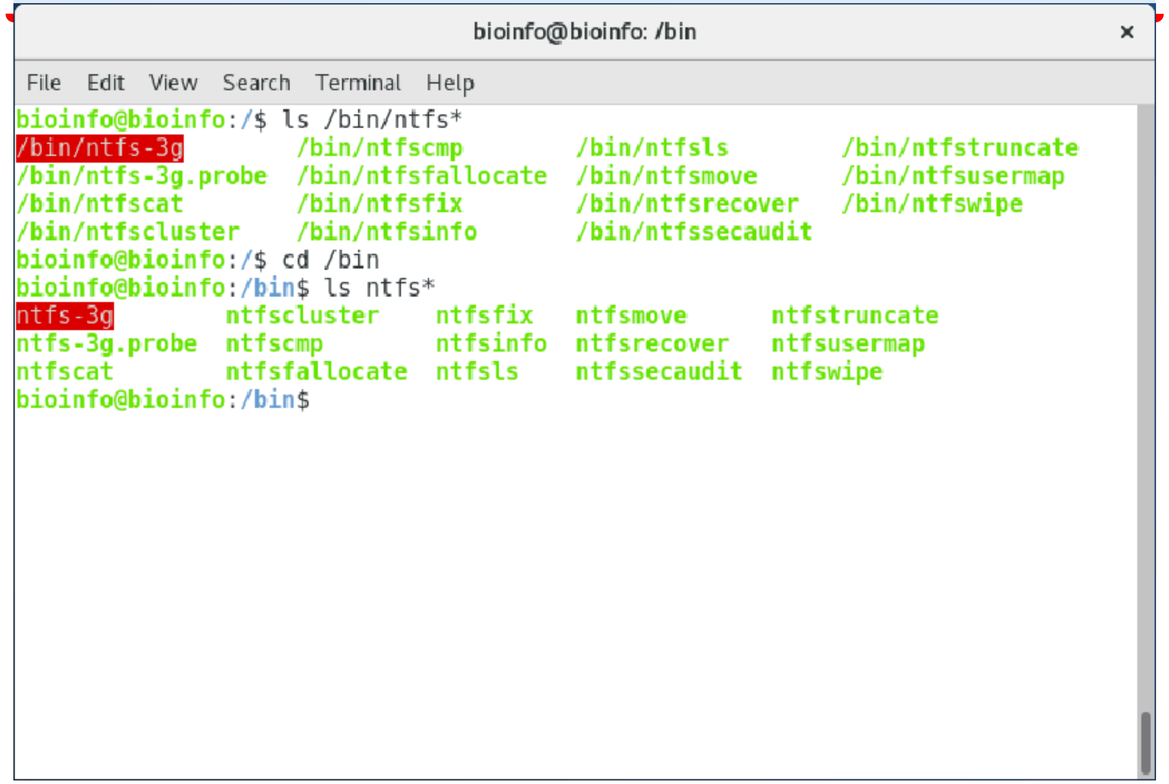
Exercise:

List files in the “/bin” directory that start with “ntfs” (do this once *without* changing your working directory, and once *with* changing your working directory).



Exercise 6:

List files in the “/bin” directory that start with “ntfs” (do this once *without* changing your working directory, and once *with* changing your working directory).



Escaping special characters

! @ $ ^ & * ~ ? . | / [ ] < > \ ` " ;# ( )

```
ls my_folder
```

list a folder

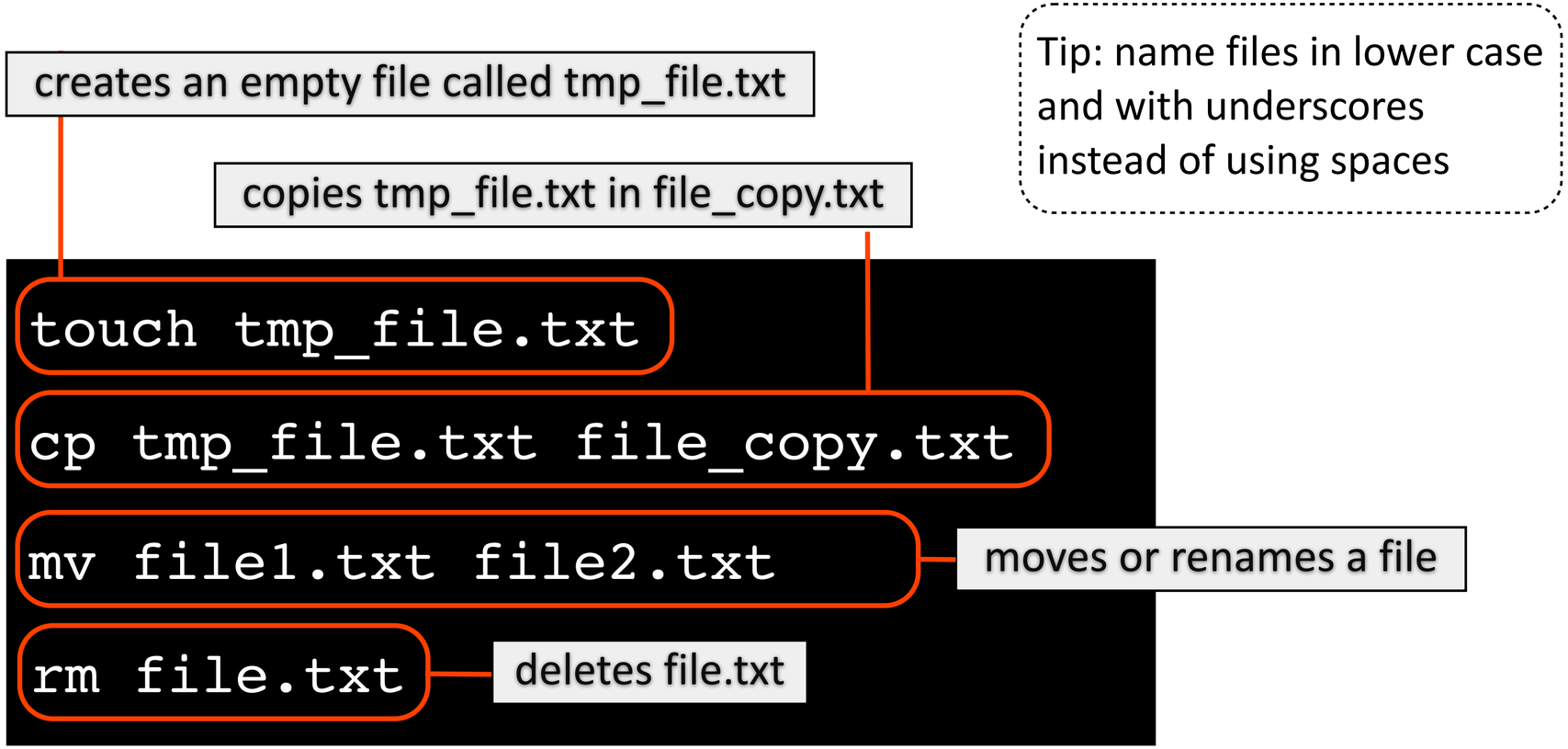
```
ls my\ folder
```

list a folder containing a space

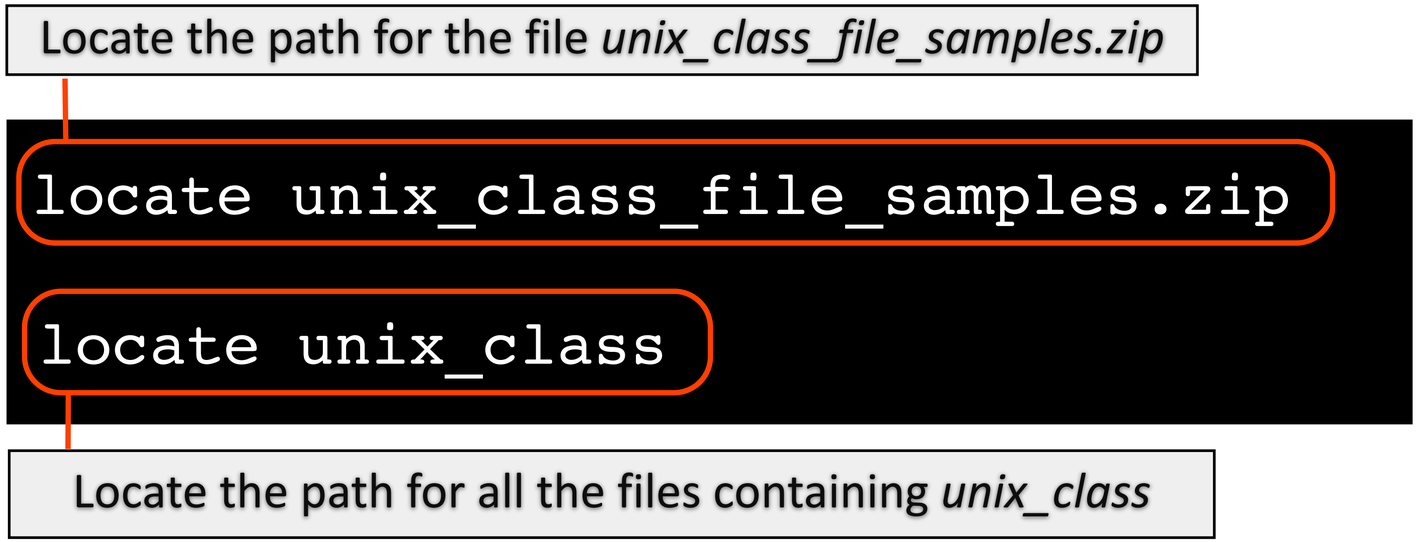
Tip: file names in lower case and with underscores instead of spaces

Use tab key to autocomplete names

Create, copy, move and delete files



Locate a file



Create, copy and delete directories

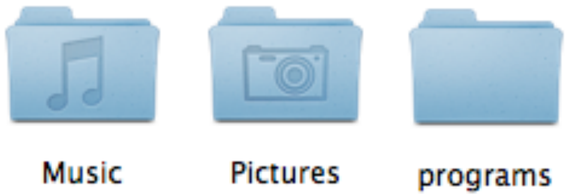
creates an empty directory called *dir_name*

deletes *dir_name* directory if it is empty

```
mkdir dir_name
rmdir dir_name
rm -r dir_name
cp -r dir_name dir_copy
```

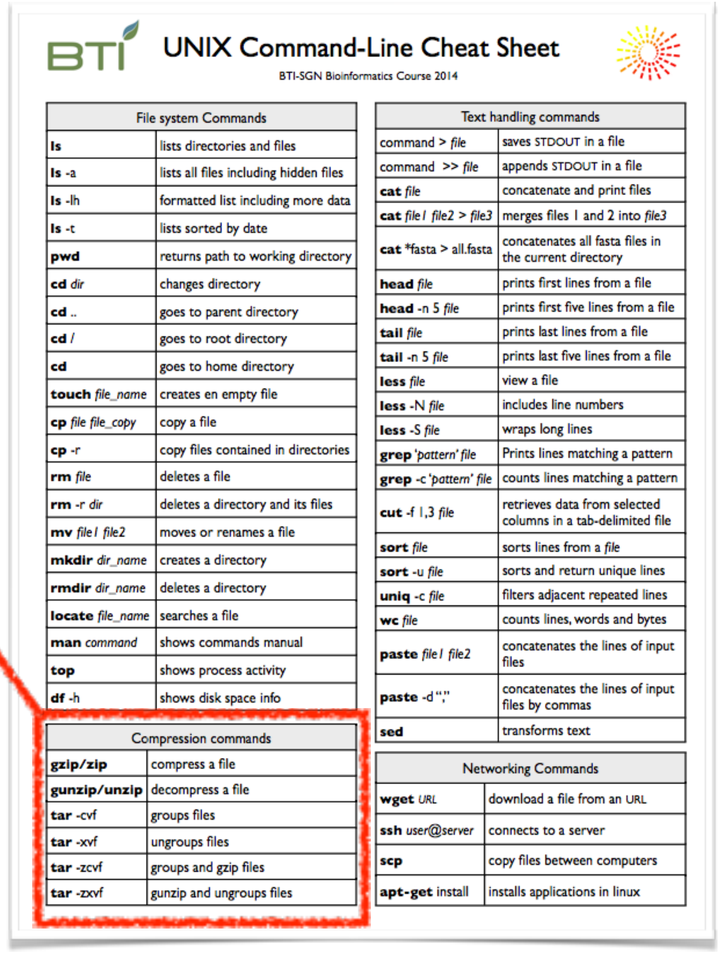
delete *dir_name* and its files

copy *dir_name* and its files in a new folder



Compression commands

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files



BTI

UNIX Command-Line Cheat Sheet

BTI-SGN Bioinformatics Course 2014

File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd *dir*	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch *file_name*	creates en empty file
cp *file file_copy*	copy a file
cp -r	copy files contained in directories
rm *file*	deletes a file
rm -r *dir*	deletes a directory and its files
mv *file1 file2*	moves or renames a file
mkdir *dir_name*	creates a directory
rmdir *dir_name*	deletes a directory
locate *file_name*	searches a file
man *command*	shows commands manual
top	shows process activity
df -h	shows disk space info

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

Text handling commands	
command > *file*	saves STDOUT in a file
command >> *file*	appends STDOUT in a file
cat *file*	concatenate and print files
cat *file1 file2 > file3*	merges files 1 and 2 into *file3*
cat *fasta > all.fasta	concatenates all fasta files in the current directory
head *file*	prints first lines from a file
head -n 5 *file*	prints first five lines from a file
tail *file*	prints last lines from a file
tail -n 5 *file*	prints last five lines from a file
less *file*	view a file
less -N *file*	includes line numbers
less -S *file*	wraps long lines
grep '*pattern*' *file*	Prints lines matching a pattern
grep -c '*pattern*' *file*	counts lines matching a pattern
cut -f 1,3 *file*	retrieves data from selected columns in a tab-delimited file
sort *file*	sorts lines from a *file*
sort -u *file*	sorts and return unique lines
uniq -c *file*	filters adjacent repeated lines
wc *file*	counts lines, words and bytes
paste *file1 file2*	concatenates the lines of input files
paste -d ","	concatenates the lines of input files by commas
sed	transforms text

Networking Commands	
wget *URL*	download a file from an URL
ssh *user@server*	connects to a server
scp	copy files between computers
apt-get install	installs applications in linux

Compression commands

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

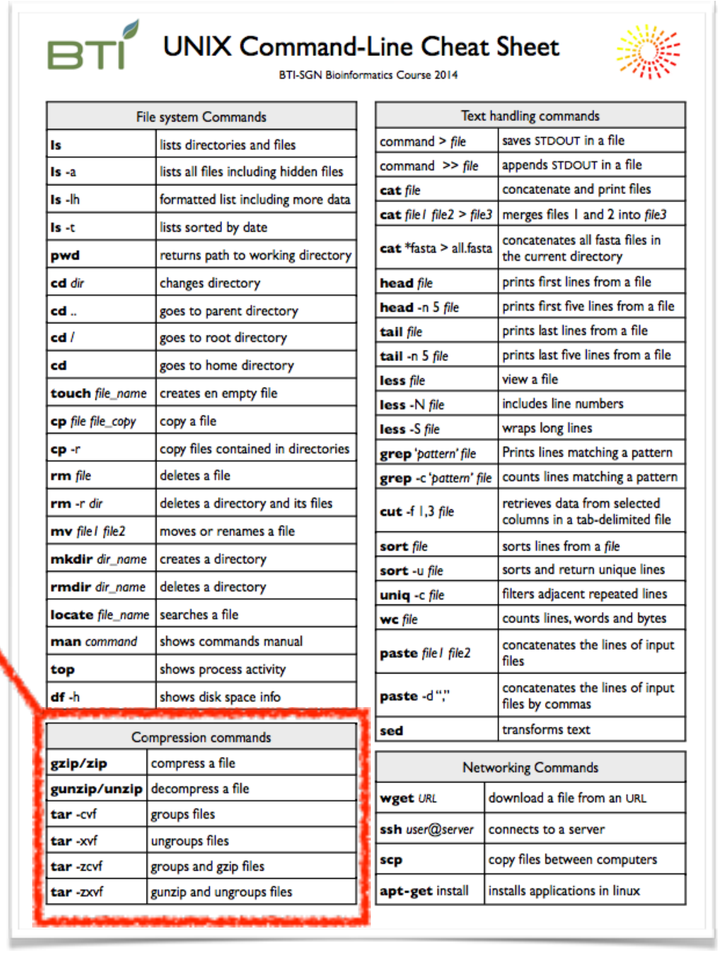
group and compress files

```
tar -zcvf file.tar.gz f1 f2
```

```
tar -zxvf file.tar.gz
```

decompress and ungroup a tar.gz file

files, directories or wildcards



BTI **UNIX Command-Line Cheat Sheet**

BTI-SGN Bioinformatics Course 2014

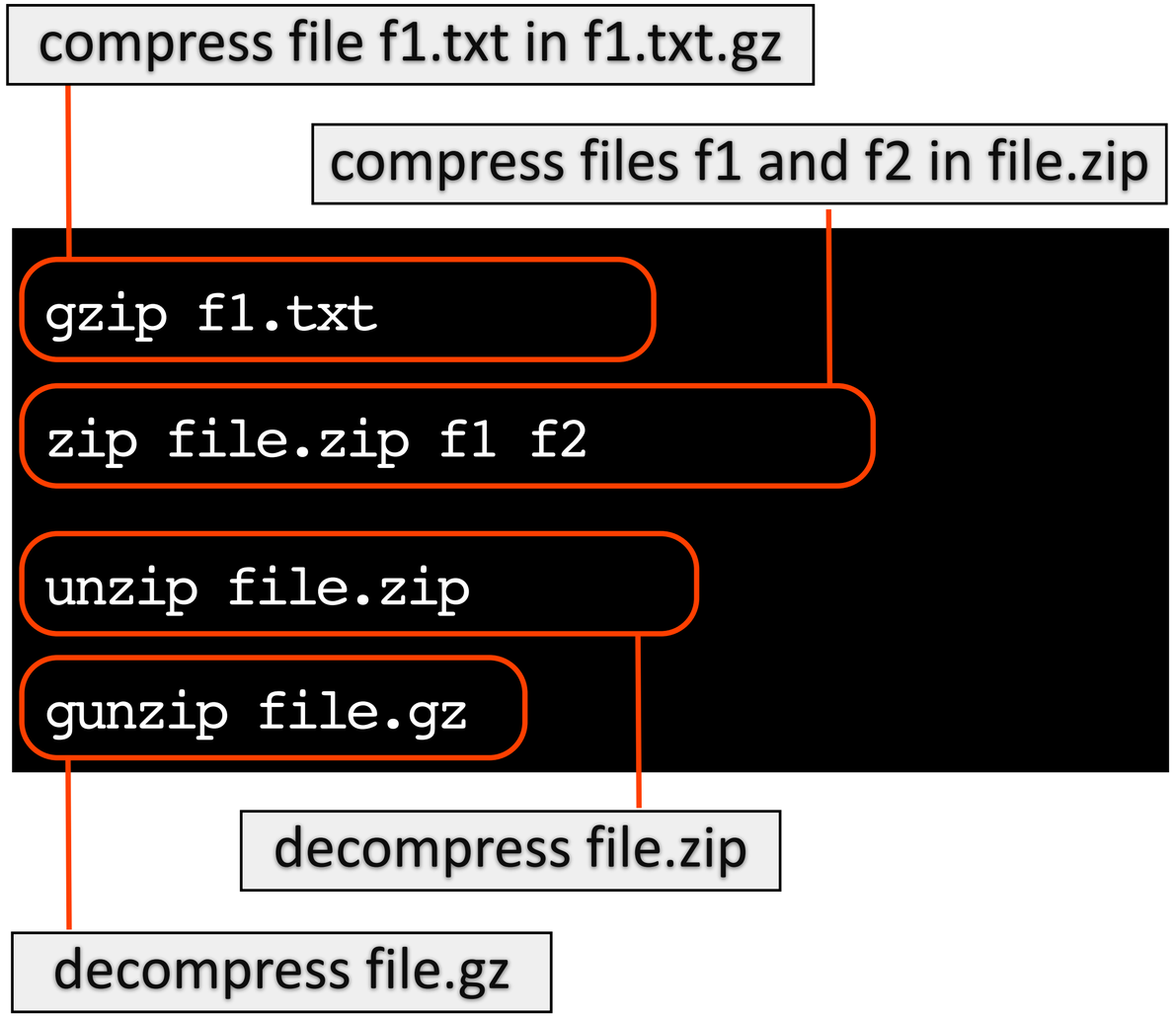
File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd *dir*	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch *file_name*	creates en empty file
cp *file file_copy*	copy a file
cp -r	copy files contained in directories
rm *file*	deletes a file
rm -r *dir*	deletes a directory and its files
mv *file1 file2*	moves or renames a file
mkdir *dir_name*	creates a directory
rmdir *dir_name*	deletes a directory
locate *file_name*	searches a file
man *command*	shows commands manual
top	shows process activity
df -h	shows disk space info

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

Text handling commands	
command > *file*	saves STDOUT in a file
command >> *file*	appends STDOUT in a file
cat *file*	concatenate and print files
cat *file1 file2* > *file3*	merges files 1 and 2 into *file3*
cat *fasta > all.fasta	concatenates all fasta files in the current directory
head *file*	prints first lines from a file
head -n 5 *file*	prints first five lines from a file
tail *file*	prints last lines from a file
tail -n 5 *file*	prints last five lines from a file
less *file*	view a file
less -N *file*	includes line numbers
less -S *file*	wraps long lines
grep *'pattern' file*	Prints lines matching a pattern
grep -c *'pattern' file*	counts lines matching a pattern
cut -f 1,3 *file*	retrieves data from selected columns in a tab-delimited file
sort *file*	sorts lines from a *file*
sort -u *file*	sorts and return unique lines
uniq -c *file*	filters adjacent repeated lines
wc *file*	counts lines, words and bytes
paste *file1 file2*	concatenates the lines of input files
paste -d ","	concatenates the lines of input files by commas
sed	transforms text

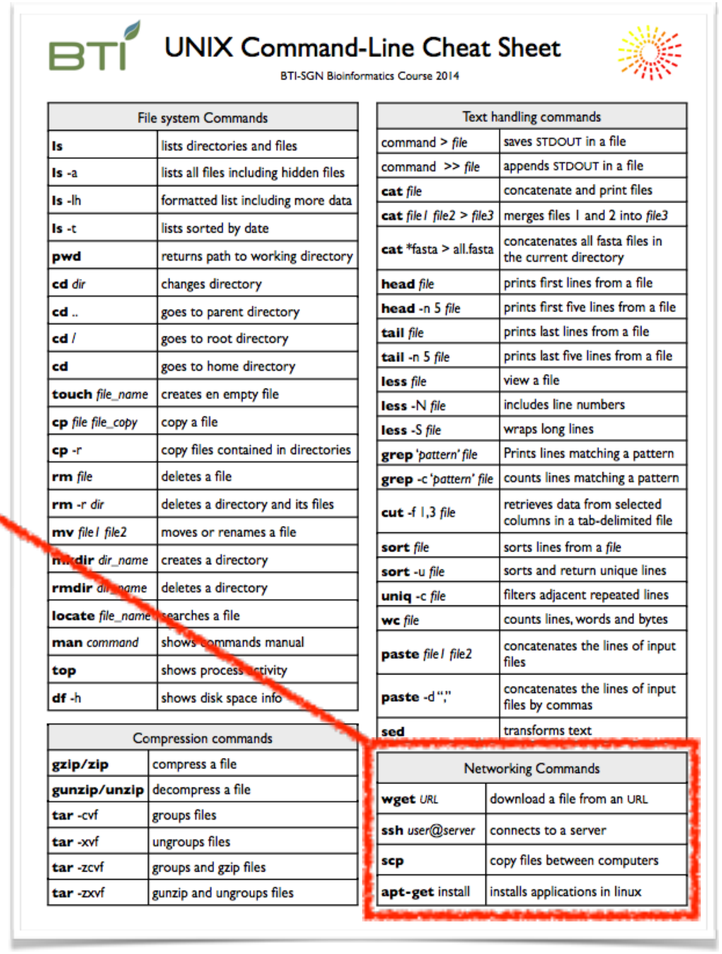
Networking Commands	
wget *URL*	download a file from an URL
ssh *user@server*	connects to a server
scp	copy files between computers
apt-get install	installs applications in linux

Compression commands



Networking Commands

- Networking commands



BTI **UNIX Command-Line Cheat Sheet**

BTI-SGN Bioinformatics Course 2014

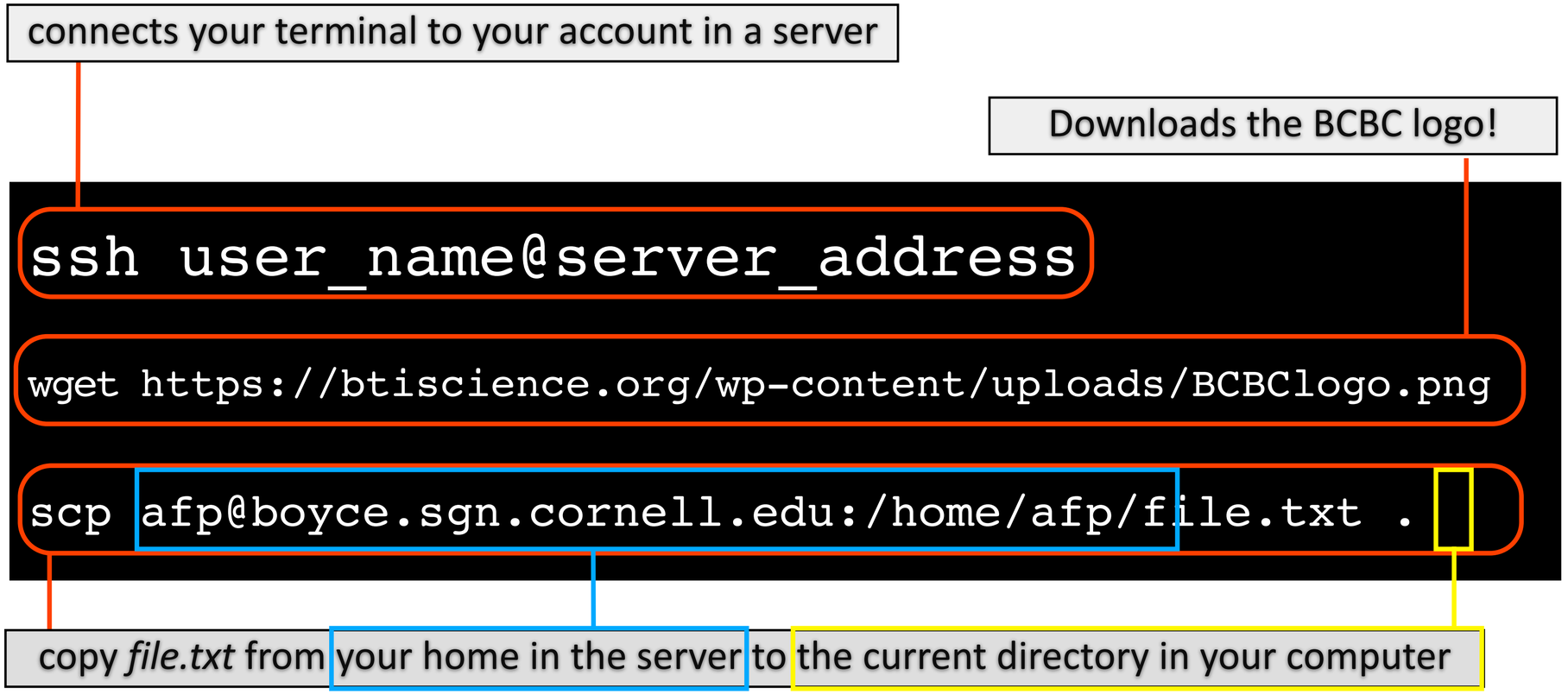
File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd *dir*	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch *file_name*	creates en empty file
cp *file file_copy*	copy a file
cp -r	copy files contained in directories
rm *file*	deletes a file
rm -r *dir*	deletes a directory and its files
mv *file1 file2*	moves or renames a file
mkdir *dir_name*	creates a directory
rmdir *dir_name*	deletes a directory
locate *file_name*	searches a file
man *command*	shows commands manual
top	shows process activity
df -h	shows disk space info

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

Text handling commands	
command > *file*	saves STDOUT in a file
command >> *file*	appends STDOUT in a file
cat *file*	concatenate and print files
cat *file1 file2* > *file3*	merges files 1 and 2 into *file3*
cat *fasta > all.fasta	concatenates all fasta files in the current directory
head *file*	prints first lines from a file
head -n 5 *file*	prints first five lines from a file
tail *file*	prints last lines from a file
tail -n 5 *file*	prints last five lines from a file
less *file*	view a file
less -N *file*	includes line numbers
less -S *file*	wraps long lines
grep '*pattern*' *file*	Prints lines matching a pattern
grep -c '*pattern*' *file*	counts lines matching a pattern
cut -f 1,3 *file*	retrieves data from selected columns in a tab-delimited file
sort *file*	sorts lines from a *file*
sort -u *file*	sorts and return unique lines
uniq -c *file*	filters adjacent repeated lines
wc *file*	counts lines, words and bytes
paste *file1 file2*	concatenates the lines of input files
paste -d ","	concatenates the lines of input files by commas
sed	transforms text

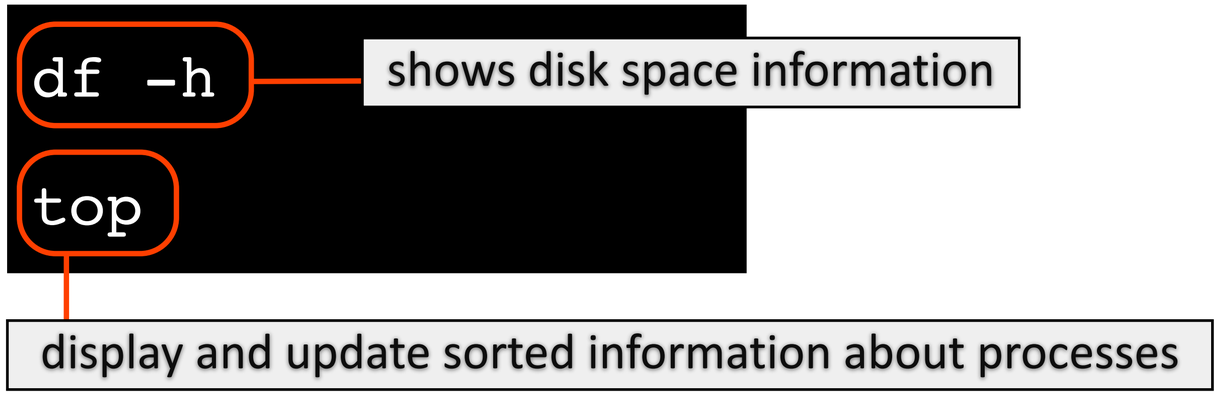
Networking Commands	
wget *URL*	download a file from an URL
ssh *user@server*	connects to a server
scp	copy files between computers
apt-get install	installs applications in linux

Networking Commands



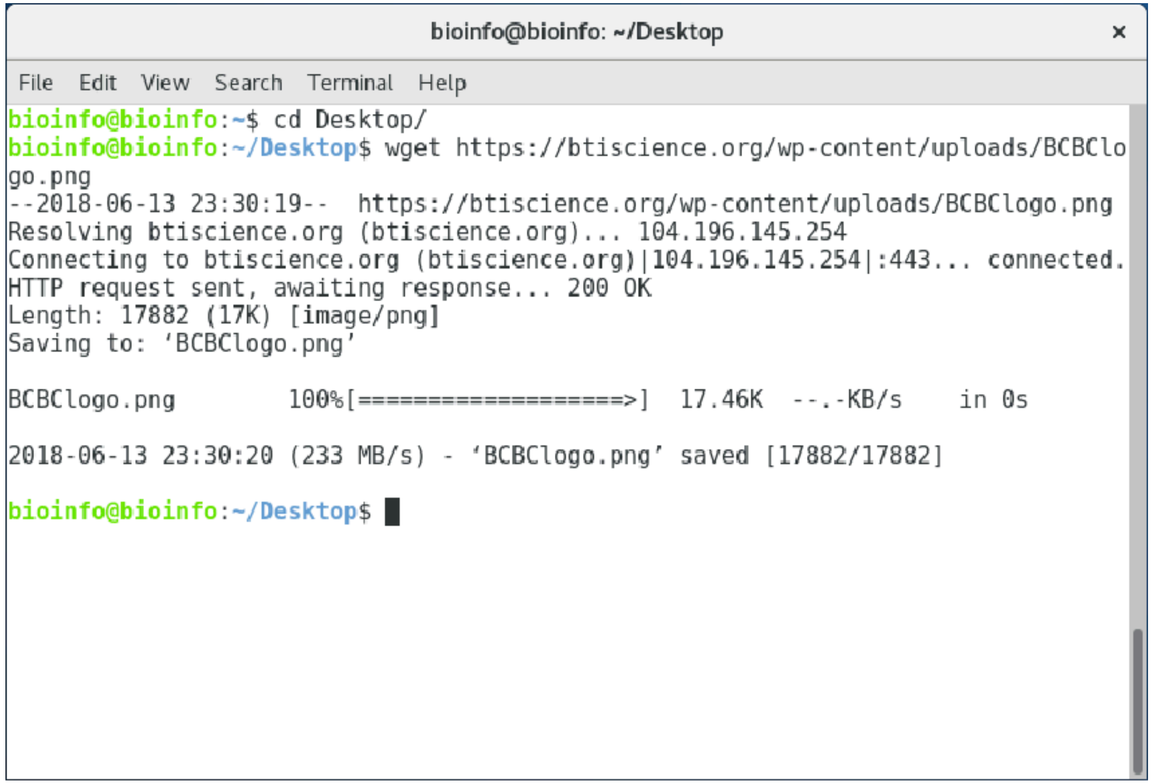
Tip: use the command `pwd` to get the path for `cp` and `scp`

Useful commands in the server



Exercise:

Download the BCBC logo (https://btiscience.org/wp-content/uploads/BCBClogo.png) to your Desktop.



“Man” pages

Man pages are the documentation for UNIX commands

- `$ man <command>`
- `$ man ls`

Searching man pages

- Use the `apropos` command
- `$ apropos “text editor”`

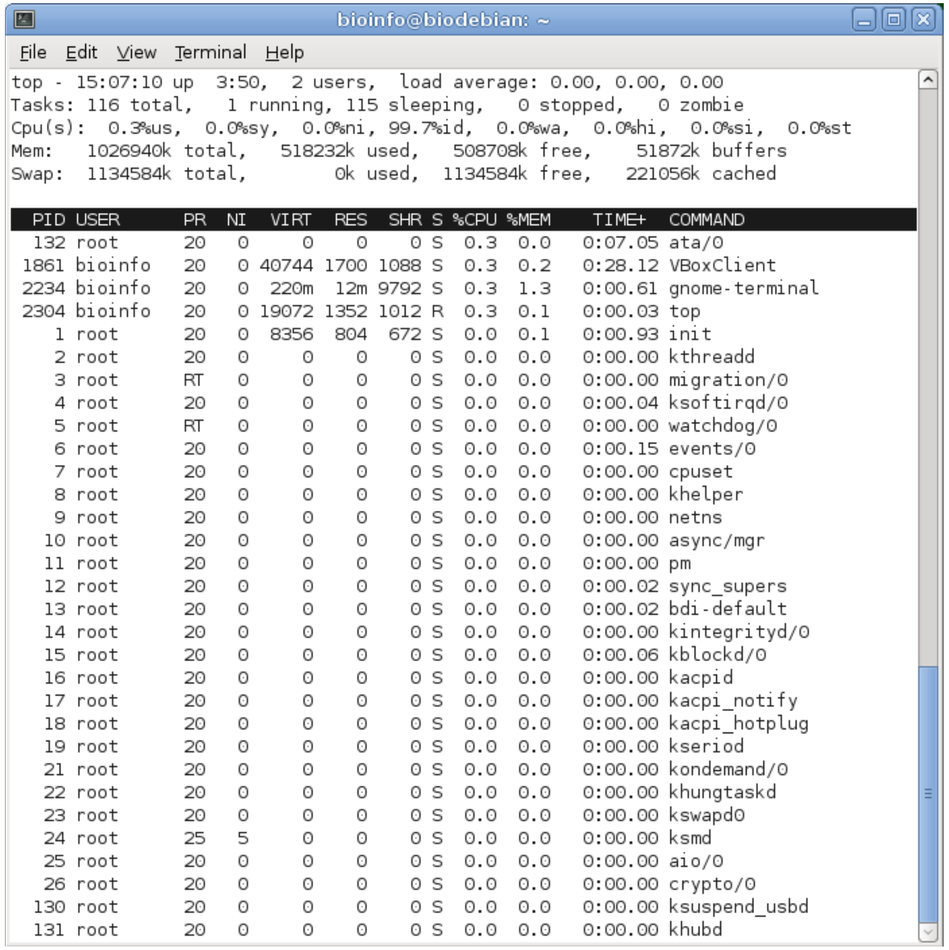
Processes

- Every running program is treated as a process
- Every process has a process ID and an “environment”
- Processes are created only from other processes through a *fork* (parent ID)
- First process is init, with process ID 1
- Viewing processes: ps , jobs , top, pstree
- Terminating processes: kill

Controlling processes

- Interrupting, terminating execution
- control-Z , control-C
- Viewing running jobs (jobs)
- Background/foreground jobs (bg , fg, &)
- Use sleep 100 to test

Top displays and updates sorted information about processes

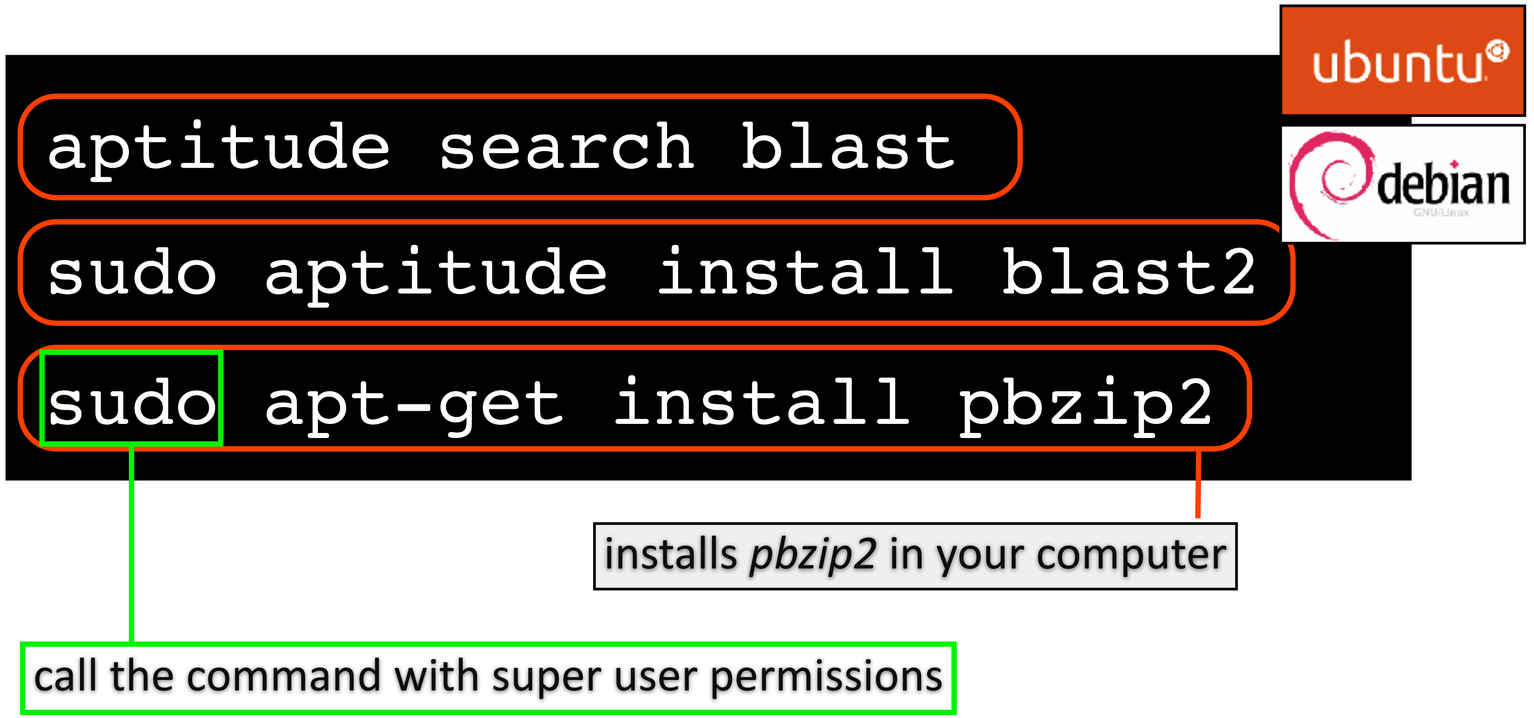


q quit

u user (top -u user)

M sort by memory usage

Commands to install software



Practice Exercises

a) Go to your Desktop directory using the command cd

b) Use the command touch to create a file called:

   Do not Use “special characters” in file names!.txt

c) Use the command rm to delete that file

d) Use the command mkdir to create a folder called unix_data in your desktop

e) Copy the file unix_class_file_samples.zip from your folder Data, in your home, to the folder unix_data, in your desktop

f) Uncompress the file unix_class_file_samples.zip in /home/bioinfo/Desktop/unix_data

g) Use the command rm with the option -r to remove the _MACOSX folder

Solutions

a) `cd Desktop` **(from your home: /home/bioinfo/) or** `cd /home/bioinfo/Desktop` **or** `cd ~/Desktop`

b) `touch Do\ not\ Use\ \“special\ characters\”\ in\ file\ names\!.txt`

c) `rm Do\ not\ Use\ \“special\ characters\”\ in\ file\ names\!.txt` **(use the tab key)**

d) `mkdir unix_data` **(from Desktop: /home/bioinfo/Desktop) or** `mkdir /home/bioinfo/Desktop/unix_data` **or** `mkdir ~/Desktop/unix_data`

e) `cp /home/bioinfo/Data/unix_class_file_samples.zip /home/bioinfo/Desktop/unix_data`

f) `cd /home/bioinfo/Desktop/unix_data` **and then** `unzip unix_class_file_samples.zip`

g) `rm -r _MACOSX`