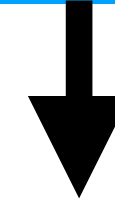


# Class 2a: Linux Command Line

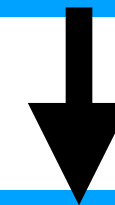
June 21, 2023

# Course Overview

Class 1: Bioinformatics Overview



Class 2: Linux Command Line  
NGS Read Mapping



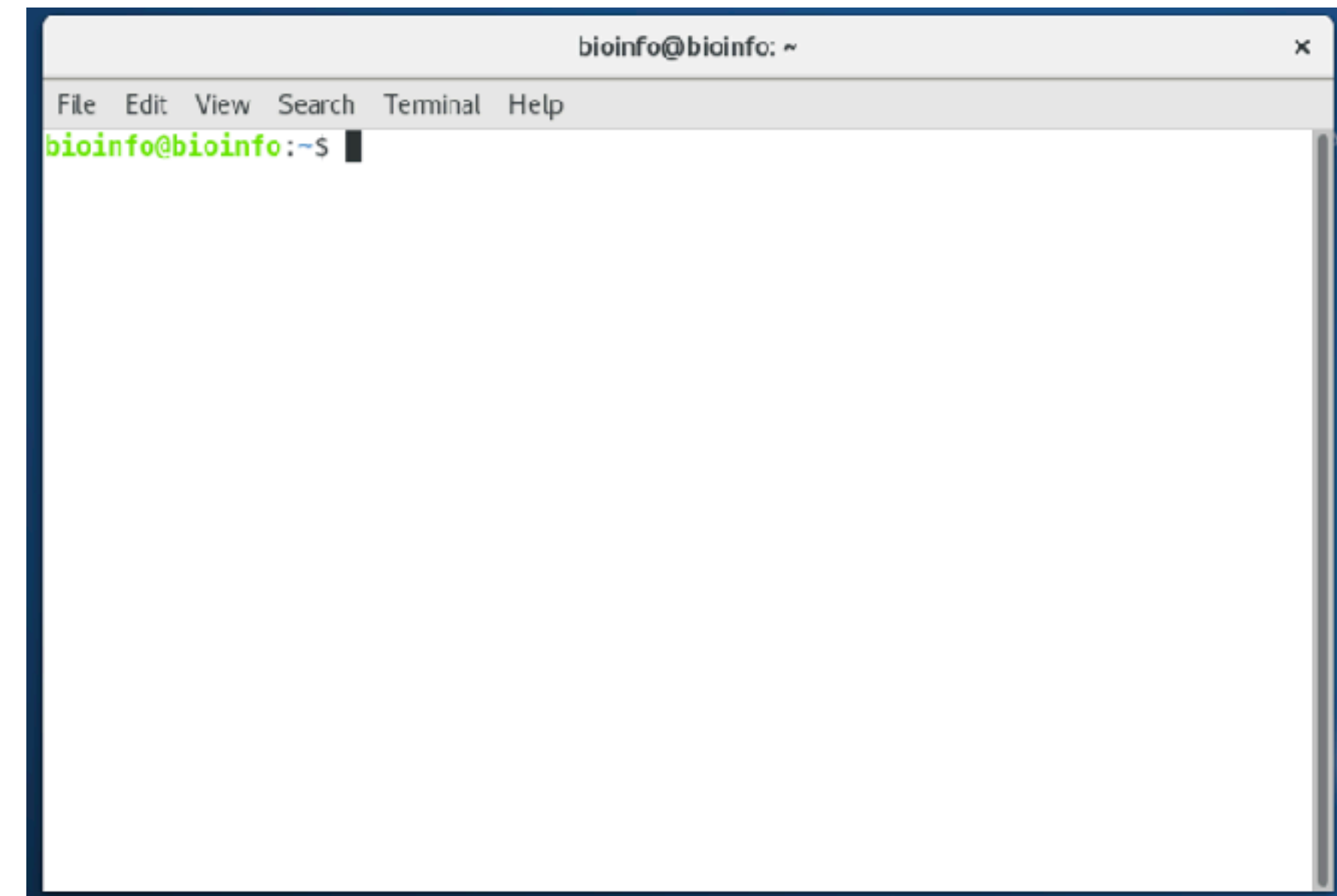
Class 3: Gene expression  
analysis in R

# 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.

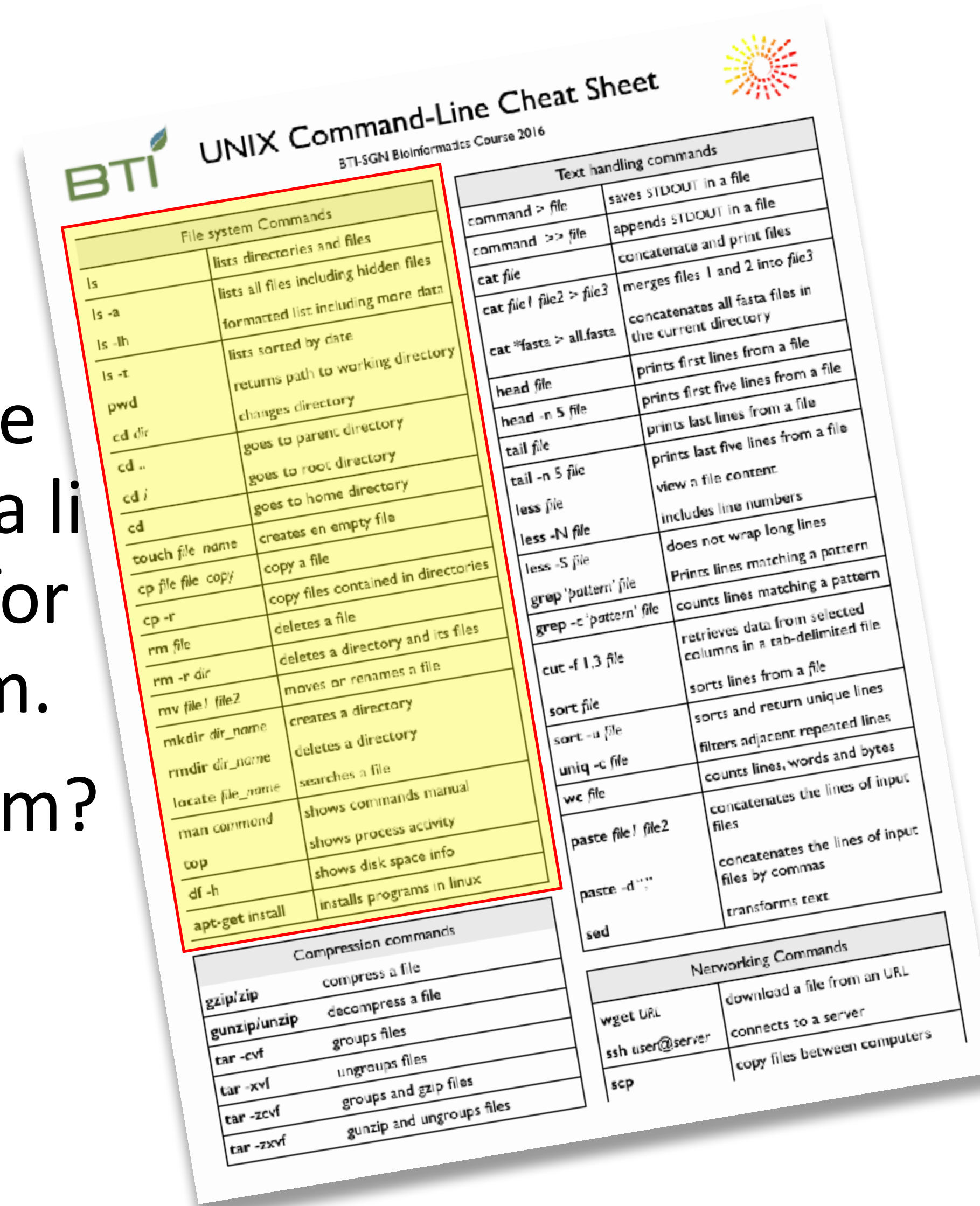


# Why use the command-line (terminal)?

- 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 list of common commands for navigating the file system.
- But what *is* the file system?



**BTI** UNIX Command-Line Cheat Sheet  
BTU-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 -rt	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
cd	creates an empty file
touch file_name	copy a file
cp file file copy	copy files contained in directories
cp -r	deletes a file
rm file	deletes a directory and its files
rm -r dir	moves or renames a file
mv file1 file2	creates a directory
mkdir dir_name	deletes a directory
rmdir dir_name	searches a file
locate file_name	shows commands manual
man command	shows process activity
top	shows disk space info
df -h	installs programs in linux
apt-get install	

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 n 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

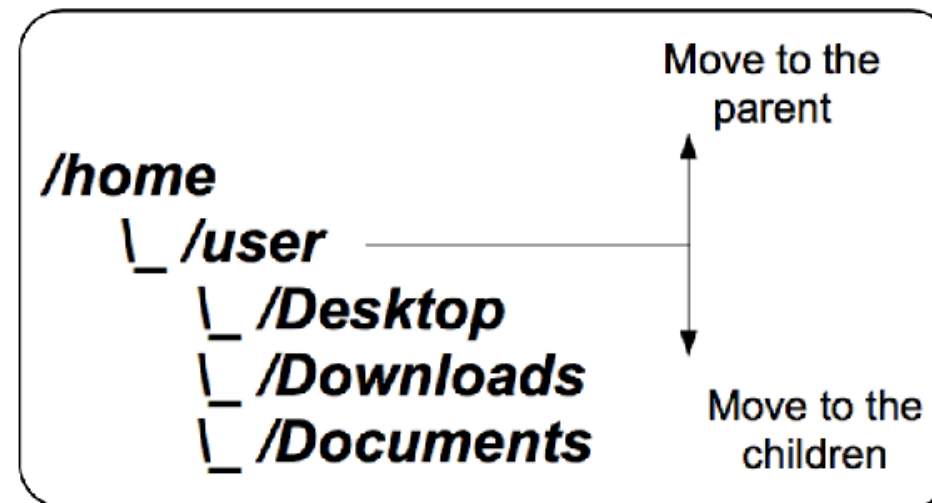
# Connect to remote server 'thompson'

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

ssh bioinfo0@thompson.sgn.cornell.edu

- #The password is bioinfo00

# The File System



The screenshot shows a terminal window on the left and a file manager window on the right. The terminal window displays the output of the command `ls -lh`, listing the contents of the home directory. The file manager window shows the same directories as icons in the Home view.

**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
```

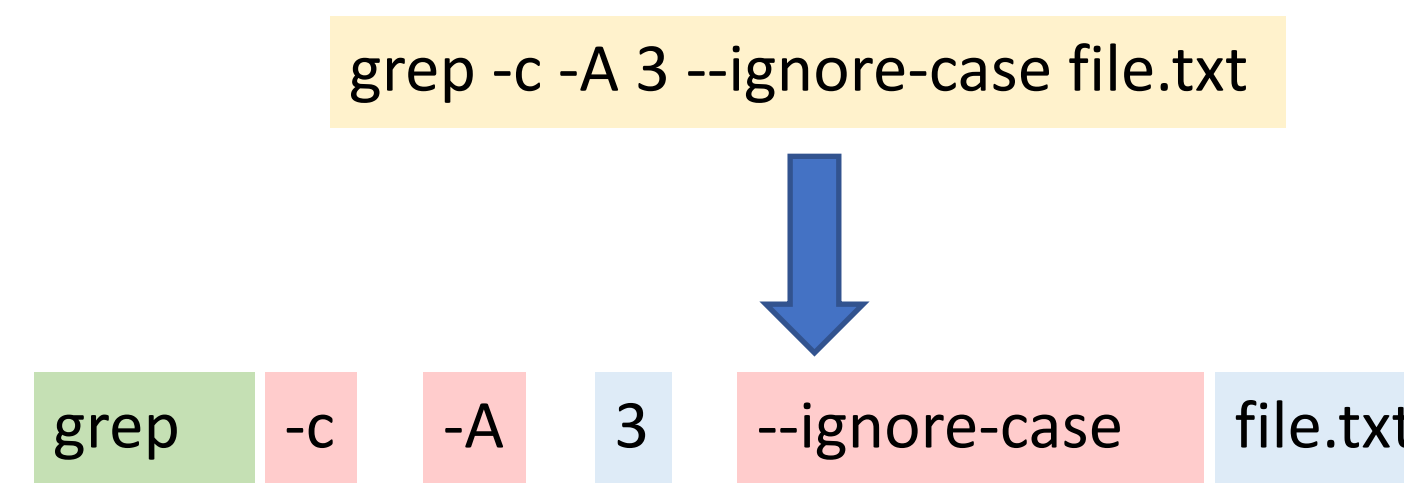
```
noe@debian-virtualbox:~$ ls -l /
total 108
drwxr-xr-x  2 root root 4096 Sep 26 2012 bin
drwxr-xr-x  3 root root 4096 Nov  9 2012 boot
drwxr-xr-x 15 root root 3140 May 31 12:46 dev
drwxr-xr-x 130 root root 12288 May 31 12:45 etc
drwxr-xr-x  5 root root 4096 Feb 28 13:54 export
drwxr-xr-x  4 root root 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 -> /lib
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

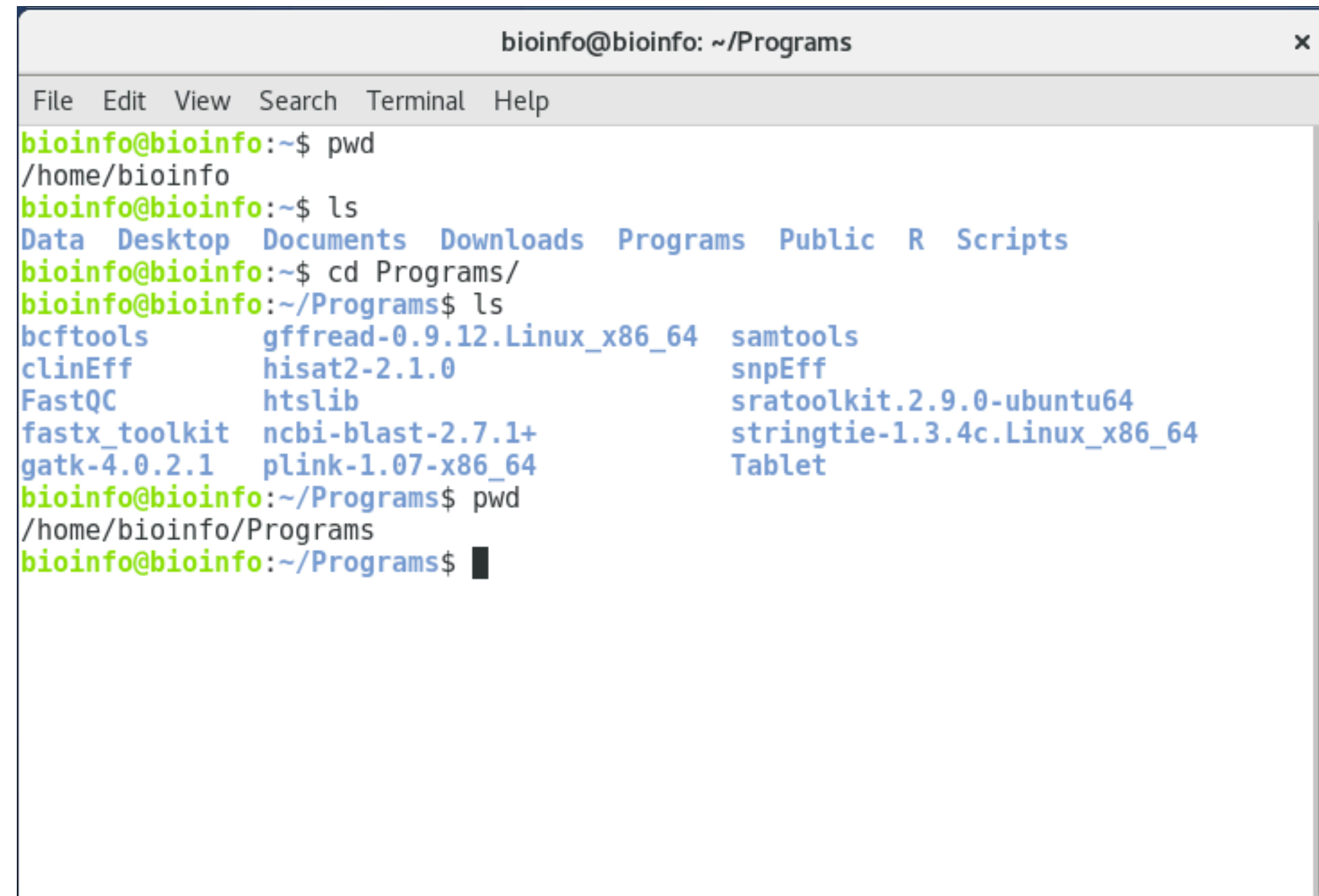
- Every UNIX command is made up of a series of space-separated strings.
- The first of these strings is always the command you will run.
- The user can also provide options (shown in red) and arguments (shown in blue).
- A command can require arguments, as can an option.
- In this example, the “-A” option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



```
bioinfo@bioinfo: ~/Programs
File Edit View Search Terminal Help
bioinfo@bioinfo:~$ pwd
/home/bioinfo
bioinfo@bioinfo:~$ ls
Data Desktop Documents Downloads Programs Public R Scripts
bioinfo@bioinfo:~$ cd Programs/
bioinfo@bioinfo:~/Programs$ ls
bcftools          gffread-0.9.12.Linux_x86_64  samtools
clinEff           hisat2-2.1.0                 snpEff
FastQC            htlib                        sratoolkit.2.9.0-ubuntu64
fastx_toolkit     ncbi-blast-2.7.1+            stringtie-1.3.4c.Linux_x86_64
gatk-4.0.2.1      plink-1.07-x86_64            Tablet
bioinfo@bioinfo:~/Programs$ pwd
/home/bioinfo/Programs
bioinfo@bioinfo:~/Programs$
```

## 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

A diagram illustrating the 'ls' command. A light gray box at the top contains the text 'list directories and files in current directory'. A vertical orange line descends from this box to a small orange circle. Inside this circle, the text 'ls' is written in white. This circle is positioned at the top-left corner of a large black rectangle, which represents a terminal window.

# ls lists files and directories

list directories and files in current directory

list all directories and files, including hidden files



```
ls
```

A terminal window with a black background. The command 'ls' is entered and highlighted with an orange rounded rectangle. A red line connects this box to the text 'list directories and files in current directory' above it.

```
ls -a
```

The command 'ls -a' is entered and highlighted with an orange rounded rectangle. A red line connects this box to the text 'list all directories and files, including hidden files' above it.



# ls lists files and directories

list directories and files in current directory

list all directories and files, including hidden files

`ls`

`ls -a`

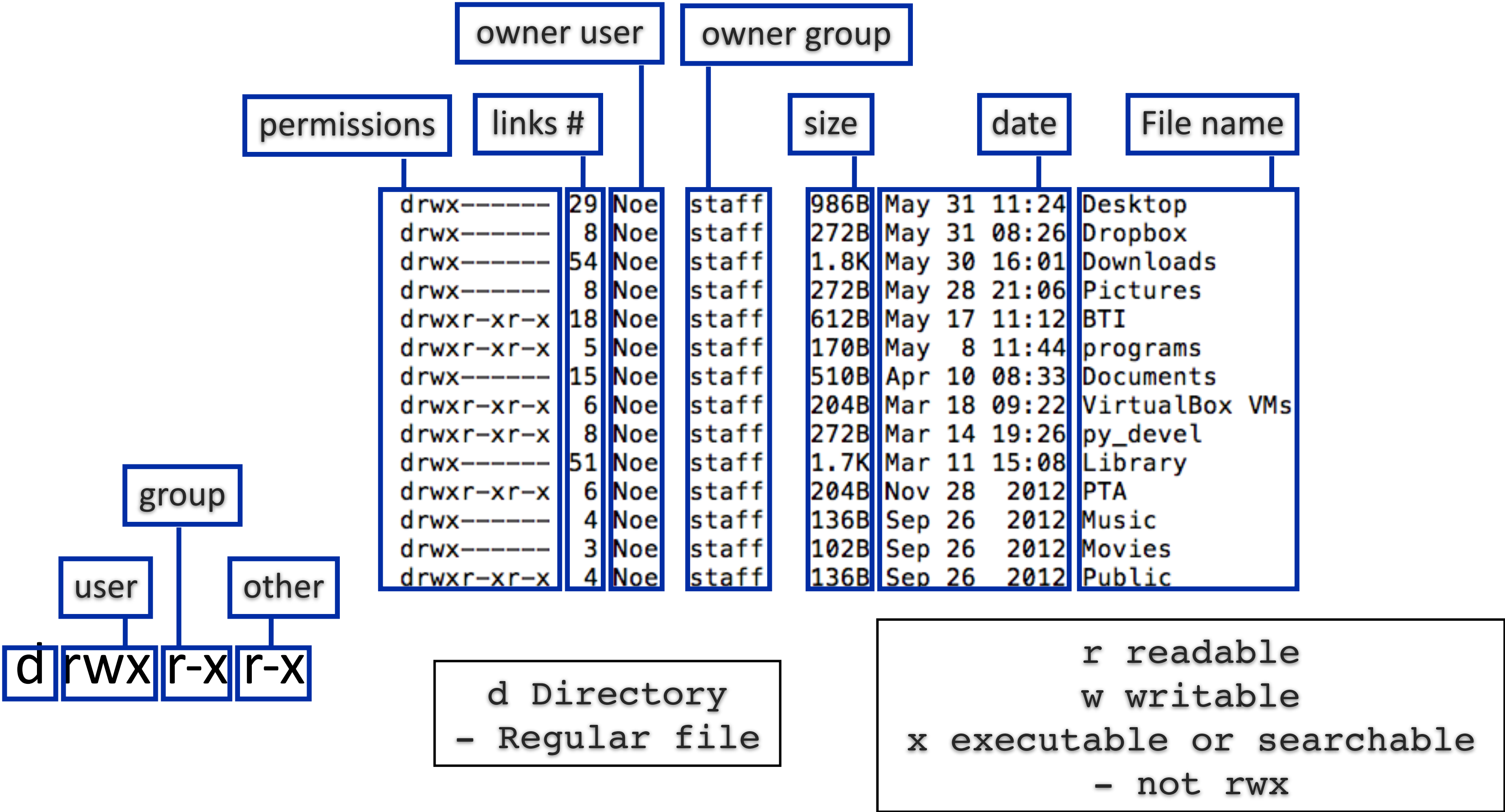
`ls -l -h`

list in long format

human readable

```
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
drwxr-xr-x   6 Noe  staff   204B Nov 28  2012 PTA
drwx-----+  4 Noe  staff   136B Sep 26  2012 Music
drwx-----+  3 Noe  staff   102B Sep 26  2012 Movies
drwxr-xr-x+   4 Noe  staff   136B Sep 26  2012 Public
Noes-MacBook-Pro:~ Noe$
```

# The ls list output



# ls lists files and directories

list directories and files in current directory

list all directories and files, including hidden files

`ls`

`ls -a`

`ls -l -h`

`ls -l -h -t`

`ls -lhS`

list in long format

human readable

size sorted

time sorted

```
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
drwxr-xr-x    6 Noe  staff   204B Nov 28 2012 PTA
drwx-----+  4 Noe  staff   136B Sep 26 2012 Music
drwx-----+  3 Noe  staff   102B Sep 26 2012 Movies
drwxr-xr-x+   4 Noe  staff   136B Sep 26 2012 Public
Noes-MacBook-Pro:~ Noe$
```



## Exercise:

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

```
bioinfo@bioinfo:~$ ls -lha /
total 104K
drwxr-xr-x 23 root root 4.0K Mar  5 2018 .
drwxr-xr-x 23 root root 4.0K Mar  5 2018 ..
drwxr-xr-x  2 root root 4.0K Feb  9 20:07 bin
drwxr-xr-x  3 root root 4.0K Feb  9 20:09 boot
drwx----- 2 root root 4.0K Mar  5 2018 .cache
drwxr-xr-x 18 root root 3.0K Mar  9 21:26 dev
drwxr-xr-x 132 root root 12K Mar  9 21:26 etc
drwxr-xr-x  3 root root 4.0K Mar  5 2018 home
lrwxrwxrwx  1 root root   29 Mar  5 2018 initrd.img -> boot/initrd.img-4.9.0-4-amd64
lrwxrwxrwx  1 root root   29 Mar  5 2018 initrd.img.old -> boot/initrd.img-4.9.0-4-amd64
drwxr-xr-x 16 root root 4.0K Mar  5 2018 lib
drwxr-xr-x  2 root root 4.0K Mar  6 2019 lib64
drwx----- 2 root root 16K Mar  5 2018 lost+found
drwxr-xr-x  3 root root 4.0K Mar  5 2018 media
drwxr-xr-x  2 root root 4.0K Mar  5 2018 mnt
drwxr-xr-x  3 root root 4.0K Mar  7 2019 opt
dr-xr-xr-x 161 root root   0 Mar  9 21:26 proc
drwx----- 7 root root 4.0K Mar  9 21:26 root
drwxr-xr-x 22 root root 660 Mar  9 21:26 run
drwxr-xr-x  2 root root 12K Feb  9 20:08 sbin
drwxr-xr-x  2 root root 4.0K Mar  5 2018 srv
dr-xr-xr-x 13 root root   0 Mar  9 21:27 sys
drwxrwxrwt 12 root root 4.0K Mar  9 21:27 tmp
drwxr-xr-x 10 root root 4.0K Mar  5 2018 usr
drwxr-xr-x 12 root root 4.0K Mar  5 2018 var
lrwxrwxrwx  1 root root   26 Mar  5 2018 vmlinuz -> boot/vmlinuz-4.9.0-4-amd64
lrwxrwxrwx  1 root root   26 Mar  5 2018 vmlinuz.old -> boot/vmlinuz-4.9.0-4-amd64
bioinfo@bioinfo:~$
```

# cd changes directory

changes directory to Desktop

goes to parent directory

```
cd Desktop
```

```
cd ..
```

```
cd /
```

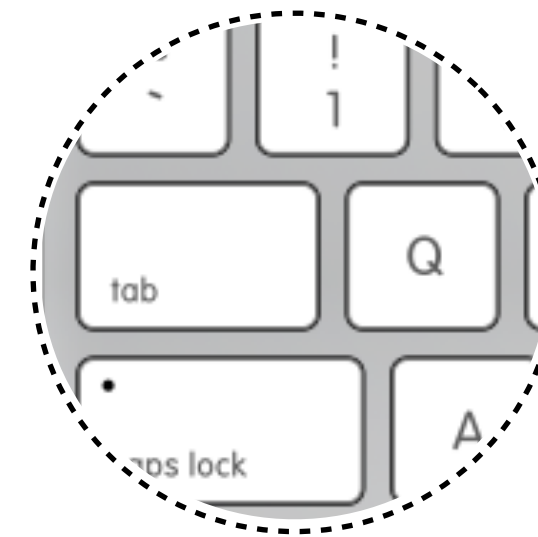
goes to root directory

```
cd
```

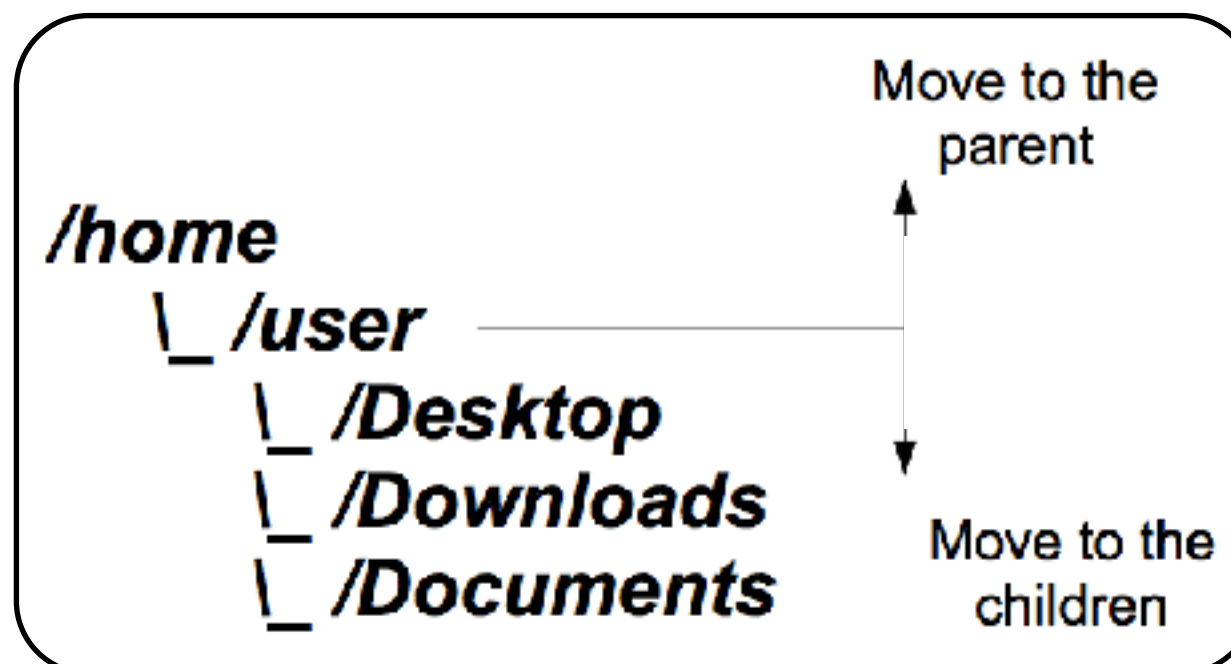
goes to home directory

```
cd -
```

goes to previous directory



Use tab key to  
autocomplete names



# Absolute and relative paths

list files in Desktop using an absolute path

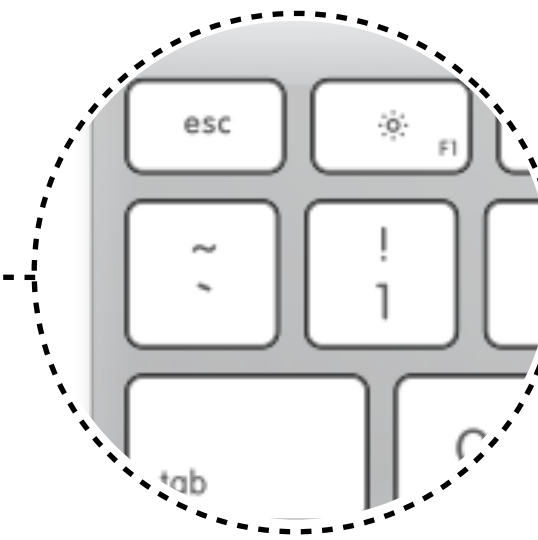
```
ls /home/user/Desktop
```

```
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 paths do not depend on where you are

```
ls /home/bioinfo/Desktop
```

```
ls ~/Desktop
```

~/ is equivalent to /home/bioinfo/

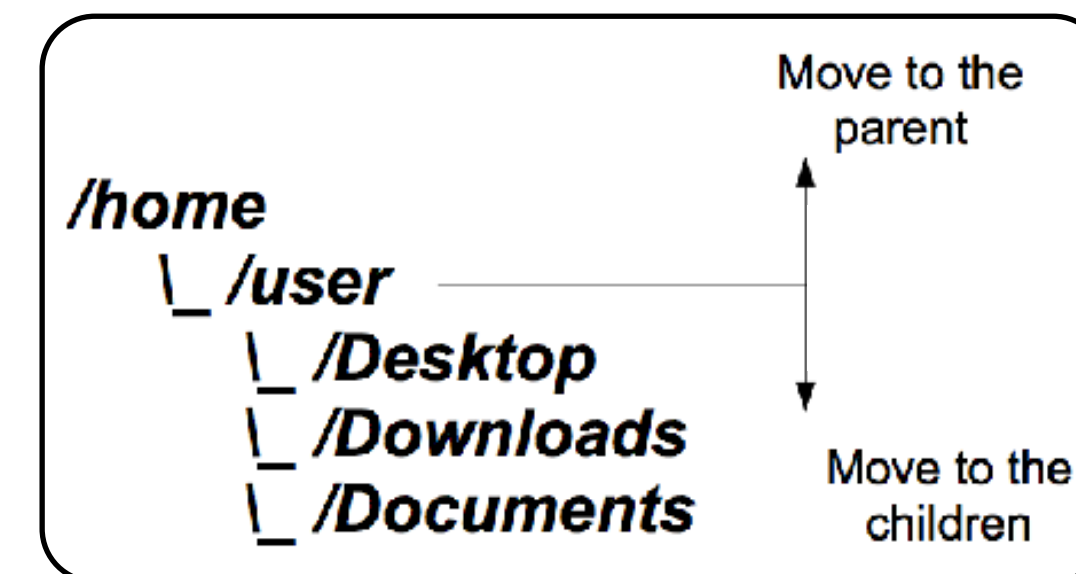
# Absolute and relative paths

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

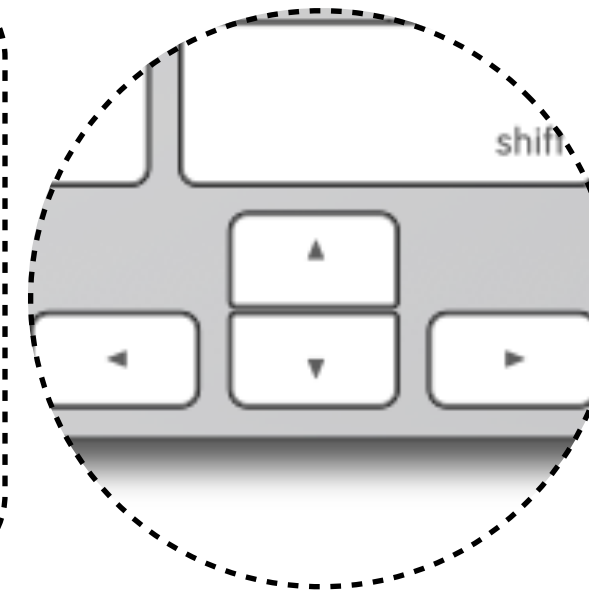
```
ls *.txt
```

list all txt files in current directory

```
ls d*v
```

list files starting with d and ending with v

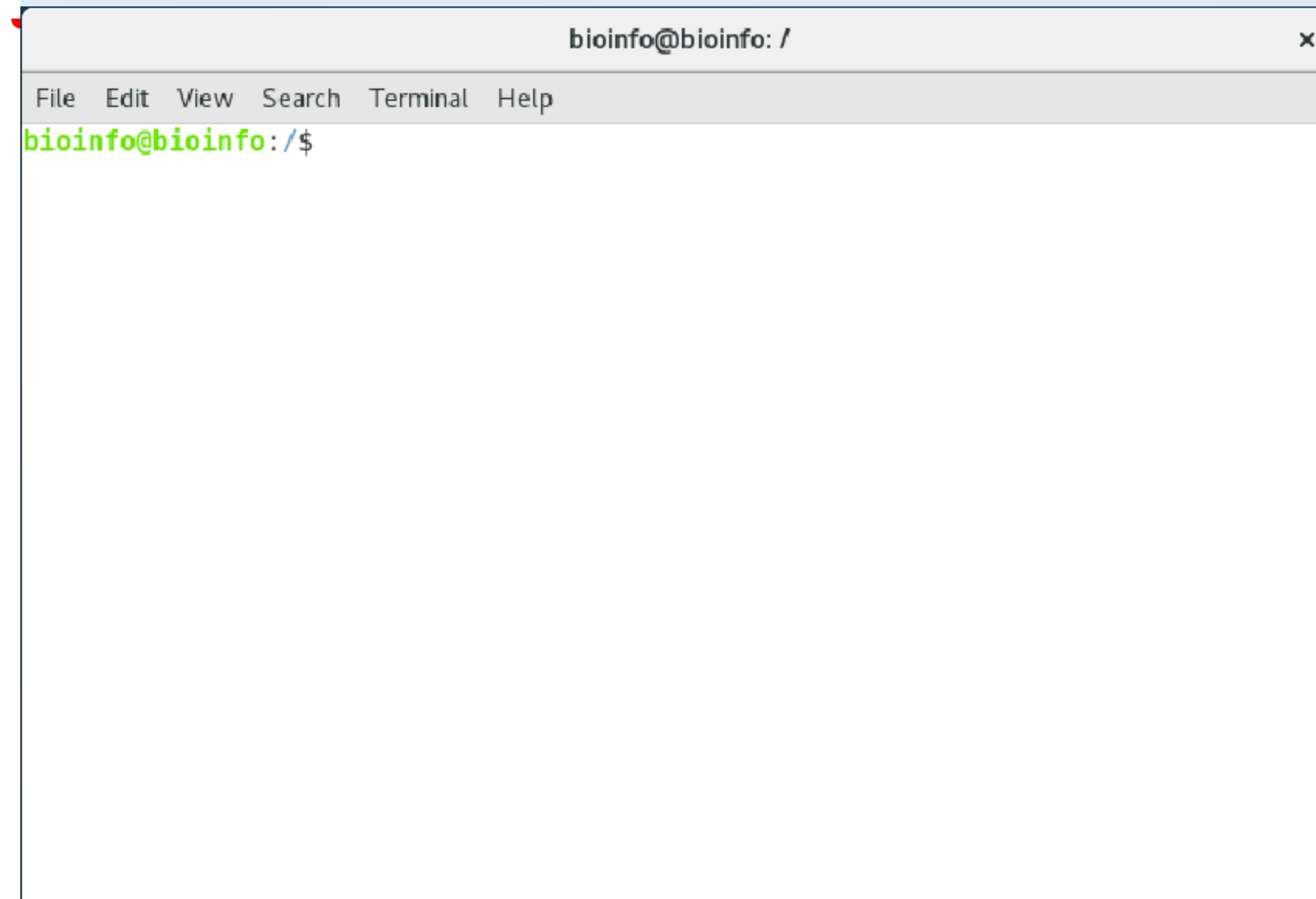
ctrl-c	stop process
ctrl-a	go to begin of line
ctrl-e	go to end of line
ctrl-r	search in command history



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).

A terminal window titled "bioinfo@bioinfo: /" with a menu bar containing "File", "Edit", "View", "Search", "Terminal", and "Help". The terminal shows a green prompt "bioinfo@bioinfo:/\$" on the first line.

```
bioinfo@bioinfo: /  
File Edit View Search Terminal Help  
bioinfo@bioinfo:/$
```



## 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).

```
bioinfo@bioinfo: /bin
File Edit View Search Terminal Help
bioinfo@bioinfo:/$ ls /bin/ntfs*
/bin/ntfs-3g      /bin/ntfscmp      /bin/ntfsls      /bin/ntfstruncate
/bin/ntfs-3g.probe /bin/ntfsfallocate /bin/ntfsmove    /bin/ntfsusermap
/bin/ntfscat      /bin/ntfsfix      /bin/ntfsrecover /bin/ntfswipe
/bin/ntfscluster  /bin/ntfsinfo     /bin/ntfssecaudit
bioinfo@bioinfo:/$ cd /bin
bioinfo@bioinfo:/bin$ ls ntfs*
ntfs-3g      ntfscluster  ntfsfix      ntfsmove     ntfstruncate
ntfs-3g.probe ntfscmp      ntfsinfo     ntfsrecover  ntfsusermap
ntfscat      ntfsfallocate ntfsls       ntfssecaudit ntfswipe
bioinfo@bioinfo:/bin$
```



# Escaping special characters

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

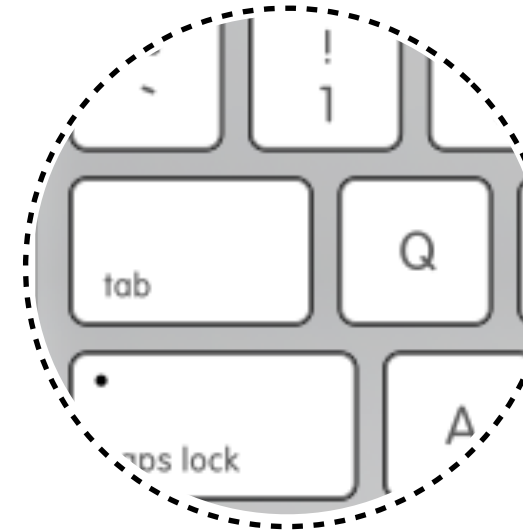
```
ls my_folder
```

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

creates an empty file called tmp\_file.txt

copies tmp\_file.txt in file\_copy.txt

Tip: name files in lower case  
and with underscores  
instead of using spaces

```
touch tmp_file.txt
```

```
cp tmp_file.txt file_copy.txt
```

```
mv file1.txt file2.txt
```

moves or renames a file

```
rm file.txt
```

deletes file.txt

# Locate a file

Locate the path for the file *unix\_class\_file\_samples.zip*

```
locate unix_class_file_samples.zip
```

```
locate unix_class
```

Locate the path for all the files containing *unix\_class*

# 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
```

delete *dir\_name* and its files

```
cp -r dir_name dir_copy
```

copy *dir\_name* and its files in a new folder



Music





Pictures



programs

# Compression commands

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

 <b>UNIX Command-Line Cheat Sheet</b> <small>BTI-SGN Bioinformatics Course 2014</small> 	
File system Commands	
<b>ls</b>	lists directories and files
<b>ls -a</b>	lists all files including hidden files
<b>ls -lh</b>	formatted list including more data
<b>ls -t</b>	lists sorted by date
<b>pwd</b>	returns path to working directory
<b>cd dir</b>	changes directory
<b>cd ..</b>	goes to parent directory
<b>cd /</b>	goes to root directory
<b>cd</b>	goes to home directory
<b>touch file_name</b>	creates an empty file
<b>cp file file_copy</b>	copy a file
<b>cp -r</b>	copy files contained in directories
<b>rm file</b>	deletes a file
<b>rm -r dir</b>	deletes a directory and its files
<b>mv file1 file2</b>	moves or renames a file
<b>mkdir dir_name</b>	creates a directory
<b>rmdir dir_name</b>	deletes a directory
<b>locate file_name</b>	searches a file
<b>man command</b>	shows commands manual
<b>top</b>	shows process activity
<b>df -h</b>	shows disk space info
Compression commands	
<b>gzip/zip</b>	compress a file
<b>gunzip/unzip</b>	decompress a file
<b>tar -cvf</b>	groups files
<b>tar -xvf</b>	ungroups files
<b>tar -zcvf</b>	groups and gzip files
<b>tar -zxvf</b>	gunzip and ungroups files
Text handling commands	
<b>command &gt; file</b>	saves STDOUT in a file
<b>command &gt;&gt; file</b>	appends STDOUT in a file
<b>cat file</b>	concatenate and print files
<b>cat file1 file2 &gt; file3</b>	merges files 1 and 2 into file3
<b>cat *fasta &gt; all.fasta</b>	concatenates all fasta files in the current directory
<b>head file</b>	prints first lines from a file
<b>head -n 5 file</b>	prints first five lines from a file
<b>tail file</b>	prints last lines from a file
<b>tail -n 5 file</b>	prints last five lines from a file
<b>less file</b>	view a file
<b>less -N file</b>	includes line numbers
<b>less -S file</b>	wraps long lines
<b>grep 'pattern' file</b>	Prints lines matching a pattern
<b>grep -c 'pattern' file</b>	counts lines matching a pattern
<b>cut -f 1,3 file</b>	retrieves data from selected columns in a tab-delimited file
<b>sort file</b>	sorts lines from a file
<b>sort -u file</b>	sorts and return unique lines
<b>uniq -c file</b>	filters adjacent repeated lines
<b>wc file</b>	counts lines, words and bytes
<b>paste file1 file2</b>	concatenates the lines of input files
<b>paste -d ","</b>	concatenates the lines of input files by commas
<b>sed</b>	transforms text
Networking Commands	
<b>wget URL</b>	download a file from an URL
<b>ssh user@server</b>	connects to a server
<b>scp</b>	copy files between computers
<b>apt-get install</b>	installs applications in linux



# Compression commands

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

BTi

UNIX Command-Line Cheat Sheet

BTI-SGN Bioinformatics Course 2014

File system Commands

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

Compression commands

<b>gzip/zip</b>	compress a file
<b>gunzip/unzip</b>	decompress a file
<b>tar -cvf</b>	groups files
<b>tar -xvf</b>	ungroups files
<b>tar -zcvf</b>	groups and gzip files
<b>tar -zxvf</b>	gunzip and ungroups files

Text handling commands

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

Networking Commands

<b>wget URL</b>	download a file from an URL
<b>ssh user@server</b>	connects to a server
<b>scp</b>	copy files between computers
<b>apt-get install</b>	installs applications in linux

files, directories or wildcards

# Compression commands

compress file f1.txt in f1.txt.gz

compress files f1 and f2 in file.zip

```
gzip f1.txt
```

```
zip file.zip f1 f2
```

```
unzip file.zip
```

```
gunzip file.gz
```

decompress file.zip

decompress file.gz

# Networking Commands

- Networking commands

BTi		UNIX Command-Line Cheat Sheet		
		BTI-SGN Bioinformatics Course 2014		
File system Commands		Text handling commands		
<b>ls</b>	lists directories and files	<b>command &gt; file</b>	saves STDOUT in a file	
<b>ls -a</b>	lists all files including hidden files	<b>command &gt;&gt; file</b>	appends STDOUT in a file	
<b>ls -lh</b>	formatted list including more data	<b>cat file</b>	concatenate and print files	
<b>ls -t</b>	lists sorted by date	<b>cat file1 file2 &gt; file3</b>	merges files 1 and 2 into file3	
<b>pwd</b>	returns path to working directory	<b>cat *fasta &gt; all.fasta</b>	concatenates all fasta files in the current directory	
<b>cd dir</b>	changes directory	<b>head file</b>	prints first lines from a file	
<b>cd ..</b>	goes to parent directory	<b>head -n 5 file</b>	prints first five lines from a file	
<b>cd /</b>	goes to root directory	<b>tail file</b>	prints last lines from a file	
<b>cd</b>	goes to home directory	<b>tail -n 5 file</b>	prints last five lines from a file	
<b>touch file_name</b>	creates an empty file	<b>less file</b>	view a file	
<b>cp file file_copy</b>	copy a file	<b>less -N file</b>	includes line numbers	
<b>cp -r</b>	copy files contained in directories	<b>less -S file</b>	wraps long lines	
<b>rm file</b>	deletes a file	<b>grep 'pattern' file</b>	Prints lines matching a pattern	
<b>rm -r dir</b>	deletes a directory and its files	<b>grep -c 'pattern' file</b>	counts lines matching a pattern	
<b>mv file1 file2</b>	moves or renames a file	<b>cut -f 1,3 file</b>	retrieves data from selected columns in a tab-delimited file	
<b>mkdir dir_name</b>	creates a directory	<b>sort file</b>	sorts lines from a file	
<b>rmdir dir_name</b>	deletes a directory	<b>sort -u file</b>	sorts and return unique lines	
<b>locate file_name</b>	searches a file	<b>uniq -c file</b>	filters adjacent repeated lines	
<b>man command</b>	shows commands manual	<b>wc file</b>	counts lines, words and bytes	
<b>top</b>	shows process activity	<b>paste file1 file2</b>	concatenates the lines of input files	
<b>df -h</b>	shows disk space info	<b>paste -d ","</b>	concatenates the lines of input files by commas	
Compression commands		<b>sed</b>	transforms text	
<b>gzip/zip</b>	compress a file	Networking Commands		
<b>gunzip/unzip</b>	decompress a file	<b>wget URL</b>	download a file from an URL	
<b>tar -cvf</b>	groups files	<b>ssh user@server</b>	connects to a server	
<b>tar -xvf</b>	ungroups files	<b>scp</b>	copy files between computers	
<b>tar -zcvf</b>	groups and gzip files	<b>apt-get install</b>	installs applications in linux	
<b>tar -zxvf</b>	gunzip and ungroups files			



# Networking Commands

connects your terminal to your account in a server

Downloads the BCBC logo!

```
ssh user_name@server_address
```

```
wget https://btiscience.org/wp-content/uploads/BCBClogo.png
```

```
scp afp@boyce.sgn.cornell.edu:/home/afp/file.txt .
```

copy *file.txt* from your home in the server to the current directory in your computer

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

# Useful commands in the server

A terminal window with a black background. Two commands are visible: 'df -h' and 'top'. The 'df -h' command is enclosed in an orange rounded rectangle, and a horizontal line connects it to a grey box containing the text 'shows disk space information'. The 'top' command is also enclosed in an orange rounded rectangle, and a vertical line connects it to a grey box below the terminal containing the text 'display and update sorted information about processes'.

```
df -h
```

shows disk space information

```
top
```

display and update sorted information about processes

## Exercise:

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

```
bioinfo@bioinfo: ~/Desktop
File Edit View Search Terminal Help
bioinfo@bioinfo:~$ cd Desktop/
bioinfo@bioinfo:~/Desktop$ wget https://btiscience.org/wp-content/uploads/BCBClogo.png
--2018-06-13 23:30:19-- https://btiscience.org/wp-content/uploads/BCBClogo.png
Resolving btiscience.org (btiscience.org)... 104.196.145.254
Connecting to btiscience.org (btiscience.org)|104.196.145.254|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 17882 (17K) [image/png]
Saving to: 'BCBClogo.png'

BCBClogo.png      100%[=====>]  17.46K  --.-KB/s    in 0s
2018-06-13 23:30:20 (233 MB/s) - 'BCBClogo.png' saved [17882/17882]

bioinfo@bioinfo:~/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

```
bioinfo@biodebian: ~
File Edit View Terminal Help
top - 15:07:10 up 3:50, 2 users, load average: 0.00, 0.00, 0.00
Tasks: 116 total, 1 running, 115 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.3%us, 0.0%sy, 0.0%ni, 99.7%id, 0.0%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 1026940k total, 518232k used, 508708k free, 51872k buffers
Swap: 1134584k total, 0k used, 1134584k free, 221056k cached

  PID USER      PR  NI  VIRT  RES  SHR  S  %CPU  %MEM    TIME+  COMMAND
 132 root        20   0    0     0     0   0  S   0.3   0.0   0:07.05 ata/0
1861 bioinfo    20   0 40744 1700 1088  S   0.3   0.2   0:28.12 VBoxClient
2234 bioinfo    20   0 220m  12m 9792  S   0.3   1.3   0:00.61 gnome-terminal
2304 bioinfo    20   0 19072 1352 1012  R   0.3   0.1   0:00.03 top
   1 root        20   0  8356   804   672  S   0.0   0.1   0:00.93 init
   2 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kthreadd
   3 root        RT    0     0     0     0  S   0.0   0.0   0:00.00 migration/0
   4 root        20   0     0     0     0  S   0.0   0.0   0:00.04 ksoftirqd/0
   5 root        RT    0     0     0     0  S   0.0   0.0   0:00.00 watchdog/0
   6 root        20   0     0     0     0  S   0.0   0.0   0:00.15 events/0
   7 root        20   0     0     0     0  S   0.0   0.0   0:00.00 cpuset
   8 root        20   0     0     0     0  S   0.0   0.0   0:00.00 khelper
   9 root        20   0     0     0     0  S   0.0   0.0   0:00.00 netns
  10 root        20   0     0     0     0  S   0.0   0.0   0:00.00 async/mgr
  11 root        20   0     0     0     0  S   0.0   0.0   0:00.00 pm
  12 root        20   0     0     0     0  S   0.0   0.0   0:00.02 sync_supers
  13 root        20   0     0     0     0  S   0.0   0.0   0:00.02 bdi-default
  14 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kintegrityd/0
  15 root        20   0     0     0     0  S   0.0   0.0   0:00.06 kblockd/0
  16 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kacpid
  17 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kacpi_notify
  18 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kacpi_hotplug
  19 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kseriod
  21 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kondemand/0
  22 root        20   0     0     0     0  S   0.0   0.0   0:00.00 khungtaskd
  23 root        20   0     0     0     0  S   0.0   0.0   0:00.00 kswapd0
  24 root        25   5     0     0     0  S   0.0   0.0   0:00.00 ksmd
  25 root        20   0     0     0     0  S   0.0   0.0   0:00.00 aio/0
  26 root        20   0     0     0     0  S   0.0   0.0   0:00.00 crypto/0
 130 root        20   0     0     0     0  S   0.0   0.0   0:00.00 ksuspend_usbd
 131 root        20   0     0     0     0  S   0.0   0.0   0:00.00 khubd
```

q quit  
u user (top -u user)  
M sort by memory usage



# Commands to install software

```
aptitude search blast
```

```
sudo aptitude install blast2
```

```
sudo apt-get install pbzip2
```

installs *pbzip2* in your computer

call the command with super user permissions

ubuntu®





# 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`