**RNASeq Workshop**

November 5-7, 2020 Florida International University

November 12-14, 2020 University of Puerto Rico

November 19-21, 2020 University of Puerto Rico

**Answers to Linux Practice Material.**

**Information:**

* **Commands to complete exercises are highlighted in yellow;**
* **Commands and outputs are included in grey boxes;**
* **Replace the word Username by your own username at the different exercises;**
* **Note different questions are answered in sections (a, b,c,…) in order to facilitate progression of answers.**

**Exercise 1:** Please list all the steps that you needed to do to login to the HPC cluster at FIU. For each step write the command and the parameters used. For windows users who use MobaXterm please provide screenshot of the MobaXterm window where you enter the login details.

Steps:

1. Open MobaXterm app and click on “session” icon on the menu bar
2. Click “SSH” on the “Session settings” window.
3. Enter wolf.cs.fiu.edu as “Remote host”
4. Check the box next to “Specify username”
5. Enter username in the text box next to “Specify username”
6. Click OK
7. type Password [It will look like it is not typing but it is storing the information]

 It will now show:

 Welcome to the new wolf.cs.fiu.edu, now running CentOS 7!

 username@wolf:~ 7%

To access worker node:

 username@wolf:~ 7% ssh  username@bioinfo-01.cs.fiu.edu

It will ask password again. Hit enter. It now shows username@bioinfo-01:~$ meaning you are now logged into node 1 of the FIU cluter. [After the 1st use of MobaXtem you can just go to “previous session” (on the left panel of the screen) and just type the worker node and proceed from there as listed above, at worker node step].



**Exercise 2:** After you login to your home directory, create a directory with name “Test\_project”.

$ mkdir test\_project

**a)** Inside this directory

cd /lclhome/ username/test\_project

 **b)** create three sub directories with names (without quotes): “test\_data”, “test\_results” and “test\_src”.

~/test\_project$ mkdir test\_data

~/test\_project$ mkdir test\_results

~/test\_project$ mkdir test\_src

**c)** Inside each of these three sub directories create a directory with name “temp” and three text files with names: <DirectoryName>\_file1.txt, <DirectoryName>\_file2.txt and <DirectoryName>\_file3.txt

 “cd” changes directory and “touch” creates file

cd /lclhome/ username/test\_project/test\_data

test\_data$ mkdir temp

cd /lclhome/ username/test\_project/test\_data/

touch test\_data\_file1.txt

touch test\_data\_file2.txt

touch test\_data\_file3.txt

cd /lclhome/ username/test\_project/test\_results/

test\_results$ mkdir temp

touch test\_results\_file1.txt

touch test\_results\_file2.txt

touch test\_results\_file3.txt

cd /lclhome/ username/test\_project/test\_src/

test\_src$ mkdir temp

touch test\_src\_file1.txt

touch test\_src\_file2.txt

touch test\_src\_file3.txt

**Exercise 3:** Go into each of the three sub directories created in Exercise 2 and run a command that

* 1. lists the directory you are in:

 username@bioinfo-01:~/test\_project/test\_data$ pwd

/lclhome/ username/test\_project/test\_data

 username@bioinfo-01:~/test\_project/test\_results$ pwd

/lclhome/ username/test\_project/test\_results

 username@bioinfo-01:~/test\_project/test\_src$ pwd

/lclhome/ username/test\_project/test\_src

* 1. ) lists all the contents of the directory

 username@bioinfo-01:~/test\_project/test\_data$ ls

temp test\_data\_file1.txt test\_data\_file2.txt test\_data\_file3.txtt

 username@bioinfo-01:~/test\_project/test\_results$ ls

temp test\_results\_file1.txt test\_results\_file2.txt test\_results\_file3.txt

 username@bioinfo-01:~/test\_project/test\_src$ ls

temp test\_src\_file1.txt test\_src\_file2.txt test\_src\_file3.txt

**a)** lists all the contents of the directory with detailed information about each file. How do you differentiate a file from a directory using the information listed? please note the command ls -l is typed with the letter l

Answer: Directory is listed with a “dr”

Detailed info commands:

 username@bioinfo-01:~/test\_project/test\_data$ ls -l [This is the letter l]

total 8

drwxr-xr-x 2 username user 4096 Jan 31 15:00 temp

-rw-r--r-- 1 username user 18 Jan 31 15:51 test\_data\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 15:55 test\_data\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 15:56 test\_data\_file3.txtt

 username@bioinfo-01:~/test\_project/test\_results$ ls -l

total 4

drwxr-xr-x 2 username user 4096 Jan 31 16:10 temp

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file3.txt

 username@bioinfo-01:~/test\_project/test\_src$ ls -l

total 4

drwxr-xr-x 2 username user 4096 Jan 31 16:11 temp

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:02 test\_src\_file3.txt

**b)** lists all the contents of the directory with detailed information about each file sorted by file modification time

 username@bioinfo-01:~/test\_project/test\_data$ ls -lt

total 8

-rw-r--r-- 1 username user 0 Jan 31 15:56 test\_data\_file3.txtt

-rw-r--r-- 1 username user 0 Jan 31 15:55 test\_data\_file2.txt

-rw-r--r-- 1 username user 18 Jan 31 15:51 test\_data\_file1.txt

drwxr-xr-x 2 username user 4096 Jan 31 15:00 temp

 username@bioinfo-01:~/test\_project/test\_data$

 username@bioinfo-01:~/test\_project/test\_results$ ls -lt

total 4

drwxr-xr-x 2 username user 4096 Jan 31 16:10 temp

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file3.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file1.txt

 username@bioinfo-01:~/test\_project/test\_src$ ls -lt

total 4

drwxr-xr-x 2 username user 4096 Jan 31 16:11 temp

-rw-r--r-- 1 username user 0 Jan 31 16:02 test\_src\_file3.txt

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file1.txt

**c)** lists all the contents of the directory with detailed information about each file in a human readable file sizes format

 username@bioinfo-01:~/test\_project/test\_data$ ls -lh

total 8.0K

drwxr-xr-x 2 username user 4.0K Jan 31 15:00 temp

-rw-r--r-- 1 username user 18 Jan 31 15:51 test\_data\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 15:55 test\_data\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 15:56 test\_data\_file3.txtt

 username@bioinfo-01:~/test\_project/test\_results$ ls -lh

total 4.0K

drwxr-xr-x 2 username user 4.0K Jan 31 16:10 temp

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:05 test\_results\_file3.txt

 username@bioinfo-01:~/test\_project/test\_src$ ls -lh

total 4.0K

drwxr-xr-x 2 username user 4.0K Jan 31 16:11 temp

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file1.txt

-rw-r--r-- 1 username user 0 Jan 31 16:01 test\_src\_file2.txt

-rw-r--r-- 1 username user 0 Jan 31 16:02 test\_src\_file3.txt

 username@bioinfo-01:~/test\_project/test\_src$

  **d)** lists all the contents of the directory including all the hidden files

 username@bioinfo-01:~/test\_project/test\_results$ ls -a

. .. temp test\_results\_file1.txt test\_results\_file2.txt test\_results\_file3.txt

 username@bioinfo-01:~/test\_project/test\_data$ ls -a

. .. temp test\_data\_file1.txt test\_data\_file2.txt test\_data\_file3.txtt

 username@bioinfo-01:~/test\_project/test\_src$ ls -a

. .. temp test\_src\_file1.txt test\_src\_file2.txt test\_src\_file3.txt

**Exercise 4**:

**a)** Change permissions for the test\_src directory created in Exercise 2 so that only the owner has read, write and execute permissions

 username@bioinfo-01:~/test\_project/test\_src$ chmod o-rwx /lclhome/ username/test\_project/test\_src

**b)** Change permissions for the test\_results directory so that only owner and group has read and write permissions.

 username@bioinfo-01:~/test\_project/test\_results$ chmod ug+rw /lclhome/ username/test\_project/test\_results

**Exercise 5:** Go to test\_data directory and create a text file using nano. Copy paste the following text and save the file with name “test\_dataset.txt”. Within the saved file using the nano editor please do the following changes. For the lines that begin with “@” symbol please delete all the characters in that line and replace them with the text: “sequence header”. Please write down all the commands (or steps) that you have used for creating the file and for modifying the text in the saved file.

~/test\_project$ nano

Copy and paste the text below inside the nano screen you see

@V300017787L4C002R0061184898/1

GTTCATTGCATTATTTAATAATATCCCGGTGTTTTCGCTGACTACTCCGCTTCCGAAATAGAAATTAATAGTACTAGTAGCTGATACTGCATCTCCATTT

+

FFFGGFAFGGEGDGFFEG@GEFFFFGGFGGGEFFGGGGFBFFGEFEGFAFFFGGGEGFB?GEFGGGEFGAGGFGFDEGAGFFF?GAGFGGFBGGFGFGFF

@V300017787L4C002R0061184920/1

CTTTATTAGAGCGCCACAATCTTTCCATACCATTACGATGTAAAGCCATTCGAGTAAGTTCAGAAAAACTTTCACGAATATTAAAATCACAGAGAGGTGA

+

F6CCDFF'BF0??8FAB=F;BF@FE0DG1FD8DCEF0FFEDFBCE=GDBFB<=F&:@G5FF@DF8F;FEFGFFE<F?CGGC>DEGEFDDCEGFF=07DF8

@V300017787L4C002R0061184952/1

ATGAGATAGGAAATCAAAGGGTATCTACAACAGCACGTATTATTCATGGAATAATTATAGTTTTTATTTTAATTTCTATAATATTAGCTATAGCAAATGT

+

?FFDFEFCFEEFFFFFEFF?FEFFFFFEF@FFEFFFFBFFFFFFFFFFDEFFFFFEFFFFEEFEBFDEFFCFFFFFFFFFFEFFFDFEFDAFC?FFFFF=

Save the file pressing keyboard keys Ctrl+o (small letter o) and hit enter

To move the cursor to a specific line and character number, use the Ctrl+\_ keys. The menu on the bottom of the screen will change. Enter the number 1 in the “Enter line number, column number:” field and hit Enter

Cursor now moved to the first line of text and you can move up and down, delete the @ like the exercise wants and replace it by “sequence header”.

Save changes to file Ctrl+o and hit enter

Exit nano with keys Ctrl+x

It will ask save modified buffer. Type Y and hit enter. You are now back to the original screen

**Exercise 6:**

**a)** Please type a command on your terminal so that you print first 6 lines of the file you have saved in the previous exercise.

 username@bioinfo-01:~/test\_project$ head -6 test\_dataset.txt

**b)** What command(s) you would use to browse the entire content of the file?

 username@bioinfo-01:~/test\_project$ cat test\_dataset.txt

**c)** Please type a command on your terminal so that you print all the lines that contain “@” in the test\_dataset.txt file

 username@bioinfo-01:~/test\_project$ grep @ test\_dataset.txt

FFFGGFAFGGEGDGFFEG@GEFFFFGGFGGGEFFGGGGFBFFGEFEGFAFFFGGGEGFB?GEFGGGEFGAGGFGFDEGAGFFF?GAGFGGFBGGFGFGFF

F6CCDFF'BF0??8FAB=F;BF@FE0DG1FD8DCEF0FFEDFBCE=GDBFB<=F&:@G5FF@DF8F;FEFGFFE<F?CGGC>DEGEFDDCEGFF=07DF8

?FFDFEFCFEEFFFFFEFF?FEFFFFFEF@FFEFFFFBFFFFFFFFFFDEFFFFFEFFFFEEFEBFDEFFCFFFFFFFFFFEFFFDFEFDAFC?FFFFF=

**d)** Please type a command on your terminal so that you print total number of lines that contain “@” in the test\_dataset.txt file

 username@bioinfo-01:~/test\_project$ grep -c @ test\_dataset.txt

3

**Exercise 7:**

**a)** Using any Linux editor, create a file with the following text with file name “samples.txt” and save it [like you did in exercise 5]

C4-45-2

C4-45-3

C4-45-4

T3-45-1

T3-45-2

T3-45-4

**b)** Sort the file in reverse order (alphanumerically)

 username@bioinfo-01:~/test\_project$ sort -n samples.txt

 username@bioinfo-01:~/test\_project$ sort -r samples.txt

  **c)** Type a command to print for each line of the samples.txt file only the first two characters before the first “-” symbol.

 username@bioinfo-01:~/test\_project$ cut -c-2 samples.txt

C4

C4

C4

T3

T3

T3

**Exercise 8:**

**a)** Copy the samples.txt file created in the previous exercise to test\_results directory.

 username@bioinfo-01:~/test\_project$ cp samples.txt /lclhome/ username/test\_project/test\_results/

**b)** Move the samples.txt file from the test\_results directory to test\_src directory

First change permission to access the folders listed in b) since in exercise 4 we had changed permission. We change to allow permission owner and group to read, write and execute, as follows.

 username@bioinfo-01:~$ chmod ug+rwx /lclhome/ username/test\_project/test\_results

 username@bioinfo-01:~$ chmod ug-rwx /lclhome/ username/test\_project/test\_src

Now you can move the file:

 username@bioinfo-01:~/test\_project$ mv samples.txt /lclhome/ username/test\_project/test\_results/

and confirm the file is now in directory test\_results:

 username@bioinfo-01:~/test\_project$ cd test\_results

 username@bioinfo-01:~/test\_project/test\_results$ ls

samples.txt temp test\_results\_file1.txt test\_results\_file2.txt test\_results\_file3.txt

**Exercise 9:** Go to test\_src directory and within that directory create a bash script to run fastqc program. Enter parameters to run the fastqc program so that it lists only the help menu and exit.

[This exercise is for advanced Linux users]

**Exercise 10**: Go to the home directory and create a tar gzip file of the test\_project directory that was created earlier. Create a directory with name “Backup” within your home directory and copy the tar gzipped file that you just created to the “Backup” directory. Within the “Backup” directory untar and unzip the file.

**a)** Go to home directory

 username@bioinfo-01:~/test\_project$ ssh bioinfo-01.cs.fiu.edu

**b)** change permission in directory test\_src in order to be able to add it to the tar gzip file

 username@bioinfo-01:~$ chmod 775 /lclhome/ username/test\_project/test\_src

**c)** Create a tar gzip file of the test\_project directory [you will see all the files and directories inside test\_project listed below your command line]

 username@bioinfo-01:~$ tar -czvf test\_project.tar.gz /lclhome/ username/test\_project

tar: Removing leading `/' from member names

/lclhome/ username/test\_project/

/lclhome/ username/test\_project/test\_src/

/lclhome/ username/test\_project/test\_src/test\_src\_file3.txt

/lclhome/ username/test\_project/test\_src/test\_src\_file1.txt

/lclhome/ username/test\_project/test\_src/test\_src\_file2.txt

/lclhome/ username/test\_project/test\_src/temp/

/lclhome/ username/test\_project/test\_project.tar.gz

/lclhome/ username/test\_project/test\_dataset.txt

/lclhome/ username/test\_project/test\_results/

/lclhome/ username/test\_project/test\_results/samples.txt

/lclhome/ username/test\_project/test\_results/test\_results\_file2.txt

/lclhome/ username/test\_project/test\_results/test\_results\_file1.txt

/lclhome/ username/test\_project/test\_results/test\_results\_file3.txt

/lclhome/ username/test\_project/test\_results/temp/

/lclhome/ username/test\_project/test\_data/

/lclhome/ username/test\_project/test\_data/test\_data\_file2.txt

/lclhome/ username/test\_project/test\_data/test\_data\_file3.txtt

/lclhome/ username/test\_project/test\_data/test\_data\_file1.txt

/lclhome/ username/test\_project/test\_data/temp/

**d)** Create a directory with name “Backup” within your home directory and copy the tar gzipped file that you just created to the “Backup” directory.

 username@bioinfo-01:~$ mkdir Backup

 username@bioinfo-01:~$ cp test\_project.tar.gz Backup

**Exercise 11:** Please download the tar gzipped file that was created in the previous exercise on to your laptop. Please upload any small text file (with size <1Mb) from your laptop to your home directory in the HPC cluster.

[This exercise will be explained by Dr. Kiran Donthu at the workshop, however you can learn how to enter into the File Transfer section in the MobaxTerm screen. Steps as follows ]

1. Click on Sessions - far left top of the MobaXterm control bar. 
2. Click on SFTP. Type in remote host: wolf.cs.fiu.edu as well as your username and click OK.
3. It will now show wolf.cs.fiu (username) in the top bar as well as all your folders in your local computer.