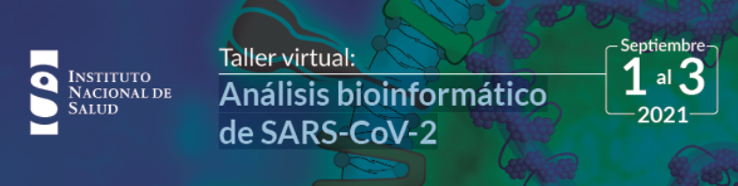
****

**TALLER 1: Introducción al uso de la consola Unix/Linux**

Orientado por: Carlos Ernesto Maldonado y Sebastián Hernández Botero.

Correos: [carlosernestomaldonado@gmail.com](mailto:carlosernestomaldonado@gmail.com) y [johanhdz03@gmail.com](mailto:johanhdz03@gmail.com)

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**INSTALACIÓN Y CONEXIÓN AL SERVIDOR**

**Pueden accedor vía shh por programas como Putty o mobaxterm**

La IP de conexión es 168.176.61.249

El nombre de usuario es el mismo del correo sin la parte del "@" y el dominio, por ejemplo. Para el correo [**apinzon@unal.edu.co**](mailto:apinzon@unal.edu.co)

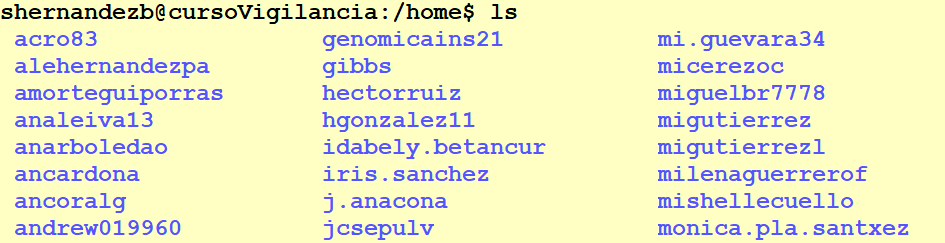
su usuario es **apinzon**, sin el @[unal.edu.co](http://unal.edu.co/)

Así para todos los usuarios.

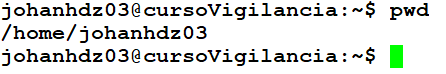
La clave en este momento es para todos la misma: **rnsg2021**

Es **URGENTE** que en la primera clase le pidan a todos los usuarios que **cambien su clave** por una personal, mediante el comando "password".

Al ingresar, en la carpeta Home, encontrará su carpeta de trabajo.



Así, por ejemplo, el usuario johanhdz03 **al ingreso se verá lo siguiente:**

****

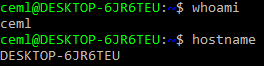
**Parte 1.** Algunos comandos básicos

Nota: los comandos a digitar en la terminal fueron escritos en fuente Currier New con el propósito de diferenciarlos del resto del texto

1. En la terminal identifique nombre de usuario y máquina en la que se encuentra la sesión

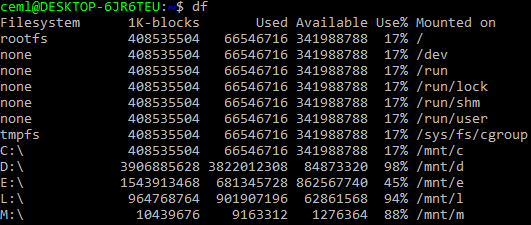
Comandos: whoami y hostname





2. Liste las unidades del sistema y el uso de los discos

Use el comando: df



3. Cuál es su ubicación actual en el sistema

Comando: pwd



4. Cree un directorio de trabajo

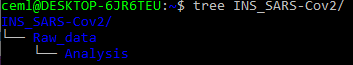
Comando: mkdir <nombre\_directorio>

Tip: Pueden crearse directorios anidados separando los nombres con el argumento -p y los nombres de los directorios separados por /



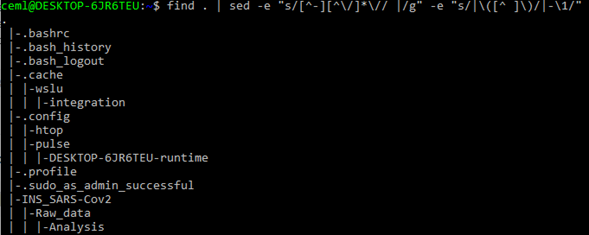
5. Despliegue la estructura del directorio

Comando: tree



Si el programa tree no está instalado puede correrse el siguiente comando integrando diferentes herramientas simultáneamente:

find . | sed -e "s/[^-][^\/]\*\// |/g" -e "s/|\([^ ]\)/|-\1/"



En el comando anterior el comando de sistema find se separa con | del programa editor sed.

6. Acceda al primer nivel del directorio creado

Comando: cd <PATH>

Tip: El PATH puede incluir múltiples niveles separando el nombre del directorio por /. La tecla Tab permite autocompletar, si escribo la letra I seguido de Tab completara el nombre del directorio o listará los directorios y archivos que empiecen por esa letra.



Acceda al tercer nivel de la estructura de directorios creada



7. Regrese al nivel del directorio INS\_SARS-Cov2

Comando: cd ../..

Tip: el comando cd sin argumentos lo lleva al directorio Home igual que cd ~. El argumento .. permitirá subir un nivel y puede usarse para subir cuantos niveles sea requerido al separarlos por /



8. En el punto 4 creamos un directorio de tres niveles, mueva el directorio Analysis al mismo nivel del directorio Raw\_data

Comando: mv PATH\_Directorio PATH\_nueva\_ubicacion

Tip: el carácter especial **.** indica “aquí”.



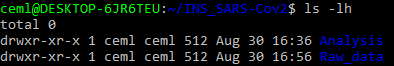
9. Haga una lista del contenido de directorio en su posición actual en el sistema.

Comando: ls



El comando ls tiene múltiples argumentos para diferentes salidas en el stdout, ejecute el comando man ls para ver el manual del comando, puede bajar con Enter y salir con q, puede desplazarse en el texto con las flechas. Puede verse la ayuda de muchos programas abriendo la ayuda con las opciones -h o --h o --help (en el caso de ls la opción -h está definida para dar valores de tamaño de archivo en unidades legibles por humanos).

Se puede hacer una lista detallada del contenido con el comando: ls -lh



10. Acceda al directorio Raw data y copie los datos crudos en ese directorio

Comando: cp <PATH\_origen> <PATH\_destino>

Tip: si se encuentra en el directorio destino puede usar el carácter especial . que indica aquí.

cp /data/setTutorial\_articFiles/barcode09/barcode09.muscle.in.fasta .

Una forma de copiar los archivos verificando la integridad de la copia es por medio del comando rsync.

Comando: rsync -av --progress PATH\_directorio\_o\_archivo\_original PATH\_directorio\_o\_archivo\_copia

rsync -av --progress /data/setTutorial\_articFiles/barcode09/barcode09.muscle.in.fasta .

TIP: Si tiene problema para encontrar este archivo, el mismo está en su carpeta de usuario.

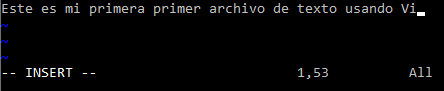
11. Cree un archivo con datos generales del experimento utilizando el editor de texto vi

Comando: vi <nombre\_archivo>

Tip: a diferencia de otros sistemas operativos los archivos creados en Linux no requieren extensión, no obstante es útil para identificar de qué tipo de archivo se trata.



Usar la tecla i para insertar texto, escribir la información que considere y salir guardando de con oprimiendo la tecla Esc y digitando :wq



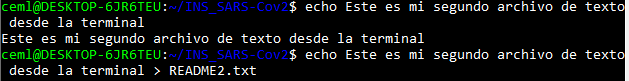
El archivo puede editarse usando el mismo comando: vi README.txt, usar la instrucción insertar (i) para modificar el texto o salir sin guardar con la tecla Esc y digitando :q!

12. Cree un archivo de una línea desde la terminal

Comando: echo texto

echo texto > nombre\_archivo

Con esta sintaxis echo imprime en el stdout el texto escrito después de él, el caracter especial > redireccionará el texto a un archivo



13. Adicione una línea de texto a un archivo existente.

Comando: echo texto >> archivo\_existente



Nota: si no se usa >> sino > se sobreescribirá el archivo dejando solo el texto seguido del comando echo.

14. Ver texto de un archivo

Comando: more nombre\_archivo   
 head nombre\_archivo

El comando head permite por defecto ver las primeras 10 líneas de un archivo, ver man head o head --help para más información.

El comando more permite abrir extensos archivos de texto Ctrl-c permite salir de un comando en ejecución. Usar man more o more --help para más opciones.

~/INS\_SARS-CoV2/Raw\_Data$ more barcode09.muscle.in.fast #salir con Ctrl-c

~/INS\_SARS-CoV2/Raw\_Data$ head -n 50 barcode09.muscle.in.fasta

~/INS\_SARS-CoV2/Raw\_Data$ head -50 barcode09.muscle.in.fasta

15. Cambiar el nombre de un archivo o directorio

Comando: mv nombre\_actual nombre\_nuevo







Revise el contenido del directorio al cual le cambió el nombre, no se afecta el contenido.

Tip: Cambiar el nombre de un archivo o directorio que tiene espacios

Comando: mv "folder 1" folder\_1

o mv folder\ 1 folder\_1

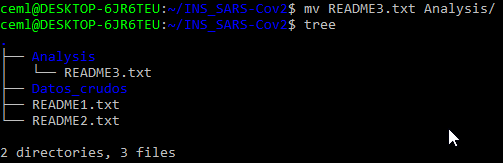
16. Copiar archivo cambiando su nombre

Comando: cp nombre\_actual nombre\_nuevo



17. Mover un archivo a otro directorio

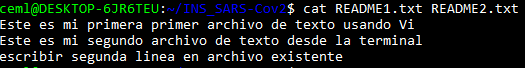
Comando: mv nombre\_archivo directorio\_destino



Tip: el caracter especial / al final del nombre del directorio indica “dentro de”, la información se copiará dentro del directorio.

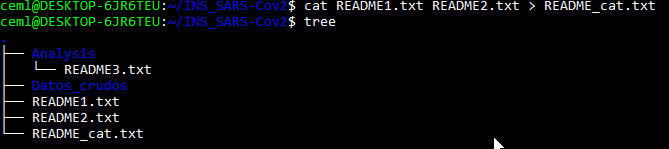
18. Concatenar archivos

Comando: cat archivo-1 archivo-2 … archivo-n



19. Concatenar archivos y guardarlo en un nuevo archivo

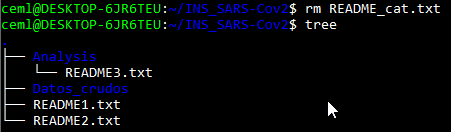
Comando: cat archivo-1 archivo-2 > nuevo\_archivo



20. Borrar un archivo

Comando: rm archivo

Nota: linux desde la terminal **no tiene** Papelera de reciclaje y no es posible deshacer al borrar directorios o archivos. Se recomienda incluir la opcion -i que preguntara antes de borrar y responder y o n (yes/no).



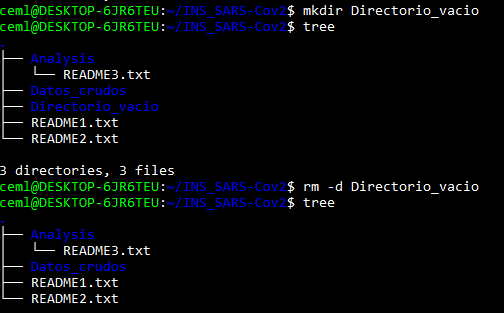


21. Borrar un directorio vacío

Comando: rm -d directorio

o rmdir directorio

Nota: linux desde la terminal no tiene Papelera de reciclaje y no es posible deshacer al borrar directorios o archivos. Se recomienda incluir la opción -i que preguntará antes de borrar y se responde con y o n (yes/no) seguido de Enter.

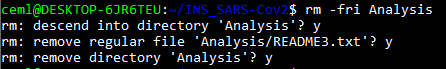




22. Borrar un directorio y su contenido

Comando: rm -fri directorio

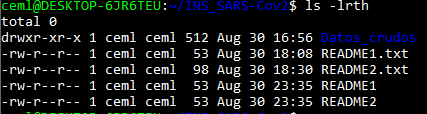
Tip: revisar manual (man) o --help del comando para ver opciones de ejecución, en este caso se hará un borrado forzado (f) y recursivo (r) preguntando antes de la eliminación del directorio.



23. Borrar todos los archivos con prefijo o sufijo común

Para este ejercicio puede hacer copias de los archivos del taller, por ejemplo:



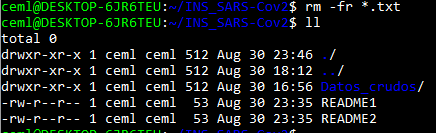


Nótese que todos los archivos presentan el prefijo común README y solo dos archivos el sufijo común .txt. Nótese también la utilización del comando ls con las opciones de lista detallada (l), ordenado de acuerdo al tiempo de creación iniciando por el más reciente (t), con listado reverso (r) para tener la lista de acuerdo al tiempo de creación de los archivos y tamaño de archivos en unidades legibles por humanos (h).

Borrado de archivos con sufijo común:

Comando: rm -fr \*sufijo

El uso de las opciones -fr es opcional y se usa para borrar sin tener en cuenta condiciones que podrían impedir el borrado.

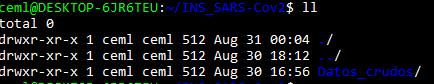


Nótese la utilización del caracter especial **\*** llamado comodin. En este comando se eliminó todo lo que terminaba en .txt

Tip: Notese también la forma de ejecutar la lista detallada donde ll es el comando corto para ls -l ; este atajo no está en todos los sistemas pero puede crearse con el comando alias (cuando se tiene permisos de administrador) de la siguiente forma: alias ll=”ls -l”. Podría usarse también para modificar la forma de ejecutar un comando ya asignado como rm para borrar con: alias rm="rm -i"

Borrado de archivos con prefijo común:

Comando: rm -fr prefijo\*



24. Algunos usos del comodín

Cómo movería todo el contenido de un directorio al directorio actual?

Comando: mv PATH/\* .

Cómo copiaría todo el contenido de un directorio a otro directorio?

Comando: cp PATH/\* .

cp PATH/\* PATH\_destino/

Cómo concatenaría todos los archivos con sufijo común de un directorio?

Comando: cat PATH/\*.txt

**Parte 2.** Algunos comandos básicos con archivos de secuencias

Se utilizaran un archivos fasta y multifasta cuya estructura es:

>nombre\_de\_la\_secuencia  
secuencia

Ejemplo:

>MZ498086.1 |B.1.2  
TAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACTTTAAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAACTAATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTT…

El archivo multifasta tiene múltiples secuencias fasta una debajo de la otra.

~/INS\_SARS-CoV2/Datos\_Crudos$ cp /data/Taller\_1/\* .

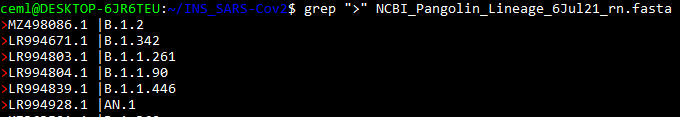
25. Cuantas lineas tiene un archivo?, en este caso un archivo multifasta

Comando: wc -l archivo.fasta



26. Cuales son los nombres de las secuencias en un archivo multifasta?

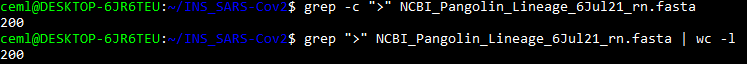
Comando: grep ">" archivo.fasta



27. Cuantas secuencias hay en un archivo multifasta?

Comando: grep -c ">" archivo.fasta

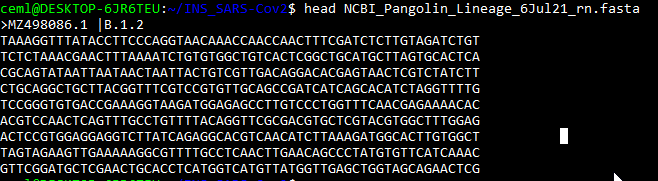
O grep ">" archivo.fasta | wc -l



28. Ver líneas de texto de un archivo multifasta

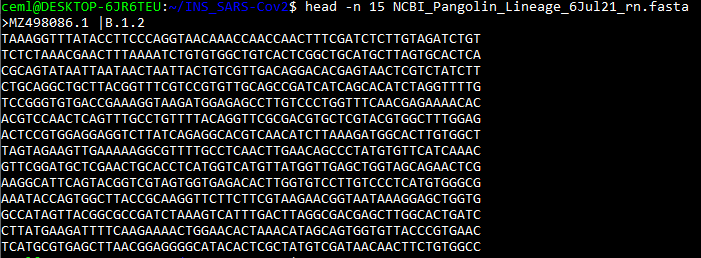
Ver las primeras 10 lineas

Comando: head archivo.fasta



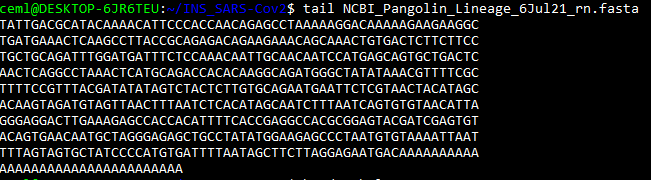
Ver las primeras n líneas

Comando: head -n NUM archivo.fasta



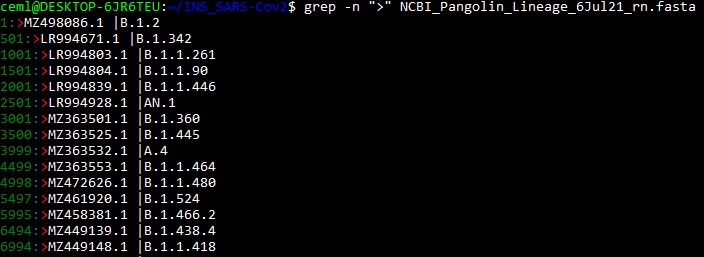
Ver las últimas 10 lineas

Comando: tail archivo.fasta



29. En que numero de linea empieza cada secuencia de un archivo multifasta?

Comando: grep -n ">" archivo.fasta



30. Con la información del punto 29, cómo extraer la primera secuencia de un archivo multifasta?

Direccionado a stdout:

Comando: head -n 500 NCBI\_Pangolin\_Lineage\_6Jul21\_rn.fasta

Guardado en un archivo:

Comando: head -n 500 NCBI\_Pangolin\_Lineage\_6Jul21\_rn.fasta > MZ498086.1\_B.1.2.fasta

30. Cuantas palabras tiene la secuencia en un archivo fasta?

Comando: cat -v ">" secuencia.fasta | wc -w



31. Cuantos caracteres tiene la porción de la secuencia en un archivo fasta?

Comando: grep -v ">" secuencia.fasta | wc -m



En el comandos anteriores se excluye la linea del nombre de la secuencia de un archivo fasta con el comando grep -v ">", excluyendo las líneas que contengan el caracter especial > y que cuente el número de caracteres con el comando wc opción -m. Tener en cuenta que si el >, siendo caracter especial, no se pone entre comillas va a guardar el resultado de grep -v en un archivo con el string escrito inmediatamente después, es decir si se ejecuta

grep -v > secuencia.fasta

se sobre-escribirá el archivo secuencia.fasta borrando toda la información del archivo, no es posible deshacer, es importante trabajar con copias de los archivos al manipularlos en la terminal.

32. Cuántos caracteres tiene la secuencia en un archivo fasta?, Cuál es la longitud de la secuencia en pb?

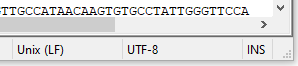
Comando: grep -v ">" secuencia.fasta | tr -d '\n' | wc -m

O grep -v ">" secuencia.fasta | awk '{ printf "%s", $0 }' | wc -m





Al comparar el resultado del punto 31 con el del punto 32 se observa que quedan en las secuencias caracteres especiales que no se ven, en este caso son los separadores de línea que corresponden a la expresión regular \n los cuales deben ser quitados antes de hacer el conteo de caracteres que dará el tamaño de la secuencia y se ejecutó con el comando tr y awk para llegar al tamaño real del genoma de referencia de SARS-Cov2. Si el archivo esta codificado como documento de Windows el tamaño del genoma sera erroneo por presencia de otro tipo de caracteres especiales, todos los archivos para trabajar en terminal deben tener la codificación para Unix. Se recomienda trabajar en editores de texto como Notepad++ o Komodo verificando siempre la codificación del archivo.



33. Como guardar el historial de comandos ejecutados?

Comando: history > history\_fecha.txt

Este archivo hará las veces de log del trabajo que haya ejecutado que se puede editar posteriormente en un editor de texto de interfaz gráfica o el programa vi o vim en línea de comandos de Linux.

Puede limitarse el número de líneas de comandos de ejecución de la última sesión integrando el comando tail con la opción -n que da el número de líneas por ejemplo:

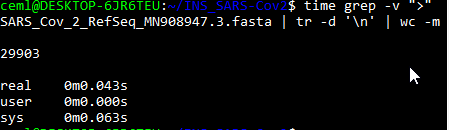
history | tail -n > history\_fecha.txt

O ver en el stdout los últimos 10 comandos con:

history | tail -n

34. Como conocer el tiempo de ejecución de una línea de comando?

Comando: time comando



Comando de utilidad para determinar en tiempo de máquina y tiempo humano en el que se obtuvo la salida de determinado proceso en una computadora con recursos determinados.