

## **PRACTICA I: INTRODUCCION A LOS SISTEMAS OPERATIVOS WINDOWS Y UNIX.**

**Objetivo general:** Conocer los aspectos básicos del manejo de los sistemas operativos Windows y Unix.

**Objetivos particulares:**

- Conocer los comandos básicos para el manejo de archivos y directorios utilizando la interfaz gráfica de MS-Windows.
- Conocer la línea de comandos de MS-DOS y los principales comandos para la manipulación de archivos y directorios.
- Conocer técnicas de acceso remoto y los comandos básicos de UNIX.

### **INTRODUCCIÓN**

Las computadoras personales son equipos de propósito general que pueden realizar una amplia variedad de tareas.

El componente principal de estos equipos es una unidad central de procesamiento (CPU) que coordina el funcionamiento de diversos dispositivos (hardware) que a su vez desempeñan funciones de entrada, almacenamiento y salida de información. La CPU de una computadora es un microprocesador (o colección de microprocesadores) que llevan a cabo las funciones aritméticas, lógicas y de control de la computadora. Además estas incluyen chips de memoria de *solo lectura* (Read Only Memory, ROM) que contienen instrucciones, grabadas en forma permanente, necesarias para el arranque y operación del equipo. La función principal de almacenamiento de una computadora, la realizan los chips de memoria de acceso aleatorio (Random Access Memory, RAM). Adicionalmente, los equipos de cómputo, incluyen una serie de dispositivos para la entrada o salida de datos, que son los que permiten la interacción con el usuario. En general, todos estos componentes son designados como Hardware.

Los programas de cómputo o software son paquetes de instrucciones que permiten que el hardware de la computadora realice diversas funciones. Los programas de software pueden dividirse en dos grandes grupos: Los sistemas operativos y las aplicaciones.

El sistema operativo es un programa básico para la computadoras, ya que este se requiere manipular y coordinar los componentes de un sistema de cómputo. Los sistemas operativos cumplen tres funciones básicas:

1. Controlar el hardware. Un sistema operativo controla las partes de un sistema de cómputo y coordina el funcionamiento de las mismas.
2. Administrar información. El sistema operativo provee formas de administrar y organizar la información almacenada en una computadora. Por ejemplo, el sistema determina como acomodar, copiar, mover, eliminar o visualizar archivos.
3. Ejecutar aplicaciones de software. El sistema operativo permite ejecutar otras aplicaciones de software, las cuales en general son diseñadas para resolver problemas específicos de los usuarios. Es importante mencionar que generalmente estas aplicaciones, son específicas para cada sistema operativo.

En la actualidad existen diversos sistemas operativos, para computadoras personales entre los cuales pueden contarse MS-Windows, Mac OS X y Linux. Los sistemas operativos MS-Windows y Mac OS X son comerciales en tanto existen muchos sistemas Linux que pueden obtenerse en forma gratuita y descargarlos directamente de Internet. Los sistemas Mac OS X y Linux actuales, tienen en común el que ambos derivan de sistemas operativos tipo UNIX, los cuales se emplean como sistemas operativos en equipos de cómputo más complejos como servidores y supercomputadoras.

Actualmente la mayoría de estos sistemas operativos cuentan con interfaces gráficas que permiten al usuario llevar a cabo las funciones básicas de todo sistema operativo. No obstante, en la mayoría de estos sistemas existen comandos avanzados, que permiten realizar funciones complejas y los cuales generalmente se pueden aplicar mejor a través de consolas de comando o terminales textuales.

En las consolas de comandos textuales el usuario escribe directamente las instrucciones que se desean ejecutar. Generalmente a través de estas se pueden ejecutar comandos con una gran variedad de parámetros, que serían difíciles de configurar mediante una interfaz gráfica. Las consolas de comandos pueden ser muy poderosas para desempeñar tareas complejas pero tienen el inconveniente de que el usuario a menudo tiene que recordar los nombres de los comandos, estos a su vez pueden constar de una larga lista de parámetros y además se debe escribirlos correctamente, lo que puede hacer tediosa la realización de ciertas tareas.

Sin embargo, para lograr obtener el máximo aprovechamiento de las tareas básicas de una computadora es apropiado tener un conocimiento básico de la forma como se estructura la información almacenada y la ejecución de comandos desde la interfaz gráfica y la textual. Por este motivo en esta práctica se ha diseñado para que el alumno aprenda algunas de las bases de cómo se almacena y organiza la información en sistemas MS-Windows y tipo UNIX (Como Linux o Mac OS X). En particular se dará énfasis a la ejecución de instrucciones mediante la consola de comandos y al uso de algunas aplicaciones que permiten el acceso remoto o otras computadoras.

## **RECURSOS INFORMÁTICOS**

<http://bioinformatica.homelinux.org/bioinformatica> (Servidor del área de bioinformática de la ENCB).

<http://www.ncbi.nlm.nih.gov> (Página del National Center for Biotechnology Information)

Software FileZilla (aplicación para el acceso remoto) (<http://filezilla-project.org/> )

Sistema Operativo Windows (95, 98, XP, Vista, Windows 7).

Sistema Operativo Tipo Unix (Linux Fedora o Mac OS X).

## **DESARROLLO**

## **Recomendaciones generales sobre la creación de archivos.**

El sistema operativo MS-DOS fue la versión anterior de las versiones actuales MS-Windows y se basa exclusivamente en una interfaz textual desde la que se escriben los comandos que se desea desempeñar. El MS-DOS original era un sistema operativo de 16-bits y admitía únicamente nombres de archivos y directorios de longitudes no mayores a 8 caracteres. Los nombres debían comenzar con una letra de la A-Z, podían contener números y el guión bajo “\_”. Los espacios y otros símbolos tales como signos de puntuación y los acentos no estaban permitidos. Adicionalmente, después del nombre del archivo se agregaba un punto y una extensión de tres letras que identificaba el tipo de archivo. Dado que algunos de los programas que actualmente se manejan en bioinformática fueron desarrollados inicialmente en dicho sistema operativo, es común que todavía algunos de ellos sigan estrictamente las reglas para nombrar archivos y directorios características de MS-DOS. La versión de Windows 3.x introdujo la Interfaz gráfica para manipular comandos básicos de MS-DOS pero el sistema aún seguía siendo de 16-bits. A partir de la versión de Windows 95 a la fecha, el sistema operativo pasó a ser de 32-bits e introdujo mejoras notables a la estructura de archivos, entre las que destacan la posibilidad de usar nombres largos de archivos y sin la restricción del uso de símbolos para ellos. Las versiones Windows más recientes son sistemas operativos de 64-bits.

## **Administración de usuarios en Windows.**

En la mayoría de los sistemas operativos actuales, se pueden tener cuentas para diversos niveles de usuario. Cada nivel restringe las actividades que el usuario puede hacer en el sistema. En general debe existir al menos un usuario administrador, el cual en principio tiene un control total sobre el sistema. Otros usuarios, tendrán cuentas limitadas los cuales tienen menos privilegios que los administradores por ejemplo para permitir la instalación de programas o para configurar el entorno gráfico. A continuación veremos como se pueden verificar o cambiar los permisos de usuario en Windows.

1. **Encender la máquina y esperar a que arranque el sistema operativo.** Dependiendo de la configuración de la máquina es posible que tenga que seleccionar el usuario correcto de una lista de los usuarios posibles o teclear el nombre de usuario. En sistemas en donde solo está dado de alta un único es posible que el acceso sea directo. Después de seleccionar el usuario es posible que tenga que proporcionar una clave de acceso (password).
2. **Verificación de los privilegios de la cuenta de usuario.** En el “Panel de control” p buscar “Cuentas de usuario” y ejecutar esta opción. Se muestra un cuadro en el que se listan las cuentas de usuario disponibles y la forma como se han configurado.

La cuenta que se está empleando debería estar configurada como “Administrador de equipo” ya que esto es necesario para poder instalar programas en el sistema. Si la cuenta no esta configurada de esta manera, se recomienda iniciar Windows en modo Administrador, y crear un usuario con tales privilegios o modificar un usuario existente. Para esto último se da un clic en la cuenta de usuario y en el cuadro de opciones que aparece presionar “Cambiar mi tipo de cuenta”. En el cuadro mostrado

se debe cambiar el tipo de cuenta de “Limitada” a “Administrador de equipo”. Reiniciar el equipo e ingresar de nuevo en la cuenta.

3. **Visualización de extensiones de archivo.** La visualización de extensiones es importante a la hora de renombrar archivos. La opción por defecto en sistemas Windows es ocultar las extensiones de tipos de archivos conocidos. Para mostrar buscar en el panel de control “Opciones de carpeta” con lo que se despliega un cuadro de dialogo. Dar clic en la carpeta “Ver” y buscar la sección de “Configuración avanzada”. En esta lista buscar con las barras de desplazamiento la opción “Ocultar las extensiones de archivo para tipos de archivo conocidos” y deshabilitarla en caso necesario dando un clic con el ratón (la palomilla debe desaparecer). Presionar el botón “aceptar”. Los archivos visualizados en el explorador deberán mostrar ahora la extensión que determina el tipo de archivo.

### **Estructura de directorios en Windows y tareas básicas para la manipulación de archivos.**

1. **Creación de la carpeta de trabajo.** Abrir el explorador desde el icono “Mi PC” y buscar el disco local (“C”). Dar doble clic en el icono con lo que se entra en el directorio principal del disco. En el panel de tareas comunes buscar “Tareas de archivo y carpeta” y presionar en la opción “Crear nueva carpeta”. Aparece una carpeta en el explorado titulada “Nueva carpeta”, cambiar el nombre a “Bioinfo”. La carpeta también puede crearse dando un clic en una zona en blanco del explorador con el botón derecho del ratón y seleccionando en las opciones mostradas “Nuevo” -> “Carpeta”. También se puede crear desde el menú “Archivo” del explorador” siguiendo las opciones “Nuevo” -> “Carpeta”. Una vez creada la carpeta dar doble clic para entrar en ella. Observe que en la barra de direcciones aparece la ruta de la carpeta, la cual debe ser C:\Bioinfo.
2. **Otras tareas dentro del explorador.** El explorador incluye tareas que permiten crear, cambiar nombres, cortar, pegar y eliminar archivos o carpetas entre otras tareas. Examinar las opciones disponibles en el menú “Archivo” o en el menú que se activa al presionar con el botón derecho del ratón.

### **Introducción al manejo de la consola de comandos de MS-DOS.**

1. **Abrir la consola de comandos.** La consola de comandos de MS-DOS puede activarse presionando en el botón “Inicio” y en el menú entrar a “Todos los programas”. Buscar el menú “Accesorios” y ahí presionar la opción “Símbolo del sistema”. Otra opción consiste en dar un clic en el botón “Inicio” y presionar “Ejecutar...”, escribir *cmd* en la línea del cuadro mostrado y presionar “Aceptar”. En cualquiera de las dos opciones aparece un cuadro en el cual se puede teclear comandos. El cursor parpadea en la línea activa. El cuadro muestra información sobre la versión de Windows y la ruta del directorio en el cual se trabaja.
2. **Unidades de almacenamiento en MS-DOS.** La unidades de almacenamiento (Disco duro, disquetes, CD, DVD y memorias USB-Flash) se identifican con letras (Normalmente A y B se usan para los disquetes, C, se reserva para el disco local, D, M y N se utilizan para discos locales adicionales, CD, DVD y memorias Flash). El comando <unidad>: sirve para cambiar de unidad de almacenamiento, por ejemplo

al escribir c: y presionar ENTER se cambia a la unidad de almacenamiento "C". Si una memoria USB tiene asignada la letra "F", el comando F: permitirá tener acceso al contenido de la unidad. Siempre se debe teclear ENTER para ejecutar un comando en MS-DOS.

- Operaciones con archivos y directorios.** Para cambiar al directorio raíz de la unidad de trabajo ejecutar el comando "**chdir \**" o bien la forma abreviada "**cd \**" (omitir las comillas al teclear los comandos). Se puede especificar el directorio raíz de una unidad de trabajo, por ejemplo para pasar a la raíz de la unidad F, escribir "**cd F:\**". Entrar a la carpeta "**Bioinfo**" recién creada con "**cd c:\Bioinfo**". El comando "**mkdir**" o su forma abreviada "**md**" permite crear directorios. En la carpeta "**Bioinfo**" ejecutar las instrucciones "**md secdna**", "**md secprot**", "**md estruct**", "**md alineam**" y "**md temporal**". Observar en el explorador de archivos que los comandos anteriores han creado cinco carpetas dentro de "**Bioinfo**". El comando "**rmdir**" o su forma abreviada "**rd**" permite eliminar directorios vacíos. Dentro de la carpeta "**Bioinfo**" ejecutar el comando "**rd temporal**". Se pregunta si se está seguro de borrar la carpeta, presionar "**S**" para confirmar la operación. El comando "**dir**" muestra los directorios, subdirectorios y archivos. Al ejecutar "**dir c:\bioinfo/s**" se muestran todos los subdirectorios y archivos contenidos dentro de bioinfo.
- Métodos abreviados.** Para cambiar desde la raíz del disco local hasta el directorio "**secdna**" contenido dentro de "**bioinfo**", se puede teclear "**cd bioinfo**" y luego "**cd secdna**" pero se puede hacer la misma operación en un solo paso escribiendo "**cd c:\bioinfo\secdna**". Como se ha indicado anteriormente "\ " indica la raíz de la unidad activa. Los dos puntos ".." representan la instrucción de ir a el directorio previo (por ejemplo "cd ..") El punto "." representa el directorio actual. Ejecutar "**cd c:\bioinfo\secdna**" con lo que se entra al directorio "**secdna**" dentro de "**Bioinfo**", luego ejecutar "**cd ..\secprot**", nótese que este comando sería equivalente a teclear "**cd c:\Bioinfo\secprot**". El asterisco "\*" se utiliza como comodín para escribir comando que permitan manipular fácilmente archivos o directorios múltiples. Por ejemplo el comando "**dir \***" muestra todos los directorios y archivos, "**dir a\***" muestra todos los directorios y archivos que empiezan con "a", "**dir \*.exe**" muestra todos los archivos con la extensión "exe".
- Otros comandos importantes.** Los comandos "**erase**" o "**md**" permiten borrar archivos, "**copy**" y "**xcopy**", permiten copiar archivos y directorios, "**move**" permite mover archivos y directorios. "**cls**" permite borrar la pantalla. Para información detallada de los comandos escribir "/?" después del nombre del comando, por ejemplo "**copy/?**" muestra las opciones disponibles para el comando "**copy**". Por ultimo el comando "**type**" puede ser útil para examinar el contenido de un archivo.
- Lista de los comandos disponibles en la consola de MS-DOS de Windows XP. El siguiente link da una lista completa de los comandos disponibles en la línea de comandos de MS-DOS que pueden emplearse en la consola:

<http://www.microsoft.com/resources/documentation/windows/xp/all/proddocs/en-us/ntcmds.mspx?mfr=true>

## Acceso remoto a archivos a servidores tipo UNIX.

FTP es el acrónimo de File Transfer Protocol, un protocolo para facilitar la comunicación entre equipos en una red, principalmente para la transferencia de archivos. En este caso se ilustrará el uso de FTP para descargar archivos a partir de servidores tipo UNIX.

1. Ejecución del programa FTP (File Transfer Protocol). Desde la consola de MS-DOS ejecutar el comando “**ftp**”. Notar que el formato de la consola cambia un poco y ya no es visible la letra de unidad del sistema. *Nota: Si en la consola de ftp ejecuta “help” aparece una lista de los comandos disponibles para “ftp”.*
2. **Ingreso al servidor FTP.** Ejecutar “open <ftp.ncbi.nlm.nih.gov>”. Se muestra un mensaje en el que se indica que se tiene acceso al servidor NCBI y se solicita una clave de usuario, teclear “**anonymous**”, el sistema pregunta ahora por una contraseña, escribir una dirección de correo electrónico personal y presionar ENTER, se ha tenido acceso remoto al servidor ftp del NCBI.
3. **Comandos FTP.** El comando “ls -l” muestra la estructura de directorios del sistema. “**cd genomes**” permite cambiar al directorio “**genomes**”. En los sistemas UNIX es muy importante respetar la forma como son escritos los nombres de archivos y directorios respetando mayúsculas y minúsculas. El comando “**get**” permite descargar archivos al directorio activo en la computadora local. Si el archivo a descargar es un programa o un archivo comprimido se debe ejecutar el comando “**binary**” antes de intentar descargar el archivo. El comando “**quit**” permite abandonar el servicio ftp. Para este ejemplo trate de ingresar al directorio genomes, de ahí ingresar a la carpeta “Bacteria”. Seleccione un genoma (por ejemplo Escherichia coli K12) ingrese al mismo. Observe los directorios. Trate de descargar un genoma (archivos con la extensión \*.fna) con el comando get <archivo>. El archivo deberá ser guardado en el directorio en que se encontraba en la consola al ejecutar el comando ftp.
4. **Acceso FTP desde navegadores WEB.** El acceso a servidores ftp también puede hacerse desde un navegador WEB como el Internet Explorer, Firefox o Netscape. Abrir el navegador WEB y teclear <ftp.ncbi.nlm.nih.gov>. Puede verse que el explorador muestra la estructura de archivos del servidor y en el panel izquierdo se muestra que se ha entrado al sistema como el usuario “**anonymous**”. No todos los servidores ftp permiten el acceso anónimo al sistema, en cuyo caso solicitaran un nombre de usuario y una contraseña. Muchos servidores ftp tienen una carpeta de acceso pública frecuentemente llamada “**pub**” la cual puede accesarse de manera anónima. Dentro de sistema de carpetas ftp pueden usarse los comandos del explorador para abrir y descargar archivos. Normalmente las operaciones de escritura en estos servidores no son permitidas y solo se llevan a cabo operaciones de descarga de archivo mediante el comando “**get** <nombre del archivo>”. El comando anterior descargará el archivo solicitado en el directorio en que se encontraba en la consola al ejecutar el programa ftp.

## Introducción al uso de sistemas tipo UNIX.

Los sistemas UNIX son ambientes que fueron concebidos desde su creación, como sistemas multiusuario. En dichos sistemas, en general existe un usuario administrador o superusuario, denominado *root* el cual tiene control total sobre el sistema y por lo mismo rara vez se tiene acceso directo al mismo. En UNIX algunos comandos tales como *sudo* o *su* permiten tener acceso a funciones del administrador y pueden requerirse especialmente para la instalación de programas. Para otras operaciones, en general se trabajará con cuentas de usuario. A continuación se hará una breve revisión de algunos comandos que son útiles en sistemas UNIX/Linux.

### 1. Instalación de un programa para acceso remoto a servidores.

La mayoría de los sistemas UNIX admiten protocolos para el acceso remoto. Dos protocolos comúnmente empleados son el FTP (File Transfer Protocol) y el SSH (Secure Shell). Para la práctica resulta conveniente emplear alguna herramienta gráfica para facilitar el acceso remoto a servidores mediante SSH o FTP:

En esta práctica se utilizará FileZilla. Existen versiones de este programa para otros sistemas operativos y pueden descargarse de <http://filezilla-project.org/>. Para los sistemas UNIX/LINUX la instalación puede llevarse a cabo en forma automática abriendo una terminal de comandos.

Para esto, una vez que se ha ingresado a una cuenta de usuario se debe ejecutar la terminal y enseguida se debe ingresar como administrador:

su –

después de ejecutarlo este solicitará la contraseña del administrador. Para instalar FileZilla se ejecuta el comando siguiente:

```
yum install filezilla
```

Una vez instalado el programa, en la carpeta de Aplicaciones de Internet debería aparecer el icono del programa. Este también puede ejecutarse desde la consola con el comando filezilla.

### 2. Conexión al sitio FTP del NCBI para la descarga de genomas.

Ejecutar FileZilla y configurar los siguientes parámetros:

**host:** ftp://ftp.ncbi.nlm.nih.gov/  
**user:** anonymous  
**password:** una dirección de correo electrónico válida.  
**port:** 21

Al realizar la conexión, el programa Filezilla nos mostrará en la ventana derecha la lista de carpetas del sitio FTP. En las ventanas a la izquierda se muestran las carpetas en la computadora de trabajo. El usuario debería seleccionar en esta última ventana una carpeta en la cual se descargarán los archivos.

En este ejemplo se va a realizar la transferencia de los archivos conteniendo los datos del genoma, transcriptoma y proteoma de una bacteria. Para esto, en el sitio FTP del NCBI se debe ingresar a:

Genomes > Bacteria > Escherichia\_coli\_K\_12\_substr\_\_DH10B\_uid58979.

En esta carpeta se deben seleccionar y descargar los archivos NC\_010473 con las extensiones fna, faa, ffn y ppt.

A continuación se da una descripción general de los contenidos de cada tipo de archivo:

fna: Secuencia de nucleótidos del genoma completo en formato FASTA

faa: Secuencia de aminoácidos de cada una de las proteínas (proteoma) en formato FASTA.

ffn: Secuencias codificantes (CDS) que codifican para cada una de las proteínas del archivo faa.

ppt: Tabla con la descripción detallada de las proteínas y CDS de los archivos faa y ffn incluyendo nombres, claves de acceso y abreviaturas de las proteínas y genes y sus posiciones con respecto al genoma.

Después de descargar los archivos, se debería verificar la integridad de los mismos abriéndolos con algún editor de texto tal como gedit o emacs.

Nota: emacs es otro editor de texto poderoso de los sistemas UNIX. Puede instalarse mediante el comando “yum install emacs”, recordando que los comandos “yum” deben ejecutarse con permisos de administrador.

### **3. Ejemplos de aplicación de comandos básicos:**

A continuación se da un resumen de comandos básicos que pueden emplearse en la consola.

#### **a) Permisos y mostrar contenido de directorios.**

Los sistemas UNIX/LINUX han sido concebidos como sistemas operativos multiusuario. Por este motivo existen diversos tipos de permisos para tener acceso a archivos o directorios.

Los archivos pueden tener asignados distintos permisos para el usuario (u), un grupo de usuarios (g) o bien para otros usuarios (o).

Los permisos ugo pueden ser w=write r=read y x=execute.

Es decir, para un usuario, en un archivo se pueden tener permisos para leerlo, modificarlo o ejecutarlo.

Los permisos de los archivos se ponen de manifiesto con los comandos “stat archivo” o “ls -l”

`stat archivo //Muestra estadísticas de acceso al archivo.`

El comando “ls” lista el contenido de un directorio. Con el atributo “l” se puede mostrar información detallada de cada archivo. Por ejemplo:

```
ls -l
```

La ejecución del comando anterior podría mostrar el contenido siguiente:

```
NC_010473.fna      -rwxrw-r--
SECUENCIAS        drwxrwxrwx
```

En el ejemplo anterior el comando ls nos muestra que el archivo NC\_010473.fna no es un directorio (los directorios se marcan con la letra *d* antes de los permisos). El usuario actual de la cuenta tiene permisos para leer, escribir y ejecutar el archivo (primera serie rwx). Sin embargo otros usuarios del grupo de trabajo pueden leer y modificar el archivo pero no ejecutarlo (serie rw-). Usuarios fuera del grupo de trabajo solo tienen permisos para leer el archivo (r--). Observe que el siguiente elemento en el listado, SECUENCIAS, está marcado como directorio, y cualquier usuario tiene permisos para leer, modificar y ejecutar dicho directorio.

## **b) Utilizar/cambiar el shell de comandos.**

El shell incorpora comandos que pueden ejecutarse en la consola que pueden emplearse para realizar operaciones complejas e incluso para para la automatización de tareas. El conjunto de comandos de cada shell es una especie de lenguaje de programación que puede emplearse para automatizar tareas por medio de la ejecución de archivos con series de estos comandos, denominados “scripts”. Cada shell tiene sus propios comandos y pueden variar de shell a shell. Los comandos que empleamos en este curso están basados en el shell “bash” que es el más común empleado en diversos sistemas Linux, no obstante existen otros tipos de shell tales como:

```
sh
bash
tsh
```

La ejecución del comando `cat /etc/shells` muestra una lista de shells disponibles en el sistema. Cada shell puede activarse desde la consola.

## **c) Comandos de uso general.**

**whoami** // Da el nombre de usuario activo.

**pwd** // Print Working Directory. Imprime la ruta actual de trabajo.

**which** *archivo* //Imprime la ruta donde se puede localizar un archivo

**cp** *archivo destino* //Copia el archivo a la dirección especificada (destino)

**mv** *archivo destino* //Mueve el archivo a la dirección especificada (destino)

#### d) Examinar/editar archivos.

**head** *nombre\_archivo* //Muestra las primeras 10 líneas del archivo

Ejemplo: abrir los archivos fna, faa, fna y ptt con el comando head, especificando 30 líneas de cada archivo.

**head -n 30** *nombre\_archivo* //Muestra las primeras 30 líneas del archivo

Note que en el comando anterior el parámetro “n” indica el número de líneas a mostrar, en caso de que 10 líneas no fuera suficiente.

**tail** *archivo* //Es similar a head pero este muestra las últimas 10 líneas

**tail -n 30** *nombre\_archivo* //muestra las últimas 30 líneas del archivo.

**cat** *archivo* //Muestra el contenido (completo) del archivo

**cat** *archivo1 archivo2 archivo3* //Muestra el contenido (concatenado) de los tres archivos.

**more** *archivo*//Ejecuta un editor de UNIX que muestra el contenido de todo el archivo. Presionando :q se puede salir del editor.

**wc** *archivo* //Muestra el número de líneas, palabras y bytes del archivo.

**wc -l** *archivo* //Muestra solo el número de líneas del archivo

**grep '>'** *archivo* //Muestra las líneas en el archivo que contienen el símbolo '>'

#### e) Tuberías.

Una característica importante de los sistemas UNIX/LINUX es la de las tuberías, que se refiere a que se pueden conectar varios comandos, con lo que la salida esperada de un comando, puede emplearse como entrada de otro comando, sin necesidad de crear un archivo intermedio.

Por ejemplo. Sin emplear tuberías deseamos contar el número de secuencias contenidas en el archivo faa. Para esto considerando que cada secuencia, tiene un título marcado con '>'. El siguiente comando enviaría la lista de los títulos de cada secuencia al archivo “títulos”.

**grep '>' NC\_010473.faa > titulos.txt**

Luego, el comando:

**wc -l titulos.txt**

mostraría el número de líneas en este archivo, que correspondería al número de secuencias en el archivo faa.

Sin embargo, empleando tuberías, el comando puede escribirse:

```
grep '>' NC_010473.faa | wc -l
```

el cual mostraría directamente el número de secuencias en el archivo, sin necesidad de crear un archivo intermedio como en el ejemplo anterior.

#### 4) Acceso remoto a otros equipos con SSH.

El protocolo SSH es muy común entre los sistemas UNIX y constituye un método bastante seguro para el intercambio de archivos (SSH también puede usarse en Windows, pero para tener acceso a equipos Windows debe instalarse un software para servidor SSH, este tópico no se cubre en esta práctica, pero si están interesado en este sistema se puede buscar el software Open SSH para Windows en Internet). Ya se ha visto como FileZilla permite el acceso a otros sistemas remotos mediante FTP. Este también puede usarse para tener acceso mediante SSH, para lo cual se debe tener acceso por el puerto 22 (por defecto).

1. **Acceso SSH mediante FileZilla.** Desde FileZilla se configura el nombre del “Host” y una clave de “usuario”. Para este ejercicio, teclear “**bioinformatica.homelinux.org**” para el nombre del Host y “**bioinfo**” para el User. El programa solicita una contraseña. Al proporcionar estos datos el sistema solicitará confirmar el acceso y mostrará la estructura de directorios de la cuenta que se ha ingresado.
2. **Acceso SSH mediante la consola.** Alternativamente, se puede tener acceso mediante una terminal mediante los comandos:

```
ssh bioinfo@bioinformatica.homelinux.org  
sftp bioinfo@bioinformatica.homelinux.org
```

ssh solo permite visualizar o editar archivos mientras sftp se utiliza para transferir archivos mediante los comandos get y put. Las operaciones de sftp, se ejecutan de forma más simple desde FileZilla.

3. **Algunos comandos básicos para LINUX.** Como ejercicio el alumno puede ensayar en el equipo remoto, algunos de los comandos básicos de Linux descritos anteriormente.

#### Ejercicios:

1. Crear un acceso en el escritorio de Windows para tener acceso inmediato a la línea de comando de MS-DOS. Considere que los accesos directos llaman a un programa y permiten pasar parámetros. Cmd.exe es el nombre del programa de la línea de comandos “**cmd.exe c:\**” permitiría abrir la ventana de comandos en el directorio raíz de la unidad “c”.
2. Con un programa como Word crear un archivo y guardarlo en la carpeta de

“Bioinfo”. Desde la línea de comandos cambiar de nombre a este archivo y copiarlo a la carpeta “secprot” de Bioinfo. Crear una carpeta con el nombre de usuario en minúsculas en el servidor ENCB y transferir el archivo usando el servicio SSH. Posteriormente eliminar el archivo en ambos sistemas.

3. Descargar genomas en formato FASTA de distintas bacterias en Linux en archivos individuales. Concatenar los archivos en un solo archivo y después ejecutar comandos típicos para contar el número de líneas y el número de archivos almacenados.
4. Si tiene acceso a un sistema Mac OS X, investigue si los comandos UNIX descritos anteriormente pueden aplicarse en la Terminal de este sistema operativo.

## **Bibliografía**

1. Gibas C, P. Jambeck, (2001): Developing Bioinformatics Computer Skills, O'Reilly, USA.
2. Constantinides A. (1987): Applied Numerical Methods with Personal Computers, McGraw-Hill international editions, Chemical Engineering Series, New York, USA, pags: 4-41.
3. Wünschiers R. (2004): Computational Biology: Unix/Linux, data processing and programming, Springer, Alemania,
4. Rathbone A. (2001). Windows XP para dummies, ST Editorial Hungry Minds, USA.
5. Microsoft Corporation, Microsoft Windows Server 2003 TechCenter, Server Help. <http://www.microsoft.com/resources/documentation/windows/xp/all/proddocs/en-us/ntcmds.msp?mfr=true>.