

## Getting Used to Linux

Part 1: You will want to keep your files organized, so you need to understand the “directory structure”.

1. Every operating system has a directory structure. Your local computer has a C drive. The desktop and document folders are somewhere on that drive, and then your actual files are nested within subfolders.
2. The root directory in Linux is the topmost directory (all other directories are nested within root). The root directory is represented by `/`.
  - a. Type **`ls /`** to show the contents of the root directory.
  - b. Most of these directories are part of the system. For instance **`/bin`** contains programs, and **`/mnt`** contains the storage devices.
3. Your home directory is very important. The home directory is also known as `~`. Typing **`cd`** or **`cd ~`** will both take you to the home directory. Your home directory will almost never be named “home”, even though there is a directory with that name. To see the actual name of your home directory (and its path), type **`pwd`**. You can also type **`echo $HOME`**, which will show the value of **`HOME`**.

Part 2: It can sometimes be convenient to work in a Linux environment on your local computer (but we won't really need to in this class).

1. On a Mac, the terminal provides a Linux-like environment. The Mac OS is built on Linux, so most of the functionality is there. There are some differences, though.
2. On Windows, it is possible to install a Linux subsystem. If you go to the Windows store and install Ubuntu, then you will have an app that launches into a Linux environment. This is a fully functional version of Linux.
3. On either Mac or Windows, it is possible to install a virtual machine running Linux. A popular, free virtual machine is Virtualbox (<https://www.virtualbox.org/>).
4. Using Linux on your local computer is slightly advanced, so only attempt it if you feel comfortable with computers.

Part 3: Make some directories.

1. In your home directory, use **`mkdir`** to make a `computer_skills` folder.
2. Move to this folder.
3. Make `week01` and `week02` folders.
4. Move to the `week01` folder.
5. Use **`nano`** to make a new text file. Type some text into the file and save it.
6. Move it to the `week02` folder.
7. Practice the use of Tab and the up arrow.

Part 4: Keyboard shortcuts.

1. Use the **Tab** key as much as you can.
2. Use **Ctrl-C** to stop a program that's running.
3. Use the **up arrow** to scroll through previous commands you have typed.
4. Other, less important shortcuts:
  - a. **Ctrl-A** goes to the beginning of the line.

- b. Ctrl-E goes to the end of the line.
- c. Ctrl-L clears the screen.
- d. Ctrl-U clears the line before the cursor.
- e. Ctrl-K clears the line after the cursor.

#### Part 5: Running programs.

1. To run a program it must be installed (obviously).
2. You just type the name of the program to run it.
3. The UI supercomputers use a module manager to manage access to the software.
  - a. Go to <https://hpc.uidaho.edu/> to get to the webpages associated with our computer.
  - b. There is a "Tutorials" link with tons of useful information.
  - c. Click on "Using Modules to load applications".
  - d. Type **module avail** to see what modules are available.
  - e. We will conduct a quick blast search just to demonstrate how to use a program.
  - f. Blast is in the ncbi-blast package, which is in the list of modules.
  - g. To use it, type **module load ncbi-blast**.
4. Getting help with a program: **man program-name** or **program -h** or **program -help** or **program --help**. It depends on the program.
5. Programs often accept arguments, and these are usually necessary to control the program's behavior.
  - a. For example, blast needs to know the filename of your query sequence and the database you would like to search.
6. Download sample\_sequence.fasta from Canvas.
7. To conduct a blast search, we need to make a local database.
  - a. Find a genome: Go to <https://ftp.ncbi.nih.gov/genomes/>.
  - b. Create a new directory for the database: ~/CSB/human\_genome/CDS
  - c. Download the human genome coding sequences:
 

```
wget "ftp://ftp.ensembl.org/pub/release-94/fasta/homo_sapiens/cds/Homo_sapiens.GRCh38.cds.all.fa.gz"
```
  - d. Decompress the file with **gunzip**.
  - e. Make a blast database:
 

```
makeblastdb -in Homo_sapiens.GRCh38.cds.all.fa.gz -dbtype nucl -parse_seqids
```
  - f. Search the blast database.
    - i. Make a new folder in the CSB folder: week02.
    - ii. Move sample\_sequence.fasta to the week02 folder.
    - iii. Run the blast command from the week02 folder.
 

```
tblastx -query sample_sequence.fasta -db ~/test/human_genome/CDS/Homo_sapiens.GRCh38.cds.all.fa
```
    - iv. It might be good to redirect the results to a text file. Rerun the command with > or with -out.
  - g. Use the **less** command to view the results. Press **q** to exit.

#### Part 6: A little more on navigating directories:

<b>cd ..</b>	Move up one directory
<b>cd /</b>	Move to the root directory
<b>cd ~</b>	Move to the home directory
<b>cd -</b>	Back
<b>pwd</b>	Print working directory
<b>ls</b>	List files in the current directory
<b>ls -a</b>	List all files
<b>ls -lh</b>	List files along with more information

#### Part 7: Basic commands

<b>cp</b>	Copy a file
<b>cp -r</b>	Copy a directory and all of its contents
<b>mv</b>	Move a file or rename a file
<b>rm</b>	Delete a file
<b>rm -r</b>	Delete a directory and all of its contents
<b>less</b>	View the contents of a file
<b>cat</b>	Print a file to the standard output
<b>wc</b>	Word count: Line, word and byte count
<b>sort</b>	Sort the lines of a file
<b>uniq</b>	Show unique lines of a file
<b>head</b>	Show the beginning of a file (use -n to specify how many lines)
<b>tail</b>	Show the end of a file (use -n as above)
<b>&gt;</b>	Redirect results to a text file
<b>&gt;&gt;</b>	Append the results to an existing text file
<b>echo</b>	Print text to the screen. Use this and > or >> to put text in a file.
<b> </b>	Pipe the results of one program into another program

#### Part 8: What are tab-delimited and comma-delimited files?

1. Download "Sample\_Data.csv".
2. Move it to the server using FileZilla.
3. Use head, tail and less to view the contents.
4. Use cut to obtain specific columns:
  - a. **cut -d "," -f 1,10,11 Sample\_Data.csv**
  - b. **-d** allows you to specify the delimiter.
  - c. **-f** specifies the columns to be cut.
  - d. **grep** can be used to get rows that match a search term. For instance, grep Alaska or grep female.
  - e. **man grep** for more info.
5. Windows versus Unix line breaks.
  - a. Use **file** to find out the type of line breaks.
  - b. Easily convert them using nano.