

Practice: Using Python on the Unix Command Line

BIOS 274: Introductory Python Programming for Genomics

Veronica Behrens

12/4/2019

WINDOWS USERS ONLY

1. Open MobaXterm

2. Install python3 for MobaXterm

`apt-get install python3`

This will take quite a while to finish.

3. Install nano for MobaXterm:

`/bin/apt-cyg install nano`

You can open a new tab within MobaXterm and run this command while python3 is installing.

4. Settings → Configuration → Terminal (it's a tab at the top) → Check the box near the bottom next to "Use Windows PATH"

5. Close MobaXterm

6. Re-open MobaXterm

WINDOWS USERS ONLY

7. Test whether python has been correctly installed

`python3 --version`

Python 3.6.8 should appear

8. Test whether nano has been correctly installed

`nano`

A different screen should pop up

To exit nano, do 'ctrl' + X

Basic Unix Commands

<code>pwd</code>	Print working directory (your current directory)
<code>cd PATH</code>	Change directory to specified PATH
<code>cd ..</code>	Move up one directory level
<code>ls</code>	List all the files/folders in the directory
<code>cp FILE DIRECTORY</code>	Copy FILE to DIRECTORY
<code>mv FILE DIRECTORY</code>	Move FILE to DIRECTORY
<code>rm FILE</code>	Delete/Remove FILE
<code>rm -r DIRECTORY</code>	Delete/Remove DIRECTORY
<code>grep 'STRING' FILE</code>	Find all instances of 'STRING' in FILE
<code>wc FILE</code>	Count the number of words in FILE
<code>wc -l FILE</code>	Count the number of lines in FILE
<code>nano</code>	Open a simple plaintext editor

REMEMBER:

Type `man COMMAND` (i.e. `man grep`) to see usage information!

In Terminal (Mac) or MobaXterm (Windows):

- 1. Change directory to wherever you downloaded your files for Day 7 from Canvas**
- 2. Print out the names of all the files in the directory**
- 3. From the terminal, view the contents of utils.py**
- 4. Enter an interactive version of python on the command line**
- 5. Use the functions in utils.py to find the complement and reverse complement of a DNA sequence**
- 6. Exit the interactive version of python**

Open `sys.argv_example.py` in Atom.

Follow the directions at the top of the script to run it.

Open `sumLinesOfFile.py` in Atom.

Follow the directions at the top of the script.

Open `shortStatureGenes.py` in Atom.

Follow the directions at the top of the script.

Can we find the diseases associated with a gene with basic Unix commands?

1. Find the diseases in `shortStatureDisorders.tsv` associated with `FLNB`.
Save them to a file called `FLNB.txt`

2. How many diseases are associated with `FLNB`?

Hint: Use a flag with either `grep` or `wc`

Do `man grep` or `man wc` to see usage information.

3. Can you do step 1 for both `FLNB` and `COL1A1` at the same time?

4. Display only the names of the disorders in `FLNB.txt` (without the column of genes).