

Unix terminal reference sheet

pwd	Print working directory. Show path of the current directory in which you are working \$ pwd users/davelunt/genetics/code
cd	Change directory. Move location of where you are working \$ cd blast \$ pwd users/davelunt/genetics/code/blast \$ cd .. # go up one level
ls	List. display the contents of the current directory \$ ls -l # list long, one item per line, all file details
cp	Copy. specify files to be copied and location to copy them
cat	Display contents of specified file to the screen, or file if directed (> myfile.txt)
head	Display the top of file specified, 10 lines is default but head -20 will display 20
tail	Display the end of file specified, 10 lines is default but tail -20 will display 20
grep	Search for pattern in location specified and report all lines containing search term \$ grep Drosophila mysequences.fasta -c will count rather than find lines containing the term
wc	Word count. Specify file to be counted \$ wc -l species.txt will count the number of lines (-l) in the file
sed	Search and replace within a file \$ sed s/transcription/translation/g ribosome_essay.txt
sort	sort the lines in a file \$ sort species.txt > sorted_species.txt
uniq	Remove duplicated lines from a sorted file (best to write > to a new data file) \$ uniq sorted_species.txt > unique_species.txt -c counts unique lines and reports number
echo	Write. \$ echo My search sequence is ATGGGTAG >> labbook.txt
 	Pipe. chains the output of one program as input for the next \$ sort mydata.fas uniq -c sort -rn head -3 >top3-species.txt
;	Semicolon. Allow multiple commands on one line \$ pwd; ls
<	From. source of data
>	To. write data to this location (overwrite if it exists)
>>	Append. Add data to this location (do not overwrite)
*	Asterisk wildcard. Represents any number of any characters *.txt represents all files ending .txt