Searching For the Things NOTES

<http://swcarpentry.github.io/shell-novice/07-find/>

* grep to select lines from text files that match simple patterns.
* Find to find files/ directories whose names match simple pattern

GREP

* grep’ is a contraction of ‘global/regular expression/print’
* grep finds and prints lines in files that match a pattern.
* Syntax. #grep pattern filename
* Option -w (used to limit the word #grep -w the mytextfile)
* grep **-w** "is not" haiku.txt. 🡪 to get the particular phrase
* grep **-w** The haiku.txt 🡪 to get the word
* grep **-n** "it" haiku.txt. 🡪 numbers lines which matches
* grep **-n** **-w** **-i** "the" haiku.txt. 🡪 Case sensitive (use -i)
* grep **-n** **-w** **-v** "the" haiku.txt 🡪 -v used for: search which **DO NOT** contain “the” that is called “invert search”

Quiz Answers

Ans

1. grep "[A+C+G+T+]" tb1.fasta | grep --color -v "^[ACGT]\*$"

A screenshot of a cell phone

Description automatically generated

2. This command prints out all the blank space in the file, or it prints nothing if there are no blank spaces (like in the example tb1.fasta) because there are no characters/pattern between ^ and $ (that means it will find blank lines/ spaces). Suppose in a text file there is code of 10 lines written and between these 10 lines, 3 lines left blank. This command will print 3 blank lines as an output.

3. grep '^\s\*$' filename. Or same as que2

4. grep -c "| GTR+G" partitionfinder\_bestscheme.txt

A screenshot of a cell phone

Description automatically generated

project management & markdown format Notes

Reg Expressions Useful

cd data

touch seqs/zmays**{**A,B,C**}**\_R**{**1,2**}**.fastq

ls seqs/

ls seqs/zmaysB**\***

ls seqs/zmays[AB]\_R1.fastq

ls seqs/zmays[A-C]\_R1.fastq

ls seqs/zmaysA\_R**{**1..2**}**.fastq

ls -lR