BIOinformatics Working Group

Taruna Schuelke
UC Santa Barbara

3 April 2019

Objectives

1. Next-Generation technologies
2. Linux terminology
3. SSH into Pod —> Command Line basics
4. Best practices
5. Github

Quick Quiz

1. _____ command is used to remove a directory.

- **a. rdir**
- **b. remove**
- **c. rd**
- **d. rmdir**

2. _____ command is used to list contents of a directory.

- **a. tar**
- **b. ls**
- **c. dir**
- **d. lp**

3. _____ command is used to grab specific strings in a file.

- **a. grep**
- **b. grab**
- **c. less**
- **d. head**

4. _____ command is used to see columns in a file.

- **a. awk**
- **b. cut**
- **c. sort**
- **d. a and b**

5. `#!/bin/sh` is called what?

Quick Quiz

1. _____ command is used to remove a directory.
   a. rdir
   b. remove
   c. rd
   d. rmdir

2. _____ command is used to list contents of a directory.
   a. tar
   b. ls
   c. dir
   d. lp

3. _____ command is used to grab specific strings in a file.
   a. grep
   b. grab
   c. less
   d. head

4. _____ command is used to see columns in a file.
   a. awk
   b. cut
   c. sort
   d. a and b

5. `#!/bin/sh` is called what?

sha-bang, hashbang, pound-bang

NGS technologies

PACBIO®

illumina®

- **Single Molecule Real Time (SMRT) technology**
- **35,000 - 70,000 good reads cell^{-1}**
- **~11% error rate (as of 2013)**
- **Read length >45 kb and >190 kb**
- **More info: https://www.pacb.com/smrt-science/smrt-sequencing/**

- **Most common**
- **millions - billions of reads lane^{-1}**
- **Low error rate**
- **Average read length ~300 bp**
- **More info: https://www.illumina.com/techniques.html**

- **COOLEST sequencer ever!**
- **Long reads (e.g., 1 mbp)**
- **High error rate (~5-15%)**
- **More info: https://nanoporetech.com/how-it-works**

Key Linux Terms

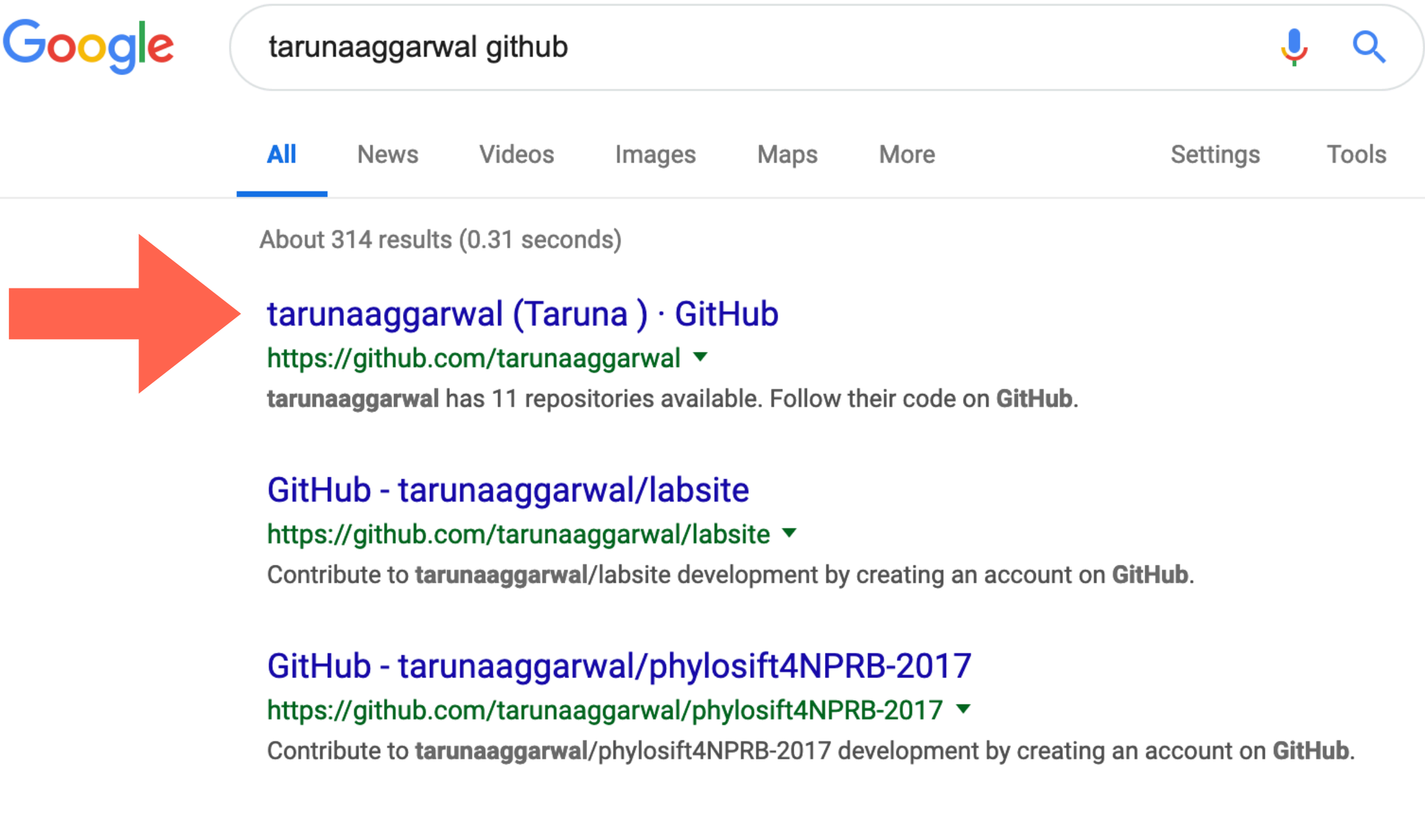
- **Linux - a type of operating system**
- **Terminal - a place where you type input and receive output**
- **Shell - the program that interprets what you type in the terminal**
- **Bash - a type of shell**
- **Directory - another name for a folder**
- **Root - *the* top level directory (denoted by /)**
- **Root - the person with admin privileges**
- **Permissions - describe who can do what to files and directories**

SSH into Pod

SSH key

https://github.com/tarunaaggarwal/bwg-s19

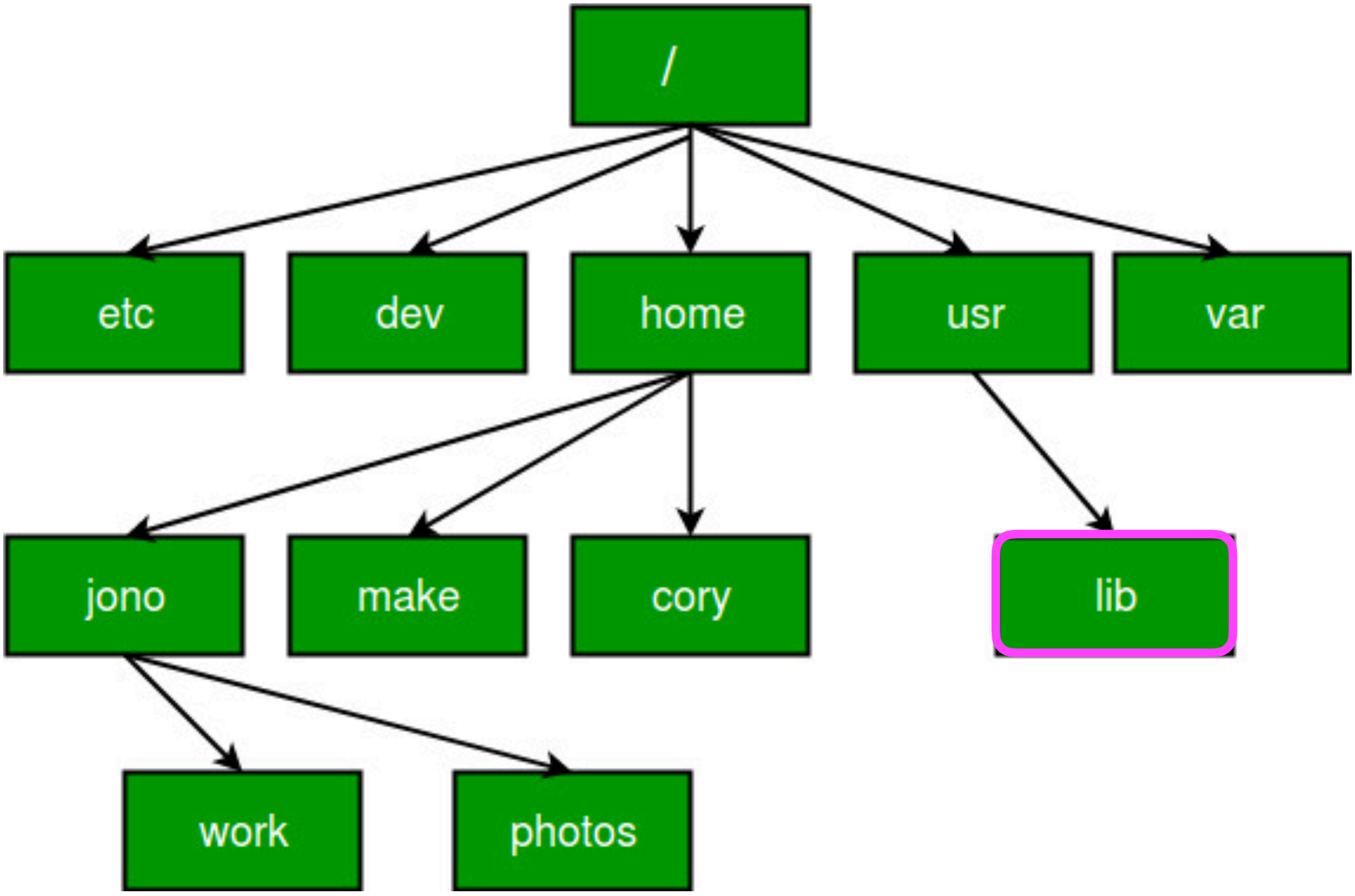
or



Command Line Basics

Command	Purpose
`$ ls`	shows a listing of dir contents
`$ pwd`	print working dir
`$ cd`	to navigate between dirs
`$ ..`	one dir above the current dir
`$ .`	current dir
`$ ~`	home dir
`$ clear`	to clear your screen
`$ rm`	to remove a file
`$ rmdir`	to remove a dir
`$ mkdir`	to create a dir

PATH



Best Practices

Post your code on Github

Use a text editor to organize commands

Use a consistent directory structure

Case MaTTerS

Avoid spaces in file and dir names

Best Practices

Post your code on Github

Use a text editor to organize commands

Use a consistent directory structure

Case MaTTerS

Backup your work! Esp. scripts!!!

Avoid spaces in file and dir names

Review

1. Next-Generation technologies
2. Linux terminology (Linux is an OS)
3. SSH into Pod —> Command Line basics
4. Best practices
5. Github

Next Time

1. Next-Generation technologies
2. Linux terminology (Linux is an OS)
3. SSH into Pod —> Command Line basics
4. Best practices
5. Github
6. **Files type and amplicon sequencing & analysis**