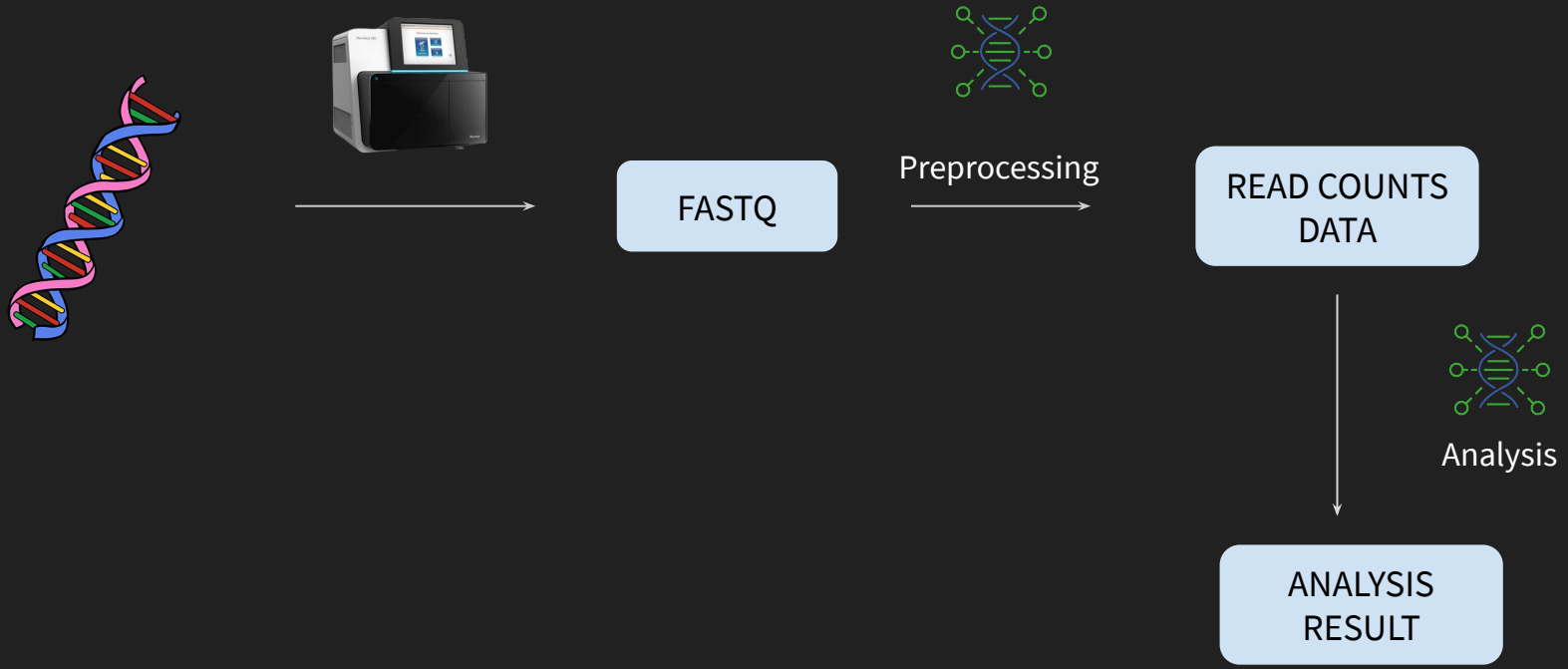
Introduction to Bioinformatics

Introduction to UNIX and Command Lines

Haries Ramdhani. 2021.

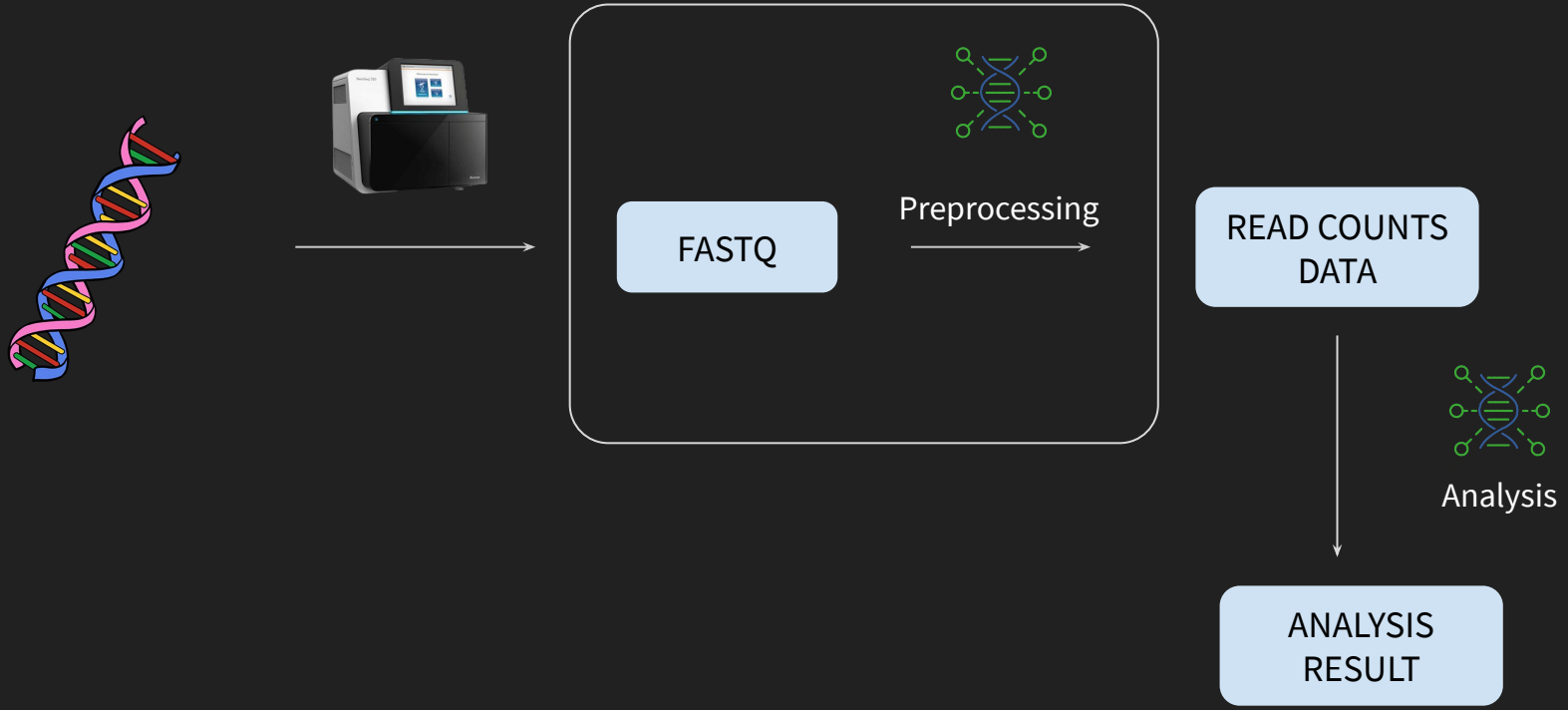
A Quick Glimpse

NGS (Bisulfite Sequencing)



A Quick Glimpse

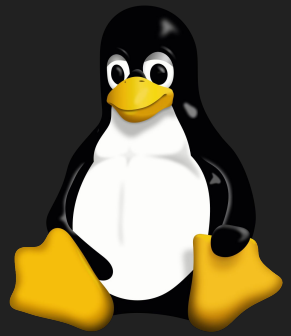
NGS (Bisulfite Sequencing)



Why do we need to learn UNIX?

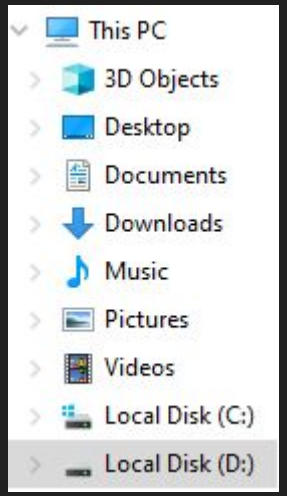
- Most of the software in the preprocessing part for NGS analysis is only available and can be run in UNIX system only
- Most high performance computing is using UNIX

What is UNIX?

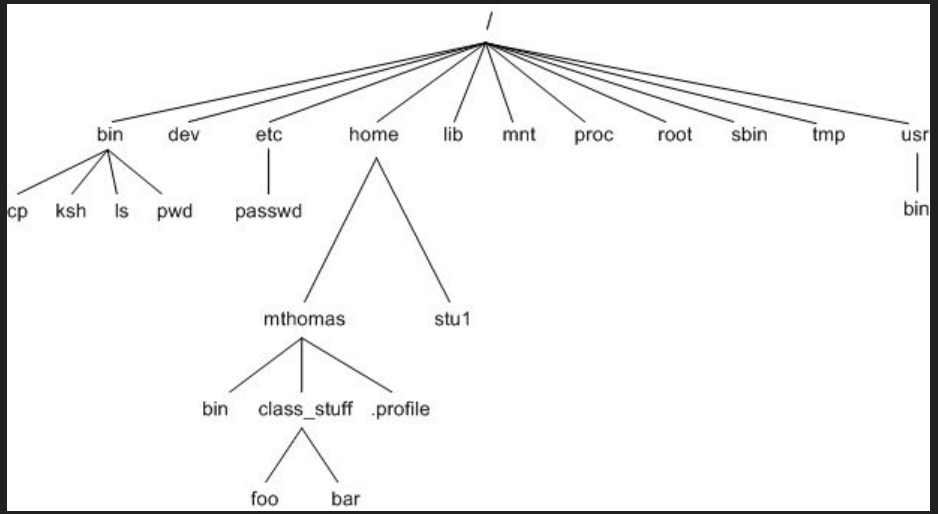


- UNIX is a family of operating system
- Majority of the time is operated using command line
- Mac and Linux are both UNIX-Like, Windows is not

Structure Difference

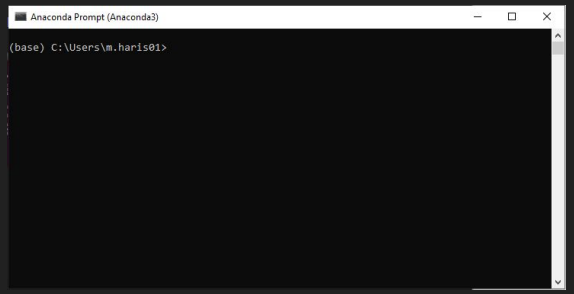


Windows

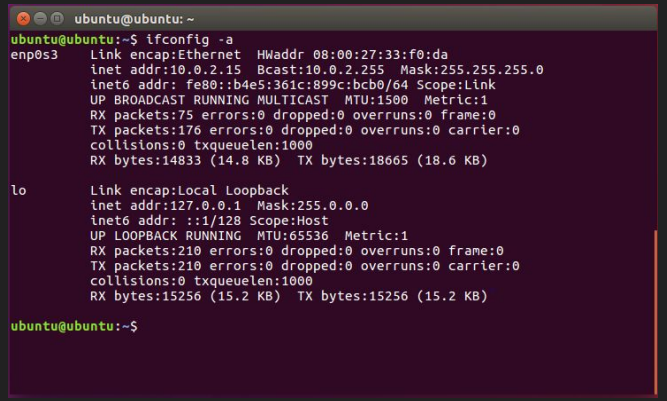


UNIX

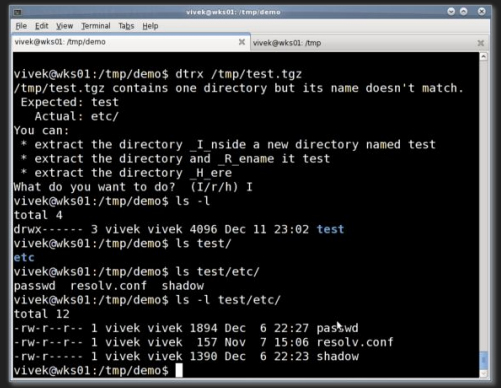
Command Lines



Anaconda prompt

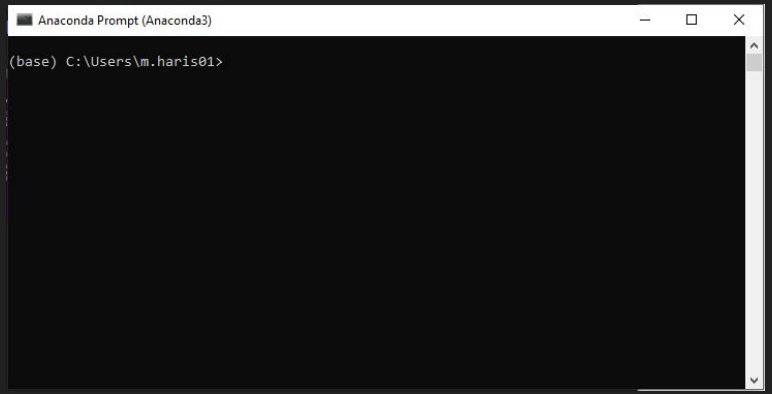


Ubuntu terminal



UNIX terminal

I’m using Windows, can I still get command line?



- Cygwin, http://www.cygwin.com/
- Git Bash, https://git-for-windows.github.io/
- Boot from a CD or USB
- Dual booting
- WSL https://docs.microsoft.com/en-us/windows/wsl/install-win10
- Connect to a cloud system

File System command

You can type different command in the command line

```
/Users/me $ ls

/Users/me $ date

/Users/me $ echo "hello there"

/Users/me $ python test.py

/Users/me $ python cmdline.py -i input -o out -d db
```

Orienting the File System

Changing directory:
$ **cd /** - go to the root directory
$ **cd /usr/home/jack/bin** - go to the user's sub-directory
$ **cd ..** - go to the upper level directory
$ **cd, or cd ~** - go to the user's home directory
$ **cd --** - go to the last visited directory

More:
$ **pwd** - print working directory
$ **ls** - list all files in the current directory
$ **ls -1** - list files in one column

Understanding the program more

$ man ls
NAME
ls -- list directory contents

SYNOPSIS
ls [-ABCFGHLOPRSTUW@abcdefghiklmnopqrstuwx1] [file ...]

DESCRIPTION
For each operand that names a file of a type other than directory, ls displays its name as well as any requested, associated information.

Direct Help for a Program

- Some programs allow you to see the information about the program directly

$ sort --help
sort --help
Usage: sort [OPTION]... [FILE]...
or: sort [OPTION]... --files0-from=F
Write sorted concatenation of all FILE(s) to standard output.

Mandatory arguments to long options are mandatory for short options too.
Ordering options:

Creating and Deleting Directories

Create Directory:
$ mkdir MyDir
$ mkdir MyDir
$ mkdir MyDir/UnderDir/DeepDir

Remove Directory:
$ rmdir MyDir
$ rmdir MyDir/UnderDir/DeepDir

Files

1. Make an empty file with
   $ **touch file**
2. Running programs can generate files
3. Direct output of a program to create a file.
   $ **echo "hello world" > hello.txt**
   $ e**cho "goodbye cruel world" > goodbye.txt**

Permissions

- ls -l will show a long version of listing output

```
lrwxrwxrwx 1 jstajich gen220     25 Sep 25 13:53 Nc20H.expr.tab
```

- There is info on the file/folder and 3 sets of permissions listed there
    - 'd' -> It is a directory, 'l' -> it is a link, empty, it is a file
    - User -> rwx or rw- permissions for the user
    - Group
    - All
- what are permissions
    - r - is it readable
    - w - is it writeable
    - x - is it executable
- Directories have to be executable to be able to be used/entered. Programs/applications need to be executable to be able to run.

Deleting Stuff

- **rm** - for removing files
- **rm -r** -- be careful removes recursively
- **rm -f** - removes without prompting you
- **rm -rf** - removes recurisvely without prompty Use with care

File Content

Want to read the content of a text file you can use these commands

- **cat** - will spit out the whole file on the screen
- **more** - a paginator - will display one page at a time (based on your screen)
- **less** - similar to more, but has additional options
- **head** - preview only the first few lines
- **tail** - preview only the last few lines

Downloading Data from the WEB

Cmdline tools curl and wget can download data from the web/ftp

```
$ wget  http://www-personal.umich.edu/~jlawler/wordlist
$ curl -O http://www-personal.umich.edu/~jlawler/wordlist
$ curl -o ucr_index.html https://www.ucr.edu # OR
$ curl  https://www.ucr.edu  > ucr_index_again.html
```

Installing Programs

You can install program in the command line using **apt**
$ sudo apt-get install python3

Sometimes not all of your programs are available in apt-get, so you have to install it from somewhere else, for example git
$ git clone https://github.com/FelixKrueger/Bismark.git

Some installation may require you to follow some steps