BIOL425 Comp Mol Bio

Part 1. Git & Unix

Professor Weigang Qiu
Hunter College of CUNY
Spring 2023

Version Control with Git (Chapter 2)

- I will use it to share files (e.g., slides), as an alternative to Blackboard
- Later (hopefully), students will be able to upload files

1. Download course repository

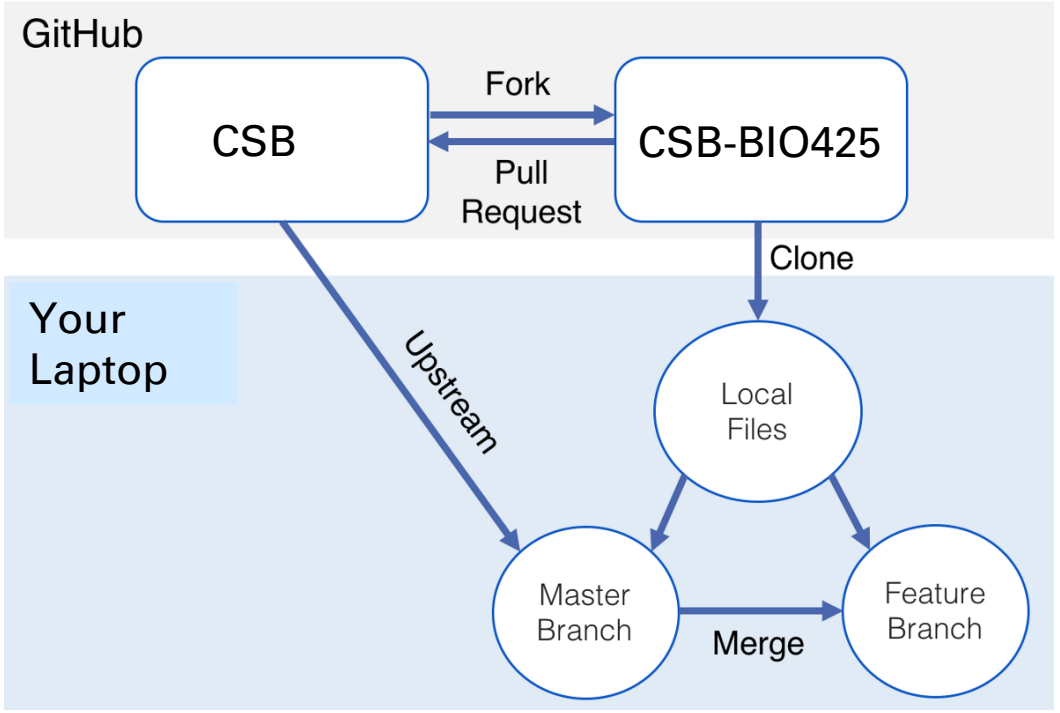
```
git clone https://github.com/weigangq/CSB-BIOL425.git
```

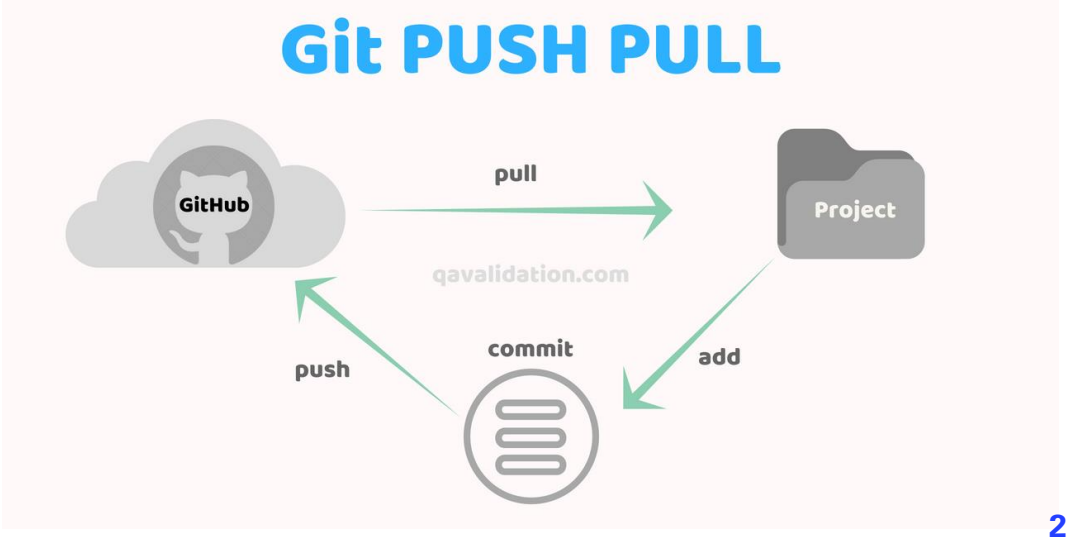
2. Pull the latest versions

```
git pull
```

3. Upload/Update file (ignore; not yet working; we will work on it later)

```
git add <filename> # add a new file
git commit –m “message”
git push
```





UNIX Basics (Chapter 1)

1. **Directory listing & shorthand**

```
ls -lrth      # long, reverse, timestamp, human-readable
pwd           # present working directory
~             # home directory
.             # current directory
..            # parent directory
```

2. **Keyboard shorthand (for command editing)**

```
Ctrl-a      # go to the beginning
Ctrl-e      # go to the end
Ctrl-l      # clear the screen
Ctrl-u      # clear text before the cursor
Ctrl-k      # clear text after the cursor
Ctrl-c      # kill the (stalled) command
```

3. **Directory navigation**

```
cd ~/CSB-BIOL425/python/data      # absolute path
Cd ../../unix/data                # relative path
cd -                              # toggle 2 directories
```

4. **Use auto-completion**

```
Tab                          # NEEVER type out a full filename
ls bad\ file\ name.txt       # NEEVER use spaces in filenames
```

Exercise 1

Intermezzo 1.1

(a) Go to your home directory.
(b) Navigate to the `sandbox` directory within the `CSB/unix` directory.
(c) Use a relative path to go to the `data` directory within the `python` directory.
(d) Use an absolute path to go to the `sandbox` directory within `python`.
(e) Return to the `data` directory within the `python` directory.

UNIX Basics (Cont'd)

1. Copy files and directories

```
cp ~/CSB-BIOL425/unix/data/Buzzard2015_about.txt ~/CSB-BIOL425/unix/sandbox/      # use absolute path
cd ~/CSB-BIOL425/unix/sandbox/
cp ../data/Buzzard2015_about.txt .                                                # use relative path
cp ../data/Buzzard2015_about.txt ./Buzzard2015_about2.txt                         # copy & rename
cp -r ../data .                                                                   # recursive copy
```

2. Move or rename a file

```
mv Buzzard2015_about2.txt ../data/                        # move file to a different directory
mv ../data/Buzzard2015_about2.txt ../data/Buzzard2015_about_new.txt      # move and rename
```

3. Remove file or directory

```
touch new_file.txt          # create an empty file (and update timestamp)
rm -i new_file.txt          # Proceed with caution; EXTREMELY destructive
mkdir -p d1/d2/d3           # make nested directories
rm -r d1                    # recursively remove a directory (and its sub-directories)
```

4. View & filter text files

```
cd  ~/CSB-BIOL425/unix/data
less Marra2014_data.fasta # spacebar to page down; b to page up; Q to quit
cat *.txt                 # concatenate all ".txt" files
wc *.txt                  # word count (all ".txt" files)
head Gesquiere2011_data.csv       # show top lines
tail -n 2 Gesquiere2011_data.csv  # show tail two lines
sort Gesquiere2011_data.csv       # sort lines in a file (alphabetically)
sort -n Gesquiere2011_data.csv    # sort lines numerically
```

Exercise 2

Intermezzo 1.2

To familiarize yourself with these basic Unix commands, try the following:

(a) Go to the `data` directory within `CSB/unix`.
(b) How many lines are in file `Marra2014_data.fasta`?
(c) Create the empty file `toremove.txt` in the `CSB/unix/sandbox` directory without leaving the current directory.
(d) List the contents of the directory `unix/sandbox`.
(e) Remove the file `toremove.txt`.