

\$ command line computing crash course

Part 2: understanding and avoiding common pitfalls

Mark Stenglein, PhD

Center for Vector-Borne Infectious Diseases

Department of Microbiology, Immunology, and Pathology

College of Veterinary Medicine and Biomedical Sciences

Colorado State University

Mark.Stenglein@ColoState.edu

StengleinLab.Org

This is what it can be like starting out on the command line

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ./blast my_sequences.fasta  
-bash: ./blast: Permission denied
```

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ls my_sequences.fasta  
ls: my_sequences.fasta: No such file or directory
```

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blastn -database nt -query my_sequences.fasta  
Error: (CArgeException::eInvalidArg) Unknown argument: "database"
```

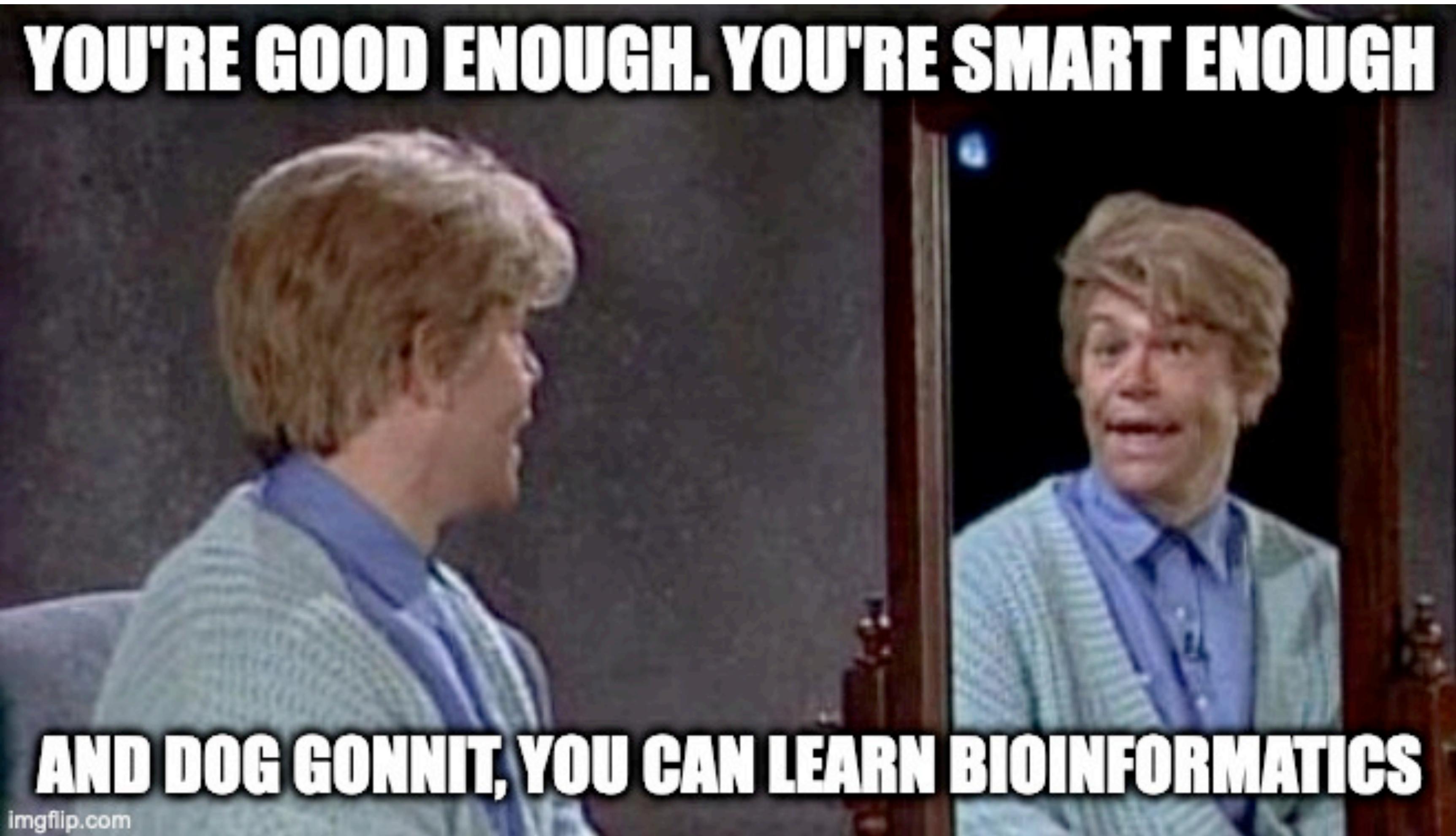
```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ OMG I hate this - I quit bioinformatics!  
-bash: OMG: command not found
```

This is what it feels like to experience those errors



Your feelings of frustration are valid

but remember...



Common command-line pitfalls

- You are not where you think you are
- The files you want to work with are not where you think they are
- You typed the wrong command name or wrong file name
- The command you are trying to run is not installed
- The command you are trying to run is installed but not in your PATH
- Permissions issues
- Problem with command arguments (missing required argument, etc)

What might be causing this “No such file or directory” error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ls my_sequences.fasta  
ls: my_sequences.fasta: No such file or directory
```

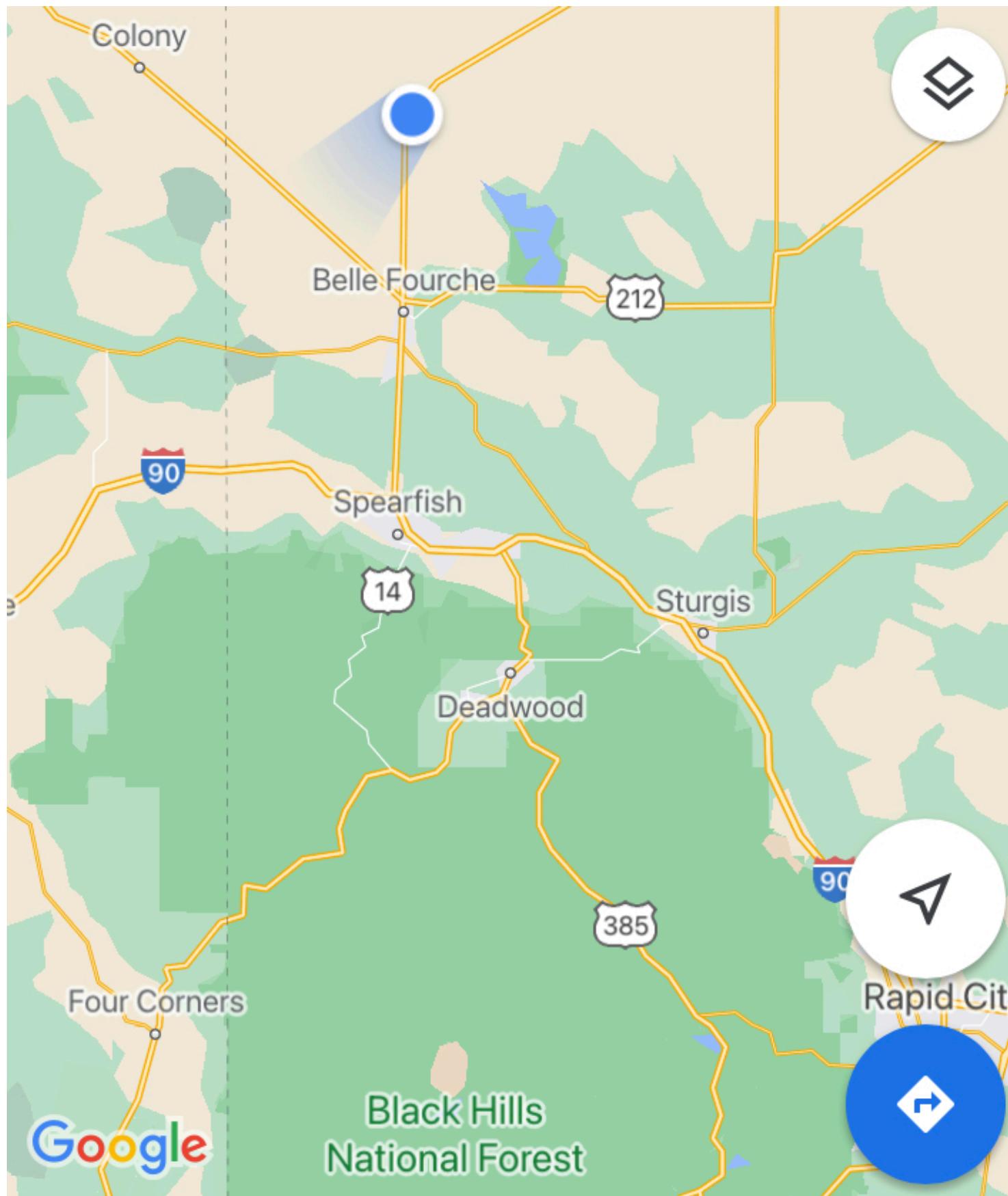
What might be causing this “No such file or directory” error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ls my_sequences.fasta  
ls: my_sequences.fasta: No such file or directory
```

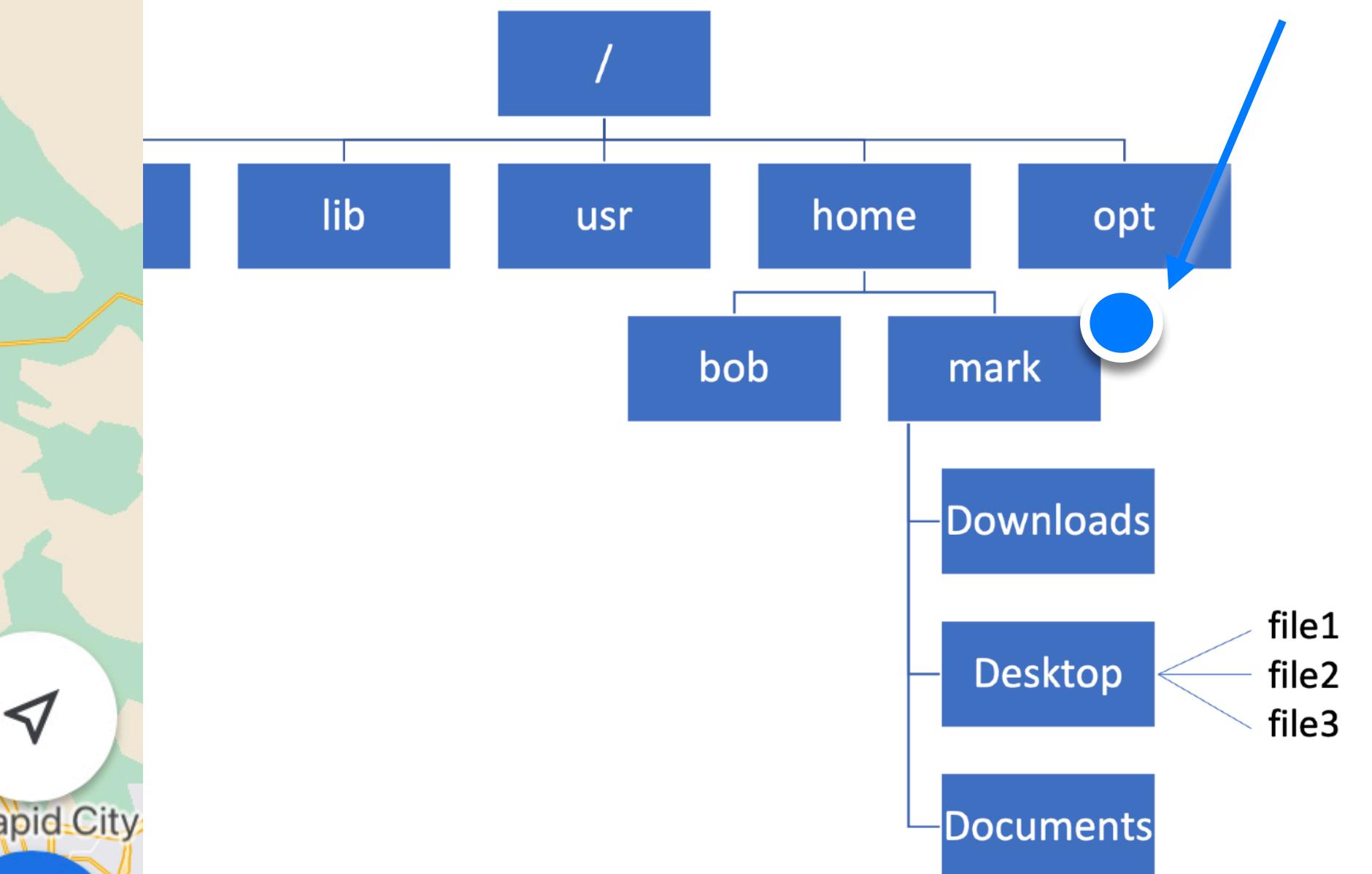
Common causes:

- You are not where you think you are
- The files you want to work with are not where you think they are
- You typed the wrong command name or wrong file name

Remember that you are always somewhere when working on the command line



Your pwd is like the blue dot on google maps



Your pwd is so important to keep track of that it's one of the few pieces of information in the CLI prompt

What does ~ mean?

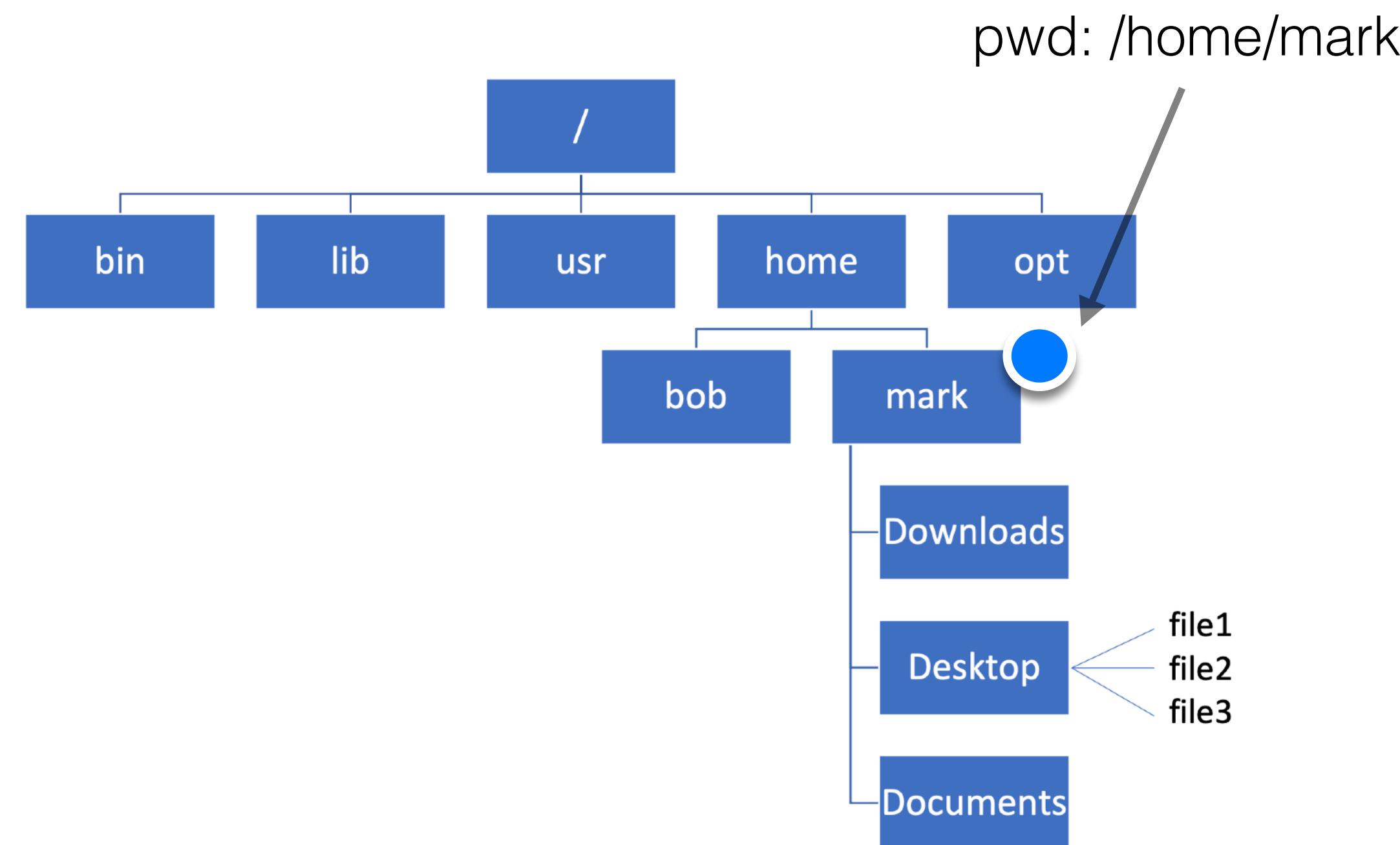


```
(miniconda3)mdstengl@cctsi-104:~$
```



```
(miniconda3)mdstengl@stngleinlab101:/home/shared_data$
```

It pays to keep track of your pwd and the location of files you are working with



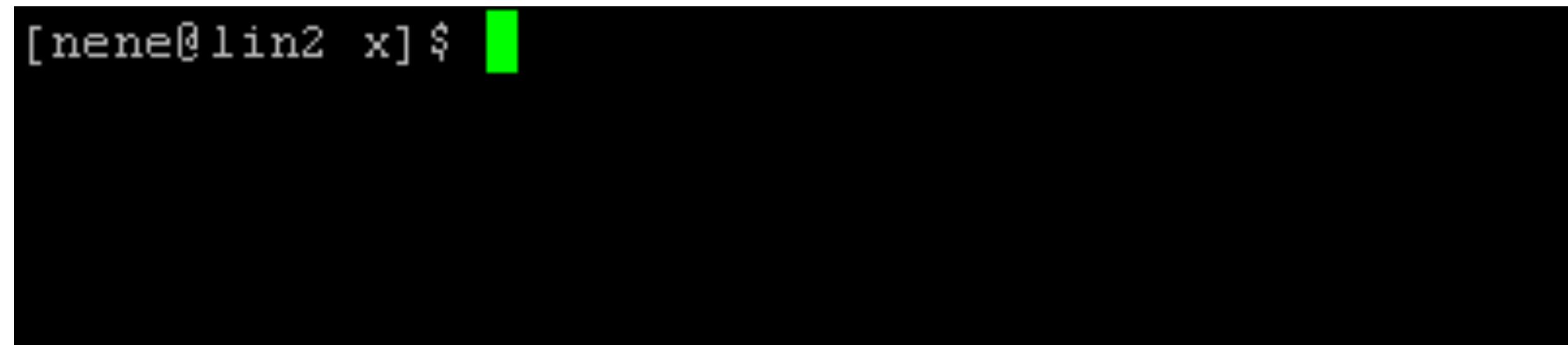
What command can you run to determine your pwd?

Given a pwd of /home/mark what would be the output of this command?

`ls -l file1`

What are different ways you could solve this problem?

Tab completion is an important tool for avoiding misspelled command and file names



```
[nene@lin2 ~] $
```

By the end of the workshop, I hope to see you all using tab completion!

What might be causing this command not found error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

What might be causing this command not found error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

Likely causes:

- blast is not installed on this computer
- blast is installed somewhere on your computer but is not in your PATH
- blast is not the name of the command

How can you distinguish between these possible causes? And correct them?

The which command tells you if a command is available to be run
and if so, where the command is located

which ls

which which

which blast

This is a very quick way to tell if a command is available

It can also be useful if there are multiple versions of some software installed

What do these two commands do?

man which
which man



What might be causing this command not found error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

Likely causes / [possible solutions](#)

- blast is not installed on this computer
 - [use which to see if a command is available](#)
 - [Install if necessary.](#)
- blast is installed somewhere on your computer but is not in your PATH
- blast is not the name of the command

Linux paths and the PATH

A path

- The location of a file or directory
- Can be absolute or relative

The PATH

- A list of directories the shell looks in to find commands
- The `which` command tells you if a command is in your PATH
- PATH is a so-called variable. You can add or remove directories to the PATH by changing its value.
- This is why people sometimes run commands by specifying the location of the command (e.g.: `./blast`) : this tells the shell exactly where the command is and the command need not be in your PATH

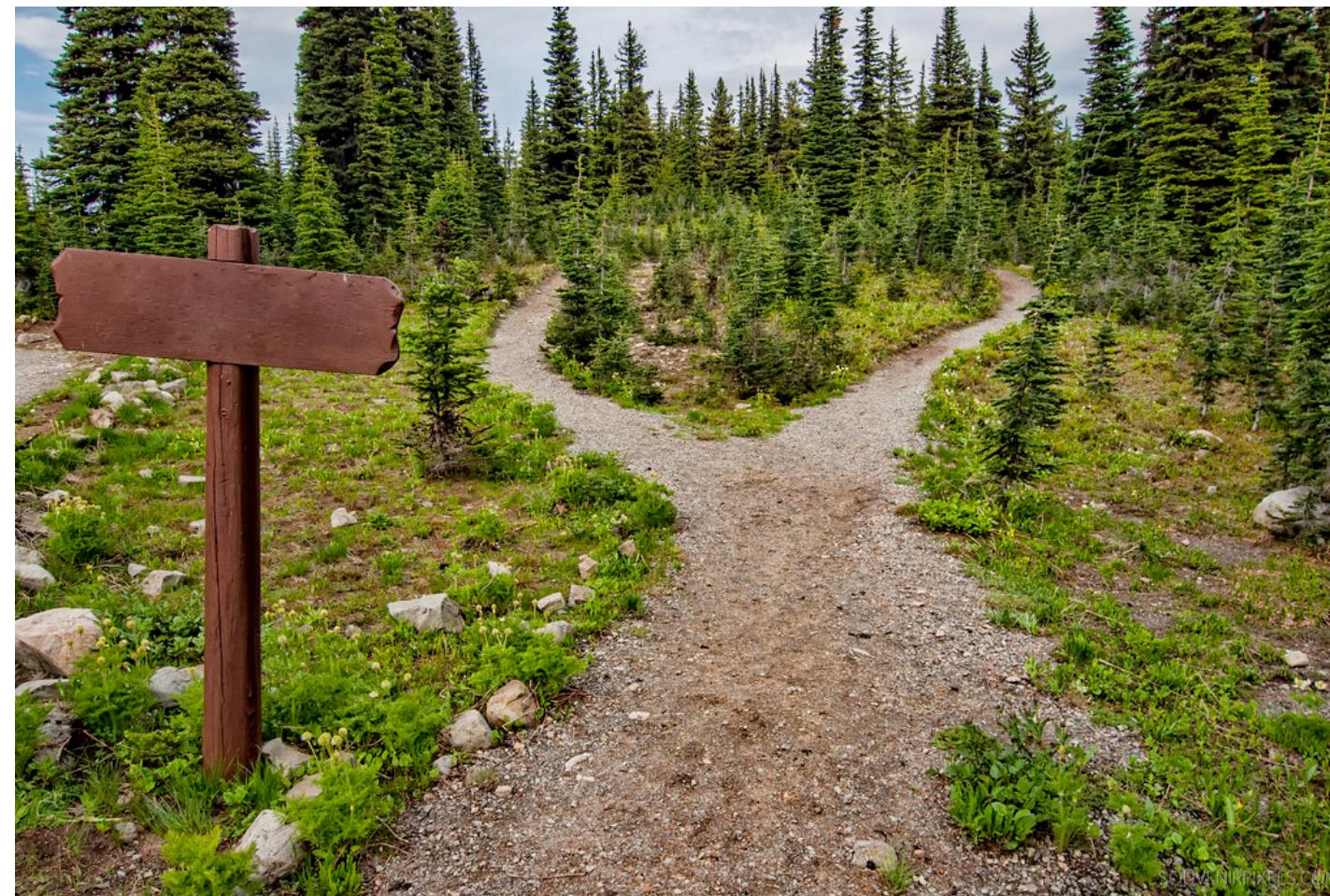


Image credit: James Wheeler (CC BY-NC-SA 2.0)

What is the PATH on your laptops?

Run this command to see the value of the PATH variable in your laptop's CLI:

```
echo $PATH
```

In the shell: \$PATH is the value of the variable named PATH

To see all of the built-in variables in your shell, run the command env

Sometimes software installation instructions tell you to modify your PATH

Add the MMseqs2 binary directory to your `PATH` environment variable by running the following command:

```
export PATH=$(pwd)/mmseqs/bin/:$PATH
```

<https://github.com/soedinglab/mmseqs2/wiki#install-mmseqs2-for-linux>

A good description of how to modify your PATH:

<https://unix.stackexchange.com/questions/26047/how-to-correctly-add-a-path-to-path>

What might be causing this command not found error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

Likely causes / solutions

- blast is not installed on this computer
 - [use which to see if a command is available.](#)
 - [Install if necessary.](#)
- blast is installed somewhere on your computer but is not in your PATH
 - [add the necessary directory to your PATH](#)
 - [specify the full path to the command](#)
- **blast is not the name of the command (??)**

What might be causing this command not found error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blast my_sequences.fasta  
-bash: blast: command not found
```

Likely causes / solutions

- blast is not installed on this computer
 - [use which to see if a command is available.](#)
 - [Install if necessary.](#)
- blast is installed somewhere on your computer but is not in your PATH
 - [add the necessary directory to your PATH](#)
 - [specify the full path to the command](#)
- blast is not the name of the command
 - [Tab completion can help](#)
 - [RTFM](#)

RTFM originated as a snarky internet response but is a legitimately good strategy



BLAST® Command Line Applications User Manual [Internet].

< Prev Next >

BLAST Help
NCBI Help Manual
 National Center for Biotechnology Information
U.S. National Library of Medicine

► Show details
Contents ▾

Search this book

Quick start

Created: June 23, 2008; Last Update: January 7, 2021.

Estimated reading time: 2 minutes

A BLAST search against a database requires at least a `-query` and `-db` option. The command:

```
blastn -db nt -query nt.fsa -out results.out
```

will run a search of `nt.fsa` (a nucleotide sequence in FASTA format) against the `nt` database, printing results to the file `results.out`. If “`-out results.out`” had been left off, the results would have been printed to `stdout` (i.e., the screen). The `blastn` application searches a nucleotide query against a nucleotide database.

<https://www.ncbi.nlm.nih.gov/books/NBK569856/>

What might be causing this (CArgException::eInvalidArg) error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blastn -database nt -query my_sequences.fasta  
Error: (CArgException::eInvalidArg) Unknown argument: "database"
```

What might be causing this (CArgException::eInvalidArg) error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blastn -database nt -query my_sequences.fasta  
Error: (CArgException::eInvalidArg) Unknown argument: "database"
```

Likely causes:

- -database is not a valid option for the blastn command

What might be causing this (CArgException::eInvalidArg) error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ blastn -database nt -query my_sequences.fasta  
Error: (CArgException::eInvalidArg) Unknown argument: "database"
```

Likely causes:

- `-database` is not a valid option for the `blastn` command

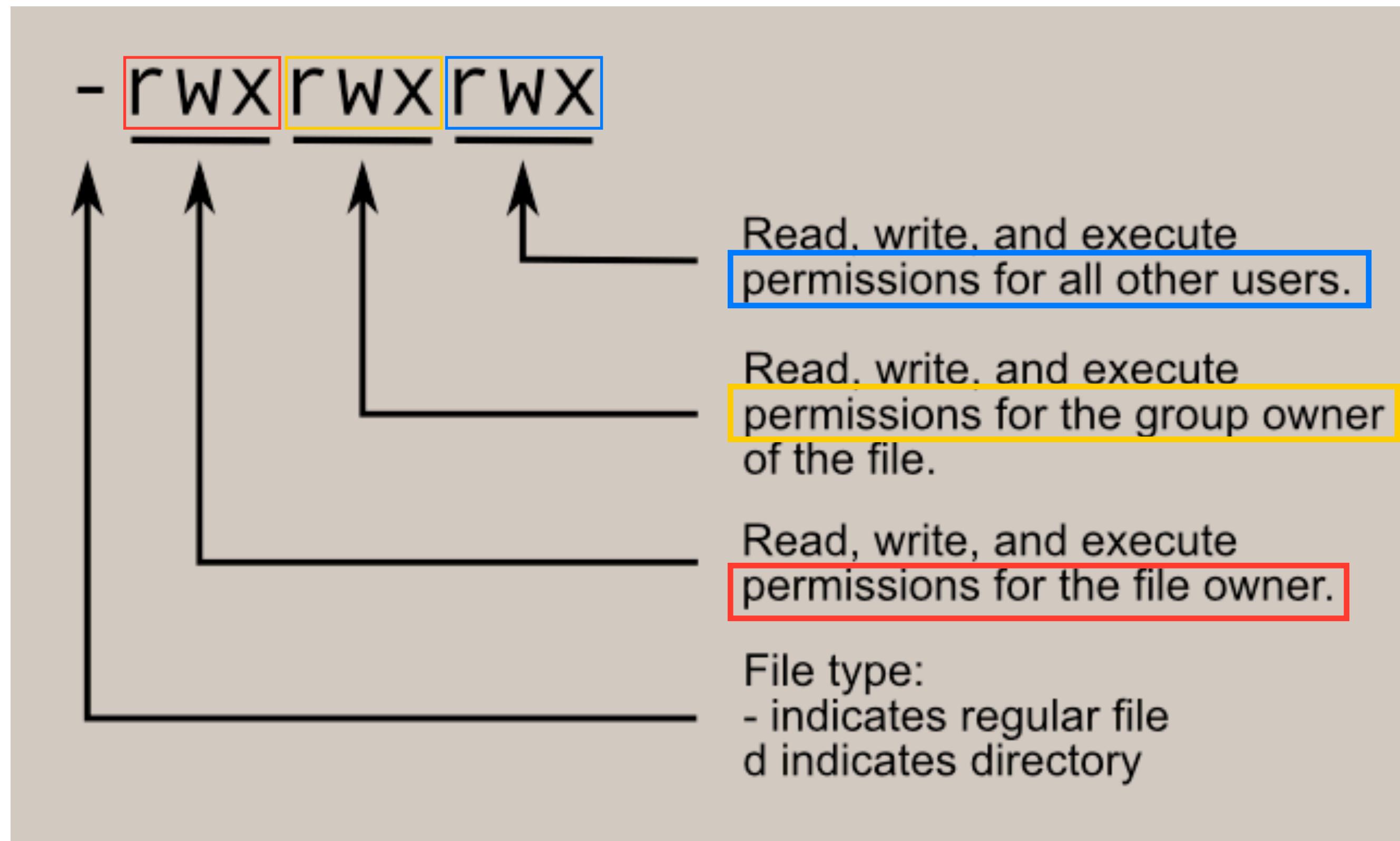
Solutions?

- Try running commands with:
 - No options
 - `-h` , `-help`, `--h`, `--help`
- RTFM!
- Try running `man blastn` (but `man` is used more for built-in linux commands)

What might be causing this error?

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ./blast my_sequences.fasta  
-bash: ./blast: Permission denied
```

File permissions in Linux define what can be done with a file and who can do it



Permission types:

Read: required to view or copy a file

Write: required to modify or delete a file

Executable: required to run a command or enter a directory

Permissions can be different for:

- **Users**
- **Groups** of users
- **All** users

Commands, including shell scripts, need to have executable permission

```
(base) mdstengl-MBP:2023_GDW_Workshop mdstengl$ ./blast my_sequences.fasta  
-bash: ./blast: Permission denied
```

If you are trying to run a command and get a Permission denied error - it likely has to do with executable permissions

In this example, there is a file named blast, but it lacks executable permissions so can't be run

Any plain text file with bash commands and executable permissions can be a “shell script”, which can be run as a command

On servers, it is common to not have permission to access other user's home directories

```
(base) mdstengl@aidlngs01:~$ ls /home/mollydvm/  
ls: cannot open directory '/home/mollydvm/': Permission denied  
(base) mdstengl@aidlngs01:~$ cd /home/mollydvm  
-bash: cd: /home/mollydvm: Permission denied
```

```
(base) mdstengl@aidlngs01:~$ ls -l /home/  
drwx----- 9 mollydvm mollydvm 4096 Mar 30 2020 mollydvm
```

Other users don't have permission to access mollydvm's home directory