

MEDICAL MICROBIOLOGY AND INFECTIOUS DISEASES BIOINFORMATICS WORKSHOP

Presents

Introduction to conda and Tool installation

INSTRUCTED BY
Grace E. Seo, MSc Student

INFORMATION FOR PARTICIPANTS

**All workshops are being recorded and posted to the
[MMID Bioinformatics Workshop – YouTube](#)**

**For live Q&A, go to [slido.com](#) and use participant
code [#9406658](#)**

2023 MMID Bioinformatics Workshop Schedule

DATE	INSTRUCTOR	TOPIC
March 2	Grace E. Seo	Introduction to the 2023 MMID Bioinformatics Workshop
March 9	Grace E. Seo	Introduction to conda and tool installation
March 16	Grace E. Seo	Introduction to genomics and viral data analysis
March 23	Jill Rumore	Bacterial Genomics
March 30	Jill Rumore	Reference Databases
April 6	Taylor Davedow	Beginner's Guide to Phylogenetic Trees
April 13	Taylor Davedow	Introduction to tree visualization and annotation using ggtree
April 20	-	Bfx workshop: Bring your own dataset!
April 27	-	Bfx workshop: Bring your own dataset!

April 20 and April 27 in-person sessions are open to the public (up to 100 people)!

Work with your colleagues/friends to analyze data together.

SET UP WI-FI (IN-PERSON PARTICIPANTS)

1. **Connect to UofM-secure (if you are a student or staff)**
- Use your @myumanitoba.ca or @umanitoba.ca login and password
2. **Connect to UofM-guest**

To access uofm-guest Wi-Fi:

1. Ensure your wireless card is active and connected to the **uofm-guest** network.
2. Open your web browser (e.g. Google Chrome, Microsoft Edge, Firefox, etc.) and browse to any website. This should redirect you to the **Acceptable Use Agreement** page.
3. Review the Acceptable Use Agreement for the unsecured wireless.
4. Select **I Agree**.

LEARNING OBJECTIVES

1. [**Describe method to install Ubuntu using PowerShell \(Windows user\)**](#)
2. [**Describe the purpose of GitHub and provide usage overview.**](#)
3. [**Provide basic overview of BASH commands.**](#)
4. [**Practice navigating through the terminal using commands.**](#)
5. [**Describe terminal permission structure for files.**](#)
6. [**Describe the purpose of conda.**](#)
7. [**Install conda environment and tools.**](#)

LEARNING OBJECTIVES

- 1. Describe method to install Ubuntu using PowerShell (Windows user)**
- 2. Describe the purpose of GitHub and provide usage overview.*
- 3. Provide basic overview of BASH commands.*
- 4. Practice navigating through the terminal using commands.*
- 5. Describe terminal permission structure for files.*
- 6. Describe the purpose of conda.*
- 7. Install conda environment and tools.*

Windows WSL PowerShell

If you have enabled WSL but cannot find Ubuntu on Microsoft Store, use command line to install.

1. Open PowerShell as administrator

2. Type the following command

> **wsl --install -d ubuntu**

```
seog@SMARTY: ~  
Welcome to Ubuntu 22.04.1 LTS (GNU/Linux 5.10.16.3-microsoft-standard-WSL2 x86_64)  
  
* Documentation:  https://help.ubuntu.com  
* Management:    https://landscape.canonical.com  
* Support:        https://ubuntu.com/advantage  
  
This message is shown once a day. To disable it please create the  
/home/seog/.hushlogin file.  
seog@SMARTY:~$
```

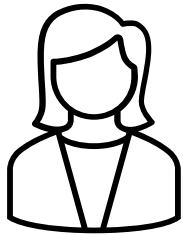
```
Administrator: Windows PowerShell  
  
Windows PowerShell  
Copyright (C) Microsoft Corporation. All rights reserved.  
  
Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows  
  
PS C:\WINDOWS\system32> wsl --install -d ubuntu_
```

LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. **Describe the purpose of GitHub and provide usage overview.**
3. *Provide basic overview of BASH commands.*
4. *Practice navigating through the terminal using commands.*
5. *Describe terminal permission structure for files.*
6. *Describe the purpose of conda.*
7. *Install conda environment and tools.*

Intro to Git and Version control

Project: build analysis pipeline



Grace



script_v0.1.sh



Pikachu



Pika_script.sh



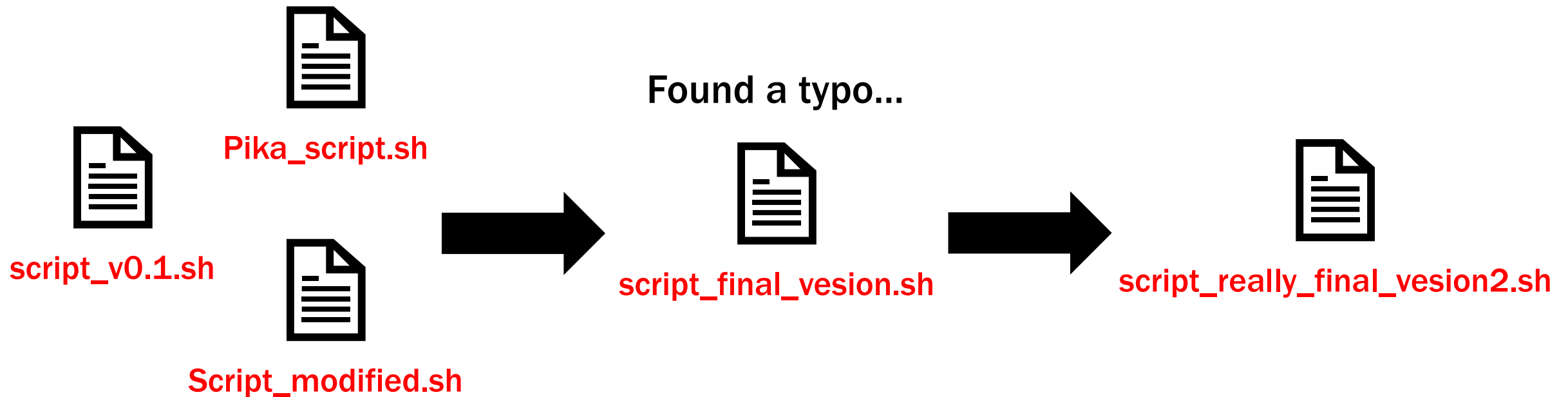
Snowball



Script_modified.sh

Intro to Git and Version control

How to merge all versions into one final script...



Intro to Git and Version control

- 1. Git is an open-source version control system software that tracks changes.**
- 2. Changes are tracked for easy comparison.**
3. You can find out what was changed in each update.



Intro to Git and Version control

SERVERPROPERTY Resource DB.sql (Working Tree) - RajendraScripts - Visual Studio Code

Message (Ctrl+Enter ...)

Changes 1

SERVER... + M

Unchanged row

Inserted rows

Removed rows

```
1 SELECT
2- SERVERPROPERTY('servername'),
3  SERVERPROPERTY('ProductVersion ' ) AS ProductVersion,
4- SERVERPROPERTY('ProductLevel') AS ProductLevel,
5
6+ SERVERPROPERTY('ResourceLastUpdateDateTime') AS ResourceLastUpd
7- SERVERPROPERTY('Collation') AS Collation;
8
9+ SERVERPROPERTY('Collation') AS Collation,
10+ SERVERPROPERTY('Collation') AS Collation,
11+ SERVERPROPERTY('ResourceLastUpdateDateTime') AS ResourceLastUpd
```

<https://code.visualstudio.com/docs/sourcecontrol/overview>

Intro to Git and Version control

1. Git is an open-source version control system software that tracks changes.
2. Changes are tracked for easy comparison.
3. **You can find out what was changed in each update.**

Intro to Git and Version control

The screenshot displays the GitHub interface for the repository `mmid-bioinformatics-workshop / 2023-03-02-Intro-to-the-2023-MMID-BFX-Workshop`. The repository is public and has 2 stars. The main branch is `main`. The commit history is filtered by date, showing commits from March 8, 2023, and March 1, 2023.

Commits on Mar 8, 2023

- Update README.md** (Verified) `f172857` `<>`
mmid-bioinformatics-workshop committed 7 hours ago
- Update README.md** (Verified) `268d671` `<>`
mmid-bioinformatics-workshop committed 7 hours ago

Commits on Mar 1, 2023

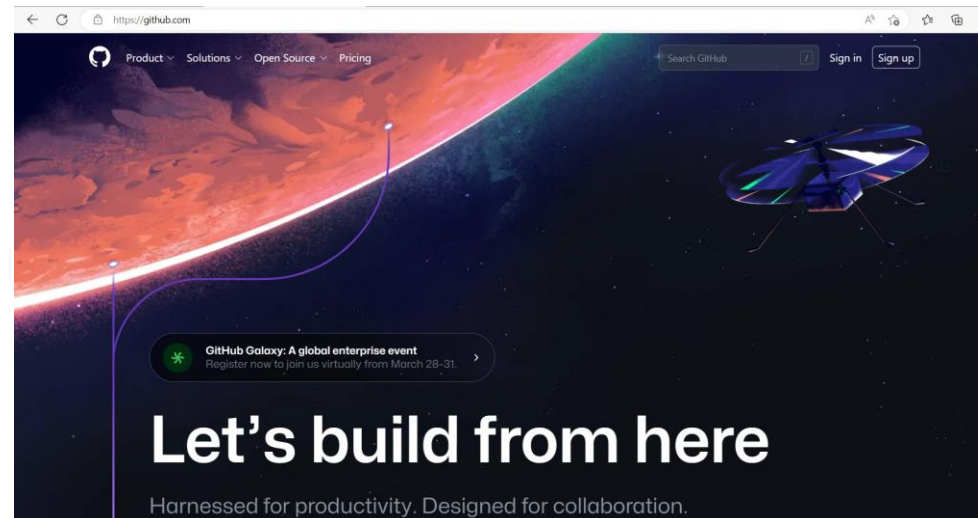
- Add files via upload** (Verified) `e926d05` `<>`
graceeseo committed last week
- Create LICENSE.md** (Verified) `4b204db` `<>`
mmid-bioinformatics-workshop committed last week
- Initial commit** (Verified) `6121f6c` `<>`
mmid-bioinformatics-workshop committed last week

Navigation buttons: `Newer` and `Older`.

Intro to Git and Version control



- GitHub is a Git repository hosting service with web-based graphical interface.
- We will use GitHub to share workshop related contents, data, scripts and PowerPoint slides.



GitHub usage

Demonstration

LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. *Describe the purpose of GitHub and provide usage overview.*
3. ***Provide basic overview of BASH commands.***
4. *Practice navigating through the terminal using commands.*
5. *Describe terminal permission structure for files.*
6. *Describe the purpose of conda.*
7. *Install conda environment and tools.*

BASH prompt

Users specify items written in ALL CAPS

USERNAME@COMPUTER_NAME:./Desktop\$

Current directory

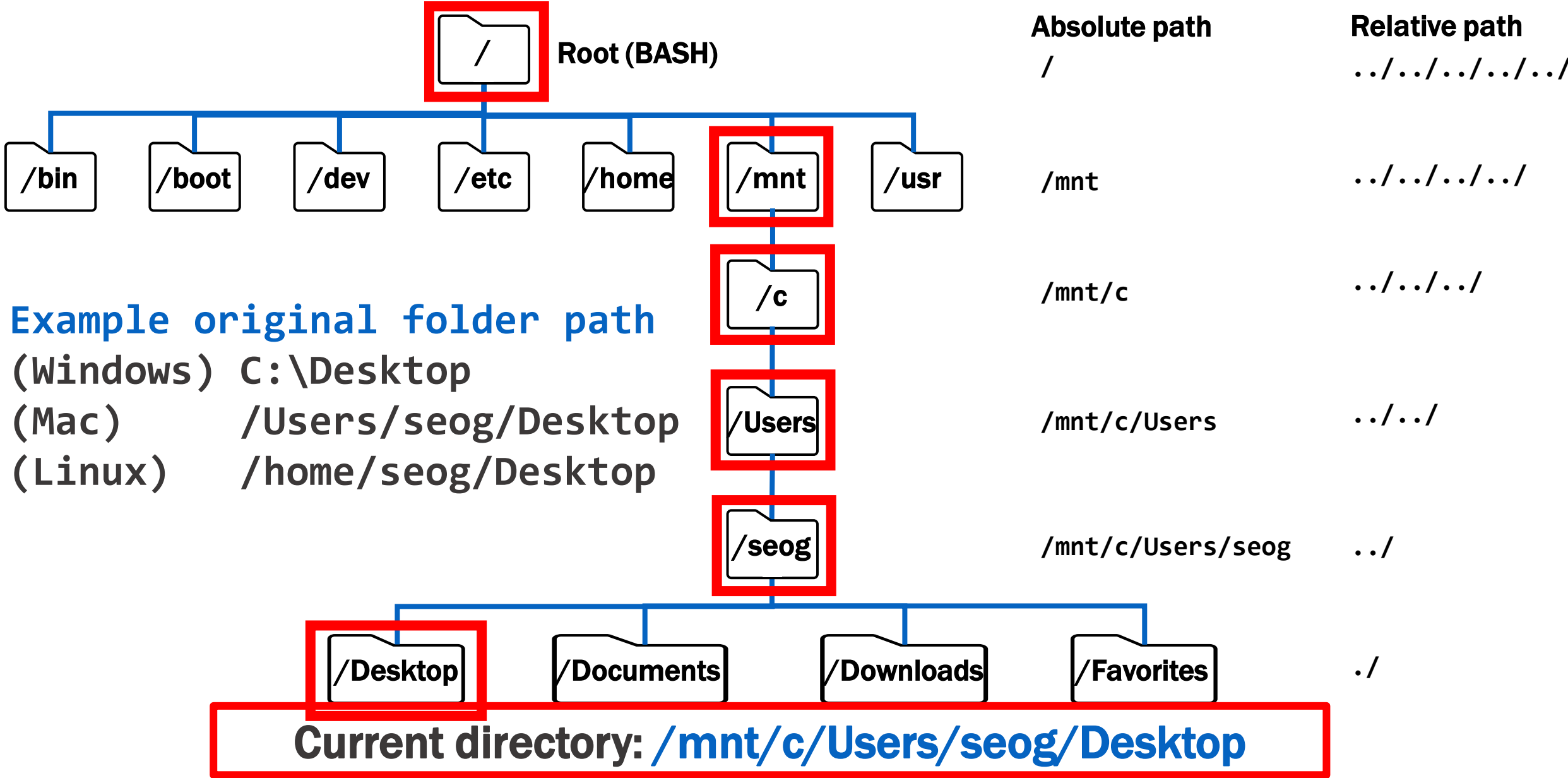
Prompt – Input command

If there is no prompt, either job is in progress or wrong command stalled your terminal

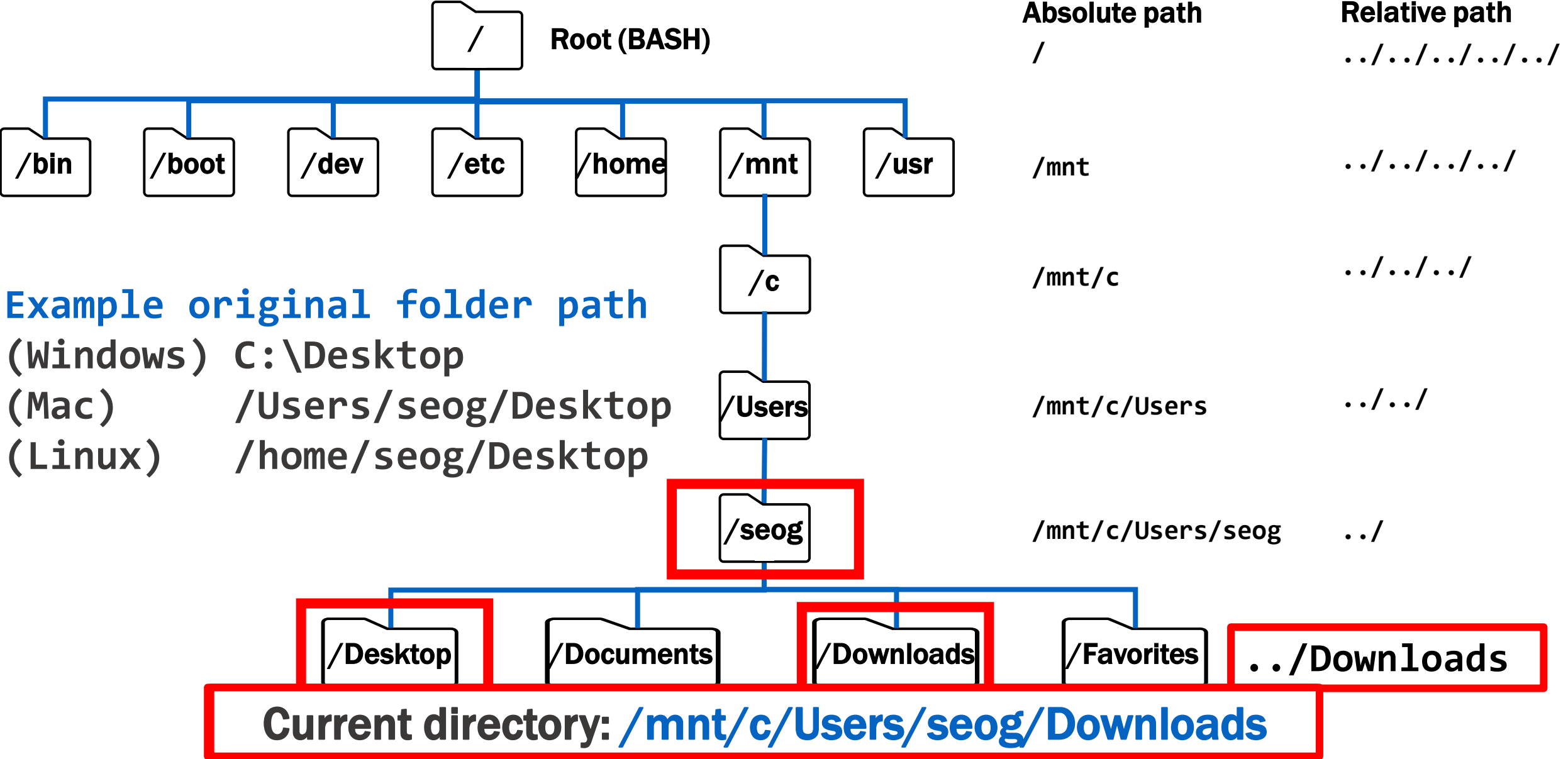
To get your prompt back
(Windows / Linux) Press ctrl-c
(Mac) Press command-c

```
seog@Smarty:.../Desktop$
```

Example Windows Subsystem for Linux BASH folder structure



Example Windows Subsystem for Linux BASH folder structure



BASH commands

- Print working directory `$ pwd`
- List directory content `$ ls` `$ ll`
- Change directory `$ cd`
- Create directory `$ mkdir`
- Create a file `$ touch`
- Rename file/directory `$ mv OLDNAME NEWNAME`
- Move file/directory `$ mv FILE NEWPLACE`
- Copy file/directory `$ cp FILE NEWCOPY`

LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. *Describe the purpose of GitHub and provide usage overview.*
3. *Provide basic overview of BASH commands.*
4. ***Practice navigating through the terminal using commands.***
5. *Describe terminal permission structure for files.*
6. *Describe the purpose of conda.*
7. *Install conda environment and tools.*

BASH Basics

Demonstration

LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. *Describe the purpose of GitHub and provide usage overview.*
3. *Provide basic overview of BASH commands.*
4. *Practice navigating through the terminal using commands.*
5. **Describe terminal permission structure for files.**
6. *Describe the purpose of conda.*
7. *Install conda environment and tools.*

File information and permission

total 0		File size		Date & time created			File name
		Owner	Group				
-rwxrwxrwx	1	seog	seog	0	Jan	9 23:14	test0001.txt
-rwxrwxrwx	1	seog	seog	0	Jan	9 23:14	test0002.txt

r: readable
w: writable
x: executable

Permission (\$ chmod ###)

No permission	- - -	- - -	- - -
Symbol:	rwx	rwx	rwx
Bit:	421	421	421
Triad:	Owner	group	others

```
$ chmod 755 test0001.txt
Owner(rwx) group(r-x) others(r-x)
```

```
$ chmod a+wrx test0001.txt
Owner(rwx) group(rwx) others(rwx)
```

File information and permission

Demonstration

LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. *Describe the purpose of GitHub and provide usage overview.*
3. *Provide basic overview of BASH commands.*
4. *Practice navigating through the terminal using commands.*
5. *Describe terminal permission structure for files.*
6. **Describe the purpose of conda.**
7. *Install conda environment and tools.*

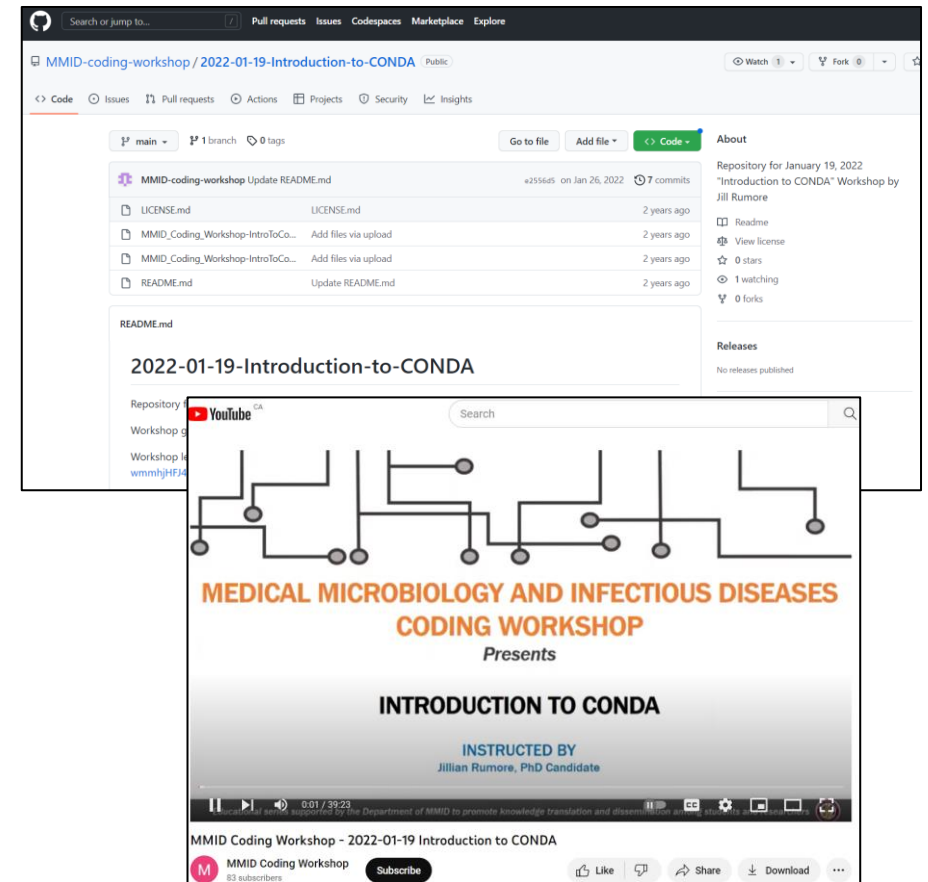
Conda guide (2022)

Jill Rumore's Conda lecture (2022):

<https://github.com/MMID-coding-workshop/2022-01-19-Introduction-to-CONDA>

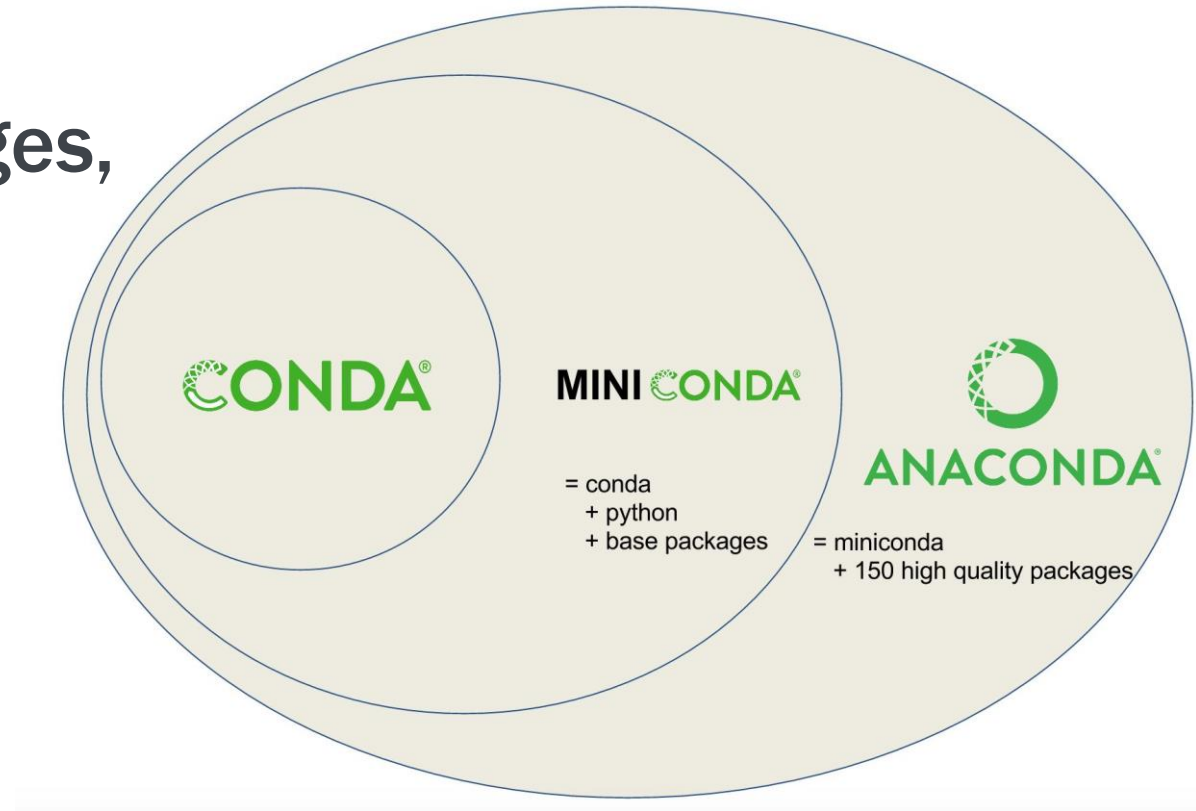
MMID Coding Workshop YouTube (2022-01-19):

<https://www.youtube.com/watch?v=r-wmmhjHFJ4>



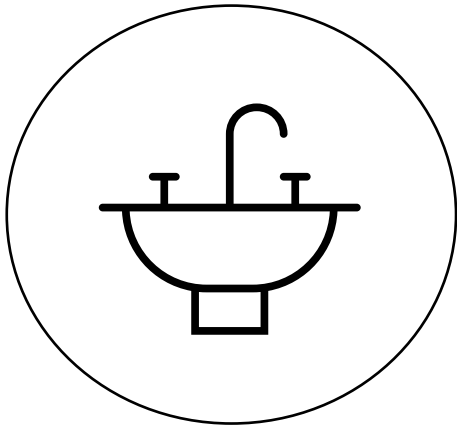
What is conda?

- conda-forge: Most dependencies (numpy, scipy, zlib, CRAN packages, etc.)
- **bioconda: Most bioinformatics packages are available in this channel i.e. (STAR, samtools, DESeq2, etc.)**
- defaults: Packages built by Anaconda Inc.

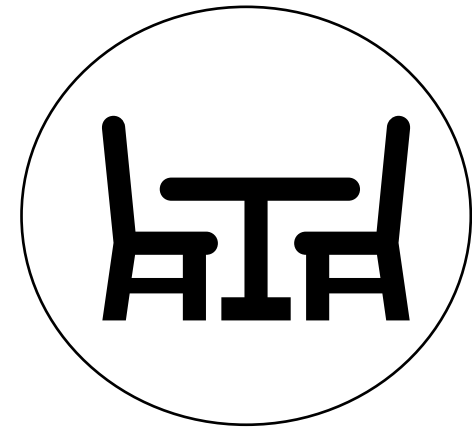
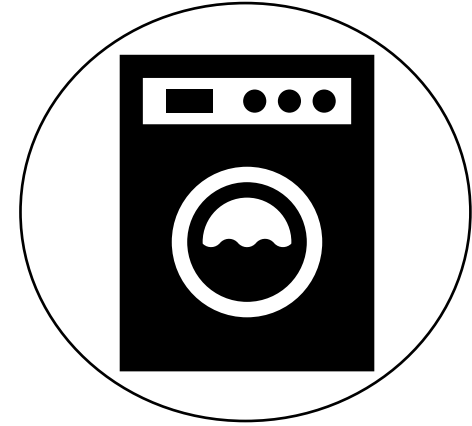


<https://bioconda.github.io/tutorials/gcb2020.html>

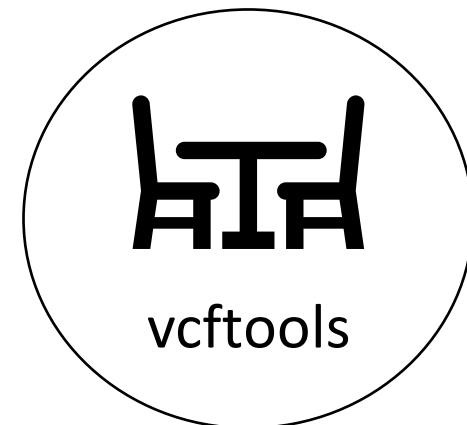
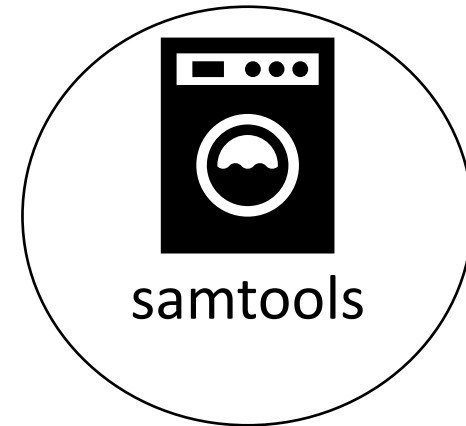
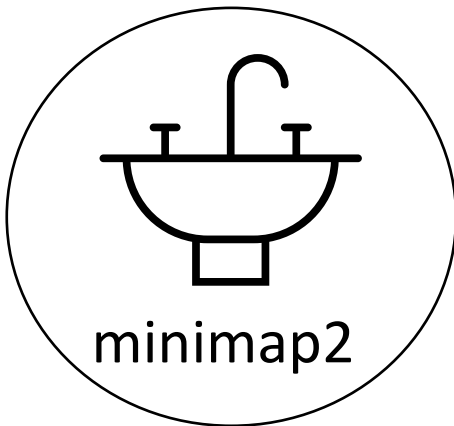
What is conda?



Share home



Activate (into the room) & deactivate (leave the room) conda environment



LEARNING OBJECTIVES

1. *Describe method to install Ubuntu using PowerShell (Windows user)*
2. *Describe the purpose of GitHub and provide usage overview.*
3. *Provide basic overview of BASH commands.*
4. *Practice navigating through the terminal using commands.*
5. *Describe terminal permission structure for files.*
6. *Describe the purpose of conda.*
7. **Install conda environment and tools.**

Install conda environment

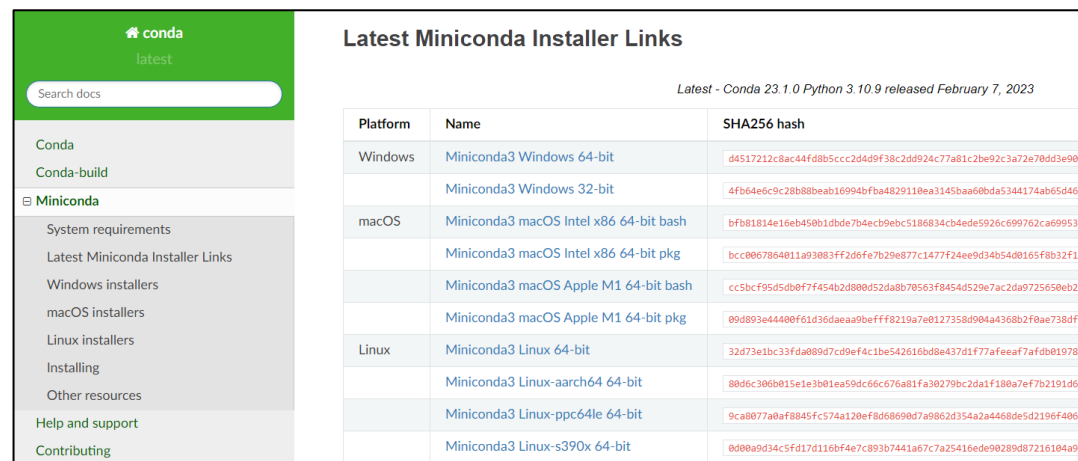
1. Update your terminal with dependencies

```
$ sudo apt update
```

```
$ sudo apt upgrade
```

2. Click **Miniconda3 Linux 64-bit** to download .sh script

<https://docs.conda.io/en/latest/miniconda.html>



The screenshot shows the Conda documentation website. On the left is a navigation sidebar with a green header containing the 'conda' logo and 'latest' text. Below the header is a search bar and a list of links: 'Conda', 'Conda-build', 'Miniconda' (expanded), 'System requirements', 'Latest Miniconda Installer Links', 'Windows installers', 'macOS installers', 'Linux installers', 'Installing', 'Other resources', 'Help and support', and 'Contributing'. The main content area is titled 'Latest Miniconda Installer Links' and includes a note: 'Latest - Conda 23.1.0 Python 3.10.9 released February 7, 2023'. Below this is a table with three columns: 'Platform', 'Name', and 'SHA256 hash'.

Platform	Name	SHA256 hash
Windows	Miniconda3 Windows 64-bit	d4517212c8ac44fd8b5ccc2d4d9f38c2dd924c77a81c2be92c3a72e70dd3e907
	Miniconda3 Windows 32-bit	4fb64e6c9c28b88beab16994bfba4829110ea3145baa60bda5344174ab65d462
macOS	Miniconda3 macOS Intel x86 64-bit bash	bfb81814e16eb450b1dbde7b4ecb9ebc5186834cb4ede5926c699762ca69953b
	Miniconda3 macOS Intel x86 64-bit pkg	bcc0067864811a93083ff2d6fe7b29e877c1477f24ee9d34b54d0165f8b32f11
	Miniconda3 macOS Apple M1 64-bit bash	cc5bcf95d5db0f7f454b2d800d52da8b78563f8454d529e7ac2da9725650eb27
	Miniconda3 macOS Apple M1 64-bit pkg	09d893e44400f61d36daea9befff8219a7e0127358d904a4368b2f0ae738df0
Linux	Miniconda3 Linux 64-bit	32d73e1bc33fda089d7cd9ef4c1be542616bd8e437d1f77afeef7afdb019787
	Miniconda3 Linux-aarch64 64-bit	80d6c306b015e13b01ea59dc66c676a81fa30279bc2da1f180a7ef7b2191d5e
	Miniconda3 Linux-ppc64le 64-bit	9ca0877a0af8845fc574a120ef8d68690d7a9862d354a2a468de5d2196f406c
	Miniconda3 Linux-s390x 64-bit	0d0e0d34c5fd17d116bf4e7c893b7441a67c7a25416ede90289d87216104a97

Install conda environment

3. Open Terminal (Ubuntu/Terminal)

```
$ cd /Downloads
```

OR

```
$ cd /mnt/c/Users/USERNAME/Downloads
```

4. Install Miniconda using bash script

```
$ bash Miniconda3-latest-Linux-x86_64.sh
```

```
seog@SMARTY:/mnt/c/Users/gsezz/Downloads$ bash Miniconda3-latest-Linux-x86_64.sh

Welcome to Miniconda3 py310_23.1.0-1

In order to continue the installation process, please review the license
agreement.
Please, press ENTER to continue
>>>
```

Install conda environment

5. Press q to exit from agreement documentation.
6. Type “yes” and press ENTER to accept the license terms.
7. Press ENTER to install the software (~1-2 mins)

```
Do you accept the license terms? [yes|no]
[no] >>> yes

Miniconda3 will now be installed into this location:
/home/seog/miniconda3

- Press ENTER to confirm the location
- Press CTRL-C to abort the installation
- Or specify a different location below

[/home/seog/miniconda3] >>> █
```

Install conda environment

8. Type “yes” to initialization and press ENTER

9. Close and re-open the terminal for the changes to take effect.

```
Do you wish the installer to initialize Miniconda3
by running conda init? [yes|no]
[no] >>> yes
no change      /home/seog/miniconda3/condabin/conda
no change      /home/seog/miniconda3/bin/conda
no change      /home/seog/miniconda3/bin/conda-env
no change      /home/seog/miniconda3/bin/activate
no change      /home/seog/miniconda3/bin/deactivate
no change      /home/seog/miniconda3/etc/profile.d/conda.sh
no change      /home/seog/miniconda3/etc/fish/conf.d/conda.fish
no change      /home/seog/miniconda3/shell/condabin/Conda.psm1
no change      /home/seog/miniconda3/shell/condabin/conda-hook.ps1
no change      /home/seog/miniconda3/lib/python3.10/site-packages/xontrib/conda.xsh
no change      /home/seog/miniconda3/etc/profile.d/conda.csh
modified       /home/seog/.bashrc

==> For changes to take effect, close and re-open your current shell. <==

If you'd prefer that conda's base environment not be activated on startup,
set the auto_activate_base parameter to false:

conda config --set auto_activate_base false

Thank you for installing Miniconda3!
seog@SMARTY:/mnt/c/Users/gsezz/Downloads$
```

Install conda environment

10. Verify installation by checking conda version.

```
$ conda --version
```

```
(base) seog@SMARTY:/mnt/c/Users/gsezz$ conda --version  
conda 23.1.0
```

Navigating through conda

- Conda help manual `$ conda --help`
 - Conda command help page `$ conda COMMAND --help`
 - List of installed environments `$ conda env list`
 - Search for conda packages `$ conda search PACKAGENAME`
-
- **Core commands:** create, list, search, install, update, remove

```
(base) seog@SMARTY:/mnt/c/Users/gsezz$ conda env list
# conda environments:
#
base                  *  /home/seog/miniconda3
```

Install conda environment

1. Install conda environment named conda_workshop

```
$ conda create --yes --name conda_workshop
```

```
(base) seog@SMARTY:/mnt/c/Users/gsezz$ conda create -y -n conda_workshop
```

```
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate conda_workshop
#
# To deactivate an active environment, use
#
#     $ conda deactivate

(base) seog@SMARTY:/mnt/c/Users/gsezz$
```

Install conda environment

2. Activate conda_workshop environment

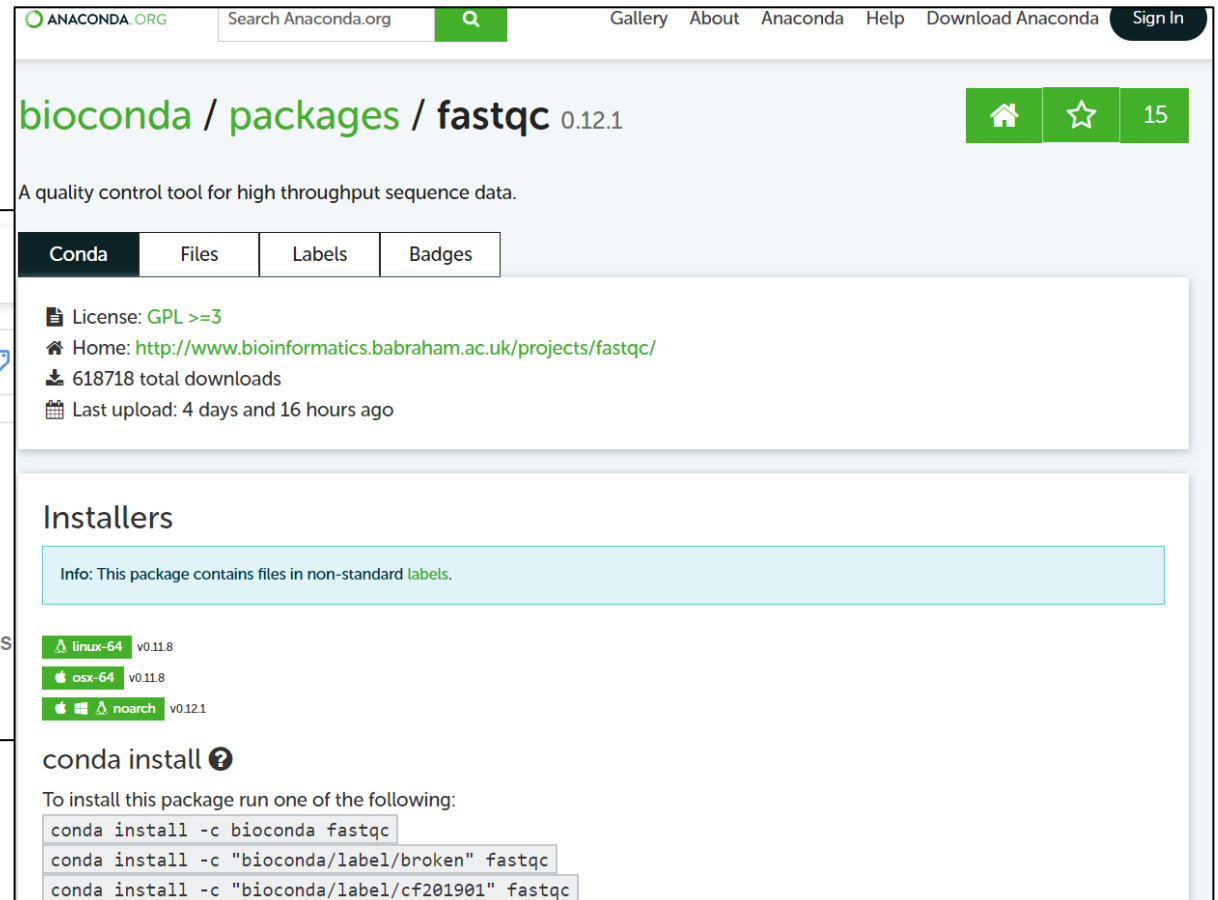
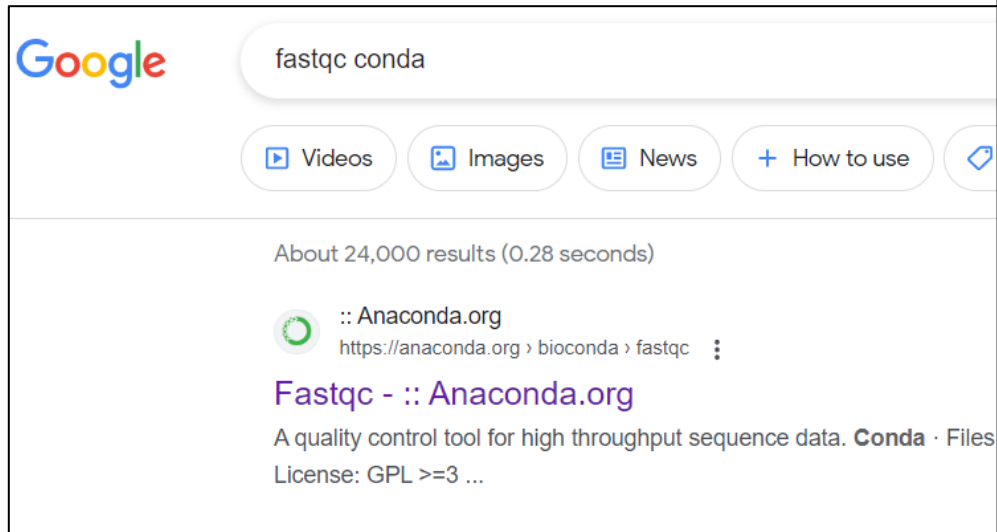
\$ conda activate conda_workshop

```
(base) seog@SMARTY:/mnt/c/Users/gsezz$ conda activate conda_workshop
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ conda env list
# conda environments:
#
base                /home/seog/miniconda3
conda_workshop      *  /home/seog/miniconda3/envs/conda_workshop

(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$
```


Install conda package - fastqc

3. Look for the tool availability on Anaconda (Google search)



Install conda package - fastqc

4. Install conda packages in an environment

```
$ conda list
```

```
$ conda install --yes --channel CHANNEL PACKAGENAME
```

```
conda install ?
```

To install this package run one of the following:

```
conda install -c bioconda fastqc
```

```
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ conda install -y -c bioconda fastqc
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

environment location: /home/seog/miniconda3/envs/conda_workshop
```

Install conda package - fastqc

5. Check the conda list to verify installation

\$ conda list

```
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ conda list
# packages in environment at /home/seog/miniconda3/envs/conda_workshop:
#
# Name                                Version                                Build      Channel
_libgcc_mutex                         0.1                                    main
_openmp_mutex                         5.1                                    1_gnu
dbus                                  1.13.18                               hb2f20db_0
expat                                 2.4.9                                 h6a678d5_0
fastqc                                0.12.1                               hdfd78af_0  bioconda
font-ttf-dejavu-sans-mono             2.37                                  hd3eb1b0_0
fontconfig                             2.14.1                               h52c9d5c_1
freetype                              2.12.1                               h4a9f257_0
gdbm                                  1.18                                 hd4cb3f1_4
```

Using the package - fastqc

6. Use the tool command to use (while inside conda environment)

```
$ fastqc --help
```

```
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ fastqc --help

FastQC - A high throughput sequence QC analysis tool

SYNOPSIS

    fastqc seqfile1 seqfile2 .. seqfileN

    fastqc [-o output dir] [--(no)extract] [-f fastq|bam|sam]
          [-c contaminant file] seqfile1 .. seqfileN

DESCRIPTION

    FastQC reads a set of sequence files and produces from each one a quality
    control report consisting of a number of different modules, each one of
    which will help to identify a different potential type of problem in your
    data.
```

Updating the package - fastqc

7. Regularly check for tool updates and update package regularly.

```
$ conda update --yes PACKAGENAME
```

```
conda update --yes myenv scipy  
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ conda update --yes fastqc  
Collecting package metadata (current_repodata.json): done  
Solving environment: done  
  
# All requested packages already installed.
```

Deactivate conda environment

8. Leave conda_workshop environment.

```
$ conda deactivate
```

```
(conda_workshop) seog@SMARTY:/mnt/c/Users/gsezz$ conda deactivate  
(base) seog@SMARTY:/mnt/c/Users/gsezz$ █
```

Practice conda package installation

Install the following tools in your conda_workshop environment before the next workshop!

- 1. fastqc (completed)**
- 2. minimap2**
- 3. vcftools**
- 4. samtools**
- 5. fastp**
- 6. checkM**
- 7. Kraken2**

LEARNING OBJECTIVES

- 1. Describe method to install Ubuntu using PowerShell (Windows user)***
- 2. Describe the purpose of GitHub and provide usage overview.***
- 3. Provide basic overview of BASH commands.***
- 4. Practice navigating through the terminal using commands.***
- 5. Describe terminal permission structure for files.***
- 6. Describe the purpose of conda.***
- 7. Install conda environment and tools.***

HELPFUL RESOURCES - BASH

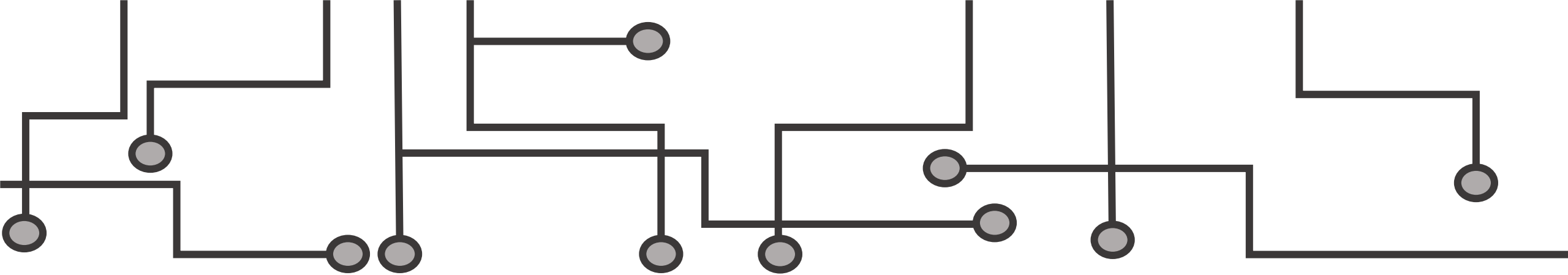
1. **What is CLI** <https://www.hostinger.com/tutorials/what-is-cli>
2. **BASH manual:** <https://www.gnu.org/software/bash/manual/bash.html>
3. **Information on Linux folder structure**
<https://www.howtogeek.com/117435/htg-explains-the-linux-directory-structure-explained/>
4. ***The BASH Guide:** <https://guide.bash.academy/>
5. ***Learn Enough Command-Line to be dangerous (free first few chapters):**
<https://www.learnenough.com/command-line-tutorial>

YouTube Videos

1. ***Joe Collins - Beginner's Guide to the Bash Terminal:**
<https://www.youtube.com/watch?v=oxuRxtrO2Ag>
2. ***Traversy Media – Shell Scripting Crash Course – Beginner Level:**
<https://www.youtube.com/watch?v=v-F3YLd6oMw>

HELPFUL RESOURCES - Conda

1. ***Jill Rumore's conda workshop (2022-01-19)**
<https://www.youtube.com/watch?v=r-wmmhjHFJ4>
2. ***Jill Rumore's conda workshop GitHub (2022-01-19)** <https://github.com/MMID-coding-workshop/2022-01-19-Introduction-to-CONDA>
3. **Conda userguide** <https://docs.conda.io/projects/conda/en/latest/user-guide/getting-started.html>
4. **Conda cheatsheet** <https://docs.conda.io/projects/conda/en/latest/user-guide/cheatsheet.html>
5. **Conda cheatsheet pdf**
https://docs.conda.io/projects/conda/en/4.6.0/_downloads/52a95608c49671267e40c689e0bc00ca/conda-cheatsheet.pdf



THANK YOU FOR ATTENDING!

Please make sure to fill out the [Exit Survey at Slido.com with #9406658](#)
We value your feedback!

More questions? Please email us at mmid.bioinformatics.workshop@gmail.com or post them to the workshop [slack channel](#)

Exit survey

Slido

Slido.com with #9406658