

# **RGGS Comparative Genomics 2 – Computational Methods (Session 2)**

Jose Barba

Gerstner Scholar in Bioinformatics & Computational Biology

## **Session 2 outline**

- **Review of basic Unix navigation tutorial**
- **Command line computing basics 1**
- **Connecting to remote servers using the terminal and a package manager**

# Review of basic navigation in the terminal

- Using the command line interface involves executing commands to navigate directories, view/edit/create files, among other tasks

Archiving and Compression	Networking and Remote Operations
<b>tar</b> - Archive files into a tarball (or extract them)	<b>curl</b> - Download and transfer data from remote servers
<b>zip/unzip</b> - Compress and extract files in zip format	<b>exit</b> - Close current shell, terminal prompt, interactive program
File and Directory Operations	<b>scp</b> - Secure copy, used for transferring files between hosts
<b>cat</b> - Concatenate and display file contents	<b>sftp</b> - Securely transfer files between computers over a network
<b>cd</b> - Change directory	<b>ssh</b> - Secure shell, used for logging into a remote machine
<b>cp</b> - Copy files or directories	<b>wget</b> - Download and transfer data from remote servers
<b>head</b> - Display the first few lines of a file	Operators
<b>less</b> - Read file one screen at a time	<b>*</b> - Mangle multiple files and directories
<b>ls</b> - List files and directories	<b>&amp;</b> - Run a command in the background
<b>mkdir</b> - Make directory	<b>&gt;</b> - Redirect the output of a command to a file
<b>mv</b> - Move or rename files or directories	<b> </b> - Connect the output of one command to the input of another
<b>pwd</b> - Print working directory	Process Management
<b>rm</b> - Remove files or directories	<b>kill</b> - Terminate processes by ID
<b>rmdir</b> - Remove directory	<b>ps</b> - Display information about running processes
<b>sort</b> - Sort the lines of a file	<b>top</b> - Interactively manage processes
<b>tail</b> - Display the last few lines of a file	Terminal Utilities
<b>touch</b> - Create an empty file or update a file's timestamp	<b>clear</b> - Clear the terminal screen
<b>tr</b> - Translate or delete characters	<b>help</b> - Show information about built-in shell commands
File and Directory Search	<b>man</b> - Display manual for a command
<b>find</b> - Search for files and directories within a hierarchy	<b>screen</b> - Managing terminal sessions
<b>grep</b> - Search for patterns within text	Text Processing
File Permissions and Ownership	<b>awk</b> - Pattern scanning and processing language
<b>chmod</b> - Change file permissions	<b>cut</b> - Extract specific sections of text
<b>chown</b> - Change file owner and group	<b>echo</b> - Display a line of text or variables
File Space and Disk Usage	<b>nano</b> - Text editor
<b>df</b> - Display disk space usage	<b>sed</b> - Stream editor for filtering and transforming text
<b>du</b> - Estimate file space usage	<b>vi</b> - Text editor

# Review of basic navigation in the terminal

- **Instructions to download the basic Unix navigation tutorial to the home directory:**

1. Open the terminal
2. Type ``cd ~``
3. Type ``wget`

[https://raw.githubusercontent.com/josebarbamontoya/rggs\\_comparative\\_genomics\\_2/main/session\\_01/basic\\_unix\\_navigation\\_tutorial.sh](https://raw.githubusercontent.com/josebarbamontoya/rggs_comparative_genomics_2/main/session_01/basic_unix_navigation_tutorial.sh)`

# Command line computing basics 1

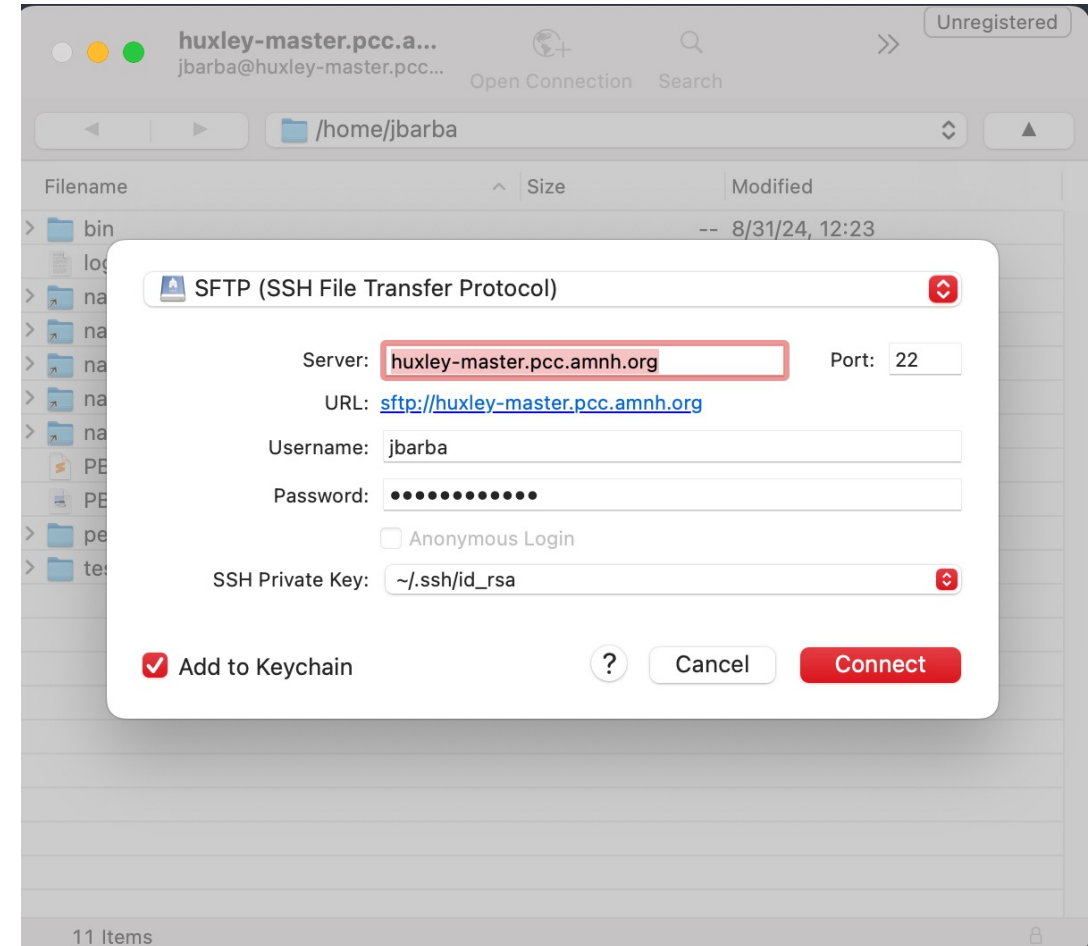
- This tutorial builds upon basic command-line navigation and file manipulation by introducing advanced operations, text processing tools, simple scripting, and filesystem management
- These tasks are crucial for data analysis, automation, and system administration
- Instructions to download the command line computing basics 1 tutorial to the home directory:
  1. Open the terminal
  2. Type ``cd ~``
  3. Type ``wget``  
[https://raw.githubusercontent.com/josebarbamontoya/rggs\\_comparative\\_genomics\\_2/main/session\\_02/command\\_line\\_computing\\_basics\\_1\\_tutorial.sh`](https://raw.githubusercontent.com/josebarbamontoya/rggs_comparative_genomics_2/main/session_02/command_line_computing_basics_1_tutorial.sh)

# Connecting to remote servers using the terminal and a package manager

- **Introduction to Huxley:**  
[https://amnh.sharepoint.com/sites/Bioinformatics/SitePages/Getting\\_Started\\_Huxley.aspx](https://amnh.sharepoint.com/sites/Bioinformatics/SitePages/Getting_Started_Huxley.aspx)
- **Instructions to download the `ssh` and `scp` commands tutorial to the home directory:**
  1. Open the terminal
  2. Type `cd ~`
  3. Type `wget`  
[https://raw.githubusercontent.com/josebarbamontoya/rggs\\_comparative\\_genomics\\_2/main/session\\_02/ssh\\_and\\_scp\\_tutorial.sh](https://raw.githubusercontent.com/josebarbamontoya/rggs_comparative_genomics_2/main/session_02/ssh_and_scp_tutorial.sh)

# Connecting to remote servers using the terminal and a package manager

- Download Cyberduck package manager (for mac):  
<https://cyberduck.io/download/>
- Instructions to connect remotely to Huxley:
  1. Open the Cyberduck
  2. Click the on the open connections tab
  3. Select SFTP
  4. Server: huxley-master.pcc.amnh.org
  5. Username: AMNH username
  6. Enter your password
  7. Select Create private key
  8. Tick Add to keychain box



# Connecting to remote servers using the terminal and a package manager

- To log in to the AMNH servers from outside the museum, a Cisco VPN connection is required.
- Instructions for installing and using Cisco VPN are available here:

[https://github.com/josebarbamontoya/rggs\\_comparative\\_genomics\\_2/blob/main/session\\_02/Duo\\_and\\_VPN\\_User\\_Guide.pdf](https://github.com/josebarbamontoya/rggs_comparative_genomics_2/blob/main/session_02/Duo_and_VPN_User_Guide.pdf)