

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

Command line computing basics 1

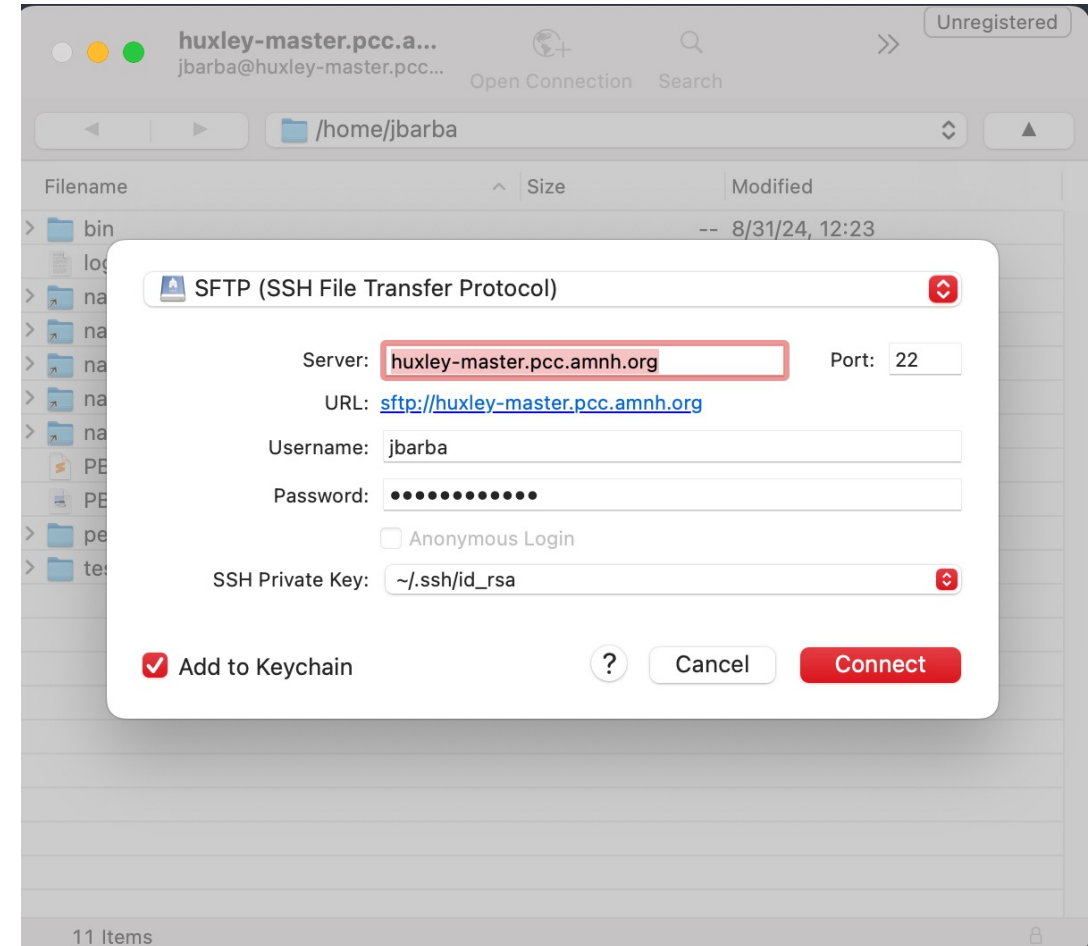
- This tutorial builds upon basic command-line navigation and file manipulation by introducing advanced operations, text processing tools, simple scripting, and filesystem management
- These tasks are crucial for data analysis, automation, and system administration
- Instructions to download the command 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
- **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

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