BIOF045: Next-Generation Sequencing Data Analysis on the Cloud

Hands-On Training - Online

October 19-22, 2020

FAES@NIH

Instructors:

Shahin Shahsavari

Vinay Swamy

Pre-Training Guide

1. Tentative Agenda

Day1 (October 19)

Introduction to Cloud Computing
Linux Basics
Shell Scripting

Day2 (October 20)

Next Generation Sequencing
DNA-Seq:
Alignment
Variant Calling
Data Visualization

Day3 (October 21)

Introduction to R and Rstudio Server
RNASeq:
Alignment
Differential Gene Expression
Data Visualization

Day4 (October 22)

Accessing Publicly Available Datasets and Resources
Exercise
Wrap-Up

Daily Workshop Schedule
9:30AM - 5:00PM EST
Lunch: 12:00 – 1:00PM

Office Hours
5:00PM - 5:30PM EST

2. Installations

Please make sure that the following softwares are installed on your computer:

1. X2Go Client

   Note: macOS users need to install XQuatz as well.

2. IGV

3. FileZilla

   a. Windows
   b. macOS

4. Zoom

5. PuTTY (Windows users ONLY)

3. Pre-Training Resources

Some great introductory resources to use before the first day of the workshop:

1. Cloud Computing

2. Linux

3. UNIX vs. Linux

4. Illumina Sequencing

4. Workshop Logistics

4.1. Local Setup

Point 1: Make sure to install all the softwares before the start of the class.

Point 2: If you have access to two computers:

You could use one laptop/desktop to connect to the Zoom session, and use the other to follow the hands-on training. However, I recommend connecting to the zoom meeting on both computers in case you need to share your screen and troubleshoot.

4.2. Remote Server

We will run an Ubuntu server for the full duration of the workshop. You will receive your unique Username/Password combination with instructions to connect to the remote server. <u>Please make sure to test your connection before the class.</u> Send me an email if you are unable to connect.

4.3. Training Material

I will publish most of the hands-on training materials on GitHub with detailed explanations. You may need to use it frequently to copy the commands during the workshop.

4.4. Support During the Training

If you face any technical difficulties during the training, feel free to send a message over chat or unmute yourself and ask. We will use screen sharing and break rooms to help you get back on track.

4.5. Office Hours

I will stay after the workshop and during breaks to answer any questions you may have.

5. Contact

If you have any issues about the workshop, online materials, Zoom meeting, remote server, or any other subjects, feel free to contact Shahin using Email or Canvas:

Email: shahsavaris2@nih.gov

6. FAQs

Question 1: Should I have any prior programming background?

A: No prior background in programming is required, though it would be very helpful if you have used a programming language in the past. We will cover what you need to work with computing clusters and analyze your data.

Question 2: Can I use my government computer?

A: It would be best if you could use your personal computer to which you have administrative access. If you must use your work computer, make sure to install the X2GoClient prior to the start of the workshop.

Question 3: How long do we have access to the server?

A: You will be provided with the server IP address and your unique username/password before the workshop. The server will be available until October 23rd 10:00PM. If you would like to transfer your files and data from the cloud to your own computer, you must do so before then. I will show you how it works on the last day.