

# **BIOF045: Next-Generation Sequencing Data Analysis**

Hands-On Training - Online

**March 22-26**

NIH/FAES

Pre-Training Guide

- 1. Tentative Agenda**
- 2. Installations**
- 3. Pre-Training Readings**
- 4. Workshop Logistics**
  - 4.1. Your Local Logistics
  - 4.2. Remote Server
  - 4.3. Training Material
  - 4.4. Support During Training
  - 4.5. Zoom Meeting
  - 4.6. Hours
- 5. Contact**
- 6. FAQs**

## **1. Tentative Agenda**

### **Day1 (March 22)**

- Introduction to Cloud Computing
- Linux Basics
- Shell Scripting
- Introduction to R and Rstudio Server

### **Day2 (March 23)**

- Next Generation Sequencing
  - Data Types
  - Quality Control
- DNA-Seq
  - Alignment
  - Variant Calling
  - Data Visualization

### **Day3 (March 24)**

- RNA-seq
  - Alignment
  - Differential Gene Expression
  - Data Visualization (PCA, tSNE, clustering)
  - Gene Set Enrichment

### **Day4 (March 25)**

- ChIP-seq
  - Alignment
  - Peak Calling
  - Genome Browser (UCSC)
  - Motif Analysis

### **Day5 (March 26)**

- Accessing Publicly Available Datasets and Resources
- Final Project

### **Daily Workshop Schedule**

10:00AM - 4:00PM EDT

### **Office Hours**

5:00PM - 5:30PM EDT

## 2. Installations

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

1. [X2Go Client](#)

Note: macOS users need to install XQuartz as well.

2. [IGV](#)

3. FileZilla

- a. [Windows](#)
- b. [macOS](#)

4. Zoom

### 3. Pre-Training Resources

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

1. [Servers vs. Desktops](#)
2. [Linux](#)
3. [Illumina Sequencing](#)
4. [RNA-seq](#)
5. [ChIP-seq](#)

## 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 Linux Server

We will run an Ubuntu server instance for the full duration of the workshop. You will receive your unique Username/Password combination with instructions to connect to the remote server. Please make sure to test your connection before the class. 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.

### 4.4. Questions and Support During training

In case you face technical issues during the training, feel free to send a message to the TA or in the chat. We will use screen share and breakout rooms to give students in-person help.

### 4.5. Office Hours

We 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 Gmail or Canvas:

Shahin Shahsavari: [shhsvri@gmail.com](mailto:shhsvri@gmail.com)

Jonathan Perrie: [jonathan.perrie@gmail.com](mailto:jonathan.perrie@gmail.com)

## 6. FAQs

Question 1: Should I have any prior programming background?

A: A background in R programming is expected. It would be very helpful if you have used another programming language in the past. We will cover what you need to work with the computing server on the first day.

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, FileZilla, and IGV 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 for one day after the training (March 27th 10:00PM).