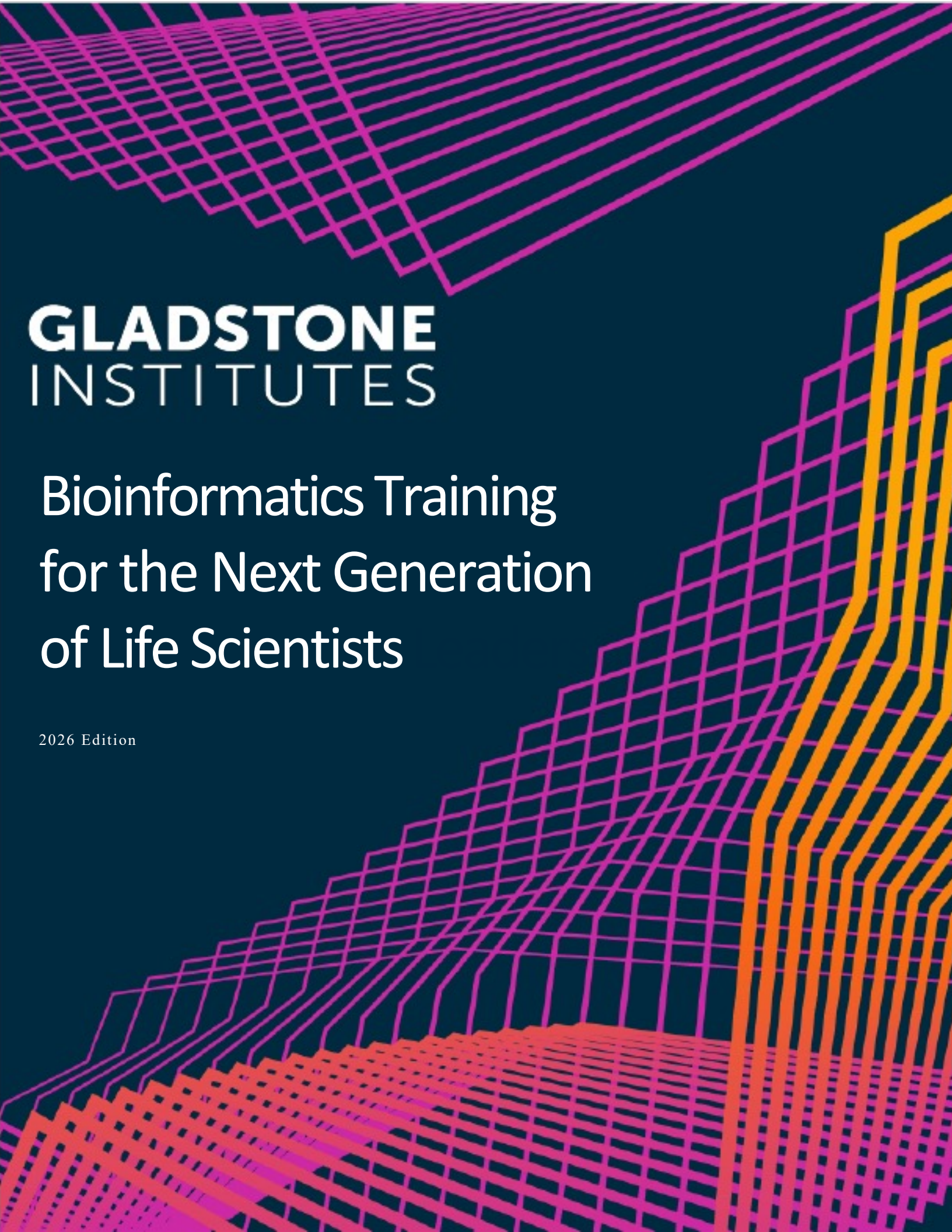




**GLADSTONE**  
INSTITUTES

# Bioinformatics Training for the Next Generation of Life Scientists

2026 Edition

# CONTENTS

## Overview

|                                      |   |
|--------------------------------------|---|
| Introduction to the Curriculum ..... | 3 |
| Organizers .....                     | 4 |

## Program Offerings

|                           |   |
|---------------------------|---|
| Overview .....            | 5 |
| Instructors .....         | 6 |
| Performance Summary ..... | 7 |
| Workshop Details .....    | 8 |

# OVERVIEW

## Introduction to the Curriculum

Gladstone Institutes is dedicated to training tomorrow's leaders in a diverse environment. We provide rigorous scientific training, personalized attention and mentoring, and a rich research environment supplemented by close interactions with our colleagues at nearby universities and organizations.

Initiated in 2018, together with UCSF Bakar Computational Health Sciences Institute, the Gladstone Bioinformatics Core is now partnering with the UCSF School of Medicine, CIRM and Genentech to provide a revitalized program of bioinformatics workshops tailored for our biomedical research community. Each year we add new topics and update materials to reflect state-of-the-art tools, methodologies, and practices. Our goal is to offer bioinformatics, biostatistics, and data science training to a diversity of researchers, spanning introductory and advanced levels.

In the subsequent pages, you will learn about our team and our curriculum. You will find a summary of our performance in recent years, as well as a detail listing of each workshop at the end of this booklet.



**TOP** Gladstone Institutes, Mission Bay, San Francisco **BOTTOM** Members of the Gladstone Bioinformatics Core.



School of Medicine | Bakar Computational Health Sciences Institute



## Organizers

### Reuben Thomas, PhD

Reuben is the Director of the Bioinformatics Core at the Gladstone Institute of Data Science and Biotechnology. With expertise in high-dimensional data statistics, he joined the Core in 2014 following his PhD from Northwestern University and postdoctoral training at the NIH and UC Berkeley. Reuben collaborates with researchers across Gladstone and UCSF on diverse projects, while also benchmarking computational techniques and developing novel methods. He is passionate about demystifying statistical and computational tools—believing that everyone should understand not just how methods work, but why certain approaches succeed or fail for specific applications.



*“We strive to empower researchers with the necessary theoretical background and tools to analyze and interpret their hard-earned biomedical data.”*

**Reuben Thomas, PhD**

### Zainab Yusuf-Sada, BS

Zainab is the Administrative Program Manager for the Gladstone Institute of Data Science and Biotechnology. She received a BS in Public Health and Social Issues from SUNY Buffalo State. Zainab has many years of experience providing a high level of administrative support to executive directors. In addition to providing day to day support for the Core, Zainab is also a community group founder of Elevated Voices focused on DEI at Gladstone



*“The Core has created an inclusive environment that inspires us to collaborate and lead.”*

**Zainab Yusuf-Sada**

# PROGRAM OFFERINGS

## Overview

The Bioinformatics Training program is a rich curriculum of workshops with didactic and hands-on components designed to help postdoctoral scholars, graduate students, and scientific staff with the skills and resources to design experiments, perform analyses, and communicate results. These are essential skills for our trainees in their careers ahead, though often neglected or unavailable earlier in their training. The curriculum is organized into eight topic areas, spanning novice, intermediate and advanced levels. Popular workshops are offered multiple times per year and altogether total over 80 hours of instruction.

## Topic Areas

### Statistics for Biologists

- *Experimental Design and Hypothesis Testing*
- *Statistics of Enrichment Analyses Methods*
- *Linear Mixed Effects Models in R*
- *Pipeline for Linear Mixed Effect Models in R*

### Computer Skills for Data Science

- *Unix command line - I,II*
- *Working on Wynton*

### Machine Learning

- *Machine Learning Using R*

### AI for Research

- *AI Fluency for Biologist*

### Data Analysis Using R

- *Introduction to R for Data Analysis - I,II*
- *Intermediate R for Data Visualization*

### RNA-Seq and ATAC-Seq Analysis

- *Introduction to RNA-seq Analysis - I,II*
- *RNA-seq II: DGE analysis using R*
- *Single cell RNA-Seq Analysis - I,II,III*
- *Advanced RNA-seq Using R*

### ATAC-Seq Analysis

- *Introduction to scATAC-seq data analysis*
- *scATAC-seq and scRNA-seq data integration*

### Pathways Analysis and Modeling

- *Introduction to Pathway Analysis*

## Instructors



**Ayushi Agrawal, MS** is a bioinformatics analyst and programmer in the Bioinformatics Core of the Gladstone Institute of Data Science and Biotechnology. Her relevant expertise is in omics data analysis, including bulk and single cell RNA-seq, and single cell ATAC-seq, as well as R programming.



**Natalie Gill, MS** a bioinformatics analyst and programmer in the Bioinformatics Core of the Gladstone Institute of Data Science and Biotechnology. Her relevant expertise is in omics data analysis, including whole-exome sequencing and bulk/single cell RNA-seq, and has expertise in R, Python, and Unix shell programming.



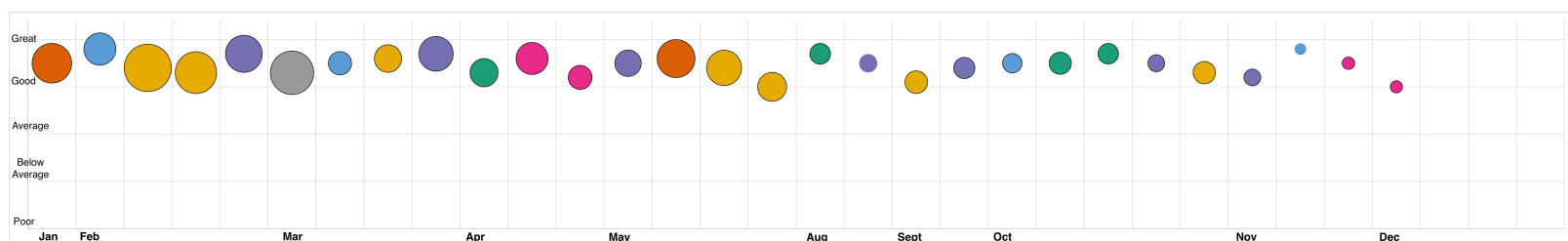
**Kristina Hanspers, MSE** is a senior research associate in the Bioinformatics Core of the Gladstone Institute of Data Science and Biotechnology. Her relevant expertise is in network and pathway analysis, and pathway modeling.



**Michela Traglia, PhD** is a bioinformatics analyst and statistician in the Bioinformatics Core of the Gladstone Institute of Data Science and Biotechnology. Her relevant expertise is in omics data analysis, including bulk and single cell RNA-seq, and single cell ATAC-seq, as well as statistical analysis.

## Performance Summary

In 2025 we registered a total of 2,328 researchers across 29 events, 80 on average per workshop. We collected survey responses for each event and tallied an overall average rank of 4.5 out of 5 for the workshops (see **Figure 1**).



**Figure 1.** Bubble plot of training events spread throughout the year (x-axis) and the overall ranking presented as a descriptive word, i.e., Poor – Great (y-axis). The size of each event bubble is proportional to the number of attendees, with minimum and maximum labeled. The color of each event bubble corresponds to a topic area (see Legend).

**Legend.** Color code for topic areas

|                                  |
|----------------------------------|
| Network and Pathway Analysis     |
| Data Analysis Using R            |
| Statistics for Biologists        |
| Computer Skills for Data Science |
| ATAC-Seq Analysis                |
| RNA-Seq Analysis                 |
| Machine Learning                 |

## Workshop Details

| Title                                              | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Speaker                                    | Duration                            | Topic Area                                                                                     |
|----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------|
| Experimental Design and Hypothesis testing using R | <p>Why do we perform experiments? What conclusions would we like to be able to draw from these experiments? Who are we trying to convince? How does the “magic” of statistics help us reach conclusions?</p> <p>As we attempt to make discoveries, we test new hypotheses using experimental data and hope that a skeptical and discerning person would start to believe the claims we make. Hypothesis testing, a branch of statistics, is a step we can rely on to bolster our claims.</p> <p>In this 2-day workshop you will gain a fundamental understanding of common hypothesis-testing concepts and terms, such as null hypothesis, alternative hypothesis, type I error, type II error, p-value, and power.</p>                                                                                                                                   | Reuben Thomas, PhD<br>Michela Traglia, PhD | Two 2-hour sessions<br>Twice a year | Statistics for Biologists<br><br><i>Prerequisites: None</i>                                    |
| Statistics of Enrichment Analyses Methods          | <p>The use of gene set or pathway enrichment methods is ubiquitous for assigning functional or mechanistic significance to results from high-throughput assays like RNA-sequencing. This workshop will focus on the theory behind some of these methods and also on the interpretation of the associated results. Even though different methods will be illustrated using R code, the emphasis will be more on the understanding the background and assumptions of the methods. See our other workshops for performing enrichment using online tools and R packages.</p>                                                                                                                                                                                                                                                                                  | Reuben Thomas, PhD<br>Michela Traglia, PhD | Two 2-hour sessions<br>Twice a year | Statistics for Biologists<br><br><i>Prerequisites: Basic stats, data analysis, and some R.</i> |
| Linear Mixed Effects Models in R                   | <p>We commonly work on experiments involving repeated measures. Examples include responses measured in cellular assays where the repeated measures may occur over multiple experimental batches, multiple plates, multiple wells within a plate etc. Data derived from behavior of mice assayed over consecutive trials would be another example of experiments involving repeated measures. The distinctions between biological and technical replicates become less clear in these scenarios. Standard statistical tests like t-tests, ANOVA are not applicable. Instead it requires the use of more sophisticated approaches, such as Linear Mixed Effect Modeling.</p> <p>In this course, we will go over the motivation for the use of these models, the underlying description and assumptions behind them, example scenarios and code using R.</p> | Reuben Thomas, PhD<br>Michela Traglia, PhD | 3 hours<br>Twice a year             | Statistics for Biologists<br><br><i>Prerequisites: None</i>                                    |

## Workshop Details

(continued)

| Title                                         | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Speaker                                    | Duration                                   | Topic Area                                                                                          |
|-----------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Pipeline for Linear Mixed Effects Models in R | This workshop will describe and demo a pipeline in R to implement all aspects of dealing with data from repeated measures experimental designs - from specifying the design and model assumptions, to testing these assumptions, identifying outlier observations and outlier groups of repeated measures, sample-size or power calculations to help design follow-up experiments and finally deriving and visualizing estimates of interest. This workshop is a follow up to the Introduction to Linear Mixed Effects Models in R workshop.                                                                         | Reuben Thomas, PhD<br>Michela Traglia, PhD | 3 hours<br>Twice a year                    | Statistics for Biologists<br><br><i>Prerequisites: Linear Mixed Effects Models in R</i>             |
| Introduction to Pathway Analysis              | Join this introduction to functional enrichment analysis and effective visualization. You'll learn about different enrichment analysis tools and methods, and how to use them. You'll also learn how to visualize data in pathways and using GO terms resulting from enrichment analysis.                                                                                                                                                                                                                                                                                                                            | Kristina Hanspers, MSE                     | 3 hours<br>Twice a year                    | Pathways Analysis and Modeling<br><br><i>Prerequisites: None</i>                                    |
| Machine Learning Using R                      | Machine learning is becoming increasingly important not only in data science, but also in bioinformatics. Machine learning gives you the opportunity to find hidden patterns in your data and draw predictive conclusions. In this workshop, you'll learn key machine learning concepts and categories, and get hands-on practice with simple examples. See how machine learning can be applied to advance your research!                                                                                                                                                                                            | Ayushi Agrawal, MS<br>Michela Traglia, PhD | 3 hours<br>Twice a year                    | Machine Learning<br><br><i>Prerequisites: Experience with R</i>                                     |
| AI Fluency for Biologist                      | Whether you're new to AI tools like Claude or ChatGPT or already using them in your research, this hands-on workshop will help you get more out of them.<br><br>You'll learn how to communicate effectively with AI tools through the "4 Ds" framework: Delegation (knowing when AI can help), Description (crafting effective prompts), Discernment (critically evaluating AI outputs), and Diligence (using AI ethically and responsibly). No coding or programming involved. While we'll use Claude for demonstrations, these skills transfer across AI tools and will remain relevant as the technology evolves. | Natalie Gill, MS                           | Two Session<br>2 hours<br><br>Twice a year | AI for Research<br><br>This is an introductory workshop. Absolute beginners are especially welcome! |

## Workshop Details

(continued)

| Title                                                | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Speaker          | Duration                                       | Topic Area                                                                            |
|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|------------------------------------------------|---------------------------------------------------------------------------------------|
| Introduction to Unix command line - parts I and II   | <p>Knowing how to use the command line of your computer can save you a lot of time. In these workshops, you will learn how to navigate the Unix terminal, process files, set permissions, etc. The only thing you need is a Mac or Linux laptop (you can use Windows if you can't access a Linux machine).</p> <p>Part I (Introductory):</p> <ul style="list-style-type: none"> <li>➤ What is the command line?</li> <li>➤ Navigating the file system</li> <li>➤ Installing software</li> <li>➤ Files &amp; permissions</li> <li>➤ Downloading files from the internet</li> <li>➤ Manipulating and searching text files</li> </ul> <p>Part II (Intermediate):</p> <ul style="list-style-type: none"> <li>➤ File Compression/decompression</li> <li>➤ System variables</li> <li>➤ Shell scripting</li> <li>➤ Other helpful commands</li> <li>➤ AWK - advanced text manipulation</li> </ul> | Natalie Gill, MS | <p>Two 2-hour sessions</p> <p>Once a year</p>  | <p>Computer Skills for Data Science</p> <p><i>Prerequisites:</i><br/>None</p>         |
| Working on Wynton                                    | <p>Are you reaching the limits of your laptop? Are you new to working on clusters and high performance computing (HPC)? New to Wynton? This workshop will start with basic definitions for computing in a cluster environment, including types of nodes, queues and containers. We will cover how to transfer files, submit jobs and run analyses. Analyses can be simple scripts, pipelines or custom-built containers.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Natalie Gill, MS | <p>3 hours</p> <p>Twice a year</p>             | <p>Computer Skills for Data Science</p> <p><i>Prerequisites:</i><br/>UCSF account</p> |
| Introduction to R for Data Analysis - parts I and II | <p>The scripting language R is considered one of the most powerful languages for quantitative analysis, statistics, and graphics. This workshop will help you get started using R to analyze your datasets and create graphs for visualization. Hands-on exercises will demystify data analysis using R.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Natalie Gill, MS | <p>Two 3-hour sessions</p> <p>Twice a year</p> | <p>Data Analysis Using R</p> <p><i>Prerequisites:</i><br/>None</p>                    |

## Workshop Details

(continued)

| Title                                              | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Speaker                                                                  | Duration                                  | Topic Area                                                                                                                                 |
|----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|-------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| Introduction to RNA-seq Analysis - parts I and II  | Gene expression is central to cell biology. Disease pathways often involve changes in the expression levels of at least some genes. RNA-seq has become one of the most popular experimental methods to quantify gene expression levels. This 2-day hands-on workshop will provide an introduction to a typical bulk RNA-seq protocol and focus on the data analysis steps for recovering actionable insights.                                                                                                                                                                                                                                                                                                                           | Ayushi Agrawal, MS<br><br>Michela Traglia, PhD                           | Two 3-hour sessions<br><br>Twice a year   | RNA-Seq Analysis<br><br><i>Prerequisites:</i><br>None                                                                                      |
| RNA-seq II: DGE analysis using R                   | The learning objectives for this workshop include 1) how to analyze RNA-seq data from raw gene counts, 2) how to model complex experimental designs and testing hypotheses to draw biological conclusions and 3) how to use generalized linear models implemented in R package edgeR.                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Michela Traglia, PhD                                                     | 3 hours<br><br>Twice a year               | RNA-Seq Analysis<br><br><i>Prerequisites:</i><br><i>Intro RNA-seq</i>                                                                      |
| Single cell RNA-Seq Analysis - parts I, II and III | This 3-session workshop will focus on the analysis of single-cell RNA-Seq data, starting from raw count matrices and progressing to network analysis of single-cell datasets. You'll learn about quality control of the raw data, normalization, feature selection, dimensionality reduction, clustering, finding marker genes, and batch correction. You'll analyze data using R.<br><ul style="list-style-type: none"> <li>➤ Day 1: Loading data, quality control, normalization, feature selection, dimensionality reduction</li> <li>➤ Day 2: Dimensionality reduction (continued), clustering, finding marker genes</li> <li>➤ Day 3: Advanced discussion on normalization, differential analysis, and batch-correction</li> </ul> | Ayushi Agrawal, MS<br><br>Michela Traglia, PhD<br><br>Reuben Thomas, PhD | Three 2-hour sessions<br><br>Twice a year | RNA-Seq Analysis<br><br><i>Prerequisites:</i><br><i>Intro RNA-seq &amp; RNA-seq II: DGE analysis using R</i>                               |
| Advanced RNA-seq Analysis Using R                  | This two session workshop will delve into three concepts/techniques used in the analysis of RNA-seq data. Normalization and batch correction followed by statistical models used for the analyses of RNA-seq data generated under repeated measures scenarios (e.g., longitudinal data) will be covered.                                                                                                                                                                                                                                                                                                                                                                                                                                | Reuben Thomas, PhD                                                       | Two 2-hour sessions<br><br>Twice a year   | RNA-Seq Analysis<br><br><i>Prerequisites:</i><br><i>Intro RNA-seq, RNA-seq II: DGE analysis using R &amp; Single cell RNA-Seq Analysis</i> |

## Workshop Details

(continued)

| Title                                     | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Speaker                                                                         | Duration                                       | Topic Area                                                                                                 |
|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| Introduction to scATAC-seq data analysis  | <p>This 2-session hands-on workshop will provide an introduction to a typical single-cell ATAC-seq protocol and focus on the data analysis steps to address unique computational challenges before the downstream analysis.</p> <p>You'll learn how to perform quality checks of raw data in the typical format provided by sequencing centers, how to process the data, perform dimensionality reduction for visualization, clustering and cell type annotation, generate a gene score matrix to identify cell types, and perform motif and peak enrichment using ArchR.</p> | <p>Ayushi Agrawal, MS</p> <p>Michela Traglia, PhD</p> <p>Reuben Thomas, PhD</p> | <p>Two 2-hour sessions</p> <p>Twice a year</p> | <p>RNA-Seq Analysis</p> <p><i>Prerequisites: Intro and intermediate RNA-seq; scRNA-seq recommended</i></p> |
| scATAC-seq and scRNA-seq data integration | <p>This hands-on workshop will provide an introduction to integrating single-cell ATAC-seq data and single-cell RNA-seq data.</p> <p>Required experience:</p> <ul style="list-style-type: none"> <li>➤ Prior attendance to scRNA-seq, bulk RNA-seq workshops or experience with analysis of transcriptomics data</li> <li>➤ Prior attendance at the Introduction to scATAC-seq data analysis workshop, or experience with analysis of scATAC-seq data</li> <li>➤ Familiarity with R and RStudio (e.g., reading in files, working with lists and dataframes)</li> </ul>        | <p>Ayushi Agrawal, MS</p> <p>Michela Traglia, PhD</p> <p>Reuben Thomas, PhD</p> | <p>One 2-hour sessions</p> <p>Twice a year</p> | <p>RNA-Seq Analysis</p> <p><i>Prerequisites: Intro and intermediate RNA-seq; scRNA-seq recommended</i></p> |

Gladstone Institutes is an independent, not-for-profit research organization affiliated with the University of California San Francisco (UCSF) and a renowned center of excellence focused on applying basic science to disease. Our science contributes to the health and wellbeing of all people through research, education, and outreach in the areas of cardiovascular disease, neurological disease, virology, immunology, and data science.

Gladstone has a vibrant group of labs and core facilities focused on emerging and highly innovative technologies for biomedical research, including stem cells, imaging, single-cell genomics, mass spectrometry, and genome editing.

The Bioinformatics Core provides expertise on experimental design and data analysis of complex data sets, with a specialization on analysis of large-scale data sets acquired from cutting-edge technologies. The core is involved in experimental design, power analysis, sequence analysis, annotation, pathway and network analysis, visualization, data integration, figure generation, and technical writing, in addition to training and workshops on topics in bioinformatics, statistics, and data science.

# GLADSTONE INSTITUTES

© 2025 Gladstone Institutes