NASA GL4U: On-Demand Bioinformatics Training Using Space Biology Omics Data

National Aeronautics and Space Administration

2024 ASGSR Educator Session 3

Amanda M. Saravia-Butler, Ph.D.
Science and GL4U Lead, NASA GeneLab
Contractor: Amentum

Biological & **P**hysical **S**ciences

GL4U Background / Objectives

Background:

- NASA's GeneLab project empowers researchers with open access to space-relevant multi-omics data through the Open Science Data Repository (OSDR)
- GeneLab for Colleges and Universities (GL4U) offers **bioinformatics training for space biology**
- GL4U was conceptualized in 2020 through a collaboration with SJSU and USRA, and the first GL4U bootcamp was held virtually in June 2021

Objectives:

- **Educate and train** the next generation of scientists to process, analyze, and interpret space-relevant 'omics data using publicly available data and bioinformatics tools
- **Maximize** the number of scientists who understand Space Biology data and utilize NASA's open-source 'omics data and tools
- **Provide educators with** the **knowledge and resources** required to train and inspire their students using GeneLab bioinformatic analyses as an entree into Space Biology

GL4U Content Design

GL4U materials are organized into introduction and omics-specific modules

Introduction Module

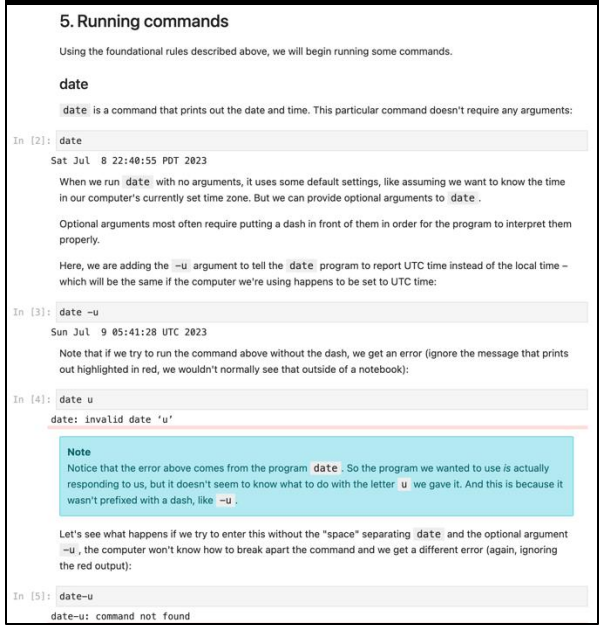
- Overview Lecture (NASA, SMD, BPS, OSDR, GeneLab)
- Jupyter Lab Tutorial
- Unix Jupyter Notebook (hands-on training)
- R Jupyter Notebook (hands-on training)

The Intro Module serves as a pre-requisite for any omics-specific module

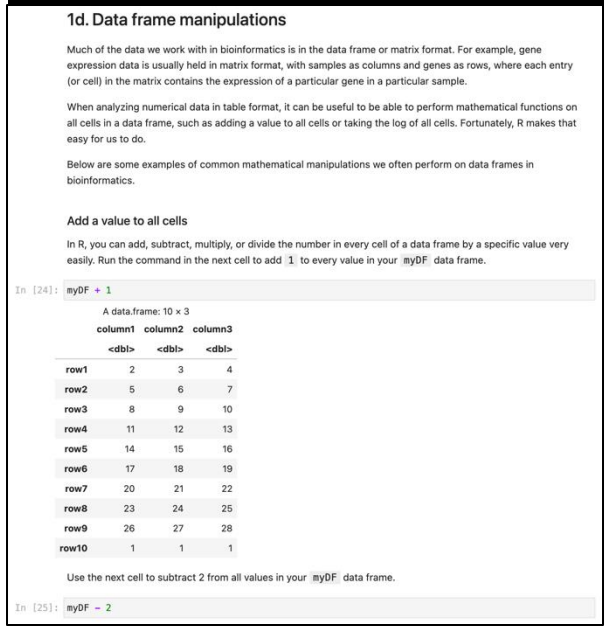
Omics-specific Modules (RNAseq, Amplicon Seq, etc.)

- Omics Data Lectures
  - Experimental design
  - Sample preparation and quality control
  - Data processing tools and visualizations
  - Results analysis and interpretation
- Hands-on data processing and analysis of an OSDR dataset via Jupyter Notebooks (JNs)

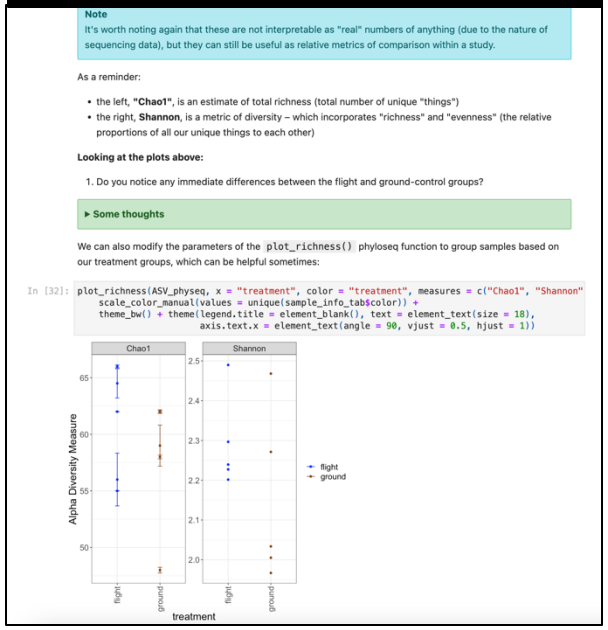
Unix Intro JN



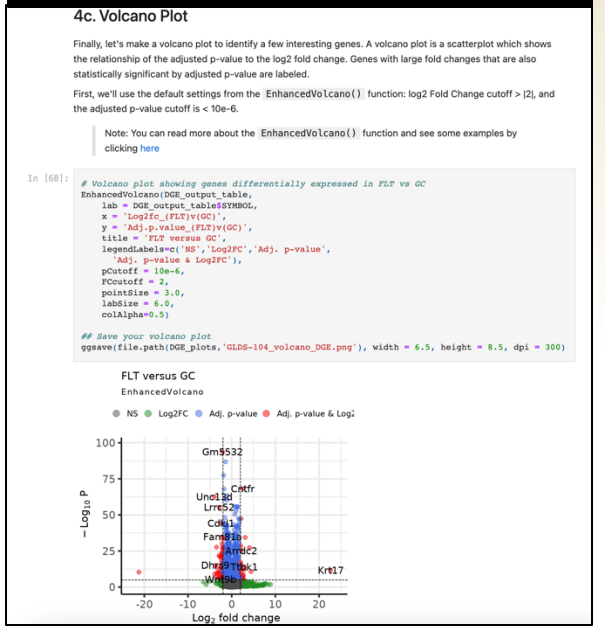
R Intro JN



AmpSeq JN



RNAseq JN



GL4U: Previous Bootcamps

GL4U Direct Approach
Students

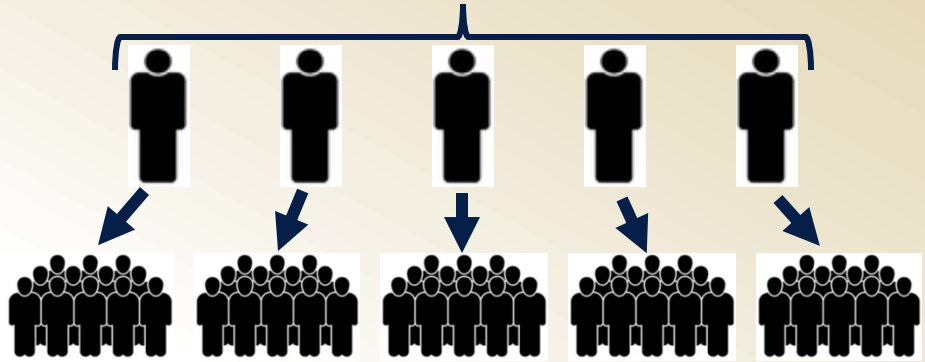
2021 – Space biology & RNAseq

- Virtual 1-week long bootcamp with 14 **SJSU** students
- Collaboration with USRA/SJSU; SJSU compute system
- Space biology- and RNAseq-specific lectures and hands-on instruction using Jupyter Notebooks (JNs)

2023 – Space biology & Amplicon Seq

- In-person 4-day bootcamp with 24 **CSULA** students
- Collaboration with JPL/CSULA; NSF ACCESS compute system
- Space biology- and Amplicon Seq-specific lectures and JNs

GL4U Indirect Approach
Educators



2022 – Space biology & RNAseq

- Virtual ~1.5-week long bootcamp
- Collaboration with JPL; SMCE compute system
- 6 professors and 4 graduate students from **4 Institutions**
- Space biology- and RNAseq-specific lectures and hands-on instruction using JNs
- Resources to teach at home institution

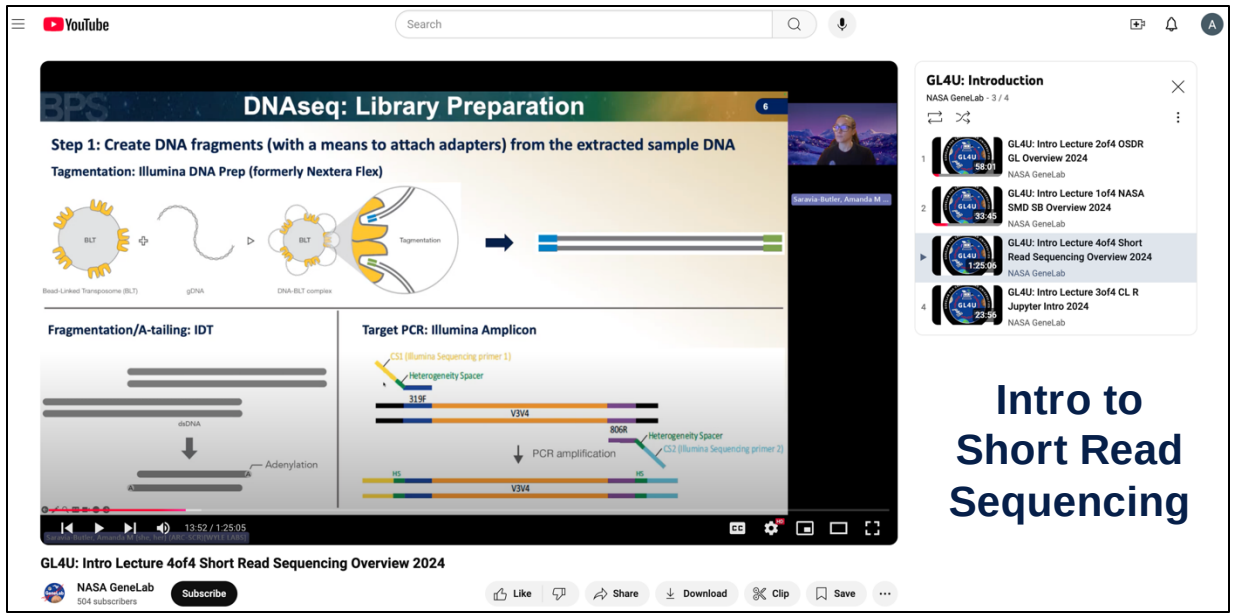
On-demand courses now available!

GL4U On-Demand

GL4U: Intro On-Demand Content

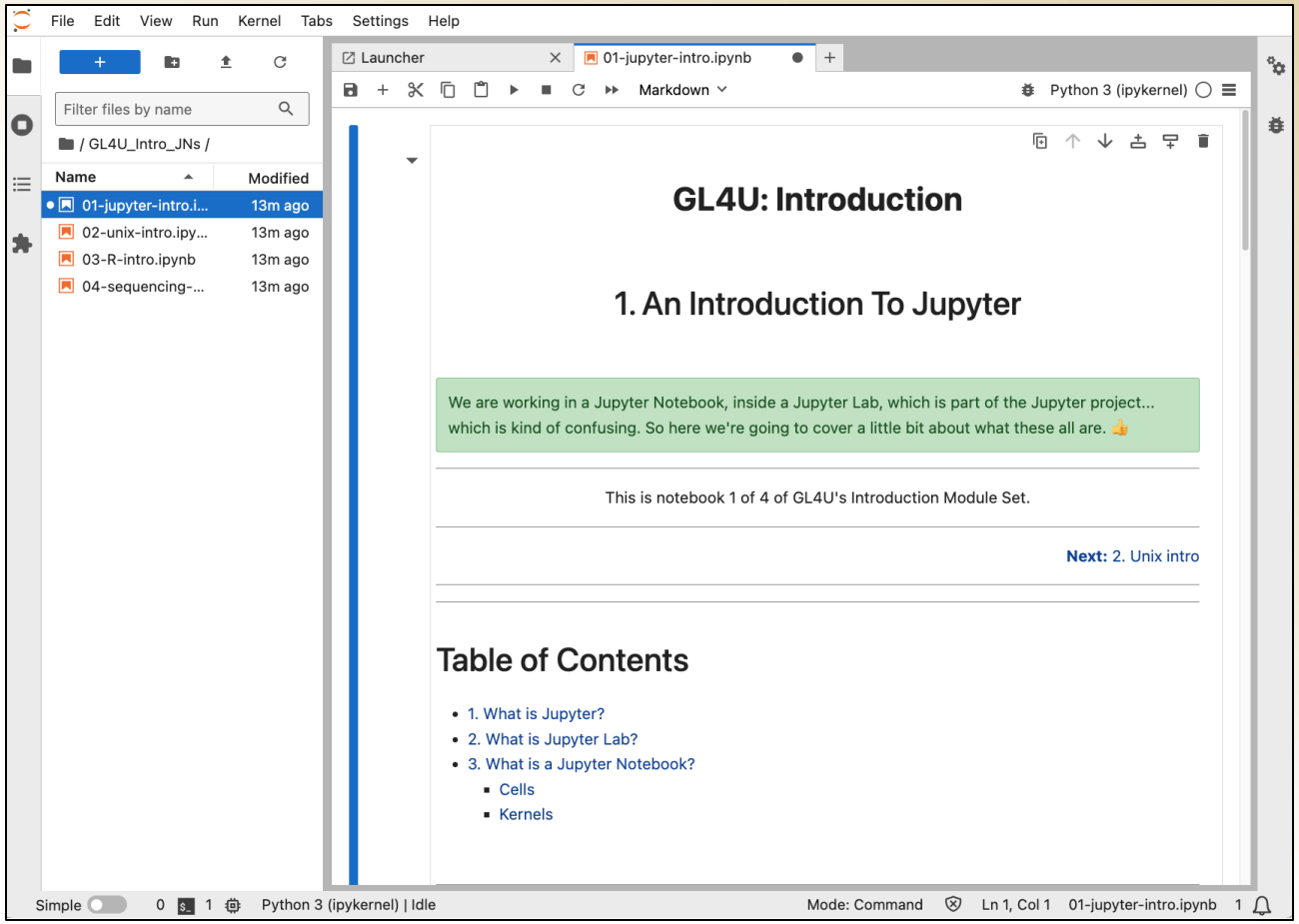
Recorded lectures on YouTube

Intro to NASA, SMD, Space Bio



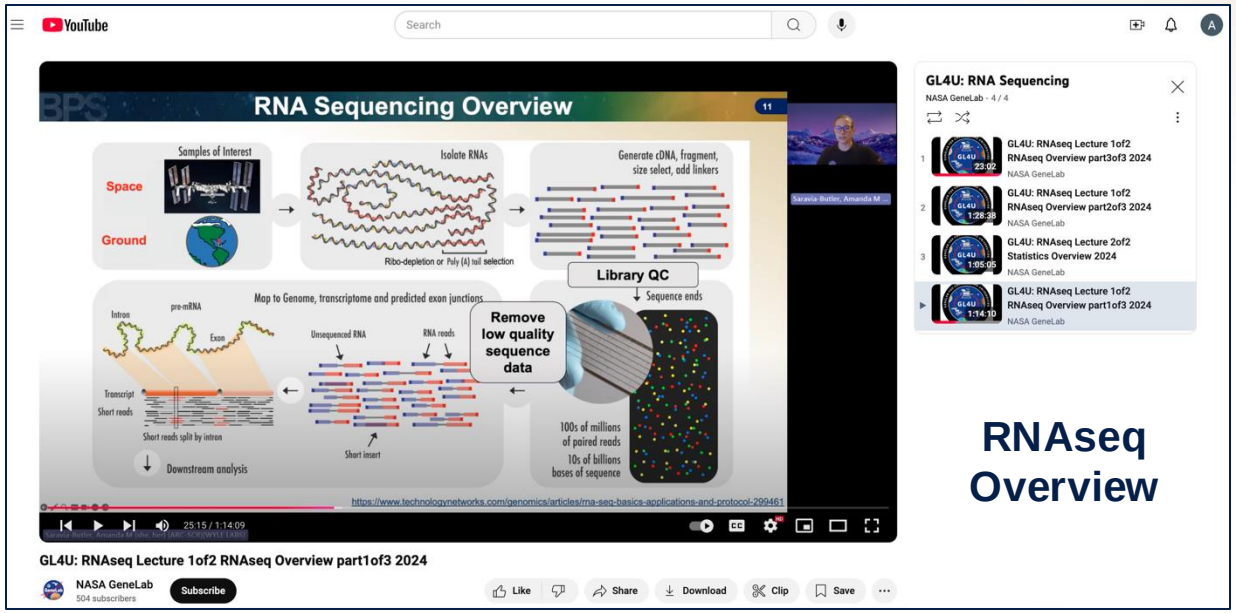
Intro to Short Read Sequencing

Hands-on activities via GitPod

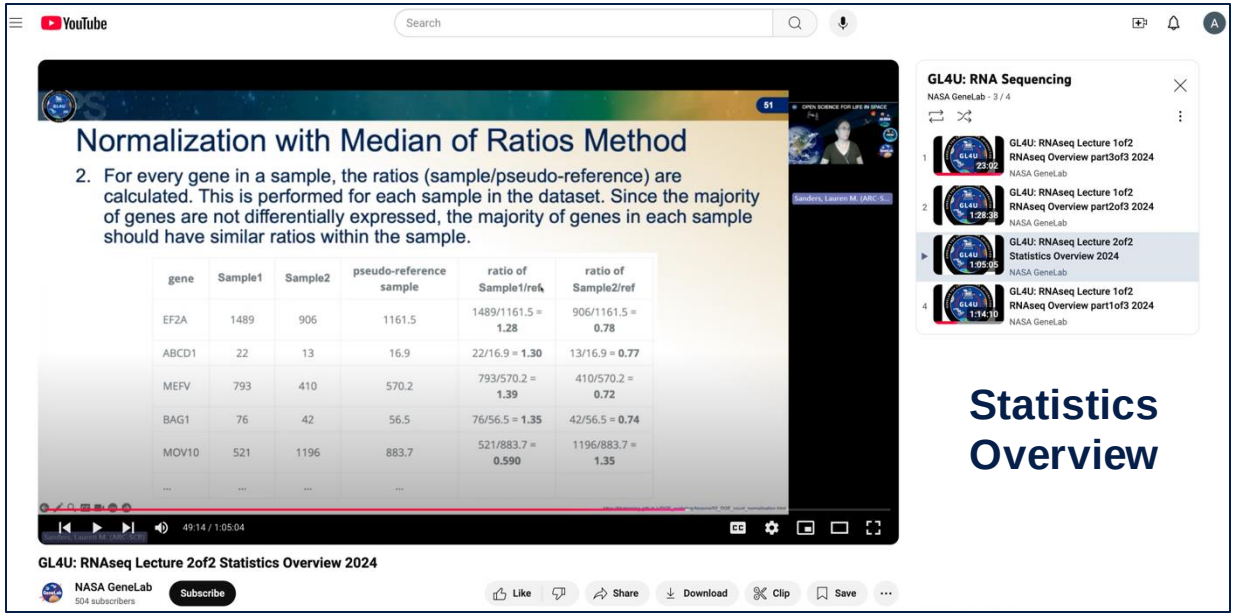


GL4U:RNAseq On-Demand Content

Recorded lectures on YouTube

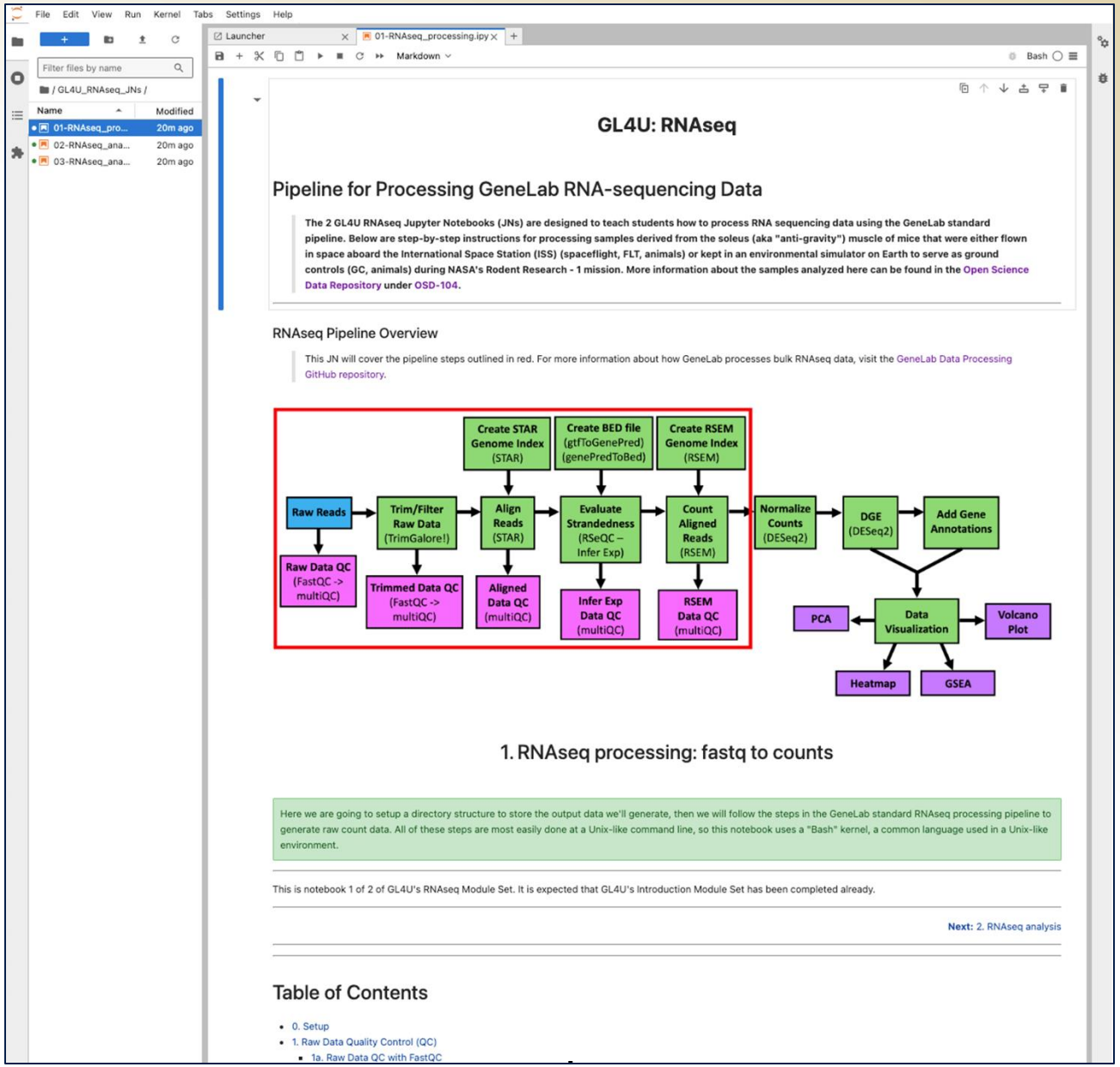


RNAseq Overview



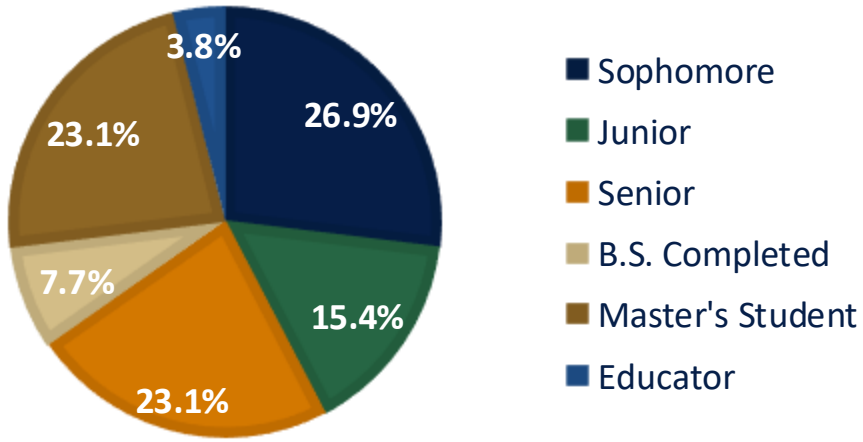
Statistics Overview

Hands-on activities via GitPod



GL4U: RNAseq Virtual Bootcamp, 9/2024

- Virtual 5-day bootcamp, 9/27 – 9/29; 10/4 – 10/5
- 34 students and 2 educators from 4 institutions
  - 14 students, 2 professors from AAMU
  - 10 students from CSULA
  - 1 student from UCSC
  - 9 students from UNM
- *Compute resources*: GitPod free compute resources
- Bootcamp covered GL4U Introduction RNA Sequencing modules

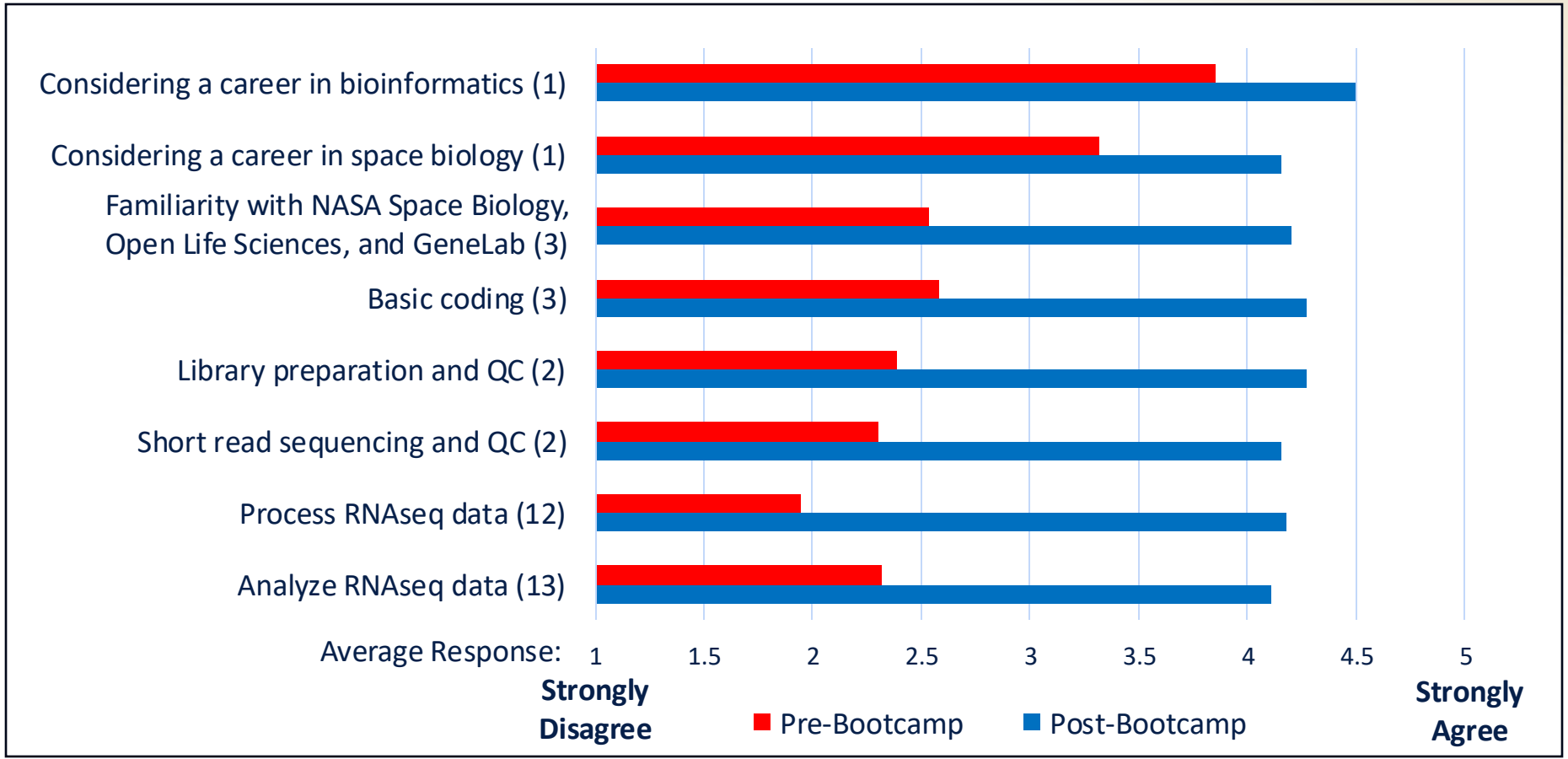


Participant Feedback

"**I learned more about bioinformatics than I have in any other higher education class**. This will greatly improve my comprehension when reading genetics papers. I cannot emphasize enough how valuable this workshop was, **it needs to continue for other students**, and I will absolutely be looking to attend more workshops such as this in the future."

"This was an amazing experience. Amanda was terrific. **It really opened up my eyes to space biology and really showed me this is something I would love to be able to pursue.**"

"When I first took the survey, I was intimidated by some of the topics because I was unfamiliar with them, but **after this bootcamp I have a better understanding of RNA sequencing**. Dr. Saravia has **inspired me to look into space biology careers**, a career field I was not considering prior to this bootcamp."



Certificates of Completion

GL4U: Introduction

GL4U: RNA Sequencing

Open Science Analysis Working Groups (AWGs)

Join an AWG!

ANIMAL
226 members

MULTI-OMICS
351 members

MICROBIAL
216 members

PLANTS
176 members

AI/ML
323 members

ALSDA
(Physiological/BioMedical)

224 members

FEMALE REPRO
88 members

HUMAN
39 members

RADLAB
67 members

Collaborate on Data Mining/Publications

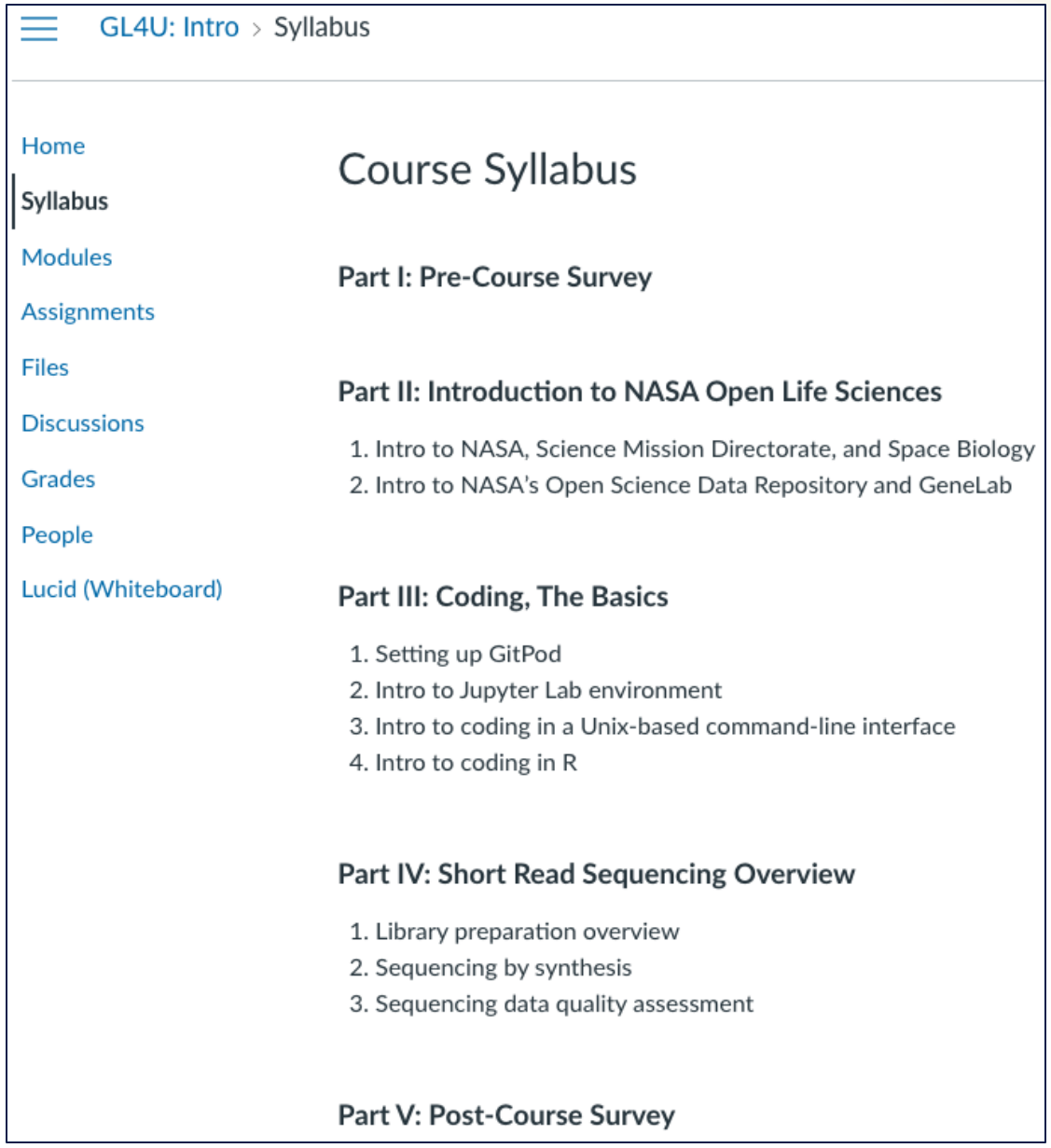
Cell Press Package 2020:
The Biology of Spaceflight
https://www.cell.com/c/the-biology-of-spaceflight

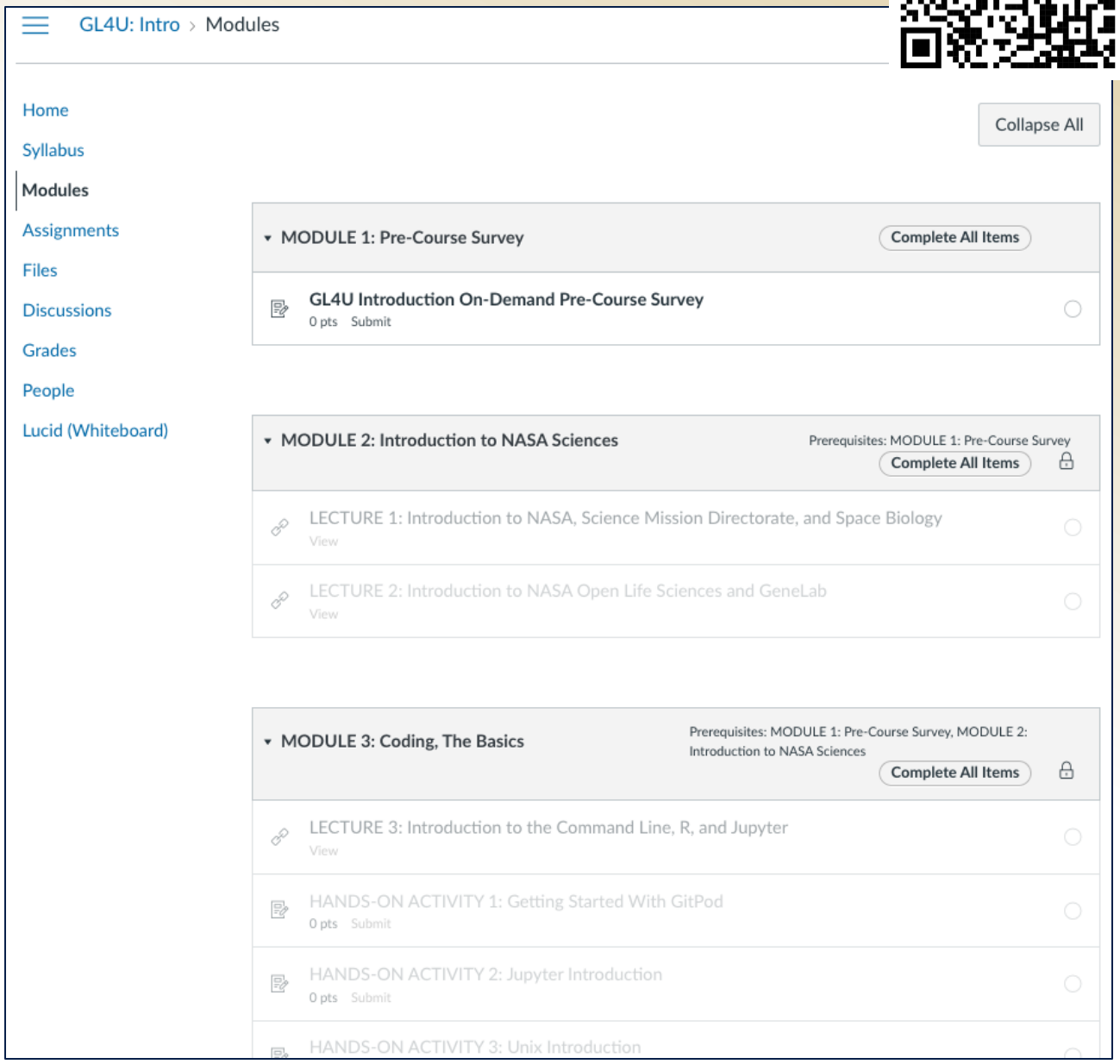
Nature Portfolio Collection 2024:
Space Omics and Medical Atlas across orbits (SOMA)
https://www.nature.com/collections/ebdbcahdgc

AWG Forum User Visits

GL4U: Intro Course Available On Canvas

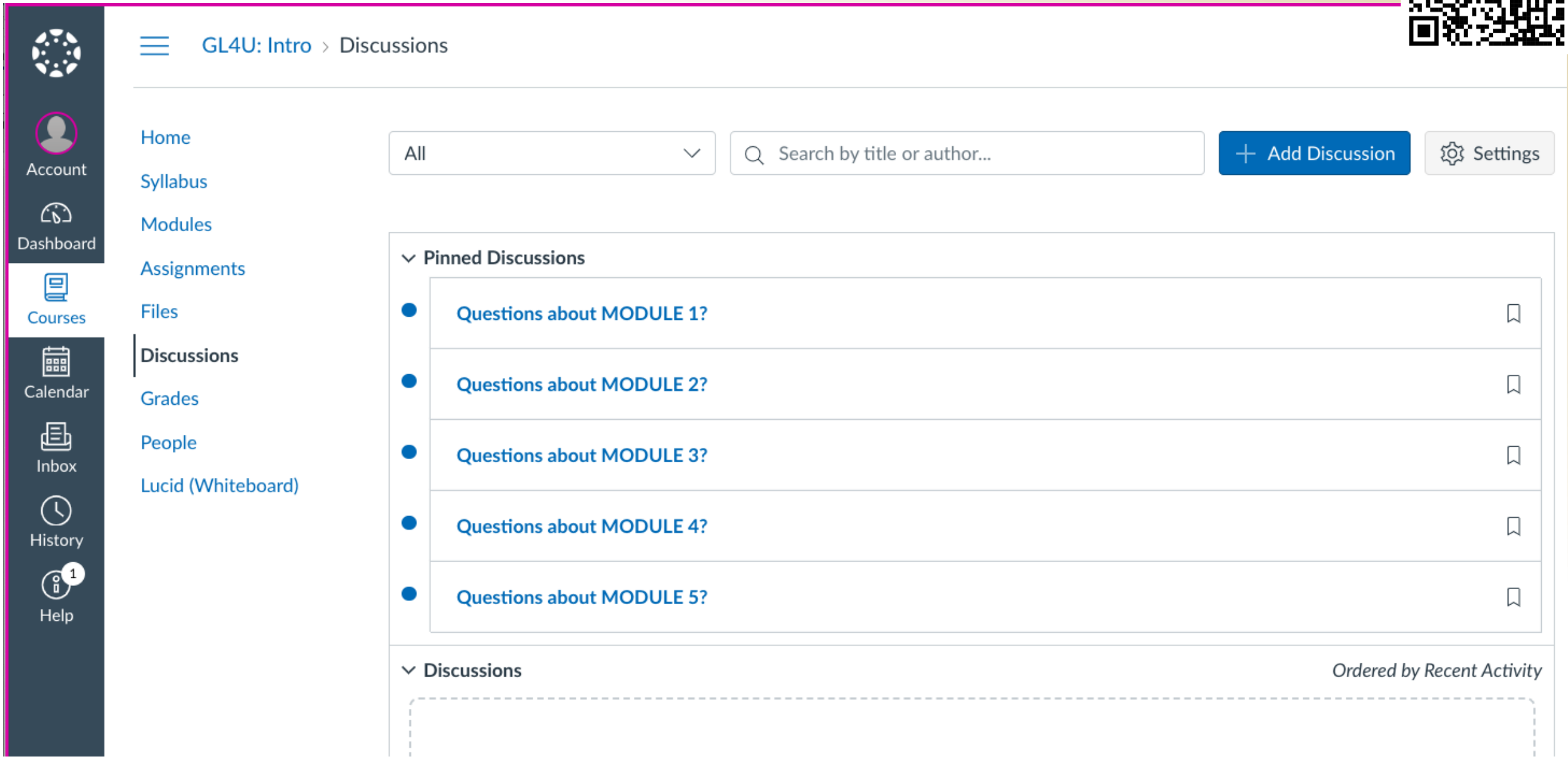
Enroll on Canvas: https://canvas.instructure.com/enroll/63FHH6





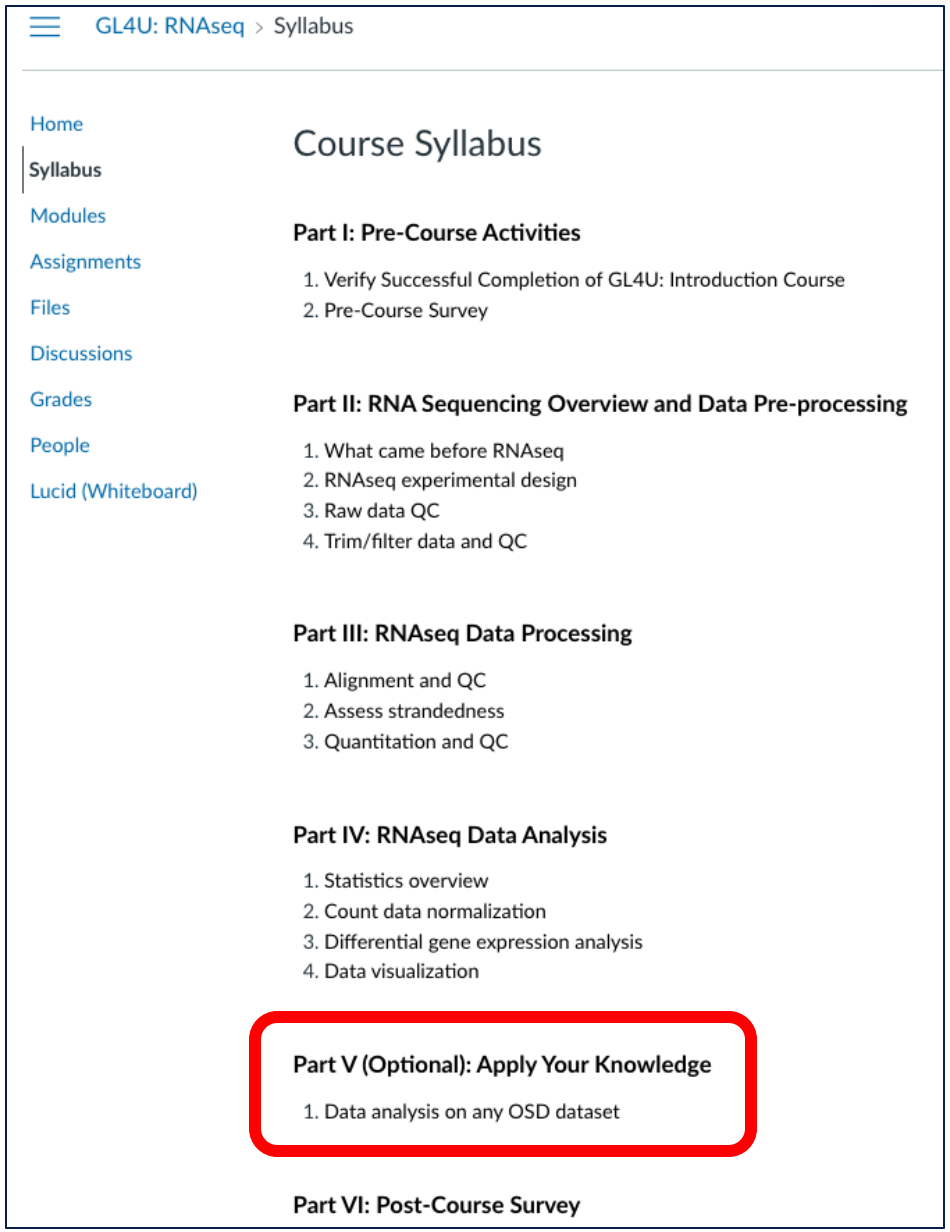
GL4U: Intro Course Available On Canvas

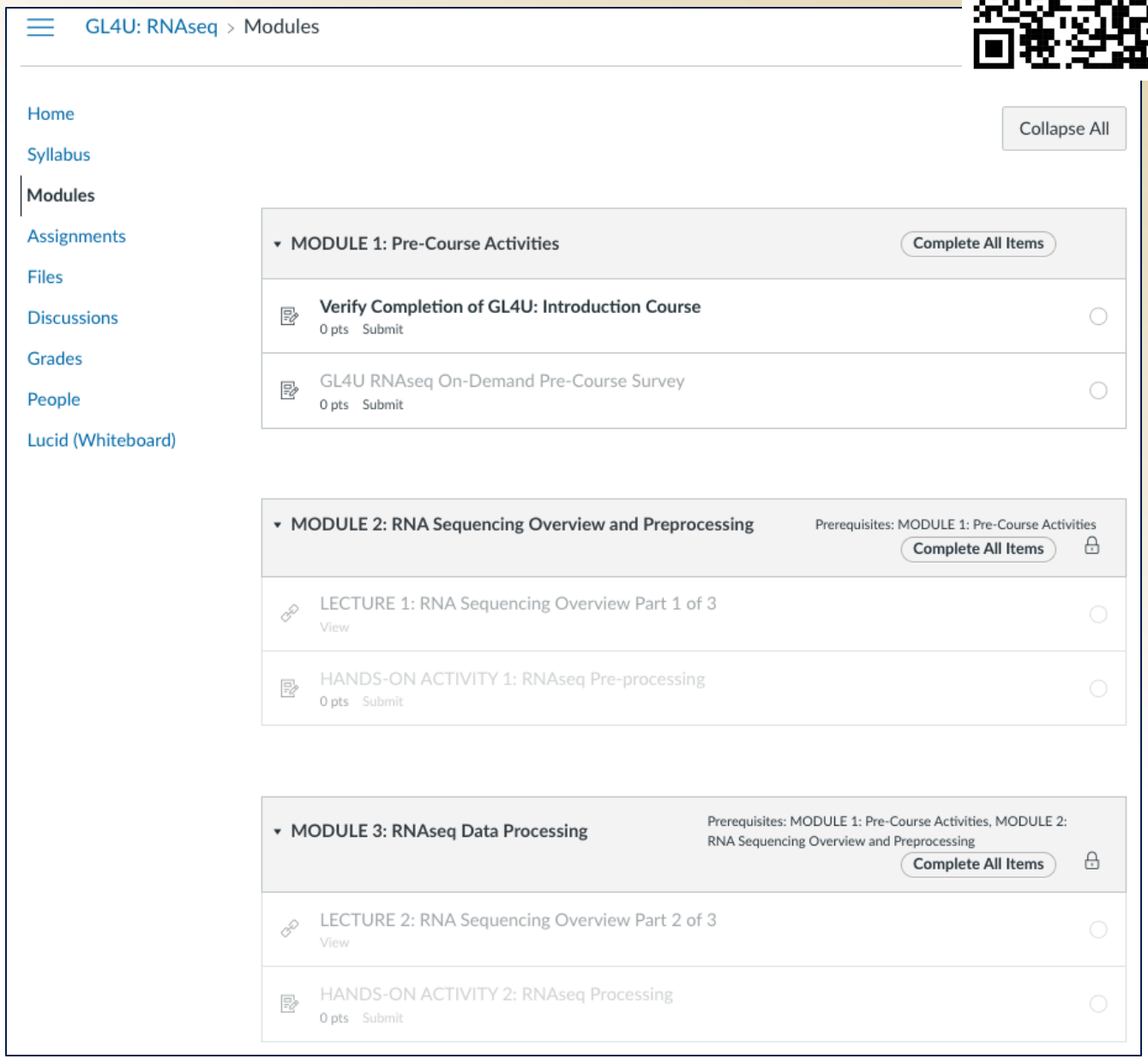
Enroll on Canvas: https://canvas.instructure.com/enroll/63FHH6



GL4U: RNAseq Course Available On Canvas

Enroll on Canvas: https://canvas.instructure.com/enroll/KYXWN6

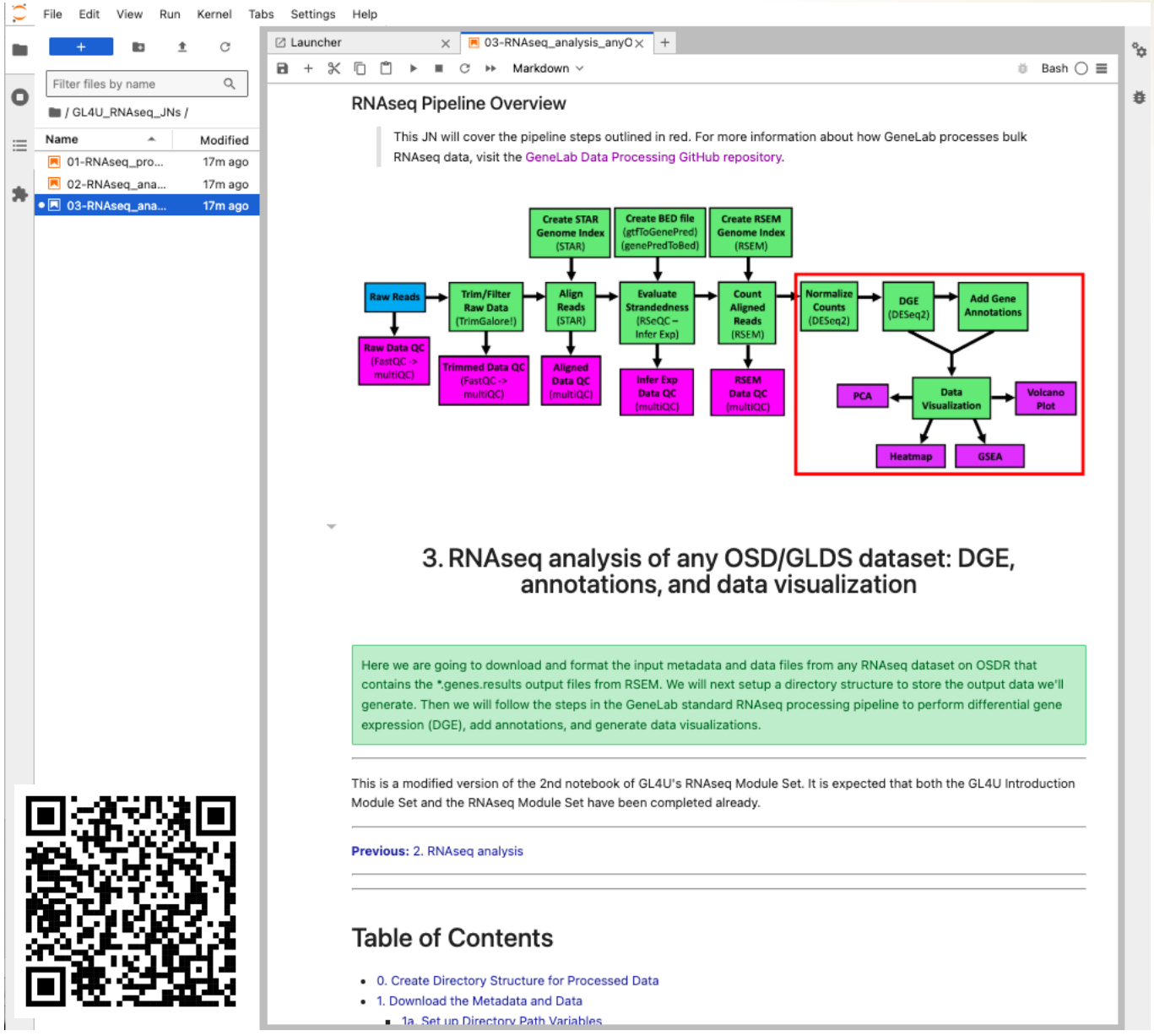


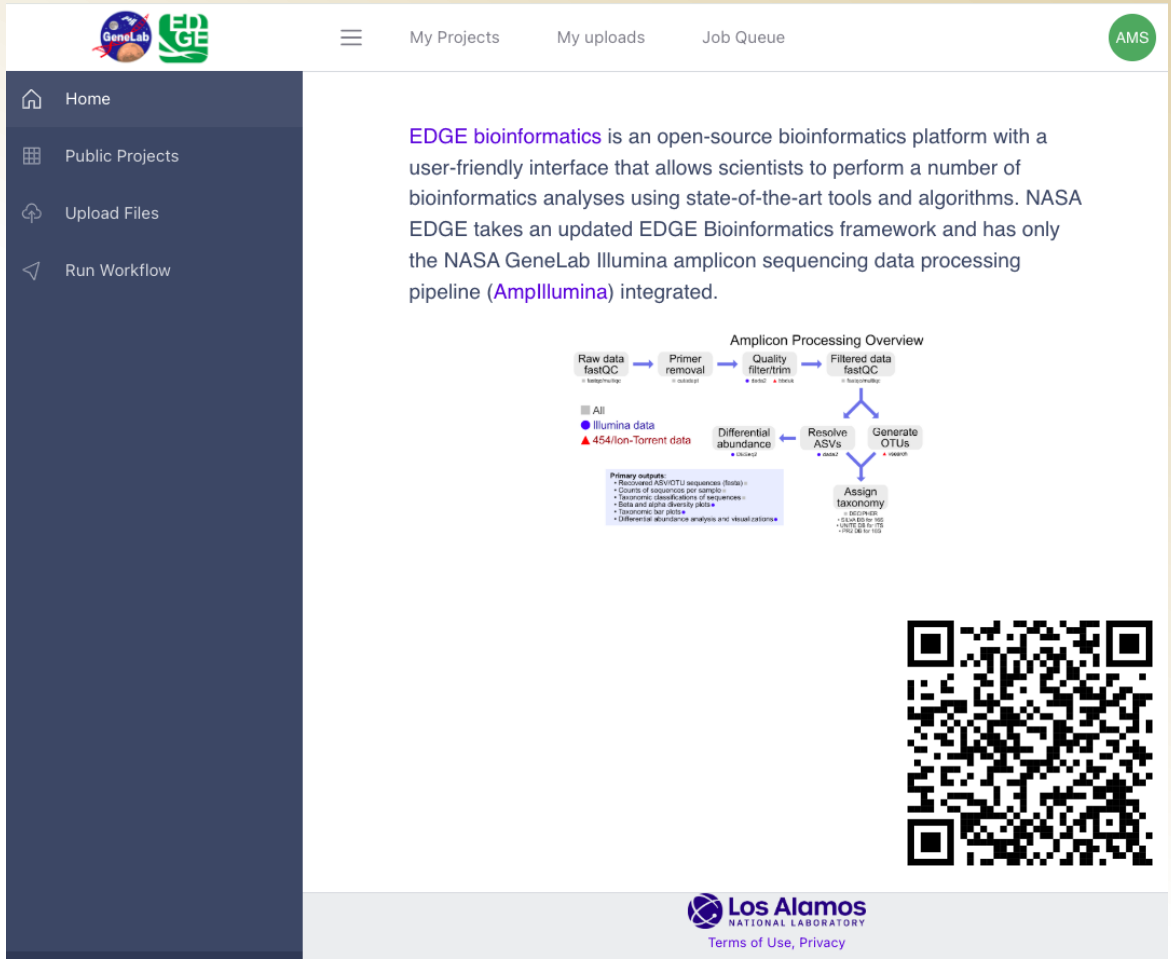


GL4U: RNAseq - Apply Your Knowledge

Enroll on Canvas: https://canvas.instructure.com/enroll/KYXWN6

https://nasa.edgebioinformatics.org/home





- Awarded ROSES F.7: Support for Open-Source Tools, Frameworks, and Libraries (3 years)
  - ➢ **Enhancing Analysis Capabilities of Biological Data With the NASA EDGE Bioinformatics Platform**

What's In Store For GL4U?

Pending funding…

- Video-recorded walk-throughs of hands-on activities
- Automatic assignment grading
- Automatic GL4U certification upon course completion
- Set-up stable Jupyter Lab environment
  - GitPod has a 30min time-out, requiring users to re-run all previous commands upon GitPod restart
- Host live (weekly / biweekly) office hours
- Add additional omics datatypes (e.g. Amplicon seq, Metagenomics, scRNAseq, etc.)
- Expand internationally
- Incorporate NASA EDGE Bioinformatics training
- Continue to host live annual bootcamps

Acknowledgements

Open Science for Life in Space Teams

Open Science Analysis Working Group Members

Support

NASA Space Biology Program
NASA Science Mission Directorate
NASA Human Research Program
NASA Biological and Physical Sciences

Questions?

For more info about GL4U:

Sign up for the GL4U mailing list

- Stay up-to-date on future GL4U events / bootcamps:
  Send and e-mail to GL4U-join@lists.nasa.gov with the Subject: **subscribe**

Enroll in GL4U: Intro On Canvas

Enroll in GL4U: RNAseq On Canvas