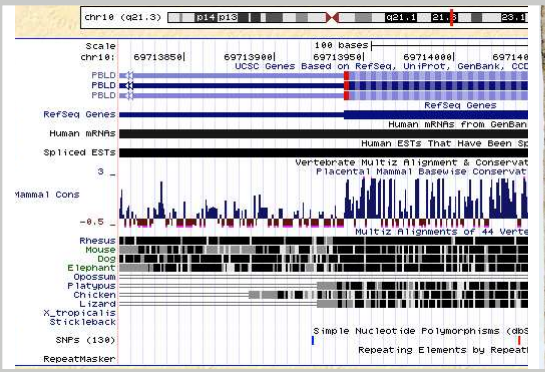
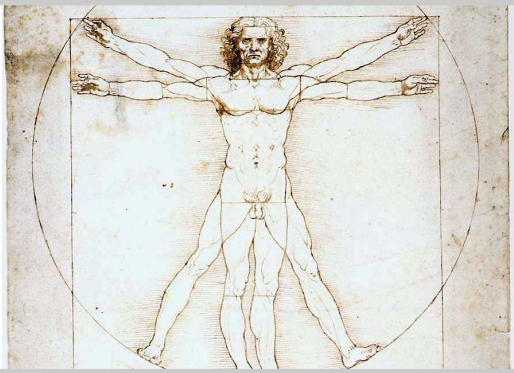
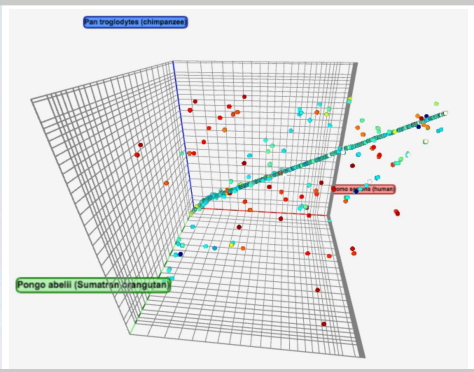
BIOL647
DIGITAL BIOLOGY

Instructor: Rodolfo Aramayo -- Biology







A PRACTICAL INTRODUCTION TO COMPUTATIONAL GENOMICS

Digital Biology is a "Hands-on-Data" experience where, students will learn the fundamentals of GNU/LINUX/UNIX, version control, scripting, and text manipulation. This knowledge will then be used in DNA/RNA Mapping, Assembly, and Quantification

Participants will use a combination of the emerging Extreme Science and Engineering Discovery Environment infrastructure, and the TAMU Supercomputer Ada

DigiBio has No Pre-Requisites

We Welcome the participation of Faculty, Post-Docs and Qualified Undergraduate Students

Limited Places Available

Questions?

Contact: raramayo@tamu.edu Subject: DigiBio

DIGITAL BIOLOGY

BIOL647

Rodolfo Aramayo raramayo@tamu.edu

Syllabus

BIOL647 - DIGITAL BIOLOGY (4 Credits)

Spring 2021

Lectures
Mon, Wed, and Fri
From 12:00hs to 14:00hs
2021/03/22 to 2021/04/23

Zoom Meeting Information

Help Sessions
Mon, Wed, and Fri
From 11:00hs to 12:00hs
2021/03/22 to 2021/04/23

Same Zoom Meeting Information as above

Instructor Information

Name:	Dr. Rodolfo Aramayo, PhD
Email address:	raramayo@tamu.edu
Office hours:	By appointment
Office location:	Department of Biology Room 412A, Biological Sciences Building West (BSBW) Texas A&M University College Station, TX 77843-3258

Course Description and Prerequisites

- *Genomics and Computational Genomics* are revolutionizing science by generating a paradigm shift in our understanding of biology. *Digital Biology* is a command-line-driven introductory course, designed to prepare students for this revolution
- *Digital Biology* aims at developing terminal-based skills to obtain, organize, manipulate, and process biological data, with a focus on mapping and assembly of transcriptome data.
- In this "Hands-on-Data" class, students will learn the fundamentals of version control, scripting, text manipulation, and emerging computational tools available on local computers, super computers, and in the cloud. Learning the fundamentals of Version Control (GIT), scripting (BASH), text manipulation (SED/AWK), as well as other essential Genomics tools including, but not restricted to SAM, BED and Picard, will allow students to ask and answer important biologically relevant questions. It will help them design, perform and analyze experiments using Mapping and Assembling software available at Cyverse and the TAMU Supercomputer Ada.
- Although DigiBio has no formal prerequisites, students should have a good, solid understanding of Genetics and Molecular Biology.
- We expect students to do independent work to supplement their background knowledge as needed, if needed. In addition, we will assume that students are familiar with the basic operational knowledge of computers and willing and able to learn how to use new software.
- Importantly, students must have a strong desire to understand biology, a fearless attitude towards computers, a strong work ethics and an open mind.
- **Students will be required to use computers extensively throughout the semester.** If you are not comfortable with computers, **do not take this course.**
- *Digital Biology* is a mature course that is being offered since 2012 and has been evolving ever since.

Learning Objectives

- Students will learn the use and logic of a set of computational tools required to manipulate genome data.
- In the end, students are expected to understand how to use genomics and genomic techniques to answer important biological questions.
- After completing this course students will be able to:
  - Master the fundamentals of Unix, including BASH Scripting

- Master the fundamentals of Version Control (Git and GitHub)
- Understand Next Generation Sequencing data
- Understand the basis of NGS data mapping
- Understand the basis of Transcriptome Assembly and Transcript level Quantification
- Understand the basics of data display and analysis
- Understand the fundamentals of experimental design
- Learn the fundamentals of both **R** and **RStudio**
- Master the use of the Supercomputer Ada

Grading Policies:

- The course will be letter graded on a curve with a median grade being somewhere in the B's
- Grades will be based on:

Assignments	Points
Attendance:	60
Participation:	40
Class Work:	600
GitHub Repository	300
Total:	1000

- Attendance (60 Points)
  - **Attendance is required.**
  - Attendance will be taken at the start and end of each class session.
  - Attendance to each class session will be graded (05 Points/Session, 05 x 12 Sessions = 60 Points).
- Participation (40 Points)
  - More than ever, participating is important as it demonstrates your active involvement in the topics being discussed.
  - Asking question, making suggestions and/or helping fellow students is both encouraged and will be rewarded.
- Class Work (Total 600 Points):
  - We will follow your class work closely. The commands you issued, the scripts you generated, the files you processed.
  - We will be paying special attention to the organization and naming of your computer files, scripts and directories.

  - We will assign small mini-projects throughout the course. Your ability to complete them on time and successfully will be graded. The value of each assignment in points will vary according to the task.

- GitHub Repository (Total 300 Points)
  - We will also follow your class work closely through the use of a GIT repository that will be assigned to you.
  - Your class work should be documented and submitted to your GIT repository at the end of each class meeting.
  - The instructor will then grade your daily submission, which should be properly documented and organized.
  - The GIT submission should include any Assignments and/or other work/scripts requested by the instructor on that particular day
  - All Scripts constructed for each topic will be graded.

Other Pertinent Course Information:

- Class meetings will take place according to the schedule below.
- Students are expected to arrive on time. Being late for class will be penalized.
- Students will be expected to actively participate and perform assigned computer work as the lecture develops.
- Students work will be reviewed on an one-on-one basis by the instructor, as the class progresses.
- **If you do not understand something, Ask.**

Course Topics and Other Important Dates:

2021-S01B (Spring): Schedule

Date	Day	Topic(s)
03-22	Mon	Introduction Accounts SetUp Introduction to GIT
03-24	Wed	Introduction to Bootstrapping Dependencies in Computational Biology Introduction to SSH Introduction to Emacs, and Org-Mode
03-26	Fri	Unix101 Unix102
03-29	Mon	Unix103 Unix104
03-31	Wed	Introduction to Next Generation Sequencing Introduction to Databases: NCBI Gene, Gene Models, and Genome Files
04-02	Fri	Introduction to Genome Browsers Genome Tools 101: Mapping and Displaying DNA NGS Data Genome Tools 102: Mapping and Displaying RNA NGS Data
04-05	Mon	Genome Tools 103: SAM Tools
04-07	Wed	Genome Tools 104: BED and Picard Tools
04-09	Fri	Ada: Introduction to Ada Ada: NGS DNA Assembly 101 Ada: NGS DNA Assembly 102
04-12	Mon	Ada: NGS RNA Mapping/Assembly/Quantification 101
04-14	Wed	Ada: NGS RNA Mapping/Assembly/Quantification 102
04-16	Fri	Project Work
04-19	Mon	Project Work
04-21	Wed	Project Work
04-23	Fri	Final Report Due at 18:00hs (Final GIT Pull)

Americans with Disabilities Act (ADA)

"The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities"

If you believe you have a disability requiring an accommodation, please contact Disability Services, currently located in the Disability Services building at the Student Services at White Creek complex on west campus or call 979-845-1637

For additional information, visit TAMU.Disability

Statement on Academic Integrity (AI)

"An Aggie does not lie, cheat, or steal, or tolerate those who do"

For additional information please visit: TAMU.Honor

TAMU Critical Incident Response Team

Step In. Stand Up Campaign

Author:

- Rodolfo Aramayo (raramayo@tamu.edu)
