BENG 183. Applied Genomic Technologies

Syllabus

Useful links:

- Course **website**
- Course **calendar**

Instructor: Sheng Zhong, Ph.D. Professor of Bioengineering.
Find me after class: 1:50-2:20 pm after every in person lecture.

This course will use a combination of in-person lectures and pre-recorded videos
In-person lectures: TuTh 12:30-1:50pm
Classroom: CENTR 214
In-person lectures can be attended on zoom. Sign-in to Zoom in required:
Meeting ID: 926 7042 4051

Discussion: Discussions will be held on zoom and will be recorded. Please ignore the discussion time on Blink and use TA's assigned time.

TAs' notes on Discussion sessions: As specified on the syllabus, we will hold all discussion sessions on zoom at the same time every Friday (from 1:30 to 2:30 PM). The link for each session will be posted here by each Thursday end of day. Below is the list of topics we'll cover in each session:

Discussion 1 - Linux (10/04), Please use this zoom link to join the session.
Discussion 2 - DNA/RNA/CHIP (10/11)
Discussion 3 - RNAseq (10/18)
Discussion 4 - MIDTERM 1 REVIEW (10/25)
Discussion 5 - Genome interaction + Prec Med (11/01)
Discussion 6 - ML & cancer classification (11/08)
Discussion 7 - MIDTERM 2 REVIEW (11/15)
Discussion 8 - ML & cancer classification, cont. (11/22)
Discussion 9 - Random Variable & Hypothesis Testing (12/06)

TAs and emails:

- Jonathan Kirkland <jokirkland@ucsd.edu>
- Yotam Voskoboynik <yvoskoboynik@ucsd.edu>
- Vicente Fajardo Rosas <vfajardorosas@ucsd.edu>
- Gaoyuan Li <gal001@ucsd.edu>

TA office hours:

- Gaoyuan Li: (Wed 2-3pm), (Zoom link)
- Jonathan Kirkland: (Thurs 10:30-11:30, (Zoom link)
- Yotam Voskoboynik: (Tues 10-11am), (Google meet)
- Vicente Fajardo Rosas: (Mon 1-2pm), (Zoom link)

Prerequisites

BIMM 100 or CHEM 114C, or consent of department.

Course Description

This course introduces genomics technologies and their biomedical applications. Students will learn and genome sequencing, RNA sequencing, and the computational methods for analyses of these data. Probability theory and applications and the basics of machine learning methods will be introduced. Students will be able to formulate hypotheses and perform statistical tests. Students will be able to design gene expression analysis experiments, transform next-generation sequencing data into a summary data matrix, identify differentially expressed genes, perform K-nearest neighbor classification, and perform hierarchical and K-means clustering. Contemporary research developments on personalized medicine and liquid biopsy analyses will be discussed.

Schedule

This course has six modules. This course will have a mixture of in-person lectures and pre-recorded lectures. Please only come to the classroom when a lecture is marked "in person".

Pre-recorded lectures will be posted on the course website on or before the lecture date and time. Please do not come to the classroom for the pre-recorded lectures.

The lecturer has requested the in-person lectures be podcasted.

- Module 1: DNA sequencing, RNA sequencing, and ChIP sequencing
    - Lecture 1 – 4: 9/26: In person.
    - 10/1: Pre-recorded lecture.
    - 10/3: In person.
    - 10/8: pre-recorded lecture.
- Module 2: Analysis of RNA-seq data
    - Lectures 5 – 7: 10/10: Pre-recorded lecture.
    - 10/15: In person.
    - 10/17: Pre-recorded lecture.
- Module 3: Genome interaction, RNA interactions and a primer to precision medicine
    - Lectures 8 – 9: 10/22, 10/24: Pre-recorded lectures.
    - Lecture 10 is an optional review. The slides are provided for your potential interest. No class time is assigned to Lecture 10.
- Midterm 1: 10/29, take home.

Note: in 2024, we will cover Module 5 before Module 4.

- Module 5: A primer to machine learning and application to cancer.
    - Lectures 14 – 15: 10/31, 11/5: In person.
- Module 4: Random variable and hypothesis testing
    - Lectures 11-13: 11/7, 11/12, 11/14: Pre-recorded.
- Midterm 2: 11/19, take home.
- Module 6: Group presentations.
    - Sessions 16 – 18: 11/21, 11/26, 12/3, 12/5. In person.

Avoid repetitive classes

Different departments build their own bioinformatics track curriculum, and thus the courses offered from different departments can have overlapping contents. If you feel the contents of

this course are repetitive with any other course(s) you have taken, you can ask your department for permission to waive this course.

Course Calendar

Please check out the course calendar from the course website.

Textbook and Other Materials

See Course Outline.

Additional Materials

In addition to the textbook listed above, there is a plethora of available information on the Internet. This information includes peer-reviewed manuscripts, Wikipedia articles, YouTube videos of lectures from UCSD and/or other universities, and much more. A course handout will be provided. Students are strongly encouraged to review additional online materials prior to each lecture.

Course Outcomes

After successfully completing this class, a student will be able to:

1. Understand the technologies of Sanger sequencing, sequencing by synthesis, RNA sequencing, and extracellular-RNA sequencing.
2. Design genomic technologies to identify gene expression biomarkers.
3. Apply probability theory to solve real-life problems.
4. Formulate statistical hypotheses and perform hypothesis testing.
5. Identify differentially expressed genes from gene expression data.
6. Cluster data from a data matrix.
7. Perform classification of data from a data matrix.
8. Analyze RNA sequencing data with bioinformatic software.