Physiology Biophysics and Systems Biology Program
Quantitative Information in Biological Systems

Course Title:
Quantitative Information in Biological Systems

Course Director:
Luce Skrabanek

Credit Hours:
15 credits (8 Fall, 7 Spring)

Office Hours and Contact Information
Luce Skrabanek las2017@med.cornell.edu
Office hours by appointment

Course Description
This core course is a year-long modular course which will integrate statistics, experimental data, quantitative methods, and interpretation of results, each within the context of a single topic or idea. Emphasis will be placed on depth of quantitative understanding. Each module, anchored on one or a few papers from the primary literature, will explore a topic in-depth, providing students with the theoretical and biological background necessary to understand the topic, the experimental techniques of data acquisition, as well as the practical aspects of approaching and analyzing the data.

Course Objectives
After completing this course, students will be able to:

- Have a deep appreciation of current DNA sequencing technologies, with an awareness of pitfalls, caveats, and confounding factors;
- Understand which technologies are appropriate for which use cases;
- Be aware of the details in deriving insights from raw data;
- Be able to critically assess next generation sequencing data and analyses, and be aware of common biases.

Course Frequency and Schedule:
Core course offered yearly to all 1st year PBSB and CBM students, and 2nd year students from other Programs.
Tuesday/Thursday 10am-12pm, Friday 12-2pm

Text and Materials
Lectures and associated readings will be made available on the course website.

Teaching Format

This course will be held in person, with Zoom availability for students in Houston.
The course is based on lectures, group problem sets, journal paper, and discussion sessions.

Method of Evaluation

This class will be graded according to the usual WCGS scale (Honors, High Pass, Low Pass, Fail). Grading will be based on class participation, deliverables from the workshops, take-home assignments and/or essays.

Grade	Criteria
Honors	Greater than Mean plus 2 standard deviations
High Pass	Equal to Mean plus/minus 1 standard deviation
Low Pass	Less than Mean minus 1 standard deviation
Fail	Less than Mean minus 2 standard deviations

Student Learning Outcomes

By the end of the course student will have an understanding of:

- Next Generation Sequencing technologies (Module 1)
- Variant calling (Module 2)
- Transcription regulation and epigenetics (Module 3)
- Metabolomics and proteomics (Module 4)
- Image analysis (Module 5)
- Single cell technologies (Module 6)

Academic Integrity

Students are expected to abide by the Weill Cornell Graduate School Code of Academic Integrity.

Teaching Assistants

Teaching Assistants will be recruited for Modules 2-5.

Teaching Assistant Duties:

- Conduct review sessions of classroom lectures
- Preparing and conducting lab/practicum sessions
- Being available to students by holding office hours
- Grading homework

Schedule

Fall 2024

Module 1: Introduction to Next-Gen Sequencing (5 weeks)

Includes an introduction to the theory, the biological background, the different short- and long-read sequencing technologies, and library preps, as well as the diverse applications of NGS. Introduces students broadly to the topics that will be covered in more detail in each module.

Module 2: DNA Sequencing (5 weeks)

Point mutation and indel detection, identification of structural variants, copy number detection, functional interpretation, meta-genomics analysis, and the application of variant calling in population genetics and GWAS.

Module 3: Transcriptional Regulation and Epigenetics (5 weeks)

How NGS can be used to explore DNA structure and chromatin accessibility, and transcriptional regulation, through techniques such as ATAC-seq, ChIP-seq , RNA-seq, PRO-seq, methylation assays, and perturbation analysis.

Spring 2025

Module 4: Metabolomics and Proteomics (4 weeks)

How assaying the metabolome and proteome differ from NGS, and how this information differs from, and complements, the quantification of nucleic acid-based technologies.

Module 5: Image Analysis (5 weeks)

Details image acquisition modalities and noise sources, spatial transcriptomics (with ML), active contouring, classic computer vision, particle tracking, source separation and image registration, thinking in k-space, deep learning approaches for pathology images.

Module 6: Single Cell Technologies (7 weeks)

Looks at the extra information that measurements from single cells afford over bulk measurements, sample and library preps, what are the advantages and disadvantages of single cell sequencing, analysis techniques like normalizing, dataset integration, clustering, trajectory inference, lineage tracing, how we can identify even rare cell types, and how that has transformed the way we look at the cells and how we explore diseases and develop therapies. Advanced techniques such as CITE-seq, scATAC-seq, Perturb-Seq, multi-modal approaches.