Computational and Quantitative Biology @UC Berkeley

Graduate courses

CCB293 / MCB 296: Doctoral seminar in Computational Biology (Syllabus**)**
This one-year interactive seminar builds skills, knowledge and community in computational biology for first year PhD and second year Designated Emphasis students. Topics covered will range from human evolutionary genomics, single cell sequencing, microbiome and plant genomics, with emphasis on statistical and computational methods and hands-on data analysis. Faculty members at UC Berkeley will participate to highlight the diversity of research topics in computational biology at Berkeley.
Instructors: Priya Moorjani (moorjani@berkeley.edu) and Ashley Wolf (awolf@berkeley.edu)

MCB280D: Genomics and Bioinformatics (Syllabus)
This course covers genomic and bioinformatic approaches for addressing questions in molecular and cell biology. It focuses on interpreting and critically evaluating large-scale data in biology, where computational and statistical concerns arise. Topics include bioinformatics and sequence analysis, functional genomics, expression profiling, genome-wide analysis of chromatin state, and other genome-scale and high-throughput experimental approaches.
Instructor: Nicholas Ingolia (ingolia@berkeley.edu)

MCB 288: Data Science for Molecular and Cell Biology (Syllabus)
Data science is rapidly becoming a critical skill for molecular and cell biologists. This course provides a survey of data science concepts and methods, including practical statistical inference and modeling, data visualization and exploration, elementary machine learning, and simulation. The course is practically oriented. Diverse real-world datasets, along with simulated data, will be used to develop skills and intuition.
Instructor: Daniel Rokhsar (dsrokhsar@gmail.com) and Michael Eisen (mbeisen@berkeley.edu)

MCB137 / 237: Physical Biology of the Cell (Syllabus)
The objective of this course is to respond to this deluge of quantitative data through quantitative models and the use of biological numeracy. The course will explore the description of a broad array of topics from modern biology using the language of physics and mathematics. One style of thinking we will emphasize imagines the kinds of simple calculations that one can do with a stick in the sand. We will draw examples from broad swaths of modern biology from our department and beyond including cell biology (signaling and regulation, cell motility), physiology (metabolism, swimming), developmental biology (patterning of body plans, how size and number of organelles and tissues are controlled), neuroscience (action potentials and ion channel gating) and evolution (population genetics) in order to develop theoretical models that make precise predictions about biological phenomena. These predictions will be tested through the hands-on analysis of experimental data and by performing numerical simulations using Python.
Instructor: Hernan Garcia (hggarcia@berkeley.edu)

INTEGBIO 201/120 COMPBIO 210: Introduction to Quantitative Methods in Biology ([Syllabus](#))

This course provides a fast-paced introduction to a variety of quantitative methods used in biology and their mathematical underpinnings. While no topic will be covered in depth, the course will provide an overview of several different topics commonly encountered in modern biological research including differential equations and systems of differential equations, a review of basic concepts in linear algebra, an introduction to probability theory, Markov chains, maximum likelihood and Bayesian estimation, measures of statistical confidence, hypothesis testing and model choice, permutation and simulation, and several topics in statistics and machine learning including regression analyses, clustering, and principal component analyses. The course includes a lab section focusing on building student skills in modern computational methods for biological data analysis using python and R

Instructors: Denis Titov (titov@berkeley.edu), John Huelsenbeck (johnh@berkeley.edu), Peter Sudmant (psudmant@berkeley.edu), Rasmus Nielsen (rasmus_nielsen@berkeley.edu)

Bioengineering 145 / 245: Introduction to Machine Learning for Computational Biology ([Syllabus](#))

Genome-scale experimental data and modern machine learning methods have transformed our understanding of biology. This course investigates classical approaches and recent machine learning advances in genomics including:

1) Computational models for genome analysis
2) Applications of machine learning to high throughput biological data
3) Machine learning for genomic data in health

This course builds on existing skills to introduce methodologies for probabilistic modeling, statistical learning, and dimensionality reduction, while grounding these methods in understanding genomic information.

Instructor: Liana Lareau (lareau@berkeley.edu)

EECS: Algorithms in Computational Biology ([Syllabus](#))

Instructors: Yun Song and Nilah Ioaniddis

STAT C245F / PB HLTH C240F: Statistical Genomics II ([Syllabus](#))**:** This course, STAT C245F/PB HLTH C240F (Statistical Genomics II), is one of the two courses offered by Statistics department and Biostatistics group on statistical genomics that concern statistical and computational methods for addressing inference problems that arise in genomic research. Particularly, this course will cover the statistics of sequencing data analysis. We will start with traditional DNA sequence analyses including assembly, alignment, BLAST and signal detection in DNA. We next will focus on (bulk and single cell) RNAseq data analysis. We will cover manifold learning and deep learning technologies that have facilitated the adoption of single-cell sequencing. We will discuss third-wave AI technologies that are transforming the way we learn from data – techniques that are being driven, in part, by the needs of biology. This course is intended for graduate students with some quantitative background and an interest in genomics and computational biology.

Instructor: Hiayan Huang and Ben Brown

Bioengineering 190/290: Functional Genomics ([Syllabus](#))
This course provides a survey of the computational analysis of genomic data, introducing the material through lectures on computational methods, presentations of primary literature, and practical bioinformatics exercises. The emphasis is on measuring the output of the genome and its regulation. Topics include modern computational and statistical methods for analyzing data from genomics experiments: high-throughput RNA sequencing data, single-cell data, and other genome-scale measurements of biological processes. Students will perform original analyses with Python and command line tools.
Instructor: Liana Lareau (lareau@berkeley.edu)

Undergraduate courses

Data 8: The Foundations of Data Science ([Details](#))
The UC Berkeley Foundations of Data Science course combines three perspectives: inferential thinking, computational thinking, and real-world relevance. Given data arising from some real-world phenomenon, how does one analyze that data so as to understand that phenomenon? The course teaches critical concepts and skills in computer programming and statistical inference, in conjunction with hands-on analysis of real-world datasets, including economic data, document collections, geographical data, and social networks. It delves into social issues surrounding data analysis such as privacy and design.
Instructors: David Wagner (daw@cs.berkeley.edu)
Ani Adhikari (adhikari@berkeley.edu)

IB 164: Human Genetics and Genomics ([Syllabus](#))
This course provides an introduction to the concepts of human genetics and genomics including: Genome organization and architecture, genetics diversity, probability, association mapping, quantitative genetics, genetic drift, natural selection, and human genetic history.
Instructors: Doris Bachtrog, dbachtrog@berkeley.edu and Peter Sudmant (psudmant@berkeley.edu)

IB 134L: Practical Genomics ([Syllabus](#))
Genome sequencing and analyses have transformed biology over the past two decades. This course provides a hands-on introduction to the world of computational biology and bioinformatics. Students will apply state-of-the-art techniques to analyze genome and microbiome data from the UC Berkeley campus fox squirrels and local California ground squirrels each week. Students will master practical bioinformatics skills and then take on their own scientific research projects, all using genomic data collected specifically for this course. In addition to learning about methods and techniques, we will explore key advances in the field of genomics over the past two decades of both humans and non-model organisms that have driven the current revolution in genome sciences.
Instructors: Peter Sudmant (psudmant@berkeley.edu) and Rebecca Tarvin (rdtarvin@berkeley.edu)

Bootcamps / Resource Building

Introduction to Python Programming ([Syllabus](#))
The Center for Computational Biology offers a 5-day condensed "Introduction to Programming for Bioinformatics Bootcamp". The online course will cover essential Python programming with particular focus on bioinformatics problems and applications. Each 8-hour day includes approximately four hour lectures and four hours of programming exercises guided by the instructors. The bootcamp will require synchronous attendance. This is offered three times in a year: Summer, Winter and Spring.

R Fundamentals ([Details](#))
The D-Lab's R Fundamentals workshop is a four-part introductory series that will teach you R from scratch with clear introductions, concise examples, and support documents. You will learn how to download and install the open-sourced R Studio software, understand data and basic manipulations, import and subset data, explore and visualize data, and understand the basics of automation in the form of loops and functions. After completion of this workshop you will have a foundational understanding to create, organize, and utilize workflows for your personal research.