

INTEGBIO 134L – Practical Genomics (4 units)

To be offered every: **Fall**

Unit Value: 4

Group: **A (Evolution and Genetics)**

Instructor: **Peter Sudmant & Becca Tarvin**

Office: **5002 VLSB / 3120 VLSB**

Course Fee: **\$25 (starting in 2021)**

Email: psudmant@berkeley.edu, rdtarvin@berkeley.edu

2021 information: **In person, TTh 11:00am-12:30pm; lab Th 3-6pm, F 9-12, F 3-6pm**

Instructional Format: **Lecture, Lab**

Ideal Course Schedule: Two lectures per week, 1.5 hours per lecture, 3-hour lab

Final Assessment: A written final exam and a final project

Target enrollment: 1 lecture section, 3 lab sections (90 students, 1 GSI each lab)

Classroom: lecture hall for computer lab for both lecture and lab

Prerequisites: Bio1A and Bio1B

Course description and aims

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.

Assessment:

Lab Assignments: 25%

Midterm: 20%

Final Project: 35%

- Pre-proposal 5%
- peer review 5%
- final proposal 5%
- report 10%
- presentation 10%

Final Exam: 20%

Final Project: Students will perform a genomic analysis on data generated by the course or publicly available data. The final project consists of four graded components.

1. Each student will submit a project proposal which will be graded on the following components: background, hypothesis, methodological approach, and expected outcomes.
2. Students will each perform a peer review of three fellow student's proposals. Reviews consist of a short summary and suggested feedback.
3. Students will, either independently, or with a partner, carry out the proposed research project, with feedback both from peer review and instructors. Students will submit a final report consisting of the methods applied and the results including at least one figure and one table.
4. Students will present a 15-min final presentation and be graded on their communication of the motivation, approach, and results of their research project.

Tentative Course Schedule

08/25/2022	History of genomics and sequencing technologies (Sudmant, lecture)
LAB	Introduction to Jupyter notebooks, bash, and datatypes
08/30/2022	Genome Sequencing I
09/01/2022	Genome Sequencing II
LAB	Mapping reads and pairwise alignment
09/06/2022	Sequence Alignment
09/08/2022	Sequence Assembly
LAB	Genome assembly + Genome annotation
09/13/2022	Sequence modeling, prediction, and annotation I
09/15/2022	Sequence modeling, prediction, and annotation II
LAB	Working with RNA-seq data/ Visualization
09/20/2022	Sequence modeling, prediction, and annotation III
09/22/2022	Population Genetics I
LAB	GWAS
09/27/2022	Population Genetics II
09/29/2022	GWAS I
LAB	Population genetics in Structure/ Project guidelines
10/04/2022	GWAS II
10/06/2022	TBD
LAB	Phylogenies and ancestral state reconstruction
10/11/2022	Review
10/13/2022	Midterm
LAB	No Lab

10/18/2022	Project Overview; Scientific Collecting
10/20/2022	Whole-genome and reduced-representation sequencing
LAB	Field Collection of Fox Squirrels and sequencing
10/25/2022	Metagenomics
10/27/2022	Transcriptomics
LAB	Project proposals due. Peer review of project proposal during lab
11/01/2022	Phylogenetics
11/03/2022	Study design
LAB	Final proposals due. Students work on projects
11/08/2022	The future of genomics
11/10/2022	Veteran's Day (no class)
LAB	No lab
11/15/2022	Open coding hour / Special topic
11/17/2022	Open coding hour / Special topic
LAB	Student open-hours
11/22/2022	Open coding hour / Special topic
11/24/2022	Thanksgiving
LAB	Thanksgiving
11/29/2022	Student presentations
12/01/2021	Student presentations
LAB	No lab
	Final Exam to take place during scheduled time (TBD)

Relevant Readings (although we will not be following these in class)

Brawand, D. *et al.* The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348 (2011).

Consortium (IWGSC), T. I. W. G. S. *et al.* Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191 (2018).

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M. & Stanke, M. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics* **32**, 767–769 (2016).

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860 (2001).

Jain, M. *et al.* Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology* **36**, 338–345 (2018).

Kunte, K. *et al.* *doublesex* is a mimicry supergene. *Nature* **507**, 229–232 (2014).

Mudge, J. M. & Harrow, J. The state of play in higher eukaryote gene annotation. *Nature Reviews Genetics* **17**, 758–772 (2016).

Novembre, J. *et al.* Genes mirror geography within Europe. *Nature* **456**, 98–101 (2008).

Prum, R. O. *et al.* A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569 (2015).

Rogers, R. L. *et al.* Genomic takeover by transposable elements in the Strawberry Poison Frog. *Mol Biol Evol* **35**, 2913–2927 (2018).

Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).

Shendure, J. *et al.* DNA sequencing at 40: past, present and future. *Nature* **550**, 345–353 (2017).

Shulse, C. N. *et al.* High-throughput single-cell transcriptome profiling of plant cell types. *Cell Rep.* **27**, 2241–2247 (2019).

Xue, Y. *et al.* Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* **348**, 242–245 (2015).