INTEGBIO 134L – Practical Genomics (4 units)

To be offered every: **Fall**
Course title: **Practical Genomics**
Abv Title: **PracticalGen**
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
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 section, 25-30 students with potential to expand in future years
Classroom: computer lab for both lecture and lab
Reading Materials: Weekly readings, noted below in the tentative course schedule and bibliography, will be assigned from recent primary literature in addition to selected classic papers that will be placed online.
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%

- proposal 10%
- peer review 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 with 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

08/27/2020	History of genomics and sequencing technologies Case Study: DNA sequencing at 40 (Shendure et al. 2018)
LAB	Learning and applying bash
09/01/2020	Genome assembly I - motivating example and question, background tech
09/03/2020	Case Study: The Human Genome, old and new (International Human Genome Sequencing Consortium 2001, Jain et al. 2018)
LAB	Squirrel mtDNA genome assembly
09/08/2020	Genome Assembly II
09/10/2020	Case Study: Wheat and Poison Frog Genomes (Consortium 2018, Rogers et al. 2018)
LAB	Field Collection of Fox Squirrels and immediately send for sequencing (20-30)
09/15/2020	Genome Annotation
09/17/2020	Case Study: Gene annotation in eukaryotes (Mudge et al. 2016, Hoff et al. 2016)
LAB	Ab initio, transcriptome-based, orthology-based
09/22/2020	Transcriptomics I
09/24/2020	Case Study: Evolution of gene expression in mammalian organs (Brawand et al. 2011)
LAB	Differential gene expression, GO terms
09/29/2020	Transcriptomics II (single-cell)
10/01/2020	Case Study: High-throughput single-cell transcriptome profiling of plant cell types (Shulse et al. 2018)
LAB	R visualization

10/06/2020	Material Review Session
10/08/2020	Midterm
LAB	No Lab
10/13/2020	Pop Gen I (intro to concepts)
10/15/2020	Case Study: European population genetics (Novembre et al. 2008)
LAB	Calling SNPs, mapping
10/20/2020	Pop Gen II (kinship, admixture, ROH)
10/22/2020	Case Study: Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding (Yue et al. 2015)
LAB	STRUCTURE
10/27/2020	Phylogenetics
10/29/2020	Case Study: A comprehensive phylogeny of birds (Prum et al. 2015)
LAB	BLAST, NCBI, reconstruct phylogeny (raxml-ng)
11/03/2020	Metagenomics
11/05/2020	Case Study: Environmental versus genetic influence on human gut microbiota (Rothschild et al 2018)
LAB	Project proposals due. Peer review of project proposal during lab
11/10/2020	Non-model genomics
11/12/2020	Case Study: Fantastic beasts and how to sequence them (Matz 2018)
LAB	Students work with instructors to form teams and work on projects
11/17/2020	How to design a study in genomics
11/19/2020	Case Study: Genomics is failing on diversity (Popejoy and Fullerton 2016)
LAB	Student open-hours
11/24/2020	The future of genomics

11/26/2020	Thanksgiving
LAB	Thanksgiving
12/01/2020	Student presentations
12/03/2020	Student presentations
LAB	no lab
12/08/2020	Student presentations
12/10/2020	Student presentations
	Final Exam to take place during scheduled time (TBD)

Readings

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).