EEB 149—Evolutionary Genomics

Spring 2021, 4 units

Professor Nandita Garud, ngarud@ucla.edu
TA Richard Wolff, rwolff@ucla.edu

	Time:	Location:
Lecture	Tues 2:00-3:50pm	Zoom
Computer lab	Thurs 2:00-3:50pm	Zoom
Section	Fri 2-2:50pm	Zoom
Office hours – Dr. Garud	Wednesdays 3:30-4:30 and after class on Tuesdays and Thursdays*	Zoom
Office hours – Richard Wolff	Monday 1:00 – 2:00 pm*	Zoom

*Also available by appointment. Please email to request a time.

Course Description:

Virtually all organisms on this planet experience evolutionary pressures. Evolutionary pressures act on genomic variation, and in turn can change the genomic composition of populations and whole species. For example, viruses and fruit flies have repeatedly acquired adaptive mutations due to the presence of anti-viral drugs and pesticides. Human genomes have changed in response to selective pressures (e.g. pathogens and changes in our diet), and past demographic events (e.g. bottlenecks, migrations, and expansions). Over long time scales, evolution can result in speciation (e.g. the divergence of humans and chimpanzees). Evolutionary genomics is the study of variation and changes in genomic sequences due to natural selection pressures.

The goal of this course is to understand how evolutionary forces of mutation, drift, selection, recombination, and migration can change genomes, and to learn how analyze genomic data to make evolutionary inferences.

Some questions that we will address in this course include:

- How do evolutionary forces of mutation, drift, selection, recombination, and migration change the genome?
- What specific features of the genome changes in response to evolutionary pressures (e.g. number of genes? Genome size? Non-synonymous polymorphisms?)
- When do genomic changes occur? Can we infer the timing of past evolutionary events?

Course Objectives:

By the end of the course you will be able to:

1. Explain how evolution impacts genomic variation
2. Make inferences from genomic data about past evolutionary histories

3. Perform basic computational genomic analyses in R

Prerequisites: LIFESCI 7A, 7B, 7C, and 23L

Course Materials: We will be reading many chapters from Graham Coop's "Population and Quantitative Genetics". This PDF is freely available here: https://github.com/cooplab/popgen-notes/blob/master/popgen_notes.pdf. In the syllabus below, readings from this PDF will be referred to as "PQG pages xx-yy"

Additionally, we will be reading primary literature, for which PDFs will be provided through CCLE. Finally, we have included short videos with relevant material on YouTube. Please consult the course schedule below for a comprehensive list of course readings and videos.

Contact and communication channels:

Course material and announcements will be made available through CCLE.

Additionally, we have created a Slack workspace in which you can talk to your classmates and form study groups (eeb149.slack.com). On the first day of class, we will make sure everyone can log in.

You are welcome to seek help with lab assignment at office hours – times listed above or by appointment. Your best resources in this class are your classmates and grading in this class is NOT curved. All emails to instructors regarding this class should include: "EEB149" in their subject line. Please allow 24 hours for a response and do not expect email responses during the weekend.

Participation: Every week there will be 1-2 reading assignments. Please come prepared to discuss the reading with your classmates in a breakout group.

A note about participation and attendance during COVID: We realize that many of you are experiencing high amounts of stress and unusual circumstances due to COVID. We are understanding of unexpected changes to your schedule, and so, we do not require attendance as long as the absence is excused. Please send us a message with a brief explanation if you are unable to attend class or hand in an assignment on time. We are happy to make accommodations. Please note that all lectures will be recorded and posted on CCLE.

Computer Lab: Participation in the computer lab is mandatory (unless there is an excused absence). Each week on Thursday you will analyze genomic data using the technique discussed in class. The software used in lab is R. We will be hosting all labs on rstudio.cloud on the web, thus please create a free account (you are also welcome to compute locally on your own computers). Additionally, you will receive a few open-ended and analytical questions every week on Tuesday to help you concretize the material learned in lectures and to prepare you for the exams. You will receive a weekly lab assignment on Thursday where you will begin working on it. If you do not finish working on the assignment during the lab session, you will be expected to continue working on it at home. You are required to upload your completed lab assignments to CCLE. **Assignments are due by Thursday at 2pm PST the following week - before the next lab begins.** Lab assignments should be handed in on time. Every late day will reduce one point from your homework grade. The first late day begins as soon as lecture starts. You are allowed to work with your peers as long as you state on your assignment who you worked with. There will be nine lab assignments to hand in and the lowest grade will be dropped so only eight

assignments will count for the final grade. The last (10^{th}) lab section will be used to review the material from the entire course in preparation for the final exam.

Friday section: Attendance for Friday section is optional, but encouraged. The TA will cover one or two of the analytical (non-computer) problems that are presented in lab on Thursday and any additional questions about lab or lecture. Section can be considered an 'extended office hours' with TA Ricky.

Exams: There will be two open-book take home exams. You are expected to follow the honor code and not consult your peers, and instead perform all the work on your own. The exams will contain open-ended and analytical questions resembling those covered in the homework and lab.

Grading:

- **Exams:** There will be two exams, each worth 20%, or 40 points each.
- **Lab and homework assignments:** There will be 9 assignments that are a combination of lab questions and homework questions. Each will be worth 5%, or 10 points each. The lowest score will be dropped, leaving 8 graded assignments.
- **Participation:** 20%, or 40 points
- Two bonus points will be added for filling out course evaluations (1 for the lecture and 1 for the lab).
- The total number of points that can be obtained are 202 out of 200.
- Grading in this class is not curved

Letter grades will be assigned accordingly:

Letter grade	Percent:
A+	99%-100%
A	93%-98.9%
A-	90%-92.9%
B+	87%-89.9%
B	83%-86.9%
B-	80%-82.9%
C+	77%-77.9%
C	73%-76.9%
C-	70%-72.9
D	60%-70%
F	<60%

Inclusive learning environment:
UCLA values diversity and inclusion. We expect everyone in this class to contribute to a respectful, welcoming, and inclusive environment to support the learning of all other members of the class. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please notify us.

Students needing academic accommodations based on a disability should contact the Center for Accessible Education (CAE) at (310) 825-1501 or in person at Murphy Hall A255. CAE will assess all requested and communicate the adjustments to your professor and/or Teaching Assistant. Any students with CAE approval for proctoring arrangements during exams will need to inform the professor and/or TA prior to exams. When possible, students should contact CAE within the first two weeks of the quarter to allow reasonable time to coordinate accommodations.

For more information, please visit the CAE website: http://www.cae.ucla.edu.

Resources are available to foster the well-being of all UCLA students as they pursue their academic goals. Any student who finds themselves in immediate distress, please call Counseling and Psychological Services (CAPS) to speak directly with a counselor 24/7 at (310) 825-0768, or please call 911. For more information, pleases visit the CAPS website: http://www.counseling.ucla.edu.

Academic Integrity:

It is our goal to help you all do well in this course. We also recognize that life is crazy in ways that it has never been before. Thus, if you are having problems keeping up with the material, please reach out to us. We will do our best to help you in this situation. Please do not cheat or engage in dishonest conduct. It is not worth it. Cheating has severe consequences. Please review the Dean of Student's Student Conduct Code (https://www.deanofstudents.ucla.edu/Individual-Student-Code).

Cheating includes, but is not limited to the following:

1. Working together on online quizzes or exams.
2. Using unauthorized aids on exams.
3. Discussing an exam with a student who has not taken that exam.
4. Plagiarism of any kind.

The UCLA student guide on academic integrity is here:
https://www.deanofstudents.ucla.edu/portals/16/documents/studentguide.pdf

You are NOT permitted to post any of the assignments or course materials to Course Hero or any other websites outside of CCLE.

Copyright Notice:
My lectures and course materials, including PowerPoint presentations, lab/discussion worksheets, tests, outlines, and similar materials, are protected by U.S. copyright law and by University policy. I am the exclusive owner of the copyright in those materials I create. You may take notes and make copies of course materials for your own use. You may also share those materials with another student who is registered and enrolled in this course. You may not reproduce, distribute or display (post/upload) lecture notes, recordings or course materials in any other way — whether or not a fee is charged — without my express written consent. You also may not allow others to do so. If you do so, you may be subject to student conduct proceedings under the UCLA Student Code of Conduct, Section 102.23. Similarly, you own the copyright in your original papers and exam essays. If I am interested in posting your answers or papers on the course web site, I will ask for your written permission.

Schedule (subject to change):

Week 1	**Lecture 1 (March 30): Evolutionary Genomics and the Genome Architecture** Learning Goals: How does genomic information inform us about evolution? What is evolution? What is a genome? **Watch:** What is genomics? https://www.youtube.com/watch?v=mmgICIg0Y1k **Lab 1 (April 1):** NCBI Genome Browser tutorial

	Read: https://www.nytimes.com/interactive/2021/03/25/magazine/genome-sequencing-covid-variants.html
Week 2	**Lecture 2 (April 6): Genome sequences and polymorphisms** Learning Goals: What is a mutation? What is a polymorphism? What might a sample of sequences tell us? How do we estimate allele frequencies? **Reading:** PQG pages 15-19 **Watch:** Short intro to R, in preparation for lab on Thursday: https://www.youtube.com/watch?v=IVKMsaWju8w **Lab 2 (April 8):** Introduction to R
Week 3	**Lecture 3 (April 13): Hardy Weinberg Equilibrium and FST** Learning goals: What is Hardy Weinberg equilibrium? What is drift? What role does large vs small population sizes play in allele frequency changes? What is FST? **Reading**: PQG pages 19-23, 59-64 **Lab 3 (April 15):** Hardy Weinberg Equilibrium and FST
Week 4	**Lecture 4 (April 20): Molecular Clock and Phylogenetic Trees** Learning goals: How do mutations accumulate over long time scales? How do we build a phylogenetic tree? **Reading:** PQG pages 91-93, 95-96 AND https://www.nytimes.com/interactive/2020/04/03/science/coronavirus-genome-bad-news-wrapped-in-protein.html **Lab 4 (April 22):** Coronavirus phylogenetics
Week 5	**Lecture 5 (April 27): Selection over long time scales** Learning Goals: dN/dS and McDonald Kreitman test for positive selection **Reading:** PQG 97-105 AND Adaptive protein evolution at the Adh locus in Drosophila (McDonald and Kreitman, 1991) **Lab week 5 (April 29):** McDonald Kreitman Test
	Take home midterm : April 29-May 4
Week 6	**Lecture 6 (May 4): Population Genetics – Positive selection** Learning goals: What do signatures of positive selection look like? What is a hard sweep? What is a soft sweep? **Reading:** PQG pages 251-253, 256-258 AND Convergent adaptation of human lactase persistence in Africa and Europe (Tishkoff et al. 2007)

	Lab week 6 (May 6): Simulating selective sweeps
Week 7	**Lecture 7 (May 11): Population Genetics – Migration and population structure** Learning goals: What is the history of human migrations? How can PCA inform us about population structure? **Reading:** PQG pages 43-50 AND Genes mirror geography within Europe (Novembre et al. 2008) **Lab 7 (May 13):** PCA
Week 8	**Lecture 8 (May 18): Population Genetics -- GWAS and linkage disequilibrium** Learning goals: GWAS and LD **Reading:** PQG pages 51-57, 119-125 **Lab 8 (May 20):** GWAS exercise in R
Week 9	**Lecture 15 (May 25): Metagenomics** Learning goals: What is metagenomics? What is 16S sequencing? What does a bacterial genome look like? What is alpha diversity? Species richness? **Reading:** Dynamics and stabilization of the human gut microbiome during the first year of life (Backhed et al. 2015) **Lab week 9 (May 27):** Analysis of microbiome data from Backhed et al. 2015.
Week 10	**Lecture 10 (June 1): Genomic variation in the microbiome** Learning goals: What does genetic variation at a SNP level look like inside host microbiomes? What is a strain? What might a selective sweep in the microbiome look like? **Reading:** Population genetics in the human microbiome (Garud and Pollard 2019) **Lab week 10 (June 3):** Review and prep for final
Week 11	**Take home final: June 4-11**

Important dates:

Midterm exam distributed: April 29
Midterm exam due via CCLE: May 4

Final exam distributed: June 4
Final exam due via CCLE: June 11