QBIO 475: Statistical and Evolutionary Genetics
Syllabus

General Information

Lecture time: TTh 12:30-1:50
Lecture location: RRI 421
Instructors: Michael “Doc” Edge (Pronouns: he/him), Jazlyn Mooney (Pronouns: she/her)
Instructor email: edgem@usc.edu, jazlynmo@usc.edu
Instructor office hours: TBD
Section time and location: T 2:00-2:50, RRI 301
Teaching Assistant: TBD
Note: Please bring a laptop and scratch paper to class.

Welcome! We are looking forward to working with you this semester. We aim to help you have as rewarding an experience learning the subjects in this course as we have had.

Course Description

This is an upper-division course designed to introduce quantitative biologists to some central ideas in the mathematical modeling and statistical analysis of genetic variation. Topics from evolutionary genetics and medical genetics will be explored using a mix of math, simulation, and data analysis. Topics may include genetic drift, natural selection, mutation, migration, population structure, and study designs for learning about genotype-phenotype relationships, such as family-based studies, analyses of gene-expression variation, and genome-wide association studies.

Textbook

Our main reading material will be Graham Coop’s notes on population and quantitative genetics, available at https://gcbias.org/population-genetics-notes/ . These notes are free to access. Other readings will be posted on the course website.

Other recommended (but not required) books include Gillespie’s *Population Genetics: A Primer* and Säll & Bengtsson’s *Understanding Population Genetics*. Some of the mathematical derivations we do in class will follow these books.

Course Notes

There is a great deal of excitement about using genetic information to predict risk for disease or understanding the sources of disease risk. Despite the widespread availability of genetic data, these goals are still hard to reach for most common health conditions. One reason is because of the way genetic variation is structured---most genetic variants are rare, and genotypes are correlated with each other in the population. In this course, we will take a close look at human genetic variation and learn some evolutionary reasons it looks the way it does. Then, we will take this same evolutionary framework and apply it to “complex” traits---traits that are influenced by genotypes in many parts of the genome, as well as environmental influences. We will finish the course by talking about some of the ways in which geneticists attempt to learn about the genetic basis of complex traits and how their interpretation rests on the evolutionary concepts developed in the earlier parts of the course.

We will flip the classroom for some of the material, meaning that you will be expected to watch a taped lecture before class, and we will spend the class time learning actively. Slides for both traditional and flipped lectures will be posted.

Learning Goals

By course's end, our aim is that you will be able to:

- Describe the major evolutionary "forces" studied by population geneticists.
- Prove some key mathematical results used in population genetics.
- Apply some major insights of theoretical population genetics to thinking about the distribution and effects of genetic variation within populations.
- Discuss the application of population-genetic and quantitative-genetic thinking to biomedical research.
- Conduct basic data analyses in population and statistical genetics
- Explain the strengths and weaknesses of major methods for learning about connections between genetic and phenotypic variation.

Prerequisites

The only required prerequisite is a course that covers probability---either QBIO 310, MATH 307, or MATH 407 will suffice. Comfort with algebra is also a must. Statistics, and some experience with statistical programming (at the level of QBIO 310) will be helpful. For biological context, the fundamentals of genetic inheritance and evolutionary biology (at the level covered in BISC 120) and some knowledge of DNA replication and meiosis (at the level of BISC 220) will also be useful, but these courses are not required.

Some of the material in this course is covered in BISC 313. This course is intended primarily for QBIO majors, and as such it places a heavier emphasis on math and computation and a lighter emphasis on biological mechanisms and processes.

Grading Policy

Your final grade will be calculated on the basis of a weighted average, with the weights

45% Homework
5% Participation
20% Midterm (split into take-home and in-person components)
30% Final Exam (split into take-home and in-person components)

If, at the end of the term, your total weighted numerical grade is higher when homework and participation are excluded from consideration, then we will exclude the low component(s) from your grade. In other words, your homework and participation grade cannot count against you; it can only help your final grade.

Take-home exams will be open-note, open-book, and open-web (not counting anything posted to the web by or in response to any member of the class—that means no posting exam questions to forums, etc.). I will ask you to sign a statement affirming that you have followed the rules for each exam.

At the end of the semester, your letter grade will be determined either by traditional cutoffs (i.e. A's in the 90s, B's in the 80s, etc.) or by a curve centered on the median grade on the exams and paper, whichever leads to higher grades.

Homework

There will be 5-6 homework assignments during the semester, due every 2-3 weeks. Doing the homework will be your most important method for learning the material. Homeworks will be graded on the following scale:

0: missing, more than 24 hours late, or less than 50% complete.
1: mediocre or poor effort, late by less than 24 hours, or incomplete.
2: Good effort, though some answers may be wrong.
3 (Bonus): Exceptional. Your TA is weeping with ecstasy.

All "2"s would give you a perfect homework score. Scores of "3" will not happen often and are considered bonus. If your score is somewhere between a "1" and a "2," your TA may give you the option to turn in a corrected version of your homework to earn a "2." The TA may also assign fractional scores (e.g., 1.5) if they choose.

You are encouraged to work collaboratively on the homework, but please write your own solutions. We will drop your lowest homework score.

Participation

Most lectures will include an activity that's worth a participation point. In case you need to miss some class sessions, we will make some substitute activities available that you can submit later. Attendance at discussion section will also count toward your participation grade.

Software

We will use R, a programming language designed for statistical computing. R is available free online from the R Project website, https://www.r-project.org/. We recommend you also use RStudio, and interactive development environment designed for use with R. RStudio is also free. (Download the open source version of RStudio Desktop from https://www.rstudio.com/products/rstudio.) RStudio requires an active R installation.

Contacting the instructors

For administrative questions that apply to everyone or any questions about course content, please post to the forum on the course website, ask your TA, or come to office hours. Please email the instructors directly for questions relating to class administration issues that apply to you specifically (e.g. disability-related accommodations, scheduling an appointment, etc.). We will do our best to reply to all emails within 48 hours, but there is no guarantee we will be able to do so earlier than that. Please include the course name or number in the subject line of any emails you send the instructors.

Course Schedule (Subject to change)

Unit I: Course introduction and review

Week 1
8/23 (Mooney) Course policies and introduction (correlation and structure of genetic variation as central problems in statistical genetics)
8/25 (Mooney) Review of basic genetics I (information content in the genome, genetic code, central dogma, mitosis/meiosis)
Reading: Coop notes, ch. 1

Week 2
8/30 (Mooney) Basic genetics II (Types of variation, ways variation affects proteins, mutation, variation in humans)
Reading: Coop notes, ch. 1
9/1 (Edge) Probability review (independence, conditional probability, Bayes' theorem, expectation, variance, covariance, commonly used distribution families)
Reading: Coop notes, appendix A

Unit 2: Evolutionary genetics models

Week 3
9/6 (Edge) Genotypes and allele frequencies (Hardy-Weinberg, descriptions of variation, inbreeding and homozygosity)
Reading: Coop notes, ch. 2
9/8 (Edge) Correlations among loci, or "linkage disequilibrium" (D, D', and r^2, recombination and decay of LD, haplotype structure)
Reading: Coop notes, section 3.0.5

Week 4
9/13 (Edge) Genetic drift I (The Wright-Fisher process, loss of diversity, θ and N_e)
Reading: Coop notes ch 4, through the end of section 4.1
9/15 (Edge) Genetic drift II (The coalescent: time to coalescence, fixation probability, heterozygosity, SFS)
Reading: Coop notes section 4.2 until end of ch 4

Week 5
9/20 (Edge) Natural Selection I (Analyzing a deterministic model of selection)
Reading: Coop notes, Ch 10 through the end of 10.1, first 3 pages of ch 11 (i.e. intro and 11.0.1)
9/22 (Edge) Natural Selection II (The W-F diffusion: selection and drift)

Reading: Coop notes, Ch 12

Week 6
9/27 (Mooney) Natural Selection III (types of selection and examples: purifying selection, nearly neutral variants, positive selection, heterozygote advantage, and polygenic selection)
Reading: none
9/29 (Mooney) Natural Selection IV (Natural selection and genetic variation: the neutral theory, dN/dS and the MK test, linked selection, background selection)
Reading: Coop notes, Ch 13 up to 13.1.2 (background selection)

Week 7
10/4 (Mooney) Midterm review
10/6 (Edge) Forensic genetics, family trees, and genetic privacy
Reading: none
Take-home midterm released after class

Unit III: Human variation

Week 8
10/11 (Mooney) Human genetic variation
Reading: Biddanda, A., Rice, D. P., & Novembre, J. (2020). Geographic patterns of human allele frequency variation: a variant-centric perspective. *BioRxiv*.
10/13 Fall recess

Week 9
10/18 (Mooney) Models and summaries of population structure (F_{ST} and the Wahlund principle, split and migration models, models of admixture)
Reading: Coop notes, chapter 3 (up to 3.0.1)
10/20 (Mooney) Methods for studying structure and admixture (STRUCTURE, PCA, etc.)
Reading: Coop notes, 3.0.2-3.0.4.

Week 10
10/25 (Guest Lecture: Vivian Link) Ancient DNA
10/27 (Guest Lecture: Adam Rutherford) The history of eugenics

Unit IV: Genotype and phenotype

Week 11
11/1 (Edge) Quantitative Genetics I: Resemblance among relatives and heritability
Reading: Coop notes, ch. 7
11/3 (Edge) Quantitative Genetics II: Response to selection, the infinitesimal model in light of contemporary GWAS
Reading: Coop notes, ch. 8
Reading: Greenwood, V. (2018). Theory suggests that all genes affect every complex trait. *Quanta*. https://www.quantamagazine.org/omnigenic-model-suggests-that-all-genes-affect-every-complex-trait-20180620/

Week 12

11/8 (Edge) Gene mapping and GWAS (Linkage studies and GWAS, LD as "tagging" causal variants, genome-wide significance)
Reading: Grady, D. (2020) Haunted by a gene. *New York Times*, March 10, 2020
https://www.nytimes.com/2020/03/10/health/huntingtons-disease-wexler.html
Reading: Bush, W. S., & Moore, J. H. (2012). Genome-wide association studies. *PLoS Comput Biol*, *8*(12), e1002822. **Read sections 1-3.**
11/10 (Edge) What causes a genotype-phenotype association? (Direct and indirect effects, stratification, assortment)
Reading: Young, A. I., Benonisdottir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype-phenotype associations in humans. *Science*, *365*(6460), 1396-1400.

Week 13

11/15 (Edge) Downstream GWAS analyses (selected topics: fine mapping, LD scores, polygenic scores)
Reading: TBD
11/17 (Guest lecture: Nick Mancuso) Gene expression and TWAS

Unit V: Beyond humans

Week 14
11/22 (Mooney) Conservation genetics
11/24 Thanksgiving, no class

Week 15
11/29 (Guest lecture: Matt Pennell) Reconstructing evolutionary history and the tree of life
12/1 (Mooney) Course wrap-up

Statement on Academic Conduct and Support Systems

Academic Conduct:

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, "Behavior Violating University Standards" policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, policy.usc.edu/scientific-misconduct.

Support Systems:

Counseling and Mental Health - (213) 740-9355 – 24/7 on call
studenthealth.usc.edu/counseling
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
suicidepreventionlifeline.org

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention and Services (RSVP) - (213) 740-9355(WELL), press "0" after hours – 24/7 on call
studenthealth.usc.edu/sexual-assault
Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED)- (213) 740-5086 | Title IX – (213) 821-8298
equity.usc.edu, titleix.usc.edu
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following *protected characteristics*: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations. The university also prohibits sexual assault, non-consensual sexual contact, sexual misconduct, intimate partner violence, stalking, malicious dissuasion, retaliation, and violation of interim measures.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298
usc-advocate.symplicity.com/care_report
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity |Title IX for appropriate investigation, supportive measures, and response.

The Office of Disability Services and Programs - (213) 740-0776
dsp.usc.edu
Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy - (213) 821-4710
uscsa.usc.edu
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101
diversity.usc.edu
Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
dps.usc.edu, emergency.usc.edu
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call
dps.usc.edu
Non-emergency assistance or information.