If you are looking for the 2023 syllabus go here

Winter 2022

HGEN 47100/BIOS 21216 Introduction to Statistical Genetics

Instructors: Xin He (xinhe@uchicago.edu), Hae Kyung (Haky) Im (haky@uchicago.edu)
Teaching Assistant: Jing Gu (jinggu@uchicago.edu)

Lectures: M-W 1:30 - 2:50 PM
Lab session: F 1:30 - 2:50 PM

Classroom: Biological Sciences Learning Center (BSLC), Room 305

Course Description

In this course we will cover the core concepts and statistical procedures that are used in the mapping of genetic traits from observational data. We will cover basic statistical techniques of genome-wide association studies and tools for ``post-GWAS'' analysis.

Prerequisites

Students are expected to have had:

- An introductory statistics course: HGEN 47400 Introduction to Probability and Statistics for Geneticists, or STAT 24400 Statistical Theory and Methods I, or equivalent. Note that STAT 24300 Statistical Models and Methods, does not cover parameter estimation and is not sufficient.
- An introductory course in genetics: BIOS 20187 Fundamental of Genetics or equivalent.
- Some basic knowledge of programming (R) and Unix command lines. Computational labs will quickly move towards using unix-command-line tools and the software package R.

Text and Materials

We will assign readings per lecture. The readings will be drawn from multiple sources including:

- **An Introduction to Statistical Genetic Data Analysis by Mills, Barban and Tropf,**
- Nature Reviews Genetics article series on “Fundamental concepts in genetics”, “Genome-wide association studies”, and “Computational Tools”, and
- Recent literature.

The book “The Fundamentals of Modern Statistical Genetics” is a bit outdated but has some relevant chapters and it’s free if you download it from the UChicago network (link).

Expectations

Your final grade will be based on the following items:

Problem Sets (70%): Typically connected to a computational lab exercise but there will also be stand-alone questions. You may work with others to complete them, but are expected to write up the material yourself and to master the material yourself. Assignments will be typically distributed on Wednesday after class and due the following week at lab sessions.

Final Presentation (20%): You will be expected to present a primary research paper (or set of papers) on a topic of your choice. Alternatively, you may carry out a small research project related to the course. In either case, please meet with the instructors to approve your plan. The presentation will be approximately 25 minutes in length.

Class Participation (10%): We encourage everyone to ask questions and add comments in class. In addition, we ask every student to post your own summary of lectures after each class. This could be just a few sentences long, and may include your own thoughts or questions. You can post them at Canvas > Discussions, under the topic created by the instructors after classes.

Honor statement

Academic integrity and honesty is a core principle of the University of Chicago and is expected at all times from its students. We ask that you add an honor statement to the end of every assignment specifying who you consulted or worked with. The statement should also include the following:

“I discussed with xx and yy for this assignment and received help from the tutor (name).

In submitting this assignment, I attest that it contains only my own independent work carried out according to the directions given to me by the instructor. I understand that academic integrity and honesty is a core principle of the University of Chicago and is expected at all times from its students.”

Communication and Office Hours

Each instructor will be available for office hours on the weeks they are teaching.

- Xin He (xinhe@uchicago.edu): 2:00 - 3:00 PM Thursday or other time by appointment. All meetings will be over zoom until further notice: https://uchicago.zoom.us/j/8048321077?pwd=UWlXQ2VjdFNhZGFFR1FPT2ExaWVjdz09

- Haky Im (haky@uchicago.edu): HKI will hold office horse after her lectures. You can also encouraged to schedule an appointment with her at other times. All meetings will be over zoom for now.
- Jing Gu(jinggu@uchicago.edu): 10:00 - 11:00 AM Thursdays. The first two meetings will be over zoom and then probably switched to in person: https://uchicago.zoom.us/j/94979712813?pwd=S1N2R2g1aFBNSDBOUC80YzNBdTV3UT09

Slack

The announcements will be made through Canvas. For other types of communications, we will use Slack. You can post your questions about lectures, homework, etc. in slack, and if it's easier for you, you can form study groups and discuss using your own slack channels. To join the slack workspace of the class, use this link:

https://join.slack.com/t/hgen471statsgen/shared_invite/zt-115hscfno-jiJNd0eoSkL2~0KzFEqYHA
Special accommodations
If you feel you need accommodation, the process is outlined on this webpage. As noted there, Student Disability Services engages with students to determine the necessary accommodations, including those temporary in nature. There is a range of accommodations they can be helpful implementing.

If you need any kind of academic accommodations, you will need a letter from Student Disability Services outlining the accommodations you are entitled to in order to receive accommodations. Feel free to contact Melissa Lindber (mlindber@bsd.uchicago.edu) if you have any questions.

Course Schedule

L	W	**Subject**	**Date**	**Who**
		Module 1: Foundations of GWAS		
1	1	Review of probability, Maximum likelihood estimation, and Bayesian inference.	1/10/2022	HKI
2	1	Fundamental concepts in statistical genetics	1/12/2022	XH
	1	Lab 1: Genotype data manipulation and simulation	1/14/2022	TA
	2	Martin Luther King Day	1/17/2022	
3	2	Association: GWAS & multiple testing.	1/19/2022	HKI

	2	Lab 2: Genome-Wide Association Studies (GWAS) with PLINK	1/21/2022	TA
4	3	Association II. QC, power.	1/24/2022	HKI
5	3	Genotype imputation	1/26/2022	XH
	3	Lab 3: Genotype Imputation and Haplotype Phasing with IMPUTE2	1/28/2022	TA
6	4	Meta-analysis. Statistical fine-mapping.	1/31/2022	XH
7	4	Population structure.	2/2/2022	HKI
	4	Lab 4: Fine-mapping using SuSiE	2/4/2022	TA
		Module 2: Polygenic Models		
8	5	Heritability: definition, estimation with family data, and with GWAS data	2/7/2022	XH
9	5	Heritability by summary statistics: LD score regression. Heritability estimation with different priors: LDAK, BSLMM.	2/9/2022	HKI
	5	Lab 5: Calculating the Heritability of Complex Traits with GCTA. (Also calculate principal components and plot)	2/11/2022	TA
10	6	Linear Mixed Model to correct for sample relatedness: BOLT-LMM. Prediction: polygenic risk score, BLUP, Lasso/Elastic net, Bayesian methods.	2/14/2022	HKI
		Module 3: Functional GWAS		
11	6	Molecular Quantitative Trait Loci (QTL) mapping	2/16/2022	XH
	6	Lab 6: tensorQTL	2/18/2022	TA
12	7	Mendelian randomization PrediXcan/TWAS	2/21/2022	HKI
13	7	Functional annotations of genetic variants. Using functional annotations in GWAS: enrichment analysis and fine-mapping.	2/23/2022	XH
	7	Lab 7: S-PrediXcan	2/25/2022	TA
14	8	Pathway analysis and cross-phenotype analysis	2/28/2022	XH

		Module 4: Student Presentations		
14	8	Student Presentations	3/2/2022	
	8	S-LDSC heritability partitioning by functional annotation	3/4/2022	
15	9	Student Presentations	3/7/2022	
	9	Student Presentations	3/9/2022	
	10	College Reading Period	3/12/2022	
	11		3/17/2022	
		Winter Quarter Ends	3/19/2022	

Suggested papers for student presentation

You are strongly encouraged to select the papers that fit your research goals and interests. Our selection may be biased towards methods, so keep that in mind.

Paper	Claimed by
Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program https://www.nature.com/articles/s41588-020-00767-x	
Functionally informed fine-mapping and polygenic localization of complex trait heritability https://www.nature.com/articles/s41588-020-00735-5	Will Tang (03/07)
Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power https://www.nature.com/articles/s41588-020-00766-y	James Li (03/09)
Improving Polygenic Prediction in Ancestrally Diverse Populations https://www.medrxiv.org/content/10.1101/2020.12.27.20248738v1	
A robust and efficient method for Mendelian randomization with hundreds of genetic variants https://www.nature.com/articles/s41467-019-14156-4	
GTEx flagship paper https://science.sciencemag.org/content/369/6509/1318	
Whole genome sequencing in the Middle Eastern Qatari population identifies genetic associations with 45 clinically relevant traits https://www.nature.com/articles/s41467-021-21381-3	Geena Woo and Niyati Jain (03/09)
Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants https://science.sciencemag.org/content/369/6503/561	Hang Chen (03/02)
A brief history of human disease genetics (Everyone should read this paper, it’s a very nicely written summary of the development in the last two decades) https://www.nature.com/articles/s41586-019-1879-7	
Current clinical use of polygenic scores will risk exacerbating health disparities https://www.nature.com/articles/s41588-019-0379-x	

Widespread signatures of natural selection across human complex traits and functional genomic categories https://www.nature.com/articles/s41467-021-21446-3	
Regulatory genomic circuitry of human disease loci by integrative epigenomics https://www.nature.com/articles/s41586-020-03145-z	
Genetic determinants of daytime napping and effects on cardiometabolic health https://www.nature.com/articles/s41467-020-20585-3	Temidayo Adeluwa & Saideep Gona (03/02)
Genetic Control of Expression and Splicing in Developing Human Brain Informs Disease Mechanisms, Cell, 2019 https://pubmed.ncbi.nlm.nih.gov/31626773/	Anna Goncalves (03/07)
Science Forum: The single-cell eQTLGen consortium https://elifesciences.org/articles/52155	
Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis, PLoS Medicine, 2018 https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002654	Hyunkyung Kim (03/02)
Comparative genetic architectures of schizophrenia in East Asian and European populations https://www.nature.com/articles/s41588-019-0512-x	Evan Wu (03/09)
~~Computationally efficient whole genome regression for quantitative and binary traits~~ ~~https://www.biorxiv.org/content/10.1101/2020.06.19.162354v2.full.pdf~~ ~~Now published in~~ ~~https://www.nature.com/articles/s41588-021-00870-7~~	
Trans Effects on Gene Expression Can Drive Omnigenic Inheritance https://www.sciencedirect.com/science/article/pii/S0092867419304003	Ethan Zhong (03/07)
Multi-ancestry Genome-wide Association Study of Varicose Veins Reveals Polygenic Architecture, Genetic Overlap with Arterial and Venous Disease, and Novel Therapeutic Opportunities https://www.medrxiv.org/content/10.1101/2022.02.22.22271350v1	
PGS-server: accuracy, robustness and transferability of polygenic score methods for biobank scale studies |	

Briefings in Bioinformatics | Oxford Academic https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbac039/6534383?login=true	
Genetic associations of protein-coding variants in human disease https://www.nature.com/articles/s41586-022-04394-w	