Zeyun Lu

Cell: (858)207-8318 **Email**: zeyunlu@usc.edu **Web**: `zeyunlu.github.io`
Address: 2001 N Soto St, Los Angeles, CA 90033 **Last Updated**: 04/28/2023

Summary

PhD candidate in Biostatistics interested in developing causal inference methods on the genetic architecture of complex diseases.

Education

PhD in Biostatistics 08/2019 - Present
University of Southern California (USC) Los Angeles, CA

MS in Statistics 09/2015 - 12/2016
University of California, San Diego (UCSD) La Jolla, CA

BS in Mathematics/Economics 09/2012 - 09/2014
University of California, Los Angeles (UCLA) Los Angeles, CA

Academic Experience

Graduate Research Assistant 08/2020 - Present
Center for Genetic Epidemiology Los Angeles, CA
Keck School of Medicine (KSOM), USC
Mentor: Dr. Nicholas Mancuso

- Developed a Bayesian gene fine-mapping method (MA-FOCUS) for complex diseases leveraging GWAS and gene expression data from multiple ancestries.
- Developed a Bayesian SNP fine-mapping method (SuShiE) to identify shared causal genetic variants of molecular QTL data.
- Developed an integrative approach (rPWAS) leveraging GWAS and protein data to test post-transcriptional regulatory mechanisms of complex diseases.

Research Internship 05/2022 - Present
Dana-Farber Cancer Institute/Harvard Medical School Online/Boston, MA
Mentor: Dr. Alexander Gusev

- Developed a statistical framework to integrate experimental perturb-seq data for *trans* effects with GWAS and eQTL data for *cis* effects to identify "core" genes that are the *trans* targets of disease-relevant genomic variations.

Statistician 04/2017 - 06/2019
Alzheimer's Therapeutic Research Institute (ATRI) San Diego, CA
KSOM, USC
Mentor: Dr. Rema Raman and Dr. Michael Donohue

- Participated as Biostatistician in study teams of several Alzheimer's Disease clinical trials and observational studies such as ADNI3, A4, and LEADS by designing and generating different study reports including but not limited to performing data analysis with mixed models, investigating, data anomalies, and implementing patient randomization.
- Designed, built, and administered R Shiny projects that created dynamic online clinical data reporting system and risk-based monitoring system for ADNI3.
- Mentored student intern in R Shiny, clinical trials, and statistical analysis.

Articles

4. *Novel insight into the etiology of ischemic stroke gained by integrative transcriptome-wide association study.* Jung J, **Lu Z**, de Smith A, Mancuso N. medRxiv pre-print.
   DOI: `https://doi.org/10.1101/2023.03.30.23287918`
3. *twas_sim, a Python-based tool for simulation and power analysis of transcriptome-wide association analysis.* Wang X, **Lu Z**, Bhattacharya A, Pasaniuc B, Mancuso N. **Bioinformatics**. April, 2023.
   DOI: `https://doi.org/10.1093/bioinformatics/btad288`
2. *Multi-ancestry fine-mapping improves precision to identify causal genes in transcriptome-wide association studies.* **Lu Z***, Gopalan S*, Yuan D, Conti D, Pasaniuc B, Gusev A, Mancuso N. **American Journal of Human Genetics**. August, 2022.
   DOI: `https://doi.org/10.1016/j.ajhg.2022.07.002`
1. *Integrative genomic analyses identify susceptibility genes underlying COVID-19 hospitalization.* Pathak GA*, Singh K*, Miller-Fleming TW, Wendt FR, Ehsan N, Hou K, Johnson R, **Lu Z**, Gopalan S, Yengo L, Mohammadi P, Pasaniuc B, Polimanti R, Davis LK, Mancuso N. **Nature Communication**. July, 2021.
   DOI: `https://doi.org/10.1038/s41467-021-24824-z`

* Equal Contribution

Awards & Fellowship

4. **Third Place Winner** 12/2021
   Video Competition for KSOM graduate students to showcase their research.
   Video Link: `https://youtu.be/Nc5SVhz7-aQ`
3. **Reviewers' Choice** 10/2021
   Abstract scored by ASHG 2021 reviewers in the top 10% of all poster abstracts.
2. **Outstanding Overall Performance** 08/2020
   In recognition of top PhD students in the annual screening exams.
1. **KSOM Graduate School Fellowship** 2019/2020
   In recognition of new PhD students' outstanding academic promise.

Invited Talks

1. *MA-FOCUS project.* Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center. 01/2023.

Posters & Abstracts

8. *Multi-ancestry fine-mapping of gene expression data identifies shared cis-regulatory mechanisms.*
   - Conferences: ProbGen 2023.
7. *Residual Proteome-wide Association Study Identifies Genes for Blood-Related Traits*
   - Conferences: ASHG 2021.
6. *Novel insight into the etiology of ischemic stroke gained by integrating human transcriptome-wide association study with rodent expression data.*
   - Conferences: ASHG 2021.
5. *Multi-ethnic fine-mapping improves precision to identify causal genes in transcriptome-wide association studies.*
   - Conferences: ProbGen 2021, ASHG 2020.

4. *Integrative analyses identify genes and their functional consequences underlying COVID-19 hospitalization.*
    - Conferences: ACMG Annual Clinical Genetics Meeting 2021.
    - Publications: Molecular Genetics and Metabolism 04/2021.
3. *Integrative analyses with large-scale COVID-19 GWAS identifies susceptibility genes underlying hospitalized outcomes.*
    - Conferences: ASHG 2020.
2. *Brain Amyloid Burden, Sleep, and Circadian Rest/Activity Rhythms: Screening Findings from A4 and LEARN.*
    - Conferences: AAIC 2017.
    - Publications: Alzheimer's & Dementia. 11/2018.
1. *An Open-source Implementation of Data Standards for AD Clinical Trials.*
    - Conferences: CTAD. 2017.

TAship

2. **Multi-omics Boot Camp of SHARPP Training**
    - Online course at Mailman School of Public Health, Columbia University.
    - Time: 06/2023, 01/2023, 06/2022, 06/2021.
1. **PM-522A: Introduction to the Theory of Statistics**
    - Graduate level course at KSOM, USC.
    - Time: Fall 2021.

Software

5. **SuShiE**
    - A command-line Python software to perform SNP fine-mapping on molecular QTL data with high-speed inference by leveraging JAX JIT compilation.
    - `https://github.com/mancusolab/sushie`
4. **TWAS Simulator**
    - A Python software to perform Transcriptome-wide Association Studies (TWAS) simulations.
    - `https://github.com/mancusolab/twas_sim`
3. **MA-FOCUS**
    - A command-line Python software to perform gene-trait fine-mapping from TWAS in either single ancestry or multi-ancestry settings.
    - `https://github.com/mancusolab/ma-focus`
2. **ADNI3 and LEADS Dynamic Web Report**
    - Two Shiny apps that transform static reports to interactive web pages for clinical trails investigators' monitoring on enrollment, primary outcomes, and safety.
    - Interval apps that are not public accessible.
1. **ADNI3, A4, and LEADS**
    - Three R packages that crawl EDC API, transform data into CDISC format, and automatically build several reports (enrollment, safety etc) for clinical research.
    - Interval packages that are not public accessible.

Reviewership	Nature Genetics
Membership	American Statistical Association. American Society of Human Genetics.
Skills	R (Tidyverse, Shiny), Python, Linux, plink, gcta, bedtools, LDSC, Git, Latex, CDISC, SAS, Stata.