Xuewei Cao

Center for Statistical Genetics in the Gertrude H. Sergievsky Center, New York, NY 10032
Email : xc2270@cumc.columbia.edu Home Page

PROFESSIONAL APPOINTMENT

Postdoctoral Research Scientist May 2023 - present
Center for Statistical Genetics in the Gertrude H. Sergievsky Center
Department of Neurology, Columbia University
Research direction: Fine-mapping and colocalization analysis, *trans*-QTL network

EDUCATION

Doctor of Philosophy in Statistics August 2018 - April 2023
Department of Mathematical Science, Michigan Technological University, MI, US
Research direction: Statistical Genetics

Master of Science in System Theory August 2015 - June 2018
School of Systems Science, Beijing Normal University, Beijing, China
Research direction: Applied complex network

Bachelor of Science in Statistics August 2011 - June 2015
Department of Mathematical Science, Heilongjiang University, Harbin, China

RESEARCH INTERESTS

Statistical genetics, Biological network analysis, Bayesian modeling, Machine learning, Functional annotation.

RESEARCH PUBLICATION

Published or Submitted

1. **Xuewei Cao**, Shuanglin Zhang, Qiuying Sha. A novel method for multiple phenotype association studies based on genotype and phenotype network. 2023+. *Submitted in August 2023*. software
2. **Xuewei Cao**†, Joyce H. Keyak†, Sigurdur Sigurdsson, Chen Zhao, Weihua Zhou, Anqi Liu, Thomas Lang, Hong-Wen Deng, Vilmundur Gudnason, Qiuying Sha. A New Hip Fracture Risk Index Derived from FEA-Computed Proximal Femur Fracture Loads and Energies-to-Failure. 2023+. *In revision at Osteoporosis International*. arXiv
3. Hongjing Xie, **Xuewei Cao**, Shuanglin Zhang, Qiuying Sha. Joint analysis of multiple phenotypes for extremely unbalanced case-control association studies using multi-layer network. 2023+. *In revision at Bioinformatics*.
4. Chen Zhao†, Joyce H. Keyak†, **Xuewei Cao**, Qiuying Sha, Li Wu, Zhe Luo, Lanjuan Zhao, Qing Tian, Chuan Qiu, Ray Su, Hui Shen, Hong-Wen Deng, Weihua Zhou. Multi-view information fusion using multi-view variational autoencoders to predict proximal femoral strength. 2023+. *Submitted in June 2023*. arXiv
5. Chen Zhao, Anqi Liu, Xiao Zhang, **Xuewei Cao**, Zhengming Ding, Qiuying Sha, Hui Shen, Hong-Wen Deng, Weihua Zhou. CLCLSA: Cross-omics Linked embedding with Contrastive Learning and Self Attention for multi-omics integration with incomplete multi-omics data. 2023+. *Submitted in August 2023*. arXiv
6. **Xuewei Cao**†, Ling Zhang†, Mingxia Zhao, Cheng He, Kui Zhang, Sanzhen Liu, Qiuying Sha, Hairong Wei. TGPred: Efficient methods for predicting target genes of a transcription factor by integrating statistics, machine learning, and optimization. NAR Genomics and Bioinformatics. 2023. *Accepted*. software
7. Meida Wang, **Xuewei Cao**, Shuanglin Zhang, Qiuying Sha. A clustering linear combination method for multiple phenotype association studies based on GWAS summary statistics. Scientific Reports. 2023; 13(1), p.3389. link
8. Hongjing Xie, **Xuewei Cao**, Shuanglin Zhang, Qiuying Sha. Joint analysis of multiple phenotypes for extremely unbalanced case-control association studies. Genetic Epidemiology. 2023. link
9. **Xuewei Cao**, Xiaoyu Liang, Shuanglin Zhang, Qiuying Sha. Gene selection by incorporating genetic networks into case-control association studies. European Journal of Human Genetics. 2022. link
10. Xiaoyu Liang, **Xuewei Cao**, Qiuying Sha, Shuanglin Zhang. HCLC-FC: a Novel Statistical Method for Phenome-Wide Association Studies. PLOS ONE. 2022; 17(11): e0276646. link
11. **Xuewei Cao**, Xuexia Wang, Shuanglin Zhang, Qiuying Sha. Gene-based association tests using GWAS summary statistics and incorporating eQTL. Scientific Reports. 2022; 12(1):3553. link

12. Nongthombam Boby†, **Xuewei Cao†**, Kelsey Williams, Shiva Kumar Goud Gadila, Monica N. Shroyer, Peter J. Didier, Sudesh K. Srivastav, Arpita Das, Kate Baker, Qiuying Sha, Bapi Pahar. Simian Immunodeficiency Virus Infection Mediated Changes in Jejunum and Peripheral SARS-CoV-2 Receptor ACE2 and Associated Proteins or Genes in Rhesus Macaques. Frontiers in Immunology. 2022; 13:835686. link
13. Nongthombam Boby†, **Xuewei Cao†**, Alyssa Ransom, Barcley T. Pace, Christopher Mabee, Monica N. Shroyer, Arpita Das, Peter J. Didier, Edith Porter, Sudesh K. Srivastav, Qiuying Sha, Bapi Pahar, Identification, characterization and transcriptional reprogramming of epithelial stem cells/intestinal enteroids in simian immunodeficiency virus infected rhesus macaques. Frontiers in Immunology. 2021; 12: 769990. link

In Preparation

1. **Xuewei Cao†**, Lirong Zhu†, Xiaoyu Liang, Shuanglin Zhang, Qiuying Sha. Constructing genotype and phenotype network helps reveal disease heritability and phenome-wide association studies.
2. Lirong Zhu, Shijia Yan, **Xuewei Cao**, Qiuying Sha, Shuanglin Zhang. Integrating external controls by regression calibration for genome-wide association study.

RESEARCH EXPERIENCE

Michigan Technological University Aug. 2018 - present
Research Assistant, Working with Prof. Qiuying Sha.

- Develop novel statistical methods and efficient bioinformatical tools to address problems from genome-wide association studies and phenome-wide association studies by incorporating genotype and phenotype network.

Michigan Technological University Summer 2021
Research Intern, Working with Prof. Hairong Wei.

- Dap-Seq and R-loop genome data analysis using bowtie2, samtools, MACS2 etc;
- Reconstruction of gene regulatory network based on the genetic network information.

Arizona State University Jan. 2018 - Feb. 2018
Research Visiting Scholar, Supervisor: Prof. Sander E. van der Leeuw

- Attend a series of orientation sessions on Arizona State University and its complex systems program.
- Take several courses, such as Introduction to Complex Networks, The role of info.processing in Sustainability, Analysis of Human Behavior using Social Networks, etc.

TEACHING EXPERIENCE

Michigan Technological University Fall 2021
Teaching Assistant, Mentored by Ann M. Humes and Prof. John Gruver.

- Prepare for four weeks course materials of Precalculus, such as lecture videos, handouts, WebAssign assignments.
- Teaching assistant of Calculus with Technology I: hold office hour, lead two lectures (the chain rule and the fundamental theory of calculus), make rubric, draft and grade exams.

Beijing Normal University Spring 2017
Teaching Assistant, Mentored by Prof. Ya Zhou and Prof. Liujun Chen.

- Teaching assistant of two undergraduate courses, Game Theory and Econometrics: lead weekly recitations, lead group project and assignment, hold office hour, grade exams and assignments.

CONFERENCES AND PRESENTATIONS

- American Society of Human Genetics 2023 (ASHG 2023), COLOC-Boost: A new gradient boosting informed colocalization algorithm improves the identification of functional disease causal variants. **Poster**, Washington, DC, Nov.3 , 2023
- The Joint Statistical Meetings (JSM 2023). Constructing genotype and phenotype network helps reveal disease heritability and phenome-wide association studies. **Poster**, Toronto, Ontario, Canada, Aug. 9, 2023

† **denote co-first author.**

- ICSA 2023 Applied Statistics Symposium (ICSA 2023). Constructing genotype and phenotype network helps reveal disease heritability and phenome-wide association studies. **Talk**, Ann Arbor, Michigan, Jun. 14, 2023
- Institute of Computing and Cybersystems (ICC) Joint Center for Biocomputing and Digital Health. Gene selection and genetic association test by incorporating the genetic networks. **Talk**, Hougthon, Michigan, Dec. 9, 2022
- American Society of Human Genetics 2022 (ASHG 2022), A novel method for multiple phenotype association studies based on genotype and phenotype network. **Poster**, Los Angeles, California, Oct. 25 - Oct. 29, 2022
- The American Society for Bone and Mineral Research Annual Meeting 2022 (ASBMR 2022), Hip Fracture Prediction using the First Principal Component Derived from FEA-Computed Fracture Loads. **Poster**, Austin, Texas, Sep. 9 - Sep. 12, 2022
- Upper Peninsula Medical Conference, The Impact of Medication Adherence on Health Care Costs for People with Diabetes. **Poster**, Houghton, Michigan, Aug. 26 - Aug. 28, 2022
- American Society of Human Genetics 2021 (ASHG 2021), Gene selection by incorporating genetic networks into case-control association studies. **Poster**, Virtual, Oct. 18 - Oct. 22, 2021
- International Genetic Epidemiology Society 2021 (IGES 2021), Gene-based association tests using GWAS summary statistics and incorporating eQTL. **Poster**, Virtual, Oct. 13 - Oct. 16, 2021
- IEEE 3rd International Conference on Big Data Analysis (ICBDA), The evolution of network topology structure of Chinese stock market. **Talk**, Shanghai, China, Mar. 9 - Mar. 12, 2018
- The 1st Chinese System Sciences Conference (CSSC2017), The structure of network topology of Chinese stock market. **Talk**, Beijing, China, May 13 - May 14, 2017

COMPUTER SKILLS

- Experience in UNIX and Windows operating system, as well in High Performance Computing (HPC).
- Proficient in Statistical Data Analysis using R, Python and MATLAB.
- Proficient in Genome Data Analysis (WGS, Dap-Seq, R-loop) using plink, QCtool and Samtools.
- Proficient in RNA-Seq Data Analysis using Tophat2, Bowtie2 and Cufflink.
- Proficient in Upstream and Downstream regulatory Analysis using DAVID and IPA.

AWARDS AND HONORS

- Outstanding Scholarship Award — Spring 2023
- Doctoral Finishing Fellowship — Spring 2023
- Portage Health Foundation (PHF) Graduate Assistantship — Fall 2022
- Health Research Institute (HRI) Graduate Fellowship — Summer 2022
- Outstanding research award graduate student at Department of Mathematical Sciences — Spring 2022
- Outstanding academic achievement in Advanced Topics in Statistics (MA6700), Mathematical Statistics II (MA5712) — Spring 2019
- Outstanding academic achievement in Linear Algebra (MA4330), Linear Models (MA5731), Mathematical Statistics I (MA5711) — Fall 2018
- Outstanding graduates of Beijing City and Beijing Normal University — June 2018
- Outstanding graduates of Heilongjiang University — June 2015