

Chong Wu

413 Statistics Building, 117 N. Woodward Ave.
P.O. Box 3064330, Tallahassee, FL, 32306-4330
✉: cwu3@fsu.edu; ☎: (612) 516-2222
🌐: <https://wuchong.org>

RESEARCH INTERESTS

Statistical genomics (TWAS, GWAS, multi-omics, integrative analysis, polygenic risk score)
Machine learning (unsupervised learning, supervised learning)
Causal inference (Mendelian randomization, instrumental variables regression)

EMPLOYMENT

Assistant Professor Aug. 2018–Present
Department of Statistics, Florida State University

EDUCATION

University of Minnesota Minneapolis, MN
Ph.D., Biostatistics Jun. 2018
Advisors: Drs. Weihua Guan & Wei Pan
Thesis: Statistical methods for high-dimensional genetic and genomic data

Huazhong University of Science & Technology Wuhan, China
B.S., Applied Mathematics Jun. 2013
Advisor: Dr. Xiaoyang Zhou

PEER-REVIEWED ARTICLES

* Corresponding author; † Co-first author; ‡ Students

—Theory & Methods

1. He, Y., Xu, G., **Wu, C.**, & Pan, W. (2021). [Asymptotically independent U-statistics in high-dimensional testing](#). *Annals of Statistics*, 49(1), 154–181.
2. **Wu, C.***, Bradley, J., Li, Y., Wu, L., & Deng, H. (2021). [A gene-level methylome-wide association analysis identifies novel Alzheimer’s disease genes](#). *Bioinformatics*, 37(14), 1933–1940.
3. Bae Y.‡, Wu, L., & **Wu, C.*** (2021). [InTACT: An adaptive and powerful framework for joint-tissue transcriptome-wide association studies](#). *Genetic Epidemiology*, 1–12.
4. **Wu, C.***, Xu, G., Shen, X., & Pan, W.* (2020). [A regularization-based adaptive test for high-dimensional generalized linear models](#). *Journal of Machine Learning Research*, 21, 1–67. (This paper won the 2019 ENAR Distinguished Student Paper Award.)
5. Xue, H., **Wu, C.**, & Pan, W. (2020). [Leveraging existing GWAS summary data to improve power for a new GWAS](#). *Genetic Epidemiology*, 44(7), 717–732.
6. **Wu, C.** (2020). [Multi-trait genome-wide analyses of the brain imaging phenotypes in UK Biobank](#). *Genetics*, 215(4), 947–958. (This paper won a platform presentation at the American Society of Human Genetics (ASHG) 2019 Annual Meeting; highlighted in *Genetics* August issue)
7. Yang, T., **Wu, C.**, Wei, P., & Pan, W. (2020). [Integrating DNA sequencing and transcriptomic data for association analyses of low-frequency variants and lipid traits](#). *Human Molecular Genetics*, 29(3), 515–526.

8. Yang, T., Kim, J., **Wu, C.**, Ma, Y., Wei, P., & Pan, W. (2020). [An adaptive test for meta-analysis of rare variant association studies](#). *Genetic Epidemiology*, 44(1), 104–116.
9. **Wu, C.***, & Pan, W.* (2020). [A powerful fine-mapping method for transcriptome-wide association studies](#). *Human Genetics*, 139, 199–213.
10. **Wu, C.***, & Pan, W.* (2019) [Integration of methylation QTL and enhancer-target gene maps with schizophrenia GWAS summary results identifies novel genes](#). *Bioinformatics*, 35(19), 3576–3583.
11. **Wu, C.***, Xu, G., & Pan, W.* (2019). [An adaptive test on high dimensional parameters in generalized linear models](#). *Statistica Sinica*, 29, 2163–2186.
12. **Wu, C.**, & Pan, W. (2018). [Integration of enhancer-promoter interactions with GWAS summary results identifies novel schizophrenia-associated genes and pathways](#). *Genetics*, 209(3), 699–709.
13. **Wu, C.**, & Pan, W. (2018). [Integrating eQTL data with GWAS summary statistics in pathway-based analysis](#). *Genetic Epidemiology*, 42(3), 303–316.
(Highlighted by International Genetic Epidemiology Society (IGES). This paper won a poster talk at the ASHG 2017 Annual Meeting.)
14. Park, J.Y., **Wu, C.**, Basu, S., McGue, M., & Pan, W. (2018). [Adaptive SNP set association testing in generalized linear mixed models with application to family studies](#). *Behavior Genetics*, 48(1), 55–66.
15. Xu, Z., **Wu, C.**, Wei, P., & Pan, W. (2017). [A powerful framework for integrating eQTL and GWAS summary data](#). *Genetics*, 207(3), 893–902.
16. Liu, B., **Wu, C.**, Shen, X., & Pan, W. (2017). [A novel and efficient algorithm for de novo discovery of mutated driver pathways](#). *Annals of Applied Statistics*, 11(3), 1481–1512.
17. Xu, Z., **Wu, C.**, Pan, W., & Alzheimer’s Disease Neuroimaging Initiative (ADNI). (2017). [Imaging-wide association study: Integrating imaging endophenotypes in GWAS](#). *NeuroImage*, 159, 159–169.
(This paper won a platform presentation at the ASHG 2017 Annual Meeting.)
18. **Wu, C.†**, Kwon, S.†, Shen, X., & Pan, W. (2016). [A new algorithm and theory for penalized regression-based clustering](#). *Journal of Machine Learning Research*, 17(188), 1–25.
19. **Wu, C.**, Chen, J., Kim, J., & Pan, W. (2016). [An adaptive association test for microbiome data](#). *Genome Medicine*, 8(1), 1–12.
(This paper won the 2016 Joint Statistical Meetings (JSM) Distinguished Student Paper Award on Statistics in Genomics and Genetics Section.)
20. **Wu, C.**, Demerath, E. W., Pankow, J. S., Bressler, J., Fornage, M., Grove, M. L., Chen, W., & Guan, W. (2016). [Imputation of missing covariate values in epigenome-wide analysis of DNA methylation data](#). *Epigenetics*, 11(2), 132–139.

—Applications & Collaborations

21. Xie, T., Xu, C., Shi, X., **Wu, C.**, Meng, R., Meng, X., Yu, G., Wang, K., Xiao, H., & Deng, H. (2021). [Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images](#). *Nature Communications*, 12, 6311.
22. **Wu, C.***, Zhu, J., King, A.‡, Tong, X., Lu, Q., Park, J. Y., ... & Wu, L.* (2021). [Novel strategy for disease risk prediction incorporating predicted gene expression and DNA methylation data: a multi-phased study of prostate cancer](#). *Cancer Communications*, 1–11.

23. Wu, L.*, Zhu, J., Liu, D., Sun, Y., & **Wu, C.*** (2021). [An integrative multiomics analysis identifies putative causal genes for COVID-19 severity](#). *Genetics in Medicine*, 1–11.
24. **Wu, C.***, Wu, L., Wang, J., Lin, L., Li, Y., Lu, Q., & Deng, H. (2021). [Systematic identification of modifiable risk factors and drug repurposing options for Alzheimer's disease: Mendelian randomization analyses](#). *Alzheimer's Dement*, 7:e12148.
25. Sun, Y., Zhou, D., Rahman, R., **Wu, C.**, Zhu, J., Cox NJ., Beach TG., Wu, C., Gamazon, ER., & Wu, L. (2021). [A transcriptome-wide association study identifies novel blood-based gene biomarker candidates for Alzheimer's disease risk](#). *Human Molecular Genetics*, ddab229.
26. Sun, Y., Zhu, J., Zhou, D., Canchi, S., **Wu, C.**, Cox NJ., Rissman, RA. Gamazon, ER., & Wu, L. (2021). [A transcriptome-wide association study of Alzheimer's disease using prediction models of related tissues identifies novel candidate susceptibility genes](#). *Genome Medicine*, 13(1), 1–11.
27. Wang, K. S., Yu, G., Xu, C., ..., **Wu, C.**, ... & Deng, H. W. (2021). [Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence](#). *BMC Medicine*, 19(1), 1–12.
28. Liu, D., Zhu, J., Zhou, D., ..., **Wu, C.**, ..., & Wu, L. (2021). [A transcriptome-wide association study identifies novel candidate susceptibility genes for prostate cancer risk](#). *International Journal of Cancer*, 1–11.
29. Zhu, J., **Wu, C.**, & Wu, L. (2021). [Associations between genetically predicted protein levels and COVID-19 severity](#). *The Journal of Infectious Diseases*, 223(1), 19–22.
30. Liu, D., Zhou, D., Sun, Y., Zhu, J., Ghoneim, D., **Wu, C.**, Yao, Q., Gamazon, E.R., Cox, N.J., & Wu, L. (2020). [A transcriptome-wide association study identifies candidate susceptibility genes for pancreatic cancer risk](#). *Cancer Research*, 80(20), 4346–4354.
31. Wu L., Yang Y., Guo X., ..., **Wu C.**, ..., & Long, J., The Practical Consortium (2020). [An integrative multi-omics analysis to identify candidate DNA methylation biomarkers related to prostate cancer risk](#). *Nature Communications*, 11(1), 1–11.
32. Song, M., Greenbaum, J., Luttrell IV, J., Zhou, W., **Wu, C.**, Shen, H., Gong P., Zhang C., & Deng, H. W. (2020). [A review of integrative imputation for multi-omics datasets](#). *Frontiers in Genetics*, 11.
33. Zhu, J., Shu, X., Guo, X., ..., **Wu, C.**, ..., & Wu L. (2020). [Associations between genetically predicted blood protein biomarkers and pancreatic cancer risk](#). *Cancer Epidemiology, Biomarkers & Prevention*, 29(7), 1501–1508.
34. Zhu, L., Li, Y., Chen, Y., Carrera, C., **Wu, C.**, & Fork, A. (2018). [Comparison between two post-dentin bond strength measurement methods](#). *Scientific Reports*, 8(1), 2350.
35. Bose, M., **Wu, C.**, Pankow, J. S., Demerath, E. W., Bressler, J., Fornage, M., Grove, M. L., Mosley, T. H., Hicks, C., North, K., Kao, W. H., Zhang, Y., Boerwinkle, E., & Guan, W. (2014). [Evaluation of microarray-based DNA methylation measurement using technical replicates: the Atherosclerosis Risk In Communities \(ARIC\) Study](#). *BMC Bioinformatics*, 15(1), 1–10.

—Peer-reviewed Proceedings

36. Park, J. Y., **Wu, C.**, & Pan, W. (2018). [An adaptive gene-level association test for pedigree data](#). *BMC Genetics*, 19(1), 68.
37. **Wu, C.**, Park, J.Y., Guan, W., & Pan, W. (2018). [A powerful gene-based test for methylation data](#). *BMC Proceedings*, 12(9), 60.

—Book Chapter

38. **Wu, C.** (2021). Using R for cell-type composition imputation in epigenome-wide association studies. *Epigenome-Wide Association Studies*, Springer, in production.

SUBMITTED MANUSCRIPTS

* Corresponding author; † Co-first author; ‡ Students

39. Zhang, Z.‡, Bae, Y.‡, Bradley, J., Wu, L., & **Wu, C.*** (2021+). SUMMIT: An integrative approach for better transcriptomic data imputation improves causal gene identification. *Nature Communications*. Under review.
(An earlier version won a poster talk (top 10% of all posters; Reviewers' Choice) at ASHG 2021).
40. King, A.‡, Wu, L., Deng, HW., & **Wu, C.*** (2021+). Polygenic risk score improves the accuracy of a clinical risk score for coronary artery disease. *JAMA Cardiology*. Under review.
41. Guo, X., Wei, W., Liu, M., Cai, T., **Wu, C.**, & Wang, J. (2021+). Assessing heterogeneous risk of type 2 diabetes associated with statins usage: Evidence from electronic health record data. *Journal of the American Statistical Association*. Under review. An earlier version is on [arXiv](#).
42. Sun, Y., Zhu, J., Yang, Y., Nowakowski, R., Long, J., **Wu, C.**, & Wu, L. (2021+). Identification of candidate DNA methylation biomarkers related to Alzheimer's disease risk by integrative multi-omics analysis. Submitted.
43. Xu X., Masca A., Zhu J., Yang T., Ghoneim D., Surendran P., Liu T., Platz E., Yao Q., Liu T., Fahle S., Butterworth A., **Wu C.***, & Wu L*. Associations between genetically predicted blood protein levels and pancreatic ductal adenocarcinoma risk: a study using protein genetic prediction models. Submitted.
44. Sun, Y.†, Bae, YE.†‡, Zhu, J., **Wu, C.***, & Wu, L.* (2021+) A splicing transcriptome-wide association study identifies novel altered splicing in Alzheimer's disease susceptibility. Submitted.
45. Wei, W., Laan, M., **Wu, C.***, & Wang, J.* (2021+). Efficient targeted learning of treatment effects for multiple subgroups in observational studies. To be submitted.
46. Zhu J., Xu X., Walker K., Ghoneim D., Zhang Z.‡, Surendran P., Fahle S., Butterworth A., **Wu, C.***, & Wu, L.* (2021+) Associations between genetically predicted protein concentrations in plasma and Alzheimer's disease risk. To be submitted.
47. Li, Y., Kang, J., **Wu, C.**, Dinov, I., Chalise, P., & Mahnken, J. (2021+). A machine-learning approach for detection of local brain networks and marginally weak signals identifies novel AD/MCI differentiating connectomic neuroimaging biomarkers. To be Submitted. An earlier version is on [bioRxiv](#).
48. Hidalgo, B., Lent, S., ..., **Wu, C.**, ..., Hivert, M.F., & Pankow J.S. for the CHARGE consortium (2021+). DNA methylation associated with glycemic traits and type 2 diabetes in multi-ethnic meta-analyses: CHARGE Consortium. *Nature Communications*. In revision.

GRANTS

—Undergoing

- [Novel statistical methods for multi-omics data integration in Alzheimer's disease](#) (Wu, C and Bradley, J) R03
National Institutes of Health (NIH) 01/01/21–12/31/22
NIA 1R03AG070669 Total cost: \$295,903

- Role: **Contact PI** 20% efforts
- Trans-omics integration of multi-omics studies for osteoporosis: Administrative supplement for COVID-19 studies (Deng, HW) U19
NIH 09/10/2021–03/31/2022
Role: Subcontract PI 5% efforts

–Pending

- Uncovering causal protein markers to improve prostate cancer etiology understanding and risk prediction in Africans and Europeans (Wu, L and Wu, C) **R01**
NIH, NCI 04/01/2022–03/31/2027
Role: **PI** 15% efforts in Years 1–2, 20% efforts in Years 3–5
Status: Pending Council Review
- Discovering causal DNA methylation and protein markers for Alzheimer’s disease (Wu, C and Wu, L) **R01**
NIH, NIA 04/01/2022–03/31/2027
Role: **Contact PI** 25% efforts
Status: Will resubmit
- A comprehensive Bridge2AI data generation project (ACB-DGP) for health care and genomics (Deng, HW)
NIH, Bridge2AI 02/01/2022–01/31/2026
Role: Co-Investigator 5% efforts
Status: Pending IRG review

—Completed

- Novel statistical methods for transcriptome-wide association studies
[First Year Assistant Professor Grant](#), Florida State University 05/08/19–08/06/19
Role: Contact PI Total cost: \$20,000
- Novel machine learning methods for Alzheimer’s disease
[The Committee on Faculty Research Support](#), Florida State University 05/07/20–06/30/20
Role: Contact PI Total cost: \$14,000

SOFTWARE

- Owner and maintainer for the following R packages:
 - [prclust](#): Penalized Regression-Based Clustering Methods; 21K downloads
 - [MiSPU](#): Microbiome Based Sum of Powered Score (MiSPU) Tests; 21K downloads
 - [aSPU2](#): A New Version of Adaptive Sum of Powered Score (aSPU) Test;
 - [GLMaSPU](#): Adaptive Tests on High Dimensional Parameters in GLMs; 19K downloads
 - [glmtp](#): Truncated Lasso Regularized Generalized Linear Models; 14K downloads
- Online software tutorial: <http://wuchong.org/software.html>
- Lab GitHub Repo: <https://github.com/ChongWuLab>

TEACHING EXPERIENCE

- [STA 4321/5323 Introduction to Mathematical Statistics](#)

- Instructor, Department of Statistics, Florida State University Jan. 2021–April. 2021
- Instructor, Department of Statistics, Florida State University Jan. 2020–May 2020
- Instructor, Department of Statistics, Florida State University Aug. 2019–Dec. 2019
- Instructor, Department of Statistics, Florida State University Jan. 2019–May 2019
- Instructor, Department of Statistics, Florida State University Aug. 2018–Dec. 2018
- [STA 5934 Statistical Genetics](#) (Topic course, redesigned by Wu, C)
 - Instructor, Department of Statistics, Florida State University Aug. 2020–Dec. 2020
 - Instructor, Department of Statistics, Florida State University Aug. 2021–Dec. 2021
- PUBH 6451 Biostatistics (Instructor: Dr. William Thomas)
 - Lab Instructor, Division of Biostatistics, University of Minnesota Jan. 2014–May 2014
- PUBH 6414 Biostatistical Methods (Instructor: Dr. Ann M. Brearley)
 - Teaching Assistant, Division of Biostatistics, University of Minnesota Sept. 2013–Dec. 2013

STUDENTS

- Ph.D. students who passed the qualifying exam: Shengjie Jiang, Ye Eun Bae, Zhuo Meng, Austin King, Hunter Melton
- Other Ph.D. students: Zichen Zhang, Madison Layfield

PRESENTATIONS

Most of my presentation slides can be found at <https://wuchong.org/talks.html>.

- “An adaptive association test for microbiome data”
 - Presentation at Eastern North American Region (ENAR) 2016 Spring Meeting
Austin, TX Mar. 2016
 - Presentation at 2016 Joint Statistical Meetings (JSM)
Chicago, IL Aug. 2016
- “Iterative PCA in epigenome-wide association studies”
 - Poster presented at American Society of Human Genetics (ASHG) 2016 Annual Meeting
Vancouver, BC, Canada Oct. 2016
- “A gene-level adaptive association test for methylation data”
 - Presentation at Genetic Analysis Workshop (GAW) 20
San Diego, CA Mar. 2017
- “An adaptive test on high dimensional parameters in GLMs”
 - Presentation at ENAR 2017 Spring Meeting
Washington, DC Mar. 2017
 - Poster presented at MSI Research Exhibition
Minneapolis, MN Apr. 2017
 - Presentation at 2017 JSM
Baltimore, MD Aug. 2017
- “Imaging-wide association study: Integrating imaging endophenotypes in GWAS”
 - Invited presentation at 2017 JSM (On behalf of Dr. Wei Pan)
Baltimore, MD Aug. 2017
 - Invited presentation at Third Annual Kliakhandler Conference (On behalf of Dr. Wei Pan)

- Houghton, MI Aug. 2017
- “Integrating eQTL data with GWAS summary statistics in pathway-based analysis”
 - Poster talk at ASHG 2017 Annual Meeting
Orlando, FL Oct. 2017
 - “Integration of methylation QTL and enhancer-target gene maps with schizophrenia GWAS summary results identifies novel genes”
 - Oral presentation at IGES 27th Annual Meeting
San Diego, CA Oct. 2018
 - “An adaptive test for high-dimensional generalized linear models with application to detect gene-environment interactions”
 - Presentation at ENAR 2019 Spring Meeting
Philadelphia, PA Mar. 2019
 - “Complex disease risk prediction via a deep learning method”
 - Presentation at JSM 2019
Denver, CO Jul. 2019
 - “Multi-trait genome-wide analyses of the brain imaging phenotypes in UK Biobank”
 - Platform presentation at ASHG 2019
Huston, TX Oct. 2019
 - “A powerful fine-mapping method for transcriptome-wide association studies”
 - Invited Presentation at JSM 2020
Virtual Aug. 2020
 - “Novel strategy for disease risk prediction incorporating predicted gene expression and DNA methylation: a multi-phased study of prostate cancer”
 - Department seminar at Florida State University School of Medicine
Virtual Sep. 2020
 - Invited presentation at International Indian Statistical Association (IISA) 2021
Virtual May 2021
 - “A gene-level methylome-wide association analysis identifies novel Alzheimer’s disease genes”
 - Poster presented at ASHG 2020
Virtual Oct. 2020
 - “A regularization-based adaptive test for high-dimensional generalized linear models”
 - Department seminar at Shanghai University of Finance and Economics, Statistics
Virtual Dec. 2020
 - Department seminar at Washington University School of Medicine, Biostatistics
Virtual Dec. 2020
 - “SUMMIT: An integrative approach for better transcriptomic data imputation improves causal gene identification”
 - Invited presentation at New Investigator in AD and AFAR Grantee Conference
Virtual Oct. 2021
 - “Accounting for winner’s curse and pleiotropy in two-sample Mendelian randomization”
 - Department seminar at Tulane University, Biomedical Informatics & Genomics Center
Virtual Dec. 2021
 - Department seminar at the Chinese University of Hong Kong, Statistics (scheduled)
Virtual Dec. 2021

HONORS & AWARDS

- Dean's Faculty Travel Award Apr. 2020
– Florida State University
- [ENAR Distinguished Student Paper Award](#) Mar. 2019
– Eastern North American Region International Biometric Society
- James R. Boen Student Achievement Award May 2018
– Division of Biostatistics, University of Minnesota
- Pre-Doctoral Trainee Award Oct. 2017
– Association of Chinese Geneticists in America (ACGA)
- Poster Award May 2017
– University of Minnesota Chapter of Sigma Xi
- Elected to Delta Omega (Public Health Honorary Society) May 2017
– University of Minnesota
- Elected to Sigma Xi (The Scientific Research Society) May 2017
– University of Minnesota
- [Distinguished Student Paper Award, Genomics and Genetics Section](#) Aug. 2016
– 2016 Joint Statistical Meetings
- [Doctoral Dissertation Fellowship](#) 2016–2017
– University of Minnesota
- Travel Award, Computational Neuroscience Summer School Jul. 2015
– Statistical and Applied Mathematical Sciences Institute
- Dean's Ph.D. Scholarship Sept. 2013
– University of Minnesota
- Honorable Mention in Mathematical Contest in Modeling Apr. 2012
– Consortium for Mathematics and Its Application
- National Scholarship Sept. 2011
– Ministry of Education, China

PROFESSIONAL ACTIVITIES

Manuscript Reviewer

- Statistics/Biostatistics: Biometrics, Journal of the American Statistical Association, Journal of Machine Learning Research, Statistics in Medicine, Electronic Journal of Statistics, Statistical Applications in Genetics and Molecular Biology, Statistics in Biosciences, Journal of Theoretical Biology, Biometrical Journal, Computational Statistics and Data Analysis
- Genetics/Genomics/Medicine/Bioinformatics: Genetic Epidemiology, Bioinformatics, Genome Biology, Frontiers in Genetics, Alzheimer's & Dementia, Briefings in Bioinformatics, PLOS Computational Biology
- Epidemiology: American Journal of Epidemiology
- Multidisciplinary: Nature Communications, PLOS One, Scientific Reports

Guest Editor, *Frontiers in Genetics*

Education Officer, mental health statistics section, American Statistical Association (ASA)

Invited Session Chair & Organizer (with Dr. Jingshen Wang), [Recent Advances of Causal Infer-](#)

[ences in Human Genetics](#), JSM 2021

Grant Reviewer

- NIA Special Emphasis panel, ZAG ZIJ-P J2, Functional genomics, 2020
- NCI ITCR study section, 2020