

Wenpin Hou, Ph.D.

Columbia University Mailman School of Public Health
Department of Biostatistics
Room 650, 722 West 168th Street
New York, NY 10032, United States

Email: wh2526@columbia.edu; Phone: (+1) 443-635-4567
Homepage: https://winnie09.github.io/Wenpin_Hou
Github: [Winnie09](#); Twitter: [@HWenpin](#)
Pronouns: she/her

RESEARCH INTERESTS

Methods: developing mathematical, statistical, and machine learning methods to model the spatio-temporal patterns in single-cell and spatial genomics and epigenomics data, Bayesian statistics, deep neural networks, functional data analysis
Theoretical models: gene regulatory networks (Boolean networks) controllability and inference, probability and optimization
Collaboration and scientific research: glioblastoma, non-small cell lung cancer, head and neck squamous cell carcinoma, pancreatic cancer, immunology, infectious disease, obesity, maternal and child health, health disparities

PROFESSIONAL POSITIONS

Assistant Professor (*tenure-track*), Columbia University, Department of Biostatistics, New York, NY, USA 7/2022-now

PROFESSIONAL TRAINING

Postdoctoral Fellow

Johns Hopkins University, Department of Biostatistics, Baltimore, MD, USA 8/2019-06/2022

Mentors: Dr. Stephanie C. Hicks, and Dr. Hongkai Ji

K99 Phase Mentors: Dr. Hongkai Ji, Dr. Stephanie C. Hicks, and Dr. Andrew P. Feinberg

K99 Phase Advisors: Dr. Xiaobin Wang, Dr. Kasper Hansen, and Dr. Gregory Hager

Johns Hopkins University, Department of Computer Science, Baltimore, MD, USA 8/2017-8/2019

Mentors: Dr. Aravinda Chakravarti, and Dr. Suchi Saria

Visiting Researcher

NYU Langone Health, Center for Human Genetics and Genomics, New York, NY, USA 12/2018-12/2018

MD Anderson Cancer Center, Department of Bioinformatics & Computational Biology, Houston, TX, USA 7/2017-8/2017

Kyoto University, Bioinformatics Center, Kyoto, Japan 5/2016-8/2016

Kyoto University, Bioinformatics Center, Kyoto, Japan 5/2015-8/2015

Soka University, Bioinformatics Lab, Tokyo, Japan 6/2014-8/2014

EDUCATION

Ph.D., Mathematics, The University of Hong Kong, Hong Kong, China 9/2013-8/2017

Mentor: Dr. Wai-Ki Ching

Thesis: Mathematical modelling and optimization in biological networks and data

B.Sc., Information and Computational Science, Sun Yat-sen University, Guangzhou, China 9/2009-7/2013

Advisor: Dr. Tianshou Zhou

GRANTS

NIH Pathway to Independence Award (1K99HG011468), NIH/NHGRI 3/2021 - 3/2026

Title: Computational Methods for Inferring Single-cell DNA Methylation and its Spatial Landscape

Role: principal investigator

HONORS AND AWARDS

Emerging Leaders in Computational Oncology Award, Memorial Sloan Kettering Cancer Center 2021

NIH Pathway to Independence Award (K99/R00), NIH/NHGRI 2021

2021 Women in Statistics and Data Science Conference Award, American Statistical Association 2021

Honorary Research Associate in Mathematics, The University of Hong Kong 2018-2020

First-place Winning Poster at the 13th Annual Symposium and Poster Session on Genomics and Bioinformatics 2019

University Postgraduate Fellowships, Philip K H Wong Foundation, The University of Hong Kong 2013-2017

Postgraduate Scholarship, The University of Hong Kong 2013-2017

Excellent Teaching Award, Department of Mathematics, The University of Hong Kong 2016

PUBLICATIONS *The * indicates the corresponding author(s).***Published / In journal revision****Methods for single-cell genomics**

1. **Hou, W.** and Ji, Z., 2022. Palo: Spatially-aware color palette optimization for single-cell and spatial data. [Bioinformatics](#), June 01, 2022. Package: [Palo](#). PMID: 35642896.
2. **Hou, W.** and Ji, Z., 2022. Single-cell unbiased visualization with SCUBI. [Cell Reports Methods](#), 100135, 2022. Package: [SCUBI](#). PMID: 35224531. PMCID: PMC8871596
3. **Hou, W.**, Ji, Z., Chen, Z., Wherry, E.J., Hicks, S.C. and Ji, H., 2021. A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples. [Preprint in bioRxiv](#). PMID: 34282418. PMCID: PMC8288148. Package: [Lamian](#). (In revision in *Nature Methods (Guided Open Access)*)
4. Caushi, J.X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E., Mog, B., **Hou, W.**, Justesen, S., Blosser, R., Tam, A., Anagnostou, V., Cottrell, T.R., Guo, H., Chan, H., Singh, D., Thapa, S., Dykema, A., Choudhury, C., Aparicio, L., Cheung, L., Lanis, M., Belcaid, Z., Asmar, M.E., Illei, P., Brock, M., Ha, J., Bush, E., Park, B., Bott, M., Naidoo, J., Marrone, K.A., Reuss, J.E., Velculescu, V.E., Chaff, J.E., Kinzler, K.W., Zhou, S., Vogelstein, B., Taube, J.M., Merghoub, T., Brahmer, J.R., Hellmann, M.D., Forde, P.M., Yegnasubramanian, S., Ji, H., Pardoll, D.M., Smith, K.N. (2021). Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. [Nature](#). 2021 Aug; 596(7870):126-132. Contributed to the pseudotime analysis and functional analysis (Figure 4g-h, Extended data Figure 12b-f). PMID: 34290408. PMCID: PMC8338555.
(To appear in *Nature Reviews Clinical Oncology* [Tumour antigen-induced T cell exhaustion – the archenemy of immune-hot malignancies](#))
5. **Hou, W.**, Ji, Z., Ji, H. and Hicks, S.C., 2020. A systematic evaluation of single-cell RNA-sequencing imputation methods. [Genome Biology](#) 21, 218. PMID: 32854757. PMCID: PMC7450705.
6. Ji, Z., Zhou, W., **Hou, W.** and Ji, H., 2020. SCATE: single-cell ATAC-seq signal extraction and enhancement. [Genome Biology](#), 21,161. PMID: 32620137. PMCID: PMC7333383. Packages: [SCATE](#). [SCATEData](#).

Controllability in Boolean networks

7. **Hou, W.**, Ruan, P., Ching, W.K. and Akutsu, T., 2019. On the number of driver nodes for controlling a Boolean network when the targets are restricted to attractors. [Journal of Theoretical Biology](#), 463, pp.1-11. PMID: 30543810.
8. **Hou, W.**, Tamura, T., Ching, W.K. and Akutsu, T., 2016. Finding and analyzing the minimum set of driver nodes in control of Boolean networks. [Advances in Complex Systems](#), 19(03), p.1650006. PMID: 30679639. PMCID: PMC6345816.

Glycosylation networks

9. **Hou, W.***, Qiu, Y., Hashimoto, N., Ching, W.K. and Aoki-Kinoshita, K.F., 2016. A systematic framework to derive N-glycan biosynthesis process and the automated construction of glycosylation networks. [BMC Bioinformatics](#), 17(7), p.240. PMID: 27454116. PMCID: PMC4965717.

Public health for low-income minority

10. Huang, W., Igusa, T., Wang, G., Buckley, J.P., Hong, X., Bind, E., Steffens, A., Mukherjee, J., Haltmeier, D., Ji, Y., Xu, R., **Hou, W.**, Fan, Z., and Wang, X., 2022. In-utero co-exposure to toxic metals and micronutrients on childhood risk of overweight or obesity: new insight on micronutrients counteracting toxic metals. [International Journal of Obesity](#).
11. **Hou, W.***, Zhang, M., Ji, Y., Hong, X., Wang, G., Liang, L., Ji, H., Saria, S., and Wang, X. *, 2021. A prospective birth cohort study of fetal exposure to cigarette smoking on child risk of obesity: concordance of self-report and maternal and cord blood biomarkers of cigarette smoking. [Preprints 2021](#), 2021100296.
12. Xu, R., Hong, X., Zhang, B., Huang, W., **Hou, W.**, Wang, G., Wang, X., Igusa, T., Liang, L. and Ji, H., 2021. DNA methylation mediates the effect of maternal smoking on offspring birthweight: a birth cohort study of multi-ethnic US mother–newborn pairs. [Clinical epigenetics](#), 13(1), pp.1-13. PMID: 33663600. PMCID: PMC7931602.
13. Ji, Y., Azuine, R.E., Zhang, Y., **Hou, W.**, Hong, X., Wang, G., Riley, A., Pearson, C., Zuckerman, B. and Wang, X., 2019. Association of cord plasma biomarkers of in utero acetaminophen exposure with risk of attention-deficit/hyperactivity disorder and autism spectrum disorder in childhood. [JAMA psychiatry](#), 77(2), pp.180-189.. PMID: 31664451. PMCID: PMC6822099. (Featured in [NIH news](#), [Reuters health](#), [MedPage Today](#), [meaww](#), [LinksMedicus](#), [TechnologyNetworks](#))

Machine learning for biomedical data

14. Jiang, H., Qiu, Y., **Hou, W.**, Cheng, X., Yim, M.Y. and Ching, W.K., 2018. Drug side-effect profiles prediction: from empirical to structural risk minimization. [IEEE/ACM transactions on computational biology and bioinformatics](#), 17(2), pp.402-410. PMID: 29994681.
15. Jiang, H., Ching, W.K., Cheung, W.S., **Hou, W.** and Yin, H., 2017. Hadamard kernel SVM with applications for breast cancer outcome predictions. [BMC Systems Biology](#), 11(7), p.138. PMID: 29322919. PMCID: PMC5763304.
16. Jiang, H., Ching, W.K. and **Hou, W.**, 2016. On orthogonal feature extraction model with applications in medical prognosis. [Applied Mathematical Modelling](#), 40(19-20), pp.8766-8776.
17. Qiu, Y., Cheng, X., **Hou, W.** and Ching, W.K., On classification of biological data using outlier detection, [The 12th International Symposium on Operations Research and its Applications \(ISORA\)](#), pp.144-150, Luoyang, China, 2015.
18. Cheng, X., Ching, W.K., **Hou, W.** and Kinoshita, K., A multiple linear regression model for structure of N-linked oligosaccharides, [The 12th International Symposium on Operations Research and its Applications \(ISORA\)](#), pp.151-157, Luoyang, China, 2015.
19. Cheng, X., Qiu, Y., **Hou, W.** and Ching, W.K., A semi-tensor product approach for probabilistic Boolean networks, [IEEE Proceedings of the 8th International Conference on Systems Biology \(ISB\)](#), pp.88-93, Qingdao, China, 2014.

Thesis

20. **Hou, W.** (2017) Mathematical modelling and optimization in biological networks and data. Ph.D. thesis. The University of Hong Kong.

PRESENTATIONS

Invited Talks

1. Johns Hopkins University, Department of Biostatistics. Baltimore, MD, USA. 3/2019.
2. MD Anderson Cancer Center, Department of Bioinformatics and Computational Biology. Houston, TX, USA. 7/2017
3. Kyoto University, Bioinformatics Center. Kyoto, Japan. 3/2016.
4. Kyoto University, Bioinformatics Center. Kyoto, Japan. 6/2015.
5. Soka University, Department of Bioinformatics. Tokyo, Japan. 8/2014.

Job Talks

1. University of Michigan, Department of Computational Medicine and Bioinformatics. 3/2022
2. Duke University, Department of Biostatistics and Bioinformatics. 2/2022
3. Columbia University, Department of Biostatistics. Virtual. 2/2022
4. Mount Sinai, Department of Genetics and Genomic Sciences. Virtual. 2/2022.
5. New York University, Department of Biostatistics. Virtual. 2/2022
6. UNC Chapel Hill, Department of Mathematics. Chapel Hill, NC, USA. 2/2022.
7. Weill Cornell Medicine, Department of Population Health Sciences. Virtual. 2/2022.
8. New York University, NYU Langone Institute for Systems Genetics. Virtual. 1/2022.
9. Johns Hopkins University, Department of Biostatistics, Statistical Genetics Working Group. Baltimore, MD, USA. 12/2021.
10. University of Pittsburgh, Department of Biological Sciences. Virtual. 12/2021.
11. University of Southern California, Division of Biostatistics at Keck School of Medicine. Virtual. 10/2021.
12. University of Southern California, Center for Craniofacial Molecular Biology. Virtual. 9/2021.

Conference Talks

1. Computational methods for analyzing gene expression and regulation dynamics at single-cell and spatial resolution. *Emerging Leaders Symposium in Computational Oncology*, Memorial Sloan Kettering Cancer Center (MSKCC), Virtual, 10/14-15/2021.
2. Lamian: a statistical framework for differential pseudotime analysis in multiple single-cell RNA-seq samples. *2021 Joint Statistical Meetings (JSM)*, Virtual, 8/8/2021.
3. A systematic evaluation of single-cell RNA-seq imputation methods. *ENAR 2020 Spring Meeting*, Virtual, 3/24/2020.
4. A systematic framework to derive N-glycan biosynthesis process and the automated construction of glycosylation networks. *Biotechnology and Bioinformatics Symposium (BIOT) 2015*, Provo, UT, USA, 12/10/2015.

Posters Sessions

1. A computational framework for differential pseudotime analysis across conditions with multiple single-cell RNA-seq samples reveals T cell immune dynamics associated with COVID-19 disease severity, *CSHL Systems Immunology 2021*, Virtual, 4/20-23/2021.
2. A systematic evaluation of single-cell RNA-seq imputation methods, *13th Annual Symposium and Poster Session on Genomics and Bioinformatics*, Johns Hopkins University, Baltimore, MD, USA, 10/17/2019. (awarded as *first-place winning poster*)
3. Causal gene regulatory network construction using single-cell RNA-seq and single-cell ATAC-seq data, [11th annual RECOMB/ISCB Conference on Regulatory & Systems Genomics with DREAM Challenges](#), New York University, New York, NY, USA, 12/8-10/2018.
4. On orthogonal feature extraction model with applications in cancer prediction, *Machine Learning Summer School*, University of Cadiz, Cadiz, Spain, 5/18/2016.

SOFTWARE

1. **Lamian**, R package, a systematic framework for differential pseudotime analysis in multiple single-cell RNA-seq
Github: <https://github.com/Winnie09/Lamian>
Role: lead author, developer
2. **SCUBI**, R package, a software for single-cell unbiased visualization
Github: <https://github.com/Winnie09/scubi>
Role: lead author, co-developer
3. **SCATE**, R package, a software for single-cell ATAC-seq signal extraction and enhancement
Github: <https://github.com/Winnie09/SCATE>
Role: co-author, co-developer
4. **Palo**, R package, a software for spatially-aware color palette optimization in single-cell and spatial data
Github: <https://github.com/Winnie09/Palo>
Role: lead author, co-developer

TEACHING EXPERIENCE

Classroom Instruction with *Certificate of Teaching and Learning in Higher Education*

Teaching Assistant and Lab Instructor (received *Excellent Teaching Award*), Department of Mathematics, The University of Hong Kong, Hong Kong, China 9/2013-2/2017

Responsibility: give supplement lectures, programming lessons, and exercise class twice per week; design and grade assignments; provide instruction to individual or small groups of students to improve academic performance; advise students about their academic and career development

- MATH 1011 University Mathematics (1 semester)
- MATH 2601&3601 Numerical Analysis (2 semesters)
- MATH 1211&2211 Multivariable Calculus (3 semesters)

Mentoring

Senior postdoc coordinator, Department of Biostatistics, Johns Hopkins University, Baltimore, MD, USA

- teach a Master's student to analyze scRNA-seq and scTCR-seq data of patients' pancreatic tumors in a collaborative research with Johns Hopkins Sidney Kimmel Comprehensive Cancer Center 3/2021 - 03/2022
- teach a Master's student how to develop a method for Phenotype Identification With Single-cell Gene Expression and Chromatin Accessibility Data 12/2019 - 8/2021

ACADEMIC SERVICE

Member

American Statistical Association (ASA), 2021-2022
Eastern North American Region (ENAR), 2019-2020
International Society for Computational Biology (ISCB), 2018-2019
The University of Hong Kong Shenzhen Institute of Research and Innovation (HKU-SIRI), 2016-2017
Society for Industrial and Applied Mathematics (SIAM), 2014-2018

Reviewer

Journals
Bioinformatics (4)
Communications Biology
Genome Biology
IEEE Transactions on Neural Networks and Learning Systems

Nature Communications (3)

Nature Portfolio

Neurocomputing

PLOS Computational Biology

Conference The 15th Asia Pacific Bioinformatics Conference (APBC 2017)

The First CCF Bioinformatics Conference (CBC 2016)

IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2016)

The 10th International Conference on Systems Biology (ISB 2016)

Judge

Poster Judge, 2019 Women in STEM Symposium, Baltimore, USA, 4/6/2019

Poster Judge, 2018 Women in STEM Symposium, Baltimore, USA, 4/7/2018

Panelist

CZI Faculty Search Bootcamp: Recent Jobseeker Panel, Virtual

8/8/2022