

Cankun Wang

BDSC, School of Public Health · UT Southwestern Medical Center, Dallas, TX
cankun.wang@utsouthwestern.edu · 614-378-4915 · github.com/Wang-Cankun · Google Scholar

Education

South Dakota State University, Brookings, SD · M.S., Plant Science	2019
Department of Agronomy, Horticulture & Plant Science	
Beijing Jiaotong University, Beijing, China · B.S., Software Engineering	2017
School of Software Engineering	

Professional Experience

BDSC, School of Public Health · UT Southwestern Medical Center	09/2026 - present
DATA SCIENTIST II	
- Develop algorithms, research software, and computational infrastructure for biomedical research. - Collaborate with researchers and bioinformatics teams to define computational problems and develop, test, and maintain analysis tools.	
Pelotonia Institute for Immuno-Oncology · OSU Comprehensive Cancer Center	07/2024 - 08/2026
BIostatISTICS SENIOR ANALYST	
- Led the AI infrastructure team at the Center for AI & Bioinformatics in Immuno-Oncology (CATION) at PIIO. Built agent-driven research infrastructure for PIIO investigators: conversational research agents over biomedical knowledge bases, reusable agent skill libraries, and a multi-provider LLM gateway with usage tracking and cost optimization. - Led computational analysis of bulk, single-cell, and spatial multi-omics datasets across immuno-oncology, cardiac development, lung disease, and neurodegeneration programs; contributing or co-first author on cross-institution studies in <i>Nature Biotechnology</i>, <i>Circulation</i>, <i>Circulation Research</i>, <i>European Heart Journal</i>, and <i>Immunity</i>. - Engineered full-stack research web applications and interactive biomedical databases with real-time data streaming, spatial visualization, and automated testing; production deployments include ssREAD (<i>Nature Communications</i> 2024). - Managed HIPAA/PHI-compliant clinical and multi-omics data for multi-institutional studies, including the LCMC3 neoadjuvant atezolizumab trial (Genentech); used OSU LifeScale (Databricks) to access EHR-linked clinical data for CATION research. - Onboarded bench scientists, clinicians, and computational teams to agent-driven research workflows through hands-on training, custom skill development, and integration support.	
Department of Biomedical Informatics · The Ohio State University	09/2019 – 07/2024
BIOMEDICAL INFORMATICS SPECIALIST	
- Developed bioinformatics methods for single-cell and spatial multi-omics analysis; first or co-first author on <i>iScience</i> (scREAD), <i>Nucleic Acids Research</i> (IRIS3), <i>Cancer Research Communications</i> (intratumoral microbe identification from ORIEN), and <i>Computational and Structural Biotechnology Journal</i> (DESSO-DB). - Designed and deployed interactive biomedical web databases and R/Python packages used by external research groups, including scREAD and DESSO-DB for motif and single-cell data analysis. - Built deep learning frameworks for transcriptional regulation and gene regulatory network inference from single-cell multi-omics data, contributing to scGNN (<i>Nature Communications</i>) and related methods work. - Analyzed EHR-derived clinical variables to study ICU patient outcomes, supporting cross-departmental clinical informatics collaborations. - Established and managed Azure and AWS HIPAA-compliant cloud infrastructure for departmental research and clinical projects.	
Department of Biomedical Informatics · The Ohio State University	05/2019 – 07/2019
STUDENT INTERN	
Developed an R package for identification of regulatory DNA motifs using ChIP-exo data.	

- Identified DNA transcription factor motifs using computational methods.
- Developed a web server for cell-type-specific regulon inference from single-cell RNA-seq data.

- Built test modules to evaluate DNA transcription factor motif finding efficiency and explored algorithm optimization.

- Developed software to automatically extract students' incorrect answers from scanned handwritten test papers.
- Curated training datasets for a deep learning NLP model.

Journal Publications

Full list: <https://scholar.google.com/citations?user=MCccvvgAAAAJ> · First and co-first author (*)

FIRST AND CO-FIRST AUTHOR

1. **Cankun Wang**, Diana Acosta, Megan McNutt, Jiang Bian, Anjun Ma, Hongjun Fu, and Qin Ma. "A Single-Cell and Spatial RNA-Seq Database for Alzheimer's Disease (ssREAD)." *Nature Communications* 15, no. 1 (2024): 4710.
2. **Cankun Wang**, Anjun Ma, Yingjie Li, Megan E. McNutt, Shiqi Zhang, Jiangjiang Zhu, Rebecca Hoyd, et al. "A Bioinformatics Tool for Identifying Intratumoral Microbes from the ORIEN Dataset." *Cancer Research Communications* 4, no. 2 (2024): 293–302.
3. Shiqiao Ye*, **Cankun Wang***, Zhaohui Xu*, Hui Lin, Xiaoping Wan, Yang Yu, Subhodip Adhichary, Joe Z. Zhang, Yang Zhou, Chun Liu, Matthew Alonzo, Jianli Bi, Angelina Ramirez-Navarro, Isabelle Deschenes, Qin Ma, Vidu Garg, Joseph C. Wu, and Ming-Tao Zhao. "Impaired Human Cardiac Cell Development due to NOTCH1 Deficiency." *Circulation Research* (2023). doi:10.1161/CIRCRESAHA.122.321398
4. Xiaoying Wang*, **Cankun Wang***, Lang Li, Qin Ma, Anjun Ma, and Bingqiang Liu. "DESSO-DB: A Web Database for Sequence and Shape Motif Analyses and Identification." *Computational and Structural Biotechnology Journal* 20 (2022): 3053–3058.
5. **Cankun Wang**, Yujia Xiang, Hongjun Fu, and Qin Ma. "Use of scREAD to Explore and Analyze Single-Cell and Single-Nucleus RNA-Seq Data for Alzheimer's Disease." *STAR Protocols* 2, no. 2 (2021): 100513.
6. Jing Jiang*, **Cankun Wang***, Ren Qi, Hongjun Fu, and Qin Ma. "scREAD: A Single-Cell RNA-Seq Database for Alzheimer's Disease." *iScience* 23, no. 11 (2020): 101769.
7. Anjun Ma*, **Cankun Wang***, Yuzhou Chang, Faith H. Brennan, Adam McDermaid, Bingqiang Liu, Chi Zhang, Phillip G. Popovich, and Qin Ma. "IRIS3: Integrated Cell-Type-Specific Regulon Inference Server from Single-Cell RNA-Seq." *Nucleic Acids Research* 48, no. W1 (2020): W275–W286.

SELECTED CONTRIBUTING AUTHOR

1. Liao, Guanrui, Tsuguhisa Nakayama, Bokai Zhu, et al. "Multi-Scaled Transcriptomics of Chronically Inflamed Nasal Epithelium Reveals Immune-Epithelial Dynamics and Tissue Remodeling in Nasal Polyp Formation." *Immunity* 58, no. 10 (2025): 2593–2608.e6.
2. Lobentanzer, Sebastian, Shaohong Feng, Noah Bruderer, Andreas Maier, **Cankun Wang**, Jan Baumbach, Jorge Abreu-Vicente, et al. "A Platform for the Biomedical Application of Large Language Models." *Nature Biotechnology* 43, no. 2 (2025): 166–169.
3. Yu, Yang, **Cankun Wang**, Shiqiao Ye, Hannah Qin, Matthew Alonzo, Angela Onorato, Aaron Argall, et al. "Common and Divergent Cellular Aetiologies Underlying Hypoplastic Left Heart Syndrome and Hypoplastic Right Heart Syndrome." *European Heart Journal* 46, no. 20 (2025): 1946–1949.
4. Alonzo, Matthew, Zhaohui Xu, Yang Yu, Shiqiao Ye, **Cankun Wang**, Jerry Wang, Megan McNutt, et al. "Cell-Free RNA Signatures in Maternal Blood with Fetal Congenital Heart Disease." *Circulation Research* 135, no. 10 (2024): 1021–1024.
5. Anjun Ma, Xiaoying Wang, Jingxian Li, **Cankun Wang**, Tong Xiao, Yuntao Liu, Hao Cheng, Juexin Wang, Yang Li, Yuzhou Chang, Jinpu Li, Duolin Wang, Yuexu Jiang, Li Su, Gang Xin, Shaopeng Gu, Zihai Li, Bingqiang Liu, Dong Xu, and Qin Ma. "Single-Cell Biological Network Inference Using a Heterogeneous Graph Transformer." *Nature Communications* 14, no. 1 (2023): 964.
6. Faith H. Brennan, Yang Li, **Cankun Wang**, Anjun Ma, Qi Guo, Yi Li, Nicole Pukos, Warren A. Campbell, Kristina G. Witcher, Zhen Guan, Kristina A. Kigerl, Jodie C. E. Hall, Jonathan P. Godbout, Andy J. Fischer, Dana M. McTigue, Zhigang He, Qin Ma, and Phillip G. Popovich. "Microglia Coordinate Cellular Interactions during Spinal Cord Repair in Mice." *Nature Communications* 13, no. 1 (2022): 4096.

7. SenNet Consortium. "NIH SenNet Consortium to Map Senescent Cells throughout the Human Lifespan to Understand Physiological Health." *Nature Aging* 2, no. 12 (2022): 1090–1100.

CONTRIBUTING AUTHOR

1. Acosta, Diana, **Cankun Wang**, Qin Ma, and Hongjun Fu. "Utilizing Single-Cell and Spatial RNA-Seq Database for Alzheimer's Disease (ssREAD) in Hypothesis-Driven Queries." *Neural Regeneration Research* 21, no. 2 (2026): 677.
2. Baoyinna, Baoyinna, Jinshan He, Jiaying Miao, et al. "Activation of USP30 Disrupts Endothelial Cell Function and Aggravates Acute Lung Injury through Regulating the S-Adenosylmethionine Cycle." *Advanced Science* 13, no. 2 (2026): e12807.
3. Hasanaj, Euxhen, Delphine Beaulieu, **Cankun Wang**, Qianjiang Hu, Marta Bueno, John C. Sembrat, Ricardo H. Pineda, et al. "SenSet, a Novel Human Lung Senescence Cell Gene Signature, Identifies Cell-Specific Senescence Mechanisms." *bioRxiv* (2024).
4. Lo Cascio, Filippa, Suhyeorn Park, Urmi Sengupta, Nicha Puangmalai, Nemil Bhatt, Nikita Shchankin, Cynthia Jerez, **Cankun Wang**, et al. "Brain-Derived Tau Oligomer Polymorphs: Distinct Aggregations, Stability Profiles, and Biological Activities." *Communications Biology* 8, no. 1 (2025): 53.
5. Li, Yang, Anjun Ma, Yizhong Wang, **Cankun Wang**, Shuo Chen, Hongjun Fu, Bingqiang Liu, and Qin Ma. "Enhancer-Driven Gene Regulatory Networks Inference from Single-Cell RNA-Seq and ATAC-Seq Data." *Briefings in Bioinformatics* 25, no. 5 (2024): bbae369.
6. Yu, Yang, **Cankun Wang**, Shiqiao Ye, Zhaohui Xu, Hui Lin, Karen Texter, Vasudha Shukla, et al. "Abnormal Progenitor Cell Differentiation and Cardiomyocyte Proliferation in Hypoplastic Right Heart Syndrome." *Circulation* 149, no. 11 (2024): 888–891.
7. Li, Jianying, Anjun Ma, Ruohan Zhang, Yao Chen, Chelsea Bolyard, Bao Zhao, **Cankun Wang**, et al. "Targeting Metabolic Sensing Switch GPR84 on Macrophages for Cancer Immunotherapy." *Cancer Immunology, Immunotherapy* 73, no. 3 (2024): 52.
8. Li, Yang, Yizhong Wang, **Cankun Wang**, Anjun Ma, Qin Ma, and Bingqiang Liu. "A Weighted Two-Stage Sequence Alignment Framework to Identify Motifs from ChIP-Exo Data." *Patterns* (2024).
9. Li, Yingjie, **Cankun Wang**, Anjun Ma, Abdul Qawee Rani, Mingjue Luo, Jenny Li, Xuefeng Liu, and Qin Ma. "Identification of HPV Oncogene and Host Cell Differentiation Associated Cellular Heterogeneity in Cervical Cancer via Single-Cell Transcriptomic Analysis." *Journal of Medical Virology* 95, no. 8 (2023): e29060.
10. Wang, Yizhong, Yang Li, **Cankun Wang**, Chan-Wang Jerry Lio, Qin Ma, and Bingqiang Liu. "CeMiG: Prediction of the Cis-Regulatory Motif Using the De Bruijn Graph from ATAC-Seq." *Briefings in Bioinformatics* 25, no. 1 (2024): bbab505.
11. Xiao, Tong, Johanna Schafer, No-joon Song, **Cankun Wang**, Payton Weltge, Xue Li, Qin Ma, and Zihai Li. "Novel Mechanisms of Androgen Receptor-Centered Transcriptional Regulatory Network in Regulating CD8+ T Cell Exhaustion and Sex Bias in Cancer." *The Journal of Immunology* 210, no. 1_Supplement (2023): 171.13–71.13.
12. Hussan, Hisham, Steven K. Clinton, Elizabeth M. Grainger, Maxine Webb, **Cankun Wang**, Amy Webb, Bradley Needleman, et al. "Distinctive Patterns of Sulfide- and Butyrate-Metabolizing Bacteria after Bariatric Surgery: Potential Implications for Colorectal Cancer Risk." *Gut Microbes* 15, no. 2 (2023): 2255345.
13. Jeffrey R. Atkinson, Andrew D. Jerome, Andrew R. Sas, Ashley Munie, **Cankun Wang**, Anjun Ma, William D. Arnold, and Benjamin M. Segal. "Biological Aging of CNS-Resident Cells Alters the Clinical Course and Immunopathology of Autoimmune Demyelinating Disease." *JCI Insight* 7, no. 12 (2022).
14. Shuo Chen, Diana Acosta, Liangping Li, Jiawen Liang, Yuzhou Chang, **Cankun Wang**, et al. "Wolframin Is a Novel Regulator of Tau Pathology and Neurodegeneration." *Acta Neuropathologica* (2022).
15. Shuo Chen, Yuzhou Chang, Liangping Li, Diana Acosta, Yang Li, Qi Guo, **Cankun Wang**, Emir Turkes, Cody Morrison, Dominic Julian, Mark E. Hester, Douglas W. Scharre, Chintda Santiskulvong, Sarah Xueying Song, Jasmine T. Plummer, Geidy E. Serrano, Thomas G. Beach, Karen E. Duff, Qin Ma, and Hongjun Fu. "Spatially Resolved Transcriptomics Reveals Genes Associated with the Vulnerability of Middle Temporal Gyrus in Alzheimer's Disease." *Acta Neuropathologica Communications* 10, no. 1 (2022): 188.
16. SenNet Consortium. "NIH SenNet Consortium to Map Senescent Cells throughout the Human Lifespan to Understand Physiological Health." *Nature Aging* 2, no. 12 (2022): 1090–1100.
17. Ying Li, Jianing Zhao, Zhaoqian Liu, **Cankun Wang**, Lizheng Wei, Siyu Han, and Wei Du. "De Novo Prediction of Moonlighting Proteins Using Multimodal Deep Ensemble Learning." *Frontiers in Genetics* 12 (2021).
18. Juexin Wang, Anjun Ma, Yuzhou Chang, Jianting Gong, Yuxu Jiang, Ren Qi, **Cankun Wang**, Hongjun Fu, Qin Ma, and Dong Xu. "scGNN Is a Novel Graph Neural Network Framework for Single-Cell RNA-Seq Analyses." *Nature Communications* 12, no. 1 (2021): 1882.
19. Ying Li, Qi Zhang, Zhaoqian Liu, **Cankun Wang**, Siyu Han, Qin Ma, and Wei Du. "Deep Forest Ensemble Learning for Classification of Alignments of Non-Coding RNA Sequences Based on Multi-View Structure Representations." *Briefings in Bioinformatics* (2020): bbaa354.

20. Mengting Niu, Jun Zhang, Yanjuan Li, **Cankun Wang**, Zhaoqian Liu, Hui Ding, Quan Zou, and Qin Ma. "CirRNAPL: A Web Server for the Identification of circRNA Based on Extreme Learning Machine." *Computational and Structural Biotechnology Journal* 18 (2020): 834–842.
21. Padmapriya Swaminathan, Michelle Ohrtman, Abigail Carinder, Anup Deuja, **Cankun Wang**, John Gaskin, Anne Fennell, and Sharon Clay. "Water Deficit Transcriptomic Responses Differ in the Invasive *Tamarix chinensis* and *T. ramosissima* Established in the Southern and Northern United States." *Plants* 9, no. 1 (2020): 86.
22. Juan Xie, Anjun Ma, Yu Zhang, Bingqiang Liu, Sha Cao, **Cankun Wang**, Jennifer Xu, Chi Zhang, and Qin Ma. "QUBIC2: A Novel and Robust Biclustering Algorithm for Analyses and Interpretation of Large-Scale RNA-Seq Data." *Bioinformatics* 36, no. 4 (2020): 1143–1149.
23. Brandon Monier, Adam McDermaid, **Cankun Wang**, Jing Zhao, Allison Miller, Anne Fennell, and Qin Ma. "IRIS-EDA: An Integrated RNA-Seq Interpretation System for Gene Expression Data Analysis." *PLOS Computational Biology* 15, no. 2 (2019): e1006792.
24. Yan Wang, Sen Yang, Jing Zhao, Wei Du, Yanchun Liang, **Cankun Wang**, Fengfeng Zhou, Yuan Tian, and Qin Ma. "Using Machine Learning to Measure Relatedness between Genes: A Multi-Features Model." *Scientific Reports* 9, no. 1 (2019).
25. Ye Xia, Seth Debolt, Qin Ma, Adam McDermaid, **Cankun Wang**, Nicole Shapiro, Tanja Woyke, and Nikos C. Kyrpides. "Improved Draft Genome Sequence of *Bacillus* sp. Strain YF23, Which Has Plant Growth-Promoting Activity." *Microbiology Resource Announcements* 8, no. 15 (2019).
26. Ye Xia, Seth Debolt, Qin Ma, Adam McDermaid, **Cankun Wang**, Nicole Shapiro, Tanja Woyke, and Nikos C. Kyrpides. "Improved Draft Genome Sequence of *Pseudomonas poae* A2-S9, a Strain with Plant Growth-Promoting Activity." *Microbiology Resource Announcements* 8, no. 15 (2019).
27. Jinyu Yang, Anjun Ma, Adam D. Hoppe, **Cankun Wang**, Yang Li, Chi Zhang, Yan Wang, Bingqiang Liu, and Qin Ma. "Prediction of Regulatory Motifs from Human ChIP-Sequencing Data Using a Deep Learning Framework." *Nucleic Acids Research* 47, no. 15 (2019): 7809–7824.
28. Adam McDermaid, Xin Chen, Yiran Zhang, **Cankun Wang**, Shaopeng Gu, Juan Xie, and Qin Ma. "A New Machine Learning-Based Framework for Mapping Uncertainty Analysis in RNA-Seq Read Alignment and Gene Expression Estimation." *Frontiers in Genetics* 9 (2018).
29. Siyu Han, Yanchun Liang, Qin Ma, Yangyi Xu, Yu Zhang, Wei Du, **Cankun Wang**, and Ying Li. "LncFinder: An Integrated Platform for Long Non-Coding RNA Identification Utilizing Sequence Intrinsic Composition, Structural Information and Physicochemical Property." *Briefings in Bioinformatics* (2018).

Presentations

BioChatter: Conversational AI for Biomedical Applications · Oral Presentation Education & Innovation in AI and Cancer Immunology. Columbus, OH.	2025
BioChatter: A Platform for the Biomedical Application of Large Language Models · Oral Presentation PIIO Research in Progress. Columbus, OH.	2025
A Single-Cell and Spatial RNA-Seq Database for Alzheimer's Disease (ssREAD) · Oral Presentation International Conference on Intelligent Biology and Medicine (ICIBM 2024). Houston, TX.	2024
A Weighted Two-Stage Sequence Alignment Framework to Identify DNA Motifs from ChIP-exo Data · Oral Presentation International Conference on Intelligent Biology and Medicine (ICIBM 2023). Tampa, FL.	2023
TriState Data Analysis Core: Advancements in Data Submission, Analysis, and Tool Development · Poster Presentation SenNet TriState Annual Conference 2023. Rochester, NY.	2023
Identifying Cell-Type-Specific Senescent Cells and Signature Genes Using Heterogeneous Graph Contrastive Learning · Poster Presentation SenNet Annual Meeting 2023. Minneapolis, MN.	2023
Analyze RNA-Seq Data Using IRIS-EDA Web Server · Oral Presentation Vitismet: Grapes at UC Davis & SLU. Columbus, OH.	2021
Computational Tool and Pipeline for RNA-Seq Data Analysis · Oral Presentation NSF PGRP Grape Grafting Annual Meeting. Columbus, OH.	2020
Identification of Regulatory DNA Motifs Using ChIP-exo Data · Poster Presentation BMI Internship Poster Session. Columbus, OH.	2019
Development of Regulatory Motif Identification Program · Poster Presentation BioSNTR Plant Science Research Day. Brookings, SD.	2019

Teaching

BMI 5750: Methods in Biomedical Informatics · Guest Lecturer	2026
Department of Biomedical Informatics, OSU. Topic: BioChatter as the biomedical application platform of large language models.	
BISR Bioinformatics Workshop · Lecturer	2024
Department of Biomedical Informatics, OSU.	
BISR CRISPR Screening Workshop · Lecturer	2021
Department of Biomedical Informatics, OSU.	
BISR RNA-Seq Workshop · Lecturer	2021
Department of Biomedical Informatics, OSU.	
Applications of Machine Learning and Artificial Intelligence · Guest Lecturer	2020
Department of Biomedical Informatics, OSU.	
Single-Cell Sequencing Data Analysis Workshop · Guest Lecturer	2020
Iowa State University.	
Applied Bioinformatics · Teaching Assistant (Dr. Xijin Ge)	2019
Department of Mathematics and Statistics, South Dakota State University.	

Mentoring

2026	Jonathan Yang (<i>Undergraduate, The Ohio State University</i>)
2023-2024	Siqi Lin (<i>Huazhong University of Science and Technology, China</i>)
2022-2023	Megan McNutt (<i>Senior Research Technician, OSU</i>)
2023	Grace Xu (<i>High school intern, Dublin, OH</i>)
2023	Kevin Wang (<i>High school intern, Columbus, OH</i>)
2022	Jiaxin Yang (<i>Undergraduate, OSU</i>)
2022	Mohnish Karthikeyan (<i>High school intern, Mason, OH</i>)

Awards

2022	Excellence in Team Science (Staff) · Department of Biomedical Informatics, OSU
-------------	--

Skills

AI AGENT DEVELOPMENT	LLM application design, Retrieval-Augmented Generation (RAG), knowledge graphs, multi-agent orchestration, agent skills, AI gateway management (LiteLLM), local and cloud LLM deployment.
BIOINFORMATICS & DATA ANALYSIS	Next-generation sequencing, single-cell multi-omics data analysis, high-performance computing, and reproducible workflows (R, Python, HPC systems).
CLOUD INFRASTRUCTURE	Establishment and management of HIPAA-compliant platforms, foundational biomedical AI model hosting, GPU cluster management, cost optimization (AWS, Azure).
FULL-STACK DEVELOPMENT	Full-stack web development (Vue.js, NestJS, TypeScript, MySQL).
DEVOPS	Automation and deployment (GitHub Actions, Azure DevOps, Docker).
MACHINE LEARNING & AI	Deep learning model development and deployment (ConvNet, Graph Transformer, AWS SageMaker).
EHR DATA MANAGEMENT	Clinical data lake platforms (OSU LifeScale / Databricks), multi-site cohort harmonization (e.g., LCMC3), HIPAA-compliant data workflows.
DATA VISUALIZATION	Interactive and publication-quality figures (Adobe Illustrator, ggplot2, Circos, Shiny).