

Huanhuan Wei

(203) 675-7138 | huanhuan.wei@yale.edu | www.linkedin.com/in/huanhuan-wei

SUMMARY

- Pioneered AI-driven **spatial omics** methods, notably a novel **LLM-based framework** for tissue niche identification, demonstrating deep **statistical thinking** in integrating biological knowledge with multimodal data.
- Led interdisciplinary teams to translate medical questions into actionable analytical plans, driving major discoveries in **lung cancer**, COPD, and silicosis through effective cross-functional communication and project management.
- Recipient of the competitive **Yale & Boehringer Ingelheim Fellowship**, demonstrating strong potential for **securing PI-level funding** and establishing an independent research program in **computational oncology**.

EDUCATION

Shanghai Jiao Tong University , Shanghai, China <i>Ph.D.</i> , Department of Bioinformatics and Biostatistics	2018 – 2023
University of Electronic Science and Technology of China , Chengdu, China BS in Bioinformatics	2014 – 2018

TECHNICAL SKILLS

- Programming Languages: Python, PyTorch, Scanpy, Seurat, R, Bash
- Models: Large Language Models (LLMs), Graph Neural Networks, Bayesian Networks, Causal Inference.
- Areas of Expertise: AI, Machine Learning, Spatial Transcriptomics Analysis, Single-Cell Data Analysis

RESEARCH EXPERIENCE

Yale University , New Haven, CT <i>Postdoctoral Fellow, Yale & Boehringer Ingelheim Fellowship</i>	2023 – Present
<ul style="list-style-type: none">• Pioneered a novel Large Language Model (LLM) framework to interpret cellular organization in spatial transcriptomic data. This approach achieved a 40% accuracy improvement in identifying functional cellular niches over standard methods, resulting in a first-author publication at ACL 2025.• Directed the spatial analysis of 48 Lung Adenocarcinoma biopsies in a collaboration with Boehringer Ingelheim. Designed a custom pipeline for deconvolution and niche identification that quantitatively distinguished "tumor core" from "tumor border" regions, revealing critical patterns of immune infiltration.• Spearheaded the primary statistical analysis for a large-scale COPD study (co-first author, Nature Genetics). Translated clinical hypotheses into executable statistical solutions to uncover disease heterogeneity.• Engineered a Human-in-the-Loop graphical interface for spatial annotation, enabling physician-scientists to interactively refine AI models without programming expertise. This tool significantly accelerated cross-disciplinary validation of pathological features in silicosis tissue.	

Shanghai Jiao Tong University , Shanghai, China <i>PhD Student</i>	2018 – 2023
<ul style="list-style-type: none">• Developed a novel causal inference method to infer gene regulatory networks from single-cell RNA sequencing data, leading to the discovery of 878 previously unknown regulatory pairs.	

University of Electronic Science and Technology of China , Chengdu, China <i>Undergraduate Research Assistant</i>	2016 – 2018
<ul style="list-style-type: none">• Authored a first-author review about machine learning methods, forecasting the critical role of deep learning and contextual information in bioinformatics, demonstrating early strategic vision in the field.	

MENTORSHIP EXPERIENCE

Yan Lab, Yale School of Medicine , Research Project Lead	2024 – Present
<ul style="list-style-type: none">• Mentored 4 Master's degree students, guiding them from literature review to publication-quality analysis.	

SELECTED PUBLICATIONS AND PRESENTATIONS

- **Wei H**, Luo X, Yu H, et al. Identifying Cellular Niches in Spatial Transcriptomics: An Investigation into the Capabilities of Large Language Models. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2025;1:9275–9289.
- Zhang Y*, **Wei H***, Nouws J*, et al. Aberrant Cellular Communities Define Disease Heterogeneity in Chronic Obstructive Pulmonary Disease. *Nature Genetics* (Accepted. *These authors contribute equally.)
- **Wei H**, Yan X. Computational methods to deconvolve low-resolution spatial transcriptomic data using single-cell RNA sequencing data. *SCIENTIA SINICA Mathematica*, 2025;55(7): 1453–1474.
- **Wei H**, Luo X, Abu HN, et al. Unraveling Tissue Microenvironments Using Integrated Spatial Transcriptomics and Multi- Modal Data through Graph Networks. AI for Drug Discovery Workshop, 22nd International Conference on Artificial Intelligence in Medicine, Utah, 2024. (Presentation & Poster)
- **Wei H**, Lu Hui, Zhao H. Inferring Time-Lagged Causality Using the Derivative of Single-Cell Expression. *Int J Mol Sci.* 2022;23(6):3348
- **Wei HH**, Yang W, Tang H, et al. The Development of Machine Learning Methods in Cell-Penetrating Peptides Identification: A Brief Review. *Curr Drug Metab.* 2019;20(3):217-223.