

Yuhui Hong

Department of Genome Sciences
University of Washington
William H. Foegen Hall, 3720 15th Ave NE
Seattle, WA 98195

✉ yuhhong@uw.edu
👤 josiehong.github.io
🔗 github.com/JosieHong
📄 Yuhui Hong
🆔 0000-0002-5647-9714

RESEARCH INTERESTS

I develop machine learning methods that advance the analysis and discovery of molecules from mass spectrometry data, including metabolites and drugs, with extension to broader problems in biology and medicine.

EDUCATION AND APPOINTMENTS

University of Washington Postdoctoral Scholar Advisor: Professor William Stafford Noble Area: Computational biology, Deep Learning, Proteomics mass spectrometry	Seattle, WA, US Aug. 2025 – Present
Indiana University Bloomington Ph.D. in Computer Science Advisor: Prof. Haixu Tang Dissertation: “Machine Learning Methods for Small Molecule Analysis in LC-MS/MS”	Bloomington, IN, US Sep. 2020 – Jul. 2025
Xidian University B.S. in Computer Science and Technology	Xi'an, Shaanxi, China Sep. 2015 – Jul. 2019

PUBLICATIONS

PEER-REVIEWED ARTICLES (FIRST AUTHOR)

- 1 **Yuhui Hong**, Sujun Li, Yuzhen Ye, & Haixu Tang (2025). “FIDDLE: a deep learning method for chemical formulas prediction from tandem mass spectra.” *Nature Communications*, 16(1), 11102. [\[Code\]](#) [\[PyPI package\]](#)
- 2 **Yuhui Hong**, Yuzhen Ye & Haixu Tang (2025). “Machine Learning in Small-Molecule Mass Spectrometry.” *Annual Review of Analytical Chemistry*, 18.
- 3 **Yuhui Hong**, Christopher J Welch, Patrick Piras, & Haixu Tang (2024). “Enhanced Structure-Based Prediction of Chiral Stationary Phases for Chromatographic Enantioseparation from 3D Molecular Conformations.” *Analytical Chemistry*, 96(6), 2351-2359. [\[Code\]](#)
- 4 **Yuhui Hong***, Mahsa Monshizadeh*, & Yuzhen Ye (2024). “Multitask Knowledge-primed Neural Network for Predicting Missing Metadata and Host Phenotype based on Human Microbiome.” *Bioinformatics Advances*, vbae203. (*Co-first author) [\[Code\]](#)
- 5 **Yuhui Hong**, Sujun Li, Christopher J Welch, Shane Tichy, Yuzhen Ye, & Haixu Tang (2023). “3DMolMS: Prediction of Tandem Mass Spectra from Three Dimensional Molecular Conformations.” *Bioinformatics*, btad354. [\[Code\]](#) [\[PyPI package\]](#) [\[Service on Koina\]](#)

PEER-REVIEWED ARTICLES (COLLABORATIVE AUTHOR)

- 6 Ludwig Lautenbacher, Kevin L. Yang, Tobias Kockmann, Christian Panse, Wassim Gabriel, Dulguun Bold, Elias Kahl, Matthew Chambers, Brendan X. MacLean, Kai Li, Fengchao Yu, Brian C. Searle, Damien Beau Wilburn, Mohammad Reza Zare Shahneh, **Yuhui Hong**, Haixu Tang, Mingxun Wang, Ralf Gabriels, Robbin Bouwmeester, Robbe Devreese, Jesse Angelis, Eduard Sabidó, Tobias K. Schmidt, Alexey I. Nesvizhskii, & Mathias Wilhelm (2025). “Koina: Democratizing machine learning for proteomics research.” *Nature Communications*, 16(1), 9933. [[Website](#)] [[Code](#)]

BOOKS & PATENTS

- 1 Qingyang Xiao, Kaiyuan Liu, **Yuhui Hong** & Haixu Tang (2024). “Neural Networks for Chemists.” *American Chemical Society*, DOI:10.1021/acsinfocus.7e8012. [[Primer](#)]
- 2 Haixu Tang, **Yuhui Hong**, & Sujun Li (2025). “Method of predicting MS/MS spectra and properties of chemical compounds.” US Patent No. US20250356958A1. [[Patent](#)]

ONGOING ARTICLES

- 1 Xianghu Wang, Abubaker Patan, Haoqi Nina Zhao, Vincent Charron-Lamoureux, Daniel Petras, **Yuhui Hong**, Benjamin P. Bowen, Trent Northen, Pieter C. Dorrestein, Mingxun Wang (2026). “Agentic AI for Structural Elucidation and Discovery of Drug Metabolites from Mass Spectrometry Data.” (Under review). [[Preprint](#)]
- 2 **Yuhui Hong**, & Haixu Tang (2025). “A Task-Specific Transfer Learning Approach to Enhancing Small Molecule Retention Time Prediction with Limited Data.” (Under review). [[Preprint](#)] [[Code](#)]
- 3 **Yuhui Hong***, Mahsa Monshizadeh*, & Yuzhen Ye (2025). “Confounder Free Predictive Models for Microbiome-based Host Phenotype Prediction.” (Under review, *co-first author). [[Preprint](#)] [[Code](#)]

CONFERENCE PRESENTATIONS

- 1 “De novo sequencing of chimeric spectra using Casanovo.” *74th Conference on Mass Spectrometry and Allied Topics*. May 31 – Jun. 4, 2026. San Diego, CA. (**Oral rate: 11.3%**). [[Slides](#)]
- 2 “A Task-Specific Transfer Learning Approach to Enhancing Small Molecule Retention Time Prediction with Limited Data.” *73rd Conference on Mass Spectrometry and Allied Topics*. Jun. 1–5, 2025. Baltimore, MD. (**Oral rate: 13.9%**). [[Slides](#)]
- 3 “Predicting Compositional Fragments of Compounds from Their Tandem Mass Spectra Using Deep Neural Networks.” *72nd Conference on Mass Spectrometry and Allied Topics*. Jun. 2–6, 2024. Anaheim, CA. (Poster presentation). [[Poster](#)]
- 4 “A Machine Learning Model for Chemical Formula Prediction Using Tandem Mass Spectra of Compounds.” *71st Conference on Mass Spectrometry and Allied Topics*. Jun. 4–8, 2023. Houston, TX. (**Oral rate: 10.7%**). [[Slides](#)]
- 5 “Prediction of Molecular Tandem Mass Spectra Using 3-Dimensional Conformers.” *70th Conference on Mass Spectrometry and Allied Topics*. Jun. 5–9, 2022. Minneapolis, MN. (Poster presentation). [[Poster](#)]

INVITED TALKS

- 1 **Invited talk**. “Beyond database search: machine learning method for small molecule identification.” *College of Science, San Diego State University*. Feb. 11, 2026. San Diego, CA.

- 2 **Invited talk.** “Enhanced Structure-Based Prediction of Chiral Stationary Phases for Chromatographic Enantioseparation from 3D Molecular Conformations.”
Research Group, Amgen. Oct. 11, 2024. Virtual.

TEACHING EXPERIENCE

Course Designer DSCI-D590, AI on Ramp	Indiana University Bloomington Fall 2025 – Present
Instructor DSCI-D590, Topics in Data Science	Indiana University Bloomington Spring 2025
Instructor INFO-I529, Machine Learning Bioinformatics	Indiana University Bloomington Fall 2024
Assistant Instructor DSCI-D351, Big Data Analytics (Instructor: Prof. Haixu Tang)	Indiana University Bloomington Aug. – Sep. 2024

SCHOLARSHIPS AND AWARDS

- **UW Data Science Fellow at the eScience Institute** 2025
University of Washington
Awarded \$1,000 fellowship for mass spectra data analysis
- **Luddy Outstanding Research Award** 2025
Indiana University Bloomington
Sole recipient university-wide for outstanding research contributions
- **Meritorious Winner of MCM (Mathematical Contest In Modeling)** 2018
Consortium for Mathematics and Its Application (COMAP)
Top 10% internationally (ranked among 8,085 participating teams)

PROFESSIONAL SERVICES

- **Reviewer:** (conferences) KDD 2026 2027, ACM BCB 2025 2026, EMBC 2026, NeurIPS 2026 AI for Drug Discovery Workshop; (journals) ACS Measurement Science Au, Journal of Chemical Information and Modeling, Pharmaceutical Research, Journal of Chromatography A, BMC Genomics, PLOS Computational Biology, PeerJ Computer Science, IEEE/ACM Transactions on Computational Biology and Bioinformatics, BMC Bioinformatics, Chemical Physics Letters, Beilstein Journal of Organic Chemistry.
- **Sub-reviewer:** (conferences) RECOMB 2022, 2023, 2025, 2026, ISMB/ECCB 2023, 2025, ACM BCB 2024; (journals) Analytical Chemistry, International Journal of Mass Spectrometry.
assisted in reviewing papers under the guidance of Professor William Noble and Professor Haixu Tang

Last updated: August 30, 2026