

# Cheng Wang, Ph.D.

Jinan, China, [chengwangsd@outlook.com](mailto:chengwangsd@outlook.com) (86) 135-0641-0587

Last updated: 12/12/2024

## MISSION

Explore the vast potential of small molecules and genomic signatures toward biomarker discovery and precision medicine.

## RESEARCH INTEREST

Systems biology (*omics* study), metabolomics, lipidomics, computational drug discovery, machine learning, cancer/chronic disease prevention.

## CURRENT POSITION

Shandong University, Shandong, China

06/2021-Present

Assistant Professor in Department of Biostatistics, School of Public health.

## SELECTED RESEARCH PROJECTS

### **Enzymatic reaction link learning by variational graph autoencoders**

- Developed a Variational Graph Autoencoders (VGAE) based framework to predict metabolite-protein interactions (MPI) in genome-scale heterogeneous enzymatic reaction networks across ten organisms.
- Incorporated molecular features of metabolites and proteins, as well as neighboring information in the MPI networks. The MPI-VGAE predictor achieved superior predictive performance compared to other machine learning methods.
- Applied the MPI-VGAE framework to reconstruct hundreds of metabolic pathways, functional enzymatic reaction networks, and a metabolite-metabolite interaction network, demonstrating the most robust performance across all scenarios.
- Implemented the MPI-VGAE framework to reconstruct disease-specific MPI networks based on disrupted metabolites and proteins in Alzheimer's disease and colorectal cancer, identifying numerous novel enzymatic reaction links.

### **Multi-omics integration for biomarker discovery in cardiovascular disease**

- Performed metagenomic sequencing on fecal samples and untargeted metabolomics analysis on fecal, plasma, and urine samples from ischemic stroke patients and healthy volunteers.
- Conducted differential analysis to identify key microbiota and metabolites associated with ischemic stroke.
- Utilized Spearman's rank correlation and linear regression analyses to study the associations between microbiota and metabolites in various metabolic mixtures.

- Systematic association analysis between the gut microbiome and metabolomics revealed that fecal metabolites have the strongest association with the gut microbiome, followed by urine and plasma metabolites.

#### **Accurate identification of lipids by combining NMR and MS**

- Designed a new two-dimensional (2D) NMR metabolite database, "COLMAR Lipids," specifically curated for hydrophobic metabolites. This database currently contains 501 compounds with accurate experimental 2D <sup>13</sup>C-<sup>1</sup>H chemical shift data measured in CDCl<sub>3</sub>.
- Applied nonuniform sampling in combination with pure shift spectroscopy to obtain 2D HSQC spectra with high spectral resolution along both <sup>13</sup>C and <sup>1</sup>H dimensions.
- Integrated the new 2D <sup>13</sup>C-<sup>1</sup>H HSQC lipid database into COLMAR, enabling the unique identification of numerous lipids present in complex lipidomics mixtures, such as Caco-2 cell and lung tissue cell extracts.

#### **Accurate and efficient determination of unknown metabolites in metabolomics by NMR-based cheminformatics approaches**

- Developed the SUMMIT method for untargeted metabolite identification by combining ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) with 3D nuclear magnetic resonance (NMR).
- Extracted individual compound information from NMR spectra using a maximal clique algorithm and derived all possible metabolite candidates from FT-ICR MS data.
- Demonstrated the method's performance for untargeted analysis of both a model mixture and *E. coli* cell lysate using 2D/3D NMR experiments in combination with Fourier transform ion cyclotron resonance MS and MS/MS data.

### **PROFESSIONAL EXPERIENCE**

**Washington University in St. Louis**, St. Louis, MO 06/2020-06/2021  
**Postdoctoral Researcher**

**Insight Data Science**, San Francisco, CA 01/2020-05/2020  
**Health Data Science Fellow**

**The Ohio State University**, Columbus, OH 05/2016-12/2019  
**Graduate Research Assistant**

### **EDUCATION**

**The Ohio State University**, Columbus, OH  
**Ph.D. Chemistry** (Bioinformatics) 2014-2019  
 Advisor: Prof. Rafael Bruschweiler

**Indiana University**, Bloomington, IN  
**M.S. Data Science** (Machine learning)  
Advisor: Ying Ding

2017-2019

**China University of Petroleum**, Qingdao, China  
**B.S. Applied Chemistry** (Cheminformatics)

2009-2013

## **RESEARCH FUNDING AND GRANTS**

1. **PI, 300,000 RMB, National Natural Science Foundation of China** 2024-2026  
"Research on Metabolomic Pseudo-mass Spectrometry Imaging Feature Screening Method Based on Variational Self-coding Nonlinear Dimensionality Reduction" (Active)
2. **PI, 150,000 RMB, Natural Science Foundation of Shandong Province** 2023-2025  
"Research on Structural Identification of Metabolites Based on Density Functional Theory and Deep Learning" (Active)
3. **PI, 500,000 RMB, Future Young Schoar of Shandong University** 2022-2026  
"Integrative Multi-omics Analysis by Deep Learning" (Active)
4. **Co-PI, 200,000 RMB, "2030 Science and Technology Innovation - 'Brain Science and Neuromimetic Research' Major Project"** 2022-2026  
"Alzheimer's Disease and Other Dementia Clinical Cohort Study"(Active)
5. **PI, 150,000 RMB, Shandong First-Medical University** 2022-2024  
"Investigation of Gastric Biomarkers by Mass Spectrometry Imaging" (Completed)

## **PUBLICATIONS**

Google scholar: <https://scholar.google.com/citations?user=UAZhchQAAAAI&hl=en>

### **First/Corresponding author papers**

1. **Wang, C.,\*** Yuan, C.,<sup>+</sup> Wang, Y.,<sup>+</sup> Shi, Y., Zhang, T. and Patti, G.J.,\* 2024. Prediction of Collision Cross-Section Values by Multimodal Graph Attention Network for Accurate Identification of Small Molecules. *Journal of Chemical Information and Modeling* (Cover Article).
2. **Wang, C.,** Yuan, C., Wang, Y., Chen, R., Shi, Y., Zhang, T., Xue, F., Patti, G.J.,\* Wei, L.\* and Hou, Q.,\* 2023. MPI-VGAE: protein-metabolite enzymatic reaction link learning by variational graph autoencoders. *Briefings in Bioinformatics*, p.bbada189.
3. **Wang, C.,** Timári, I., Zhang, B., Li, D.W., Leggett, A., Amer, A.O., Bruschweiler-Li, L., Kopec, R.E. and Brüscheiler, R.,\* 2020. COLMAR Lipids Web Server and Ultrahigh-Resolution Methods for Two-Dimensional Nuclear Magnetic Resonance-and Mass Spectrometry-Based Lipidomics. *Journal of proteome research*, 19(4), pp.1674-1683.
4. **Wang, C.,** Zhang, B., Timári, I., Somogyi, Á., Li, D.W., Adcox, H.E., Gunn, J.S., Bruschweiler-Li, L. and Brüscheiler, R.,\* 2019. Accurate and efficient determination of unknown metabolites in metabolomics by NMR-based molecular motif identification. *Analytical chemistry*, 91(24), pp.15686-15693.

5. **Wang, C.**,<sup>+</sup> He, L.,<sup>+</sup> Li, D.W.,<sup>+</sup> Bruschweiler-Li, L., Marshall, A.G. \* and Brüscheweiler, R.,\* 2017. Accurate identification of unknown and known metabolic mixture components by combining 3D NMR with fourier transform ion cyclotron resonance tandem mass spectrometry. *Journal of proteome research*, 16(10), pp.3774-3786.
6. Zhao, L.,<sup>+</sup> **Wang, C.**,<sup>+</sup> Peng, S., Zhu, X., Zhang, Z., Zhao, Y., Zhang, J., Zhao, G., Zhang, T.,\* Heng, X.\* and Zhang, L.,\* 2022. Pivotal interplays between fecal metabolome and gut microbiome reveal functional signatures in cerebral ischemic stroke. *Journal of Translational Medicine*, 20(1), pp.1-15.
7. Leggett, A.,<sup>+</sup> **Wang, C.**,<sup>+</sup> Li, D.W., Somogyi, A., Bruschweiler-Li, L. and Brüscheweiler, R.,\* 2019. Identification of unknown metabolomics mixture compounds by combining NMR, MS, and cheminformatics. *Methods in enzymology* (Vol. 615, pp. 407-422). Academic Press.
8. Wang, B.,<sup>+</sup> **Wang, C.**,<sup>+</sup> and Hanks, A.,\* 2024. Where Are They and How Do They Perform? Measuring Long-term Career Outcomes of Public Health Doctoral Recipients, *Journal of Public Health*, fdae031.
9. Shi, Y.,<sup>+</sup> Xu, B.,<sup>+</sup> Wang, Z.,<sup>+</sup> Chen, Q., Chai, J. \*, & **Wang, C.\***. PhenoMultiOmics: An enzymatic reaction inferred multi-omics network visualization web server. *Bioinformatics*, 2024, 40(11), p.btae623.
10. Lin, X., Lv, J., Zhang, S., Ma, X., Zhang, X., **Wang, C.\*** and Zhang, T.\*, 2024. Healthy Lifestyle Behaviors Attenuate the Effect of Poor Sleep Patterns on Chronic Kidney Disease Risk: A Prospective Study from the UK Biobank. *Nutrients*, 16(23), p.4238.
11. Xu B, Shi Y, Yuan C, Wang Z, Chen Q, **Wang C\***, Chai J.\* Integrated gene-metabolite association network analysis reveals key metabolic pathways in gastric adenocarcinoma. *Heliyon*. 2024 Sep 15;10(17).
12. Che J, Zhao Y, Gu B, Li S, Li Y, Pan K, Sun T, Han X, Lv J, Zhang S, Fan B., Li, C., **Wang,C.,** \* Wang, J. \* and Zhang, T., \* 2023. Untargeted serum metabolomics reveals potential biomarkers and metabolic pathways associated with the progression of gastroesophageal cancer. *BMC cancer*. Dec 15;23(1):1238.
13. Chen, R., Li, X., Yang, Y., Song, X., **Wang, C.\*** and Qiao, D.,\* 2022. Prediction of protein-protein interaction sites in intrinsically disordered proteins. *Frontiers in Molecular Biosciences*, 9.

#### Co-author papers

14. Lv, J., Pan, C., Cai, Y., Han, X., **Wang, C.**, Ma, J., Pang, J., Xu, F., Wu, S., Kou, T. and Ren, F., 2024. Plasma metabolomics reveals the shared and distinct metabolic disturbances associated with cardiovascular events in coronary artery disease. *Nature Communications*, 15(1), p.5729.
15. Li, J., Lv, J., Yu, F., Zhang, Y., Wang, Y., Yan, L., Xiao, Q., Li, Q., **Wang, C.**, Wang, X. and Hou, Y., 2024. Temporal changes in plasma metabolic signatures to predict immune response of antiretroviral therapy among people living with HIV. *Journal of Medical Virology*, 96(8), p.e29798.
16. Li, S., Che, J., Gu, B., Li, Y., Han, X., Sun, T., Pan, K., Lv, J., Zhang, S., **Wang, C.** and Zhang, T., 2024. Metabolites, Healthy Lifestyle, and Polygenic Risk Score Associated with Upper Gastrointestinal Cancer: Findings from the UK Biobank Study. *Journal of Proteome Research*, 23(5), pp.1679-1688.
17. Gu, B., Zhang, S., Fan, Z., Che, J., Li, S., Li, Y., Pan, K., Lv, J., **Wang, C.**, Zhang, T. and Wang, J., 2023. Prognostic model construction and immune microenvironment analysis of esophageal cancer based on gene expression data and microRNA target genes. *Translational Cancer Research*, 12(5), p.1165.
18. Yuan, C.,\* **Wang, C.**, Zhu, K., Li, S. and Miao, Z.,\* 2022. Measles epidemiology and viral nucleoprotein gene evolution in Shandong Province, China. *Journal of Medical Virology*, 94(10), pp.4926-4933.

19. Wang, Y., Stancliffe, E., Fowle-Grider, R., Wang, R., **Wang, C.**, Schwaiger-Haber, M., Shriver, L.P. and Patti, G.J.,\* 2022. Saturation of the mitochondrial NADH shuttles drives aerobic glycolysis in proliferating cells. *Molecular cell*, 82(17), pp.3270-3283.
20. Hansen, A.L., Kupčė, E., Li, D.W., Bruschweiler-Li, L., **Wang, C.** and Brüschweiler, R.,\* 2021. 2D NMR-based metabolomics with HSQC/TOCSY NOAH supersequences. *Analytical Chemistry*, 93(15), pp.6112-6119.
21. Knobloch, T.J., Ryan, N.M., Bruschweiler-Li, L., **Wang, C.**, Bernier, M.C., Somogyi, A., Yan, P.S., Cooperstone, J.L., Mo, X., Brüschweiler, R.P. and Weghorst, C.M.,\* 2019. Metabolic regulation of glycolysis and AMP activated protein kinase pathways during black raspberry-mediated oral cancer chemoprevention. *Metabolites*, 9(7), p.140.
22. Timári, I., **Wang, C.**, Hansen, A.L., Costa dos Santos, G., Yoon, S.O., Bruschweiler-Li, L. and Brüschweiler, R.,\* 2019. Real-time pure shift HSQC NMR for untargeted metabolomics. *Analytical chemistry*, 91(3), pp.2304-2311.
23. Yuan, J., Zhang, B., **Wang, C.** and Brüschweiler, R.,\* 2018. Carbohydrate background removal in metabolomics samples. *Analytical chemistry*, 90(24), pp.14100-14104.
24. Hansen, A.L., Li, D., **Wang, C.** and Brüschweiler, R.,\* 2017. Absolute Minimal Sampling of Homonuclear 2D NMR TOCSY Spectra for High-Throughput Applications of Complex Mixtures. *Angewandte Chemie*, 129(28), pp.8261-8264.
25. Li, D.W., **Wang, C.** and Brüschweiler, R.,\* 2017. Maximal clique method for the automated analysis of NMR TOCSY spectra of complex mixtures. *Journal of biomolecular NMR*, 68(3), pp.195-202.

## **CONFERENCES AND PRESENTATIONS**

1. Oral talk, 18th Annual Conference of the Metabolomics Society, Online, 06/2021  
“Protein–metabolite Enzymatic Reaction Link Learning by Graph Neural Network”
2. Oral talk, 2nd Annual MANA Conference, Seattle, WA, 09/2020  
“COLMAR Lipids Web Server and Ultrahigh-Resolution Methods for Two-Dimensional Nuclear Magnetic Resonance-and Mass Spectrometry-Based Lipidomics”
3. Oral talk, 3rd Gateway NMR Conference, Pittsburgh, PA, 11/2018  
“Accurate Identification of Known and Unknown Metabolites by Multidimensional NMR and Customized Metabolite Database”
4. Oral talk, 14th Annual Conference of the Metabolomics Society, Seattle, WA, 06/2018  
“Accurate Identification of Known and Unknown Metabolites by Multidimensional NMR and Customized Metabolite Database”
5. Oral talk, 2nd Annual Ohio Mass Spectrometry and Metabolomics Symposium, 05/2018  
“Accurate Identification of Known and Unknown Metabolites in Gallbladder Bile by Multidimensional NMR and Customized Metabolite Database”

## **ACADEMIC SERVICE AND TEACHING EXPERIENCE**

**Guest editor:** *Metabolites* (IF:5.581), *Journal of Personalized Medicine* (IF:3.508)

**Reviewer of journals:** *Nature machine intelligence*, *Briefings in bioinformatics*, *eLife*, *BMC Bioinformatics*,

*Metabolites, Nutrients, IJMS, Applied Sciences, Journal of personalized medicine, etc.*

**Shandong University**, Jinan, China,

06/2021-present

- Instructor of data science course to biomedical data science major students.
- Lead courses including big data mining, data structure, machine learning.

**The Ohio State University**, Columbus, OH

**Graduate Teaching Assistant, Physical Chemistry**

08/2015-05/2016

- Instructor of recitation for physical chemistry course to chemical engineering major students.
- Taught basics of quantum mechanics, calculus, linear algebra and fundamentals of probability.

## **OTHER SKILLS**

**Language:** Proficient in English, oral, writing and reading, native speaker of Chinese.

**Programming Skills:** Proficient in Python, Pytorch, Keras, Tensorflow, Cloud based web application deployment.

**Data Analysis Skills:** Data visualization, Bayesian inference, machine learning, and deep learning modeling.

## **HONORS AND AWARDS**

**Food for Health Graduate Student Fellowship**, The Ohio State University

2017-2018

**Graduate Student Fellowship for Data Science Program**, Indiana University

2017-2019

**Student Travel Award**, 14th International Conference of Metabolomics Society

06/2018