

Guojie Zhong

PhD candidate, Department of Systems Biology, Columbia University

1130 St. Nicholas Avenue, New York, NY, USA | Phone: (+01)7187305993 | Email: gz2294@cumc.columbia.edu

Website: <https://zhongguojie1998.github.io/GuojieZhong/>

RESEARCH INTERESTS

My research is driven by the ultimate goal to bridge the gap between complex human diseases and effective treatments. To achieve this, I develop biologically-inspired machine learning algorithms that can be applied to high-throughput genomics data to uncover disease mechanisms.

EDUCATION

PhD in Systems Biology, Integrated program in cellular, molecular and biomedical sciences 08/2019 – present
Department of Systems Biology, **Columbia University**, New York, NY, USA.

Thesis Advisor: [Yufeng Shen](#)

Committee: Wendy Chung, Mohammad AlQurashi, David Knowles

Exchange Student in Berkeley Bioscience Study Abroad Program (BBSA) 08/2017 – 12/2017
Department of Integrated Biology, **University of California, Berkeley**, CA, USA.

Bachelor of Science, Integrated Science Program 09/2015 – 07/2019
Yuanpei College, **Peking University**, Beijing, China.

Thesis Advisor: [Zemin Zhang](#)

Thesis: Reconstruction of cell spatial organization from single-cell RNA sequencing data based on ligand-receptor mediated self-assembly

PUBLICATIONS

(# = co-first author)

Deep Learning and missense variant effect prediction

-
- ◆ **Zhong, G.**, Zhao, Y., Zhuang, D., Chung, W.K. and Shen, Y. PreMode predicts mode-of-action of missense variants by deep graph representation learning of protein sequence and structural context. *bioRxiv*, 2024.02.20.581321 (2024).
 - ◆ Zhao, Y., **Zhong, G.**, Hagen, J., Pan, H., Chung, W.K., and Shen, Y. A probabilistic graphical model for estimating selection coefficient of nonsynonymous variants from human population sequence data. *medRxiv*, 2023.12.11.23299809 (2023).

Statistical genetics and rare genetic diseases analysis

-
- ◆ **Zhong, G.**, Choi, Y.A. and Shen, Y. VBASS enables integration of single cell gene expression data in Bayesian association analysis of rare variants. *Commun Biol* **6**, 774 (2023).
 - ◆ **Zhong, G.**, and Shen, Y. Statistical models of the genetic etiology of congenital heart disease. *Curr Opin Genet Dev* **76**, 101967 (2022).
 - ◆ **Zhong, G.**, Ahimaz, P., Edwards, N.A., Hagen, J.J., Faure, C., Lu, Q., Kingma, P., Middlesworth, W., Khlevner, J., El Fiky, M., et al. Identification and validation of candidate risk genes in endocytic vesicular trafficking associated with esophageal atresia and tracheoesophageal fistulas. *HGG Adv* **3**, 100107 (2022).
 - ◆ Edwards, N., **Zhong, G.**, Ahimaz, P., Kenny, A., Kingma, P., Wells, J., Shen, Y., Chung, W., and Zorn, A. Discovering the Developmental Basis of Trachea-Esophageal Birth Defects: Evidence for Endosome-opathies. *The FASEB Journal* **36** (2022).

Single cell and spatial transcriptomics

-
- ◆ Su, J., Reynier, J.B., Fu, X., **Zhong, G.**, Jiang, J., Escalante, R.S., Wang, Y., Aparicio, L., Izar, B., Knowles, D.A., and Rabadan, R. Smoother: a unified and modular framework for incorporating structural dependency in spatial omics data. *Genome Biol* **24**, 291(2023).
 - ◆ Ren, X.#, **Zhong, G.**#, Zhang, Q., Zhang, L., Sun, Y., and Zhang, Z. Reconstruction of cell spatial organization from single-cell RNA sequencing data based on ligand-receptor mediated self-assembly. *Cell Res* **30**, 763-778 (2020).
 - ◆ Zhang, Q., He, Y., Luo, N., Patel, S.J., Han, Y., Gao, R., Modak, M., Carotta, S., Haslinger, C., Kind, D., Peet G.W., **Zhong, G.**, Lu, S., Zhu, W., Mao, Y., Xiao, M., et al. Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. *Cell* **179**, 829-845 e20 (2019).

CONFERENCES

- ◆ **Zhong, G.**, and Shen, Y. (2023). Predicting mode of action of missense variants by graph representation of protein structural context. *American Society of Human Genetics 2023 Annual Meeting, Washington, D.C.*
- ◆ **Zhong, G.**, and Shen, Y. (2022). Representation of missense variants for predicting modes of action. *Machine Learning in Structural Biology, Workshop at the 36th Conference on Neural Information Processing Systems.*
- ◆ **Zhong, G.**, Choi, Y.A., and Shen, Y. (2022). Integration of gene expression data in Bayesian association analysis of rare variants. *American Society of Human Genetics 2022 Annual Meeting, Los Angeles, CA.*
- ◆ **Zhong, G.**, Wang, J., He, S. and Fu, X. (2021). Towards better understanding of developmental disorders from integration of spatial single-cell transcriptomics and epigenomics. *The 2021 ICML Workshop on Computational Biology, Virtual meeting.*

AWARDS

Dean's fellowship, Graduate School of Arts and Science, Columbia University	2019
2014-2015 academic year Outstanding Freshman Scholarship, Peking University	2015
The 28th National Olympiad in Chemistry in Provinces, 1st prize, Zhejiang Province, China	2014

MENTORSHIP EXPERIENCES

Andrew Lee, summer visiting high school student	06/2021 – 08/2021
Department of Systems Biology	
Columbia University Irving Medical Center, New York, NY, USA	
<hr/>	
Demi Zhuang, summer visiting undergraduate student	06/2022 – 08/2022
	06/2023 – 08/2023
Department of Systems Biology	
Columbia University Irving Medical Center, New York, NY, USA	

PROFESSIONAL EXPERIENCES

Research Assistant	08/2019 – present
Department of Systems Biology	
Columbia University Irving Medical Center, New York, NY, USA	
<hr/>	
Research Assistant	03/2017 – 07/2019
Biomedical Pioneering Innovation Center (BIOPIC)	
Peking University, Beijing, China	