Zhenmiao Zhang (张真苗), Ph.D.

Postdoctoral Scholar in Computer Science, University of California, San Diego

Email: zhz142@ucsd.edu   Homepage: https://zhangzhenmiao.github.io

Research Interests

- **Computational Biology & AI for Science:** Developing novel computational and deep learning frameworks to address fundamental challenges in genomics and metagenomics.
- **Core Expertise:** De novo genome and metagenome assembly, metagenomic contig binning, comparative genomics, and deep learning for biological sequence analysis. My research bridges algorithmic innovation with modern AI to enable high-quality genome reconstruction and exploration of microbial dark matter.

Education & Academic Appointments

Postdoctoral Scholar, Department of Computer Science & Engineering   *03/2024 – Present*
University of California, San Diego (UCSD), USA
Advisor: Prof. Pavel A. Pevzner

Ph.D. student, Department of Computer Science   *09/2019 – 02/2024*
Hong Kong Baptist University, Hong Kong, China
Advisor: Prof. Lu Zhang
Thesis: *Advanced Computational Methods to Decipher Microbial Dark Matter Using Metagenomic Sequencing.*

Visiting Scholar, Department of Computer Science & Engineering   *02/2023 – 11/2023*
University of California, San Diego (UCSD), USA
Advisor: Prof. Pavel A. Pevzner

B.Eng. in Computer Science & B.Sc. in Mathematics   *09/2015 – 06/2019*
Renmin University of China (RUC), China

Publications (Selected)

Zhenmiao Zhang*, Marcus W. Fedarko, Anton Bankevich, Pavel A. Pevzner. *Haplotype-mixed assembly of long and accurate reads.* Under review at **Genome Biology**.

Bohao Zou, **Zhenmiao Zhang (co-first)**, Xiaohan Wang, Rong Tao, Nianzhen Gu, Karsten Kristiansen, Mo Han, Lu Zhang*. *CompleteBin: A transformer-based framework unlocks microbial dark matter through improved short contig binning.* Under review at **Nature Biotechnology**.

Jingjing Wang, **Zhenmiao Zhang (co-first)**, Yang Chen, Xi Zhou, Jiaxin Xiang, Chao Yang, Dmitry A Rodionov, Andrei L Osterman, Qinwei Qiu, Yusheng Deng, Yanmin Liu, Chengrui Wang, Xiaoxiao Shang, Li Huang, Chen Sun, Jianwen Guo, Zhimin Yang, Lijuan Han, Lixiang Zhai, Zhaoxiang Bian, Wei Jia*, Xiaodong Fang*, Lu Zhang*. *Exploring Population Differences in the Human Gut Microbiome: from Microbial Abundance to Single Nucleotide Polymorphisms.* **Genome Biology**, in press.

Zhenmiao Zhang*, Ishaan Gupta, Pavel A. Pevzner. *GenomeDecoder: Inferring Segmental Duplications in Highly-Repetitive Genomic Regions.* **Bioinformatics**, 2025.

T2T Consortium (contributing author). *Complete sequencing of ape genomes.* **Nature**, 2025.

Zhenmiao Zhang, Jin Xiao, Hongbo Wang, Chao Yang, Yufen Huang, Zhen Yue, Yang Chen, Lijuan Han, Kejing Yin, Aiping Lyu, Xiaodong Fang, Lu Zhang*. *Exploring high-quality microbial genomes by assembling short reads with long-range connectivity.* **Nature Communications**, 2024.

Chao Yang, **Zhenmiao Zhang (co-first)**, Yufen Huang, Xuefeng Xie, Herui Liao, Jin Xiao, Werner Pieter Veldsman, Kejing Yin, Xiaodong Fang, Lu Zhang*. *LRTK: a platform-agnostic toolkit for linked-read analysis of both human genome and metagenome.* **GigaScience**, 2024.

Zhenmiao Zhang, Chao Yang, Werner Pieter Veldsman, Xiaodong Fang, Lu Zhang*. *Benchmarking genome assembly methods on metagenomic sequencing data.* **Briefings in Bioinformatics**, 2023.

Zhenmiao Zhang, Lu Zhang*. *METAMVGL: a multi-view graph-based metagenomic contig binning algorithm by integrating assembly and paired-end graphs.* **BMC Bioinformatics**, 2021.

Conferences

- *Improve metagenomic binning by recovering short contigs using deep learning*, RECOMB 2023 (Poster).
- *Exploring high-quality microbial genomes by assembly of linked-reads with high barcode specificity using deep learning*, RECOMB-SEQ 2023 (Oral).
- *METAMVGL: a multi-view graph-based metagenomic contig binning algorithm by integrating assembly and paired-end graphs*, APBC 2021 (Oral).

Software & Algorithmic Contributions

- **Mosaic Genome Assembler** — De novo haplotype-mixed genome assembler for HiFi reads
- **GenomeDecoder** — Framework for inferring synteny blocks in complete genomes
- **CompleteBin** — Contig binning via dynamic contrastive learning with pretrained language models
- **Pangaea** — Metagenome assembly with long-range connectivity using deep variational autoencoder
- **METAMVGL** — Contig binning using multi-view graph leaning integrating assembly and PE graphs

Technical Skills

- **Programming**: C++, Python, R
- **AI/ML**: PyTorch

Teaching Experience

- COMP7250 Machine Learning
- COMP1007 Introduction to Python (multiple semesters)

Academic Service

- **Reviewer**: *Microbiome, Gut Microbes, NAR Genomics and Bioinformatics, Scientific Reports, Computational Biology and Chemistry, BMC Methods, BIBM*, and others
- **PC Member**: BIBM 2025
- **Workshop PC Member**: BIBM 2022 workshops on *Algorithmic Advances for Single-Cell and Spatial Omics Data Analysis*
- **Workshop PC Member**: BIBM 2025 workshops on *Computational methods to characterize genomic variants using high-throughput sequencing data*

Grants & Honors

- Open Project Grant (BGI-Shenzhen; main contributor): *Novel algorithms of metagenome assembly and haplotype phasing using single-tube Long Fragment Read sequencing technology*
- Excellent Teaching Assistant Performance Award (COMP1007 Python), 2020–2023
- Meritorious Winner, International Mathematical Contest in Modeling, 2018