
Haohong Zhang

Bioinformatics / Foundation models / AI for biology

Ph.D. student · Huazhong University of Science and Technology · Wuhan, China

haohongzh@gmail.com · ludenszhang.github.io · GitHub · Scholar

ORCID 0000-0001-6267-4244

Education

2023 - Present Ph.D. in Bioinformatics · Direct-track

Huazhong University of Science and Technology · Advisor: Prof. Kang Ning

2019 - 2023 Bachelor of Science

Huazhong University of Science and Technology · Dengfeng Program, Class of 1901

Selected peer-reviewed publications

MGM as a Large-Scale Pretrained Foundation Model for Microbiome Analyses in Diverse Contexts

H Zhang, Y Zhang, Z Kang, J Xiong, R Yang, K Ning · Advanced Science 13(24), e13333 (2026).

10.1002/adv.202513333

Large-scale pretraining / Microbiome / Transfer learning

Deep learning enabled integration of tumor microenvironment microbial profiles and host gene expressions for interpretable survival subtyping in diverse types of cancers

H Zhang, X Xiong, M Cheng, L Ji, K Ning · mSystems 9(12), e01395-24 (2024).

10.1128/msystems.01395-24

Multi-omics integration / Cancer survival / Interpretable deep learning

Highly accurate diagnosis of pancreatic cancer by integrative modeling using gut microbiome and exposome data

Y Zhang, H Zhang, B Liu, K Ning · iScience 27(3) (2024).

10.1016/j.isci.2024.109294

Pancreatic cancer / Gut microbiome / Exposome integration

Tracing human life trajectory using gut microbial communities by context-aware deep learning

H Zhang, H Chong, Q Yu, Y Zha, M Cheng, K Ning · Briefings in Bioinformatics 24(1), bbac629 (2023).

10.1093/bib/bbac629

Life trajectories / Gut microbiome / Context-aware learning

Honors

Outstanding Graduate, 2023 · Huazhong University of Science and Technology

Haohong Zhang

Academic CV

Preprints

MGM2 as a Unified Foundation Model for Microbiome World Exploration

H Zhang, Y Zhang, Y Qi, T Liu, R Yang, K Ning · bioRxiv (2026).

[10.64898/2026.07.20.739063](#)

[Foundation models](#) / [Microbiome](#) / [Representation learning](#)

MetaClaw: an auditable AI agent for end-to-end, multi-directional metagenomic and multi-omics analysis

H Zhang, Z Li, P N P Lagniton, Z Wang, L Zhao, W Li, P Duan, X Jiang, K Ning · bioRxiv (2026).

[10.64898/2026.07.21.739769](#)

[AI agents](#) / [Metagenomics](#) / [Reproducible workflows](#)

Conference presentation

A discrete token language of prokaryotic genomes learned by MicroVQVAE

Haohong Zhang, Kang Ning · Poster, Cold Spring Harbor Asia: AI and Biology.

April 20-24, 2026 · Suzhou, China.

[Conference program](#) · [Poster 41](#)

Research software

MGM2 [Repository](#)

A unified foundation model for microbiome representations, prediction and interpretation.

MGM [Repository](#)

Large-scale pretraining for transferable microbiome analysis across diverse contexts.

MicroVQVAE [Repository](#)

A genome foundation model that learns discrete, context-aware genome tokens from ordered prokaryotic protein sequences.

EXPERT-lightning [Repository](#)

Context-aware microbial source tracking with transfer learning and biome ontology.

DeepMicroCancer [Repository](#)

Microbiome-based models for the diagnosis of a broad spectrum of cancer types.

ASD-cancer [Repository](#)

Interpretable survival subtyping from microbial profiles and host gene expression.

Technical skills

Machine learning. Python, PyTorch, PyTorch Lightning, Hugging Face Transformers and Datasets; masked language modeling, transfer learning and autoencoders.

Biological data. Biopython, FASTA processing, genome dataset preparation, microbiome abundance profiles and discrete sequence representations.

Analysis & tooling. NumPy, pandas, SciPy, scikit-learn, scikit-bio, plotnine; PCA/PCoA, command-line tools, Python packaging and TensorBoard.