Yicheng Gao

Department of Bioinformatics, Tongji University, China

gao.yicheng.98@gmail.com https://gaoyichengtj.github.io Google Scholar

Education

Tongji University (985, 211), Shanghai, China **Sep. 2020 – Mar. 2026**
Bioinformatics PhD student supervised by Prof. Qi Liu

Technical University of Munich & Helmholtz Munich, Munich, Germany **Aug. 2025 – Feb. 2026**
Visiting Scholar supervised by Prof. Fabian J. Theis

Huazhong Agricultural University (211), Wuhan, China **Sep. 2016 – Jun. 2020**
Bachelor of Bioinformatics (with Highest Honor), supervised by Prof. Wen Zhang

Main Research Experience

Algorithm design for causal disentanglement for single-cell data **2025**
Advisor: Dr. Dongsheng Li and Dr. Caihua Shan *Mircosoft*

- Developed a new causal disentanglement representation framework for single-cell data, called CausCell.
- Proposed a new optimization objective function for this framework.
- CausCell was implemented based on PyTorch and achieved superior or comparable performance on disentanglement and reconstruction, deployed on github https://github.com/bm2-lab/CausCell.
- Paper was accepted on Nature Communications.

Algorithm design for genetic perturbation prediction **2024**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new subtask decomposition-based genetic perturbation prediction model, called STAMP.
- Proposed a three-subtasks-based benchmark evaluation strategy.
- STAMP was implemented based on PyTorch and achieved superior or comparable performance in this task, deployed on github https://github.com/bm2-lab/STAMP.
- Paper was published on Nature Computational Science.

Algorithm design for multi-modal analysis in T-cells **2023**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new low-resource-aware multi-modal representation learning for cross-modality integration and analysis of T-cell receptor and T-cell transcriptomes in a unified way, called UniTCR.
- UniTCR was used for an array of downstream tasks, including single modality analysis, modality gap analysis, epitope-TCR binding prediction and cross-modality generation task.
- UniTCR was implemented based on PyTorch and achieved superior or comparable performance in these tasks, deployed on github https://github.com/bm2-lab/UniTCR.
- Paper was published on Cell Genomics.

Algorithm design for predicting peptide-TCR binding **2022**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new meta-learning framework combined with the ideas of meta learning and neural turning machine (NTM) for tackling the peptide-TCR binding prediction task, called PanPep.
- Based on the NTM, we proposed a disentanglement distillation module for generalizing few-shot learning to the zero-shot learning.
- PanPep was implemented with PyTorch and achieved SOTA performance in this task, deployed on github https://github.com/bm2-lab/PanPep.
- Paper was published on Nature Machine Intelligence.

Selected Publications

- **Causal disentanglement for single-cell representations and controllable counterfactual generation**
  Yicheng Gao*, Kejing Dong*, Caihua Shan, Dongsheng Li, Qi Liu
  Nature Communications, 2025
- **Toward subtask decomposition-based learning and benchmarking for genetic perturbation outcome prediction and beyond**
  Yicheng Gao*, Zhiting Wei*, Kejing Dong, Jingya Yang, Guohui Chuai, Qi Liu
  Nature Computational Science (Research Highlight), 2024
- **Unified cross-modality integration and analysis of T-cell receptors and T-cell transcriptomes**
  Yicheng Gao*, Kejing Dong*, Yuli Gao, Xuan Jin, Qi Liu
  Cell Genomics (Featured Article), 2024
- **Pan-Peptide Meta Learning for T-cell receptor–antigen binding recognition**
  Yicheng Gao*, Yuli Gao*, Yuxiao Fan, Chengyu Zhu, Zhiting Wei, Chi Zhou, Guohui Chuai, Qinchang Chen, He Zhang, Qi Liu
  Nature Machine Intelligence (Research Highlight, ESI highly cited), 2023
- **Delineating the cell types with transcriptional kinetics**
  Yicheng Gao, Qi Liu
  Nature Computational Science, 2024
- **Weakly-supervised peptide-TCR binding prediction facilitates neoantigen identification**
  Yuli Gao*, **Yicheng Gao***, Siqi Wu*, Danlu Li, Chi Zhou, Fangliangzi Meng, Kejing Dong, Xueying Zhao, Ping Li, Aibin Zhang, Qi Liu
  Cell Systems, 2025
- **Weakly-supervised peptide-TCR binding prediction facilitates neoantigen identification**
  Zhiting Wei*, Duanmiao Si*, Bin Duan*, **Yicheng Gao***, Qian Yu, Ling Guo, Qi Liu
  Nucleic Acid Research, 2024

Selected Honors and Awards

- Chinese Association for Artificial Intelligence (CAAI) **"Potential Qingyuan Scholar"** (15 people in China), 2025
- Shanghai Interdisciplinary Innovation Foundation for Intelligent Science & Technology (50K RMB), 2025
- **National Natural Science Foundation of China (NSFC)** Support for PhD student (300K RMB), 2025
- **Academic Pioneer Award** at Tongji University (Top 0.05%), **ranked 1st**, 2024
- **Stars of Tomorrow** in Microsoft Research (Top 10% interns in MSR), 2024
- **National scholarship** for PhD students in China (Top 0.2%), 2024
- **National scholarship** for PhD students in China (Top 0.2%), 2023
- Tongji University **Scholarship for Outstanding Ph.D. Freshman** (Top 5%), 2020
- The 16th **'Top Ten Students'** of Huazhong Agricultural University (Top 0.1%), **ranked 1st**, 2020
- Huazhong Agricultural University 2020 **Outstanding Graduates** (Top 5%), 2020
- The 9th MathorCup University Mathematical Modeling Challenge, **First Prize**, 2019

Intern Experience

Microsoft | *AI/ML group Research Intern* **Dec. 2023 – Oct. 2024**

- Mentors: Dr. Dongsheng Li and Dr. Caihua Shan
- Disentanglement representation method design for single-cell data
- Accepted on Nature Communications

BGI Genomics | *Algorithm Engineer* **Jul. 2020 – Sep. 2020**

- Algorithm optimization for detection of structural variation

BGI Genomics | *Algorithm Engineer* **Jul. 2019 – Sep. 2019**

- Algorithm optimization for detection of SNP variation.