

Yumeng Zhang

Yumeng.Zhang1@monash.edu | [Github](#) | [Google Scholar](#) | [LinkedIn](#)

Education

Monash University 2024/09 – 2028/03 (Expected)

Ph.D. Candidate in Computational Biology, Department of Biochemistry and Molecular Biology Melbourne, VIC, Australia

- Supervisors: Jiangning Song & Tony Purcell.

Shanghai Jiao Tong University 2021/09 – 2024/06

M.Sc. in Biology, GPA 3.81/4.0 Shanghai, China

- Courses: Omics Big Data, Structural Bioinformatics, Statistical Methods in Data Science, etc.
- Supervisor: Hong-Yu Ou.

Shanghai Jiao Tong University 2017/09 – 2021/06

B.Sc. in Bioinformatics, GPA 3.58/4.0 (86.90/100) Shanghai, China

- Courses: Data Structures & Algorithms, Bioinformatics, Biostatistical Modeling, Molecular Biology, etc.

Work Experience

Monash University 2023/07 – 2024/04

Research Student Melbourne, VIC, Australia

- Conduct a research project on predicting TCR-antigen recognition using a deep-learning approach.

Shanghai Artificial Intelligence Laboratory 2021/06 – 2021/09

Research Intern Shanghai, China

- Develop a siamese graph attention network to predict variations in protein thermostability upon point mutations.
- Identify potential variants of *Streptomyces mobaraensis* transglutaminase (smTG) with improved thermostability.

Research Projects

Antigen-specific TCR Design 2024/11 – 2026/04

Biomedicine Discovery Institute, Monash University Melbourne, VIC, Australia

- Develop a cross-organism and pan-MHC TCR-pMHC binding model trained on multiple TCR binding specificity datasets.
- Design and optimize CDR3 sequences of TCRs given the cognate pMHC target using a conditional diffusion model.
- Validate the TCR design pipeline on the MAGE-A3 epitope via luciferase and activation assays.

TCR-antigen Recognition Prediction 2023/07 – 2024/09

Biomedicine Discovery Institute, Monash University Melbourne, VIC, Australia

- Predict binding specificity and interactions between T-cell receptors (TCRs) and epitopes presented by MHC-I molecules.
- Interpret spike-specific T-cell responses upon SARS-CoV-2 exposure using predicted TCR specificity and binding ranks.
- Elucidate T-cell cross-reactivity in autoimmune disease and cancer immunotherapy via residue-level contacts and simulating peptide library screening.

Bacterial Secreted Protein Prediction 2022/03 – 2023/10

State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University Shanghai, China

- Extend the prediction of bacterial secreted proteins to five major types in Gram-negative bacteria.
- Explore the taxonomic distribution of bacterial secretion systems and secreted proteins and provide an analytic platform.

Type IV Secreted Effector Prediction 2020/11 – 2022/01

State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University Shanghai, China

- Utilize a pre-trained model for the universal representation of protein sequences to predict type IV secreted effectors.
- Achieve a comparable performance with existing prediction tools and significantly enhance the computation efficiency.

Pan-genome Assembly 2019/12 – 2021/04

Department of Bioinformatics and Biostatistics, Shanghai Jiao Tong University Shanghai, China

- Develop a generic human pan-genome analysis procedure to assemble and analyze the pan-genomes of Tibetans.
- Construct Chinese Tibetan and Han Chinese pan-genomes and extract unique contigs and potential genes.

Cryo-EM Image Denoising 2020/02 – 2020/09

Department of Computer Science and Engineering, Shanghai Jiao Tong University Shanghai, China

- Utilize pre-trained ResNet and siamese neural network to extract features from cryo-EM images effectively.
- Assign pre-labels to the cryo-EM images by clustering and iteratively update model parameters to achieve denoising.

Skills

Programming Languages: Python, R, Bash

Tech Skills: Deep Learning, NGS Data Analysis, PyMOL, SQL, LATEX

Language: TOEFL iBT: 107 (Speaking: 23; Listening: 29; Writing: 25; Reading: 30.)

Awards

Excellent Graduate of Shanghai <i>Shanghai Municipal Education Commission</i>	2024
National Scholarship <i>Ministry of Education, China</i>	2022
The International Genetically Engineered Machine Competition Silver Medal <i>iGEM Foundation</i>	2020
China Undergraduate Mathematical Contest in Modeling (Shanghai) Second Prize <i>CSIAM</i>	2019
Mathematical Contest In Modeling Honorable Mention <i>COMAP</i>	2019

Publications

First author or Co-first author

1. R Han†, **Y Zhang**†, X Liu†, L Fu†, et al. (2026) Generalizable mutation-effect prediction across adaptive immune recognition via unified multimodal framework. *Nat Mach Intell*, 8, 913–929.
2. **Y Zhang**, J Song & C de la Fuente-Nunez. (2025) Deep learning speeds the search for new antibiotic scaffolds. *Nat Biotech*, 44, 1101–1103.
3. **Y Zhang** & J Song. (2025) Toward fair AI-driven medical text generation. *Nat Comput Sci*, 6, 361–362.
4. **Y Zhang**, Z Wang, Y Jiang, et al (2024) Epitope-anchored contrastive transfer learning for paired CD8+ T cell receptor-antigen recognition. *Nat Mach Intell*, 6, 1344–1358.
5. **Y Zhang**, J Guan, C Li, et al. (2023) DeepSecE: a deep learning-based framework for multi-class prediction of secreted proteins in Gram-negative bacteria. *Research*, 6, 0258.
6. H Gong†, **Y Zhang**†, C Dong, et al. (2023). Unbiased curriculum learning enhanced global-local graph neural network for protein thermodynamic stability prediction. *Bioinformatics*, 39(10), btad589.
7. **Y Zhang**†, Y Zhang†, Y Xiong, et al. (2022) T4SEfinder: a bioinformatics tool for genome-scale prediction of bacterial type IV secreted effectors using pre-trained protein language model. *Brief Bioinform*, 23(1), bbab420.

Others

1. X Li, H Gong, Y Wang, ..., **Y Zhang**, et al. (2026) De novo multi-mechanism antimicrobial peptide design via multimodal deep learning *Adv Sci* 13(28), e15835.
2. J Ma, **Y Zhang**, L Zhang, H Liu & J Song. (2026) HLABrew for human leukocyte antigen class I-presented epitope recognition and mimotope discovery. *J Chem Inf Model*, 66(9), 5544–5557.
3. Y Shen, H Lun, **Y Zhang**, et al. (2026) DposFinder: an interpretable transformer model for predicting phage-derived polysaccharide depolymerases and their host capsular serotypes. *Genome Med*, 18, 75.
4. T Pan, X Cui, HY Koh, ..., **Y Zhang**, et al. (2026) A geometric deep learning framework for genome-wide prediction of enzyme turnover number. *Genome Biol*, 27, 92.
5. Y Chen, **Y Zhang**, M Niu, et al. (2026) UniSyn: a multi-modal framework with knowledge transfer for anti-cancer drug synergy prediction. *Genome Biol*, 27, 74.
6. LC Shen, **Y Zhang**, Z Wang, et al. (2025) Self-iterative multiple instance learning enables the prediction of CD4+ T cell immunogenic epitopes. *Nat Mach Intell*, 7, 1250–1265.
7. Z Sun, J Xu, **Y Zhang**, et al. (2025) Multimodal geometric learning for antimicrobial peptide identification by leveraging alphafold2-predicted structures and surface features, *Brief Bioinform*, 26(3), bbaf261.
8. L Gao, **Y Zhang**, F Ge, et al. (2025) Structure-directed pan-specific T-cell receptor–peptide-major histocompatibility complex interaction prediction. *J Chem Inf Model*, 65(9), 4674–4686.
9. LC Shen, L Yan, Z Liu, **Y Zhang**, et al. (2025) Supervised contrastive learning enhances MHC-II peptide binding affinity prediction. *Expert Syst Appl*, 269, 126463.
10. Z Wang, **Y Zhang**, Y Xu, S Imoto, H Chen & J Song (2025) Histo-genomic knowledge distillation for cancer prognosis from histopathology whole slide images. *IEEE Trans Med Imaging*, 44(5), 2170–2181.
11. Y Tang, J Zhang, J Guan, ..., **Y Zhang**, et al. (2024) Transposition with Tn 3-family elements occurs through interaction with the host β -sliding clamp processivity factor. *Nucleic Acids Res*, 52(17), 10416–10430.

12. Z Yan, F Ge, Y Liu, **Y Zhang**, et al. (2024) TransEFVP: a two-stage approach for the prediction of human pathogenic variants based on protein sequence embedding fusion. *J Chem Inf Model*, 64(4), 1407–1418.
13. Y Li, B Wu, **Y Zhang**, L Liu, L Bai & T Shi. (2024) Enhanced thermostability of *Streptomyces mobaraensis* transglutaminase via computation-aided site-directed mutations and structural analysis. *New J Chem*, 48(2), 591-602.

Preprints

1. **Y Zhang**[†], W Liang[†], S Xu, et al. (2026) Generative design of antigen-specific T-cell receptor sequences with a conditional diffusion model. *bioRxiv*, 2026.06.10.730756.
2. S Hu, T Pan, J Wang, ..., **Y Zhang**, et al. (2026) ACENet: a graph neural network for predicting enzyme pH_{min} using structure and sequence insights. *bioRxiv*, 2026.03.11.711222.
3. Z Meng, Z Bai, K Yuan, J Song, **Y Zhang**, et al. (2026) ExplainBind: explainable physicochemical determinants of protein–ligand binding via non-covalent interactions. *bioRxiv*, 2026.03.03.707476.
4. M Xu, D Wang, Q Liu, ..., **Y Zhang**, et al. (2025) Multimodal AI decodes extreme environment functional dark matter beyond homology. *bioRxiv*, 2025.11.15.688615.