

# Yumeng Zhang

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

## Education

---

**Monash University** 2024/09 – 2028/03 (Expected)  
Ph.D. 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 – Present  
*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.

**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. **Y Zhang** & J Song. (2025) Toward fair AI-driven medical text generation. *Nature Computational Science*, 6, 361–362.
2. **Y Zhang**, Z Wang, Y Jiang, DR Littler, M Gerstein, AW Purcell, J Rossjohn, HY Ou & J Song. (2024) Epitope-anchored contrastive transfer learning for paired CD8+ T cell receptor-antigen recognition. *Nature Machine Intelligence*, 6, 1344–1358.
3. **Y Zhang**, J Guan, C Li, Z Wang, Z Deng, RB Gasser, J Song & HY Ou. (2023) DeepSecE: a deep learning-based framework for multi-class prediction of secreted proteins in Gram-negative bacteria. *Research*, 6, 0258.
4. H Gong†, **Y Zhang**†, C Dong, Y Wang, G Chen, B Liang, H Li, L Liu, J Xu & G Li. (2023). Unbiased curriculum learning enhanced global-local graph neural network for protein thermodynamic stability prediction. *Bioinformatics*, 39(10), btad589.
5. **Y Zhang**†, Y Zhang†, Y Xiong, H Wang, Z Deng, J Song & HY Ou. (2022) T4SEfinder: a bioinformatics tool for genome-scale prediction of bacterial type IV secreted effectors using pre-trained protein language model. *Briefings in Bioinformatics*, 23(1), bbab420.

*Others*

1. LC Shen, **Y Zhang**, Z Wang, DR Littler, L Yan, J Tang, J Rossjohn, DJ Yu & J Song. (2025) Self-iterative multiple instance learning enables the prediction of CD4+ T cell immunogenic epitopes. *Nature Machine Intelligence*, doi:10.1038/s42256-025-01073-z.
2. Z Sun, J Xu, **Y Zhang**, Y Zhang, Z Wang, X Wang, S Li, Y Guo, HH Shen & J Song. (2025) Multimodal geometric learning for antimicrobial peptide identification by leveraging alphafold2-predicted structures and surface features, *Briefings in Bioinformatics*, 26(3), bbaf261.
3. L Gao, **Y Zhang**, F Ge, S Li, Y Guo, J Song & DJ Yu. (2025) Structure-directed pan-specific T-cell receptor–peptide-major histocompatibility complex interaction prediction. *Journal of Chemical Information and Modeling*, 65(9), 4674-4686.
4. LC Shen, L Yan, Z Liu, **Y Zhang**, Z Wang, Y Guo, J Rossjohn, J Song & DJ Yu. (2025) Supervised contrastive learning enhances MHC-II peptide binding affinity prediction. *Expert Systems with Applications*, 269, 126463.
5. 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 Transactions on Medical Imaging*, 44(5), 2170-2181.
6. Y Tang, J Zhang, J Guan, W Liang, MT Petassi, **Y Zhang**, X Jiang, M Wang, W Wu, HY Ou & JE Peters. (2024) Transposition with Tn 3-family elements occurs through interaction with the host  $\beta$ -sliding clamp processivity factor. *Nucleic Acids Research*, 52(17), 10416-10430.
7. Z Yan, F Ge, Y Liu, **Y Zhang**, F Li, J Song & DJ Yu. (2024) TransEFVP: a two-stage approach for the prediction of human pathogenic variants based on protein sequence embedding fusion. *Journal of Chemical Information and Modeling*, 64(4), 1407–1418.
8. 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 Journal of Chemistry*, 48(2), 591-602.

*Preprints*

1. T Pan, X Cui, HY Koh, Y Bi, X Wang, **Y Zhang**, S Hu, GI Webb, L Kurgan, G Zhang & J Song. (2025) KcatNet: a geometric deep learning framework for genome-wide prediction of enzyme catalytic efficiency. *bioRxiv*, 2025.03.09.642294.