Ruihao Zhang

https://thuzrh.github.io thuzhangruihao@gmail.com (+86) 15737468186

EDUCATION

Tsinghua University (985, 211), Beijing, China September 2023 - June 2026
- Master of Engineering, Shenzhen International Graduate School
- Specialization: Materials and Chemical Engineering
- GPA: overall 3.91/4.0

Tianjin University (985, 211), Tianjin, China September 2019 - June 2023
- Bachelor of Engineering, Institute of Electrical Automation and Information Engineering
- Specialization: Electronic Information Engineering
- Award: Outstanding Graduate of Tianjin University

MIT MicroMasters Program 2025
- Data Analysis: Statistical Modeling and Computation in Applications (94/100)
- Fundamentals of Statistics (91/100)
- Machine Learning with Python-From Linear Models to Deep Learning (90/100)
- Probability - The Science of Uncertainty and Data (95/100)

RESEARCH EXPERIENCE

A Scalable Multi-Architecture Model for Epitope-Specific T Cell Receptor de novo Design 2024 - 2025
- Designed LSMTCR, a scalable framework that separates epitope specificity from immunogenetic constraints to generate de novo, epitope-conditioned full-length TCR αβ. Built a diffusion-enhanced BERT for time-conditioned epitope encoding, conditional GPT decoders (CDR3β→CDR3α transfer) with cross-modal conditioning and temperature control, and a gene-aware Transformer for V/J prediction and full-chain assembly.
- Outperformed baselines on GLIPH, TEP, MIRA, McPAS, and a curated dataset: higher predicted binding, better positional/length grammars, and tunable diversity. α-chain transfer improved binding and realism; full-length assembly preserved k-mer spectra, kept low edit distances, and paired α/β co-modeling improved pTM/ipTM.
- Delivered an end-to-end epitope-to-receptor pipeline combining diffusion-style encoding, conditional autoregression, shared epitope–CDR3 embeddings, and retrieval-augmented, soft-constrained decoding—yielding diverse, gene-contextualized TCR designs for high-throughput screening and practical constructability.

Classification of autoimmune diseases from TCR repertoires by multi-instance learning 2024 - 2025
- EAMIL pioneers a multimodal MIL framework that integrates sample-level and instance-level optimization, leveraging PrimeSeq for high frequency sequence selection and ESMonehot for feature extraction. This approach effectively addresses challenges in TCR repertoire scale, witness rates, and weak-labeling scenarios.
- Gating and spatial attention modules enable precise identification of disease-associated sequences and genes, facilitating holistic and local feature learning. These mechanisms enhance performance and interpretability, even with limited sample-level annotations.
- EAMIL achieves state-of-the-art performance in SLE and RA classification, SLEDAI score stratification, and marks the first effective solution for diagnosing organ-specific disease damage in SLE, demonstrating its transformative potential for clinical diagnostics.

Lightweight contrastive TCR-pMHC specificity learning with context-aware prompt 2023 - 2024
- Assisted in developing LightCTL, a lightweight contrastive learning framework for TCR-pMHC binding specificity prediction. By integrating lightweight encoding with contrastive learning and context-aware prompting modules, the framework significantly enhanced prediction accuracy and generalization performance.
- Spearheaded lightweight encoding module design and contributed to baseline model evaluation and large-scale protein binding dataset curation and preprocessing.

RESEARCH PROJECTS

Multi-Sensor Mobile Robot for Environmental Perception
- Provincial Project Leader May 2021 - May 2022
- Led development of a tracked robot platform integrating IMU, LiDAR, and stereo vision sensors for autonomous navigation. Implemented multi-sensor fusion algorithms enabling simultaneous localization and mapping in complex environments.
- Acquired proficiency in Linux, ROS framework, SLAM algorithms, and embedded systems programming.

Autonomous Mechanical Handling Robot with Vision-Based Recognition

- First Prize in New Engineering Competition August 2021 - December 2021
- Developed a mobile robotic system with autonomous grasping and placement capabilities using microcontroller programming and computer vision. Integrated camera-based global vision with mechanical arm flexibility to enable operation in diverse scenarios.
- Learned OpenMV and STM32 development, including color recognition, shape detection, and mechanical design.

PUBLICATIONS

- **Zhang, R.**, Chen, M., Ye, F., Meng, D., Huang, Y., & Liu, X. (2025). Classification of autoimmune diseases from peripheral blood TCR repertoires by multimodal multi-instance learning. arXiv.https://arxiv.org/abs/2507.04981.
- **Zhang, R.**, Liu, X. (2025). LSMTCR: A scalable multi-architecture model for epitope-specific T cell receptor de novo design. arXiv. https://arxiv.org/abs/2509.07627.
- Ye, F., Chen, M., Huang, Y., **Zhang, R.**, Li, X., Wang, X., ... & Liu, X. (2025). LightCTL: lightweight contrastive TCR-pMHC specificity learning with context-aware prompt. Briefings in Bioinformatics, 26(3).

AWARDS

Tsinghua University Comprehensive Excellence Scholarship (School-Level 2024)
Most Technically Innovative Award, Tsinghua University SDG Open Innovation Marathon Challenge 2023
Second Prize, CUHK Greater Bay Area Innovation Challenge 2023
Outstanding Graduate of Tianjin University 2023
Tianjin University Three-Good Student Award 2022
National Inspirational Scholarship 2022
First Prize, Tianjin New Engineering Competition 2021

OTHER PROJECTS

Brain Cell Classification Analysis - MIT project

- Accurately identify distinct brain cell types with 96%+ accuracy.
- Reveal cellular subtypes within major categories.
- Optimize analytical parameters for maximum biological insight.
- Validate biological theories through data-driven approaches.

WorkAssist - APP Development

- Localized Storage: Customizable data save paths
- Privacy Protection: Complete offline operation, data never leaves device
- Intelligent Analysis: Automatic generation of learning/work efficiency reports
- Lightweight Design: Install and use immediately, no complex configuration required

Alora AI - Startup

- Beta testing core platform functionalities and user workflows
- Providing feedback on avatar personality consistency and authenticity
- Testing cross-platform integration capabilities
- Evaluating monetization potential and creator tools effectiveness

PRESENTATIONS

Host Immunity in Multidrug-Resistant Bacterial Infections

- Authored complete MDR chapter covering bacterial resistance mechanisms and clinical implications.
- Analyzed four major antimicrobial resistance mechanisms and their interactions with host immune systems.
- Developed scientific communication skills through collaborative manuscript creation on antimicrobial resistance.

COMMUNITY INVOLVEMENT

- SIGS Dragon Boat Team Member September2024 - June2025
- SIGS Haishu Drama Club Member September2024 - June2025
- SIGS Student Safety Officer September2024 - June2025
- Tsinghua University Zijing Volunteer June2024 - June2025