Yuchen (Ethan) Shen

✉ ethanshen111@gmail.com

 github.com/EthanShenx

RESEARCH INTERESTS

Primarily interested in elucidating the genetic, epigenetic, and 3D chromatin regulatory mechanisms that underlie aging, cancer, and development, through the integration of computational approaches and functional genomics. My previous work lies at the interface of the integrative use of single-cell-level multi-omics strategies and advanced experimental systems such as organoid models with the ultimate aim of accelerating biomarker discovery and thus yielding clinical translational insights.

EDUCATION

- **Zhejiang University, School of Medicine** — China
  B.S. in Biomedical Informatics; GPA: 3.8/4.0; Rank: 3/58 — *Sep 2023 - Jun 2027*
  Courses (ZJU): Genomics and Proteomics, Algorithm Design, Statistics in R, Databases | **Online Courses (Coursera):** Machine Learning Specialization, Deep Learning Specialization

RESEARCH EXPERIENCE

- **High-Fat Diet Impact on Mammary Gland Development** — *Jul 2024 - Present*
  Chaochen Wang Lab, *Zhejiang University* — *Manuscript completed*
  Project Leader

  - **Objective**: To dissect the molecular mechanisms by which High-Fat Diet (HFD) impairs pubertal mammary gland epithelial development using snRNA-seq and wet-lab techniques.
  - **Mouse Model Construction**: Established Normal Diet (ND)- and HFD-fed mouse models covering the 3–10 week puberty window; performed whole-mount staining and snRNA-seq to capture developmental dynamics.
  - **Key Gene Module Profiling**: Applied Weighted Gene Co-expression Network Analysis (WGCNA) and Non-negative Matrix Factorization (NMF) to identify HFD-induced dysregulated gene co-expression modules.
  - ***Areg–Ghr* Regulatory Mechanism Discovery**: Deciphered *Areg*'s regulatory effect on mammary *Ghr* expression by blocking ADAM17-mediated cleavage of *Areg* via drug injection; quantified *Ghr* expression using qPCR and Western blot.
  - **Stromal Subset Characterization**: Integrated the results from NicheNet analysis, gene set scoring and enrichment analysis and identified $Map2^+$ $Adam12^+$ fibroblasts as HFD-activated collagen-producing stromal cells; localized *in vivo* via immunohistochemistry.
  - **Adipocyte–Microenvironment Crosstalk Analysis**: Performed pseudotime analysis revealing altered adipo-progenitor-to-adipocyte differentiation under HFD; identified *Nox4* as a potential intervention target and profiled intercellular signaling using CellChat.

- **Skin scRNA-seq Database Project** — *Nov 2024 - Present*
  Chaochen Wang & **Ying Xiao Lab**, *Zhejiang University* — *In progress*
  Atlas Constructor, UI Designer

  - **Objective**: To build a million-scale, state-of-the-art skin transcriptomic atlas and user-friendly web platform that pools resources for skin research and identifies novel disease-related biomarkers and establishes a new paradigm for understanding skin aging.
  - **GEO Datasets Scraping Tool**: Developed python web scraping programs to extract skin sc/snRNA-seq GEO datasets and their descriptions from NCBI. Manually curated each entry (>260) to update the information for each sample.
  - **Preprocessing Pipeline**: Created 8 versions of pre-processing pipelines, each tailored to a specific combination of sequencing platform, species, and sequencing method, from quality control to transfer learning-based cell type annotation.
  - **Downstream Analysis**: Conducted DEG analysis, functional enrichment analysis, cell-cell communication analysis, and gene regulatory network analysis based on Scanpy framework.
  - **Frontend UI and Web Function Design**: Optimized web data visualization through dynamic, responsive interfaces with HTML5, CSS, and JavaScript. Extended website functionality by enabling users to upload custom datasets for cell-cell communication analysis and data integration.

Enrichment Simplification (R Package: MAPA) *Mar 2025 - Aug 2025*

- ***Xiaotao Shen Lab**, Nanyang Technological University* *Preprint online*
  Software Contributor
  - **Objective**: To build an LLM-powered R package (MAPA) that resolves redundancy in pathway enrichment in transcriptomic and metabolomic data analysis.
  - **Metabolite GSEA Function**: Developed a metabolite-set GSEA algorithm implementing MAPA's pipeline for metabolites and cross-database (KEGG, HMDB) set result integration.
  - **Interactive Tutorials (Quarto/Shiny)**: Authored a 17-chapter interactive web tutorial with Quarto for both the R package and Shiny app/Web of MAPA, featuring hands-on examples, live demonstrations, and purpose-driven guides to streamline user onboarding.

PUBLICATIONS

†, (co–)first author; ∗, corresponding author.

- Ge, Y.[†], Zhang, F.[†], Liu, Y.[†], Jiang, C., Gao, P., Tan, N.S., Zhang, S., **Shen, Y.**, Zhou, Q., Zhou, X., Wang, C∗, Shen, X∗. Leveraging Large Language Models for Redundancy-Aware Pathway Analysis and Deep Biological Interpretation. bioRxiv preprint (2025). doi: 10.1101/2025.08.23.671949

QUALIFICATIONS

- **IELTS Academic Test** *Aug 2025*
  Overall 8.0 (Listening 9.0, Reading 9.0, Writing 7.0, Speaking 7.0)

HONORS

- **Zhejiang University Scholarship Third Prize** *Award rate: 9%* *2025*
- **ZJE Academic Scholarship Second Prize** *Award rate: 7%* *2025*
- **Zhejiang University Scholarship Third Prize** *Award rate: 9%* *2024*
- **ZJE Academic Scholarship Second Prize** *Award rate: 7%* *2024*

SKILLS

- **Programming Languages**: R (advanced), Python (advanced), Bash (advanced), SQL, Java, LaTeX, MATLAB
- **Software**: Prism, ImageJ, PyMOL, IGV, UCSF Chimera, Git, Anaconda, Adobe Illustrator, Adobe Photoshop, MS Office