Chanyue Charlotte Hu

UCLA PhD Candidate

PhD candidate in Bioinformatics with a Computer Science background, specializing in machine-learning-driven scientific modeling and large-scale biological data analysis. Experienced in developing reproducible Python-based ML pipelines, statistical models, and fine-tuning foundation models for high-dimensional biological datasets. Extensive experience analyzing single-cell and spatial transcriptomics data (scRNA-seq, CosMx, Xenium), integrating multi-omics datasets, and building computational frameworks to extract biological insights from complex cellular systems. Strong interest in applying AI/ML methods to accelerate discovery in biomedical research and drug development.

@ chanyuehero@gmail.com www.linkedin.com/in/chanyue-hu-believe https://github.com/DaisyCuttie

SKILLS

Programming Languages:
Python, R, SQL, Java

AI/Computational Tools:
Hugging Face, Pytorch, LLM, transformer, encoder-decoder, scikit-learn

Spatial Transcriptomics:
NanoString CosMx, 10x Xenium

Single-cell & Bioinformatics:
Seurat, Scanpy/AnnData, scRNA-seq integration, cell type annotation, differential expression, age prediction, visualization, batch correction, dimensionality reduction

EDUCATION

Ph.D. in Bioinformatics

UCLA Sep. 2021 – Jun. 2026 (expected) Los Angeles

Bachelor of Computer Science

UCI 2015 – 2019 Irvine

PUBLICATIONS

- **Hu, C.**, Pellegrini, M. *Determining the age of single cells using scBayesAge* **Submitted to npj Aging**, 2026. **bioRxiv**
- **Hu, C.**, ... *The heterogeneity between Th17 and Treg* **In Preparation**, 2026.
- **Hu, C.**, ... *Using Large Language Model for Single-cell Age Prediction* **In Preparation**, 2026.
- **Do, TH.**, ... *TREM2 macrophages induced by human lipids drive inflammation in acne lesions* **Sci Immunol**, 2022.
- **Farrell, C.**, ... *The Epigenetic Pacemaker is a more sensitive tool than penalized regression for identifying moderators of epigenetic aging* **Frontiers in Bioinformatics**, 2024.
- **Mohapatra, S.**, ... *Novel Taxol-Derivative, STO-1, Induces Selective Anti-Tumor Immunity and Sustained Remission of Glioblastoma Without Triggering Autoimmune Reactions* **Cells**, 2025.

PH.D RESEARCH PROJECTS

Single-cell Transcriptomics & Aging Study using a Large Foundational Model

UCLA Sep 2025 – Present Los Angeles **Github Repo Link** **Hugging Face Link**

Keywords: *Aging, Geroscience, LLM, Machine Learning*

- Fine-tuned a large language model (LLM) on single-cell RNA sequencing data to predict biological age at single-cell resolution in mouse datasets using LoRA
- Designed a tokenization and representation strategy for high-dimensional gene expression profiles, enabling effective adaptation of foundation models to scRNA-seq data
- Achieved preliminary prediction accuracy of ($Accuracy \approx 0.8$) across all organ tissue, demonstrating strong generalizations

Single-cell & Spatial Transcriptomics (CosMx & Xenium)

UCLA and Tunisia Jun 2020 – Present Los Angeles

Keywords: *Single-cell RNA, Spatial Transcriptomics, Disease Study, Immunology*

- Built end-to-end spatial transcriptomics pipelines for NanoString CosMx and 10x Xenium, including cell segmentation, QC, spatial visualization, and marker mapping using Seurat and custom Python workflows
- Integrated scRNA-seq references with spatial data for cell type annotation and spatial analysis in dermatological disease samples (FFPE $>$ 200,000 cells and $>$ 10 samples)
- Identified and spatially localized key immune and macrophage markers (including TREM2), linking spatial organization to inflammatory phenotypes in human tissue
- Performed heterogeneity and antimicrobial program analysis using single-cell spatial co-analysis approaches
- Mentored undergraduate researchers and led development of spatial analysis workflows for a Xenium-based project

Mapping the Aging Process with Single-cell RNA Sequencing

UCLA Sep 2023 – Oct 2025 Los Angeles

Keywords: *Single-cell RNA, Machine Learning, Aging, Geroscience, Python, Statistic Model*

- Implemented a pipeline of data processing, machine learning model training and visualization
- Built a novel machine learning framework for biological age prediction at single-cell resolution, achieving ($R^2 \approx 0.8$) with the best model
- Integrated transcriptomic age modeling with cell-type-specific analysis for cross-tissue aging comparisons
- Performed feature selection using gene–age correlations and statistical modeling for biological interpretation

Mediation of an Epigenetic Clock by Phenotypes

UCLA Jun 2020 – Sep 2023 Los Angeles

Keywords: *Epigenetics, Python, Aging, Geroscience, Statistic Model*

- Implemented a pipeline of data processing, machine learning model training and visualization
- Discovered the difference of epigenetic ages between men and women using the same CpG methylation sites
- Utilized various statistical tools, p-values, t-test, Generalized Additive Model (GAM) and violin plots
- Team-worked with a postdoc and undergraduate students

RELATED EXPERIENCE AND INTERNSHIPS

Staff Research Associate II

UCLA Apr 2020 – Aug 2021 UCLA

- Researched in analyzing and visualizing the relationships between aging and bisulfite sequencing
- Implemented Python built-in models and a novel statistical model measuring epigenetic states
- Integrated data analysis with experimental statistical methods
- Computed sequencing pipeline to generate bisulfite sequencing information

Backend Web Developer Intern

Huawei Technologies Co. Ltd. Jun 2019 – Sept 2019 Guangdong, China

- Programmed in Java with Spring framework to assemble backend data to connect with frontend interface
- Maintained daily functionality of file-processing modules
- Improved and optimized the battery warning system with Spring framework
- Reduced the latency of retrieving data of web assets by using cache

AI Image Intern

Beijing Elensdata Technology Co., LLC. Jul 2018 – Sep 2018 Beijing, China

- Developed a deep learning method to classify facial images of thousands of Chinese celebrities with a 94% accuracy rate
- Improved quality of image by filtering bad-quality images with a pre-trained CNN model

Teaching Experience

UCLA March.2025 – Present Los Angeles, CA

- Taught students about DNA sequence and manual correction of the sequence
- Give speeches about scRNA sequencing and spatial sequencing
- Walked through the DNA $\rightarrow$ protein structure/funtional analysis with students

AWARDS

Advancing Women in Technology *2016*

LANGUAGES

Chinese
Fluent

English
Fluent