Hongru Hu

Linkedin Profile | Google Scholar | hrhu.compbio@gmail.com | +1 (216)375-5885

PROFESSIONAL SUMMARY

Hongru is an applied machine learning researcher with interdisciplinary expertise in computational science, statistical learning, genomics, and neuroscience. His research is centered on developing machine learning frameworks tailored to multiscale and multimodal biomedical data. He focuses on designing generative frameworks for multimodal alignment, compositional data synthesis, large-scale graph co-representation, LLM adapters, and interpretable methodologies for AI-driven scientific research (AI4Science).

EXPERIENCES

Microsoft Research **June 2024 – Sept 2024**
Research Intern, LLM Innovation for Genomics *Redmond, WA*

- Utilized gene embeddings derived from bio-molecular and natural Large Language Models (LLMs) to prototype adapters, enabling context-aware feature representation learning which boosts the gene module discovery in single-cell studies.
- Constructed sample-feature co-embedding graph frameworks to improve biomarker discovery implicitly.

Sanofi **July 2023 – Dec 2023**
Machine Learning co-op Scientist, Generative AI for mRNA Design *Waltham, MA*

- Designed and optimized mRNA sequences with desired properties such as high stability and tailored protein expression to enable rapid vaccine prototyping by incorporating both predictive models and generative models into mRNA vaccine platforms.

Chan Zuckerberg CELLxGENE Discover **Jan 2023 – Present**
Core computational member, Human Liver Cell Atlas *International Collabration*

- Collaborated with cross-functional teams within the Human Cell Atlas (HCA) consortium as a core member of the computational group, focusing on multi-platform large-scale data integration using generative representation learning frameworks to develop the integrated cell atlas of the human liver in health and disease.

University of California, Davis **Sept 2020 – Present**
Research Assistant, Computational Biology *Davis, CA*

- Developed a multi-view representation learning framework for single-cell multimodal data integration, enabling cross-modality alignment and inference of regulatory programs in a cell state-specific manner.
- Designed a deconvolution framework utilizing deep generative models to estimate cell type proportions in bulk samples and project expression profiles into cell type resolution.
- Created a novel generative approach to eliminate batch effects, improving genotype prediction in high-content phenotypic imaging data.
- Conducted integrative computational analysis to explore the relationship between genetic variants, gene expression, and cellular behaviors, and quantified the heritability of psychiatric disorders in relation to electrophysiological responses.

EDUCATION

University of California **Davis, CA, USA**
Ph.D. Candidate in Computational Biology (Quon Lab); GPA 4.0 *Sept 2020 – Present*

University of Oxford **Oxford, UK**
Machine Learning Summer School (ML Fundamentals and ML X Health) *June - August 2022*

Florida State University **Tallahassee, FL, USA**
M.Sc. in Neuroscience (Arbeitman Lab) *July 2020*

Case Western Reserve University **Cleveland, OH, USA**
M.Sc. in Molecular Genetics (Mizutani and Sousa-Neves Lab) *May 2017*

Northwest A&F University **Shannxi, P.R.China**
B.Sc. in Biotechnology *July 2015*

HONORS and AWARDS

Human Genetics Pilot Grant (Genome Center, UC Davis) **Davis, CA**
Awarded $12,000 in direct funding for independent research in AI-drive drug repurposing *May 2023*

American Society of Human Genetics (ASHG) Meeting 2022 **Los Angeles, CA**
Reviewer's Choice Award *Oct 2022*

Genetic Society of America (GSA) Drosophila Meeting 2019 **Dallas, TX**
First Place of Graduate Trainee Presenters *Mar 2019*

PUBLICATIONS

Hu, H., Zhang, S., Choi, Y., Malladi, V.S., and Quon, G., 2025. sciLaMA: A Single-Cell Representation Learning Framework to Leverage Prior Knowledge from Large Language Models. **bioRxiv** [Paper]

Hu, H., and Quon, G., 2024. Boosting single cell multimodal analysis by leveraging single cell atlases and implicit feature selection. **Nature Communications** Code [Paper]

Napoliello, R.*, Hu, H.*, Ferns, L., and Quon, G., 2024. Multi-modal input integration using convolutional neural networks to predict CRISPRi efficacy. **NeurIPS 2024 WiML workshop**

Hu, H., Valdarrago, R.M., Dennis, M.Y., and Quon, G., 2024. Disentangling batch effects and morphological alterations in zebrafish through compositional generative models and registration. **(under review)**

Johansen, N.*, Hu, H.*, and Quon, G., 2023. Projecting RNA measurements onto single cell atlases to extract cell type-specific expression profiles using scProjection. **Nature Communications** Code [Paper]

Wang, Y., Zhang, W., Wu, Y., Qu, C., Hu, H., Lee, T., Lin, S., Zhang, J., Lam, K.S. and Wang, A., 2023. Protocol for vision transformer-based evaluation of drug potency using images processed by an optimized Sobel operator. **STAR protocols** [Paper]

Wang, Y., Zhang, W., Yip, H., Qu, C., Hu, H., Chen, X., Lee, T., Yang, X., Yang, B., Kumar, P. and Lee, S.Y., 2023. SIC50: Determining drug inhibitory concentrations using a vision transformer and an optimized Sobel operator. **Patterns** [Paper]

Choi, Y.*, Hu, H.*, and Quon, G., 2023. Promoters2Gene: Flexible transformer-based framework for decoding cell type-specific transcriptomic regulation. **(under review)**

Ren, Y., Wang, X., Zhang, S., Hu, H., Quicksall, Z., Lee, S., Morganti, J., Johnson, L., Asmann, Y., Zhao, N., 2023. Deconvolution reveals cell-type-specific transcriptomic changes in the aging mouse brain **Scientific Reports** [Paper]

Li, X.*, Hu, H.*, Johansen, N., and Quon, G., 2022. Cell adhesion molecules play subclass-specific roles in electrophysiological response and Schizophrenia risk. **bioRxiv** [Paper]

Brovero, S.G., Fortier, J.C., Hu, H.*, Lovejoy, P.C., Newell, N.R., Palmateer, C.M., Tzeng, R.Y., Lee, P.T., Zinn, K. and Arbeitman, M.N., 2021. Investigation of Drosophila fruitless neurons that express Dpr/DIP cell adhesion molecules. **Elife** [Paper]

Zhao, P., Wang, K., Zhang, W., Liu, H.Y., Du, L.P., Hu, H.R. and Ye, X.G., 2017. Comprehensive analysis of differently expressed genes and proteins in albino and green plantlets from a wheat anther culture. **Biologia plantarum** [Paper]

* indicates co-first-authorship