Yang Xu (ORCID: https://orcid.org/0000-0003-4173-4337)

Contact

Phone: +1 (865) 299-8647   Email: yxu71@vols.utk.edu

Education

2017-2022	**University of Tennessee, Knoxville**, Knoxville, TN
	Ph.D. in Life Sciences (concentration in Computational Biology)
2011-2015	**Huazhong Agricultural University**, Wuhan, China
	B.S. in Biotechnology

Academic experience

2022-present   Machine learning scientist I,   **Broad institute of MIT and Harvard**, Cambridge, MA

1. Integration of multimodal single-cell data
2. Inferring transcriptional dynamics of single-cell data via deep learning

2017-2022   Graduate teaching/research assistant,   **University of Tennessee, Knoxville**, Knoxville, TN

1. Development of CoSTA, unsupervised convolutional neural network clustering for spatial transcriptomics analysis.
2. Development of SMILE, contrastive learning for multimodal integration of single-cell omics data.
3. Development of sciCAN, adversarial approach to integrate single-cell gene expression and chromatin accessibility data.

2020-2020   Visiting scientist intern,   **Broad institute of MIT and Harvard**, Cambridge, MA

Development of MACA: a marker-based automatic cell-type annotation tool for single cell expression data

2016-2017   Research assistant,   **Zhejiang University**, Hangzhou, China

Development of Microwell-seq platform for high throughput single-cell RNA-seq.

2014-2015   Undergraduate research assistant,   **Huazhong Agricultural University**, Wuhan, China

Molecular examination of MPK12 in guard cell CO_2 signaling.

Industrial experience

2021-2021   Intern, Statistics, and algorithm,   **Adaptive Biotechnologies**, Seattle, WA

1. Reframed internal error correction algorithm for T-cell receptor sequencing
2. New algorithm corrected PCR and sequencing errors with substantial improvement

2020-2020   Bioinformatics summer intern,   **Bayer Pharmaceuticals**, Cambridge, MA

3. Development of MACA: a marker-based automatic cell-type annotation tool for single-cell expression data
4. Feat. a flash talk at Human Cell Atlas Asia Meeting 2020, and a poster talk at RECOMB 2021

2019-2019   Bioinformatics summer intern,   **Engine Biosciences**, San Carlos, CA

5. Topic modeling to identify genetic pathways and associated genes from biomedical publications.
6. Multiple classification models with topic as feature to target biomedical publications.
7. Comprehension tasks by Google BERT: bio-term tagging and gene interaction extraction.

Peer review

2021   Reviewer for ***Nature communications***

First-author publications

1. **Xu, Y.** & McCord, R. P. (2022). Diagonal integration of multimodal single-cell data: potential pitfalls and paths forward. ***Nature communications***.
2. **Xu, Y.**, Baumgart, S. J., Stegmann, C., & Hayat, S. (2021). MACA: Marker-based automatic cell-type annotation for single cell expression data. ***Bioinformatics***. (co-corresponding author)
3. **Xu, Y.**, Das, P., & McCord, R. P. (2021). SMILE: Mutual Information Learning for Integration of Single Cell Omics Data. ***Bioinformatics***.
4. **Xu, Y.**, & McCord, R. P. (2021). CoSTA: Unsupervised Convolutional Neural Network Learning for Spatial Transcriptomics Analysis. ***BMC bioinformatics***.
5. Lai, S., Huang, W., **Xu, Y.**, Jiang, M., Chen, H., Cheng, C., ... & Han, X. (2018). Comparative transcriptomic analysis of hematopoietic system between human and mouse by Microwell-seq. ***Cell discovery***. (co-first author)

Co-author publications

1. McCord, R.P., **Xu, Y.**, Li, H., Das, P., & San Martin, R. (2022). Snapshot: Chromosome organization. ***Molecular cell***.
2. Sanders, J.T., Golloshi, R., Das, P., **Xu, Y.**, Terry, P.H., Nash, D.G., Dekker, J. & McCord, R.P. (2022). Loops, TADs, Compartments, and Territories are Elastic and Robust to Dramatic Nuclear Volume Swelling. ***Scientific reports***.
3. San Martin, R., Das, P., Marques, R. D. R., **Xu, Y.**, & McCord, R. P. (2021). Alterations in chromosome spatial compartmentalization classify prostate cancer progression. ***Journal of cell biology***.
4. Sanders, J. T., Freeman, T. F., **Xu, Y.**, Golloshi, R., Stallard, M. A., Hill, A. M., ... & McCord, R. P. (2020). Radiation-induced DNA damage and repair effects on 3D genome organization. ***Nature communications***.
5. Jiang, M., Chen, H., Lai, S., Wang, R., Qiu, Y., Ye, F., Fei, L., Sun, H., **Xu, Y.**, ... & Guo, G. (2018). Maintenance of human haematopoietic stem and progenitor cells in vitro using a chemical cocktail. ***Cell discovery***.
6. Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., Huang, D., **Xu, Y.**, ... & Guo, G. (2018). Mapping the mouse cell atlas by microwell-seq. ***Cell***.
7. Han, X., Yu, H., Huang, D., **Xu, Y.**, Saadatpour, A., Li, X., ... & Guo, G. (2017). A molecular roadmap for induced multi-lineage trans-differentiation of fibroblasts by chemical combinations. ***Cell research***.
8. Jakobson, L., Vaahtera, L., Tõldsepp, K., Nuhkat, M., Wang, C., Wang, Y. S., Hõrak, H., Valk, E., Pechter, P., Sindarovska, Y., Tang, J., Xiao, C., **Xu, Y.**, ... & Brosche, M. (2016). Natural variation in Arabidopsis Cvi-0 accession reveals an important role of MPK12 in guard cell CO_2 signaling. ***PLoS biology***.

Non-research publications

1. **Xu, Y.** (2022). Quitting grad school felt like failing—but it gave me an opportunity to grow. ***Science***.

Preprints

1. **Xu, Y.**, Kramann, R., McCord, R. P., & Hayat, S. (2022). Fast model-free standardization and integration for single-cell transcriptomics data. ***bioRxiv***.
2. **Xu, Y.**, Begoli, E., & McCord, R. P. (2021). sciCAN: Single-cell chromatin accessibility and gene expression data integration via Cycle-consistent Adversarial Network. ***bioRxiv***.
3. **Xu, Y.**, & Strick, A. J. (2021). Integration of Unpaired Single-cell Chromatin Accessibility and Gene Expression Data via Adversarial Learning. ***arXiv***.
4. **Xu, Y.**, Shen, T., & McCord, R. P. (2019). 3D Genome Structure Variation Across Cell Types Captured by Integrating Multi-omics. ***bioRxiv***.
5. Lai, S., **Xu, Y.**, Huang, W., Jiang, M., Chen, H., Ye, F., ... & Guo, G. (2017). Mapping human hematopoietic hierarchy at single cell resolution by Microwell-seq. ***bioRxiv***.

Talks and posters

1. Marker-assisted standardization and integration for single-cell transcriptomics data. scientific talk at **Novo Nordisk**

2. sciCAN: Single-cell chromatin accessibility and gene expression data integration via Cycle-consistent Adversarial Network. talk and poster at **Rocky Mountain bioinformatic conference 2021**, Aspen, CO
3. MACA: Integration and standardization of cell-type annotation across multiple single-cell expression data. poster at **RECOMB 2021** and **Human Cell Atlas Asia Meeting 2021** (Virtual)
4. SMILE: Mutual Information Learning for Integration of Single Cell Omics Data. SciTalk at **Keystone Symposia: Single Cell Biology 2021** (Virtual)
5. MACA: Marker-Based Automatic Cell-Type Annotation for Single Cell Expression Data. flash talk at **Human Cell Atlas Asia Meeting 2020** (Virtual)
6. Capturing Dynamic Genome Organization. poster at **4DN Annual Meeting 2019**, Washington, D.C.