

Yuan Xue (Soso)

Single-cell biology, genomics, data science, 

Bioinformatics Scientist @ Bluestar Genomics

 yuanxue@stanford.edu

 [github/xuesoso](https://github.com/xuesoso)

 [linkedin/yuan-xue](https://www.linkedin.com/in/yuan-xue)

 xuesoso.github.io

Summary

- Enthusiastic and self-motivated scientist working on non-invasive cfDNA diagnostic test for early-stage cancer detection.
- Developed several efficient cloud computing pipelines on Unix system for high volume NGS data processing, single-nucleotide mutations analysis, transcriptome assembly from long-read sequencing, functional annotation, *de novo* microbial identification (Nextflow, Snakemake, FastQC, Samtools, BWA, Bowtie2, Minimap2, STAR, HTSeq-count, Salmon, tBLASTx, Kraken).
- Proficient in statistical and machine learning approaches to derive biological insights from high-dimensional data. Experienced in both unsupervised learning of data structure (PCA, CCA, t-SNE, UMAP), supervised regression techniques, and classification tasks (Naive Bayes, SVM, Random Forest, XGBoost, LightGBM).
- Expertise in single-cell transcriptomics (dimensionality reduction, clustering, trajectory analysis, cross-experiment alignment), *de novo* assembly tasks (long-read Nanopore and Pacbio sequencing), and functional genomics analysis (GSEA, GSOA, KEGG, GO). Developed and implemented cutting edge analytical methods.
- Hands-on experience with molecular biology and NGS techniques (10X scRNA-seq, SMART-seq2, Illumina library index).

Experience

Bluestar Genomics

Bioinformatics Scientist

2022 – now

San Mateo, CA, USA

Education

Stanford University

Ph.D. Bioengineering. Thesis advisers: Stephen Quake, John Boothroyd

2015 – 2021

Stanford, CA, USA

Stanford University

M.S. Bioengineering

2015 – 2017

Stanford, CA, USA

Reed College

B.A. Biology

2010 – 2014

Portland, OR, USA

Skills

Languages English, Cantonese, Mandarin, Japanese, Python, R, Bash, \LaTeX

Visualization matplotlib, plotly, seaborn, bokeh, networkX, graphviz, graph-tool

Machine Learning numpy, pandas, sklearn, scikit-learn, tensorflow, keras, statsmodels

Workflows nextflow, snakemake, cloud computing (e.g. AWS, slurm), docker container

Web development Flask-REST backend, Apache2

Bioinformatics STAR, minimap2, htseq-count, salmon, velocity, samtools, scanpy, Kraken

Maintained Packages

singleCell_snake A snakemake pipeline for local/Slurm submission of single cell data alignment and transcript counting.

nheatmap A python package to generate multi-level heatmap with extensive configuration options.

DensityPlot A python package to generate density scatter plot.

bag_of_velocity A bash / python script for parallel submission of RNA velocity alignment on Slurm.

Selected publications

1. Trung Pham*, **Yuan Xue***, Susan Brewer, Liliana Massis, Stephen R. Quake, Denise Monack. Single-cell profiling reveals functional diversity of granuloma macrophages during persistent Salmonella infection. **Under review (2022)**.
2. **Yuan Xue**, Ido Braslavsky, Stephen R. Quake. Temperature effect on DNA polymerase fidelity. **Journal of Biological Chemistry (2021)**. [bioRxiv preprint](#)
3. Pengyang Li, Dania Nanes Sarfati*, **Yuan Xue***, Xi Yu, Alexander Tarashansky, Stephen R. Quake, Bo Wang. Single-cell analysis of Schistosoma mansoni reveals a conserved genetic program controlling germline stem cell fate. **Nature Communications (2020)**. [bioRxiv preprint](#)
4. Suchita Rastogi, **Yuan Xue**, Stephen R. Quake*, John Boothroyd*. Differential Impacts on Host Transcription by ROP and GRA Effectors from the Intracellular Parasite Toxoplasma gondii. **mBio (2020)**. [bioRxiv preprint](#)
5. **Yuan Xue**, Terence Theisen, Suchi Rastogi, Abel Ferrel, Stephen R. Quake*, John Boothroyd*. A single-parasite transcriptional atlas of Toxoplasma gondii reveals novel control of antigen expression. **eLife (2020)**. [bioRxiv preprint](#)
6. Alexander Tarashansky, **Yuan Xue**, Pengyang Li, Stephen R. Quake, Bo Wang. Self-assembling Manifolds in Single-cell RNA Sequencing Data. **Elife (2019)**. [bioRxiv preprint](#)
7. **The Tabula Muris Consortium**, Stephen R. Quake, Tony Wyss-Coray, Spyros Darmanis. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. **Nature (2018)**. [bioRxiv preprint](#)
8. **Yuan Xue**, Jossef Osborn, Anand Panchal, Jay L. Mellies. The RpoE stress response pathway mediates reduction of enteropathogenic Escherichia coli virulence by zinc. **Applied and Environmental Microbiology (2015)**.
9. Jing Zhou, Shi-Hao Tan, Valerie Nicolas, Chantal Bauvy, Nai-Di Yang, Jianbin Zhang, **Yuan Xue**, Patrice Codogno, Han-Ming Shen. Activation of lysosomal function in the course of autophagy via mTORC1 suppression and autophagosome-lysosome fusion. **Cell Research (2013)**.

*equal contributions

Awards & Honors

Stanford Bio-X Travel Award	2019
Stanford Bio-X SIGF Fellow	2018
> One of 11 students awarded with a three-year fellowship to conduct interdisciplinary research on the topics of parasitology and single-cell bioinformatics.	
Reed College Larry Ruben Postbac. Research Fellow	2014
Reed College Summer Experience Awardee	2013
Reed College Independent Research Awardee	2012
iGEM Competition Team Gold Medalist	2009