

Kyung Hoi (Joseph) Min

COMPUTATIONAL BIOLOGIST · SOFTWARE ENGINEER

✉ kmin@mit.edu | 📄 Lioscro | 🌐 lioscro | 🐦 @lioscro

Education

MIT

M.S. IN EECS

- Research Assistant in the Weissman Lab at the Whitehead Institute

Cambridge, MA

Aug 2020 - Jun 2022

Caltech

B.S. IN BIOLOGY AND COMPUTER SCIENCE

- Biology Courses: Regulation of Gene Expression, Cell Biology, Genetics, Bioinformatics
- CS Courses: Decidability and Tractability, Algorithms, Learning Systems

Pasadena, CA

Sept 2016 - Jun 2020

Experience

Software Engineer

GINKGO BIOWORKS

- Worked as an individual contributor on the Solution Engineering team identifying gaps in core software, proposing and designing technical solutions, driving cross-team initiatives in an Agile environment, and unblocking high-impact projects.

Boston, MA

Aug 2020 - Present

Research Assistant

WEISSMAN LAB

- Studied cancer evolution via synthetic lineage reconstruction using a CRISPR-enabled lineage tracer in a lung cancer mouse model.
- Developed a tool to preprocess single cell metabolic labeling data to quantify RNA splicing kinetics.
- Developed new approaches of integrating multi-modal single-cell measurements (gene expression, lineage, spatial)

Cambridge, MA

Dec 2020 - Jun 2022

Software Solution Engineering Intern

GINKGO BIOWORKS

- Worked on migrating Ginkgo's mass spectrometry data pipeline to AWS and Airflow.
- Was responsible for the full development cycle of the project—from conceptualization, making architectural and design decisions to implementation, validation, and release.

Boston, MA

Jun 2020 - Aug 2020

Undergraduate Researcher

PACHTER LAB

- Worked with **Sina Boeshaghi** and **Prof Lior Pachter** on single-cell RNA-seq projects.
- Developed a friendly wrapper around a popular scRNA-seq pre-processing pipeline (kb-python).
- Developed a user interface for the open-source fraction collector (Colosseum) using PyQt5 and later as a web application with React.js.

Pasadena, CA

Sept 2019 - Jun 2020

Software Engineering Intern

GINKGO BIOWORKS

- Improved Ginkgo's in-house data analysis pipeline and various backend services.
- Worked with other software engineers in an Agile environment to deliver stable, scalable and tested products.

Boston, MA

Jun. 2019 - Sept 2019

Undergraduate Researcher

STERNBERG LAB

- Worked with **David Angeles PhD.** and **Prof Paul W. Sternberg** on projects involving the model organism *C. elegans*.
- Developed an automatic, complete RNA-seq pipeline for nematodes (Alaska).

Pasadena, CA

Jan 2017 - Jun 2019

Projects

Spateo

WEISSMAN LAB

- Python package for analysis of high-resolution spatial transcriptomics.
- Developed novel RNA-based cell segmentation approaches.

 [aristoteleo/spateo-release](https://github.com/aristoteleo/spateo-release)

Dec 2021 - Jun 2022

Tumor Spatial Transcriptomics and Lineage Tracing

WEISSMAN LAB

May 2021 - Jun 2022

- Multi-modal analysis of tumor spatial transcriptomics and single-cell lineage tracing.

ngs-tools

WEISSMAN LAB

 [Lioscro/ngs-tools](#)

Apr 2021 - Jun 2022

- A reusable Python library for working with next-generation sequencing (NGS) data

Cassiopeia

WEISSMAN LAB

 [YosefLab/Cassiopeia](#)

Feb 2021 - Jun 2022

- A package for processing and analysis of Cas9-enabled single-cell lineage tracing data.
- Implemented support for spatial lineage tracing data.
- Major improvements to preprocessing pipeline, including optimization and parallelization.

Dynast

WEISSMAN LAB

 [aristoteleo/dynast-release](#)

Oct 2020 - Jun 2022

- Pipeline for complete splicing and labeling quantification from metabolic labeling scRNA-seq experiments.
- Ability to perform statistical correction of labeled/unlabeled RNA using a Bayesian inference model.

Copy Number Variation (CNV) Inference and Analysis

WEISSMAN LAB

Nov 2020 - May 2022

- Performed CNV inference from tumor single-cell gene expression using inferCNV.
- Validated the inferred tumor phylogenetic trees and subclonal expansions using the CNV patterns of each cell.

Mass Spectrometry Pipeline

GINKGO BIOWORKS

Jun 2020 - Aug 2020

- Migrated mass spectrometry pipeline, which depended on an on-prem Windows machine, to Linux on AWS and Airflow.
- Implemented the ability to call Windows DLLs within Linux with Wine and IronPython.
- Retired an internal fork of ProteoWizard by contributing back to the project.

Colosseum

PACHTER LAB

 [pachterlab/colosseum](#)

Sept 2019 - Jun 2020

- A low-cost, modular, open-source automated fluid sampling device for scalable microfluidic applications.
- Can be built for less than \$100 using off-the-shelf components in less than an hour.
- Developed its user interface using PyQt5 and installable from Pypi; later migrated to a fully-featured web application written with React.js.

kb-python

PACHTER LAB

 [pachterlab/kb_python](#)

Oct 2019 - Jun 2022

- A Python wrapper around the kallisto | bustools pipeline for scRNA-seq pre-processing.
- Pre-processes scRNA-seq data into gene count matrices and RNA velocity matrices.
- Provides pre-built reference indices for pseudoalignment.

NGS Pipeline

GINKGO BIOWORKS

Jun 2019 - Sept 2019

- Migrated in-house NGS pipeline to the cloud (AWS S3, Batch, DynamoDB).
- Implemented pipeline metrics collection with Elasticsearch and Grafana.
- Implemented contamination screening via taxonomy analysis with Centrifuge.

Alaska

STERNBERG LAB

 [lioscro/alaska-parse](#)

Jan 2017 - Aug 2019

- Developed a framework for automated, complete RNA-seq analysis.
- Parse Server backend. Frontend portal in Javascript, HTML/CSS. Containerization with Docker.
- Presented a poster at the 21st International *C. elegans* Conference.

Publications

2024	Quantifying orthogonal barcodes for sequence census assays , author	Bioinformatics Adv.
2023	kallisto, bustools, and kb-python for quantifying bulk, single-cell, and single-nucleus RNA-seq , author	bioRxiv
2023	Metadata retrieval from genomics database with ffq , author	Bioinformatics
2022	Spateo: multidimensional spatiotemporal modeling of single-cell spatial transcriptomics , author	bioRxiv
2022	Dynast: Inclusive and efficient quantification of metabolically labeled transcripts in single cells , thesis	MIT
2022	Inferring gene regulation from stochastic transcriptional variation across single cells at steady state , author	PNAS
2022	Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution , author	Cell
2022	Mapping Transcriptomic Vector Fields of Single Cells , author	Cell
2021	Low-cost, scalable, and automated fluid sampling for fluidics applications , author	HardwareX
2021	Modular and efficient pre-processing of single-cell RNA-seq , author	Nature Biotech.

Skills

Programming	Python, Javascript, Ruby on Rails, C# (.NET Framework), C/C++, HTML/CSS
Software	Docker, Django, Apache Airflow, AWS (S3, Batch, RDS, Lambda), React.js, GraphQL, Kubernetes, Kafka, Snowflake
Data Analysis	Bulk RNA-seq, Single-cell RNA-seq, Spatial Transcriptomics
Molecular Biology	PCR, Cloning, Purification