

# Kyung Hoi (Joseph) Min

COMPUTATIONAL BIOLOGIST · SOFTWARE ENGINEER

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

## Education

### MIT

PHD. STUDENT IN EECS

- Working as a research assistant in the Weissman Lab at the Whitehead Institute

Cambridge, MA

Aug 2020 - Present

### 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

### Research Assistant

WEISSMAN LAB

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

Cambridge, MA

Dec 2020 - Present

### 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

### Tumor Spatial Transcriptomics and Lineage Tracing

WEISSMAN LAB

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

May 2021 - Present

### Cassiopeia

WEISSMAN LAB

- 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.

 [YosefLab/Cassiopeia](#)

Feb 2021 - Present

## Copy Number Variation (CNV) Inference and Analysis

WEISSMAN LAB

Nov 2020 - Present

- 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.

## dynast

 [aristoteleo/dynast-release](#)

WEISSMAN LAB

Oct 2020 - Present

- 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.

## 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

 [pachterlab/colosseum](#)

PACHTER LAB

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

 [pachterlab/kb\\_python](#)

PACHTER LAB

Oct 2019 - Jun 2020

- 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

 [lioscro/alaska-parse](#)

STERNBERG LAB

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

- |      |   |                                 |
|------|---|---------------------------------|
| 2021 | <b>Molecular Recording Reveals the Phylodynamics, Plasticity, and Paths of Tumor Evolution</b> , author | <a href="#">Under review</a>    |
| 2021 | <b>Mapping Transcriptomic Vector Fields of Single Cells</b> , author                                    | <a href="#">Under review</a>    |
| 2021 | <b>Low-cost, scalable, and automated fluid sampling for fluidics applications</b> , author              | <a href="#">HardwareX</a>       |
| 2021 | <b>Modular and efficient pre-processing of single-cell RNA-seq</b> , author                             | <a href="#">Nature Biotech.</a> |

## Presentations

### CSHL Single Cell Analyses Conference

[Cold Spring Harbor, NY](#)

POSTER PRESENTER FOR <LINEAGE RECORDING REVEALS TUMOR SUBCLONAL DYNAMICS, GENE PROGRAMS, AND METASTATIC ORIGINS>

Nov 2021

### SURF Seminar Day

[Pasadena, CA](#)

PRESENTER FOR <A FRAMEWORK FOR AUTOMATED, COMPLETE RNA-SEQ ANALYSIS>

Oct 2018

### 21st International *C. elegans* Conference

[Los Angeles, CA](#)

POSTER PRESENTER FOR <A FRAMEWORK FOR AUTOMATED, COMPLETE RNA-SEQ ANALYSIS>

Jun 2017

## Awards & Honors

---

- 2018 **Semi-finalist**, Peripall Speaking Competition  
2018 **Fellow**, Samuel P. and Frances Krown SURF Fellowship  
2017 **Fellow**, SURF Fellowship

*Pasadena, CA*  
*Pasadena, CA*  
*Pasadena, CA*

## Skills

---

<b>Programming</b>	Python, Javascript, C/C++, HTML/CSS
<b>Software</b>	Docker, Django, Apache Airflow, AWS S3/Batch/DynamoDB, Elasticsearch, React.js, CI/CD, Numpy, Pandas
<b>Data Analysis</b>	Bulk RNA-seq, Single-cell RNA-seq
<b>Molecular Biology</b>	PCR, Cloning, Purification

## Extracurriculars

---

### **The Big T Yearbook**

*Pasadena, CA*

BUSINESS MANAGER, DESIGNER, PHOTOGRAPHER; EDITOR-IN-CHIEF SEPT 2019 - JUN 2020

*Sept 2016 - Jun 2020*

- (Editor-in-Chief) Responsible for setting the theme, layout, content, and publishing of the 2019-2020 yearbook.
- Established a new advertisement system, shared cloud storage system, job organization and assignment system, and payment reimbursement system.
- Designed pages and took/edited photos with Adobe Creative Cloud software.

### **Caltech & American Red Cross**

*Pasadena, CA*

HEALTH ADVOCATE & EMERGENCY MEDICAL RESPONDER

*Sept 2017 - Jun 2020*

- American Red Cross certified emergency medical responder and basic life support provider.
- Worked as the dormitory health advocate and provided first aid to students.

### **Student Government, Lloyd House**

*Pasadena, CA*

HISTORIAN

*Jun 2017 - May 2019*

- Responsible for recording Lloyd House (dormitory) events by taking photos.
- Established a new shared cloud storage system to distribute these photos.