

# Kyung Hoi (Joseph) Min

COMPUTATIONAL BIOLOGIST · SOFTWARE ENGINEER

☎ (626) 240-9851 | ✉ kmin@mit.edu | 🌐 Lioscro | 🔗 lioscro | 🐦 @lioscro

## Education

### Caltech

B.S. IN BIOLOGY AND COMPUTER SCIENCE

*Pasadena, CA*

*Sept 2016 - Jun 2020*

- Cum. GPA: 4.0 / 4.3
- Biology Courses: Regulation of Gene Expression, Cell Biology, Genetics, Bioinformatics
- CS Courses: Decidability and Tractability, Algorithms, Learning Systems

## Experience

### Ginkgo Bioworks

SOFTWARE SOLUTION ENGINEERING INTERN

*Boston, MA*

*Jun 2020 - Aug 2020*

- 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.

### Pachter Lab

UNDERGRADUATE RESEARCHER

*Pasadena, CA*

*Sept 2019 - Jun 2020*

- 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).

### Ginkgo Bioworks

SOFTWARE ENGINEERING INTERN

*Boston, MA*

*Jun. 2019 - Sept 2019*

- Worked on improving 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.

### Sternberg Lab

UNDERGRADUATE RESEARCHER

*Pasadena, CA*

*Jan 2017 - Jun 2019*

- 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).

## Projects

### lfq

HOBBY PROJECT

 [Lioscro/lfq](#)

*Aug 2020 - Present*

- New binary encoding for FASTQ sequence data.
- Developed a cross-platform command-line tool to encode and decode FASTQ files to and from this new data format.

### 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.

### kb-python

PACHTER LAB

 [pachterlab/kb\\_python](#)

*Oct 2019 - Present*

- 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

2020 **Feature demultiplexing for single-cell genomics assays**, co-first author

*In preparation*

2020 **Modular and efficient pre-processing of single-cell RNA-seq**, author

*Under review*

## Presentations

### SURF Seminar Day

*Pasadena, CA*

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

*Oct 2018*

- Presented work done on Alaska during summer fellowship.

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

*Los Angeles, CA*

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

*Jun 2017*

- Introduced Alaska, an RNA-seq analysis pipeline for nematodes.

## Awards & Honors

2018 **Semi-finalist**, Perpall Speaking Competition

*Pasadena, CA*

2018 **Fellow**, Samuel P. and Frances Krown SURF Fellowship

*Pasadena, CA*

2017 **Fellow**, SURF Fellowship

*Pasadena, CA*

## Skills

**Programming** Python, Javascript, C/C++, HTML/CSS

**Software** Docker, Django, Apache Airflow, AWS S3/Batch/DynamoDB, Elasticsearch, React.js, CI/CD, Numpy, Pandas

**Data Analysis** Bulk RNA-seq, Single-cell RNA-seq

**Molecular Biology** 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.