

# Kyung Hoi (Joseph) Min

COMPUTATIONAL BIOLOGIST · SOFTWARE ENGINEER

☎ (626) 240-9851 | ✉ kmin@caltech.edu | 🏠 lioscro.github.io/homepage | 📱 Lioscro | 🌐 lioscro | 🐦 @lioscro

## Education

### Caltech

B.S. IN BIOLOGY AND COMPUTER SCIENCE

Pasadena, CA

Sept 2016 - Present

- Expected Graduation: 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

### Pachter Lab

UNDERGRADUATE RESEARCHER

Pasadena, CA

Sept 2019 - Present

- Working with **Sina Boosaghghi** and **Prof Lior Pachter** on single-cell RNA-seq projects.
- Developing a friendly wrapper around a popular scRNA-seq pre-processing pipeline (kb-python).
- Developing a repository for scRNA-seq data that will allow aggregation across multiple datasets (Commons Cell Atlas).

### 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.
- Developed software directly affecting production workflows.

### 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

### Commons Cell Atlas

PACHTER LAB

🔒 (private)

Sept 2019 - Present

- Repository for all the world's single-cell RNA-seq data.
- Arbitrary aggregation and comparison of data across multiple datasets by using a variational autoencoder to simplify multi-dimensional data.
- Web application written in React.js and Django.

### 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

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

---

- 2019 **Feature demultiplexing for single-cell genomics assays**, co-first author  
2019 **Modular and efficient pre-processing of single-cell RNA-seq**, author

*In preparation*  
*Under review*

## Presentations

---

### Ginkgo Bioworks

*Boston, MA*

#### INTERNSHIP PRESENTATION

*Sept 2019*

- Technical and non-technical presentations to software team and stakeholders.
- Presented NGS pipeline improvements and migration to the cloud.

### 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  
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, HTML/CSS, C, C++
<b>Software</b>	Docker, Django, Apache Airflow, AWS S3/Batch/DynamoDB, Elasticsearch, React.js, CI/CD, Celery, RabbitMQ, 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 - PRESENT

*Sept 2016 - Present*

- (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 - Present*

- 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.