

Justin Hong

500 W. 120th St.
351H Engineering Terrace
New York NY 10027

978-339-3108
justin.hong@columbia.edu
justinhong.me

Education

- 2022 – **Ph.D. Computer Science (Computational Biology Track)**
COLUMBIA UNIVERSITY, 4.14/4.00
Advisor: Elham Azizi
- 2019 – 2020 **M.S. Electrical Engineering and Computer Science**
UNIVERSITY OF CALIFORNIA, BERKELEY, 4.00/4.00
Advisors: Yun Song, Kannan Ramchandran
Thesis: A Likelihood-based Deconvolution of Bulk Gene Expression Data Using Single-cell References [T1].
- 2015 – 2019 **B.A. Computer Science & Molecular and Cellular Biology (emph. Immunology)**
UNIVERSITY OF CALIFORNIA, BERKELEY, 3.95/4.00
Graduated with Highest Distinction and EECS Department Honors.

Research Experience

- 2021 – 2022 Research Engineer. YOSEF LAB, UC BERKELEY
Advisor: Nir Yosef. Developed open-source software for deep probabilistic modeling of single-cell omics data, scvi-tools [J2]. Contributed to novel research work concerning RNA velocity inference and meta-analysis of large-scale scRNA-seq datasets [W2, J1].
- 2019 – 2020 Graduate Student Researcher. SONG LAB, UC BERKELEY
Advisor: Yun Song. Developed a method for the deconvolution of bulk RNA-seq samples using a single-cell RNA-seq reference [J3].
- 2018 – 2019 Undergraduate Student Researcher. BLISS LAB, UC BERKELEY
Advisor: Kannan Ramchandran. Developed a robust method for the federated learning regime in the presence of adversaries [W3].
- 2016 – 2017 Research Assistant. BREM LAB, UC BERKELEY
Advisor: Rachel Brem. Trained in experimental wet lab protocols concerning gene transformations and knockouts in yeast.

Honors & Awards

- 2023 1st Place GSK.ai CausalBench Challenge [W1]
- 2020 Outstanding Graduate Student Instructor Award, UC Berkeley
- 2019 Graduation with Highest Distinction, UC Berkeley (equiv. *summa cum laude*)
- 2019 EECS Department Honors, UC Berkeley
- 2018 Jim and Donna Gray Scholarship, UC Berkeley
- 2017 Upsilon Pi Epsilon, UC Berkeley

Publications

Star symbol (*) denotes equal contributions as a co-first author.

JOURNAL ARTICLES

[J1] Adam Gayoso, Philipp Weiler, Mohammad Lotfollahi, Dominik Klein, **Justin Hong**, Aaron Streets, Fabian J Theis, and Nir Yosef. “Deep generative modeling of transcriptional dynamics for RNA velocity analysis in single cells”. In: *Nature Methods* (2023), pp. 1–10. [\[URL\]](#).

[J2] Adam Gayoso*, Romain Lopez*, Galen Xing*, Pierre Boyeau, Valeh Valiollah Pour Amiri, **Justin Hong**, Katherine Wu, Michael Jayasuriya, Edouard Mehlman, Maxime Langevin, Yining Liu, Jules Samaran, Gabriel Misrachi, Achille Nazaret, Oscar Clivio, Chenling Xu, Tal Ashuach, Mariano Gabitto, Mohammad Lotfollahi, Valentine Svensson, Eduardo da Veiga Beltrame, Vitalii Kleshchevnikov, Carlos Talavera-López, Lior Pachter, Fabian J. Theis, Aaron Streets, Michael I. Jordan, Jeffrey Regier, and Nir Yosef. “A Python library for probabilistic analysis of single-cell omics data”. In: *Nature Biotechnology* 40.2 (2022), pp. 163–166. [\[URL\]](#).

[J3] Dan D Erdmann-Pham*, Jonathan Fischer*, **Justin Hong**, and Yun S Song. “Likelihood-based deconvolution of bulk gene expression data using single-cell references”. In: *Genome Research* 31.10 (2021), pp. 1794–1806. [\[URL\]](#).

ARTICLES IN CONFERENCE PROCEEDINGS

[J1] Achille Nazaret*, **Justin Hong***, Elham Azizi, and David Blei. “Stable Differentiable Causal Discovery”. In: *Proceedings of the 41st International Conference on Machine Learning (ICML), to appear* (2024). [\[URL\]](#).

REFEREED WORKSHOP PAPERS

[W1] Achille Nazaret and **Justin Hong**. “BetterBoost-Inference of Gene Regulatory Networks with Perturbation Data”. In: *Machine Learning for Drug Discovery (MLDD) Workshop at ICLR, Oral Presentation* (2023). [\[URL\]](#). 1st place submission of GSK.ai Causal-Bench Challenge.

[W2] Pierre Boyeau*, **Justin Hong***, Adam Gayoso, Michael Jordan, Elham Azizi, and Nir Yosef. “Deep generative modeling for quantifying sample-level heterogeneity in single-cell omics”. In: *Machine Learning in Computational Biology (MLCB), Oral presentation* (2022). [\[URL\]](#).

[W3] Avishek Ghosh*, **Justin Hong***, Dong Yin, and Kannan Ramchandran. “Robust Federated Learning in a Heterogeneous Environment”. In: *ICML Workshop on Privacy and Security in ML* (2019). [\[URL\]](#).

PREPRINTS

[P1] Ross M Giglio, Nicholas Hou, Adeya Wyatt, **Justin Hong**, Lingting Shi, Mathini Vaikunthan, Henry Fuchs, Jose Pomarino Nima, Seth W Malinowski, Keith L Ligon, et al. “A heterogeneous pharmaco-transcriptomic landscape induced by targeting a single oncogenic kinase”. In: *bioRxiv* (2024). [\[URL\]](#).

[P2] Jiacheng Gu, ... **Justin Hong** ..., and Jellert T Gaublomme. “CRISPRmap: Sequencing-free optical pooled screens mapping multi-omic phenotypes in cells and tissue”. In: *bioRxiv* (2023). [\[URL\]](#).

THESIS WORK

[T1] **Justin Hong**, Dan D Erdmann-Pham, Jonathan Fischer, and Yun S Song. “A Likelihood-based Deconvolution of Bulk Gene Expression Data Using Single-cell References”. Master’s Thesis. University of California, Berkeley, 2021. [\[URL\]](#).

Presentations

CONTRIBUTED TALKS

Nov 2022 Machine Learning in Computational Biology Workshop (MLCB), Oral Presentation
 May 2022 CZI Assembling Tissue References Workshop

POSTERS

Oct 2023 Single Cell Genomics
 Jul 2023 ICML Workshop on Computational Biology

OTHER MEETINGS AND EVENTS

Dec 2023 Irving Institute for Cancer Dynamics Monthly Research Meeting
 Dec 2021 UC Berkeley Computational Biology Skills Seminar

Software

2023 [SDCD](#) – A method for inferring causal graphs from labeled interventional data.
 2022 [MrVI](#) – A library for deep sample-level meta-analysis of single-cell omics data.
 2021 – 2022 [scvi-tools](#) – A library for deep probabilistic analysis of single-cell omics data.
 2019 – 2020 [RNA-Sieve](#) – A package for deconvolution of bulk RNA-seq data with single-cell RNA-seq references.

Teaching Experience

2023 Teaching Assistant. COLUMBIA UNIVERSITY
Course Title: Probabilistic Models and Machine Learning (STCS 6701)
Responsibilities: Held office hours and recitations, developed and graded assignments.

- 2019–2020 Head Graduate Student Instructor. UC BERKELEY
Course Title: Probability and Random Processes (EE 126)
Responsibilities: Developed course content, assignments, and exams.
 Organized staff of over ten student instructors. Lectured as a substitute.
 Coordinated the course transition to fully online during the COVID-19 pandemic.
- 2018–2019 Undergraduate Student Instructor. UC BERKELEY
Course Title: Probability and Random Processes (EE 126)
Responsibilities: Taught discussions sections, led office hours, created course content.
- 2016–2017 Undergraduate Student Instructor / Course Tutor. UC BERKELEY
Course Title: Structure and Interpretation of Computer Programs (CS 61A)
Responsibilities: Taught discussion sections, led office hours, graded exam content.

Journal and Conference Reviewing

- Nov 2023 Machine Learning in Computational Biology Workshop (MLCB)
 Jul 2023 ICML Workshop on Computational Biology

Mentorship

- Summer 2024 Léa Bohbot, IICD ALLIANCE PROGRAM SUMMER INTERN, ÉCOLE POLYTECHNIQUE
 Spring 2024 Khushi Desai, MASTER'S STUDENT, COLUMBIA UNIVERSITY
 Summer 2023 Tu Duyen Nguyen, IICD ALLIANCE PROGRAM SUMMER INTERN, ÉCOLE POLYTECHNIQUE

Industry Experience

- 2020 – 2021 Software Engineer. NURO
 Developed infrastructure software for the evaluation and introspection of the autonomy software stack. Mentored new hires and a summer intern.
- Summer 2019 Software Engineer Intern. NURO
 Developed software for remote large-scale bot fleet management. Co-inventor of a patent relevant to this project.
- Summer 2018 Software Engineer Intern. PALANTIR
 Developed code editing and execution software within the Foundry platform.
- Summer 2017 Software Engineer Intern. AFFINITY
 Developed software for customer relationship management in the venture capital space.