

Romain Lopez

STANFORD UNIVERSITY
Department of Genetics
Clark Center, 318 Campus Drive
Stanford, CA 94305

+1 (650) 422-9703
romainl@stanford.edu
<https://romain-lopez.github.io/>

Education

- Aug 2016 **Ph.D. Electrical Engineering and Computer Sciences**
–May 2021 UNIVERSITY OF CALIFORNIA, BERKELEY, USA
Dissertation title: *Charting Cellular States, One Cell at a Time: Computational, Inferential and Modeling Perspectives* [PDF]
Affiliations: Berkeley Artificial Intelligence Research, Center for Computational Biology
Advisors: Michael I. Jordan & Nir Yosef
Additional committee members: Sandrine Dudoit & Jennifer Listgarten
- Aug 2013 **Diplome d'Ingénieur; M.S. Applied Mathematics**
–Aug 2016 ÉCOLE POLYTECHNIQUE, FRANCE
Affiliation: Centre de Mathématiques Appliquées
Advisors: Laurent Massoulié & Erwan Le Pennec

Academic Appointments

- Aug 2021 GENENTECH RESEARCH & EARLY DEVELOPMENT, *Postdoctoral Scholar*
– Present STANFORD UNIVERSITY, *Visiting Postdoctoral Scholar*
Hosted by Aviv Regev and Jonathan Pritchard. Exploration of pooled combinatorial genetic screening technologies with single-cell readout.
- Feb 2020 WEIZMANN INSTITUTE OF SCIENCE, *Visiting Research Fellow, Rehovot, Israel*.
Hosted by Ido Amit. Exploration of spatial transcriptomics technologies with a computational angle. Research paper on multi-resolution deconvolution of 10x Visium data, applied to murine lymph nodes and mouse tumor models (DestVI).
- April 2016 HARVARD MEDICAL SCHOOL, *Visiting Research Scholar*
–Aug 2016 Hosted by Allon Klein. Understanding cell fate decisions based on statistical methods for analyzing single-cell RNA sequencing data. Research paper on doublet detection for single-cell transcriptomics data (Scrublet).

Awards & Fellowships

2022	National Interest Waiver award: an immigrant visa that grants lawful US permanent residency to a foreign national without the requirement of an employment offer, based on exceptional abilities. Awarded by the US Immigration Services.
2021	Key Personnel of an Essential Open Source Software for Science grant (PI: Nir Yosef). Awarded by the Chan-Zuckerberg Initiative in support for scvi-tools (\$400k for two years).
2021	Best Paper Award Honorable Mention (with Inderjit Dhillon and Michael I. Jordan), for “Learning From eXtreme Bandit Feedback” in AAAI Conference on Artificial Intelligence.
2019	Travel award. NeurIPS Learning Meaningful Representations of Life Workshop.
2019	Best Student Poster Award (with Achille Nazaret) for “A Joint Model of Unpaired Data from scRNA-seq and Spatial Transcriptomics for Imputing Missing Gene Expression Measurements” in ICML Computational Biology Workshop.
2019	Amazon Web Service Cloud Credit for Research Award (\$20k).
2016	UC Berkeley EECS Departmental Graduate Fellowship. Awarded by the William Oldham Fellowship Fund in Electrical Engineering and Computer Sciences.
2016	Monahan Foundation Fellowship (within the Fulbright-France network). Awarded to students of French scientific institutions to attend graduate school in the US.
2016	Carnot Foundation Fellowship, annually awarded to 2 students from École polytechnique for pursuing graduate studies in the US.
2015	Outstanding Investment Medal, École polytechnique. Awarded annually by the school authorities to 10% of students for their dedication to the student body.
2014	French National Defence Medal, Bronze Echelon. Awarded for public service, helping the young French overseas’ population who face difficulties in building their future.
2011	Languedoc Roussillon Merit Fellowship. Awarded by the French territorial authority to highest honors graduating high-school students, and based on family income.
2011	Black belt, French Judo Federation.

Publications

JOURNAL ARTICLES

Romain Lopez*, Baoguo Li*, Hadas Keren-Shaul*, Pierre Boyeau, Merav Kedmi, David Pilzer, Adam Jelinski, Ido Yofe, Eyal David, et al. “DestVI identifies continuums of cell types in spatial transcriptomics data”. In: *Nature Biotechnology* (2022). [\[PDF\]](#)

Adam Gayoso*, **Romain Lopez***, Galen Xing*, Pierre Boyeau, Katherine Wu, Michael Jayasuriya, Edouard Mehlman, Maxime Langevin, Yining Liu, et al. “A Python library for probabilistic analysis of single-cell omics data”. In: *Nature Biotechnology* (2022). Received an Essential Open Source Software for Science grant from CZI. [\[PDF\]](#)

Adam Gayoso*, Zoë Steier*, **Romain Lopez**, Jeffrey Regier, Kristopher L. Nazor, Aaron Streets, and Nir Yosef. “Joint probabilistic modeling of single-cell multi-omic data with totalVI”. In: *Nature Methods* (2021). [\[PDF\]](#)

Chenling Xu*, **Romain Lopez***, Edouard Mehlman*, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. “Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models”. In: *Molecular Systems Biology* (2021). In the top 20 most downloaded papers of the journal in 2021. [\[PDF\]](#)

Samuel L. Wolock, **Romain Lopez**, and Allon M. Klein. “Scrublet: computational identification of cell doublets in single-cell transcriptomic data”. In: *Cell Systems* (2019). [\[PDF\]](#)

Romain Lopez, Jeffrey Regier, Michael B. Cole, Michael I. Jordan, and Nir Yosef. “Deep generative modeling for single-cell transcriptomics”. In: *Nature Methods* (2018). Presented in a News & Views section. Received a F1000 Exceptional recommendation. [\[PDF\]](#)[\[N&V\]](#)

ARTICLES IN HIGHLY SELECTIVE CONFERENCE PROCEEDINGS

Romain Lopez, Inderjit Dhillon, and Michael I. Jordan. “Learning from eXtreme bandit feedback”. In: *AAAI Conference on Artificial Intelligence* (2021). Selected for a Best Paper Award Honorable Mention. [\[PDF\]](#)

Romain Lopez, Pierre Boyeau, Nir Yosef, Michael I. Jordan, and Jeffrey Regier. “Decision-making with auto-encoding variational Bayes”. In: *Advances in Neural Information Processing Systems* (2020). [\[PDF\]](#)

Romain Lopez, Chenchen Li, Xiang Yan, Junwu Xiong, Michael I. Jordan, Yuan Qi, and Le Song. “Cost-effective incentive allocation via structured counterfactual inference”. In: *AAAI Conference on Artificial Intelligence* (2020). [\[PDF\]](#)

Romain Lopez, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. “Information constraints on auto-encoding variational Bayes”. In: *Advances in Neural Information Processing Systems* (2018). [\[PDF\]](#)

REVIEW ARTICLES

Romain Lopez, Adam Gayoso, and Nir Yosef. “Enhancing scientific discoveries in molecular biology with deep generative models”. In: *Molecular Systems Biology* (2020). [\[PDF\]](#)

MANUSCRIPTS IN SUBMISSION

Romain Lopez, Jan-Christian Hütter, Jonathan K. Pritchard, and Aviv Regev. “Large-Scale Differentiable Causal Discovery of Factor Graphs”. In: *arXiv* (2022). [\[PDF\]](#)

Pierre Boyeau, Jeffrey Regier, Adam Gayoso, Michael I. Jordan, **Romain Lopez***, and Nir Yosef*. “An empirical Bayes method for differential expression analysis of single cells with deep generative models”. In: *bioRxiv* (2022). [\[PDF\]](#)

REFEREED WORKSHOP PAPERS

Khalil Ouardini, **Romain Lopez**, Matthew G. Jones, Sebastian Prillo, Richard Zhang, Michael I. Jordan, and Nir Yosef. “Reconstructing unobserved cellular states from paired single-cell lineage tracing and transcriptomics data”. In: *ICML Workshop in Computational Biology* (2021). Selected for a contributed talk award. [\[PDF\]](#)

Pierre Boyeau, **Romain Lopez**, Jeffrey Regier, Adam Gayoso, Michael I. Jordan, and Nir Yosef. “Deep generative models for detecting differential expression in single cells”. In: *Machine Learning in Computational Biology (MLCB)* (2019). [\[PDF\]](#)

Oscar Clivio, **Romain Lopez**, Jeffrey Regier, Adam Gayoso, Michael I. Jordan, and Nir Yosef. “Detecting zero-inflated genes in single-cell transcriptomics data”. In: *Machine Learning in Computational Biology (MLCB)* (2019). Selected for a spotlight talk. [\[PDF\]](#)

Adam Gayoso, **Romain Lopez**, Zoë Steier, Jeffrey Regier, Aaron Streets, and Nir Yosef. “A joint model of RNA expression and surface protein abundance in single cells”. In: *Machine Learning in Computational Biology (MLCB)* (2019). [\[PDF\]](#)

Romain Lopez*, Achille Nazaret*, Maxime Langevin*, Jules Samaran*, Jeffrey Regier*, Michael I. Jordan, and Nir Yosef. “A joint model of unpaired data from scRNA-seq and spatial transcriptomics for imputing missing gene expression measurements”. In: *ICML Workshop in Computational Biology* (2019). Selected for a spotlight talk and a best student poster award. [\[PDF\]](#)

Maxime Langevin, Edouard Mehlman, Jeffrey Regier, **Romain Lopez**, Michael I. Jordan, and Nir Yosef. “A deep generative model for semi-supervised classification with noisy labels”. In: *Bay Area Machine Learning Symposium* (2018). Selected for an oral presentation. [\[PDF\]](#)

Romain Lopez, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. “A deep generative model for gene expression profiles from single-cell RNA sequencing with application to differential expression”. In: *NeurIPS Machine Learning workshop in Computational Biology* (2017). Selected for a spotlight talk. [\[PDF\]](#)

Romain Lopez, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. “A deep generative model for gene expression profiles from single-cell RNA sequencing”. In: *Bay Area Machine Learning Symposium* (2017). Selected for an oral presentation. [\[PDF\]](#)

Presentations

INVITED TALKS, SEMINARS & KEYNOTES

June 2022	Stanford Statistics Seminar
Jan 2022	Microsoft Research New England, Machine Learning Seminar
Dec 2021	NeurIPS Deep Generative Models and Downstream Applications Workshop, Keynote
Sept 2021	Walter and Eliza Hall Institute of Medical Research, Machine Learning Group, Seminar

April 2021	PyData, AI & Single-cell Genomics, ImmunAI Special Meeting
Dec 2020	Genentech Research and Early Development, Seminar
Nov 2020	Delft University of Technology, Bioinformatics Seminar
May 2020	Amazon, Machine Learning Search Team, Research Presentation
Feb 2020	10x Genomics, Journal Club
Nov 2019	Chan Zuckerberg Initiative & NY Genome Center, Normalization Workshop
Nov 2019	Broad Institute of MIT and Harvard, Models, Inference & Algorithms Special Seminar
Nov 2019	Dana Farber Cancer Institute, Data Science Departmental Seminar
Nov 2019	Pfizer, Machine Learning Seminar
Nov 2019	Celsius Therapeutics, Seminar
Oct 2019	Google Brain Paris, Seminar
March 2019	Deep Learning for Biomedicine conference, UCSF, Guest Speaker
Feb 2017	Two Sigma Investments, Guest Speaker

CONTRIBUTED TALKS

Nov 2020	Seed Networks Annual Meeting, Chan Zuckerberg Initiative, Software Demonstration
Oct 2019	Beyond the Cell Atlas Meeting, Lightning Talk
June 2019	ICML workshop in Computational Biology, Lightning Talk
Oct 2018	Northern California Computational Biology Symposium, Oral Presentation
Dec 2017	NeurIPS Machine Learning workshop in Computational Biology, Lightning Talk
Oct 2017	Northern California Computational Biology Symposium, Oral Presentation
Oct 2017	Bay Area Machine Learning Symposium, Oral Presentation

GROUP MEETINGS

July 2021	BEEHIVE (B. Engelhardt) @ Princeton
May 2021	Marks Lab @ Harvard Medical School
April 2021	Kundaje Lab @ Stanford University
March 2021	Amit Lab @ Weizmann Institute of Science
Dec 2020	Pritchard Lab @ Stanford University
Oct 2020	Battle Lab @ Johns Hopkins University
Oct 2020	Morris Lab @ Memorial Sloan Kettering Cancer Center
July 2020	Applied Bayesian Group (J. Regier) @ University of Michigan
Nov 2019	Regev Lab @ Broad Institute of MIT and Harvard
Sept 2018	Biostatistics Lab (Dudoit, Purdom) @ UC Berkeley

POSTERS

July 2021	ICML Workshop in Computational Biology
Jan 2021	AAAI Conference in Artificial Intelligence
Dec 2020	Advances in Neural Information Processing Systems
Feb 2020	AAAI Conference in Artificial Intelligence
Dec 2019	NeurIPS Workshop on Learning Meaningful Representations of Life
Dec 2019	Machine Learning in Computational Biology meeting
Oct 2019	Probabilistic Modeling In Genomics
Sept 2019	Single-cell Genomics conference
Jun 2019	ICML Workshop in Computational Biology
Jun 2019	UC-wide AI in Biomedicine Symposium
Dec 2018	Advances in Neural Information Processing Systems
Oct 2018	Single-cell Genomics conference
March 2018	Single-cell Biology conference
Dec 2017	NeurIPS Machine Learning workshop in Computational Biology

Teaching & Mentoring Experience

2017 - 2021	<p>SELECTED GUEST LECTURES</p> <p>Caltech, <i>Representation Learning for Science</i> (CS159, Spring 2022)</p> <p>MIT, <i>Deep Learning in the Life Sciences</i> (6.874, Spring 2021)</p> <p>Yale University, <i>Deep Learning Theory and Applications</i>, (CPSC663, Spring 2020)</p> <p>UC Berkeley, <i>Machine Learning and Statistics meet Biology</i>, (CS294, Spring 2017)</p>
May 2018 –May 2021	<p>SINGLE-CELL VARIATIONAL INFERENCE TEAM, UC BERKELEY,</p> <p><i>Visiting Student Recruiting & Mentoring.</i></p> <p>Recruitment and mentoring of 7 master’s students and 2 undergraduate students visiting the Yosef Lab while working on their thesis (full-time five-month internships). Designed screening exams, conducted interviews, provided projects and organized regular working group with students.</p> <p>The team’s work resulted in 7 peer-reviewed publications, 2 outstanding research internship award from École polytechnique, 1 best student poster award at ICML WCB 2019 and 1 contributed talk award at ICML WCB 2021. 6 students accepted PhD positions in computer science departments of top academic institutions (including ENS Paris, UC Berkeley, Columbia and Oxford university).</p>
Spring 2019 – Fall 2019	<p>ELECTRICAL ENGINEERING 127 / 227A, <i>University of California, Berkeley</i></p> <p>Advanced undergraduate and graduate course in convex optimization (250 students).</p> <p>Head Graduate Student Instructor (Fall 2019).</p> <p>Graduate Student Instructor (Spring 2019).</p>

Industry Experience

Sept 2019 –Apr 2020	AMAZON, <i>Applied Scientist Intern, Berkeley</i> . Hosted by Inderjit Dhillon. Research paper on counterfactual inference with extremely large action spaces. Application to search algorithms for Amazon online platform.
June 2018 –Aug 2018	ANT FINANCIAL SERVICES GROUP, <i>Research-based Software Engineer Intern, Hangzhou</i> . Hosted by Le Song. Research paper on counterfactual inference for estimating responses to economical incentives. Application to efficient coupon allocation for mobile marketing campaigns.
Aug 2017 –Jan 2018	CODI (FORMERLY HIVEN), <i>Entrepreneurship project, UC Berkeley</i> . Customer discovery and prototype at early stage of the company. Codi connects remote workers with home-based workspaces right in their neighborhood. The company raised \$7M in 2020.
Sept 2015 –Feb 2016	CARDIOLOGS, <i>Data Scientist Intern, Paris</i> . Hosted by Jia Li, Co-founder & CSO. Cardiologs develops a FDA-cleared AI based EKG analysis software and got acquired by Phillips in 2021. Reconstructed EKGs 3D signal from a 2D projection using convolutional neural networks.
June 2015 –Aug 2015	AXA LIFE JAPAN, <i>Actuarial Intern, Tokyo</i> . Hosted by Takashi Nojima, Head of Pricing and Product development. AXA Life Japan was the second most important subsidiary of AXA group regarding medical insurance in 2015. Predictive modeling, pricing sheets, stress tests and technical reports.

Professional Service

WORKSHOP ORGANIZATION COMMITTEE MEMBERSHIP

NeurIPS LMRL - Learning Meaningful Representations of Life (2021)

JOURNAL & CONFERENCE REVIEWING

Science, 2022– (1[†] paper)

Nature Biotechnology, 2022– (1 paper)

Foundations and Trends in Machine Learning, 2021– (1 paper)

Nature Review Genetics, 2021– (1[†] paper)

Science Advances, 2020– (1 paper)

Bioinformatics, 2020– (1 paper)

Nature Methods, 2020– (1[†] paper)

International Conference on Artificial Intelligence and Statistics (AISTATS), 2022
 AAAI Conference on Artificial Intelligence (AAAI), 2021
 International Conference on Learning Representations (ICLR), 2021
 Neural Information Processing Systems (NeurIPS), 2019, 2020, 2021
 International Conference on Machine Learning (ICML), 2019, 2021

ICML Workshop in Computational Biology (WCB), 2020, 2022
 Machine Learning in Computational Biology (MLCB), 2019, 2020, 2021

†: assisting a senior reviewer.

Leadership & Outreach

Nov 2021 –May 2022	PROMOTING INCLUSIVITY IN COMPUTING, SFSU & GENENTECH, <i>Course Contributor</i> . Diverse feedback and contributions on the coursework materials designed for San Francisco State University’s undergraduate Certificate in Data Science and Machine Learning for Biotechnology.
Oct 2019 –Aug 2020	DIVERSIFYING ACCESS TO RESEARCH IN ENGINEERING, UC BERKELEY, <i>Student Mentor</i> . Provide undergraduate students from under-represented background with research opportunities in electrical engineering and computer science to promote diversity.
Sept 2017 –Aug 2019	FRENCH ALUMNI BERKELEY, <i>Founder</i> . Connecting Berkeley students that share a part of their education in France with Alumni. Organized monthly meetings with startups in San Francisco, bi-monthly networking events on campus. Collaboration with the French consulate, industry and diverse associations.
March 2015	FRENCH TOURNAMENT OF YOUNG MATHÉMATIENNES ET MATHÉMATIENS, <i>Member of the Final Jury</i> . Participated as a jury member of the finale for the national mathematics tournament for high school students, hosted by École polytechnique.
Oct 2014 –Jul 2015	FRESHMAN WEEKEND OF ÉCOLE POLYTECHNIQUE, <i>Treasurer & Vice-President</i> . In charge of the \$160k budget and co-organising the event for 600 students.
Oct 2013 –Apr 2014	FRENCH MINISTRY OF DEFENCE, <i>Officer Cadet, Reunion Island, Indian Ocean</i> . Military training for underprivileged youth towards the job market. Supervised the military recruit training of thirty people and their five supervisors.
2008 – 2011	JUDO CLUB, <i>Montpellier Area, France</i> . Participated in national level Judo tournaments and trained young judokas: awarded black belt at the age of fourteen. Released an open source Judo scoreboard software to empower

youth to refereeing. Prepared a computer park to be used for educational purposes during a humanitarian mission during twenty days in Senegal with my Judo Club.

Media Coverage

PRESS

[“The convergence of deep neural networks and immunotherapy”](#). In: *Tech Crunch* (January 2022).

[“CZI awards \\$16 million for foundational open source software tools essential to biomedicine”](#). In: *Chan Zuckerberg Initiative Newsroom* (August 2021).

[“Two Amazon papers were runners-up for best-paper awards at AAAI”](#). In: *Amazon Science* (March 2021).

[“Bayesian deep learning for single-cell analysis”](#). In: *Nature Methods* (Nov 2018).

PODCASTS

[“scVI with Romain Lopez and Gabriel Misrachi”](#). In: *The Bioinformatics Chat* (Sept 2019).

BLOG POSTS

[“Behind the paper: DestVI identifies continuums of cell types in spatial transcriptomics data”](#). In: *Nature Portfolio Bioengineering* (Apr 2022).

[“Experiments with scVI”](#). In: *Saket Choudhary’s personal blog* (Dec 2020).

[“Hyperparameter search for scVI”](#). In: *YosefLab Blog* (July 2019).

[“Should we zero-inflate scVI?”](#) In: *YosefLab Blog* (June 2019).

[“Building gene expression atlases with deep generative models for single-cell transcriptomics”](#). In: *Berkeley Artificial Intelligence Research Blog* (Dec 2018).

[“Integrating scRNA-seq and spatial STARmap data from mouse frontal cortex with scVI”](#). In: *What Do You Mean “Heterogeneity”?* (Oct 2018).

[“Count based autoencoders and the future for scRNA-seq analysis”](#). In: *What Do You Mean “Heterogeneity”?* (Apr 2018).