Curriculum Vitae

Date Prepared: May 14, 2020

Name: Bo Li

Office Address: Massachusetts General Hospital
149 13th Street, Room 8214
Charlestown, MA 02129

Work Phone: 617/724-2641

Work Email: bli28@mgh.harvard.edu

Education:

9/2004-7/2008	BE	Computer Science and Engineering	Shanghai Jiao Tong University
9/2008-5/2010	MS	Computer Science	University of Wisconsin-Madison
9/2008-7/2013	PhD	Computer Science (Colin Dewey, PhD)	University of Wisconsin-Madison

Postdoctoral Training:

8/2013-6/2017	Postdoctoral Researcher	Computational Biology (Lior Pachter, PhD)	University of California, Berkeley
7/2017-4/2019	Postdoctoral Associate	Computational Biology (Aviv Regev, PhD)	Broad Institute of MIT and Harvard

Faculty Academic Appointments:

4/2020-	Assistant Professor	Medicine	Harvard Medical School
5/2019-3/2020	Member of the Faculty	Medicine	Harvard Medical School

Appointments at Hospitals/Affiliated Institutions:

5/2019-	Research Staff (Assistant Investigator)	Center for Immunology and Inflammatory Diseases, Division of Rheumatology, Allergy and Immunology	Massachusetts General Hospital
5/2019-	Associate Scientist	Klarman Cell Observatory	Broad Institute of MIT and Harvard

Major Administrative Leadership Positions:

Local

5/2019-	Director, Bioinformatics and Computational Biology Program, Center for Immunology and Inflammatory Diseases	Massachusetts General Hospital

Committee Service:

Local

2014-2015	Postdoc Industry Exploration Program (PIEP)	University of California, Berkeley Committee Member

International

2016	Program Committee Member, RECOMB Satellite Workshop on Massively Parallell Sequencing (RECOMB-SEQ)	International Conference on Research in Computational Molecular Biology (RECOMB)
2016	Program Committee Member	International Joint Conferences on Artificial Intelligence

Professional Societies:

2012 2014	International Society for Computational Biology

Editorial Activities:

Ad hoc Reviewer

BMC Bioinformatics
Bioinformatics
Cell Reports
GigaScience
Nucleic Acids Research

Honors and Prizes:

2004	First Place, Association for Computing Machinery – International Collegiate Programming Contest (ACM-ICPC) Asia Programming Contest, Beijing Site	ACM-ICPC Committee	Programming Contest
2005 & 2006	Scholarship	Orient Overseas Container Line Corporation, China	

2005	Fifth Place, ACM-ICPC Asia Programming Contest, Taipei Site	ACM-ICPC Committee	Programming Contest
2005	Second Place, ACM-ICPC Asia Programming Contest, Beijing Site	ACM-ICPC Committee	Programming Contest
2006	Thirteenth Place, Google Code Jam, China	Google, China	Programming Contest
2007	State Scholarship	Ministry of Education, China	
2008	Alumni Scholarship	Computer Sciences Department, University of Wisconsin-Madison	
2010-2012	Morgridge Institute for Research Support for Senior Graduate Students in Computation and Informatics in Biology and Medicine	Morgridge Institute for Research	Research
2013	Finalist, Lane Fellows in Computational Biology	Carnegie Mellon University	Postdoctoral Fellowship

Report of Funded and Unfunded Projects

Funding Information:

Current

2019-2022 Li Start-up Sundry
PI
Laboratory start-up institutional support from Massachusetts General Hospital

Projects Submitted for Funding

Pending Cumulus: a cloud-based data analysis framework for large-scale single-cell and single-nucleus genomics
NHGRI/NIH R35 RFA-HG-18-006
PI ($1,500,000; total direct costs)
This grant proposes to develop Cumulus as a cloud-based framework for analyzing large scale single-cell and single-nucleus genomics data. Th end product, Cumulus, will support the analysis of large-scale single-cell/nucleus RNA-seq (sc/snRNA-seq), single-cell ATAC-seq (scATAC-seq) and single-cell immune repertoire sequencing (scIR-seq) data, which are three of the most important single-cell and single-nucleus genomics data types.

Pending Single-cell genomics dissection of common immune networks driving autoimmunity

JDRF, LRA & NMSS Joint RFA: Decoding Immune Mediated Diseases – Novel Approaches for Therapeutic Insights
Co-PI ($227,790; total direct costs)
This grant proposes to characterize tissue lesions and matched blood across patients diagnosed with lupus nephritis, multiple sclerosis, rheumatoid arthritis, Crohn's disease, as well as Type I diabetes using single-cell multi-omics technologies. The goal is to establish cellular and molecular definition of autoimmunity and identify specifically the cellular states impacted by these susceptibility variants that should be targeted therapeutically by analyzing the large-scale dataset that we plan to generate. My team is responsible for processing and analyzing all single-cell multi-omics data that will be generated from this proposal.

Report of Local Teaching and Training

Teaching of Students in Courses:

Teaching Prior to Current Appointment

2006	Programming Practice of Computer Algorithms 2nd year computer science students	Shanghai Jiao Tong University 1 hour session/week for 26 weeks
2007	Project Workshop Compiler 2nd year computer science students	Shanghai Jiao Tong University 1 hour session/month for 4 months
2008	Introduction to Bioinformatics Senior undergraduates and 1st year graduate students in computer science or biostatistics and medical informatics	University of Wisconsin-Madison Teaching Assistant 1 hour office sessions/week for 14 weeks

Research Supervisory and Training Responsibilities:

2018	Supervision of visiting undergraduate student	Broad Institute of MIT and Harvard Dr. Aviv Regev's Laboratory Daily mentorship for one month
2019-	Supervision of computational biologist	Massachusetts General Hospital Weekly mentoring (15 hours/week)
2019-	Bulk and single-cell genomics data analysis training CIID research fellows and students	Massachusetts General Hospital 1:1 supervision; 56 hours/year

Local Invited Presentations:

No presentations below were sponsored by outside entities

2015	Core Skills in Computational Biology / Lecture RNA-Seq Transcript Quantification with RSEM: A Detailed Tutorial with Common Use Cases, University of California, Berkeley, Berkeley, CA
2016	PROBer: A General Toolkit for Analyzing Sequencing-based 'Toeprinting' Assays / Invited Speaker Center for RNA Systems Biology, University of California, Berkeley, Berkeley, CA

2018 Our Journey Towards the Human Immune Cell Atlas / Invited Speaker (co-presented with Dr. Alexandra-Chloé Villani)
Annual Meeting of the Scientific Advisory Board, Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA

2018 Scaling Computational Pipeline for Massive Datasets / Invited Speaker
Human Tumor Atlas Pilot Project Site Visit, Broad Institute of MIT and Harvard, Cambridge, MA

2018 Human Immune Cell Atlas -- An Overview / Invited Speaker
Symposium for NIH Director Francis Collins, Broad Institute of MIT and Harvard, Cambridge, MA

2018 Nuclei-Hashing for Single-Nucleus RNA-Seq Experiments / Invited Speaker
Comprehensive Center for Mouse Brain Cell Atlas Biannual Symposium, Broad Institute of MIT and Harvard, Cambridge, MA

2019 Computational Analysis of Cell Hashing/Nucleus Hashing/CITE-Seq Data / Lecture
Single Cell Working Group, Broad Institute of MIT and Harvard, Cambridge, MA

2019 scCloud: Cloud-Based Data Analysis for Large-Scale Single-Cell and Single-Nucleus Genomics / Invited Speaker
7th Annual Klarman Cell Observatory Retreat, Broad Institute of MIT and Harvard, Cambridge, MA

2019 Scale up Analysis to Millions of Single Cells, Faster and Cheaper / Invited Speaker
Bill & Melinda Gates Medical Research Institute Visit, Broad Institute of MIT and Harvard, Cambridge, MA

2019 scCloud: Cloud-Based Data Analysis for Large-Scale Single-Cell and Single-Nucleus Genomics / Invited Speaker
Bristol-Myers Squibb Visit, Broad Institute of MIT and Harvard, Cambridge, MA

2019 scCloud: Cloud-Based Data Analysis for Large-Scale Single-Cell and Single-Nucleus Genomics / Invited Speaker
ImmGen 2019 Computational Workshop, Harvard Medical School, Boston, MA

Report of Regional, National and International Invited Teaching and Presentations

No presentations below were sponsored by outside entities

Regional

2018 How to build a successful human cell atlas? Lessons learned from 1.7 million single immune cells / Invited Speaker
Immunology Seminar Series, Massachusetts General Hospital, Boston, MA

2019 Nucleus Hashing: Multiplexing snRNA-Seq with Barcoded Antibodies / Invited Speaker
Comprehensive Center for Mouse Brain Cell Atlas Virtual Symposium, BRAIN Initiative Cell Census Network (online meeting)

2020 Cumulus: cloud-based data analysis framework for large-scale single-cell and single-nucleus RNA-seq / Invited Speaker

Cell Circuits and Epigenomics & Immunology Program, Broad Institute of MIT and Harvard, Cambridge, MA (online meeting)

National

2013 Computational Analysis of RNA-Seq Data in the Absence of a Known Genome: from Transcript Quantification to De Novo Transcriptome Assembly Evaluation / Invited Speaker
Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA

2017 Taming Big Sequencing Data for RNA Biology: From Transcript Abundance Estimation to 'Epitranscriptomic' Mark Detection / Invited Speaker
Toyota Technological Institute at Chicago, Chicago, IL

2017 Taming Big Sequencing Data for RNA Biology: From Transcript Abundance Estimation to 'Epitranscriptomic' Mark Detection / Invited Speaker
Department of Computer Science, University of Illinois at Urbana-Champaign, IL

2017 Taming Big Sequencing Data for RNA Biology: From Transcript Abundance Estimation to 'Epitranscriptomic' Mark Detection / Invited Speaker
Computer Science Colloquium, Duke University, Durham, NC

2019 Immunology Research in the Era of Single-Cell Genomics - Lessons Learned from 1.7 Million Single Immune Cells / Invited Speaker
Biostatistics & Medical Informatics Department Seminar Series, University of Wisconsin-Madison, Madison, WI

2019 scCloud: Cloud-Based Data Analysis for Large-Scale Single-Cell and Single-Nucleus Genomics / Invited Speaker
Human Cell Atlas Analysis Community Meeting (online meeting)

2019 Cumulus: cloud-based data analysis framework for large-scale single-cell and single-nucleus RNA-seq / Invited Speaker
Duke StatGen Seminar Series, Duke University, Durham, NC

2020 Cumulus: cloud-based data analysis framework for large-scale single-cell and single-nucleus RNA-seq / Invited Speaker
HTAN sc/snRNA-seq subgroup meeting, National Cancer Institute (online meeting)

2020 Cumulus: cloud-based data analysis framework for large-scale single-cell and single-nucleus RNA-seq / Invited Speaker
GeneLab, National Aeronautic and Space Administration (online meeting)

International

2016 Quantifying RNA Information from Transcriptome-Wide Chemical Probing Experiments / Invited Speaker
Next Generation Sequencing (NGS) Data Analysis and Informatics Conference, San Diego, CA

2016 PROBer: A General Toolkit for Analayzing Sequencing-based 'Toeprinting' Assays (selected oral abstract)
21st Annual Meeting of the RNA Society, Kyoto, Japan

Report of Technological and Other Scientific Innovations

RSEM software for bulk RNA-seq data gene and isoform expression estimation	Developed in 2010 and 2011 / http://deweylab.github.io/RSEM/ RSEM (RNA-Seq by Expectation Maximization) is a software I had developed when I was a graduate student. It now becomes the gold standard of quantifying expression levels for bulk RNA-seq and plate-based SMART-seq2 single-cell RNA-seq data and is widely used internationally.
CSEM software for allocating multi-mapping reads in ChIP-seq data	Developed in 2011 / http://deweylab.biostat.wisc.edu/csem/ CSEM (ChIP-Seq multi-read allocation using Expectation-Maximization) is a software I had developed when I was a graduate student. CSEM was the first multi-mapping read allocator designed for ChIP-seq data and was discussed in In Brief section of *Nature Reviews Genetics*. 2011; 12:588.
RSEM-EVAL software for reference-free evaluation of de novo transcriptome assembly	Developed in 2014 as part of the DETONATE software suite / http://deweylab.biostat.wisc.edu/detonate/ RSEM-EVAL is a software tool I had developed based on RSEM. It is the first reference-free evaluator of *de novo* transcriptome assemblies and is used internationally for choosing the best quality *de novo* transcriptome assemblies.
Methods and compositions for multiplexing single cell and single nuclei sequencing	US Patent Application, 62/770,580, filed December 7, 2018, pending As a member of the Regev lab at the Broad Institute of MIT and Harvard, my colleagues and I developed methods to multiplexing single cells and single nuclei from different samples experimentally and demultiplex sequencing data to sample-specific cells and nuclei computationally. These methods has the potential of significantly reducing the experimental cost of single cell and single nucleus RNA-Seq experiments internationally. My contribution is the development of the computational method (DemuxEM) for demultiplexing the pooled sequencing data.
Cumulus software/framework for cloud-based single-cell and single-nucleus RNA-seq data analysis	Developed in 2019 / https://cumulus.readthedocs.io/en/latest/ Cumulus is a software / framework that I have developed at the Broad Institute. It is the first comprehensive cloud-based large-scale single-cell and single-nucleus RNA-seq data analysis framework. Currently, it has been used in several big cell atlas consortia, such as Human Immune Cell Atlas and Human Tumor Atlas Pilot Project.

Report of Scholarship

Peer-Reviewed Scholarship in print or other media:

Research Investigations

1. **Li B**, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* 2010; 26(4):493-500.

2. Chung D, Kuan PF, **Li B**, Sanalkumar R, Liang K, Bresnick EH, Dewey C, Keleş S. Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data. *PLoS Comput Biol*. 2011; 7(7):e1002111.
    - In Brief. *Nature Reviews Genetics*. 2011; 12:588.

3. **Li B**, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; 12:323.

4. Haas BJ*, Papanicolaou A*, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, **Li B**, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 2013; 8(8):1494-512.

5. **Li B***, Fillmore N*, Bai Y, Collins M, Thomson JA, Stewart R, Dewey CN. Evaluation of *de novo* transcriptome assemblies from RNA-Seq data. *Genome Biol*. 2014; 15(12):553.

6. Zeng X, **Li B**, Welch R, Rojo C, Zheng Y, Dewey CN, Keleş S. Perm-seq: Mapping Protein-DNA Interactions in Segmental Duplication and Highly RepetitiveRegions of Genomes with Prior-Enhanced Read Mapping. *PLoS Comput Biol*. 2015; 11(10):e1004491.

7. Choudhary K, Shih NP, Deng F, Ledda M, **Li B**, Aviran S. Metrics for rapid quality control in RNA structure probing experiments. *Bioinformatics* 2016; 32(23):3575-3583.

8. **Li B**, Tambe A, Aviran S, Pachter L. PROBer Provides a General Toolkit for Analyzing Sequencing-Based Toeprinting Assays. *Cell Syst*. 2017; 4(5):568-574.

9. Gaublomme JT*†, **Li B***, McCabe C, Knecht A, Yang Y, Drokhlyansky E, Van Wittenberghe N, Waldman J, Dionne D, Nguyen L, De Jager PL, Yeung B, Zhao X, Habib N, Rozenblatt-Rosen O†, Regev A†. Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat Commun*. 2019; 10: 2907.

10. Popescu DM*, Botting RA*, Stephenson E*, Green K, Webb S, Jardine L, Calderbank EF, Polanski K, Goh I, Efremova M, Acres M, Maunder D, Vegh P, Gitton Y, Park JE, Vento-Tormo R, Miao Z, Dixon D, Rowell R, McDonald D, Fletcher J, Poyner E, Reynolds G, Mather M, Moldovan C, Mamanova L, Greig F, Young MD, Meyer KB, Lisgo S, Bacardit J, Fuller A, Millar B, Innes B, Lindsay S, Stubbington MJT, Kowalczyk MS, **Li B**, Ashenberg O, Tabaka M, Dionne D, Tickle TL, Slyper M, Rozenblatt-Rosen O, Filby A, Carey P, Villani AC, Roy A, Regev A, Chédotal A, Roberts I, Göttgens B, Behjati S, Laurenti E†, Teichmann SA†, Haniffa M†. Decoding human fetal liver haematopoiesis. *Nature* 2019; 574:365-371.

11. Haas BJ†, Dobin A, **Li B**, Stransky N, Pochet N, Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* 2019; 20(1):213.

12. Sen P, Wilkie AR, Ji F, Yang Y, Taylor IJ, Velazquez-Palafox M, Vanni EAH, Pesola JM, Fernandez R, Chen H, Morsett LM, Abels ER, Piper M, Lane RJ, Hickman SE, Means TK, Rosenberg ES, Sadreyev RI, **Li B**, Coen DM, Fishman JA, El Khoury J. Linking indirect effects of cytomegalovirus in transplantation to modulation of monocyte innate immune function. *Sci Adv.* 2020; 6(17):eaax9856.

13. Slyper M*, Porter CBM*, Ashenberg O*, Waldman J, Drokhlyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, Vigneau S, Jané-Valbuena J, Tickle TL, Napolitano S, Su M, Patel AG, Karlstrom A, Gritsch S, Nomura M, Waghray A, Gohil SH, Tsankov AM, Jerby-Arnon L, Cohen O, Klughammer J, Rosen Y, Gould J, Nguyen L, Hofree M, Tramontozzi PJ, **Li B**, Wu CJ, Izar B, Haq R, Hodi FS, Yoon CH, Hata AN, Baker SJ, Suvà ML, Bueno R, Stover EH, Clay MR, Dyer MA, Collins NB, Matulonis UA, Wagle N, Johnson BE, Rotem A, Rozenblatt-Rosen O†, Regev A†. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat Med.* 2020; 26:792-802.

* contributed equally
† corresponding author

Thesis:

Li B. Computational analysis of RNA-Seq data in the absence of a known genome [dissertation]. Madison, Wisconsin: University of Wisconsin-Madison; July 2013.

Narrative Report

I am a Principal Investigator in the Center for Immunology and Inflammatory Disease (CIID) at Massachusetts General Hospital. I am also Director of the Bioinformatics and Computational Biology Program in the CIID, and hold an Associate Scientist appointment at the Klarman Cell Observatory at the Broad Institute of MIT and Harvard. I currently commit 70% of my time to my Area of Excellence – Investigation, 20% to training and mentoring students, and 10% to Administration & Institutional Service.

I have over 12 years of experience working in the field of bioinformatics and computational biology. I received my bachelor's degree in computer science and engineering from Shanghai Jiao Tong University. I received my Ph.D. in Computer Sciences from University of Wisconsin-Madison, focusing on developing fundamental tools for RNA-sequencing (RNA-Seq) data analysis. I developed the world-renowned RNA-Seq transcript abundance estimation software, RSEM (Li et al., *Bioinformatics* 2010; Li and Dewey, *BMC Bioinform.* 2011), which has been cited 8,080 times (Google Scholar). RSEM is selected as the most preferable bulk RNA-Seq expression-level estimation tool by nationally funded consortia such as ENCODE (ENCyclopedia Of DNA Elements) and TCGA (The Cancer Genome Atlas), and is recommended by Human Cell Atlas (HCA) as a five-star single-cell RNA-Seq (scRNA-Seq) expression quantification tool for plate-based SMART-Seq2 data. Besides RSEM, I also developed the first reference-free *de novo* transcriptome assembly evaluator, RSEM-EVAL (Li*, Fillmore* et al., *Genome Biol.* 2014), that can select the best assembly without knowing the ground truth transcript sequences, and the first multi-mapping read allocator for ChIP-Seq data, CSEM (Chung, Kuan, Li et al., *PLOS Comput. Biol*. 2011; Zeng, Li et al., *PLOS Comput. Biol.* 2015). I completed my first postdoctoral training with Dr. Lior Pachter at University of California, Berkeley, focusing on developing tools that are critical for the study of post-transcriptional dynamics of RNAs, such as RNA secondary structures, RNA modifications and RNA-protein interactions (Li et al., *Cell Syst.* 2017). I completed my second postdoctoral training with Dr. Aviv Regev at the Broad Institute of MIT and Harvard, focusing on large-scale single-cell/single-nucleus RNA-Seq (sc/sn RNA-Seq) data analysis and systems immunology. I lead the computational efforts of the human Immune Cell Atlas (ICA) project, which aims at identifying all cell types and cell states of the human immune system and thus providing the community a comprehensive roadmap of the immune system to better dissect the drivers of immune disorders and responses to treatments. We have profiled the transcriptomes of 1.7 million human immune cells so far and obtained over 5 terabytes of sequencing data. To process datasets of such volumes, I developed Cumulus (under review in *Nature Methods*), the first sc/sn RNA-Seq analysis framework that combines the power of cloud computing with algorithm and implementation improvements to achieve high scalability, low cost and user-friendliness. Besides the ICA project, Cumulus is also used in National Cancer Institute's Human Tumor Atlas Pilot Project. In addition, Jellert Gaublomme and I developed nucleus hashing, an antibody-barcoding-based snRNA-Seq multiplexing protocol that effectively reduces batch effects and experimental costs, and DemuxEM, a demultiplexing algorithm that accurately detects inter-sample multiplets and assigns singlets to their sample of origin (Gaublomme*†, Li* et al., *Nat Commun*. 2019).

As an MGH faculty member, my research program aims at developing standard operating procedures (SOPs) for scalable single-cell and single-nuclei genomics data analysis. Currently, we are continuing our efforts to develop and maintain the Cumulus framework as an open source software for the community. The new features we are working on include a) benchmarking and developing novel tools for gene-count matrix construction that can further lower the cloud-based computational cost; b) enabling interactive data analysis and visualization on the cloud so that Cumulus is user-friendly to biologists without programming experience; c) extending Cumulus to analyze single-cell ATAC-Seq (scATAC-Seq) data for chromatin accessibility and single-cell immune profiling data (scIP) for constructing B cell and T cell receptors; d) developing novel algorithms to mine biology of interest from single-cell multi-omics data, consisting of scRNA-Seq, scATAC-Seq, CITE-Seq, and scIP data. Our efforts will result in a powerful cloud-based analysis framework that will free immunologists from complicated computation and help them better focus on deciphering the mechanisms of

immunology disorders thus enhancing patient care. Since the bioinformatics challenges we address are ubiquitous in the broader field of genomics, the tools my lab will have significant impact on the field of genomics. Besides developing fundamental tools, my lab is actively collaborating with immunologists on addressing key immunology questions. Dr. Villani and I co-lead the human blood atlas effort, which is part of the Human Cell Atlas Initiative and aims at mapping all cell types with a frequency of at least ~ 0.1%. In addition, my lab is collaborating with Drs. Chloe Villani, Andrew Luster, Jim Moon, Ben Medoff and Joe El Khoury on projects covering a variety of topics such as immune-related adverse events, asthma and allergic airway inflammation, and virus and fungi infection in heart transplantation.

Throughout my training, I was awarded several prizes and scholarships and invited to deliver presentations in local, regional, national and international forums. I have also established a good track record of teaching and mentoring. I started teaching as a junior undergraduate student and was responsible for designing syllabi, giving lectures, and evaluating programming homework and exams in the first course I taught. Since 2015, I have been giving lectures on topics related to bulk and single-cell RNA-Seq data analysis. During my second postdoctoral training, I mentored an undergraduate intern daily for one month, providing guidance on large-scale data analysis using the cloud environment. At MGH, I mentor and train a computational scientist daily. I also conduct a bulk and single-cell genomics training program for CIID research fellows and students. In addition, I have a strong interest in teaching courses related to computational biology/bioinformatics at undergraduate and/or graduate levels. As an instructor for bioinformatics courses, I will supplement the classic topics, such as sequence alignment and phylogeny, with recent algorithmic developments in single-cell genomics.

I have served on the committee of the Postdoc Industry Exploration Program (PIEP) at UC Berkeley. PIEP aims at helping postdocs to better connect to the industry through organizing workshops and industrial site visits. As a committee member, I organized two industrial site visits (30 postdocs/visit) at Facebook and ThermoFisher Scientific. These efforts empowered me to further develop my leadership skills.

In summary, my training has prepared me for a successful career as an independent investigator, building an impactful bioinformatics infrastructure that will benefit the HMS immunology community.