Shilpa Nadimpalli Kobren

+1 978-407-8042 • shilpa_kobren@hms.harvard.edu • http://shilpakobren.com

Education

2018 **Ph.D. in Computer Science** • Princeton University
Thesis: Detecting and Analyzing Variation in Protein Interaction Interfaces
Advisor: Mona Singh

2013 **M.Sc. in Computer Science** • Princeton University

2011 **B.Sc. in Biology & Computer Science**, *summa cum laude* • Tufts University
Thesis, Awarded Highest Honors: Correcting Protein Multiple Structural Alignments by Sequence Peeking
Advisor: Lenore Cowen

Awards & Honors

10/2021–04/2022 Amazon Web Services (AWS) Cloud Credit for Research Award *totalling $32,000*

10/2021 Outstanding Postdoctoral Researcher *awarded annually to one member of each department at Harvard University*

08/2015 **Siebel Scholar, Class of 2016** *awarded annually for academic excellence and demonstrated leadership to 93 top students from the world's leading graduate schools*

09/2011–09/2015 **NSF Graduate Research Fellowship** *for graduate study in STEM disciplines*

09/2011–09/2015 **Princeton Gordon Wu Fellowship** *for engineering graduate study*

05/2011 **Tufts James Schmolze Award** *for top undergraduate achievement in computer science*

05/2010 **Google Anita Borg Memorial Scholarship** *for women in computer science*

08/2007–05/2011 Tufts Neubauer Scholarship *for aptitude in research; grants to fully cover all four years of undergraduate student loans (1 in 10 of ∼1500)*

04/2011 Tufts Class of 1882 Prize Scholarship *for intellectual leadership (1 in 4 of ∼5000)*

04/2010 Tufts Audrey Butvay Gruss Award *for scientific achievement (1 in 4 of ∼5000)*

06/2009–10/2009 CRA-W/CDC Distributed Research Experiences for Undergraduates Fellowship

09/2008–05/2009 CRA-W Multidisciplinary Research Opportunities for Women Fellowship

06/2008–09/2008 NSF/NIH Southern California Bioinformatics Summer Institute Fellowship

05/2006 Massachusetts State Science Fair, 1st Place Award

Publications

Under Review

SN Kobren*, M Moldovan*, RM Reimers, D Traviglia, X Li, D Barnum, A Veit, RI Corona, GVC Neto, J Willett, M Berselli, W Ronchetti, SF Nelson, JA Martinez-Agosto, R Sherwood, JB Krier, IS Kohane, Undiagnosed Diseases Network, SR Sunyaev. "Joint, multifaceted genomic analysis enables diagnosis of diverse, ultra-rare monogenic presentations." *bioRxiv*.

E Alsentzer*, MM Li*, **SN Kobren**, Undiagnosed Diseases Network, IS Kohane, M Zitnik. "Few shot learning for phenotype-driven diagnosis of patients with rare genetic diseases." *medRxiv*.

L Fitzsimmons, B Beaulieu-Jones†, **SN Kobren**†. "Phenotypic overlap between rare disease patients and variant carriers in a large population cohort informs biological mechanisms." *medRxiv*.

R Yin, A Gutierrez, Undiagnosed Diseases Network, **SN Kobren**†, P Avillach†. "VarPPUD: Variant post prioritization for undiagnosed genetic disorders." *medRxiv*.

S Fazal, H Dashnow, M Dohrn, J Schatzman, L Hiatt, MC Danzi, I Xu, C Toro, D Adams, K Usdin, B Hayward, **SN Kobren**, SR Sunyaev, RC Spillmann, V Shashi, Undiagnosed Diseases Network, M Tekin, AR Quinlan, S Züchner. "A genotype-first approach for the diagnosis of repeat expansion disorders in the Undiagnosed Diseases Network cohort." *medRxiv*.

J Wen, S Zeng, C-L Bonzel, **SN Kobren**, J Du, Y Chai, H Wang, M Zhu, S Chen, F Leng, HG Zhang, KP Liao, K Cho, IS Kohane, M Zitnik, A Pereira, JS Liu, T Cai. "Phenotypic prediction of missense variants via deep contrastive learning." *medRxiv*.

In Preparation A Gutierrez, A Serret-Larmande, R Yin, S Makwana, T DeSain, N Chu, JP Stedman, D Pillion, JC Lyons, **SN Kobren**, D Traviglia, C Esteves, K LeBlanc, E Hughes, Undiagnosed Diseases Network, S Churchill, SR Sunyaev, AT McCray, IS Kohane, P Avillach. "A patient-centric information commons for a national undiagnosed diseases network."

2024 S Fazal, M Danzi, I Xu, **SN Kobren**, S Sunyaev, C Reuter, S Marwaha, MT Wheeler, E Dolzhenko, F Lucas, S Wuchty, M Tekin, S Zuchner, V Aguiar-Pulido. "RExPRT: a machine learning tool to predict pathogenicity of tandem repeat loci." *Genome Biol*, 25: 39.

S Chen, ALM Tan, C Perry, S Churchill, M Vella, J Mao, V Viswanadham, **SN Kobren**, IS Kohane. "Polygenic risk scores for autoimmune related diseases are significantly different and skewed in cancer exceptional responders." *npj Precision Oncology*.

2023 E Alsentzer, SG Finlayson, MM Li, **SN Kobren**†, IS Kohane†. "Simulation of undiagnosed patients with novel genetic conditions." *Nat Commun*, 14: 6403.

RJ Tinker, L Bastarache, KM Ezell, **SN Kobren**, C Esteves, JA Rosenfeld, EF Macnamara, R Hamid, JD Cogan, D Rinker, S Mukharjee, I Glass, K Dipple, JA Phillips III, Undiagnosed Diseases Network (2023). "The contribution of mosaicism to genetic diseases and de novo pathogenic variants." *American Journal of Medical Genetics Part A*, 191: 2482–2492.

2021 **SN Kobren**, D Baldridge, M Velinder, JB Krier, K LeBlanc, C Esteves, BN Pusey, S Züchner, E Blue, H Lee, A Huang, L Bastarache, A Bican, J Cogan, S Marwaha, A Alkelai, DR Murdock, P Liu, DJ Wegner, AJ Paul, Undiagnosed Diseases Network, SR Sunyaev, IS Kohane (2021). "Commonalities across computational workflows for uncovering explanatory variants in undiagnosed cases." *Genetics in Medicine*. 23: 1075–1085.

BK Beaulieu-Jones, C Darabos, D Kim, A Verma, **SN Kobren** (2021). "Innovative methodological approaches for data integration to derive patterns across diverse, large-scale biomedical datasets." *Pac Symp Biocomput*, 26: 256–260.

SN Kobren (2021). "How medical mysteries push back the frontier of genomics knowledge." *UDN PEER Newsletter*, Summer 2021.

2020 **SN Kobren**, B Chazelle, M Singh. (2020). "PertInInt: An integrative, analytical approach to rapidly uncover cancer driver genes with perturbed interactions and functionalities." *Cell Systems.* 11: 63–74.

SN Kobren, BK Beaulieu-Jones, C Darabos, D Kim, A Verma (2020). "Ongoing challenges and innovative approaches for recognizing patterns across large-scale, integrative biomedical datasets." *Pac Symp Biocomput*, 25: 286–294.

2019 **SN Kobren**, M Singh. (2019). "Systematic domain-based aggregation of protein structures highlights DNA-, RNA- and other ligand-binding positions." *Nucleic Acids Research*. 47: 582–593.

2015 **S Nadimpalli**, AV Persikov, M Singh. (2015). "Pervasive variation of transcription factor orthologs contributes to regulatory network divergence." *PLoS Genetics*. 11: e1005011.

2012 N Daniels, **S Nadimpalli**, L Cowen. (2012). "Formatt: Correcting protein structural alignments by incorporating sequence alignment." *BMC Bioinformatics*. 13: 259–266.

2011 **S Nadimpalli***, N Daniels*, L Cowen. (2011). "Formatt: Correcting protein structural alignments by sequence peeking." *Proceedings of the 2011 ACM Conference on Bioinformatics, Computational Biology, and Biomedicine*. 2: 315–319.

2009 J Rieffel, F Sauders, **S Nadimpalli**, H Zhou, S Hassoun, J Rife, B Trimmer. (2009). "Evolving soft robotic locomotion in PhysX." *Proceedings of the 2009 ACM Conference on Genetic and Evolutionary Computation*. 11: 2499–2504.

Research & Work Experience

07/2024 – Present **Associate Director of Rare Disease Analysis**, Dept. of Biomedical Informatics, Harvard Medical School • *Boston, MA*
Teaching and leading computational research analysis of genotypic and phenotypic patient data to derive insights into rare disease etiologies.

08/2018 – 06/2024 **Postdoctoral Research Fellow**, Harvard Medical School • *Boston, MA*
Developing computational approaches to identify recurrent disease etiologies from cohort-level genomic and clinical data to improve diagnoses of rare diseases under Drs. Isaac S. Kohane and Shamil R. Sunyaev

03/2012–06/2018 **Graduate Research Assistant**, Princeton University • *Princeton, NJ*
"Detecting and Analyzing Variation in Protein Interaction Interfaces"
Investigated the evolution, natural variation, and disease-related mutation of cellular networks through analysis of protein interaction interfaces under Prof. Mona Singh

Undergraduate Research Assistant, Tufts University • *Medford, MA*

01/2010–05/2011 *"Formatt: Correcting Protein Multiple Structural Alignments by Sequence Peeking"*
Improved the Matt protein structural aligner by incorporating a sequence alignment correction step and validating using objective measures under Prof. Lenore J. Cowen

09/2008–05/2009 *"Modeling the Neuro-Mechanical Control of a Soft-Tissue Organism"*
Evolved a solution representing the firing of neurons for muscle contraction using genetic algorithms to successfully elicit motion in a SoftBot under Prof. Soha Hassoun

06/2009–10/2009 **Summer Intern, Bioinformatics Dept**, Virginia Tech • *Blacksburg, VA*
"Cataloging Animal Retrocopies for Retrotransposon Model"
Developed a genome annotation independent algorithm to identify and classify retrocopies and chimeric genes given a transcriptome under Prof. Liqing Zhang

06/2008–09/2008 **Summer Intern, Biochemistry Dept**, University of California, Los Angeles • *Westwood, CA*
"Discovery of Novel Metabolic Types of Bacterial Microcompartments"
Implemented a pipeline to analyze oceanic metagenomic sequencing data for bacterial microcompartment (BMC) shell proteins to identify potentially novel metabolic BMC types under Prof. Todd Yeates

05/2007–08/2007 **Summer Intern, Preclinical Dept**, Shire Pharmaceuticals • *Cambridge & Lexington, MA*
Set up a secure database containing clinical and research reports for drugs in production or under testing.

06/2006–09/2006 **Summer Intern, Microbiology Dept**, Univ. of Massachusetts Medical School • *Worcester, MA*
"Identification of Genes Involved in Expansion of Chromosomal Repeat in E. Coli"
Developed a series of knockout bacterial strains and tested viability of bacteria and amplification of a specific chromosomal repeat under Prof. Anthony Poteete

Teaching & Advising Experience

08/2024–11/2024 **Course Co-Director**, Harvard Medical School • *Boston, MA*
(Upcoming) Design and lead a new 7-week elective course for Master's Program in Biomedical Informatics: *BMI711: Integrative Analyses for Rare Genetic Disease Diagnostics*

01/2022–05/2024 **Master's Capstone Project Adviser**, Harvard Medical School • *Boston, MA*
Primary adviser for year-long capstone projects for student in the School of Public Health and for student in the Department of Biomedical Informatics

03/14/2024 **Instructor for CME Accredited Course: Long-Read Sequencing: Diagnosing the Undiagnosed through Comprehensive Clinical Genetic Testing**, American College of Medical Genetics (ACMG) Annual Clinical Genetics Meeting • *Toronto, Canada*
Lecture introducing long-read sequencing technologies and the diagnostic potential over traditional sequencing technologies

05/2023, 05/2024 **Instructor for CME Accredited Course: Evaluation of Undiagnosed and Rare Conditions**, Harvard Medical School • *Boston, MA*
Gave 1-hour lecture on clinical genomic sequencing and research analytic pipelines

11/15/2018 **Guest Lecturer in Computational Biology**, Tufts University • *Medford, MA*
Gave 1-hour lecture on DNA motif finding algorithms to undergraduate and graduate students in *CS167*.

01/2016–09/2017 **Coadviser for First-Year Graduate Students**, Princeton University • *Princeton, NJ*
Fall 2016, Spring 2016 & Spring 2017 Coadvised three Ph.D. students in Computational Biology and Computer Science on separate projects related to protein domains, protein binding and cancer

06/2013–07/2013 **Coadviser for Undergraduate Students**, Princeton University • *Princeton, NJ*
Summer 2013 Coadvised three undergraduates on a summer project through the Summer Programming Experiences program to build a secure voting mobile application

09/2012–05/2013 **Assistant Instructor**, Princeton University • *Princeton, NJ*
Fall 2012 & Spring 2013 Taught precepts twice weekly for *COS126: General Computer Science*, held office hours (∼10 hours/week), graded assignments and proctored midterm and final exams.
*Top rated (4.5/5 on average) by students for all precepts taught over both semesters.

09/2009–05/2010 **Teaching Assistant**, Tufts University • *Medford, MA*
Fall 2009 & Spring 2010 Attended lectures, assisted in labs, held office hours (∼10-18 hours/week) and graded assignments for two courses *CS40: Machine Structure and Assembly Language Programming* and *CS80: Programming Languages*.

02/2009–06/2009 **High School Bioinformatics Instructor**, Somerville High School • *Somerville, MA*
Analyzed soil metagenomic data to discover new microbes and led four spring seminars for students about gene sequencing and alignment algorithms. Project funded via an HHMI ARRAYS grant.

Presentations

Invited Talks

02/20/2024 Sampson Lab for Kidney Genomics at Boston Children's Hospital, *Boston, MA*
02/08/2024 Medical and Population Genetics (MPG) Seminar at the Broad Institute, *Cambridge, MA*
08/10/2023 Undiagnosed Diseases Network (UDN) Grand Rounds, *Boston, MA*
04/25/2023 American Academy of Neurology Child Neurology Section Spotlight, *Boston, MA*
10/24/2022 St. Jude Pediatric Translational Neuroscience Initiative Ultrarare Disease Workshop, *Memphis, TN*
09/21/2021 Broad Institute's Statistical Genetics Seminar Series, *Virtual Meeting*
09/10/2021 International Session of Undiagnosed & Rare Diseases at China Rare Diseases Summit, *Virtual Meeting*
08/13/2021 Tufts University Bioinformatics & Computational Biology (BCB) Research Group, *Virtual Meeting*
08/04/2021 Stanford Center for Undiagnosed Diseases Seminar Series, *Virtual Meeting*
07/14/2021 Undiagnosed Diseases Network Annual In-Person Meeting, *Virtual Meeting*
07/13/2021 Boston Children's Hospital Seminar Series, *Virtual Meeting*
06/23/2021 Illumina Clinical Bioinformatics Workshop, *Virtual Conference*
12/05/2019 Quantitative and Computational Biology Colloquium at the Univ. of Southern California, *Los Angeles, CA*
11/14/2019 Undiagnosed Diseases Network (UDN) Grand Rounds, *Boston, MA*
11/11/2019 Janelia Women in Computational Biology Conference at HHMI-Janelia Research Campus, *Ashburn, VA*
10/25/2018 'Rising Stars in Computer Science' Symposium at Tufts University, *Medford, MA*
03/08/2018 Open Insights Seminar Series at Harvard Medical School, Dept. of Biomedical Informatics, *Boston, MA*
01/09/2018 Seminar at the Flatiron Institute, Center for Computational Biology (CCB), *New York City, NY*
11/03/2017 Lewis-Sigler Institute for Integrative Genomics 2017 Retreat, *Princeton, NJ*
09/26/2017 Data Science Seminar Series at the University of Massachusetts, *Amherst, MA*
09/13/2017 'Rising Stars in Data Science' Symposium at the University of Chicago, *Chicago, IL*
01/06/2017 Princeton Area Alumni Association First Friday Series, *Princeton, NJ*
01/06/2016 Icahn Institute Think & Drink Symposium, *Princeton, NJ*

Platform Presentations

09/16/2024 Nanopore Community Meeting, *Boston, MA*
09/29/2023 Critical Assessment of Genome Interpretation (CAGI)** Workshop, *Boston, MA*
10/20/2021 American Society of Human Genetics (ASHG) Annual Meeting, *Virtual Conference*
08/02/2017 NY Area Meeting in Quantitative Biology, *Cold Spring Harbor, NY*
04/02/2011 Tufts 13th Annual Undergraduate Research Symposium, *Medford, MA*

Posters

11/03/2023 American Society of Human Genetics (ASHG) Annual Meeting, *Washington, DC*
07/10/2023 Gordon Research Conference on Human Genetics & Genomics, *Waterville Valley, NH*
11/15/2015 RECOMB/ISCB Conference on Regulatory and Systems Genomics, *Philadelphia, PA*
03/20/2014 Systems Biology Meeting: Global Regulation of Gene Expression, *Cold Spring Harbor, NY*
04/09/2011 3rd Annual New England Undergraduate Computing Symposium, *Cambridge, MA*
10/09/2009 Midwest Women in Computing Conference, *Chicago, IL*

Panels

11/02/2023 "The Role of AI & Data Technology in Overcoming the Challenges of Rare Diseases" at the North America Rare Disease Summit, *New York City, NY*
07/15/2021 "Identifying and engaging the patient" at St. Jude Pediatric Translational Neuroscience Initiative Ultrarare Diseases Workshop, *Memphis, TN*
07/15/2021 "Doctoral Career Trajectories" at Harvard DBMI Summer Institute in Biomedical Informatics, *Boston, MA*
01/06/2021 "Pattern Recognition in Biomedical Data for Discovery" at PSB, *Virtual Conference*
07/16/2020 "Doctoral Career Trajectories" at Harvard DBMI Summer Institute in Biomedical Informatics, *Boston, MA*
01/06/2020 "Pattern Recognition in Biomedical Data: Challenges in Putting Big Data to Work" at PSB, *Waimea, HI*
11/11/2019 "Challenges and Opportunities with Emerging Technologies in Molecular and Cell Biology" at Janelia Women in Computational Biology Conference, *Ashburn, VA*
07/23/2019 "Doctoral Career Trajectories" at Harvard DBMI Summer Institute in Biomedical Informatics, *Boston, MA*
08/22/2017 "Women in Bioinformatics" at ACM Conf. on Bioinformatics, Comp Bio & Biomedicine, *Boston, MA*
10/02/2015 "Pursuing Graduate Studies in Computer Science" at Princeton Univ. Computer Sci. Dept., *Princeton, NJ*
03/28/2013 "Maximizing the Utility of Office Hours" at Princeton University McGraw Teaching Center, *Princeton, NJ*
04/09/2011 "Graduate School Application Process and Visit Experience" at 3rd Annual New England Undergraduate Computing Symposium, *Boston, MA*

Service

Conferences

07/2024 ISMB 2024 referee, *Montreal, Canada*
01/2019–01/2021 Pacific Symposium on Biocomputing 2020 & 2021, Session Co-Chair and referee, *Big Island of Hawaii*
12/2018 Machine Learning for Health (ML4H) at NeurIPS 2018, referee, *Montreal, Canada*
01/2017–08/2017 ACM Bioinformatics & Computational Biology 2017 Student Activity Co-Chair and referee, *Boston, MA*
05/2017, 05/2019 RECOMB 2017 & 2019 referee, *Hong Kong & Washington, DC*
07/2010, 07/2014 ISMB 2010 & 2014 Student Volunteer and referee, *Boston, MA*

Leadership

02/2024–Present Undiagnosed Diseases Network, Tool Building Coalition Working Group, *Chair*
10/2018–02/2024 Undiagnosed Diseases Network, Tool Building Coalition Working Group, *Primary Organizer*
05/2012–02/2015 Princeton Graduate Student Government, *Computer Science Representative, Events Board Member*
07/2012–05/2014 Princeton Jewish Graduate Students and Young Professionals, *President*

Mentoring and Outreach

08/2023–09/2023 CAGI Salon for trainees, *Organizer and participant*
01/2019–05/2020 Harvard Women in STEM (WiSTEM) Mentoring Program, *Undergraduate Student Mentor*
09/2011–05/2018 Princeton Graduate Women in Science and Engineering, *Mentoring Program Participant*
08/2011–Present Tufts Alumni Admissions Program, *Applicant Interviewer*
09/2008–05/2011 Tufts Admissions Office, *April Open House Host, Voices Host, Speaker at Engineering Open House*
09/2007–05/2011 Tufts Association for Computing Machinery, Women, *Mentor & Outreach Contact*