

Abhishek Sarkar

EMAIL aksarkar@alum.mit.edu

URL <https://aksarkar.github.io>

Research interests

Machine learning, single cell genomics, epigenomics, gene regulation, complex traits

Appointments held

2023– **Vesalius Therapeutics, Inc.**
Senior Scientist
Platform Group

2021–2023 **Vesalius Therapeutics, Inc.**
Scientist II
Computational Biology and Data Science Group

2017–2021 **University of Chicago**
Postdoctoral Scholar
Matthew Stephens (PI)
Department of Human Genetics

Education

2017 Ph.D. Computer Science, Massachusetts Institute of Technology

2013 M.S. Computer Science, Massachusetts Institute of Technology

2011 B.S. Computer Science with Highest Honors, University of North Carolina at Chapel Hill

Pre-prints

1. Peter Carbonetto, Kaixuan Luo, **Abhishek Sarkar**, Anthony Hung, Karl Tayeb, Sebastian Pott, Matthew Stephens. “Interpreting structure in sequence count data with differential expression analysis allowing for grades of membership.” *BioRxiv* (2023).
2. Fabio Morgante, Peter Carbonetto, Gao Wang, Yuxin Zou, **Abhishek Sarkar**, Matthew Stephens. “A flexible empirical Bayes approach to multivariate multiple regression, and its improved accuracy in predicting multi-tissue gene expression from genotypes.” *BioRxiv* (2022).
3. Peter Carbonetto, **Abhishek Sarkar**, Zihao Wang, and Matthew Stephens. “Non-negative matrix factorization algorithms greatly improve topic model fits.” *arXiv* (2021).

Peer-reviewed publications

1. **Abhishek Sarkar** and Matthew Stephens. “Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis.” *Nature Genetics* (2021).
2. Michelle C. Ward, Nicholas E. Banovich, **Abhishek Sarkar**, Matthew Stephens, Yoav Gilad. “Dynamic effects of genetic variation on gene expression revealed following hypoxic stress in cardiomyocytes.” *eLife* (2021).

3. Benjamin Jung Fair, Lauren Blake, **Abhishek Sarkar**, Bryan Pavlovic, Claudia Cuevas, and Yoav Gilad. "Gene expression variability in human and chimpanzee populations share common determinants." *eLife* (2020).
4. Gao Wang, **Abhishek Sarkar**, Peter Carbonetto, and Matthew Stephens. "A simple new approach to variable selection in regression, with application to genetic fine-mapping." *Journal of the Royal Statistical Society B* (2020).
5. **Abhishek Sarkar***, Po-Yuan Tung* et al. "Discovery and characterization of variance QTLs in human induced pluripotent stem cells." *PLoS Genetics* (2019). **Equal contribution*
6. Yaping Liu, **Abhishek Sarkar**, Pouya Kheradpour, Jason Ernst, and Manolis Kellis. "Evidence of reduced recombination rate in human regulatory domains." *Genome Biology* (2017).
7. Felix Day, ..., **Abhishek K. Sarkar**, et al. "Genomic analyses for age at menarche identify 389 independent signals and indicate BMI-independent effects of puberty timing on cancer susceptibility." *Nat Genet* (2017). *Author 9/215*
8. Roadmap Epigenomics Consortium et al. "Integrative analysis of 111 reference human epigenomes." *Nature*, 518(7539), 317–330 (2015). *Integrative analysis lead (equal contributor)*.

Technical reports

1. Yongjin Park, **Abhishek Sarkar**, Khoi Nguyen, and Manolis Kellis. "Causal Mediation Analysis Leveraging Multiple Types of Summary Statistics Data." *arXiv* (2019).
2. Yongjin Park*, **Abhishek Sarkar***, et al. "A Bayesian approach to mediation analysis predicts 206 causal target genes in Alzheimer's disease." *BioRxiv* (2017). **Equal contribution*
3. Yongjin Park*, **Abhishek K. Sarkar***, et al. "Multi-tissue polygenic models for transcriptome-wide association studies." *BioRxiv* (2017). **Equal contribution*
4. Kunal Bhutani*, **Abhishek Sarkar***, et al. "Modeling prediction error improves power of transcriptome-wide association studies." *BioRxiv* (2017). **Equal contribution*
5. **Abhishek K. Sarkar**, Lucas D. Ward, Manolis Kellis. "Functional enrichments of disease variants across thousands of independent loci in eight diseases." *BioRxiv* (2016).

Presentations

1. **Abhishek Sarkar** and Matthew Stephens. "A general statistical framework clarifies confusion in scRNA-seq analysis." Machine Learning in Computational Biology, Vancouver, Canada. 2019 (poster)
2. **Abhishek Sarkar** and Matthew Stephens. "Common pitfalls in the analysis of scRNA-seq data." Probabilistic Models in Genomics, Aussois, France. 2019 (poster)
3. **Abhishek Sarkar**, et al. "Variance QTL mapping in human induced pluripotent stem cells." Probabilistic Models in Genomics, Cold Spring Harbor, New York, USA. 2018 (poster)
4. **Abhishek Sarkar**, et al. "Variance QTL mapping in human induced pluripotent stem cells." 71st meeting of the American Society of Human Genetics, San Diego, California, USA. 2018 (poster)
5. Yongjin Park*, **Abhishek Sarkar***, Liang He* and Manolis Kellis. "Variational Bayes inference algorithm for causal multivariate mediation with linkage disequilibrium." Workshop on Machine Learning in Computational Biology, Thirty-first Annual Conference on Neural Information Processing Systems, Los Angeles, California, USA. 2017. (talk; **Equal contribution*)
6. Yongjin Park*, **Abhishek Sarkar***, Liang He* and Manolis Kellis. "Identification transcriptomic and epigenetic mediators in Alzheimer's disease: Bayesian inference and causal mediation analysis of regulatory programs in GWAS statistics." 69th meeting of the American Society of Human Genetics, Orlando, Florida, USA. 2017 (talk; **Equal contribution*)

7. Yongjin Park*, **Abhishek Sarkar***, Kunal Bhutani, Manolis Kellis. "Multi-tissue polygenic models for transcriptome-wide association studies" (talk). Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. 2017 (talk; **Equal contribution*)
8. **Abhishek Sarkar**, Yongjin Park, Manolis Kellis. "Dissecting the non-infinitesimal architecture of complex traits using group spike-and-slab priors". Workshop on Machine Learning in Computational Biology, Thirtieth Annual Conference on Neural Information Processing Systems, Barcelona, Spain. 2016 (talk)
9. **Abhishek Sarkar**, Yongjin Park, Manolis Kellis. "Dissecting the non-infinitesimal architecture of complex traits". 68th meeting of the American Society of Human Genetics, Vancouver, Canada. 2016 (poster)
10. **Abhishek Sarkar**, Luke Ward, Manolis Kellis. "Functional enrichments of disease variants across thousands of independent loci in eight diseases". Leena Peltonen School of Human Genomics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. 2016 (talk)
11. Yongjin Park, **Abhishek Sarkar**, Nick Mancuso, Alexander Gusev, Bogdan Pasaniuc, Manolis Kellis. "Computational discovery of epigenetic mediators in Alzheimer's disease from imputed methyome-wide association statistics". The Biology of Genomes, Cold Spring Harbor, NY, USA. 2016 (poster)
12. Kunal Bhutani*, **Abhishek Sarkar***, Yongjin Park, Manolis Kellis, Nicholas Schork. "Propagating uncertainty of predicted expression in transcriptome-wide association studies". The Biology of Genomes, Cold Spring Harbor, NY, USA. 2016 (poster; **Equal contribution*)
13. **Abhishek Sarkar**, Lucas D. Ward, Manolis Kellis. "Regulatory annotations implicate thousands of independent loci". 67th meeting of the American Society of Human Genetics, Baltimore, MD, USA. 2015 (poster)
14. **Abhishek K. Sarkar**, Lucas D. Ward, Manolis Kellis. "Genome-wide enrichments for regulatory regions across thousands of unlinked disease-associated variants". 65th meeting of the American Society of Human Genetics, Boston, MA, USA. 2013 (poster)
15. Vineeta Agarwala, **Abhishek Sarkar**, Kyle Gaulton. "Using the Epigenome Roadmap data to analyze genetic studies of Type 2 Diabetes". 65th meeting of the American Society of Human Genetics, Boston, MA, USA. 2013 (talk)
16. **Abhishek K. Sarkar**, Lucas D. Ward, Manolis Kellis. "Systematically interpreting GWAS using regions from Roadmap". Epigenomics: A Roadmap to the Living Genome, Boston, MA, USA. 2013 (poster)
17. Lucas Ward, **Abhishek Sarkar**, Manolis Kellis. "Global contributions of regulatory elements to disease risk and evolutionary fitness in the human population". 5th annual RECOMB Conference on Regulatory and Systems Genomics, San Francisco, CA, USA. 2012 (poster)
18. **Abhishek Sarkar**. "Functional GWAS enrichments across tens of thousands of enhancer elements". Cell Circuits and Epigenomics Seminar Series, Broad Institute, Cambridge, MA, USA. 2011 (talk)

Teaching experience

2014 Teaching assistant, "Advanced Computational Biology: Genomes, Networks, Evolution" (fall)

Service

PC Machine Learning in Computational Biology (virtual). 2020

PC Machine Learning in Computational Biology, Vancouver, BC, Canada. 2019

REVIEWER Nature Genetics, Nature Communications, PLoS Genetics, Genome Biology, BMC Genomics

Honors

2016 Accepted to Leena Peltonen School of Human Genomics

2011 Awarded NSF Graduate Research Fellowship

2011 Inducted into Phi Beta Kappa honors fraternity

References

Matthew Stephens, University of Chicago (mstephens@uchicago.edu)

Yoav Gilad, University of Chicago (y_gilad@uchicago.edu)

Manolis Kellis, MIT (manoli@mit.edu)

Alkes Price, Harvard School of Public Health (aprice@hsph.harvard.edu)