Anika Gupta

408.858.0485 | anikagupta@g.harvard.edu | linkedin.com/in/anika18 | gupta-anika.github.io

EDUCATION

Harvard University 2018 – 2023 (Expected)
Doctor of Philosophy, Bioinformatics and Integrative Genomics Cambridge, MA

Massachusetts Institute of Technology June 2018
Bachelor of Science, Computer Science and Molecular Biology Cambridge, MA

Relevant Coursework: Machine Learning, Human Pathology, Deep Learning for Biomedical Data, Computational and Functional Genomics, Design and Analysis of Algorithms, Genetics, Concepts in Genome Analysis, Healthcare Ventures, Strategic Decision Making in the Life Sciences; **Teaching Assistant:** Applied Machine Learning, Deep Learning

RESEARCH EXPERIENCE

Broad Institute of Harvard and MIT, *PI: Eric Lander* Jan 2019 – Present
Genomics Data Science Researcher Cambridge, MA
Inferring causal gene regulatory interactions from observational data to systematically identify cellular pathways. Predicting from high-throughput single cell transcriptional and temporal labeling; goal is to learn gene programs without perturbing cells.

Stanford University, *PI: Dennis Wall* June 2017 – June 2018
Autism Data Science Researcher Palo Alto, CA
Implemented coalitional game theory and unsupervised machine learning techniques on 4,610 autism spectrum disorder genomes to map the genotype-phenotype bridge and enable a causative, molecular understanding for more precise treatments.

Broad Institute of Harvard and MIT, *PIs: Aviv Regev and Kasper Lage* Sep 2015 – June 2018
Statistical Genomics Researcher Cambridge, MA
Detected cancer vulnerabilities through a robust statistical framework that capitalizes on the identification of protein interaction networks under purifying selection in 4,700 cancer genomes.

Foundation Medicine May 2016 – May 2017
Cancer Genomics Researcher Cambridge, MA
Characterized the therapeutically targetable gene alteration landscapes of lung cancers through analysis of >20,000 patient tumors to help make personalized medicine a reality.

Merck May – Aug 2015
Bioinformatics Researcher Palo Alto, CA
Analyzed gene expression and signaling pathway data for 416 genes to develop immunotherapy combination signature-predicting algorithms for synergistic therapeutic delivery against melanoma.

PROJECTS

The Data Pulse, Podcast Host Jul 2020 - Present
I host a podcast at the intersection of data science & biomedical innovation, ranging from molecular to clinical and population scales, and spanning academia and industry. Season 1 has >5k downloads and >500 repeat listeners: linktr.ee/thedatapulse.

5AM Ventures, Fellow Jul – Dec 2019
One of two fellows, guiding strategic decisions in the venture creation arm. Designed the bioinformatics platform for an immuno-oncology company, outlined the translational biomarker strategy for a neurodegenerative disease company, and conducted scientific diligence in the regenerative medicine and RNA splicing domains for two potential new company builds.

RA Capital, Business of Biotechnology Advanced Concepts Short Course June 2019
Chosen from >100 applicants to learn and debate principles of drug pricing, company formation, and biotech investment landscapes. Developed frameworks for thinking about value creation and financing strategies.

Undiagnosed Diseases Network, *Bioinformatics Case Analysis Team* Oct - Dec 2018
Standardized and automated (for variant classification) a proof of concept genomic analysis pipeline for the Brigham Genomic Medicine team's efforts to uncover the genetic causes of four previously intractable, rare patient disease cases.

Flagship Pioneering, *VentureLabs Innovation Intern* Jan – May 2017
Built quantitative models to enable key strategic decisions that could maximize value creation for a mitochondrial-based Series A biotech company being built within this venture firm.

LEADERSHIP

Models, Inference, and Algorithms, *Steering Committee Member* Aug 2020 - Present
Guide the dialogue on machine learning within the Broad Institute & broader community as a team member, inviting researchers to share their latest work, with a focus on the theoretical underpinnings and a discussion-based forum to foster collaborations.

MIT Biotechnology Group, *Founder and Co-President* June 2015 – June 2017
Established the undergraduate chapter of MIT's first (now >1,000-member) biotech initiative as a liaison between students and the biotech industry. Organize pitch competitions, speaker panels, interactive workshops, mentorship programs, due diligence groups, and interactive symposiums to foster student entrepreneurship within the biotech industry.

Residential Associate Adviser, *to Professor Maria Zuber (MIT VP of Research)* Sep 2015 – June 2018
Mentored MIT freshmen as they navigate through their first year academic and personal pursuits at the institute.

PUBLICATIONS

Sun M*, **Gupta A***, Varma M, Paskov K, Jung J, Stockham N, Wall D (2019). Coalitional game theory facilitates identification of non-coding variants associated with autism. *Journal of Biomedical Informatics Insights*.

Gupta A*, Sun M*, Paskov K, Stockham N, Jung J, Wall D (2018). Coalitional game theory as a promising approach to identify candidate autism genes. *Pac Symp Biocomputing 2018*. pp. 436-447 (https://doi.org/10.1142/9789813235533_0040).

Gupta A*, Horn H*, Razaz P, Kim A, Lawrence M, Getz G, Lage K (2017). Detecting cancer vulnerabilities through gene networks under purifying selection in 4,700 cancer genomes (https://www.biorxiv.org/content/early/2017/11/21/222687.1).

Gupta A, Connelly C, Frampton G, Chmielecki J, Ali S, Suh J, Schrock A, Ross J, Stephens P, Miller V (2017). The druggable mutation landscape of lung adenocarcinoma. *Journal of Thoracic Oncology*. Volume 12, Issue 1, S977.

CONFERENCE PRESENTATIONS

Gupta A*, Siraj L*, Jones TR, Bloemendal A, Subramanian V, Lander ES (2019). Systematically learning cellular programs from single-cell transcriptional and chromatin accessibility data. *33rd Conference on Neural Information Processing Systems: Learning Meaningful Representations for Life workshop.* Vancouver, CA.

Gupta A*, Siraj L*, Jones TR, Bloemendal A, Subramanian V, Lander ES (2019). Systematically learning biological pathway signatures from single cell gene expression data. *3rd Annual Women in Data Science Conference.* Cambridge, MA.

Gupta A, Horn H, Razaz P, Kim A, Lawrence M, Getz G, Lage K (2018). Detecting cancer vulnerabilities through gene networks under purifying selection. *American Society of Human Genetics Annual Meeting.* San Diego, CA.

Gupta A*, Sun M*, Paskov K, Stockham N, Jung J, Wall D (2018). Identifying candidate autism genes via coalitional game theory. *American Medical Informatics Association Informatics Summit*. San Francisco, CA.

Gupta A*, Sun M*, Paskov K, Stockham N, Jung J, Wall D (2018). Coalitional game theory as a promising approach to identify candidate autism genes. *23rd Annual Pacific Symposium on Biocomputing*. Big Island, HI.

Frampton G, Hartmaier R, Sokol E, **Gupta A**, Greenbow J, Roels S, Gay L, Stephens P (2018). Novel CDH1 mutations in breast invasive lobular carcinoma. *American Association for Cancer Research Annual Meeting*. Chicago, IL.

Gupta A, Chalmers Z, Connelly C, Frampton G, Chmielecki J, Ali S, Suh J, Schrock A, Ross J, Stephens P, Miller V (2016). The druggable mutation landscape of lung cancer. *IASLC 17th World Conference on Lung Cancer.* Vienna, Austria.

Gupta A, Horn H, Lawrence M, Getz G, Lage K (2015). Identifying and targeting gene networks under purifying selection. *11th Annual Broad Institute Symposium***.** Cambridge, MA.

AWARDS

Google Ventures, Travel Award for NeurIPS 2019	2019
National Human Genome Research Institute, T32 Grant	2018
Pacific Symposium on Biocomputing, National Library of Medicine/National Institutes of Health Travel Award	2017
Grace Hopper Conference for Women in Computing, Microsoft Scholarship Recipient	2016
Intel Science Talent Search Competition, National Semifinalist	2014
Siemens Competition in Math, Science, and Technology, National Semifinalist	2012

SKILLS

Computer Science Languages: Python (keras, scikit-learn, numpy, pandas, scipy), R, Java, HTML, Unix/Linux, LaTex
Databases Familiar With Analyzing: gNOMAD (population-level mutation frequencies), The Cancer Genome Atlas, Human Protein Atlas, Project Achilles (cancer vulnerabilities), Genome Aggregation Consortium, Broad Mutation Signatures Database, ClinicalTrials.gov, genome-wide expression (bulk and single cell RNA-seq), transcription factor binding (ChIP-seq), chromatin accessibility (ATAC-seq), and sequencing data (UK Biobank and proprietary; .bam and .bed formats)
Languages: English (native), Hindi (fluent)