

# ANIKA GUPTA

anikagupta@g.harvard.edu | [in](#): anika18 | [gupta-anika.github.io](#) | [twitter](#): @anikagupta18

---

I'm a PhD student working at the intersection of data science and high-throughput biology through the lens of single cell genomics. I'm passionate about data-driven biomedical innovation through venture creation.

## EDUCATION

### Harvard University

Doctor of Philosophy,  
Bioinformatics and  
Integrative Genomics

2018 - present

### Massachusetts Institute of Technology

Bachelor of Science,  
Computer Science and  
Molecular Biology

2014 - 2018

---

## Programming Languages

Python (keras, scikit-learn,  
pandas, numpy, scipy), R,  
HTML, Unix/Linux, LaTeX

## Teaching Experience

Applied Machine Learning

Deep Learning for  
Biomedical Data

## Select Courses

### Computational:

- Machine Learning
- Design and Analysis of Algorithms
- Computational Genomics

### Biomedical:

- Genetics
- Principles & Applications of Genetic Engineering
- Human Pathology (HMS)
- Concepts in Genome Analysis

### Biotech Entrepreneurship:

- Strategic Decision Making in the Life Sciences (Sloan)
- Business of Biotech (RA Capital)

## RESEARCH EXPERIENCE

### Broad Institute of Harvard and MIT, PI: Eric Lander

Jan 2019 - Present

Inferring causal gene regulatory interactions from high-throughput, time stamped single cell transcriptomic data. Goal is to systematically identify cellular pathways without perturbing cells.

### Stanford University, PI: Dennis Wall

June 2017 - June 2018

Implemented coalitional game theory and unsupervised machine learning on the genomes of 4,610 individuals with autism spectrum disorder to map the combinatorial effect of genotype on ASD phenotypes.

### Broad Institute, PIs: Aviv Regev, Kasper Lage

Sep 2015 - June 2018

Detected cancer vulnerabilities through a statistical framework that targets protein interaction networks under purifying selection—those that are rarely mutated due to their key role in tumor proliferation—in 4,700 cancer genomes

### Foundation Medicine, Cancer Genomics Team

May 2016 - May 2017

Characterized the therapeutically targetable gene alteration landscapes—driver mutations, tumor mutational burden, and checkpoint inhibitors—of lung cancers, through analysis of >20,000 patient tumors towards precision medicine.

### Merck, Bioinformatics Team

May - Aug 2015

Analyzed gene expression and signaling pathway data for 416 genes to predict immunotherapy combination gene signatures and enable 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 >10k downloads and >500 repeat listeners.

### 5AM Ventures, Fellow

Jul - Dec 2019

As one of two fellows, I guided 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 company builds.

### Undiagnosed Diseases Network, Case Analysis Team

Oct - Dec 2018

Standardized and automated a genomic analysis (variant classification) 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 Intern

Jan - May 2017

Built quantitative models to enable key strategic decisions that could maximize value creation for a mitochondrial therapy-based Series A biotech company build.

## LEADERSHIP

### **Models, Inference, and Algorithms, Steering Committee**

**Aug 2020 - Present**

Guide the dialogue on machine learning within the Broad Institute & broader community as a committee member, inviting researchers to share their latest work, with a focus on the theoretical underpinnings and a discussion-based forum to foster collaborations at the interface of machine learning and biomedicine.

### **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. Organized 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

Horn H, Fagre C, **Gupta A**, Tsafou K, Fornelos N, Neal JT, Lage K (2020). Using protein interaction networks to identify cancer dependencies from tumor genome data. *bioRxiv*.

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

Frampton GM, Hartmaier R, Sokol E, **Gupta A**, Greenbowe J, Roels S, Gay L, Stephens PJ (2018). Novel CDH1 mutations in breast invasive lobular carcinoma. *Cancer Research*. Volume 79, pp. 2363.

**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*. pp. 436-447.

**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. *bioRxiv*

## 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. *33<sup>rd</sup> 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. *3<sup>rd</sup> 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. *23<sup>rd</sup> 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.

Frampton G, **Gupta A**, Connelly C, Fabrizio D, et al (2017). Hybrid-capture based comprehensive genomic profiling of lung adenocarcinoma identifies patients who may benefit from targeted therapies as well as immunotherapies using tumor mutational Burden (TMB), a new predictive biomarker of response to immune checkpoint inhibitors. *Laboratory Investigation*.

**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 17<sup>th</sup> 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. *11<sup>th</sup> Annual Broad Institute Symposium*. Cambridge, MA.