ANUJ GUPTA

ag.anuj1991@gmail.com |+1-404-579-4985 | https://www.linkedin.com/in/guptaanuj1991/

EDUCATION

Georgia Institute of Technology (Ph.D. and M.S) **Atlanta, USA**
Bioinformatics and Machine Learning 2014- Present
Computational Biology Faculty Research Award (full-tuition scholarship)

Indian Institute of Technology (IIT) Kanpur (B.Tech) **Kanpur, India**
Biological Sciences and Bioengineering 2009-2013
SSPC Award (full-tuition scholarship); Hostel President (+500 students); Squash Team Captain

Skills*:* Python, R, Matlab, C/C++, JS, D3, SQL, Shell, TensorFlow, sckit-learn, Pandas, Keras

PUBLICATIONS

- **Gupta A,** Styczynski MP, the MaHPIC consortium, Voit EO. (Manuscript submitted). *Dramatic Transcriptomic Differences in Macaca mulatta and M. fascicularis with Plasmodium knowlesi Infections*. Scientific Reports (Nature).
- **Gupta A,** I. King Jordan, Lavanya Rishishwar. (2017). *stringMLST: a fast k-mer based tool for multilocus sequence typing*. Bioinformatics (Oxford University Press).
- **Gupta A**, Tang Y, Garimalla S, MaHPIC consortium, Galinski MR, Styczynski MP, Fonseca LL, Voit EO. (2018). *Metabolic Modeling Helps Interpret Transcriptomic changes during a Complex Disease*. Molecular Basis of Disease(Elsevier).
- Downs DM, Bazurto JV, **Gupta A**, Fonseca LL, Voit EO. (2018). *The Three-Legged Stool of Understanding Metabolism: Integrating Metabolomics with Biochemical Genetics and Computational Modeling*. AIMS Microbiology(AIMS press).

PROFESSIONAL EXPERIENCE

- Graduate Research Assistant at Voit Lab, Georgia Tech and Emory School of Medicine, Atlanta (*May'16-Present)*
- Bioinformatics Intern, La-Hoffman Roche, Basel, Switzerland *(July'18-Jan'19)*
- Graduate Research Assistant, Applied Bioinformatics Laboratory(ABiL) / Georgia Tech, Atlanta *(Aug'14-Apr'16)*
- Bioinformatics Intern, Philips Healthcare Research, NY *(May'15-July'15)*
- Product Engineer & Software Developer (Java), eGain Communications Corporation, India *(May'13-July'14)*

SELECT PROJECTS

Multi-omics comparison between susceptibility and resilience of macaque hosts against *Plasmodium* pathogen
Mentor: E.O.Voit (Georgia Tech) Transcriptomics | Metabolomics | Dynamic Modeling | Machine Learning

- Developed quantitative methodologies linking evolution with immune response to identify key drug targets
- Designed heuristic algorithm to estimate gene-specific evolutionary distances, identifying immune signaling genes that mutated to alter pathogen detection and thereby delayed immune response
- Created a statistical test based on differential enrichment of biological functions to identify key signatures in fundamental biological processes responsible for small but divergent immune response
- Built prediction methodology from gene expression to metabolic changes to evaluate effects of mutant genes and points of intervention by altering enzymatic activity in kinetic metabolic models (power-law/Michaelis-Menten)

Microbiome characterization and modelling using metagenomics data for IBD treatment
La-Hoffman Roche Metagenomics | Metatranscriptomics | Dynamic Modeling | Machine Learning

- Built novel methodology to predict changes in gut microbiome from nutritional media *(Most Innovative Project Award)*
- Designed a feature specific parametric mixture model to define distance matrix, characterize and cluster gut microbiome to identify essential and variable species that affect the progression of diseases
- Created constraint-based metabolic models to form a symbiotic system and identify the role of variable species
- Developed correlational models for these constraints to nutrition using Monte Carlo simulations

Development of bacterial typing tool for fast and accurate predictions
Applied Bioinformatics Laboratory Algorithm development | Computational Genomics | Machine Learning

- Designed and implemented a novel bacterial classification algorithm for CDC *(20K downloads, 50+ citations)*
- Developed maximum parsimony inspired algorithm followed by a Bayesian inference model to achieve 100% real outbreak prediction accuracy while achieving 20% less runtime and disk usage than gold standard methods
- Proposed evolution focused, tree-inspired data structure for efficient performance and storage of database

Development of contact tracing tool for nosocomial infection at Westchester Medical Center
Philips Healthcare Research Algorithm development | Computational Genomics | Machine Learning

- Developed tool to track infection spread in a hospital environment using maximum likelihood model
- Identified signature mutations and developed linear regression models for contact tracing
- Predicted the course of existing MRSA spread (with p-value significance of 0.001), hence isolating the root cause