Mohit Thakur

46 Progress Pl • Voorhees, NJ 08043
mohitmthakur@yahoo.com • (856) 304-2822 • www.iMohit.com

EDUCATION

M.S. Bioinformatics	**Georgia Tech. University, Atlanta** *(Completion: Dec'2018)*	**2017-Present**
• Switched from pursuing M.D. to M.S. Bioinformatics		GPA: 3.8/4.0
M.D.	**Rutgers NJ Medical School, Newark**	**2013-2016**
• Completed both preclinical years and first clinical year		
B.S. (Biology) & M.D. Joint Program	**The College of New Jersey, Ewing**	**2010-2013**
• Member of the Phi Beta Kappa Honor Society		GPA: 3.9/4.0

SKILLS

Technical Skills: Python, PERL, Bash scripting, SQL, R, PHP, Java, C, HTML, CSS, Linux, MacOS

Bioinformatics: Genomic file formats (FASTA, PED, BED, VCF, FASTQ, BCL, etc.), Databases (NCBI, 1000 Genomes, PDB)**,** Software tools (BLAST, BioPerl, JMOL, Bowtie2, SAMtools, vcftools, FASTQC, Cytoscape).

Statistical analysis: Continuous, discrete, and bivariate distributions, probability, maximum likelihood estimation

Data analysis libraries*: NumPy, Pandas, Matplotlib, SciPy

Database and configuration management*: Git, Subversion, MySQL

Laboratory Skills: Biological and Microbiological techniques such as western blots, maintaining bacterial cultures & cytometry.

PROJECTS

Local Ancestry Pipeline: Developed a processing pipeline to validate STRUCTUREpainter, a novel local ancestry inference algorithm. STRUCTUREpainter can identify source populations with the unique advantage that, unlike currently available tools, multiple reference populations aren't required. **Python, R, Bash, Git, C, VCFtools, ADMIXTURE, Plink**

NGS variant calling: Given BCL files as NGS output data, convert to FASTQ and use Bowtie2 to map fragments onto a reference. After mapping, perform variant calling using SAMtools to locate variants in pathogenic regions. **Perl, FASTQC, Bowtie2, SAMtools**

Global and Local Sequence Alignment: Construct a scoring matrix in Python and backtrace all possible alignment scores between two sequences to identify the optimal global and local sequence alignments, based on Needleman-Wunsch and Smith-Waterman algorithms respectively. **Python, Bash**

Properties of PPI network: Given a list of protein-protein interaction (PPI) pairs, create a PPI network that represents all interactions in a single graph. Generate a histogram of the degree distributions and clustering coefficient of the network to identify the dominant interactions. Visualize with Cytoscape. **Perl, Cytoscape.**

Database Modeling: Team project developing and implementing a database model as a property management website. Used a PHP front end and MySQL back end. **PHP, MySQL, SQL, Git**

Protein Interfaces: Given a chain of amino acids in a 3d space, determine all interfaces. Each interface is defined as a set of contiguous points separated by a threshold distance. Visualize with JMOL. **Perl, JMOL**

ADDITIONAL PROJECTS

NAACCR Cancer Informatics Hackathon	**University of Pittsburg Medical Center**	**June'18**
• Project involved natural language processing to identify cancer type from registries, using Python and Java (3rd place winner)		
Rocky Mountain Genomics HackCon	**Biofrontiers Institute, CO**	**June'18**
• Simulating free diffusion by way of experimental data of fluorescence accumulation after DNA damage, using Python		

WORK EXPERIENCE

Graduate Research Student	**LaChance Lab., GA**	**Sep'17-Present**
• Worked on Local Ancestry Pipeline (see "Projects" above)		
Licensed Pharmacy Technician	**Walmart Pharmacy, NJ**	**Oct'16-Jun'17**
	CVS Pharmacy, NJ	**Jun'12-Jul'13**
• Handled all pharmacy tasks such as pharmacist support, customer and insurance company interactions		
Resident Advisor	**TCNJ, NJ**	**Aug'11-May'13**
Tutor	**Kumon Math and Reading Center, NJ**	**Sep'06-Aug'10**

Online Certification Courses

Fundamentals of Computing Specialization (7 Cert. Courses)	**Rice University @ Coursera.org**	**May'2017**
Introduction to R for Data Science	**Microsoft @ edX.org**	**Nov'2016**
HTML & CSS for Beginners	**udemy.com**	**Sep'2016**

Citizenship: U.S. Citizen