Karan Ajay Kapuria

510-309-0140 789 Hammond Drive, Atlanta GA kapuriakaran@gmail.com

PROFILE

Bioinformatics professional with 3 years of experience in next-generation sequencing (NGS) pipeline development, validation, and implementation.

EDUCATION

Georgia Institute of Technology, Atlanta, GA 2016 – 2017
Master of Science, Bioinformatics

Guru Nanak Khalsa College, Mumbai, India 2014 – 2016
Master of Science, Bioinformatics

Mithibai College, Mumbai, India 2011 –2014
Bachelor of Science, Chemistry

PROFESSIONAL EXPERIENCE

Laboratory Corporation of America *(MNG Labs became part of LabCorp)*
Atlanta, GA
Biostatistician March 2019 – Present

- Evaluate analytical performance, maintain and develop new modules for all MNG's NGS analysis workflows (WES/WGS/Targeted Panels/Mitochondrial DNA)
- Provide bioinformatics support to reporting specialists with case specific queries
- Work closely with Director of R&D and Director of Reporting to update and improve bioinformatics reporting functionalities
- Automate MNG's flagship NGS pipeline with the goal of combining multiple-cloud technologies to reduce workflow debts and yield a flexible platform in link with other databases.
- Work directly with the LabCorp's NGS team to present MNG's bioinformatics capabilities to LabCorp, addressing their needs.

MNG Laboratories **Atlanta, GA**
Bioinformatics Developer March 2018 – March 2019

- Provided first-in-class bioinformatics services for a CAP/CLIA accredited clinical NGS lab specializing in neurogenetics utilizing Illumina technologies (MiSeq, HiSeq 1500/2500, NovaSeq 6000, BaseSpace Sequencing Hub)
- Lead bioinformatics development and provide support on bringing new NGS products (WGS/Express Exome) to market, working closely with Director of Bioinformatics, CMO, and CIO
- Develop / validate / and lead implementation of NGS pipelines on Amazon Web Services (AWS), including:
    - Whole genome sequencing: constitutional whole genome variant, copy number, expansion repeat, uniparental disomy, Mitochondrial heteroplasmy / deletion / depletion analysis
    - Whole transcriptome sequencing: differential expression, transcript, junction, variant analysis
    - Whole exome sequencing: constitutional exome variant, copy number, uniparental disomy analysis
    - Custom genetics panel: constitutional variant, single exon resolution copy number analysis
    - Mitochondrial analysis: heteroplasmy (1%) variant analysis, break point resolution deletion analysis with quantification

Georgia Institute of Technology (Dr. Brown's Lab) **Atlanta, GA**
Bioinformatics Data Analyst April 2017 – December 2017
Graduate Research Assistant August 2016 – April 2017

- Developed and managed a metagenomics OTU classification pipeline using QIIME and DADA2 for 16s rRNA sequencing data.
- Performed PCA, machine learning and regression to evaluate co-occurrence patterns of in microbiome abundance data.
- Creating visualizations using R package, ggplot2 and phyloseq. fitting the dataset to generalized lotka-volterra model.
- Performed network analysis in lung microbiome using SparCC and Pearson co-relations.
- Performed machine learning on cross sectional abundance matrix obtained from lung microbiome in Cystic Fibrosis (CF) affected cohort.

Guru Nanak Institute of Research and Development **Mumbai, India**
Research Assistant June 2014 - March 2016

- Designed anti-psoriatic flavonoid analogues using Chem Sketch and docked against protein target using GOLD (Genetic Optimization for Ligand Docking) software.
- Performed and lead the QSAR study on best binding ligands obtained from GOLD using VLifeMDS software for predicting models with best biological activities.

Karan Ajay Kapuria

510-309-0140 789 Hammond Drive, Atlanta GA kapuriakaran@gmail.com

SKILLS

Technology

- Software Languages: Bash, R Statistical Programing Language, Python, SQL, C++
- Software Tools: Illumina Base Space Sequencing Hub (BSSH), Genome Analysis Toolkit (GATK), Dragen, NextGENe, Burrows-Wheeler Aligner (BWA), Integrative Genomics Viewer (IGV), VarScan, Picard, bedtools, vcftools, bcftools, samtools, Annovar, Sun Grid Engine, Variant Effect Predictor (VEP)
- Amazon Web Services (AWS) : EC2, S3, ECR, AWS Batch, ECS, SNS, IAM, CLI, boto3 (python for AWS)
- Online Resources / Databases: GTEx, Online Mendelian Inheritance in Man (OMIM), ClinVar, Catalogue Of Somatic Mutations In Cancer (COSMIC), Ensembl, 1000 Genomes, genome Aggregation Database (gnomAD), Exome Aggregation Consortium (ExAC), dbSNP, dbVAR, MITOMAP, STRING (Protein-Protein Interaction Networks), International Mouse Phenotyping Consortium (IMPC), SIFT, PolyPhen, UCSC Genome Browser
- Management Tools and Version Control: Jira, Bitbucket, Github

SCHOLARSHIPS and AWARDS

- Computational Biology Faculty Research Awards – Georgia Institute of Technology
  For research proposal on **'Network Analysis of Microbial Communities in Cystic Fibrosis Lung'**. *Fall 2017*
- Computational Biology Faculty Research Awards – Georgia Institute of Technology
  For research proposal on **'Inferring Interactions in the Cystic Fibrosis Microbiome with Dynamic Modeling and Machine Learning'**. *Spring 2017*

PUBLICATIONS and POSTERS

Co-author. **Determination of the phylogenetic origins of the Árpád Dynasty based on Y chromosome sequencing of Béla the Third**. https://doi.org/10.1038/s41431-020-0683-z, *Nature*, 07 July 2020

Co-presented poster. **The Role of Whole Genome Sequencing in Clinical Diagnosis**. 2018. The American Society of Human Genetics.

Co-presented poster. **Complementary Test Methods Needed to Improve Clinical Sensitivity of Genetic Testing**. 2018. The American Society of Human Genetics.

Co-presented poster. **Applying Machine Learning Techniques on Cystic Fibrosis Lung Microbiome**. 2017. Metagenomics Symposium, Georgia Institute of Technology.