KHUSHBU PATEL

Bioinformatics Scientist
The children's Hospital of Philadelphia

Hamilton, NJ 08619 • (646) 508-6872 • khushbu_p@hotmail.com

SUMMARY

A bioinformatics scientist, in the field over three years with comprehensive background in computational cancer genomics and transcriptomic data. Competent working in Linux and HPC computing environments. Skilled in developing and implementing computational methods and workflow to analyze high throughput genomic data. Strong background in cancer genomics, cancer biology, genetics and molecular biology.

GitHub: github.com/kpatel427
Personal Website: kpatel427.github.io
LinkedIn: www.linkedin.com/in/khushbu-patel-b1a196b5/

SKILLS

o **Programming Languages:** R, Bash, Python, Perl
o **Tools & Software:** GATK, bcftools, samtools, STAR, StringTie, Kallisto, MACS2, Homer, RShiny, BBMap, MEME suite, VEP, Snakemake, Docker
o **Methods & concepts:** Gene set enrichment analysis (GSEA), Gene Ontology enrichment analysis (ToppGene, Enrichr), Differential Gene expression analysis (DESeq2, EdgeR), Dimensionality reduction methods (t-SNE, PCA, hierarchical clustering), Sequence Assembly & Alignment (BWA/STAR/BowTie), Gene abundance estimations (HTSeq-counts), Survival Analysis, Calling and Annotating mutations and genetic aberrations (VarScan2, CNVkit, ClinVar, COSMIC), Statistical Analysis (Pearson/Spearman correlation, Student t-test, Wilcox rank sum test, ANOVA, Kruskal-Wallis)

PROFESSIONAL EXPERIENCE

The Children's Hospital of Philadelphia, *Philadelphia PA • Bioinformatics Scientist I*   *07/2018 - Present*
Key Qualifications and Responsibilities:

o Experience analyzing multi-omics (WGS, WES, RNA-Seq, ChIP-Seq, Microarray, DNA methylation) data and perform large scale genomic analysis including quality control, data preprocessing, running pipelines, downstream analysis and result visualization.
o Proficient coding skills in R and bash to wrangle, analyze and visualize data.
o Sound knowledge of statistical concepts and implementation of statistical tests for hypothesis driven research.
o Employing various unsupervised machine learning methods to cluster and reduce dimensionality of a high dimensional dataset.
o Parallelization of jobs over HPC to run routine pipelines for large number of samples in addition to troubleshooting and development of pipelines.
o Containerizing pipelines (Docker) for efficient reproducibility of pipelines.
o Maintaining best practices and using version controlling system (git) to maintain code.
o Developed multiple web applications (HTML, CSS, RShiny) for efficient data visualization, analysis and sharing for internal use.
o Ability to effectively communicate results in the form of reports/spreadsheets/visualizations.
o Keeping abreast of latest algorithms and data generation platforms.

o   Collaborations with diverse research groups to provide computational support as well as contributions to grants and manuscripts.

Centers for Disease Control and Prevention, *Atlanta GA • Bioinformatics Fellow*   *02/2018 - 07/2018*
Key Qualifications and Responsibilities:

o   Performed routine high-quality SNP analysis on the sequences submitted to PulseNet in the Enteric Diseases Laboratory Branch (EDLB) from laboratories in all 50 states as well as several local public health labs and federal agencies (USDA and FDA) in order to sub-type pathogenic enteric bacteria.
o   Performed next generation sequencing data analysis which includes quality control, sequence alignment, SNP calling and downstream phylogenetic analysis.
o   Performed hierarchical clustering of hqSNPs to distinguish between related strains and unrelated isolates in a food borne outbreak investigation.
o   Implemented routine quality control checks at various checkpoints and contributed scripts to the existing bioinformatics pipeline with additional metrics to be included in the next version of the pipeline.
o   Written scripts in Python, Perl and Bash to generate summary metrics which helped troubleshoot inconsistencies in sequencing data.
o   Employed contamination determination steps across samples (Kraken) submitted by labs across the country.

EDUCATION

MS Bioinformatics • *Georgia Institute of Technology, Atlanta GA*
Graduation Year (2016 - 2017)

MSc Bioinformatics • *University of Mumbai, Mumbai IN*
Graduation Year (2014 - 2016)

BSc Microbiology • *University of Mumbai, Mumbai IN*
Graduation Year (2011 - 2014)

PUBLICATIONS

1.   Rokita, J. L., Rathi, K. S., Cardenas, M. F., Upton, K. A., Jayaseelan, J., Cross, K. L., Pfeil, J., Egolf, L. E., Way, G. P., Farrel, A., Kendsersky, N. M., **Patel, K.**, Gaonkar, K. S., Modi, A., Berko, E. R., Lopez, G., Vaksman, Z., Mayoh, C., Nance, J., McCoy, K., … Maris, J. M. (2019). Genomic Profiling of Childhood Tumor Patient-Derived Xenograft Models to Enable Rational Clinical Trial Design. *Cell reports*, *29*(6), 1675–1689.e9. https://doi.org/10.1016/j.celrep.2019.09.071
2.   Upton, K., Modi, A., **Patel, K**. *et al.* Epigenomic profiling of neuroblastoma cell lines. *Sci Data* 7, 116 (2020). https://doi.org/10.1038/s41597-020-0458-y