

K H U S H B U P A T E L

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: khushbupatel.info

LinkedIn: www.linkedin.com/in/khushbu-patel-b1a196b5/

SKILLS

- **Programming Languages:** R, Bash, Python, Perl
- **Tools & Software:** GATK, bcftools, samtools, STAR, StringTie, Kallisto, MACS2, Homer, RShiny, BBMap, MEME suite, VEP, Snakemake, Docker
- **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, Wilcoxon rank sum test, ANOVA, Kruskal-Wallis)

PROFESSIONAL EXPERIENCE

Children's Hospital of Philadelphia, Philadelphia PA • Bioinformatics Scientist II

07/2018 - Present

Key Qualifications and Responsibilities:

- 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.
- Proficient coding skills in R and bash to wrangle, analyze and visualize data.
- Sound knowledge of statistical concepts and implementation of statistical tests for hypothesis driven research.
- Employing various unsupervised machine learning methods to cluster and reduce dimensionality of a high dimensional dataset.
- Parallelization of jobs over HPC to run routine pipelines for large number of samples in addition to troubleshooting and development of pipelines.
- Containerizing pipelines (Docker) for efficient reproducibility.
- Maintaining best practices and using version controlling system (git) to maintain code.
- Developed multiple web applications (HTML, CSS, RShiny) for efficient data visualization, analysis and sharing for internal use.
- Ability to effectively communicate results in the form of reports/spreadsheets/visualizations.

- Keeping abreast of latest algorithms and data generation platforms.
- 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 • ORISE Bioinformatics Fellow 02/2018 - 07/2018

Key Qualifications and Responsibilities:

- 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.
- Performed next generation sequencing data analysis which includes quality control, sequence alignment, SNP calling and downstream phylogenetic analysis.
- Performed hierarchical clustering of hqSNPs to distinguish between related strains and unrelated isolates in a food borne outbreak investigation.
- 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.
- Written scripts in Python, Perl and Bash to generate summary metrics which helped troubleshoot inconsistencies in sequencing data.
- 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

2021

- Kendzersky, N. M., Lindsay, J., Kolb, E. A., Smith, M. A., Teicher, B. A., Erickson, S. W., Earley, E. J., Mosse, Y. P., Martinez, D., Pogoriler, J., Krytska, K., **Patel, K.**, Groff, D., Tsang, M., Ghilu, S., Wang, Y., Seaman, S., Feng, Y., Croix, B. S., . . . Maris, J. M. (2021). The B7-H3–Targeting Antibody–Drug Conjugate m276-SL-PBD Is Potently Effective Against Pediatric Cancer Preclinical Solid Tumor Models. *Clinical Cancer Research*, 1078–0432. <https://doi.org/10.1158/1078-0432.ccr-20-4221>

2020

- Raman, S., Buongervino, S. N., Lane, M., Zhelev, D. V., Zhu, Z., Cui, H., Martinez, B., Martinez, D., Wang, Y., Upton, K. A., **Patel, K.**, Rath, K. S., Navia, C. T., Harmon, D. B., Pawel, B., Dimitrov, D., Maris, J. M., Julien, J. P., & Bosse, K. (2020). The D3-GPC2-PBD ADC is Potently Efficacious Against Neuroblastoma and SCLC Via Engagement of a Conformational GPC2 Epitope. *SSRN Electronic Journal*. Published. <https://doi.org/10.2139/ssrn.3673604>
- 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>

2019

- Rokita, J. L., Rath, K. S., Cardenas, M. F., Upton, K. A., Jayaseelan, J., Cross, K. L., Pfeil, J., Egolf, L. E., Way, G. P., Farrel, A., Kendersky, 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>

PROFESSIONAL ACTIVITIES

- **Reviewer** of ICCBB 2021 (5th Internal Conference on Computational Biology and Bioinformatics)
- **Reviewer** of The Journal of Medical Internet Research
- **Reviewer** of ICBE2021 (The 10th International Conference on Biomedical Engineering and Biotechnology 2021, 18 papers)
- **Reviewer** of the 4th International Conference on Applied Biochemistry and Biotechnology (ABB 2021, 2 papers)
- **Reviewer** of the 7th International Conference on Agricultural and Biological Sciences (ABS 2021, 6 papers)
- **Reviewer** of the 7th International Conference on Fuzzy Systems and Data Mining (FSDM2021, 8 papers)
- **Reviewer** of the 3rd International Conference on Machine Learning and Intelligent Systems (MLIS 2021, 1 paper)
- **Reviewer** of 2021 International Conference on Medical Imaging Science and Technology (MIST 2021, 2 papers)
- **Member of International Program Committee** at CSBio2021 (The 12th International Conference on Computational Systems-Biology and Bioinformatics)
- **Program Committee member** of HIS 2021 (10th International Conference on Health Information Science 2021)
- **International Advisory Committee member** of International Institute of Engineers & Researchers

PRESENTATIONS

- ***Talk: Bioinformatics Tools and Applications***
Introduction to Applied Bioinformatics in Research organized by The Penn Center for Global Genomics & Health Equity at the University of Pennsylvania with the University of the Southern Caribbean, March 2021