KHUSHBU PATEL

Bioinformatics Scientist
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: www.github.com/kpatel427
Personal Website: www.khushbupatel.info
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

Children's Hospital of Philadelphia, Philadelphia PA • *Bioinformatics Scientist II*   *02/2021 - 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 Analyzed multidimensional high content data sets, integration of these data sets with clinical metadata, and assessment of correlations across datasets to identify potential therapeutic targets in neuroblastoma and other childhood cancers.
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) to ensure efficient computational reproducibility in research.
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.

Children's Hospital of Philadelphia, *Philadelphia PA • Bioinformatics Scientist I*   *07/2018 – 02/2021*
Key Qualifications and Responsibilities:

- o Provided research support by analyzing data of high complexity using sound statistical and commonly accepted bioinformatics methods to -omics data under supervision from mentors and peers.
- o Worked with bioinformatics group members to develop functioning pipelines upstream of customized analysis workflows using various programming languages like R, Perl and Bash.
- o Analyzed high throughput sequencing data from childhood cancers from available data sources and public repositories to identify potential targets associated with better prognostic outcomes.
- o Maintained pipelines by staying current with evolving algorithms and standards and incorporating into pipeline.
- o Implemented and maintained commercial and open-source software for analyzing genomic and transcriptomic data derived from other pediatric cancers.
- o Collaborating with research groups and sharing research outcomes using reports, visualizations and user interactive web portals.
- o Delivered analytical endpoints such as feature identification, annotation, characterization, prioritization, etc.
- o Contributed to presentations, research grants and manuscripts under supervision by mentors and peers.

Centers for Disease Control and Prevention, *Atlanta GA • ORISE 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

2021

- Kendsersky, 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. V., Zhelev, D. V., Zhu, Z., Cui, H., Martinez, B., Martinez, D., Wang, Y., Upton, K., **Patel, K.**, Rathi, K. S., Navia, C. T., Harmon, D. B., Li, Y., Pawel, B., Dimitrov, D. S., Maris, J. M., Julien, J. P., & Bosse, K. R. (2021). A GPC2 antibody-drug conjugate is efficacious against neuroblastoma and small-cell lung cancer via binding a conformational epitope. Cell reports. Medicine, 2(7), 100344. https://doi.org/10.1016/j.xcrm.2021.100344
- 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., 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

PROFESSIONAL ACTIVITIES

- Reviewer at the 10th International Conference on Biomedical Engineering and Biotechnology (ICBEB 2021) - 20 reviews
- Reviewer at the 7th International Conference on Fuzzy Systems and Data Mining (FSDM2021) - 18 reviews
- Reviewer at the 7th International Conference on Agricultural and Biological Sciences (ABS 2021) - 6 reviews
- Reviewer at the 4th International Conference on Applied Biochemistry and Biotechnology (ABB 2021) - 3 reviews
- Reviewer at the 3rd International Conference on Machine Learning and Intelligent Systems (MLIS 2021) - 2 reviews
- Reviewer at International Conference on Medical Imaging Science and Technology (MIST 2021) - 2 reviews
- Reviewer for Journal of Big Data - 3 reviews
- Reviewer for the Journal of Medical Internet Research – 3 reviews
- Reviewer and Program Committee member at 10th International Conference on Health Information Science 2021 (HIS 2021) - 1 review

- o Member of International Program Committee at CSBio2021 (The 12th International Conference on Computational Systems-Biology and Bioinformatics)
- o International Advisory Committee member of International Institute of Engineers & Researchers

POSTERS/ABSTRACTS

- o Vaksman Z, Blauel E, Egolf LE, Lee AL, Kaufman R, Modi A, **Patel K**, Farrel A, Basta PV, Olshan AF, Maris JM, Diskin SJ. Discovering the genetic basis of neuroblastoma initiation and progression. The Gabriella Miller Kids First Pediatric Research Program (Kids First) Poster Session at ASHG: Accelerating Pediatric Genomics Research through Collaboration. Houston, TX, USA (Poster) October 2019.

PRESENTATIONS

- o *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