KHUSHBU PATEL

khushbu_p@hotmail.com · khushbupatel.info · https://github.com/kpatel427
Atlanta, GA 30041
[redacted]

BIOINFORMATICS SCIENTIST

Experienced bioinformatics scientist with 5+ years' expertise in computational analysis of diverse biological data. Proficient in leveraging advanced algorithms and high-performance computing, adept at unraveling complex genomic and transcriptomic landscapes. Demonstrated success in developing and implementing cutting-edge methodologies to unravel intricate biological mechanisms. Adept in cancer genomics, genetics, and molecular biology, contributing to significant advancements in the understanding of biological systems. A collaborative leader driving impactful advancements in bioinformatics research.

TECHNICAL PROFICIENCIES

Languages: R, Bash, Python, Perl
Version Control: Git, GitHub
High-throughput Data: WGS/WXS, RNA-Seq, Microarray, Single-cell multiomics, ATAC-Seq, ChIP-Seq
Tools & Software: GATK, bcftools, samtools, STAR, StringTie, Kallisto, MACS2, Homer, RShiny, BBMap, MEME suite, VEP, Docker, conda/mamba, renv, Snakemake
Methods & Concepts: Single-cell multi-omics Analysis (Seurat, Signac, Monocle, slingshot, MOFA+, Harmony, DoubletFinder), Gene set enrichment (GSEA), Gene Ontology enrichment (ToppGene, Enrichr), Differential Gene expression (DESeq2, EdgeR), Dimensionality reduction (t-SNE, PCA, hierarchical clustering), Sequence Assembly & Alignment (BWA/STAR/BowTie), Gene abundance estimation (HTSeq-counts), Survival Analysis, Mutation and genetic aberration analysis (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 **02/2021 - Present**
Bioionformatics Scientist II

- 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.
- 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.
- 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) to ensure efficient computational reproducibility in research.
- 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.

Children's Hospital of Philadelphia, Philadelphia PA **07/2018 – 02/2021**
Bioionformatics Scientist I

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

Centers for Disease Control and Prevention, Atlanta GA **02/2018 - 07/2018**
ORISE Bioinformatics Fellow

- 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

Georgia Institute of Technology, Atlanta GA **2016-2017**
MS Bioinformatics

University of Mumbai, Mumbai IN **2014-2016**
MSc Bioinformatics

University of Mumbai, Mumbai IN **2011-2014**
BSc Microbiology

PUBLICATIONS

2023

- Farrel, A., Li, P., Veenbergen, S., **Patel, K**., Maris, J. M., & Leonard, W. J. (2023). ROGUE: an R Shiny app for RNA sequencing analysis and biomarker discovery. BMC bioinformatics, 24(1), 303. https://doi.org/10.1186/s12859-023-05420-y

2022

- Makvandi, M., Samanta, M., Martorano, P., Lee, H., Gitto, S. B., **Patel, K**., Groff, D., Pogoriler, J., Martinez, D., Riad, A., Dabagian, H., Zaleski, M., Taghvaee, T., Xu, K., Lee, J. Y., Hou, C., Farrel, A., Batra, V., Carlin, S. D., Powell, D. J., Jr, … Maris, J. M. (2022). Pre-clinical investigation of astatine-211-parthanatine for high-risk neuroblastoma. Communications biology, 5(1), 1260. https://doi.org/10.1038/s42003-022-04209-8

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

POSTERS

- **Patel K**, Rathi K, Farrel A, Raman P, Maris JM. Pediatric Cancer Web Portal: From Data to Discovery, Enhancing Pediatric Cancer Research at MidAtlantic Bioinformatics Conference, Philadelphia PA, USA (Poster) October 2023.

- 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/OUTREACH

Talk: 'Upward Bound' Lecturer (2022)
Introduction to Bioinformatics (Day 1) for Pharmacology Graduate Opportunities for Outreach and Development (Pharm4GOOD), organized by the graduate students at the Perelman School of Medicine at the University of Pennsylvania, July 2022.

Workshop: Virtual hands-on workshop with Bioridll India (2021)
Virtual hands-on workshop demonstrating Bioinformatics tools and applications by performing genomic analysis on transcriptomic data derived from pediatric cancers organized by Bioriidl in their conference Darwin, October 2021.

Talk: Bioinformatics Tools and Applications (2021)
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.

EXTERNAL LINKS

- Personal Website: https://khushbupatel.info/
- Github: https://github.com/kpatel427
- LinkedIn: https://www.linkedin.com/in/khushbupatel17/
- Google Scholar: https://scholar.google.com/citations?user=nzixRq8AAAAJ&hl=en

References are available on request.