

SKILLS

Data Science & Programming

- Adept in Python (Numpy, Pandas, Scikit-learn), SQL, R, Data Visualization (Seaborn, Plotly), and Git
- Proficient in statistical Machine Learning and Big Data analysis using Spark (MapReduce and ML)
- Well-versed in Deep Learning (Tensorflow, Keras, transfer learning) and data workflows on AWS

Multi-Omic Proficiency and Bioinformatic Expertise

- **Expertise in transcriptomic and genomic data**
 - Knowledge of germline, somatic, and RNA seq and WES NGS workflows and pipelines
- **Experience in molecular assay development and proteomic assays**
 - High-throughput screening platform for mutant p53 reactivating compounds
 - CRISPR-mediated gene editing: Rapid knockout and tagging tools
 - LC-Mass Spectrometry analysis and Multicolor flow cytometry (BD FACSDIVA), ELISA
- **Conversant in scRNA-seq and scATAC-seq methodologies**

Project Management

- Skilled in using JIRA, Bitbucket, and Confluence for project organization, repository management, and documentation
 - Responsible for cross-departmental communication with stakeholders in R&D, clinical diagnostics, assay validation, and clinical lab
-

EXPERIENCE

Bioinformatic Scientist III

Ambry Genetics

Dec 2019 - Present

- Analyzed NGS panel probe coverage, sequencing run consistency, coverage uniformity, detection sensitivity and specificity for improved results.
- Provided insights for regions with insufficient coverage, assay detection limits, and final validation of the HGMDNext panel, which includes 4,500 genes
- Evaluated and validated pipeline performance for oncology panels, including high-volume CancerNext and Somatic panels.
- Advised on reducing CNV calling QC failures and false positive rates during a capture chemistry transition.
- Established sample DNA concentration thresholds using Logistic Regression for the bait-capture automation project, involving the Tecan-robot DNA M200 workflow.
- Maintained weekly CNV calling statistics, ensuring high-quality CNV calling.
- Set statistical thresholds on allele frequency for reducing laboratory Sanger workload.
- Offered pipeline maintenance and troubleshooting solutions to support clinical operations.
- Produced crucial data for variant assessment, including SNV, indel, CNV, Mobile Elements, and processed pseudogenes.
- Participated in the BIMM project to compile key sequencing performance indicators across Ambry panels.
- Employed GradientBoostingRegressor for estimating the number of CNVs per sample and determining feature importance (including sequencing, coverage, and augmented metrics) in high false positive CNV investigations.

Data Science Fellowship

The Data Incubator

June 2019 - Sept 2019

- Identified degree of connection, page rank, and best friends in “New York Social Diary” archive using Beautiful Soup, **Pandas**, Regex, Web-scraping, and NetworkX
- Analyzed New York restaurant inspection data with **SQL**: table joining, aggregation, and clean up
- Predicted Yelp star ratings based on venue features through **Machine Learning** models
- Utilized Vectorizers for **Natural Language Processing** on Yelp review text to predict star ratings

Postdoctoral Researcher

UC Irvine

Jan 2014 - Dec 2019

- Developed **CRISPR gene editing** (tagging and knockout) and viral derived vectors to investigate PP2A L309 methylation during SAM-checkpoint activation
- Purified HBTH-tagged PP2A (SILAC) for LC-MS analysis, resulted in identifying PP2A interaction map during methionine stress
- Developed bio-screening platform in mammalian cells for p53 reactivating compounds, resulted in a publication in **Nature Communication**

EDUCATION

University of California, Irvine

Sept 2007 – Jan 2014

Ph.D. Biological Sciences

Georgia Institute of Technology

M.S. OMSCS with Machine Learning Specialization

In progress

AWARDS & PUBLICATIONS

Submitted | Nutrient control of splice site selection contributes to methionine dependence of cancer 2023

Submitted | Sphingosine and anti-neoplastic sphingosine analogs activate PP2A and inhibit nuclear import in parallel by engaging PPP2R1A and importins 2023

Cell Chemical Biology | Discovery of compounds that reactivate p53 mutants in vitro and in vivo 2022

Journal of Lipid Research | Lipid remodeling in response to methionine stress in MDA-MBA-468 triple-negative breast cancer cells 2021

Data Science Fellowship | The Data Incubator - San Francisco 2019

Journal of Biological Chemistry | Microhomology based CRISPR tagging tools for protein tracking, purification, and depletion 2019

Methionine Dependence of Cancer and Aging | Isolation and characterization of methionine-independent clones from methionine-dependent cancer cells 2019

U.S. Patent | Chembridge Small Molecules that could enhance p53 activity 2015

Journal of Cell Science | SAM limitation induces p38 mitogen-activated protein kinase and triggers cell cycle arrest in G1 2014

Nature Communication | Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53 2013

Cell Cycle | Downregulation of Cdc6 and pre-replication complexes in response to methionine stress in breast cancer cells 2012

Predoctoral Fellowship | California Institute for Regenerative Medicine 2012

Journal of Biological Chemistry | Transforming growth factor β up-regulates cysteine-rich protein 2 in vascular smooth muscle cells via activating transcription factor 2 2008

Genes to Cells | Identification of a putative human mitochondrial thymidine monophosphate kinase associated with monocytic/macrophage terminal differentiation 2008