

PATRICK CHERRY, PHD

✉ patrick [at] pcherry dot io | 📞 upon request | 🏢 Senior Scientist | Oncology | 🏠 BillionToOne | 🏡 San Francisco, California |
🌐 pcherry.io | 🐙 github.com/pdcherry | in linkedin.com/in/p-cherry | 🇺🇸 United States Citizen | Updated on 2025-11-15

PhD data scientist and professional tool-builder and science communicator. I bring the incisive critical thinking of an -omics-trained scientist to data science. Skilled at data visualization, pipeline building, and custom tool building and validation. I am dedicated to serving as the bridge between complex data problems and action-oriented leadership seeking data-driven strategy.

🏢 INDUSTRY EXPERIENCE

Current
|
Sep 2024

● Senior Scientist, Oncology

BillionToOne, Inc.

📍 Menlo Park, California

- Implemented department-wide copy number database using extract-transform-load (ETL) procedure on production data
- Ideated, coded, and validated [patented](#) method for tumor fraction estimation from plasma aneuploidy data using Gaussian mixture models, custom model selection metric, and custom consensus logic
- Solved DNA variant detection sensitivity issue by designing and implementing a multiplexed amplicon NGS assay with public data sets and internal database of variant data
- Productionized RNA-seq by writing a custom Nextflow pipeline for analysis and automated quality control reporting.
- Answered business and product design questions from leadership about un-met customer need with internal reports on Notion written in reproducible coding notebooks using Python and R in a HIPAA-manner environment
- As hiring manager, recruited and hired four Research Associates; managed the Statistics interview

Apr 2024
|
Dec 2021

● Senior Scientist, Genomics

Twist Bioscience, Inc.

📍 South San Francisco, California

- Invented and introduced multiple new products to market, yielding \$1M+ in new revenue as *Tech Lead*, including: *Pan-cancer cfDNA* (v1 & v2) ([patented](#)), *CNV Controls*, *RNA Fusion Controls*, RNA-seq, and [patented adapter removal](#)
- Answered feasibility questions with [Design of Experiment](#) principles and data analysis using R and Python.
- Wrote, documented, and maintained production QC pipelines for sub-components and finished good QC; trained end-users in Manufacturing and Quality on their use, interpretation, and reporting.
- Maintained production code (`git` version control with code review and team best practices) with CD (Github actions); operated in AWS & Databricks high-performance compute environment using Unix shell
- Answered manufacturing & business questions using SQL / Snowflake database of production events and customer data

Jan 2021
|
Apr 2019

● Scientist I, Next Generation Sequencing (NGS)

Zymergen, Inc.

📍 Emeryville, California

- Boosted pan-project NGS genotyping success by 45% using [Design of Experiment](#) and Estimated Marginal Means modeling for causal inference; Guided demanding and diverse internal customers on complex NGS experiments
- Built and disseminated reproducible statistical templates in Rmarkdown and JMP for self-serve NGS data exploration
- Rendered strain build and plasmid build reports from LIMS SQL database in MariaDB and R / `ggplot2`
- Designed, built, and demonstrated a [patented](#) counterselectable marker for the non-model microbe *Bacillus subtilis*

🎓 EDUCATION

Apr 2019
|
Aug 2013

● PhD in Molecular Biology

University of Colorado School of Medicine

📍 Aurora/Denver, Colorado

- Ph.D. in Molecular Biology. Advisor: Jay Hesselberth, PhD. Key courses: 6606 Statistics; 7621 Genome Analysis Workshop.
- Thesis: RNA terminus chemistry potentiates decay events targeting *HAC1* mRNA during the Unfolded Protein Response

May 2013
|
Aug 2009

● BA in Molecular Biology with Distinction, minor in Mathematics

Hendrix College

📍 Conway, Arkansas

- Senior Capstone Project showing changes of key stress response gene expression to oxidative stress in liver cells
- Key courses: Probability & Statistics; Discrete Mathematics; Adv Biological Chemistry, Analytical Chemistry