PATRICK CHERRY

✉ pcherry [at] pm dot me | 📞 upon request | 🗂 Senior Scientist | Genomics | 🏢 Twist Bioscience | 📖 San Francisco, California |
🌐 pdcherry.github.io | ⭘ github.com/pdcherry | in linkedin.com/in/p-cherry | 🪪 United States Citizen | Updated on 2024-05-22

PhD scientist skilled in data visualization, statistical modeling, bioinformatics, next-generation sequencing (NGS), and tool-building. I've coded reproducible and rigorous pipelines for high-throughput experimental designs and multi-omic analyses for communication to technical and non-technical audiences. I've launched best-in-class oncology reference standards and analyzed NGS data from Illumina, MGI, and Element. Originally trained in Molecular Biology, I am passionate about advancing data science and bioinformatics to improve human health and understand biology.

🎓 EDUCATION

PhD
2019 | 2013

University of Colorado School of Medicine

📍 Aurora/Denver, Colorado

- Ph.D. in Molecular Biology. Advisor: Jay Hesselberth, PhD
- Thesis: RNA terminus chemistry impact decay events that target *HAC1* mRNA during the Unfolded Protein Response

BA
2013 | 2009

Hendrix College

📍 Conway, Arkansas

- Biochemistry and Molecular Biology, with Distinction. Advisor: Andres Caro, PhD
- Senior Capstone Project showing key stress response gene expression changes to oxidative stress in liver cells
- Minor in Mathematics; PI: Lars Seme; Project: Newton's method as a fractal chaotic dynamical system

💼 INDUSTRY EXPERIENCE

Senior Scientist, Genomics
2024 | 2021

Twist Bioscience

📍 South San Francisco, California

- Invented and introduced multiple new products to market yielding millions of dollars in new revenue as *Tech Lead*, including: *Pan-cancer cfDNA* (v1 & v2), *CNV Controls*, *RNA Fusion Controls*, *Fragmentome Controls*, and RNA-seq
- Answered biological and business quesitons with reproducible exploratory analysis using R and the libraries: `dplyr`, `dbplyr`, `DBI`, `purrr`, `ggplot2`, `Bioconductor`, `DEseq2`, `seurat`, `tidymodels`, `glm`, `nls`, `lme4`, and more
- Edited and mainteined production QC pipelines with automated reporting using `Python` and the libraries: `numpy`, `pandas`, `polars`, `seaborn`, `statsmodels`, `biopython`, `pybedtools`, `pysam`, `vcfpy`, `scanpy`, and others
- Mintained production code (`git` version control with code review) and operated in high-performance compute environment `aws` `s3` & `ec2` and `Databricks` using Unix shell / `bash` tools, like `STAR-fusion`
- Authored internal R package "`twistcolorpal`" on Github with help files to help scientists style plots to Twist branding

Scientist, Genomics
2024 | 2021

Twist Bioscience

📍 South San Francisco, California

- As Tech Lead, launched the Twist *Pan-Cancer Reference Standard*, an ISO-13485 synthetic positive control with 458 unique variants among 84 cancer-associated genes at six QC'd VAFs, plus a WT control; Launched in Nov of 2021, and earned $1 million in new revenue in first year
- Identified and optimized compatible ddPCR assays using web scraping in R with `rvest` for cfDNA quality control
- Used Python to author automated data / QC reports using `Jinja2` (with `CSS`) and `WeasyPrint`
- Managed data and ran bioinformatic analyses on Illumina RNA & DNA seq using bash tools `awk`, `grep`, `scp`, `wget`, `ssh`, `bwa`, `bedtools`, `samtools`, `vcftools`, `UMI-tools/fgbio`
- Wrote original SQL queries for `Snowflake`-based database to answer production and business questions.

Scientist I, NGS & NPI-Build
2021 | 2019

Zymergen, Inc.

📍 Emeryville, California

- Boosted *NGS Core* genotyping success by 45% using data-driven decision-making and teaching; Guided demanding and diverse internal customers on complex NGS experiments
- Built and disseminated `Rmarkdown` notebook for autonomous NGS data exploration
- Coded build and QC experiments using Zymergen's `Drawbridge` alembic API Python library to LIMS
- Rendered strain build and plasmid build reports from LIMS SQL database in `MySQL` Workbench

SELECTED INTELLECTUAL PROPERTY

Libraries for mutational analysis
4/9/21
Twist Bioscience
⦿ South San Francisco, California

· US Prov. Pat. Ref No 823.101
· Configuration and fabrication of synthetic DNA & RNA reference standards and synthetic variant sequences

Method for counterselection in microorganisms
3/25/21
Zymergen, Inc.
⦿ Emeryville, California

· US 2021_0087586 A1
· Demonstration of novel counterselection mechanism in non-model *Bacillus* microbe for genetic modification

SELECTED PUBLICATIONS

Twist cfDNA Pan-Cancer Reference Standard v2 Technical Guidance
1/4/24
Twist Bioscience Product Sheet & FAQ
⦿ South San Francisco, California

· Patrick Cherry, Lydia Bonar, & Mike Bocek

Characteristics and specificity of the wild-type / 0% VAF reference material
4/22/22
Twist Bioscience
⦿ South San Francisco, California

· Patrick Cherry & Mike Bocek

Multiple decay events target HAC1 mRNA during splicing to regulate the unfolded protein response
3/19/19
eLife
⦿ Aurora/Denver, Colorado

· Cherry, P., Peach, S., & Hesselberth, J.

Genetic bypass of essential RNA repair enzymes in budding yeast
12/6/17
RNA
⦿ Aurora/Denver, Colorado

· Cherry, P., White, L., York, K., & Hesselberth, J.

SELECTED PRESENTATIONS & TALKS

Twist pan-cancer reference standard V2: Enhanced precision and reduced errors in ctDNA analysis
2/6/24
Advances in Genome Biology and Technology
⦿ Orlando, Florida

Twist pan-cancer synthetic RNA fusion control for assay development
4/16/23
American Association for Cancer Researchers, Twist Bioscience
⦿ Orlando, Florida

Use of synthetic CNV fragments to mimic copy number alterations for ctDNA reference standards
2/7/23
Advances in Genome Biology and Technology
⦿ Hollywood, Florida

Twist pan-cancer synthetic reference materials for cell-free DNA (cfDNA) assay development
4/12/22
American Association for Cancer Researchers
⦿ New Orleans, Louisiana

RNA processing regulates the unfolded protein response
8/24/17
CSHL: mRNA Processing Meeting Talk
⦿ Cold Spring Harbor, New York

R use at Zymergen
6/16/20
Z-Tech Talk
⦿ Emeryville, CA

Data-driven troubleshooting of NGS experiments
4/20/20
Data Science Talk
⦿ Emeryville, CA