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 2024-06-10

PhD scientist and professional critical thinker, problem-solver, and science communicator. At the interface of data science, genomics, and statistics; skilled at data visualization, bioinformatics, next-generation sequencing (NGS), and tool-building. I've coded reproducible pipelines for high-throughput experimental designs and genomic analyses. 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 to understand biology.

🎓 EDUCATION

PhD
University of Colorado School of Medicine

2019 | 2013

📍 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
Hendrix College

2013 | 2009

📍 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
Twist Bioscience

2024 | 2021

📍 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 questions with DoE and data analysis using R and the libraries: `dplyr`, `dbplyr`, `DBI`, `purrr`, `ggplot2`, Bioconductor, `DEseq2`, `seurat`, `tidymodels`, `glm`, `nls`, `lme4`, `AlgDesign`, and more
- Edited and maintained production QC pipelines with automated reporting using `Python` and the libraries: `NumPy`, `pandas`, `polars`, `seaborn`, `statsmodels`, `biopython`, `pybedtools`, `pysam`, `vcfpy`, `scanpy`, and others
- Maintained production code (`git` version control with code review) with CD (github actions); operated in high-performance compute environment `aws s3` 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
Twist Bioscience

2024 | 2021

📍 South San Francisco, California

- Tech Lead of Nov. 2021 launch of Twist *Pan-Cancer Reference Standard*, an ISO-13485 synthetic control with 458 variants among 84 cancer-associated genes at six QC'd VAFs, plus a WT control, earning $1 million in revenue in year 1
- 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`, documented on Jira
- Managed data and ran bioinformatic analyses on Illumina RNA & DNA seq using bash tools like: `awk`, `grep`, `scp`, `wget`, Illumina BaseSpace CLI, `ssh`, `bwa`, `bedtools`, `samtools`, `vcftools`, `UMI-tools/fgbio`; Confluence documentation
- Wrote original SQL queries for `Snowflake`-based database to answer production and business questions.

Scientist I, NGS & NPI-Build
Zymergen, Inc.

2021 | 2019

📍 Emeryville, California

- Boosted *NGS Core* genotyping success by 45% using DoE experimentation, 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 plasmid and strain build and QC experiments using Zymergen's alembic Python API to LIMS database
- Rendered strain build and plasmid build reports from LIMS SQL database in `MySQL` Workbench

💡 SELECTED INTELLECTUAL PROPERTY

4/9/21

Libraries for mutational analysis
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

3/25/21

Method for counterselection in microorganisms
Zymergen, Inc.
📍 Emeryville, California

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

📖 SELECTED PUBLICATIONS

1/4/24

Twist cfDNA Pan-Cancer Reference Standard v2 Technical Guidance
Twist Bioscience Product Sheet & FAQ
📍 South San Francisco, California

- Patrick Cherry, Lydia Bonar, & Mike Bocek

4/22/22

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

- Patrick Cherry & Mike Bocek

3/19/19

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

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

12/6/17

Genetic bypass of essential RNA repair enzymes in budding yeast
RNA
📍 Aurora/Denver, Colorado

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

🖥️ SELECTED PRESENTATIONS & TALKS

2/6/24

Twist pan-cancer reference standard V2: Enhanced precision and reduced errors in ctDNA analysis
Advances in Genome Biology and Technology
📍 Orlando, Florida

4/16/23

Twist pan-cancer synthetic RNA fusion control for assay development
American Association for Cancer Researchers, Twist Bioscience
📍 Orlando, Florida

2/7/23

Use of synthetic CNV fragments to mimic copy number alterations for ctDNA reference standards
Advances in Genome Biology and Technology
📍 Hollywood, Florida

4/12/22

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

6/16/20

R use at Zymergen
Z-Tech Talk
📍 Emeryville, CA

4/20/20

Data-driven troubleshooting of NGS experiments
Data Science Talk
📍 Emeryville, CA

8/24/17

RNA processing regulates the unfolded protein response
CSHL: mRNA Processing Meeting Talk
📍 Cold Spring Harbor, New York