PATRICK CHERRY, PHD

✉ pcherry [at] pm dot me | 📞 upon request | 💼 Senior Scientist | Oncology | 🏬 BillionToOne | 🏙 San Francisco, California |
🌐 pdcherry.github.io | github.com/pdcherry | in linkedin.com/in/p-cherry | 🪪 United States Citizen | Updated on 2025-01-14

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 am passionate about advancing data science and bioinformatics to improve human health and to understand biology.

💼 INDUSTRY EXPERIENCE

Senior Scientist, Oncology

Current | Sep 2024

BillionToOne, Inc.

📍 Menlo Park, CA

- Designed and implemented multiplexed amplicon NGS assay with `Bioconductor` NCBI tools and `primer3` primer selection; generated 500+ automated ddPCR designs with pipeline in `snakemake` using `s3` database of variant data.
- Informed key product decisions with analyses using PostgreSQL and s3 data sources and linear mixed-effects modeling
- Circulated internal research reports on Notion written in Quarto notebooks with reproducible Python and R code.

Senior Scientist, Genomics

Apr 2024 | Dec 2021

Twist Bioscience, Inc.

📍 South San Francisco, California

- Invented and introduced multiple new products to market, yielding $1M+ in new revenue as *Tech Lead*, including: *Pancancer cfDNA* (v1 & v2), *CNV Controls*, *RNA Fusion Controls*, *Fragmentome Controls*, RNA-seq, and primer / adapter removal methods for highly-multiplexed dsDNA pools
- 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
- Wrote and published reports in Rmarkdown and LaTeX. Edited and maintained production QC pipelines with automated reporting using `Python` and the libraries: NumPy, pandas, `polars`, `seaborn`, `statsmodels`, biopython, `pybedtools`, `pysam`, vcfpy, `scanpy`, `Jinja2` (with CSS), `WeasyPrint`, 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 `ssh`, STAR–fusion, GATK, BaseSpace CLI, bwa, `bedtools`, `samtools`, vcftools, UMI–tools/fgbio; Confluence documentation
- Authored internal R package "`twistcolorpal`" on Github with help files to help scientists style plots to Twist branding
- Answered production and business questions with original SQL queries for `Snowflake`-based database

Scientist I, NGS Core & Strain-Build Process Development

Jan 2021 | Apr 2019

Zymergen, Inc.

📍 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` and JMP statistical templates 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

🎓 EDUCATION

PhD

Apr 2019 | Aug 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
- Key courses: BIOS 6606 Statistics for Basic Sciences; MOLB 7621 Genome Analysis Workshop.

BA

May 2013 | Aug 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
- Key courses: MATH 310 Probability & Statistics; MATH 240 Discrete Mathematics; CHEM 335 Adv Biological Chemistry.

SELECTED INTELLECTUAL PROPERTY

Apr 2021

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

Mar 2021

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

Jan 2024

Twist cfDNA Pan-Cancer Reference Standard v2 Technical Guidance
Twist Bioscience Product Sheet & FAQ

📍 South San Francisco, California

• Patrick Cherry, Lydia Bonar, & Mike Bocek

Apr 2022

Characteristics and specificity of the wild-type / 0% VAF reference material
Twist Bioscience

📍 South San Francisco, California

• Patrick Cherry & Mike Bocek

Mar 2019

Multiple decay events target HAC1 mRNA during splicing to regulate the unfolded protein response
eLife

📍 Aurora/Denver, Colorado

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

Dec 2017

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

Feb 2024

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

📍 Orlando, Florida

Apr 2023

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

📍 Orlando, Florida

Feb 2023

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

📍 Hollywood, Florida

Apr 2022

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

📍 New Orleans, Louisiana

Jun 2020

R use at Zymergen
Z-Tech Talk

📍 Emeryville, CA

Apr 2020

Data-driven troubleshooting of NGS experiments
Data Science Talk

📍 Emeryville, CA

Aug 2017

RNA processing regulates the unfolded protein response
CSHL: mRNA Processing Meeting Talk

📍 Cold Spring Harbor, New York