

Manuel Razo-Mejia

Postdoctoral Scholar · Stanford University

327 Campus Drive, Stanford, CA 94305

✉ manuel.razo.m@gmail.com 🏠 <https://mrzomej.github.io> 🌐 [mrzomej](#) 📄 [Manuel Razo-Mejia](#) 📍 [Manuel Razo-Mejia](#) 📄

0000-0002-9510-0527

I am a computational scientist specializing in machine learning for biological data analysis. I develop novel computational approaches combining dynamical systems theory, biophysical modeling, and deep learning techniques to extract actionable insights from complex multi-scale biological datasets. My research focuses on creating robust, scalable workflows that bridge fundamental cellular principles with rigorous data-driven decision making—all implemented with strong software engineering practices to ensure reproducibility and deployment readiness.

Professional Experience

Statistical Consultant

Superfluid Dx, Inc.

Reference: [John J. Sninsky, PhD](#)

February 2024 - April 2024

Conducted statistical consulting to design a statistical method for the analysis of high-dimensional sequencing data in the medical diagnostics space.

- Proposed a novel statistical normalization method for bulk RNA-sequencing data from clinical samples, developing interpretable approaches to account for uncertainty due to variable sequencing depth and batch effects.
- Presented a scalable variational inference algorithm for implementation of the proposed normalization method, demonstrating expertise in probabilistic modeling techniques relevant to biological data analysis.

Postdoctoral Scholar & [Schmidt Science Fellow](#)

Stanford University

Department of Biology | [Dmitri Petrov Lab](#)

October 2021 - Present

Combining theoretical, statistical, and deep learning approaches to gain insights into the structure of the genotype-phenotype-fitness map determining the evolution of living systems from high-dimensional data.

Highlighted work: [Razo-Mejia, M., et al. \(2024\)](#) [DOI: [10.1371/journal.pcbi.1011937](#)] Scales up Bayesian inference via variational methods for high-throughput DNA barcode tracking in competition experiments.

- Developed and implemented a novel statistical framework for analyzing high-throughput DNA barcode sequencing data, significantly enhancing the quantification of uncertainty in insights derived from complex, high-dimensional datasets.
- Authored a high-quality open-source software package ([BarBay.jl](#)), employing rigorous software engineering practices for reproducible analysis of evolutionary dynamics data.

PhD Student

California Institute of Technology

Division of Biology and Biological Engineering | [Rob Phillips Lab](#)

September 2014 - August 2021

Explored the information processing capabilities of cells from first principles, investigating how organisms gather environmental information, generate internal representations, and respond to maximize survival.

Highlighted work: [Razo-Mejia, M., et al. \(2020\)](#) [DOI: [10.1103/PhysRevE.102.022404](#)] Develops a physical model of how a simple genetic circuit integrates and processes information, allowing for the prediction of the number of bits that cells can reliably encode from the environment.

- Combined insights from *statistical physics* and *information theory* to develop a theoretical model for the information processing capabilities of a common regulatory motif in bacteria.
- Integrated diverse information sources with rigorous uncertainty quantification to achieve robust predictions of experimental outcomes, ensuring all model parameters were effectively constrained by the data.

Technical Skills

Programming	Python, Julia, Docker, bash, Linux, Git & GitHub, GitHub Actions, GPU Kernel Programming
Data Science	Jax, Pandas, NumPy, Matplotlib, Seaborn, SciPy, Jupyter, NumPyro
Math & Statistics	Deep Learning, Differential Geometry, Bayesian Inference, Variational Inference, Statistical Physics, Differential Equations, Linear Algebra, Probability Theory, Stochastic Processes
Data Analysis	Bulk & Single-Cell RNA Sequencing, Quantitative Microscopy, Flow Cytometry, DNA Barcode Technologies

Scientific Software

SCRIBE

A Python package for analyzing single-cell RNA sequencing (scRNA-seq) data using variational inference and MCMC sampling based on NumPyro. [GitHub](#) | [Docs](#)

AutoEncoderToolkit.jl

A Julia package for training variational autoencoders and their extensions. Includes utilities for geometric analysis of the latent spaces. [GitHub](#) | [Docs](#)

BarBay.jl

A Julia package for statistical analysis of high-throughput pooled assays in experimental evolution using variational Bayesian inference. [GitHub](#) | [Docs](#)

Education

PhD Biochemistry & Molecular Biophysics
California Institute of Technology
2014 - 2021 | Pasadena, CA | Advisor: [Rob Phillips](#)

BSc Biotechnological Engineering
Instituto Politecnico Nacional
2009 - 2014 | Silao, Gto, MX

Selected Publications

A full list is available via [Google Scholar](#) and my [personal website](#).

Razo-Mejia, M., Mani, M., Petrov, D. (2025). *Learning the Shape of Evolutionary Landscapes: Geometric Deep Learning Reveals Hidden Structure in Phenotype-to-Fitness Maps*. bioRxiv . [DOI](#)

- **Summary:** A method for transforming complex biological fitness landscapes into interpretable geometric spaces using cutting-edge deep learning, unlocking hidden evolutionary patterns and adaptive trajectories that traditional methods cannot capture.
- **Applied Skills:** Geometric deep learning, evolutionary biology, phenotype-to-fitness maps, high-dimensional data analysis, dimensionality reduction.

Razo-Mejia, M. (2024). *AutoEncoderToolkit.jl: A Julia package for training (Variational) Autoencoders*. JOSS . [DOI](#)

- **Summary:** A Julia software package that provides a flexible interface for training variational autoencoders (VAEs) and their extensions, with utilities for geometric analysis of the learned latent spaces.
- **Applied Skills:** Scientific software development, generative model implementation, representation learning, geometric data analysis, high-dimensional data analysis.

Razo-Mejia, M., Mani, M., Petrov, D. (2024). *Bayesian inference of relative fitness on high-throughput pooled competition assays*. PLoS Comput Biol, 20, e1011937 . [DOI](#)

[Paper website](#) | [GitHub repository](#)

- **Summary:** A Bayesian workflow for inferring microbial fitness from lineage tracking data that handles multiple noise sources, propagates uncertainties systematically, and scales to large datasets via variational inference. The framework supports multi-environment assays and genotype-linked barcodes, with open-source software for experimental evolution applications.
- **Applied Skills:** Probabilistic modeling, variational inference methods, scalable algorithms for biological data, scientific software development and deployment.

Razo-Mejia, M., et al. (2020). *First-principles prediction of the information processing capacity of a simple genetic circuit*. Phys. Rev. E , 022404, 1–14 . [DOI](#)

[Paper website](#) | [GitHub repository](#)

- **Summary:** A first-principles model of a simple genetic circuit that predicts the information processing capacity of bacterial cells, accounting for stochastic gene expression. The model, built entirely from independently published parameters, successfully captures how cell-to-cell variability scales with biophysical parameters and accurately predicts the circuit's ability to transmit information about environmental states, with implications for cellular survival in fluctuating environments.
- **Applied Skills:** Theoretical modeling, information theory, statistical physics, Bayesian inference.

Razo-Mejia, M.^①, Barnes, S. L.^①, Belliveau, N. M.^①, Chure, G.^①, Einav, T.^①, Lewis, M., Phillips, R. (2018). *Tuning Transcriptional Regulation through Signaling: A Predictive Theory of Allosteric Induction*. Cell Systems , ^①. These authors contributed equally . [DOI](#)

[Paper website](#) | [GitHub repository](#)

- **Summary:** Developed a predictive theory of allosteric transcriptional regulation connecting tunable parameters to gene expression responses. The model accurately captured experimental data from bacterial systems and provided analytic expressions for key regulatory properties, advancing our understanding of allosteric regulation.
- **Applied Skills:** Theoretical modeling, statistical mechanics, Bayesian inference, team leadership.