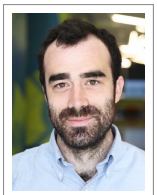
Rory Donovan-Maiye

615 Westlake Ave N
Seattle, WA 98109
✉ donovanr@alumni.reed.edu
donovanr.github.io



Employment

2017– **Scientist**, *Allen Institute for Cell Science*.
Integrative methods for cell modeling, primarily focused on deep learning techniques

2016–2017 **Postdoctoral Fellow**, *Institute for System Biology*.
Hodd & Price Lab – personalized medicine & wellness, mouse and human genomics

Education

2011–2016 **Ph.D.**, *Carnegie Mellon–University of Pittsburgh Program in Computational Biology*.
Systems Biology & Machine Learning

2004–2005 **M.S.**, *University of Washington*.
Physics

2000–2004 **B.A**, *Reed College*.
Physics

Research

Cell states beyond transcriptomics.
Quantitative co-analysis of RNA abundance and sarcomere organization in single cells and an integrated framework to predict subcellular organization states from gene expression. We establish a framework for multi-dimensional analysis of single cells to study the relationships between gene expression and subcellular organization and to develop a more nuanced description of cell states.

Integrated Cell Modeling.
I've been working in collaboration with Greg Johnson to build integrated models of single cells. We use conditional generative adversarial networks (GANs) and variational autoencoders (VAEs) to fuse data from multiple fluorescence microscopy experiments into a coherent model of sub-cellular structure localization in single cells.

Sparse Time-Series Models.
This project aims to integrate data across time-points to build sparse regression models for time-series data, such that the sparse regressors at neighboring time points vary smoothly. This would be useful for e.g. RNA-seq experiments with multiple time-points, if you wanted to predict the set genes driving a phenotype, and see how that set changes over time.

Integrated Workflow for Transcription Factor Binding Site Prediction.
In an effort to find the best candidate transcriptions factors to input to the Price Lab's transcriptional regulatory network inference tools, I constructed a machine learning pipeline to integrate an array of genome-scale data and predictive tools to output a single high confidence prediction of transcriptional activity at arbitrary sites across the genome.

Graphical Models for Free Energy Estimation.
Computationally estimating how strongly two biomolecules bind together is often either overly time consuming (e.g. molecular dynamics) or overly empirical (e.g. docking). Alternatively, using graphical models of proteins to compute the Bethe free energy of binding can be both fast and accurate.

Weighted Ensemble Systems Biology.
By applying weighted ensemble sampling to stochastic models of biological processes, we achieved orders of magnitude speed-ups in simulating events of interest in these complex systems. Our first paper dealt with non-spatial models, and our second paper addresses models of spatially resolved cellular processes.

Quantitative Evolution and the ATP Synthase.
Using simple state-based models of proton transport and free energy transduction, we probe the optimality of the curiously engineered rotary mechanism of the ATP synthase, using models that are entirely agnostic to structure and are optimized over all unknown parameters that are thermodynamically permissible.

Stochastic Models of Cellular Heterogeneity.
Simple stochastic models can recapitulate the population-level heterogeneity of protein abundance found in, for example, colonies of E. coli. This work employs a non-spatial model of gene expression, stochastically simulated with a modified Gillespie algorithm that takes into account cell division.

Publications

Kaytlyn A Gerbin, Tanya Grancharova, Rory Donovan-Maiye, Melissa C Hendershott, Jackson Brown, Stephanie Q Dinh, Jamie L Gehring, Matthew Hirano, Gregory R Johnson, Aditya Nath, Angelique Nelson, Charles M Roco, Alex B Rosenberg, M Filip Sluzewski, Matheus P Viana, Calysta Yan, Rebecca J Zaunbrecher, Kimberly R Cordes Metzler, Vilas Menon, Sean P Palecek, Georg Seelig, Nathalie Gaudreault, Theo Knijnenburg, Susanne M Rafelski, Julie A Theriot, and Ruwanthi N Gunawardane. Cell states beyond transcriptomics: integrating structural organization and gene expression in hipsc-derived cardiomyocytes. *bioRxiv*, 2020.

Cory C Funk, Segun Jung, Matthew A Richards, Alex Rodriguez, Paul Shannon, Rory Donovan, Ben Heavner, Kyle Chard, Yukai Xiao, Gustavo Glusman, Nilufer Erteskin-Taner, Todd Golde, Arthur Toga, Leroy Hood, John D Van Horn, Carl Kesselman, Ian Foster, Seth Ament, Ravi Madduri, and Nathan D Price. Atlas of transcription factor binding sites from encode dnase hypersensitivity data across 27 tissue types. *bioRxiv*, 2018.

Gregory R Johnson, Rory M Donovan-Maiye, and Mary M Maleckar. Building a 3d integrated cell. *bioRxiv*, 2017.

Rory M Donovan-Maiye, Christopher J Langmead, and Daniel Zuckerman. Systematic testing of belief-propagation estimates for absolute free energies in atomistic peptides and proteins. *Journal of chemical theory and computation*, 2017.

Gregory R Johnson, Rory M Donovan-Maiye, and Mary M Maleckar. Generative modeling with conditional autoencoders: Building an integrated cell. *arXiv preprint arXiv:1705.00092*, 2017.

Ramu Anandakrishnan, Zining Zhang, Rory Donovan-Maiye, and Daniel M Zuckerman. Biophysical comparison of atp synthesis mechanisms shows a kinetic advantage for the rotary process. *Proceedings of the National Academy of Sciences*, 113(40):11220–11225, 2016.

Rory M. Donovan, Jose-Juan Tapia, Devin P. Sullivan, James R. Faeder, Robert F. Murphy, Markus Dittrich, and Daniel M. Zuckerman. Unbiased rare event sampling in spatial stochastic systems biology models using a weighted ensemble of trajectories. *PLoS Computational Biology*, 2016.

Andrew J. Sedgewick, Ivy Shi, Rory M. Donovan, and Panatiotis V. Benos. Learning mixed graphical models with separate sparsity parameters and stability-based model selection. *BMC Genomics*, 2014.

Rory M. Donovan, Andrew J. Sedgewick, James R. Faeder, and Daniel M. Zuckerman. Efficient stochastic simulation of chemical kinetics networks using a weighted ensemble of trajectories. *The Journal of Chemical Physics*, 139(11):115105, 2013.

Rory M. Donovan. *Quantization of constrained systems*. Bachelor's thesis, Reed College, 2004.

Invited Talks

2017 **Reed College**, *Joint Physics, Biology, and Statistics Seminar*.
Deep Learning the Integrated Cell

Computational Skills

Languages, *Python, Julia, R, Matlab, Mathematica, Fortran, C, C++, Bash, HTML & CSS, LaTeX*.
Scientific programming with a focus on simulation algorithms, data analysis, computational efficiency, machine learning, and GPU-driven deep learning. Strong abilities and interests in:

- Python (Pytorch), Matlab, R, and Julia, with additional proficiency in lower-level compiled languages.
- Machine learning and the automated extraction of knowledge from data.
- Web design and the visual display of quantitative information.