

Christopher R. Cotter

cotter@sciencesundries.com • www.sciencesundries.com • www.github.com/cottersci

SUMMARY

My work is not done when the algorithm works on test data, it is just beginning. Answering data science and machine learning problems in the laboratory means working with noisy, imperfect data, developing biologically sound hypotheses, and working within the limits of current experimental techniques. I'm interested in solving real-world problems by turning vague ideas and lots of data into knowledge, understanding, and actionable conclusions. Particularly, I enjoy problems that have multi-disciplinary origins encompassing biology, computer science, statistics, and mathematics.

PUBLISHED WORKS

Cotter, Schüttler, Igoshin, & Shimkets. Data-driven modeling reveals cell behaviors controlling self-organization during *Myxococcus xanthus* development. PNAS. 2017.

Presents the first fully quantified model of cell coordination in our model organism of choice (*Myxococcus xanthus*), identified a cell behavior previously thought not to occur in *M. xanthus*, and predicted mutant phenotypes now being explored in the laboratory. It was featured on Nature's Biofilms and Microbiomes blog (<http://go.nature.com/2pgm8sw>).

Babbitt G.A. and **Cotter C.R.** Functional conservation of nucleosome formation selectively biases presumably neutral molecular variation in yeast genomes. Genome Biol Evol. 2011.

Identified an inverse correlation between the genome-wide nucleotide substitution frequency and the average energetic effect on nucleosome formation. Results suggest that nucleotide substitutions are biased to conserve the ability of DNA to efficiently wrap around histones, and that such bias could help explain nucleotide substitution choices in areas of elevated or suppressed nucleosome formation.

EDUCATION

University of Georgia
Athens, GA
Degree: Ph.D. Microbiology
December 2017

Rochester Institute of Technology
Rochester, NY
Degree: B.S. Bioinformatics
May 2011

EXPERIENCE

Postdoctoral Researcher

January 2018 – Present

Ph.D. Dissertation

August 2011 – December 2017

University of Georgia, Athens GA

Mentor: Dr. Lawrence Shimkets

Combined computer science, statistics, machine learning, and biological science to turn high-dimensional and noisy experimental data into understanding about biological systems. Projects include:

Feature Extraction and Clustering of Time-Lapse Movies Using Deep Learning

Developed semi-supervised convolutional neural networks to extract global phenotypic features from a dataset of highly variable time-lapse movies of *Myxococcus xanthus* development. Dataset contains over 250 different genotypes.

- Designed a mutual information theory centric training approach able to distinguish salient variation from noise.
- Used cutting-edge method of estimating unknown probability distributions with generative adversarial networks.
- Work performed in PyTorch, utilizing packages such as Jupyter Notebooks, NumPy, and scikit-learn.
- Models trained on the AWS EC2 cloud platform as well as a local compute cluster.
- Applied biological domain knowledge to identify custom dataset augmentation methods.
- Confirmed algorithm works on test datasets (MNIST, SVHN), tuning for experimental data is in progress.
- Continually implementing and evaluating usefulness of new neural network models published on arXiv.

Image Acquisition, Quantification, Feature Extraction, and Hypothesis Testing

Extracted almost 40 quantitative features from time-lapse fluorescent microscopy data and performed model driven hypothesis testing to understand cell coordination in a bacterial model of multicellular patterning.

- Identified 4 cell-behavior cues that allow cells to create structures hundreds of times larger than themselves.
- Developed custom image segmentation and multiple object tracking algorithms.
- Utilized unsupervised machine learning techniques, such as nearest neighbor regression and ReliefF.
- Denoised and categorized highly variable cell trajectories using extended Kalman filters.

- Tested identified-cell-behavior predictors using in silico simulations directly driven by experimental data.
- Created advanced visualizations for presentations and publication (example at: sciencesundries.com/publications).
- Performed strain construction (cloning, PCR, sequencing, etc.) and time-lapse fluorescent microscopy.

Using Neural Networks for Model Fitting

Designed neural networks to fit partial differential equation models to multivariate and temporal experimental data.

- Proof-of-concept models are accurate on test data within 5% or within 20% when 20% gaussian noise is added.

Noted Outcomes

- Publication in high-impact journal: **Cotter**, Schüttler, Igoshin, & Shimkets. Data-driven modeling reveals cell behaviors controlling self-organization during *Myxococcus xanthus* development. PNAS. 2017.
- Fostered cross-disciplinary collaborations with experts in physics, systems biology, and microbiology.

NSF REU in Prokaryotic Biology

University of Georgia, Athens GA

June 2010 – August 2010

Mentor: Dr. Margie Lee

Inferred metabolic state of bacterial species in chicken gut microbiome samples from RNA sequencing data.

Undergraduate Research

Rochester Institute of Technology, Rochester NY

10/2009 - 6/2011

Mentor: Dr. Gregory Babbitt

Developed Perl wrappers for Fortran code to analyze nucleosome positioning in yeast genomes.

Website Design and Consulting

Elk County Community Foundation, St. Marys, PA

July 2010 – September 2010

Created a website front end (HTML/CSS/JavaScript) and backend framework (PHP/MySQL) based on customer's needs.

RELEVANT TECHNICAL EXPERIENCE

Neural Networks: Convolutional/Linear Networks, Variational Autoencoders, Generative Adversarial Networks

Programming: Python, C++, Java, MATLAB, PyTorch, Relational Databases, Jupyter Notebooks, NumPy, scikit-learn, matplotlib, Parallel Programming, AWS EC2, MapReduce, Unit Testing, Web development (e.g. sciencesundries.com)

Statistical Learning: Clustering, tSNE, PCA, Cross-Validation, Classification, Bootstrapping

AWARDS

2015: UGA Graduate School Travel Award to attend 42nd Myxobacteria meeting.

2014: UGA Graduate School Travel Award to attend q-Bio meeting.

2014: UGA Graduate School Travel Award to attend South Eastern Society for Developmental Biology Conference.

2013: UGA Innovative and Interdisciplinary Research Grant.

2011: University of Georgia: Graduate School Assistantship.

2011: Stony Brook University: Lauffer Center for Physical and Quantitative Biology Fellowship (Declined).

PRESENTATIONS

2017: Invited talk for UGA Department of Physics, Center for Simulation Physics Lunch Seminars.

2017: 44th International Conference of the Biology of Myxobacteria

2016: 2016 iPoLS Annual Meeting

2015: 42nd International Conference of the Biology of Myxobacteria

2014: 8th q-Bio Conference.

2014: South Eastern Society for Developmental Biology Conference.

2013: UGA Developmental Biology Alliance Retreat.

MENTORING

2013-2014: Mentored undergraduate student participating in the UGA Center for Undergraduate Research Opportunities (CURO). Student went on to pursue a Ph.D. in biological sciences.

EXTRACURRICULAR

RIT Bioinformatics Journal Club

August 2010 - May 2011

Rochester Institute of Technology, Rochester, NY

Founded journal club to foster collaboration and interaction between bioinformatics majors at RIT.

Applied Deep Learning Notes

January 2018 - Present

Communication of notes and code from analysis and implementation of deep learning techniques. The first of these communications, on generative adversarial networks (GANs), is available at <https://www.sciencesundries.com/sundries> with code at <https://github.com/cottersci/ModelTests>