

Miriam Shiffman

shiffman@mit.edu

Github:// [meereuum](#)

Twitter:// [@meereuum](#)

Website:// [mit.edu/~shiffman](#)

EDUCATION

MIT — 2016-present
Ph.D. candidate (GPA 5.0)
Computational & Systems Biology

RELEVANT COURSES:

Machine Learning
Bayesian Modeling & Inference
Fundamentals of Probability
Algorithms for Machine Learning
Picturing Science & Engineering
Advanced Computational Biology
Machine Learning for Healthcare
Deep Learning for Life Sciences
Frontiers in Health & Disease
Molecular Biology

Pomona College — 2009-2013
B.A., *Magna cum laude*
Molecular Biology

RELEVANT COURSES:

Linear Algebra • Probability
Deterministic Operations Research
Monte Carlo Methods • Statistics
Computation & Cognition
Foundations of Digital Design

EXPERIENCE

Ph.D. student, MIT CSAIL & Broad Institute
Cambridge, MA — 2016-present

- ⊆ Bayesian methods for learning from single-cell studies
- ⊆ Advised by [Tamara Broderick](#) and [Aviv Regev](#)

Research engineer intern, Fast Forward Labs
New York City — 2016

- ⊆ Published [series of whitepapers](#) for the FFL blog exploring the theory and applications of Variational Autoencoders (VAEs)
- ⊆ Developed open-source TensorFlow implementation: [vae-tf](#)

Programmer, Recurse Center
New York City — 2015

- ⊆ Received grant to pursue projects in machine learning, data visualization, and image processing

Fulbright scholar, Australian Centre for Ecogenomics,
University of Queensland
Brisbane, Australia — 2014-2015

- ⊆ Computational microbiology research, using next-generation DNA sequencing to study marsupial gut microbiomes, from community-level to individual population genomes
- ⊆ Led project to collect and analyze large-scale (10-100 Gb) metagenomic data using Python, command-line tools, and R
- ⊆ First-author publication, 2017

Cheese microbiologist, Cellars at Jasper Hill
Greensboro, VT — 2013

- ⊆ Inaugural microbiologist for an award-winning cheesemaker

Researcher, Pomona College
Claremont, CA — 2010-2013

- ⊆ Molecular genetics research, studying the evolution of selfish DNA (homing endonucleases) through bioinformatics and experiment
- ⊆ Senior research thesis; first-author manuscript

Intern, Harvard FAS Center for Systems Biology
Cambridge, MA — 2012

- ⊆ Studying cheese rind microbes as model microbial ecosystems
- ⊆ Poster presented at microbiology conference (ISME) in Seoul, 2014

PUBLICATIONS & MANUSCRIPTS

[Shiffman M](#), WT Stephenson, G Schiebinger, J Huggins, T Campbell, A Regev, and T Broderick (2018) Reconstructing probabilistic trees of cellular differentiation from single-cell RNA-seq data. [arXiv:1811.11790](#).

[Shiffman M](#), RM Soo, PG Dennis, M Morrison, GW Tyson, and P Hugenholtz (2017) Gene and genome-centric analyses of koala and wombat fecal microbiomes point to metabolic specialization for Eucalyptus digestion. *PeerJ* 5(e4075).

[Shiffman M](#) and BM Charalambous (2012) The search for archaeal pathogens. *Reviews in Medical Microbiology* 23(3): 45-51.

HONORS

NSF Graduate Research Fellowship
U.S. Fulbright Scholarship (Australia)

Bertsch Prize in Molecular Biology
Distinction in the Senior Thesis

Phi Beta Kappa

Sigma Xi

Jaeger Mathematics Prize

Shiffman M and LM Seligman (manuscript) Evolutionary dynamics of homing endonucleases: Characterizing related inteins in functionally distinct host genes in hyperthermophilic archaea. In prep.

RECENT PRESENTATIONS

Reconstructing probabilistic trees of cellular differentiation (work with: A Regev, T Broderick, et al)

[Talks]

- † Spotlight, *Learning Meaningful Representations of Life* workshop @ NeurIPS (2019) Vancouver.
- † Contributed, *Women in Computational Biology* (2019) HHMI Janelia, Ashburn, VA.
- † Contributed, *BNP12* (2019) Oxford. [ISBA travel award]
- † Invited, *Phyloseminar.org*: series on nontraditional data (2019).
- † Contributed, *Bayesian Nonparametrics* workshop @ NeurIPS (2018) Montréal.
- † Invited primer, *Models, Inference, and Algorithms* (2018) Broad Institute, Cambridge, MA.

[Poster #1]: *initial description of the model.*

- † *Approximate Bayesian Inference* and *Machine Learning in Computational Biology* workshops @ NeurIPS (2017) Long Beach, CA.
- † *Machine Learning Summer School* (2018) Universidad Torcuato Di Tella, Buenos Aires.
- † *Women in Data Science* conference (2018) Microsoft Research New England, Cambridge, MA.
- † Broad Institute annual retreat (2017) Boston, MA.
- † MIT Computational & Systems Biology retreat (2018) Kennebunkport, ME.

[Poster #2]: *development of inference, simulation experiments showing recovery of latent trajectories.*

- † *Bayesian Nonparametrics* workshop @ NeurIPS (2018) Montréal. [Google travel award]
- † *Women in Machine Learning* workshop (2018) Montréal.
- † MIT College of Computing launch (2019) Cambridge, MA.
- † Broad Institute annual retreat (2018) Boston, MA.

[Poster #3]: *creative probabilistic programming for biology.*

- † *Learning Meaningful Representations of Life* workshop @ NeurIPS (2019) Vancouver. [NIH travel award]

OTHER PROJECTS

lda2vec-tf: simultaneous inference of document, topic, and word embeddings via *lda2vec*, a hybrid of **latent Dirichlet allocation** and **word2vec**

- ↔ Ported the **original model** (in Chainer) to the first published version in TensorFlow
- ↔ Adapted to analyze 25,000 microbial genomes (80 million genes) to learn **microbial gene and genome embeddings**, and predict the function of non-annotated genes

k-meanz: image segmentation (TensorFlow)

- ↔ **Visualization** of clustering dynamics

Set solver: team project to solve the card game “Set”

- ↔ **Computer vision** algorithms to detect winning sets in photos of dealt cards
- ↔ Interactive Twitter bot: **@ProfessorSet**

Pangea3: interactive map (D3) of global continental drift and fossil paleolocation over geologic time (230 million years)

- ↔ Team hackathon at the American Museum of Natural History (covered by **NY Observer**)

cinematic / lsbeer / mit-lunch: command-line tools for everyday life

TEACHING EXPERIENCE

Teaching Assistant, MIT 7.03 (undergraduate genetics), with Profs. Aviv Regev & Peter Reddien — 2018