Jake Crawford

CONTACT
INFORMATION

Email: jjc2718@gmail.com
Web: http://jjc2718.github.io
GitHub: https://github.com/jjc2718

OVERVIEW

I am a third-year computational biology PhD student at Penn, working with Casey Greene. My current work primarily focuses on applying machine learning to genomic and transcriptomic data in cancer. More generally, I'm interested in ML/statistical inference for *-omics data, and in combining tabular data with structured biological priors such as networks and pathways to improve prediction and interpretability.

EDUCATION

University of Pennsylvania, Philadelphia, PA

Ph.D. Candidate, Genomics and Computational Biology
Started Fall 2018 | Advisor: Casey Greene

Tufts University, Medford, MA

M.S. in Computer Science, awarded Spring 2017 (Advisor: Lenore Cowen)
B.S. in Biology, awarded Spring 2012

EMPLOYMENT

Microsoft Research New England, Cambridge, MA   May 2017 - July 2018
Research Assistant

Wayfair, Boston, MA   January 2014 - August 2015
Software Engineer

PUBLICATIONS

M. Tomasoni, S. Gomez, **J. Crawford**, W. Zhang, S. Choobdar, D. Marbach, S. Bergmann. MONET: a toolbox integrating top-performing methods for network modularisation. *Bioinformatics*, 2020. (applications note)

J. Crawford, C.S. Greene. Incorporating biological structure into machine learning models in biomedicine. *Current Opinion in Biotechnology*, 2020. (review article)

S. Choobdar*, M. E. Ahsen*, **J. Crawford***, M. Tomasoni, D. Lamparter, J. Lin, B. Hescott, X. Hu, J. Mercer, T. Natoli, R. Narayan, The DREAM Module Identification Challenge Consortium, A. Subramanian, G. Stolovitzky, Z. Kutalik, K. Lage, D. Slonim, J. Saez-Rodriguez, L. J. Cowen, S. Bergmann, D. Marbach. Assessment of network module identification across complex diseases. *Nature Methods*, 2019.
(* = co-first author)

J. Listgarten, M. Weinstein, B. Kleinstiver, A. A. Sousa, J. K. Joung, **J. Crawford**, K. Gao, M. Elibol, L. Hoang, J. Doench, N. Fusi. Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering*, 2018.

PREPRINTS

J. Crawford, C.S. Greene. Graph biased feature selection of genes is better than random for many genes. *bioRxiv*. (no peer-reviewed journal submission currently planned for this work)

TALKS/POSTERS

(Lightning Talk/Poster) "Open Community Challenge Reveals Molecular Network Modules with Key Roles in Diseases." Intelligent Systems for Molecular Biology (ISMB), 2018.

(Talk) "Detangling PPI Networks to Uncover Functionally Meaningful Clusters." Workshop on Computational Network Biology: Modeling, Analysis and Control (CNB-MAC), 2017.

(Talk/Poster) "A Double Spectral Approach to Disease Module Identification." RE-COMB/ISCB Conference on Regulatory and Systems Genomics (RSG), 2016.

(Poster) "A Double Spectral Approach to Disease Module Identification." Microsoft Research Computational Aspects of Biological Information (CABI), 2016.

TEACHING EXPERIENCE	**Teaching Assistant**, Tufts University	
	COMP 167: Computational Biology	Spring 2017
	COMP 40: Machine Structure and Assembly Language	Fall 2015, Spring 2016

SKILLS AND INTERESTS

Research Interests:
- Bioinformatics/computational biology
- Application and development of ML/statistical methods
- Biological networks and network science

Programming Languages (from most to least comfortable):
Python, R, C/C++, Bash, JavaScript, PHP, Scheme

Software (in no particular order):
Vim, Git, LATEX, NumPy, SciPy, scikit-learn, pandas, PyTorch, Jupyter Lab

RELEVANT COURSEWORK

University of Pennsylvania:
- Machine Learning (CIS 520)
- Statistical Inference (STAT 431)
- Bayesian Modeling and Computation (STAT 927)
- Numerical Linear Algebra (AMCS 602)
- Signaling Pathways in Cancer (BMB 585)
- Regulation of the Genome (BIOM 555)

Tufts University:
- Probabilistic Systems Analysis (EE 104)
- Algorithms (COMP 160)

SERVICE AND SOCIETIES

GCB Advising Committee, student representative (2020-present)
Reviewer, *Bioinformatics* (2019-2020)
Reviewer, *Computational and Structural Biotechnology* (2020)
Member, International Society for Computational Biology (ISCB)
Member, Association for Computing Machinery (ACM)