MENGTING GU

Mountain View, CA • 203-397-6155 • gumengting333@gmail.com

EDUCATION

School of Arts & Sciences, Yale University 2013-2019
Ph.D, Computational Biology and Bioinformatics
Concentrations: Algorithms and Machine Learning, Functional Genomics

School of Engineering, Yale University 2014-2016
M.S., Computer Science, GPA 4.0

Tsinghua University 2009-2013
B.A., School of Economics and Management
B.S., School of Life Sciences and Biotechnology

The University of Hong Kong 2011-2012
Exchange

EXPERIENCES

Visa Research 06/2019 - Present
Staff Research Scientist *Palo Alto, CA*

· Develop graph neural network algorithms; led network modeling for transaction graphs
· Leverage natural language processing techniques for sequence modeling
· Improve Visa fraud detection system over the financial data warehouse at Visa
· Develop predictive models for essential parties in the payment and commerce system

Yale University 08/2013 - 05/2019
Graduate Student Researcher *New Haven, CT*

· Developed convolution-based method to detect gene regulators at genome-wide scale
· Software adopted by scientific research consortiums like ENCODE to analyze large-scale genomic data
· Co-first author manuscript published in ***Nature Methods***
· Using deep-learning models to predict psychiatric disease risk based on individual genome
· Co-first author manuscript published in ***Science***, front cover.

StatLab, Yale University 07/2015 - 07/2018
Consultant *New Haven, CT*

· Support all levels of data analysis at Yale University, including consultations on research plans, experiments designs, and data analysis strategies.
· Develop materials for and lead university-wide workshops like workshops on R, Python, Matlab, etc.

TEACHING

Biomedical Data Science, Mining and Modeling 2017 Spring
· Instructor: Dr. Mark Gerstein, Albert L. Williams Professor of Biomedical Informatics

Designing the Digital Economy 2016 Fall
· Instructor: Dr. E. Glen Weyl, Microsoft CTO Political Economist & Social Technologist

SELECTED PUBLICATIONS

A Sethi*, **M Gu***, E Gumusgoz, L Chan, KK Yan, J Rozowsky, I Barozzi, V Afzal, J Akiyama, I Plajzer-Frick, C Yan, C Novak, M Kato, T Garvin, Q Pham, A Harrington, B Mannion, E Lee, Y Fukuda-Yuzawa, A Visel, DE. Dickel, K Yip, R Sutton, L.A. Pennacchio and M Gerstein†. Supervised enhancer prediction with epigenetic pattern recognition and targeted validation. *Nature Methods* 17.8 (2020): 807-814

J Zhang*, D Lee*, V Dhiman*, P Jiang*, J Xu*, P McGillivray*, H Yang*, J Liu, W Meyerson, D Clarke, **M Gu**, S Li, S Lou, J Xu, L Lochovsky, M Ung, L Ma, S Yu, Q Cao, A Harmanci, KK Yan, A Sethi, G Gürsoy, M R Schoenberg, J Rozowsky, J Warrell, P Emani, Y T Yang, T Galeev, X Kong, S Liu, X Li, J Krishnan, Y Feng, J Rivera-Mulia, J Adrian, J R Broach, M Bolt, J Moran, D Fitzgerald, V Dileep, T Liu, S Mei, T Sasaki, C Trevilla-Garcia, S Wang, Y Wang, C Zang, D Wang, R J Klein, M Snyder, D M Gilbert, K Yip, C Cheng, F Yue, X S Liu, K P White, M Gerstein†. An integrative ENCODE resource for cancer genomics. *Nature Communications*

EYK Ho, Q Cao, **M Gu**, RWL Chan, Q Wu, M Gerstein, KY Yip. Shaping the nebulous enhancer in the era of high-throughput assays and genome editing. *Briefings in bioinformatics* 21.3 (2020): 836-850

FCP Navarro, H Mohsen, C Yan, S Li, **M Gu**, W Meyerson, M Gerstein. Genomics and data science: an application within an umbrella. *Genome biology* 20.1 (2019): 109

A Moro*, T P Driscoll*, L C Boraas, W Armero, D M Kasper, N Baeyens, C Jouy, V Mallikarjun, J Swift, S J Ahn, Donghoon Lee, J Zhang, **M Gu**, M Gerstein, M Schwartz, S Nicoli. MicroRNA-dependent regulation of biomechanical genes establishes tissue stiffness homeostasis. *Nature cell biology* 21.3 (2019): 348-358.

D Wang*, S Liu*, J Warrell*, H Won*, X Shi*, F Navarro*, D Clarke*, **M Gu***, P Emani*, M Xu, YT Yang, JJ Park, SK Rhie, K Manakongtreecheep, H Zhou, A Nathan, J Zhang, M Peters, E Mattei, D Fitzgerald, T Brunetti, J Moore, PsychENCODE Consortium‡, N Sestan, AE Jaffe, K White, Z Weng, DH Geschwind†, J Knowles†, M Gerstein†. Comprehensive functional genomic resource and integrative model for the adult brain. *Science* (2018), 362 (6420)

P McGillivray, D Clarke, W Meyerson, J Zhang, D Lee, **M Gu**, S Kumar, H Zhou, M Gerstein†. Network Analysis as a Grand Unifier in Biomedical Data Science. *Annual Review of Biomedical Data Science* (2018), 1: 153-180.

V Despic, M Dejung, **M Gu**, J Krishnan, J Zhang, L Herzel, K Straube, MB Gerstein, F Butter, KM Neugebauer†. Dynamic RNA-protein interactions underlie the zebrafish maternal-to-zygotic transition. *Genome Research* (2017), 27:1184-1194

L Guan, Q Yang, **M Gu**, L Chen, X Zhang†. Exon expression QTL (eeQTL) analysis highlights distant genomic variations associated with splicing regulation. *Quantitative Biology* (2014), 2(2):71-79.

* These authors contribute equally

ACADEMIC SERVICE

Program Committee:

Machine Learning in Computational Biology (MLCB) 2020
ICML 2020 Workshop on Computational Biology
International Workshop on Data Mining in Bioinformatics (BIOKDD) 2020
Grace Hopper Celebration (GHC 2020) Artificial Intelligence
ENCODE Consortium Enhancer Prediction Challenges Steering Committee

Reviewer:

IEEE SMARTCOMP 2019; PLOS One; Smart Health; PLOS Computational Biology

INVITED SEMINARS/ CONFERENCE PRESENTATIONS

NEC lab (Princeton) seminar	05/2019
University of Massachusetts Boston, Computer Science Department seminar	04/2019
IBM Research seminar (Thomas J. Watson Research Center, NY)	04/2019
Microsoft Research seminar (New England)	01/2019
ISCB conference on Regulatory & Systems Genomics (RSGDREAM 2018, New York, NY)	12/2018
CSH-Asia Systems Biology of Gene Regulation & Genome Editing (Suzhou, China)	10/2018
ENCODE Consortium Annual Meeting 2018 (Palo Alto, CA)	02/2018
American Society of Human Genetics Annual Meeting (ASHG 2017, Orlando, FL)	10/2017

SKILLS

Programming	Python, R, C, C++, Perl, Matlab, Mathematica
Operating Systems	Mac OSX, Linux and Window systems
Databases	MySQL, Pig, Hadoop, Spark
library and Tools	Numpy, Pandas, PyTorch, Tensorflow, Vim, Emacs
Languages	Chinese (Native), English (Professional working proficiency)