

# MENGTING GU

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

## EDUCATION

---

<b>School of Arts &amp; Sciences, Yale University</b> <i>Ph.D., Computational Biology and Bioinformatics</i> Concentrations: Machine Learning, Sequence Analysis, Network Analysis	2013-2019
<b>School of Engineering, Yale University</b> <i>M.S., Computer Science, GPA 4.0</i>	2014-2016
<b>Tsinghua University</b> <i>B.A., School of Economics and Management</i> <i>B.S., School of Life Sciences and Biotechnology</i>	2009-2013
<b>The University of Hong Kong</b> <i>Exchange</i>	2011-2012

## EXPERIENCES

---

<b>Visa Research</b> <i>Staff Research Scientist</i>	06/2019 - Present <i>Palo Alto, CA</i>
<ul style="list-style-type: none"><li>· Lead network modeling for transaction graphs; Develop novel graph neural network algorithms;</li><li>· Develop recurrent neural network structures and attention based models for transaction data modeling</li><li>· Improve Visa fraud detection system over the financial data warehouse at Visa</li><li>· Develop predictive models for essential parties in the payment and commerce system</li></ul>	
<b>Yale University</b> <i>Graduate Student Researcher</i>	08/2013 - 05/2019 <i>New Haven, CT</i>
<ul style="list-style-type: none"><li>· Developed convolution-based method to detect gene regulators at genome-wide scale</li><li>· Software adopted by scientific research consortiums like ENCODE to analyze large-scale genomic data</li><li>· Co-first author manuscript published in <i>Nature Methods</i></li><li>· Using deep-learning models to predict psychiatric disease risk based on individual genome</li><li>· Co-first author manuscript published in <i>Science</i>, front cover.</li></ul>	
<b>StatLab, Yale University</b> <i>Consultant</i>	07/2015 - 07/2018 <i>New Haven, CT</i>
<ul style="list-style-type: none"><li>· Support all levels of data analysis at Yale University, including consultations on research plans, experiments designs, and data analysis strategies.</li><li>· Develop materials for and lead university-wide workshops like workshops on R, Python, Matlab, etc.</li></ul>	

## TEACHING

---

<b>Biomedical Data Science, Mining and Modeling</b> · Instructor: Dr. Mark Gerstein, Albert L. Williams Professor of Biomedical Informatics	2017 Spring
<b>Designing the Digital Economy</b> · Instructor: Dr. E. Glen Weyl, Microsoft CTO Political Economist & Social Technologist	2016 Fall

## 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<sup>†</sup>. 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 Harman, 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<sup>†</sup>. 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<sup>†</sup>, N Sestan, AE Jaffe, K White, Z Weng, DH Geschwind<sup>†</sup>, J Knowles<sup>†</sup>, M Gerstein<sup>†</sup>. 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<sup>†</sup>. 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<sup>†</sup>. 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<sup>†</sup>. 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:

The Web Conference (WWW), 2021  
International Joint Conference on Artificial Intelligence (IJCAI), 2021  
International Conference on Machine Learning (ICML) WCB workshop, 2020  
International Workshop on Data Mining in Bioinformatics (BIOKDD), 2020  
Grace Hopper Celebration (GHC) Artificial Intelligence, 2020  
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

---

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