

# Daifeng Wang

Phone: (631) 638-1329

Email: [daifeng.wang@stonybrookmedicine.edu](mailto:daifeng.wang@stonybrookmedicine.edu)

Address: HSC L3-045H, 101 Nicolls Road, Stony Brook NY 11794

Website: <https://daifengwang.com>

## RESEARCH AREAS

Bioinformatics, Machine learning, Computational systems biology, Biomedical data science, Computational medicine, Network science

## PROFESSIONAL EXPERIENCE

August 2016 – Present, *Assistant Professor*, Department of Biomedical Informatics and *Faculty member* of Stony Brook Cancer Center, Stony Brook University

February 2017 – Present, *Affiliated Assistant Professor*, Departments of Electrical and Computer Engineering, Applied Mathematics and Statistics, and Computer Science, Stony Brook University

August 2015 – July 2016, *Associate Research Scientist*, Department of Molecular Biophysics and Biochemistry, Yale University

## EDUCATION AND TRAINING

*Postdoctoral Associate*, Computational Biology & Bioinformatics Program and Department of Molecular Biophysics and Biochemistry, Yale University, 2012-2015

Mentor: Dr. Mark B. Gerstein, Albert L. Williams Professor of Biomedical Informatics

*Ph.D.*, Electrical and Computer Engineering, The University of Texas at Austin, 2011

*Certificate*, Statistical Foundations, Division of Statistics and Scientific Computation, The University of Texas at Austin, 2011

*Master of Science*, Electrical and Computer Engineering, The University of Texas at Austin, 2006

*Bachelor of Science*, Electronics and Information Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China, 2004

## RESEARCH GRANTS

- *Identification of gene regulatory networks for direct conversion of fibroblasts into bladder epithelia*, Interdisciplinary pilot grant of Cancer Center, College of Engineering and Applied Sciences at Stony Brook University, PI with Co-PI Dr. Flaminia Talos, 07/2017 – 06/2019, Success rate 20%
- *Large-Scale Comparative Regulatory Network Analysis in Photosynthetic Organisms*, Stony Brook University – Brookhaven National Laboratory seed grant, PI with Co-PI Dr. Ian Blaby, 07/2017 – 12/2018, Success rate 13%
- *Discovery and Validation of Neuronal Enhancers Associated with the Development of Psychiatric Disorders*, NIMH 1U01MH116492, Co-I (Site Project Director), 07/2018 – 06/2023, PI: Dr. Zhiping Weng, UMass Medical School

## PEER-REVIEWED PUBLICATIONS (\*co-first-author, equal contribution; #corresponding author)

- Adam P Arkin, Robert W Cottingham, Christopher S Henry, Nomi L Harris, Rick L Stevens, Sergei Maslov, ..., **Daifeng Wang**, Fangfang Xia, Hyunseung Yoo, Shinjae Yoo, Dantong Yu, KBase: The United States Department of Energy Systems Biology Knowledgebase, *Nature Biotechnology*, 36(7):566-569, 2018
- Zongdong Li, Natasha M. Nesbitt, Lisa E. Malone, Dimitri V. Gnatenko, Song Wu, **Daifeng Wang**, Wei Zhu, Geoffrey D. Girnun, Wadie F. Bahou, Heme degradation enzyme biliverdin IXB reductase is required for stem cell glutamine metabolism, *Biochemical Journal*, 475(6):1211-1223, 2018

- Pedro Alves, Shuang Liu, **Daifeng Wang**<sup>#</sup> and Mark Gerstein<sup>#</sup>, Multiple-Swarm Ensembles: Improving the Predictive Power and Robustness of Predictive Models and Its Use in Computational Biology, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(3):926-933, 2018
- **Daifeng Wang**<sup>#</sup>, John D. Haley and Patricia Thompson<sup>#</sup>, Comparative gene co-expression network analysis of epithelial to mesenchymal transition reveals lung cancer progression stages, *BMC Cancer*, 17:830, 2017
- **Daifeng Wang**, Fei He, Sergei Maslov, Mark Gerstein, DREISS: Using state-space models to infer the dynamics of gene expression driven by external and internal regulatory networks, *PLoS Computational Biology*, 12(10): e1005146, 2016
- **Daifeng Wang**, Koon-Kiu Yan, Joel Rozowsky, Eric Pan, Mark Gerstein, Temporal dynamics of collaborative networks driven by large scientific consortia, *Trends in Genetics*, 32, 251-253, 2016
- Koon-Kiu Yan\*, **Daifeng Wang**\*, Anurag Sethi, Robert Kitchen, Paul Muir, Chao Cheng, Mark Gerstein, Cross-Disciplinary Network Comparison: Matchmaking between Hairballs, *Cell Systems*, 2(3):147-157, 2016 (featured front-matter)
- Paul Muir, Shantao Li, Shaoke Lou, **Daifeng Wang**, Daniel Spakowicz, Leonidas Salichos, Jing Zhang, Farren Isaacs, George M. Weinstock, Joel Rozowsky, Mark Gerstein, The real cost of sequencing: scaling computation to keep pace with data generation, *Genome Biology*, 17:53, 2016
- Fei He, Shinjae Yoo, **Daifeng Wang**, Sunita Kumari, Mark Gerstein, Doreen Ware, Sergei Maslov, Large-scale atlas of microarray data reveals biological landscape of gene expression in Arabidopsis, *The Plant Journal*, 86(6), 472-480, 2016
- The PsychENCODE Consortium including **Daifeng Wang**, The PsychENCODE Project Consortium, *Nature Neuroscience*, 18, 1707-1712, 2015
- **Daifeng Wang**, Koon-Kiu Yan, Cristina Sisu, Chao Cheng, Joel Rozowsky, William Meyerson, Mark Gerstein, Logic: A method to characterize the cooperative logic of regulatory factors, *PLoS Computational Biology* 11(4): e1004132, 2015 (featured article)
- Chao Cheng, Erik Andrews, Koon-Kiu Yan, Matthew Ung, **Daifeng Wang**, Mark Gerstein, An Approach for Determining and Measuring Network Hierarchy: Application to Comparing the Phosphorylome and the Regulome, *Genome Biology*, 16:63, 2015
- Shuang Liu, Anjali Datta, Derek Ho, Jordan Dwelle, **Daifeng Wang**, Thomas E. Milner, H. Grady Rylander III, Mia K. Markey, Effect of image registration on longitudinal analysis of retinal nerve fiber layer thickness of non-human primates using Optical Coherence Tomography (OCT), *BMC Eye and Vision*, 2:3, 2015
- Mark Gerstein\*, Joel Rozowsky\*, Koon-Kiu Yan\*, **Daifeng Wang**\*, Chao Cheng\*, ..., Steven Brenner, Brenton Graveley, Susan Celniker, Thomas Gingeras, and Robert Waterston, Comparative Analysis of the Transcriptome across Distant Species, *Nature* 512, 445–448, 2014 (News & Views, Nature)
- Alan P. Boyle\*, Carlos L. Araya\*, ..., **Daifeng Wang**, ..., Robert H. Waterston, Mark Gerstein, Kevin P. White, Manolis Kellis, Michael Snyder, Comparative analysis of regulatory information and circuits across diverse species, *Nature* 512, 453–456, 2014 (News & Views, Nature)
- Koon-Kiu Yan\*, **Daifeng Wang**\*, Joel Rozowsky, Henry Zheng, Mark Gerstein, OrthoClust: An orthology-based network framework for expression clustering across multiple species, *Genome Biology*, 15:R100, 2014
- Cristina Sisu, Baikang Pei, ..., Rachel Harte, **Daifeng Wang**, Michael Rutenberg Schoenberg, Wyatt Clark, Mark Diekhans, Joel Rozowsky, Tim Hubbard, Jennifer Harrow, Mark Gerstein, Comparative analysis of pseudogenes across three phyla, *Proceedings of the National Academy of Sciences* (PNAS), vol. 111, no. 37, pp. 13361–13366, 2014
- **Daifeng Wang**, Mia K. Markey, Claus O. Wilke and Ari Arapostathis, Eigen-genomic System Dynamic-pattern Analysis (ESDA): Modeling mRNA degradation and self-regulation, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 9, no. 2, pp. 430-437, 2012

- **Daifeng Wang**, Ari Arapostathis, Claus O. Wilke, and Mia K. Markey, Principal-Oscillation-Pattern Analysis of Gene Expression, *PLoS ONE* 7(1): e28805, 2012
- Sean X. Shi, Anand Ramalingam, **Daifeng Wang**, and David Z. Pan, Latch Modeling for Statistical Timing Analysis, Proceedings of the IEEE Conference on Design, Automation and Test in Europe 2008
- Wei Wu, **Daifeng Wang**, Ari Arapostathis and K. Davey, Optimal Power Generation Scheduling of a Shipboard Power System, Proceedings of the IEEE Electric Ship Technologies Symposium 2007
- **Daifeng Wang** and Brian L. Evans, Codebook Design for Noncoherent MIMO Communications Via Reflection Matrices, Proceedings of the IEEE Global Telecommunications Conference 2006

## BOOK REVIEWS AND CHAPTERS

- **Daifeng Wang**, Systems Biology: Constraint-Based Reconstruction and Analysis by Bernhard O. Palsson, The Quarterly Review of Biology, 92:3, 303-304, 2017
- **Daifeng Wang**, Epigenomics in Health and Disease by Mario Fraga and Agustin F. Fernandez, The Quarterly Review of Biology, 93:2, 163-164, 2018
- **Daifeng Wang**, Chao Cheng, Chapter 27: Genomics and Systems Biology, *Cooperative and Graph Signal Processing*, Editors: Petar M. Djuric and Cedric Richard, Elsevier, 2018, ISBN: 978-0-12-813677-5

## HONORS AND AWARDS

- Faculty of 1000 Best Poster Presentation Award, 10<sup>th</sup> Great Lakes Bioinformatics Conference, International Society for Computational Biology 2015
- The Extreme Science and Engineering Discovery Environment Startup Allocation Grant 10/2015
- Graduate Student Professional Development Award, The University of Texas at Austin 2009
- Nominated for Texas Exes Teaching Awards to Teaching Assistant 2008
- Graduate Student Travel Award, IEEE GLOBECOM 2006
- Advanced Class of Elite Range (ACER), Huazhong University of Science and Technology
- The Top Academic Outstanding Student, Huazhong University of Science and Technology

## MEDIA COVERAGE

- “Shared genetics in humans, worms and flies”, [Yale Daily News](#), 2014, personal Interview
- “Evolution used similar molecular toolkits to shape flies, worms, and humans”, [Yale News](#), 2014
- “Tiny, Vast Windows Into Human DNA”, [The New York Times](#), September 2, 2014
- “Scientists looking across human, fly and worm genomes find shared biology”, [National Institutes of Health \(NIH\)](#), 2014

## SELECTED RESEARCH EXPERIENCE

August 2016 – Present: *Assistant Professor*, Stony Brook University

- Interpretable machine learning models for single-cell transcriptomic deconvolution and genotype-phenotype prediction to reveal molecular mechanisms and network architectures in brain disorders (PsychENCODE)
- Translational bioinformatics and computational medicine
- Systems modeling, prediction and analysis of gene regulatory networks and circuits with applications to developmental biology, cancer, neuroscience and bioenergy

January 2012 – July 2016: *Postdoctoral Associate* and *Associate Research Scientist*, Yale University

- Bioinformatics, genome informatics and computational systems biology in the projects of Encyclopedia of DNA Elements (ENCODE/modENCODE/PsychENCODE), Extracellular RNA Communication (ERC), The Cancer Genome Atlas (TCGA), Pan-Cancer analysis of Whole Genomes (PCAWG) and 1000 Genomes
- Plant gene network analysis as leading developer in DOE Systems Biology Knowledgebase (KBase)

- Academic social networks driven by large scientific consortia (Big Science)

December 2007 – December 2011: *Graduate Research Assistant*, University of Texas at Austin.

- Principal dynamic characteristics of high dimensional temporal gene expression data
- Functional annotations & pathway analysis of mouse brain neuro-genomic data, in collaboration with the Waggoner Center for Alcohol and Addiction Research, UT-Austin
- Clinical decision support systems about tissue expansion implant breast reconstruction

## **SELECTED TEACHING EXPERIENCE**

*Assistant Professor*, Stony Brook University

- BMI 511 Translational Bioinformatics – modeling, mining and learning in biomedical data science (Fall 2017)
- CSE 523/524 Advanced Project in Computer Science I/II (Faculty advisor)
- CSE 393 Introduction to Biomedical Informatics (Spring 2018, guest lecturer)
- PhD students: Nam Nguyen (Computer Science, 07/2017 –), Ting Jin (Biomedical Informatics, 08/2017 –)
- Masters students: Sanjeevani Choudhery (Electrical & Computer Engineering, 01/2018 –), Neha Mane and Rahul Bhansali (Computer Science, 01/2018 –), Alisha Kamat (Computer Science, 01-12/2017)
- Rotation: Manojit Mosur Swamynathan (Molecular and Cellular Biology, 01-04/2017)
- Undergraduate: So Yeon Min (MIT EECS, funded by Google Summer of Code 2017)
- Dissertation committee: Dr. Cheng Chang (Electrical & Computer Engineering, 03/2018)

*Graduate Teaching Assistant*, University of Texas at Austin

- Undergraduate Courses: EE331 Electric Circuits, Electronics and Machinery, EE411 Circuit Theory, EE445S Real-Time Digital Signal Processing Lab
- Graduate Courses: EE382V-3 VLSI Communication/Digital Signal Processing Systems (this course was designed for working professionals of high-tech companies in Austin, TX).

## **SELECTED WORK EXPERIENCE**

- June-August 2008: *Graduate Research Intern*, Qualcomm Research Center, San Diego, CA.
- June 2007-January 2008: *Software Engineer Co-op*, Cisco Systems, Austin, TX.
- May-August 2005: *Software Engineer Intern*, National Instruments, Austin, TX.

## **INVITED TALKS AND ORAL PRESENTATIONS**

- “Comparative gene network analysis of epithelial to mesenchymal transition reveals lung cancer progression stages”, International Conference on Intelligent Biology & Medicine, Los Angeles, CA, Jun 2018
- “Using state-space models to infer the dynamics of gene expression driven by external and internal regulatory networks”, Systems Genomics Workshop of the International Plant and Animal Genome Conference XXV, San Diego, CA, January 2017
- NIMH PsychENCODE Workshop, Society for Neuroscience conference, San Diego, CA, November 2016
- “Temporal dynamics of collaborative networks driven by large scientific consortia”, Intelligent Systems for Molecular Biology (ISMB), Orlando, FL, July 2016 (Highlights talk)
- “Systematic multi-scale modeling and analysis for gene regulation”, 1st SysMod SIG meeting on Computational modeling of biological systems, Intelligent Systems for Molecular Biology (ISMB), Orlando, FL, July 2016
- “DREISS: dynamics of gene expression driven by external and internal regulatory networks based on state space model”, the eighth Annual RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges, Philadelphia, PA, November 2015

- “Using state-space models to infer the dynamics of gene expression driven by external and internal regulatory networks”, Yale Institute for Network Science and Kavli Institute for Neuroscience, Yale University, November 2015
- “Loregic: A method to characterize the cooperative logic of regulatory factors”, 10<sup>th</sup> Great Lakes Bioinformatics Conference, International Society for Computational Biology, Purdue University, May 2015
- NHGRI ENCODE consortium meeting, Boston, Massachusetts, 2012
- “Development of a Bayesian Network to Model Decisions about Tissue Expansion Implant Breast Reconstruction”, 3<sup>rd</sup> Annual International Conference in Computational Surgery, Houston, Texas, 2011

## **POSTER PRESENTATIONS AND ABSTRACTS**

- Sanjeev Choudhery (Master’s research project), “Machine learning analysis reveals functional developmental and predictive gene modules associated with cortical thickness changes for Autism Spectrum Disorder (ASD)” on Society of Biological Psychiatry Annual Meeting, New York City, 2018
- Computational Aspects of Biological Information 2016 at Microsoft Research New England, Cambridge, MA
- The New York Area Meeting in Quantitative Biology, Computational and Statistical Genomics: Applications to Disease, Cold Spring Harbor Laboratory, New York, 2016
- Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, New York, 2016
- The eighth Annual RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges, Philadelphia, Pennsylvania, November 15-18, 2015
- The Great Lakes Bioinformatics Conference, West Lafayette, Indiana, May 18-20, 2015
- The seventh Annual RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges and Cytoscape Workshops, San Diego, California, 2014
- Biological Data Science, Cold Spring Harbor Laboratory, New York, 2014
- 18<sup>th</sup> Annual International Conference on Research in Computational Molecular Biology (RECOMB 2014), Pittsburgh, Pennsylvania, 2014
- Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, New York, 2014
- CSHL Genome Informatics 2013, Cold Spring Harbor Laboratory, New York, 2013
- 4<sup>th</sup> ACM Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB 2013), Washington DC, 2013
- Plant Biology 2013, American Society of Plant Biologists, Providence, Rhode Island, 2013
- The Biology of Genomes 2013, Cold Spring Harbor Laboratory, New York, 2013
- Department of Energy Joint Genome Institute User Meeting, Walnut Creek, California, 2013
- ENCODE consortium meeting, Boston, Massachusetts, 2012
- American Medical Informatics Association Annual Symposium, Washington D.C., 2011
- Biomedical Engineering Society (BMES) 2011 Annual Meeting, Hartford, CT, 2011
- The third Annual International Conference in Computational Surgery, Houston, Texas, 2011

## **PEER-REVIEW EXPERIENCE**

Proposals: NIH/NHLBI P01 (ad hoc)

Journals and conferences: PLoS Genetics, Scientific Reports, PLoS One, Information Sciences, Evolutionary Bioinformatics, Algorithms for Molecular Biology, Cancer Management and Research, Cancer Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Journal of Bioinformatics and Computational Biology, 2005 IEEE Signal Processing Systems Workshop, IEEE Transactions on Signal Processing, IEEE International Symposium on Circuits and Systems, RECOMB 2009