

Daifeng Wang

Phone: (608) 262-8567

Office: Room 517, 1500 Highland Ave, Madison, WI 53705

Email: daifeng.wang@wisc.edu

Website: <http://daifengwanglab.org>

SUMMARY

My research focuses on developing interpretable machine learning approaches and bioinformatics tools for understanding functional genomics and gene regulation in complex biological systems.

PROFESSIONAL EXPERIENCE

- **Assistant Professor**, Department of Biostatistics and Medical Informatics and Department of Computer Sciences (affiliate), University of Wisconsin – Madison, 07/2019 – present
- **Investigator**, Waisman Center, University of Wisconsin – Madison, 07/2019 – present
- **Assistant Professor**, Department of Biomedical Informatics, Stony Brook University, 08/2016 – 06/2019
- **Associate Research Scientist**, Department of Molecular Biophysics and Biochemistry, Yale University, 08/2015 – 07/2016

EDUCATION AND TRAINING

- **Postdoctoral Associate**, Computational Biology and Bioinformatics Program and Department of Molecular Biophysics and Biochemistry, Yale University, 01/2012 – 07/2015
- **Ph.D.**, Electrical and Computer Engineering, The University of Texas at Austin, 12/2011
- **Master of Science**, Electrical and Computer Engineering, The University of Texas at Austin, 06/2006
- **Bachelor of Science**, Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, 06/2004

RESEARCH GRANTS

- **Project Title: Understanding the molecular mechanisms that contribute to neuropsychiatric symptoms in Alzheimer Disease**
Agency and Project ID: NIH/NIA R01AG067025
Role: MPI
Project period: 09/2019 – 08/2024
- **Project Title: A new bladder cancer model based on tissue reprogramming and gene targeting**
Agency and Project ID: NIH/NCI R21CA237955
Role: MPI
Project period: 01/2020 – 12/2021
- **Project Title: Discovery and Validation of Neuronal Enhancers Associated with the Development of Psychiatric Disorders**
Agency and Project ID: NIH/NIMH U01MH116492
Role: Co-Investigator
Project period: 04/2018 – 03/2023
- **Project Title: Quantitative Plant Science Initiative**

Agency and Project ID: DOE 0000250595

Role: Co-PI

Project period: 02/2020 – 08/2023

PUBLICATIONS

(*co-first-author, equal contribution; #co-corresponding author; last author as corresponding author)

Revision/Under Review/Preprint:

- Mufang Ying, Peter Rehani, Panos Roussos, **Daifeng Wang**, Integrative multi-omics analyses identify cell-type disease genes and regulatory networks across schizophrenia and Alzheimer's disease, under review, [Biorxiv](#)
- Yathindar Giffin-Rao, Bennett Strand, Leslie Huang, Margaret Medo, Bradley Levesque, Roger A Daley Jr., Lindsey Amundson, Aratrika Keshan, Rebecca Reese, **Daifeng Wang**, Su-Chun Zhang, Anita Bhattacharyya, Altered patterning of interneuron progenitors in Down syndrome, [Biorxiv](#)
- Kunling Huang, Yuchang Wu, Junha Shin, Ye Zheng, Alireza Fotuhi Siahpirani, Yupei Lin, Zheng Ni, Jiawen Chen, Jing You, Sunduz Keles, **Daifeng Wang**, Sushmita Roy, Qiongshi Lu, Transcriptome-wide transmission disequilibrium analysis identifies novel risk genes for autism spectrum disorder, [Biorxiv](#)

Published:

- Nam D Nguyen, Ting Jin, **Daifeng Wang**, Varmole: A biologically drop-connect deep neural network model for prioritizing disease risk variants and genes, btaa866, *Bioinformatics*
- Ting Jin, Nam D Nguyen, Flaminia Talos, **Daifeng Wang**, ECMarker: Interpretable machine learning model identifies gene expression biomarkers predicting clinical outcomes and reveals molecular mechanisms of human disease in early stages, accepted, *Bioinformatics*
- Yu Gao, Minjie Shen, Jose Carlos Gonzalez, Qiping Dong, Sudharsan Kannan, Johnson Hoang, Brian E. Eisinger, Qiang Chang, **Daifeng Wang**[#], Linda Overstreet-Wadiche, Xinyu Zhao, RGS6 mediates effects of voluntary running on adult hippocampal neurogenesis, *Cell Reports*, 32(5):107997, 2020
- Koon-Kiu Yan*, **Daifeng Wang***, Kun Xiong, Mark Gerstein, Comparing technological development and biological evolution from a network perspective, *Cell Systems*, 10(3), P219-222, 2020
- Jing Zhang et al. including **Daifeng Wang**, An integrative ENCODE resource for cancer genomics, *Nature Communications*, 11, 3696, 2020
- Nam D Nguyen, **Daifeng Wang**, Multi-view learning for understanding functional multiomics, *PLoS Computational Biology*, 16(4): e1007677, 2020
- Nam Nguyen, Ian K. Blaby, **Daifeng Wang**, ManiNetCluster: A Manifold Learning Approach to Reveal the Functional Linkages Across Multiple Gene Networks, *BMC Genomics*, 20, 1003, 2019
- John A Haley, Christian F Ruiz, Emily D Montal, **Daifeng Wang**, John D Haley, Geoffrey D Girnun, Decoupling of Nrf2 Expression Promotes Mesenchymal State Maintenance in Non-Small Cell Lung Cancer, *Cancers*, 2019, 11(10), 1488
- Mohamed Salama, Mengling Liu, Christopher J. Clarke, Mel Pilar Espailat, John Haley, Ting Jin, **Daifeng Wang**, Lina M. Obeid, Yusuf A. Hannun, PKC α is required for Akt-mTORC1 activation in Non-Small Cell Lung Carcinoma (NSCLC) with EGFR mutation, *Nature Oncogene*, 2019

- **Daifeng Wang***, Shuang Liu*, ..., PsychENCODE Consortium, Nenad Sestan, Andrew E. Jaffe, Kevin White, Zhiping Weng, Daniel H. Geschwind, James Knowles, Mark Gerstein, Comprehensive functional genomic resource and integrative model for the human brain, **Science**, 362, 1266, 2018
- Michael J Gandal, Pan Zhang, ..., **Daifeng Wang**, ..., PsychENCODE Consortium, ..., Daniel H. Geschwind, Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder, **Science**, 362, 1265, 2018
- Mingfeng Li*, Gabriel Santpere*, ..., **Daifeng Wang**, ..., BrainSpan Consortium, PsychENCODE Consortium: Developmental Subgroup, ..., Nenad Sestan, Integrative Functional Genomic Analysis of Human Brain Development and Neuropsychiatric Risk, **Science**, 362, 1264, 2018
- PsychENCODE consortium including **Daifeng Wang**, Revealing the brain's molecular architecture, **Science**, 2018 Dec 14;362(6420):1262-1263
- Adam P Arkin, Robert W Cottingham, ..., **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**[#] and Mark Gerstein[#], 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**[#], John D. Haley and Patricia Thompson[#], 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
- 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, Loregic: A method to characterize the cooperative logic of regulatory factors, *PLoS Computational Biology* 11(4): e1004132, 2015
- 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
- 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
- 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

- Justin Snider, **Daifeng Wang**, Daniel F. Bogenhagen, John D. Haley, Pulse SILAC Approaches to the Measurement of Cellular Dynamics, Advances in Experimental Medicine and Biology book series (volume 1140), pp 575-583

- **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

SOFTWARE AND COMPUTATIONAL INFRASTRUCTURE

- Data science core in Waisman Center for bioinformatics and biostatistical analyses with AWS cloud computing and high performance computing platforms, <https://www.dropbox.com/s/o8owobhul62xkqu/WaismanCenter-AWS.pdf?dl=0>
- ManiNetCluster, R package for revealing the functional links between gene networks using manifold learning, <https://github.com/daifengwanglab/ManiNetCluster>
- ECMarker, Python package (Pytorch), an interpretable neural network model for identifying gene expression biomarkers predicting clinical phenotypes and outcomes and revealing underlying gene regulatory mechanisms, <https://github.com/daifengwanglab/ECMarker>
- Varmole, Python package (Pytorch), a biologically drop-connect deep neural network model for prioritizing disease risk variants and genes, <https://github.com/daifengwanglab/Varmole>
- scGRN, R package, a computational pipeline for predicting gene regulatory networks via multi-omics data (e.g., cell-type gene regulatory networks), <https://github.com/daifengwanglab/scGRN>

RESEARCH EXPERIENCE

- **Assistant Professor**, University of Wisconsin – Madison, 07/2019 – present
 - Interpretable machine learning models for understanding functional genomics and gene regulation in complex biological systems, especially at the single cell level
 - Neurodevelopmental, neurodegenerative and neuropsychiatric diseases
 - Data science core for the Waisman Center
- **Assistant Professor**, Stony Brook University, 08/2016 – 06/2019
 - Systems identification of gene regulatory networks in human diseases (cancers, brain disorders) by machine learning and network science approaches
 - Co-director of Biostatistics & Bioinformatics Shared Resource in Stony Brook Cancer Center
- **Postdoctoral Associate and Associate Research Scientist**, Yale University, 01/2012 – 07/2016
 - Bioinformatics and Functional genomics in the projects of Encyclopedia of DNA Elements (ENCODE/modENCODE), Extracellular RNA Communication (ERC), Pan-Cancer analysis of Whole Genomes (PCAWG) and 1000 Genomes
 - Gene network analysis as leading developer in DOE Systems Biology Knowledgebase
 - Academic social networks driven by large scientific consortia (Big Science)
- **Graduate Research Assistant**, University of Texas at Austin, 12/2006 – 12/2011
 - 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, in collaboration with MD Anderson Cancer Center

TEACHING EXPERIENCE

- **Assistant Professor**, University of Wisconsin – Madison, July 2019 – present
 - CS 540 Introduction to Artificial Intelligence (Fall 2020, co-teaching with Drs. Yin Li and Anthony Gitter)
 - STAT/BMI 877 Statistical Methods for Molecular Biology (Fall 2020, guest lecture)
 - BMI/CS 776 Advanced Bioinformatics (Spring 2020)
- **Assistant Professor**, Stony Brook University, August 2016 – June 2019
 - BMI 511 Translational Bioinformatics (Fall 2017, Fall 2018)
 - CSE 523/524 Advanced Project in Computer Science I/II (Faculty advisor)
 - CSE 393 Introduction to Biomedical Informatics (Spring 2018, guest lecturer)

HONORS AND AWARDS

- Best Poster Award, ACM Conference on Bioinformatics, Computational Biology & Health Informatics 2018
- Faculty of 1000 Best Poster Presentation Award, 10th Great Lakes Bioinformatics Conference, International Society for Computational Biology 2015
- NSF 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

PEER-REVIEW EXPERIENCE

- **Grants:** NIH P01 (ad hoc 2018, special emphasis panelist 2019), NSF (virtual panelist 2019), AHA Data Science (2019, 2020), UW Health Animal (2020)
- **Journals:** Genome Biology, Patterns (Cell), Genome Medicine, PLoS Genetics, PLoS Computational Biology, Genomics, Proteomics & Bioinformatics, Schizophrenia Research, Nucleic Acids Research, Oxford Bioinformatics, Cancer Research, Molecular Cancer Research, Cancers, Journal of Visualized Experiments, Scientific Reports, PLoS One, Information Sciences, Evolutionary Bioinformatics, BMC Bioinformatics, Algorithms for Molecular Biology, Cancer Management and Research, Cancer Informatics, IEEE/ACM TCBB

PROFESSIONAL MEMBERSHIPS

- International Society for Computational Biology
- Institute of Electrical and Electronics Engineers
- Society for Neuroscience

INVITED TALKS AND ORAL PRESENTATIONS

- Annual RECOMB Conference on Regulatory and Systems Genomics, with DREAM Challenges, 2020
- Biomedical Data Science Visit Day, University of Wisconsin – Madison, Madison, WI, 02/2020
- BMI student training seminar, University of Wisconsin – Madison, Madison, WI, 12/2019
- Gene discovery session, Directors meeting of Eunice Kennedy Shriver Intellectual & Developmental Disabilities Research Centers, NICHD, Seattle, WA, 11/2019
- Board of Visitors, Waisman Center, University of Wisconsin – Madison, Madison, WI, 10/2019
- PsychENCODE symposium, 2019 World Congress of Psychiatric Genetics, Anaheim, CA, 10/2019
- NIMH PsychENCODE Consortium In-Person Workshop, University of Chicago, Chicago, IL, 10/2019
- Computation and Informatics in Biology and Medicine (CIBM) Seminar, University of Wisconsin – Madison, Madison, WI, 10/2019
- Institute of Brain Science (IOBS), Fudan University, Shanghai, China, 09/2019
- College of Computer Science, Central South University, Changsha, Hunan, China, 09/2019
- International Conference on Intelligent Biology and Medicine (ICIBM 2019), Columbus, OH, 06/2019
- Network Biology 2019 at Cold Spring Harbor Laboratory, NY, 03/2019
- American College of Neuropsychopharmacology Annual Meeting, Hollywood, FL, 12/2018
- American Society of Human Genetics Annual Meeting, San Diego, CA, 10/2018
- International Conference on Intelligent Biology & Medicine, Los Angeles, 06/2018
- NIMH PsychENCODE Consortium In-Person Workshop, Yale University, New Haven, CT, 05/2018
- “Using state-space models to infer the dynamics of gene expression driven by external and internal regulatory networks”, International Plant & Animal Genome Conference, San Diego, CA, 01/2017
- NIMH PsychENCODE Workshop, Society for Neuroscience conference, San Diego, CA, 11/2016
- Intelligent Systems for Molecular Biology (ISMB), Orlando, FL, 07/2016 (Highlights talk)
- 1st SysMod SIG meeting, Intelligent Systems for Molecular Biology (ISMB), Orlando, FL, 07/2016
- Annual RECOMB Conference on Regulatory and Systems Genomics, with DREAM Challenges, Philadelphia, PA, 11/2015
- Yale Institute for Network Science and Kavli Institute for Neuroscience, Yale University, 11/2015
- 10th Great Lakes Bioinformatics Conference, International Society for Computational Biology, Indiana, 05/2015
- NHGRI ENCODE consortium meeting, Boston, Massachusetts, 2012
- 3rd Annual International Conference in Computational Surgery, Houston, Texas, 2011

POSTER PRESENTATIONS AND ABSTRACTS (recent five years)

- Alzheimer’s Association International Conference Neuroscience Next, 2020
- Annual RECOMB Conference on Regulatory and Systems Genomics, 2020
- International Conference on Intelligent Systems for Molecular Biology (ISMB), 2020
- “Computational Identification of Disease Genes at the Cellular Resolution for Alzheimer’s Disease and Schizophrenia”, Alzheimer’s Disease and Related Disorders Research Day, 2020 (Best Undergraduate Student Poster by Peter Rehani)
- ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB), 2018

- “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, 2018 (Sanjeev Choudhery, Master’s research project)
- Computational Aspects of Biological Information 2016 at Microsoft Research New England
- The New York Area Meeting in Quantitative Biology, Computational and Statistical Genomics: Applications to Disease, Cold Spring Harbor Laboratory, 2016
- Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, 2016
- The eighth Annual RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges, 2015
- The Great Lakes Bioinformatics Conference, 2015