

CONTENTS

1	Background	3
1.1	Education	3
1.2	Employment History	3
1.3	Awards and Honors	4
1.3.1	Awards to Advised Students	4
1.4	Professional Memberships	4
2	Scholarship	4
2.1	Journal Publications	4
2.2	Book Chapters (refereed)	7
2.3	Conference Proceedings (refereed)	7
2.4	Conference Workshops (refereed)	7
2.5	Scholarship in progress	8
2.6	Patents & Inventions	8
2.6.1	Patents	8
2.6.2	Invention Disclosures	8
2.7	Presentations	8
2.7.1	Invited Presentations	8
2.7.2	Research/Contributed Presentations	9
2.7.3	Poster Presentations	10
2.8	Press	10
3	Teaching	10
3.1	Courses Developed at UMass Amherst	10
3.2	Courses Taught at UMass Amherst	10
3.3	Independent Study Courses Taught at UMass Amherst	10
3.4	Courses Developed at WPI	10
3.5	Courses Taught at WPI	11
3.6	Independent Study Courses Taught at WPI	11
3.7	Guest Lectures	11
3.8	Other Teaching Experience	11
3.9	Students Supervised	11
3.9.1	Postdoctoral Scholars	11
3.9.2	PhD Students	11
3.9.3	MS Students	12
3.9.4	Senior Honors Thesis and Major Qualifying Project Students	12
3.9.5	Undergraduate Research Students	12
3.9.6	Academic Advising	12
4	Sponsored Research	13
4.1	Active Support	13
4.2	Completed Support	13
5	Service	14
5.1	Department Service at UMass Amherst	14
5.2	College Service at UMass Amherst	14
5.3	Thesis Committees	15
5.3.1	PhD	15

5.3.2	MS	15
5.3.3	Undergraduate Honors	16
5.4	Professional Service	16
5.4.1	Leadership	16
5.4.2	Peer Review	16
5.4.3	Grant Review	16
5.5	K-12 Outreach Activities	16
6	Professional Development	17
7	Consulting	17

Patrick J. Flaherty

Department of Mathematics & Statistics
University of Massachusetts, Amherst
Lederle Graduate Research Tower Rm 1444
710 N. Pleasant St., Amherst MA 01003
Phone: (413) 545-1796
flaherty@math.umass.edu
<http://people.math.umass.edu/~flaherty/>

BACKGROUND

EDUCATION

- 2007 – 2012** **Postdoctoral Fellowship Stanford University**
Department of Biochemistry
Advisors: Ronald W. Davis & Hanlee Ji
- 2003 – 2006** **Ph.D. University of California Berkeley**
Electrical Engineering and Computer Science
Designated Emphasis in Communication, Computation, and Statistics
Designated Emphasis in Computational and Genomic Biology
Dissertation: “A Kinetic Model for G protein-coupled Signal Transduction in Macrophage Cells”
Committee: Michael I. Jordan (advisor), Adam P. Arkin (co-advisor), Richard M. Karp
- 2000 – 2003** **M.S. University of California Berkeley**
Electrical Engineering and Computer Science
Dissertation: “A Study of In-vivo Drug Interactions in Haploinsufficient *Saccharomyces cerevisiae*”
Committee: Michael I. Jordan (advisor), Adam P. Arkin (co-advisor)
- 1996 – 2000** **B.S. Rochester Institute of Technology**

EMPLOYMENT HISTORY

- 2015 – present** University of Massachusetts Amherst
Assistant Professor
Department of Mathematics & Statistics
affiliated: Bioinformatics and Integrative Biology, UMass Medical School (since 2012)
affiliated: Center for Data Science, UMass Amherst (since 2015)
affiliated: Institute for Applied Life Sciences (IALS) Models to Medicine (M2M) (since 2016)
member: Molecular and Cellular Biology Graduate Program, UMass Amherst (since 2019)
adjunct assistant professor: College of Computer and Information Sciences
- 2012 – 2015** WPI
Assistant Professor
Biomedical Engineering
affiliated: Bioinformatics and Computational Biology Program, Worcester Polytechnic Institute (WPI)
affiliated: Bioinformatics and Integrative Biology, UMass Medical School

AWARDS AND HONORS

- 2019 Nominated for UMass University Spotlight Scholar Award
- 2018 Nominated for UMass University Distinguished Teaching Award
- 2017 UMass Lilly Fellowship for Teaching Excellence
- 2015 UMass Mellon Mutual Mentoring Grant
- 2014 Best Poster Award, Research in Computational and Molecular Biology (RECOMB)
- 2014 Best Flash Presentation Award, Great Lakes Bioinformatics Conference (GLBIO)
- 2013 Sigma Xi Scientific Research Society, Full Member
- 2008 Ruth L. Kirschstein National Research Service Award (NRSA T32)
- 2008 Postdoctoral Travel Fellowship, American Society of Biochemistry and Molecular Biology
- 2005 Outstanding Student Paper Award, Neural Information Processing Symposium (NIPS)
- 2000 Tau Beta Pi, Engineering Honor Society
- 2000 RIT Presidential Scholarship
- 2000 Nathaniel Rochester Society Scholarship
- 2000 Outstanding Undergraduate Scholar Award
- 2000 Maletta Foundation Scholarship

Awards to Advised Students

- 2014 Timothy DeFreitas – WPI Salisbury Prize
- 2014 Hachem Saddiki – first place in university-wide undergraduate senior thesis research contest at Al Akhawayn University
- 2014 Timothy DeFreitas – Goldwater Scholarship Honorable Mention

PROFESSIONAL MEMBERSHIPS

- 2018 – present International Society of Bayesian Analysis
- 2013 – present American Statistical Association
- 2013 – present International Society of Computational Biology
- 2013 – 2014 Sigma Xi

SCHOLARSHIP

JOURNAL PUBLICATIONS

1. Giguere, Collin, Harsh Dubey, Vishal Sarsani, Hachem Saddiki, Shai He, and Patrick Flaherty. “SC-SIM: Jointly simulating correlated single-cell and bulk next-generation DNA sequencing data.” *BMC Bioinformatics* 21, no. 215 (2020): 1–10. doi:10.1186/s12859-020-03550-1. biorxiv: biorxiv:10.1101/2020.02.03.930354.
2. He, Shai, Aaron Schein, Vishal Sarsani, and Patrick Flaherty. “A Bayesian Nonparametric Model for Inferring Subclonal Populations from Structured DNA Sequencing Data.” (accepted *Annals of Applied Statistics*), 2020. doi:10.1101/2020.11.10.330183. eprint: biorxiv:10.1101/2020.11.10.330183.
3. Tsurumi, Amy, Yok-Ai Que, Colleen M. Ryan, Patrick J. Flaherty, Marianna Almpiani, Arunava Bandyopadhyaya, Asako Ogura, Yashoda V. Dhole, Laura F. Goodfield, Ronald G. Tompkins, and Laurence G. Rahme. “Multi-biomarker Prediction Models for Multiple Infection Episodes Following Blunt Trauma.” Pre-print published online on Oct 7, 2020, *iScience*, 2020, 1–32. doi:10.1016/j.isci.2020.101659.

4. Wang, Xin, Tao Liu, Hai Xin Song, Shaoyang Cui, Gang Liu, Andrea Christoforou, Patrick Flaherty, Xun Luo, Lisa Wood, and Qing Wang. “Targeted metabolomic profiling reveals an association between elevated arginine and poor functional recovery after stroke.” *Frontiers in Neurology-Neurorehabilitation*, 2020.
5. Ding, Xinxin, Leah M. Pervere, Carl Bascom, Jeffrey P. Bibeau, Sakshi Khurana, Allison M. Butt, Robert G. Orr, Patrick J. Flaherty, Magdalena Bezanilla, and Luis Vidali. “Conditional Genetic Screen in *Physcomitrella Patens* Reveals a Novel Microtubule Depolymerizing-End-Tracking Protein.” *PLOS Genetics* 14, no. 5 (2018): e1007221. doi:10.1371/journal.pgen.1007221.
6. Trapp, Andrew C., Chao Li, and Patrick Flaherty. “Recovering All Generalized Order-Preserving Submatrices: New Exact Formulations and Algorithms.” *Annals of Operations Research* 263, nos. 1-2 (2018): 385–404. doi:10.1007/s10479-016-2173-9.
7. Zhang, Fan, and Patrick Flaherty. “Variational Inference for Rare Variant Detection in Deep, Heterogeneous Next-Generation Sequencing Data.” *BMC Bioinformatics* 18, no. 1 (2017): 45. doi:10.1186/s12859-016-1451-5.
8. DeFreitas, Timothy, Hachem Saddiki, and Patrick Flaherty. “GEMINI: a Computationally-Efficient Search Engine for Large Gene Expression Datasets.” *BMC Bioinformatics* 17, no. 1 (2016): 102. doi:10.1186/s12859-016-0934-8.
9. Wong, Lai H., Sunita Sinha, Julien R. Bergeron, Joseph C. Mellor, Guri Giaever, Patrick Flaherty, and Corey Nislow. “Reverse Chemical Genetics: Comprehensive Fitness Profiling Reveals the Spectrum of Drug Target Interactions.” *PLOS Genetics* 12, no. 9 (2016): e1006275. doi:10.1371/journal.pgen.1006275.
10. Flaherty, Patrick, and Ronald W. Davis. “Robust Optimization of Biological Protocols.” *Technometrics* 57, no. 2 (2015): 234–244. doi:10.1080/00401706.2014.915890.
11. He, Yuting, Fan Zhang, and Patrick Flaherty. “RVD2: an Ultra-Sensitive Variant Detection Model for Low-Depth Heterogeneous Next-Generation Sequencing Data.” *Bioinformatics* 31, no. 17 (2015): 2785–2793. doi:10.1093/bioinformatics/btv275.
12. Mor, Visesato, Antonella Rella, Amir M. Farnoud, Ashutosh Singh, Mansa Munshi, Arielle Bryan, Shamoon Naseem, James B. Konopka, Iwao Ojima, Erika Bullesbach, Alan Ashbaugh, Michael J. Linke, Melanie Cushion, Margaret Collins, Hari Krishna Ananthula, Larry Sallans, Pankaj B. Desai, Nathan P. Wiederhold, Annette W. Fothergill, William R. Kirkpatrick, Thomas Patterson, Lai Hong Wong, Sunita Sinha, Guri Giaever, Corey Nislow, Patrick Flaherty, Xuewen Pan, Gabriele Vargas Cesar, Patricia de Melo Tavares, Susana Frases, Kildare Miranda, Marcio L. Rodrigues, Chiara Luberto, Leonardo Nimrichter, and Maurizio Del Poeta. “Identification of a New Class of Antifungals Targeting the Synthesis of Fungal Sphingolipids.” *mBio* 6, no. 3 (2015): nil. doi:10.1128/mbio.00647-15.
13. Telli, Melinda L., Kristin C. Jensen, Shaveta Vinayak, Allison W. Kurian, Jafi A. Lipson, Patrick J. Flaherty, Kirsten Timms, Victor Abkevich, Elizabeth A. Schackmann, Irene L. Wapnir, Robert W. Carlson, Pei-Jen Chang, Joseph A. Sparano, Bobbie Head, Lori J. Goldstein, Barbara Haley, Shaker R. Dakhil, Julia E. Reid, Anne-Renee Hartman, Judith Manola, and James M. Ford. “Phase II Study of Gemcitabine, Carboplatin, and Iniparib As Neoadjuvant Therapy for Triple-Negative and BRCA1/2 Mutation-Associated Breast Cancer With Assessment of a Tumor-Based Measure of Genomic Instability: Precog 0105.” *Journal of Clinical Oncology* 33, no. 17 (2015): 1895–1901. doi:10.1200/jco.2014.57.0085.

14. Yan, Shuangchun, Amy Tsurumi, Yok-Ai Que, Colleen M. Ryan, Arunava Bandyopadhyaya, Alexander A. Morgan, Patrick J. Flaherty, Ronald G. Tompkins, and Laurence G. Rahme. "Prediction of Multiple Infections After Severe Burn Trauma." *Annals of Surgery* 261, no. 4 (2015): 781–792. doi:10.1097/sla.0000000000000759.
15. Nadauld, Lincoln D, Sarah Garcia, Georges Natsoulis, John M Bell, Laura Miotke, Erik S Hopmans, Hua Xu, Reetesh K Pai, Curt Palm, John F Regan, Hao Chen, Patrick Flaherty, Akifumi Ootani, Nancy R Zhang, James M Ford, Calvin J Kuo, and Hanlee P Ji. "Metastatic Tumor Evolution and Organoid Modeling Implicate TGFBR2 as a Cancer Driver in Diffuse Gastric Cancer." *Genome Biology* 15, no. 8 (2014): 428. doi:10.1186/s13059-014-0428-9.
16. Saddiki, Hachem, Jon McAuliffe, and Patrick Flaherty. "GLAD: a Mixed-Membership Model for Heterogeneous Tumor Subtype Classification." *Bioinformatics* 31, no. 2 (2014): 225–232. doi:10.1093/bioinformatics/btu618.
17. Cushing, Anna, Patrick Flaherty, Erik Hopmans, John M Bell, and Hanlee P Ji. "RVD: a Command-Line Program for Ultrasensitive Rare Single Nucleotide Variant Detection Using Targeted Next-Generation Dna Resequencing." *BMC Research Notes* 6, no. 1 (2013): 206. doi:10.1186/1756-0500-6-206.
18. Lee, HoJoon, Patrick Flaherty, and Hanlee P Ji. "Systematic Genomic Identification of Colorectal Cancer Genes Delineating Advanced From Early Clinical Stage and Metastasis." *BMC Medical Genomics* 6, no. 1 (2013): 54. doi:10.1186/1755-8794-6-54.
19. Flaherty, Patrick, Georges Natsoulis, Omkar Muralidharan, Mark Winters, Jason Buenrostro, John Bell, Sheldon Brown, Mark Holodniy, Nancy Zhang, and Hanlee P. Ji. "Ultrasensitive Detection of Rare Mutations Using Next-Generation Targeted Resequencing." *Nucleic Acids Research* 40, no. 1 (2011): e2–e2. doi:10.1093/nar/gkr861.
20. Schiffman, J. D., J. G. Hodgson, S. R. VandenBerg, P. Flaherty, M.-Y. C. Polley, M. Yu, P. G. Fisher, D. H. Rowitch, J. M. Ford, M. S. Berger, H. Ji, D. H. Gutmann, and C. D. James. "Oncogenic BRAF Mutation With CDKN2A Inactivation Is Characteristic of a Subset of Pediatric Malignant Astrocytomas." *Cancer Research* 70, no. 2 (2010): 512–519. doi:10.1158/0008-5472.can-09-1851.
21. Flaherty, Patrick, Mala L. Radhakrishnan, Tuan Dinh, Robert A. Rebres, Tamara I. Roach, Michael I. Jordan, and Adam P. Arkin. "A Dual Receptor Crosstalk Model of G-Protein-Coupled Signal Transduction." *PLoS Computational Biology* 4, no. 9 (2008): e1000185. doi:10.1371/journal.pcbi.1000185.
22. Flaherty, P., G. Giaever, J. Kumm, M. I. Jordan, and A. P. Arkin. "A Latent Variable Model for Chemogenomic Profiling." *Bioinformatics* 21, no. 15 (2005): 3286–3293. doi:10.1093/bioinformatics/bti515.
23. Lee, William, Robert P. St.Onge, Michael Proctor, Patrick Flaherty, Michael I. Jordan, Adam P. Arkin, Ronald W. Davis, Corey Nislow, and Guri Giaever. "Genome-Wide Requirements for Resistance To Functionally Distinct DNA-Damaging Agents." *PLoS Genetics* 1, no. 2 (2005): e24. doi:10.1371/journal.pgen.0010024.
24. Giaever, G., P. Flaherty, J. Kumm, M. Proctor, C. Nislow, D. F. Jaramillo, A. M. Chu, M. I. Jordan, A. P. Arkin, and R. W. Davis. "Chemogenomic Profiling: Identifying the Functional Interactions of Small Molecules in Yeast." *Proceedings of the National Academy of Sciences* 101, no. 3 (2004): 793–798. doi:10.1073/pnas.0307490100.

- Giaever, Guri, Angela M. Chu, Li Ni, Carla Connelly, Linda Riles, Steeve Véronneau, Sally Dow, Ankuta Lucau-Danila, Keith Anderson, Bruno André, Adam P. Arkin, Anna Astromoff, Mohamed El Bakkoury, Rhonda Bangham, Rocio Benito, Sophie Brachat, Stefano Campanaro, Matt Curtiss, Karen Davis, Adam Deutschbauer, Karl-Dieter Entian, Patrick Flaherty, Françoise Foury, David J. Garfinkel, Mark Gerstein, Deanna Gotte, Ulrich Güldener, Johannes H. Hegemann, Svenja Hempel, Zelek Herman, Daniel F. Jaramillo, Diane E. Kelly, Steven L. Kelly, Peter Kötter, Darlene LaBonte, David C. Lamb, Ning Lan, Hong Liang, Hong Liao, Lucy Liu, Chuanyun Luo, Marc Lussier, Rong Mao, Patrice Menard, Siew Loon Ooi, Jose L. Revuelta, Christopher J. Roberts, Matthias Rose, Petra Ross-Macdonald, Bart Scherens, Greg Schimmack, Brenda Shafer, Daniel D. Shoemaker, Sharon Sookhai-Mahadeo, Reginald K. Storms, Jeffrey N. Strathern, Giorgio Valle, Marleen Voet, Guido Volckaert, Ching-yun Wang, Teresa R. Ward, Julie Wilhelmy, Elizabeth A. Winzeler, Yonghong Yang, Grace Yen, Elaine Youngman, Kexin Yu, Howard Bussey, Jef D. Boeke, Michael Snyder, Peter Philippsen, Ronald W. Davis, and Mark Johnston. “Functional Profiling of the *Saccharomyces Cerevisiae* Genome.” *Nature* 418, no. 6896 (2002): 387–391. doi:10.1038/nature00935.

BOOK CHAPTERS (REFEREED)

- Zhang, Fan, Chuangqi Wang, Andrew C. Trapp, and Patrick Flaherty. “A global optimization algorithm for sparse mixed membership matrix factorization.” In *Contemporary Biostatistics with Biopharmaceutical Application*, edited by Lanju Zhang, Ding-Geng Chen, Hongmei Jiang, Gang Li, and Quan Hui. 2019.

CONFERENCE PROCEEDINGS (REFEREED)

- Greenberg, Craig, Nicholas Monath, Ari Kroben, Patrick Flaherty, Ji Ah Lee, Andrew McGregor, and Andrew McCallum. “Compact Representation of Uncertainty in Hierarchical Clustering.” In *Proceedings of AISTATS2021*. (accepted to AISTATS2021). 2021.
- Greenberg, Craig, Nicholas Monath, Ari Kroben, Patrick Flaherty, Andrew McGregor, and Andrew McCallum. “Compact Representation of Uncertainty in Clustering.” In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*. 2018.
- Carvalho, Paulo, A Kesari, Sean Weaver, Patrick Flaherty, and Gregory S Fischer. “Robotic Assistive Device for Phlebotomy.” In *Proceedings of the ASME IDETC/CIE*. 2015.
- Saddiki, Hachem, and Patrick Flaherty. “Subtype Prediction Intervals for Samples with Intratumor Heterogeneity.” In *Proceedings of the Great Lakes Bioinformatics Conference (GLBIO)*. 2015.
- Flaherty, Patrick, Michael I. Jordan, and Adam P. Arkin. “Robust Design of Biological Experiments.” In *Proceedings of the Neural Information Processing Systems (NIPS) Meeting*. 2005.
- Flaherty, Patrick, and Edward Chung. “Autonomous Video Feedback Controlled Surveillance using an Embedded DSP System.” In *Proceedings of the Texas Instruments DSPfest Conference*. 2000.

CONFERENCE WORKSHOPS (REFEREED)

- Flaherty, Patrick, Pitchaya Wiratchotisatian, Ji Ah Lee, and Andrew C. Trapp. *MAP Estimation for the Gaussian Mixture Model via Mixed Integer Nonlinear Programming*. Sets and Partitions NeurIPS Workshop, 2019.
- Saddiki, Hachem, Andrew C. Trapp, and Patrick Flaherty. *A Deterministic Global Optimization Method for Variational Inference*. NIPS workshop on Advances in Approximate Inference, 2016. eprint: arXiv:1703.07169 (stat.ML).

- Schein, Aaron, Patrick Flaherty, Daniel Sheldon, Mingyuan Zhou, and Hanna Wallach. *Beta Tucker Decomposition for DNA methylation Data*. NIPS workshop on Machine Learning in Computational Biology. **full presentation (20% acceptance rate)**, 2016.

SCHOLARSHIP IN PROGRESS

- Greenberg, Craig, Sebastian Macaluso, Nicholas Monath, Patrick Flaherty, Kyle Cranmer, Andrew McGregor, and Andrew McCallum. “Cluster Trellis: Data Structures & Algorithms for Exact Inference in Hierarchical Clustering.” (submitted to UAI2021), 2021.
- Sarsani, Vishal, Berent Aldikacti, Shai He, Rilee Zienert, Peter Chien, and Patrick Flaherty. “Model-based identification of conditionally-essential genes from transposon-insertion sequencing data.” (submitted to GLBIO2021/PLOS Computational Biology), 2021.
- Schein, Aaron, Anjali Nagulpaly, Patrick Flaherty, Hanna Wallach, and Patrick Flaherty. “Doubly Non-Central Beta Matrix Factorization for DNA Methylation Data.” (submitted to UAI2021), 2021.
- Flaherty, Patrick, and Vishal Sarsani. “Latent read allocation for donor-derived cell-free DNA quantification.” (in preparation with CareDx, Inc.) 2020.
- Flaherty, Patrick, Pitchaya Wiratchotisation, Ji Ah Lee, and Andrew C. Trapp. “MAP Estimation for the Gaussian Mixture Model via Mixed Integer Nonlinear Programming.” 2020. arXiv: 1911.04285 [stat-ml].
- Schein, Aaron, Patrick Flaherty, Daniel Sheldon, Mingyuan Zhou, and Hanna Wallach. “Beta Tucker Decomposition for DNA Methylation Data.” (in preparation), 2020.

PATENTS & INVENTIONS

Patents

- 2019** (joint UMass/MGH provisional patent in process) “Multi Biomarker-based Prediction for Risk Assessment of Infections and Related Outcomes in Trauma and in Trauma-related Clinical Trials”

Invention Disclosures

- 2019** UMass 19-073 / MGH 25737 “Multi Biomarker-based Prediction for Risk Assessment of Infections and Related Outcomes in Trauma and in Trauma-related Clinical Trials”
- 2018** UMass 18-026 “Systems and Methods for Measuring Non-host Cell-free DNA fraction”

PRESENTATIONS

Invited Presentations

- 2020** “Learning Genetic Subpopulations from Structured DNA Sequencing Data” UMass Veterinary and Animal Sciences Seminar
- 2020** “MAP Clustering under the Gaussian Mixture Model via Mixed Integer” UC Davis Department of Mathematics
- 2020** “MAP Clustering under the Gaussian Mixture Model via Mixed Integer” Colorado State University
- 2019** “MAP Estimation for the Gaussian Mixture Model via Modern Optimization Methods” Fields Institute, Toronto

- 2019** “Bayesian Nonparametric Modeling for Integrating Single-cell and Bulk Sequencing Data from Grouped Experiments” Program in Bioinformatics & Computational Biology, UMass Medical School
- 2019** “Statistics for Fundamental and Translational Molecular Biology” Molecular & Cellular Biology Program, University of Massachusetts Amherst
- 2019** “MAP Estimation for the Gaussian Mixture Model via Modern Optimization Methods” Department of Statistics, University of Connecticut
- 2019** “MAP Estimation for the Gaussian Mixture Model via Modern Optimization Methods” NESS, University of Connecticut
- 2018** “Nonparametric Bayes Modeling for Single-cell DNA Sequencing Count Data” Joint Statistical Meeting
- 2018** “A Global Optimization Algorithm for Sparse Mixed Membership Matrix Factorization” New England Statistics Symposium (NESS), University of Massachusetts Amherst
- 2017** “A Nonparametric Bayesian Model for Single-cell Variant Calling” Machine Learning Friends Lunch (Computer Science), University of Massachusetts Amherst
- 2017** “A Nonparametric Bayesian Model for Single-cell Variant Calling” Biostatistics Department , Yale University
- 2016** “TBA” Quality and Productivity Research Conference—Big Data Session, University of Connecticut
- 2015** “Scalable Statistical Inference Methods for Large-scale Genomic Data from Heterogeneous Tumor Samples” University of Massachusetts Amherst Mathematics & Statistics Seminar
- 2015** “TBA” Advances in Analyzing Genomic Data Session at the International Chinese Statistical Association (ICSA) Applied Statistics Symposium, Fort Collins, CO
- 2014** “Statistical Inference with Genomic Data from Heterogeneous Populations” Bioinformatics Group, MD Anderson Cancer Center
- 2014** “From Data to the Clinic: Statistical Inference for Heterogeneous Tumor Genomic Data” EGR390 Special Topics in Engineering: Engineering and Cancer, Smith College
- 2014** “Statistical Inference with Genomic Data from Heterogeneous Populations” Bioinformatics Group, University of Connecticut
- 2014** “The analysis of very large molecules: the human genome” Chemistry and Biochemistry Department, Worcester Polytechnic Institute
- 2013** “Biomedical Informatics and Computational Biology” Invited panelist, Wellesley College
- 2013** “Next Generation Sequencing Technology” EGR390 Special Topics in Engineering: Engineering and Cancer, Smith College
- 2013** “Improving Genomic Diagnostics” Life Sciences and Bioengineering Celebration of Science, Worcester Polytechnic Institute
- 2010** “Robust Optimization of a Biological Protocol” UC Berkeley
- 2009** “Robust Optimization of a Biological Protocol” Biostatistics Group, Stanford University
- 2007** “Modeling the G protein signal transduction system” Rochester Section of the Institute of Electrical and Electronics Engineers
- 2006** “Modeling the G protein signal transduction system” Information Theory and Application Conference, UC San Diego

Research/Contributed Presentations

- 2019** “Globally Optimal Model-based Clustering via Mixed Integer Nonlinear Programming” NSF TRIPODS PI Meeting, Alexandria, VA
- 2015** “Latent Read Allocation” CareDx, Inc.

Poster Presentations

2020 “Exact Inference on Hierarchical Clustering in Particle Physics and Cancer Genomics” NeurIPS Workshop on Machine Learning and the Physical Sciences, Vancouver, BC

2019 “MAP Estimation for the Gaussian Mixture Model via Mixed Integer Nonlinear” NeurIPS Sets and Partitions Workshop, Vancouver, BC

PRESS

2019-11-19 “Statistician Flaherty, Molecular Biologist Chien Join Forces” UMass News

2019-11-12 “Computer Science-Math-Engineering Team Forms New NSF Institute” UMass News

TEACHING

COURSES DEVELOPED AT UMASS AMHERST

Summer 2020 Introduction to Foundations of Data Science
Fall 2018 STAT190F/CS190F: Foundations of Data Science
Spring 2018 STAT535: Statistical Computing
Fall 2016 STAT697ML: Statistical Machine Learning

COURSES TAUGHT AT UMASS AMHERST

Spring 2021 STAT535: Statistical Computing (41 students)
Fall 2020 STAT697ML: Statistical Machine Learning (8 students)
STAT535: Statistical Computing (17 students)
Spring 2020 STAT535: Statistical Computing (40 students)
Fall 2019 STAT535: Statistical Computing (23 students)
STAT697ML: Statistical Machine Learning (10 students)
Spring 2019 STAT535: Statistical Computing (40 students)
Fall 2018 STAT535: Statistical Computing (21 students)
STAT190F/CS190F: Foundations of Data Science (30 students, new course)
Spring 2018 STAT535: Statistical Computing (37 students, new course)
Fall 2017 STAT697ML: Statistical Machine Learning (9 students)
Course Release: Lilly Fellowship
Spring 2017 (Parental Leave)
Fall 2016 STAT697ML: Statistical Machine Learning(46 students, new course)
Spring 2015 STAT525: Linear Models (24 students)
Fall 2015 STAT516: Statistics II (32 students)

INDEPENDENT STUDY COURSES TAUGHT AT UMASS AMHERST

Spring 2021 Harsh Dubey
Fall 2020 Vishal Sarsani, Ji Ah Lee, Anjali Nagulpally
Spring 2020 Vishal Sarsani, Ji Ah Lee
Fall 2019 Ji Ah Lee
Fall 2018 Shai He, Collin Giguere, Fusheng Yang, Jacinth David
Spring 2018 Michael Shlisselberg

COURSES DEVELOPED AT WPI

Spring 2014 BME595L: Machine Learning for Biomedical Informatics

D term 2013 BME2211: Biomedical Data Analysis and Programming
D term 2013 BME3014: Signal Processing Laboratory

COURSES TAUGHT AT WPI

D Term 2015 BME2211: Biomedical Data Analysis and Programming
D Term 2015 BME3014: Signal Processing Laboratory
Fall 2015 BME591: Biomedical Engineering Graduate Seminar

D term 2014 BME2211: Biomedical Data Analysis and Programming
Spring 2014 BME595L: Machine Learning for Biomedical Informatics

D term 2013 BME2211: Biomedical Data Analysis and Programming
D term 2013 BME3014: Signal Processing Laboratory

INDEPENDENT STUDY COURSES TAUGHT AT WPI

2015 D Erin Esco
2014 D Xavier Jackson
2013 D Hachem Saddiki

GUEST LECTURES

2015 BMEX100: Mathematical Biology, Smith College, “SIR Disease Models”
2014 CS534: Artificial Intelligence, WPI, “Bayesian Hierarchical Models in Biology”
2014 BMEX100: Mathematical Biology, Smith College, “Analysis of Messy Data”
2014 BME1001: Introduction to Bioengineering, WPI, “From Data to the Clinic: Statistical Inference for Heterogeneous Tumor Genomic Data”
2014 BME2210: Biomedical Signals, Instruments and Measurements, WPI, “Next-generation Sequencing”
2012 BCB4001: Bioinformatics, WPI, “Identifying Rare Variants”
2005 CS294: Practical Machine Learning, University of California Berkeley

OTHER TEACHING EXPERIENCE

2008 – 2011 Undergraduate Research Mentor, Stanford University;
Ashley Davis (Summer 2008), Anna Cushing (2011 – 2013)
2000 – 2005 Graduate Student Instructor, University of California at Berkeley;
Probability & Random Processes, Intro. to Microelectronics, Modeling Biological Systems
1999 – 2000 Teaching Assistant, Rochester Institute of Technology;
Intro. to Electronics, Electrical and Magnetic Fields

STUDENTS SUPERVISED

Postdoctoral Scholars

2021 – present Tingting Zhao (UMass Amherst/NSF TRIPODS)

PhD Students

2020 – present Harsh Dubey (UMass Amherst)
2020 – present Anjali Nagulpally (UMass Amherst)
2019 – present Vishal Sarsani (UMass Amherst)

- 2019 – present** Ji Ah Lee (UMass Amherst)
2018 – present Shai He (UMass Amherst)
2013 – 2016 Fan Zhang, “Robust Statistical Methods for Rare Variant Detection using Next-generation Sequencing” (WPI) now at Harvard/Broad Institute

MS Students

- 2019 – 2020** Harsh Dubey (UMass Amherst)
2014 – 2018 Hachem Saddiki (now at UMass Biostatistics)
2013 – 2014 Yuting He, “RVD2: An ultra-sensitive variant detection model for low-depth heterogeneous next-generation sequencing data” now at Foundation Medicine

Senior Honors Thesis and Major Qualifying Project Students

- 2019 – 2020** Fu Y. “Towards Defining New Subclasses of Cis-Regulatory Elements via Reformulated Dimension Reduction” (UMass Amherst Commonwealth Honors College)
2014 – 2015 DeFreitas T. “GEMINI: The Genomic Search Engine” (WPI)
2014 – 2015 Pervere L. “Mapping Conditional Loss of Growth Mutations in *Physcomitrella patens*” (WPI)
2014 – 2015 Sadraei A, Sacks J, Lacarra E, Day M “Theranostics of Glioblastoma Multiforme: In Vitro Characterization of Targeted Nanoemulsions and Creation of a 3D Statistical Heatmap to Visualize Nanoemulsion Uptake” (WPI)
2014 – 2015 Weaver S., Kesari A., Carvalho P. “Phlebot: The robotic phlebotomist” (WPI) (**awarded Provost’s MQP Award**)
2013 – 2014 Saddiki H “A Mixed-Membership Model for Heterogenous Subtype Classification Based on an Integrated Analysis of both Gene Expression and Copy Number Variation” (Al Akhawayn University) (**awarded first place in AUI Undergraduate Research Contest 2014**)
2013 – 2014 Miraglia S, Gammel T, Newton B, Ericson J. “Mobile DNA Sequencing.” (WPI)
2013 – 2014 Erb J, Pepo M, Rebh A, Maung W. “DNA Sequence Analysis and Error Correction” (WPI)
2013 – 2014 Ding X. “Isolation of Temperature-sensitive Mutants in the Moss *Physcomitrella patens* and Mapping of their Causal Mutation by Genome Sequencing of Pooled Segregants” (WPI) (**awarded Salisbury Prize 2014**)

Undergraduate Research Students

- Summer 2018** Collin Giguere (UMass Amherst REU student), Fusheng Yang (UMass Amherst REU student)
Summer 2016 Jeanie Lim (UMass Amherst REU student)
D Term 2015 Erin Esco (WPI)
D Term 2014 Xavier Jackson (WPI)
Summer 2013 Timothy DeFreitas (WPI), Hachem Saddiki (WPI), Rachel Blakely (WPI)
B Term 2012 Spencer Keilich (WPI)
2011 –2012 Anna Cushing (Stanford University)
Summer 2010 Ashley Davis (Stanford University)

Academic Advising

- 2019 – 2020** 15 undergraduate academic advisees
2018 – 2019 13 undergraduate academic advisees
2017 – 2018 12 undergraduate academic advisees
2016 – 2017 9 undergraduate academic advisees

2014 – 2015 23 undergraduate academic advisees
2012 – 2013 12 undergraduate academic advisees

SPONSORED RESEARCH

Amounts listed are direct + indirect.

ACTIVE SUPPORT

2019–2022 NIH 1R01GM135931-01 “Learning Conditionally Essential Networks in the Protein Homeostasis System” (Role: PI) \$582,883
ABSTRACT: The goal of this project is to elucidate and validate conditionally essential networks in the protein homeostasis system by developing computationally efficient and statistically accurate methods for analyzing deep DNA sequencing data from massively parallel mutagenesis experiments.

2019–2022 NSF HDR TRIPODS 1934846 “Institute for Integrated Data Science: A Transdisciplinary Approach to Understanding Fundamental Trade-offs and Theoretical Foundations”(Role: co-PI) \$1,500,000
ABSTRACT: The main goal of the proposed work is to establish a new institute at the University of Massachusetts focusing on the theoretical foundations of data science. In the first phase, our research focus is on developing an integrated approach across the lifecycle of typical data science problems: experimental design and data acquisition, analysis, validation, and prediction, which then drives further data acquisition.

2019–2022 Shriner’s Hospital (Role: sub-PI) \$10,000 to UMass
ABSTRACT: The goal of this project is to develop a predictive biomarker panel for multiple infection risk after severe trauma injury.

2020–2025 NIH T32GM008515 “UMass Chemistry-Biology Training Program” (Role: Preceptor & Academic Committee) \$268,464/yr
ABSTRACT: THIS T32 TRAINING PROGRAM SUPPORTS GRADUATE STUDENTS IN CHEMISTRY AND BIOLOGY AT UMASS AMHERST.

COMPLETED SUPPORT

2020 Google Cloud “Google Cloud Platform Educational Grant - STAT190F: Foundations of Data Science” (Role: PI) \$3,400
ABSTRACT: This project supports the computational infrastructure for stat190f. Students use jupyter notebooks hosted on computing infrastructure on the platform for homeworks and labs. This enables them to participate in the class with minimal barriers—only access to a web browser.

2016 – 2018 CareDx Research Contract “Algorithm Improvement for estimating donor-derived cfDNA in transplant patients on a NGS platform” (Role: PI) \$10,000
ABSTRACT: CareDx’s cfDNA assay is an NGS assay to quantify low amount of donor-derived-cfDNA (dd-cfDNA) in transplant recipients without independent genotyping. To further improve the sensitivity, accuracy and precision of the assay, some improvements to the current algorithm can be used. Specific Aim 1: Introducing dependency among SNP markers within same individual. Specific Aim 2: Context-based background estimation for SNP markers in the panel. Specific Aim 3. Weighted contribution of SNP markers based on their empirical performance.

2018 UMass Amherst IALS Midi Grant (Role: PI) \$9,420
ABSTRACT: In September 2017, we submitted a proposal in response to a joint NSF/NIH initiative for promoting research at the interface of the biological and mathematical sciences. That proposal was reviewed, but was not funded. A major weakness the reviewers identified was that we did not have a demonstrated track record of collaboration and sufficient preliminary data. The goal of this M2M midi-grant proposal is to address these weaknesses in preparation for our planned resubmission this year.

- 2016 – 2016** UMass Amherst Mellon Micro Grant “Data Science and Genomic Life Science Mentoring” (Role: PI) \$1,200
 ABSTRACT: The Mutual Mentoring program supports faculty in developing robust professional networks that include a variety of mentoring partners within and outside the UMass Amherst campus, and at a wide variety of career stages.
- 2015 – 2015** NSF “SCH: INT: Collaborative Research: Wearable Devices for In-home Monitoring of Patients with Heart Failure” (Role: co-PI) (awarded but released due to move to UMass Amherst)
- 2014 – 2015** PhRMA Foundation Informatics Starter Grant “Ultrasensitive Detection of Rare Mutations in Heterogeneous Cell Populations using Next-Generation Sequencing” (Role: PI) \$100,000
- 2013 – 2013** WPI Summer Undergraduate Research Fellowship to Timothy DeFreitas “GEMINI: A Genomic Data Search Engine” (Role: PI)
- 2008 – 2012** NIH P01 HG000205 “Genetic Dissection of Complex Traits in Clinical Isolates of *S. cerevisiae*” (Role: Postdoctoral Fellow, PI: Ronald W. Davis)
 ABSTRACT: This study examines quantitative and complex traits in *S. cerevisiae* using a new method called allelic footprinting. The aim is to enumerate, identify and verify all causative loci related to a quantitative phenotype in a model organism.
- 2008 – 2010** NIH T32 CA121940, “Genetic Markers for Risk Stratification of Pediatric Gliomas” (Role: Postdoctoral Fellow; PI: Hanlee P. Ji)
 ABSTRACT: This study examines how to use large quantities of sequence data to discriminate causative effects from random variation and resolve genetic differences between childhood brain cancer tumors.

SERVICE

DEPARTMENT SERVICE AT UMASS AMHERST

- 2020 – 2021** Research Computing Facility Committee
 Data Science Liason
 Undergraduate Curriculum Committee
 Applied Statistics Exam Committee
- 2019 – 2020** Data Science & Scientific Computing Search Committee
 Data Science Liason
 Applied Statistics Exam Committee
- 2018 – 2019** Statistics Graduate Admissions Chair
 Probability & Statistics Seminar Coordinator
 Applied Statistics Exam Committee
- 2017 – 2018** Mathematical Biology Search Committee Member
 Basic Statistics Exam Committee Member
- 2016 – 2017** Probability & Statistics Seminar Coordinator
- 2015 – 2016** Undergraduate Affairs Committee
 Math Distinguished Lecture Series Committee Member
 Applied Statistics Exam Committee Member

COLLEGE SERVICE AT UMASS AMHERST

- 2019 – 2020** CNS Research Review Council
 Microbiology Faculty Search Committee (Quantitative Microbiology)

- 2018 – 2019** CNS Research Review Council
Statistical and Computational Data Science Distinguished Lecture Coordinator
- 2017 – 2018** CNS Research Review Council
Statistical and Computational Data Science Distinguished Lecture Coordinator
- 2015 – 2016** Institute for Applied Life Sciences (IALS) Computational Biology Search Committee Member

THESIS COMMITTEES

PhD

- 2021** Trappit Bansal “TBD” College of Information and Computer Sciences, UMass Amherst (Advisor: Andrew McCallum, Role: External Committee Member)
- 2020** Soumyabrata Pal “Statistical Reconstruction of Combinatorial Structures” College of Information and Computer Sciences, UMass Amherst (Advisor: Arya Mazumdar, Role: External Committee Member)
- 2020** Nils Pilotte “Improved Molecular Diagnostics for Soil-Transmitted Helminths” Molecular & Cellular Biology Program, UMass Amherst (Advisor: Steven Williams, Role: External Committee Member)
- 2020** Daniel Cohen “Neural Methods for Answer Passage Retrieval over Sparse Collections” College of Computer and Information Sciences, UMass Amherst (Advisor: Bruce Croft, Role: External Committee Member)
- 2020** Craig Greenberg “Compact Representations of Uncertainty in Clustering” College of Computer and Information Sciences, UMass Amherst (Advisor: Andrew McCallum, Role: External Committee Member)
- 2020** Dilay Hazal Ayhan “Evolutionary genomics of *Fusarium oxysporum*” Molecular Biology and Biochemistry, UMass Amherst (Advisor: Li Jun Ma, Role: External Committee Member)
- 2020** Ying Zhou “TBD” Biology & Biotechnology, WPI (Advisor: Scarlet Shell, Role: External Committee Member)
- 2020** Toni Delory “Host-fungal pathogen interactions: A study of *Candida albicans* and mammalian macrophage and epithelial cells at the transcriptional level” Biology & Biotechnology, WPI (Advisor: Reeta Rao, Role: External Committee Member)
- 2020** Hang Su “TBD” Computer Science, UMass Amherst (Advisor: Erik Learned-Miller, Role: External Committee Member)
- 2019** Garrett Bernstein “Noise-Aware Inference for Differential Privacy via Sufficient Statistics Perturbation” Computer Science, UMass Amherst (Advisor: Dan Sheldon, Role: External Committee Member)
- 2018** Aaron Schien “Allocative Poisson Factorization for Computational Social Science” Computer Science, UMass Amherst (Advisor: Hanna Wallach, Role: External committee member)
- 2018** Roy Adams “Machine Learning Methods for Activity Detection in Wearable Sensor Data Streams” Computer Science, UMass Amherst (Advisor: Ben Marlin, Role: External Committee Member)
- 2018** Tao Sun “Learning with Aggregate Data” Computer Science, UMass Amherst (Advisor: Dan Sheldon, Role: External Committee Member)
- 2016** Kostis Gourgoulis Mathematics and Statistics, UMass Amherst (Advisor: Markos Katsoulakis, Role: Committee Member)
- 2016** Yuan (Benny) Yin “Synergistic modulation of microRNAs for the treatment of Glioblastoma Multiforme” Biomedical Engineering, WPI (Role: External Committee Member)

MS

- 2016** Kristen Warren “Multichannel Pulse Oximetry: Effectiveness in Reducing HR and SpO2 error due to Motion Artifacts” Biomedical Engineering, WPI

Undergraduate Honors

- 2020** Yu Fu “TBD” Computer Science, Commonwealth Honors College, UMass Amherst (Chair: Patrick Flaherty, Role: Committee Chair)
- 2019** Qin Li “Single Neuron Reconstruction using the Point Set Registration method” Psychology, Commonwealth Honors College, UMass Amherst (Chair: Joseph Bergan, Role: Committee Member)

PROFESSIONAL SERVICE

Leadership

- 2016 – 2020** Associate Editor for *Journal of Classification*
- 2018 – 2020** Program Chair (elected) for International Society of Bayesian Analysis (ISBA) Section on Bayesian Education and Research
- 2016 – 2020** Associate Editor for *Computational Statistics*
- 2017 – 2020** Area Chair for Neural Information Processing Symposium (NIPS)
- 2018** Review Chair for New England Statistics Symposium (NESS) 2018 Review Chair
- 2016** Session Organizer and Chair for JSM “Large-Scale Variational Bayesian Inference”
- 2015** Senior Program Committee (SPC) Member for International Joint Conference on Artificial Intelligence (IJCAI)
- 2016** Topic-contributed Session Organizer for Joint Statistical Meeting (JSM)

Peer Review

- 2020** Classification Journal, Annals of Applied Statistics
- 2019** Annals of Applied Statistics, Bioinformatics, New England Statistics Symposium
- 2018** International Conference on Machine Learning (ICML), Bioinformatics, New England Statistics Symposium
- 2017** International Conference on Machine Learning (ICML), Bioinformatics, New England Statistics Symposium
- 2016** PLOS Computational Biology, Annals of Applied Statistics, Bioinformatics
- 2015** International Conference on Machine Learning (ICML), BMC Bioinformatics, Bioinformatics, Neural Information Processing Symposium (NIPS), Science Reports, PLOS Computational Biology
- 2014** Bulletin of Mathematical Biology, Journal of Applied Statistics, PLOS One
- 2013** Bulletin of Mathematical Biology, Journal of Applied Statistics
- 2012** BMC Bioinformatics
- 2010** Data Modeling and Network Discovery
- 2008** Journal of Applied Statistics

Grant Review

- 06/18/2020** NIH BioData Management and Analysis Study Section
- confidential** Joint DMS/NLM Initiative on Generalizable Data Science Methods for Biomedical Research
- 2019, 2020** Army Research Office
- 10/24/2018** NIH BioData Management and Analysis Study Section
- 2015** National Science Foundation

K-12 OUTREACH ACTIVITIES

- Summer 2020** Introduction to Foundations of Data Science Summer Course (3 weeks)
- Summer 2018** DNA Sequencing Workshop, Eureka by Girls, Inc of Holyoke, MA

- 2014** DNA Sequencing Workshop, Millbury High School, Auburn, MA
2014 “From Cancer to the Flu: Understanding Genomics Through Bioinformatics”, WPI from University Research to the K-12 Classroom Teacher Lecture Series

PROFESSIONAL DEVELOPMENT

- 2018** National Research Mentoring Network: Entering Mentoring Curriculum
2016 UMass GSGW Grant Writer Workshop “Statistical Identification of Persistence Mechanisms in *P. aeruginosa*”
2010 CTL 312: Science and Engineering Course Design, Stanford University
2010 LAW638: Mediation, Stanford University
2010 LAW615: Negotiation, Stanford University
2006 Advanced Bacterial Genetics Course, Cold Spring Harbor

CONSULTING

- 2016** CareDx, Brisbane, CA
2010 Solumn Inc., Palo Alto, CA; Engineering Consultant