

Derek Cole Aguiar

Curriculum Vitae

January 2024

📍 School of Computing, University of Connecticut (UConn), Storrs, CT 06269
Institute for Systems Genomics, UConn, Storrs, CT 06269
🏠 www.derekaguiar.com
📞 (508) 642-8082
✉ derek.aguiar@uconn.edu

Research Interests

I am interested in advancing the foundations of computer science and machine learning and their applications including hierarchical and Bayesian nonparametric latent variable modelling, variational inference and graph theoretic algorithms, and the modelling of haplotypes, transcription, disease genetics, and population genomics.

Academic Experience

UNIVERSITY OF CONNECTICUT, School of Computing, Storrs, CT 2018-
Assistant Professor
Affiliated with the Institute for Systems Genomics

PRINCETON UNIVERSITY, Department of Computer Science, Princeton, NJ 2014-2018
Postdoctoral Researcher
Advised by Barbara Engelhardt, performed research on probabilistic machine learning and Bayesian nonparametric methods for isoform reconstruction and haplotype modelling.

Education

BROWN UNIVERSITY, Department of Computer Science, Providence, RI 2008-2014
Ph.D. (2014) and M.S. (2010), Computer Science
Dissertation Title: "Genome-wide algorithms for haplotype assembly, haplotype phasing, and IBD inference"
Advisor: Sorin Istrail

UNIVERSITY OF RHODE ISLAND, Department of Computer Science, Kingston, RI 2003-2007
Bachelor of Science, Computer Engineering, 2007
Bachelor of Science, Computer Science, 2007
summa cum laude, President's Awards for Excellence (CE & CS)

Publications

Journal manuscripts are annotated with impact factors; conference manuscripts included in [CS rankings](#) conferences are indicated by a 🌟. All Impact Factors are reported according to Clarivate's Journal Citation Reports 5 Year Impact Factor collected on July 2023. Advised UConn or UCHC students are indicated by bolding.

Preprints and Works Under Review

1. **J. Deng, A. Palmer, R. Mahmood, E. Rathbun, J. Bi, K. Mahmood, and D. Aguiar.** Distilling Adversarial Robustness using Heterogeneous Teachers (2023). Under review at (csrankings conference) AAAI 2023.
2. **J. Deng, X. Zhou, H. Tian, Z. Pan, and D. Aguiar.** GBSD: Generative Bokeh with Stage Diffusion. *arXiv preprint arXiv:2306.08251* (2023).
3. **S. Legitime, K. Prabhu, D. McConnell, B. Wang, D. K. Dey, and D. Aguiar.** Improving Opioid Use Disorder Risk Modelling through Behavioral and Genetic Feature Integration (2023). Under second round of review at Journal of Biomedical Informatics (IF: 6.9).
4. **K. Udeh, D. W. Wanik, D. Cerrai, E. Anagnostou, and D. Aguiar.** Probabilistic Storm and Electric Utility Customer Outage Prediction (2023). Under first round of review at IEEE Transactions on Power Delivery (IF: 4.9).

5. **M. Hosseini, D. McConnell**, and D. Aguiar. Bayesian Reconstruction and Differential Testing of Excised mRNA. *arXiv preprint arXiv:2211.07105* (2022). Under first round of review at *Genome Research* (IF: 11.8).
6. R. De Vito, I. N. Grabski, D. Aguiar, L. M. Schneper, A. Verma, J. C. Fernandez, C. Mitchell, J. Bell, S. McLanahan, D. A. Notterman, et al. Differentially methylated regions and methylation QTLs for teen depression and early puberty in the Fragile Families Child Wellbeing Study. *bioRxiv* (2021), 2021–05.

Refereed Papers

1. **J. Deng**, X. Zhou, H. Tian, Z. Pan, and D. Aguiar. Smooth and Stepwise Self-Distillation for Object Detection. *IEEE International Conference on Image Processing (ICIP)* (2023).
2. **M. Hosseini, A. Palmer, W. Manka**, P. G. Grady, **V. Patchigolla**, J. Bi, R. O'Neill, Z. Chi, and D. Aguiar. Deep statistical modelling of nanopore sequencing translocation times reveals latent non-B DNA structures. *Bioinformatics* **39**(Supplement_1) (2023), i242–i251. Also in *proceedings of ISMB 2023* (✎).
3. **M. Hosseini**, A. J. Sabet, S. He, and D. Aguiar. Interpretable fake news detection with topic and deep variational models. *Online Social Networks and Media* **36** (2023), 100249. (IF: 7.61, journal-reported).
4. **M. Hutchison**, B. S. Russell, **A. Leander**, N. Rickles, D. Aguiar, X. S. Cong, O. Harel, and A. V. Hernandez. Trends and Barriers of Medication Treatment for Opioid Use Disorders: A Systematic Review and Meta-Analysis. *Journal of Drug Issues* (2023).
5. **A. Palmer**, Z. Chi, D. Aguiar, and J. Bi. Auto-Encoding Goodness of Fit. In: *The Eleventh International Conference on Learning Representations (ICLR)*. 2023. <https://openreview.net/forum?id=JjCAdMulu9v>. (✎).
6. **B. A. Alpay**, M. Gosink, and D. Aguiar. Evaluating molecular fingerprint-based models of drug side effects against a statistical control. *Drug Discovery Today* (2022), 103364. (IF: 8.1).
7. C. Li*, L. Qu*, **A. J. Matz***, P. A. Murphy, Y. Liu, A. W. Manichaikul, D. Aguiar, S. S. Rich, D. M. Herrington, D. Vu, et al. AtheroSpectrum Reveals Novel Macrophage Foam Cell gene Signatures Associated with Atherosclerotic Cardiovascular Disease Risk. *Circulation* (2022). (IF: 33.2).
8. **K. Udeh**, D. W. Wanik, D. Cerrai, D. Aguiar, and E. Anagnostou. Autoregressive Modeling of Utility Customer Outages with Deep Neural Networks. In: *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*. 2022, pp.0406–0414.
9. **D. J. McConnell, J. Zhu**, S. Pandya, and D. Aguiar. Case-level prediction of motion outcomes in civil litigation. In: *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*. 2021, pp.99–108.
10. **B. A. Alpay**, P. Demetci, S. Istrail, and D. Aguiar. Combinatorial and statistical prediction of gene expression from haplotype sequence. *Bioinformatics* **36**(Supplement_1) (2020), i194–i202. Also in *proceedings of ISMB 2020* (✎).
11. D. Aguiar, L.-F. Cheng, B. Dumitrascu, F. Mordelet, A. Pai, and B. Engelhardt. Bayesian nonparametric discovery of isoforms and individual specific quantification. *Nature Communications* **9**(1) (2018), 1681. (IF: 17.0).
12. D. Aguiar, E. M. Morrow, and S. Istrail. Tractatus: An Exact and Subquadratic Algorithm for Inferring Identical-by-Descent Multi-shared Haplotype Tracts. In: *Annual International Conference on Research in Computational Molecular Biology (RECOMB)*. Springer. 2014, pp.1–17. (✎).
13. D. Aguiar, W. S. Wong, and S. Istrail. Tumor haplotype assembly algorithms for cancer genomics. In: *Pacific Symposium on Biocomputing*. NIH Public Access. 2014, pp.3.
14. I. C. McDowell, C. Nikapitiya, D. Aguiar, C. E. Lane, S. Istrail, and M. Gomez-Chiarri. Transcriptome of American oysters, *Crassostrea virginica*, in response to bacterial challenge: insights into potential mechanisms of disease resistance. *PLoS One* **9**(8) (2014), e105097. (IF: 3.8).
15. D. Aguiar and S. Istrail. Haplotype assembly in polyploid genomes and identical by descent shared tracts. *Bioinformatics* **29**(13) (2013), i352–i360. Also in *proceedings of ISMB 2013* (✎).
16. S. Tulin*, D. Aguiar*, S. Istrail, and J. Smith. A quantitative reference transcriptome for *Nematostella vectensis* earlyembryonic development: a pipeline for de novo assembly in emergingmodel systems. *EvoDevo* **4**(1) (2013), 16. (IF: 3.1).

* denotes co-first authorship when ambiguous

17. D. Aguiar, B. V. Halldórsson, E. M. Morrow, and S. Istrail. DELISHUS: an efficient and exact algorithm for genome-wide detection of deletion polymorphism in autism. *Bioinformatics* **28**(12) (2012), i154–i162. Also in *proceedings of ISMB 2012* (✎).
18. D. Aguiar and S. Istrail. HapCompass: a fast cycle basis algorithm for accurate haplotype assembly of sequence data. *Journal of Computational Biology* **19**(6) (2012), 577–590. (IF: 1.6).
19. B. V. Halldórsson*, D. Aguiar*, and S. Istrail. “Haplotype phasing by multi-assembly of shared haplotypes: phase-dependent interactions between rare variants”. In: *Pacific Symposium on Biocomputing*. 2011, pp.88–99.
20. B. V. Halldórsson*, D. Aguiar*, R. Tarpine, and S. Istrail. The Clark phase-able sample size problem: long-range phasing and loss of heterozygosity in GWAS. In: *Annual International Conference on Research in Computational Molecular Biology (RECOMB)*. Springer. 2010, pp.158–173. Note: Journal version published in *Journal of Computational Biology*. (✎).

Book Chapters

21. S. Istrail, R. Tarpine, K. Schutter, and D. Aguiar. “Practical Computational Methods for Regulatory Genomics: A cisGRN-Lexicon and cisGRN-Browser for Gene Regulatory Networks”. In: *Computational Biology of Transcription Factor Binding*. Totowa, NJ: Humana Press, 2010, pp. 369–399.

Ph.D. Thesis

22. D. Aguiar. “Genome-wide algorithms for haplotype assembly, haplotype phasing, and IBD inference”. PhD thesis. Brown University, 2014.

Works in Progress

23. D. Aguiar, G. Gliner, L. T. Elliott, **A. Palmer**, **M. Hosseini**, Y. W. Teh, and B. E. Engelhardt. *The fragmentation coagulation process haplotype cluster graph model*. 2023. (In preparation for *Genome Biology*).
24. **J. Deng** and D. Aguiar. *Distilling Adversarial Robustness using Heterogeneous Teachers*. 2023. (In preparation for *AAAI*).
25. **M. Hosseini** and D. Aguiar. *Haplotype Assembly and Phasing Unification*. 2023. (In preparation for *RECOMB*).
26. D. McErlean*, D. Aguiar*, and S. Istrail. *A globally optimal algorithm for computing maximum likelihood haplotype phasings: expo-posynomial optimization by graph homomorphism*. 2023. (In preparation for *PNAS*).
27. **A. Palmer**, Z. Chi, J. Bi, and D. Aguiar. *Supervised Bayesian Nonparametric Clustering*. 2023. (In preparation for *AISTATS*).

Funded Proposals

My share of total funding is \$1,291,504.28 as of December 2023. My share of funding is given in parenthesis when ambiguous.

Derek Aguiar (PI of subaward) 12/01/2023-11/30/2027
NIH

ASCVD risk assessment in T2DM facilitated by novel computational immunology programs
Role: Principal Investigator of subaward, \$182,294.00 total award

Derek Aguiar (PI) 10/01/2023–06/30/2024
DEFENSEWERX

Develop AI/ML Tools
Role: Principal Investigator, \$175,000 total award (my share: \$58,333.33)

Derek Aguiar (PI) 08/15/2023–06/30/2024
The Bushnell Performing Arts Center
Bushnell Virtual Mentor Digital Learning Project: Development of Conversational AI and Assessment Systems
Role: Principal Investigator, \$43,880.00 total award

Derek Aguiar (PI) 07/15/2023–06/30/2028
National Science Foundation
CAREER: Practical algorithms and high dimensional statistical methods for multimodal haplotype modelling
Role: Principal Investigator, \$548,346.00 total award

Derek Aguiar (co-PI) 05/05/2023–07/31/2025
Travelers Companies, Inc.
Creating insurance-specific Transformers for representation learning from large-scale unstructured claim text
Role: co-Principal Investigator, \$584,528 total award (my share: \$184,865.33)

Derek Aguiar (Collaborator) 01/01/2023-12/31/2024
American Heart Association
miR-150 regulated adipose tissue B cells in obesity-induced health risk
Role: Collaborator, \$65,106.00 total award (my share: \$0.00)
Note: Relinquished after the proposal was awarded F30 from NIH/NHLBI

Derek Aguiar (Senior Personnel) 01/01/2023-5/31/2023
American Heart Association
Assessing cardiovascular risk using diabetes-induced transcriptomic networks
Role: Senior Personnel, \$65,106.00 total award (my share: \$6,510.60)
Note: Relinquished after the proposal was awarded F30 from NIH/NHLBI

Derek Aguiar (PI) 09/01/2022-08/01/2023
UConn Research Excellence Program
Probabilistic Learning Algorithms for Haplotype Phasing and Haplotype Assembly Unification
Role: Principal Investigator, \$24,974.38 total award

Derek Aguiar (PI) 04/01/2022-03/01/2023
DEFENSEWERX
Artificial Intelligence / Machine Learning (AIML) Development
Role: Principal Investigator, \$250,000 total award (my share: \$145,790.04)

Derek Aguiar (PI) 03/01/2021-02/01/2023
Google
Supporting math and science teacher learning to address societal problems with data science and artificial intelligence
Role: Principal Investigator, \$80,000 total award

Derek Aguiar (senior personnel) 08/01/2020-02/01/2022
The Institute for Collaboration on Health, Intervention, and Policy (InCHIP)
A Pilot Investigation Understanding the Connecticut experience on what factors most affect patient adherence to Medication Assisted Treatments (MAT) for Opioid Use Disorder.
Role: Senior Personnel, \$25,000 total award (my share: \$0.00)

Derek Aguiar (PI) 09/01/2019-06/01/2021
 The Institute for Collaboration on Health, Intervention, and Policy (InCHIP)
Unifying clinical support methods for preventing and managing opioid use disorder
 Role: Principal Investigator, \$25,000 total award (my share: \$10,000.00)

Scored Proposals Without Official Decisions

Derek Aguiar (co-Investigator) 12/01/23-11/30/28
 National Institutes of Health
Novel Computational Immunology Algorithms for ASCVD Risk Assessment in T2DM
 Impact Score: 27, Percentile: 6
 Role: co-Investigator, \$2,847,700 total award (my share: \$250,092)

Teaching Experience

UNIVERSITY OF CONNECTICUT, Storrs, CT 2018-

Instructor, Undergraduate Core Course with a Lab, *Introduction to Computing for Engineers*: 2019 (307 students), 2020 (1051 students)

- Designed and Taught CSE1010 – Introduction to Computing for Engineers. Complete re-design and development of CSE1010, Introduction to Computing for Engineering. Creation of course materials, lab assignments, homeworks, a group project and presentation, a COVID-19 data science tracking project, lecture videos for 14 weeks using the lightboard room, new exams, exam reviews, and lecture notes.
- Integrated data science as a 3 week unit into an introductory computer science class.

Instructor, Undergraduate Core Course, *Algorithms and Complexity*: 2019 (48 students), 2021 (80 students), 2022 (95 students), 2023 (54 students), 2024 (TBD – includes honors section)

- Median scores (out of 5) for the question “What is your overall rating of the instructor’s teaching?” were 4, 4, 5, and 4 for 2019, 2021, 2022, and 2023, respectively.
- Taught CSE3500 – Algorithms and Complexity, which entailed creating lesson plans, syllabus, problem sets, programming assignments, tests, in-class demonstrations (FFT, GPU programming, optimization algorithms), honors projects, and worksheets.
- Topics included: asymptotic notations, solving summations, recurrences, invariants, proof techniques, dynamic programming, divide-and-conquer, data structures, probabilistic algorithms, competitive analysis, greedy algorithms, graph algorithms, intractability and NP-completeness, parallel computing, approximation algorithms, and optimization algorithms.

Instructor, Graduate Course, *Bayesian Machine Learning*: 2018 (7 students), 2019 (24 students), 2021 (18 students), 2022 (21 students), 2023 (28 students)

- Median scores (out of 5) for the question “What is your overall rating of the instructor’s teaching?” were 5’s for each of the five years.
- Developed the course CSE5095 – Bayesian Machine Learning, which entailed creating lesson plans, a syllabus, and project assignments.
- We covered the three fundamental components of Bayesian Machine Learning: probabilistic modeling, inference algorithms, and model checking.
- The course culminates in research-quality projects related to each student’s research interests that implements some probabilistic framework from problem specification to model checking.

Instructor, Undergraduate Core Course, *Senior Design Project*: 2021 (5 students)

- ▶ Median scores (out of 5) for the question “What is your overall rating of the instructor’s teaching?” was a 5.
- ▶ Led Senior Design Project team along with the sponsor Pfizer Inc. on a drug discovery machine learning project that culminated in a research manuscript (Alpay *et al.* 2022).

Instructor, Undergraduate Independent Study (NASA), *Seminar*: 2022 (5 students)

- ▶ Mentored 5 CSE students taking part in the NASA Big Idea Challenge team. Topics included algorithms and machine learning.

Instructor, Graduate Course, *Seminar*: 2023 (10 students), 2024 (10 students)

- ▶ Bioinformatics seminar covering research in bioinformatics, machine learning, and algorithms.

BROWN UNIVERSITY, Providence, RI

2009-2014

Head Teacher’s Assistant, *Algorithmic Foundations of Computational Biology*: 2011, 2013-2014

- ▶ Assisted teaching of Algorithmic Foundations of Computational Biology to classes of intellectually diverse undergraduate students from the Computer Science and Applied Mathematics departments.
- ▶ Developed and delivered a week of lectures on Haplotype Assembly. Administered online discussion platform to encourage student participation.
- ▶ Mentored students participating in interdisciplinary CYRENE project on cis-regulatory genomics.
- ▶ Co-developed the syllabus, created the course webpage, developed homework assignments and exams, curated lecture notes.

Head Teacher’s Assistant, *Medical Bioinformatics*: 2009-2010, 2012-2013

- ▶ Co-created the Medical Bioinformatics course designed for computational and life sciences graduate students.
- ▶ Created the Genomathica, a set of interactive genomics algorithms written in Mathematica, and used Genomathica in lectures and homeworks enabling life science students to tinker and develop data analysis software pipelines without the burden of learning a programming language.
- ▶ Prepared and presented a week of lectures on haplotype phasing and genomic deletion inference in genome-wide association study data.
- ▶ Co-developed the syllabus, created the course webpage, developed homework assignments and exams, constructed lecture notes, and graded homework and exams.

Teacher’s Assistant, *Algorithmic Foundations of Computational Biology*: 2010

- ▶ Taught Algorithmic Foundations of Computational Biology to an intellectually diverse class of undergraduate students from the Computer Science and Applied Mathematics departments.
- ▶ Maintained course webpage, graded homework assignments and exams, and curated lecture notes.

UNIVERSITY OF RHODE ISLAND, Kingston, RI

2007

Teacher’s Assistant, *Computer Communications*: 2007

- ▶ Provided instruction for Computer Communications, a computer engineering and computer science senior undergraduate course.
- ▶ Independently researched the topics covered in the course with the objective of assisting in the redesign of the curriculum. The five lesson plans I created, including programming and homework assignments, incorporated: working at the frame level with Application, Transport, and Internet layer protocols (DNS, SMTP, TELNET, ARP, TCP/IP), and C Socket programming.

Invited Talks & Panels

1. The 31st Annual Intelligent Systems For Molecular Biology and the 22nd Annual European Conference on Computational Biology “Deep statistical modelling of nanopore sequencing translocation times reveals latent non-B DNA structures.” July 2023
2. Nutrition Sciences Colloquium, University of Connecticut “Deep statistical modelling of nanopore sequencing reveals latent non-B DNA structures” February 2023
3. 12th John A. Speziale Symposium, Quinnipiac University “[Artificial Intelligence in Law](#)” November 2022
4. ImmunoCardio Group Meeting, University of Connecticut, “Machine Learning, Genomics, Transcriptomics, Bash and R Tutorial.” May 2022
5. Panelist, exploreCSR Workshop, University of Rhode Island, “Life after Grad School.” April 2022
6. Connecticut Computer Science Teachers Association Chapter Meeting, “An Update on our SMARTAI Education Module.” February 2022
7. 35th Anniversary Lecture Series, University of Connecticut, “Bayesian machine learning and the isoform reconstruction problem.” January 2022
8. Computer Science Colloquium, University of Rhode Island, “Bayesian nonparametric modelling and scalable inference in genomics and transcriptomics data.” November 2021
9. Panelist, [UConn STEM Seminar Series](#), Interdisciplinary Engineering Research, “An Interdisciplinary Tour of Machine Learning in Biology, Law, and More.” September 2021
10. Bioinformatics Seminar Series, University of Connecticut, “Bayesian Networks and Nonparametric Haplotype Clustering.” September 2021
11. Connecticut Computer Science Teachers Association Chapter Meeting, “Preparing Students for a Data-Centric Computer Science Curriculum.” May 2021
12. Panelist, exploreCSR Workshop, University of Rhode Island, “Life after Grad School.” April 2021
13. Panelist, Computational Modeling Meet & Speak, CT Institute for the Brain and Cognitive Sciences, “Bayesian machine learning in large models with applications in genomics.” May 2020
14. Department of Statistics Colloquia, University of Connecticut, “Bayesian Nonparametric Modelling and Scalable Inference In Large-Scale Genomics Data.” January 2020
15. New York Area Population Genomics Workshop, “Efficient construction of haplotype cluster graphs and estimating the age of mutations.” January 2017
16. American Society of Human Genetics (ASHG) Annual Meeting, “Haplotype phasing using cluster graphs.” October 2015
17. NIPS Workshop on Machine Learning in Computational Biology (MLCB) “A Bayesian nonparametric statistical framework for haplotype phasing.” December 2014
18. 18th Annual International Conference on Research in Computational Molecular Biology “Tractatus: an exact and subquadratic algorithm for inferring identical-by-descent multi-shared haplotype tracts.” April 2014
19. 19th Pacific Symposium on Biocomputing “Tumor Haplotype Assembly Algorithms for Cancer Genomics.” January 2014
20. Brown University IPP Symposium: Putting Big Data to Work. “Ome sweet ome: the genome as a model for big data.” April 25, 2013
21. 21st Annual International Conference on Intelligent Systems for Molecular Biology “Haplotype assembly in polyploid genomes and identical by descent shared tracts.” July 2013
22. HiT-Seq workshop in the 21st Annual International Conference on Intelligent Systems for Molecular Biology “Haplotype assembly in the presence of hemizyosity, haplotype sharing, polyploidy, and viral quasispecies.” July 2013

23. 20th Annual International Conference on Intelligent Systems for Molecular Biology “DELISHUS: An efficient and exact algorithm for Genome-Wide detection of deletion polymorphism in autism.” July 2012
24. Brown Computational Biology Open House “SNPs and haplotypes and GWAS oh my!” Feb. 27, 2012
25. Second Annual IEEE ICCABS CANGS Workshop “Robust algorithms for inferring haplotype phase and deletion polymorphism from high throughput whole genome sequence data.” Feb. 24, 2012
26. Brown Computer Science - Research Exchange Seminars with Tea (REST) “Computational Challenges in Genome-wide Association Studies.” Nov. 30, 2010
27. Fourteenth International Conference on Research in Computational Molecular Biology “The Clark Phase-able Sample Size Problem Long-range Phasing and Loss of Heterozygosity in GWAS.” August 12, 2010

Extended Abstracts and Posters

1. Bai, Y., Campbell, T., Legitime, S., Aguiar, D., Staples, M., Corricelli, J. “Curriculum Design and Improvement: Integrating AI Concepts and Societal Problems in a Secondary Science Module.” Paper for the 2024 National Association for Research in Science Teaching (NARST) International Convention. March 17-20, Denver, CO.
2. Alyssa J Matz*, Keaton Karlinsey, Anthony T. Vella, Derek Aguiar, Laura Haynes, Beiyan Zhou, “Novel multi-omics strategies to examine the impact of aging on B cell responses to influenza vaccination.” CHARGE 2023, San Antonio, Texas.
3. Alyssa J Matz*, Keaton Karlinsey, Anthony T. Vella, Derek Aguiar, Laura Haynes, Beiyan Zhou, “Novel Function Annotation Tool Revealed Obesity-Induced Inflammation is Exacerbated by miR-150 Regulated B Cell Receptor Signaling.” American Heart Association’s annual Scientific Sessions 2023, Philadelphia, Pennsylvania.
4. Chuan Li*, Lili Qu*, Alyssa J Matz*, Patrick A. Murphy, Yongmei Liu, Ani Manichaikul, Derek Aguiar, Stephen S Rich, David M Herrington, David Vu, Craig W Johnson, Jerome I Rotter, Wendy S Post, Anthony T. Vella, Annabelle Rodriguez, Beiyan Zhou, “AtheroSpectrum reveals pathogenic foaming program and provides gene signatures for CVE prediction.” CHARGE 2022, Philadelphia, Pennsylvania.
5. Alyssa J Matz, Chuan Li, Lili Qu, Patrick A. Murphy, Yongmei Liu, Ani W Manichaikul, Derek Aguiar, Stephen S Rich, David M Herrington, David Vu, W. Craig Johnson, Jerome I. Rotter, Wendy S. Post, Anthony T. Vella, Annabelle Rodriguez-Oquendo and Beiyan Zhou. AtheroSpectrum reveals novel macrophage foam cell gene signatures associated with atherosclerotic cardiovascular disease risk. 18th Society of Chinese Bioscientists in America (SCBA) International Symposium. July 2022.
6. Alyssa J Matz, Chuan Li, Lili Qu, Patrick A. Murphy, Derek Aguiar, Anthony T. Vella, Annabelle Rodriguez, Beiyan Zhou. AtheroSpectrum reveals novel macrophage foaming programs. Virtual Event. Society for Leukocyte Biology 54th Annual Event. October 2021.
7. Marjan Hosseini*, Devin McConnell*, Derek Aguiar, “Hierarchical Bayesian Admixture Modelling of Intron Excision Structure.” Workshop on Missing Data Challenges in Computation, Statistics and Applications 2020, Institute for Advanced Study, Princeton, New Jersey.
8. Derek Aguiar, Li-Fang Cheng, Bianca Dumitrascu, Fantine Mordelet, Barbara E. Engelhardt, “BIISQ: Bayesian nonparametric discovery of Isoforms and Individual Specific Quantification from RNA-seq data.” RECOMB 2016, Santa Monica, CA.
9. Derek Aguiar, Li-Fang Cheng, Bianca Dumitrascu, Fantine Mordelet, Barbara E. Engelhardt, “BIISQ: Bayesian nonparametric discovery of Isoforms and Individual Specific Quantification from RNA-seq data.” Probabilistic Modeling in Genomics 2016, Oxford, UK.
10. Derek Aguiar, Li-Fang Cheng, Bianca Dumitrascu, Fantine Mordelet, Athma Pai, Barbara E. Engelhardt, “Bayesian nonparametric isoform discovery and quantification of many RNA-seq samples.” American Society of Human Genetics Annual Meeting 2016, Baltimore, MD.

11. Derek Aguiar, Lloyd T. Elliott, Yee Whye Teh, and Barbara E. Engelhardt, "Haplotype phasing using cluster graphs." American Society of Human Genetics Annual Meeting 2015, Baltimore, MD.
12. Derek Aguiar, Lloyd T. Elliott, Yee Whye Teh, and Barbara E. Engelhardt, "Tractable haplotype phasing and imputation with nonparametric fragmentation-coagulation processes." Probabilistic Modeling in Genomics 2015, Cold Spring Harbor Laboratory, NY.
13. Derek Aguiar, Lloyd T. Elliott, Yee Whye Teh, and Barbara E. Engelhardt, "A statistical framework for modeling genetic data as haplotype cluster graphs with application to haplotype phasing, association mapping, and whole-genome compression." The Biology of Genomes 2015, Cold Spring Harbor Laboratory, NY.
14. Derek Aguiar, Lloyd T. Elliott, Yee Whye Teh, and Barbara E. Engelhardt, "A Bayesian nonparametric statistical framework for haplotype phasing." Workshop: Machine Learning in Computational Biology in NIPS 2014, Montreal, Canada. (selected for oral presentation)
15. Derek Aguiar, Austin Huang, Rami Kantor, Eric Morrow and Sorin Istrail, "Haplotype assembly in the presence of hemizyosity, haplotype sharing, polyploidy, and viral quasispecies." HiT-Seq workshop in the 21st Annual International Conference on Intelligent Systems for Molecular Biology, July 2013, Berlin, Germany. (selected for oral presentation)
16. Sarah Tulin, Derek Aguiar, Sorin Istrail, and Joel Smith, "Nematostella reference transcriptome and high throughput gene regulatory network construction." SDB 71st Annual Meeting, July 2012, Montreal, Canada.
17. Derek Aguiar and Sorin Istrail, "HAPCOMPASS: A fast cycle basis algorithm for accurate haplotype assembly of next-generation sequence data." 20th Annual Intelligent Systems for Molecular Biology, July 2012, Long Beach, CA.
18. Derek Aguiar, Ryan Tarpine, Fumei Lam, Bjarni V. Halldorsson, Eric Morrow, and Sorin Istrail, "Long-Range Haplotype Phasing by Multi-Assembly of Shared Haplotypes: Phase-Dependent Interactions Between Rare Variants." Genomics of Common Disease, Wellcome Trust Sanger Genome Center, Cambridge UK, December 2011
19. Derek Aguiar, Ryan Tarpine, Fumei Lam, Bjarni V. Halldorsson, Eric Morrow, and Sorin Istrail, "Long-Range Haplotype Phasing by Multi-Assembly of Shared Haplotypes: Phase-Dependent Interactions Between Rare Variants." The Gordon Research Conference on Human Genetics and Genomics, July 17-22, 2011, Salve Regina University, Newport, RI.
20. Ryan Tarpine, James Hart, Timothy Johnstone, Derek Aguiar, Sorin Istrail, "Report on the Cyrene Project: A cis-Lexicon Containing the Regulatory Architecture of 557 Regulatory Genes Experimentally Validated Using the *Davidson Criteria*." The Developmental Biology of the Sea Urchin Meeting, April 27-30, 2011, Woods Hole, MA.
21. Derek Aguiar, Ryan Tarpine, Eric Ruggieri, Jessica Nadel, David Moskowitz, Sorin Istrail, "Beyond GWAS: Robust Computational Analysis of the Multiple Sclerosis Genetic Consortium Data." Fourth Annual Center for Computational Biology Poster Session, April 28, 2010, Brown University, RI.

Software

1. D. Aguiar, L.-F. Cheng, B. Dumitrascu, and B. Engelhardt. *Bayesian nonparametric discovery of Isoforms and Individual Specific Quantification with BIISQ*. Version 1.0.0. [\[code\]](#). 2017.
2. D. Aguiar and S. Istrail. *HapCompass: fast cycle basis algorithms for accurate diploid and polyploid haplotype assembly*. Version 0.8.2. [\[code\]](#). 2016.
3. D. Aguiar and S. Istrail. *Tractatus: an exact and subquadratic algorithm for inferring identity-by-descent multi-shared haplotype tracts*. Version 0.0.1. [\[code\]](#). 2013.
4. D. Aguiar, S. Tulin, S. Istrail, and J. Smith. *A quantitative reference transcriptome for Nematostella vectensis earlyembryonic development: a pipeline for de novo assembly in emerging model systems*. Version 1.0.0. [\[code\]](#). 2013.

5. D. Aguiar and S. Istrail. *DELISHUS: an efficient and exact algorithm for genome-wide detection of deletion polymorphism in autism*. Version 0.1.0. [\[code\]](#). 2012.
6. R. Tarpine, D. Aguiar, and S. Istrail. *CYRENE cis-Regulatory Genome Browser*. Version 1.0.0. [\[code\]](#). 2012.

Academic Advising

Ph.D. Advising (current)

- Major Advisor, Marjan Hosseini (expected graduation in AY 2023-2024), Computer Science;
- Major Advisor, Chandra Sekhar Dronavajjala, Computer Science;
- Major Advisor, Jieren Deng (expected graduation in AY 2023-2024), Computer Science;
- Major Advisor, Diego Nogueras Rivera, Computer Science;
- Major Advisor, William Manka, Computer Science;
- Associate Advisor, Piyush Shrivastav, Computer Science;
- Associate Advisor, Keaton Karlinsey, Immunology;
- Associate Advisor, Alyssa Matz, Immunology;

Ph.D. Advising (alumni)

- Major Advisor, Kingsley Udeh, Computer Science; now a *Postdoctoral Research Associate*, Eversource Energy Center, University of Connecticut
- Associate Advisor, Pariksheet Nanda, Molecular and Cell Biology;
- Associate Advisor, Lina Kloub, Computer Science;
- Associate Advisor, Tianyu Wang, Computer Science;
- Associate Advisor, Fahad AlQahtani, Computer Science;
- Associate Advisor, Md Abdullah Al Fahim, Computer Science;
- Associate Advisor, Abhijit Mondal, Computer Science;

MS and MEng Masters Advising (current)

- Major Advisor (Thesis track), William Manka, Computer Science;

MS and MEng Masters Advising (alumni)

- Major Advisor (Thesis track), Sybille Legitime, Computer Science; now a *Senior Machine Learning Engineer*, Prudential Financial
- Major Advisor, Jackie Ramirez, Computer Science;
- Major Advisor, Dongping Zhu, Computer Science;
- Major Advisor, Devin McConnell, Computer Science;
- Major Advisor, Kaustubh Prabhu, Computer Science;
- Major Advisor, Shubhangi Shubhangi, Computer Science;
- Associate Advisor, Saurav Dhar, Computer Science;
- Associate Advisor, Mingwan Zhang, Computer Science;
- Associate Advisor, Nachiket Deo, Computer Science;

Honors Program Advising

- Major Advisor, Berk Alpay, University Scholar, Computer Science and Engineering; now a *Ph.D. Candidate* at Harvard University
- Associate Advisor, Alexey Pozdnyakov, University Scholar, Computer Science and Engineering;

MS Data Science Masters Advising

- Major Advisor, Jennifer Minh, Computer Science;
- Major Advisor, Madison Beebe, Computer Science;

Individualized Majors Advising

- Undergraduate Advisor, Matthew Jane, Computational Neuroscience;
- Undergraduate Advisor, Patrick Cummins, Samuel Hughes, Anthony Zeimbekakis, Michael Zheng, Shannon Yeung, Data Science;

Other Advising

- Advanced Research Mentorship (ARM) program AY 2023, Naavya Jain (expected) and Aarav Mehta (expected), Glastonbury Highschool;
- Advanced Research Mentorship (ARM) program AY 2022, Arya Patel and Heather Newandee, Glastonbury Highschool;
- Advanced Research Mentorship (ARM) program AY 2021, Rachel Ha, Glastonbury Highschool;

- Advanced Research Mentorship (ARM) program AY 2019, Dhanush Kalangi, Glastonbury Highschool, received 2nd honors at Connecticut Science & Engineering Fair;
- University of Texas at El Paso BUILD Scholar Mentor (Summer 2021), Isaiah Perez
- Thesis Review Committee: Chao Shang, Timothy Becker
- Undergraduate Students Advising Total: ≈ 60
- Work Study, Christina Gorbenko, Computer Science and Engineering;

Academic Service

Conference and Workshop Organizing Committees

Organizing Committee Member, [SorinFest: Phase Transitions in Computer Science and Computational Biology](#), Brown University (2023)

Organizing Committee Member, [Rise Against the Machines: Challenges and Opportunities for AI in Industry](#), CT Advanced Computing Center, Organizing Committee Member (2023)

Organizing Committee Member, [Computer Science Teachers Association New England Annual Meeting: Building CS Pathways](#), Organizing Committee Member (2023)

Grant Proposal Review

Panel Reviewer, National Institutes of Health, SBIR/STTR Program, Study Section ZMD1 JT (J1) (2023)

Reviewer, Natural Sciences and Engineering Research Council of Canada (NSERC) (2022)

International Committees

Program Committee, ISMB (2018), BNP Workshop at NIPS (2018), RECOMB (2020, 2021, 2022, 2023), RECOMB Highlights (2023), BIBM (2019, 2020), AAAI (2021,2022), IJCAI (2021,2022,2024), ISBRA (2021, 2022, 2023), ICCABS (2023)

Session Chair, RECOMB (2020)

Editorial Boards

Journal of Computational Biology, 2021-

University Committees

Dean of Engineering Search 2023;

Departmental Committees

Workload Committee 2022-2023;

Undergraduate C&C Committee 2021-2022;

UG Life Committee 2021-2022;

Search Committee, Data Science RAP, Stamford 2021-2022;

MS DS Concentration Advisors 2021-2023;

Professional Education Committee 2020-2021;

Grad Life Committee 2020-2022;

Graduate C&C Committee 2020-2021;

Web and Publicity 2019-2023;

Open House 2019-2021;

Faculty Search Open 2018-2019, Systems and ML 2018-2019, Open, Security-SYF 2019-2020;

Peer-Review

Nature Communications, 2017, 2022; *Genome Research*, 2016, 2021; *Bioinformatics*, 2012-2019; *Proceedings of the National Academy of Sciences (PNAS)*, 2011; *Journal of Computational Biology (JCB)*, 2010-; *IEEE Transactions on Computational Biology and Bioinformatics*, 2020; *BMC Bioinformatics*, 2020; *Frontiers in Genetics* 2021; *Machine Learning* 2021-2022, *IEEE CCWC* 2022; *AI4TS2022 (IJCAI Workshop)* 2022, 2024; *BMC Supplements* 2023

Other Service

School of Engineering Mentoring Workshop Panelist. August 2022, 2023

University of Connecticut ENGR 1000 Talk. "Painting with cats and other nonsense" December 2022

Academic Web Administration and Development

- Built and maintained the [Princeton BEEHIVE](#), 2014-2018
- Built and maintained the [Istrail Lab Web](#); in 2012 rebuilt in PHP, 2009-2014
- Managed Brown University's Center for Computational Molecular Biology (CCMB) website, 2009-2011

Organizational Support

- Produced presentations and brochures for Brown University's CCMB Symposium, 2010
- Developed presentations and movies using Google Earth, Google SketchUp, and KML to support creation of the Ph.D. program for the Brown University's CCMB; Ph.D. program was approved in 2009

Other professional activities

- Senior Design Project Judge, UConn CSE, 2019, 2020, 2022, 2023
- North Carolina Student Academy of Science State Competition Judge, 2016-2017

Scholarships & Honors

- Student Nominated Graduate Faculty Career Mentor of the Year Finalist from the UConn Center for Career Development, 2023
- University Teaching Innovation Award Finalist from the Center for Excellence in Teaching & Learning, UConn, 2023
- Excellence in Mentoring Award, Glastonbury High School Advanced Research Mentorship, 2022
- Undergraduate Faculty Career Mentor of the Year Award Finalist from the UConn Center for Career Development, 2021
- Graduate Faculty Career Mentor of the Year Award Finalist from the UConn Center for Career Development, 2021
- 2x Mentorship Excellence Award Finalist from the UConn Office of Undergraduate Research, 2020, 2021
- Charles J. Epstein Trainee Awards for Excellence in Human Genetics Research semifinalist, 2015
- Travel Awards: Research in Computational Molecular Biology (RECOMB) Student Travel Fellowship, 2010, 2014; Intelligent Systems for Molecular Biology (ISMB) Student Travel Fellowship, 2012, 2013; National Science Foundation (NSF) Established Program to Stimulate Competitive Research (EPSCoR), 2013; IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS) Computational Advances for Next Generation Sequencing (CANGS) Workshop, 2012
- Outstanding Graduating Senior in Computer Engineering, URI, 2007
- Outstanding Junior in Computer Engineering, URI, 2006

Community Outreach

McNair Faculty Mentor, University of Connecticut	2021-
William J. Johnston Middle School Career Fair	2023
University of Texas at El Paso BUILD Scholar Mentor, University of Connecticut	Summer 2021
Advanced Research Mentorship program, Glastonbury High School, Glastonbury, CT	AYs 19, 21, 22

- Responsibilities include assisting mentee students in developing an appropriate and feasible research project.

Explore Engineering Young Engineering Science Scholars Program, School of Engineering, [UCONN](#), Storrs, CT 2019

- [Explore Engineering](#) is a one-week STEM summer camp for high school juniors and seniors where they are introduced to engineering fields.

- ▶ Students in my CS section were introduced to Programming, Data Science, Machine Learning and Artificial Intelligence through a series of lectures, activities, Jupyter notebooks, and programming exercises.
- ▶ On a scale of 1 (Unacceptable) to 5 (Outstanding), students rated the CS section an average score of 4.77.

Senior Design Project Judge, School of Engineering, UConn, Storrs, CT 2019-2020, 2022

- ▶ [Senior design projects](#) are the culmination of a two-semester course where students team up with faculty and industry sponsors to work collaboratively to solve a real-world problem.

Invention Convention Judge, Connecticut Invention Convention, Storrs, CT 2019-2022

- ▶ The [Connecticut Invention Convention](#) is an internationally recognized nonprofit educational program that encourages interest in STEM fields and develops critical thinking skills in children in grades K-12 through invention and innovation.

AI4All Postdoctoral Mentor, Fragile Families group, PRINCETON UNIVERSITY, Princeton, NJ 2018

- ▶ [AI4All](#) is a high school summer camp aimed at popularizing AI among groups underrepresented in the field.
- ▶ Responsibilities include mentoring the Fragile Families project team; development of curriculum and Jupyter notebooks detailing machine learning applications to the Fragile Families data.

Volunteer for Kids Run Durham, Bull City Running Company, Durham, NC 2015-2016

- ▶ [Kids Run Durham](#) is a program meant to encourage building healthy lifestyle habits, self-esteem, and character development, goal setting, and management in young children.
- ▶ Responsible for registration of children into their age-appropriate race-series and organizing the children into their respective groups at the race starting line.

Industry Experience

RAYTHEON INTEGRATED DEFENSE SYSTEMS (IDS), Portsmouth, RI 2006-2008

Software Engineer, 2007-2008

- ▶ Developed and tested data control and management software for the Zumwalt Total Ship Computing Environment Infrastructure.

Software Engineer, Mission Innovation (MI) Group, 2006-2007

- ▶ Applied company technology and resources to address world-threatening issues.
- ▶ Used Google Earth, KML, SQL, and .NET to create web-based collaborative disaster planning tool (Joint Rapid Integrated Planning Service); listed as co-inventor on Raytheon patent.

Memberships & Activities

International Society for Computational Biology (ISCB), 2012-

Institute of Electrical and Electronics Engineers (IEEE), 2006-

Association for Computing Machinery (ACM), 2006-

American Association for the Advancement of Science (AAAS), 2019-

Six Sigma Specialist, 2007

Phi Eta Sigma Honor Society, 2006

Tau Beta Pi Honor Society, 2006