Signature:

Name: Robert Patro
Date: May 5, 2025

Curriculum Vitae

I. Personal Information

I.A. UID, Last Name, First Name, Middle Name, Contact Information

UID: 107535276
Last Name: Patro
First Name: Robert
Middle Name: Scott
Mailing address:
3220 Brendan Iribe Center for Computer Science and Engineering
8125 Paint Branch Dr, College Park, MD 20740
email: rob@cs.umd.edu
Google Scholar: https://scholar.google.com/citations?user=H36hOqEAAAAJ

I.B. Academic Appointments at UMD

Associate Professor, Department of Computer Science, University of Maryland
Dates : July 2020 — current

Assistant Professor, Department of Computer Science, University of Maryland
Dates : July 2019 — July 2020

I.D. Other Employment

Assistant Professor, Department of Computer Science, Stony Brook University Dates: August 2014 — July 2019

Postdoctoral Research Associate, Lane Center for Computational Biology (now the Department of Computational Biology), School of Computer Science, Carnegie Mellon University
Dates: August 2012 — July 2014

Graduate Research Assistant, Department of Computer Science, University of Maryland
Dates: June 2007 — August 2012

Graduate Teaching Assistant, Department of Computer Science, University of

Maryland
Dates: September 2006 — June 2007

I.E. Educational Background

Degree: **PhD** in Computer Science
Institute*:* University of Maryland at College Park
Dates: 2006 — 2012

Degree*:* **BS** in Computer Science (cum laude & with departmental honors)
Institute: University of Maryland at College Park
Dates: 2003 — 2006

Degree*:* **Attended**
Institute*:* Worcester Polytechnic Institute
Dates*:* 2002

I.G. Professional Certifications, Licenses, and Memberships

International Society for Computational Biology (ISCB) Member
Dates: 2012—2014, 2016—current

ACM
Dates: 2023—current

ACM SigBIO Member
Dates: 2016, 2017

II. Research, Scholarly, Creative and/or Professional Activities

II.C. Refereed Journals

II.C.1 Refereed Journal Articles

Noor Pratap Singh, Euphy Y. Wu, Jason Fan, Michael I. Love, and Rob Patro. "Tree-based differential testing using inferential uncertainty for RNA-Seq." *Genome Research*, gr.279981.124. https://doi.org/10.1101/gr.279981.124. (online access ahead of Oct. issue).

Yuan Gao, Peng Jiang, Rob Patro, Collapsible tree: interactive web app to present collapsible hierarchies, *Bioinformatics*, 2024;, btae645, https://doi.org/10.1093/bioinformatics/btae645

Campanelli, Alessio, Giulio Ermanno Pibiri, Jason Fan, and Rob Patro. Where the Patterns Are: Repetition-Aware Compression for Colored de Bruijn Graphs. *Journal* *of Computational Biology*. 2024 Oct;31(10):1022-1044. doi: 10.1089/cmb.2024.0714. Epub 2024 Oct 9. PMID: 39381838.

Jamshed Khan, Tobias Rubel, Erin Molloy, Laxman Dhulipala, Rob Patro. (2024). Fast, parallel, and cache-friendly suffix array construction. *Algorithms for Molecular Biology*, *19*(1), 16. (substantially extended journal version of WABI 2023 conference paper)

Jason Fan, Jamshed Khan, Noor Pratap Singh, Giulio Ermanno Pibiri, and Rob Patro. "Fulgor: A fast and compact k-mer index for large-scale matching and color queries." *Algorithms for Molecular Biology* 19, no. 1 (2024): 1-21. (substantially extended journal version of WABI 2023 conference paper)

Simone Tiberi, Joël Meili, Peiying Cai, Charlotte Soneson, Dongze He, Hirak Sarkar, Alejandra Avalos-Pacheco, Rob Patro, Mark D Robinson, “*DifferentialRegulation*: a Bayesian hierarchical approach to identify differentially regulated genes”, *Biostatistics*, 2024;, kxae017, https://doi.org/10.1093/biostatistics/kxae017

Moein Karami, Aryan Soltani Mohammadi, Marcel Martin, Barış Ekim, Wei Shen, Lidong Guo, Mengyang Xu, Giulio Ermanno Pibiri, Rob Patro, Kristoffer Sahlin, “Designing efficient randstrobes for sequence similarity analyses”, *Bioinformatics*, Volume 40, Issue 4, April 2024, btae187, https://doi.org/10.1093/bioinformatics/btae187

Welles Robinson, Joshua K Stone, Fiorella Schischlik, Billel Gasmi, Michael C Kelly, Charlie Seibert, Kimia Dadkhah, E Michael Gertz, Joo Sang Lee, Kaiyuan Zhu, Lichun Ma, Xin Wei Wang, S Cenk Sahinalp, Rob Patro, Mark DM Leiserson, Curtis C Harris, Alejandro A Schäffer, Eytan Ruppin. “Identification of intracellular bacteria from multiple single-cell RNA-seq platforms using CSI-Microbes.” *Science advances*, *10*(27), eadj7402.

Dongze He, and Rob Patro. "simpleaf: A simple, flexible, and scalable framework for single-cell data processing using alevin-fry." *Bioinformatics* (2023): btad614.

Noor Pratap Singh, Michael I. Love, and Rob Patro. "TreeTerminus—creating transcript trees using inferential replicate counts." *iScience* 26.6 (2023).

Euphy Y. Wu, , Noor P. Singh, Kwangbom Choi, Mohsen Zakeri, Matthew Vincent, Gary A. Churchill, Cheryl L. Ackert-Bicknell, Rob Patro, and Michael I. Love. "SEESAW: detecting isoform-level allelic imbalance accounting for inferential uncertainty." *Genome biology* 24, no. 1 (2023): 165.

Lukas Heumos, Anna C. Schaar, Christopher Lance, Anastasia Litinetskaya, Felix Drost, Luke Zappia, Malte D. Lücken, Daniel C. Strobl, Juan Henao, Fabiola Curion, Hananeh Aliee, Meshal Ansari, Maren Büttner, Emma Dann, Daniel Dimitrov, Leander Dony, Amit Frishberg, Dongze He, Soroor Hediyeh-zadeh, Leon Hetzel, Ignacio L.

Ibarra, Matthew G. Jones, Mohammad Lotfollahi, Laura D. Martens, Christian L. Müller , Mor Nitzan , Johannes Ostner , Giovanni Palla, Rob Patro, Zoe Piran, Ciro Ramírez-Suástegui , Julio Saez-Rodriguez , Hirak Sarkar, Benjamin Schubert, Lisa Sikkema, Avi Srivastava, Jovan Tanevski, Isaac Virshup, Philipp Weiler, Herbert Schiller, and Fabian J. Theis. "Best practices for single-cell analysis across modalities". *Nature Reviews Genetics* (2023): 1-23.

Jamshed Khan, Marek Kokot, Sebastian Deorowicz, and Rob Patro. "Scalable, ultra-fast, and low-memory construction of compacted de Bruijn graphs with Cuttlefish 2." *Genome biology* 23, no. 1 (2022): 1-32.

Dongze He, Mohsen Zakeri, Hirak Sarkar, Charlotte Soneson, Avi Srivastava, Rob Patro. Alevin-fry unlocks rapid, accurate, and memory-frugal quantification of single-cell RNA-seq data. *Nature Methods* 19, no. 3 (2022): 316-322.

Fatemeh Almodaresi, Jamshed Khan, Sergey Madaminov, Prashant Pandey, Michael Ferdman, Rob Johnson, Rob Patro. An Incrementally Updatable and Scalable System for Large-Scale Sequence Search using the Bentley-Saxe Transformation. Bioinformatics, 2022; btac142, https://doi.org/10.1093/bioinformatics/btac142

Fan, Jason, Skylar Chan, and Rob Patro. Perplexity: evaluating transcript abundance estimation in the absence of ground truth. Algorithms for Molecular Biology 17, no. 1 (2022): 1-23. *(Extended journal version of WABI conference paper of the same title.)*

Giorgos Skoufos, Fatemeh Almodaresi, Mohsen Zakeri, Joseph N Paulson, Rob Patro, Artemis G Hatzigeorgiou, Ioannis S Vlachos. AGAMEMNON: an Accurate metagenomics and metatranscriptomics quantification analysis suite. Genome biology 23, no. 1 (2022): 1-27.

Wancen Mu, Hirak Sarkar, Avi Srivastava, Kwangbom Choi, Rob Patro, Michael I Love, Airpart: interpretable statistical models for analyzing allelic imbalance in single-cell datasets, *Bioinformatics*, Volume 38, Issue 10, 15 May 2022, Pages 2773–2780, https://doi.org/10.1093/bioinformatics/btac212

Fatemeh Almodaresi, Mohsen Zakeri, Rob Patro. PuffAligner: a fast, efficient and accurate aligner based on the Pufferfish index, *Bioinformatics*, 2021;, btab408, https://doi.org/10.1093/bioinformatics/btab408

Preprocessing choices affect RNA velocity results for droplet scRNA-seq data. Charlotte Soneson, Avi Srivastava, Rob Patro, Michael B. Stadler (2021). PLoS Comput Biol 17(1): e1008585. https://doi.org/10.1371/journal.pcbi.1008585

Compression of quantification uncertainty for scRNA-seq counts.
Scott Van Buren, Hirak Sarkar, Avi Srivastava, Naim U. Rashid, Rob Patro,
And Michael I. Love (2021). , *Bioinformatics*, 2021;
btab001, https://doi.org/10.1093/bioinformatics/btab001

Avi Srivastava, Laraib Malik, Hirak Sarkar, Mohsen Zakeri, Fatemeh Almodaresi, Charlotte Soneson, Mike Love, Carl Kingsford, Rob Patro (2020). Alignment and mapping methodology influence transcript abundance estimation. Genome biology, 21(1), 1-29.

Michael I. Love, Charlotte Soneson, Peter F. Hickey, Lisa K. Johnson, N. Tessa Pierce, Lori Shepherd, Martin Morgan, Rob Patro. "Tximeta: Reference sequence checksums for provenance identification in RNA-seq." PLoS computational biology 16.2 (2020): e1007664. *PLoS Computational Biology, 16(2), e1007664, 2020.*

Anqi Zhu, Avi Srivastava, Joseph G. Ibrahim, Rob Patro, Michael I. Love. Nonparametric expression analysis using inferential replicate counts. *Nucleic Acids Research*, gkz622, 2019.

Koen Van Den, Katharina M. Hembach, Charlotte Soneson, Simone Tiberi, Lieven Clement, Michael I. Love, Rob Patro, and Mark D. Robinson. "RNA sequencing data: hitchhiker's guide to expression analysis." *Annual Review of Biomedical Data Science* 2, 2019.

Guillaume Marçais, Brad Solomon, Rob Patro, and Carl Kingsford. Sketching and sublinear data structures in genomics. *Annual Review of Biomedical Data Science*, *2*, 2019.

Avi Srivastava, Laraib Malik, Tom Smith, Ian Sudbery, and Rob Patro. Alevin efficiently estimates accurate gene abundances from dscrna-seq data. *Genome Biology*, 20(1):65, 2019

Charlotte Soneson, Michael I Love, Rob Patro, Shobbir Hussain, Dheeraj Malhotra, and Mark D Robinson. A junction coverage compatibility score to quantify the reliability of transcript abundance estimates and annotation catalogs. *Life science alliance*, 2(1):e201800175, 2019.

Michael I. Love, Charlotte Soneson, and Rob Patro. Swimming downstream: statistical analysis of differential transcript usage following Salmon quantification. *F1000Research*, 7:952, Oct. 2018

Moriah L Jacobson, Lydia A Kim, Robert Patro, Barbara Rosati, and David McKinnon. Common and differential transcriptional responses to different models of traumatic stress exposure in rats. *Nature Translational Psychiatry*, 8(1):165, 2018

Laraib Malik, Fatemeh Almodaresi, and Rob Patro. Grouper: graph-based clustering and annotation for improved de novo transcriptome analysis. *Bioinformatics*, 34(19):3265–3272, 2018.

Laraib Malik and Rob Patro. Rich Chromatin Structure Prediction from Hi-C Data.

IEEE/ACM Transactions on Computational Biology and Bioinformatics, pages 1–12, 2018. (**extended** journal version of ACM-BCB paper by the same name)

Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and Rob Patro. Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index. *Cell Systems*, 7, no. 2 (2018): 201-207. (Journal version of RECOMB 2018 paper)

Gruning, B., et al. "Bioconda: sustainable and comprehensive software distribution for the life sciences." *Nature Methods* 15.7 (2018): 475-476. (**listed as part of "The Bioconda Team" available in the Consortium author list)**

Rob Patro, Geet Duggal, Michael I Love, Rafael A Irizarry, and Carl Kingsford. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4):417–419, March 2017

Prashant Pandey, Michael A Bender, Rob Johnson, and Rob Patro. Squeakr: an exact and approximate k-mer counting system. Bioinformatics, 34(4):568–575, 2017.

Hirak Sarkar and Rob Patro. Quark enables semi-reference-based compression of RNA-seq data. *Bioinformatics*, 33(21):3380–3386, 2017

Avi Srivastava, Hirak Sarkar, Nitish Gupta, and Rob Patro. RapMap: a rapid, sensitive and accurate tool for mapping RNA-seq reads to transcriptomes. *Bioinformatics*, 32(12):i192–i200, June 2016. Proceedings of ISMB 2016

Rob Patro, Raquel Norel, Robert J. Prill, Julio Saez-Rodriguez, Peter Lorenz, Felix Steinbeck, Bjoern Ziems, Mitja Luštrek, Nicola Barbarini, Alessandra Tiengo, Riccardo Bellazzi, Hans-Jürgen Thiesen, Gustavo Stolovitzky, and Carl Kingsford. A computational method for designing diverse linear epitopes including citrullinated peptides with desired binding affinities to intravenous immunoglobulin. *BMC Bioinformatics*, 17(1):1–13, 2016

Richard Smith-Unna, Chris Boursnell, Rob Patro, Julian M Hibberd, and Steven Kelly. TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Research*, 26(8):1134–1144, 2016

Carl Kingsford and Rob Patro. Reference-based compression of short-read sequences using path encoding. *Bioinformatics*, 31(12):1920–1928, 2015

Rob Patro and Carl Kingsford. Data-dependent bucketing improves reference-free compression of sequencing reads. *Bioinformatics*, 31(17):2770–2777, 2015.

Runxuan Zhang, Cristiane PG Calixto, Nikoleta A Tzioutziou, Allan B James, Craig G Simpson, Wenbin Guo, Yamile Marquez, Maria Kalyna, Rob Patro, Eduardo Eyras, et al. AtRTD–a comprehensive reference transcript dataset resource for accurate quantification of transcript-specific expression in arabidopsis thaliana. *New*

Phytologist, 2015

Darya Filippova*, Rob Patro*, Geet Duggal*, and Carl Kingsford. Identification of Alternative Topological Domains in Chromatin. *Algorithms for Molecular Biology*, 9(1):1, 2014

Rob Patro, Stephen M Mount, and Carl Kingsford. Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. *Nature Biotechnology*, 32(5):462–464, 2014.

Geet Duggal, Rob Patro, Emre Sefer, Hao Wang, Darya Filippova, Samir Khuller, and Carl Kingsford. Resolving spatial inconsistencies in chromosome conformation measurements. *Algorithms for Molecular Biology*, 8(1):8, 2013. (**extended** journal version of WABI 2012 paper which is not listed here).

Rob Patro, John Dickerson, Sujal Bista, Satyandra Gupta, and Amitabh Varshney. Speeding Up Particle Trajectory Simulations under Moving Force Fields using GPUs. *ASME Journal of Computing and Information Science in Engineering*, 12(2), 2012.

Rob Patro and Carl Kingsford. Global network alignment using multiscale spectral signatures. *Bioinformatics*, 28(23):3105–3114, 2012.

Rob Patro, Emre Sefer, Justin Malin, Guillaume Marçais, Saket Navlakha, and Carl Kingsford. Parsimonious reconstruction of network evolution. *Algorithms for Molecular Biology*, 7(1):25, 2012. (**extended** journal version of WABI 2011 paper below).

Youngmin Kim, Rob Patro, Cheuk Yiu Ip, Dianne P. O'Leary, and Andriy Anishkin. Salient Frame Detection for Molecular Dynamics Simulations. In *Scientific Visualization: Interactions, Features, Metaphors*, volume 2 of *Dagstuhl Follow-Ups*, pages 160–175. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2011.

Andre Maximo, Rob Patro, Amitabh Varshney, and Ricardo Farias. A Robust and Rotationally Invariant Local Surface Descriptor with Applications to Non-local Mesh Processing. *Graphical Models*, 73(5):231–242, 2011.

Rob Patro, Cheuk Yiu Ip, Sujal Bista, and Amitabh Varshney. Social Snapshot: A System for Temporally Coupled Social Photography. *Computer Graphics and Applications, IEEE*, 31(1):74–84, 2011.

Rob Patro, Cheuk Yiu Ip, and Amitabh Varshney. Saliency Guided Summarization of Molecular Dynamics Simulations. In *Scientific Visualization: Advanced Concepts*, volume 1 of *Dagstuhl Follow-Ups*, pages 321–335. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2010.

Aswin C. Sankaranarayanan, Rob Patro, Pavan Turaga, Amitabh Varshney, and Rama

Chellappa. Modeling and Visualization of Human Activities for Multicamera Networks. *EURASIP Journal on Image and Video Processing*, vol. 2009(Article ID 259860), 2009.

II.D. Published Conference Proceedings

II.D.1 Refereed Conference Proceedings (*denotes equal contribution)

Alessio Campanelli, Giulio Ermanno Pibiri, and Rob Patro. Fast Pseudoalignment Queries on Compressed Colored de Bruijn Graphs. In 25th International Conference on Algorithms for Bioinformatics (WABI 2025). Leibniz International Proceedings in Informatics (LIPIcs), Volume 344, pp. 6:1-6:21, Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2025) https://doi.org/10.4230/LIPIcs.WABI.2025.6

Zahra Zare Jousheghani, Noor Pratap Singh, Rob Patro, Oarfish: enhanced probabilistic modeling leads to improved accuracy in long read transcriptome quantification, Bioinformatics, Volume 41, Issue Supplement_1, July 2025, Pages i304–i313, https://doi.org/10.1093/bioinformatics/btaf240 (Proceedings of ISMB 2025)

Noor Pratap Singh, Jamshed Khan, Rob Patro, Alevin-fry-atac enables rapid and memory frugal mapping of single-cell ATAC-seq data using virtual colors for accurate genomic pseudoalignment, Bioinformatics, Volume 41, Issue Supplement_1, July 2025, Pages i237–i245, https://doi.org/10.1093/bioinformatics/btaf234 (Proceedings of ISMB 2025)

Lorraine Ayad, Gabriele Fici, Ragnar Groot Koerkamp, Grigorios Loukides, Rob Patro, Giulio Ermanno Pibiri and Solon Pissis. “A Universal Indexing Framework for Matching Long Patterns” Schloss Dagstuhl – Leibniz-Zentrum für Informatik. https://doi.org/10.4230/LIPICS.SEA.2025.4. (Preceedings of the Symposium on Experimental Algorithms (SEA) 2025)

Dongze He, Yuan Gao, Spencer Skylar Chan, Natalia Quintana-Parrilla, and Rob Patro. "Forseti: A mechanistic and predictive model of the splicing status of scRNA-seq reads." *Bioinformatics* 40, no. Supplement_1 (2024): i297-i306. (Proceedings of ISMB 2024)

Giulio Ermanno Pibiri, Jason Fan, and Rob Patro. “Meta-colored compacted de Bruijn graphs”.In *International Conference on Research in Computational Molecular Biology* (pp. 131-146) (Proceedings of RECOMB 2024)..

Jamshed Khan, Tobias Rubel, Laxman Dhulipala, Erin Molloy, and Rob Patro. "Fast, Parallel, and Cache-Friendly Suffix Array Construction." Dagstuhl Research Online Publication Server, edited by Djamal Belazzougui and Aïda Ouangraoua, vol. 273, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, pp. 16:1-21, https://doi.org/10.4230/LIPIcs.WABI.2023.16 (Proceedings of WABI 2023)

Jason Fan, Noor Pratap Singh, Jamshed Khan, Giulio Ermanno Pibiri, Rob Patro. "Fulgor: A Fast and Compact k-mer Index for Large-Scale Matching and Color Queries." Dagstuhl Research Online Publication Server, edited by Djamal Belazzougui and Aïda Ouangraoua, vol. 273, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, pp. 18:1–21, https://doi.org/10.4230/LIPIcs.WABI.2023.18. (Proceedings of WABI 2023)

Hossein Asghari, Ehsan Haghshenas, Roby Thomas, Eric Schultz, Rob Patro, Stan Skrzypczak, Carl Kingsford; Abstract 1400: Novel expression biomarkers via prediction of response to FOLFIRINOX (FFX) treatment for PDAC. Cancer Res 1 April 2023; 83 (7_Supplement): 1400. https://doi.org/10.1158/1538-7445.AM2023-1400 (AACR 2023)

Jason Fan, Jamshed Khan, Giulio Ermanno Pibiri, and Rob Patro. "Spectrum preserving tilings enable sparse and modular reference indexing." In *International Conference on Research in Computational Molecular Biology*, pp. 21-40. Cham: Springer Nature Switzerland, 2023. (Proceedings of RECOMB 2023)

TreeTerminus - Creating transcript trees using inferential replicate counts
Noor Pratap Singh, Michael I. Love, Rob Patro
bioRxiv 2022.11.01.514769; doi: https://doi.org/10.1101/2022.11.01.514769. (Appeared in proceedings of RECOMB-SEQ 2023; **won best paper award**)

Jamshed Khan, Rob Patro, Cuttlefish: fast, parallel and low-memory compaction of de Bruijn graphs from large-scale genome collections, Bioinformatics, Volume 37, Issue Supplement_1, July 2021, Pages i177–i186, https://doi.org/10.1093/bioinformatics/btab309. (appeared at ISMB 2021; **won best paper award**)

Hirak Sarkar, Avi Srivastava, Héctor Corrada Bravo, Michael I Love, Rob Patro, Terminus enables the discovery of data-driven, robust transcript groups from RNA-seq data, Bioinformatics, Volume 36, Issue Supplement_1, July 2020, Pages i102–i110, https://doi.org/10.1093/bioinformatics/btaa448. (appeared at ISMB 2020)

Avi Srivastava, Laraib Malik, Hirak Sarkar, Rob Patro, A Bayesian framework for inter-cellular information sharing improves dscRNA-seq quantification, Bioinformatics, Volume 36, Issue Supplement_1, July 2020, Pages i292–i299, https://doi.org/10.1093/bioinformatics/btaa450. (appeared at ISMB 2020)

Hirak Sarkar, Avi Srivastava, and Rob Patro Minnow: a principled framework for rapid simulation of dscRNA-seq data at the read level. *Bioinformatics* 35.14 (2019): i136-i144. (*ISMB 2019 Special Issue).*

Fatemeh Almodaresi, Prashant Pandey, Michael Ferdman, Rob Johnson, and Robert Patro. An efficient, scalable and exact representation of high-dimensional color information enabled via de bruijn graph search. In Lenore J. Cowen, editor, *Research in Computational Molecular Biology - 23rd Annual International Conference, RECOMB 2019, Washington, DC, USA, May 5-8, 2019, Proceedings*, volume 11467 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2019. (appeared at RECOMB 2019)

Fatemeh Almodaresi, Hirak Sarkar, Avi Srivastava, and Rob Patro. A space and time-efficient index for the compacted colored de Bruijn graph. *Bioinformatics*, 34(13):i169–i177, 2018. (special Bioinformatics issue with proceedings of ISMB 2018)

Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and Rob Patro. Mantis: A Fast, Small, and Exact Large-Scale Sequence Search Index. volume 10812 of *Lecture Notes in Computer Science*, pages 271–273. Springer, 2018. (Proceedings of RECOMB 2018; published as extended abstract, since invited to publication in Cell Systems).

Hirak Sarkar, Mohsen Zakeri, Laraib Malik, and Rob Patro. Towards Selective-Alignment: Bridging the Accuracy Gap Between Alignment-Based and Alignment-Free Transcript Quantification. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, BCB '18, pages 27–36, New York, NY, USA, 2018. ACM.

Laraib Malik and Rob Patro. Rich chromatin structure prediction from Hi-C data. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 184–193. ACM, 2017. (**won best paper award**).

Prashant Pandey, Michael A. Bender, Rob Johnson, and Rob Patro. A General-Purpose Counting Filter: Making Every Bit Count. In *Proceedings of the 2017 ACM International Conference on Management of Data*, SIGMOD '17, pages 775–787, New York, NY, USA, 2017. ACM

Prashant Pandey, Michael A Bender, Rob Johnson, and Rob Patro. deBGR: an efficient and near-exact representation of the weighted de bruijn graph. *Bioinformatics*, 33(14):i133–i141, 2017. (special Bioinformatics issue with proceedings of ISMB 2017)

Mohsen Zakeri, Avi Srivastava, Fatemeh Almodaresi, and Rob Patro. Improved data-driven likelihood factor- izations for transcript abundance estimation. *Bioinformatics*, 33(14):i142–i151, 2017. (special Bioinformatics issue with proceedings of ISMB 2017)

Rob Patro and Carl Kingsford. Predicting protein interactions via parsimonious network history inference. *Bioinformatics*, 29(13):i237–i246, 2013. (special

Bioinformatics issue with proceedings of ISMB/ECCB 2013)

Hao Wang, Geet Duggal, Rob Patro, Michelle Girvan, Sridhar Hannenhalli, and Carl Kingsford. Topological Properties of Chromosome Conformation Graphs Reflect Spatial Proximities Within Chromatin. In *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics*, BCB'13, pages 306:306–306:315, New York, NY, USA, 2013. ACM

Robert Patro*, Geet Duggal*, Emre Sefer, Hao Wang, Darya Filippova, and Carl Kingsford. The missing models: a data-driven approach for learning how networks grow. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 42–50. ACM, 2012.

Rob Patro, Cheuk Yiu Ip, Sujal Bista, Samuel Cho, Dave Thirumalai, and Amitabh Varshney. MDMap : A system for data-driven layout and exploration of molecular dynamics simulations. In *Proceedings of the 1st IEEE symposium on biological data visualization*, pages 111–118, Providence, RI, USA, 2011.

II.E. Conferences, Workshops, and Talks

II.E.2. Invited Talks

Counting is not easy: Assessing and quantifying uncertainty in abundance inferences from high-throughput sequencing data. Invited Talk at the Models, Inference & Algorithms (**MIA**) seminar series at the *Broad Institute* (February 2025). (talk can be viewed at: https://youtu.be/cAJNqu50YkM?si=BAGeDs8121yWGuUi)

The Matrix Has You : The Importance of Accurate, Efficient and Open Methods Upstream of Single-Cell Count Matrices. Invited talk: Vancouver Bioinformatics User Group (VanBUG; February 2025) (talk can be viewed at: https://youtu.be/fQD7JggiSl0?si=t9_RDz1yNNB5kbKz&t=1077)

Rob Patro. Toward efficient, accurate, scalable, and reproducible pre-processing of high-throughput single-cell sequencing data. **Keynote talk**: scVerse Conference (Technical University of Munich; Sept. 2024)

Rob Patro. Counting isn't easy: Computational successes and inferential challenges in the processing of single-cell RNA-seq data. Colloquium at the Center for Computational Mathematics; Flatiron institute (April 2024).

Rob Patro. What could be next(flow): Workflows as augmented execution contexts to improve provenance tracking and interoperability. **Keynote address:** Nextflow summit 2022. (talk can be viewed at: https://www.youtube.com/watch?v=vNrKFT5eT8U)

Rob Patro. Keeping k-mers in check—Building fast, small, and composable indices based on the De Bruijn graph. Biological Data Science Meeting, Cold Spring Harbor Laboratory. 2022.

Rob Patro. Computational Successes and Challenges in the Analysis of Single-cell RNA Sequencing Data. Genetics, Bioinformatics, & Systems Biology Colloquium at University of California, San Diego. 2022.

Accurate, fast, and memory-frugal methods for processing single-cell RNA-seq data. AI Summer School (Özyeğin University; remote zoom talk). 202

Computational tools for accurate, fast and memory-frugal analysis of bulk and single-cell RNA-seq data. BrazWebinars (Brazilian Webinar series in Bioinformatics and Computational Biology). 2022 Can be viewed at: https://youtu.be/xtizmT0U6z8

Rob Patro. Judging a k-mer by The Company it Keeps: Compressing the de Bruijn Graph and Signals That Live On It. *Compression + Computation 2022 workshop.*

Rob Patro. Algorithms you can count on: Accurate, fast, and memory-frugal single-cell processing with alevin-fry. *Tufts University, Bioinformatics and Computational Biology Research Group, 2021.*

Rob Patro. Tell me something I don't know : The role of inferential uncertainty in bulk and single-cell RNA-seq analysis. *CCB Seminar, Duke University, 2021.*

Rob Patro. Tell me something I don't know : The role of inferential uncertainty in RNA-seq analysis. *Quantitative Science Seminar, Johns Hopkins University, 2021.*

Rob Patro. The known unknowns : Addressing uncertainty in bulk and single-cell RNA-seq quantification. *Bioinformatics Brownbag Seminar, Pacific Biosciences, 2020.* (postponed due to COVID-19 pandemic)

Rob Patro. Single-cell transcriptome analysis: opportunities and computational challenges. *Molecular Medicine Tri-Conference, 2020.*

Rob Patro. The known unknowns : Addressing uncertainty in bulk and single-cell RNA-seq quantification. *TorBUG Seminar, 2020*

Rob Patro. Alignment and Mapping Methodology Influence Transcript Abundance Estimation. *Genome Informatics, 2019*

Rob Patro. Mapping and Modeling Methodology Matter: Making the Most of your RNA-seq Data. *CBBG Seminar Series, 2019*

Rob Patro. Alignment: Back to the Future. *CBBG Program Retreat, 2019*

Rob Patro. Alignment: Back to the Future. *BioC conference, 2019*

Rob Patro. Improving Computational Methods for Single-Cell Gene Expression Analysis. *Gloria and Mark Snyder Symposium on Cancer Medicine*, 2019

Rob Patro. From single cells to thousands of genomes: computational challenges and algorithmic solutions in high-throughput genomics. *UMD*, February 2019

Rob Patro. The past, present and future of transcript quantification: recent progress & some remaining challenges. *EMBL-EBI Industry Workshop: The past, present and future of RNA-seq technology and its application to drug discovery*, November 2018

Rob Patro. Drinking from the firehose: lightweight methods and data structures for indexing, searching and analyzing genomic data. *UMD (Computational Biology Seminar)*, October 2018

Rob Patro. Drinking from the firehose: lightweight methods and data structures for indexing, searching and analyzing genomic data. *USC (Computational Biology Seminar)*, October 2018

Rob Patro. Deconvolution, dictionaries and de Bruijn graphs : algorithm and data structure design for modern genomics. *Penn State Workshop on Emerging Methods for Sequence Analysis*, June 2018

Rob Patro. When less enables more: making models and methods for modern genomics. *MIT (Computational Biology Seminar)*, March 2018

Rob Patro. Faster, Smaller, Better: How improved methods can facilitate better biology. *University of Basel (Computational Biology Seminar)*, November 2017

Rob Patro. Algorithmic advances in high-throughput transcriptome analysis. *Calico*, July 2017

Rob Patro. Algorithmic advances in high-throughput transcriptome analysis. *GRAIL*, July 2017

Rob Patro. Don't count on it: Pragmatic and theoretical concerns and best practices for mapping and quantifying RNA-seq data. *DIBSI 2017 training (UC Davis)*, July 2017

Rob Patro. Algorithmic advances in transcript quantification. *HiCOMB 2017 (IEEE International Workshop on High Performance Computational Biology)*, May 2017

Rob Patro. Accurate, fast and bias-aware transcript quantification with Salmon. *Biostatistics & Medical Informatics Seminar Series, University of Wisconsin, Madison*, September 2016

Rob Patro. A New Deal for Transcriptomics: Better Biology through Better Algorithms. *Laufer Center for Quantitative Biology Seminar*, 2016

Rob Patro. Fast Algorithms for Improved Transcriptome Analysis. *Genomics JHU seminar series, Johns Hopkins University*, 2016

Rob Patro. Rethinking the big BAM: Analysis-efficient computing for improved transcriptome analysis. *Weill Cornell Medical School*, 2016

Rob Patro. Rethinking the big BAM: Fast algorithms for improved transcriptome analysis. *Personal Genome Diagnostics*, 2016.

Rob Patro. The wait is over: lightweight methods for accurate, fast and scalable analysis of transcriptomes. *New York Genome Center,* September 2015.

Rob Patro, Stephen Mount, and Carl Kingsford. Salmon spawns the next the next generation of RNAseq analysis with versatile and ultrafast transcript quantification. *Simons Center for Quantitative Biology Seminar*, March 2015.

Rob Patro, Stephen Mount, and Carl Kingsford. Sailfish: Expression Estimates Shouldn't Take Longer than a Cup of Coffee. *Automated Personal Genome Analysis for Clinical Advisors*, 2013.

II.E.3. Refereed Presentations

I include below only the refereed presentations which I have given, not those for which I am a co-author but which were presented by another co-author.

Rob Patro. Fulgor: A fast and compact k-mer index for large-scale matching and color queries. WABI 2023 Conference

Mohsen Zakeri, Avi Srivastava, Fatemeh Almodaresi, and Rob Patro. Improved data-driven likelihood factorizations for transcript abundance estimation. *ISMB 2017*.

Avi Srivastava, Larib Malik, and Rob Patro. Accurate, Fast and Lightweight Clustering of *de novo* Transcriptomes using Fragment Equivalence Classes. *RECOMB-Seq 2016.*

Komal Sanjeev, Nitish Gupta, Tim Wall, Carl Kingsford, and Rob Patro. Efficient Index Maintenance Under Dynamic Genome Modification. *RECOMB-Seq 2016.*

Rob Patro, Stephen Mount, and Carl Kingsford. Sailfish and Salmon enable alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. *RECOMB 2015 Highlight Talk*, April 2015.

Rob Patro, Stephen Mount, and Carl Kingsford. Sailfish — RNA-seq Expression Estimates Need Not Take Longer Than a Cup of Coffee. *Genome Informatics*, October

2013.

Rob Patro and Carl Kingsford. Predicting protein interactions via parsimonious network history inference. ISMB 2013*.*

Rob Patro and Carl Kingsford. Global Network Alignment Using Multiscale Spectral Signatures. *GLBIO Conference*, 2013.

Rob Patro, Cheuk Yiu Ip, Sujal Bista, Samuel Cho, Dave Thirumalai, and Amitabh Varshney. MDMap : A system for data-driven layout and exploration of molecular dynamics simulations. *IEEE BioVis Conference*, 2011.

Rob Patro and Carl Kingsford. Learning from Diversity: Epitope Prediction with Sequence and Structure Features using an Ensemble of Support Vector Machines. *RECOMB Systems Biology Satellite Conference*, 2010. (presenting our prediction algorithm that won the DREAM5 Challenge 1 competition)

II.E.4. Refereed Workshop Papers (*denotes equal contribution)

Jason Fan, Skylar Chan, and Rob Patro. "Perplexity: Evaluating Transcript Abundance Estimation in the Absence of Ground Truth." 21st International Workshop on Algorithms in Bioinformatics (WABI 2021). Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.

Fatemeh Almodaresi, Prashant Pandey, and Rob Patro. Rainbowfish: A Succinct Colored de Bruijn Graph Representation. In Russell Schwartz and Knut Reinert, editors, 17th International Workshop on Algorithms in Bioinformatics (WABI 2017), volume 88 of Leibniz International Proceedings in Informatics (LIPIcs), pages 18:1–18:15, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik

Komal Sanjeev, Nitish Gupta, Tim Wall, Carl Kingsford, and Rob Patro. Efficient Index Maintenance Under Dynamic Genome Modification. *Proceedings of RECOMB-Seq 2016*, April 2016

Avi Srivastava, Larib Malik, and Rob Patro. Accurate, Fast and Lightweight Clustering of *de novo* Transcriptomes using Fragment Equivalence Classes. *Proceedings of RECOMB-Seq 2016*, April 2016

Zhong Sichen, Lu Zhao, Yan Liang, Mohammadzaman Zamani, Rob Patro, Rezaul Chowdhury, Esther M Arkin, Joseph SB Mitchell, and Steven Skiena. Optimizing Read Reversals for Sequence Compression. In *Algorithms in Bioinformatics* (WABI 2015), pages 189–202. Springer, 2015.

Darya Filippova*, Rob Patro*, Geet Duggal*, and Carl Kingsford. Multiscale Identification of Topological Domains in Chromatin. In *Algorithms in Bioinformatics* (WABI 2013), pages 300–312. Springer, 2013.

Rob Patro, Emre Sefer, Justin Malin, Guillaume Marçais, Saket Navlakha, and Carl Kingsford. Parsimonious Reconstruction of Network Evolution. In Teresa Przytycka and Marie-France Sagot, editors, *Algorithms in Bioinformatics* (WABI 2011), volume 6833 of *Lecture Notes in Computer Science*, pages 237–249. Springer Berlin / Heidelberg, 2011.

II.E.6. Refereed Posters

Noor Pratap Singh, Michael I Love, Rob Patro. Tree Terminus—Creating transcript trees using inferential replicate counts. CSHL Biological Data Science 2022.

Noor Pratap Singh, Euphy Wu, Michael I Love, Rob Patro. Tree-based differential testing using inferential replicate counts for RNASeq data. CSHL Biological Data Science 2022.

Hirak Sarkar, Avi Srivastava, Mohsen Zakeri, Scott Van Buren, Naim U Rashid, Michael Love, Rob Patro. Accurate, efficient, and uncertainty-aware expression quantification of single-cell RNA-seq data. CSHL Biological Data Science 2020.

Giorgos Skoufos, Fatemeh Almodaresi, Mohsen Zakeri, Joseph Paulson, Rob Patro, Artemis Hatzigeorgiou, Ioannis Vlachos. Time and space-efficient k-mer-based microbial abundance quantification at strain resolution. HSCBB 2019.

Avi Srivastava, Laraib Iqbal Malik, Tom Smith, Ian Sudbery and Rob Patro.
Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *BioC* conference 2019.

Fatemeh Almodaresi, Mohsen Zakeri and Rob Patro
Cedar: scalable, accurate and fast metagenomic abundance estimation. *RECOMB* 2019.

Avi Srivastava, Laraib Iqbal Malik, Tom Smith, Ian Sudbery and Rob Patro.
Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *RECOMB* 2019.

Fatemeh Almodaresi, Prashant Pandey, and Rob Patro. Rainbowfish: A Succinct Colored de Bruijn Graph Representation. *WABI*, 2017.

Fatemeh Almodaresi, Hirak Sarkar, and Rob Patro. A space and time-efficient index for the compacted colored de Bruijn graph. *ACM BCB*, 2017.

Laraib Malik and Rob Patro. Graph regularized, semi-supervised learning improves annotation of de novo transcriptomes. *CSHL BioData*, 2016.

Rob Patro and Carl Kingsford. Global Network Alignment Using Multiscale Spectral Signatures. *GLBIO Conference*, 2013.

Rob Patro and Carl Kingsford. Global Network Alignment Using Multiscale Spectral Signatures. *Systems Biology: Networks (Cold Spring Harbor Laboratory Meeting)*, 2013.

Rob Patro and Carl Kingsford. Global Network Alignment Using Multiscale Spectral Signatures. *ISMB*, 2012.

Rob Patro and Carl Kingsford. Learning from Diversity: Epitope Prediction with Sequence and Structure Features using an Ensemble of Support Vector Machines. *RECOMB Systems Biology Satellite Conference*, 2010.

II.F. Professional and Extension Publications

II.F.2. Pre-print / Working Paper *(Not Work in Progress)*

mim: A lightweight auxiliary index to enable fast, parallel, gzipped FASTQ parsing
Rob Patro, Siddhant Bharti, Prajwal Singhania, Rakrish Dhakal, Thomas J. Dahlstrom, Ragnar Groot Koerkamp
bioRxiv 2025.11.24.690271; doi: https://doi.org/10.1101/2025.11.24.690271

Victor Levallois, Yoshihiro Shibuya, Bertrand Le Gal, Rob Patro, Pierre Peterlongo, Giulio Ermanno Pibiri “Kaminari: a resource-frugal index for approximate colored k -mer queries”. *bioRxiv* (2025); url: https://doi.org/10.1101/2025.05.16.654317

Jamshed Khan, Laxman Dhulipala, Rob Patro. “Fast and Scalable Parallel External-Memory Construction of Colored Compacted de Bruijn Graphs with Cuttlefish 3”. *bioRxiv* (2025); url: https://doi.org/10.1101/2025.02.02.636161

Zahra Zare Jousheghani, Noor Pratap Singh, Rob Patro, “A replicable and modular benchmark for long-read transcript quantification methods”. *bioRxiv* (2024), 2024-07. url: https://www.biorxiv.org/content/10.1101/2024.07.30.605821v1

Dongze He, Stephen M. Mount, and Rob Patro. "scCensus: Off-target scRNA-seq reads reveal meaningful biology." bioRxiv (2024): 2024-01.

Dongze He, Charlotte Soneson, Rob Patro. “Understanding and evaluating ambiguity in single-cell and single-nucleus RNA-sequencing”
bioRxiv 2023.01.04.522742; url: https://doi.org/10.1101/2023.01.04.522742

A like-for-like comparison of lightweight-mapping pipelines for single-cell RNA-seq data pre-processing. Mohsen Zakeri, Avi Srivastava, Hirak Sarkar, Rob Patro. bioRxiv 2021.02.10.430656; doi: https://doi.org/10.1101/2021.02.10.430656

Accounting for fragments of unannotated origin reduces the “effects of transcriptional noise on estimates of gene and transcript expression in RNA sequencing experiments". Avi Srivastava, Mohsen Zakeri, Hirak Sarkar, Charlotte Soneson, Carl Kingsford, and Rob Patro. bioRxiv 2021.01.17.426996; doi: https://doi.org/10.1101/2021.01.17.426996.

Mikhail Lipatov, Komal Sanjeev, Rob Patro, Krishna R Veeramah. Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. bioRxiv 023374; doi: https://doi.org/10.1101/023374

II.G. Book Reviews, Notes, and Other Contributions

Dongze He, Avi Srivastava, Hirak Sarkar, Rob Patro. (2023). Raw Data Processing. In L. Heumos and A. Schaar (Ed.). *Single-cell best practices*. (book chapter; book available at https://www.sc-best-practices.org/preamble.html)

Algorithms meet sequencing technologies– 10th edition of the RECOMB-Seq workshop. Rob Patro, and Leena Salmela. *iScience.* 24(1):101956. 2021.

II.G.6. Manuals

User manual and documentation for the oarfish software.
Available at: https://combine-lab.github.io/oarfish/

User manual and documentation for the alevin-fry software.
Available at: https://alevin-fry.readthedocs.io/en/latest/

User manual and documentation for the pyroe software.
Available at: https://pyroe.readthedocs.io/en/latest/

User manual and documentation for the salmon software.
Available at: https://salmon.readthedocs.io/en/latest/

User manual and documentation for the sailfish software.
Available at: https://sailfish.readthedocs.io/en/latest/

User manual and documentation for the simpleaf software.
Available at: https://simpleaf.readthedocs.io/en/latest/

User manual and documentation for the GHOST software.
Available at:
http://www.cs.cmu.edu/~ckingsf/software/ghost/downloads/UserGuide.pdf

II.H. Completed Creative Works and Scholarship

II.H.5. Software and Applications

The following software was either developed within my lab, or for papers on which I am the senior author (with significant contributions from myself and my students):

alevin-fry: An efficient, scalable, and flexible framework for processing single-cell sequencing data
Software available at: https://github.com/COMBINE-lab/alevin-fry

simpleaf: A simple, flexible, and scalable framework for single-cell data processing using alevin-fry
Software available at: https://github.com/COMBINE-lab/simpleaf

roers: A rust library for preparing expanded transcriptome references for quantification with alevin-fry
Software available at: **https://github.com/COMBINE-lab/roers**

grangers: A rust library for parsing and manipulating genomic annotations.
Software available at: **https://github.com/COMBINE-lab/grangers**
QCatch: Quality Control downstream of alevin-fry / simpleaf.
Software available at: https://github.com/COMBINE-lab/QCatch

Cuttlefish: Cuttlefish is a fast, parallel, and very lightweight memory tool to construct the compacted de Bruijn graph from genome references
Software available at: https://github.com/COMBINE-lab/cuttlefish

Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index
Software available at: https://github.com/splatlab/mantis

oarfish: Accurate transcript quantification from long-read RNA-seq data
Software available at: https://github.com/COMBINE-lab/oarfish

Piscem: A next-generation tool for indexing and mapping against large-scale reference collections using the compacted de Bruijn graph and efficient reference tilings.
Software available at: https://github.com/COMBINE-lab/piscem

piscem-infer: An efficient and modular tool for quantification, abundance and uncertainty inference from lightweight mapping data.
Software available at: https://github.com/COMBINE-lab/piscem-infer

Pufferfish2: a reference based index for exact k-mer queries designed to be a successor to Pufferfish
Software available at: https://github.com/COMBINE-lab/pufferfish2

TreeTerminus: a program for grouping transcripts into trees based on their inferential uncertainty in RNA-seq experiments.

Software available at: https://github.com/COMBINE-lab/TreeTerminus

Perplexity: evaluating transcript abundance estimation in the absence of ground truth.
Software available at: https://github.com/COMBINE-lab/perplexity

pyroe: Python library for preparing splici reference sequences for alevin-fry and for importing alevin-fry data into AnnData objects for use in ScanPy and related tools.
Software available at: https://github.com/COMBINE-lab/pyroe

roe: R library for preparing splici reference sequences for alevin-fry and for importing alevin-fry data into Bioconductor SingleCellExperiment objects.
Software available at: https://github.com/COMBINE-lab/roe

Usefulaf: functions and utilities that are useful for preparing data for processing with alevin-fry, as well as for reading alevin-fry data into other packages for downstream analysis; pipelines for simplified analysis.
Software available at: https://github.com/COMBINE-lab/usefulaf

Squeakr: An Exact and Approximate k -mer Counting System
Software available at: https://github.com/splatlab/squeakr

deBGR: An Efficient and Near-Exact Representation of the Weighted deBruijn Graph
Software available at: https://github.com/splatlab/debgr

Salmon: Fast and bias-aware quantification of transcript expression
Software available at: https://github.com/COMBINE-lab/salmon

Terminus: Robust and accurate data-driven transcriptome analysis
Software available at: https://github.com/COMBINE-lab/terminus

Pufferfish: An efficient index for the colored, compacted, de Bruijn Graph
Software available at: https://github.com/COMBINE-lab/pufferfish

Wasabi: Prepare Sailfish and Salmon output for downstream analysis
Software available at: https://github.com/COMBINE-lab/wasabi

RapMap: Rapid, sensitive and accurate read mapping via quasi-mapping
Software available at: https://github.com/COMBINE-lab/RapMap

Grouper: Clustering and quantification of *de novo* transcriptomes
Software available at: https://github.com/COMBINE-lab/grouper

Minnow: A principled framework for dscRNA-seq simulation at the read level
Software available at: https://github.com/COMBINE-lab/minnow

Quark: Semi-reference-based short-read compression
Software available at: https://github.com/COMBINE-lab/quark

LCMLKin: Low-coverage maximum likelihood relatedness estimation
Software available at:
https://github.com/COMBINE-lab/maximum-likelihood-relatedness-estimation

Rainbowfish: A succinct colored deBruijn Graph Representation
Software available at: https://github.com/COMBINE-lab/rainbowfish

Matryoshka: Automated discovery of hierarchically-structured chromatin domains
Software available at: https://github.com/COMBINE-lab/matryoshka

RapClust: Accurate, Lightweight Clustering of de novo Transcriptomes using Fragment Equivalence Classes
Software available at: https://github.com/COMBINE-lab/RapClust

SkipPatch: A dynamically-updatable index for text search
Software available at: https://github.com/COMBINE-lab/SkipPatch

FQFeeder: A simple multi-threaded FASTA/FASTQ parser
Software available at: https://github.com/rob-p/FQFeeder

I either led or substantially contributed to the development of the following software:

Sailfish: Rapid Mapping-based Isoform Quantification from RNA-Seq Reads
Software available at: https://github.com/kingsfordgroup/sailfish

Armatus: Multiresolution domain calling software for chromosome conformation capture interaction matrices
Software available at: https://github.com/kingsfordgroup/armatus

GrowCode: A non-parametric framework for learning how networks grow
Software available at: https://github.com/geetduggal/growcode

PARANA2: Ancestral network reconstruction by summing over parsimonious histories
Software available at: https://github.com/kingsfordgroup/parana2

PARANA: Parsimonious Ancestral Reconstruction And Network Analysis
Software available at:
http://www.cs.cmu.edu/~ckingsf/software/parana/downloads/parana.tgz

Mince: Reference-free compression of sequencing reads
Software available at: https://github.com/Kingsford-Group/mince

GHOST: Global Network Alignment using Multiscale Spectral Signatures
Software available at: https://github.com/Kingsford-Group/ghost

II.J Sponsored Research and Programs – Administered by the Office of Research Administration (ORA)

II.J.1 Grants

Pending

Title: Collaborative Research: CSR: Medium: Dynamic and Distributed Data Structures on the GPU
PIs: Prashant Pandey, Michael Bender, Martin Farach-Colton, John Owens, Robert Patro
Agency: National Science Foundation
Award Number: N/A
Amount: **$**1,200,000 ($240,000 UMD portion)

Title: Collaborative Research: EAGER: SEQBase: A Public Archive of Sequence Data
PI: Robert Patro, Michael Ferndman
Agency: National Science Foundation
Award Number: N/A
Amount: $299,260 ($149,630 UMD portion)

Title: An optimized FarmGTEx pipeline and its applications using high-performance computing
PD: Li Ma
Co-PD: Rob Patro, George Liu, Lingzhao Fang
Agency: United States Department of Agriculture
Award Number: N/A
Amount: $650,000

Funded

Title: A Modular Framework for Accurate, Interpretable, and Reproducible Analysis of Long Read RNA-Seq Data
PIs: Michael Love, Rob Patro
Agency: National Institutes of Health
Award Number: 2R01 HG009937 07
Amount: $3,717,445 ($1,840,064 Patro's portion)

Title: Unifying and enhancing open-source transcriptomics analysis
PI: Robert Patro
Agency: Chan Zuckerberg Initiative

Award Number: N/A
Amount: $400,000

Title: DSFAS: MASH - Machine learning and advanced
‘omics’ data Analysis for improved food Safety and
public Health
PI: Abani K Pradhan
Agency: USDA
Award Number: N/A
Amount: $250,000 (co-PI Patro’s portion)

Title: Expanding and Modernizing the Salmon Ecosystem
PI: Robert Patro
Agency: Chan Zuckerberg Initiative
Award Number: 252586
Duration: 12/01/2022 - 12/01/2024
Amount: $350,000

Title: Salmon: Improving RNA-seq quantification & building an inclusive
community
PI: Robert Patro
Agency: Chan Zuckerberg Initiative
Award Number: CZIF2020-004893
Duration: 01/01/2021 - 12/31/2021
Amount: $235,000
Note: Awarded through Ocean Genomics, Inc.

Title: Hyperspectral terahertz imaging for cancer tumor delineation through deep
neural network learning and geometry-aware curvelet analysis
PI: Hassan Arbab
co-PI: Robert Patro
Agency: Stony Brook Office of the Vice President of Research Seed Funding Program
Duration: 02/25/2019 — 02/24/2020
Amount: $12,000

Title: A Modular Framework for Accurate, Efficient, and Reproducible Analysis of
RNA-seq Data
PI/PD: Robert Patro
co-PI/PD: Michael Love
Agency: NIH
Award Number: R01 HG009937
Duration: 08/18/2018 – 6/30/2023
Amount: $1,538,195

Title: CSR: Medium: Approximate Membership Query Data Structures in
Computational Biology and Storage

PI: Robert Patro
co-PIs: Michael Bender, Michael Ferdman
Agency: NSF
Award Number: 1763680
Duration: 08/15/2018 – 7/31/2022
Amount: $1,999,995

Title: Efficient tools for quantifying and simulating transcript-level abundance in single-cell RNA-seq
PI: Robert Patro
Agency: SVCF
Award Number: 182752
Duration: 03/01/2018 – 09/30/2019
Amount: $132,708.00

Title: CAREER: A Comprehensive and Lightweight Framework for Transcriptome Analysis
PI: Robert Patro
Agency: NSF
Award Number: 1750472
Duration: 02/01/2018 – 01/31/2023
Amount: $625,000.00

Title: Bilateral BBSRC-NSF/BIO: ABI Innovation: Data-driven hierarchical analysis of de novo transcriptomes PI: Robert Patro (co-PI: Julian Hibberd)
Agency: NSF
Award Number: 1564917
Duration: 07/01/2016 – 06/30/2019
Amount: $639,767.80 ($310,627 at Stony Brook, £205,713 at Cambridge University)

II.N. Patents

Other

Kingsford C., Patro R. (2025, Sept). *TUMOR MUTATIONAL BURDEN (TMB) ESTIMATE, 12,417,818.* Domestic Patent, Issued

II.O. Entrepreneurial, Technology Transfer, and Public Engagement Activities

Co-founder and **Chief Technology Officer** of Ocean Genomics Inc.
Website: https://oceangenomics.com/

II.P. Research Fellowships, Prizes, and Awards

2025 ICBS Frontiers of Science Award (declined) — “The International Congress for Basic Science honors top research, with an emphasis on achievements from the past five years which are both excellent and of outstanding scholarly value. The goal of this award is to encourage young scholars to look to the frontiers of basic science, set goals to obtain breakthrough results as early as possible, and contribute wisdom and energy to humankind's study of the mysteries of the natural world.”

Honoree — 2025 Maryland Research Excellence Celebration; “identified by (your) Dean as part of a select group of researchers from your college who have demonstrated outstanding research excellence.”

RECOMB-Seq 2023 — “Tree Terminus: Creating transcript trees using inferential replicate counts” received the best paper award.

University of Maryland Department of Computer Science Excellence in Teaching Award (professor category), 2022.

CMU Allen Newell Award for Research Excellence 2021 — Awarded for work on efficient methods for transcript quantification (i.e. Sailfish). Co-awardees include Carl Kingsford and Steve Mount.

ISMB2021 — “Cuttlefish” paper received Ian Lawson Van Toch Memorial Award (best paper with a student first author)

Dean’s Millionaires Club (2018, Stony Brook, raising over $1,000,000 in external funding as PI within a calendar year)

Best Paper Award, ACM BCB ’17, for the paper Rich chromatin structure prediction from Hi-C data. (2017)

ISMB Student Travel Fellowship (2012, 2013)

1st place in DREAM5 (Dialogue on reverse-engineering assessment and methods) Epitope Antibody Recognition Challenge (2010)

Goldhaber Travel Fellowship (2010)

i3D Student Travel Fellowship (2009)

Verizon Fellowship (2007—2008)

Block Fellowship, Department of Computer Science, University of Maryland (2006—2008)

III. Teaching, Extension, Mentoring, and Advising

III.A. Courses Taught

The following courses were taught at the University of Maryland

CMSC423: Bioinformatics Algorithms, Databases and Tools : Fall 2025 52 students

CMSC701 : Computational Genomics : Spring 2025 30 students

BIOI607 : Computational Genomics : Spring 2025 13 students

CMSC423: Bioinformatics Algorithms, Databases and Tools : Fall 2024 52 students

BIOI506 : Algorithms & Data Structures for Bioinformatics : Spring 2024 6 students

CMSC701 : Computational Genomics : Spring 2024 40 students,

CMSC423 : Bioinformatics Algorithms, Databases and Tools : Fall 2023 30 students

CMSC701 : Computational Genomics : Spring 2023 40 students

CMSC423 : Bioinformatics Algorithms, Databases and Tools : Fall 2022 36 students

CMSC858D : Algorithms, Data Structures and Inference for High-throughput Genomics (graduate-level) : Spring 2022, 27 students

CMSC423 : Bioinformatics Algorithms, Databases and Tools : Fall 2021 75 students

CMSC858D : Algorithms, Data Structures and Inference for High-throughput Genomics (graduate-level) : Fall 2020, 10 students

CMSC423 : Bioinformatics Algorithms, Databases and Tools : Spring 2020, 94 students

CMSC858D : Algorithms, Data Structures and Inference for High-throughput Genomics (graduate-level) : Fall 2019, 17 students

The following courses were taught at Stony Brook University:

CSE373 : Analysis of Algorithms (Undergraduate-level) : Spring 2019, ~110 students
CSE642 : Seminar in Algorithms : Fall 2019, ~50 students (co-taught with other faculty seminar leaders)
CSE549 : Computational Biology (Graduate-level) : Fall 2018, ~65 students
CSE642 : Seminar in Algorithms : Fall 2018, ~45 students (co-taught with other faculty seminar leaders)
CSE373 : Analysis of Algorithms (Undergraduate-level) : Spring 2018, ~160

students
CSE642 : Seminar in Algorithms : Spring 2018, ~35 students (co-taught with other faculty seminar leaders)
CSE549 : Computational Biology (Graduate-level) : Fall 2017, ~100 students
CSE642 : Seminar in Algorithms : Fall 2017, ~30 students (co-taught with other faculty seminar leaders)
CSE549 : Computational Biology (Graduate-level) : Fall 2016, ~60 students
CSE549 : Computational Biology (Graduate-level) : Fall 2015, ~75 students
CSE548 : Analysis of Algorithms (Graduate-level) : Spring 2015, ~10 students
CSE549 : Computational Biology (Graduate-level) : Fall 2014, ~60 students

III.B Teaching Innovations

III.B.6. Course or Curriculum Development

BIOI607: Data Structures and Algorithms for Bioinformatics (graduate-level)
I have created all of the content for this class, including the lectures, the exams, and the (automatically graded) programming assignments (there are 5). While there is overlap of the material between BIOI607 and both CMSC701 and CMSC423, both the lectures and programming assignments have a different nature, as this class is offered as part of a professional MS program within the Science Academy, and the background of the students is highly varied. For example, some have no background in computer science at all, while others have successfully completed an undergraduate degree in computer science. Some have a budding interest in bioinformatics, while others are working bioinformaticiains in industry or government. Thus, this course includes more introductory and background material, and the projects are scaled down appropriately to meet the needs of the students.

CMSC701: Computational Genomics (graduate-level)
I have substantially updated the curriculum for this class, merging the relevant content from my special topics class CMSC858D. In addition, I have revised the content to focus on newer and more modern developments in computational genomics (e.g. including sketching, hashing, and novel indexing techniques based on the de Bruijn graph). Finally, I have re-designed and systematized the course programming projects to operate on the Gradescope autograder.

CMS423: Bioinformatic Algorithms, Databases, and Tools
My iteration of this course includes the same core curriculum used by the other faculty on the field committee who teach the course, however, I have also added content on more advanced topics on genomics algorithms, including content on computational transcriptomics and RNA-sequencing analysis. Further, I completely revised the programming assignments for the course to create a series of interconnected assignments, all submitted and evaluated automatically via Gradescope, that build upon each other and culminate in the students writing their own “toy” sequence aligner.

CMSC858D : Algorithms, Data Structures and Inference for High-throughput Genomics (graduate-level)

I have developed the curriculum for this new advanced topics graduate-level course, which is being offered for the first time in Fall 2019. While there is some topical overlap with **CSE549** (see below), this course is substantially different, and sacrifices some breadth to provide considerably more depth on topics in algorithmic and computational genomics. The curriculum contains material on state-of-the-art developments in algorithms, data structures and statistical inference procedures that are used in contemporary tools and methods for genomics analysis.

The following courses had curricula developed by me while at Stony Brook University:

CSE549 : Computational Biology (Graduate-level)

While this course existed before I joined the department, I developed the curriculum during all semesters where I taught the course, and the curriculum differed significantly from when the course was previously taught by other faculty. I also updated the curriculum in each year to account for recent research developments in the field.

CSE548 : Analysis of Algorithms (Graduate-level)

While this course existed before I joined the department, I developed the curriculum when I taught the course, and the curriculum differed non-trivially from when the course was previously taught by other faculty.

CSE373 : Analysis of Algorithms (Undergraduate-level)

While this course existed before I joined the department, I developed the curriculum during all semesters where I taught the course. The curriculum differed non-trivially from when the course is taught by other faculty, and I, in fact, use a different book ("Algorithm Design" by Kleinberg and Tardos vs. "The Algorithm Design Manual" by Skiena).

III.C Advising: Research or Clinical

III.C.1 Undergraduate

Ongoing advising:

Nan Jiang (UMD CS undergraduate)

Independent research advisor, Fall 2025

Maya Popova (UMD CS undergraduate)

Independent research advisor, Summer/Fall 2025

Maya Popova (UMD CS undergraduate)
Independent research advisor, Summer/Fall 2025

Archita Naik (UMD CS undergraduate)
Independent research advisor, Summer/Fall 2025

Ved Kothavade (UMD CS undergraduate)
Independent research advisor, Spring 2025

Joel Haftel (UMD CS undergraduate)
Independent research advisor, Spring 2025

Siddhant Kumar (UMD CS undergraduate)

Independent research advisor, Spring 2025

Completed advising:

Sophie Robertson (NSF BRIDGE REU)
Research Advisor, Summer 2024

Skylar Rinderknecht (NSF BRIDGE REU)
Research Advisor, Summer 2024

Noah Cape (NSF BRIDGE REU)
Research Advisor, Summer 2023

Natalia Quintana Parrilla (NSF BRIDGE REU)
Research Advisor, Summer 2023

Nathan Joseph Jacobi (NSF BRIDGE REU)
Research Advisor, Summer 2022

Sean Markey (NSF BRIDGE REU)
Research Advisor, Summer 2022

Spiro Razis (Stony Brook)
Undergraduate (non-matriculating) researcher advisor (2016,2017)
First placement: programmer/analyst for NCI (NIH) subcontractor (IMS)

Kyla Zurlage (UMD CS student)
Research advisor (2022)

The following is ongoing advising / co-advising (alphabetical):

Skylar Spencer Chan (UMD CS Honors student)
Research advisor (2020—current)

Noah Eby (UMD CS student)
Research advisor (2024)

III.C.2 Master's

Geoffrey Churchill (Stony Brook)
Advisor (2017)
First placement: Google (software engineer)
Nitish Gupta (Stony Brook)
Advisor, MS Thesis committee member (2014—2016)
First placement: Bloomberg LP (software engineer)
Komal Sanjeev (Stony Brook)
Advisor, MS Thesis committee member (2014—2016)
First placement: Amazon (software engineer)

III.C.3 Doctoral

Completed Doctoral Advising:

Noor Pratap Singh
Advisor (2019—2025)**(defended March 2025)**
First placement: Research Scientist, Illumina

Hyeon Jin Cho
co-advised with Dan Larson at NIH
Advisor (2019—2025)**(defended March 2025)**
First placement: Continuing as researcher at NIH

Jamshed Khan
Advisor (2019—2024)**(defended December 2024)**
First placement: Research Scientist, UMIACS
Currently: Khoury College Distinguished Postdoctoral Fellow

Dongze He
Advisor (2020—2024)**(defended March 2024)**
First placement: Research Scientist, Altos Labs

Jason Fan
Advisor (2020—2023) (**defended Aug. 2023)**
First placement: Computational Biologist, Fulcrum Genomics

Mohsen Zakeri

Advisor (2015—2021) (**defended Nov. 2021)**
First placement: Postdoctoral Research Associate, Department of Computer Science, Johns Hopkins University (postdoctoral mentor Ben Langmead)

Fatemeh Almodaresi
Advisor, Research Proficiency Exam Committee Member (2016—current)
Advisor, Committee Member (2016—2020) (**defended July. 2020)**
Note: Awarded Larry S. Davis Doctoral Dissertation Award
First placement: Postdoctoral Research Associate, Ontario Institute for Cancer Research (postdoctoral mentor Lincoln Stein)

Hirak Sarkar
Advisor, Research Proficiency Exam Committee Member
Advisor, Committee Member (2014—2020) (**defended July. 2020)**
First placement: Postdoctoral Research Associate, Department of Biomedical Informatics, Harvard Medical School (postdoctoral mentor Peter Kharchenko)
Currently: Assistant Professor, College of Connected Computing, Vanderbilt University

Laraib Malik (Stony Brook)
Advisor, Research Proficiency Exam Committee Member, Preliminary PhD Exam
Advisor, Committee Member (2014—2020) (**defended Jan. 2020**)
First placement: Senior Scientist, AbbVie
Currently: Assistant Professor, Syed Babar Ali School of Science and Engineering, Lahore University of Management Sciences, Pakistan

Avi Srivastava (Stony Brook)
Advisor, Research Proficiency Exam Committee Member, Preliminary PhD Exam
Advisor , Committee Member (2014—2019)(**defended Dec. 2019**)
First placement: Postdoctoral Research Associate, New York Genome Center (postdoctoral mentor Rahul Satija).
Currently: Assistant Professor, Wistar Institute (and Affiliate Assistant Professor, UPenn)

Thu Nguyen (Stony Brook)
co-Advisor (co-advised by Dr. Alexi Soares at Brookhaven National Lab) (2018—2020)

The following is ongoing advising / co-advising (alphabetical):

Yuan Gao
Advisor, co-advised with Jian Peng at NIH (2022—current)

Zahra Zare Jousheghani
Advisor (2022—current)

III.C.4 Other Directed Research *(e.g. K-12 Interactions)*

The following is ongoing advising / co-advising

Completed advising (alphabetical):

Geoffrey Glass (Simons Summer Research Program Fellow)
Advisor
First placement: Brown University

Anjalie Kini (Simons Summer Research Program Fellow)
Advisor
First placement: Massachusetts Institute of Technology

Alice Yeh (Simons Summer Research Program Fellow; BISV Senior Project)
Advisor

III.E. Advising: Other than Directed Research
III.E.1 Undergraduate

Lise Ho (Stony Brook, practicum in teaching)
Kyeongsoo Kim (Stony Brook, practicum in teaching)
Freeman Lou (Stony Brook, practicum in teaching)
Matthew Weston (Stony Brook, practicum in teaching)

III.E.2 Master's

***The following students were advised as part of CSE523/524 at Stony Brook University, and performed research under my direction, but were not MS thesis students*:**

Srikant Aggarwal
Eshita Bheda
Anuja Bawaskar
Siddhartha Chhabra
Revati Damle
Komal Dhuri
Shreyas Harisha
Nikhil Junneti
Pavan Kottapalli
Anand Kulkarni
Nikhil Mehta
Nikhil Mohan
Rasika Pohankar
Mridul Ranjan
Manu Rao

Aravind Reddy Ravula
Debjyoti Roy
Ayush Sengupta
Siddarth Harinarayanan
Akshat Singhal
Aditya Srivastava
Alok Thatikunta
Shravya Rani Thatipally
Kevin Wasiluk

III.F. Professional and Extension Education

III.F.4. Guest Lectures

CBMG688Y (Cell Biology and Molecular Genetics Graduate Program; 2019)
  Guest lecture on methods for single-cell transcriptome analysis
PQB 558 (Seminar in Physical & Quantitative Biology at Stony Brook; 2017)
  Guest lecture on sequence alignment and read mapping
BGE 510 (Graduate Genetics at Stony Brook; 2017)
  Guest lecture on RNA-seq read mapping
  Guest lecture on RNA-seq assembly and quantification
BGE 510 (Graduate Genetics at Stony Brook; 2016)
  Guest lecture on high-throughput sequencing and assembly
  Guest lecture on RNA-seq quantification

III.H. Other Teaching Activities

Guest lecturer at CSHL Course "Statistical Methods for Functional Genomics" 2019
Gave lecture covering the theory and practice of transcript quantification and a description of common technical biases in RNA-seq data. Presented a detailed assessment of the effect of different mapping and alignment methodologies on the accuracy of transcript abundance estimation.

Guest lecturer at CSHL Course "Statistical Methods for Functional Genomics" 2018
Gave lecture covering aspects of read alignment and mapping, detailed coverage of the theory of transcript quantification and common technical biases in RNA-seq data, and transcript assembly for working in the presence of an incomplete annotation.

Instructor at Next-Gen Sequence Analysis Workshop (ANGUS) 2017
URL: https://angus.readthedocs.io/en/2017/
Gave lecture covering aspects of read alignment and mapping and quantification using RNA-seq, and delivered practical tutorial (co-taught with Mike Love) on quantification and differential expression analysis.

Instructor at Next-Gen Sequence Analysis Workshop (ANGUS) 2016

URL: https://angus.readthedocs.io/en/2016/
Created course modules (lecture and practical lab) covering read alignment and mapping and quantification using RNA-seq.

IV. Service and Outreach

IV.A. Editorships, Editorial Boards, and Reviewing Activities

IV.A.1. Editorships

Oxford Bioinformatics: Editorial Board Member (Nov. 2019 — current)
PLOS Computational Biology (guest editor; 2019, 2021)

IV.A.3. Reviewing Activities for Journals and Presses (in alphabetical order)

Algorithms
BMC Bioinformatics
BMC Genomics
Cell Systems
F1000
Genome Biology
Genome Research
IEEE Transactions on Computational Biology and Bioinformatics
Journal of Bioinformatics and Computational Biology
Nature Communications
Nature Genetics
Nature Methods
Nucleic Acids Research
Nucleic Acids Research Genomics and Bioinformatics
Oxford Bioinformatics
PeerJ
PLOS Computational Biology
PLOS One
Proceedings of the National Academy of Sciences of the United States of America

IV.A.4. Reviewing Activities for Agencies and Foundations (in alphabetical order)

American Heart Association Panelist (2018)
National Institutes of Health Panelist (BDMA study section, NHGRI study section, recurring)
National Science Foundation Panelist (CISE, BIO, recurring)

IV.A.5 Reviewing Activities for Conferences (in alphabetical order)

ACM BCB Conference
ACM SIGGRAPH Conference

Asia Pacific Bioinformatics Conference
IEEE Intl. Workshop on High Performance Computational Biology (HICOMB)
IEEE Visualization Conference
Intelligent Systems for Molecular Biology (ISMB)
Research in Computational Molecular Biology (RECOMB)
RECOMB Satellite Workshop on Massively Parallel Sequencing (RECOMB-Seq)
SIAM Conference on Applied and Computational Discrete Algorithms (ACDA)
Workshop on Algorithms in Bioinformatics (WABI)

IV.B. Committees, Professional & Campus Service

IV.B.1 Campus Service – Department

CS Department Council member (2025/2026)
UMIACS Steering Committee member (2025/2026)
Learning Outcomes Committee (2024/2025)
CS Department, Graduate Admissions Committee (2024/2025)
UMIACS APT Committee, member (2023/2024)
Learning Outcomes Committee, **(chair)** (2022/2023)
UMIACS Peer-institutes fact-finding committee, member (2023)
UMIACS self-assessment committee, member (2022/2023)
CS Department Council member, (2022/2023)
Department self-assessment, Undergraduate Instruction sub-committee (**chair**) (2021/2022)
Instructor Search Committee (2021/2022)
Learning Outcomes Committee (2020/2021, 2021/2022)
Teaching Evaluation Committee (2020/2021, 2021/2022)
Faculty Hiring committee, Computer Science Department (2020/2021)
Graduate admissions committee (2020 — current)
Colwell Fellowship selection committee (2019 — current)
Graduate admissions committee (at Stony Brook; 2014 — 2019)
Graduate curriculum committee (at Stony Brook; 2017 — 2019)
Faculty hiring committee (at Stony Brook; 2014, 2015, 2016, 2018)

Ongoing committee membership (alphabetical):

Ataberk Donmez (UMD, Department of Computer Science)
Committee Member, PhD Preliminary Exam (2025—)

Yunheng Han (UMD, Department of Computer Science)
Committee Member, PhD Preliminary Exam (2024—)

Completed committee membership (alphabetical):

Dr. Rana Momtaz Abdelsamie Zaki Khali (UMD, Department of Computer Science)

Committee Member, PhD Defense (March 2025)

Dr. Harihara Subrahmaniam Muralidharan (UMD, Department of Computer Science)
Committee Member, Preliminary PhD Exam (2022)
Committee Member, PhD Defense (2024)

Dr. Bixuan Wang (UMD, Department of Biological Sciences)
Committee member, PhD Preliminary Exam & PhD Defense (2022-2023)

Dr. Xinyang Huang (UMD, Department of Nutrition and Food Science)
Committee Member (Dean's Representative), PhD Defense (2023)

Dr. Jeremiah Emidih (UMD, Department of Mathematics)
Committee Member (Dean's Representative), PhD Defense (2023)

Dr. Shaopeng Zhu (UMD, Department of Computer Science)
Committee Member, PhD Defense (2023)

Dr. Kiran Gajanan Javkar (UMD, Department of Computer Science)
Committee Member, Preliminary PhD Exam & PhD Defense (2020—2022)

Dr. Mohammad Amin (Stony Brook, Department of Computer Science)
Research Proficiency Exam Committee Member, Preliminary PhD Exam Committee Member, PhD Defense Committee Chair (2017—2019)
First placement: Tenure-track assistant professor of computer science at Fordham University

Dr. Victoria Paz Cepeda Espinoza (UMD, Department of Computer Science)
Committee Member, PhD Defense (2020)
First placement: Postdoctoral Research Associate, Department of Biomedical Data Science, Stanford University (postdoctoral advisor, Carlos Bustamante)

Dr. Ahmed Elgohary Ghoneim (UMD, Department of Computer Science)
Committee Member, PhD Defenese (2021)
Committee Member, PhD Preliminary Proposal (2019)

Dr. Hayan Lee (Stony Brook, Department of Computer Science)
PhD Defense Committee Member (2015)
First placement: Postdoc at DOE Joint Genome Institute, Lawrence Berkeley National Laboratory

Dr. Brian David Ondov (UMD, Department of Computer Science)
Committee Member, PhD Defense (2020)

Dr. Prashant Pandey (Stony Brook, Department of Computer Science)

Committee Member, PhD Defense Committee (2018)
First placement: Postdoctoral Research Associate, Department of Computational Biology, Carnegie Mellon University (postdoctoral advisor, Carl Kingsford)

Dr. Welles Robinson (UMD, Department of Computer Science)
Committee Member, Preliminary Exam (2021)
Committee Member, PhD Defense (2021)

Dr. Nidhi Shaha (UMD, Department of Computer Science)
Committee Member, PhD Defense (2021)

IV.B.2 Campus Service – College

CMNS Interim Faculty Advisory Council, Computer Science Department Representative (2024/2025)

CMNS Associate Dean for Research search committee, member (2022)

CBBG Program: Graduate admissions committee (2020, 2021, 2022)

Participant in CMNS "Future Center Leaders" program (inaugural 2021 class)

Ongoing committee membership outside of the department (alphabetical):

Oladipupo Ridwan Bello (UMD, Animal and Avian Sciences)
Preliminary Exam Committee Member (2024)
PhD Defense Committee Member (2025)

Nour El Husseini (UMD, Biological Science Graduate Program)
PhD Defense Committee Chair (2025)

Completed committee membership outside of the department (alphabetical):

Dr. Al-Fahad Mohammed Al-Qadhi (UMD, Mathematics)
PhD Defense Committee Member (Dean's Representative) (2025)

Dr. Bixuan Wang (UMD, Biological Science Graduate Program)
Preliminary Exam Committee Member (2021)
Qualifying Exam Committee Member (2021)
PhD Defense Committee Member (2023)

Dr. Theresa Alexander, (UMD, Biological Science Graduate Program)
Preliminary Exam Committee Member (2020)
Qualifying Exam Committee Member (2021)
PhD Defense Committee Member (2023)

Dr. Muzi Li (UMD, Biological Science Graduate Program)
Preliminary Exam Committee Member (2019)
Qualifying Exam Committee Member (2020)
PhD Defense Committee Member (2022)

Dr. Xuan ("Cindy") Li (UMD, Biological Science Graduate Program)
PhD Defense (Dean's Representative) (2024)
Preliminary Exam Committee Member (2020)

Dr. Domenick Braccia (UMD, Biological Science Graduate Program)
Preliminary Exam Committee Member (2020)
Qualifying Exam Committee Member (2020)
PhD Defense Committee Member (2020)

Bihua Yu (Stony Brook, Department of Applied Mathematics and Statistics)
Preliminary PhD Exam Committee Member (2017)

Dr. Han Fang (Stony Brook, Department of Applied Mathematics and Statistics)
Preliminary PhD Exam Committee Member, PhD Defense Committee Member (2015—2017)
First placement: Facebook Data Scientist

Dr. Thomas Pranzatelli (UMD, Biological Science Graduate Program)
Preliminary Exam Committee Member (2020)
Qualifying Exam Committee Member (2020)
Dissertation Exam Committee Member (2024)

Dr. Sanju Sinha (UMD, Biological Science Graduate Program)
Committee Member (Dean's representative), PhD Defense (2021)

IV.B.3. Campus Service – University

Simons Summer Research Program admissions committee (at Stony Brook; 2018, 2019)

IV.B.5 Campus Service — Other

Completed committee membership outside of the college (alphabetical):

Dr. Jason O'rawe (Stony Brook, Department of Genetics)
Preliminary PhD Exam Committee Member, PhD Defense Committee Member (2016—2017)
First placement: Data Science Associate at Open Data Science Conference

IV.B.6. Inter-institutional and Regional *(e.g. inter-library agencies, regional consortia)*

SUNY Research Foundation Germination Space for Engineering Driven Medicine participant (at Stony Brook; 2018)

Ongoing committee membership at different universities (alphabetical):

Keffy Kehrli (Stony Brook, Department of Genetics)
PhD Committee Member (2017—)

Completed committee membership at different universities (alphabetical):

Dr. Camille Scott (UC Davis, Department of Computer Science)
Qualifying Exam Committee Member & Dissertation Committee Member (2020—2022)

Dr. Luiz Iber (UC Davis, Department of Computer Science)
Qualifying Exam Committee Member (2019)

Dr. Alisa Yurovsky (Stony Brook, Department of Computer Science)
PhD Committee Member (2018—2020)

IV.B.7 Offices and Committee Memberships

Management Committee for scVerse organization (2024 —)
url: https://scverse.org/people/

Program committee for ACM BCB (2017—2019)

Program committee for Asia Pacific Bioinformatics Conference (2016, 2017)

Program committee for IPDPS (2019)

Program committee for ISMB (2017 — 2025)

Program committee for RECOMB (2019, 2022 — 2025)

Program committee for RECOMB-Seq (2016 — 2018, 2021—2025)

Program committee for SCIAM ACDA (2021)

Program committee for Workshop on Algorithms in Bioinformatics (2017, 2018, 2023)

IV.B.8. Leadership Roles in Meetings and Conferences

Conference **co-Chair** for WABI 2025

Area **co-Chair** for ISMB 2023 (Genome Sequence Analysis area)

CSHL Biological Data Science Meeting 2022 **session co-Chair** (Algorithmics)

Area **co-Chair** for ISMB 2022 (Genome Sequence Analysis area)

Highlight Track **co-Chair** for ACM-BCB 2020

Program Committee **co-Chair** for RECOMB-Seq 2020

Poster Committee **Chair** for RECOMB 2019

Genome Informatics 2019 Meeting **discussion leader** (transcriptomics session)

IV.D. Media Contributions

IV.D.6. Blogs

RoBlog (personal blog discussing topics related to computational genomics and computer science)
URL: http://robpatro.com/blog/