

### Education

- 2006—2012 Doctor of Philosophy, Department of Computer Science, University of British Columbia, Vancouver, Canada, GPA – 4.0/4  
Advisors: Ján Maňuch, Cedric Chauve and Arvind Gupta  
Thesis title: Variants of the consecutive-ones property motivated by the reconstruction of ancestral species
- 2004—2006 Master of Science, School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada, GPA – 3.9/4  
Advisors: Eugenia Ternovska and Arvind Gupta  
Thesis title: Grounding for model expansion in  $k$ -guarded formulas
- 1999—2003 Bachelor of Computer Science (Honours), Jodrey School of Computer Science, Acadia University, Wolfville, Nova Scotia, Canada, GPA – 3.8/4 (top 5%)  
Advisors: Raymond J. Spiteri and Jim Diamond  
Thesis title: Implementing Runge-Kutta solvers in Java

### Work Experience

- 2020— Assistant Professor, Department of Computer Science, Georgia State University, Atlanta, Georgia, USA  
Teaching both undergraduate and graduate courses in computer science, carrying out a research program in bioinformatics, and serving the department, and the university
- 2019—2020 Visiting Assistant Professor, Department of Computer Science and Engineering, Fairfield University, Fairfield, Connecticut, USA  
Teaching both undergraduate and graduate courses in computer science, as well as carrying out my research program in bioinformatics
- 2016—2019 Postdoctoral Researcher, Experimental Algorithmics Lab (AlgoLab), Informatics Department, Università degli Studi di Milano-Bicocca, Milano, Italy  
Advisors: Gianluca Della Vedova and Paola Bonizzoni  
Performing research on haplotype assembly, cancer phylogenetics and their application
- 2014—2016 Postdoctoral Researcher, Laboratoire de Biométrie et Biologie Évolutive (LBBE), Université Claude Bernard Lyon 1, Villeurbanne, France  
Advisors: Daniel Kahn and Vincent Daubin  
Studying the co-evolution of (metabolic) function, and how it can be used to help the construction of phylogenetic trees

- 2013—2014 ERCIM Postdoctoral Research Fellow, Life Sciences Group, Centrum Wiskunde Informatica (CWI), Amsterdam, The Netherlands  
 Advisors: Alexander Schönhuth and Gunnar W. Klau  
 Improving the state-of-the-art haplotype assembly methods for next-generation genome sequencing data
- 2012—2013 ERCIM Postdoctoral Research Fellow, INRIA Rhône-Alpes and the Laboratoire de Biométrie et Biologie Évolutive, Université Claude Bernard Lyon 1, Villeurbanne, France  
 Advisor: Eric Tannier  
 Performing a study of the co-evolution of genes, and how it can be used to help the construction of phylogenetic trees

## Teaching Experience

- Fall 2025 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4510 / 6510 — a senior year / masters-level course on automata theory
- Fall 2024 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4520 / 6520 — a senior year / masters-level course on algorithm design / analysis, and  
 ○ CSC 8540 — a grad-level course on advanced algorithms in bioinformatics
- Fall 2024 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Fall 2024 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Spring 2023 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC / DSCI 4810 / 6810 — a senior year / master-level course on artificial intelligence
- Fall 2022 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Spring 2022 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Fall 2021 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4710 / 6710 — a senior year / masters-level course on database systems
- Spring 2021 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Fall 2020 Georgia State University, Atlanta, Georgia, USA  
 ○ CSC 4330 / 6330 — a senior year / masters-level programming language concepts course
- Spring 2020 Fairfield University, Fairfield, Connecticut, USA  
 ○ CS 131 — an introductory programming course in Java, and  
 ○ CS 354 — a junior year theory of programming languages course using Haskell and Prolog

- Fall 2019 Fairfield University, Fairfield, Connecticut, USA
- CS 131 — an introductory programming course in Java
  - CS 231 — an intermediate (sophomore year) programming course, focusing on the object-oriented features of Java, and
  - SW 402 — a senior-year / masters-level course on database management systems
- 2015 Instructor, Department of Mathematics, Université Claude Bernard Lyon 1, Villeurbanne, France
- Section head: Raymond Lobry
- Teaching a section of mathematics for the life sciences to first-year university students

### Advising

- Fall 2025 I supervised PhD students Prakash Chourasia, Avais Jan, Ayotomiwa Ezekiel Adeniyi and Mansoor Ahmed, and Master's student of the Biology Department James Dang
- Mansoor Ahmed won an MBD fellowship, funding the remainder of his study.
  - Prakash received the 2025 MBD Outstanding Fellow Award.
- Spring 2025 I supervised PhD students Prakash Chourasia, Avais Jan, Ayotomiwa Ezekiel Adeniyi and Mansoor Ahmed
- Prakash received the 2025 Outstanding Graduate Student Award in the Department of Computer Science.
- Fall 2024 I supervised PhD students Sarwan Ali, Prakash Chourasia, Avais Jan, Ayotomiwa Ezekiel Adeniyi and Mansoor Ahmed, Master's student Sai Roopesh Mandava, and Undergraduate student Ramita Rathore
- Sai Roopesh Mandava successfully completed his Master's project titled: "Kickstarting football skills: A reinforcement learning approach".
  - Sarwan Ali successfully defended his PhD dissertation titled: "Designing representation learning methods for molecular sequences analysis".
  - Sarwan Ali received a Georgia State dissertation grant award.
- Spring 2024 I supervised PhD students Zahra Tayebi, Sarwan Ali, Prakash Chourasia, Taslim Murad, Avais Jan and Ayotomiwa Ezekiel Adeniyi, Master's students Raghuveer Gummadi and Sai Roopesh Mandava, and Undergraduate students Ramita Rathore and Anderson Hernandez
- Taslim Murad successfully defended her PhD dissertation titled: "Designing methods for representation learning of molecular sequences and its application in analysis tasks".
  - Zahra Tayebi successfully defended her PhD dissertation titled: "Clustering, classification, and simulation of evolutionarily related sequences in the context of cancer and viral spread".
  - Sarwan Ali received the 2024 Graduate Research Award in the Department of Computer Science.
  - Ramita Rathore won a 1st Place Novice Research Award at the 2024 GSU Summer Undergraduate Research Symposium.

1 Park Place, Room 636 – Atlanta, Georgia 30303 (USA)

☎ +1 475 330 0978 • ✉ mpatterson30@gsu.edu

🌐 murraypatterson.github.io

- Fall 2023 I supervised PhD students Zahra Tayebi, Sarwan Ali, Prakash Chourasia, Taslim Murad, Avais Jain and Ayotomiwa Ezekiel Adeniyi, and Master's students Raghuveer Gummadi and Eunice Olorunshola.
- Prakash Chourasia won an MBD fellowship, funding the remainder of his study.
  - Zahra Tayebi successfully completed her dissertation proposal titled: "Efficient clustering and classification of protein sequences".
  - Sarwan Ali successfully completed his dissertation proposal titled: "Designing representation learning methods for molecular sequences analysis".
  - Eunice Olorunshola successfully defended her Master's thesis titled: "Classifying different cancer types based on transcriptomics data using machine learning algorithms".
- Spring 2023 I supervised PhD students Zahra Tayebi, Sarwan Ali, Prakash Chourasia, Taslim Murad and Avais Jain, and Master's students Raghuveer Gummadi, Eunice Olorunshola, Sai Venkata Sumanth Gaddam, Kushal Khatri, Surya Subrahmanya Sai Chandra Sekhar Kolli, Samvidha Reddy Mannem, Fatma Noor, Ashwin Pilgaonkar, Ria C. Thazhe Punathil, Sainikhila Vemulapalli and Karishma Mahendra Wadhe.
- Taslim Murad successfully completed her dissertation proposal titled: "Designing methods for representation learning of biological sequences and its application in analysis tasks".
- The following Master's projects were graduated:
- Sai Gaddam: "Diabetic retinopathy detection using deep learning".
  - Kushal Khatri: "A stratified subsampler for hierarchically structured data".
  - Surya Kolli: "Recommendations using graph databases".
  - Fatma Noor: "Can community detection build a good phylogeny?".
  - Ashwin Pilgaonkar: "Robinson Country Intelligence Index (RCII) – A country risk analysis tool".
  - Ria Punathil: "CoVScan – A tool for viral host specification from coronavirus spike sequences".
  - Sainikhila Vemulapalli: "Vehicle image classification and damage analysis using deep learning model".
  - Karishma Wadhe: "Drive safe – A driver drowsiness alert system".
- Fall 2022 I supervised PhD students Zahra Tayebi, Sarwan Ali, Prakash Chourasia and Taslim Murad, and Master's students Supriya Arun, Kaustubh Khandai, Kushal Khatri, Prashanth Annam, Babitha Reddy Devireddy, Indu Munnangi, Venkata Krishna Chaitanya Nuthalapati, Dhruval Patel, Shraddha Sharma and Atharv Utkarsh Yeolekar.
- Zahra Tayebi won a B&B fellowship, funding the remainder of her study.
- The following Master's projects were graduated:
- Supriya Arun: "Comparative analysis of Covid classifiers".
  - Kaustubh Khandai: "Determining significant correlation between pairs of extant characters in a small parsimony framework".

- Prashanth Annam: “Neobook: A social medial application using Flutter and Fire-base”.
  - Babitha Reddy Devireddy: “Chaos engineering using Gremlin and GC analyzer”.
  - Indu Munnangi: “My recipe book: A MERN stack web application”.
  - Venkata Krishna Chaitanya Nuthalapati: “Tsuba: Web security scanner”.
  - Dhruval Patel: “Best actor detection”.
  - Shraddha Sharma: “Sunsetting customer identification service”.
  - Atharv Utkarsh Yeolekar: “Data condensation using generative adversarial networks”.
- Spring 2022 I supervised PhD students Zahra Tayebi, Sarwan Ali, Prakash Chourasia and Taslim Murad, and Master’s students Babatunde Bello, Supriya Arun and Kaustubh Khandai.
- Babatunde Bello graduated with Master’s thesis titled: “Functional enrichment analysis of transcriptomics data of breast cancer RNA-seq”.
- Fall 2021 PhD students Zahra Tayebi, Sarwan Ali and Prakash Chourasia, Master’s students Babatunde Bello and Kaustubh Khandai, and Honors student Yijing Zhou.
- Sarwan Ali won an MBD fellowship, funding the remainder of his study.
  - Yijing Zhou graduated with honors thesis titled: “Towards a complete workflow for inferring tumor phylogeny from raw single-cell sequencing reads”
- Spring 2021 I supervised PhD students Zahra Tayebi and Sarwan Ali
- Fall 2020 I supervised PhD student Zahra Tayebi

## Volunteer Experience

- 2017—2019 Research Advisor, Lycée Stendhal, Milan, Italy  
 Advising and guiding students on an after-school mathematics project which is part of the larger European program: Math en Jeans. There is an annual conference where the students go to present their work on the year’s project.

## Grants

- 2024 Using AI techniques to uncover transmission dynamics and hotspots from HIV big data, Research Internal Grant (RIG), GSU, Amount: \$20,000 USD (1 year)  
 PI: Alexander Kirpich
- To collect HIV incidence data and numerically encode it for analysis for trends in transmission with machine learning algorithms
- 2024 REU site: Multidisciplinary graph data analytics, National Science Foundation, Amount: \$372,720 USD (3 years)  
 PI: Esra Akbas, Mehmet Emin Aktas
- To establish Georgia State University as a Research Experiences for Undergraduates (REU) site by hiring paid undergraduate summer interns to work together with faculty and graduate students on their research projects.

- 2021 Global education online module design, Hanyang University ERICA, Amount: \$24,255 USD (3 years)  
PI: Zhipeng Cai, Co-I: Armin Mikler, Rajshekhar Sunderraman, Pavel Skums, Murray Patterson, Rolando Estrada, Yingshu Li, Jennifer Sherer, Geoffrey Graybeal  
○ For the development of online modules in data science and business intelligence, which could be offered as a course or training program.

### Contracts

- 2021 AI Learning Alliance: Broadening participation, Meta, Inc., Amount: \$50,000 USD (2 years)  
PI: Anu Bourgeois, Co-I: Murray Patterson  
○ To support work related to providing education at the frontier of R&D and to increase the representation of historically marginalized researchers and engineers by preparing them for jobs in the artificial intelligence field.

### Under Review

- 2025 SCH: Tracing semantic shifts in textual and genomic data: A new approach for understanding the co-evolution of human behavior and pathogen phenotypic traits, National Science Foundation, Amount: \$899,924 (3 years)  
PI: Murray Patterson, Co-I: Alexander Zelikovsky, Mario Kubek, Alexander Kirpich and Gerardo Chowell  
○ Human activity such as drug use and homelessness results in elevated susceptibility to infectious diseases, representing a critical global public health concern. Importantly, there is an ongoing co- evolution between human host behavior and pathogens: certain activities can accelerate pathogen transmission, thereby expanding opportunities for viral diversification and fostering the emergence of novel, potentially more transmissible or drug-resistant variants that threaten the broader population. Understanding and predicting such co-evolutionary events is essential for public health preparedness. We hypothesize that these events may manifest as synchronized semantic shifts in human-generated text (e.g., social media) and in the organizational or informational structure of viral genomes. Detecting these shifts independently is challenging, and their joint detection presents an even greater methodological problem. To address this, we propose a cross-domain language model (LM)-based framework designed to identify semantic shifts in both social discourse and epistatically-shaped viral genomes, and to infer associations between them.
- 2025 CAREER: Learning new patterns of evolution from sequencing big data, National Science Foundation, Amount: \$651,363 USD (5 years)  
PI: Murray Patterson

- The current amount of molecular sequencing data available is orders of magnitude larger than in any other time throughout history, and has revealed evidence of evolutionary events such as saltations (evolutionary jumps) and reversions (back mutations). These patterns seem to contradict a century of development of evolutionary theory, suggesting a gap in our understanding of evolution, particularly at the smaller scale of viruses and tumors, where these events are often being observed. Better computational methods, models, and their resulting tools for the analysis of this big data resource are needed to close this gap. This would result in better surveillance, screening, and therapeutic strategies, since some viral strains of concern are the result of a saltation event from a more benign strain, while sequencing data from tumors reveal the widespread presence of reversions in any phylogenetic reconstruction from these sequences. This proposal develops methods based on phylo- and population genetics techniques and deep learning for better anticipating saltations and modeling reversions.
- 2025 An interactive visual tool for exploring tree and graph structures, National Science Foundation, Amount: \$635,759 USD (3 years)  
PI: Murray Patterson, Co-I: Rajshekhar Sunderraman and Anu Bourgeois
- A recurring theme in many CS courses is the hidden structures—such as trees or graphs—that underlie many foundational CS concepts, e.g., an SQL query can be represented by an expression tree. Instructors often convey such hidden structures in the classroom with diagrams of expressions, recursive function calls, state transitions, etc. A key challenge in this approach is that such structures are often dynamic: changing with each step of the associated computation that it represents. While some expert instructors produce animations of such diagrams, these can be time-consuming to generate, and difficult to scale and customize. This proposal aims to develop an easily extendable open-source toolbox for the interactive and visual exploration of the various tree and graph structures underlying concepts such as function composition and program modularization, collections of nested objects, database querying, tree and graph algorithms, automata, and non-imperative programming styles such as function and logic programming. This will involve the construction of seven modules, organized into the following three aims: 1. Exploring expression and computation trees for programming languages 2. Interactively stepping through algorithms and automata 3. Developing a visual interface for exploring and building code.
- 2025 A comprehensive multi-center epidemiological study on childhood acute leukemia in Ukraine: Exploring the impact of Chernobyl and the ongoing war using AI methods, National Institute of General Medical Sciences, Amount: \$1,304,929.00 USD (3 years)  
PI: Murray Patterson, Co-I: Alexander Kirpich and Alexander Zelikovsky

- This project employs advanced data science techniques to gain deeper insights into childhood leukemia by analyzing surveillance data aggregated from multiple partnering centers across Ukraine. It aims to retrospectively examine leukemia-focused data from regional medical centers over the past several decades, including the critical context of the last three years of the full-scale war between Russia and Ukraine. The ultimate goal of this project is to extrapolate findings from this extensive historical dataset to the broader population, while addressing the recent challenges and emerging trends related to leukemia.

## Awards and Honors

- 2023 Nomination for membership to Sigma Xi, the scientific research honor society
- 2011 The Marie Curie Alain Bensoussan (ABCDE) Postdoctoral Fellowship of the European Research Consortium for Informatics and Mathematics (ERCIM)  
Award of 60,300 € and 5000 € for travel, in order to fully fund 2 years of postdoctoral study in the European ERCIM member countries
- 2006 The Postgraduate Scholarship (PGS) of the Natural Sciences and Engineering Research Council (NSERC) of Canada  
Award of \$63,000 CAD in order to fully fund 3 years of postgraduate (PhD) research
- 2006 Simon Fraser University Computing Science Graduate Fellowship  
Award of \$3000 CAD towards funding postgraduate research
- 2005 Simon Fraser University Computing Science Graduate Fellowship  
Award of \$3000 CAD towards funding postgraduate research
- 2002 The Undergraduate Student Research Award (USRA) of the Natural Sciences and Engineering Research Council (NSERC) of Canada  
Award of \$6400 CAD in order to fully fund a summer of undergraduate research

## Technical Skills

advanced	C/C++, Python, $\LaTeX$	7+ years experience
intermediate	Java, Unix tools, Shell programming	4+ years experience
basic	Matlab, R, Prolog, Haskell	2+ years experience

## Languages

English	maternal language	
French	second language	conversational and technical fluency
Italian	intermediate	comprehension and basic conversational skills
Spanish	intermediate	comprehension and basic conversational skills
中文	basic	basic words and phrases



## Publications

### Journal Articles

1. Sarwan Ali, Prakash Chourasia, and Murray Patterson. Anderson acceleration for molecular sequencing. *Neural Computing and Applications*, 2025. to appear.
2. Sarwan Ali, Prakash Chourasia, Bipin Koirala, and Murray Patterson. Nearest neighbor CCP-based molecular sequence analysis. *Transactions on Computational Biology and Bioinformatics*, pages 1–14, 2025. doi:10.1109/TCBBIO.2025.3621138.
3. Simone Ciccolella, Murray Patterson, Iman Hajirasouliha, and Gianluca Della Vedova. Cancer progression inference using a finite-state model to allow recurrences and losses of mutations. *Neural Computing and Applications*, 2025. doi:10.1007/s00521-025-11474-1.
4. Sarwan Ali, Tamkanat E. Ali, and Murray Patterson. Hash-binding: DNA-protein binding prediction using hashing-based embedding. *Neural Computing and Applications*, 2025. doi:10.1007/s00521-025-11349-5.
5. Sarwan Ali, Tamkanat E. Ali, Haris Mansoor, Prakash Chourasia, and Murray Patterson. Hist2Vec: a histogram and kernel-based embedding method for molecular sequence analysis. *Expert Systems with Applications*, 273:126859, 2025. doi:10.1016/j.eswa.2025.126859.
6. Sarwan Ali, Madiha Shabbir, Haris Mansoor, Prakash Chourasia, and Murray Patterson. Elliptic geometry-based kernel matrix for improved biological sequence classification. *Knowledge-Based Systems*, 304:112479, 2024. doi:10.1016/j.knosys.2024.112479.
7. Zahra Tayebi, Sarwan Ali, and Murray Patterson. TCellR2Vec: Efficient feature selection for TCR sequences for cancer classification. *PeerJ Computer Science*, 10:e2239, 2024. doi:10.7717/peerj-cs.2239.
8. Sarwan Ali, Tamkanat E. Ali, Taslim Murad, Haris Mansoor, and Murray Patterson. Molecular sequence classification using efficient kernel based embedding. *Information Sciences*, 679:121100, 2024. doi:10.1016/j.ins.2024.121100.
9. Sarwan Ali, Prakash Chourasia, and Murray Patterson. From PDB files to protein features: a comparative analysis of PDB bind and STCRDAB datasets. *Medical & Biological Engineering & Computing*, 62:2449–2483, 2024. doi:10.1007/s11517-024-03074-3.
10. Zahra Tayebi, Sarwan Ali, Taslim Murad, Imdad Ullah Khan, and Murray Patterson. PseAAC2Vec protein encoding for TCR protein sequence classification. *Computers in Biology and Medicine*, 170:107956, 2024. doi:10.1016/j.combiomed.2024.107956.
11. Sarwan Ali, Prakash Chourasia, and Murray Patterson. When protein structure embedding meets large language models. *MDPI Genes*, 15(1):25, 2023. doi:10.3390/genes15010025.

12. Sarwan Ali and Murray Patterson. Improving IOSOMAP efficiency with RKS: a comparative study with t-distributed neighbor embedding on protein sequences. *MDPI J — Multidisciplinary Scientific Journal*, 6(4):579–591, 2023. doi:10.3390/j6040038.
13. Taslim Murad\*, Sarwan Ali\*, Imdad Ullah Khan, and Murray Patterson. Spike2CGR: an efficient method for spike sequence classification using chaos game representation. *Machine Learning*, 112:3633–3658, 2023. doi:10.1007/s10994-023-06371-4.
14. Sarwan Ali, Prakash Chourasia, Zahra Tayebi, Babatunde Bello, and Murray Patterson. ViralVectors: compact and scalable alignment-free virome feature generation. *Medical & Biological Engineering & Computing*, 2023. doi:10.1007/s11517-023-02837-8.
15. Taslim Murad, Sarwan Ali, and Murray Patterson. Exploring the potential of GANs in biological sequence analysis. *MDPI Biology*, 12(6):854, 2023. doi:10.3390/biology12060854.
16. Bikram Sahoo\*, Sarwan Ali, Pin-Yu Chen, Murray Patterson, and Alexander Zelikovskiy\*. Assessing the resilience of machine learning classification algorithms on SARS-CoV-2 genome sequences generated with long-read specific errors. *Biomolecules*, 13(6):934, 2023. doi:10.3390/biom13060934.
17. Danushka Bandara, Karen Exantus, Cristian Navarro-Martinez, Murray Patterson, and Ashley Byun. Identifying distinguishing acoustic features in felid vocalizations based on call type and species classification. *Acoustics Australia*, 2023. doi:10.1007/s40857-023-00298-5.
18. Sarwan Ali, Babatunde Bello, and Murray Patterson. Solvent accessibility of coronaviridae spike proteins through the lens of information gain. *MDPI J — Multidisciplinary Scientific Journal*, 6(2):236–247, 2023. doi:10.3390/j6020018.
19. Sarwan Ali, Bikram Sahoo, Alexander Zelikovskiy, Pin-Yu Chen, and Murray Patterson. Benchmarking machine learning robustness in covid-19 genome sequence classification. *Scientific Reports*, 13(1):4154, 2023. doi:10.1038/s41598-023-31368-3.
20. Prakash Chourasia, Sarwan Ali, Simone Ciccolella, Gianluca Della Vedova, and Murray Patterson. Reads2Vec: Efficient embedding of raw high-throughput sequencing reads data. *Journal of Computational Biology*, 2022. doi:10.1089/cmb.2022.0424.
21. Sarwan Ali, Babatunde Bello, Zahra Tayebi, and Murray Patterson. Characterizing SARS-CoV-2 spike sequences based on geographical location. *Journal of Computational Biology*, 2022. doi:10.1089/cmb.2022.0391.
22. Sarwan Ali, Bikram Sahoo, Muhammad Asad Khan, Alexander Zelikovskiy, Imdad Ullah Khan\*, and Murray Patterson\*. Efficient approximate kernel based spike sequence classification. *Transactions on Computational Biology and Bioinformatics*, 2022. doi:10.1109/TCBB.2022.3206284.
23. Kaustubh Khandai, Cristian Navarro-Martinez, Brendan Smith, Rebecca Buonopane, S. Ashley Byun\*, and Murray Patterson\*. Determining significant correlation between pairs of extant characters in a small parsimony framework. *Journal of Computational Biology*, 29(10):1132–1154, 2022. doi:10.1089/cmb.2022.0141.

24. Sarwan Ali, Babatunde Bello, Prakash Chourasia, Ria Thazhe Punathil, Yijing Zhou, and Murray Patterson. PWM2Vec: An efficient embedding approach for viral host specification from coronavirus spike sequences. *MDPI Biology*, 11(3):418, 2022. doi:10.3390/biology11030418.
25. Sarwan Ali, Yijing Zhou, and Murray Patterson. Efficient analysis of COVID-19 clinical data using machine learning models. *Medical & Biological Engineering & Computing*, 60(7):1881–1896, 2021. doi:10.1007/s11517-022-02570-8.
26. Zahra Tayebi, Sarwan Ali, and Murray Patterson. Robust representation and efficient feature selection allows for effective clustering of SARS-CoV-2 variants. *Algorithms*, 14(12):348, 2021. doi:10.3390/a14120348.
27. Sarwan Ali, Simone Ciccolella, Lorenzo Lucarella, Gianluca Della Vedova, and Murray Patterson. Simpler and faster development of tumor phylogeny pipelines. *Journal of Computational Biology*, 28(11):1142–1155, 2021. doi:10.1089/cmb.2021.0271.
28. Andrew Melnyk, Fatemeh Mohebbi, Sergey Knyazev, Bikram Sahoo, Roya Hosseini, Pavel Skums, Alex Zelikovsky\*, and Murray Patterson\*. From alpha to zeta: Identifying variants and subtypes of SARS-CoV-2 via clustering. *Journal of Computational Biology*, 28(11):1113–1129, 2021. doi:10.1089/cmb.2021.0302.
29. Simone Ciccolella\*, Murray Patterson\*, Paola Bonizzoni, and Gianluca Della Vedova. Effective clustering for single cell sequencing cancer data. *IEEE Journal of Biomedical and Health Informatics*, 25(11):4068–4078, 2021. doi:10.1109/JBHI.2021.3081380.
30. Simone Ciccolella, Mauricio Soto, Murray D. Patterson, Gianluca Della Vedova, Iman Hajirasouliha, and Paola Bonizzoni. Gpps: An ILP-based approach for inferring cancer progression with mutation losses from single cell data. *BMC Bioinformatics*, 21(1):413, 2020. doi:10.1186/s12859-020-03736-7.
31. Simone Ciccolella, Camir Ricketts, Mauricio Soto Gomez, Murray Patterson, Dana Silverbush, Paola Bonizzoni, Iman Hajirasouliha, and Gianluca Della Vedova. Inferring cancer progression from single-cell sequencing while allowing mutation losses. *Bioinformatics*, 37(3):326–333, 2020. doi:10.1093/bioinformatics/btaa722.
32. Raffaella Rizzi, Stefano Beretta, Murray Patterson, Yuri Pirola, Marco Previtali, Gianluca Della Vedova, and Paola Bonizzoni. Overlap graphs and de Bruijn graphs: data structures for de novo genome assembly in the big data era. *Quantitative Biology*, 7(4):278–292, 2019. doi:10.1007/s40484-019-0181-x.
33. Stefano Beretta\*, Murray Patterson\*, Simone Zaccaria, Gianluca Della Vedova, and Paola Bonizzoni. HapCHAT: Adaptive haplotype assembly for efficiently leveraging high coverage in long reads. *BMC Bioinformatics*, 19(1):252, 2018. doi:10.1186/s12859-018-2253-8.
34. Wandrille Duchemin, Yoann Anselmetti, Murray Patterson, Yann Ponty, S  verine B  rard, Cedric Chauve, Celine Scornavacca, Vincent Daubin, and Eric Tannier. DeCoSTAR: Reconstructing the ancestral organization of genes or genomes using reconciled phylogenies. *Genome Biology and Evolution*, 9(5):1312–1319, 2017. doi:10.1093/gbe/evx069.

35. Andrea Bracciali, Marco Aldinucci, Murray Patterson, Tobias Marschall, Nadia Pisanti, Ivan Merelli, and Massimo Torquati. PWhatsHap: Efficient haplotyping for future generation sequencing. *BMC Bioinformatics*, 17(11):342, 2016. doi:10.1186/s12859-016-1170-y.
36. Murray Patterson\*, Tobias Marschall\*, Nadia Pisanti, Leo van Iersel, Leen Stougie, Gunnar W. Klau<sup>†</sup>, and Alexander Schoenhuth<sup>†</sup>. WhatsHap: Weighted haplotype assembly for future-generation sequencing reads. *Journal of Computational Biology*, 22(6):498–509, 2015. doi:10.1089/cmb.2014.0157.
37. Murray Patterson, Gergely Szöllősi, Vincent Daubin, and Eric Tannier. Lateral gene transfer, rearrangement, reconciliation. *BMC Bioinformatics*, 14(15):S4, 2013. doi:10.1186/1471-2105-14-S15-S4.
38. Mohammed El-Kebir\*, Tobias Marschall\*, Inken Wohlers\*, Murray Patterson, Jaap Heringa, Alexander Schoenhuth, and Gunnar W. Klau. Mapping proteins in the presence of paralogs using units of coevolution. *BMC Bioinformatics*, 14(15):S18, 2013. doi:10.1186/1471-2105-14-S15-S18.
39. Ján Maňuch\*, Murray Patterson\*, Roland Wittler\*, Cedric Chauve, and Eric Tannier. Linearization of ancestral multichromosomal genomes. *BMC Bioinformatics*, 13(19):S11, 2012. doi:10.1186/1471-2105-13-S19-S11.
40. Ján Maňuch\*, Murray Patterson\*, and Cedric Chauve. Hardness results on the gapped consecutive-ones property problem. *Discrete Applied Mathematics*, 160(18):2760–2768, 2012. doi:10.1016/j.dam.2012.03.019.
41. Ján Maňuch\* and Murray Patterson\*. The complexity of the gapped consecutive-ones property problem for matrices of bounded maximum degree. *Journal of Computational Biology*, 18(9):1243–1253, 2011. doi:10.1089/cmb.2011.0128.
42. Roland Wittler, Ján Maňuch\*, Murray Patterson\*, and Jens Stoye. Consistency of sequence-based gene clusters. *Journal of Computational Biology*, 18(9):1023–1039, 2011. doi:10.1089/cmb.2011.0083.

## Conference Proceedings

43. Chandra Kiran Guntupalli, Thomas Böhme, Murray Patterson, and Mario M. Kubek. Meaning in motion: tracking temporal semantic drift through incremental representation spaces. In the 8th International Conference on Natural Language Processing (ICNLP, Xi'an, China, 2026). 2026. to appear.
44. Ahmed Luqman, Khuzemah Qazi, Murray Patterson, Malik Jahan Khan, and Imdadullah Khan. A data-free analytical quantization scheme for deep learning models. In the 2025 IEEE International Conference on Data Mining (IEEE ICDM, Washington DC, USA, 2025). 2025. to appear.
45. Sarwan Ali, Taslim Murad, Prakash Chourasia, Haris Mansoor, Imdad Ullah Khan, Pin-Yu Chen, and Murray Patterson. Position specific scoring is all you need? Revisiting protein sequence classification tasks. In the 22nd RECOMB Conference on Comparative Genomics (RECOMB-CG, Seoul, South Korea, 2025), 202–215. 2026. doi:10.1007/978-3-031-94928-9\_15.
46. Sarwan Ali, Haris Mansoor, Prakash Chourasia, Imdad Ullah Khan, and Murray Patterson. Computing Gram matrix for SMILES strings using RDKFingerprint and Sinkhorn-Knopp algorithm. In the 13th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Atlanta, Georgia, 2025), 264–276. 2026. doi:10.1007/978-3-032-02489-3\_20.
47. Qasim Zia, Avais Jan, Zafar Iqbal, Mumtaz Muhammad Ali, Mukarram Ali, and Murray Patterson. EfficientNet in digital twin-based cardiac arrest prediction and analysis. In the 13th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Atlanta, Georgia, 2025), 317–328. 2026. doi:10.1007/978-3-032-02489-3\_24.
48. Taslim Murad, Prakash Chourasia, Sarwan Ali, Avais Jan, and Murray Patterson. Neuromorphic spiking neural network based classification of COVID-19 spike sequences. In the 13th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Atlanta, Georgia, 2025), 345–356. 2026. doi:10.1007/978-3-032-02489-3\_26.
49. Avais Jan, Qasim Zia, and Murray Patterson. Enhancing privacy preservation and reducing analysis time with federated transfer learning in digital twins-based computed tomography scan analysis. In the 13th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Atlanta, Georgia, 2025), 277–289. 2026. doi:10.1007/978-3-032-02489-3\_21.
50. Mansoor Ahmed, Sarwan Ali, Avais Jan, Imdad Ullah Khan, and Murray Patterson. Improved graph-based antibody-aware epitope prediction with protein language model-based embeddings. In the 13th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Atlanta, Georgia, 2025), 290–302. 2026. doi:10.1007/978-3-032-02489-3\_22.

51. Sarwan Ali, Tamkanat E. Ali, Imdad Ullah Khan, and Murray Patterson. Hilbert curve based molecular sequence analysis. In the 29th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD, Sydney, Australia, 2025), volume 15871 of LNCS, 405–417. 2025. doi:10.1007/978-981-96-8173-0\_32.
52. Sarwan Ali, Prakash Chourasia, Haris Mansoor, Bipin Koirala, and Murray Patterson. MIK: Modified isolation kernel for biological sequence visualization, classification, and clustering. In the 4th Machine Learning for Health Symposium (ML4H, Vancouver, Canada, 2024), volume 259 of Proceedings of Machine Learning Research, 30–47. 2025. URL: <https://proceedings.mlr.press/v259/ali25a.html>.
53. Prakash Chourasia, Heramb Lonkar, Sarwan Ali, and Murray Patterson. EPIC: Enhancing privacy through iterative collaboration. In the 11th International Conference on Information Management and Big Data (SIMBig, Ilo, Peru, 2024), volume 2496 of Communications in Computer and Information Science, 20–33. 2025. doi:10.1007/978-3-031-91428-7\_2.
54. Prakash Chourasia, Tamkanat E. Ali, Sarwan Ali, and Murray Patterson. DWFL: Enhancing federated learning through dynamic weighted averaging. In the 11th International Conference on Information Management and Big Data (SIMBig, Ilo, Peru, 2024), volume 2496 of Communications in Computer and Information Science, 34–50. 2025. doi:10.1007/978-3-031-91428-7\_3.
55. Sarwan Ali, Haris Mansoor, Prakash Chourasia, Imdad Ullah Khan, and Murray Patterson. Preserving hidden hierarchical structure: Poincaré distance for enhanced genomic sequence analysis. In the 11th International Conference on Information Management and Big Data (SIMBig, Ilo, Peru, 2024), volume 2496 of Communications in Computer and Information Science, 3–19. 2025. doi:10.1007/978-3-031-91428-7\_1.
56. Sarwan Ali, Prakash Chourasia, and Murray Patterson. DeepPWM-BindingNet: Unleashing binding prediction with combined sequence and PWM features. In the 31st International Conference on Neural Information Processing (ICONIP, Auckland, New Zealand, 2024), volume 2286 of Communications in Computer and Information Science, 148–162. 2025. doi:10.1007/978-981-96-6960-8\_11.
57. Taslim Murad, Prakash Chourasia, Sarwan Ali, and Murray Patterson. Advancing protein-DNA binding site prediction: Integrating sequence models and machine learning classifiers. In the 31st International Conference on Neural Information Processing (ICONIP, Auckland, New Zealand, 2024), volume 2286 of Communications in Computer and Information Science, 356–371. 2025. doi:10.1007/978-981-96-6960-8\_25.
58. Bhashithe Abeysinghe, Murray Patterson, and Rajshekhar Sunderraman. An interactive lambda calculus interpreter and visualization tool. In the 17th Annual ACM India Compute Conference (COMPUTE, Gandhinagar, India, 2024), volume 2400 of Communications in Computer and Information Science, 113–125. 2024. doi:10.1007/978-3-031-84391-4\_9.

59. Sarwan Ali, Haris Mansoor, Prakash Chourasia, Yasir Ali, and Murray Patterson. Gaussian Beltrami-Klein model for protein sequence classification: A hyperbolic approach. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Kunming, China, 2024), volume 14954 of LNCS, 52–62. 2024. doi:10.1007/978-981-97-5128-0\_5.
60. Sarwan Ali, Tamkanat E. Ali, Prakash Chourasia, and Murray Patterson. A universal non-parametric approach for improved molecular sequence analysis. In the 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD, Taipei, Taiwan, 2024), volume 14648 of LNCS, 194–206. 2024. doi:10.1007/978-981-97-2238-9\_15.
61. Taslim Murad, Sarwan Ali, and Murray Patterson. Weighted chaos game representation for molecular sequence classification. In the 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD, Taipei, Taiwan, 2024), volume 14648 of LNCS, 234–245. 2024. doi:10.1007/978-981-97-2238-9\_18.
62. Taslim Murad, Sarwan Ali, Prakash Chourasia, Haris Mansoor, and Murray Patterson. Circular arc length-based kernel matrix for protein sequence classification. In the 2023 IEEE International Conference on Big Data (IEEE BigData, Sorrento, Italy, 2023), 1429–1437. 2023. doi:10.1109/BigData59044.2023.10386477.
63. Sarwan Ali, Prakash Chourasia, and Murray Patterson. Expanding chemical representation with  $k$ -mers and fragment-based fingerprints for molecular fingerprinting. In the 10th International Conference on Information Management and Big Data (SIM-Big, Mexico City, Mexico, 2023), volume 2142 of Communications in Computer and Information Science, 17–29. 2023. doi:10.1007/978-3-031-63616-5\_2.
64. Prakash Chourasia, Taslim Murad, Zahra Tayebi, Sarwan Ali, Imdad Ullah Khan, and Murray Patterson. Efficient classification of SARS-CoV-2 spike sequences using federated learning. In the 10th International Conference on Information Management and Big Data (SIMBig, Mexico City, Mexico, 2023), volume 2142 of Communications in Computer and Information Science, 80–96. 2023. doi:10.1007/978-3-031-63616-5\_6.
65. Zahra Tayebi, Akshay Juyal, Alex Zelikovsky, and Murray Patterson. Simulating tumor evolution from scDNA-seq as an accumulation of both SNVs and CNAs. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 530–540. 2023. doi:10.1007/978-981-99-7074-2\_43.
66. Sarwan Ali\*, Haris Mansoor\*, Prakash Chourasia\*, and Murray Patterson. Hist2Vec: Kernel-based embeddings for biological sequence classification. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 387–397. 2023. doi:10.1007/978-981-99-7074-2\_30.
67. Sarwan Ali\*, Prakash Chourasia\*, and Murray Patterson. PDB2Vec: Using 3D structural information for improved protein analysis. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 376–386. 2023. doi:10.1007/978-981-99-7074-2\_29.

68. Prakash Chourasia\*, Taslim Murad\*, Sarwan Ali\*, and Murray Patterson. Enhancing t-SNE performance for biological sequencing data through kernel selection. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 442–452. 2023. doi:10.1007/978-981-99-7074-2\_35.
69. Usama Sardar\*, Sarwan Ali\*, Muhammad Sohaib Ayub, Muhammad Shoaib, Khurram Bashir, Imdad Ullah Khan, and Murray Patterson. Sequence-based nanobody-antigen binding prediction. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 227–240. 2023. doi:10.1007/978-981-99-7074-2\_18.
70. Sarwan Ali, Pin-Yu Chen, and Murray Patterson. Unveiling the robustness of machine learning models in classifying COVID-19 spike sequences. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 1–15. 2023. doi:10.1007/978-981-99-7074-2\_1.
71. Sarwan Ali, Usama Sardar, Imdad Ullah Khan, and Murray Patterson. Efficient sequence embedding for SARS-CoV-2 variants classification. In the 19th International Symposium on Bioinformatics Research and Applications (ISBRA, Wrocław, Poland, 2023), volume 14248 of LNCS, 16–30. 2023. doi:10.1007/978-981-99-7074-2\_2.
72. Mansoor Ahmed, Usama Sardar, Sarwan Ali, Shafiq Alam, Murray Patterson, and Imdad Ullah Khan. Robust brain age estimation via regression models and MRI-derived features. In Advances in Computational Collective Intelligence (ICCCI, Budapest, Hungary, 2023), volume 1864 of Communications in Computer and Information Science. 2023. doi:10.1007/978-3-031-41774-0\_52.
73. Zahra Tayebi\*, Sarwan Ali\*, Prakash Chourasia, Taslim Murad, and Murray Patterson. T cell receptor protein sequences and sparse coding: A novel approach to cancer classification. In the 2023 International Conference on Neural Information Processing (ICONIP, Changsha, China, 2023), Communications in Computer and Information Science. 2023. doi:10.1007/978-981-99-8141-0\_17.
74. Sarwan Ali, Babatunde Bello, Prakash Chourasia, Ria Thazhe Punathil, Pin-Yu Chen, Imdad Ullah Khan, and Murray Patterson. Virus2Vec: Viral sequence classification using machine learning. In the 2023 Conference on Health, Inference, and Learning (CHIL, Cambridge, MA, USA, 2023), volume 209 of Proceedings of Machine Learning Research, 6–18. 2023. URL: <https://proceedings.mlr.press/v209/ali23a.html>.
75. Sarwan Ali, Prakash Chourasia, and Murray Patterson. When biology has chemistry: Solubility and drug subcategory prediction using SMILES strings. In the 11th International Conference on Learning Representations (ICLR, Kigali, Rwanda, 2023), Tiny Papers Track. 2023. URL: <https://openreview.net/forum?id=28si4RXwDt1>.
76. Taslim Murad, Sarwan Ali, and Murray Patterson. A new direction in membranolytic anticancer peptides classification: Combining spaced  $k$ -mers with chaos game representation. In the International Neural Network Society Workshop on Deep Learning Innovations and Applications (INNS DLIA, Queensland, Australia, 2023), volume 222 of Procedia Computer Science, 666–675. 2023. doi:10.1016/j.procs.2023.08.204.

1 Park Place, Room 636 – Atlanta, Georgia 30303 (USA)

☎ +1 475 330 0978 • ✉ [mpatterson30@gsu.edu](mailto:mpatterson30@gsu.edu)

🌐 [murraypatterson.github.io](https://murraypatterson.github.io)



77. Prakash Chourasia\*, Zahra Tayebi\*, Sarwan Ali\*, and Murray Patterson. Empowering pandemic response with federated learning for protein sequence data analysis. In the 2023 International Joint Conference on Neural Networks (IJCNN, Queensland, Australia, 2023), 1–8. 2023. doi:10.1109/IJCNN54540.2023.10191721.
78. Sarwan Ali\*, Taslim Murad\*, and Murray Patterson. PCD2Vec: A Poisson correction distance based approach for viral host classification. In the 2023 International Joint Conference on Neural Networks (IJCNN, Queensland, Australia, 2023), 1–8. 2023. doi:10.1109/IJCNN54540.2023.10191311.
79. Sarwan Ali, Usama Sardar, Murray Patterson, and Imdad Ullah Khan. BioSequence2Vec: Efficient embedding generation for biological sequences. In the 27th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD, Osaka, Japan, 2023), volume 13936 of LNCS, 173–185. 2023. doi:10.1007/978-3-031-33377-4\_14.
80. Sarwan Ali, Taslim Murad, and Murray Patterson. PSSM2Vec: A compact alignment-free embedding approach for coronavirus spike sequence classification. In the 29th International Conference on Neural Information Processing (ICONIP, Indore, India, 2022), volume 1794 of Springer CCIS, 420–432. 2023. doi:10.1007/978-981-99-1648-1\_35.
81. Prakash Chourasia, Sarwan Ali, and Murray Patterson. Informative initialization and kernel selection improves t-SNE for biological sequences. In the 2022 IEEE International Conference on Big Data (IEEE BigData, Osaka, Japan, 2022), 101–106. IEEE Computer Society, 2022. doi:10.1109/BigData55660.2022.10020217.
82. Taslim Murad\*, Prakash Chourasia\*, Sarwan Ali\*, and Murray Patterson. Hashing2Vec: Fast embedding generation for SARS-CoV-2 spike sequence classification. In the 14th Asian Conference on Machine Learning (ACML, Hyderabad, India, 2022), volume 189 of Proceedings of Machine Learning Research, 754–769. 2022. URL: <https://proceedings.mlr.press/v189/taslim23a.html>.
83. Prakash Chourasia, Sarwan Ali, Simone Ciccolella, Gianluca Della Vedova, and Murray Patterson. Clustering SARS-CoV-2 variants from raw high-throughput sequencing reads data. In the 11th International Conference on Computational Advances in Bio and medical Sciences (ICCABS, Virtual Conference, 2021), volume 13254 of Springer LNBI, 133–148. 2022. doi:10.1007/978-3-031-17531-2\_11.
84. Sarwan Ali, Taslim Murad, Prakash Chourasia, and Murray Patterson. Spike2Signal: Classifying coronavirus spike sequences with deep learning. In the 8th IEEE International Conference on Big Data Service and Applications (IEEE BDS, San Francisco Bay Area, USA, 2022), 81–88. 2022. doi:10.1109/BigDataService55688.2022.00020.
85. Sarwan Ali, Bikram Sahoo, Naimat Ullah, Alex Zelikovsky, Murray Patterson\*, and Imdad Ullah Khan\*. A  $k$ -mer based approach for SARS-CoV-2 variant identification. In the 17th International Symposium on Bioinformatics Research and Applications (ISBRA, Shenzhen, China, 2021), volume 13064 of LNCS, 153–164. 2021. doi:10.1007/978-3-030-91415-8\_14.

86. Brendan Smith, Cristian Navarro-Martinez, Rebecca Buonopane, S. Ashley Byun\*, and Murray Patterson\*. Correlated evolution in the small parsimony framework. In the 17th International Symposium on Bioinformatics Research and Applications (ISBRA, Shenzhen, China, 2021), volume 13064 of LNCS, 608–619. 2021. doi:10.1007/978-3-030-91415-8\_51.
87. Sarwan Ali and Murray Patterson. Spike2Vec: An efficient and scalable embedding approach for COVID-19 spike sequences. In the 2021 IEEE International Conference on Big Data (IEEE BigData, Virtual Conference, 2021), 1533–1530. 2021. doi:10.1109/BigData52589.2021.9671848.
88. Sarwan Ali, Tamkanat E. Ali, Muhammad Asad Khan, Imdad Ullah Khan, and Murray Patterson. Effective and scalable clustering of SARS-CoV-2 sequences. In the 5th International Conference on Big Data Research (ICBDR, Tokyo, Japan, 2021), 42–49. 2021. doi:10.1145/3505745.3505752.
89. Andrew Melnyk, Fatemeh Mohebbi, Sergey Knyazev, Bikram Sahoo, Roya Hosseini, Pavel Skums, Alex Zelikovsky\*, and Murray Patterson\*. Clustering based identification of SARS-CoV-2 subtypes. In the 10th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS, Virtual Conference, 2020), volume 12686 of LNCS, 127–141. 2021. doi:10.1007/978-3-030-79290-9\_11.
90. Simone Ciccolella\*, Murray Patterson\*, Paola Bonizzoni, and Gianluca Della Vedova. Effective clustering for single cell sequencing cancer data. In the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB, Niagara Falls, NY, USA, 2019), ACM, 437–446. 2019. doi:10.1145/3307339.3342149.
91. Giulia Bernardini\*, Paola Bonizzoni\*, Gianluca Della Vedova\*, and Murray Patterson\*. A rearrangement distance for fully-labelled trees. In the 30th Annual Symposium on Combinatorial Pattern Matching (CPM, Pisa, Italy, 2019), volume 128 of LIPIcs, 1–15. 2019. doi:10.4230/LIPIcs.CPM.2019.28.
92. Gianluca Della Vedova\*, Murray Patterson\*, Raffaella Rizzi\*, and Mauricio Soto\*. Character-based phylogeny construction and its application to tumor evolution. In the 13th Conference on Computability in Europe (CiE, Turku, Finland, 2017), volume 10307 of LNCS, 3–13. 2017. doi:10.1007/978-3-319-58741-7\_1.
93. Marco Aldinucci, Andrea Bracciali, Tobias Marschall, Murray Patterson, Nadia Pisanti, and Massimo Torquati. High-performance haplotype assembly. In the 11th International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics (CIBB, Cambridge, UK, 2014), volume 8623 of LNCS, 245–258. 2015. doi:10.1007/978-3-319-24462-4\_21.
94. Murray Patterson\*, Tobias Marschall\*, Nadia Pisanti, Leo van Iersel, Leen Stougie, Gunnar W. Klau†, and Alexander Schoenhuth†. WhatsHap: Haplotype assembly for future-generation sequencing reads. In the 18th Annual International Conference on Research in Computational Molecular Biology (RECOMB, Pittsburgh PA, USA, 2014), volume 8394 of LNCS, 237–249. 2014. doi:10.1007/978-3-319-05269-4\_19.

95. Cedric Chauve\*, Murray Patterson\*, and Ashok Rajaraman\*. Hypergraph covering problems motivated by genome assembly questions. In the International Workshop on Combinatorial Algorithms (IWOCA, Rouen, France, 2013), volume 8288 of LNCS, 428–432. 2013. doi:10.1007/978-3-642-45278-9\_37.
  96. Cedric Chauve\*, Ján Maňuch\*, Murray Patterson\*, and Roland Wittler\*. Tractability results for the consecutive-ones property with multiplicity. In the 22nd Annual Symposium on Combinatorial Pattern Matching (CPM, Palermo, Italy, 2011), volume 6661 of LNCS, 90–103. 2011. doi:10.1007/978-3-642-21458-5\_10.
  97. Ján Maňuch\*, Murray Patterson\*, and Arvind Gupta. Towards a characterization of the generalized cladistic character compatibility problem for non-branching character trees. In the 7th International Symposium on Bioinformatics Research and Applications (ISBRA, Changsha, China, 2011), volume 6674 of LNCS, 440–451. 2011.
  98. Ján Maňuch\*, Murray Patterson\*, Sheung-Hung Poon\*, and Chris Thachuk\*. Complexity of finding non-planar rectilinear drawings of graphs. In the 18th International Symposium on Graph Drawing (GD, Konstanz, Germany, 2010), volume 6502 of LNCS, 305–316. 2010. doi:10.1007/978-3-642-18469-7\_28.
  99. Cedric Chauve\*, Ján Maňuch\*, and Murray Patterson\*. On the gapped consecutive-ones property. In the European Conference on Combinatorics, Graph Theory and Applications (EuroComb, Bordeaux, France, 2009), volume 34 of ENDM, 121–125. 2009. doi:10.1016/j.endm.2009.07.020.
  100. Ján Maňuch\*, Murray Patterson\*, and Arvind Gupta. On the generalised character compatibility problem for non-branching character trees. In the 15th Annual International Computing and Combinatorics Conference (COCOON, Niagara Falls, USA, 2009), volume 5609 of LNCS, 268–276. 2009. doi:10.1007/978-3-642-02882-3\_27.
  101. Murray Patterson, Yongmei Liu, Eugenia Ternovska, and Arvind Gupta. Grounding for model expansion in  $k$ -guarded formulas with inductive definitions. In the 20th International Joint Conferences on Artificial Intelligence (IJCAI, Hyderabad, India, 2007), 161–166. 2007.
- note In any publication listing containing a \*, this indicates that all authors whose names are followed by a \* have contributed equally to this publication. In order to be consistent, this means that for papers that list the names of authors alphabetically (which is a common practice amongst computer science and theoretical bioinformatics papers), there is a \* following every name.