MidSouth Computational Biology and Bioinformatics Society
Presents
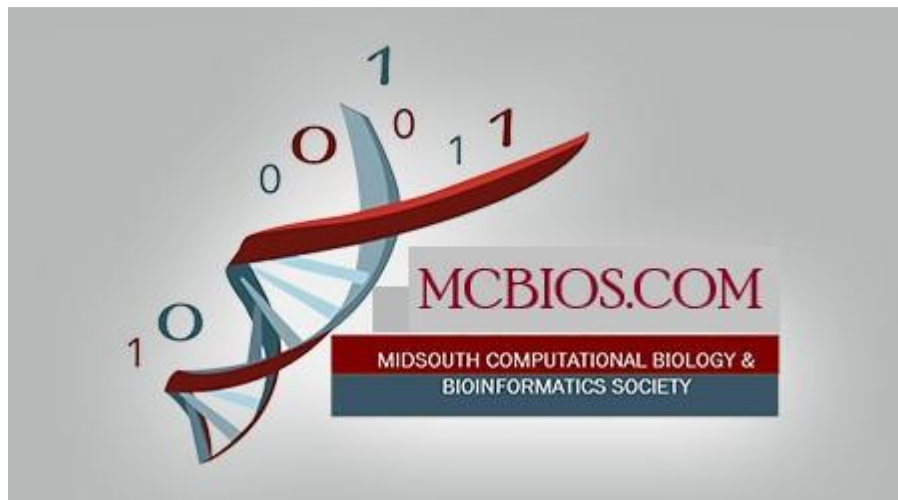
MCBIOS 2024

The 20th Annual Meeting of MCBIOS

March 22-24, 2024

R. Randall Rollins Building
Emory University

From codes to cells to care, transforming health care with AI

Index

Maps

Conference venue: Randall R. Rollins Building

Street address:
**1534 Clifton Road, Atlanta GA 30322 (Google Map).**
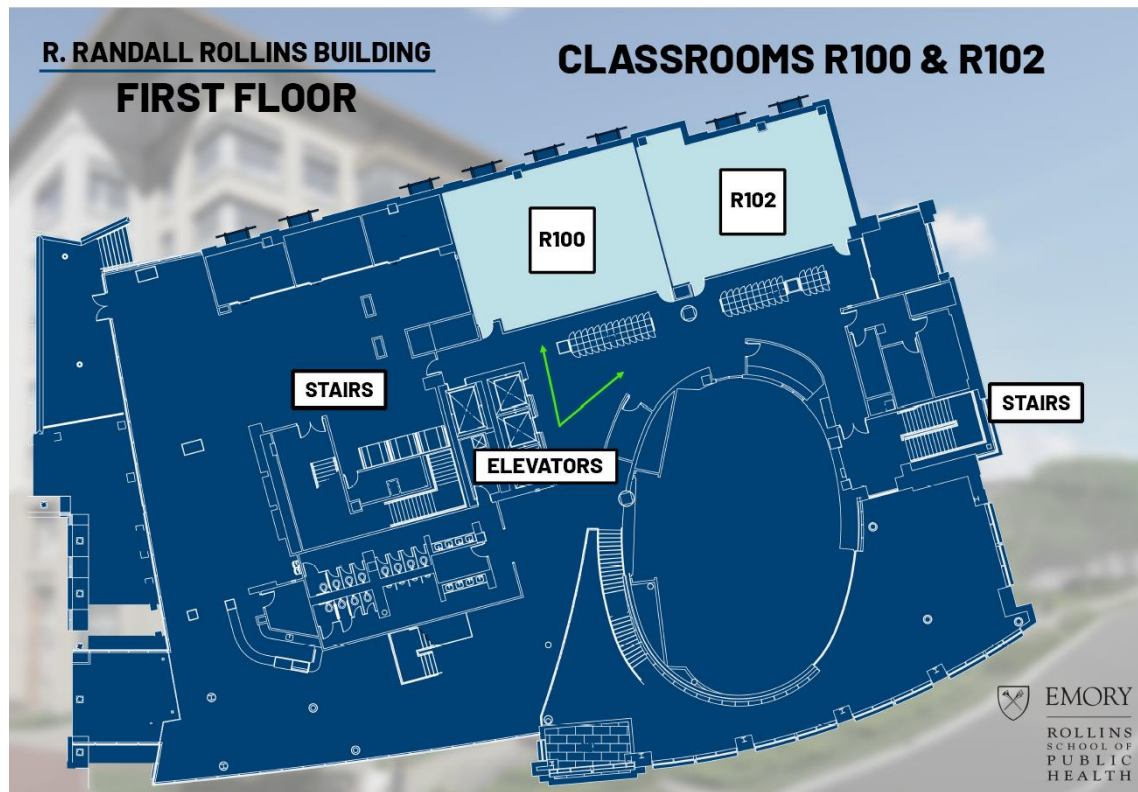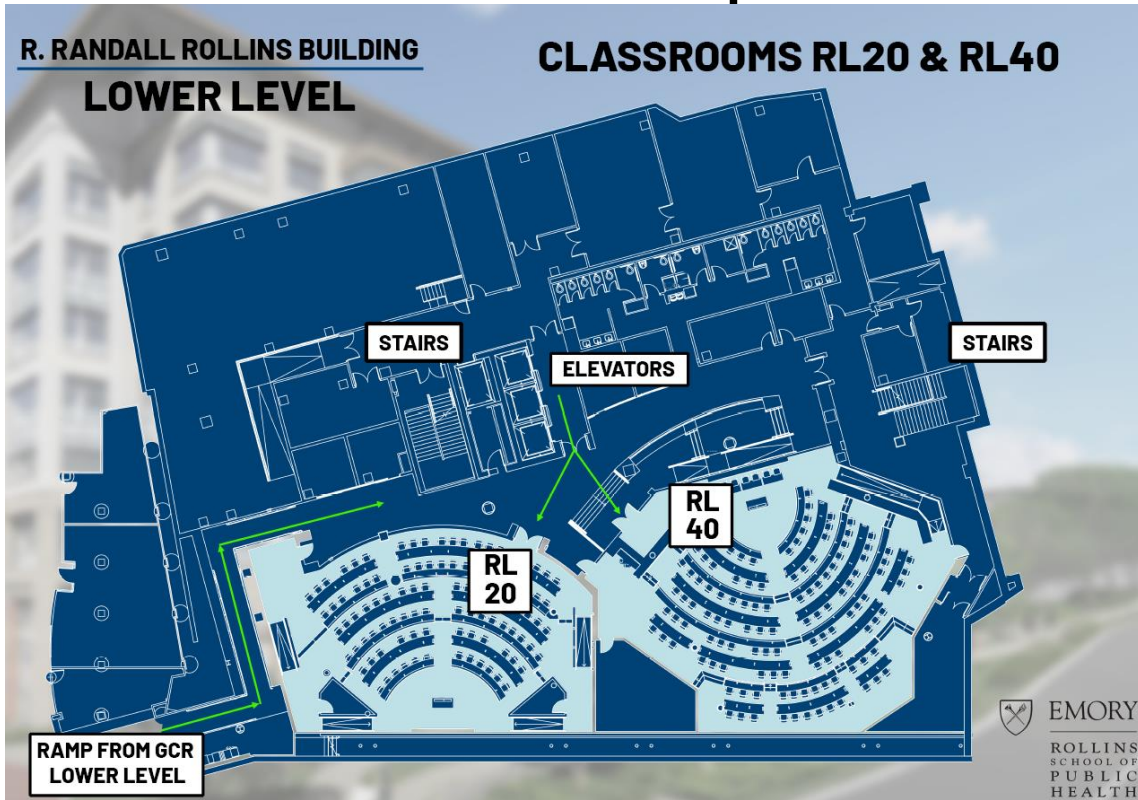**1518 Clifton Rd NE, Atlanta, GA (Bing Map)**

Parking is available at nearby Michael Street Parking Deck. Please use visitor parking on Friday afternoon (free parking validation available). Please use regular parking on Saturday and Sunday.

Walking direction from Emory Conference Center hotel to Randall R. Rollins Building:
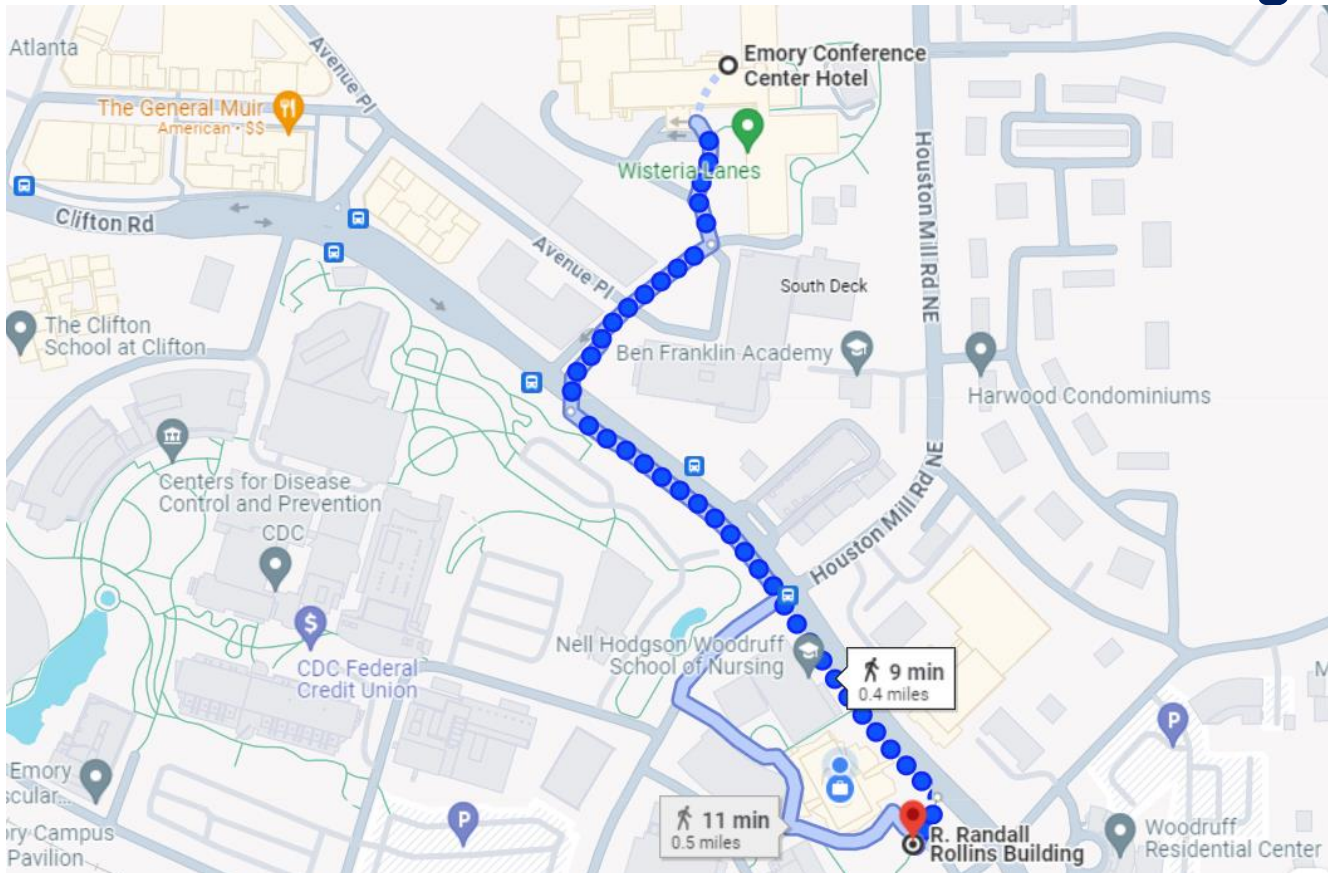
Classroom Maps

Welcome from MCBIOS President and Organizing Committee co-chairs

Greetings.

On behalf of the organizing committee, we would like to personally welcome you to the 2024 annual scientific conference of the MidSouth Computational Biology and Bioinformatics Society (MCBIOS). MCBIOS is a non-profit organization found in 2003, started as a grassroots effort spearheaded by a small group of researchers at the FDA NCTR (located in Little Rock, Arkansas). It has since been expanded significantly. It is now one of only two regional societies in North America (with GLBIO) affiliated with ISCB, the biggest international bioinformatics society. The mission of MCBIOS is to foster networking and collaboration and to promote the professional development of members and supporting our student members. The annual conference is the flagship event for MCBIOS, which attracts talented scientists from across the nation who gather together for the sharing of scientific information to advance our understanding in nature and health. The annual conferences have enjoyed a long history.

MCBIOS 2024 features more than 80 world-class researchers to present in 24 invited technical sessions and workshops. Additionally, there will be Young Scientist Excellence Award (YSEA) sessions, poster sessions, hackathon and many more fun activities. All registered attendees are eligible to submit their research papers to MCBIOS 2024 proceedings to be published as special issues of high-quality Bioinformatics journals. We are looking forward to a great Bioinformatics conference with lasting impact. Please join us!

Steve Qin and Chang Su

Program at a Glance

Day 1: March 22, 2024

1:00– 1:15	Welcome remarks
1:15 – 2:15	Keynote speech (Anant Madabhushi)
2:15 – 2:30	Coffee Break
2:30 – 3:45	Ashish Sharma Memorial Session
3:45 – 4:00	Coffee Break
4:00 – 5:15	YSEA sessions (students and postdocs)
5:15 – 7:00	Poster session and reception

Day 2: March 23, 2024

7:30 – 8:30	Breakfast
8:30 – 9:30	Keynote speech (Greg Gibson)
9:30 – 9:45	Coffee Break
9:45 – 12:30	Tutorial: Introduction to Tidymodels
9:45 –11:00	Breakout sessions I (101,102,103,104)
11:00 – 11:15	Coffee Break
11:15 – 12:30	Breakout sessions II (201, 202, 203, 204)
12:30 – 1:30	Lunch
1:30 – 2:45	Breakout sessions III (301, 302)
2:45 – 3:00	Coffee Break
3:00 – 4:15	Breakout sessions IV (401, 402, 403, 404)
4:15 – 4:30	Coffee Break
4:30 – 5:30	Plenary speech (Nancy Cox)
5:30 – 7:30	Election and banquet

Hackathon	All day

Day 3: March 24, 2024

7:30 – 8:30	Breakfast
8:30 – 9:30	Keynote speech (Heng Li)
9:30 – 9:45	Coffee Break
9:45 – 11:00	Breakout sessions V (501, 502, 503, 504, 505)
11:00 – 11:15	Coffee Break
11:15 – 12:30	Breakout sessions VI (601, 602, 603, 604, 605)
12:30 – 1:00	Awards and closing

Session summary

ID	Title	Session chair
PS101	Statistical methods for genetics and omics data	Chang Su
PS102	Predictive Analysis of Disease Risk Factors from Genetic Complexity	Jian Hu
PS103	Potentials of Artificial Intelligence in Risk Assessment at Gastrointestinal Tract, Microbiome Research, and antibiotic resistance	Sangeeta Khare
PS104	Computational methods for single cell spatial omics data	Zhana Duren
PS201	Computational methods for single cell multi-omics data	Xiuwei Zhang
PS202	Computational Frontiers in Biomolecular Analysis: Navigating Sequence Data Landscapes	Inimary Toby
PS203	Unraveling Intricate Molecular Pathways in Complex Diseases: Insights from Recent Statistical and Computational Advances	Yuehua Cui
PS204	A New Frontier in Public Health: Leveraging Bioinformatics and Generative AI for Advanced Disease Surveillance and Response	Kun Zhao
PS301	Advances in Computational Immuno-Oncology	Aik Choon Tan
PS302	Glycoscience Informatics: Maximizing Insights with Database and Software Tools	Sushil Mishra Robert J. Doerksen
PS401	Statistical Approaches for Interpreting Dynamic Biological Signals from Tissue Molecular and Histological Profiles	Jian Hu
PS402	Dawn of Foundation Models for Physiological Data	Xiao Hu
PS403	Advanced AI methods for improving precision medicine in complex diseases	Li Chen
PS404	Artificial Intelligence Applications in Biomedical Research	Huixiao Hong
PS501	Applications of AI and ML in Human Health at Emory	Julia Wrobel
PS502	Recent Insights Using Machine Learning in Genetics and Genomic Studies	Xiaoxi Shen
PS503	Computational Approaches for Disease Prediction and Biomarker Selection in Multi-Scale Biological Systems	Xiunan Wang
PS504	Advanced AI methods for analyzing single-cell and spatial omics data	Li Chen
PS505	Lightning talks	Nisha Pillai
PS601	Advanced statistical methods for omics data analysis	Steve Qin
PS602	Application of AI in Infectious Diseases/Biological Data Sets	Tomojit Ghosh
PS603	Network Medicine and Drug Discovery	Zongliang Yue
PS604	Delineate multi-level mechanisms of complex diseases under specific cellular contexts.	Jianrong Wang
PS605	Lightning talks	Qianqian Song

Day 1, Friday, March 22, 2024

12:00noon – 5:00pm	Registration and Poster Setup

Welcome Remarks	
Margaret H. Rollins Room	
1:00pm – 1:05pm	**Steve Qin** MCBIOS President
1:05pm – 1:10pm	**Robert Krafty** Rollins Professor and Chair Department of Biostatistics and Bioinformatics Emory University
Keynote Speech	
Margaret H. Rollins Room	
1:10pm – 1:15pm	Introduction to keynote speaker **Chang Su** Assistant Professor Department of Biostatistics and Bioinformatics Emory University
1:15pm – 2:15pm	 Walking the talk: Retrospective and Prospective Clinical Trial Validation for AI in Pathology **Anant Madabhushi, Emory and Georgia Tech**
2:15pm – 2:30pm	Coffee Break

Ashish Sharma Memorial Session	
Margaret H. Rollins Room	

2:30pm – 2:45pm	
	
	1. Tribute to a Pioneer in Computational Imaging
	   Joel Saltz, Stony Brook University
	2. Deep Learning Workflow for Scalable Detection of Tau Inclusions in Digital Pathology
	   David Gutman, Emory University
	3. Elastic Net Modeling of Urinary Extracellular Vesicle RNA for Detection of Aggressive Prostate Cancer
	   Carlos Moreno, Emory University
	4. Books on the shelf: Answering the right question
	   Lance Waller, Emory University

3:45pm – 4:00pm	Coffee Break

Young Scientist Excellence Award (YSEA) Session I

Room: RL 40

4:00pm – 5:15pm	1. Machine Learning and 3-D Modeling of CDR3 Sequences from T-Cell Receptor β in Acute Respiratory Distress Syndrome **Tri Ha,** University of Dallas 2. Predicting the biological activity of SWEET genes in *Pisum Sativum* using bioinformatics and molecular approaches **Sophia Palafox,** University of Dallas 3. Discovering Novel Glaucoma Neural Repair Genes Using Computational Approaches Based on Single-Cell RNA-Seq Data **Yeganeh Madadi,** University of Tennessee Health Science Center 4. Pilot study for the role of bacteriophages in viral infections among kidney transplant patients **Tyrome Sweet,** Emory University

Young Scientist Excellence Award (YSEA) Session II

Room: RL 20

4:00pm – 5:15pm	1. Multiomics integration to explore Triple Negative Breast Cancer Before and After Chemotherapy for Enhanced Immunotherapy Insights **David Otohinoyi,** LSU Health New Orleans 2. Therapeutic Efficiency and Uptake of Gold Nanoparticle-Conjugated GE-11 Peptide Loaded with Rapamycin on Cell Proliferation and Apoptosis in Pancreatic Cancer Cells **Adeolu Oluremi,** University of Arkansas at Little Rock 3. Investigating spatial dynamics in spatial omics data with StarTrail **Jiawen Chen,** UNC Chapel Hill 4. Exploring Mutation Effects in Sodium and Calcium Channels: A Data-Driven Approach for Machine Learning Applications **Amin Akbari Ahangar,** University of Mississippi

	5. Mining Social Media Data About Xylazine Associated Wounds **JaMor Hairston,** Emory University 6. Uncertainty-Aware Pre-Trained Foundation Models for Patient Risk Prediction via Gaussian Process **Jiaying Lu,** Emory University 7. EquiPPIS: a robust and accurate protein-protein interaction site prediction method powered by E(3)-equivariant graph neural networks **Rahmatullah Roche,** Virginia Tech 8. MammOnc-DB: A proteo-genomic platform for comprehensive breast cancer data analysis, integration and visualization **Santhosh Kumar Karthikeyan,** University of Alabama at Birmingham
Reception and Poster session	
RRR Second Floor Terrace and Hallway	
5:15pm – 7:00pm	55 posters

Day 2, Saturday, March 23, 2024

8:00am – 5:00pm	Registration

Hackathon Organizer : Steve Pittard, Matthew Reyna, James Weigle, Mazen Istanbouli	
Room: R201	
9:30am – 7:30pm	

7:30am – 8:30am	Breakfast
Keynote Speech	
Room: Margaret H. Rollins Room	
8:25am – 8:30am	Introduction to keynote speaker **Zhaohui Qin** Professor Department of Biostatistics and Bioinformatics Emory University
8:30am – 9:30am	 Scientific and clinical implications of pervasive polygenic score-by-environment interactions **Greg Gibson,** Georgia Tech
9:30am – 9:45am	Coffee Break

Workshop: Using tidymodels for Building Better Models Instructor : Emil Hvitfeldt, Posit	
Room: Margaret H. Rollins Room	
9:45am – 12:30pm	

Breakout Session PS101: Statistical methods for integrative analysis of genetics and omics data Session Chair: Chang Su	
Room: R 100	
9:45am – 11:00am	1. Large-scale imputation models for multi-ancestry proteome-wide association analysis **Chong Wu**, MD Anderson Cancer Center 2. SAIGE-QTL: scalable and accurate expression quantitative trait locus mapping for single-cell studies **Wei Zhou**, Massachusetts General Hospital 3. A robust cis-Mendelian randomization method with application to drug target discovery **Zhaotong Lin**, Florida State University

Breakout Session PS102: Predictive Analysis of Disease Risk Factors from Genetic Complexity Session Chair: Jian Hu, Ph. D	
Room: RL 40	
9:45am – 11:00am	1. CompDA for Novel Health-Microbiome Associations and Controlling False Discoveries **Siyuan Ma,** Vanderbilt University 2. Graph-informed high-dimensional analysis for risk factor discovery and disease classification. **Jin Jin,** University of Pennsylvania 3. Pervasive correlations between causal disease effects of proximal SNPs vary with functional annotations and implicate stabilizing selection **Martin Zhang**, Carnegie Mellon University 4. MultiSTAAR: A statistical framework for powerful multi-trait rare variant analysis in large-scale whole-genome sequencing studies **Xihao Li**, The University of North Carolina at Chapel Hill

Breakout Session PS103: Potentials of Artificial Intelligence in Risk Assessment at Gastrointestinal Tract, Microbiome Research, and antibiotic resistance
Session Chair: Sangeeta Khare

Room: R 102

9:45am – 11:00am	1. Bile Acid Conjugations to Neuroactive Amines in the Human Gut Microbiome **Michael W. Mullowney,** University of Chicago 2. Gut bacteria and immune markers are associated with heat-related illness symptoms in agricultural workers **Madelyn C. Houser,** Emory University 3. Computational and high throughput screening approach to identify potential candidate for drug discovery against bacterial enzyme **Kuppan Gokulan,** NCTR/US Food and Drug Administration 4. Designing Synthetic Microbiomes: A Bayesian Optimization Approach **Nisha Pillai,** Mississippi State University

Breakout Session PS104: Computational methods for single cell spatial omics data
Session Chair: Zhana Duren

Room: RL 20

9:45am – 11:00am	1. Integrating transcriptomic and pathomic features to reconstruct 3D tissue maps with super-resolution **Mingyao Li,** University of Pennsylvania 2. Statistical identification of cell type-specific spatially variable genes in spatial transcriptomics **Xiang Zhou,** University of Michigan 3. Spatial location encoded in gene expression: a new analytical approach to spatial transcriptomic **Saurabh Sinha**, Georgia Institute of Technology & Emory University 4. Improving cell type deconvolution for spatial transcriptomics by leveraging temporal information **Xiuwei Zhang**, Georgia Institute of Technology

11:00am – 11:15am	Coffee Break

Breakout Session PS201: Computational methods for single cell multi-omics data
Session Chair: Xiuwei Zhang

Room: RL 20

11:15am – 12:30pm	1. Building interpretable regulatory maps using graph-embedding applied to single-cell multiomics data **Luca Pinello,** Massachusetts General Hospital & Harvard Medical School 2. PerturbNet predicts single-cell response to chemical and genetic perturbations **Joshua Welch,** University of Michigan 3. G3DC: a Gene-Graph-Guided selective Deep Clustering method for single cell RNA-seq data **Tianwei Yu,** Chinese University of Hong Kong in Shenzhen 4. Inferring gene regulatory networks from single cell multiome data using atlas-scale external data **Zhana Duren,** Clemson University

Breakout Session PS202: Computational Frontiers in Biomolecular Analysis: Navigating Sequence Data Landscapes
Session Chair: Inimary Toby-Ogundeji

Room: RL 40

11:15am – 12:30pm	1. Exploring the T-Cell immunome for pattern discovery in Acute Respiratory Distress Syndrome **Inimary Toby-Ogundeji,** University of Dallas 2. Integrating Multiscale Quantum Mechanical/Machine Learning Approaches Towards Rational Drug Design **Prajay Patel,** University of Dallas 3. Unraveling Poultry Salmonella Resistance to Diverse Antimicrobials with Genomic Sequences and Machine Learning **Moses Ayoola,** Mississippi State University 4. HyperGRN: Elucidate Gene Regulatory Networks in the Spatial Context **Yi Jiang**, The Ohio State University

Breakout Session PS203: Unraveling Intricate Molecular Pathways in Complex Diseases: Insights from Recent Statistical and Computational Advances
Session Chair: Yuehua Cui

Room: R 100

11:15am – 12:30pm	1. A Kernel-Based Neural Network for High-dimensional Risk Prediction on Massive Genetic Data **Qing Lu**, University of Florida 2. Advanced Learning Algorithms for Genetics and Genomics Data **Fei Zou**, University of North Carolina at Chapel Hill 3. High-resolution 3D reconstruction of spatial conformations for the human genome decodes the context-specific mechanisms of long-range genetic associations. **Jianrong Wang**, Michigan State University 4. Analyzing CITE-seq Data via a Quantum Algorithm **Ping Ma**, University of Georgia

Breakout Session PS204: A New Frontier in Public Health: Leveraging Bioinformatics and Generative AI for Advanced Disease Surveillance and Response
Session Chair: Kun Zhao

Room: R 102

11:15am – 12:30pm	1. A Comparative Study on Multiagent Systems for Advanced Metadata Retrieval with Large Language Models **Melchizedek Mashiku,** Centers for Disease Control and Prevention 2. Genomic and Geographic Disease Clusters of Invasive Pneumococcal Strains in the United States **Yuan Li,** Centers for Disease Control and Prevention

	3. Bioinformatics feasibility for accessing the presence of poliovirus in the Sequence Read Archive database **Katie Farrell,** Centers for Disease Control and Prevention

Lunch: 12:30pm-1:30pm

Breakout Session PS301: Advances in Computational Immuno-Oncology
Session Chair: Aik Choon Tan; Riyue Bao

Room: RL 20

1:30pm – 2:45pm	1. Harnessing Public Cancer Omics Data to Uncover Mechanisms of Immune Exclusion **Riyue Bao,** University of Pittsburgh 2. Computational and integrative understanding of the single-cell space for tumor microenvironment **Yi Zhang**, Duke University 3. Learning immune-related adverse events from real-world data **Aik Choon Tan**, University of Utah

Breakout Session PS302: Workshop: Glycoscience Informatics: Maximizing Insights with Database and Software Tools
Session co-chairs: Sushil K. Mishra & Robert J. Doerksen

Room: R 100

1:30pm – 2:45pm	This workshop aims to explore the latest advancements in database and software technologies tailored specifically for glycoscience research.
2:45pm – 3:00pm	Coffee Break

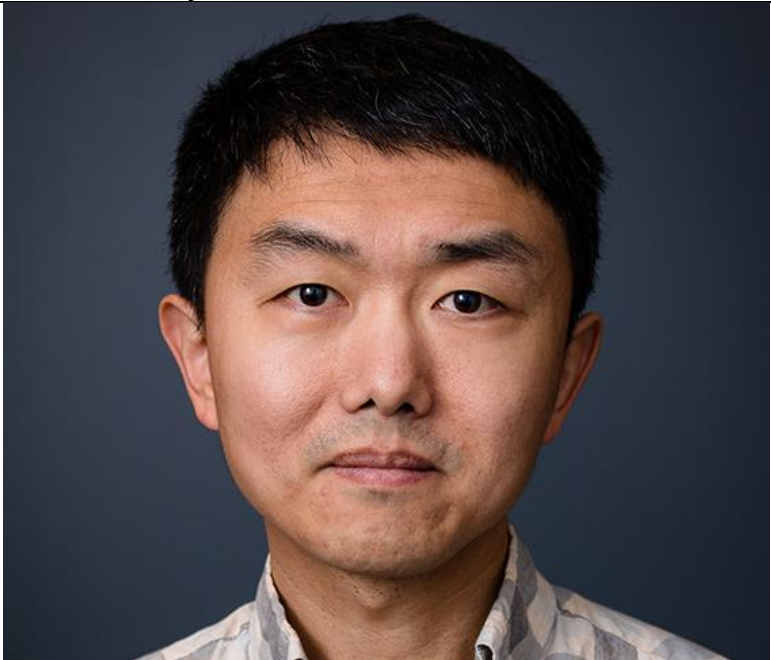| **Breakout Session PS401: Statistical Approaches for Interpreting Dynamic Biological Signals from Tissue Molecular and Histological Profiles** **Session Chair: Jian Hu**	
Room: RL 40	
3:00pm – 4:15pm	1. Interpretable models for uncovering dynamic biological signals in single-cell RNA-seq data **Rhonda Bacher**, University of Florida 2. Accurate and Efficient Integrative Reference-Informed Spatial Domain Detection for Spatial Transcriptomics. **Ying Ma**, Brown University 3. GeneSegNet: a deep learning framework for cell segmentation by integrating gene expression and imaging **Zhicheng Ji**, Duke University
Breakout Session PS402: Dawn of Foundation Models for Physiological Data **Session Chair: Xiao Hu**	
Room: RL 20	
3:00pm – 4:15pm	1. Do General Large Language and Vision Models Posses Pattern Recognition Abilities for Real-World Physiological Signals? **Darren Liu**, Emory University 2. Adapting a Generative Pretrained Transformer Achieves SOTA Performance in Assessing Diverse Physiological Functions Using Only Photoplethysmography Signals: A GPT-PPG Approach **Zhaoliang Chen**, Emory University 3. SiamQuality: A ConvNet-Based Foundation Model for Imperfect Physiological Signals **Cheng Ding**, Emory University 4. Using Irregular Time Series Transformers to Make Sense of Physiological Monitor Alarms **Nirbhay Modhe**, Emory University 5. Normalizing flow generative models by JKO scheme

	Chen Xue, The Georgia Institute of Technology

Breakout Session PS403: Advanced AI methods for improving precision medicine in complex diseases **Session Chair: Li Chen**	
Room: R 100	
3:00pm – 4:15pm	1. Harnessing big data for precision oncology **Leng Han,** Indiana University 2. Personalized cell-type-specific omics Profile Deconvolution and Inference **Hao Feng**, Case Western Reserve University 3. Advancing Systems Biology Approaches for Precision Medicine in Complex Diseases **Chi Zhang**, Indiana University

Breakout Session PS404: Artificial Intelligence Applications in Biomedical Research **Session Chair: Huixiao Hong**	
Room: RL 102	
3:00pm – 4:15pm	1. Deep Learning Approaches to Single and Multi-omics Imputation **Chaoyang Zhang**, University of Southern Mississippi 2. RxNorm Roulette: Unveiling the Challenges and Solutions in Drug Nomenclature Quality for Effective Safety Surveillance **Wenjing Guo**, National Center for Toxicological Research 3. An Effective Data Segmentation Method for Improving the Performance of Genotype Imputation **Lei Huang**, University of Southern Mississippi 4. Machine Learning and Deep Learning Approaches for Enhanced Prediction of hERG Blockade in Compound Assessment for Improved Drug Safety **Huixiao Hong**, National Center for Toxicological Research
4:15pm – 4:30pm	Coffee Break

Plenary Speech

Room: Margaret H. Rollins Room

4:30pm – 4:35pm	Introduction to keynote speaker **Yijuan Hu** Professor Department of Biostatistics and Bioinformatics Emory University
4:35pm – 5:35pm	Why polygenic scores for years of education has a negative slope across age cohorts in US biobank, despite increased rates of college education over the same age cohorts **Nancy Cox** Vanderbilt University Medical Center

Election and Banquet

Room: Margaret H. Rollins Room

5:35pm – 7:00pm	

Day 3, Sunday, March 24, 2024

8:00am – 12:00noon	Registration

7:30am – 8:30am	Breakfast
Keynote Speech	
Room: Margaret H. Rollins Room	
8:25am – 8:30am	Introduction to keynote speaker **Julia Wrobel** Assistant Professor Department of Biostatistics and Bioinformatics Emory University
8:30am – 9:30am	 **The assembly of a human pangenome** **Heng Li** Dana Farber Cancer Institute and Harvard Medical School
9:30am – 9:45am	Coffee Break

Breakout Session PS501: Applications of AI and ML in Human Health at Emory
Session Chair: Julia Wrobel

Room: RL 20

9:45am – 11:00am	1. Statistical inference of cell-type-specific gene co-expression networks with single cell and bulk RNA-seq data **Chang Su**, Emory University 2. Generalizable Machine Learning Methods for Electrophysiology **Samaneh Nasiri**, Emory University 3. TBA **Manoj Bhasin**, Emory University 4. Learning the Language of Life from DNA and Proteins **Yana Bromberg**, Emory University

Breakout Session PS502: Recent Insights Using Machine Learning in Genetics and Genomic Studies
Session Chair: Xiaoxi Shen

Room: R 102

9:45am – 11:00am	1. An Association Test Based on Kernel-Based Neural Networks for Complex Genetic Association Analysis **Tingting Hou**, University of Florida 2. High-Dimensional Genetic Survival Analysis with Kernel-Based Neural Networks **Rongzi Liu**, University of Florida 3. GAN-based Instrumental Variable Method for Causal Analysis of Omics Data **Yuan Zhou**, University of Florida 4. Applications of Neural Networks in Genetic Association Studies **Xiaoxi Shen**, Texas State University

Breakout Session PS503: Computational Approaches for Disease Prediction and Biomarker Selection in Multi-Scale Biological Systems
Session Chair: Xiunan Wang

Room: RL 40

9:45am – 11:00am	1. Multi-Scale Physics-Guided Graph Attention Networks to Predict COVID-19 Trends **Yu Liang**, University of Tennessee at Chattanooga
	2. Learning from Class-Imbalanced Data for Clinical Prediction Models: Cases of Borrelia and Lassa Infections **Winnie Mkandawire**, UMass Chan Medical School
	3. Using outbreak games to learn about real outbreaks **Andres Colubri**, Umass Chan Medical School
	4. Sparse Linear Centroid-Encoder: A Biomarker Selection tool for High Dimensional Biological Data **Tomojit Ghosh**, University of Tennessee at Chattanooga
	5. From policy to prediction: Forecasting COVID-19 Dynamics Under Imperfect Vaccination **Xiunan Wang**, University of Tennessee at Chattanooga

Breakout Session PS504: Advanced AI methods for analyzing single-cell and spatial omics data
Session Chair: Li Chen

Room: R 100

9:45am – 11:00am	1. Graph representation learning of single-cell and spatial transcriptomics data **Qin Ma,** Ohio State University
	2. AI-powered insights into the spatial variability of drug sensitivity **Qianqian Song**, University of Florida
	3. MorphLink: Mapping the Molecular Landscape to Cell Morphology in Spatial multi-Omics **Jian Hu**, Emory University

Breakout Session PS505: Lightning talks session 1
Session Chair: Nisha Pillai

Room: Margaret H. Rollins Room

9:45am – 11:00am	1. Unified somatic calling and machine learning-based classification enhance the discovery of clonal hematopoiesis of indeterminate potential **Shulan Tian**, Mayo Clinic 2. Data science-driven drug discovery and design for treatment of neurodegenerative diseases **Samuel Kakraba**, Tulane University 3. scMultiSim: simulation of multi-modality single cell data guided by cell-cell interactions and gene regulatory networks **Hechen Li**, Georgia Institute of Technology 4. MODEST: Modeling of Epigenomics and Spatial Transcriptomics Data **Ishita Debnath**, Clemson University 5. scaDA: A Novel Statistical Method for Differential Analysis of Single-Cell Chromatin Accessibility Sequencing Data **Fengdi Zhao**, University of Florida 6. Data-Driven Approach for Automatic Brain Layer Structures Annotation from Spatial Transcriptomics **Shihan Liu**, Emory University 7. High-resolution CNS-specific spatial architectures of chromosomes in Drosophila reveal higher order organization patterns of gene regulation **Xinrui Yu**, Michigan State University 8. scDisInFact: disentangled learning for integration and prediction of multi-batch multi-condition single-cell RNA-sequencing data **Ziqi Zhang**, Georgia Institute of Technology 9. From Static to Dynamic: Temporal Visualization of Gene Expression in Cancer Cell Line **Ehsan Saghapour**, University of Alabama at Birmingham 10. DrugPair2Vec: Drug Repurposing Networks via Data Mining for Exploring Drug Pair Combinations

	Kevin Song, University of Alabama at Birmingham 11. Low Rank Mixup Augmentations for Contrastive Learning of Phenotypes from Functional Connectivity **Anton Orlichenko**, Tulane University 12. Learning the language of metagenomes to understand molecular function **Prabakaran Ramakrishnan**, Emory University
11:00am – 11:15am	Coffee Break

Breakout Session PS601: Advanced statistical methods for omics data analysis Session Chair: Zhaohui "Steve" Qin	
Room: RL 20	
11:15am – 12:30pm	1. Pattern analysis on relational data unravels mRNA-protein dynamics over development time **Jun Xie**, Purdue University 2. Toward precise predicting phenotype of missense mutations in voltage-gated sodium channels **Jing Li**, University of Mississippi 3. A general testing method for inference of microbial networks with compositional data **Yi-Juan Hu**, Emory University
Breakout Session PS602: Application of AI in Infectious Diseases/Biological Data Sets Session Chair: Tomojit Ghosh	
Room: RL 40	
11:15am – 12:30pm	1. Classifying the SARS-CoV-2 variants with deep learning and exploring their functional regions in deep learning **Parisa Hatami**, University of Tennessee at Chattanooga 2. Discrete Inverse Method for Extracting Disease Transmission Rates from Accessible Infection Data **Xiunan Wang**, University of Tennessee at Chattanooga 3. Deep Learning Approaches for the Protein Scaffold Filling Problem **Richard Annan**, North Carolina A & T State University 4. Biomarker Selection using Masked Sparse Bottleneck Centroid-Encoder **Tomojit Ghosh**, University of Tennessee at Chattanooga
Breakout Session PS603: Network Medicine and Drug Discovery Session Chair: Zongliang Yue	
Room: R 100	

11:15am – 12:30pm	1. Decoding Cancer: Mapping Protein-Protein Interaction Networks to Identify Drug Targets **Ece Uzun**, Brown University 2. Network medicine framework reveals generic herb-symptom effectiveness of Traditional Chinese Medicine. **Xiao Gan,** Nanjing University 3. Molecular differentiation between complete and incomplete responders to neoadjuvant therapy in rectal cancer **Fengyuan Huang**, University of Alabama at Birmingham 4. TodenE: Topology-Based and Density-Based Ensembled Clustering for the Development of Super-PAG in Functional Genomics using PAG Network and LLM **Qi Li**, Auburn University

Breakout Session PS604: Delineate multi-level mechanisms of complex diseases under specific cellular contexts
Session Chair: Jianrong Wang

Room: R 102

11:15am – 12:30pm	1. One-sample Mendelian randomization for causal inference with MR-SPLIT **Yuehua Cui,** Michigan State University 2. Systematic study of stress response pathways using spatial transcriptomics data **Sha Cao,** Indiana University School of Medicine 3. Correcting intrinsic biases in bulk and single-cell chromatin profiling data **Chongzhi Zang,** University of Virginia 4. Bayesian integration of 3D chromatin interactions reveals complex long-range genetic associations **Jiwoong Kim,** University of South Florida

Breakout Session PS605: Lightning talks session 2
Session Chair: Qianqian Song

Room: Margaret H. Rollins Room

11:15am – 12:30pm	1. MedGCN: Query-time Character Level Embedding for Biomedical Named Entity

Recognition via Bidirectional Graph Convolutional Neural Networks
Islam Ebeid, Texas Woman's University
2. Fish oil supplementation modifies the genetic potential for blood lipids
Kaixiong Ye, University of Georgia
3. On Embedding Microbiome Abundances for Infection Detection
Santhana Krishnan Boopalan, Mississippi State University
4. Pan-cancer mutational signature surveys correlated racial disparity with geospatial environmental exposures, and viral infections
Judy Bai, Greenhills School
5. A Novel Neuro-Symbolic Frameworks for Bioactivity Classification and Drug Efficacy Prediction in Triple-Negative Breast Cancer
Delower Hossain, The University of Alabama at Birmingham
6. UAB Systems Pharmacology AI Research Center (SPARC)
Zhandos Sembay, The University of Alabama at Birmingham
7. Exploring Pathogen Prediction in Food Safety through Transformer-Based Models and Enhanced Attention Mechanism Explainability
Athish Ram Das, Mississippi State University
8. Latent class analysis-based identification of non-alcoholic fatty liver disease subtypes with distinct metabolic and genetic signatures
Tahmina Sultana Priya, Virginia Tech
9. Molecular Dynamics and Machine Learning Reveal Altered Binding Dynamics of Belzutifan to Threonine 324-phosphorylated HIF-2α:ARNT
Vishva Natarajan, Georgia Institute of Technology
10. Aicardi–Goutières syndrome (AGS) - orthologous mutation in RNaseH2C subunit

	mimics embedded ribonucleotides patterns of mutation in RNaseH2B PIP-box **Deepali Kundnani**, Georgia Institute of Technology 11. Learning Control Policies of Hodgkin-Huxley Neuronal Dynamics **Malvern Madondo**, Emory University
Awards and Closing	
Room: Margaret H. Rollins Room	
12:30pm – 1:00pm	

Poster information

1. BIOSequenceSearchAPI: A Comprehensive API for Genome Sequence Analysis
   Pelumi Abimbola
   Mississippi State University

2. Predicting Osteoporosis Risk from Imbalance Data Using Deep Learning Methods
   Mahid Ahmed
   Computing Science & Computer Engineering, University of Southern Mississippi

3. Exploring Mutation Effects in Sodium and Calcium Channels: A Data-Driven Approach for Machine Learning Applications
   Amin Akbari Ahangar
   The University of Mississippi

4. Batch mapping and out-of-reference discovery with heterogeneous single-cell ATAC-seq features
   Yuqi Cheng
   Georgia Tech

5. MODEST: Modeling of Epigenomics and Spatial Transcriptomics Data
   Ishita Debnath
   Clemson University Center for Human Genetics

6. Predicting the evolution of virus genome with deep learning
   Yibo Dong
   Mississippi State University

7. stPularity improves clustering for spot-based spatial transcriptomics data
   Chandra Sekhar Reddy Edula
   Georgia Institute of Technology

8. Network medicine framework reveals generic herb-symptom effectiveness of Traditional Chinese Medicine
   Xiao Gan
   Nanjing University of Information Science & Technology

9. Identification of PTSD Subtypes through Pathway-Based Molecular Biomarkers
   Maziar Ganji
   Department of Biological Sciences, The University of Memphis, TN, USA

10. Counterfactual Explanations in Microbiome Analysis
    GANGA GIREESAN
    Mississippi State University

11. Genetic and social risk factors for type 2 diabetes health disparities: a test of the Rose hypothesis
    Sonali Gupta
    Applied Bioinformatics Laboratory (ABiL)

12. A Novel Neuro-Symbolic Frameworks for Bioactivity Classification and Drug Efficacy Prediction in Triple-Negative Breast Cancer
   Delower Hossain
   The University of Alabama at Birmingham

13. Logic Tensor Networks: A Neuro-Symbolic AI Approach for Precise hERG Cardiotoxicity Assessment in Drug Development
   Md Delower Hossain
   The University of Alabama at Birmingham

14. A Neuro-symbolic AI Approach to Identifying Potent DPP-4 Inhibitors for Diabetes Treatment
   Delower Hossain
   UAB Systems Pharmacology AI Research Center (SPARC)

15. Acute Stress Disorder: Epigenomic Comparisons Across Cohorts
   Mazen Istanbouli
   University of Memphis, WRAIR ORISE Fellow

16. Automatic Generation of Survivorship Care Plans
   Neha Jain
   Georgia Institute of Technology

17. Self Organizing Map For Interpretable Genomic Clustering for Autoimmune Phenotype Prediction
   Neha Jain
   Georgia Institute of Technology

18. The Effects of Adipokines Leptin and Adiponectin on Cardiovascular Health
   Jordan Johnson
   Mississippi Valley State University

19. Exploring Autoimmune Diseases using Gene Ontology Analyses
   Asma Khimani
   Bio-MIBLab at Georgia Tech

20. Aicardi–Goutières syndrome (AGS) - orthologous mutation in RNaseH2C subunit mimics embedded ribonucleotides patterns of mutation in RNaseH2B PIP-box
   Deepali Kundnani
   School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA

21. scMultiSim: simulation of multi-modality single cell data guided by cell-cell interactions and gene regulatory networks
   Hechen Li
   Georgia Institute of Technology

22. Data-Driven Approach for Automatic Brain Layer Structures Annotation from Spatial Transcriptomics
   Shihan Liu
   Emory University

23. Improving cell type deconvolution for spatial transcriptomics by leveraging temporal information
Macrina Lobo
School of Computational Science and Engineering, Georgia Institute of Technology

24. Deep5hmC: Predicting genome-wide 5-Hydroxymethylcytosine landscape via multimodal deep learning model
Xin Ma
Department of Biostatistics, University of Florida

25. Single-cell RNA sequencing analysis reveals prognostic markers for lung adenocarcinoma
Jialu Ma
University of Arkansas at Little Rock

26. Prediction of Chromatin Accessibility from Personal Genetic Variants and Transcription Factor Expression
Hui Ma
Clemson University

27. Learning Control Policies of Hodgkin-Huxley Neuronal Dynamics
Malvern Madondo
Emory University, Department of Computer Science

28. A Neural Ordinary Differential Equation Approach to Glucose-Insulin Control
Malvern Madondo
Emory University, Department of Computer Science

29. Query and Prediction of Drug-Drug Interactions
Alaa Marwan Abu Taha
UAMS Graduate School

30. Creating an Adaptive Epigenetics Machine Learning Model for Disease Prevention
Sreya Modepalli
University of Michigan-Flint

31. Build a pipeline to integrate single-cell RNA-seq & ATAC-seq data by benchmarking current popular methods
FNU Naqing
Center for Human Genetics of Clemson University

32. Byte Pair Encoding: Enhancing Tokenization for LLM-Based Genomic Sequence Analysis
Ramak Nassiri
Mississippi State University

33. Molecular Dynamics and Machine Learning Reveal Altered Binding Dynamics of Belzutifan to Threonine 324-phosphorylated HIF-2α:ARNT
Vishva Natarajan
Georgia Institute of Technology

34. Structural Insights Into A2A Adenosine Receptor Inhibition Through Virtual Screening and Molecular Dynamic Simulations
Lance Nuique

Bowie State University

35. Development of novel therapeutic medication for Amyotrophic Lateral Sclerosis (ALS) using AI and Molecular Modeling
Anulika Nwashili
AR-INBRE, University of Arkansas at Little Rock

36. Drug-UGT interactions as Predictors of Drug-Induced Liver Injury
Ayooluwa Olubamiwa
National Center for Toxicological Research (NCTR), U.S. Food and Drug Administration

37. Low Rank Mixup Augmentations for Contrastive Learning of Phenotypes from Functional Connectivity
Anton Orlichenko
Tulane University

38. LinRace: Cell Division History Reconstruction of Single Cells Using Paired Lineage Barcode and Gene Expression Data
Xinhai Pan
Georgia Institute of Technology

39. A More Informative Approach to Identify Cancer Survival Subgroups using Integrative Omics Data
Imran Parvez
Augusta University

40. Latent class analysis-based identification of non-alcoholic fatty liver disease subtypes with distinct metabolic and genetic signatures
Tahmina Sultana Priya
Virginia Tech

41. Macrophage-Mediated Progression of Breast Cancer Bone Metastasis: A Mouse Gene Array Analysis
Vinay Raj
University of Arkansas at Pine Bluff

42. Learning the language of metagenomes to understand molecular function
Prabakaran Ramakrishnan
Department of Biology, Emory University

43. UAB Systems Pharmacology AI Research Center (SPARC)
Zhandos Sembay
UAB

44. Diversity in the All of Us Research Program: race, ethnicity, and genetic ancestry
Shivam Sharma
Georgia Institute of Technology

45. Two-Stage Procedure for the Detection of Differentially Methylated Regions Using WGBS Data
Kathryn Sinha
Augusta University

46. DrugPair2Vec: Drug Repurposing Networks via Data Mining for Exploring Drug Pair Combinations
Kevin Song
University of Alabama at Birmingham

47. Enhancing Transcription Factor Binding Site Prediction with Epigenetic Data: A BERT-Based Machine Learning Innovation
Tairan Song
Michigan State University

48. PharmAlchemy: A Comprehensive Knowledge Base for Mining Gene-Drug-Disease Relationships
Kevin Song
University of Alabama at Birmingham

49. Advancing Antimicrobial Resistance Prediction through Residue Graph Learning in Graph Neural Networks
Ramyasri Veerapaneni
Computer Science and Engineering, Mississippi State University

50. The Art of Feature Embedding: Boosting Transformer Capabilities for Superior Structured Data Analysis
Aditi Virupakshaiah
Mississippi State University

51. Adverse Drug Reaction Profiling in Antidepressant Drug Labeling by Utilizing FDALabel
Lan Ying
FDA NCTR

52. High-resolution CNS-specific spatial architectures of chromosomes in Drosophila reveal higher order organization patterns of gene regulation
Xinrui Yu
Michigan State University

53. scDisInFact: disentangled learning for integration and prediction of multi-batch multi-condition single-cell RNA-sequencing data
Ziqi Zhang
Georgia Institute of Technology

54. scaDA: A Novel Statistical Method for Differential Analysis of Single-Cell Chromatin Accessibility Sequencing Data
Fengdi Zhao
University of Florida

Introduction To TidyModels

Emil Hvitfeldt
Software Engineer at Posit PBC
Saturday, March 23rd
9:45 AM – 12:15 PM
Via Zoom @ Margaret H. Rollins Room
R. Randall Rollins Building

MCBIOS is delighted to be hosting this introductory workshop on Tidymodels, which provides a cohesive ecosystem built on tidy principles, offering a streamlined approach to modeling that emphasizes clarity, flexibility, and reproducibility. This workshop will teach core Tidymodels packages and their uses: data splitting/resampling with rsample, model fitting with parsnip, measuring model performance with yardstick, and basic pre-processing with recipes. You'll learn Tidymodels syntax as well as the process of predictive modeling for tabular data. Time permitting, you'll be introduced to model optimization using the tune package.

Knowledge Level: Basic knowledge of tidyverse packages to read, transform and reshape data, and create basic graphs using ggplot2. Previous exposure to basic statistical concepts such as linear models and associated model diagnostic concepts. Attendees are encouraged to attend on-site but remote Zoom access is available.

Instructor: Emil Hvitfeldt is co-author or "Supervised Machine Learning for Text Analysis in R". His interests include developing tools for natural language processing, machine learning using Tidymodels, education, and the use of colors in data visualizations.

Questions: wsp@emory.edu
https://zoom.us/j/94613421671?pwd=RkptRTlaOW01a2h5dW5YbUtTSm1zZz09

Hackathon

We are pleased to announce the 2024 MCBIOS Hackathon which is open to all attendees. The effort provides an opportunity to sharpen your skills on important data while also learning from your peers in a friendly collegial environment. This year's hackathon will leverage the ongoing George B. Moody PhysioNet Challenge which "invites teams to develop algorithms for digitizing and classifying electrocardiograms (ECGs) captured from images or paper printouts"

From the Moody Challenge website:

> "Despite recent advances in digital ECG devices, paper or physical ECGs remain common, especially in the Global South. These paper ECGs document the history and diversity of cardiovascular diseases (CVDs), and algorithms that can digitize and classify these images have the potential to improve our understanding and treatment of CVDs, especially for underrepresented and underserved populations"

> We ask participants to design and implement working, open-source algorithms that, based only on the provided ECG images, reconstruct the waveforms and/or classify or diagnose the images. The teams with the best scores for these tasks on the hidden test set win the Challenge"

While official registration into the larger, ongoing Moody Physionet Challenge is not required it is helpful to track your team name and members.

The MCBIOS Hackathon is an economized version of the Moody Challenge in terms of duration in recognition that participants will also wish to attend conference events and experience networking opportunities. The conference will provide a breakout room for use by hackathon participants and teams which is in close proximity to the main meeting room.

Attendees will develop code using personally available computational resources which may (or not) include use of GPUs although your submission should make clear if your code is reliant upon that.

To get started please consult the "rules" section for specifics on objectives, data sources, deliverables and how to appropriately format your submission. The deadline for your submission is 9PM on Saturday March 23rd after which evaluation will take place followed by formal recognition of winners on Sunday the 24th.

Organizing Committee

Zhaohui "Steve" Qin (co-chair)	**Emory University**
Chang Su (co-chair)	**Emory University**
Hong Qin	U of Tennessee Chattanooga
Huixiao Hong	NCTR /FDA
Jake Chen	U of Alabama Birmingham
Jie Chen	Augusta University
Kun Zhao	CDC
Manoj Bhasin	Emory University
Mark Borodovsky	Georgia Tech
Matthew A. Reyna	Emory University
Paul Schliekelman	University of Georgia
Peng Qiu	Georgia Tech
Sangeeta Khare	NCTR/FDA
Steve Pittard	Emory University
Xiao Hu	Emory University
Xiuwei Zhang	Georgia Tech
Yana Bromberg	Emory University
Yanqing Zhang	Georgia State University
Yongsheng Bai	Eastern Michigan University
Zhana Duren	Clemson University
Zongliang Yue	Auburn University

Acknowledgement
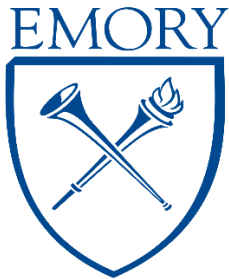
Provost office
Emory University

Rollins School of Public health
Dean's office
Emory University

Department of Biostatistics and Bioinformatics

Department of Human Genetics

Center for AI Learning

The Center of Computational and Quantitative Genetics (CCQG)

EMORY Artificial Intelligence in Genomics

Sponsors

The Department of Computer Science and Engineering at the University of Tennessee at Chattanooga (UTC) offers the following graduate degrees:

Master of Science (MS) in Computer Science: This program is designed for graduate students seeking to deepen their knowledge in computer science with advanced coursework in areas such as artificial intelligence, databases, networking, and software engineering. It also offers opportunities for research projects and hands-on experience.

Master of Science (MS) in AI and Data Science: This program focuses on artificial intelligence, machine learning, data science, and statistical modeling, teaching students how to derive meaningful insights from large datasets.

Master of Science (MS) in Cyber Security: Dedicated to the field of cybersecurity, this program equips professionals with the skills needed to protect digital infrastructures and networks. The curriculum covers cryptography, network security, and ethical hacking, among other subjects.

Doctoral Program (PhD): Specializing in Computational Science with a concentration in Computer Science, this program is designed for those aiming to contribute to the field through research and advanced technical knowledge.

Notes

Notes

Notes