

ACM-BCB 2020 Advance Program

All sessions will become available live to all registered attendees, via the ACM-BCB 2020 Virtual Conference Platform: <https://acm-bcb-virtual.org/>

Day 0 (Sep. 21; Monday), Workshops & Tutorials

Tutorials (Chairs: Mukul Bansal, Ercument Cicek)

- | | |
|-----------|---|
| 2-5pm | <i>Homomorphic encryption and its application to privacy-preserving genotype imputation</i>
Gamze Gursoy, Michael Maniatakos, Eduardo Chielle, Oleg Mazonka |
| 3-5pm | <i>Phylogenetic Tree Reconciliation from Practice to Theory</i>
Ran Libeskind-Hadas |
| 12-2pm | <i>Privacy-Preserving Genomic Data Sharing</i>
Erman Ayday, Xiaoqian Jiang |
| 9am-12pm | <i>From Multi-omics Data to Knowledge Networks</i>
Rabie Saidi, Maryam Abdollahyan, Maria J Martin |
| 11am-12pm | <i>SC1: Interactive web-based single cell RNA-seq analysis</i>
Marmar Moussa, Ion Mandoiu |
| 1-3pm | <i>Computational Analysis of Viral Genomes and Outbreaks</i>
Pavel Skums, Alex Zelikovsky, Sergey Knyazev |
| 9-11am | <i>Fundamentals of Alignment-free Sequence Analysis: k-mer hashing</i>
Sven Rahmann, Jens Zentgraf |

Workshops (Chairs: Xia Ning, Wei Zhang)

11am-1:30pm **MODI: Workshop on Machine Learning Models for Multi-omics Data Integration**
Organizers: Abed Alkhateeb, Luis Rueda

Diagnosing COVID-19 in X-ray Images Using HOG Image Feature and Artificial Intelligence Classifiers.

Faten Kharbat, Tarik Ekamsy, Nuha H. Hamada

Convolutional Neural Network Strategy for Skin Cancer lesions Classification and Detections.

Abdala Nour, Boubakeur Boufama

Mining Frequent Dense Subgraphs from Multiple Gene Coexpression Networks Using Post-Processing Summarization.

San Ha Seo, Saeed Salem

Link Analysis to Discover Insights from Structured and Unstructured Data on COVID-19.

Ying Zhao, Charles Zhou

Machine Learning Model to Track SARS-CoV-2 Viral Mutation, Evolution and Speciation Using Next-generation Sequencing Data. (poster)

Iulian Derecichei, Govindaraja Atikukke

Deep Learning Approach for Breast Cancer InClust 5 Prediction based on Multiomics Data Integration.

Abedalrhmman Alkhateeb, Li Zhou, Ashraf Abou Tabl, Luis Rueda

8:45am-6:10pm **CAME: 9th Workshop on Computational Advances in Molecular Epidemiology**
Organizers: Yury Khudyakov, Ion Mandoiu, Pavel Skums, Alex Zelikovsky
(access via: <https://alan.cs.gsu.edu/came20/>)

8:45 Workshops Chairs Welcome

9:00 Session 1: Molecular Epidemiology

Ethical aspects of genetic testing for public health.

Lilia Ganova-Raeva, Yury Khudyakov

Scale-free Spanning Trees and their Applications in Genomic Epidemiology.

Yury Orlovich, Kirill Kukhareenko, Volker Kaibel, Pavel Skums

A genetically guided strategy for optimal contact tracing of viral infections.

David S. Campo, Yury Khudyakov

Signals from Noise: Enabling Real-Time Disease Surveillance from Clinical Data Sources.

Wade Schulz

10:00 GatherTown Discussion

11:00 Session 2: SARS-CoV-2

Characterizing SARS-CoV-2 viral diversity within and between hosts.

Mohammed El-Kebir

Reconstructing transmission networks of SARS-CoV-2: Methodology analysis.

Pavel Skums

Inferring the international COVID-19 transmission network using TNet.

Saurav Dhar, Ion Mandoiu, Mukul Bansal

Phylogeny networks for SARS-CoV-2 genomes.

Sergei Knyazev, Harman Singh, Varuni Sarwal, Ram Ayyala, Roya Hosseini, Serghei Mangul, Alex Zelikovsky and Daniel Novikov

12:00 GatherTown Discussion and Lunch

1:00 Session 3: SARS-CoV-2

The emergence of SARS-CoV-2 in Europe and North America.

Joel Wertheim

On regularized predictor-corrector algorithm for modeling and forecasting COVID-19.

Brian Pidgeon, Alexandra Smirnova

Clonal interference of D614G mutation in Sars-Cov-2 using a Monte Carlo and Epidemic Modelling.

Bart Weimer

1:45 GatherTown Discussion

2:15 Session 4: Cancer Genomics

Quantifying the contribution of somatic mutational phenotypes to cancer progression risk.

Russel Schwartz

Scaling state-of-the-art character-based phylogeny reconstruction methods to the dimension of the latest datasets.

Robert Harrison

Structural characterization and modulation of protein interaction networks.

Ilya Vakser

Automated construction/extraction of DNA/RNA/Protein-based biomolecular markers from NGS data.

James Lara, Yury Khudyakov

pyTCR: Comprehensive platform for TCR-Seq data analysis using interactive notebooks to facilitate reproducibility of immunogenomics research.

Kerui Peng, Serghei Mangul

5:00 GatherTown Discussion

5:30 Session 6: Machine Learning

Machine Learning for improved prioritization of Variants of Uncertain Significance.

Daniel Mahecha, Haydemar Nuñez, Maria Claudia Lattig, Jorge Duitama

Automated Genomic featurE Engineering and Machine LearnINg (GREMLIN).

Walker Gussler, David S. Campo, Pavel Skums and Yury Khudyakov

Using genetics and systems biology to inform precision medicine for lung disorders.

Murray Cairns, William Reay, Sahar El Shair, Michael Geaghan

10-11:55am ParBio: 9th Workshop on Parallel and Cloud-based Bioinformatics and Biomedicine

Organizers: Giuseppe Agapito, Wes Lloyd, Mario Cannataro

10:00 *An Investigation on Public Cloud Performance Variation for an RNA Sequencing Workflow.*

10:20 *Interactive exploration of population scale pharmacoepidemiology dataset.*

10:40 *SparkBeagle: Scalable Genotype Imputation from Distributed Whole-Genome Reference Panels in the Cloud.*

11:00 *Hadoop-CNV-RF: A Scalable Copy Number Variation Detection Tool for Next-Generation Sequencing Data.*

11:20 *Asymptotically Stable Privacy Protection Technique for fMRI Shared Data over Distributed Computer Networks.*

9am-3pm **CSBW: Computational Structural Bioinformatics Workshop**

Organizers: Filip Jagodzinski, Kevin Molloy, Brian Chen

9:00 Session 1: Cryo-EM

A Divide and Conquer Algorithm for Electron Microscopy Segmentation.

Ruba Jebiril, Yingde Zhu, Wei Chen, Kamal Al Nasr

Combine Cryo-EM Density Map and Residue Contact for Protein Structure Prediction – A Case Study.

Maytha Alshammari, Jing He

Segmentation based Feature Extraction for Cryo Electron Microscopy at Medium Resolution.

Lin Chen, Ruba Jebiril, Kamal Al Nasr

10:10 Break

10:20 Session 2: Machine Learning-based Approaches

From Interatomic Distances to Protein Tertiary Structures with a Deep Convolutional Neural Network.

Yuanqi Du, Anowarul Kabir, Liang Zhao, Amarda Shehu

ProLanGO2: Protein Function Prediction with Ensemble of Encoder-Decoder Networks.

Kyle Hippe, Sola Gbenro, Renzhi Cao

Efficient Exploration of Protein Conformational Pathways using RRT and MC.*

Fatemeh Afrasiabi, Nurit Haspel

11:30 Break

11:50 Session 3: Panel

Industry and Academia Careers in Bioinformatics: Current Trends and Prospects for the Future

Naomi Fox, Kameron Decker Harris, Anna Ritz, Amarda Shehu

12:30 Break

1:30 Session 4: Protein Structure and Function

Using Guided Motion Planning to Study Binding Site Accessibility.

Diane Marie Bernard Uwacu, Abigail Ren, Shawna Thomas, Nancy M. Amato

HMMeta: Protein Function Prediction using Hidden Markov Models.

Sola Gbenro, Kyle Hippe, Renzhi Cao

Interpretable Molecule Generation via Disentanglement Learning.

Yuanqi Du, Xiaojie Guo, Liang Zhao, Amarda Shehu

Impactful Mutations in Mpro of the SARS-CoV-2 Proteome.

Gideon Wolfe, Othmane Belhoussine, Anais Dawson, Maxwell Lisaius, Filip Jagodzinski

3:00 Break

3:20 Session 5: Docking, Mutations, and Protein Binding

Binding Free Energy of the Novel Coronavirus Spike Protein and the Human ACE2 Receptor: An MMGB/SA Computational Study.

Negin Forouzes

Using Curriculum Learning in Pattern Recognition of 3-dimensional Cryoelectron Microscopy Density Maps.

Yangmei Deng, Yongcheng Mu, Salim Sazzed, Jiangwen Sun, Jing He

Assessing Drug Resistance Due to Mutations via Energy Minimization Profiles.

Edward Thompson, Tess Thackray, Cecilia Kalthoff, Ryan Rapoport, Filip Jagodzinski

Using player generated data to elucidate molecular docking.

Torin Adamson, Selina Bauernfeind, Bruna Jacobson, Lydia Tapia

1-4pm

HPC-BOD: High Performance Computing, Big Data Analytics & Integration for Multi-Omics Biomedical Data

Organizers: Fahad Saeed, Serdar Bozdog

- 1:00 Mining representative approximate frequent coexpression subnetworks.
San Ha Seo and Saeed Salem
Fusion Transcript Detection from RNA-Seq using Jaccard Distance.
Hamid Reza Mohebbi, Joyce Quach, Nurit Haspel
CanMod: A computational model to identify co-regulatory modules in cancer.
Duc Do, Serdar Bozdog
Large-Scale Machine Learning and Optimization for Bioinformatics Data Analysis.
Jianlin Cheng
Strategies to integrate multi-omics data for patient survival prediction.
Lana Garmire

3:30 Q&A and discussion

Day 1 (Sep. 22; Tuesday), Cancer Theme

8:40-10:30 Session 1A: Sequences & Networks I

(Chair: Jaroslaw Zola)

- 8:40 *Augmenting Signaling Pathway Reconstructions.*
Tobias Rubel Janssen and Anna Ritz.
- 9:00 *Identifying Evolutionary Origins of Repeat Domains in Protein Families.*
Chaitanya Aluru and Mona Singh.
- 9:20 *Modularity Analysis of Bipartite Networks and Multivariate ANOVA for Identification of Differentially Expressed Proteins in a Mouse Model of Down Syndrome.*
Ali Jazayeri, Sara Pajouhanfar, Sadaf Saba and Christopher Yang.
- 9:40 **(Research Highlight)** *Pairwise Versus Multiple Global Network Alignment.*
Vipin Vijayan, Shawn Gu, Eric Krebs, Lei Meng and Tijana Milenkovic.
- 10:00 *Transforming the Language of Life: Transformer Neural Networks for Protein Prediction Tasks.*
Ananthan Nambiar, Maeve Heflin, Simon Liu, Sergei Maslov, Mark Hopkins and Anna Ritz.
- 10:12 Live Q&A

8:40-10:30 Session 1B: Cancer Omics I

(Chair: Luis Rueda)

- 8:40 *Automated Classification of Acute Rejection from Endomyocardial Biopsies.*
Felipe Giuste, Mythreye Venkatesan, Conan Zhao, Li Tong, Yuanda Zhu, Shriprasad Deshpande and May Wang.
- 9:00 *CTDPathSim: Cell line-tumor deconvoluted pathway- based similarity in the context of precision medicine in cancer.*
Banabithi Bose and Serdar Bozdog
- 9:20 *Integrative Deep Learning for PanCancer Molecular Subtype Classification Using Histopathological Images and RNAseq Data.*
Fatima Zare, Javad Noorbakhsh, Tianyu Wang, Jeffery Chuang and Sheida Nabavi.
- 9:40 *Functional Enrichment Analysis of Deregulated Long Non-Coding RNAs in Cancer Based on their Genomic Neighbors.*
Gulden Olgun and Ozgur Tastan
- 10:00 *Fusion Lasso and Its Applications to Cancer Subtype and Stage Prediction.*
Zhong Chen, Andrea Edwards and Kun Zhang.
- 10:12 Live Q&A

8:40-10:30 Session 1C: Regulatory Genomics

(Chair: Niina Haiminen)

- 8:40 **(Research Highlight)** *How to build regulatory networks from single-cell gene expression data?*
Aditya Pratapa, Amogh Jalihal, Jeffrey Law, Aditya Bharadwaj and T. M. Murali.
- 9:00 *The impact of sample size and tissue type on the reproducibility of gene co-expression networks.*
Katie Ovens, Brian Eames and Ian McQuillan.

- 9:20 *A Generalized Robinson-Foulds Distance for Clonal Trees, Mutation Trees, and Phylogenetic Trees and Networks.*
Mercè Llabrés, Francesc Rosselló and Gabriel Valiente.
- 9:40 *Three Co-expression Pattern Types across Microbial Transcriptional Networks of Plankton in Two Oceanic Waters.*
Ruby Sharma, Xuye Luo, Sajal Kumar and Mingzhou Song.
- 10:00 *Translocator: local realignment and global remapping enabling accurate translocation detection using single-molecule sequencing long reads.*
Ye Wu, Ruibang Luo, Tak-Wah Lam, Hing-Fung Ting and Junwen Wang.
- 10:12 Live Q&A
- 10:30-10:40 Break
- 10:40-12:00 Plenary Session 1**
- 10:40 Conference Remarks by General Chairs and Program Chairs
- 10:55 Introduction (Chair: Yan Liu)
- 11:00 Keynote Address: Olivier Elemento**
Using Multi-omics and AI to Accelerate Precision Medicine
- 12:00-12:30 Poster Session 1 (13 posters; 2 min each): Session Chairs: Chirag Jain, Jaric Zola**
- 12:30-12:50 Break
- 12:50-1:55 Featured Session on Funding (Chair: May Wang)**
Speakers:
Dr. Jean Yuan, NIH Program Director on Translational Bioinformatics
Dr. Veerasamy “Ravi” Ravichandran, NIH Program Director on Bioinformatics & Computational Biology
Dr. Susan Gregurick, NIH, Associate Director for Data Science
- 2:00-3:50 Session 2A: Sequences & Networks II**
(Chair: Dan DeBlasio)
- 2:00 *A Supervised Machine Learning Approach for Distinguishing Between Additive and Replacing Horizontal Gene Transfers.*
Abhijit Mondal, Misagh Kordi and Mukul S. Bansal.
- 2:20 **(Research Highlight)** *SMART: SuperMaximal Approximate Repeats Tool.*
Lorraine Ayad, Panagiotis Charalampopoulos and Solon Pissis.
- 2:40 **(Research Highlight)** *NeTFactor, a framework for identifying transcriptional regulators of gene expression-based biomarkers.*
Mehmet Eren Ahsen, Yoojin Chun, Alexander Grishin, Galina Grishina, Gustavo Stolovitzky, Gaurav Pandey and Supinda Bunyavanich.
- 3:00 *HMSC: a Hybrid Metagenomic Sequence Classification Algorithm.*
Subrata Saha, Zigeng Wang and Sanguthevar Rajasekaran.
- 3:12 *Structural representations of DNA regulatory substrates can enhance sequence-based algorithms by associating functional sequence variants.*
Jan Zrimec.
- 3:24 *Characterization of S. cerevisiae Protein Complexes by Representative DDI Graph Planarity.*

William Gasper, Kate Cooper, Nathan Cornelius, Sanjukta Bhowmick and Hesham Ali.

3:36 Live Q&A

2:00-3:50 Session 2B: Cancer Omics II

(Chair: Gaurav Pandey)

2:00 **(Research Highlight)** *Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images.*

Bino Varghese, Frank Chen, Darryl Hwang, Suzanne Palmer, Andre Luis De Castro Abreu, Osamu Ukimura, Monish Aron, Manju Aron, Inderbir Gill, Vinay Duddalwar and Gaurav Pandey.

2:20 *TLSurv: Integrating Multi-Omics Data by Multi-Stage Transfer Learning for Cancer Survival Prediction.*

Yixing Jiang, Kristen Alford, Frank Ketchum, Li Tong and May D. Wang.

Rhabdomyosarcoma Histology Classification using Ensemble of Deep Learning

2:40 *Networks.*

Saloni Agarwal, Mohamedelfatih Eltigani Osman Abaker, Xinyi Zhang, Ovidiu Daescu, Donald A. Barkauskas, Erin R Rudzinski and Patrick Leavey.

3:00 *A multi-context feature learning approach to identify disease-specific gene neighborhoods*

Sudhir Ghandikota and Anil Jegga.

3:12 *Population-scale Genomic Data Augmentation Based on Conditional Generative Adversarial Networks.*

Junjie Chen, Mohammad Mowlaei and Xinghua Shi.

3:24 Live Q&A

2:00-3:50 Session 2C: Structural Bioinformatics

(Chair: Yu Zhang)

2:00 *Variational Autoencoders for Protein Structure Prediction.*

Fardina Alam and Amarda Shehu.

2:20 *Protein evolution is structure dependent and non-homogeneous across the tree of life.*

Akanksha Pandey and Edward Braun.

2:40 **(Research Highlight)** *Boosting the accuracy of protein secondary structure prediction through nearest neighbor search and method hybridization.*

Spencer Krieger and John Kececioglu.

3:00 *Predicting protein secondary structure by an ensemble through feature-based accuracy estimation.*

Spencer Krieger and John Kececioglu.

3:20 *Deep Ranking in Template-free Protein Structure Prediction.*

Xiao Chen, Nasrin Akhter, Zhiye Guo, Tianqi Wu, Jie Hou, Amarda Shehu and Jianlin Cheng.

3:40 Live Q&A

3:50-4:00 Break

4:15-4:55 SIGBIO General Meeting (chair: Srinivas Aluru)

5:00-6:00 Virtual Tour of Atlanta

Day 2 (Sep. 23; Wednesday), Theme: COVID-19

8:40-10:30 **Session 3A: COVID-19 I**

(Chair: Hongmei Jiang)

8:40 *A Preliminary Investigation in the Molecular Basis of Host Shutoff Mechanism in SARS-CoV.*

Niharika Pandala, Casey A. Cole, Devaun McFarland, Anita Nag and Homayoun Valafar.

9:00 *A Qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19.*

David Oniani and Yanshan Wang.

9:20 *Global Surveillance of COVID-19 by mining news media using a multi-source dynamic embedded topic model.*

Yue Li, Pratheeksha Nair, Zhi Wen, Imane Chafi, Anya Okhmatovskaia, Guido Powell, Yannan Shen and David Buckeridge.

9:40 *Performance Evaluation of Viral Infection Diagnosis using T-Cell Receptor Sequence and Artificial Intelligence.*

Tim Kosfeld, Jonathan McMillan, Richard DiPaolo, Jie Hou and Tae-Hyuk Ahn.

10:00 *EXAM: An Explainable Attention-based Model for COVID-19 Automatic Diagnosis.*

Wenqi Shi, Li Tong, Yuchen Zhuang, Yuanda Zhu and May Wang.

10:12 Live Q&A

8:40-10:30 **Session 3B: HTS Data I**

(Chair: TM Murali)

8:40 **(Research Highlight)** *Avocado: Deep tensor factorization characterizes the human epigenome via imputation of tens of thousands of functional experiments.*

Jacob Schreiber, Timothy Durham, Jeffrey Bilmes and William Noble.

9:00 **(Research Highlight)** *iSOM-GSN: An Integrative Approach for Transforming Multi-omic Data into Gene Similarity Networks via Self-organizing Maps*

Luis Rueda and Nazia Fatima.

9:20 *Zero-shot imputations across species are enabled through joint modeling of human and mouse epigenomics.*

Jacob Schreiber, Deepthi Hegde and William Noble.

9:40 *Unsupervised manifold alignment for single-cell multi-omics data.*

Ritambhara Singh, Pinar Demetci, Giancarlo Bonora, Vijay Ramani, Choli Lee, He Fang, Zhijun Duan, Xinxian Deng, Jay Shendure, Christine Disteche and William S. Noble.

10:00 *Correlation Imputation in Single cell RNA-seq using Auxiliary Information and Ensemble Learning.*

Luqin Gan, Giuseppe Vinci and Genevera Allen.

10:12 Live Q&A

10:30-10:55 Break

10:55-12:00 **Plenary Session 2**

10:55 Introduction (Chair: Laxmi Parida)

Keynote Address: Madhav Marathe

11:00 **Real-time Computational Science for COVID-19 Pandemic Planning and Response**

12:00-1:00 Break

1:00-2:20 COVID-19 Special Speaker Session

Introduction (Chair: Srinivas Aluru)

Gregory Abowd, Using Digital Technologies to Support COVID-19 Response on Campus: A case study in the use of WiFi data

Joshua Weitz, Dynamics of COVID-19: Near- and Long-Term Challenges

2:30-4:20 Session 4A: COVID-19 II

(Chair: Nurit Haspel)

2:30 *Bio-JOIE: Joint Representation Learning of Biological Knowledge Bases.*

Junheng Hao, Chelsea J.-T. Ju, Muhao Chen, Yizhou Sun, Carlo Zaniolo and Wei Wang. *GPU-Accelerated Drug Discovery with Docking on the Summit Supercomputer:*

2:50 *Porting, Optimization, and Application to COVID-19 Research.*

Scott Le Grand, Aaron Scheinberg, Andreas Tillack, Mathialakan Thavappiragasam, Josh Vermaas, Rupesh Agarwal, Jeff Larkin, Duncan Poole, Diogo Santos-Martins, Leonardo Solis-Vasquez, Andreas Koch, Stefano Forli, Oscar Hernandez, Jeremy Smith and Ada Sedova.

3:10 *Cross-Global Attention Graph Kernel Network Prediction of Drug Prescription.*

Hao-Ren Yao, Der-Chen Chang, Ophir Frieder, Wendy Huang, I-Chia Liang and Chi-Feng Hung.

3:30 *GANDALF: Peptide Generation for Drug Design using Sequential and Structural Generative Adversarial Networks.*

Allison Rossetto and Wenjin Zhou.

3:50 *Predicting Criticality in COVID-19 Patients.*

Roger Hallman, Anjali Chikkula and Temiloluwa Prioleau.

4:02 Live Q&A

2:30-4:20 Session 4B: Mining & Scalable Tools

(Chair: Ananth Kalyanaraman)

2:30 *Abstract Mining.*

Ellie Small, Javier Cabrera and John Kostis.

2:50 *Collaborative Cloud Computing Framework for Health Data with Open Source Technologies.*

Fatemeh Rouzbeh, Ananth Grama, Paul Griffin and Mohammad Adibuzzaman.

3:10 *MeSH Indexing Using the Biomedical Citation Network.*

William Gasper, Dario Ghersi and Parvathi Chundi.

3:30 *A Dynamics-based Approach for the Target Control of Boolean Networks.*

Cui Su and Jun Pang.

3:32 Live Q&A

4:20-4:30 Break

4:30-5:30 Student Mentoring (Chairs: Anna Ritz, Dan DeBlasio)

Day 3 (Sep. 24; Thursday), Theme: Informatics

8:40-10:30 Session 5A: Genomics

(Chair: Mukul Bansal)

8:40 **(Research Highlight)** *DeCoDe: degenerate codon design for complete protein-coding DNA libraries.*

Tyler Shimko, Polly Fordyce and Yaron Orenstein.

9:00 *An Integer Linear Programming Solution for the Most Parsimonious Reconciliation Problem under the Duplication-Loss-Coalescence Model.*

Morgan Carothers, Joseph Gardi, Gianluca Gross, Tatsuki Kuze, Nuo Liu, Fiona Plunkett, Julia Qian and Yi-Chieh Wu.

9:20 *Linearization of Ancestral Genomes with Duplicated Genes.*

Pavel Avdeyev and Max Alekseyev.

9:40 *Re-balancing Variational Autoencoder Loss for Molecule Sequence Generation.*

Chaochao Yan, Sheng Wang, Jinyu Yang, Tingyang Xu and Junzhou Huang.

9:52 Live Q&A

8:40-10:30 Session 5B: Medical Informatics I

(Chair: Xiuzhen Huang)

8:40 *ELMV: an Ensemble-Learning Approach for Analyzing Electrical Health Records with Significant Missing Values.*

Lucas Jing Liu, Hongwei Zhang, Jianzhong Di and Jin Chen.

9:00 *A deep learning fusion model for brain disorder classification: Application to distinguishing schizophrenia and autism spectrum disorder.*

Yuhui Du, Bang Li, Yuliang Hou and Vince D Calhoun.

9:20 *Joint Grid Discretization for Biological Pattern Discovery.*

Jiandong Wang, Sajal Kumar and Mingzhou Song.

9:40 *Smart Computational Approaches with Advanced Feature Selection Algorithms for Optimizing the Classification of Mobility Data in Health Informatics.*

Elham Rastegari, Donovan Orn and Hesham Ali.

9:52 *CNN Based Segmentation of Infarcted Regions in Acute Cerebral Stroke Patients From Computed Tomography Perfusion Imaging.*

Luca Tomasetti, Kjersti Engan, Mahdieh Khanmohammadi and Kathinka Dæhli Kurz.

10:04 Live Q&A

10:30-10:55 Break

10:55-12:00 Plenary Session 3

10:55 Introduction (Chair: May Wang)

Keynote Address: Mihaela Pertea

11:00 **The Human Gene Catalogue: Are we there yet?**

12:00-12:30 Break

12:30-2:00 Featured Session on Diversity and Inclusion (Chair: May Wang)

Speakers:

Mgavi Elombe Brathwaite, New York University

2:00-2:15 Break

2:15-4:15 Session 6A: HTS Data II

(Chair: John Kececioglu)

2:15 *Staging Epileptogenesis with Deep Neural Networks.*

Diyuan Lu, Sebastian Bauer, Valentin Neubert, Lara Sophie Costard, Felix Rosenow and Jochen Triesch.

2:35 *Submodular sketches of single-cell RNA-seq measurements.*

Wei Yang, Jeff Bilmes and William Stafford Noble.

2:55 **(Research Highlight)** *Processing millions of single cells by SHARP.*

Shibiao Wan, Junil Kim, Yiping Fan and Kyoung Jae Won.

3:15 **(Research Highlight)** *Benchmarking of computational error-correction methods for next-generation sequencing data.*

Jaqueline J. Brito and Serghei Mangul.

3:35 *MinIsoClust: Isoform clustering using minhash and locality sensitive hashing.*

Sairam Behera, Jitender S. Deogun and Etsuko N. Moriyama.

3:47 Live Q&A

2:15-4:15 Session 6B: Medical Informatics II

(Chair: Yanshan Wang)

2:15 *Efficiently mining rich subgraphs from vertex-attributed graphs.*

Riyad Hakim and Saeed Salem.

2:35 *Multi-Site Assessment of Pediatric Bone Age Using Deep Learning.*

Aly Valliani, John Schwartz, Varun Arvind, Amir Taree, Jun Kim and Samuel Cho.

2:55 *Ir-Man: An Information Retrieval Framework for Marine Animal Necropsy Analysis.*

Alexander Carmichael, Deepayan Bhowmik, Johanna Baily, Andrew Brownlow, George Gunn and Aaron Reeves.

3:15 *A Novel Pupillometric-Based Application for the Automated Detection of ADHD Using Machine Learning.*

William Das and Shubh Khanna.

3:27 Live Q&A

4:15-4:30 Break

Recognitions and Awards Ceremony

4:30-5:00 (Chairs: May Wang, Srinivas Aluru, Ananth Kalyanaraman)

5:00 Closing Remarks

ACM-BCB 2020 Posters (Chairs: Chirag Jain, Jaroslaw Zola)

(all posters accessible via the virtual platform)

Filtering STARR-Seq Peaks for Enhancers with Sequence Models.

Ronald Nowling, Rafael Reple Geromel, Benjamin Halligan

Prediction of Large for Gestational Age Infants in Overweight and Obese women at approximately 20 Gestational Weeks.

Yuhan Du, John Mehegan, Fionnuala M McAuliffe, Catherine Mooney

PeakMatcher: Matching Peaks Across Genome Assemblies.

Ronald Nowling, Christopher Beal, Scott Emrich, Susanta Behura, Marc Halfon, Molly Duman-Scheel

Using patient information for the prediction of caregiver burden in Amyotrophic Lateral Sclerosis.

Anna Markella Antoniadis, Miriam Gavin, Mark Heverin, Orla Hardiman, Catherine Mooney

A data-centric pipeline using convolutional neural network to select better multiple sequence alignment method.

Mengmeng Kuang

Towards CNN Representations for Small Mass Spectrometry Data Classification.

Khawla Seddiki, Isabelle Fournier, Arnaud Droit, Frédéric Precioso

A Unified Cloud-Native Architecture For Heterogeneous Data Aggregation And Computation.

Fatemeh Rouzbeh, Ananth Grama, Paul Griffin, Mohammad Adibuzzaman

Representing Cellular Lines with SVM and Text Processing.

Ivan Carrera, Inês Dutra, Eduardo Tejera Puente

Comparing Type 2 Diabetes Self-Management Apps Against the Needs of Low-Income Minority Patients: Is There An Implicit Functionality Bias?

Wayne Zachary, Hita Gupta

Prefix/Suffix Variation in Retinoic Acid Response Elements.

Yuan Zhuang, Kara Cerveney, Anna Ritz

Graphery: a Biological Network Algorithm Tutorial Webservice.

Heyuan Zeng, Anna Ritz

A Protein-Protein Interactome for an African Cichlid.

Gabriel Preising, Joshua Faber-Hammond, Suzy Renn, Anna Ritz

Multi-omics Data Integration in the Cloud: Analysis of Statistically Significant Associations Between Clinical and Molecular Features in Breast Cancer.

Kawther Abdilleh, Boris Aguilar, Ross Thomson

Multi-omics characterization of Microtubule-actin cross linking factor 1 (MACF1) using the ISB-Cancer Genomics Cloud.

Alexis McClary, Kawther Abdilleh, Boris Aguilar, Dondra Bailey

Novel Generated Peptides for COVID-19 Targets.

Allison Rossetto, Wenjin Zhou

Refinement of G protein-coupled receptor structure models: Improving the prediction of loop conformations and the virtual ligand screening performances.

Bhumika Arora

Deep Neural Network Modeling for Phenotypic Prediction of Metagenomic Samples.

Yassin Mreyoud, Tae Hyuk Ahn

Learning the regulatory grammar of DNA for gene expression engineering.

Jan Zrimec

Beyond B-Cell Epitopes: Curating Positive Data on Antipeptide Paratope Binding to Support Development of Computational Tools for Vaccine Design and Other Translational Applications.

Salvador Eugenio Caouli

Multimodal Learning for Cardiovascular Risk Prediction using EHR Data.

Ayoub Bagheri, T. Katrien J. Groenhof, Wouter B. Veldhuis, Pim A. de Jong, Folkert Asselbergs, Daniel Oberski

RA2Vec: Distributed Representation of Protein Sequences with Reduced Alphabet Embeddings.

Rajitha Yasas Wijesekara, Ashwin Lahorkar, Kunal Rathore, Jayaraman Valadi

FINDER1: an automated gene finder from multiple sources of data.

Sagnik Banerjee, Priyanka Bhandary, Margaret Woodhouse, Roger Wise, Carson Andorf