# ACM-BCB 2015 Program

## REGISTRATION

Wednesday (9/9) 7:30 am–5:00 pm / Thursday (9/10) 8:00 am–11:00 am / Friday (9/11) 8:00 am–11:00 am

<table>
<tr><th colspan="8">Wednesday, September 9, 2015</th></tr>
<tr><th></th><th>Classroom 331</th><th>Auditorium 235</th><th>Classroom 323</th><th>Classroom 324</th><th>Classroom 233</th><th>Classroom 328</th><th>Classroom 330</th></tr>
<tr><td>8–9</td><td rowspan="5">CompPath (8 am–1 pm)</td><td rowspan="8">BigLS (8:25 am–12 pm) (1:30 pm–5 pm)</td><td></td><td rowspan="8">CNB-MAC (8:45 am–11:30 am) (1 pm–6 pm)</td><td rowspan="8">pSALSA(8am-6pm)</td><td rowspan="3">CHS (8:30 am–1 pm)</td><td></td></tr>
<tr><td>9–10</td><td rowspan="2">GMDM (9 am–12 pm)</td><td rowspan="4">ParBio (9 am–1:15 pm)</td></tr>
<tr><td>10–11</td></tr>
<tr><td>11–12</td><td></td></tr>
<tr><td>12–13</td><td rowspan="3">MMBD (1:30 pm–5 pm)</td></tr>
<tr><td>13–14</td><td></td><td></td></tr>
<tr><td>14–15</td><td></td><td></td><td></td></tr>
<tr><td>15–16</td><td></td><td></td><td></td></tr>
<tr><td>16–17</td><td></td><td></td><td></td></tr>
<tr><td>17–18</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

## WORKSHOPS

**CompPath:** The Computational Pathology Workshop: Linking Tissue Phenotypes with Genomics and Clinical Outcomes
Organizer: Lee Cooper and David Gutman

**GMDM:** Genomics and Metagenomics Data Mining Workshop
Organizer: Quoc-Nam Tran and Hairong Wei

**MMBD:** Novel enabling technologies in mining massive biomedical data
Organizer: Zhaohui S Qin and Tianwei Yu

**CNB-MAC:** 2nd Workshop on Computational Network Biology: Modeling, Analysis, and Control
Organizer: Byung-Jun Yoon, Xiaoning Qian, and Tamer Kahveci

**BigLS:** 3rd Workshop on: Big Data in Life Sciences
Organizer: Ananth Kalyanaraman and Jaroslaw Zola

**CHS:** 1st workshop on Computational Health Science
Organizer: Sherry Emery and Christophe Giraud-Carrier

**ParBio:** 4th workshop on Parallel and Cloud-based Bioinformatics and Biomedicine
Organizer: Mario Cannataro and John A. Springer

**pSALSA**: Parallel Software Libraries for Sequence Analysis
Organizer: Srinivas Aluru and Yongchao Liu

## TUTORIALS: Room 222

**T1:** Analysis methods for 3D nucleome data: from processing basics to three-dimensional modeling
Contact: Ferhat Ay (ferhatay@uw.edu); Geet Duggal (geetduggal@gmail.com); Ming Hu (Ming.Hu@nyumc.org); Emre Sefer (esefer@andrew.cmu.edu)

**T2:** Computational Construction of Intra-Cellular Networks
Contact: Tolga Can (tcan@ceng.metu.edu.tr)

**T3:** Functional Module Identification in Biological Networks
Contact: Xiaoning Qian (xqian@ece.tamu.edu)

**T4:** When imaging informatics meets bioinformatics: How multiple modality image-omics can integrate with genomics data analysis for personalized medicine
Contact: Lin Yang (lin.yang@bme.ufl.edu)

**T5:** Computational methods for analyzing intra-tumor heterogeneity using next-generation sequencing
Contact: Iman Hajirasouliha (imanh@stanford.edu)

| Thursday, September 10, 2015 | | | |
|---|---|---|---|
| 7:00–8:15 | **Continental Breakfast** | | |
| 8:15–8:30 | **Opening Remarks (Location: Room 236)**<br>General Chairs: Srinivas Aluru, Georgia Institute of Technology & May D. Wang, Georgia Institute of Technology and Emory University | | |
| 8:30–9:30 | **Keynote Talk 1 (Location: Room 236)**<br>**From Meaningful use to Precision Medicine: Lessons Learned**<br>**Douglas B. Fridsma, President and Chief Executive Officer of American Medical Informatics Association**<br>Session Chairs: Mark Braunstein and May D. Wang, Georgia Institute of Technology | | |
| 9:30-10 | **Morning Break** | | |
| | **Session 1A**<br>**Location: Room 222**<br>**Genome and Sequence Analysis I**<br>Session Chair: Carl Kingsford, Carnegie Mellon University | **Featured Sessions**<br>**Location: Rooms 334–335**<br>Session Chairs:<br>Greg Gibson & May D. Wang, Georgia Tech | **Session 1B**<br>**Location: Room 233**<br>**Data Privacy and Epidemiological Modeling**<br>Session Chair: Pierangelo Veltri, Univ. of Magna Graecia of Catanzaro |
| 10:00–12:00 | L: 10:00–10:24 am<br>Alexej Gossmann, Shaolong Cao and Yu-Ping Wang. Identification of significant genetic variants via SLOPE, and its extension to Group SLOPE.<br>L: 10:24–10:48 am<br>Aika Terada, Hanyoung Kim and Jun Sese. High-speed Westfall-Young permutation procedure for genome-wide association studies.<br>L: 10:48–11:12 am<br>Veronika Strnadova-Neeley, Aydin Buluc, Jarrod Chapman, John Gilbert, Joseph Gonzalez and Leonid Oliker. Efficient Data Reduction for Large-Scale Genetic Mapping.<br>L: 11:12–11:36 am<br>Segun Jung,Hongjian Jin and Ramana Davuluri. Identification of Candidate Regulatory SNPs by Integrative Analysis for Prostate Cancer Genome Data.<br>L: 11:36–12:00 am<br>Shaolong Cao, Huaizhen Qin, Alexej Gossmann, Hong-Wen Deng and Yu-Ping Wang. Unified tests for fine scale mapping and identifying sparse high-dimensional sequence associations | **NIH/NIGMS Trainee Forum: Computational Biology and Medical Informatics at Georgia Tech**<br><br>Professor Greg Gibson<br>Director of Integrative Genomics, Georgia Tech<br>and Four NIH Trainees | L: 10:00–10:24 am<br>Dingcheng Li, Majid Rastegar, Ravikumar Komandur Elayavilli, Yanshan Wang, Yue Yu, Saeed Mehrabi, Naveed Afzal, Sunghwan Sohn, Yanpeng Li and Hongfang Liu. A Frequency-filtering Strategy of Obtaining PHI-free Sentences from Clinical Data Repository.<br>L: 10:24–10:48 am<br>Feng Chen, Noman Mohammed, Shuang Wang, Wenbo He, Samuel Cheng and Xiaoqian Jiang. Cloud-Assisted Distributed Private Data Sharing.<br>L: 10:48–11:12 am<br>Carl Gunter, Muhammad Naveed, Jean-Pierre Hubaux, Erman Ayday, Ellen Clayton, Xiaofeng Wang, Brad Malin and Jacques Fellay. Privacy in the Genomic Era. **(H)**<br>L: 11:12–11:36 am<br>Suruchi Deodhar, Jiangzhuo Chen, Mandy Wilson, Keith Bisset, Bryan Lewis, Chris Barrett and Madhav Marathe. EpiCaster: An Integrated Web Application For Forecasting of Global Epidemics.<br>L: 11:36–12:00 am<br>Sherif Abdelhamid, Chris Kuhlman, Gizem Korkmaz, Madhav Marathe and S. S. Ravi. EDISON: A Web Application for Computational Health Informatics At Scale. |
| 12:00-12:30 | | **NIH/NIGMS Funding Opportunity in Computational Sciences**<br>**Dr. Veerasamy "Ravi" Ravichandran** | |

| | | | |
|---|---|---|---|
| 12:30-13:15 | **Lunch** | | |
| 13:15–14:15 | **Editors-In-Chief Plenary (Location: Room 236)**<br>**Dr. Ying Xu (TCBB EiC) and Briefing from J-BHI EiC** | | |
| | **Session 2A**<br>**Location: Room 222**<br>**Genome and Sequence Analysis II**<br>Session Chair: Ananth Kalyanaraman, Washington State University | **Featured Sessions** | **Session 2B**<br>**Location: Room 233**<br>**Machine Learning and Applications**<br>Session Chair: Umit Catayurek, Ohio State University |
| 14:15–15:30 | L: 2:15–2:40 pm<br>Qingming Tang, Sheng Wang, Jian Peng, Jianzhu Ma and Jinbo Xu. Bermuda: Bidirectional de novo assembly of transcripts with new insights for handling uneven coverage.<br>L: 2:40–3:05 pm<br>Darya Filippova and Carl Kingsford. Rapid Separable Compression Enables Fast Analyses of Sequence Alignments.<br>L: 3:05–3:30 pm<br>Dan Deblasio and John Kececiogu. Ensemble Multiple Sequence Alignment via Advising. | | L: 2:15–2:40 pm<br>Asish Ghoshal, Ananth Grama, Saurabh Bagchi and Somali Chaterji. An Ensemble SVM Model for the Accurate Prediction of Non-Canonical MicroRNA Targets.<br>L: 2:40–3:05 pm<br>Chanchala Kaddi and May D. Wang. Developing Robust Predictive Models for Head and Neck Cancer across Microarray and RNA-seq Data.<br>L: 3:05–3:30 pm<br>Dong Nie. A Deep Framework for Bacteria Image Segmentation and Classification. |
| 15:30–16:00 | **Afternoon Break** | | |
| 16:00–18:00 | **ACM SIGBio General Meeting**<br>**Location: Room 236** | | |
| 18:00–20:00 | **Poster Reception** | | |

| Friday, September 11, 2015 | | | |
|---|---|---|---|
| 7:00–8:30 | **Continental Breakfast** | | |
| 8:30–9:30 | **Keynote Talk 2 (Location: Room 236)**<br>**Computational biology in the 21st century: Algorithms that scale**<br>**Bonnie Berger, Professor of Applied Math and Computer Science at MIT, and head of the Computation and Biology group at MIT's Computer Science and AI Lab**<br>Session Chairs: T. M. Murali, Virginia Tech & Christopher C. Yang, Drexel University | | |
| 9:30-10:00 | **Morning Break** | | |
| | **Session 3A**<br>**Location: Room 222**<br>**Networks and Ontologies**<br>Session Chair: Predrag Radivojac, Indiana University | **Featured Sessions**<br>**Location: Room 225**<br>**Industry Track: Big Data Analytic Technology for Bioinformatics and Health Informatics**<br>Session Chairs: Xin Deng, Microsoft & Donghui Wu, Knowledgent | **Session 3B**<br>**Location: Room 233**<br>**Clinical and Health Decision Support Systems**<br>Session Chair: Saeid Belkasim, Georgia State University |
| 10:00–12:00 | L: 10:00–10:24 am<br>Shi Qiao, Mehmet Koyuturk and Z. Meral Özsoyoğlu. Integrated Querying of Disparate Association and Interaction Data in Biomedical Applications.<br>L: 10:24–10:48 am<br>Haitham Gabr, Alin Dobra and Tamer Kahveci. Estimating Reachability in Dense Biological Networks.<br>L: 10:48–11:12 am<br>Cen Wan and Alex A. Freitas. Two Methods for Constructing a Gene Ontology-based Feature Network for a Bayesian Network Classifier and Applications to Datasets of Aging-related Genes.<br>L: 11:12–11:36 am<br>Shufan Ji, Xing Tian and Jin Chen. Improving Biological Signicance of Gene Expression Biclusters with Key Missing Genes.<br>L: 11:36–12:00 am<br>Gaurav Pandey, Sonali Arora, Sahil Manocha and Sean Whalen. Enhancing the Functional Content of Eukaryotic Protein Interaction Networks. **(H)** | 10:00–10:20 am<br>Building and Implementing Data Science Products to Change Healthcare IT. Vincent A. Emanuele II, Ph.D, Wellcentive & David Lloyd, Wellcentive.<br>10:20–10:40 am<br>Senior Health Management through Internet of Things and Real-Time Big Data Analytics. Xin Deng, Ph.D., Microsoft & Donghui Wu, Ph.D., Knowledgent.<br>10:40–11:00 am<br>A Bayesian Nonparametric Approach for Latent Class Regression Analysis. Nong Shang, Ph.D., Division of Bacterial Diseases, NCIRD.<br>11:00–11:20 am<br>Learning Causal Structures with Background Knowledge in Health Data. Yiheng Liang, Ph.D., University of North Texas & Armin R. Mikler, Ph.D., University of North Texas.<br>11:20–11:40 am<br>Development of a mobile app to address a healthcare gap about laboratory testing. Julie Taylor, Ph.D., CDC.<br>11:40–12:00 pm: PANEL DISCUSSION<br>Big Data and Predictive Modeling Topics in Healthcare. Chairs: Xin Deng, Ph.D., Donghui Wu, Ph.D., | L: 10:00–10:24 am<br>Nikhil Yadav, Christian Poellabauer, Louis Daudet, Tomas Collins, Shane McQuillan, Patrick Flynn and Sandra Schneider. Portable Neurological Disease Assessment Using Temporal Analysis of Speech.<br>L: 10:24–10:48 am<br>Michael Uelschen and Heinz-Josef Eikerling. A Mobile Sensor System for Gait Analysis supporting the Assessment of Rehabilitation Measures.<br>L: 10:48–11:12 am<br>Xun Lu, Aston Zhang, Carl A. Gunter, Daniel Fabbri, David Liebovitz and Brad Malin. Discovering De Facto Diagnosis Specialties.<br>L: 11:12–11:36 am<br>Daniel Riofrio, Shuang Luan, Jun Zhou and Lijun Ma. Particle Swarm Optimization for Radiation Therapy Planning. |
| 12:00-13:30 | **Lunch** | **Women in Bioinformatics Panel**<br>**Location: Room 158**<br>Chair: Wei Wang,<br>Univ. of California Los Angeles | **Lunch** |

| | **Session 4A**<br>**Location: Room 222**<br>**Genome and Sequence Analysis III**<br>Session Chair: Vasant Honavar, Pennsylvania State University | **Featured Sessions**<br>**Location: Room 225**<br>**Special Panel: National Surveys of Population Health: A CDC Panel**<br>Session Chair: Bruce Schatz, PhD University of Illinois at Urbana-Champaign | **Session 4B**<br>**Location: Room 233**<br>**Proteins, Structures, and Models**<br>Session Chair: Mehmet Koyuturk, Case Western Reserve University |
|---|---|---|---|
| 13:30–15:30 | L: 1:30–1:54 pm<br>Dan He and Laxmi Parida. SAME: A Sampling-based Multi-locus Epistasis Algorithm for Quantitative Genetic Trait Prediction.<br>L: 1:54–2:18 pm<br>Marzieh Ayati and Mehmet Koyuturk. Assessing The Collective Disease Association of Multiple Genomic Loci.<br>L: 2:18–2:42 pm<br>Aseel Awdeh, Hilary Phenix, Mads Kaern and Theodore Perkins. The Potential Power of Dynamics in Epistasis Analysis.<br>L: 2:42–3:06 pm<br>Darya Filippova, Rob Patro, Geet Duggal and Carl Kingsford. Identification of alternative topological domains in chromatin. **(H)**<br>L: 3:06–3:30 pm<br>David Ream, Asma Riyaz and Iddo Friedberg. An event-driven approach for studying gene block evolution in bacteria. **(H)** | **Speakers**<br>**Machell Town, PhD**, Branch Chief, Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion, CDC<br>**Paula Yoon, ScD**, Director, Division of Health Informatics and Surveillance, Center for Surveillance, Epidemiology, and Laboratory Services, CDC<br>**Leslie Lenert, MD**, Chief Research Information Officer, Medical University of South Carolina; Chief Medical Information Officer, Health Sciences South Carolina; keynote speaker for Health Informatics BCB 2014; former Director of National Center for Public Health Informatics, CDC<br>**Bruce Schatz, PhD**, Head, Department of Medical Information Science, University of Illinois at Urbana-Champaign; program chair for Health Informatics at BCB 2014, and panel organizer | L: 1:30–1:54 pm<br>Chao Ji, Yong Li, Earl Bellinger, Sujun Li, Randy Arnold, Predrag Radivojac and Haixu Tang. A maximum-likelihood approach to absolute protein quantification in mass spectrometry.<br>L: 1:54–2:18 pm<br>Xuefeng Cui, Hiroyuki Kuwahara, Shuai Cheng Li and Xin Gao. Compare Local Pocket and Global Protein Structure Models by Small Structure Patterns.<br>L: 2:18–2:42 pm<br>Deukhyun Cha, Qin Zhang, Alexander Rand, Jesmin Jahan Tithi, Rezaul Chowdhury and Chandrajit Bajaj. Accelerated Molecular Mechanical and Solvation Energetics on Multicore CPUs and Manycore GPUs.<br>L: 2:42–3:06 pm<br>Minghan Chen, Fei Li, Kartik Subramanian, John Tyson and Yang Cao. Two-dimensional Model of PopZ Bipolarization in Caulobacter crescentus.<br>L: 3:06–3:30 pm<br>Jin Chen and Lei Xu. Plant photosynthesis phenomics data quality control. **(H)** |
| 15:30–16:00 | **Afternoon Break** | | |
| 16:00–18:00 | **NSF Sponsored Student Research Forum**<br>**Location: Room 225** | | |
| 18:30–20:30 | **Banquet** | | |

| Saturday, September 12, 2015 | | | |
|---|---|---|---|
| 7:00–8:30 | **Continental Breakfast** | | |
| 8:30–9:30 | **Keynote Talk 3 (Location: Room 236)**<br>**What is a gene? How and why computer science helps in answering this Question?**<br>**Paola Bonizzoni, Professor of Computer Science at the Università di Milano-Bicocca**<br>Session Chair: Helene Touzet, CNRS France and Srinivas Aluru, Georgia Tech | | |
| 9:30–10:00 | **Morning Break** | | |
| | **Session 5A**<br>**Location: Room 222**<br>**Genome and Sequence Analysis IV**<br>Session Chair: John Kececioglu, University of Arizona | **Featured Sessions**<br>**Location: Room 225**<br>**Student Education Session**<br>Chair: Marzieh Ayati Case Western Reserve Univ.<br>Co-Chairs: Hang Wu, Li Tong, Georgia Tech | **Session 5B**<br>**Location: Room 233**<br>**Predictive Modeling and Text Mining**<br>Session Chair: May D Wang, Georgia Tech and Emory U. |
| 10:00–12:00 | L: 10:00-10:24<br>Christopher Hoobin, Trey Kind, Christina Boucher and Simon Puglisi. Fast and Efficient Compression of High-Throughput Sequencing Reads.<br>L: 10:24-10:48<br>Kun Wang, Kan Cao and Sridhar Hannenhalli. Chromatin and Genomic determinants of alternative splicing.<br>L: 10:48-11:12<br>Cheng Yang, Po-Yen Wu, Li Tong, John Phan and May Wang. The impact of RNA-seq aligners on gene expression estimation.<br>L: 11:12-11:36<br>Zihe Chen, Danyang Chen, Xiangyu Wang, Andrew Fritz, Nitasha Sehgal, Ronald Berezney, Jinhui Xu and Hu Ding. Mining k-Median Chromosome Association Graphs from a Population of Heterogeneous Cells.<br>L: 11:36-12:00<br>Sharma V. Thankachan, Sriram Chockalingam, Yongchao Liu, Alberto Apostolico and Srinivas Aluru. The k-Mismatch Maximal Common Substring Problem: Efficient Algorithms and Applications. **(H)** | How to write a technical paper (by communication director at Georgia Tech)?<br><br>How to prepare for a career in academia?<br>Prof. Meral Ozsoyoglu, Case Western Reserve Univ.<br>Prof. Aidong Zhang, NSF/CISE/IIS and SUNY at Buffalo<br><br>How to land your first job in industry?<br>Dr. Xin Deng, Microsoft<br>Dr. Donghui Wu, Knowledgent | L: 10:00–10:24 am<br>Hui Li, Xiaoyi Li, Xiaowei Jia, Murali Ramanathan and Aidong Zhang. Bone Disease Prediction and Phenotype Discovery using Feature Representation over Electronic Health Records.<br>L: 10:24–10:48 am<br>Nhathai Phan, Dejing Dou, Hao Wang, David Kil and Brigitte Piniewski. Ontology-based Deep Learning for Human Behavior Prediction in Health Social Networks.<br>L: 10:48–11:12 am<br>Chih-Wen Cheng and May Wang. Improving Personalized Clinical Risk Prediction Based on Causality-Based Association Rules.<br>L: 11:12–11:36 am<br>Sanghoon Lee, Yanjun Zhao, Mohamed Eid Mahmoud Masoud, Maria Valero and Saeid Belkasim. Domain Specific Information Retrieval and Text Mining in Medical Document.<br>L: 11:36-12:00<br>Anthony Rios and Ramakanth Kavuluru. Convolutional Neural Networks for Biomedical Text Classification: Application in Indexing Biomedical Articles. |
| 12:00-13:30 | **Lunch** | | |
| | | | |

| | **Session 6A**<br>**Location: Room 222**<br>**Genome and Sequence Analysis V**<br>Session Chair: Filip Jagodzinski, Western Washington University | **Session 6B**<br>**Location: Room 233**<br>**Networks and Modules**<br>Session Chair: Christina Boucher, Colorado State University |
|---|---|---|
| 13:30–15:30 | L: 1:30–1:54 pm<br>Dan Gusfield. Persistent Phylogeny: A Galled-Tree and Integer Linear Programming Approach.<br>L: 1:54–2:18 pm<br>Elisabetta Bergamini, Romina D'Aurizio, Mauro Leoncini and Marco Pellegrini. CNVScan: detecting borderline copy number variations in NGS data via scan statistics.<br>L: 2:18–2:42 pm<br>Chen Peihua, Weiheng Huang and Hongmin Cai. Discrimination of recurrent CNVs from individual ones from multisample aCGH by jointly constrained minimization.<br>L: 2:42–3:06 pm<br>Subrata Saha and Sanguthevar Rajasekaran. HECTOR: A Novel Paradigm for Correcting Short Reads.<br>L: 3:06–3:30 pm<br>Hamid Mirebrahim, Timothy Close and Stefano Lonardi. De novo meta-assembly of ultra-deep and single-cell sequencing data. **(H)** | L: 1:30–1:54 pm<br>Md Abdul Alim, Ahmet Ay, Md Mahmudul Hasan, My Thai and Tamer Kahveci. Multiple Reference Networks Improve Accuracy of Signaling Network Construction.<br>L: 1:54–2:18 pm<br>Andrei Todor, Alin Dobra and Tamer Kahveci. Counting Motifs in Probabilistic Biological Networks.<br>L: 2:18–2:42 pm<br>Ayat Hatem, Kamer Kaya, Jeffrey Parvin, Kun Huang and Umit Catalyurek. MICA: MicroRNA Integration for Active Module Discovery.<br>L: 2:42–3:06 pm<br>Guopeng Wei, Connor Walsh, Irina Cazan and Radu Marculescu. Molecular Tweeting: Unveiling the Social Network behind Heterogeneous Bacteria Populations.<br>L: 3:06–3:30 pm<br>Adib Shafi, Michele Donato and Sorin Draghici. A systems biology approach for the identification of significantly perturbed genes. |