ACM BCB 2014

The 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics

Newport Beach, CA, U.S.A.
September 20 - 23, 2014
**ACM-BCB Organization Committee**

**Steering Committee Chair**  
Aidong Zhang  
State University of New York at Buffalo

**General Chairs**  
Pierre Baldi  
University of California, Irvine  
Wei Wang  
University of California, Los Angeles

**Program Chairs**  
Terry Gaasterland  
University of California, San Diego  
Ümit V. Çatalyürek  
The Ohio State University  
Bruce Schatz  
University of Illinois at Urbana-Champaign

**Workshop Chair**  
Tamer Kahveci  
University of Florida

**Tutorial Chairs**  
Vasant Honavar  
Penn State University  
Amarda Shehu  
George Mason University

**Poster Chairs**  
Dongxiao Zhu  
Wayne State University

**PhD Forum Chair**  
Jean Gao  
University of Texas at Arlington

**Translational Bioinformatics Panel Chairs**  
May D. Wang  
Georgia Institute of Technology  
Orly Alter  
University of Utah

**Women in Bioinformatics Panel Chair**  
May D. Wang  
Georgia Institute of Technology

**Exhibit/System Demo Chair**  
Richard Lathrop  
University of California, Irvine

**Registration Chair**  
Preetam Ghosh  
Virginia Commonwealth University

**Proceedings Chairs**  
Jing Gao  
State University of New York at Buffalo  
Xinghua Mindy Shi  
University of North Carolina

**Industry Chairs**  
Anastasia Christianson  
Bristol-Myers Squibb  
Bharat Rao  
KPMG LLP  
Bruce Schatz  
Poheme Inc.

**Publicity Chair**  
Jianlin Jack Cheng  
University of Missouri, Columbia

**Local Arrangement Chairs**  
Xiaohui Xie  
University of California, Irvine  
Janet Ko  
University of California, Irvine

**Student Travel Award Chair**  
Catie Welsh  
Rhodes College

**Web Masters**  
Kang Li  
State University of New York at Buffalo  
Xiaowei Jia  
State University of New York at Buffalo
# Conference Schedule

<table>
<thead>
<tr>
<th></th>
<th>Saturday Sep. 20</th>
<th>Sunday Sep. 21</th>
<th>Monday Sep. 22</th>
<th>Tuesday Sep. 23</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8:30am – 6:30pm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Workshop Sessions</strong> &amp; Tutorials</td>
<td>(1-7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lunch is not provided</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>8:30am – 6:30pm</strong></td>
<td>Noon – 1:00pm Lunch</td>
<td>Noon – 1:00pm Lunch &amp; Panel: Women in Bioinformatics</td>
<td>Noon – 1:00pm Lunch &amp; Panel: Women in Bioinformatics</td>
<td>Noon – 1:00pm Lunch &amp; PhD Forum</td>
</tr>
<tr>
<td><strong>9:00am – 10:00am</strong></td>
<td>9:00am – 10:00am Keynote Talk 1</td>
<td>9:00am – 10:00am Keynote Talk 2</td>
<td>9:00am – 10:00am Keynote Talk 3</td>
<td></td>
</tr>
<tr>
<td><strong>10:00am – 10:30am</strong></td>
<td>10:00am – 10:30am Break</td>
<td>10:00am – 10:30am Break</td>
<td>10:00am – 10:30am Break</td>
<td></td>
</tr>
<tr>
<td><strong>10:30am – 11:30am</strong></td>
<td>10:30am – 12:00pm Paper Session 2 (S3,S4)</td>
<td>10:30am – 12:00pm Industrial Session 2 &amp; Translational Bioinformatics Panel</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>11:30am – 12:00pm</strong></td>
<td>Noon – 1:00pm Lunch</td>
<td>Noon – 1:00pm Lunch &amp; Panel: Women in Bioinformatics</td>
<td>Noon – 1:00pm Lunch &amp; Panel: Women in Bioinformatics</td>
<td>Noon – 1:00pm Lunch &amp; PhD Forum</td>
</tr>
<tr>
<td><strong>2:30pm – 2:45pm</strong></td>
<td></td>
<td>2:30pm – 3:00pm Break</td>
<td>2:30pm – 3:00pm Break</td>
<td></td>
</tr>
<tr>
<td><strong>3:00pm – 4:00pm</strong></td>
<td></td>
<td>3:00pm – 4:30pm Paper Session 7 (S13, S14)</td>
<td>4:30pm – 4:45pm Closing</td>
<td></td>
</tr>
<tr>
<td><strong>4:00pm – 6:00pm</strong></td>
<td></td>
<td>4:00pm – 6:00pm ACM SIG-BIO Community Meeting</td>
<td>4:30pm – 6:00pm Paper Session 5 (S9, S10)</td>
<td></td>
</tr>
<tr>
<td><strong>6:00pm – 8:00pm</strong></td>
<td></td>
<td>6:00pm – 8:00pm Reception/Poster session/Demo/Exhibit</td>
<td>6:30pm – 8:30pm Banquet &amp; Best Paper/Poster Awards</td>
<td></td>
</tr>
</tbody>
</table>
Saturday, September 20, 2014

Workshops

Workshop 1

8:45am - 5:00pm
Location: Salon 3
Big Data in the Life Sciences (BigLS 2014)
Ananth Kalyanaraman, Washington State University; Jaroslaw Zola, SUNY Buffalo

8:45am – 10:30am Session I: Invited Talks

• 8:45-9:00 Opening remarks. Jaroslaw Zola and Ananth Kalyanaraman.
• 9:00-9:45 Feature subset selection for inferring relative importance of taxonomy. Gregory Ditzler and Gail Rosen.
• 9:45-10:30 Promises and challenges in analysis of biological big data. Jason McDermott.

10:30am – 11:00am Coffee Break

11:00am – 12:30pm Session II: Keynote Address

• Quantifying Your Superorganism Body Using Big Data Supercomputing. Larry Smarr, Founding Director of Calit2, Harry E. Gruber Professor in Computer Science and Engineering, University of California, San Diego, CA.

12:30pm – 2:00pm Lunch

2:00pm – 3:30pm Session III: Peer Reviewed Papers

• 2:00-2:30 Big Data challenges for estimating genome assembler quality. Abhishek Biswas, David Gauthier, Desh Ranjan, and Mohammad Zubair.
• 2:30-3:00 RNA-Seq gene and transcript expression analysis using the BioExtract Server and iPlant Collaborative. Etienne Gnimpieba, Chango Abalo, and Carol M. Lushbough.
• 3:00-3:30 AlignMR: Mass spectrometry peak alignment using Hadoop MapReduce. Urmi Bhayani, and John Springer.

3:30pm – 5:00pm Session IV: Student Posters and Student Mentoring Session

• 3:30-4:00 Student poster presentations
• 4:00-5:00 Mentoring interaction with faculty/researchers and students. Coffee/refreshments will be available between 4-4:30pm. List of poster presenters to be announced after student travel scholarships have been awarded. Please look for updates on the BigLS workshop webpage: http://www.bigls.org
Workshop 2

9:20am - 4:40pm
Location: Salon 5
Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC 2014)
Byung-Jun Yoon, Texas A&M University; Xiaoning Qian, Texas A&M University

9:20am – 9:30am Opening Remarks

9:30am – 10:30am Invited Keynote Talk by Dr. Fengzhu Sun
• Molecular Networks and the Internet: Are They Similar?

10:30am – 11:30am Highlight Talks
• Combinatorial Therapy Discovery using Mixed Integer Linear Programming. Zhandong Liu.
• Mathematical Design and Validation of Synergistic Anti-Cancer Drug Combinations. Ranadip Pal.

11:30am – 1:00pm Lunch Break

1:00pm – 2:00pm Invited Tutorial by Dr. Lori Dalton
• Optimal Bayes Classification

2:00pm – 2:20pm Five-Minute Lightning Talks for Posters
• Biological network clustering by robust NMF. Yijie Wang and Xiaoning Qian.
• A Novel Context-Sensitive Random Walk Model for Estimating Node Correspondence Between Two Biological Networks. Hyundoo Jeong and Byuong-Jun Yoon.
• Network-regularized Bi-clique Finding For Tumor Stratification. Amin Ahmadi Adl and Xiaoning Qian.
• Computational Identification of Functional Network Modules that are Associated with the Pathogenicity of Fusarium verticillioides. Mansuck Kim, Huan Zhang, Charles Woloshuk, Won-Bo Shim, and Byuong-Jun Yoon.

2:20pm – 3:30pm Poster Presentations & Coffee Break

3:30pm – 4:30pm Paper Presentations
• Simultaneous Identification of Robust Synergistic Subnetwork Markers for Effective Cancer Prognosis. Navadon Khunlertgit and Byung-Jun Yoon.

4:30pm – 4:40pm Closing Remarks
**Workshop 3**

8:30am - 12:30pm  
Location: Salon 3  
Computational Structural Bioinformatics (CSBW 2014)  
Jing He, Old Dominion; Amarda Shehu, George Mason University; Nurit Haspel, University of Massachusetts Boston; Brian Chen, Lehigh University

---

### 8:30am – 8:40am Opening Remarks

### 8:40am – 10:20am Session I

- **8:40-9:00** *Orientations of β-strand Traces and Near Maximum Twist.* Dong Si and Jing He.
- **9:00-9:20** *A Computational Model for Data Acquisition in SAXS.* Hui Dong and Gregory Chirikjian.
- **9:20-9:40** Highlight talk – Nurit Haspel, University of Massachusetts at Boston.
- **9:40-10:00** *An Analysis of Conformational Changes upon RNA-Protein Binding.* Kannan Sankar, Rasna R. Walia, Carla M. Mann, Robert L. Jernigan, Vasang G. Honavar, and Drena Dobbs.
- **10:00-10:20** Highlight Talk – Jianlin Cheng, University of Missouri.

### 10:20am – 11:00am Morning Break & Poster Session

### 11:00am – 12:20pm Session II

- **11:00-11:20** *Variational Bayesian Clustering on Protein Cavity Conformations for Detecting Influential Amino Acids.* Ziyi Guo and Brian Chen.
- **11:20-11:40** *Construction of Protein Backbone Pieces using Segment based FBCCD and Cryo-EM Skeleton.* Kamal Al Nasr and Jing He.
- **11:40-12:00** *Improving Decoy Databases for Protein Folding Algorithms.* Aaron Lindsey, Hsin-Yi Yeh, Chih-Peng Wu, Shawna Thomas, and Nancy Amato.
- **12:00-12:20** Highlight talk – Amarda Shehu, George Mason University

### 12:20pm – 12:30pm Concluding Remarks
8:30am – 9:30am Invited Keynote Talk by Dr. Christian Schoenbach

- Bridging the gaps between prediction of T-cell epitopes and what impacts immunity

9:30am – 10:30am Session I: Paper Presentations

- Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. Sinu Paul, John Sidney, Bjoern Peters, and Alessandro Sette.

10:30am – 11:00am Morning Break

11:00am – 12:30pm Session II: Paper & Poster Presentations

- Relapsing-Remitting Multiple Sclerosis and the role of Vitamin D: an agent based model. Francesco Pappalardo, Marzio Pennisi, Abdul Mateen Rajput, Ferdinando Chiacchio, and Santo Motto.

12:30pm – 2:00pm Lunch

2:00pm – 2:30pm Session III: Poster Presentations

- Towards the characterization of normal peripheral immune cells with data from ImmPort. Sandra Andorf, Jennifer B. Bollyky, Patrick Dunn, Jeffrey A. Wiser, Sanchita Bhattacharya, and Atul J. Butte.

2:30pm – 4:00pm Session IV: Tutorial: Bring your laptop for hands-on exercises!

- RImmPort: enabling ready-for-analysis immunology research data & Hands-on tutorial using RImmPort to access publicly-available individual-level immunology research data. Ravi Shankar, Sandra Andorf, Sanchita Bhattacharya, Jeffrey Wiser, and Atul Butte.
Workshop 5

4:30pm - 6:00pm
Location: Salon 1 & 2

Biomolecular Network Analysis (IWBNA 2014)
Young-Rae Cho, Baylor University; Pietro H. Guzzi and Pierangelo Veltri, University Magna Graecia of Catanzaro, Italy

4:30pm – 4:45pm
• Incremental Network Querying in Biological Networks. Md Mahmudul Hasan and Tamer Kahveci.

4:45pm – 5:00pm
• On the Impact of Data Integration and Edge Enrichment in Mining Significant Signals from Biological Networks. Sean West and Hesham Ali.

5:00pm – 5:15pm
• Evidence of post translational modification bias extracted from the tRNA and corresponding amino acid interplay across a set of diverse organisms. Oliver Bonham-Carter, Ishwor Thapa, and Dhundy Bastola.

5:15pm – 5:30pm
• Heuristic Parallelizable Algorithm for Similarity Based Biosystems Comparism. Mathialakan Thavappiragasam, Carol Lushbough, and Etienne Gnimpieba.

5:30pm – 5:45pm
• An R-based tool for miRNA data analysis and correlation with clinical ontologies. Francesca Cristiano and Pierangelo Veltri.

5:45pm – 6:00pm Closing Remarks. Pierangelo Veltri
Workshop 6
2:00pm - 4:00pm
Location: Salon 1&2
Parallel and Cloud-based Bioinformatics and Biomedicine (ParBio 2014)
Mario Cannataro, University Magna Graecia of Catanzaro, Italy; John A. Springer, Purdue University

2:00pm – 4:00pm
• 2:00-2:10 Opening Remarks. Mario Cannataro and John A. Springer.
• 2:35-3:00 Pheno2GRN: a workflow for phenotype to gene network studies and reverse engineering comparison. Nick Weinandt, Etienne Z. Gnimpieba, Laura Jackson, and Carol Lushbough.
• 3:00-3:25 A taxonomy for bioinformatics tools: exploiting semantics, parallelism, and services for analyzing omics data. Mario Cannataro, and Pietro Hiram Guzzi.
• 3:25-3:50 A system for ubiquitous distributed acquisition of voice alternation samples through a mobile application. Pietro Cinaglia, Manuela Macrì, Barbara Calabrese, Patrizia Vizza, Giuseppe Tradigo, and Pierangelo Veltri.
• 3:50-4:00 Closing Remarks. Mario Cannataro, and John A. Springer.

Workshop 7
2:00pm - 4:00pm
Location: Salon 4
Deep Learning in Bio- and Health Informatics (DL 2014)
Pierre Baldi, University of California, Irvine

2:00pm – 4:00pm
2:00-2:30 Deep Learning to Predict Chemical Properties and Reactions. Pierre Baldi.
2:30-3:00 Deep Learning for Protein Structure Prediction. Jianlin Cheng.
3:00-3:30 Applications of Deep Learning to Large-Scale Genomic Data. Xiaohui Xie.
3:30-4:00 A Multi-label Deep Belief Network for Predicting Multiple Diseases. Aidong Zhang and Hui Li.
Integrated Analysis of Next-gen sequencing data analysis using variant tools
Bo Peng, University of Texas; Suzanne Leal, Baylor College of Medicine

Abstract: Calling, annotating, filtering, and analyzing millions of genetic variants from whole-genome and whole-exome studies can be difficult due to the availability of a wide array of data formats, tools and annotation sources, as well as the sheer size of the data files. A big trunk of a bioinformatician’s time can be wasted on writing and maintaining scripts to convert data between different file formats, handle annotations from different sources, and connect inputs and outputs of various tools to create data processing pipelines. variant tools is a flexible data analysis toolset that provides a powerful command line interface to import and manipulate genetic variants and genotypes, to annotate variants using a large number of annotation databases, and to locate disease predisposing variants using more than 20 rare-variant association tests implemented in an association analysis framework called variant association tools (VAT). Using a sample whole exome sequencing project, this tutorial demonstrates how to perform quality control, annotation, variant selection, and association analysis for next-gen sequencing studies using variant tools and VAT.

Informatics approaches to Evidence-Based Medicine, with emphasis on Systematic Reviews
Aaron M. Cohen, Oregon Health & Science University; Neil R. Smalheiser, University of Illinois at Chicago

Abstract: National policy, insurance reimbursements, and funding agencies increasingly focus on the need for evidence-based recommendations to support clinical guidelines and therapies. Systematic reviews and meta-analyses play a central role in informing evidence-based medicine practice and policy. Standardized methods for comprehensive collection of relevant evidence, and careful filtering for quality are applied by experts to assess the overall state of knowledge about a medical question. These questions may address, for example, confidence in the relative and absolute efficacy and safety of treatments for a specific disease. This process requires a large investment of time and manual team effort. Dr. Cohen will focus in detail on the process of assembling evidence and writing and updating systematic reviews. He will identify steps which are particularly cumbersome and inefficient, which are good candidates for being assisted via automated informatics models and tools. He will discuss recent approaches to improving the systematic review process, ranging from machine learning tools to streamline effort,
Dr. Smalheiser will describe a pipeline of three machine learning-based tools designed to re-engineer the literature retrieval and triage steps: a metasearch engine for collecting relevant articles with high recall, a publication type tagger to identify randomized controlled trial articles with high accuracy, and a tool to identify distinct articles based on the same underlying trial.

**Tutorial 3**

2:00pm – 4:00pm
Location: Baycliff

**Robot Motion Planning Methods for Modeling Structures and Motions of Biomolecules**
Amarda Shehu, George Mason University; Nurit Haspel, University of Massachusetts Boston

**Abstract:** In the last two decades, great progress has been made in molecular modeling through robotics-inspired computational treatments of biological molecules. Deep mechanistic analogies between articulated robots and biomolecules have allowed robotics researchers to bring forth methods originally developed to address the robot motion planning problem in robotics to address and elucidate the relationship between macromolecular structure, dynamics, and function in computational structural biology. Tight coupling of approaches based on robot motion planning with computational physics and statistical mechanics have resulted in powerful methods capable of elucidating protein-ligand binding, order of secondary structure formation in protein folding, kinetic and thermodynamic properties of folding and equilibrium fluctuations in proteins and RNA, loop motions in proteins, small-scale and large-scale motions in multimodal proteins transitioning between different stable structures, and more. The objective of this tutorial is to introduce the broad community of researchers and students at ACM-BCB to robotics-inspired treatments and methodologies for modeling structures and motions in biomolecules. A comprehensive review of the current state of the art, ranging from the probabilistic roadmap approach to tree-based approaches, will be accompanied with specific detailed highlights and software demonstrations of powerful and recent representative robotics-inspired methods for peptides, proteins, and RNA.

**Tutorial 4**

4:30pm – 6:30pm
Location: Baycliff

**Network Approaches in Aging Research with Focus on Biological Network Alignment**
Tijana Milenkovic, University of Notre Dame; Fazle Elahi Faisa, University of Notre Dame

**Abstract:** Genes (proteins) interact with each other to keep us alive. And this is exactly what biological networks (BNs) model. Therefore, BN research is promising to revolutionize our biological understanding. Because susceptibility to diseases increases with age, studying human aging is important. But studying human aging experimentally is hard. Hence, aging-related knowledge needs to be transferred from model species. This transfer has traditionally been carried out by genomic sequence alignment. But because sequence data and BN data can give complementary insights, sequence alignment alone can limit the knowledge...
transfer. Thus, BN alignment can be used to transfer aging-related knowledge between topologically and functionally conserved network regions of different species. Gene expression research has also been indispensable for investigating aging, but it typically ignores genes' interconnectivities. Thus, analyzing genes' topologies in BNs could contribute to our understanding of aging. However, current methods for analyzing systems-level BNs deal with their static representations, although cells are dynamic. Because of this, and because different data can give complementary biological insights, current static BNs can be integrated with aging-related expression data to form dynamic, age-specific BNs. Then, cellular changes with age can be studied from such BNs. This tutorial will review state-of-the-art BN research of aging.

### Tutorial 5

**11:00am – 1:00pm**  
**Location: Baycliff**

**Computational Prediction of Protein-Protein Interfaces with Emphasis on Partner-Specific Protein-Protein Interactions**

Vasant G. Honavar, Pennsylvania State University; Li Xue, Pennsylvania State University

**Abstract:** Protein-protein interactions play a central role in formation of complexes and pathways that carry out virtually all major cellular processes. Both the distortion of protein interfaces in obligate complexes and aberrant recognition in transient complexes can lead to disease. Because of the difficulties and cost associated with experimental determination of protein complexes, there is an urgent need for reliable computational methods for predicting protein-protein interfaces from sequence and/or structure of a protein, and when available, its putative binding partner. Although most protein-protein interactions, in particular, transient interactions, are partner-specific, most existing protein interface predictors are not. Our group has recently shown that in partner-specific protein-protein interface prediction reveal that in the case of complexes resulting from transient protein-protein interactions, interfaces are highly conserved across homologous complexes and exploited this finding to design reliable partner-specific protein-protein interface predictors. We will briefly review the current state of the art in computational methods for protein-protein interface prediction. We will introduce both sequence homology based as well as machine learning based partner-specific protein-protein interface predictors. We also will also discuss the challenges of evaluating the performance of such predictors using commonly used metrics such as AUC and offer some alternatives.
Sunday, September 21, 2014

Opening Remarks
8:45am – 9:00am
Location: Salon 1-5
Session Chair: Pierre Baldi, University of California, Irvine

Keynote Talk 1
9:00am – 10:00am
Location: Salon 1-5
Session Chair: Pierre Baldi, University of California, Irvine

Fine-Grained Phenotypes, Comorbidities and Disease Trajectories from Data Mining of Electronic Patient Records
Søren Brunak, Technical University of Denmark & University of Copenhagen

Abstract: Electronic patient records remain a rather unexplored, but potentially rich data source for discovering correlations between diseases, drugs and genetic information in individual patients. Such data makes it possible to compute fine-grained disease co-occurrence statistics, and to link the comorbidities to the treatment history of the patients. A fundamental issue is to resolve whether specific adverse drug reaction stem from variation in the individual genome of a patient, from drug/environment cocktail effects, or both. Here it is essential to perform temporal analysis of the records for identification of ADRs directly from the free text narratives describing patient disease trajectories over time. We can then characterize the similarity of ADR profiles of approved drugs using drug-ADR networks and report on the relationship between the chemical similarity of drugs and their ADRs. Given the availability of longitudinal data covering long periods of time we can extend the temporal analysis to become more life-course oriented. We describe how the use of an unbiased, national registry covering 6.2 million people from Denmark can be used to construct disease trajectories which describe the relative risk of diseases following one another over time. We show how one can "condense" millions of trajectories into a smaller set which reflect the most frequent and most populated ones. This set of trajectories then represent a temporal diseaseome as opposed to a static one computed from non-directional comorbidities.

10:00am – 10:30am Morning Break
Abstract: The ever-increasing output of genome sequencing instruments has recently resulted in a sub $1,000 genome inclusive of the informatics to map the reads and call variants. As researchers switch their focus from single genome analysis to cohort-level genomic analysis, and with more comprehensive genomic characterization of each subject (eg, DNA, transcriptome, epigenome, microbiome, etc), the informatics challenges are moving into the realm of Big Data analytics. This presentation will put these trends into context and will provide a perspective on how these challenges are being addressed and where future improvements are needed.
**Paper Session I**

1:00pm – 3:30pm

### Session S1:
**Common Methods and Systems**
**Location:** Salon 4 & 5
**Session Chair:** Amarda Shehu, George Mason University

1. **L: 1:00-1:30pm**
   *Graph-Theoretic Analysis of Epileptic Seizures on Scalp EEG Recordings.* Nimit Dhulekar, Basak Oztan, Bülent Yener, Haluk O. Bingol, Gulcin Irim, Berrin Aktekin, and Canan Aykut Bingol.

2. **L: 1:30-2:00pm**
   *A Hadoop-Galaxy adapter for user-friendly and scalable data-intensive bioinformatics in Galaxy.* Luca Pireddu, Simone Leo, Nicola Soranzo, and Gianluigi Zanetti.

3. **S: 2:00-2:15pm**
   *Deep Autoencoder Neural Networks for Gene Ontology Annotation Predictions.* Davide Chicco, Peter Sadowski, and Pierre Baldi.

4. **S: 2:15-2:30pm**
   *An Improved Algorithm for the Sorting by Reversals and Transpositions Problem.* Ulisses Dias, Andre Rodrigues Oliveira, and Zanoni Dias.

5. **S: 2:30-2:45pm**
   *Data Mining to Aid Beam Angle Selection for Intensity-Modulated Radiation Therapy.* Stuart Price, Bruce Golden, Edward Wasil, and Hao Zhang.

6. **S: 2:45-3:00pm**
   *omniClassifier: a Desktop Grid Computing System for Big Data AccurMsd: A Machine Learning Approach to Predicting...*

### Session S2:
**Bioinformatics**
**Location:** Salon 1-3
**Session Chair:** Leonard McMillan, University of North Carolina, Chapel Hill

1. **L: 1:00-1:30pm**
   *IPED2: Inheritance Path based Pedigree Reconstruction Algorithm for Complicated Pedigrees.* Dan He, Zhanyong Want, Laxmi Parida, and Eleazar Eskin.

2. **L: 1:30-2:00pm**

3. **S: 2:00-2:15pm**
   *Mining massive SNP data for identifying associated SNPs and uncovering gene relationships.* Amy Webb, Aaron Albin, Zhan Ye, Majid Rastegar-Mojarad, Kun Huang, Jeffrey Parvin, Wolfgang Sadee, Lang Li, Simon Lin, and Yang Xiang.

4. **S: 2:15-2:30pm**

5. **S: 2:30-2:45pm**
   *Focus: A New Multilayer Graph Model for Short Read Analysis and Extraction of Biologically Relevant Features.* Julia Warnke, and Hesham Ali.

6. **S: 2:45-3:00pm**
   *AccuRmsd: A Machine Learning Approach to Predicting...*
**Prediction Modeling**

7. **S: 3:00-3:15pm**

8. **S: 3:15-3:30pm**
Dynamic Coordinate Registration Method for Image-Guided Surgery. Xi Wen, Hong Wang, and Weiming Zhai.

**Structure Similarity of Docked Protein Complexes.** Bahar Akbal-Delibas, Marc Pomplun, and Nurit Haspel.

7. **S: 3:00-3:15pm**
Large Highly Connected Clusters in Protein-Protein Interaction Networks. Suzanne Gallagher and Debra Goldberg.

8. **S: 3:15-3:35pm**

---

**3:30pm – 4:00pm Afternoon Break**

---

**ACM SIG-BIO Community Meeting**

4:00pm – 6:00pm
Location: Salon 1-5
Session Chair: Aidong Zhang, State University of New York at Buffalo

---

**Reception/Poster Session/ Demo/Exhibit**

6:00pm – 8:00pm
Location: Cardiff, Del Mar, Laguna, and Sunset
Session Chairs: Yu-Ping Wang, Tulane University; Xiaohui Xie, State University of California, Irvine; Dongxiao Zhu, Wayne State University

**DEMO 1:**
A Schema-Matching Tool for Alzheimer’s Disease Data Integration.
Peeho Dewan, Naveen Ashish, and Arthur Toga,
University of Southern California, Los Angeles, CA, USA
DEMO 2:
The DOE Systems Biology Knowledgebase (KBase): Progress Towards a System for Collaborative and Reproducible Inference and Modeling of Biological Function.
Robert Cottingham, Oak Ridge National Laboratory, Oak Ridge, TN, USA
URL: http://kbase/us/

List of accepted posters are at the end of this brochure
Abstract: Uncovering and interpreting genotype/phenotype relationships are among the most challenging open questions in disease studies. In cancer, uncovering these relationships is complicated even further due to the heterogeneous nature of the disease. Over the years, we have developed several algorithms that help to analyze heterogeneous cancer data in the context of uncovering genotype-phenotype relations, identification of dysregulated pathways, and cancer classification. These approaches span a large spectrum of algorithmic techniques including optimization-based techniques and mixture models. Taken together, these approaches help to leverage datasets collected through TCGA and other initiatives for better understanding of cancer and cancer diversity.

1. 10:30-11:00am

1. 10:30-11:00am
CNVnet: Combining Sparse learning and Biological Networks to Capture Joint Effect of Copy Number Variants. Zhiyong Wang, Jinbo Xu, and Xinghua Shi.
2. **L: 11:00-11:30am**  

3. **S: 11:30-11:45am**  

4. **S: 11:45am-12:00pm**  

---

**Noon – 1:00pm Lunch**  
**Location: Rose Garden**

---

**Panel: Women in Bioinformatics**

**12:00 – 1:00pm**  
**Location: Salon 1-3**  
Panel Chair: May D. Wang, Georgia Institute of Technology

**WiB Panel Discussion Topics List**

- As a woman working in biomedical informatics, what would be the top three factors that you consider to be important for a productive career?
- Do you have a mentor in your career? Or do you want to serve as a mentor for young women in biomedical informatics?
- What are some challenged women are facing in biomedical informatics?
- What are the essential skills students or post-docs should acquire if they want to pursue research in biomedical informatics?
- What are the emerging areas in biomedical informatics?
- What are the pros and cons of working in academia versus industry?
- Are there different skill set required for industry and academic positions?
- How to balance work and family as a professional women working in biomedical informatics?
- How can women play more leadership in biomedical informatics?
Paper Session III

1:00pm – 2:30pm

Session S5:
Bioinformatics

Location: Salon 4 & 5
Session Chair: Stefano Lonardi, University of California, Riverside

1. L: 1:00-1:30pm
   Unconstrained Gene Tree Diameters for Deep Coalescence.
   Pawel Gorecki, Jaroslaw Paszek, and Oliver Eulenstein.

2. L: 1:30-2:00pm
   Extracting phylogenetic signals from gene trees and its significance for species tree construction.
   Rasha Elhesha, Tamer Kahveci, Gordon Burleigh, and Ryan Cobb.

3. S: 2:00-2:15pm
   Strand: Fast Sequence Comparison using MapReduce and Locality Sensitive Hashing.
   Jake Drew, and Michael Hahsler.

4. S: 2:15-2:30pm
   A Noise- Aware Method for Building Radiation Hybrid Maps.
   Raed Seetan, Anne Denton, Omar Al-Azzam, Ajay Kumar, M. Javed Iqba,l, and Shahryar Kianian.

Session S6:
Computational and Translational Bioinformatics

Location: Salon 1-3
Session Chair: T.M. Murali, Virginia Tech

1. L: 1:00-1:30pm
   Improving indentification of key players in aging via network de-noising.
   Boyoung Yoo, Huili Chen, Fazle Faisal, and Tijana Milenkovic.

2. L: 1:30-2:00pm
   Integrated miRNA and mRNA Analysis of Time Series Microarray Data.
   Julian Dymacek, and Nancy L. Guo.

3. S: 2:00-2:15pm
   Community Detection-based Features for Sequence Classification.
   Karthik Tangirala, and Doina Caragea.

4. S: 2:15-2:30pm
   One Feature Doesn’t Fit All: Characterizing Topological Features of Targets in Signaling Networks.
   Huey-Eng Chua, Sourav S Bhowmick, and Lisa Tucker-Kellogg.

2:30pm – 2:45pm Afternoon Break
# Paper Session IV

2:45pm – 4:15pm

<table>
<thead>
<tr>
<th>Session S7: Bioinformatics</th>
<th>Session S8: Health Informatics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location: Salon 4 &amp; 5</td>
<td>Location: Salon 1-3</td>
</tr>
<tr>
<td>Session Chair: Filip Jagodzinski, Central Washington University</td>
<td>Session Chair: Nurit Haspel, University of Massachusetts Boston</td>
</tr>
</tbody>
</table>

1. **L: 2:45-3:15pm**  
   *Fast Dendrogram-based OUT Clustering using Sequence Embedding.* Thuy Diem Nguyen, Bertil Schmidt, and Chee Keong Kwoh.

2. **L: 3:15-3:45pm**  
   *Learning Parameter Advisors for Multiple Sequence Alignment.* Dan Deblasio and John Kececioglu.

3. **S: 3:45-4:00pm**  
   *COuplet Supertree by Equivalence Partitioning of taxa set and DAG formation.* Sourya Bhattacharyya and Jayanta Mukhopadhyay.

4. **S: 4:00-4:15pm**  

1. **L: 2:45-3:15pm**  
   *A Novel Classification Method for Predicting Acute Hypotensive Episodes in Critical Care.* Sakyajit Bhattacharya, Vaibhav Rajan, and Vijay Huddar.

2. **L: 3:15-3:45pm**  

3. **S: 3:45-4:00pm**  
   *MotionTalk: Personalized home rehabilitation system for assisting patients with impaired mobility.* Janani Venugopalan, Chih-Wen Cheng, and May Wang.

4. **S: 4:00-4:15pm**  
   *SideEffectPTM: An Unsupervised Topic Model to Mine Adverse Drug Reactions from Health Forums.* Sheng Wang, Yanen Li, Duncan Ferguson, and Chengxiang Zhai.

4:15pm – 4:30pm Afternoon Break
### Paper Session V

**Session S9:** Bioinformatics  
**Location:** Salon 4 & 5  
**Session Chair:** Brian Chen, Lehigh University

1. **L:** 4:30-5:00pm  
   *Scaled Sparse High Dimensional Tests for Localizing Sequence Variants.* Shaolong Cao, Huaizhen Qin, Jian Li, Hong-Wen Deng, and Yu-Ping Wang.

2. **L:** 5:00-5:30pm  
   *Prioritization of Genomic Locus Pairs for Testing Epistasis.* Marzieh Ayati and Mehmet Koyuturk.

3. **S:** 5:30-5:45pm  
   *A Multiscale Hybrid Evolutionary Algorithm to Obtain Sample-based Representations of Multi-basin Protein Energy Landscapes.* Rudy Clausen and Amarda Shehu.

4. **S:** 5:45-6:00pm  

### Session S10: Health Informatics  
**Location:** Salon 1-3  
**Session Chair:** Hongfang Liu, Mayo Clinic

1. **L:** 4:30-5:00pm  

2. **L:** 5:00-5:30pm  
   *Joint Inference for End-to-End Coreference Resolution for Clinical Notes.* Prateek Jindal, Dan Roth, and Carl A. Gunter.

3. **L:** 5:30-6:00pm  
   *Understanding User Intents in Online Health Forums.* Thomas Zhang, Hyun Duk Cho, and Chengxiang Zhai.

### Banquet & Best Paper/Poster Awards

**6:30pm – 8:30pm**  
**Location:** Rose Garden  
**Session Chair:** Wei Wang, University of California, Los Angeles
Tuesday, September 23, 2014

Keynote Talk 3

9:00am – 10:00am

Location: Salon 1-5

Session Chair: Bruce Schatz, University of Illinois at Urbana-Champaign

The Fractal-Like Architecture of the Learning Health System

Leslie Lenert, Medical University of South Carolina

Abstract: A recent report from the Institute of Medicine entitled "The Best Care at Lower Costs" maps out the escape path for the United States from the paradox of its high-cost low-value health system. That path runs through the transformation of the nation's healthcare business sector into a Learning Health System. At MUSC and HSSC, Dr. Lenert is working to create a Learning Health System that spans individual practices, the research laboratories, clinics, and universities of the state of South Carolina. This talk will focus on the enterprise architecture of that system and how its structure is like a fractal, with parallels at each level of organization that make organizational learning, as well as scientific insights possible from the data generated by the routine care of patients. Creating a Learning Health System for an academic health center requires creating linkages between genomic and proteomic databases, tissue repositories, electronic medical records systems and data warehouses, and mobile and web based tools for capture of personal sensor data and patient reported outcomes. In South Carolina, linkages across systems are facilitated by a statewide research master person index system that is available as a web service. Another critical component of South Carolina's architecture is the HSSC Clinical Data Warehouse, which has transaction level data on patients on 3.2 million patients from the four largest health systems in the state and makes these data available in pseudo-anonymized form for research through an i2b2 database. Work with this database is supported by statewide human subjects permissions system that facilitates Institutional Review Board (IRB) reliance. A unique governance mechanism for access to statewide data protects the interests of organizational data contributors and those of patients, creating a collaborative environment for learning.

10:00am – 10:30am Morning Break
Wearable Sensors: Moving from the Quantified Self to the Understood Self
Steven Steinhubl, Director of Digital Health, Scripps Healthcare
&
Rapid Learning Using Privacy-Preserving Distributed Data-Mining
Balaii Krishnapuram. Director of Health Analytics, Siemens Healthcare

Abstract: Through progressively miniaturized and increasingly powerful mobile computing capabilities, individuals now have the capability to monitor, track and transmit important health metrics continuously and in real time. Taking advantage of this, a wide spectrum of novel technologies has been developed to allow for personalized wellness, acute disease diagnostics, and chronic condition management from home that would otherwise have required an office or hospital visit. Wrist-worn sensor technologies currently available or in late-stage development may best exemplify both the capabilities and analytic challenges these remarkable advances are bringing to the healthcare setting. Over a dozen important health and wellness parameters are capable of being monitored continuously with a watch-like device during routinely daily activities; parameters currently only available in an Intensive Care Unit setting such as beat-to-beat blood pressure, cardiac output, ECG, oxygen saturation, and more. Beyond being just a more convenient way for vital signs to be monitored, these multiple, continuous data streams offer tremendous opportunities to understand an individual’s unique and personalized physiologic responses to daily stressors, and most importantly, help guide healthy responses to them. Transforming these numerous, vast and inter-related data streams into understandable and actionable information for the individual and their healthcare team is a critical requirement for mobile sensor technology to achieve its potential to improve the health and wellness of all of us.

Abstract: Recent advances in technology and data acquisitions costs are leading to an explosion of electronic data for bioinformatics and clinical research. At the same time we are witnessing a new paradigm for clinical research based on secondary use of data from Electronic Medical Records (EMR). In this talk, we describe a Health IT system that supports Rapid Learning across hospitals in Germany, Belgium and Netherlands. We describe that technological developments that enable us to conduct clinical research by bringing the computation to the data in federated databases in each hospital. The system avoids centralizing the data and preserves patient privacy, thus overcoming ethical, political and legal challenges to sharing patient data. Further, it uses ontologies and machine learning methods to dramatically reduce administrative and IT costs for collecting, normalizing and exchanging information across disparate source systems that use different languages, clinical protocols, database schema. We demonstrate the impact of the system on translational clinical research based on a case study across 5 hospitals in 4 countries.
Panel: Translational Bioinformatics

10:30am – 12:00pm
Location: Salon 1-3
Panel Chairs: Orly Alter, University of Utah; May Wang, Georgia Institute of Technology

10:30am – 10:45am Discovery of Principles of Nature from Matrix and Tensor Modeling of Large-Scale Molecular Biological Data
Orly Alter. USTAR Associate Professor of Bioengineering and Human Genetics Scientific Computing and Imaging (SCI) Institute, University of Utah. http://www.alterlab.org/
Abstract: I will briefly describe the use of matrix and tensor decompositions in the simultaneous modeling of different types of large-scale molecular biological data, from different studies of cell division and cancer and from different organisms, to computationally predict previously unknown physical, cellular, and evolutionary mechanisms that govern the activity of DNA and RNA. I will briefly present novel multi-matrix and multi-tensor generalizations of the singular value decomposition as well as experimental verification and validation of some of the computational predictions. Last, I will briefly note on a laboratory test based on one discovery, which is on the verge of being implemented in a clinical setting. These models bring physicians a step closer to one day being able to predict and control the progression of cell division and cancer as readily as NASA engineers plot the trajectories of spacecraft today.

10:45am – 11:00am Integrating Data from Discovery Research, Preclinical Studies and Clinical Trials
Abstract: Biological discovery is only the first step in the process of drug development. Even if a scientific idea has been well-developed in the lab, integrating data from in vitro, in vivo and clinical sources into a coherent picture remains a major challenge of translational bioinformatics. I will describe our efforts to develop biomarkers for an oncology indication, discuss the difficulties facing the field, and present future possible directions for maximizing the value of clinical trial data.

11:00am – 11:15am Rank-Rank Hypergeometric Overlap (RRHO) Gene Expression Signature Comparison and the Benefits of Cross-Species Analysis
Thomas G. Graeber. Associate Professor of Molecular and Medical Pharmacology Crump Institute for Molecular Imaging and UCLA Metabolomics Center, UCLA. http://systems.crump.ucla.edu
Abstract: The Rank-Rank Hypergeometric Overlap (RRHO) bioinformatic algorithm was developed for gene expression signature comparison in cases where the similarity is relatively weak but of statistical significance. The RRHO approach provides a statistical measure of overlap and a graphical map of the pattern of correlation between expression profiles. Previous techniques involve choosing a proper fixed threshold of differential expression. RRHO uses the full list of genes ranked by their degree of differential expression. RRHO is a two-dimensional analog of the Gene Set Enrichment Analysis (GSEA). Translational bioinformatic uses of RRHO in pharmaceutical research include drug response comparisons to guide drug development, molecular signature-based
validation of a mouse model of a human disease, comparisons of neuron developmental stages, and prioritization of leads from genome-wide association studies (GWAS). RRHO is available at http://systems.crump.ucla.edu/rankrank

11:15am – 11:30am Comprehensive RNA-Seq Data Analysis Pipeline Investigation for Translational Genomics
May D. Wang, Associate Professor, Kavli Fellow, and Georgia Research Alliance Distinguished Cancer Scholar
Coulter Department of Biomedical Engineering, Department of Electrical and Computer Engineering, and the Winship Cancer Institute, Georgia Institute of Technology and Emory University http://www.bio-miblab.org

Abstract: As RNA-seq technology becomes available for translational genomics, finding the proper data analysis pipelines remains a critical challenge. At the FDA’s Sequencing Quality Control (SEQC) Consortium, we investigated 278 RNA-seq data analysis pipelines to determine their impact on gene expression accuracy and precision, sensitivity in detecting low-expression genes, specificity in detecting differentially expressed genes, and downstream prediction performance. We found that the quality of gene expression and the statistical power of downstream analysis were significantly impacted by the interaction among multiple pipeline components. We established a guideline for selecting RNA-seq data analysis pipelines for improved reproducibility, and effective decision making.

11:30am – 12:00pm Panel Discussion

Noon – 1:00pm Lunch
Location: Seaview Terrace

PhD Forum
Noon – 1:00pm
Location: Cardiff, Del Mar, Laguna, & Sunset Room
Session Chair: Jean Gao, University of Texas at Arlington

The Ph.D. Forum at the ACM-BCB conference provides a dynamic academic and career information-exchange platform for students and young researchers. Research results will be presented in a variety of formats including poster presentations, real-time computer demonstrations, and face-to-face discussions. We expect the forum to help stimulate enthusiasm for new research directions, trigger new research ideas, and share experiences on research and other topics, such as preparation for interdisciplinary jobs and career paths in academia and industry.
### Paper Session VI

**1:00pm – 2:30pm**

**Session S11:**
Health Informatics  
**Location:** Salon 4 & 5  
**Session Chair:** ChengXiang Zhai, University of Illinois

1. **L: 1:00-1:30pm**  
*Towards a Natural Walking Monitor for Pulmonary Patients using Simple Smart Phones.* Joshua Juen, Qian Cheng, and Bruce Schatz.

2. **S: 1:30-1:45pm**  
*Using Mobile Phones to simulate Pulse Oximeters: Gait Analysis predicts Oxygen Saturation.* Qian Cheng, Joshua Juen, and Bruce Schatz.

3. **S: 1:45-2:00pm**  

4. **S: 2:00-2:15pm**  

**Session S12:**
Computational and Translational Bioinformatics  
**Location:** Salon 1-3  
**Session Chair:** Xiaohui Xie, University of California, Irvine

1. **L: 1:00-1:30pm**  
*Docking Features for Predicting Binding Loss due to Protein Mutation.* Norman Goodacre, Dr. Nathan Edwards, Dr. Mark Danielsen, Dr. Peter Uetz, and Dr. Cathy Wu.

2. **L: 1:30-2:00pm**  
*Leveraging Hierarchy in Medical Codes for Predictive Modeling.* Anima Singh, Girish Nadkarni, John Guttag, and Erwin Bottinger.

3. **S: 2:00-2:15pm**  
*A Flexible Volumetric Comparison of Protein Cavities can Reveal Patterns in Ligand Binding Specificity.* Ziyi Guo, Trevor Kuhlengel, Steven Stinson, Seth Blumenthal, Soutir Bandyopadhyay, and Brian Chen.

4. **S: 2:15-2:30pm**  
*Knowledge-based Search and Multi-objective Filters: Proposed Structural Models of GPCR Dimerization.* Irina Hashmi, Daniel Veltri, Nadine Kabbani, and Amarda Shehu.

**2:30pm – 3:00pm Afternoon Break**
Paper Session VII

Session S13:
Health Informatics
Location: Salon 4 & 5
Session Chair: Bruce Schatz,
University of Illinois at Urbana-Champaign

1. L: 3:00-3:30pm
   Resolving Healthcare Forum Posts via Similar Thread Retrieval.
   Hyun Duk Cho, Parikshit Sondhi, Chengxiang Zhai, and Bruce Schatz.

2. L: 3:30-4:00pm
   Automatic Risk of Bias Assessment for Clinical Trials.
   Iain Marshall, Joel Kuiper, and Byron Wallace.

3. S: 4:00-4:15pm
   Modeling climate-dependent population dynamics of mosquito to guide public health policies.
   Aditya Vaidya, Angel Bravo-Salgado, and Armin Mikler.

4. S: 4:15-4:30pm
   A Method for Reducing the Severity of Epidemics by Allocating Vaccines According to Centrality.
   Krzysztof Drewniak, Joseph Helsing, and Armin Mikler.

Session S14:
Computational and Translational Bioinformatics
Location: Salon 1-3
Session Chair: Pierangelo Veltri,
Laboratory of Bioinformatics, University of Catanzaro

1. L:3:00-3:30pm
   Analysing the distribution of synaptic vesicles using a spatial point process model.
   Mahdieh Khanmohammadi, Rasmus Waagepetersen, Nicoletta Nava, Jens Nyengaard, and Jon Sporring.

2. L: 3:30-4:00pm
   Automated Ranking of Stem Cell Colonies By Translating Biological Rules to Computational Models.
   Adele Peskin, Steve Lund, Ya-Shian Li-Baboud, Michael Halter, Anne Plant, and Peter Bajcsy.

3. S: 4:00-4:15pm
   Utilizing twilight zone sequence similarities to increase the accuracy of protein 3D structure comparison.
   Aleksandar Poleksic, and Paul Gray.

4. S: 4:15-4:30pm
   Challenges in Adapting Text Mining for Full text Articles to Assist Pathway Curation.
   Ravikumar Komandur Elayavilli, Kavishwar Wagholikar, and Hongfang Liu.
1. Dynamic networks reveal key players in aging. Fazle Faisal and Tijana Milenkovic.

2. Global network alignment in the context of aging. Fazle Faisal, Han Zhao and Tijana Milenkovic.


4. DTC genetic testing and consumer comprehension. Scott McGrath and Dhundy Kiran Bastola.


10. The UniFrac Significance Test Generates Different Outputs Given Semantically Equivalent Inputs. Jeffrey Long, Qingxiang Yan, Brett Trost and Anthony Kusalik.


14. Validation and implementation of whole-exome sequencing bioinformatics processes for clinical applications. Rimma Shakhshtyan, Himanshu Sharma, Ellen Tsai, Mark Bowser, Birgit Funke and Matthew Lebo.


17. CLARK, Accurate and Efficient Classification of DNA Sequences. Rachid Ounit, Stefano Lonardi, Timothy Close and Steve Wananaker.


22. Identifying causal variants at loci with multiple signals of association. Farhad Hormozdiari, Emrah Kostem, Eun Yong Kang, Bogdan Pasaniuc and Eleazar Eskin.


30. An Automated Pipeline for Discovering Gene Expression Patterns Associated with Increased Cancer Survival Time. Jeffrey A. Thompson, Christine Duarte, Peter Marks and Clare Bates Congdon.


32. GraphSpace: Sharing and collaborating through networks on the web. Craig Estep, Jaeil Yi, Anna Ritz and T. M. Murali.


36. A Collaborative Filtering Approach to Assess Individual Condition Risk Based on Patients’ Social Network Data. Xiang Ji, Soon Ae Chun and James Geller.

37. DDI2PPI: an integrated web server for protein-protein interaction and residue contact matrix predictions. Tianchuan Du, Alvaro Gonzalez, Qinghua Wang, Hongzhan Huang, Li Liao and Cathy Wu.


41. Predicting Protein Contact Maps by Bagging Decision Trees. Chuqiao Ren and Brian King.


43. Improving identification of key players in aging via network de-noising. Boyoung Yoo, Huili Chen, Fazle Faisal and Tijana Milenkovic.

44. Learning Parameter Sets for Alignment Advising. Dan Deblasio and John Kececioglu.


46. Exploring in silico about the neuroprotective action of erythropoietin enhanced by flavonoids: a study that reveal the importance of cooperative ligands in the cerebral ischemia case. David Chaupis-Meza, Eduardo Gushiken and Carlos Bueno.

47. A Flexible Volumetric Comparison of Protein Cavities can Reveal Patterns in Ligand Binding Specificity. Ziyi Guo, Trevor Kuhlengel, Steven Stinson, Seth Blumenthal, Soutir Bandyopadhyay and Brian Chen.

Demos/Exhibits

1. A Schema-Matching Tool for Alzheimer’s Disease Data Integration

2. The DOE Systems Biology Knowledgebase (KBase): Progress Towards a System for Collaborative and Reproducible Inference and Modeling of Biological Function
Program Chairs

**Program Chairs**

Terry Gaasterland  
University of California, San Diego

Ümit V. Çatalyürek  
The Ohio State University

Bruce Schatz  
University of Illinois at Urbana-Champaign

**Bioinformatics Track Chairs**

Marilyn Ritchie  
Penn State

Leonard McMillan  
University of North Carolina, Chapel Hill

**Computational and Translational Bioinformatics Track Chairs**

Attila Gursoy  
Koc University, Turkey

Kun Huang  
The Ohio State University

**Health Informatics Track Chairs**

ChengXiang Zhai  
University of Illinois

Santosh Kumar  
University of Memphis

**Common Methods and Systems Track Chairs**

T.M. Murali  
Virginia Tech

Christophe Giraud-Carrier  
Brigham Young University
## Program Committee Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mario Albrecht</td>
<td>University Medicine Greifswald</td>
</tr>
<tr>
<td>Nancy Amato</td>
<td>Texas A&amp;M University</td>
</tr>
<tr>
<td>Dhundy Bastola</td>
<td>University of Nebraska at Omaha</td>
</tr>
<tr>
<td>Asa Ben-Hur</td>
<td>Colorado State University</td>
</tr>
<tr>
<td>Jay Bernhardt</td>
<td>University of Texas</td>
</tr>
<tr>
<td>Paola Bonizzoni</td>
<td>Università di Milano-Bicocca</td>
</tr>
<tr>
<td>Serdar Bozdag</td>
<td>Marquette University</td>
</tr>
<tr>
<td>C. Titus Brown</td>
<td>Michigan State University</td>
</tr>
<tr>
<td>Yang Cao</td>
<td>Virginia Tech</td>
</tr>
<tr>
<td>Greg Carter</td>
<td>The Jackson Laboratory</td>
</tr>
<tr>
<td>Frederic Cazals</td>
<td>INRIA Sophia</td>
</tr>
<tr>
<td>Aristotelis Chatziioanou</td>
<td>National Hellenic Research Foundation</td>
</tr>
<tr>
<td>Yi-Ping Phoebe Chen</td>
<td>La Trobe University</td>
</tr>
<tr>
<td>Jianlin Cheng</td>
<td>University of Missouri Columbia</td>
</tr>
<tr>
<td>Munmun Choudhury</td>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td>Mark Clement</td>
<td>Brigham Young University</td>
</tr>
<tr>
<td>Hector Corrada Bravo</td>
<td>University of Maryland</td>
</tr>
<tr>
<td>Carlos Cotta</td>
<td>University of Malaga</td>
</tr>
<tr>
<td>Lenore Cowen</td>
<td>Tufts University</td>
</tr>
<tr>
<td>Dana Crawford</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Dina Demner Fushman</td>
<td>National Library Medicine, NIH</td>
</tr>
<tr>
<td>Emre Ertin</td>
<td>The Ohio State University</td>
</tr>
<tr>
<td>James Foster</td>
<td>University of Idaho</td>
</tr>
<tr>
<td>Alex Freitas</td>
<td>University of Kent</td>
</tr>
<tr>
<td>Kavita Ganesan</td>
<td>3M Research</td>
</tr>
<tr>
<td>Andrew Gentles</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Nurit Haspel</td>
<td>University of Massachusetts Boston</td>
</tr>
<tr>
<td>Jing He</td>
<td>Old Dominion University</td>
</tr>
<tr>
<td>Lenwood Health</td>
<td>Virginia Tech</td>
</tr>
<tr>
<td>Vasant Honavar</td>
<td>Pennsylvania State University</td>
</tr>
<tr>
<td>Jun Huan</td>
<td>University of Kansas</td>
</tr>
<tr>
<td>Filip Jagodzinski</td>
<td>Central Washington University</td>
</tr>
<tr>
<td>Janina Jeff</td>
<td>Icahn School of Medicine at Mount Sinai</td>
</tr>
<tr>
<td>Xiaojian Jiang</td>
<td>University of California, San Diego</td>
</tr>
<tr>
<td>Vladimir Jojic</td>
<td>University of North Carolina</td>
</tr>
<tr>
<td>Kamer Kaya</td>
<td>The Ohio State University</td>
</tr>
<tr>
<td>Zia Khan</td>
<td>University of Maryland, College Park</td>
</tr>
<tr>
<td>Dokyoon Kim</td>
<td>Pennsylvania State University</td>
</tr>
<tr>
<td>Mehmet Koyuturk</td>
<td>Case Western Reserve University</td>
</tr>
<tr>
<td>Xiaoli Li</td>
<td>Institute for Infocomm Research</td>
</tr>
<tr>
<td>Simon Lin</td>
<td>Marshfield Clinic Research Foundation</td>
</tr>
<tr>
<td>Hongfang Liu</td>
<td>Mayo Clinic</td>
</tr>
<tr>
<td>Jinze Liu</td>
<td>University of Kentucky</td>
</tr>
<tr>
<td>Stefano Lonardi</td>
<td>UC Riverside</td>
</tr>
<tr>
<td>Hui Lu</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td>Xinghua Lu</td>
<td>University of Pittsburgh</td>
</tr>
<tr>
<td>Name</td>
<td>Affiliation</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Jonathan Silverstein</td>
<td>NorthShore University HealthSystem</td>
</tr>
<tr>
<td>Zhiyong Lu</td>
<td>National Center Biotechnology Information, NIH</td>
</tr>
<tr>
<td>Shaun Mahony</td>
<td>Pennsylvania State University</td>
</tr>
<tr>
<td>Brett Mckinney</td>
<td>University of Tulsa</td>
</tr>
<tr>
<td>Qiaozhu Mei</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>Genevieve Melton-Meaux</td>
<td>University of Minnesota</td>
</tr>
<tr>
<td>Jason Moore</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Baldo Oliva</td>
<td>Pompeu Fabra University</td>
</tr>
<tr>
<td>Gaurav Pandey</td>
<td>Mount Sinai School of Medicine</td>
</tr>
<tr>
<td>Jyotishman Pathak</td>
<td>Mayo Clinic</td>
</tr>
<tr>
<td>Sarah Pendergrass</td>
<td>The Pennsylvania State University</td>
</tr>
<tr>
<td>Andy Perkins</td>
<td>Mississippi State University</td>
</tr>
<tr>
<td>Hoifung Poon</td>
<td>Microsoft Research</td>
</tr>
<tr>
<td>Ivo Sbalzarini</td>
<td>Max Planck Institute of Molecular Cell Biology and Genetics</td>
</tr>
<tr>
<td>Amarda Shehu</td>
<td>George Mason University</td>
</tr>
<tr>
<td>Yao Shen</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Xinghua Shi</td>
<td>University of North Carolina at Charlotte</td>
</tr>
<tr>
<td>Jimeng Sun</td>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td>Nurcan Tuncbag</td>
<td>Massachusetts Institute of Technology</td>
</tr>
<tr>
<td>Pierangelo Veltri</td>
<td>Laboratory of Bioinformatics, University of Catanzaro</td>
</tr>
<tr>
<td>Jason Wang</td>
<td>New Jersey Institute of Technology</td>
</tr>
<tr>
<td>Ryen White</td>
<td>Microsoft Research</td>
</tr>
<tr>
<td>Christopher Yang</td>
<td>Drexel University</td>
</tr>
<tr>
<td>Zhan Ye</td>
<td>Marshfield Clinic Research Foundation</td>
</tr>
<tr>
<td>Yinyin Yuan</td>
<td>The Institute of Cancer Research, UK</td>
</tr>
<tr>
<td>Aidong Zhang</td>
<td>State University of New York at Buffalo</td>
</tr>
<tr>
<td>Bin Zhang</td>
<td>Mount Sinai School of Medicine</td>
</tr>
<tr>
<td>Jie Zhang</td>
<td>Ohio State University</td>
</tr>
<tr>
<td>Liqing Zhang</td>
<td>Virginia Tech</td>
</tr>
<tr>
<td>Ping Zhang</td>
<td>IBM Thomas J. Watson Research Center</td>
</tr>
<tr>
<td>Jie Zhang</td>
<td>Ohio State University</td>
</tr>
</tbody>
</table>
Newport Beach Marriott Hotel & Spa

900 Newport Center Drive, Newport Beach, CA 92660
Telephone (949) 640-4000
Fax (949) 640-5055