

Message from the General Chairs



Welcome to the Second IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS 2012)! Computational techniques are revolutionizing the way in which research is conducted in science and engineering. Unsurpassed advances have been made in myriads of application domains. This is particularly true in the areas of biology, medicine, and drug discovery. Even though a number of conferences exist today in the general area of bioinformatics, they focus on computational biology to a large extent. ICCABS has the goal of bringing together scientists in all the three areas and hence serving as a platform for bridging the research efforts in these areas. The response from these communities to the ICCABS call for papers has been indeed gratifying. We are honored to serve these international communities by bringing together researchers, scientists, and students from academia, laboratories, and industry to this premier meeting held in the ever exciting city of Las Vegas, NV.

We are thankful to the steering committee members (Srinivas Aluru, Reda A. Ammar, Tao Jiang, Vipin Kumar, Ming Li, John Reif, and Sartaj Sahni) who have been instrumental in getting this conference approved by IEEE.

We would like to thank Program Chairs Ion Mandoiu, Mihai Pop, and John Spouge for assembling an outstanding technical program. The program committee members have done a wonderful job of reviewing all the submitted papers thoroughly within a short period of time. We are thankful to them as well as all the authors who have expressed their faith in ICCABS by submitting the fruits of their hard work. ICCABS features two keynote speeches (from Pavel Pevzner and David Sankoff), and eight invited talks (from Srinivas Aluru, Tanya Berger-Wolf, Vladimir Filkov, Bjarni Halldórsson, Maricell Kann, Vipin Kumar, Rachel O'Neill, and Roded Sharon). We are grateful to these well-established and highly successful researchers for taking time out of their busy schedules to make ICCABS a great success.

Thanks are also due to: Financial Chairs (Reda Ammar and Yufeng Wu), Local Arrangements Chair (Marty Schiller), Workshops Chair (Alex Zelikovsky), Proceedings Chairs (Sahar Al Seesi and Mai Hamdalla), and Webmaster (Sherif Tolba) for their tireless efforts. Ion Mandoiu and Alex Zelikovsky deserve our special thanks for putting together the Second Workshop on Computational Advances for Next Generation Sequencing (CANGS 2012).

Selected papers from ICCABS 2012 will appear as special issues in BMC Genomics and BMC Bioinformatics. We are grateful to the editors of these journals and Sarah Bauer for their efforts in this regard. Last but not least, we would like to thank UConn and the Booth Engineering Center for Advanced Technology (BECAT) for the constant administrative support rendered to ICCABS. In particular, we offer our special thanks to Debra Mielczarek for her incessant support.

We hope you'll continue to support ICCABS in future! Have a wonderful time in the conference!

Sorin Istrail, Brown University and *Sanguthevar Rajasekaran*, University of Connecticut

Message from the Program Chairs



On behalf of the Program Committee, we would like to welcome you to the 2nd IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS 2012), held in Las Vegas, Nevada on February 23-25, 2012. This year the technical program includes 25 extended abstracts selected by the Program Committee from a number of 55 submissions received in response to the call for papers. The program also includes 8 invited talks and a poster session, and features keynote talks by two distinguished speakers. David Sankoff from University of Ottawa speaks on fractionation, the process that results in the loss of one copy from almost all pairs of duplicated genes created by whole genome duplication, and Pavel Pevzner from the University of California at San Diego speaks on *de novo* sequencing of novel peptide antibiotics by tandem mass spectrometry. Additionally, the technical program includes 15 invited talks presented as part of the 2nd Workshop on Computational Advances for Next Generation Sequencing, held in conjunction with ICCABS 2012.

We would like to thank the Program Committee members and external reviewers for volunteering their time to review and discuss submitted abstracts. We would like to extend special thanks to the General Chairs, Sorin Istrail and Sanguthevar Rajasekaran, for their leadership, and to the Workshop, Local Organization, Finance, and Proceedings Chairs, for their hard work in making ICCABS 2012 a successful event. Last but not least we would like to thank all authors for presenting their work at the conference.

We hope you will find the ICCABS 2012 technical program interesting and thought provoking and that attending the conference will provide you with stimulating ideas and ample opportunities to meet other researchers from around the world. Enjoy!

Ion Mandoiu, University of Connecticut

Mihai Pop, University of Maryland

John Spouge, NCBI/NIH

Message from CANGS Workshop Chairs



Welcome to the 2nd Workshop on Computational Advances for Next Generation Sequencing (CANGS 2012), held on February 24-25, 2012 in Las Vegas, Nevada, in conjunction with ICCABS 2012. Massively parallel DNA and RNA sequencing have become widely available, reducing the cost by several orders of magnitude and placing the capacity to generate gigabases to terabases of sequence data into the hands of individual investigators. These next-generation technologies have the potential to dramatically accelerate biological and biomedical research by enabling the comprehensive analysis of genomes and transcriptomes to become inexpensive, routine and widespread.

The CANGS workshop aims to accelerate progress in the field by bringing together experts to discuss new directions of research and exchange ideas on the many mathematical and computational challenges presented by analysis of the exploding volume of next-generation sequencing data. This year's workshop program will span 3 sessions comprising invited talks on a variety of current research topics, including clustering and assembly of metagenomic data, graph compression approaches to assembly, discovery of mutated pathways in cancer, ecoinformatics, read error correction, gene fusion detection, haplotype inference and characterization of structural variation from low-coverage sequencing data, virus quasispecies reconstruction from shotgun and amplicon sequencing data, transcriptome assembly and inference of isoform expression levels, including allele specific expression, from RNA-Seq data. The workshop will conclude with a roundtable discussion of future research directions in the field.

We would like to thank ICCABS General Chairs Sorin Istrail and Sanguthevar Rajasekaran for the opportunity to organize CANGS 2012, and all speakers for presenting their work at the workshop.

Ion Mandoiu, University of Connecticut
Alex Zelikovsky, Georgia State University

Conference Organization

Steering Committee

Srinivas Aluru, Iowa State University
Reda A. Ammar, University of Connecticut
Tao Jiang, U.C. Riverside
Vipin Kumar, University of Minnesota
Ming Li, University of Waterloo
S. Rajasekaran, University of Connecticut (Chair)
John Reif, Duke University
Sartaj Sahni, University of Florida

General Chairs

Sorin Istrail, Brown University
S. Rajasekaran, University of Connecticut

Program Chairs

Ion Mandoiu, University of Connecticut
Mihai Pop, University of Maryland
John Spouge, NCBI/NIH

Finance Chairs

Reda A. Ammar, University of Connecticut
Yufeng Wu, University of Connecticut

Local Arrangements Chair

Martin Schiller, University of Nevada Las Vegas

Workshops Chair

Alex Zelikovsky, Georgia State University

Proceedings Chairs

Sahar Al Seesi, University of Connecticut
Mai Hamdalla, University of Connecticut

Webmaster

Sherif Tolba, University of Connecticut

Program Committee Members

Irina Astrovskaya
David A. Bader
Daniel Brown
Titus Brown
Doina Caragea
Kun-Mao Chao
Matthias Chung
Matteo Comin
Juan Cui
Ovidiu Daescu
Bhaskar Dasgupta
Jorge Duitama
Richard Edwards
Scott Emrich
Oliver Eulenstein
Guillaume Fertin
Vladimir Filkov
Liliana Florea
Martin Frith
Osamu Gotoh
Katia Guimaraes

Robert Harrison
Matthew He
Steffen Heber
Manuela Helmer-Citterich
Yongsheng Huang
Hasan Jamil
Sumit Kumar Jha
Lars Kaderali
Ananth Kalyanaraman
Yury Khudyakov
Danny Krizanc
Henry Lin
Stefano Lonardi
Bin Ma
Leonardo Marino-Ramirez
Niranjan Nagarajan
Radhakrishnan Nagarajan
Bogdan Pasaniuc
Andrei Paun
Nadia Pisanti
Teresa Przytycka

Ben Raphael
Daniel Schwartz
Russell Schwartz
Joao Setubal
Ileana Streinu
Raj Sunderraman
Wing-Kin Sung
Vishal Thapar
Kannan Tharakaraman
Jerzy Tiuryn
Todd Treangen
Ugo Vaccaro
Matthew Vaughn
Jianxin Wang
Li-San Wang
Fang Xiang Wu
Yuzhen Ye
Yanbin Yin
Shaojie Zhang

Additional Reviewers

Laurent Bulteau
David S. Campo
Diana Clausznitzer
Freddy Cliquet
Peter Clote
Raj Dutta
Ming Fang
Giulio Genovese
Katsuhisa Horimoto
Faraz Hussain
Hans-Michael Kaltenbach

Chadi Kari
Narsis Aftab Kiani
Bettina Knapp
James Lara
Audrey Lee-St.John
Yuk Yee Leung
Yuan Li
Liang Liu
Alexey Nefedov
Layla Oesper
Aureliano Rama

Paul Ryvkin
Pavel Skums
Bismita Srichandan
Cheemeng Tan
Fabio Vandin
Xiyin Wang
Bartek Wilczynski
Noah Zaitlen
Marko Zivanic
Cuncong Zhong

ICCABS 2012 Program

Thursday, February 23, 2012

8:30am-
8:40am **Opening Remarks**

8:40am-
10:10am **Plenary Session**
Chair: *Sanguthevar Rajasekaran*

Vipin Kumar (University of Minnesota), *Discovering Combinatorial Biomarkers (invited talk)*
Tanya Berger-Wolf (University of Illinois), *Computational Behavioral Ecology (invited talk)*
Maricel Kann (University of Maryland Baltimore County), *A Domain-Centric Approach of Translational Bioinformatics (invited talk)*

10:10am-
10:30am **Coffee Break**

10:30am-
12:10pm **Parallel Sessions**

Session 1A: Phylogenetics and Genomics
Chair: *Vipin Kumar*

- Chunfang Zheng, Victor Albert, Eric Lyons and David Sankoff. *Ancient Angiosperm Hexaploidy Meets Ancestral Eudicot Gene Order*
- Margareta Ackerman, Daniel Brown and David Loker. *Effects of Rooting via Outgroups on Ingroup Topology in Phylogeny*
- Wei Zhang, Erliang Zeng, Dan Liu, Stuart Jones and Scott Emrich. *A Machine Learning Framework for Trait Based Genomics*
- Yuri Pirola, Gianluca Della Vedova, Stefano Biffani, Alessandra Stella and Paola Bonizzoni. *A fast and practical approach to genotype phasing and imputation on a pedigree with erroneous and incomplete information*

Session 1B: Modeling
Chair: *Rachel O'Neill*

- Faraz Hussain, Raj Gautam Dutta, Sumit Kumar Jha, Christopher James Langmead and Susmit Jha. *Parameter Discovery for Stochastic Biological Models against Temporal Behavioral Specifications using an SPRT based Metric for Simulated Annealing*
- Eren Gultepe and Ilias Tagkopoulos. *A Bayesian network for early diagnosis of sepsis patients: a basis for a clinical decision support system*
- Ming-Chih Shih, Shou-Hsuan Stephen Huang, Youli Zu, Rachel Donohue and Chun-Che Jeff Chang. *Automatic B Cell Lymphoma Detection Using Flow Cytometry Data*
- Yam Ki Cheung, Ovidiu Daescu, Steven Kirtzic and Lech Papiezy. *Volume matching with application in medical treatment planning*

12:10pm-
1:30pm **Lunch**

1:30pm-
3:00pm **Plenary Session**
Chair: *John Spouge*

David Sankoff (University of Ottawa), *Aspects of Fractionation (keynote talk)*
Vladimir Filkov (UC Davis), *Identifying Mutations from TILLING Experiments (invited talk)*

**3:00pm-
3:20pm** Coffee Break

**3:20pm-
4:30pm** Parallel Sessions

Session 2A: Poster Overviews: Sequence Analysis and Systems Biology

Chair: Sahar Al Seesi

- Yuan Zhang, Dhananjai Rao, Jens Mueller, Praveen Kumar, Chun Liang and John Karro. *De Novo Expression Fragment Assembly from Spanning Trees (EAST)*
- James Lindsay, Ion Mandoiu, Hamed Salooti and Alex Zelikovsky. *Accurate Scaffolding of Large Genomes using Integer Programming and Non-Serial Dynamic Programming*
- Tomoya Tagami, Tsuyoshi Hachiya and Yasubumi Sakakibara. *In Silico Microarray algorithm: Accurate and fast taxonomic profiling from short read sequences*
- James Morton, Patricia Abrudan, Chun Liang and John Karro. *Sequence Classification of homoPolymer Emissions (SCOPE)*
- Jerlin Merlin and Hieu Dinh. *Randomized Algorithms for Planted Motif Search*
- Paul Yenerall, Yuanyuan Jiang and Leming Zhou. *MIGL: A Database for Identifying the Mechanisms of Intron Gain and Loss*
- Behzad Bokanizad. *Signaling Pathway Impact Analysis Robustness*

Session 2B: Poster Overviews: Imaging, and Medical Applications

Chair: Mai Hamdalla

- Ernur Saka and Eric Rouchka. *Image Registration and Visualization Tool For In-Situ Gene Expression Images*
- Maryam Rahimi and Mohamadali Oghabian. *Evaluation of the level set method by measure similarity (KCC) for clustering and analysis of fMRI data*
- Italo Zoppis, Massimiliano Borsani, Erica Gianazza, Clizia Chinello, Giancarlo Albo, Francesco Rocco, André M. Deelder, Yuri E. M. Van Der Burgt, Marco Antoniotti, Fulvio Magni and Giancarlo Mauri. *Characterization of Distinguishing Regions for Renal Cell Carcinoma Discrimination*
- Lingfei Zhi, Zhao Wang and Aman Behal. *Parameter Estimation for a Reduced Order Neuronal Spiking Model with Relevance to In Vitro Embryonic Rat Motoneuronal Data*
- Khaled Dawoud, Ala Qabaja, Shang Gao, Reda Alhaji, and Jon Rokne. *Identifying Cancer Biomarkers by Knowledge Discovery from Medical Literature*
- Fabiola Pinheiro and Mu-Hsing Kuo. *Applying the Data Mining Algorithms to Early Detection of Liver Cancer*
- Hye-Kyeong Kwon, Dong-Yu Kim, Jang-Soo Chun and Zee-Yong Park. *Mass Spectrometric Protein Profiling Analyses of Pathological and Physiological Hypertrophy Cardiac Muscle Tissues*
- Marissa Smith, Jesse Port, Jim Wallace and Elaine Faustman. *Public Health Applications of Metagenomic Data using Publicly Available Computing Frameworks*

**4:30pm-
6:00pm** Poster Session

Friday, February 24, 2012

8:30am- Plenary Session

9:30am Chair: *Ion Mandoiu*

Srinivas Aluru (Iowa State University), *Reverse Engineering Whole-Genome Networks (invited talk)*
Bjarni Halldórsson (Reykjavik University), *Algorithmic challenges in DNA sequencing and disease association (invited talk)*

9:30am- Coffee Break

9:50am

9:50am- Parallel Sessions

12:20pm

Session 3A: Systems Biology

Chair: *Bjarni Halldórsson*

- Mai Hamdalla, David Grant, Ion Mandoiu, Dennis Hill, Sanguthevar Rajasekaran and Reda Ammar. *The Use of Graph Matching Algorithms to Identify Biochemical Substructures in Synthetic Chemical Compounds: Application to Metabolomics*
- Linh Huynh and Ilias Tagkopoulos. *A robust, library-based, optimization-driven method for automatic gene circuit design*
- Peikai Chen, Y.S. Hung, Ching C. Lau, Yubo Fan, and Stephen T.-C. Wong. *A gene signature based method for identifying subtypes and subtype-specific drivers in cancer with an application to medulloblastoma*
- Martin Strauch, Julia Rein and C. Giovanni Galizia. *Signal extraction from movies of honeybee brain activity by convex analysis*
- Fabrizio F. Borelli, Raphael Y. Camargo, David Correa Martins Jr, Beatriz Stransky and Luiz C. S. Rozante. *Accelerating Gene Regulatory Networks Inference through GPU/CUDA Programming*
- Arup K. Ghosh, Faraz Hussain, Sumit K. Jha, Christopher J. Langmead and Sumit Jha. *Decision Procedure Based Discovery of Rare Behaviors in Stochastic Differential Equation Models of Biological Systems*

Session 3B: CANGS Workshop

Chair: *Alex Zelikovsky*

- Srinivas Aluru. *Taxonomic Clustering of Large-scale Metagenomic Data*
- Hieu Dinh and Sanguthevar Rajasekaran. *An efficient data structure for exact-match overlap graphs and next generation sequence assembly*
- Jason Pell, Arend Hintze, Rosangela Canino-Koning, Adina Howey, James M. Tiedjeyz, and C. Titus Brown. *Graph compression approaches to assembly*
- Bo Liu and Mihai Pop. *Comparative Assembly of Metagenomic Sequences*
- Fabio Vandin, Eli Upfal, and Benjamin J. Raphael. *Algorithms for Discovery of Mutated Pathways in Cancer*
- Scott Emrich. *Opportunities and challenges of non-model ecoinformatics*

12:20pm- Lunch

1:40pm

**1:40pm-
2:40pm** **Plenary Session**
Chair: Sanguthevar Rajasekaran

Pavel Pevzner (UCSD), *De Novo Sequencing of Novel Peptide Antibiotics by Tandem Mass Spectrometry (keynote talk)*

**2:40pm-
3:00pm** **Coffee Break**

**3:00pm-
5:30pm** **Parallel Sessions**

Session 4A: Sequence analysis

Chair: John Spouge

- Junjie Li, Sanjay Ranka and Sartaj Sahni. *Pairwise Sequence Alignment for Very Long Sequences on GPUs*
- John Spouge, Leo Marino-Ramirez and Sergey Sheetlin. *The Ruzzo-Tompa Algorithm Can Find Maximal Paths Between Vertices in Weighted One-Dimensional Lattice Graphs*
- Shibdas Bandyopadhyay, Sartaj Sahni and Sanguthevar Rajasekaran. *PMS6: A Faster Algorithm for Motif Discovery*
- Juan Carlos Francisco, Fred Cohan and Danny Krizanc. *Demarcation of Bacterial Ecotypes from DNA Sequence Data: A Comparative Analysis of Four Algorithms*
- Cuncong Zhong, Justen Andrews and Shaojie Zhang. *Discovering Non-coding RNA Elements in Drosophila 3' Untranslated Regions*
- Andrew Thrasher, Zachary Musgrave, Douglas Thain and Scott Emrich. *Shifting the Bioinformatics Computing Paradigm: A Case Study using Genome Annotation*

Session 4B: CANGS Workshop

Chair: Srinivas Aluru

- Rachel O'Neill, Ion Mandoiu, Mazhar I. Khan, Craig Obergfell, Hongjun Wang, Andrew Bligh, Bassam Tork, and Nicholas Mancuso, and Alexander Zelikovsky. *Bioinformatics methods for reconstruction of Infectious Bronchitis Virus quasispecies from next generation sequencing data*
- Konrad Scheffler, N Lance Hepler, Martin D Smith, Wayne Delpert, Art FY Poon, Douglas Richman and Sergei L Kosakovsky Pond. *Error correction, noise filtering, and phylogenetic analysis of HIV sequences using the 454 platform*
- Nicholas Mancuso, Bassam Tork, Pavel Skums, Lilia Ganova-Raeva, Ion Mandoiu, Alex Zelikovsky. *A Maximum Likelihood Method For Quasispecies Spectrum Assembly*
- Derek Aguiar and Sorin Istrail. *Robust algorithms for inferring haplotype phase and deletion polymorphism from high-throughput whole genome sequence data*
- Jin Zhang and Yufeng Wu. *Calling Structural Variation with Low-coverage Sequencing Data by Mapping to Focal Region*

**7:00pm-
9:00pm** **Banquet**

Saturday, February 25, 2012

8:30am- Plenary Session

9:30am Chair: *John Spouge*

Rachel O'Neill (University of Connecticut), *Using Novel Small RNAs To Understand Genome Structure And Function (invited talk)*

Roded Sharan (Tel Aviv University), *From Protein Networks to Disease Mechanisms (invited talk)*

9:30am- Coffee Break 9:50am

9:50am- Parallel Sessions 11:55pm

Session 5A: Structure analysis

Chair: *Roded Sharan*

- Jason Gallia, Katelyn Lavrich, Rachel McGough, Anna Tan-Wilson and Patrick Madden. *Filtering of MS/MS Data for Peptide Identification*
- Ricardo Henao, J. Will Thompson, M. Arthur Moseley, Geoffrey S. Ginsburg, Lawrence Carin and Joseph E. Lucas. *Hierarchical Factor Modeling of Proteomics Data*
- Pamela Clark, Jessica Grant, Samantha Monastra, Filip Jagodzinski and Ileana Streinu. *Periodic Rigidity of Protein Crystal Structures*
- Naomi Fox and Ileana Streinu. *Towards accurate modeling for protein rigidity analysis*
- Yuan Li and Shaojie Zhang. *Finding consensus stable local optimal structures for aligned RNA sequences*

Session 5B: CANGS Workshop

Chair: *Ion Mandoiu*

- Onur Sakarya and Heinz Breu. *Gene Fusion Detection with Short and Medium Length Next Generation Sequencing*
- Wei Li and Tao Jiang. *Transcriptome Assembly and Isoform Expression Level Estimation from Biased RNA-Seq Reads*
- Serghei Mangul, Adrian Caciula, Ion Mandoiu and Alex Zelikovsky. *Novel Transcript Reconstruction from Paired-End RNA-Seq Reads Using Fragment Length Distribution*
- Sahar Al Seesi and Ion Mandoiu. *Inference of allele specific expression levels from RNA-Seq data*

Notes

