ICCABS 2013

2013 IEEE 3rd International Conference on Computational Advances in Bio and medical Sciences (ICCABS)

June. 12-14, 2013,
Crowne Plaza,
New Orleans Airport, LA
http://www.iccabs.org

Table Of Contents

Message from the General Chairs

Welcome to the Third IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS 2013)! 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 New Orleans, LA.

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 Knut Reinert and Vladimir Filkov 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 three keynote speeches (from Dan Gusfield, Oliver Kohlbacher and Satoru Miyano), and three invited talks (from Mehmet Koyutürk, Yuri Khudyakov, and Ion I. Moraru). 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 (Subrata Saha), Workshops Chairs (Ananth Kalyanaraman, Yury Khudyakov, Ion Mandoiu, Alex Zelikovsky, and Jaroslaw Zola), Proceedings Chair (Sudipta Pathak), and Webmaster (Priya Periaswamy) for their tireless efforts.

Selected papers from ICCABS 2013 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 and Karen Slater for their incessant support. We hope you’ll continue to support ICCABS in future! Have a wonderful time in the conference!

Srinivas Aluru, Iowa State 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 3rd IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS 2013) in New Orleans, Louisiana on June 12-14, 2013. This year the technical program includes 18 extended abstracts selected by the Program Committee from a number of 43 submissions received in response to the call for papers. The program also includes 3 invited talks and 6 poster presentations, and features keynote talks by three distinguished speakers: Dan Gusfield from the University of California Davis, Oliver Kohlbacher from the University of Tübingen, Germany and Satoru Miyano from the University of Tokyo. Additionally, the technical program includes invited talks in the CANGS, CAME and BigLS workshops.

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, Srinivas Aluru and Sanguthevar Rajasekaran, for their leadership, and to the Workshop, Local Organization, Finance, and Proceedings Chairs, for their hard work in making ICCABS 2013 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 2013 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!

Vladimir Filkov, University of California at Davis
Knut Reinert, Freie Universität Berlin

Message from CANGS Workshop Chairs

Welcome to the 3rd Workshop on Computational Advances for Next Generation Sequencing (CANGS 2013), held on June 12-13, 2013 in New Orleans, Louisiana, in conjunction with ICCABS 2013. 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 four sessions (including a joint session with the 2nd Workshop on Computational Advances in Molecular Epidemiology) comprising 15 invited talks on a variety of current research topics, including space-efficient indexing, retrieval, and assembly of next-generation sequencing reads; genotyping deletions and insertions with population NGS reads; ribosomal footprinting; reconstructing the organization of cancer genomes using NGS; transcriptome reconstruction, isoform frequency estimation, and differential gene expression analysis from RNA-Seq reads; deconvolution of gene expression data from heterogeneous samples; viral quasispecies reconstruction; and pooling strategies for NGS-based molecular surveillance of viral diseases.

We would like to thank ICCABS General Chairs Srinivas Aluru and Sanguthevar Rajasekaran for the opportunity to organize CANGS 2013, and all speakers for presenting their work at the workshop.

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

Message from CAME Workshop Chairs

Welcome to the 2nd Workshop on Computational Advances in Molecular Epidemiology (CAME 2013), held on June 13, 2013 in New Orleans, Louisiana, in conjunction with ICCABS 2013. Molecular epidemiology is essentially an integrative scientific discipline that considers molecular biological processes in specific epidemiological settings. It relates molecular biological events to etiology, distribution and prevention of disease in human populations. Over years, molecular epidemiology has become extensively fused with mathematical and computational science and immensely benefitted from this tight association. The CAME workshop aims to provide a forum for presentation and discussion of the latest computational research in molecular epidemiology. This multidisciplinary workshop will bring together field practitioners of molecular epidemiology, molecular evolutionists, population geneticists, medical researchers, bioinformaticians, statisticians and computer scientists interested in the latest developments in algorithms, mining, visualization, modeling, simulation and other methods of computational, statistical and mathematical analysis of genetic and molecular data in the epidemiological context.

This year's workshop program will span two sessions (including a joint session with the 3rd Workshop on Computational Advances for Next Generation Sequencing) comprising 9 invited talks on a variety of current research topics, including viral quasispecies reconstruction from both shotgun and amplicon next-generation sequencing (NGS) reads; pooling strategies for NGS-based molecular surveillance of viral diseases; studies of mutational robustness and intra-host adaptation of hepatitis C virus; HCV genetic association to rate of liver fibrosis progression; novel approaches to drug resistance prediction including target structure; and methods for nucleic acids sequence reconstruction from mass spectrometry data.

We would like to thank ICCABS General Chairs Srinivas Aluru and Sanguthevar Rajasekaran for the opportunity to organize CAME 2013, and all speakers for presenting their work at the workshop.

Yury Khudyakov, CDC/CCID/NCHHSTP
Ion Mandoiu, University of Connecticut
Alex Zelikovsky, Georgia State University

Message from the BigLS Workshop Chairs


Welcome to the inaugural edition of the International Workshop on Big Data in Life Scienes (BigLS 2013), held on June 13, 2013 in New Orleans, Louisiana, in conjunction with ICCABS 2013. With the ever-increasing volume, velocity and variety biological and biomedical data collections continue to pose new challenges and increasing demands on computing and data management. The inherent complexity of this big data has forced us to rethink how we collect, store, combine and analyze it. The BigLS workshop is dedicated to the broad theme of big data analytics in life sciences. The goal of the workshop is to bring together leading researchers and practitioners working on a diverse range of big data problems related to biology and medicine, and engage them in a discussion about current big data problems, the state of computational tools and analytics, current and imminent challenges, and the future trends within life sciences. Specific topics in this inaugural edition of the BigLS workshop include:

Next-generation sequencing and assembly; Metagenomics and metaproteomics; Systems biology and informatics; Network science and analytics; Statistical and integrative approaches to big data biology and informatics; Scientific workflows and platforms; High performance computing tools and challenges.

This workshop constitutes a series of 8 invited talks, including 2 keynote talks by luminaries in the field of big data biological analytics, followed by a joint panel with focus on addressing key questions that relate to the identification of key applications areas within biology that need big data analytics, and obtaining a fundamental understanding and characterization of key challenges involved in developing the next generation of algorithms and software for this important and emerging domain.

We would like to thank the ICCABS General Chairs Profs. Srinivas Aluru and Sanguthevar Rajasekaran for the opportunity and support to organize BigLS 2013, and all speakers for participating at the workshop.

Ananth Kalyanaraman, School of EECS, Washington State University
Jaroslaw Zola, Rutgers Discovery Informatics Institute, Rutgers 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
John Reif - Duke University
Sartaj Sahni - University of Florida
S. Rajasekaran (Chair) - University of Connecticut

General Chairs

Srinivas Aluru - Iowa State University
S. Rajasekaran, University of Connecticut

Program Chairs

Vladimir Filkov - University of California, Davis
Knut Reinert - Freie Universität, Berlin

Finance Chairs

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

Local Arrangements Chair

Subrata Saha, University of Connecticut

Proceedings Chair

Sudipta Pathak, University of Connecticut

Webmaster

Priya Periaswamy, University of Connecticut

Program Committee Members

Tatsuya Akutsu
Irina Astrovskaya
David A. Bader
Daniel Brown
Doina Caragea
Kun-Mao Chao
Matthias Chung
Matteo Comin
Ovidiu Daescu
Bhaskar Dasgupta
Jorge Duitama
Richard Edwards
Leonardo Marino-Ramirez
Victor Missirian
T. M. Murali
Radhakrishnan Nagarajan
Niranjan Nagarajan
Bogdan Pasaniuc
Andrei Paun
Nadia Pisanti
Natasa Przulj
Teresa Przytycka
Knut Reinert
Daniel Schwartz
Joao Setubal
Steven Skiena
John Spouge
Ileana Streinu
Raj Sunderraman
Wing-Kin Sung
Sing-Hoi Sze
Ilias Tagkopoulos
Vishal Thapar
Jerzy Tiuryn
Todd Treangen
Ugo Vaccaro
Matthew Vaughn
Li-San Wang
Jianxin Wang
Fang Xiang Wu
Yanbin Yin
Alex Zelikovsky
Shaojie Zhang

ICCABS 2013 Program

Wednesday, June 12, 2013

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

8:40am-10:20am **Parallel Sessions**

Session 1A: ICCABS 1
Chair: *Knut Reinert*

- Shibdas Bandyopadhyay, Sartaj Sahni and Sanguthevar Rajasekaran. PMS6MC: *A Multicore Algorithm for Motif Discovery*
- Junjie Li, Sanjay Ranka and Sartaj Sahni. *Multicore and GPU Algorithms for Nussinov RNA Folding*
- Nam Vo, Quang Tran, Nobal Niraula and Vinhthuy Phan. *A Randomized Algorithm for Aligning DNA Sequences to Reference Genomes*
- Robert Gysel, Kristian Stevens and Dan Gusfield. *Triangulation Heuristics for Maximum Character Compatibility*

Session 1B: CANGS 1
Chair: *Ion Mandoiu*

- Chong Chu, Jin Zhang and Yufeng Wu. *An SVM-based Approach for Genotyping Deletions and Insertions with Population Sequence Reads*
- Meznah Almutairy, Jordan Fish, and C. Titus Brown. *Space-efficient read indexing and retrieval based on compressed de Bruijn graph structures*
- Farhad Hormozdiari and Eleazar Eskin. *Memory Efficient Assembly of Human Genome*
- Sanguthevar Rajasekaran. *Efficient Algorithms for Sequence Assembly*

10:20am-10:50am **Coffee Break**

10:50am-11:50am **Plenary Session**
Chair: ***Sanguthevar Rajasekaran***

Dan Gusfield (University of California Davis) keynote talk

11:50am-1:40pm **Lunch**

1:40pm-3:00pm **Parallel Sessions**

Session 2A: ICCABS 2
Chair: *Jaroslaw Zola*

- Guoxing Fu, Amit Sabnis and Robert Harrison. *A Deterministic-Stochastic Crossover Algorithm for Simulation of Complex Biochemical Systems*
- Inna Rytsareva and Ananth Kalyanaraman. *Scalable heuristics for clustering biological graphs*
- Karthik Tangirala and Doina Caragea. *Extraction of Gene Regulatory Networks from Biological Literature*

Session 2B: CANGS 2
Chair: *Alex Zelikovsky*

- Karen Merchante, Qiwen Hu, Anna N. Stepanova, Jose M. Alonso and Steffen Heber: *Deep sequencing of ribosomal footprints for studying* genome-wide mRNA translation in plants
- Layla Oesper and Benjamin Raphael: *Reconstructing the Organization of Cancer Genomes*
- Serghei Mangul, Sahar Al Seesi, Adrian Caciula, Dumitru Brinza, Ion Mandoiu and Alex Zelikovsky: *Transcriptome Assembly and Quantification from Ion Torrent RNA-Seq Data*

3:00pm-3:30pm **Coffee Break**

3:30pm-5:10pm **Parallel Sessions**

Session 3A: ICCABS 3
Chair: *Sing-Hoi Sze*

- Rittika Shamsuddin, Milka Doktorova, Sheila Jaswal, Audrey Lee-St.John and Kathryn McMenimen. *Computational Prediction of Hinge Axes in Proteins*
- Mohammad Shamsur Rahman and Alioune Ngom. FAC-PIN: *Fast Agglomerative Clustering Method for Functional Modules and Protein Complex Identification in PINs*
- Jason Kinser, Stephen Lockett, Thomas Turbyville, Karlyne Reilly and John Beutler. *Comparing Analysis Engines for Generated Micro-patterned, Actin Images*

Session 3B: CANGS 3
Chair: *Pavel Skums*

- Sing-Hoi Sze and Aaron Tarone: *A memory-efficient algorithm to obtain splicing graphs and de novo expression estimates from de Bruijn graphs of RNA-Seq data*
- Adrian Caciula, Alex Zelikovsky, Serghei Mangul, James Lindsay and Ion Mandoiu: *Monte-Carlo Regression Algorithm for Isoform Frequency Estimation from RNA-Seq Data*
- Ei-Wen Yang and Tao Jiang. *Differential gene expression analysis from RNA-Seq read counts*
- James Lindsay, Craig Nelson, and Ion Mandoiu: *Towards Whole Transcriptome Deconvolution Using Single-cell Data*

Thursday, June 13, 2013

8:40am- 9:40am	**Plenary Session** **Chair:** ***Knut Reinert***	**BigLS**

Oliver Kohlbacher (Universität Tübingen), keynote talk

Session 4: BigLS 1
Chair: *Ananth Kalyanaraman*

- Kamesh Madduri. *Parallel analysis of large graph-structured data in genomics and proteomics*
- Benjamin Langmead. *Practical software for big genomics data*
- Fabio Vandin. *Identifying significant mutations in large cohorts of cancer genomes*

9:40am- 10:10am	**Coffee Break**	**10:00am - 10:30am**	**BigLS Coffee Break**

10:10am - 11:50am	**Parallel Sessions**

Session 5A: ICCABS 4
Chair: *Folker Meyer*

- Emily Flynn, Filip Jagodzinski, Sharon Pamela Santana and Ileana Streinu. *Rigidity and Flexibility of Protein–Nucleic Acid Complexes*
- Kevin Molloy, Jennifer Van, Daniel Barbara and Amarda Shehu. *Higher-order Representations of Protein Structure Space*
- Ionel Rata, Kyle Wessells and Yaohang Li. *An Improved Statistics-based Backbone Torsion Potential Energy for Protein Loop Structure Modeling*
- Ashraf Yaseen and Yaohang Li. *Template-based Prediction of Protein 8-state Secondary Structures*

Session 5B: CANGS 4/CAME1
Chair: *Yuri Khudyakov*

- Armin Töpfer, Francesca Di Giallonardo, Karin J. Metzner, Huldrych F. Günthard and Niko Beerenwinkel: *QuasiRecomb: Probing of viral diversity by global haplotype prediction*
- Alexander Artyomenko, Nicholas Mancuso, Pavel Skums, Ion Măndoiu and Alex Zelikovsky*: kGEM: An Expectation Maximization Error Correction Algorithm for Next Generation Sequencing of Amplicon-based Data*
- Nicholas Mancuso, Pavel Skums, Bassam Tork, Ion Mandoiu, and Alex Zelikovsky. *Estimation of Viral Population Structure from Amplicon-Based Reads*
- Pavel Skums, Olga Glebova, Alex Zelikovsky, Ion Mandoiu and Yury Khudyakov: *Optimizing pooling strategies for the massive next-generation sequencing of viral samples*

Session 5C: BigLS 2, Keynotes
Chair: *Jaroslaw Zola*

- Srini Parthasarathy. *Recent advances in stochastic flow clustering*
- Srinivas Aluru. *Large-scale Metagenomic Clustering via Quasi Clique Enumeration and Read Assignment Ambiguity Resolution*

11:50am-
1:30pm **Lunch**

1:30pm-
3:00pm **Plenary Session**
Chair: *Knut Reinert*

Yuri Khudyakov, (CDC, Atlanta), invited talk
Ion Moraru, (U Connecticut), invited talk

Session 6: BigLS 3
Chair: *Ananth Kalyanaraman*

- Folker Meyer. *Surviving the NGS data deluge - A scalable architecture for next generation metagenomics*
- Vladimir Filkov. *Big data from plant genomics: elucidating regulatory networks in poplar trees*
- Ilkay Altintas. *Workflow-driven programming paradigms for distributed analysis of biological big data*

3:00pm -
3:30pm **Coffee Break**

3:30pm -
4:30pm **Plenary session**
Chair: *Vladimir Filkov*

Session 7: BigLS 4

Satoru Miyano (Tokyo university), keynote talk

- BigLS Panel: Big Data in Life Sciences - fad, fiction or fact?
  Moderators: A. Kalyanaraman, J. Zola

4:30pm-
6:35pm **Parallel Sessions**

Session 8A: ICCABS Posters
Chair: *Sanguthevar Rajasekaran*

- Sanghoon Lee, Saeid Belkasim,: *Poster : User-Centered Query Expansion Model for Health Disparities Research*.
- George Mathew, Zoran Obradovic. *Poster: Auto-reduction of Features for Containing Communication Costs in a Distributed Privacy-Preserving Clinical Decision Support System.*
- Pinheiro, Fabiola, Kuo, Mu-Hsing, Thomo, Alex, Barnett, Jeff. *Extracting association rules from liver cancer data using the FP-growth algorithm.*
- Raed I. Seetan, Anne M. Denton, Omar Al-Azzam, Ajay Kumar, M. Javed Iqbal and Shahryar F. Kianian. *Poster: Clustering-based Approach for Constructing Reliable Radiation Hybrid Framework Maps.*
- Rafael Toledo F. de Souza, José Luiz Rybarczyk Filho, Ney Lemke, Emerson L. de Santa-Helena, Suzana V. Schönwald, Günther J. L. Gerhardt. *Poster: Quantification of correlations between sleep spindles in EEG for patients with sleep apnea.*
- Claudia Vilo, Daniel Kunz, Xiang Gao and Qunfeng Dong. *POSTER: "Pseudomonas fluorescens Strain NCIMB 11764, a unique bacterium with cyanide adaptation features"*

Session 8B: CAME 2
Chair: *Alex Zelikovsky*

- David S. Campo, Zoya Dimitrova, Pavel Skums and Yury Khudyakov: *Mutational robustness of hepatitis C virus intra-host variants*
- Pavel Skums, Leonid Bunimovich, David S. Campo Rendon, Yury Khudyakov. *Computational analysis and modeling of intra-host adaptation of hepatitis C virus: the role of immune cross-reactivity of HCV intra-host variants*
- James Lara, F. Xavier López-Labrador, Fernando Gonzales Candelas and Yury Khudyakov. *HCV genetic association to rate of liver fibrosis progression.*
- Robert Harrison, Xiaxia Yu, and Irene Weber. *Using Triangulation to Include Target Structure Improves Drug Resistance Prediction Accuracy*

Gabriel Ilie, Alex Zelikovsky, and Ion Mandoiu. *Reference Assisted Nucleic Acid Sequence Reconstruction from Mass Spectrometry Data*

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

Friday, June 14, 2013

8:40am- **Plenary Session**
9:20am **Chair: *Vladimir Filkov***

Mehmet Koyuturk (Case Western Reserve University), invited talk

9:20am-
9:50am **Coffee Break**

9:50am- 11:30am	**Regular Session**

Session 9: ICCABS 5
Chair: ***Vladimir Filkov***

- Wei Zhang, Erliang Zeng, Josh Livermore, Dan Liu, Stuart Jones and Scott Emrich. *Predicting Bacterial Functional Traits from Whole Genome Sequences Using Random Forest*.
- Xiaxia Yu, Robert Harrison and Irene Weber. HIV drug resistance prediction using multiple regression
- Irena Lanc and Scott Emrich. *An Unsupervised Learning Approach to Assembly Validation*
- Elena Randou, Daniel Veltri and Amarda Shehu. *Systematic Analysis of Global Features and Model Building for Recognition of Antimicrobial Peptides*

Notes