

# **2012 IEEE 2nd International Conference on Computational Advances in Bio and Medical Sciences**

**(ICCABS 2012)**

**Las Vegas, Nevada, USA  
23 – 25 February 2012**



**IEEE Catalog Number: CFP12CAB-PRT  
ISBN: 978-1-4673-1320-9**

# TABLE OF CONTENTS

## KEYNOTE TALK ABSTRACTS

<b>Keynote: De Novo Sequencing of Novel Peptide Antibiotics by Tandem Mass Spectrometry</b> .....	1
<i>P. Pevzner</i>	
<b>Keynote: Aspects of Fractionation</b> .....	2
<i>D. Sankoff</i>	
<b>Invited: Using Novel Small RNAs To Understand Genome Structure And Function</b> .....	3
<i>R. O'Neill, M. Longo, C. Zhang, S. Trusiak, L. Hall, J. Brown, C. Obergfell</i>	

## INVITED TALK ABSTRACTS

<b>Invited: Reverse Engineering Whole-Genome Networks</b> .....	4
<i>S. Aluru</i>	
<b>Invited: Computational Behavioral Ecology</b> .....	5
<i>T. Berger-Wolf</i>	
<b>Invited: Identifying Mutations from TILLING Experiments</b> .....	6
<i>V. Filkov</i>	
<b>Invited: Algorithmic Challenges In DNA Sequencing And Disease Association</b> .....	7
<i>B. Halldorsson</i>	
<b>Invited: A Protein Domain-Centric Approach to Translational Bioinformatics</b> .....	8
<i>M. Kann</i>	
<b>Invited: Discovering Combinatorial Biomarkers</b> .....	9
<i>G. Atluri, S. Dey, G. Fang, S. Landman, V. Paunic, W. Wang, M. Steinbach, V. Kumar</i>	
<b>Invited: From Protein Networks To Disease Mechanisms</b> .....	11
<i>R. Sharan</i>	

## EXTENDED ABSTRACTS

<b>Effects of Rooting via Outgroups on Ingroup Topology in Phylogeny</b> .....	12
<i>M. Ackerman, D. Brown, D. Loker</i>	
<b>PMS6: A Fast Algorithm for Motif Discovery</b> .....	19
<i>S. Bandyopadhyay, S. Sahni, S. Rajasekaran</i>	
<b>Accelerating Gene Regulatory Networks Inference through GPU/CUDA Programming</b> .....	25
<i>F. Borelli, R. Camargo, D. Martins-Jr., B. Stransky, L. Rozante</i>	
<b>A Gene Signature Based Method For Identifying Subtypes And Subtype-Specific Drivers In Cancer With An Application To Medulloblastoma</b> .....	31
<i>P. Chen, Y. Hung, T. Man, C. Lau, Y. Fan, S. Wong</i>	
<b>Volume Matching With Application In Medical Treatment Planning</b> .....	37
<i>Y. Cheung, O. Daescu, S. Kirtzic, L. Papiez</i>	
<b>Periodic Rigidity of Protein Crystal Structures</b> .....	43
<i>P. Clark, J. Grant, S. Monastra, F. Jagodzinski, I. Streinu</i>	
<b>Towards Accurate Modeling For Protein Rigidity Analysis</b> .....	49
<i>N. Fox, I. Streinu</i>	
<b>Demarcation of Bacterial Ecotypes from DNA Sequence Data: A Comparative Analysis of Four Algorithms</b> .....	55
<i>J. Francisco, F. Cohan, D. Krizanc</i>	
<b>Filtering of MS/MS Data for Peptide Identification</b> .....	61
<i>J. Gallia, K. Lavrich, R. McGough, A. Tan-Wilson, P. Madden</i>	
<b>Decision Procedure Based Discovery of Rare Behaviors in Stochastic Differential Equation Models of Biological Systems</b> .....	67
<i>A. Ghosh, F. Hussain, S. Jha, C. Langmead, S. Jha</i>	
<b>A Bayesian Network For Early Diagnosis Of Sepsis Patients: A Basis For A Clinical Decision Support System</b> .....	73
<i>E. Gultepe, H. Nguyen, T. Albertson, I. Tagkopoulos</i>	

<b>The Use of Graph Matching Algorithms to Identify Biochemical Substructures in Synthetic Chemical Compounds: Application to Metabolomics</b> .....	78
<i>M. Hamdalla, D. Grant, I. Mandoiu, D. Hill, S. Rajasekaran, R. Ammar</i>	
<b>Hierarchical Factor Modeling of Proteomics Data</b> .....	84
<i>R. Henao, J. Thompson, M. Moseley, G. Ginsburg, L. Carin, J. Lucas</i>	
<b>A Robust, Library-Based, Optimization-Driven Method For Automatic Gene Circuit Design</b> .....	90
<i>L. Hunyh, I. Tagkopoulos</i>	
<b>Parameter Discovery for Stochastic Biological Models against Temporal Behavioral Specifications using an SPRT based Metric for Simulated Annealing</b> .....	96
<i>F. Hussain, G. Dutta, S. Jha, C. Langmead, S. Jha</i>	
<b>Pairwise Sequence Alignment for Very Long Sequences on GPUs</b> .....	102
<i>J. Li, S. Ranka, S. Sahni</i>	
<b>Finding Consensus Stable Local Optimal Structures For Aligned RNA Sequences</b> .....	108
<i>Y. Li, S. Zhang</i>	
<b>A Fast And Practical Approach To Genotype Phasing And Imputation On A Pedigree With Erroneous And Incomplete Information</b> .....	114
<i>Y. Pirola, G. Vedova, S. Biffani, A. Stella, P. Bonizzoni</i>	
<b>Automatic B Cell Lymphoma Detection Using Flow Cytometry Data</b> .....	120
<i>M. Shih, S. Huang, Y. Zu, R. Donohue, C. Chang</i>	
<b>The Ruzzo-Tompa Algorithm Can Find the Maximal Paths in Weighted, Directed Graphs on a One-Dimensional Lattice</b> .....	126
<i>J. Spouge, L. Marino-Ramirez, S. Sheetlin</i>	
<b>Signal Extraction From Movies Of Honeybee Brain Activity By Convex Analysis</b> .....	132
<i>M. Strauch, J. Rein, C. Galizia</i>	
<b>Shifting the Bioinformatics Computing Paradigm: A Case Study in Parallelizing Genome Annotation Using MAKER and Work Queue</b> .....	138
<i>A. Thrasher, D. Thain, S. Emrich, Z. Musgrave</i>	
<b>A Machine Learning Framework for Trait Based Genomics</b> .....	144
<i>W. Zhang, E. Zeng, D. Liu, S. Jones, S. Emrich</i>	
<b>Ancient Angiosperm Hexaploidy Meets Ancestral Eudicot Gene Order</b> .....	150
<i>C. Zheng, V. Alber, E. Lyons, D. Sankoff</i>	
<b>Discovering Non-coding RNA Elements in Drosophila 3' Untranslated Regions</b> .....	156
<i>C. Zhong, J. Andrews, S. Zhang</i>	

## **POSTER ABSTRACTS**

<b>Identifying Cancer Biomarkers by Knowledge Discovery from Medical Literature</b> .....	162
<i>K. Dawoud, A. Qabaja, S. Gao, R. Alhajj, J. Rokne</i>	
<b>Poster Title (Mass Spectrometric Protein Profiling Analyses of Pathological and Physiological Hypertrophy Cardiac Muscle Tissues)</b> .....	163
<i>H. Kwon, J. Chun, Z. Park, D. Kim</i>	
<b>Poster: Accurate Scaffolding Of Large Genomes Using Integer Programming And Non-Serial Dynamic Programming</b> .....	164
<i>J. Lindsay, I. Mandoiu, H. Salooti, A. Zelikovsky</i>	
<b>Poster: Randomized Algorithms for Planted Motif Search</b> .....	165
<i>J. Merlin, H. Dinh</i>	
<b>Poster: Sequence Classification of homoPolymer Emissions (SCOPE)</b> .....	166
<i>J. Morton, P. Abrudan, C. Liang, J. Karro</i>	
<b>Poster: Applying Data Mining Algorithms to Early Detection of Liver Cancer</b> .....	167
<i>F. Pinheiro, M. Kuo</i>	
<b>Poster Evaluation of the Level Set Method by using Similarity criterion (KCC) for Clustering and Analysis of Functional MRI Data</b> .....	168
<i>M. Rahimi, M. Oghabian</i>	
<b>Poster: Image Registration and Visualization Tool For In-Situ Gene Expression Images</b> .....	169
<i>E. Saka, E. Rouchka</i>	
<b>Poster: Public Health Applications of Metagenomic Data using Publicly Available Computing Framework</b> .....	170
<i>M. Smith, J. Port, J. Wallace, E. Faustman</i>	
<b>Poster: In Silico Microarray Algorithm: Accurate And Fast Taxonomic Profiling From Short Read Sequences</b> .....	171
<i>T. Tagami, T. Hachiya, Y. Sakakibara</i>	

<b>Poster: MIGL - A Database for Identifying the Mechanisms of Intron Gain and Loss</b> .....	172
<i>P. Yenerall, Y. Jiang, L. Zhou</i>	
<b>Poster: de novo Expression Fragment Assembly from Spanning Trees (EAST)</b> .....	173
<i>Y. Zhang, D. Rao, M. Ozden, J. Mueller, P. Kumar, C. Liang, J. Karro</i>	
<b>Poster: Parameter Estimation for a Reduced Order Spiking Neuron Model with Relevance to In Vitro Embryonic Rat Motoneuronal Data</b> .....	174
<i>L. Zhi, Z. Wang, A. Behal</i>	
<b>Poster: Characterization of Distinguishing Regions for Renal Cell Carcinoma Discrimination</b> .....	175
<i>I. Zoppis, M. Borsani, E. Gianazza, C. Chinello, G. Albo, F. Rocco, A. Deelder, Y. Burgt, M. Antoniotti, F. Magni, G. Mauri</i>	

## **CANGS WORKSHOP ABSTRACTS**

<b>Workshop: Robust Algorithms For Inferring Haplotype Phase And Deletion Polymorphism From High-Throughput Whole Genome Sequence Data</b> .....	176
<i>D. Aguiar, S. Istrail</i>	
<b>Workshop: Inference Of Allele Specific Expression Levels From RNA-Seq Data</b> .....	177
<i>S. Seesi, I. Mandoiu</i>	
<b>Workshop: Taxonomic Clustering of Large-scale Metagenomic Data</b> .....	178
<i>S. Aluru</i>	
<b>Workshop: An Efficient Data Structure For Exact-Match Overlap Graphs And Next Generation Sequence Assembly</b> .....	179
<i>H. Dinh, S. Rajasekaran</i>	
<b>Workshop: Opportunities And Challenges Of Non-Model Ecoinformatics</b> .....	180
<i>S. Emrich</i>	
<b>Workshop: Transcriptome Assembly and Isoform Expression Level Estimation from Biased RNA-Seq Reads</b> .....	181
<i>W. Li, T. Jiang</i>	
<b>Workshop: Comparative Assembly of Metagenomic Sequences</b> .....	182
<i>B. Liu, M. Pop</i>	
<b>Workshop: A Maximum Likelihood Method For Quasispecies Spectrum Assembly</b> .....	183
<i>N. Mancuso, B. Tork, P. Skums, L. Ganova-Raeva, I. Mandoiu, A. Zelikovsky</i>	
<b>Workshop : Novel Transcript Reconstruction from Paired-End RNA-Seq Reads Using Fragment Length Distribution</b> .....	184
<i>S. Mangul, A. Caciula, I. Mandoiu, A. Zelikovsky</i>	
<b>Workshop: Bioinformatics Methods For Reconstruction Of Infectious Bronchitis Virus Quasispecies From Next Generation Sequencing Data</b> .....	185
<i>R. O'Neill, I. Mandoiu, M. Khan, C. Oberfell, H. Wang, A. Bligh, B. Tork, N. Mancuso, A. Zelikovsky</i>	
<b>Workshop: Graph Compression Approaches In Assembly</b> .....	186
<i>J. Pell, A. Hintze, R. Canino-Koning, A. Howe, J. Tiedje, C. Brown</i>	
<b>Workshop: Gene Fusion Detection with Short and Medium Length Next Generation Sequencing</b> .....	188
<i>O. Sakarya, H. Brey</i>	
<b>Workshop: Error Correction, Noise Filtering, And Phylogenetic Analysis Of HIV Sequences Using The 454 Platform</b> .....	189
<i>K. Scheffler, N. Hepler, M. Smith, W. Delpont, A. Poon, D. Richman, S. Pond</i>	
<b>Workshop: Algorithms for Discovery of Mutated Pathways in Cancer</b> .....	190
<i>F. Vandin, E. Upfal, B. Raphael</i>	
<b>Workshop: Calling Structural Variation with Low-coverage Sequencing Data by Mapping to Focal Region</b> .....	191
<i>J. Zhang, Y. Wu</i>	
<b>Author Index</b>	