

# **Proceedings 2012 IEEE International Workshop on Genomic Signal Processing and Statistics**

**(GENSIPS 2012)**

**Washington, DC, USA  
2 – 4 December 2012**



**IEEE Catalog Number: CFP12GEN-PRT  
ISBN: 978-1-4673-5234-5**

# Table of Contents

## Mon-Session: Integrative Genomic Data Analysis I

Integrative analysis of methylation and gene expression data in TCGA .....	1
<i>Yihua Liu, Peng Qiu</i>	
Hierarchical Bayesian Methods for Integration of Various Types of Genomics Data .....	5
<i>Elizabeth Jennings, Jeffrey Morris, Raymond Carroll, Ganiraju Manyam, Veerabhadran Baladandayuthapani</i>	
Anticancer Drug Sensitivity Analysis: An integrated approach applied to Erlotinib sensitivity prediction in the CCLE Database .....	9
<i>Ranadip Pal, Noah Berlow, Saad Haider</i>	
Integration of Multiple Data Sources for Identifying Functional Modules Using Bayesian Network .....	13
<i>Jinlian Wang, Hongyan Yuan, Mahlet Tadesse, Habtom Ressom</i>	
Integrating Multi-Platform Genomic Data Using Hierarchical Bayesian Relevance Vector Machines .....	18
<i>Sanvesh Srivastava, Wenyi Wang, Pascal Zinn, Rivka Colen, Veerabhadran Baladandayuthapani</i>	
Integration of Imaging and Systems Biology to Study Treatment of Medulloblastoma .....	22
<i>Sean Wang, Olga Rodriguez, Ye Tian, Shaozhen Ye, Emanuel Petricoin, Chris Albanese</i>	

## Mon-Session: Next-Generation Sequencing Data Analysis I

Fast DNA Barcode Generating Algorithm Using Radix Coding Method .....	26
<i>Amir Nikooienejad, Richard Metz, Byung-Jun Yoon, Charles Johnson</i>	
Design of Short Barcodes for Next Generation Sequencing of DNA and RNA .....	31
<i>Steffen Schober, Katharina Mir, Klaus Neuhaus, Martin Bossert</i>	
A Message Passing Algorithm for Reference-Guided Sequence Assembly from High-Throughput Sequencing Data .....	35
<i>Xiaohu Shen, Haris Vikalo</i>	
A Comparative Study of Methods for Detecting Small Somatic Variants in Disease-Normal Paired Next Generation Sequencing Data .....	38
<i>Qingguo Wang, Zhongming Zhao</i>	
A Bayesian Model for SNP Discovery Based on Next-Generation Sequencing Data .....	42
<i>Yanxun Xu, Xiaofeng Zheng, Yuan Yuan, Marcos Estecio, Jean-Pierre Issa, Yuan Ji, Shoudan Liang</i>	
Utilizing RNA-Seq Data for Cancer Network Inference .....	46
<i>Ying Cai, Bernard Fendler, Gurinder Atwal</i>	

## Mon-Session: Interaction Networks and Controls

Optimal Therapeutic Methods with Random-Length Response in Probabilistic Boolean Networks . . . . .	50
<i>Mohammadmahdi Rezaei Yousefi, Aniruddha Datta, Edward Dougherty</i>	
A New Phenotypically Constrained Control Policy in Boolean Networks based on Basins of Attraction	54
<i>Xiaoning Qian</i>	
Combination Therapy Design for Targeted Therapeutics from a Drug-Protein Interaction Perspective .	58
<i>Saad Haider, Noah Berlow, Ranadip Pal, Lora Davis, Charles Keller</i>	
Genome-Wide Meta-Regression of Gene-Environment Interaction . . . . .	62
<i>Xiaoxiao Xu, Gang Shi, Arye Nehorai</i>	
Feature Ranking based on Synergy Networks to Identify Prognostic Markers in DPT-1 . . . . .	66
<i>Amin Ahmadi Adl, Xiaoning Qian, Ping Xu, Kendra Vehik, Jeffrey Krischer</i>	

## Mon-Session: Next-generation Sequencing Data Analysis II

Improving the Flexibility of RNA-Seq Data Analysis Pipelines . . . . .	70
<i>John H. Phan, Po-Yen Wu, May D. Wang</i>	
Accurate Identification of Significant Aberrations in Contaminated Cancer Genome . . . . .	74
<i>Xuchu Hou, Bai Zhang, Xiguo Yuan, Ie-Ming Shih, Zhen Zhang, Robert Clark, Subha Madhavan, Guoqiang Yu</i>	
A Comparison of Periodicity Profile Methods for Sequence Analysis . . . . .	78
<i>Manas Bellani, Julien Epps, Gavin Huttley</i>	
NMF Based Approach for Finding Recurrent Aberrations in DNA Copy Number Data . . . . .	82
<i>Mohammed Elhenawy, Guoqiang Yu</i>	
MiB: A Comparative Assembly Processing Pipeline . . . . .	86
<i>Bilal Wajid, Erchin Serpedin, Mohamed Nounou, Hazem Nounou</i>	
Signal Transducers and Activators of Transcription (STATs) Act as Genomic Enhancer Organizers . . . .	N/A
<i>Golnaz Vahedi, Hayato Takahashi, Shingo Nakayamada, Hong-wei Sun, Vittorio Sartorelli, Yuka Kanno, John J. O'Shea</i>	

## Mon-Session: Poster Session

OBSGene: Optimal Barcode Generator for NGS Experiments . . . . .	90
<i>Esmail Atashpaz-Gargari, Mohammad Shahrokh Esfahani, Richard Metz, Charles D. Johnson</i>	
Determining the relative prevalence of different subpopulations in heterogeneous cancer tissue . . . . .	95
<i>Anwoy Mohanty, Aniruddha Datta, Jijayanagaram Venkatraj</i>	
Detection and Clustering of Palindrome in Genome Sequences by DCT-II Method . . . . .	97
<i>AhsanRizviand C. Bhattacharya.</i>	
Cell Type Specific Analysis of Human Transcriptome Data . . . . .	99
<i>Xiaoxiao Xu, Arye Nehorai, Joseph Dougherty</i>	

Comprehensive Analyses of Tumor Suppressor Genes and Oncogenes in Protein-protein Interaction Networks: A Topological Perspective .....	101
<i>Min Zhao, Jingchun Sun, Zhongming Zhao</i>	
An investigation of clinical outcome prediction from integrative genomic profiles in ovarian cancer .....	103
<i>Lin Zhang, Hui Liu, Tzu-Hung Hsiao, Yidong Chen, Yufei Huang</i>	
Detection of Short Exons in DNA Sequences Using Complex Wavelet Transform of Structural Features .....	107
<i>Ivo Provaznik, Vladimira Kubickov, Helena Skutkova, Jir Nedved, Ewaryst Tkacz, Petr Babula, Ren Kizek</i>	
Multiclass cancer-microarray classification algorithm with Pair-Against-All redundancy .....	111
<i>Mattia Bosio, Philippe Salembier, Pau Bellot, Albert Oliveras-Vergs</i>	
Sparse Bayesian Graphical Models for RPPA Time Course Data .....	113
<i>Riten Mitra, Peter Mueller, Yuan Ji, Gordon Mills, Yiling Lu</i>	
A Thermodynamics-Based Model for Identifying Splicing Regulatory Elements and Their Interactions .	N/A
<i>Ji Wen, Zhibin Chen, Xiaodong Cai</i>	

## **Tue-Session: Integrative Genomic Data Analysis II**

A new algorithm for predicting competing endogenous RNAs .....	118
<i>Mario Flores, Yufei Huang</i>	
Logistic Principal Component Analysis for Rare Variants in Gene-Environment Interaction Analysis ...	122
<i>Meng Lu, Hye-Seung Lee, David Hadley, Jianhua Z. Huang, Xiaoning Qian</i>	
A Non-Parametric Approach for Estimating Stromal Contamination in Cancer Samples .....	126
<i>Mohammed Elthenawy, Xuchu Hou, Guoqiang Yu</i>	
Latent Feature Decompositions for Integrative Analysis of Diverse High-throughput Genomic Data .....	130
<i>Karl Gregory, Kevin Coombes, Amin Momin, Luc Girard, Lauren Byers, Steve Lin, Micheal Peyton, John Heymach, John Minna, Veerabhadran Baladandayuthapani</i>	
A Bayesian Graphical Model for Integrative Analysis of TCGA Data .....	135
<i>Yanxun Xu, Jie Zhang, Yuan Yuan, Riten Mitra, Peter Muller, Yuan Ji</i>	

## **Tue-Session: Marker Identification and Tumor Classification**

Network-based Methods to Identify Highly Discriminating Subsets of Biomarkers .....	139
<i>SeyedJavad Sajjadi, Xiaoning Qian, Bo Zeng</i>	
Information Theoretic Feature Selection for High Dimensional Metagenomic Data .....	143
<i>Gregory Ditzler, Robi Polikar, Gail Rosen</i>	
Finding Robust Pathway Markers for Cancer Classification .....	147
<i>Navadon Khunlertgit, Byung-Jun Yoon</i>	
A Statistical Test for Intrinsically Multivariate Predictive Genes .....	151
<i>Ting Chen, Ulisses Braga-Neto</i>	
Identification of Functional Genomic Regions with Copy Number Alteration in Liver Cancer .....	155
<i>Tzu-Hung Hsiao, Hung-I Harry Chen, Yidong Chen</i>	

## Tue-Session: Genetic Regulatory Networks

Inferring gene functions from metabolic reactions .....	159
<i>Gunhan Gulsoy, Tamer Kahveci</i>	
Optimal Bayesian Classification and Its Application to Gene Regulatory Networks .....	164
<i>Lori Dalton, Edward Dougherty</i>	
Inferring Weighted and Directed Gene Interaction Networks from Gene Expression Data Using the Phi-Mixing Coefficient .....	168
<i>Nitin Singh, Mehmet Ahsen, Shiva Mankala, Michael White, Mathukumalli Vidyasagar</i>	
Inference of Time-Varying Gene Networks using Constrained and Smoothed Kalman Filtering .....	172
<i>Ghulam Rasool, Nidhal Bouaynaya</i>	
Determining Gene Function in Boolean Networks using Boolean Satisfiability .....	176
<i>Pey-Chang Lin, Sunil Khatri</i>	
External Control Of Probabilistic Boolean Networks Using Inhomogeneous Markov Chains .....	180
<i>Timothy Josserand</i>	

## Tue-Session: Methods in Proteomics

Toward exhaustive gating of flow cytometry data .....	183
<i>Peng Qiu</i>	
Normalization of LC-MS data using Gaussian process .....	187
<i>Mohammad R. Nezami Ranjbar, Mahlet Tadesse, Yue Wang, Habtom Ressom</i>	
Robust MS serum sample classification in proteomics by the use of inverse problems .....	191
<i>Pascal Szacherski, Jean-Francois Giovannelli, Laurent Gerfault, Audrey Giremus, Pierre Grangeat</i>	
Variability assessment of LC-MS experiments and its application to experimental design and difference detection .....	195
<i>Yi Zhao, Tsung-Heng Tsai, Cristina Di Poto, Lewis Pannell, Mahlet Tadesse, Habtom Ressom</i>	
An SILAC Quantification Algorithm for LC-FTMS .....	199
<i>Jian Cui, Xuepo Ma, Xiaolin Zhang, William Haskins, Jianqiu Zhang</i>	
An Evaluation for Merging Signaling Pathways by Using Protein-Protein Interaction Data .....	203
<i>Xiaogang Wu, Jake Chen</i>	

## Author Index