

Pacific Symposium on Biocomputing 2013

**Kohala Coast, Hawaii, USA
3 - 7 January 2013**

Editors:

**Russ B. Altman
A. Keith Dunker
Lawrence Hunter**

**Tiffany Murray
Teri E. Klein**

ISBN: 978-1-62748-016-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2012) by World Scientific Publishing Co., Inc.
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact World Scientific Publishing Co., Inc.
at the address below.

World Scientific Publishing Co., Inc.
27 Warren Street, Suite 401-402
Hackensack, NJ 17601

Phone: (800) 227-7562
Fax: (201) 487-9656

wspc@wspc.com

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

COMPUTATIONAL DRUG REPOSITIONING

<i>Session Introduction</i>	1
Zhiyong Lu, Pankaj Agarwal, and Atul Butte	
<i>Evaluation of Analytical Methods for Connectivity Map Data</i>	5
Jie Cheng, Qing Xie, Vinod Kumar, Mark Hurle, Johannes M. Freudenberg, Lun Yang, Pankaj Agarwal	
<i>Predictive Systems Biology Approach to Broad-Spectrum, Host-Directed Drug Target Discovery in Infectious Diseases</i>	17
Ramon M. Felciano, Sina Bavari, Daniel R. Richards, Jean-Noel Billaud, Travis Warren, Rekha Panchal, Andreas Krämer	
<i>A Novel Multi-Modal Drug Repurposing Approach for Identification of Potent ACK1 Inhibitors</i>	29
Sharangdhar S. Phatak, Shuxing Zhang	
<i>Protein-chemical Interaction Prediction via Kernelized Sparse Learning SVM</i>	41
Yi Shi, Xinhua Zhang, Xiaoping Liao, Guohui Lin, Dale Schuurmans	
<i>Drug Target Predictions Based on Heterogeneous Graph Inference</i>	53
Wenhui Wang, Sen Yang, Jing Li	

EPIGENOMICS

<i>Session Introduction</i>	65
Alexander J. Hartemink, Manolis Kellis, William Stafford Noble, Zhiping Weng	
<i>A Powerful Statistical Method for Identifying Differentially Methylated Markers in Complex Diseases</i>	69
Surin Ahn, Tao Wang	
<i>Using DNase Digestion Data to Accurately Identify Transcription Factor Binding Sites</i>	80
Kaixuan Luo, Alexander J. Hartemink	
<i>Epigenomic Model of Cardiac Enhancers with Application to Genome Wide Association Studies</i>	92
Avinash Das Sahu, Radhouane Aniba, Yen-Pei Christy Chang, Sridhar Hannenhalli	

IDENTIFICATION OF ABERRANT PATHWAY AND NETWORK ACTIVITY FROM HIGH-THROUGHPUT DATA

<i>Session Introduction</i>	103
Rachel Karchin, Michael F. Ochs, Joshua M. Stuart, Trey Ideker, Joel S. Bader	
<i>From Uncertain Protein Interaction Networks to Signaling Pathways Through Intensive Color Coding</i>	111
Haitham Gabr, Alin Dobra, Tamer Kahveci	
<i>Identifying Master Regulators of Cancer and Their Downstream Targets by Integrating Genomic and Epigenomic Features</i>	123
Olivier Gevaert, Sylvia Plevritis	
<i>Module Cover – A New Approach to Genotype-Phenotype Studies</i>	135
Yoo-Ah Kim, Raheleh Salari, Stefan Wuchty, Teresa M. Przytycka	
<i>Next-Generation Analysis of Cataracts: Determining Knowledge Driven Gene-Gene Interactions Using Biofilter, and Gene-Environment Interactions Using the PhenX Toolkit</i>	147
Sarah A. Pendergrass, Shefali S. Verma, Emily R. Holzinger, Carrie B. Moore, John Wallace, Scott M. Dudek, Wayne Huggins, Terrie Kitchner, Carol Waudby, Richard Berg, Catherine A. McCarty, Marylyn D. Ritchie	
<i>Interpreting Personal Transcriptomes: Personalized Mechanism-Scale Profiling of RNA-seq Data</i>	159
Alan Perez-Rathke, Haiquan Li, Yves A. Lussier	
PERSONALIZED MEDICINE: FROM GENOTYPES AND MOLECULAR PHENOTYPES TOWARDS THERAPY	
<i>Session Introduction</i>	171
Oliver Stegle, Steven E. Brenner, Quaid Morris, Jennifer Listgarten	
<i>AMP: Assembly Matching Pursuit</i>	175
Surojit Biswas, Vladimir Jojic	
<i>Characterization of the MetaboChip in Diverse Populations from the International HapMap Project in the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) Project</i>	188
Dana C. Crawford, Robert Goodloe, Kristin Brown-Gentry, Sarah Wilson, Jamie Roberson, Niloufar B. Gillani, Marylyn D. Ritchie, Holli H. Dilks, William S. Bush	
<i>Insights into Diseases of Human Telomerase from Dynamical Modeling</i>	200
Samuel Coulbourn Flores, Christina Waldsich	
<i>Spectral Clustering Strategies for Heterogeneous Disease Data</i>	212
Grace T. Huang, Kathryn I. Cunningham, Panayiotis V. Benos, Chakra S. Chennubhotla	

Systematic Identification of Risk Factors for Alzheimer's Disease Through Shared Genetic Architecture and Electronic Medical Records.....224
Li Li, David Ruau, Rong Chen, Susan Weber, Atul J. Butte

A Correlated Meta-Analysis Strategy for Data Mining “Omic” Scans.....236
Michael A. Province, Ingrid B. Borecki

PHYLOGENOMICS AND POPULATION GENOMICS: MODELS, ALGORITHMS, AND ANALYTICAL TOOLS

Session Introduction.....247
Luay Nakhleh, Noah Rosenberg, Tandy Warnow

Inferring Optimal Species Trees Under Gene Duplication and Loss.....250
M. S. Bayzid, Siavash Mirarab, Tandy Warnow

Evaluating Variations on the STAR Algorithm for Relative Efficiency and Sample Sizes Needed to Reconstruct Species Trees.....262
James H. Degnan

The Behavior of Admixed Populations in Neighbor-Joining Inference of Population Trees.....273
Naama M. Kopelman, Lewi Stone, Olivier Gascuel, Noah A. Rosenberg

Maximum Likelihood Phylogenetic Reconstruction from High-Resolution Whole-Genome Data and a Tree of 68 Eukaryotes.....285
Yu Lin, Fei Hu, Jijun Tang, Bernard M.E. Moret

An Analytical Comparison of Multilocus Methods Under the Multispecies Coalescent: The Three-Taxon Case.....297
Sebastien Roch

POST-NGS: INTERPRETATION AND ANALYSIS OF NEXT GENERATION SEQUENCING DATA FOR BASIC AND TRANSLATIONAL SCIENCE

Session Introduction.....307
Gurkan Bebek, Mehmet Koyuturk, Thomas LaFramboise, Benjamin J. Raphael, Mark R. Chance

LSHPlace: Fast Phylogenetic Placement Using Locality-Sensitive Hashing.....310
Daniel G. Brown, Jakub Truszkowski

ChIPModule: Systematic Discovery of Transcription Factors and Their Cofactors from ChIP-seq Data.....320
Jun Ding, Xiaohui Cai, Ying Wang, Haiyan Hu, Xiaoman Li

Using BioBin to Explore Rare Variant Population Stratification.....332

Carrie B. Moore, John R. Wallace, Alex T. Frase, Sarah A. Pendergrass, Marylyn D. Ritchie

Detecting Highly Differentiated Copy-Number Variants from Pooled Population Sequencing. .344
Daniel R. Schrider, David J. Begun, Matthew W. Hahn

MetaSeq: Privacy Preserving Meta-Analysis of Sequencing-Based Association Studies.....356
Angad Pal Singh, Samreen Zafer, Itsik Pe'er

TEXT AND DATA MINING FOR BIOMEDICAL DISCOVERY

Session Introduction.....368
Graciela H. Gonzalez, Kevin Bretonnel Cohen, Casey Greene, Udo Hahn, Maricel G. Kann, Robert Leaman, Nigam Shah, Jieping Yie

Enabling High-Throughput Genotype-Phenotype Associations in the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) Project as Part of the Population Architecture Using Genomics and Epidemiology (PAGE) Study.....373
William S. Bush, Jonathan Boston, Sarah A. Pendergrass, Logan Dumitrescu, Robert Goodloe, Kristin Brown-Gentry, Sarah Wilson, Bob McClellan Jr., Eric Torstenson, Melissa A. Basford, Kylee L. Spencer, Marylyn D. Ritchie, Dana C. Crawford

ATHENA: A Tool for Meta-Dimensional Analysis Applied to Genotypes and Gene Expression Data to Predict HDL Cholesterol Levels.....385
Emily R Holzinger, Scott M. Dudek, Alex T. Frase, Ronald M. Krauss, Marisa W. Medina, Marylyn D. Ritchie

Statistical Epistasis Networks Reduce the Computational Complexity of Searching Three-Locus Genetic Models.....397
Ting Hu, Angeline S. Andrew, Margaret R. Karagas, Jason H. Moore

Evaluation of Linear Classifiers on Articles Containing Pharmacokinetic Evidence of Drug-Drug Interactions.....409
Artemy Kolchinsky, Analia Lourenco, Lang Li, Luis M. Rocha

Incorporating Expert Terminology and Disease Risk Factors into Consumer Health Vocabularies.....421
Michael Sedorff, Kevin J. Peterson, Laurie A. Nelsen, Cristian Cocos, Jennifer B. McCormick, Christopher G. Chute, Jyotishman Pathak

Detection of Protein Catalytic Sites in the Biomedical Literature.....433
Karin Verspoor, Andrew MacKinlay, Judith D. Cohn, Michael E. Wall

WORKSHOPS

Modeling cell heterogeneity: from single-cell variations to mixed cells populations.....445

Eric Batchelor, Maricel G. Kann, Teresa M. Przytycka, Benjamin J. Raphael, and Damian Wojtowicz

Computational Biology in the Cloud: Methods and New Insights from Computing at Scale.....451
Peter M. Kasson

***Computational Challenges of Mass Phenotyping*.....454**
Lawrence Hunter

***The Future of Genome-Based Medicine*.....456**
Quaid Morris, Steven E. Brenner, Jennifer Listgarten, Oliver Stegle