Pacific Symposium on Biocomputing 2014

Kohala Coast, Hawaii, USA
3-7 January 2014

Editors:

Russ B. Altman	**A. Keith Dunker**
Lawrence Hunter	**Marylyn D. Ritchie**
Teri E. Klein	**Tiffany Murray**

ISBN: 978-1-63266-343-6

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© (2014) by World Scientific Publishing Co., Inc.
All rights reserved.

Printed by Curran Associates, Inc. (2014)

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

Preface..v

CANCER PANOMICS: COMPUTATIONAL METHODS AND INFRASTRUCTURE FOR INTEGRATIVE ANALYSIS OF CANCER HIGH-THROUGHPUT "OMICS" DATA

Session Introduction..1
Søren Brunak, Francisco M. De La Vega, Gunnar Rätsch, Joshua M. Stuart

Tumor Haplotype Assembly Algorithms for Cancer Genomics..3
Derek Aguiar, Wendy S.W. Wong, Sorin Istrail

Extracting Significant Sample-Specific Cancer Mutations Using Their Protein Interactions..15
Liviu Badea

The Stream Algorithm: Computationally Efficient Ridge-Regression via Bayesian Model Averaging, and Applications to Pharmacogenomic Prediction of Cancer Cell Line Sensitivity..27
Elias Chaibub Neto, In Sock Jang, Stephen H. Friend, Adam A. Margolin

Sharing Information to Reconstruct Patient-Specific Pathways in Heterogeneous Diseases..39
Anthony Gitter, Alfredo Braunstein, Andrea Pagnani, Carlo Baldassi, Christian Borgs, Jennifer Chayes, Riccardo Zecchina, Ernest Fraenkel

Detecting Statistical Interaction Between Somatic Mutational Events and Germline Variation from Next-Generation Sequence Data..51
Hao Hu, Chad D. Huff

Systematic Assessment of Analytical Methods for Drug Sensitivity Prediction from Cancer Cell Line Data..63
In Sock Jang, Elias Chaibub Neto, Juistin Guinney, Stephen H. Friend, Adam A. Margolin

Integrative Analysis of Two Cell Lines Derived from a Non-Small-Lung Cancer Patient – A Panomics Approach..75
Oleg Mayba, Florian Gnad, Michael Peyton, Fan Zhang, Kimberly Walter, Pan Du, Melanie A. Huntley, Zhaoshi Jiang, Jinfeng Liu, Peter M. Haverty, Robert C. Gentleman, Ruiqiang Li, John D. Minna, Yingrui Li, David S. Shames, Zemin Zhang

An Integrated Approach To Blood-Based Cancer Diagnosis And Biomarker Discovery..87
Martin Renqiang Min, Salim Chowdhury, Yanjun Qi, Alex Stewart, Rachel Ostroff

Multiplex Meta-Analysis of Medulloblastoma Expression Studies with External Controls..99
Alexander A. Morgan, Matthew D. Li, Achal S. Achrol, Purvesh J. Khatri, Samuel H. Cheshier

COMPUTATIONAL APPROACHES TO DRUG REPURPOSING AND PHARMACOLOGY

Session Introduction..110
S. Joshua Swamidass, Zhiyong Lu, Pankaj Agarwal, Atul Butte

Challenges in Secondary Analysis of High Throughput Screening Data..114
Aurora S. Blucher, Shannon K. McWeeney

Drug Intervention Response Predictions with Paradigm (DIRPP) Identifies Drug Resistant Cancer Cell Lines and Pathway Mechanisms of Resistance.........125
Douglas Brubaker, Analisa Difeo, Yanwen Chen, Taylor Pearl, Kaide Zhai, Gurkan Bebek, Mark Chance, Jill Barnholtz-Sloan

Anti-Infectious Drug Repurposing Using an Integrated Chemical Genomics and Structural Systems Biology Approach.........136
Clara Ng, Ruth Hauptman, Yinliang Zhang, Philip E. Bourne, Lei Xie

Drug-Target Interaction Prediction by Integrating Chemical, Genomic, Functional and Pharmacological Data.........148
Fan Yang, Jinbo Xu, Jianyang Zeng

Prediction of Off-Target Drug Effects Through Data Fusion.........160
Emmanuel R. Yera, Ann E. Cleves, Ajay N. Jain

Exploring the Pharmacogenomics Knowledge Base (PharmGKB) for Repositioning Breast Cancer Drugs by Leveraging Web Ontology Language (OWL) and Cheminformatics Approaches.........172
Qian Zhu, Cui Tao, Feichen Shen, Christopher Chute

DETECTING AND CHARACTERIZING PLEIOTROPY: NEW METHODS FOR UNCOVERING THE CONNECTION BETWEEN THE COMPLEXITY OF GENOMIC ARCHITECTURE AND MULTIPLE PHENOTYPES

Session Introduction.........183
Anna L. Tyler, Dana C. Crawford, Sarah A. Pendergrass

Using the Bipartite Human Phenotype Network to Reveal Pleiotropy and Epistasis Beyond the Gene.........188
Christian Darabos, Samantha H. Harmon, Jason H. Moore

Environment-Wide Association Study (EWAS) for Type 2 Diabetes in the Marshfield Personalized Medicine Research Project Biobank.........200
Molly A. Hall, Scott M. Dudek, Robert Goodloe, Dana C. Crawford, Sarah A. Pendergrass, Peggy Peissig, Murray Brilliant, Catherine A. McCarty, Marylyn D. Ritchie

Dissection of Complex Gene Expression Using the Combined Analysis of Pleiotropy and Epistasis.........212
Vivek M. Philip, Anna L. Tyler, Gregory W. Carter

PERSONALIZED MEDICINE: FROM GENOTYPES AND MOLECULAR PHENOTYPES TOWARDS THERAPY

Session Introduction.........224
Jennifer Listgarten, Oliver Stegle, Quaid Morris, Steven E. Brenner, Leopold Parts

PATH-SCAN: A Reporting Tool for Identifying Clinically Actionable Variants.........229
Roxana Daneshjou, Zachary Zappala, Kim Kukurba, Sean M Boyle, Kelly E Ormond, Teri E Klein, Michael Snyder, Carlos D Bustamante, Russ B Altman, Stephen B Montgomery

Imputation-based Assessment of Next Generation Rare Exome Variant Arrays.......................................241
Alicia R. Martin, Gerard Tse, Carlos D. Bustamante, Eimear E. Kenny

Utilization of an EMR-Biorepository to Identify the Genetic Predictors of Calcineurin-Inhibitor Toxicity in Heart Transplant Recipients...253
Matthew Oetjens, William S. Bush, Kelly A. Birdwell, Holli H. Dilks, Erica A. Bowton, Joshua C. Denny, Russell A. Wilke, Dan M. Roden, Dana C. Crawford

Robust Reverse Engineering of Dynamic Gene Networks Under Sample Size Heterogeneity..................265
Ankur P. Parikh, Wei Wu, Eric P. Xing

Variant Priorization and Analysis Incorporating Problematic Regions of the Genome............................277
Anil Patwardhan, Michael Clark, Alex Morgan, Stephen Chervitz, Mark Pratt, Gabor Bartha, Gemma Chandratillake, Sarah Garcia, Nan Leng, Richard Chen

Bags of Words Models of Epitope Sets: HIV Viral Load Regression with Counting Grids........................288
Alessandro Perina, Pietro Lovato, Nebojsa Jojic

Joint Association Discovery and Diagnosis of Alzheimer's Disease by Supervised Heterogeneous Multiview Learning...300
Shandian Zhe, Zenglin Xu, Yuan Qi, Peng Yu

TEXT AND DATA MINING FOR BIOMEDICAL DISCOVERY

Session Introduction..312
Graciela H. González, Kevin Bretonnel Cohen, Robert Leaman, Casey Greene, Nigam Shah, Maricel G. Kann, Jieping Ye

Vector Quantization Kernels for the Classification of Protein Sequences and Structures........................316
Wyatt T. Clark, Predrag Radivojac

Combining Heterogenous Data for Prediction of Disease Related and Pharmacogenes.........................328
Christopher S. Funk, Lawrence E. Hunter, K. Bretonnel Cohen

A Novel Profile Biomarker Diagnosis for Mass Spectral Proteomics..340
Henry Han

Towards Pathway Curation Through Literature Mining – A Case Study Using PharmGKB...................352
Ravikumar K.E., Kavishwar B. Wagholikar, Hongfang Liu

Sparse Generalized Functional Linear Model for Predicting Remission Status of Depression Patients..364
Yashu Liu, Zhi Nie, Jiayu Zhou, Michael Farnum, Vaibhav A Narayan, Gayle Wiittenberg, Jieping Ye

Development of a Data-Mining Algorithm to Identify Ages at Reproductive Milestones in Electronic Medical Records..376
Jennifer Malinowski, Eric Farber-Eger, Dana C. Crawford

An Efficient Algorithm to Integrate Network and Attribute Data for Gene Function Prediction.............388
Shankar Vembu, Quaid Morris

Matrix Factorization-Based Data Fusion for Gene Function Prediction in Baker's Yeast and Slime Mold....................400
Marinka Zitnik, Blaz Zupan

WORKSHOPS

Applications of Bioinformatics to Non-Coding RNAs in the Era of Next-Generation Sequencing...........412
Chao Cheng, Jason Moore, Casey Greene

Building the Next Generation of Quantitative Biologists....................417
Kristine A. Pattin, Anna C. Greene, Russ B. Altman, Lawrence E. Hunter, David A. Ross, James A. Foster, Jason H. Moore

Uncovering the Etiology of Autism Spectrum Disorders: Genomics, Bioinformatics, Environment, Data Collection and Exploration, and Future Possibilities....................422
Sarah A. Pendergrass, Santhosh Girirajan, Scott Selleck