COMPUTATIONAL APPROACHES TO UNDERSTANDING THE EVOLUTION OF MOLECULAR FUNCTION

Session Introduction......................................................................................................................................................... 1
Yana Bromberg, Matthew W. Hahn, Predrag Radivojac

Identification and Analysis of Bacterial Genomic Metabolic Signatures................................................................. 3
Nathan Bowerman, Nathan Tintle, Matthew DeJongh, Aaron A. Best

When should we NOT transfer functional annotation between sequence paralogs? ........................................... 15
Mengfei Cao, Lenore J. Cowen

ProSNet: integrating homology with molecular networks for protein function prediction........................................... 27
Sheng Wang, Meng Qu, Jian Peng

On the power and limits of sequence similarity based clustering of proteins into families .................................. 39
Christian Wiwie, Richard Röttger

IMAGING GENOMICS

Session Introduction.......................................................................................................................................................... 51
Li Shen, Lee A.D. Cooper

Adaptive testing of SNP-brain functional connectivity association via a modular network analysis ........... 58
Chen Gao, Junghi Kim, Wei Pan

Exploring Brain Transcriptomic Patterns: A Topological Analysis Using Spatial Expression Networks........... 70
Zhana Kuncheva, Michelle L. Krishnan, Giovanni Montana

Integrative Analysis for Lung Adenocarcinoma Predicts Morphological Features Associated with Genetic Variations................................................................................................................................................ 82
Chao Wang, Hai Su, Lin Yang, Kun Huang

Identification of Discriminative Imaging Proteomics Associations in Alzheimer's Disease via a Novel Sparse Correlation Model............................................................................................................. 94
Jingwen Yan, Shannon L. Risacher, Kwangsik Nho, Andrew J. Saykin, Li Shen

Enforcing Co-expression in Multimodal Regression Framework ................................................................................. 105
Pascal Zille, Vince D. Calhoun, Yu-Ping Wang

METHODS TO ENSURE THE REPRODUCIBILITY OF BIOMEDICAL RESEARCH

Session Introduction......................................................................................................................................................... 117
Konrad J. Karczewski, Nicholas P. Tatonetti, Chirag J. Patel, Arjun K. Manrai, C. Titus Brown,
John P.A. Ioannidis

Exploring the Reproducibility of Probabilistic Causal Molecular Network Models .......................................... 120
Ariella Cohain, Aparna A. Divaraniya, Kuixi Zhu, Joseph R. Scarpa, Andrew Kasarskis, Jun Zhu, Rui Chang, Joel T. Dudley, Eric E. Schadt

Reproducible Drug Repurposing: When Similarity Does Not Suffice ............................................................................. 132

Emre Guney

Empowering Multi-Cohort Gene Expression Analysis to Increase Reproducibility .............................................................. 144

Winston A. Haynes, Francesco Vallania, Charles Liu, Erika Bongen, Aurelie Tomczak, Marta Andres-Terrè, Shane Lofgren, Andrew Tam, Cole A. Deisseroth, Matthew D. Li, Timothy E.Sweeney, and Purvesh Khatri

Rabix: An Open-Source Workflow Executor Supporting Recomputability and Interoperability of Workflow Descriptions .......................................................................................................................... 154

Gaurav Kaushik, Sinisa Ivkovic, Janko Simonovic, Nebojsa Tijanic, Brandi Davis-Dusenbery, Deniz Kural

Data Sharing and Reproducible Clinical Genetic Testing: Successes and Challenges ............................................................ 166

Shan Yang, Melissa Cline, Can Zhang, Benedict Paten and Stephen E. Lincoln

PATTERNS IN BIOMEDICAL DATA—HOW DO WE FIND THEM?

Session Introduction ............................................................................................................................................................................. 177

Anna Okula Basile, Anurag Verma, Marta Byrska-Bishop, Sarah P. Pendergrass, Christian Darabos, H. Lester Kirchner

Learning Attributes of Disease Progression from Trajectories of Sparse Lab Values .............................................................. 184

Vibhu Agarwal, Nigam H. Shah

Computer Aided Image Segmentation and Classification for Viable and Non-Viable Tumor Identification in Osteosarcoma .................................................................................................................................... 195

Harish Babu Arunachalam, Rashika Mishra, Bogdan Armaselu, Ovidiu Daescu, Maria Martinez, Patrick Leavey, Dinesh Rakheja, Kevin Cederberg, Anita Sengupta, Molly Ni'suilleabhain

Missing Data Imputation in the Electronic Health Record Using Deeply Learned Autoencoders ........................................... 207


A Deep Learning Approach for Cancer Detection and Relevant Gene Identification ............................................................. 219

Padideh Danaee, Reza Ghaeini, David Hendrix

Development and Performance of Text-Mining Algorithms to Extract Socioeconomic Status from De-Identified Electronic Health Records ......................................................................................... 230

Brittany M. Hollister, Nicole A. Restrepo, Eric Farber-Eger, Dana C. Crawford, Melinda C. Aldrich, Amy Non

Genome-Wide Interaction with Selected Type 2 Diabetes Loci Reveals Novel Loci for Type 2 Diabetes in African Americans .................................................................................................................... 242

Jacob M. Keaton, Jacklyn N. Hellwege, Maggie C. Y. Ng, Nicholette D. Palmer, James S. Pankow, Myriam Fornage, James G. Wilson, Adolfo Correa, Laura J. Rasmussen-Torvik, Jerome

DeMo Dashboard: Visualizing and Understanding Genomic Sequences Using Deep Neural Networks ................................................................................................................. 254
Jack Lanchantin, Ritambhara Singh, Beilun Wang, Yanjun Qi

Meta-Analysis of Continuous Phenotypes Identifies a Gene Signature that Correlates with COPD Disease Status .............................................................................................. 266
Madeleine Scott, Francesco Vallania, and Purvesh Khatri

Predictive Modeling of Hospital Readmission Rates Using Electronic Medical Record-Wide Machine Learning: A Case-Study Using Mount Sinai Heart Failure Cohort ......................................................................................................................... 276
Khader Shameer, Kipp W. Johnson, Alexandre Yahi, Riccardo Miotto, Li Li, Doran Ricks, Jebakumar Jebakaran, Patricia Kovatch, Partho P. Sengupta, Annetteine Gelijns, Alan Moskovitz, Bruce Darrow, David L. Reich, Andrew Kasarskis, Nicholas P. Tatonetti, Sean Pinney, Joel T. Dudley

Learning Parsimonious Ensembles for Unbalanced Computational Genomics Problems ......................................................................................................................... 288
Ana Stanescu, Gaurav Pandey

Methods for Clustering Time Series Data Acquired from Mobile Health Apps ................................................................................................................................. 300
Nicole Tignor, Pei Wang, Nicholas Genes, Linda Rogers, Steven G. Hershman, Erick R. Scott, Micol Zweig, Yu-Feng Yvonne Chan, Eric E. Schadt

A New Relevance Estimator for the Compilation and Visualization of Disease Patterns and Potential Drug Targets ................................................................................ 312
Modest von Korff, Tobias Fink, Thomas Sander

Network Map of Adverse Health Effects Among Victims of Intimate Partner Violence ......................................................................................................................... 324
Kathleen Whiting, Larry Y. Liu, Mehmet Koyuturk, Gunnur Karakurt

Discovery of Functional and Disease Pathways by Community Detection in Protein-Protein Interaction Networks ................................................................................................................. 336
Stephen J. Wilson, Angela D. Wilkins, Chih-Hsu Lin, Rhonald C. Lua, Olivier Lichtarge

PRECISION MEDICINE: DATA AND DISCOVERY FOR IMPROVED HEALTH AND THERAPY

Session Introduction ......................................................................................................................... 348
Alexander A. Morgan, Dana C. Crawford, Josh C. Denny, Sean D. Mooney, Bruce J. Aronow, Steven E. Brenner

Opening the Door to the Large Scale Use of Clinical Lab Measures for Association Testing: Exploring Different Methods for Defining Phenotypes ................................................................................................................. 356
Christopher R. Bauer, Daniel Lavage, John Snyder, Joseph Leader, J. Matthew Mahoney, Sarah A. Pendergrass

A Powerful Method for Including Genotype Uncertainty in Tests of Hardy-Weinberg Equilibrium ......................................................................................................................... 368
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Order of Disease Pairs Affects Subsequent Disease Trajectories: The Case of Diabetes and Sleep Apnea</td>
<td>380</td>
</tr>
<tr>
<td>Mette Beck, David Westergaard, Leif Groop and Soren Brunak</td>
<td></td>
</tr>
<tr>
<td>MicroRNA-Augmented Pathways (mirAP) and Their Applications to Pathway Analysis and Disease Subtyping</td>
<td>390</td>
</tr>
<tr>
<td>Diana Diaz, Michele Donato, Tin Nguyen, Sorin Draghici</td>
<td></td>
</tr>
<tr>
<td>Frequent Subgraph Mining of Personalized Signaling Pathway Networks Groups Patients with Frequently Dysregulated Disease Pathways and Predicts Prognosis</td>
<td>402</td>
</tr>
<tr>
<td>Arda Durmaz, Tim A.D. Henderson, Douglas Brubaker, Gurkan Bebek</td>
<td></td>
</tr>
<tr>
<td>Human Kinases Display Mutational Hotspots at Cognate Positions Within Cancer</td>
<td>414</td>
</tr>
<tr>
<td>Jonathan Gallion, Angela D. Wilkins, Olivier Lichtarge</td>
<td></td>
</tr>
<tr>
<td>MUSE: A Multi-locus Sampling-based Epistasis Algorithm for Quantitative Genetic Trait Prediction</td>
<td>426</td>
</tr>
<tr>
<td>Dan He, Laxmi Parida</td>
<td></td>
</tr>
<tr>
<td>ceRNA Search Method Identified a MET-activated Subgroup Among EGFR DNA Amplified Lung Adenocarcinoma Patients</td>
<td>438</td>
</tr>
<tr>
<td>Halla Kabat, Leo Tunkle, Inhan Lee</td>
<td></td>
</tr>
<tr>
<td>Improved Performance of Gene Set Analysis on Genome-Wide Transcriptomics Data when Using Gene Activity State Estimates</td>
<td>449</td>
</tr>
<tr>
<td>Thomas Kamp, Micah Adams, Craig Disselkoen, Nathan Tintle</td>
<td></td>
</tr>
<tr>
<td>methylDMV: Simultaneous Detection of Differential DNA Methylation and Variability with Confounder Adjustment</td>
<td>461</td>
</tr>
<tr>
<td>Pei Fen Kuan, Junyan Song, Shuyao He</td>
<td></td>
</tr>
<tr>
<td>Identify Cancer Driver Genes Through Shared Mendelian Disease Pathogenic Variants and Cancer Somatic Mutations</td>
<td>473</td>
</tr>
<tr>
<td>Meng Ma, Changchang Wang, Benjamin Glicksberg, Eric Schadt, Shuyu D. Li, Rong Chen</td>
<td></td>
</tr>
<tr>
<td>Identifying Cancer Specific Metabolic Signatures Using Constraint-Based Models</td>
<td>485</td>
</tr>
<tr>
<td>André Schultz, Sanket Mehta, Chenyue W. Hu, Fieke W Hoff, Terzah M. Horton, Steven M. Kornblau, Amina A. Qutub</td>
<td></td>
</tr>
<tr>
<td>Differential Pathway Dependency Discovery Associated with Drug Response across Cancer Cell Lines</td>
<td>497</td>
</tr>
<tr>
<td>Gil Speyer, Divya Mahendra, Hai J. Tran, Jeff Kiefer, Stuart L. Schreiber, Paul A. Clemons, Harshil Dhruv, Michael Berens, Seungchan Kim</td>
<td></td>
</tr>
<tr>
<td>A Methylation-to-Expression Feature Model for Generating Accurate Prognostic Risk Scores and Identifying Disease Targets in Clear Cell Kidney Cancer</td>
<td>509</td>
</tr>
<tr>
<td>Jeffrey A. Thompson, Carmen J. Marsit</td>
<td></td>
</tr>
<tr>
<td>De Novo Mutations in Autism Implicate the Synaptic Elimination Network</td>
<td>521</td>
</tr>
</tbody>
</table>
Identifying Genetic Associations with Variability in Metabolic Health and Blood Count Laboratory Values: Diving into the Quantitative Traits by Leveraging Longitudinal Data from an EHR

Strategies for Equitable Pharmacogenomic-Guided Warfarin Dosing Among European and African American Individuals in a Clinical Population

SINGLE-CELL ANALYSIS AND MODELLING OF CELL POPULATION HETEROGENEITY

Production of a Preliminary Quality Control Pipeline for Single Nuclei RNA-Seq and Its Application in the Analysis of Cell Type Diversity of Post-Mortem Human Brain Neocortex

Tracing Co-Regulatory Network Dynamics in Noisy, Single-Cell Transcriptome Trajectories

An Updated Debarcoding Tool for Mass Cytometry with Cell Type-Specific and Cell Sample-Specific Stringency Adjustment

Mapping Neuronal Cell Types Using Integrative Multi-Species Modeling of Human and Mouse Single Cell RNA Sequencing

A Spatiotemporal Model to Simulate Chemotherapy Regimens for Heterogeneous Bladder Cancer Metastases to the Lung

Scalable Visualization for High-dimensional Single-cell Data

WORKSHOPS

Harnessing Big Data for Precision Medicine: Infrastructures and Applications

Kun-Hsing Yu, Steven N. Hart, Rachel Goldfeder, Qiangfeng Cliff Zhang, Stephen C. J. Parker, Michael Snyder
The Training of Next Generation Data Scientists in Biomedicine.................................................................640
   Lana Garmire, Stephen Gliske, Quynh C Nguyen, Jonathan H. Chen, Shamim Nemati, John D.
   Van Horn, Jason H. Moore, Carol Shreffler, Michelle Dunn

No-Boundary Thinking in Bioinformatics........................................................................................................646
   Jason H. Moore, Steven F. Jennings, Casey S. Greene, Lawrence E. Hunter, Andy D. Perkins,
   Clarlynda Williams-Devane, Donald C. Wunsch, Zhongming Zhao, Xiuzhen Huang

Open Data for Discovery Science.......................................................................................................................649
   Philip R.O. Payne, Kun Huang, Nigam H. Shah, Jessica Tenenbaum