

Pacific Symposium on Biocomputing 2017

Fairmont Orchid, Hawaii, USA
3 - 7 January 2017

ISBN: 978-1-7138-2793-1

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© (2017) by World Scientific Publishing Co., Inc.
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

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 07601

Phone: 1-201-487-9655
Fax: 1-201-487-9656

www.worldscientific.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

TABLE OF CONTENTS

COMPUTATIONAL APPROACHES TO UNDERSTANDING THE EVOLUTION OF MOLECULAR FUNCTION	1
<i>Y. Bromberg, M.W. Hahn, P. Radivojac</i>	
IDENTIFICATION AND ANALYSIS OF BACTERIAL GENOMIC METABOLIC SIGNATURES	4
<i>N. Bowerman, N. Tintle, M. Dejongh, A.A. Best</i>	
WHEN SHOULD WE NOT TRANSFER FUNCTIONAL ANNOTATION BETWEEN SEQUENCE PARALOGS?	15
<i>M. Cao, L.J. Cowen</i>	
PROSNET: INTEGRATING HOMOLOGY WITH MOLECULAR NETWORKS FOR PROTEIN FUNCTION PREDICTION	27
<i>S. Wang, M. Qu, J. Peng</i>	
ON THE POWER AND LIMITS OF SEQUENCE SIMILARITY BASED CLUSTERING OF PROTEINS INTO FAMILIES	39
<i>C. Wiwie, R. Rottger</i>	
IMAGING GENOMICS	51
<i>L. Shen, L.A.D. Cooper</i>	
ADAPTIVE TESTING OF SNP-BRAIN FUNCTIONAL CONNECTIVITY ASSOCIATION VIA A MODULAR NETWORK ANALYSIS	59
<i>C. Gao, J. Kim, W. Pan</i>	
EXPLORING BRAIN TRANSCRIPTOMIC PATTERNS: A TOPOLOGICAL ANALYSIS USING SPATIAL EXPRESSION NETWORKS	71
<i>Z. Kuncheva, M.L. Krishnan, G. Montana</i>	
INTEGRATIVE ANALYSIS FOR LUNG ADENOCARCINOMA PREDICTS MORPHOLOGICAL FEATURES ASSOCIATED WITH GENETIC VARIATIONS	83
<i>C. Wang, H. Su, L. Yang, K. Huang</i>	
IDENTIFICATION OF DISCRIMINATIVE IMAGING PROTEOMICS ASSOCIATIONS IN ALZHEIMER'S DISEASE VIA A NOVEL SPARSE CORRELATION MODEL.....	95
<i>J. Yan, S.L. Risacher, K. Nho, A.J. Saykin, L. Shen</i>	
ENFORCING CO-EXPRESSION IN MULTIMODAL REGRESSION FRAMEWORK.....	106
<i>P. Zille, V.D. Calhoun, Y.P. Wang</i>	
METHODS TO ENSURE THE REPRODUCIBILITY OF BIOMEDICAL RESEARCH	118
<i>K.J. Karczewski, N.P. Tatonetti, A. Manrai, C.J. Patel, C.T. Brown, J.P.A. Ioannidis</i>	
EXPLORING THE REPRODUCIBILITY OF PROBABILISTIC CAUSAL MOLECULAR NETWORK MODELS.....	121
<i>A. Cohain, A.A. Divaraniya, K. Zhu, J.R. Scarpa, A. Kasarskis, J. Zhu, R. Chang, J.T. Dudley, E.E. Schadt</i>	
REPRODUCIBLE DRUG REPURPOSING: WHEN SIMILARITY DOES NOT SUFFICE	133
<i>E. Guney</i>	

EMPOWERING MULTI-COHORT GENE EXPRESSION ANALYSIS TO INCREASE REPRODUCIBILITY	145
<i>W.A. Haynes, F. Vallania, C. Liu, E. Bongen, A. Tomczak, M. Andres-Terre, S. Lofgren, A. Tam, C.A. Deisseroth, M.D. Li, T.E. Sweeney, P. Khatri</i>	
RABIX: AN OPEN-SOURCE WORKFLOW EXECUTOR SUPPORTING RECOMPUTABILITY AND INTEROPERABILITY OF WORKFLOW DESCRIPTIONS.....	155
<i>G. Kaushik, S. Ivkovic, J. Simonovic, N. Tijanic, B. Davis-Dusenbery, D. Kural</i>	
DATA SHARING AND REPRODUCIBLE CLINICAL GENETIC TESTING: SUCCESSES AND CHALLENGES	167
<i>S. Yang, M. Cline, C. Zhang, B. Paten, S.E. Lincoln</i>	
PATTERNS IN BIOMEDICAL DATA-HOW DO WE FIND THEM?.....	178
<i>A.O. Basile, A. Verma, M. Byrska-Bishop, S.A. Pendergrass, C. Darabos, H.L. Kirchner</i>	
LEARNING ATTRIBUTES OF DISEASE PROGRESSION FROM TRAJECTORIES OF SPARSE LAB VALUES.....	185
<i>V. Agarwal, N.H. Shah</i>	
COMPUTER AIDED IMAGE SEGMENTATION AND CLASSIFICATION FOR VIABLE AND NON-VIABLE TUMOR IDENTIFICATION IN OSTEOSARCOMA.....	196
<i>H.B. Arunachalam, R. Mishra, B. Armaselu, O. Daescu, M. Martinez, P. Leavey, D. Rakheja, K. Cederberg, A. Sengupta, M. Ni'Suilleabhain</i>	
MISSING DATA IMPUTATION IN THE ELECTRONIC HEALTH RECORD USING DEEPLY LEARNED AUTOENCODERS	208
<i>B.K. Beaulieu-Jones, J.H. Moore</i>	
A DEEP LEARNING APPROACH FOR CANCER DETECTION AND RELEVANT GENE IDENTIFICATION	220
<i>P. Danaee, R. Ghaeini, D.A. Hendrix</i>	
DEVELOPMENT AND PERFORMANCE OF TEXT-MINING ALGORITHMS TO EXTRACT SOCIOECONOMIC STATUS FROM DE-IDENTIFIED ELECTRONIC HEALTH RECORDS	231
<i>B.M. Hollister, N.A. Restrepo, E. Farber-Eger, D.C. Crawford, M.C. Aldrich, A. Non</i>	
GENOME-WIDE INTERACTION WITH SELECTED TYPE 2 DIABETES LOCI REVEALS NOVEL LOCI FOR TYPE 2 DIABETES IN AFRICAN AMERICANS	243
<i>J.M. Keaton, J. Hellwege, M.C.Y. Ng, N.D. Palmer, J.S. Pankow, M. Fornage, J.G. Wilson, A. Correa, L.J. Rasmussen-Torvik, J.I. Rotter, Y.D.I. Chen, K.D. Taylor, S.S. Rich, L.E. Wagenknecht, B.I. Freedman, D.W. Bowden</i>	
DEEP MOTIF DASHBOARD: VISUALIZING AND UNDERSTANDING GENOMIC SEQUENCES USING DEEP NEURAL NETWORKS.....	255
<i>J. Lanchantin, R. Singh, B. Wang, Y. Qi</i>	
META-ANALYSIS OF CONTINUOUS PHENOTYPES IDENTIFIES A GENE SIGNATURE THAT CORRELATES WITH COPD DISEASE STATUS	267
<i>M. Scott, F. Vallania, P. Khatri</i>	
PREDICTIVE MODELING OF HOSPITAL READMISSION RATES USING ELECTRONIC MEDICAL RECORD-WIDE MACHINE LEARNING: A CASE-STUDY USING MOUNT SINAI HEART FAILURE COHORT.....	277
<i>K. Shameer, K.W. Johnson, A. Yahi, R. Miotto, L. Li, D. Ricks, J. Jebakaran, P. Kovatch, P.P. Sengupta, A. Gelijns, A. Moskovitz, B. Darrow, D.L. Reich, A. Kasarskis, N.P. Tatonetti, S. Pinney, J.T. Dudley</i>	

LEARNING PARSIMONIOUS ENSEMBLES FOR UNBALANCED COMPUTATIONAL GENOMICS PROBLEMS	289
<i>A. Stanescu, G. Pandey</i>	
METHODS FOR CLUSTERING TIME SERIES DATA ACQUIRED FROM MOBILE HEALTH APPS	301
<i>N. Tignor, P. Wang, N. Genes, L. Rogers, S.G. Hershman, E.R. Scott, M. Zweig, Y.F.Y. Chan, E.E. Schadt</i>	
A NEW RELEVANCE ESTIMATOR FOR THE COMPILATION AND VISUALIZATION OF DISEASE PATTERNS AND POTENTIAL DRUG TARGETS	313
<i>M. Von Korff, T. Fink, T. Sander</i>	
NETWORK MAP OF ADVERSE HEALTH EFFECTS AMONG VICTIMS OF INTIMATE PARTNER VIOLENCE	325
<i>K. Whiting, L.Y. Liu, M. Koyuturk, G. Karakurt</i>	
DISCOVERY OF FUNCTIONAL AND DISEASE PATHWAYS BY COMMUNITY DETECTION IN PROTEIN-PROTEIN INTERACTION NETWORKS	337
<i>S.J. Wilson, A.D. Wilkins, C.H. Lin, R.C. Lua, O. Lichtarge</i>	
PRECISION MEDICINE: DATA AND DISCOVERY FOR IMPROVED HEALTH AND THERAPY	349
<i>A.A. Morgan, D.C. Crawford, J.C. Denny, S.D. Mooney, B.J. Aronow, S.E. Brenner</i>	
OPENING THE DOOR TO THE LARGE SCALE USE OF CLINICAL LAB MEASURES FOR ASSOCIATION TESTING: EXPLORING DIFFERENT METHODS FOR DEFINING PHENOTYPES	357
<i>C.R. Bauer, D. Lavage, J. Snyder, J. Leader, J.M. Mahoney, S.A. Pendergrass</i>	
A POWERFUL METHOD FOR INCLUDING GENOTYPE UNCERTAINTY IN TESTS OF HARDY-WEINBERG EQUILIBRIUM	369
<i>A. Beck, A. Luedtke, K. Liu, N. Tintle</i>	
TEMPORAL ORDER OF DISEASE PAIRS AFFECTS SUBSEQUENT DISEASE TRAJECTORIES: THE CASE OF DIABETES AND SLEEP APNEA	381
<i>M.K. Beck, D. Westergaard, A.B. Jensen, L. Groop, S. Brunak</i>	
MICRORNA-AUGMENTED PATHWAYS (MIRAP) AND THEIR APPLICATIONS TO PATHWAY ANALYSIS AND DISEASE SUBTYPING	391
<i>D. Diaz, M. Donato, T. Nguyen, S. Draghici</i>	
FREQUENT SUBGRAPH MINING OF PERSONALIZED SIGNALING PATHWAY NETWORKS GROUPS PATIENTS WITH FREQUENTLY DYSREGULATED DISEASE PATHWAYS AND PREDICTS PROGNOSIS	403
<i>A. Durmaz, T.A.D. Henderson, D. Brubaker, G. Bebek</i>	
HUMAN KINASES DISPLAY MUTATIONAL HOTSPOTS AT COGNATE POSITIONS WITHIN CANCER	415
<i>J. Gallion, A.D. Wilkins, O. Lichtarge</i>	
MUSE: A MULTI-LOCUS SAMPLING-BASED EPISTASIS ALGORITHM FOR QUANTITATIVE GENETIC TRAIT PREDICTION	427
<i>D. He, L. Parida</i>	

CERNA SEARCH METHOD IDENTIFIED A MET-ACTIVATED SUBGROUP AMONG EGFR DNA AMPLIFIED LUNG ADENOCARCINOMA PATIENTS	439
<i>H. Kabat, L. Tunkle, I. Lee</i>	
IMPROVED PERFORMANCE OF GENE SET ANALYSIS ON GENOME-WIDE TRANSCRIPTOMICS DATA WHEN USING GENE ACTIVITY STATE ESTIMATES	450
<i>T. Kamp, M. Adams, C. Disselkoen, N. Tintle</i>	
METHYLDMMV: SIMULTANEOUS DETECTION OF DIFFERENTIAL DNA METHYLATION AND VARIABILITY WITH CONFOUNDER ADJUSTMENT	462
<i>P.F. Kuan, J. Song, S. He</i>	
IDENTIFY CANCER DRIVER GENES THROUGH SHARED MENDELIAN DISEASE PATHOGENIC VARIANTS AND CANCER SOMATIC MUTATIONS.....	474
<i>M. Ma, C. Wang, B.S. Glicksberg, E.E. Schadt, S.D. Li, R. Chen</i>	
IDENTIFYING CANCER SPECIFIC METABOLIC SIGNATURES USING CONSTRAINT- BASED MODELS	486
<i>A. Schultz, S. Mehta, C.W. Hu, F.W. Hoff, T.M. Horton, S.M. Kornblau, A.A. Qutub</i>	
DIFFERENTIAL PATHWAY DEPENDENCY DISCOVERY ASSOCIATED WITH DRUG RESPONSE ACROSS CANCER CELL LINES	498
<i>G. Speyer, D. Mahendra, H.J. Tran, J. Kiefer, S.L. Schreiber, P.A. Clemons, H. Dhruv, M. Berens, S. Kim</i>	
A METHYLATION-TO-EXPRESSION FEATURE MODEL FOR GENERATING ACCURATE PROGNOSTIC RISK SCORES AND IDENTIFYING DISEASE TARGETS IN CLEAR CELL KIDNEY CANCER.....	510
<i>J.A. Thompson, C.J. Marsit</i>	
DE NOVO MUTATIONS IN AUTISM IMPLICATE THE SYNAPTIC ELIMINATION NETWORK.....	522
<i>G.R. Venkataraman, C. O'Connell, F. Egawa, D. Kashaf-Haghighi, D.P. Wall</i>	
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	534
<i>S.S. Verma, A.M. Lucas, D.R. Lavage, J.B. Leader, R. Metpally, S. Krishnamurthy, F. Dewey, I. Borecki, A. Lopez, J. Overton, J. Penn, J. Reid, S.A. Pendergrass, G. Breitwieser, M.D. Ritchie</i>	
STRATEGIES FOR EQUITABLE PHARMACOGENOMIC-GUIDED WARFARIN DOSING AMONG EUROPEAN AND AFRICAN AMERICAN INDIVIDUALS IN A CLINICAL POPULATION.....	546
<i>L.K. Wiley, J.P. VanHouten, D.C. Samuels, M.C. Aldrich, D.M. Roden, J.F. Peterson, J.C. Denny</i>	
SINGLE-CELL ANALYSIS AND MODELLING OF CELL POPULATION HETEROGENEITY	558
<i>N. Samusik, N. Aghaeepour, S. Bendall</i>	
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	566
<i>B. Aevermann, J. McCarrison, P. Venepally, R. Hodge, T. Bakken, J. Miller, M. Novotny, D.N. Tran, F. Diez-Fuertes, L. Christiansen, F. Zhang, F. Steemers, R.S. Lasken, E. Lein, N. Schork, R.H. Scheuermann</i>	

TRACING CO-REGULATORY NETWORK DYNAMICS IN NOISY, SINGLE-CELL TRANSCRIPTOME TRAJECTORIES	578
<i>P. Cordero, J.M. Stuart</i>	
AN UPDATED DEBARCODING TOOL FOR MASS CYTOMETRY WITH CELL TYPE- SPECIFIC AND CELL SAMPLE-SPECIFIC STRINGENCY ADJUSTMENT	590
<i>K.I. Fread, W.D. Strickland, G.P. Nolan, E.R. Zunder</i>	
MAPPING NEURONAL CELL TYPES USING INTEGRATIVE MULTI-SPECIES MODELING OF HUMAN AND MOUSE SINGLE CELL RNA SEQUENCING.....	601
<i>T. Johnson, Z. Abrams, Y. Zhang, K. Huang</i>	
A SPATIOTEMPORAL MODEL TO SIMULATE CHEMOTHERAPY REGIMENS FOR HETEROGENEOUS BLADDER CANCER METASTASES TO THE LUNG	613
<i>K.R.K. Winner, J.C. Costello</i>	
SCALABLE VISUALIZATION FOR HIGH-DIMENSIONAL SINGLE-CELL DATA.....	625
<i>J. Kim, N. Russell, J. Peng</i>	
HARNESSING BIG DATA FOR PRECISION MEDICINE: INFRASTRUCTURES AND APPLICATIONS.....	637
<i>K.H. Yu, S.N. Hart, R. Goldfeder, Q.C. Zhang, S.C.J. Parker, M. Snyder</i>	
THE TRAINING OF NEXT GENERATION DATA SCIENTISTS IN BIOMEDICINE	642
<i>L.X. Garmire, S. Gliske, Q.C. Nguyen, J.H. Chen, S. Nemati, J.D. Van Horn, J.H. Moore, C. Shreffler, M. Dunn</i>	
NO-BOUNDARY THINKING IN BIOINFORMATICS	648
<i>J.H. Moore, S.F. Jennings, C.S. Greene, L.E. Hunter, A.D. Perkins, C. Williams-Devane, D.C. Wunsch, Z. Zhao, X. Huang</i>	
OPEN DATA FOR DISCOVERY SCIENCE	651
<i>P.R.O. Payne, K. Huang, N.H. Shah, J. Tenenbaum</i>	

Author Index