

# **Pacific Symposium on Biocomputing 2022**

Big Island, Hawaii, USA  
3 - 7 January 2022

ISBN: 978-1-7138-4498-3

**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.**

The full collection of papers in these proceedings is licensed under a Creative Commons Attribution 4.0 International License.

License details: <http://creativecommons.org/licenses/by/4.0/>.

The individual papers are © 2020 The Authors. Distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License

License details: <https://creativecommons.org/licenses/by-nc/4.0/>.

Printed with permission by Curran Associates, Inc. (2022)

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](http://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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## **AI-DRIVEN ADVANCES IN MODELING OF PROTEIN STRUCTURE**

SESSION INTRODUCTION.....	1
<i>Krzysztof Fidelis, Sergei Grudinin</i>	
TRAINING DATA COMPOSITION AFFECTS PERFORMANCE OF PROTEIN STRUCTURE ANALYSIS ALGORITHMS .....	10
<i>Alexander Derry, Kristy A. Carpenter, Russ B. Altman</i>	
TRANSFERABILITY OF GEOMETRIC PATTERNS FROM PROTEIN SELF-INTERACTIONS TO PROTEIN-LIGAND INTERACTIONS .....	22
<i>Antoine Koehl, Milind Jagota, Dan D. Erdmann-Pham, Alexander Fung, Yun S. Song</i>	
INTERPRETING POTTS AND TRANSFORMER PROTEIN MODELS THROUGH THE LENS OF SIMPLIFIED ATTENTION .....	34
<i>Nicholas Bhattacharya, Neil Thomas, Roshan Rao, Justas Dauparas, Peter K. Koo, David Baker, Yun S. Song, Sergey Ovchinnikov</i>	
SIDE CHAIN PACKING USING SE(3)-TRANSFORMER.....	46
<i>Akhil Jindal, Yimin Zhu, Sergey Kotelnikov, Rezaul Chowdhury, Sandor Vajda, Dzmitry Padhorny, Dima Kozakov</i>	
DEEPVASP-E: A FLEXIBLE ANALYSIS OF ELECTROSTATIC ISOPOTENTIALS FOR FINDING AND EXPLAINING MECHANISMS THAT CONTROL BINDING SPECIFICITY .....	56
<i>Felix M. Quintana, Zhaoming Kong, Lifang He, Brian Y. Chen</i>	

## **BIG DATA IMAGING GENOMICS**

SESSION INTRODUCTION.....	68
<i>Peter Kochunov, Li Shen, John Darrell van Horn, Paul M. Thompson</i>	
A NEW MENDELIAN RANDOMIZATION METHOD TO ESTIMATE CAUSAL EFFECTS OF MULTIVARIABLE BRAIN IMAGING EXPOSURES.....	73
<i>Chen Mo, Zhenyao Ye, Hongjie Ke, Tong Lu, Travis Canida, Song Liu, Qiong Wu, Zhiwei Zhao, Yizhou Ma, L. Elliot Hong, Peter Kochunov, Tianzhou Ma, Shuo Chen</i>	
EFFICIENT DIFFERENTIALLY PRIVATE METHODS FOR A TRANSMISSION DISEQUILIBRIUM TEST IN GENOME WIDE ASSOCIATION STUDIES .....	85
<i>Akito Yamamoto, Tetsuo Shibuya</i>	
IDENTIFYING IMAGING GENETIC ASSOCIATIONS VIA REGIONAL MORPHOMETRICITY ESTIMATION .....	97
<i>Jingxuan Bao, Zixuan Wen, Mansu Kim, Andrew J. Saykin, Paul M. Thompson, Yize Zhao, Li Shen</i>	
IDENTIFYING HIGHLY HERITABLE BRAIN AMYLOID PHENOTYPES THROUGH MINING ALZHEIMER'S IMAGING AND SEQUENCING BIOBANK DATA.....	109
<i>Jingxuan Bao, Zixuan Wen, Mansu Kim, Xiwen Zhao, Brian N. Lee, Sang-Hyuk Jung, Christos Davatzikos, Andrew J. Saykin, Paul M. Thompson, Dokyoon Kim, Yize Zhao, Li Shen</i>	

EFFECTS OF APOE4 AND APOE2 GENOTYPES ON SUBCORTICAL MAGNETIC SUSCEPTIBILITY AND MICROSTRUCTURE IN 27,535 PARTICIPANTS FROM THE UK BIOBANK.....	121
<i>Talia M. Nir, Alyssa H. Zhu, Iyad Ba Gari, Daniel Dixon, Tasfiya Islam, Julio E. Villalon-Reina, Sarah E. Medland, Paul M. Thompson, Neda Jahanshad</i>	

SEPARATING CLINICAL AND SUBCLINICAL DEPRESSION BY BIG DATA INFORMED STRUCTURAL VULNERABILITY INDEX AND ITS IMPACT ON COGNITION: ENIGMA DOT PRODUCT.....	133
<i>Peter Kochunov, Yizhou Ma, Kathryn S. Hatch, Lianne Schmaal, Neda Jahanshad, Paul M. Thompson, Bhim M. Adhikari, Heather Bruce, Joshua Chiappelli, Andrew Van der vaart, Eric L. Goldwaser, Aris Sotiras, Tianzhou Ma, Shuo Chen, Thomas E. Nichols, L. Elliot Hong</i>	

GENERALIZING FEW-SHOT CLASSIFICATION OF WHOLE-GENOME DOUBLING ACROSS CANCER TYPES.....	144
<i>Sherry Chao, David Belanger</i>	

**HUMAN INTRIGUE: META ANALYSIS APPROACHES FOR BIG QUESTIONS WITH BIG DATA**

SESSION INTRODUCTION.....	156
<i>Carly A. Bobak, Meghan E. Muse, Kristine A. Giffin, Derek A. Williamson, Casey S. Greene, Jason H. Moore, Dennis P. Wall</i>	

MULTITASK GROUP LASSO FOR GENOME WIDE ASSOCIATION STUDIES IN DIVERSE POPULATIONS.....	163
<i>Asma Nourira, Chloé-Agathe Azencottt</i>	

MIXED EFFECTS MACHINE LEARNING MODELS FOR COLON CANCER METASTASIS PREDICTION USING SPATIALLY LOCALIZED IMMUNO-ONCOLOGY MARKERS.....	175
<i>Joshua J. Levy, Carly A. Bobak, Mustafa Nasir-Moin, Eren M. Veziroglu, Scott M. Palisoul, Rachael E. Barney, Lucas A. Salas, Brock C. Christensen, Gregory J. Tsongalis, Louis J. Vaickus</i>	

IMPROVING QSAR MODELING FOR PREDICTIVE TOXICOLOGY USING PUBLICLY AGGREGATED SEMANTIC GRAPH DATA AND GRAPH NEURAL NETWORKS.....	187
<i>Joseph D. Romano, Yun Hao, Jason H. Moore</i>	

CAMML: MULTI-LABEL IMMUNE CELL-TYPING AND STEMNESS ANALYSIS FOR SINGLE-CELL RNA-SEQUENCING.....	199
<i>Courtney Schiebout, H. Robert Frost</i>	

RECONCILING SIGNALING PATHWAY DATABASES WITH NETWORK TOPOLOGIES.....	211
<i>Tobias Rubel, Pramesh Singh, Anna Ritz</i>	

**PRECISION MEDICINE: USING ARTIFICIAL INTELLIGENCE TO IMPROVE DIAGNOSTICS AND HEALTHCARE**

SESSION INTRODUCTION.....	223
<i>Roxana Daneshjou, Steven E Brenner, Jonathan H Chen, Dana C Crawford, Samuel G Finlayson, Lukasz Kidziński, Martha L Bulyk</i>	

INTERPRETABLE DEEP LEARNING PREDICTION OF 3D ASSESSMENT OF CARDIAC FUNCTION.....	231
<i>Grant Duffy, Ishan Jain, Bryan He, David Ouyang</i>	
PREDICTING VISUO-MOTOR DISEASES FROM EYE TRACKING DATA .....	242
<i>Kailas Vodrahalli, Maciej Filipkowski, Tiffany Chen, James Zou, Yaping Joyce Liao</i>	
IDENTIFYING CELL TYPE-SPECIFIC CHEMOKINE CORRELATES WITH HIERARCHICAL SIGNAL EXTRACTION FROM SINGLE-CELL TRANSCRIPTOMES.....	254
<i>Sherry Chao, Michael P. Brenner, Nir Hacohen</i>	
HIERARCHICAL GAUSSIAN PROCESSES AND MIXTURES OF EXPERTS TO MODEL COVID-19 PATIENT TRAJECTORIES .....	266
<i>Sunny Cui, Elizabeth C. Yoo, Didong Li, Krzysztof Laudanski, Barbara E. Engelhardt</i>	
DE NOVO PREDICTION OF CELL-DRUG SENSITIVITIES USING DEEP LEARNING-BASED GRAPH REGULARIZED MATRIX FACTORIZATION .....	278
<i>Shuangxia Ren, Yifeng Tao, Ke Yu, Yifan Xue, Russell Schwartz, Xinghua Lu</i>	
CLINICAL RECOMMENDER ALGORITHMS TO SIMULATE DIGITAL SPECIALTY CONSULTATIONS .....	290
<i>Morteza Noshad, Ivana Jankovic, Jonathan H. Chen</i>	
A NOVEL APPROACH FOR TIME SERIES FORECASTING OF INFLUENZA-LIKE ILLNESS USING A REGRESSION CHAIN METHOD.....	301
<i>Nooriyah Poorawala-Lohani, Patricia Riddle, Mehnaz Adnan, Jörg Wicker</i>	
A METHOD FOR LOCALIZING NON-REFERENCE SEQUENCES TO THE HUMAN GENOME.....	313
<i>Brianna Sierra Chrisman, Kelley M Paskov, Chloe He, Jae-Yoon Jung, Nate Stockham, Peter Yigitcan Washington, Dennis Paul Wall</i>	
NETCRS: NETWORK-BASED COMORBIDITY RISK SCORE FOR PREDICTION OF MYOCARDIAL INFARCTION USING BIOBANK-SCALED PHEWAS DATA.....	325
<i>Yonghyun Nam, Sang-Hyuk Jung, Anurag Verma, Vivek Sriram, Hong-Hee Won, Jae-Seung Yun, Regeneron Genetics Center, Dokyoon Kim</i>	
CLOUDPRED: PREDICTING PATIENT PHENOTYPES FROM SINGLE-CELL RNA-SEQ.....	337
<i>Bryan He, Matthew Thomson, Meena Subramaniam, Richard Perez, Chun Jimmie Ye, James Zou</i>	
NONLINEAR POST-SELECTION INFERENCE FOR GENOME-WIDE ASSOCIATION STUDIES .....	349
<i>Lotfi Slim, Clément Chatelain, Chloé-Agathe Azencott</i>	
EVALUATION OF SEX-AWARE PREDIXCAN MODELS FOR PREDICTING GENE EXPRESSION.....	361
<i>Emily Mahoney, Vaibhav Janve, Timothy J. Hohman, Logan Dumitrescu</i>	
MULTI-OMIC GRAPH TRANSFORMERS FOR CANCER CLASSIFICATION AND INTERPRETATION .....	373
<i>Emily Kaczmarek, Amoon Jamzad, Tashifa Imtiaz, Jina Nanayakkara, Neil Renwick, Parvin Mousavi</i>	

AN INVESTIGATION OF THE KNOWLEDGE OVERLAP BETWEEN PHARMACOGENOMICS AND DISEASE GENETICS.....	385
<i>Binglan Li, Michelle Whirl-Carrillo, Matt W. Wright, Larry Babb, Heidi L. Rehm, Teri E. Klein</i>	

## **WORKSHOPS**

EMERGING TOPICS IN CANCER EVOLUTION .....	397
<i>Mohammed El-Kebir, Quaid Morris, Layla Oesper, S. Cenk Sahinalp</i>	
GENOMIC AND SYNTHETIC BIOLOGY DIGITAL BIOSECURITY .....	402
<i>Corey M. Hudson, Nicholas D. Pattengale, Ravishankar K. Iyer, Zbigniew T. Kalbarczyk, Nina Alli</i>	
IMAGE-BASED PROFILING: A POWERFUL AND CHALLENGING NEW DATA TYPE .....	407
<i>Gregory P. Way, Hannah Spitzer, Philip Burnham, Arjun Raj, Fabian Theis, Shantanu Singh, Anne E. Carpenter</i>	
PACKAGING BIOCOMPUTING SOFTWARE TO MAXIMIZE DISTRIBUTION AND REUSE .....	412
<i>William S. Bush, Nicholas Wheeler, Christian Darabos, Brett Beaulieu-Jones</i>	
SOCIAL, TECHNICAL, AND ETHICAL CHALLENGES IN BIOMEDICAL DATA PRIVACY .....	417
<i>Gamze Gursoy, Bradley Malin, Steven E. Brenner</i>	

## **Author Index**