

PACIFIC SYMPOSIUM ON BIOCOMPUTING 2011

2011 marks the 16th Pacific Symposium on Biocomputing. The impact of two major biomedical research trends are clearly seen in this year's conference. First, the national push towards "translational research" for moving discovery from bench to bedside is manifest in the sessions on integration of biological & clinical data and on personal genomics. Second, the revolution in DNA sequencing similarly impacts our sessions on data integration, genome-wide association studies, microbiomes, personal genomics and many of the others too! Thus, we present a conference in which biocomputation is at the forefront of work aimed at bringing the fruits of the genome projects to practical applications. Other sessions focus on the emerging fields of synthetic biology and multiscale modeling. It is an exciting time for the application of computational and informatics technologies to the key problems facing biomedical science.

Indeed, computation has become a recognized component of virtually all major biomedical research efforts, as a cadre of scientists dually trained in biology & medicine as well as computer science, statistics and engineering approach problems of data analysis, fusion, and the generation of new knowledge. The NIH recently renewed the "National Centers for Biomedical Computation" program (<http://www.ncbcs.org/>). This program grew out of the 1999 "BISTI Report" (Biomedical Information Science and Technology Initiative) which recommended (1) a National Centers program, (2) a program on the principles of information storage, curation, analysis and retrieval, (3) the provision of additional resources for investigators creating and apply biomedical computing tools, (4) the creation of a scaleable national computational infrastructure. The Centers have a dual role of performing outstanding research in methods for biomedical computation, while also disseminating software and data they produce to others via training sessions, workshops and collaborative research relationships. A key function of the centers is to provide a milieu in which biocomputing professional can develop. The PSB meeting is proud to also contribute to the creation of a cadre of skilled professional scientists and engineers, but providing pre-meeting tutorials, travel support for students and post-doctoral fellowships, and opportunities for "bottom up" organization of new sessions.

We would like to thank our keynote speakers. Dr. Vijay Pande, Associate Professor of Chemistry will talk about recent progress in large scale simulation of biological macromolecules. Our keynote in the area of Ethical, Legal and Social Implications of Technology will be Ellen Wright Clayton, the Rosalind E. Franklin Professor of Genetics & Health Policy, and Professor of Law and Pediatrics, at Vanderbilt University. Professor Clayton is a leader in law and genetics.

PSB provides sessions focusing on emerging areas in biomedical computation. These sessions are often conceived at the meeting as people discuss the opportunities for new and exciting sessions. The efforts of a dedicated group of leaders has produced an outstanding set of sessions, with associated introductory tutorials. These organizers provide the scientific core of PSB, and their sessions are as follows:

Computational Methods Integrating Diverse Biological and Clinical Data for Translational Science

Gurkan Bebek, Mark Chance, Mehmet Koyuturk, Nathan D. Price

Genome-wide association mapping and rare alleles: from population genomics to personalized medicine

Francisco M. De La Vega, Carlos D. Bustamante, Suzanne M. Leal

Microbiome studies: Understanding how the dominant form of life affects us

James Foster, Jason Moore

Multi-scale Modelling of Biosystems: from Molecular to Mesoscale

Julie Bernauer, Samuel Flores, Xuhui Huang, Seokmin Shin, Ruhong Zhou

Personal Genomics

Can Alkan, Emidio Capriotti, Fereydoun Hormozdiari, Eleazar Eskin, Maricel G. Kann

Reverse Engineering and Synthesis of Biomolecular Systems

Gil Alterovitz, Silvio Cavalcanti, May Wang, and Marco F. Ramoni

With regards to the *Reverse Engineering and Synthesis of Biomolecular Systems* session, we were saddened by the unexpected death of our friend and colleague Dr. Marco Ramoni. Marco was an internationally known computer scientist and Bayesian theorist whose contributions ranged from new understanding of the genetic mechanisms of stroke and asthma to developing novel methodologies. He was a senior faculty member of the Children's Hospital Informatics Program and an Associate Professor of Pediatrics and Medicine at Harvard Medical School and Associate Director of Bioinformatics in Harvard Partners Center for Genetics and Genomics. He will be missed.

We are also pleased to present three workshops, in which investigators with a common interest come together to exchange results and new ideas in a format that is more informal than the peer-reviewed sessions. For this year, the workshops and their organizers are:

Mining the Pharmacogenomics Literature

Kevin Bretonnel Cohen, Yael Garten, Udo Hahn, Nigam H. Shah

Identification of Aberrant Pathway and Network Activity from High-Throughput Data

Michael Ochs, Rachel Karchin, Habtom Ressom, Robert Gentleman

Validation and Modeling of Electron Cryo-microscopy Structures of Biological Nanomachines

Wah Chiu, Helen Berman, Steven Ludtke, Gerard Kleywegt

Finally, we are happy to welcome a new group for a birds-of-a-feather meeting on "Systems Pharmacogenomics." This session is sponsored by the NIH Pharmacogenomics Research Network Statistical Analysis Resource (P-STAR), which is lead by Dr. Marylyn Ritchie under NIH HL065962.

Tiffany Murray continues expertly to manage the peer review process and assembly of the proceedings. We thank the National Institutes of Health and the International Society for Computational Biology (ISCB) for travel grant support. We are particularly grateful to BJ McKay-Morrison at ISCB for her assistance. We also acknowledge the many busy researchers who reviewed the submitted manuscripts on a very tight schedule. The partial list following this preface does not include many who wished to remain anonymous, and of course we apologize to any who may have been left out by mistake.

We look forward to a great meeting once again.

Aloha!

Pacific Symposium on Biocomputing Co-Chairs,
September 29, 2010

Russ B. Altman

Departments of Bioengineering, Genetics & Medicine, Stanford University

A. Keith Dunker

Department of Biochemistry and Molecular Biology, Indiana University School of Medicine

Lawrence Hunter

Department of Pharmacology, University of Colorado Health Sciences Center

Teri E. Klein

Department of Genetics, Stanford University

Thanks to the reviewers...

Finally, we wish to thank the scores of reviewers. PSB requires that every paper in this volume be reviewed by at least three independent referees. Since there is a large volume of submitted papers, paper reviews require a great deal of work from many people. We are grateful to all of you listed below and to anyone whose name we may have accidentally omitted or who wished to remain anonymous.

Gary Bader	Hui Jiang	Ben Raphael
Ranjit Bahadur	Eun Yong Kang	Soumya Raychaudhuri
Gurkan Bebek	Ulas Karaoz	David Reif
Julie Bernauer	Parminder Kaur	Marylyn Ritchie
Greg Bowman	Alon Keinan	Masaki Sasai
Sharon Browning	Sadik Khuder	Dan Schaid
Mike Brudno	Ken Kidd	Ursel Schuette
John Bunge	Sungwan Kim	Josiah Seaman
Anita Burgun	Carl Kingsford	Dong Shen
Jun Cai	Rob Knight	Jana Shen
Ali Cakmak	Mehmet Koyutürk	Seokmin Shin
Sukmohinder Chahal	Charu Gupta Kumar	Daniel Silva
Jack Chen	Sadhir Kumar	Mitchell Sogin
Xian-Ming Chen	Douglas Lauffenburger	Jean-Marc Stayert
Cheng Cheng	Jose Lazano	Yan Sun
Jacques Chomilier	Seunghak Lee	Min-Han Tan
Bradley Coe	Jing Li	Nelson Tang
Sarah Cohen Boulakia	Qian-Zhong Li	Nicholas Tatonetti
Markus Covert	Xin Li	Cheryl Thompson
Phuong Dao	Liming Liang	Duygu Ucar
Payel Das	Pu Liu	Jerome Waldispuhl
Mariza De Andrade	Ivan Viktorovich Maly	Kai Wang
James Faeder	Elena Marchiori	Kesheng Wang
Michael Feig	Eden Martin	George Weinstock
Juan Fernandez-Recio	Andrew McPherson	Dahlia Weiss
Pawel Gajer	Jason Mezey	Scott Weiss
Jean Garnier	Karen Nelson	Zhen Xia
Christopher Haiman	Mohammed Orloff	Lei Xie
Farhad Hormozdiari	Badri Padhukasahasram	Gokhan Yavas
Bryan Howie	Andrew Pakstis	Seungtai Yoon
Joanna Howson	John Pani	Deniz Yorukoglu
Xuhui Huang	Vishal Patel	Ye Yuzhen
Yin-Fu Huang	Lynette Phillips	GQ Zhang
Jui-Hung Hung	Mihai Pop	Zemin Zhang
Thomas Ioerger	Nathan D. Price	Hoghyu Zhao
Andy Itsara	Kriti Puniyani	Jasmine Zhou
Asif Javed	Jeroen Raes	

Preface.....	v
--------------	---

COMPUTATIONAL METHODS INTEGRATING DIVERSE BIOLOGICAL AND CLINICAL DATA FOR TRANSLATIONAL SCIENCE

<i>Session Introduction</i>	1
Gurkan Bebek, Mehmet Koyutuk, Mark Chance, Nathan D. Price	
<i>Towards Integrative Gene Prioritization in Alzheimer's Disease</i>	4
Jang Lee, Graciela Gonzalez	
<i>Systems Biology Analyses of Gene Expression and Genome Wide Association Study Data in Obstructive Sleep Apnea</i>	14
Yu Liu, Sanjay R. Patel, Rod K. Nibbe, Sean Maxwell, Salim A. Chowdhury, Mehmet Koyuturk, Xiaofeng Zhu, Emma Larkin, Sarah Buxbaum, Naresh Punjabi, Sina Gharib, Susan Redline, Mark R. Chance	
<i>Finding Most Likely Haplotypes in General Pedigrees through Parallel Search with Dynamic Load Balancing</i>	26
Lars Otten, Rina Dechter	
<i>Dynamic, Multi-Level Network Models of Clinical Trials</i>	38
Marco D. Sorani, Geoffrey T. Manley, J. Claude Hemphill, Sergio E. Baranzini	
<i>Mining Functionally Relevant Gene Sets for Analyzing Physiologically Novel Clinical Expression Data</i>	50
Sevin Turcan, Xintao Wei, Jill Maron, Douglas E. Vetter, Donna K. Slonim	
<i>Genotype Phenotype Mapping in RNA Viruses - Disjunctive Normal Form Learning</i>	62
Chuang Wu, Andrew S. Walsh, Roni Rosenfeld	
GENOME-WIDE ASSOCIATION MAPPING AND RARE ALLELES: FROM POPULATION GENOMICS TO PERSONALIZED MEDICINE	
<i>Session Introduction</i>	74
Francisco M. De La Vega, Carlos D. Bustamante, Suzanne M. Leal	
<i>An Application and Empirical Comparison of Statistical Analysis Methods for Associating Rare Variants to a Complex Phenotype</i>	76
Vikas Bansal, Ondrej Libiger, Ali Torkamani, Nicholas J. Schork	
<i>Haplotype Phasing by Multi-Assembly of Shared Haplotypes: Phase-Dependent Interactions Between Rare Variants</i>	88
Bjarni V. Halldorsson. Derek Aguiar, Sorin Istrail	

<i>An Evaluation of Power to Detect Low-frequency Variant Associations Using Allele-Matching Tests that Account for Uncertainty</i>	100
Eleftheria Zeggini, Jennifer L. Asimit	
<i>Penalized Regression for Genome-wide Association Screening of Sequence Data</i>	106
Hua Zhou, David Alexander, Mary Sehl, Janet Sinsheimer, Kenneth Lange	
MICROBIOME STUDIES: UNDERSTANDING HOW THE DOMINANT FORM OF LIFE AFFECTS US	
<i>Session Introduction</i>	118
James Foster, Jason Moore	
<i>Estimating the Number of Species with Catchall</i>	121
John Bunge	
<i>A Framework for Analysis of Metagenomic Sequencing Data</i>	131
A. Murat Eren, Michael J. Ferris, Christopher M. Taylor	
<i>Visualization and Statistical Comparisons of Microbial Communities Using R Packages on Phylochip Data</i>	142
S. Holmes, A. Alekseyenko, A. Timme, T. Nelson, P.J. Pasricha, A. Spormann	
<i>Human Microbiome Visualization Using 3D Technology</i>	154
Jason H. Moore, Richard Cowper-Sal.Lari, Douglas Hill, Patricia Hibberd, Juliette Madan	
<i>Comparing Bacterial Communities Inferred from 16S rRNA Gene Sequencing and Shotgun Metagenomics</i>	165
Neethu Shah, Haixu Tang, Thomas G. Doak, Yuzhen Ye	
MULTI-SCALE MODELLING OF BIOSYSTEMS: FROM MOLECULAR TO MESOSCALE	
<i>Session Introduction</i>	177
Julie Bernauer, Samuel Flores, Xuhui Huang, Seokmin Shin, Ruhong Zhou	
<i>Computational Generation Inhibitor-Bound Conformers of P38 Map Kinase and Comparison with Experiments</i>	181
Ahmet Bakan, Ivet Bahar	
<i>Molecular Dynamics Simulations of the Full Triple Helical Region of Collagen Type I Provide an Atomic Scale View of the Protein's Regional Heterogeneity</i>	193
Dale L Bodian, Randall J Radmer, Sean Holbert, Teri E Klein	
<i>Structural Insights Into Pre-Translocation Ribosome Motions</i>	205
Samuel Coulbourn Flores, Russ B. Altman	

New Conformational Search Method Using Genetic Algorithm and Knot Theory for Proteins217
Yoshitake Sakae, Tomoyuki Hiroyasu, Mitsunori Miki, Yuko Okamoto

PERSONAL GENOMICS

Session Introduction.....229
Can Alkan, Emidio Capriotti, Eleazar Eskin, Fereydoun Hormozdiari, Maricel G. Kann

The Reference Human Genome Demonstrates High Risk of Type 1 Diabetes and Other Disorders231
Rong Chen, Atul J. Butte

Matching Cancer Genomes to Established Cell Lines for Personalized Oncology243
Joel T. Dudley, Rong Chen, Atul J. Butte

Use of Biological Knowledge to Inform the Analysis of Gene-Gene Interactions Involved in Modulating Virologic Failure with Efavirenz-Containing Treatment Regimens in ART-Naïve ACTG Clinical Trials Participants253
Benjamin J. Grady, Eric C. Torstenson, Paul J. McLaren, Paul W. De Bakker, David W. Haas, Gregory K. Robbins, Roy M. Gulick, Richard Haubrich, Heather Ribaud, Marylyn D. Ritchie

Visual Integration of Results from a Large DNA Biobank (BioVU) Using Synthesis-View.....265
Sarah Pendergrass, Scott M. Dudek, Dan M. Roden, Dana C. Crawford, Marylyn D. Ritchie

Multivariate Analysis of Regulatory SNPs: Empowering Personal Genomics by Considering Cis-Epistasis and Heterogeneity276
Stephen D. Turner, William S. Bush

Haplotype Inference from Single Short Sequence Reads Using a Population Genealogical History Model288
Jin Zhang, Yufeng Wu

REVERSE ENGINEERING AND SYNTHESIS OF BIOMOLECULAR SYSTEMS

Session Introduction.....300
Gil Alterovitz, Silvio Cavalcanti, Taro M. Muso, Marco F. Ramoni, May Wang

Binary Counting with Chemical Reactions.....302
Aleksandra Kharam, Hua Jiang, Marc Riedel

Defining the Players in Higher-Order Networks: Predictive Modeling for Reverse Engineering Functional Influence Networks314
Jason E. McDermott, Michelle Archuleta, Susan L. Stevens, Mary P. Stenzel-Poore,
Antonio Sanfilippo

Rate-Independent Constructs for Chemical Computation326
Philip Senum, Marc Riedel

Towards Real-Time Control of Gene Expression: Controlling the HOG Signaling Cascade338
Jannis Uhlendorf, Samuel Bottani, Francois Fages, Pascal Hersen, Gregory Batt

Identifying Targets for Intervention by Analyzing Basins of Attraction350
Michael P Verdicchio, Seungchan Kim

MINING THE PHARMACOGENOMICS LITERATURE

Workshop Introduction362
Kevin Bretonnel Cohen, Yael Garten, Udo Hahn, Nigam H. Shah

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

Workshop Introduction364
Michael Ochs, Rachel Karchin, Habtom Resson, Robert Gentleman

VALIDATION AND MODELING OF ELECTRON CRYO-MICROSCOPY STRUCTURES OF BIOLOGICAL NANOMACHINES

Workshop Introduction369
Steven Ludtke, Catherine Lawson, Gerard Kleywegt, Helen Berman, Wah Chiu