Series Editors
Sorin Istrail, Brown University, Providence, RI, USA
Pavel Pevzner, University of California, San Diego, CA, USA
Michael Waterman, University of Southern California, Los Angeles, CA, USA

Volume Editors
Eleazar Eskin
Ben Raphael
University of California
Department of Computer Science
San Diego, CA, USA
E-mail: {eeskin,braphael}@cs.ucsd.edu

Trey Ideker
Christopher Workman
University of California
Department of Bioengineering
San Diego, CA, USA
E-mail: {trey,cworkman}@bioeng.ucsd.edu

Library of Congress Control Number: 2006940071

CR Subject Classification (1998): F.2, G.3, E.1, H.2.8, J.3

LNCS Sublibrary: SL 8 – Bioinformatics

ISSN 0302-9743

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11916338 06/3142 5 4 3 2 1 0
Preface

It has become increasingly evident that the use of large-scale experimental data and the application of principles from systems biology are gaining widespread acceptance in mainstream biology. Systems biology involves the use of global cellular measurements, i.e., genomic, proteomic, and metabolomic, to construct computational models of cellular processes and disease. These approaches involve an integration of experimental and computational techniques and may include: 1) developing models of cellular processes, 2) measuring the response to perturbations of model components, and 3) iteratively formulating and testing new hypotheses for unexpected observations.

A research area that is particularly important to systems biology is the study of gene regulatory networks. Although genome sequencing efforts have been tremendously successful, much is unknown about the regulation of these sequenced genomes. Automatic methods for helping decipher the regulatory mechanism are crucial for understanding the regulatory network as a whole. However, many new challenges are presented when analyzing complete genomes. These challenges include motif discovery in large genomes, leveraging information from multiple genomes, detection of weak signals and incorporating different types of genomic data such as protein localization data and gene expression. Novel methodology will be particularly relevant given the hypothesis that the observed phenotypic differences between organisms with very similar genomes may be due to variations in the gene regulation.

The amount of research in both of these areas has exploded in recent years, as witnessed by the number of research presentations at meetings such as RECOMB, ISMB, PSB, the Biopathways Consortium, and ICSB. The jointly held RECOMB Satellite on Systems Biology and the RECOMB Satellite on Regulatory Genomics provide a forum for addressing these challenges.

This year’s workshops also included a special session on Computational Developmental Biology organized by David Gifford. Unique challenges posted by developmental biology include: 1) computational model representations that can express the execution of complex natural programs over time, 2) the identification of key developmental state variables and experimental methods to reliably observe these variables, and 3) the use of computational models to understand developmentally related disease, and to help develop therapeutics including the programming of stem cells as therapeutic agents. Invited speakers at this year’s special session addressed the question: “What key developmental biology problems can now be examined from a systems biology perspective, and what data are necessary to do so?” The goal of this special session was to help computational and systems biologists understand both the challenge and excitement of working on development.

August 2006

The Organizing Committee
Organization

Steering Committee: Regulatory Genomics

Pierre Baldi University of California, Irvine
Michael Eisen Lawrence Berkeley National Lab
Eleazar Eskin University of California, San Diego
Pavel Pevzner University of California, San Diego

Steering Committee: Systems Biology

Leroy Hood Institute for Systems Biology
Trey Ideker University of California, San Diego
Douglas Lauffenburger MIT
Satoru Miyano University of Tokyo
Ron Shamir Tel Aviv University

Organizing Committee

Eleazar Eskin University of California, San Diego
David Gifford Massachusetts Institute of Technology
Trey Ideker University of California, San Diego
Teresa M. Przytycka NIH/NLM/NCBI
Ben Raphael University of California, San Diego
Cenk Sahinalp Simon Fraser University
Samantha Smeraglia University of California, San Diego
Christopher Workman University of California, San Diego

Program Committee

John Aitchison Institute for Systems Biology
Adam Arkin University of California, Berkeley
Gary Bader Memorial Sloan-Kettering Cancer Center
Pierre Baldi University of California, Irvine
Yoseph Barash Hebrew University
Mathieu Blanchette McGill University
Gal Chechik Stanford University
Francis Chin University of Hong Kong
Michael Cusick Harvard Medical School
Eric Davidson CalTech
Michael Eisen Lawrence Berkeley National Lab
Eleazar Eskin University of California, San Diego
Nir Friedman               Hebrew University
Irit Gat-Vicks            Tel Aviv University
Mikhail Gelfand           Moscow State University
David Gifford             Massachusetts Institute of Technology
Sridhar Hannenhalli       University of Pennsylvania
Jeff Hasty                University of California, San Diego
Leong Hon Wai             National University of Singapore
Leroy Hood                Institute for Systems Biology
Trey Ideker               University of California, San Diego
Richard Karp              University of California, Berkeley
Uri Keich                 Cornell University
Douglas Lauffenburger     Massachusetts Institute of Technology
Christina Leslie          Columbia University
Mike Levine               University of California, Berkeley
Hao Li                    University of California, San Francisco
Nick Luscombe             European Bioinformatics Institute
Satoru Miyano             University of Tokyo
Dana Pe’er                Harvard Medical School
Pavel Pevzner             University of California, San Diego
Tzachi Pilpel             Weizmann Institute of Science
Teresa M. Przytycka        NIH/NLM/NCBI
Ben Raphael               University of California San Diego
Bing Ren                  University of California, San Diego
Aviv Regev                Harvard Medical School
Mireille Regnier          INRIA
Marie-France Sagot        INRIA
Cenk Sahinalp             Simon Fraser University
Eran Segal                Rockefeller University
Ron Shamir                Tel Aviv University
Roded Sharan              Tel Aviv University
Amos Tanay                Tel Aviv University
Alfonso Valencia          Centro Nacional de Biotecnologia
Marc Vidal                Harvard Medical School
Christopher Workman       University of California, San Diego
Eric Xing                 Carnegie Mellon University
Zohar Yakhini             Agilent
Ralf Zimmer               LMU, Institut für Informatik

Sponsoring Institutions

Industry-University Cooperative Research Program, The UC Discovery Grant
California Institute for Telecommunications and Information Technology, Cal-(IT)²
# Table of Contents

An Interactive Map of Regulatory Networks of *Pseudomonas aeruginosa* Genome .......................................................... 1
    Weihui Wu, Yongling Song, Shouguang Jin, and Su-Shing Chen

The Pathalyzer: A Tool for Analysis of Signal Transduction Pathways .......................................................... 11
    David L. Dill, Merrill A. Knapp, Pamela Gage, Carolyn Talcott, Keith Laderoute, and Patrick Lincoln

Decomposition of Overlapping Protein Complexes: A Graph Theoretical Method for Analyzing Static and Dynamic Protein Associations ....... 23
    Elena Zotenko, Katia S. Guimarães, Raja Jothi, and Teresa M. Przytycka

Comparison of Protein-Protein Interaction Confidence Assignment Schemes .......................................................... 39
    Silpa Suthram, Tomer Shlomi, Eytan Ruppin, Roded Sharan, and Trey Ideker

Characterization of the Effects of TF Binding Site Variations on Gene Expression Towards Predicting the Functional Outcomes of Regulatory SNPs .......................................................... 51
    Michal Lapidot and Yitzhak Pilpel

Comparative Systems Biology of the Sporulation Initiation Network in Prokaryotes .................................................... 62
    Michiel de Hoon and Dennis Vitkup

Improvement of Computing Times in Boolean Networks Using Chi-square Tests .................................................... 70
    Haseong Kim, Jae K. Lee, and Taesung Park

Build a Dictionary, Learn a Grammar, Decipher Stegoscripts, and Discover Genomic Regulatory Elements ............................ 80
    Guandong Wang and Weixiong Zhang

Causal Inference of Regulator-Target Pairs by Gene Mapping of Expression Phenotypes .......................................................... 95
    David Kulp and Manjunatha Jagalur

 Examination of the tRNA Adaptation Index as a Predictor of Protein Expression Levels .......................................................... 107
    Orna Man, Joel L. Sussman, and Yitzhak Pilpel
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved Duplication Models for Proteome Network Evolution</td>
<td>119</td>
</tr>
<tr>
<td>Gürkan Bebek, Petra Berenbrink, Colin Cooper, Tom Friedetzky,</td>
<td></td>
</tr>
<tr>
<td>Joseph H. Nadeau, and S. Cenk Sahinalp</td>
<td></td>
</tr>
<tr>
<td>Application of Expectation Maximization Clustering to Transcription</td>
<td>138</td>
</tr>
<tr>
<td>Factor Binding Positions for cDNA Microarray Analysis</td>
<td></td>
</tr>
<tr>
<td>Chih-Yu Chen, Von-Wun Soo, and Chi-Li Kuo</td>
<td></td>
</tr>
<tr>
<td>Combinatorial Genetic Regulatory Network Analysis Tools for High</td>
<td>150</td>
</tr>
<tr>
<td>Throughput Transcriptomic Data</td>
<td></td>
</tr>
<tr>
<td>Elissa J. Chesler and Michael A. Langston</td>
<td></td>
</tr>
<tr>
<td>Topological Robustness of the Protein-Protein Interaction Networks</td>
<td>166</td>
</tr>
<tr>
<td>Chien-Hung Huang, Jywe-Fei Fang, Jeffrey J.P. Tsai, and Ka-Lok Ng</td>
<td></td>
</tr>
<tr>
<td>A Bayesian Approach for Integrating Transcription Regulation and</td>
<td>178</td>
</tr>
<tr>
<td>Gene Expression: Application to Saccharomyces Cerevisiae Cell Cycle</td>
<td></td>
</tr>
<tr>
<td>Sudhakar Jonnalagadda and Rajagopalan Srinivasan</td>
<td></td>
</tr>
<tr>
<td>Probabilistic in Silico Prediction of Protein-Peptide Interactions</td>
<td>188</td>
</tr>
<tr>
<td>Wolfgang Lehrach, Dirk Husmeier, and Christopher K.I. Williams</td>
<td></td>
</tr>
<tr>
<td>Improved Pattern-Driven Algorithms for Motif Finding in DNA</td>
<td>198</td>
</tr>
<tr>
<td>Sequences</td>
<td></td>
</tr>
<tr>
<td>Sing-Hoi Sze and Xiaoyan Zhao</td>
<td></td>
</tr>
<tr>
<td>Annotation of Promoter Regions in Microbial Genomes Based on DNA</td>
<td>212</td>
</tr>
<tr>
<td>Structural and Sequence Properties</td>
<td></td>
</tr>
<tr>
<td>Huiquan Wang and Craig J. Benham</td>
<td></td>
</tr>
<tr>
<td>An Interaction-Dependent Model for Transcription Factor Binding</td>
<td>225</td>
</tr>
<tr>
<td>Li-San Wang, Shane T. Jensen, and Sridhar Hannenhalli</td>
<td></td>
</tr>
<tr>
<td>Computational Characterization and Identification of Core Promoters</td>
<td>235</td>
</tr>
<tr>
<td>of MicroRNA Genes in C. elegans, H. sapiens and A. thaliana</td>
<td></td>
</tr>
<tr>
<td>Xuefeng Zhou, Jianhua Ruan, Guandong Wang, and Weixiong Zhang</td>
<td></td>
</tr>
<tr>
<td>A Comprehensive Kinetic Model of the Exocytotic Process: Evaluation</td>
<td>249</td>
</tr>
<tr>
<td>of the Reaction Mechanism</td>
<td></td>
</tr>
<tr>
<td>Aviv Mezer, Eran Bosis, Uri Ashery, Esther Nachliel, and Menachem Gutman</td>
<td></td>
</tr>
<tr>
<td>Author Index</td>
<td>259</td>
</tr>
</tbody>
</table>