

Lecture Notes in Bioinformatics

7632

Edited by S. Istrail, P. Pevzner, and M. Waterman

Editorial Board: A. Apostolico S. Brunak M. Gelfand

T. Lengauer S. Miyano G. Myers M.-F. Sagot D. Sankoff

R. Shamir T. Speed M. Vingron W. Wong

Subseries of Lecture Notes in Computer Science

Tetsuo Shibuya Hisashi Kashima
Jun Sese Shandar Ahmad (Eds.)

Pattern Recognition in Bioinformatics

7th IAPR International Conference, PRIB 2012
Tokyo, Japan, November 8-10, 2012
Proceedings



Springer

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

Tetsuo Shibuya
The University of Tokyo, Institute of Medical Science
4-6-1 Shirokanedai, Minato-ku, 108-8639 Tokyo, Japan
E-mail: tshibuya@hgc.jp

Hisashi Kashima
The University of Tokyo, Department of Mathematical Informatics
7-3-1 Hongo, Bunkyo-ku, 113-8654 Tokyo, Japan
E-mail: kashima@mist.i.u-tokyo.ac.jp

Jun Sese
Tokyo Institute of Technology, Department of Computer Science
2-12-1 Ookayamama, Meguro-ku, 152-8550 Tokyo, Japan
E-mail: sesejun@cs.titech.ac.jp

Shandar Ahmad
National Institute of Biomedical Innovation
7-6-8 Saito-Asagi, Suita, 567-0085 Osaka, Japan
E-mail: shandar@nibio.go.jp

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-34122-9 e-ISBN 978-3-642-34123-6
DOI 10.1007/978-3-642-34123-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012948870

CR Subject Classification (1998): J.3, I.5, F.2.2, I.2, I.4, H.3.3, H.2.8

LNCS Sublibrary: SL 8 – Bioinformatics

© Springer-Verlag Berlin Heidelberg 2012

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.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Pattern Recognition in Bioinformatics (PRIB) meeting was established in 2006 under the auspices of the International Association for Pattern Recognition (IAPR) to create a focus for the development and application of pattern recognition techniques in the biological domain. Since its establishment, PRIB has brought together top researchers, practitioners, and students from around the world to discuss the applications of pattern recognition methods in the field of bioinformatics so as to solve problems in the life sciences.

The seventh PRIB conference was held in Tokyo, Japan, during November 8–10, 2012. This year we received 33 high-quality submissions from various countries around the world, and 24 of them were finally included in these conference proceedings. Their topics range widely from fundamental techniques, sequence analysis to biological network analysis. We were fortunate to have three leading researchers for invited speakers; Hwanjo Yu from Postech, Kwong-Sak Leung from The Chinese University of Hong Kong, and Takayuki Aoki from Tokyo Institute of Technology.

We would like to thank all authors for submitting the high-quality papers, the reviewers for their efforts to keeping the high quality of this conference, and the sponsors for generously providing financial support. Finally, we are grateful to Springer for their professional support in preparing these proceedings and for the continued support of PRIB.

November 2012

Tetsuo Shibuya
Hisashi Kashima
Jun Sese
Shandar Ahmad

Organization

Organizing Committee

General Chairs

Jun Sese	Tokyo Institute of Technology
Shandar Ahmad	National Institute of Biomedical Innovation

Program Chairs

Tetsuo Shibuya	The University of Tokyo
Hisashi Kashima	The University of Tokyo

Publicity Chairs

Tsuyoshi Kato	Gunma University
Koji Tsuda	National Institute of Advanced Industrial Science and Technology

Invited Speaker Chairs

Masakazu Sekijima	Tokyo Institute of Technology
Tsuyoshi Hachiya	Keio University

Program Committee

Jesus S. Aguilar-Ruiz	Zhenyu Jia	Esa Pitkänen
Tatsuya Akutsu	Yoshimoto Junichiro	Clara Pizzuti
Kiyoko Aoki-Kinoshita	Giuseppe Jurman	Beatriz Pontes
Sebastian Böcker	R. Krishna Murthy	Miguel Rocha
Jaume Bacardit	Karuturi	Juho Rousu
Hideo Bannai	Tsuyoshi Kato	Yvan Saeys
Rainer Breitling	Tetsuji Kuboyama	Hiroto Saigo
Frederic Cazals	Xuejun Liu	Taro Saito
Florence D'Alché-Buc	Marco Loog	Guido Sanguinetti
Dick De Ridder	Elena Marchiori	Masakazu Sekijima
Tjeerd Dijkstra	Francesco Masulli	Kim Seyoung
Federico Divina	Alison Motsinger-Reif	Evangelos Simeonidis
Richard Edwards	Vadim Mottl	Johan Suykens
Maurizio Filippone	Sach Mukherjee	Roberto Tagliaferri
Rosalba Giugno	Allioune Ngom	Koji Tsuda
Jin-Kao Hao	Josselin Noirel	Alexey Tsymbal
Morihiro Hayashida	Carlotta Orsenigo	Hong Yan
Tom Heskes	Andrea Passerini	Haixuan Yang
Seiya Imoto	Thang Pham	

Sponsors

International Association for Pattern Recognition (IAPR)

Tokyo Institute of Technology

Japanese Society for Bioinformatics (JSBi)

Table of Contents

Generic Methods – I

Robust Community Detection Methods with Resolution Parameter for Complex Detection in Protein Protein Interaction Networks	1
<i>Twan van Laarhoven and Elena Marchiori</i>	
Machine Learning Scoring Functions Based on Random Forest and Support Vector Regression	14
<i>Pedro J. Ballester</i>	
A Genetic Algorithm for Scale-Based Translocon Simulation	26
<i>Sami Laroum, Béatrice Duval, Dominique Tessier, and Jin-Kao Hao</i>	
A Framework of Gene Subset Selection Using Multiobjective Evolutionary Algorithm	38
<i>Yifeng Li, Alioune Ngom, and Luis Rueda</i>	

Generic Methods – II

Multiple Tree Alignment with Weights Applied to Carbohydrates to Extract Binding Recognition Patterns	49
<i>Masae Hosoda, Yukie Akune, and Kiyoko F. Aoki-Kinoshita</i>	
A Unified Adaptive Co-identification Framework for High-D Expression Data	59
<i>Shuzhong Zhang, Kun Wang, Cody Ashby, Bilian Chen, and Xiuzhen Huang</i>	
Protein Clustering on a Grassmann Manifold	71
<i>Chendra Hadi Suryanto, Hiroto Saigo, and Kazuhiro Fukui</i>	

Visualization, Image Analysis, and Platforms

Improving the Portability and Performance of jViz.RNA – A Dynamic RNA Visualization Software	82
<i>Boris Shabash, Kay Wiese, and Edward Glen</i>	
A Novel Machine Learning Approach for Detecting the Brain Abnormalities from MRI Structural Images	94
<i>Lavneet Singh, Girija Chetty, and Dharmendra Sharma</i>	

An Open Framework for Extensible Multi-stage Bioinformatics Software	106
<i>Gabriel Keeble-Gagnère, Johan Nyström-Persson, Matthew I. Bellgard, and Kenji Mizuguchi</i>	

Applications of Pattern Recognition Techniques

An Algorithm to Assemble Gene-Protein-Reaction Associations for Genome-Scale Metabolic Model Reconstruction	118
<i>João Cardoso, Paulo Vilaça, Simão Soares, and Miguel Rocha</i>	
A Machine Learning and Chemometrics Assisted Interpretation of Spectroscopic Data – A NMR-Based Metabolomics Platform for the Assessment of Brazilian Propolis	129
<i>Marcelo Maraschin, Amélia Somensi-Zeggio, Simone K. Oliveira, Shirley Kuhnen, Maíra M. Tomazzoli, Ana C.M. Zeri, Rafael Carreira, and Miguel Rocha</i>	
Principal Component Analysis for Bacterial Proteomic Analysis.....	141
<i>Y-h. Taguchi and Akira Okamoto</i>	

Protein Structure and Docking

Application of the Multi-modal Relevance Vector Machine to the Problem of Protein Secondary Structure Prediction	153
<i>Nikolay Razin, Dmitry Sungurov, Vadim Mottl, Ivan Torshin, Valentina Sulimova, Oleg Seredin, and David Windridge</i>	
Cascading Discriminant and Generative Models for Protein Secondary Structure Prediction.....	166
<i>Fabienne Thomarat, Fabien Lauer, and Yann Guermeur</i>	
Improvement of the Protein-Protein Docking Prediction by Introducing a Simple Hydrophobic Interaction Model: An Application to Interaction Pathway Analysis	178
<i>Masahito Ohue, Yuri Matsuzaki, Takashi Ishida, and Yutaka Akiyama</i>	
Representation of Protein Secondary Structure Using Bond-Orientational Order Parameters.....	188
<i>Cem Meydan and Osman Ugur Sezerman</i>	

Complex Data Analysis

Diagnose the Premalignant Pancreatic Cancer Using High Dimensional Linear Machine	198
<i>Yifeng Li and Alioune Ngom</i>	

Predicting V(D)J Recombination Using Conditional Random Fields	210
<i>Raunaq Malhotra, Shruthi Prabhakara, and Raj Acharya</i>	

A Simple Genetic Algorithm for Biomarker Mining	222
<i>Dusan Popovic, Alejandro Sifrim, Georgios A. Pavlopoulos, Yves Moreau, and Bart De Moor</i>	

Sequence Analysis

Finding Conserved Regions in Protein Structures Using Support Vector Machines and Structure Alignment	233
<i>Tatsuya Akutsu, Morihiro Hayashida, and Takeyuki Tamura</i>	

Aligning Discovered Patterns from Protein Family Sequences	243
<i>En-Shiun Annie Lee, Dennis Zhuang, and Andrew K.C. Wong</i>	

Application of the Burrows-Wheeler Transform for Searching for Approximate Tandem Repeats	255
<i>Agnieszka Danek, Rafal Pokrzywa, Izabela Makatowska, and Andrzej Polański</i>	

Pattern Recognition for Subfamily Level Classification of GPCRs Using Motif Distillation and Distinguishing Power Evaluation	267
<i>Ahmet Sinan Yavuz, Bugra Ozer, and Osman Ugur Sezerman</i>	

Author Index	277
-------------------------------	-----