German Conference on Bioinformatics 2012

GCB'12, September 19–22, 2012, Jena, Germany

Edited by

Sebastian Böcker
Franziska Hufsky
Kerstin Scheubert
Jana Schleicher
Stefan Schuster

Editors

Sebastian Böcker
sebastian.boecker@uni-jena.de
Franziska Hufsky
franziska.hufsky@uni-jena.de
Kerstin Scheubert
kerstin.scheubert@uni-jena.de

Jana Schleicher
jana.foerster@uni-jena.de
Stefan Schuster
stefan.schuster@uni-jena.de

Chair of Bioinformatics
Faculty of Mathematics and Computer Science
Friedrich-Schiller-University Jena

Department of Bioinformatics
Faculty of Biology and Pharmacy
Friedrich-Schiller-University Jena

ACM Classification 1998
J.3 Life and Medical Sciences

ISBN 978-3-939897-44-6

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at http://www.dagstuhl.de/dagpub/978-3-939897-44-6.

Publication date
September, 2012

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

License
This work is licensed under a Creative Commons Attribution-NoDerivs (BY-NC-ND) license: http://creativecommons.org/licenses/by-nd/3.0/legalcode
In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.
- No derivation: It is not allowed to alter or transform this work.

The copyright is retained by the corresponding authors.

Digital Object Identifier: 10.4230/OASIcs.GCB.2012.i

ISBN 978-3-939897-44-6 **ISSN 2190-6807** **http://www.dagstuhl.de/oasics**

OASIcs – OpenAccess Series in Informatics

OASIcs aims at a suitable publication venue to publish peer-reviewed collections of papers emerging from a scientific event. OASIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

Editorial Board

- Daniel Cremers (TU München, Germany)
- Barbara Hammer (Universität Bielefeld, Germany)
- Marc Langheinrich (Università della Svizzera Italiana – Lugano, Switzerland)
- Dorothea Wagner (*Editor-in-Chief*, Karlsruher Institut für Technologie, Germany)

ISSN 2190-6807

www.dagstuhl.de/oasics

Contents

Preface
Sebastian Böcker, Franziska Hufsky, Kerstin Scheubert, Jana Schleicher, and Stefan Schuster i

ModeScore: A Method to Infer Changed Activity of Metabolic Function from Transcript Profiles
Andreas Hoppe and Hermann-Georg Holzhütter 1

Comparing Fragmentation Trees from Electron Impact Mass Spectra with Annotated Fragmentation Pathways
Franziska Hufsky and Sebastian Böcker 12

Finding Characteristic Substructures for Metabolite Classes
Marcus Ludwig, Franziska Hufsky, Samy Elshamy, and Sebastian Böcker 23

A Two-Step Soft Segmentation Procedure for MALDI Imaging Mass Spectrometry Data
Ilya Chernyavsky, Theodore Alexandrov, Peter Maass, and Sergey I. Nikolenko ... 39

Building and Documenting Workflows with Python-Based Snakemake
Johannes Köster and Sven Rahmann 49

Online Transitivity Clustering of Biological Data with Missing Values
Richard Röttger, Christoph Kreutzer, Thuy Duong Vu, Tobias Wittkop, and Jan Baumbach 57

ConReg: Analysis and Visualization of Conserved Regulatory Networks in Eukaryotes
Robert Pesch, Matthias Böck, and Ralf Zimmer 69

Designing q-Unique DNA Sequences with Integer Linear Programs and Euler Tours in De Bruijn Graphs
Marianna D'Addario, Nils Kriege, and Sven Rahmann 82

Polyglutamine and Polyalanine Tracts Are Enriched in Transcription Factors of Plants
Nina Kottenhagen, Lydia Gramzow, Fabian Horn, Martin Pohl, and Günter Theißen 93

Computation and Visualization of Protein Topology Graphs Including Ligand Information
Tim Schäfer, Patrick May, and Ina Koch 108

Unbiased Protein Interface Prediction Based on Ligand Diversity Quantification
Reyhaneh Esmaielbeiki and Jean-Christophe Nebel 119

German Conference on Bioinformatics 2012 (GCB'12).
Editors: S. Böcker, F. Hufsky, K. Scheubert, J. Schleicher, S. Schuster

OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Preface

This volume contains papers presented at the German Conference on Bioinformatics (GCB 2012) held in Jena, Germany, September 19 – 22, 2012. The German Conference on Bioinformatics is an annual, international conference, which provides a forum for the presentation of current research in bioinformatics and computational biology. The GCB 2012 was organized by the Jena Center of Bioinformatics (JCB) in cooperation with the German Society for Chemical Engineering and Biotechnology (DECHEMA) and the Society for Biochemistry and Molecular Biology (GBM). The conference was open to all fields of bioinformatics and theoretical systems biology.

Five satellite workshops that took place on 19 September 2012 placed thematic emphasis on diverse aspects of systems biology: "Systems Biology of Aging" organized by J. Sühnel, "Organ-oriented Systems Biology" by D. Driesch and R. Mrowka, "Network Reconstruction and Analysis in Systems Biology" by W. Wiechert and T. Lengauer, "Computational Proteomics and Metabolomics" by S. Böcker, and "Image-based Systems Biology" by M.T. Figge.

Six leading scientists accepted our invitation to give keynote lectures at the conference:

- Claude dePamphilis (Pennsylvania State University, University Park, USA)
  The draft genome sequence of *Amborella trichopoda* sheds light on the ancestral angiosperm genome
- Oliver Fiehn (University of California, Davis, USA)
  Comprehensive metabolomic databases and annotation workflows: The U.S. West Coast Metabolomics Center
- Arndt von Haeseler (Max F. Perutz Laboratories, Vienna, Austria)
  Exploring the sampling universe of RNA-seq
- Tom Kirkwood (Newcastle University, Newcastle, GB)
  Probing the deep complexity of ageing
- Erik van Nimwegen (University of Basel, Basel, Switzerland)
  A democracy of transcription factors: Inferring transcription regulatory interactions from high-throughput data
- Ruth Nussinov (National Cancer Institute, Frederick, USA)
  Structural proteome scale prediction of protein-protein interactions using interfaces

With the topics of these talks the meeting indeed succeeded in '***J***oining ***E***volution, ***N***etworks, and ***A***lgorithms', according to this year's conference motto.

From 39 submissions, the program committee selected 10 highlight papers and 11 regular papers as contributed talks for the conference. Additionally, about 95 poster abstracts were accepted for presentation. All regular papers are collected in this volume. The highlight papers, the abstracts from the invited speakers, and the poster abstracts are collected in a separate volume available online (www.gcb2012-jena.de).

We thank all members of the program committee as well as all local organizers and helpers for their efforts. We are also very grateful to the participants who presented their work at the lively panel sessions and poster party. Our special thanks go to the sponsors who supported the conference financially.

August 2012,
Sebastian Böcker, Franziska Hufsky, Kerstin Scheubert, Jana Schleicher, and Stefan Schuster

German Conference on Bioinformatics 2012 (GCB'12).
Editors: S. Böcker, F. Hufsky, K. Scheubert, J. Schleicher, S. Schuster
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Program Committee

Program Chairs
Sebastian Böcker
Stefan Schuster

Program Committee
Mario Albrecht
Rolf Backofen
Jan Baumbach
Michael Beckstette
Niko Beerenwinkel
Tim Beissbarth
Sebastian Böcker
Thomas Dandekar
Dmitrij Frishman
Georg Fuellen
Robert Giegerich
Ivo Große
Reinhard Guthke
Andreas Hildebrandt
Ivo Hofacker
Daniel Huson
Christoph Kaleta
Gunnar W. Klau
Ina Koch
Oliver Kohlbacher
Stefan Kurtz
Hans-Peter Lenhof
Manja Marz
Burkhard Morgenstern
Steffen Neumann
Kay Nieselt
Stefan Posch
Sven Rahmann
Matthias Rarey
Knut Reinert
Uwe Scholz
Dietmar Schomburg
Falk Schreiber
Michael Schroeder
Stefan Schuster
Joachim Selbig
Jens Stoye
Korbinian Strimmer
Andrew Torda
Martin Vingron
Arndt Von Haeseler
Edgar Wingender
Ralf Zimmer

Additional Referees
Nicola Bonzanni
Stefan Canzar
Christian Colmsee
Simona Constantinescu
Fabrizio Costa
Simone Daminelli
Mohammed El-Kebir
Andre Gohr
Giorgio Gonnella
Niels Grabe
Udo Hahn
Volker Heun
Christian Höner Zu Siederdissen
Benny Kneissl
Tina Koestler
Frank Kramer
Andreas Leha
Ioana Lemnian
Manuel Landesfeind
Fernando Meyer
Bui Quang Minh
Konstantin Riege
Fabian Schmich
Heiko Schmidt
Juliane Siebourg
Peter Stadler
Sascha Steinbiss
Patrick Trampert
George Tsatsaronis
Claus Weinholdt
Stephan Weise
Dirk Willrodt

German Conference on Bioinformatics 2012 (GCB'12).
Editors: S. Böcker, F. Hufsky, K. Scheubert, J. Schleicher, S. Schuster
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Supporters and Sponsors

Supporting Scientific Institutions

DECHEMA Gesellschaft für Chemische Technik und Biotechnologie e.V.
http://www.dechema.de/

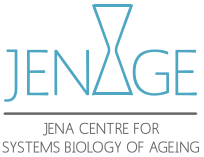
GBM Gesellschaft für Biochemie und Molekularbiologie e.V.
http://www.gbm-online.de/

Jena Centre for Bioinformatics
http://www.jcb-jena.de/

Hans-Knöll-Institute Jena
http://www.hki-jena.de/

Leibniz Institute for Age Research – Fritz Lipmann Institute
http://www.fli-leibniz.de/

Friedrich-Schiller-Universiät Jena
http://www.uni-jena.de/



Jena Centre for Systems Biology of Ageing
http://www.jenage.de/

Jena School for Microbial Communication
http://www.jsmc.uni-jena.de/

International Leibniz Research School for Microbial and Biomolecular Interactions
http://www.ilrs.hki-jena.de/

German Conference on Bioinformatics 2012 (GCB'12).
Editors: S. Böcker, F. Hufsky, K. Scheubert, J. Schleicher, S. Schuster
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Sponsors and Donors

BioControl Jena GmbH
http://www.biocontrol-jena.com/

Jenaer Universitäts-Buchhandlung Thalia
http://www.thalia.de/

TimeLogic biocomputing solutions
http://www.timelogic.com/

HMK Supercomputing GmbH
http://conveycomputer.com/lifesciences/

antibodies-online GmbH
http://www.antikoerper-online.de/

Index of Authors

A
Alexandrov, Theodore 39

B
Baumbach, Jan 57
Böck, Matthias 69
Böcker, Sebastian 12, 23

C
Chernyavsky, Ilya 39

D
D'Addario, Marianna 82

E
Elshamy, Samy 23
Esmaielbeiki, Reyhaneh 119

G
Gramzow, Lydia 93

H
Holzhütter, Hermann-Georg 1
Hoppe, Andreas 1
Horn, Fabian 93
Hufsky, Franziska 12, 23

K
Koch, Ina 108
Kottenhagen, Nina 93
Kreutzer, Christoph 57
Kriege, Nils 82
Köster, Johannes 49

L
Ludwig, Marcus 23

M
Maass, Peter 39
May, Patrick 108

N
Nebel, Jean-Christophe 119
Nikolenko, Sergey 39

P
Pesch, Robert 69
Pohl, Martin 93

R
Rahmann, Sven 49, 82
Röttger, Richard 57

S
Schäfer, Tim 108

T
Theißen, Günter 93

V
Vu, Thuy Duong 57

W
Wittkop, Tobias 57

Z
Zimmer, Ralf 69

German Conference on Bioinformatics 2012 (GCB'12).
Editors: S. Böcker, F. Hufsky, K. Scheubert, J. Schleicher, S. Schuster
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany