

METHODS IN MOLECULAR BIOLOGY™

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For other titles published in this series, go to
www.springer.com/series/7651

Biomedical Informatics

Edited by

Vadim Astakhov

*Biomedical Informatics Research Network Coordination Center, The Center for Research
in Biological Systems (CRBS), UCSD, La Jolla, CA*

 Humana Press

Editor

Vadim Astakhov
Biomedical Informatics
Research Network Coordination Center
The Center for Research
in Biological Systems (CRBS)
UCSD, La Jolla
CA

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-934115-63-3 e-ISBN 978-1-59745-524-4
DOI 10.1007/978-1-59745-524-4

Library of Congress Control Number: 2009926171

© Humana Press, a part of Springer Science+Business Media, LLC 2009

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Humana Press, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013 USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

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

Preface

In recent decades, bioinformatics has emerged as a dynamic area producing a wide spectrum of new approaches and playing an important role in modern biotechnological development. This book provides an overview of novel cyberinfrastructures which are currently under development in various biomedical centers around the world. The first three chapters demonstrate various architectures for large-scale collaboration which integrate scientific accord across multiple centers. The next five chapters demonstrate modern approaches currently used in various areas of bioinformatics. The final four chapters illustrate the software challenges and strategies to resolve those challenges for large-scale biomedical informatics projects.

Unlike many other research areas which use standard techniques, development of biomedical cyberinfrastructure requires considerable expertise. The knowledge compiled in this book is contributed by experts in their fields who have many years of experience not only in using well-known techniques but also in development and troubleshooting them. This book might be helpful for working specialists as well as for students taking courses in bioinformatics and computer science. The chapters are arranged in a logical order, progressing from an overview of general architecture to detailed aspects of bioinformatics research. Consequently, each chapter contains a step-by-step description of the methodology and provides invaluable tips and tricks to safeguard against potential mishaps and pitfalls. It is a book that will be an essential manual for newcomers to this area, as well as a valuable addition to the laboratories and offices of more experienced researchers.

I would like to express my gratitude to all the authors who have made this book possible, and thank Professor John Walker, the Series Editor, for the invitation to edit this volume. My thanks also go to Ms. Meghan Kennelly for her assistance during the final stages of editing.

La Jolla, CA

Vadim Astakhov

Contents

<i>Preface</i>	<i>v</i>
<i>Contributors</i>	<i>ix</i>
1. Management of Information in Distributed Biomedical Collaboratories <i>David B. Keator</i>	1
2. Enabling Public Data Sharing: Encouraging Scientific Discovery and Education <i>Christine Fennema-Notestine</i>	25
3. Mediator Infrastructure for Information Integration and Semantic Data Integration Environment for Biomedical Research <i>Jeffrey S. Grethe, Edward Ross, David Little, Brian Sanders, Amarnath Gupta, and Vadim Astakhov</i>	33
4. System Biology of Gene Regulation <i>Michael Baitaluk</i>	55
5. Current Computational Methods for Prioritizing Candidate Regulatory Polymorphisms <i>Stephen Montgomery</i>	89
6. Methods of Information Geometry in Computational System Biology (Consistency between Chemical and Biological Evolution) <i>Vadim Astakhov</i>	115
7. Protein Structure Prediction Based on Sequence Similarity <i>Lukasz Jaroszewski</i>	129
8. Applications of Bioinformatics to Protein Structures: How Protein Structure and Bioinformatics Overlap <i>Gye Won Han, Chris Rife, and Michael R. Sawaya</i>	157
9. Knowledge Discovery via Machine Learning for Neurodegenerative Disease Researchers <i>I. Burak Özyurt and Gregory G. Brown</i>	173
10. Brain Model of Text Animation as a Data Mining Strategy <i>Tamara Astakhova and Vadim Astakhov</i>	197
11. Processes Parallel Execution Using Grid Wizard Enterprise <i>Marco Ruiz</i>	219
12. Single Sign-On in a Grid Portal <i>Ramil V. Manansala</i>	239
<i>Index</i>	263

Contributors

- VADIM ASTAKHOV • *Biomedical Informatics Research Network Coordination Center, The Center for Research in Biological Systems (CRBS), UCSD, La Jolla, CA, USA*
- TAMARA ASTAKHOVA • *Joint Center for Structural Genomics, University of California – San Diego, La Jolla, CA, USA*
- MICHAEL BAITALUK • *San Diego Supercomputer Center, University of California – San Diego, La Jolla, CA, USA*
- GREGORY G. BROWN • *Department of Psychiatry, University of California – San Diego, La Jolla, CA, USA*
- CHRISTINE FENNEMA-NOTESTINE • *Departments of Psychiatry and Radiology, University of California – San Diego, La Jolla, CA, USA*
- JEFFREY S. GRETHE • *Biomedical Informatics Research Network, La Jolla, CA, USA*
- AMARNATH GUPTA • *San Diego Supercomputer Center, University of California – San Diego, La Jolla, CA, USA*
- GYE WON HAN • *Joint Center for Structural Genomics (JCSG), The Scripps Research Institute, La Jolla, CA, USA*
- LUKASZ JAROSZEWSKI • *The Burnham Institute, La Jolla, CA, USA*
- DAVID B. KEATOR • *Brain Imaging Center, University of California – Irvine, Irvine, CA, USA*
- DAVID R. LITTLE • *The Center for Research in Biological Systems (CRBS), La Jolla, CA, USA*
- RAMIL V. MANANSALA • *Center for Research in Biological Systems, University of California – San Diego, La Jolla, CA, USA*
- STEPHEN MONTGOMERY • *Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK*
- I. BURAK ÖZYURT • *Department of Psychiatry, University of California – San Diego, La Jolla, CA, USA*
- CHRIS RIFE • *Joint Center for Structural Genomics (JCSG), Stanford Synchrotron Radiation Laboratory, Stanford University, Menlo Park, CA, USA*
- EDWARD ROSS • *Biomedical Informatics Research Network, San Diego, CA, USA*
- MARCO RUIZ • *Biomedical Informatics Research Network, San Diego, CA, USA*
- BRIAN SANDERS • *The Center for Research in Biological Systems (CRBS), La Jolla, CA, USA*
- MICHAEL R. SAWAYA • *Howard Hughes Medical Institute, UCLA-DOE Institute of Genomics and Proteomics, Los Angeles, CA, USA*