

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

INTEGRATING OMICS DATA

In most modern biomedical research projects, application of high-throughput genomic, proteomic, and transcriptomic experiments has gradually become an inevitable component. Popular technologies include microarray, next generation sequencing, mass spectrometry, and proteomics assays. As the technologies have become mature and the price affordable, omics data are rapidly generated, and the problem of information integration and modeling of multi-lab and/or multi-omics data is becoming a growing one in the bioinformatics field.

This book provides comprehensive coverage of these topics and will have a long-lasting impact on this evolving subject. Each chapter, written by a leader in the field, introduces state-of-the-art methods to handle information integration, experimental data, and database problems of omics data.

George Tseng completed his ScD in biostatistics with a concentration in genomics from the Harvard School of Public Health. He is currently a professor of biostatistics, human genetics, and computational and systems biology at the University of Pittsburgh. His research interests focus on statistical and computational method development for analyzing high-throughput omics data.

Debashis Ghosh is professor and chair of the Department of Biostatistics and Informatics at the Colorado School of Public Health, located at the University of Colorado Anschutz Medical Campus. Ghosh was involved with the development of the statistical methods for ONCOMINE, an online data-mining platform used in cancer research and genetics. His research in bioinformatics methodology has been funded by NIH and NSF grants over the last ten years. He has published more than 160 peer-reviewed articles, commentaries, and book chapters in statistical and scientific literature.

Xianghong Jasmine Zhou completed her PhD at the Swiss Federal Institute of Technology (ETH Zurich) and conducted her post-doc training at Harvard University. She is currently Director and Professor of the Computational Biology and Bioinformatics program at the University of Southern California. Dr. Zhou is the PI of the NIH center for knowledge base on disease connections within the MAPGen consortium. She heads the laboratory of computational integrative genomics, addressing the “Big Data” challenges brought by the enormous amount of extremely diverse genomic data in public repositories. She was a recipient of several awards, including an Alfred Sloan fellowship and a NSF Career award.

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

INTEGRATING OMICS DATA

Edited by

GEORGE TSENG

University of Pittsburgh

DEBASHIS GHOSH

University of Colorado

XIANGHONG JASMINE ZHOU

University of Southern California



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

CAMBRIDGE
UNIVERSITY PRESS

32 Avenue of the Americas, New York, NY 10013-2473, USA

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107069114

© Cambridge University Press 2015

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2015

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication Data

Tseng, George.

Integrating omics data / George Tseng, University of Pittsburgh, Debashis Ghosh, University of Colorado, Xianghong Jasmine Zhou, University of Southern California.

pages cm

ISBN 978-1-107-06911-4 (hardback)

1. Genomics – Statistical methods. 2. Meta-analysis. I. Title.

QH438.4.S73T74 2015

572.8'6–dc23 2014048684

ISBN 978-1-107-06911-4 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate.

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

Contents

| | |
|---|-----------------|
| <i>Contributors</i> | <i>page vii</i> |
| Introduction | 1 |
| Part A: Horizontal Meta-Analysis | |
| 1. Meta-Analysis of Genome-Wide Association Studies: A Practical Guide <i>Wei Chen</i> | 9 |
| 2. MetaOmics: Transcriptomic Meta-Analysis Methods for Biomarker Detection, Pathway Analysis and Other Exploratory Purposes <i>SungHwan Kim, Zhiguang Huo, Yongseok Park, and George C. Tseng</i> | 39 |
| 3. Integrative Analysis of Many Biological Networks to Study Gene Regulation <i>Wenyuan Li, Chao Dai, and Xianghong Jasmine Zhou</i> | 68 |
| 4. Network Integration of Genetically Regulated Gene Expression to Study Complex Diseases <i>Zhidong Tu, Bin Zhang, and Jun Zhu</i> | 88 |
| 5. Integrative Analysis of Multiple ChIP-X Data Sets Using Correlation Motifs <i>Hongkai Ji and Yingying Wei</i> | 110 |
| Part B: Vertical Integrative Analysis (General Methods) | |
| 6. Identify Multi-Dimensional Modules from Diverse Cancer Genomics Data <i>Shihua Zhang, Wenyuan Li, and Xianghong Jasmine Zhou</i> | 135 |
| 7. A Latent Variable Approach for Integrative Clustering of Multiple Genomic Data Types <i>Ronglai Shen</i> | 155 |
| 8. Penalized Integrative Analysis of High-Dimensional Omics Data <i>Jin Liu, Xingjie Shi, Jian Huang, and Shuangge Ma</i> | 174 |

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

vi

Contents

- | | | | |
|---|---|--|-----|
| 9. | A Bayesian Graphical Model for Integrative Analysis of TCGA Data: BayesGraph for TCGA Integration | <i>Yanxun Xu, Yitan Zhu, and Yuan Ji</i> | 205 |
| 10. | Bayesian Models for Flexible Integrative Analysis of Multi-Platform Genomics Data | <i>Elizabeth J. McGuffey, Jeffrey S. Morris, Ganiraju C. Manyam, Raymond J. Carroll, and Veerabhadran Baladandayuthapani</i> | 221 |
| 11. | Exploratory Methods to Integrate Multisource Data | <i>Eric F. Lock and Andrew B. Nobel</i> | 242 |
| | | | |
| Part C: Vertical Integrative Analysis (Methods Specialized to Particular Data Types) | | | |
| 12. | eQTL and Directed Graphical Model | <i>Wei Sun and Min Jin Ha</i> | 271 |
| 13. | MicroRNAs: Target Prediction and Involvement in Gene Regulatory Networks | <i>Panayiotis V. Benos</i> | 291 |
| 14. | Integration of Cancer Omics Data into a Whole-Cell Pathway Model for Patient-Specific Interpretation | <i>Charles Vaske, Sam Ng, Evan Paull, and Joshua Stuart</i> | 310 |
| 15. | Analyzing Combinations of Somatic Mutations in Cancer Genomes | <i>Mark D. M. Leiserson and Benjamin J. Raphael</i> | 337 |
| 16. | A Mass-Action-Based Model for Gene Expression Regulation in Dynamic Systems | <i>Guoshou Teo, Christine Vogel, Debashis Ghosh, Sinae Kim, and Hyungwon Choi</i> | 362 |
| 17. | From Transcription Factor Binding and Histone Modification to Gene Expression: Integrative Quantitative Models | <i>Chao Cheng</i> | 380 |
| 18. | Data Integration on Noncoding RNA Studies | <i>Zhou Du, Teng Fei, Myles Brown, X. Shirley Liu, and Yiwen Chen</i> | 403 |
| 19. | Drug-Pathway Association Analysis: Integration of High-Dimensional Transcriptional and Drug Sensitivity Profile | <i>Cong Li, Can Yang, Greg Hather, Ray Liu, and Hongyu Zhao</i> | 425 |
| | | | |
| | <i>Index</i> | | 445 |
| <i>Color plates follow page 134</i> | | | |

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

Contributors

Veerabhadran Baladandayuthapani, Department of Biostatistics, UT MD Anderson Cancer Center, Houston, TX

Panayiotis V. Benos, Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA

Myles Brown, Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA

Raymond J. Carroll, Department of Statistics, Texas A&M University, College Station, TX

Wei Chen, Department of Pediatrics, University of Pittsburgh, Pittsburgh, PA

Yiwen Chen, Department of Bioinformatics and Computational Biology, Division of Quantitative Sciences, UT MD Anderson Cancer Center, Houston, TX

Chao Cheng, Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH; Institute for Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth, Lebanon, NH; Norris Cotton Cancer Center, Geisel School of Medicine at Dartmouth, Lebanon, NH

Hyungwon Choi, Saw Swee Hock School of Public Health, National University of Singapore

Chao Dai, Molecular and Computational Biology, University of Southern California, Los Angeles, CA

Zhou Du, Howard Hughes Medical Institute, Program in Cellular and Molecular Medicine, Boston Children's Hospital, and Department of Genetics, Harvard Medical School, Boston, MA

Teng Fei, Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA

Debashis Ghosh, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Denver

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

viii

Contributors

Min Jin Ha, Department of Biostatistics, MD Anderson Cancer Center, Houston, TX

Greg Hather, Takeda Pharmaceuticals International Co., Cambridge, MA

Jian Huang, Department of Statistics and Actuarial Science, University of Iowa

Zhiguang Huo, Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA

Hongkai Ji, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Yuan Ji, Program of Computational Genomics and Medicine, NorthShore University HealthSystem; Department of Public Health Sciences, University of Chicago

Sinae Kim, Department of Biostatistics, School of Public Health, Rutgers University

SungHwan Kim, Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA

Mark D. M. Leiserson, Department of Computer Science and Center for Computational Molecular Biology, Brown University, Providence, RI

Cong Li, Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT

Wenyuan Li, Molecular and Computational Biology, University of Southern California, Los Angeles, CA

Jin Liu, Centre for Quantitative Medicine, Duke-NUS Graduate Medical School

Ray Liu, Takeda Pharmaceuticals International Co., Cambridge, MA

X. Shirley Liu, Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA

Eric F. Lock, Division of Biostatistics, University of Minnesota, Minneapolis, MN

Shuangge Ma, Department of Biostatistics, Yale University School of Statistics; Capital University of Economics and Business, China

Ganiraju C. Manyam, Department of Bioinformatics and Computational Biology, UT MD Anderson Cancer Center, Houston, TX

Elizabeth J. McGuffey, Mathematics Department, United States Naval Academy, Annapolis, MD

Jeffrey S. Morris, Department of Biostatistics, UT MD Anderson Cancer Center, Houston, TX

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

Contributors

ix

Sam Ng, Department of Biomolecular Engineering, Center for Biomolecular Science and Engineering, University of California at Santa Cruz, Santa Cruz, CA

Andrew B. Nobel, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC

Yongseok Park, Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA

Benjamin J. Raphael, Department of Computer Science and Center for Computational Molecular Biology, Brown University, Providence, RI

Ronglai Shen, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

Xingjie Shi, Department of Statistics, Nanjing University of Finance and Economics, China School of Statistics and Management, Shanghai University of Finance and Economics, China

Joshua Stuart, Department of Biomolecular Engineering, Center for Biomolecular Science and Engineering, University of California at Santa Cruz, Santa Cruz, CA

Wei Sun, Department of Biostatistics, Department of Genetics, University of North Carolina, Chapel Hill, Chapel Hill, NC

Guoshou Teo, Saw Swee Hock School of Public Health, National University of Singapore

George C. Tseng, Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA

Zhidong Tu, Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

Charles Vaske, NantOmics, Culver City, CA

Christine Vogel, Center for Systems Biology, Department of Biology, New York University

Yingying Wei, Department of Statistics, The Chinese University of Hong Kong, Hong Kong

Yanxun Xu, Department of Applied Mathematics and Statistics, Johns Hopkins University

Can Yang, Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

Bin Zhang, Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

Cambridge University Press

978-1-107-06911-4 - Integrating Omics Data

Edited by George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou

Frontmatter

[More information](#)

x

Contributors

Shihua Zhang, National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing

Hongyu Zhao, Department of Biostatistics, Yale School of Public Health, New Haven, CT; Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT

Xianghong Jasmine Zhou, Molecular and Computational Biology, University of Southern California, Los Angeles, CA

Jun Zhu, Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

Yitan Zhu, Program of Computational Genomics and Medicine, NorthShore University HealthSystem