

METHODS IN MOLECULAR BIOLOGY™

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

For further volumes:
<http://www.springer.com/series/7651>

Next Generation Microarray Bioinformatics

Methods and Protocols

Edited by

Junbai Wang

Department of Pathology, Oslo University Hospital, Radium Hospital, Montebello, Oslo, Norway

Aik Choon Tan

*Division of Medical Oncology, Department of Medicine, School of Medicine,
University of Colorado Anschutz Medical Campus, Aurora, CO, USA*

Tianhai Tian

School of Mathematical Sciences, Monash University, Melbourne, VIC, Australia

 **Humana Press**

Editors

Junbai Wang, Ph.D.
Department of Pathology
Oslo University Hospital
Radium Hospital
Montebello, Oslo, Norway
junbai.wang@rr-research.no

Aik Choon Tan, Ph.D.
Division of Medical Oncology
Department of Medicine School of Medicine
University of Colorado Anschutz Medical Campus
Aurora, CO, USA
aikchoon.tan@ucdenver.edu

Tianhai Tian, Ph.D.
School of Mathematical Sciences
Monash University
Melbourne, VIC, Australia
tianhai.tian@monash.edu

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-61779-399-8 e-ISBN 978-1-61779-400-1
DOI 10.1007/978-1-61779-400-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011943561

© Springer Science+Business Media, LLC 2012

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.

Printed on acid-free paper

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

Preface

The twenty-first century is the time of excitement and optimism for biomedical research. Since the completion of the human genome project in 2001, we are entering into the postgenome era where the key research efforts are now interpreting and making sense of these massive genomic data, in order to translate into disease treatment and management. Over the past decade, DNA-based microarrays have been the assays of choice for high-throughput studies of gene expression. Microarray-based expression profiling was provided, for the first time, by means of monitoring genome-wide gene expression changes in a single experiment. Though microarray technology has been widely employed to reveal molecular portraits of gene expression in various cancers' subtypes and correlations with disease progression as well as response to drug treatments, it is not limited to measure gene expression. As the technology became established in early 2000, researchers began to use microarrays to measure other important biological phenomena. For example, (1) Microarrays are being used to genotype single-nucleotide polymorphisms (SNPs) by hybridizing the DNA of individuals to arrays of oligonucleotides representing different polymorphic alleles. The SNP microarray has accelerated genome-wide association studies over the last 5 years, and many loci that are associated with diseases have been discovered and validated. Similarly, another innovative application of the SNP microarray is to interrogate allele-specific expression for identifying disease-associated genes. (2) Array-comparative genomic hybridization (aCGH) is being used to detect genomic structural variations, such as segments of the genome that have varying numbers of copies in different individuals. (3) Epigenetic modifications such as methylation at CpG sites can also be assessed by microarray. (4) Using ChIP-chip assay, genome-wide protein–DNA interactions and chromatin modifications can be profiled by microarrays. (5) More recently, microarray has been used to measure genome-wide microRNA expression patterns to reveal the regulatory role of these noncoding RNAs in disease states. Obviously, the progress of microarray applications is tightly associated with the development of novel computational and statistical methods to analyze and interpret these data sets.

Recent improvements in the efficiency, quality, and cost of genome-wide sequencing have prompted biologists and biomedical researchers to move away from microarray-based technology to ultrahigh-throughput, massively parallel genomic sequencing (Next Generation Sequencing, NGS) technology. NGS technology opens up new research avenues for the investigation of a wide range of biological and medical questions across the entire genome at single base resolution; for example, sequencing of several human genomes, monitoring of genome-wide transcription levels (RNA-seq), understanding of epigenetic phenomena, DNA–protein interactions (ChIP-seq), and de novo sequencing of several genomes. Despite the differences in the underlying sequencing technologies of various NGS machines, the common output from them are the capability to generate tens of millions of short reads (tags) from each experimental run. Thus, NGS technology shifts the bottleneck in sequencing processes from experimental data production to computationally intensive informatics-based data analysis. As in the early days of microarray data analysis, novel computational and statistical methods tailored to NGS are urgently needed for drawing meaningful and accurate conclusions from the massive short reads. Furthermore, it is expected that NGS technology may eventually replace microarray technology in the

next decade, which will grow from a pioneering method applied by innovators at the cutting edge research to a ubiquitous technique that will allow researchers to investigate “big-picture” questions in biology at much higher resolution.

This book, Next Generation Microarray Bioinformatics, is our attempt to bring together current computational and statistical methods in analyzing and interpreting both microarray and NGS data. Here, we have compiled and edited 26 chapters that cover a wide range of methodological and application topics in microarray and NGS bioinformatics. These chapters are organized into five thematic sections: (1) Resources for Microarray Bioinformatics; (2) Microarray Data Analysis; (3) Microarray Bioinformatics in Systems Biology; (4) Next Generation Sequencing Data Analysis; and (5) Emerging Applications of Microarray and Next Generation Sequencing. Each chapter is a self-contained review of a specific methodological or application topic. Every chapter typically starts with a brief review of a particular subject, then describes in detail the computational and statistical techniques used to solve the biological questions, and finally discusses the computational results generated by these bioinformatics tools. Therefore, the reader need not read the chapters in a sequential manner. We expect this book would be a valuable methodological resource not only to molecular biologists and computational biologists who are interested in understanding the principle of these methods and designing future research project, but also to computer scientists and statisticians who work in a microarray core facility or other similar organizations that provide service for the high-throughput experiment community.

The first section of this book contains three important resource chapters of microarray and NGS bioinformatics community. The introductory chapter provides an overview on the current state of microarray technologies and is contributed by Kuo and colleagues. The second chapter is contributed by the KEGG group. The KEGG database represents one of the earliest databases to store, manage, integrate, and visualize genomics data. In this chapter, Kotera and colleagues provide the latest developments of the KEGG efforts in analyzing and interpreting omics data. The NCBI Gene Expression Omnibus (GEO) group writes the third chapter in this section, which is one of the major data repositories for high-throughput microarray and next-generation sequencing data. White and Barrett describe various strategies to explore functional genomics data sets in the GEO database.

The second section of this book consists of eight chapters that describe methods to analyze microarray data from the top down approach. The first chapter, contributed by Van Loo and colleagues, that described a novel R-package ASCAT specifically designed to delineate genomic aberration in cancer genomes from SNP microarrays. Then Cheung, Meng, and Huang wrote the following two chapters of advanced machine learning methods in investigating disease classification and time-series microarray data analysis, respectively. Lin and colleagues provide a tutorial on a novel R-package, GeneAnswers, to perform gene-concept network analysis in the next chapter. Nair contributed the next chapter, which emphasizes the utility of R/Bioconductor, an open source software for bioinformatics, in the analysis and interpretation of splice isoforms in microarray. The next three chapters focusing on cross-platform comparisons of microarray data and integrative approaches for microarray data analysis were delivered by Li et al., Hovig et al., and Huttenhower et al., respectively.

The third section of this book concentrates on the bottom-up approaches for establishing different types of models based on microarray expression datasets in which the number of genes is much larger than that of samples. The first chapter written by Yu and colleagues discussed a general profiling method to estimate parameters in the ordinary differential

equation models from the time-course gene expression data. To deal with inhomogeneity and nonstationarity in temporal processes, Husmeier and colleagues described the inhomogeneous dynamic Bayesian networks which allow the network structure to change over time in the second chapter. Castelo and Roverato contributed the third chapter that introduced an R package of a graphic approach for inferring regulatory networks from microarray datasets. Wang and Tian contribute the final chapter of this section. They introduced a nonlinear model, which can be used to infer the transcriptional factor activities from the microarray expression data of the target genes as well as to predict the regulatory relationship between transcriptional factors and their target genes.

The fourth section of this book contains six chapters, specifically devoted to NGS data analysis. It starts from an overview of the NGS data analysis by Gogol-Döring and Chen, which includes the basic steps for analyzing NGS such as quality check and mapping to a reference genome. The second chapter is written by Sandber and colleagues, where the authors provide a detailed illustration of how to analyze gene expression using RNA-Sequencing data through several real examples. Lin and colleagues contributed to the third chapter that introduces the low level ChIP-seq data analysis such as preprocessing, normalization, differential identification, and binding pattern characterization. The fourth chapter is contributed by Xu and Sung, in which reader will find how to use Hidden Markov Model to identify differential histone modification sites from ChIP-seq data. The last two chapters describe two software packages (SISSRs developed by Narlikar and Jothi and ChIPMotifs developed by Jin and colleagues) that are designed to study protein–DNA interactions (e.g., peak finder and de novo motif discovery) by analyzing ChIP-based high-throughput experiments.

The final section of this book contains five methodological chapters that cover the emerging applications of microarray and next-generation sequencing in biomedical researchers. In Wei's chapter, it describes Hidden Markov Models for controlling false-discovery rate in genome-wide association analysis. Tan describes Gene Set Top Scoring Pairs (GSTSP), a novel machine learning method in identifying discriminative gene set classifier, based on the relative expression concept. In the next chapter, Wu and Ji focus on JAMIE, a software tool that can perform jointly analysis on multiple ChIP-chip experiments. In the chapter written by Pelligrini and Ferrari, they described an overview on bioinformatics methods in analyzing epigenetic data. The final chapter is a bioinformatics workflow for the analysis and interpretation of genome-wide shRNA synthetic lethal screen based on next-generation sequencing written by Kim and Tan.

We would like to acknowledge the contribution of all authors to the conception and completion of this book. We would like to thank Prof. John M. Walker, the Methods in Molecular Biology series editor, for entrusting and giving us this opportunity to edit this volume. We also like to thank the staff at the Humana Press and Springer publishing company for their professional assistance in preparing this volume. Finally, we would like to thank our families for their love and support.

Oslo, Norway
Aurora, CO, USA
Melbourne, VIC, Australia

Junbai Wang
Aik Choon Tan
Tianhai Tian

Contents

<i>Preface</i>	v
<i>Contributors</i>	xiii

PART I INTRODUCTION AND RESOURCES FOR MICROARRAY BIOINFORMATICS

1 A Primer on the Current State of Microarray Technologies	3
<i>Alexander J. Trachtenberg, Jae-Hyung Robert, Azza E. Abdalla, Andrew Fraser, Steven Y. He, Jessica N. Lacy, Chiara Rivas-Morello, Allison Truong, Gary Hardiman, Lucila Ohno-Machado, Fang Liu, Eivind Hovig, and Winston Patrick Kuo</i>	
2 The KEGG Databases and Tools Facilitating Omics Analysis: Latest Developments Involving Human Diseases and Pharmaceuticals	19
<i>Masaaki Kotera, Mika Hirakawa, Toshiaki Tokimatsu, Susumu Goto, and Minoru Kanehisa</i>	
3 Strategies to Explore Functional Genomics Data Sets in NCBI's GEO Database	41
<i>Stephen E. Wilhite and Tanya Barrett</i>	

PART II MICROARRAY DATA ANALYSIS (TOP-DOWN APPROACH)

4 Analyzing Cancer Samples with SNP Arrays	57
<i>Peter Van Loo, Gro Nilsen, Silje H. Nordgard, Hans Kristian Moen Vollan, Anne-Lise Børresen-Dale, Vessela N. Kristensen, and Ole Christian Lingjærde</i>	
5 Classification Approaches for Microarray Gene Expression Data Analysis	73
<i>Leo Wang-Kit Cheung</i>	
6 Biclustering of Time Series Microarray Data	87
<i>Jia Meng and Yufei Huang</i>	
7 Using the Bioconductor GeneAnswers Package to Interpret Gene Lists	101
<i>Gang Feng, Pamela Shaw, Steven T. Rosen, Simon M. Lin, and Warren A. Kibbe</i>	
8 Analysis of Isoform Expression from Splicing Array Using Multiple Comparisons	113
<i>T. Murlidharan Nair</i>	
9 Functional Comparison of Microarray Data Across Multiple Platforms Using the Method of Percentage of Overlapping Functions	123
<i>Zhiguang Li, Joshua C. Kwekel, and Tao Chen</i>	
10 Performance Comparison of Multiple Microarray Platforms for Gene Expression Profiling	141
<i>Fang Liu, Winston P. Kuo, Tor-Kristian Jenssen, and Eivind Hovig</i>	
11 Integrative Approaches for Microarray Data Analysis	157
<i>Levi Waldron, Hilary A. Collier, and Curtis Huttenhower</i>	

PART III MICROARRAY BIOINFORMATICS IN SYSTEMS BIOLOGY
(BOTTOM-UP APPROACH)

12 Modeling Gene Regulation Networks Using Ordinary
Differential Equations 185
Jiguo Cao, Xin Qi, and Hongyu Zhao

13 Nonhomogeneous Dynamic Bayesian Networks in Systems Biology 199
Sophie Lèbre, Frank Dondelinger, and Dirk Husmeier

14 Inference of Regulatory Networks from Microarray Data
with R and the Bioconductor Package qgraph 215
Robert Castelo and Alberto Roverato

15 Effective Non-linear Methods for Inferring Genetic Regulation
from Time-Series Microarray Gene Expression Data 235
Junbai Wang and Tianhai Tian

PART IV NEXT GENERATION SEQUENCING DATA ANALYSIS

16 An Overview of the Analysis of Next Generation Sequencing Data 249
Andreas Gogol-Döring and Wei Chen

17 How to Analyze Gene Expression Using RNA-Sequencing Data 259
Daniel Ramsköld, Ersen Kavak, and Rickard Sandberg

18 Analyzing ChIP-seq Data: Preprocessing, Normalization,
Differential Identification, and Binding Pattern Characterization 275
Cenny Taslim, Kun Huang, Tim Huang, and Shili Lin

19 Identifying Differential Histone Modification Sites from ChIP-seq Data 293
Han Xu and Wing-Kin Sung

20 ChIP-Seq Data Analysis: Identification of Protein–DNA Binding
Sites with SISSRs Peak-Finder 305
Leelavati Narlikar and Raja Jothi

21 Using ChIPMotifs for De Novo Motif Discovery of OCT4
and ZNF263 Based on ChIP-Based High-Throughput Experiments 323
*Brian A. Kennedy, Xun Lan, Tim H.-M. Huang,
Peggy J. Farnham, and Victor X. Jin*

PART V EMERGING APPLICATIONS OF MICROARRAY AND
NEXT GENERATION SEQUENCING

22 Hidden Markov Models for Controlling False Discovery Rate
in Genome-Wide Association Analysis 337
Zhi Wei

23 Employing Gene Set Top Scoring Pairs to Identify Deregulated
Pathway-Signatures in Dilated Cardiomyopathy from Integrated
Microarray Gene Expression Data 345
Aik Choon Tan

24 JAMIE: A Software Tool for Jointly Analyzing Multiple
ChIP-chip Experiments 363
Hao Wu and Hongkai Ji

25 Epigenetic Analysis: ChIP-chip and ChIP-seq 377
Matteo Pellegrini and Roberto Ferrari

26 BiNGS!SL-seq: A Bioinformatics Pipeline for the Analysis
and Interpretation of Deep Sequencing Genome-Wide
Synthetic Lethal Screen 389
Jihye Kim and Aik Choon Tan

Index 399

Contributors

- AZZA E. ABDALLA · *Department of Biology, University of South Carolina, Columbia, SC, USA*
- TANYA BARRETT · *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA*
- ANNE-LISE BØRRESEN-DALE · *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway; Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway*
- JIGUO CAO · *Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada*
- ROBERT CASTELO · *Research Program on Biomedical Informatics, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, and Institut Municipal d'Investigació Mèdica, Barcelona, Spain*
- TAO CHEN · *Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR, USA*
- WEI CHEN · *Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Berlin, Germany*
- LEO WANG-KIT CHEUNG · *Bioinformatics Core, Department of Preventive Medicine and Epidemiology, Stritch School of Medicine, Loyola University Medical Center, Maywood, IL, USA*
- HILARY A. COLLER · *Department of Molecular Biology, Princeton University, Princeton, NJ, USA*
- FRANK DONDELINGER · *Biomathematics and Statistics Scotland, Scotland, UK School of Informatics, University of Edinburgh, Edinburgh, UK*
- PEGGY J. FARNHAM · *Department of Biochemistry & Molecular Biology, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA*
- GANG FENG · *Biomedical Informatics Center, Clinical and Translational Sciences Institute, Northwestern University, Chicago, IL, USA*
- ROBERTO FERRARI · *Department of Biological Chemistry, University of California, Los Angeles, CA, USA*
- ANDREW FRASER · *Department of Allergy and Inflammation, BIDMC, Boston, MA, USA*
- ANDREAS GOGOL-DÖRING · *Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Berlin, Germany*
- SUSUMU GOTO · *Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan*
- GARY HARDIMAN · *Department of Allergy and Inflammation, BIDMC, Boston, MA, USA*
- STEVEN Y. HE · *Department of Medicine, University of California San Diego, San Diego, CA, USA*

- MIKA HIRAKAWA · *Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan*
- EIVIND HOVIG · *Departments of Tumor Biology and Medical Informatics, Institute for Cancer Research, Norwegian Radium Hospital, Montebello, Oslo, Norway*
- KUN HUANG · *Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA*
- TIM H.-M. HUANG · *Department of Molecular Virology, Immunology & Medical Genetics, The Ohio State University, Columbus, OH, USA*
- YUFEI HUANG · *Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX, USA; Greehey Children's Cancer Research Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*
- DIRK HUSMEIER · *Biomathematics and Statistics Scotland, Scotland, UK*
- CURTIS HUTTENHOWER · *Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA*
- TOR-KRISTIAN JENSSEN · *PubGene AS, Vinderen, Oslo, Norway*
- HONGKAI JI · *Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA*
- VICTOR X. JIN · *Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA*
- RAJA JOTHI · *National Institutes of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA*
- MINORU KANEHISA · *Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan*
- ERSEN KAVAK · *Department of Cell and Molecular Biology, Karolinska Institutet and Ludwig Institute for Cancer Research, Stockholm, Sweden*
- BRIAN A. KENNEDY · *Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA*
- WARREN A. KIBBE · *Biomedical Informatics Center, Clinical and Translational Sciences Institute, Northwestern University, Chicago, IL, USA*
- JIHYE KIM · *Division of Medical Oncology, Department of Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA*
- MASAAKI KOTERA · *Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan*
- VESSELA N. KRISTENSEN · *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway; Institute for Clinical Medicine, Institute for Clinical Epidemiology and Molecular Biology (EpiGen), Akershus University Hospital, Faculty of Medicine, University of Oslo, Nordbyhagen, Norway*
- WINSTON PATRICK KUO · *Harvard Catalyst – Laboratory for Innovative Translational Technologies, Harvard Medical School, Boston, MA, USA; Department of Developmental Biology, Harvard School of Dental Medicine, Boston, MA, USA*
- JOSHUA C. KWEKEL · *Division of System Biology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR, USA*

- JESSICA N. LACY · *Harvard Catalyst – Laboratory for Innovative Translational Technologies, Harvard Medical School, Boston, MA, USA*
- XUN LAN · *Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA*
- SOPHIE LÈBRE · *Université de Strasbourg, LSIIT – UMR 7005, Strasbourg, France*
- ZHIGUANG LI · *Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR, USA*
- SHILI LIN · *Department of Statistics, The Ohio State University, Columbus, OH, USA*
- SIMON M. LIN · *Biomedical Informatics Center, Clinical and Translational Sciences Institute, Northwestern University, Chicago, IL, USA*
- OLE CHRISTIAN LINGJÆRDE · *Biomedical Research Group, Department of Informatics, Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway*
- FANG LIU · *Department of Tumor Biology, Institute for Cancer Research, Norwegian Radium Hospital, Montebello, Oslo, Norway; PubGene AS, Vinderen, Oslo, Norway*
- JIA MENG · *Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX, USA*
- T. MURLIDHARAN NAIR · *Departments of Biological Sciences, Computer Science/Informatics, Indiana University South Bend, Bloomington, IN, USA*
- LEELAVATI NARLIKAR · *National Institutes of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA; Centre for Modeling and Simulation, University of Pune, Pune, Maharashtra, India*
- GRO NILSEN · *Biomedical Research Group, Department of Informatics, Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway*
- SILJE H. NORDGARD · *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway*
- LUCILA OHNO-MACHADO · *Division of Biomedical Informatics, University of California San Diego, San Diego, CA, USA*
- MATTEO PELLEGRINI · *Department of Molecular, Cell and Developmental, University of California, Los Angeles, CA, USA*
- XIN QI · *School of Public Health, Yale University, New Haven, CT, USA*
- DANIEL RAMSKÖLD · *Department of Cell and Molecular Biology, Karolinska Institutet and Ludwig Institute for Cancer Research, Stockholm, Sweden*
- CHIARA RIVAS-MORELLO · *Harvard Catalyst – Laboratory for Innovative Translational Technologies, Harvard Medical School, Boston, MA, USA*
- JAE-HYUNG ROBERT · *Department of Developmental Biology, Harvard School of Dental Medicine, Boston, MA, USA*
- STEVEN T. ROSEN · *Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL, USA*
- ALBERTO ROVERATO · *Department of Statistical Science, Università di Bologna, Bologna, Italy*
- RICKARD SANDBERG · *Department of Cell and Molecular Biology, Karolinska Institutet and Ludwig Institute for Cancer Research, Stockholm, Sweden*

- PAMELA SHAW · *Galter Health Sciences Library, Northwestern University, Chicago, IL, USA*
- WING-KIN SUNG · *Department of Computational and Mathematical Biology, Genome Institute of Singapore, Singapore, Singapore; School of Computing, National University of Singapore, Singapore, Singapore*
- AIK CHOON TAN · *Division of Medical Oncology, Department of Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA*
- CENNY TASLIM · *Department of Molecular Virology, Immunology & Medical Genetics, The Ohio State University, Columbus, OH, USA; Department of Statistics, The Ohio State University, Columbus, OH, USA*
- TIANHAI TIAN · *School of Mathematical Sciences, Monash University, Melbourne, VIC, Australia*
- TOSHIAKI TOKIMATSU · *Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan*
- ALEXANDER J. TRACHTENBERG · *Harvard Catalyst – Laboratory for Innovative Translational Technologies, Harvard Medical School, Boston, MA, USA*
- ALLISON TRUONG · *Department of Biology, University of California Los Angeles, Los Angeles, CA, USA*
- PETER VAN LOO · *Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; Department of Molecular and Developmental Genetics, VIB, Leuven, Belgium; Department of Human Genetics, University of Leuven, Leuven, Belgium*
- HANS KRISTIAN MOEN VOLLAN · *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway; Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway; Division of Surgery and Cancer, Department of Breast and Endocrine Surgery, Oslo University Hospital Ullevål, Oslo, Norway*
- LEVI WALDRON · *Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA*
- JUNBAI WANG · *Department of Pathology, Oslo University Hospital, Radium Hospital, Montebello, Oslo, Norway*
- ZHI WEI · *Department of Computer Science, New Jersey Institute of Technology, Newark, NJ, USA*
- STEPHEN E. WILHITE · *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA*
- HAO WU · *Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, USA*
- HAN XU · *Department of Computational and Mathematical Biology, Genome Institute of Singapore, Singapore, Singapore*
- HONGYU ZHAO · *School of Public Health, Yale University, New Haven, CT, USA*