Preeti Arivaradarajan · Gauri Misra
Editors

Omics Approaches, Technologies And Applications

Integrative Approaches For Understanding OMICS Data

Springer

Omics Approaches, Technologies And Applications

Preeti Arivaradarajan • Gauri Misra
Editors

Omics Approaches, Technologies And Applications

Integrative Approaches For Understanding OMICS Data

Springer

Editors
Preeti Arivaradarajan
Amity Institute of Biotechnology
Amity University
Noida, Uttar Pradesh, India

Gauri Misra
Amity Institute of Biotechnology
Amity University
Noida, Uttar Pradesh, India

ISBN 978-981-13-2924-1 ISBN 978-981-13-2925-8 (eBook)
https://doi.org/10.1007/978-981-13-2925-8

Library of Congress Control Number: 2018965925

© Springer Nature Singapore Pte Ltd. 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, 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.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Dedicated to
The lotus feet of Shri Radha Krishna
and
Goverdhannathji

Foreword

This book examines omics technologies and introduces the subject for new readers. It is particularly suitable for students (graduate, postgraduate, and doctoral), analytical scientists, and lab technicians. However, it has enough information to allow established scientists in related fields to understand the power and limitations of omics technology.

Omics is a relatively new area of study and cuts across all the biological disciplines and is relevant to all biological sciences. It attempts to look at biological systems in a holistic way and to account for all the interactions between, genes, proteins, RNA, and metabolites. It is divided into various disciplines which are discussed in the book. The chapters span from genomics that studies the structure and function of the whole genomes of organisms, proteomics describing the expressed proteins in a cell or tissue, transcriptomics dealing with the RNA present in a cell or tissue leading to an understanding about differential gene expression under particular conditions to metabolomics which provides a glimpse of the end products of metabolism contributing most to phenotypes. The book also contributes toward an in-depth understanding of the microbiome that encompasses the total genes of all the microorganisms in a particular ecological niche. Humans have their own microbiome (as do almost all environments on Earth), and there are more microbial cells in a human than there are human cells. The microbiome has gathered a lot of scientific attention recently for exploitation of its therapeutic relevance in humans, other animals, and plants; thus, it is sometimes considered as a separate discipline.

The book is edited by two exciting young scientists, one of whom I had the privilege of supervising for part of her PhD. They have invited the experts in the thematic areas to design a contributed volume delving into different omics branches.

Peter Mullany

Eastman Dental Institute
University College London
London, UK

Preface

Our exposure to the omics world during doctoral program in India and abroad laid the foundation for this book. The advent of highly parallel assays led to the transition of biological research from discrete knowledge of gene/transcript/protein/metabolite to a complete interlinked biological picture. This field has gained momentum with its immense usefulness across various dimensions including disease prognosis, therapeutics, personalized medicine, and drug discovery. Therefore, it is imperative to design a resource that will not only be useful for beginners but also for the experts seeking the advancement of their knowledge in this field.

The present book is divided into seven chapters with an aim to address the fundamental questions of diverse audience pertaining to the interdisciplinary field of omics. The introductory chapter describes the scope of omics, experimental design in omics research, its applications, and the usage of R language for analyzing high-throughput omics data. The second chapter outlines a coherent view of the human genome architecture, DNA sequencing approaches, and new technological advances in genomics. The third chapter discusses the principle of transcriptomics, technologies (expression sequence tag, serial/cap analysis of gene expression, microarray, RNA-seq) used to study transcriptomes, and applications of transcriptomics in disease profiling, ecology, evolution, and gene function annotation. The fourth chapter details different types of proteomics and advanced proteomic techniques such as two-dimensional electrophoresis, isotype-coded affinity tag peptide labeling, mass spectrometry, and multidimensional protein identification technique. The fifth chapter describes metabolome, its applications, and integrated platforms for analysis and interpretation of the metabolomics data. The sixth chapter provides an insight to soil, plant, marine, and human microbiome. A special emphasis is laid on human coinhabitants, wherein microbiome of various niches such as the gut, skin, oral, and urine is discussed in detail. The last chapter on bioinformatics resources gives an in-depth description about various bioinformatics approaches available to analyze genomics, transcriptomics, proteomics, and metabolomics data.

We believe our effort will be a priceless treasure for the general audience. The text has been enriched with the help of appropriate annotations, tables, and further readings. A positive feedback and scientific appreciation will be the true reward that the editors genuinely seek.

Noida, Uttar Pradesh, India

Preeti Arivaradarajan
Gauri Misra

Contents

Editors and Contributors

About the Editors

Preeti Arivaradarajan is Assistant Professor at Amity University, Noida, Uttar Pradesh, India. Previously, she completed Ph.D. from the School of Biological Sciences, Madurai Kamaraj University, Tamil Nadu, India. Her research inclination is towards the study of microbiome associated with pristine environments, with a special interest for human oral microbiome. She has been conferred with various prestigious awards, notably, Commonwealth Split-Site Doctoral Scholarship by Commonwealth Scholarship Commission in United Kingdom, Innovation in Scientific Pursuit for Inspired Research (INSPIRE) fellowship by Department of Science and Technology, Government of India, and gold medal for the best outgoing student of the year during master's program by Sri Ramachandra University, Chennai, India. Furthermore, she is a member of the Indian Science Congress Association, and is serving as a reviewer for numerous international journals. She has also published a number of research articles in peer-reviewed international journals and has authored book chapters.

Gauri Misra is currently working as Assistant Professor at Amity University, Noida, Uttar Pradesh (UP), India. Before joining Amity University, she worked as Assistant Professor at Hygia Institute of Pharmaceutical Education and Research, Lucknow (UP), India. From 2010 to 2011, she worked as a postdoctoral fellow at CHUL Research Centre, Quebec, Canada, where she worked towards understanding the role of androgen receptor in the growth and proliferation of breast cancer cells using various structural biology approaches. During her doctoral studies at the Central Drug Research Institute, Lucknow, she made an innovative contribution towards understanding the structural and functional characterization of the *Plasmodium falciparum* proteins that are involved in the transit peptide-mediated pathway. Her research interests are in the field of pathogen biology and cancer. She is striving for a healthy future for mankind through her research efforts in the field of structure based drug design paving way for drug discovery. She is the recipient of many awards, including the Young Scientist best scientific presentation award at the International Conference on “Trends in Biochemical and Biomedical Research,” Banaras Hindu University, Banaras, India, in 2018, the prestigious Eli-Lilly Asia Outstanding Thesis Award (first prize) in 2009, and best oral presentation award at the “National Seminar on Crystallography-37,” held in Jadavpur University, Kolkata, in 2008. She has been an outstanding performer, receiving gold medals and honors at various stages of her academic journey. Previously, she was selected as visiting scientist under the INSA bilateral exchange program to visit the Israel Structural Proteomics Center situated at Weizmann Institute of Science, Rehovot, Israel, in 2014.

She is serving as the reviewer for various renowned international journals. Furthermore, she is a member of many scientific societies, including the Indian Biophysical Society, the Indian Science Congress Association, and the Indian Crystallographic Association. To date, she has authored and co-authored 13 articles in various peer-reviewed journals. She has successfully edited a book on biophysics with Springer in 2017. During the past 8 years, she has been actively involved in both research and teaching graduate and postgraduate students.

Contributors

Sudeep Bose Amity University, Noida, Uttar Pradesh, India

Sangeeta Choudhury Research Department, Sir Ganga Ram Hospital, New Delhi, India

Manali Datta Amity Institute of Biotechnology, Amity University Rajasthan, Jaipur, Rajasthan, India

Neetu Jabalia Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh, India

S. V. Kirthanashri Amity Institute of Molecular Medicine & Stem Cell Research, Amity University, Noida, Uttar Pradesh, India

Priyanka Narad Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh, India

Debarati Paul Amity University, Noida, Uttar Pradesh, India

Jyotika Rajawat Molecular and Human genetics laboratory, Department of Zoology, University of Lucknow, Lucknow, Uttar Pradesh, India

Abhishek Sengupta Amity University, Noida, Uttar Pradesh, India

Desh Deepak Singh Amity Institute of Biotechnology, Amity University Rajasthan, Jaipur, Rajasthan, India

Vivek Tanavde Bioinformatics Institute, Agency for Science Technology and Research (A*STAR), Singapore, Singapore

Division of Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Ahmedabad, India

Candida Vaz Bioinformatics Institute, Agency for Science Technology and Research (A*STAR), Singapore, Singapore

Chapter 1
Introduction to Omics

Priyanka Narad and S. V. Kirthanashri

Abstract Omics technologies also referred as high-dimensional biology encompasses the cells, tissues, and organisms in a manner that integrates the data from various platforms and helps in its interpretation. It primarily detects the genes (genomics), mRNAs (transcriptomics), proteins (proteomics), and metabolites (metabolomics) in a nontargeted and non-biased manner. The integration and interrelationships between networks of biological processes is termed as systems biology. The approach provides hope for unravelling the intricate details in various aspects of biology and accelerates innovation in healthcare. Understanding the various dimensions encompassing not only the three levels constituting the central dogma of life but also the intermediate metabolites is significant for the scientists to cover new horizons in drug discovery and disease regulation. This chapter outlines the scope of omics, experimental design in omics research, and its applications. It will also provide an overview to the usage of languages like R for analyzing high-throughput data from all branches of "omics" technologies.

The primary focus is to understand omics approaches that enable the validation of large-scale data that is generated from various experimental platforms. Systems biology and omics data are way apart from hypothesis-driven traditional studies. The systems biology experiments generate hypothesis by employing all data that needs to be further analyzed.

Omics technology applied majorly for accurate understanding of normal physiological processes and gaining knowledge related to disease processes which involves screening, diagnosis, and prognosis that provides an understanding of the etiology of diseases.

Keywords Omics · Systems biology · R language

P. Narad
Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh, India
e-mail: pnarad@amity.edu

S. V. Kirthanashri (✉)
Amity Institute of Molecular Medicine & Stem Cell Research, Amity University,
Noida, Uttar Pradesh, India
e-mail: svkirthanashri@amity.edu

P. Arivaradarajan, G. Misra (eds.), *Omics Approaches, Technologies And Applications*, https://doi.org/10.1007/978-981-13-2925-8_1

1.1 Background

In biology the suffix -omics refers to huge biological molecules; the broad analysis of large biological molecules was needed to be studied in detail as the conclusion of human genome project (HGP) in 2001. The HGP revealed that the human genome contained lesser number of genes and biological process were regulated not particularly on DNA sequence but involved various other processes, and with this evolved the new branch of study termed the omics (Hood and Rowen 2013).

This technology deciphered the cell, tissue, and organism in a holistic way around central dogma for the detection of genes (genomics), mRNA (transcriptomics), proteins (proteomics), and metabolites (metabolomics) in the samples (specific biological component). Since the technology is non-biased, they are also referred to as high-dimensional biology, while the integration of these is the systems biology. Following the discovery of DNA structure by Watson-Crick in 1953, a series of inventions and discoveries followed. The development of PCR by Kary Mullis opened all possible channels in molecular biology research. The progress in *Omics* started from the development of genomics further followed by transcriptomics and finally the proteomics, and the term was coined in 1994 by Marc Wilkins. This was possible because of advanced development in techniques like high-resolution two-dimensional electrophoresis. The cascade of events in *Omics* is depicted in Fig. 1.1. The advantage of the omics study is that they reveal specific results that promote understanding. As the omics technology is of immense potential, they have been explored in various branches of medical and health science. This technology can help to understand the etiology of disease condition through the process of screening, diagnosis, and prognosis and also for the biomarker discovery to be made easy as they involve simultaneous investigation of multiple molecules (Poisot et al. 2013). Further *Omics* is of great use in drug discovery and toxicity assessment. Pharmacogenomics deals with the connection of genomics and pharmacology to examine the role of inheritance in individual variation in drug response utilized to individualize and optimize drug therapy. They help in the field of oncology to evaluate rigorous systemic toxicity and unpredictable efficacies that are

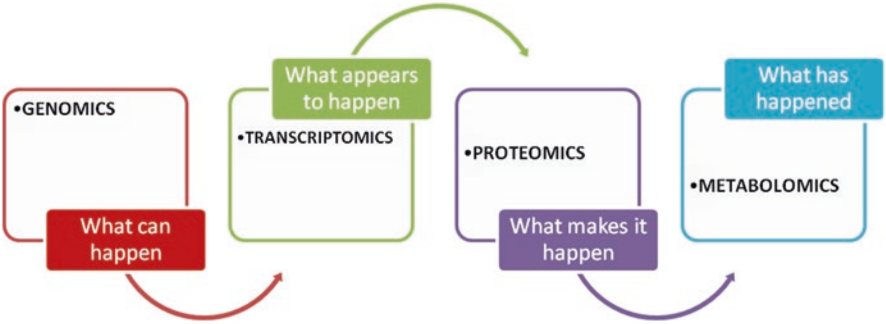

Fig. 1.1 Cascade of *Omics* development

hallmarks of cancer therapies. These technologies are helpful in selecting novel targets for the treatment including conditions like cancer, cardiovascular disease, and obesity. In the future, systems biology promises to develop new approaches that will be predictive, preventive, and personalized (Sagner et al. 2017). Research in the field of obstetrics and gynecology is currently taking advantage of these possibilities which can be used to solve the problems related to fertility. This review aims to provide a complete overview of various omics technologies available.

1.2 Overview of Omics

The omics technology can be classified into various types depending on their function. Figure 1.2 highlights a few of various omics technologies that are presented in detail in the following chapters.

1.2.1 Genomics

This refers to the interdisciplinary study based on evaluating the structure and function and mapping of the genomes. In short this is the study of a set of genes, the inheritance substance. The term genomics was coined by Tom Roderick in 1986 on mapping the human gene. The possible and highly researched areas under genomics include the functional genomics, metagenomics, and epigenomics (Feinberg 2010).

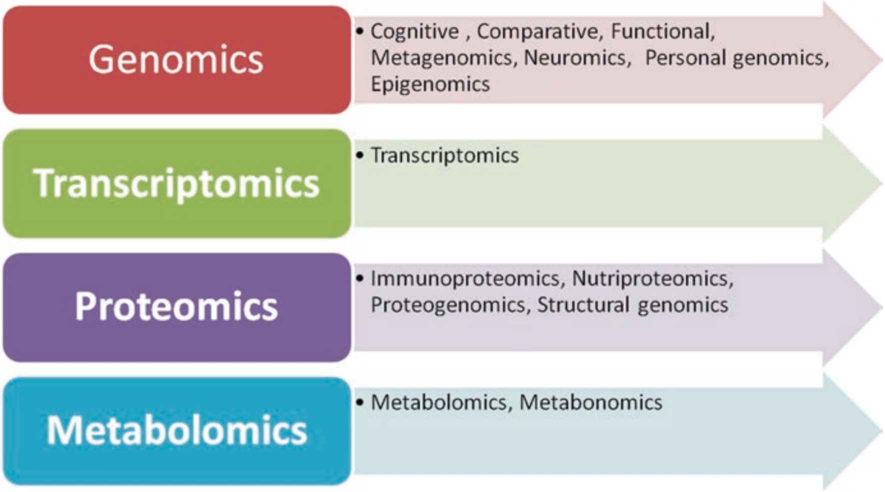

Fig. 1.2 Various omics technologies

1.2.2 Transcriptomics

The presence of mRNA in the sample reflects the abundance level of the corresponding gene. Gene expression involves the detection and classification of mRNA mixture in a specific sample. The goal of gene expression profiling is to differentiate the mRNA mixtures from different samples. Contrary to genotyping, gene expression categorizes the level of gene expression. The variation of the transcriptome can be seen over time between cell types and change according to environmental conditions (Hubank 2004).

1.2.3 Proteomics

The proteome refers to the total proteins expressed by a cell, tissue, or organism. The proteome is highly variable over time, shows species variation, and alters depending on environmental conditions. Proteomics is defined as the study that provides insights to protein functions in biological systems. Due to the variability and abundance of proteins in biological samples, there is a requirement to develop technologies to detect a wide range of proteins in samples of different origins. Currently exploited proteomic technologies are mass spectrometry (MS) and protein microarrays using capturing agents such as antibodies. However, the high dynamic range (abundance and concentration) of proteins complicates this type of proteomic analysis (Chandramouli and Qian 2009).

1.2.4 Metabolomics

The small molecules (e.g., lipids or vitamins) referred to as metabolites constitute the metabolome. The interaction between genetic, environmental, lifestyle, and other factors results in metabolic phenotypes. Interaction of metabolome with other biological macromolecules in the cell results in metabolic pathways. The metabolic profiles of biological sample represent the metabolomics which are changeable and time dependent and had a wide range of chemical structures (Bino et al. 2004).

1.3 Overview of Systems Biology

1.3.1 Systems Biology

The genome is the total DNA of a cell in the organism. The human genome contains about 3.2 billion bases with 30,000–40,000 protein-coding genes. The microarray technique enables quick analysis of the genes and also helps in examining the