METHODS IN MOLECULAR BIOLOGY

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

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

Disease Gene Identification

Methods and Protocols

Second Edition

Edited by

Johanna K. DiStefano

Translational Genomics Research Institute, Phoenix, AZ, USA

Humana Press

Editor
Johanna K. DiStefano
Translational Genomics Research Institute
Phoenix, AZ, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-7470-2 ISBN 978-1-4939-7471-9 (eBook)
https://doi.org/10.1007/978-1-4939-7471-9

Library of Congress Control Number: 2017964097

Printed on acid-free paper

This Humana Press imprint is published by Springer Nature
The registered company is Springer Science+Business Media, LLC
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

Completion of the Human Genome Project (HGP) not only yielded a greater appreciation of the role of DNA in shaping species development and evolution, biology, and disease susceptibility, but also helped spawn technological advances that have revolutionized the field of human genetics. Perhaps the most significant impact of this endeavor has been on the manner in which researchers investigate the causes of complex human diseases. Efforts to characterize the genetic variation in the human genome have led directly to the development and application of a diverse range of technological and bioinformatics approaches to identify the roles of both rare and common alleles in complex disease. Such strategies range from genome-wide association studies to whole genome sequencing, and everything in between.

Over three thousand genetic mutations have been identified that contribute to the pathogenesis of highly penetrant human diseases. Efforts to uncover the genomic basis of rare conditions have been successful due to the less complicated genetics of monogenic diseases compared to complex disorders. In rare conditions, a single mutation, inherited in a simple manner between generations in affected families, is typically sufficient to cause disease. In contrast, complex diseases, such as diabetes, heart disease, neurological disorders, and most kinds of cancer, result from a complicated interaction of multiple genetic and environmental determinants, none of which are amenable to identification and characterization using the traditional approaches to monogenic disease gene discovery. Recent efforts to characterize genetic variation in the human genome, coupled with the rapidly developing field of genomics, have led directly to the development of new and innovative approaches to the identification of genes contributing to complex human diseases. This volume was prepared to present molecular methodologies used in the process of identifying a disease gene, from the initial stages of study design to locus identification and target characterization. The need for such a book derives from the intellectual revolution in biomedical science and the realization that the molecular determinants of human diseases are rapidly becoming identifiable through well-planned, technologically advanced approaches.

While descriptions of the technical procedures described here are available in the literature, there is generally a dearth of practical detail in these publications, particularly in terms of modifications developed from personal experience and discussions of optimal study design or potential problems that may be encountered throughout the protocol, as well as ways to avoid them. The structure of this volume is unique in that it aims to address these deficiencies.

This text is written at a level accessible to graduate students, postdoctoral researchers, and bench scientists in the fields of molecular genes and molecular biology. The primary aim of this volume is to present detailed laboratory procedures in an easy to follow format that can be carried out with success by competent investigators lacking previous exposure to a specific research method. The book's main focus is on the application of molecular approaches to disease gene identification, but overviews and case studies are also presented.

The volume begins with six introductory chapters, which provide overviews of strategies for disease gene identification and functional characterization, and include introductions to microbiome sequencing methods for studying human disease, as well as the emerging role of long noncoding RNAs in human disease. The next section of the text contains chapters presenting methods for identifying potential susceptibility loci, including practical

procedures for genome-wide association analysis, whole genome, whole exome, and single-cell library construction, and methylation profiling.

The volume follows with a section on current applications in human genomics, which provide tools for target validation and functional assessment. These protocols are useful for characterizing disease-associated loci and include methods in quantitative polymerase chain reaction, lentiviral-mediated CRISPR-cas9, RNA interference, and luciferase reporter assay.

We end with four discursive chapters providing examples of disease gene identification and application. The first chapter in this section is related to physiologic interpretation of genome-wide association signals, using type 2 diabetes as a model. The following two chapters present overviews of disease gene discovery in two distinct disorders: hereditary hemochromatosis and small cell carcinoma of the ovary. A discussion of the reemergence of linkage analysis, as an adjunct to association studies, concludes this section.

Completion of this volume would not have been possible without the support and contributions of many individuals. In particular, I thank Dr. John M. Walker, the series editor, who provided expert guidance and oversight at each step of bringing this volume to fruition. I also appreciate the efforts of the authors, all of whom contributed outstanding chapters. It was a pleasure working with this expert team of scientists. It is my hope that this volume leads to the identification and characterization of many more disease-related genes, which may someday pave the way toward more accurate and improved methods for disease diagnosis, as well as novel and effective strategies for disease treatment and prevention.

Phoenix, AZ, USA *Johanna K. DiStefano*

Contents

Contributors

Jonathan Adkins • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*

Scott Alper • *Department of Biomedical Research, Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA; Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, CO, USA*

Javier Apfeld • *Department of Biology, Northeastern University, Boston, MA, USA*

Leslie J. Baier • *Diabetes Molecular Genetics Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, AZ, USA*

Taylor M. Benson • *Department of Biomedical Research, Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA*

John Blangero • *South Texas Diabetes and Obesity Institute, School of Medicine, University of Texas Rio Grande Valley, Edinburg, TX, USA*

Austin Christofferson • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*

Joanne E. Curran • *South Texas Diabetes and Obesity Institute, School of Medicine, University of Texas Rio Grande Valley, Edinburg, TX, USA*

Lori Cuyugan • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*

Rebecca M. Davidson • *Department of Biomedical Research, Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA*

Johanna K. DiStefano • *Translational Genomics Research Institute, Phoenix, AZ, USA*

Robert Durruthy-Durruthy • *Fluidigm Corporation, South San Francisco, CA, USA*

L. Elaine Epperson • *Department of Biomedical Research, Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA*

Jamie L. Everman • *Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA*

Glenn S. Gerhard • *Department of Medical Genetics and Molecular Biochemistry, 960 Medical Education and Research Building (MERB), Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

David A. Greenberg • *Battelle Center for Mathematical Medicine, Nationwide Children's Hospital, Columbus, OH, USA; Department of Pediatrics, Wexner Medical Center, Ohio State University, Columbus, OH, USA*

Haiyong Han • *Translational Genomics Research Institute, Phoenix, AZ, USA*

Adrienne Helland • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*

William P.D. Hendricks • *Translational Genomics Research Institute, Phoenix, AZ, USA*

Jonathan J. Keats • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*

Mahdieh Khosroheidari • *Illumina, Inc., San Diego, CA, USA*

Christopher B. Kingsley • *Roche Sequencing Solutions, Inc., San Jose, CA, USA*

Satish Kumar • *South Texas Diabetes and Obesity Institute, School of Medicine, University of Texas Rio Grande Valley, Edinburg, TX, USA*

Jessica D. Lang • *Translational Genomics Research Institute, Phoenix, AZ, USA*

Fatjon Leti • *Department of Biomedical Research, Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA*

Winnie S. Liang • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*
Ettie M. Lipner • *Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA; Department of Pharmacology, School of Medicine, University of Colorado Denver, Aurora, CO, USA*
Lorida Llaci • *Neurogenomics Division, Translational Genomics Research Institute, Phoenix, AZ, USA*
Ivana Malenica • *Division of Biostatistics, University of California, Berkeley, Berkeley, CA, USA*
Anup K. Nair • *Diabetes Molecular Genetics Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, AZ, USA*
Brian P. O'Connor • *Genomics Facility, Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA*
Barbara V. Paynton • *Department of Medical Genetics and Molecular Biology, The Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*
Rupesh Kanchi Ravi • *Pfizer, Inc., San Diego, CA, USA*
Manisha Ray • *Fluidigm Corporation, South San Francisco, CA, USA*
Cydney Rios • *Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA*
Dorit Schleinitz • *IFB Adiposity Diseases, Leipzig University Medical Center, University of Leipzig - Medical Faculty, Leipzig, Germany; Clinic and Policlinic for Endocrinology and Nephrology, Leipzig University Medical Center, Leipzig, Germany*
Max A. Seibold • *Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA; Department of Pediatrics, National Jewish Health, Denver, CO, USA*
Kristi Stephenson • *Translational Genomics Research Institute (TGen), Phoenix, AZ, USA*
Kendra Walton • *Genomics Facility, Center for Genes, Environment, and Health, National Jewish Health, Denver, CO, USA*
Richard M. Watanabe • *Departments of Preventive Medicine and Physiology & Biophysics, Keck School of Medicine of USC, Los Angeles, CA, USA*
Tobias Wohland • *IFB Adiposity Diseases, Leipzig University Medical Center, University of Leipzig - Medical Faculty, Leipzig, Germany*