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

Next Generation Sequencing

Methods and Protocols

Editors

Steven R. Head

Next Generation Sequencing and Microarray Core Facility,
The Scripps Research Institute,
La Jolla, CA, USA

Phillip Ordoukhanian

Next Generation Sequencing and Microarray Core Facility,
The Scripps Research Institute,
La Jolla, CA, USA

Daniel R. Salomon

Department of Molecular Medicine, The Scripps Research Institute,
Center for Organ and Cell Transplantation,
La Jolla, CA, USA

Humana Press

Editors
Steven R. Head
Next Generation Sequencing and Microarray
Core Facility
The Scripps Research Institute
La Jolla, CA, USA

Phillip Ordoukhanian
Next Generation Sequencing and Microarray
Core Facility
The Scripps Research Institute
La Jolla, CA, USA

Daniel R. Salomon
Department of Molecular Medicine
The Scripps Research Institute
Center for Organ and Cell Transplantation
La Jolla, CA, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-7512-9 ISBN 978-1-4939-7514-3 (eBook)
https://doi.org/10.1007/978-1-4939-7514-3

Library of Congress Control Number: 2017960416

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.

Dedication

This volume is dedicated to the memory of our friend, colleague, mentor, and surf buddy, Daniel R. Salomon, M.D.

Preface

The revolution in high-throughput sequencing technology—Next Generation Sequencing (NGS)—has transformed the science of genomics in the last decade. The resulting deluge of new methods, protocols, and techniques to answer fundamental questions in biology has enabled highly efficient strategies for addressing problems of DNA, RNA, and their interactions with proteins. Each chapter in this *Methods in Molecular Biology* book describes a current state of the art in NGS application and is intended as a resource for researchers with all levels of NGS experience who wish to expand their knowledge and practical skills in high-throughput sequencing. Examples of the challenges of the role of NGS in basic research protocols data analysis, as well as clinical applications are included.

This book covers a wide range of various fields of research, with the common thread being NGS-related methods and applications, as well as some analysis and interpretation of the data obtained. The first two chapters focus on the highly dynamic processes of translational and transcriptional profiling of a cell. In the first chapter, polysome and ribosome isolation by sucrose gradient is used to investigate translational activity, both approaches are described, and the information obtained from each is discussed. In the second chapter, isolation of cell nuclei and the nascent RNA transcripts therein allow a look at the nascent transcriptome of a cell. The third chapter focuses on a method to detect copy number alterations (CNAs) using whole genome amplification and low pass whole genome sequencing. Chapters 4 and 5 touch on more targeted sequencing applications. The first uses small "bait" oligonucleotides to "fish" out long pieces of DNA (~2 Kb) from the genome combined with long read sequencing using the MinION (Oxford Nanopore, Inc.) allowing for the interrogation of these unknown flanking regions not contained in the baits. Chapter 5 deals with an extremely efficient and cost-effective method called "Hi-Plex" to characterize known polymorphic loci using a highly multiplexed amplicon-based approach to look for genetic variants.

Chapter 15 focuses on DNA structural rearrangements using a method called, "Hi-C," which maps chromatin interactions in nuclei using NGS. The NGS libraries were generated for resting and activated human CD4 T cells to study activation-induced chromatin structural rearrangement. Another chapter (Chapter 13) dealing with chromosomal changes uses CRISPR-based knockout libraries in genome-wide screens to systematically investigate gene function in biological systems. Specifically, the Genome-Scale CRISPR Knock Out (GeCKO v2) library is used and methods for NGS library generation and sequencing are discussed.

There are several chapters (Chapters 7, 8, 10, and 14) dealing with a single cell of DNA or RNA in NGS. Chapter 7 describes a post-bisulfite treatment adapter tagging strategy to generate single-cell data looking at genome-wide cytidine methylation states. Chapter 10 describes the use of modified adapters in a small RNA protocol termed "CleanTag" (TriLink Biotechnologies Inc.) to prepare NGS libraries. The use of these adapters prevents adapter dimer formation, the principal artifactual product in a standard small RNA library prep. The authors demonstrate that the technique can work with very low inputs down to the single-cell level (~10 pg). Chapter 14 uses a single-cell analysis approach to identify and

characterize rare circulating CD4 T cells using a Biomark HD and profiles 96 different genes by quantitative PCR (qPCR) and generates NGS libraries using the Biomark HD Access Array. Chapter 8 describes a protocol for isolation, extraction, and sequencing of single bacterial and archaeal cells using FACS, MDA, a 16S rRNA screen, and a computational approach for quality assurance.

Two other chapters (Chapters 9 and 12) have more of a computational focus. Chapter 9 uses low pass whole genome sequencing and reference-guided assembly of non-model organisms for SNP discovery allowing for the genotyping of populations. Chapter 12 describes a computational pipeline for RNAseq analysis from tissue samples containing a complex heterogeneous population of cell types (e.g., a blood sample). The pipeline is able to estimate cell type composition and other statistical analysis information generated from bulk RNAseq profiles.

Another important component to almost all NGS-related work, especially those involving microbiome studies, is contamination. To this end, Chapter 11 focuses on the best practices and approaches for sample handling, DNA and/or RNA extraction, and library preparation from microbial and viral samples to generate NGS libraries. Chapter 6 discusses a novel method for generating sequencing libraries from viral RNA termed, "Clickseq," which uses "Click Chemistry" to attach one of the NGS adapters preventing artifactual generation of chimeras, and consequently, greatly increasing the ability to detect rare recombination events in viral RNA. The last chapter in the book (Chapter 16) describes another RNAseq library prep method for profiling reverse transcription termination sites. It is an efficient protocol and generates good NGS library yields from low RNA inputs generated from protocols such as nascent RNA sequencing, RNA Immunoprecipitation (RIPseq), and 5′-RACE structural probing.

Next Generation Sequencing technology has brought together disparate fields of research from bacterial and viral studies to the study of plants, non-model organisms, and of course human disease. This requires collaboration between the users of NGS data and those needed to design and perform the enzymatic procedures for the preparation of sequencing libraries and ensure the desired target is being captured, targeted, or amplified. Other critical players are the engineers that develop microfluidics and sequencing hardware systems or the bioinformatics experts necessary to ensure data generated is being analyzed correctly and translated into a meaningful analysis. We hope you enjoy this book and find it informative and a useful reference.

La Jolla, CA, USA

Steven R. Head
Phillip Ordoukhanian
Daniel R. Salomon

Contents

Contributors

Brian Abe • *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA*

Robert M. Bowers • *Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA*

Fei Xavier Chen • *Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA*

Anne Costanzo • *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA*

Dieter De Coninck • *Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium*

Dieter Deforce • *Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium*

Lieselot Deleye • *Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium*

Yu Deng • *Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA*

Andrew D. Foote • *Molecular Ecology and Fisheries Genetics Laboratory, School of Biological Sciences, Bangor University, Bangor, Gwynedd, UK*

Louis Gioia • *Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA*

Fleur Hammet • *Genomic Technologies Group, Genetic Epidemiology Laboratory, Department of Medicine, The University of Melbourne, Parkville, VIC, Australia*

Steven R. Head • *Next Generation Sequencing and Microarray Core Facility, The Scripps Research Institute, La Jolla, CA, USA*

Jordana M. Henderson • *Research and Development, Cell and Molecular Biology, TriLink BioTechnologies, LLC, San Diego, CA, USA*

Christopher M. Hill • *Department of Computer Science, University of Maryland, College Park, MD, USA*

Marie Holt • *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA*

Elizabeth Jaworski • *Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA*

Hyun Yong Jin • *Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA, USA*

Timokratis Karamitros • *Department of Zoology, University of Oxford, Oxford, UK; Department of Medical Microbiology, Hellenic Pasteur Institute, Athens, Greece*

Scott T. Kelley • *Department of Biology, San Diego State University, San Diego, CA, USA*

Aimee R. Lang • *Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, CA, USA*

Heather J. Lee • *Epigenetics Programme, Babraham Institute, Babraham Research Campus, Cambridge, UK; School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, The University of Newcastle, Callaghan, NSW, Australia*
Janey Lee • *Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA*
Marie Louis • *Scottish Oceans Institute, University of St Andrews, St Andrews, UK*
Gkikas Magiorkinis • *Department of Zoology, University of Oxford, Oxford, UK*
Stacy A. Marshall • *Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA*
Anton P. McCaffrey • *Research and Development, Cell and Molecular Biology, TriLink BioTechnologies, LLC, San Diego, CA, USA*
Xiangzhi Meng • *Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA, USA*
Phillip A. Morin • *Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, CA, USA*
Tu Nguyen-Dumont • *Genomic Technologies Group, Genetic Epidemiology Laboratory, Department of Medicine, The University of Melbourne, Parkville, VIC, Australia*
Jessica Nichols • *Next Generation Sequencing and Microarray Core Facility, The Scripps Research Institute, La Jolla, CA, USA*
Phillip Ordoukhanian • *Next Generation Sequencing and Microarray Core Facility, The Scripps Research Institute, La Jolla, CA, USA*
Daniel Jonathan Park • *Melbourne Bioinformatics, The University of Melbourne, Parkville, VIC, Australia; Genomic Technologies Group, Genetic Epidemiology Laboratory, Department of Medicine, The University of Melbourne, Parkville, VIC, Australia*
Bernard J. Pope • *Melbourne Bioinformatics, The University of Melbourne, Parkville, VIC, Australia*
Tariq M. Rana • *Solid Tumor Therapeutics Program, Moores Cancer Center, University of California, San Diego, La Jolla, CA, USA; Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA; Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA*
Nicole Riley • *Department of Inflammation Biology, La Jolla Institute of Allergy and Immunology, La Jolla, CA, USA*
Andrew Routh • *Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA; Sealy Centre for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX, USA*
Casey P. Shannon • *Prevention of Organ Failure (PROOF) Centre of Excellence, Centre for Heart Lung Innovation, University of British Columbia, St. Paul's Hospital, Vancouver, BC, Canada*
Siddhartha Sharma • *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA*
Sabrina Shore • *Research and Development, Cell and Molecular Biology, TriLink BioTechnologies, LLC, San Diego, CA, USA*
Benoit Simon-Bouhet • *Centre d'Etudes Biologiques de Chizé, UMR 7372 CNRS-Université de La Rochelle, Villiers-en-Bois, France*
Sébastien A. Smallwood • *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland; Epigenetics Programme, Babraham Institute, Babraham Research Campus, Cambridge, UK*

Andrew I. Su • *Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA*
Scott J. Tebbutt • *Prevention of Organ Failure (PROOF) Centre of Excellence, Centre for Heart Lung Innovation, University of British Columbia, St. Paul's Hospital, Vancouver, BC, Canada; Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada*
Luc Teyton • *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA*
Ryan Thompson • *Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA*
Sun Tianjiao • *Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA*
Pedro J. Torres • *Department of Biology, San Diego State University, San Diego, CA, USA*
Filip Van Nieuwerburgh • *Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium*
Tanja Woyke • *Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA*
Changchun Xiao • *Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA, USA*
Chen Xi Yang • *Prevention of Organ Failure (PROOF) Centre of Excellence, Centre for Heart Lung Innovation, University of British Columbia, St. Paul's Hospital, Vancouver, BC, Canada*
Edwin H. Yau • *Division of Hematology-Oncology, Department of Internal Medicine, University of California, San Diego, La Jolla, CA, USA; Solid Tumor Therapeutics Program, Moores Cancer Center, University of California, San Diego, La Jolla, CA, USA*