METHODS IN MOLECULAR BIOLOGY™ 439

Genomics Protocols

Second Edition

Edited by

Mike Starkey
Ramnath Elaswarapu

Humana Press

Genomics Protocols

Second Edition

Edited by

Mike Starkey
Animal Health Trust,
Lanwades Park, Newmarket,
Suffolk, United Kingdom

Ramnath Elaswarapu
Oxford Gene Technology Ltd.,
Oxford, United Kingdom

Humana Press

Editors
Mike Starkey
Animal Health Trust
Lanwades Park, Newmarket
Suffolk, United Kingdom

Ramnath Elaswarapu
Oxford Gene Technology Ltd.
Oxford, United Kingdom

Series Editor
John M. Walker
University of Hertfordshire
Hatfield, Herts
United Kingdom

ISBN: 978-1-58829-871-3 e-ISBN: 978-1-59745-188-8

Library of Congress Control Number: 2007938729

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Genomics Protocols

METHODS IN MOLECULAR BIOLOGY™

John Walker, SERIES EDITOR

439. **Genomics Protocols:** *Second Edition*, edited by *Mike Starkey and Ramnath Elaswarapu, 2008*
438. **Neural Stem Cells: Methods and Protocols**, *Second Edition*, edited by *Leslie P. Weiner, 2008*
437. **Drug Delivery Systems**, edited by *Kewal K. Jain, 2008*
436. **Avian Influenza Virus**, edited by Erica Spackman, 2008
435. **Chromosomal Mutagenesis**, edited by *Greg Davis and Kevin J. Kayser, 2008*
434. **Gene Therapy Protocols: Volume 2:** Design and Characterization of Gene Transfer Vectors, edited by Joseph M. LeDoux, 2008
433. **Gene Therapy Protocols: Volume 1:** Production and In Vivo Applications of Gene Transfer Vectors, edited by *Joseph M. LeDoux, 2008*
432. **Organelle Proteomics**, edited by *Delphine Pflieger and Jean Rossier, 2008*
431. **Bacterial Pathogenesis: Methods and Protocols**, edited by *Frank DeLeo and Michael Otto, 2008*
430. **Hematopoietic Stem Cell Protocols**, edited by *Kevin D. Bunting, 2008*
429. **Molecular Beacons: Signalling Nucleic Acid Probes, Methods and Protocols**, edited by *Andreas Marx and Oliver Seitz, 2008*
428. **Clinical Proteomics: Methods and Protocols**, edited by *Antonio Vlahou, 2008*
427. **Plant Embryogenesis**, edited by *Maria Fernanda Suarez and Peter Bozhkov, 2008*
426. **Structural Proteomics: High-Throughput Methods**, edited by *Bostjan Kobe, Mitchell Guss, and Huber Thomas, 2008*
425. **2D PAGE:** Sample Preparation and Fractionation, Volume 2, edited by *Anton Posch, 2008*
424. **2D PAGE:** Sample Preparation and Fractionation, Volume 1, edited by *Anton Posch, 2008*
423. **Electroporation Protocols: Preclinical and Clinical Gene Medicine**, edited by *Shulin Li, 2008*
422. **Phylogenomics**, edited by *William J. Murphy, 2008*
421. **Affinity Chromatography: Methods and Protocols**, Second Edition, edited by *Michael Zachariou, 2008*
420. **Drosophila: Methods and Protocols**, edited by *Christian Dahmann, 2008*
419. **Post-Transcriptional Gene Regulation**, edited by *Jeffrey Wilusz, 2008*
418. **Avidin–Biotin Interactions: Methods and Applications**, edited by *Robert J. McMahon, 2008*
417. **Tissue Engineering**, Second Edition, edited by *Hannsjörg Hauser and Martin Fussenegger, 2007*
416. **Gene Essentiality: Protocols and Bioinformatics**, edited by *Svetlana Gerdes and Andrei L. Osterman, 2008*
415. **Innate Immunity**, edited by Jonathan Ewbank and Eric Vivier, 2007
414. **Apoptosis in Cancer: Methods and Protocols**, edited by *Gil Mor and Ayesha Alvero, 2008*
413. **Protein Structure Prediction**, Second Edition, edited by *Mohammed Zaki and Chris Bystroff, 2008*
412. **Neutrophil Methods and Protocols**, edited by *Mark T. Quinn, Frank R. DeLeo, and Gary M. Bokoch, 2007*
411. **Reporter Genes: A Practical Guide**, edited by *Don Anson, 2007*

Preface

The genomics research world has moved on since the first volume of *Genomics Protocols* in the Methods in Molecular Biology series was published in 2001. This progression is reflected in the themes of the chapters that adorn the current volume. It is true that there is the same apparently eclectic mixture of subjects as in the former volume. However, undeniably, the emphasis has switched from gene identification to functional genomics and the characterization of genes and gene products.

Although essentially a volume of "wet lab" techniques, in silico approaches also are represented through three chapters addressing the detection of genome sequence variation and the prediction of gene function, respectively.

We make no apologies for including texts that may appear in volumes devoted to other of the "omics." Rather, we feel that the majority of these relative newcomers are derivatives of genomics, and from the perspective of being able to see "the bigger picture," a great value is to be gained from considering these research disciplines collectively.

For some of the specialist techniques, it is difficult to avoid the use of a specific/particular commercial platform or piece of equipment. For these procedures, the authors have focused on the generic issues (e.g., data analysis), which are all too often given inadequate coverage in user manuals. While not wishing to advocate particular systems, by including these topics we acknowledge the importance, and possibly the uniqueness, of the technologies involved.

Through the choice of some of the topics, we tried to remember that not everyone is engaged in investigation of organisms for which an annotated sequenced genome or a high-density genetic map is available. Similarly, lest we forget that not all researchers are affiliated to well-equipped laboratories, Chapter 5 describes a relatively inexpensive and low-tech approach to genetic mapping using microsatellites.

To our minds, the jewel in the crown of the Methods in Molecular Biology series remains the "Notes" section of each chapter. In spite of the complexity of some of the subject matter, the authors have endeavored to ensure that as little technical detail as possible is left to either the reader's imagination or to the extensivity of their organization's library.

The first chapter addresses an issue that all too often is a prerequisite for comprehensive genome analysis: high fidelity whole genome amplification. Chapters 2–4 focus on a spectrum of procedures that are a prelude to genetic linkage analysis (Chapter 5).

The molecular profiling of genomic DNA in a variety of guises is encompassed in Chapters 6–9. Chapters 10–12 are devoted to transcriptional profiling, and focusing on small samples and the use of formalin-fixed archival tissue, deal with two of the thornier issues confronting researchers. The profiling of microRNAs (Chapter 12) is an exciting area of interest in the study of the regulation of gene expression in disease.

The line between proteomics and functional genomics can be a little fuzzy, but Chapters 12–17 describe alternative strategies for protein profiling, while Chapters 18–26 arguably constitute approaches for gene characterization and the identification of protein interaction networks. The volume concludes with a detailed tutorial (Chapter 27) about an exciting genomics technology useful in the elucidation of gene function and with huge potential therapeutic applications.

So there you have it. All that remains is for us to thank a group of redoubtable authors, the series editor, John Walker, and all those involved at Humana Press. We truly hope that you find the book informative and a valuable addition to your laboratory bookshelf.

Mike Starkey
Ramnath Elaswarapu

Contributors

Lutgarde Arckens
Laboratory of Neuroplasticity and Neuroproteomics, Katholieke Universiteit Leuven, Leuven, Belgium

Larissa Belov
Medsaic Pty Ltd, National Innovation Centre, Eveleigh, Australia

Marina Bibikova
Illumina Inc., San Diego, CA

Helen Butler
Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

Lina Cekaite
Department of Tumor Biology, Institute for Cancer Research, Rikshopitalet–Radiumhospitalet Medical Center, Montebello, Oslo, Norway

Lynda Chin
Center for Applied Cancer Science, Belfer Institute for Innovative Cancer Science, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

Byung-Kwan Cho
Department of Bioengineering, University of California–San Diego, La Jolla, CA

Richard I. Christopherson
School of Molecular and Microbial Biosciences, University of Sydney, Sydney, Australia

John E. Collins
Wellcome Trust Sanger Institute, Cambridge, United Kingdom

Louise A. Dawson
Institute of Cellular and Molecular Biology, University of Leeds, Leeds, United Kingdom

Wen Ding
Institute for Biological Sciences, National Research Council, Ottawa, Ontario, Canada

Nicole Draper
Oxford Gene Technology Ltd, Yarnton, Oxford, United Kingdom

Per O. Ekstrøm
The Norwegian Radium Hospital, Oslo, Norway

Peter Ellmark
Department of Immunotechnology, Lund University, Lund, Sweden

S. Fabio Falsone
Institute of Pharmaceutical Sciences, University of Graz, Graz, Austria

Jian-Bing Fan
Illumina Inc., San Diego, CA

Bin Feng
Center for Applied Cancer Science, Belfer Institute for Innovative Cancer Science, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

Toni Gabaldón
Bioinformatics Department, CIPF, Valencia, Spain

Bernd Gesslbauer
Institute of Pharmaceutical Sciences, University of Graz, Graz, Austria

Opher Gileadi
Structural Genomics Consortium, Botnar Research Centre, University of Oxford, Oxford, United Kingdom

Stephen D. Ginsberg
Center for Dementia Research, Nathan Kline Institute, and Departments of Psychiatry and Physiology and Neuroscience, New York University School of Medicine, Orangeburg, NY

Arsalan S. Haqqani
Institute for Biological Sciences, National Research Council, Ottawa, Ontario, Canada

Runtao He
National Microbiology Laboratory, Public Health Agency of Canada Manitoba, Manitoba, Canada

Lan Huang
Department of Physiology & Biophysics and Department of Developmental and Cell Biology, University of California Irvine, Irvine, CA

Trupti Joshi
Digital Biology Laboratory, Computer Science Department and Christopher S. Bond Life Sciences Center, University of Missouri–Columbia, Columbia, MO

Peter Kaiser
Department of Biological Chemistry, College of Medicine, University of California Irvine, Irvine, CA

John F. Kelly
Institute for Biological Sciences, National Research Council, Ottawa, Ontario, Canada

Eric M. Knight
Department of Bioengineering, University of California–San Diego, La Jolla, CA

Andreas J. Kungl
Institute of Pharmaceutical Sciences, University of Graz, Graz, Austria

Sigrun Langbein
Department of Urology, Uro-Oncology, Academic Medic Centrum (AMC), University of Amsterdam, Amsterdam, The Netherlands

Xuguang Li
Centre for Biologics Research, Biological and Genetic Therapies Directorate, Health Canada, Ottawa, Canada

Guan Ning Lin
Digital Biology Laboratory, Computer Science Department, Christopher S. Bond Life Sciences Center, University of Missouri–Columbia, Columbia, MO

David Meierhofer
Department of Biological Chemistry, College of Medicine, University of California Irvine, Irvine, CA

Cathryn Mellersh
Animal Health Trust, Suffolk, United Kingdom

Snehal Naik
Molecular Imaging Center, Mallinckrodt Institute of Radiology, and Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO

Bernhard Ø. Palsson
Department of Bioengineering, University of California–San Diego, La Jolla, CA

Arturas Petronis
The Krembil Family Epigenetics Laboratory, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

David Piwnica-Worms
Molecular Imaging Center, Mallinckrodt Institute of Radiology, and Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO

Alexei Protopopov
Center for Applied Cancer Science, Belfer Institute for Innovative Cancer Science, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA

Jiannis Ragoussis
Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

Janna Saarela
National Public Health Institute, Helsinki, Finland

Axel Schumacher
Epigenetics Lab, Department of Medicine II, Klinikum rechts der Isar, Technical University Munich, Munich, Germany

Binesh Shrestha
Structural Genomics Consortium, Botnar Research Centre, University of Oxford, Oxford, United Kingdom

Kaisa Silander
National Public Health Institute, Helsinki, Finland

Mouldy Sioud
Department of Immunology, Institute for Cancer Research, Rikshopitalet - Radiumhospitalet Medical Center, Montebello, Oslo, Norway

Carol Smee
Structural Genomics Consortium, Botnar Research Centre, University of Oxford, Oxford, United Kingdom

Zhao Song
Digital Biology Laboratory, Computer Science Department, Christopher S. Bond Life Sciences Center, University of Missouri–Columbia, Columbia, MO

Danica B. Stanimirovic
Institute for Biological Sciences, National Research Council, Ottawa, Ontario, Canada

Minoru Takata
Shinshu University School of Medicine, Matsumoto, Japan

Badar A. Usmani
Institute of Cellular and Molecular Biology, University of Leeds, Leeds, United Kingdom

Gert Van den Bergh
Laboratory of Neuroplasticity and Neuroproteomics, Katholieke Universiteit Leuven, Leuven, Belgium

Victor Villalobos
Molecular Imaging Center, Mallinckrodt Institute of Radiology, and Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO

Xiaorong Wang
Department of Physiology and Biophysics and Department of Developmental and Cell Biology, University of California Irvine, Irvine, CA

Jessica Wang-Rodriguez
University of California, San Diego, Department of Pathology and the Veterans Administration San Diego Healthcare System, San Diego, CA

Andreas Weinhäusl
ARCS, Austrian Research Center, Seibersdorf, Austria

Adrian Woolfson
University of Cambridge School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, United Kingdom

Dong Xu
Digital Biology Laboratory, Computer Science Department, Christopher S. Bond Life Sciences Center, University of Missouri–Columbia, Columbia, MO

Joanne M. Yeakley
Illumina, Inc., San Diego, CA

Chao Zhang
Digital Biology Laboratory, Computer Science Department, Christopher S. Bond Life Sciences Center, University of Missouri–Columbia, Columbia, MO

Contents

Chapter 1
Whole Genome Amplification with Phi29 DNA Polymerase to Enable Genetic or Genomic Analysis of Samples of Low DNA Yield

Kaisa Silander and Janna Saarela

Abstract In many large genetic studies, the amount of available DNA can be one of the criteria for selecting samples for study. In the case of large population cohorts, selecting samples based on their DNA yield can lead to biased sample selection. In addition, many valuable clinical and research sample collections exist in which the amount of DNA is very small. Unbiased whole genome amplification (WGA) of such unique samples enables genomewide scale genetic studies that would have been impossible otherwise. Multiply primed rolling circle amplification (MPRCA) and multiple displacement amplification (MDA) methods are based on the same principle. The DNA amplification is non-PCR based and uses Φ29 DNA polymerase and random hexamer primers for unbiased whole genome amplification. MDA is used for linear DNA molecules, such as genomic DNA. This chapter reviews the various applications in which whole genome amplified DNA can be used, the types of commercial kits available, and the quality control steps necessary before using the DNA in the genetic studies.

Keywords multiply primed rolling circle amplification; multiple displacement amplification; WGA; DNA yield; genotyping

1 Introduction

1.1 Multiple Displacement Amplification and Other Methods for Whole Genome Amplification

A gold standard method for creating infinite DNA source is the immortalization of peripheral blood lymphocytes by transformation with Epstein-Barr virus. However, the method is laborious and expensive and thus not applicable to large-scale studies. Moreover, it does not allow amplification of the already existing genomic DNA sample collections. Several strategies have been utilized to amplify and preserve existing DNA samples. Most of the earlier methods have been based on cyclic amplification with Taq polymerase and utilized primers

From: Methods in Molecular Biology, Vol. 439: *Genomics Protocols: Second Edition*
Edited by Mike Starkey and Ramnath Elaswarapu

directed to repeated sequence elements, such as human Alu1-repeats [1], or random (PEP [2]) or partly degenerate (DOP [3]) primers. Both the originally described primer extension preamplification (PEP) and degenerate oligonucleotide primed polymerase chain reaction (DOP-PCR) techniques have been further developed and refined to increase the fidelity and the length of the amplification products [4, 5]. Another subgroup of whole-genome amplification methods relies on first generating fragments of the genomic DNA by restriction enzyme cleavage [6–8], random shearing of genomic DNA [9] or nick translation [10], followed by ligation of adaptors, which allow amplification using universal PCR primers. However, all these methods are limited by features of the Taq polymerase: typical amplification fragment length of <3 kb and the error rate of 3×10^{-5} (9×10^{-6} in combination with Pfu polymerase, reviewed in 11). These methods also suffer from incomplete coverage and uneven amplification of the genomic loci; 10^2–10^4- and 10^3–10^6-fold amplification biases have been described using PEP and DOP-PCR methods, respectively.

The multiple displacement amplification (MDA) method is based on isothermal amplification using the strand displacement activity of the Φ29 DNA polymerase [12].

MDA provides several improvements to the previously described methods. High processivity and fidelity of the Φ29 DNA polymerase allow highly uniform amplification across the genome: Dean et al. [12] observed less than a threefold amplification bias among the eight chromosomal loci monitored. However, their results indicated that repetitive sequences were underrepresented in the MDA product. The average amplification was 10,000-fold with the average product length of >10 kb. The MDA-amplified genomic DNA was shown to be suitable for several common genetic analyses, including single nucleotide polymorphisms (SNP) genotyping, chromosome painting, Southern blotting, restriction length polymorphism analysis, subcloning, and DNA sequencing [11]. In addition to high quality genomic DNA, the amplification can be carried out directly from biological samples, like whole blood or tissue culture cells [12, 13], plasma or serum [14], and laser-capture microdissected tissue [15]. Further, MDA amplification of low-copy-number and highly degraded DNA samples for forensic testing has been evaluated [16]. In addition to human samples, MDA method has also been utilized for canine [17], insect [18], and microbial samples [19]. Figure 1.1 shows the principle of WGA with Φ29 DNA polymerase and random hexamer primers.

1.2 *Applications in Which Whole Genome Amplified DNA Have Been Used*

Table 1.1 summarizes some of the subsequent studies evaluating the usability of MDA products in different genetic analyses. MDA-amplified genomic DNA has been successfully utilized in several downstream analyses, including sequencing, mutation detection, SNP and STR genotyping, Southern blot analysis, copy