NUCLEIC ACIDS HYBRIDIZATION

Nucleic Acids Hybridization
Modern Applications

Edited by

ANTON BUZDIN

Shemyakin-Ovchinnikov,
Institute of Bioorganic Chemistry,
Moscow, Russia

and

SERGEY LUKYANOV

Shemyakin-Ovchinnikov,
Institute of Bioorganic Chemistry,
Moscow, Russia

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-6039-7 (HB)
ISBN 978-1-4020-6040-3 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Cover idea is the courtesy of Dr. Lilia M. Ganova-Raeva

All Rights Reserved
© 2007 Springer
No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

CONTENTS

Contributors xi

Acknowledgments xv

Preface xvii

1. Nucleic Acids Hybridization: Potentials and Limitations 1
   Anton A. Buzdin
   1. Introduction 2
   2. Cloning the Differences: Subtractive Hybridization 4
      2.1 Birth of a Method 4
      2.2 PCR-assisted Subtractive Hybridization 8
      2.3 First Worldwide Success: Representational Differences Analysis 10
      2.4 Further Improvements: Suppression Subtractive Hybridization, Polymerase Chain Reaction Suppression Effect, and Normalization of cDNA Libraries 13
      2.5 Covalently Hybridized Subtraction, Primer Extension Enrichment Reaction, and Other Promising Approaches in Subtractive Hybridization 18
   3. Finding Common DNA: Coincidence Cloning 20
   4. Hybridization in Solution for the Recovery of Genomic Polymorphisms 22
   5. Conclusions 24
   References 25

2. Selective Suppression of Polymerase Chain Reaction and Its Most Popular Applications 29
Sergey A. Lukyanov, Konstantin A. Lukyanov, Nadezhda G. Gurskaya, Ekaterina A. Bogdanova, Anton A. Buzdin
1. Introduction 30
2. Selective PCR Suppression 30
3. Preparation of ITR-Containing DNA Samples 32
4. Strategy of Employment of SSP 33
5. SSP-based Procedures 33
5.1 Construction of cDNA Libraries from a Small Amount of Biological Material 35
5.2 Detection of Differentially Expressed Genes 36
5.3 Search for 5′- and 3′-Terminal Fragments of cDNA 40
5.4 Search for Promoter Sites (Chromosome Walking) 40
5.5 *In Vitro* Cloning 41
5.6 Multiplex PCR 42
6. Detailed Protocols 43
6.1 Materials 43
6.2 Methods 45
References 48

3. Suppression Subtractive Hybridization 53
Sergey A. Lukyanov, Denis Rebrikov, Anton A. Buzdin
1. Introduction 54
2. The Principle of Suppression Subtractive Hybridization 55
3. The Principle of Mirror-Oriented Selection 57
4. Technical Comments and Considerations 59
4.1 Normalization Step and the Efficiency of SSH 59
4.2 Differential Screening of the Subtracted Libraries 61
4.3 Starting RNA Materials 61
4.4 Size of cDNA Fragments 62
5. Detailed Protocols: Materials 63
5.1 Oligonucleotides 63
5.2 Buffers and Enzymes 63
6. Detailed Protocols: Methods 65
6.1 Preparation of Subtracted cDNA or Genomic DNA Library 65
6.2 Mirror Orientation Selection 70
6.3 Cloning of Subtracted cDNAs 74
6.4 Differential Screening of the Subtracted cDNA Library 74
7. Notes 77
References 83

4. Stem-Loop Oligonucleotides as Hybridization Probes and Their Practical Use in Molecular Biology and Biomedicine 85
Anton A. Buzdin, Sergey A. Lukyanov
1. Introduction 86
2. Molecular Beacons 87
3. Stem-Loop DNA Probes on Microarrays 92
4. Conclusions 93
References 94

5. Normalization of cDNA Libraries 97
Alex S. Shcheglov, Pavel A. Zhulidov, Ekaterina A. Bogdanova, Dmitry A. Shagin
1. Introduction 98
1.1 Normalized cDNA Libraries: What are they needed for? 98
1.2 Evaluation of cDNA Library Normalization Efficacy 100
1.3 Basic Approaches to Generate Normalized cDNA Libraries 100
2. Methods of cDNA Normalization 102
2.1 cDNA Normalization by Means of Hydroxyapatite Column Chromatography 102
2.2 Generation of Normalized Full-Length-Enriched cDNA Libraries Using DNA Immobilization on a Solid Support 104
2.3 Normalization of Full-Length cDNA with the use of Biotinylated RNA 105
2.4 Normalization of Fragmented cDNA by Means of Selective Amplification 107
2.5 cDNA Normalization Using Frequent-Cutter Restriction Enzymes 109
2.6 Normalization of Full-Length-Enriched cDNA with Duplex-Specific Nuclease (DSN Normalization) 109
3. cDNA Preparation for DSN Normalization 111
3.1 RNA Requirements 111
3.2 cDNA Synthesis 112
3.3 cDNA Purification 113
4. DSN Normalization Protocol 113
4.1 Materials 113
4.2 cDNA Precipitation 114
4.3 Hybridization 115
4.4 DSN Treatment 116
4.5 First Amplification of Normalized cDNA 117
4.6 Preliminary Analysis of the Normalization Results 119
4.7 Second Amplification of Normalized cDNA 121
References 122

6. Primer Extension Enrichment Reaction (PEER) and Other Methods for Difference Screening 125
Lilia M. Ganova-Raeva
1. Introduction 127
2. Primer Extension Enrichment Reaction 128
2.1 Method Outline 128
2.2 Discussion 132
2.3 PEER Protocol 135
3. Other Subtraction and Hybridization Based Methods for Difference Screening 137
3.1 Differential Screening 137
3.2 Subtractive Hybridization 139
3.3 Subtractive Cloning 140
3.4 Differential Display 140
3.5 AFLP, SAGE/CAGE, GSTs, and DARFA 142
3.6 Representational Differences Analysis 143
3.7 SPAD–RDA 146
3.8 Enzymatic Degradation Subtractions (EDS, LCS, DSC, NSC, UDG/USA, and CODE) 146
3.9 Suppression Subtraction Hybridization 150
3.10 Selective Amplification Via Biotin and Restriction-Mediated Enrichment 153
3.11 DNA Enrichment by Allele-Specific Hybridization 154
3.12 Methods Combining the use of SSH and Microarrays 155
3.13 Conclusions 156
References 157

7. Subtractive Hybridization with Covalently Modified Oligonucleotides 167
Shi-Lung Lin, Donald Chang, Joseph D. Miller, Shao-Yao Ying
1. Introduction 168
2. Subtractive Hybridization Methods 169
3. Covalent Modification 171
4. Subtractive Hybridization with Covalently Modified Subtracters 176
5. Applications 179
6. Protocols 180
6.1 Preparation of Subtracter and Tester DNA Libraries 180
6.2 Covalent Modification of Subtracter DNAs 180
6.3 Subtractive Hybridization and CHS–PCR Amplification 181

6.4 Covalent Binding Efficiency and Subtractive Stringency of CHS 183
6.5 Identification of Genomic Deletion Using CHS 184
References 185

8. Coincidence Cloning: Robust Technique for Isolation of Common Sequences 187
Anton A. Buzdin
1. Introduction 188
2. Cloning Selection of Heteroduplexes 191
3. Physical Separation of Hybrids 193
4. PCR-only-based Approaches 193
5. Future Prospects 202
6. Protocols 202
6.1 Cloning Similarities in Genomic DNAs 202
6.2 Cloning and Presice Mapping of Transcribed Repetitive Elements 204
6.3 Finding Methylated or Unmethylated CpGs in Large Genomic Contigs 207
References 209

9. DNA Hybridization in Solution for Mutation Detection 211
Anton A. Buzdin
1. Introduction 212
2. Chemical Approaches 215
3. Enzymatic Approaches 217
3.1 Nuclease-Based Mutation Scanning 218
3.2 Allele-Specific PCR-Based Approaches 229
3.3 Other Enzymatic Approaches for Mutation Scanning 230
4. Physical Approaches 233
5. Bioinformatical Approaches 235
References 236

10. Current Attempts to Improve the Specificity of Nucleic Acids Hybridization 241
Anton A. Buzdin
1. Introduction 242
2. Improving Hybridization Kinetics 244
2.1 Simplification of Hybridizing Mixtures 246
2.2 Chemical Modifications 248
3. Improving Selection of Perfectly Matched Hybrids 249

4. Protocols 255
4.1 Targeted Genomic Difference Analysis 255
4.2 Using Competitor DNA to Decrease the Background of Genomic Repeats 258
4.3 Mispaired DNA Rejection 260
References 262

11. Concepts on Microarray Design for Genome and Transcriptome Analyses 265
Helder I. Nakaya, Eduardo M. Reis, Sergio Verjovski-Almeida
1. Building a Microarray Chip 266
1.1 Spotted DNA Microarrays 268
1.2 *In situ* Synthesis 271
2. Selecting the Probes 277
2.1 Gene-Oriented Arrays 278
2.2 Epigenomic Microarrays 279
2.3 Tiling Arrays 279
3. Specific Question, Specific Chip 280
3.1 Transcriptional Profiling 280
3.2 Comparative Genome Hybridization 282
3.3 Alternative Splicing 284
3.4 Transcriptome Annotation 285
3.5 Small MicroRNA Profiling 289
3.6 Methylation Pattern 289
3.7 ChIP-Chip 293
3.8 Genotyping 295
3.9 Intronic Transcription 297
4. Conclusions 299
References 300

Index 309

CONTRIBUTORS

Ekaterina A. Bogdanova
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Anton A. Buzdin
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Nadezhda G. Gurskaya
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Konstantin A. Lukyanov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Sergey A. Lukyanov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Denis Rebrikov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Dmitry A. Shagin
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Alex S. Shcheglov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Pavel A. Zhulidov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
Moscow 117997, Miklukho-Maklaya 16/10, Russia

Donald Chang
Department of Cell and Neurobiology, Keck School of Medicine,
University of Southern California,
1333 San Pablo Street, BMT-403,
Los Angeles, CA 90033, USA

Shi-Lung Lin
Department of Cell and Neurobiology, Keck School of Medicine,
University of Southern California,
1333 San Pablo Street, BMT-403,
Los Angeles, CA 90033, USA

Joseph Miller
Department of Cell and Neurobiology, Keck School of Medicine,
University of Southern California,
1333 San Pablo Street, BMT-403,
Los Angeles, CA 90033, USA

Shao-Yao Ying
Department of Cell and Neurobiology, Keck School of Medicine,
University of Southern California,
1333 San Pablo Street, BMT-403,
Los Angeles, CA 90033, USA

Lilia M. Ganova-Raeva
Centers for Disease Control and Prevention
Division of Viral Hepatitis
1600 Clifton Rd. NE, MS A-33
Atlanta, Georgia 30329, USA

Helder I. Nakaya
Departamento de Bioquimica, Instituto de Quimica,
Universidade de São Paulo,
05508-900 São Paulo, SP, Brazil

Eduardo M. Reis
Departamento de Bioquimica, Instituto de Quimica,
Universidade de São Paulo,
05508-900 São Paulo, SP, Brazil

Sergio Verjovski-Almeida
Departamento de Bioquimica, Instituto de Quimica,
Universidade de São Paulo,
05508-900 São Paulo, SP, Brazil

ACKNOWLEDGMENTS

Anton A. Buzdin:
Many thanks to Professor Eugene D. Sverdlov for his fruitful discussion, innovative ideas, and overall support of this project. Thanks to my colleagues, friends, and family members for their help, patience, and understanding. A. Buzdin was funded by the Molecular and Cellular Biology Program of the Presidium of the Russian Academy of Sciences, by the personal grant from the President of the Russian Federation, and by Russian Foundation for Basic Research grants Nos. 05-04-48682-a, 2006.20034.

Sergey A. Lukyanov, Alex Shcheglov, Pavel Zhulidov, Ekaterina Bogdanova, Dmitry Shagin:
Work on this book was supported by the Molecular and Cellular Biology Program of the Russian Academy of Sciences and Evorogen JSC (Moscow, Russia).

Lilia M. Ganova-Raeva:
Special thanks to Dr. Y. Khudyakov who contributed most to the PEER backbone idea and has been relentlessly resourceful, helpful, and patient throughout the development of the method. Thanks to Dr. H. Fields in whose lab the PEER testing was initiated. Thanks to Dr. X. Zhang for great help with the library screenings and to Dr. F. Cao for introducing better enzymes in the protocol.

Shi-Lung Lin, Donald Chang, Joseph D. Miller, Shao-Yao Ying:
This study was supported by NIH/NCI grant CA-85722. Rhw CHS technology is the property of University of Southern California and protected by US patent numbers, 5,928,872 and 6,130,040.

Helder I. Nakaya, Eduardo M. Reis, Sergio Verjovski-Almeida:
The work in the authors' laboratory was supported by grants and fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), and Conselho Nacional do Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

PREFACE

Watson–Crick hybridization of complementary sequences in nucleic acids is one of the most important fundamental processes necessary for molecular recognition *in vivo*, as well as for nucleic acid identification and isolation *in vitro*. This book is devoted to a large family of *in vitro* DNA hybridization-based experimental techniques. A wide spectrum of experimental tasks covered by these approaches includes finding differential sequences in both genomic DNAs and mRNAs, genome walking, multiplex PCR, cDNA library construction starting from minute amount of total RNA, rapid amplification of cDNA 5′- and 3′-ends, effective smoothing of the concentrations of rare and abundant transcripts in cDNA libraries, recovery of promoter active repeats and differentially methylated genomic DNA, identification of common sequences in genomic or cDNA sources, new gene mapping, finding evolutionary conserved DNA and both single-nucleotide and extended mutation discovery, or large-scale monitoring. Several approaches, such as microarray hybridization, have become extremely popular tools for specialists in biochemistry and biomedicine, whereas the potential of many other advantageous techniques seems to be underestimated now.

Analysis of differential gene expression requires application of global approaches that represent a leading tool in postgenomic studies and include transcriptome and proteome analysis, as well as methods allowing population-wide sequence and functional polymorphism analysis. Central to these new technologies are DNA chips designed for quantitative and qualitative uses. Although they are very useful and widely distributed, many popular DNA microarray techniques share a number of shortcomings:

1. The analysis is limited by a number of cDNAs/synthetic oligonucleotides applied on the chip. This number is usually significantly lower than the total

gene quantity of the organisms under study. It creates, therefore, a problem, that many genes escape such an analysis.
2. General transcriptome-wide chip techniques in their actual state hardly distinguish between different gene splice forms.
3. The expression of genes transcribed at low levels cannot be detected by using standard microarray approaches.
4. cDNA-based chips do not differentiate between many gene family members and/or between many transcripts containing repetitive DNA.
5. Microarray chips lack many natural RNAi, cDNAs, or synthetic oligonucleotides and, therefore, cannot be used for the comprehensive study of gene expression regulation at the level of RNA interference by small interfering RNAs.

However, most of these concerns can be effectively addressed by using specific variants of microchip technology, thus making microarrays a truly universal technique (see Chapter 11). Probably, the most important disadvantage of closed systems such as microarrays is that they require preliminary genomic sequence information in order to identify differentially expressed transcripts.

Open systems have the flexibility of identifying uncatalogued sequences. Related experimental techniques, based on DNA hybridization in solution, may be advantageous for many applications, starting from representative cDNA library construction for expressed sequence tag (EST) sequencing, to the identification of evolutionary conserved sequences, differentially expressed genes, or genomic deletions. Unique characteristics of many such techniques make them powerful competitors for well-known approaches that are appreciated worldwide like microarray and competitive genomic hybridizations. Nucleic acid hybridization in solution has few general advantages over hybridization with solid carrier-immobilized nucleic acids: faster hybridization kinetics, better discrimination of proper hybrids, and their availability for further PCR amplification and cloning. Among such in-solution hybridization methods, subtractive hybridization is undoubtedly the most popular technique.

Many techniques have a low efficiency of identifying rare genes that are differentially expressed. This problem is exacerbated when the change in expression level of rare transcripts is small. Since genes expressed at low levels also play a role in establishing differentiated phenotypes, their identification is essential for a complete mechanistic understanding of cellular changes. The major advantage of subtractive hybridization lies in the ability to identify differentially transcribed genes, irrespective of the level of expression, in the absence of sequence information. In addition to preparation of differential cDNA libraries, subtractive hybridization is also extremely useful for identification of genomic DNA differences.

Coincidence cloning, on the contrary, makes it possible to identify sequences which are common for all samples under comparison; cDNA normalization, which is used for smoothing of rare and frequent transcript content in cDNA libraries, may be extremely useful for representative EST library construction. Moreover, several techniques deal with the large-scale DNA polymorphism recovery, including identification of single nucleotide polymorphisms.

The international team of the authors of this book has tried both to elucidate the current state of the art in hybridization techniques and to help the readers in choosing an appropriate method for performing an experiment in the most efficient way. Enclosed experimental protocols along with both comprehensive and detailed method descriptions make this truly universal book useful to all those interested in the modern life science methodologies.