

Methods in Molecular Biology™

VOLUME 192

PCR Cloning Protocols

Second Edition

Edited by

Bing-Yuan Chen

Harry W. Janes



HUMANA PRESS

METHODS IN MOLECULAR BIOLOGY™

John M. Walker, SERIES EDITOR

210. **MHC Protocols**, edited by *Stephen H. Powis and Robert W. Vaughan, 2003*
209. **Transgenic Mouse Methods and Protocols**, edited by *Marten Hofker and Jan van Deursen, 2002*
208. **Peptide Nucleic Acids: Methods and Protocols**, edited by *Peter E. Nielsen, 2002*
207. **Human Antibodies for Cancer Therapy: Reviews and Protocols**, edited by *Martin Welschof and Jürgen Krauss, 2002*
206. **Endothelin Protocols**, edited by *Janet J. Maguire and Anthony P. Davenport, 2002*
205. **E. coli Gene Expression Protocols**, edited by *Peter E. Vaillancourt, 2002*
204. **Molecular Cytogenetics: Methods and Protocols**, edited by *Yao-Shan Fan, 2002*
203. **In Situ Detection of DNA Damage: Methods and Protocols**, edited by *Vladimir V. Didenko, 2002*
202. **Thyroid Hormone Receptors: Methods and Protocols**, edited by *Aria Baniahmad, 2002*
201. **Combinatorial Library Methods and Protocols**, edited by *Lisa B. English, 2002*
200. **DNA Methylation Protocols**, edited by *Ken I. Mills and Bernie H. Ramsahoye, 2002*
199. **Liposome Methods and Protocols**, edited by *Subhash C. Basu and Manju Basu, 2002*
198. **Neural Stem Cells: Methods and Protocols**, edited by *Tanja Zigova, Juan R. Sanchez-Ramos, and Paul R. Sanberg, 2002*
197. **Mitochondrial DNA: Methods and Protocols**, edited by *William C. Copeland, 2002*
196. **Oxidants and Antioxidants: Ultrastructural and Molecular Biology Protocols**, edited by *Donald Armstrong, 2002*
195. **Quantitative Trait Loci: Methods and Protocols**, edited by *Nicola J. Camp and Angela Cox, 2002*
194. **Posttranslational Modifications of Proteins: Tools for Functional Proteomics**, edited by *Christoph Kannicht, 2002*
193. **RT-PCR Protocols**, edited by *Joseph O'Connell, 2002*
192. **PCR Cloning Protocols, 2nd ed.**, edited by *Bing-Yuan Chen and Harry W. Janes, 2002*
191. **Telomeres and Telomerase: Methods and Protocols**, edited by *John A. Double and Michael J. Thompson, 2002*
190. **High Throughput Screening: Methods and Protocols**, edited by *William P. Janzen, 2002*
189. **GTPase Protocols: The RAS Superfamily**, edited by *Edward J. Manser and Thomas Leung, 2002*
188. **Epithelial Cell Culture Protocols**, edited by *Clare Wise, 2002*
187. **PCR Mutation Detection Protocols**, edited by *Bimal D. M. Theophilus and Ralph Rapley, 2002*
186. **Oxidative Stress and Antioxidant Protocols**, edited by *Donald Armstrong, 2002*
185. **Embryonic Stem Cells: Methods and Protocols**, edited by *Kursad Turksen, 2002*
184. **Biostatistical Methods**, edited by *Stephen W. Looney, 2002*
183. **Green Fluorescent Protein: Applications and Protocols**, edited by *Barry W. Hicks, 2002*
182. **In Vitro Mutagenesis Protocols, 2nd ed.**, edited by *Jeff Braman, 2002*
181. **Genomic Imprinting: Methods and Protocols**, edited by *Andrew Ward, 2002*
180. **Transgenesis Techniques, 2nd ed.:** *Principles and Protocols*, edited by *Alan R. Clarke, 2002*
179. **Gene Probes: Principles and Protocols**, edited by *Marilena Aquino de Muro and Ralph Rapley, 2002*
178. **Antibody Phage Display: Methods and Protocols**, edited by *Philippa M. O'Brien and Robert Aitken, 2001*
177. **Two-Hybrid Systems: Methods and Protocols**, edited by *Paul N. MacDonald, 2001*
176. **Steroid Receptor Methods: Protocols and Assays**, edited by *Benjamin A. Lieberman, 2001*
175. **Genomics Protocols**, edited by *Michael P. Starkey and Ramnath Elasarapu, 2001*
174. **Epstein-Barr Virus Protocols**, edited by *Joanna B. Wilson and Gerhard H. W. May, 2001*
173. **Calcium-Binding Protein Protocols, Volume 2: Methods and Techniques**, edited by *Hans J. Vogel, 2001*
172. **Calcium-Binding Protein Protocols, Volume 1: Reviews and Case Histories**, edited by *Hans J. Vogel, 2001*
171. **Proteoglycan Protocols**, edited by *Renato V. Iozzo, 2001*
170. **DNA Arrays: Methods and Protocols**, edited by *Jang B. Rampil, 2001*
169. **Neurotrophin Protocols**, edited by *Robert A. Rush, 2001*
168. **Protein Structure, Stability, and Folding**, edited by *Kenneth P. Murphy, 2001*
167. **DNA Sequencing Protocols, Second Edition**, edited by *Colin A. Graham and Alison J. M. Hill, 2001*
166. **Immunotoxin Methods and Protocols**, edited by *Walter A. Hall, 2001*
165. **SV40 Protocols**, edited by *Leda Raptis, 2001*
164. **Kinesin Protocols**, edited by *Isabelle Vernos, 2001*
163. **Capillary Electrophoresis of Nucleic Acids, Volume 2: Practical Applications of Capillary Electrophoresis**, edited by *Keith R. Mitchelson and Jing Cheng, 2001*
162. **Capillary Electrophoresis of Nucleic Acids, Volume 1: Introduction to the Capillary Electrophoresis of Nucleic Acids**, edited by *Keith R. Mitchelson and Jing Cheng, 2001*
161. **Cytoskeleton Methods and Protocols**, edited by *Ray H. Gavin, 2001*
160. **Nuclease Methods and Protocols**, edited by *Catherine H. Schein, 2001*
159. **Amino Acid Analysis Protocols**, edited by *Catherine Cooper, Nicole Packer, and Keith Williams, 2001*
158. **Gene Knockout Protocols**, edited by *Martin J. Tymms and Ismail Kola, 2001*
157. **Mycotoxin Protocols**, edited by *Mary W. Trucksess and Albert E. Pohland, 2001*
156. **Antigen Processing and Presentation Protocols**, edited by *Joyce C. Solheim, 2001*
155. **Adipose Tissue Protocols**, edited by *Gérard Ailhaud, 2000*
154. **Connexin Methods and Protocols**, edited by *Roberto Bruzzone and Christian Giaume, 2001*
153. **Neuropeptide Y Protocols**, edited by *Ambikaipakan Balasubramaniam, 2000*
152. **DNA Repair Protocols: Prokaryotic Systems**, edited by *Patrick Vaughan, 2000*

METHODS IN MOLECULAR BIOLOGY™

PCR Cloning Protocols

Second Edition

Edited by

Bing-Yuan Chen

and

Harry W. Janes

*Rutgers University,
New Brunswick, NJ*


Humana Press  Totowa, New Jersey

© 2002 Humana Press Inc.
999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

www.humanapress.com

All rights reserved. No part of this book 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. *Methods in Molecular Biology™* is a trademark of The Humana Press Inc.

All papers, comments, opinions, conclusions, or recommendations are those of the author(s), and do not necessarily reflect the views of the publisher.

This publication is printed on acid-free paper. 
ANSI Z39.48-1984 (American Standards Institute) Permanence of Paper for Printed Library Materials.

Cover illustration: The Subtracted cDNA amplification of PBMCs stimulated with PHA. *See Fig. 2 on page 107.*

Cover design by Patricia F. Cleary.

Production Editor: Mark J. Breaugh.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: humana@humanapr.com; Website: <http://humanapress.com>

Photocopy Authorization Policy:

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Humana Press Inc., provided that the base fee of US \$10.00 per copy, plus US \$00.25 per page, is paid directly to the Copyright Clearance Center at 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license from the CCC, a separate system of payment has been arranged and is acceptable to Humana Press Inc. The fee code for users of the Transactional Reporting Service is [0-89603-969-2/02 \$10.00 + \$00.25].

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging in Publication Data

Main entry under title: *Methods in molecular biology™*.

PCR cloning protocols: second edition / edited by Bing-Yuan Chen and Harry W. Janes.--2nd ed.
p. cm. -- (Methods in molecular biology ; 192)

Includes bibliographical references and index.

ISBN 0-89603-969-2 (hb : alk. paper) -- ISBN 0-89603-973-0 (comb. : alk. paper)

1. Molecular cloning--Laboratory manuals. 2. Polymerase chain reaction--Laboratory manuals. I. Chen, Bing-Yuan. II. Janes, Harry W. III. *Methods in molecular biology* (Clifton, N.J.) ; v. 192

QH442.2 .P37 2002
572.8'6--dc21

2001039702

Preface

PCR is probably the single most important methodological invention in molecular biology to date. Since its conception in the mid-1980s, it has rapidly become a routine procedure in every molecular biology laboratory for identifying and manipulating genetic material, from cloning, sequencing, mutagenesis, to diagnostic research and genetic analysis. What's astounding about this invention is that new and innovative applications of PCR have been generated with stunning regularity; its potential has shown no signs of leveling off. New applications for PCR are literally transforming molecular biology. In the post-genomic era, PCR has especially become the method of choice to clone existing genes and generate a wide array of new genes by mutagenesis and/or recombination within the genes of interest. The fast and easy availability of these genes is essential for the study of functional genomics, gene expression, protein structure–function relationships, protein–protein interactions, protein engineering, and molecular evolution.

PCR Cloning Protocols was prepared in response to the need to have an up-to-date compilation of proven protocols for PCR cloning and mutagenesis. It builds upon the best-selling first edition, *PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering*, a book in the *Methods in Molecular Biology*TM series published in 1997. We divided the new edition into five parts. Part I. **Performing and Optimizing PCR**, contains basic PCR methodology, including PCR optimization and computer programs for PCR primer design and analysis, as well as novel variations for cloning genes of particular characteristics or origins, emphasizing long-distance PCR and GC-rich template amplification. Part II. **Cloning PCR Products**, presents both conventional and novel enzyme-free and restriction site-free procedures to clone PCR products into various vectors, either directionally or non-directionally. Part III. **Mutagenesis and Recombination**, addresses the use of PCR to facilitate DNA mutagenesis and recombination in various innovative approaches to generate a wide array of mutants. Part IV. **Cloning Unknown Neighboring DNA**, contains a comprehensive collection of protocols to fulfill the frequent and challenging task of cloning uncharacterized DNA flanking a known DNA fragment. Finally, Part V. **Library Construction and Screening**, addresses particular applications of PCR in library and sublibrary generation and screening. Each part also contains an overview, which summarizes the current methods available and their underlying

strategies, advantages, and disadvantages for that particular topic. These reviews are especially helpful to new researchers to orient themselves with the field and to guide them to choose a procedure that is most suitable for their experiments.

We hope that *PCR Cloning Protocols* will provide readily reproducible laboratory protocols that researchers in the field will follow closely and thereby increase their success rate in their experiments.

We are indebted to Mirah Riben for her superb help during the editing of the book. We also thank Prof. John M. Walker, the series editor, for his help, advice, and guidance.

Bing-Yuan Chen

Harry W. Janes

Contents

Preface	v
Contributors	xi
PART I. PERFORMING AND OPTIMIZING PCR	
1 Polymerase Chain Reaction: <i>Basic Principles and Routine Practice</i> Lori A. Kolmodin and David E. Birch	3
2 Computer Programs for PCR Primer Design and Analysis Bing-Yuan Chen, Harry W. Janes, and Steve Chen	19
3 Single-Step PCR Optimization Using Touchdown and Stepdown PCR Programming Kenneth H. Roux	31
4 XL PCR Amplification of Long Targets from Genomic DNA Lori A. Kolmodin	37
5 Coupled One-Step Reverse Transcription and Polymerase Chain Reaction Procedure for Cloning Large cDNA Fragments Jyrki T. Aatsinki	53
6 Long Distance Reverse-Transcription PCR Volker Thiel, Jens Herold, and Stuart G. Siddell	59
7 Increasing PCR Sensitivity for Amplification from Paraffin-Embedded Tissues Abebe Akalu and Juergen K. V. Reichardt	67
8 GC-Rich Template Amplification by Inverse PCR: <i>DNA Polymerase and Solvent Effects</i> Alain Moreau, Da Shen Wang, Steve Forget, Colette Duez, and Jean Dusart	75
9 PCR Procedure for the Isolation of Trinucleotide Repeats Teruaki Tozaki	81
10 Methylation-Specific PCR Haruhiko Ohashi	91

11	Direct Cloning of Full-Length Cell Differentially Expressed Genes by Multiple Rounds of Subtractive Hybridization Based on Long-Distance PCR and Magnetic Beads Xin Huang, Zhenglong Yuan, and Xuetao Cao	99
PART II. CLONING PCR PRODUCTS		
12	Cloning PCR Products: <i>An Overview</i> Baotai Guo and Yuping Bi	111
13	Using T4 DNA Polymerase to Generate Clonable PCR Products Kai Wang	121
14	Enzyme-Free Cloning of PCR Products and Fusion Protein Expression Brett A. Neilan and Daniel Tillett	125
15	Directional Restriction Site-Free Insertion of PCR Products into Vectors Guo Jun Chen	133
16	Autosticky PCR: <i>Directional Cloning of PCR Products with Preformed 5' Overhangs</i> József Gál and Miklós Kálmán	141
17	A Rapid and Simple Procedure for Direct Cloning of PCR Products into Baculoviruses Tamara S. Gritsun, Michael V. Mikhailov, and Ernest A. Gould	153
PART III. MUTAGENESIS AND RECOMBINATION		
18	PCR Approaches to DNA Mutagenesis and Recombination: <i>An Overview</i> Binzhang Shen	167
19	In-Frame Cloning of Synthetic Genes Using PCR Inserts James C. Pierce	175
20	Megaprimer PCR Sailen Barik	189
21	PCR-Mediated Recombination: <i>A General Method Applied to Construct Chimeric Infectious Molecular Clones</i> Guowei Fang, Barbara Weiser, Aloise Visosky, Timothy Moran, and Harold Burger	197
22	PCR Method for Generating Multiple Mutations at Adjacent Sites Jiri Adamec	207

23	A Fast Polymerase Chain Reaction-Mediated Strategy for Introducing Repeat Expansions into CAG-Repeat Containing Genes Franco Laccone	217
24	PCR Screening in Signature-Tagged Mutagenesis of Essential Genes Dario E. Lehoux and Roger C. Levesque	225
25	Staggered Extension Process (StEP) In Vitro Recombination Anna Marie Aguinaldo and Frances Arnold	235
26	Random Mutagenesis by Whole-Plasmid PCR Amplification Donghak Kim and F. Peter Guengerich	241
PART IV. CLONING UNKNOWN NEIGHBORING DNA		
27	PCR-Based Strategies to Clone Unknown DNA Regions from Known Foreign Integrants: <i>An Overview</i> Eric Ka-Wai Hui, Po-Ching Wang, and Szecheng J. Lo	249
28	Long Distance Vectorette PCR (LDV PCR) James A. L. Fenton, Guy Pratt, and Gareth J. Morgan	275
29	Nonspecific, Nested Suppression PCR Method for Isolation of Unknown Flanking DNA ("Cold-Start Method") Michael Lardelli	285
30	Inverse PCR: <i>cDNA Cloning</i> Sheng-He Huang	293
31	Inverse PCR: <i>Genomic DNA Cloning</i> Ambrose Y. Jong, Anna T'ang, De-Pei Liu, and Sheng-He Huang	301
32	Gene Cloning and Expression Profiling by Rapid Amplification of Gene Inserts with Universal Vector Primers Sheng-He Huang, Hua-Yang Wu, and Ambrose Y. Jong	309
33	The Isolation of DNA Sequences Flanking Tn5 Transposon Insertions by Inverse PCR Vincent J. J. Martin and William W. Mohn	315
34	Rapid Amplification of Genomic DNA Sequences Tagged by Insertional Mutagenesis Martina Celerin and Kristin T. Chun	325
35	Isolation of Large Terminal Sequences of BAC Inserts Based on Double-Restriction-Enzyme Digestion Followed by Anchored PCR Zhong-Nan Yang and T. Erik Mirkov	337

36	A “Step Down” PCR-Based Technique for Walking Into and the Subsequent Direct Sequence Analysis of Flanking Genomic DNA Ziguo Zhang and Sarah Jane Gurr	343
PART V. LIBRARY CONSTRUCTION AND SCREENING		
37	Use of PCR in Library Screening: <i>An Overview</i> Jinbao Zhu	353
38	Cloning of Homologous Genes by Gene-Capture PCR Renato Mastrangeli and Silvia Donini	359
39	Rapid and Nonradioactive Screening of Recombinant Libraries by PCR Michael W. King	377
40	Rapid cDNA Cloning by PCR Screening (RC-PCR) Toru Takumi	385
41	Generation and PCR Screening of Bacteriophage λ Sublibraries Enriched for Rare Clones (the “Sublibrary Method”) Michael Lardelli	391
42	PCR-Based Screening for Bacterial Artificial Chromosome Libraries Yuji Yasukochi	401
43	A 384-Well Microtiter-Plate-Based Template Preparation and Sequencing Method Lei He and Kai Wang	411
44	A Microtiter-Plate-Based High Throughput PCR Product Purification Method Ryan Smith and Kai Wang	417
	Index	423

Contributors

- JYRKI T. AATSINKI • *Institute of Dentistry, University of Oulu, Finland*
- JIRI ADAMEC • *Mayo Clinic and Foundation, Rochester, MN*
- ANNA MARIE AGUINALDO • *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA*
- ABEBE AKALU • *Institute for Genetic Medicine, USC School of Medicine, Los Angeles, CA*
- FRANCES ARNOLD • *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA*
- SAILEN BARIK • *Department of Biochemistry and Molecular Biology, University of South Alabama, Mobile, AL*
- YUPING BI • *Institute of Plant Biotechnology, Shangdong Academy of Agricultural Sciences, Jinan, China*
- DA' ID E. BIRCH • *Roche Molecular Systems, Alameda, CA*
- HAROLD BURGER • *Wadsworth Center, Albany, NY*
- XUETAO CAO • *Department of Immunology, Second Military Medical University, Shanghai, China*
- MARTINA CELERIN • *Department of Biology, Indiana University, Bloomington, IN*
- BING-YUAN CHEN • *Department of Plant Science, Rutgers University, New Brunswick, NJ*
- GUO JUN CHEN • *F. Hoffmann La-Roche, Basel, Switzerland*
- STEVE CHEN • *NetOsprey Inc., Berkeley, CA*
- KRISTIN T. CHUN • *Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN*
- SILVIA DONINI • *Istituto di Ricerca Cesare Serono, Rome, Italy*
- COLETTE DUEZ • *Centre D'Ingénierie des Protéines, Université de Liège, Liege, Belgium*
- JEAN DUSART • *Centre D'Ingénierie des Protéines, Université de Liège, Liege, Belgium*
- GUOWEI FANG • *Wadsworth Center, Albany, NY*
- JAMES A. L. FENTON • *Department of Molecular Oncology, University of Leeds, Leeds, UK*
- STEVE FORGET • *Sainte-Justine Hospital Research Center, Montreal, Canada*
- JÓZSEF GÁL • *Institute for Biotechnology, Bay Zoltán Foundation for Applied Research, Szeged, Hungary*

- ERNEST A. GOULD • *CEH Oxford, Oxford, UK*
- TAMARA S. GRITSUN • *CEH Oxford, Oxford, UK*
- F. PETER GUENGERICH • *Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, TN*
- BAOTAI GUO • *Institute of Plant Biotechnology, Laiyang Agricultural College, Shandong, China*
- SARAH JANE GURR • *Department of Plant Sciences, University of Oxford, Oxford, UK*
- LEI HE • *PhenoGenomics Corp., Bothell, WA*
- JENS HEROLD • *SWITCH-Biotech AG, Martinsried, Germany*
- SHENG-HE HUANG • *Department of Pediatrics, University of Southern California, Los Angeles, CA*
- XIN HUANG • *Department of Immunology, Second Military Medical University, Shanghai, China*
- ERIC KA-WAI HUI • *Department of Microbiology, Immunology and Molecular Genetics, University of California Los Angeles, Los Angeles, CA*
- HARRY W. JANES • *Department of Plant Science, Rutgers University, New Brunswick, NJ*
- AMBROSE Y. JONG • *Department of Pediatrics, University of Southern California, Los Angeles, CA*
- MIKLÓS KÁLMÁN • *Institute for Biotechnology, Bay Zoltán Foundation for Applied Research, Szeged, Hungary*
- DONGHAK KIM • *Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, TN*
- MICHAEL W. KING • *Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Terre Haute, IN*
- LORI A. KOLMODIN • *Roche Molecular Systems, Pleasanton, CA*
- FRANCO LACCONE • *Institute of Human Genetics, University of Goettingen, Goettingen, Germany*
- MICHAEL LARDELLI • *Department of Molecular Biosciences, Adelaide University, Australia*
- DARIO E. LEHOUX • *Health and Life Sciences Research Center, Université Laval, Sainte-Foy, Québec, Canada*
- ROGER C. LEVESQUE • *Health and Life Sciences Research Center, Université Laval, Sainte-Foy, Québec, Canada*
- DE-PEI LIU • *Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*
- SZECHEUNG J. LO • *Institute of Microbiology and Immunology, National Yang-Ming University, Taipei, Taiwan, ROC*

- VINCENT J. J. MARTIN • *Department of Chemical Engineering, University of California, Berkeley, CA*
- RENATO MASTRANGELI • *Istituto di Ricerca Cesare Serono, Rome, Italy*
- MICHAEL V. MIKHAILOV • *CEH Oxford, Oxford, UK*
- T. ERIK MIRKOV • *Department of Plant Pathology and Microbiology, The Texas A&M University Agricultural Experiment Station, Weslaco, TX*
- WILLIAM W. MOHN • *Department of Microbiology and Immunology, University of British Columbia, Vancouver, Canada*
- TIMOTHY MORAN • *Wadsworth Center, Albany, NY*
- ALAIN MOREAU • *Sainte-Justine Hospital Research Center, Montreal, Canada*
- GARETH J. MORGAN • *Department of Molecular Oncology, University of Leeds, Leeds, UK*
- BRETT A. NEILAN • *School of Microbiology and Immunology, the University of New South Wales, Sydney, Australia*
- HARUHIKO OHASHI • *Nagoya National Hospital, Nagoya, Japan*
- JAMES C. PIERCE • *University of the Sciences in Philadelphia, Philadelphia, PA*
- GUY PRATT • *Department of Molecular Oncology, University of Leeds, Leeds, UK*
- JUERGEN K.V. REICHARDT • *Institute for Genetic Medicine, USC School of Medicine, Los Angeles, CA*
- KENNETH H. ROUX • *Department of Biological Science, Florida State University, Tallahassee, FL*
- BINZHANG SHEN • *Department of Molecular Biology, Massachusetts General Hospital, Boston, MA*
- STUART G. SIDDELL • *Institute of Virology and Immunology, University of Würzburg, Würzburg, Germany*
- RYAN SMITH • *PhenoGenomics Corp., Bothell, WA*
- ANNA T'ANG • *Department of Pathology, University of Southern California, Los Angeles, CA*
- TORU TAKUMI • *Osaka Bioscience Institute, Osaka, Japan*
- VOLKER THIEL • *Institute of Virology and Immunology, University of Würzburg, Würzburg, Germany*
- DANIEL TILLET • *School of Microbiology and Immunology, University of New South Wales, Sydney, Australia*
- TERUAKI TOZAKI • *Department of Molecular Genetics, Laboratory of Racing Chemistry, Utsunomiya, Tochigi, Japan*
- ALOISE VISOSKY • *Wadsworth Center, Albany, NY*
- DA SHEN WANG • *Sainte-Justine Hospital Research Center, Montreal, Canada*
- KAI WANG • *PhenoGenomics Corp., Bothell, WA*
- PO-CHING WANG • *Department of Medicine, National Yang-Ming University, Taipei, Taiwan, ROC*

BARBARA WEISER • *Wadsworth Center, Albany, NY*

HUA-YANG WU • *Department of Pediatrics, University of Southern California, Los Angeles, CA*

ZHONG-NAN YANG • *Department of Plant Pathology and Microbiology, The Texas A&M University Agricultural Experiment Station, Weslaco, TX*

YUJI YASUKOCHI • *National Institute of Agrobiological Sciences, Ibaraki, Japan*

ZHENGLONG YUAN • *Department of Immunology, Second Military Medical University, Shanghai, China*

ZIGUO ZHANG • *Department of Plant Sciences, University of Oxford, Oxford, UK*

JINBAO ZHU • *Department of Genetics and Plant Breeding, China Agricultural University, Beijing, China*

I _____

PERFORMING AND OPTIMIZING PCR

Polymerase Chain Reaction

Basic Principles and Routine Practice

Lori A. Kolmodin and David E. Birch

1. Introduction

1.1. PCR Definition

The polymerase chain reaction (PCR) is a primer-mediated enzymatic amplification of specifically cloned or genomic DNA sequences (*1*). This PCR process, invented more than a decade ago, has been automated for routine use in laboratories worldwide. The template DNA contains the target sequence, which may be tens or tens of thousands of nucleotides in length. A thermostable DNA polymerase such as *Taq* DNA polymerase, catalyzes the buffered reaction in which an excess of an oligonucleotide primer pair and four deoxynucleoside triphosphates (dNTPs) are used to make millions of copies of the target sequence. Although the purpose of the PCR process is to amplify template DNA, a reverse transcription step allows the starting point to be RNA (*2–5*).

1.2. Scope of PCR Applications

PCR is widely used in molecular biology and genetic disease studies to identify new genes. Viral targets, such as HIV-1 and HCV, can be identified and quantified by PCR. Active gene products can be accurately quantitated using RNA-PCR. In such fields as anthropology and evolution, sequences of degraded ancient DNAs can be tracked after PCR amplification. With its exquisite sensitivity and high selectivity, PCR has been used in wartime human identification and validation in crime labs for mixed-sample forensic casework. In the realm of plant and animal breeding, PCR techniques are used to screen for traits and to evaluate living four-cell embryos. Environmental and food pathogens can be quickly identified and quantitated at high sensitivity in complex matrices with simple sample preparation techniques.

1.3. PCR Process (see Note 1)

The PCR process requires a repetitive series of the three fundamental steps that defines one PCR cycle: double-stranded DNA template denaturation, annealing of two oligonucleotide primers to the single-stranded template, and enzymatic extension of the primers to produce copies that can serve as templates in subsequent cycles. The target copies are double-stranded and bounded by annealing sites of the incorporated primers. The 3' end of the primer should complement the target exactly, but the 5' end can actually be a noncomplementary tail with restriction enzyme and promoter sites that will also be incorporated. As the cycles proceed, both the original template and the amplified targets serve as substrates for the denaturation, primer annealing, and primer extension processes. Since every cycle theoretically doubles the amount of target copies, a geometric amplification occurs. Given an efficiency factor for each cycle, the amount of amplified target Y produced from an input copy number X after n cycles is

$$Y = X(1 = \text{efficiency})^n \quad (1)$$

With this amplification power, 25 cycles could produce 33 million copies. Every extra 10 cycles produces 1024 more copies. Unfortunately, the process becomes self-limiting and amplification factors are generally between 10^5 - and 10^9 -fold. Excess primers and dNTPs help drive the reaction that commonly occurs in 10 mM Tris-HCl buffer, pH 8.3 (at room temperature). In addition, 50 mM KCl is present to provide proper ionic strength and magnesium ion is required as an enzyme cofactor (6).

The denaturation step occurs rapidly at 94–96°C. Primer annealing depends on the T_m , or melting temperature, of the primer:template hybrids. Generally, one uses a predictive software program to compute the T_m s based on the primer's sequence, their matched concentrations, and the overall salt concentration. The best annealing temperature is determined by optimization. Extension occurs at 72°C for most templates. PCR can also easily occur with a two-temperature cycle consisting of denaturation and annealing/extension.

1.4. Carryover Prevention

PCR has the potential sensitivity to amplify single molecules, so PCR products that can serve as templates for subsequent reactions must be kept isolated after amplification. Even tiny aerosols can contain thousand of copies of carried-over target molecules that can convert a true negative into a false positive. In general, dedicated pipetors, pipet tips with filters, and separate work areas should be considered and/or designated for RNA or DNA sample preparation, reaction mixture assemblage, the PCR process, and the reaction product analysis. As with any high sensitivity technique, the judicious and frequent use of positive and negative controls is required for each amplification (7–9). Through the use of dUTP instead of dTTP for all PCR samples, it is possible to design an internal biochemical mechanism to attack the PCR carryover problem. These PCR products are dU-containing and can be cloned, sequenced, and analyzed as usual. Pretreatment of each PCR reaction with uracil-N glycosylase (UNG), which catalyzes the removal of uracil from single- and double-

stranded DNA, will destroy any PCR product carried over from previous reactions, leaving the native T-containing sample ready for amplification (10).

1.5. Hot Start

PCR is conceptualized as a process that begins when thermal cycling ensues. The annealing temperature sets the specificity of the reaction, assuring that the primary primer binding events are the ones specific for the target in question. In preparing a PCR amplification on ice or at room temperature, however, the reactants are all present for nonspecific primer annealing to any single-stranded DNA present. Because DNA polymerases have some residual activity even at lower temperatures, it is possible to extend these misprimed hybrids and begin the PCR process at the wrong sites. To prevent this mispriming/misextension, a number of “Hot Start” strategies have been developed. In Hot Start PCR, a key reaction component essential for polymerase activity is withheld or separated from the reaction mixture until an elevated temperature is reached (11,12).

To separate an essential component from the reaction mixture in order to delay amplification, the following techniques can be utilized:

1.5.1. Manual Hot Start

In Manual Hot Start, a key reaction component such as *Taq* DNA polymerase or $MgCl_2$ is withheld from the original amplification mixture and added to the reaction when the temperature within the tube exceeds the optimal annealing temperature, i.e., above 65°–70°C.

1.5.2. Physical Barrier Hot Start, i.e., AmpliWax® PCR Gems from Applied Biosystems

In AmpliWax PCR gem-facilitated Hot Start, reaction components are divided into two mixes, and separated by a solid wax layer within the reaction tube (11). During the initial denaturation step, the wax layer melts at 75°–80°C allowing the two reaction mixes to combine through thermal convection.

1.5.3. Monoclonal Antibodies to DNA Polymerases Hot Start, i.e., PfuTurbo® Hotstart DNA polymerase from Stratagene or TaqStart from Clontech

In polymerase-antibody Hot Start, a PCR preincubation step is added, during which a heat-sensitive antibody attaches to the DNA polymerase [*Taq* or recombinant *Thermus thermophilus* (*rTth*)] inactivating the enzyme within the reaction mixture. As the temperature within the tubes rises, the antibody detaches and is inactivated, setting the polymerase free to begin polymerization.

1.5.4. Modified DNA Polymerases for Hot Start, i.e., AmpliTaq Gold® from Applied Biosystems

With AmpliTaq Gold, Hot Start is achieved with a chemically modified *Taq* DNA polymerase. The modification blocks the polymerase activity until it is reversed by a high temperature, pre-PCR incubation (e.g., 95°C for >10 min). The pre-PCR incuba-

tion links directly to the denaturation step of the first PCR cycle. So, the reaction mixture never sees active polymerase below the optimal primer annealing temperature. If the pre-PCR incubation is omitted, the modification is reversed during the PCR cycling, and polymerase activity increases slowly. In addition to a Hot Start, this provides a time release effect, where polymerase activity builds as the DNA substrate accumulates (12).

1.5.5. Oligonucleotide Inhibitors of DNA Polymerases for Hot Start

In polymerase-inhibitor Hot Start, DNA polymerase-binding oligonucleotides are added to the PCR amplification, keeping the enzyme inactive at ambient temperatures. Increasing the temperature dissociates the inhibitor from the enzyme, setting it free to begin polymerization. Moreover, inhibition is thermally reversible (13–16).

1.6. PCR Achievements

PCR has been used to speed the human genome discovery and for early detection of viral diseases. Single sperm cells to measure crossover frequencies can be analyzed and four-cell cow embryos can be typed. Trace forensic evidence of even mixed samples can be analyzed. Single-copy amplification requires some care, but is feasible for both DNA and RNA. True needles in haystacks can be found simply by amplifying the needles. PCR facilitates cloning of DNA sequences and forms a natural basis for cycle sequencing by the Sanger method (17). In addition to generating large amounts of template for cycle sequencing, PCR has been used to map chromosomes and to analyze both large and small changes in chromosome structure.

1.7. PCR Enzymes

The choice of the DNA polymerase is determined by the aims of the experiment. There are a variety of commercially available enzymes to choose from that differ in their thermal stability, processivity, and fidelity as depicted in **Table 1**. The most commonly used and most extensively studied enzyme is *Taq* DNA polymerase, e.g., AmpliTaq® DNA polymerase.

1.7.1. AmpliTaq DNA Polymerase

AmpliTaq DNA Polymerase (Applied Biosystems, Foster City, CA) is a highly characterized recombinant enzyme for PCR. It is produced in *Escherichia coli* (*E. coli*) from the *Taq* DNA polymerase gene, thereby assuring high purity. It is commonly supplied and used as a 5 U/μL solution in buffered 50% (v/v) glycerol (18).

1. *Biophysical Properties*. The enzyme is a 94-kDa protein with a 5'-3' polymerization activity that is most efficient in the 70°–80°C range. This enzyme is very thermostable, with a half-life at 95°C of 35–40 min. In terms of thermal cycling, the half-life is approx 100 cycles. PCR products amplified using AmpliTaq DNA polymerase will often have single base overhangs on the 3' ends of each polymerized strand, and this artifact can be successfully exploited for use with T/A cloning vectors.
2. *Biochemical Reactions*. DNA Polymerase requires magnesium ion as a cofactor and catalyzes the extension reaction of a primed template at 72°C. The four dNTPs (consisting of

Table 1
Some Commercially Available DNA Polymerases and Associated Properties (18)

DNA Polymerase	Source	Commercial Name	95°C Half-life	Exonuclease Activity		Extension Rate (nucleosides/s)
				5'–3'	3'–5'	
<i>Taq</i>	<i>Thermus aquaticus</i>	AmpliTaq	40 min	+	–	75
<i>Pwo</i>	<i>Pyrococcus woesei</i>		?	–	+	?
<i>Pfu</i>	<i>Pyrococcus furiosus</i>		>120 min	–	+	60
<i>rTth</i>	<i>Thermus thermophilus</i>		20 min	+	–	60
<i>Tfl</i>	<i>Thermus flavus</i>		?	–	–	?
<i>Tli</i>	<i>Thermus litoris</i>	Vent	400 min	–	+	67
<i>Tma</i>	<i>Thermotoga maritima</i>		>50 min	–	+	?

dATP, dCTP, dGTP, and dTTP or dUTP) are used according to the basepairing rule to extend the primer and thereby to copy the target sequence. Modified nucleotides (ddNTPs, biotin-11-dNTP, dUTP, deaza-dGTP, and fluorescently labeled dNTPs) can be incorporated into PCR products.

3. *Associated Activities.* AmpliTaq DNA Polymerase has a fork-like structure-dependent, polymerization enhanced, 5'–3' nuclease activity. This activity allows the polymerase to degrade downstream primers and indicates that circular targets should be linearized before amplification. In addition, this nuclease activity has been employed in a fluorescent signal-generating technique for PCR quantitation (19). AmpliTaq DNA Polymerase does not have an inherent 3'–5' exonuclease or proofreading activity, but produces amplicons of sufficient high fidelity for most applications.

1.7.2. AmpliTaq Gold

AmpliTaq Gold (Applied Biosystems, Foster City, CA) is chemically modified AmpliTaq DNA polymerase. The reversible modification keeps the enzyme inactive at room temperature. High temperature and low pH promote the reversal, restoring the enzyme activity. These conditions occur in a Tris-buffered PCR at 92°–95°C (Tris-Cl formulated to pH 8.3 at 25°C drops below pH 7.0 above 90°C). AmpliTaq Gold is formulated to perform the same as 5 U/μL AmpliTaq DNA polymerase. Therefore, a hot start can be added to most PCRs optimized with AmpliTaq DNA polymerase by substituting AmpliTaq Gold and adding a 10-min, 95°C, pre-PCR, activation step. The same results can be achieved without the pre-PCR activation step by adding an additional 10 or more PCR cycles. Under these conditions, the enzyme is activated incrementally during the PCR denaturation steps.

1.8. Primers

PCR Primers are short oligodeoxyribonucleotides, or oligomers, that are designed to complement the end sequences of the PCR target amplicon. These synthetic DNAs are usually 15–25 nucleotides long and have approx 50–60% G + C content. Because each of the two PCR primers is complementary to a different individual strand of the target sequence duplex, the primer sequences are not related to each other. In fact, special care must be taken to assure that the primer sequences do not form duplex structures with each other or hairpin loops within themselves. The 3' end of the primer must match the target in order for polymerization to be efficient, and allele-specific PCR strategies take advantage of this fact. In screening for potential sequences and their homology, primer design software packages such as Oligo[®] (National Biosciences, Plymouth, NC) and online search sites such as BLAST (NCBI, www.ncbi.nlm.nih.gov/BLAST/), can be utilized. To screen for mutants, a primer complementary to the mutant sequence is used and results in PCR positives, whereas the same primer will be a mismatch for the wild type and does not amplify. The 5' end of the primer may have sequences that are not complementary to the target and that may contain restriction sites or promoter sites that are also incorporated into the PCR product. Primers with degenerate nucleotide positions every third base may be synthesized in order to allow for amplification of targets where only the amino acid sequence is known. In this case, early PCR cycles are performed with low, less stringent annealing temperatures, followed by later cycles with high, more stringent annealing temperatures.

A PCR primer can also be a homopolymer, such as oligo (dT)₁₆, which is often used to prime the RNA PCR process. In a technique called RAPDS (randomly amplified polymorphic DNAs), single primers as short as decamers with random sequences are used to prime on both strands, producing a diverse array of PCR products that form a fingerprint of a genome (20). Often, logically designed primers are less successful in PCR than expected, and it is usually advisable to try optimization techniques for a practical period of time before trying new primers frequently designed near the original sites.

1.8.1. T_m Predictions

DNA duplexes, such as primer-template complexes, have a stability that depends on the sequence of the duplex, the concentrations of the two components, and the salt concentration of the buffer. Heat can be used to disrupt this duplex. The temperature at which half the molecules are single-stranded and half are double-stranded is called the T_m of the complex. Because of the greater number of intermolecular hydrogen bonds, higher G+C content DNA has a higher T_m than lower G+C content DNA. Often, G + C content alone is used to predict the T_m of the DNA duplex, however, DNA duplexes with the same G + C content may have different T_m values. A simple, generic formula for calculating the T_m is: $T_m = 4(G+C) + 2(A+T) ^\circ\text{C}$. A variety of software packages are available to perform more accurate T_m predictions using sequence information (nearest neighbor analysis) and to assure optimal primer design, e.g., Oligo, BLAST, or Melt (Mt. Sinai School of Medicine, New York, NY).

Because the specificity of the PCR process depends on successful primer binding events at each amplicon end, the annealing temperature is selected based on the consensus of the melting temperatures (within 2–4°C) of the two primers. Usually, the annealing temperature is chosen a few degrees below the consensus annealing temperatures of the primers (1). Different strategies are possible, but lower annealing temperatures should be tried first to assess the success of amplification to find the stringency required for best product specificity.

1.9. PCR Samples

1.9.1. Types

The PCR sample type may be single- or double-stranded DNA of any origin—animal, bacterial, plant, or viral. RNA molecules, including total RNA, poly (A+) RNA, viral RNA, tRNA, or rRNA, can serve as templates for amplification after conversion to so-called complementary DNA (cDNA) by the enzyme reverse transcriptase (either MuLV or recombinant, *rTth* DNA polymerase) (21,22).

1.9.2. Amount

The amount of starting material required for PCR can be as little as a single molecule, compared to the millions of molecules needed for standard cloning or molecular biological analysis. As a basis, up to nanogram amounts of DNA cloned template, up to microgram amounts of genomic DNA, or up to 10⁵ DNA target molecules are best for initial PCR testing.

1.9.3. Purity

Overall, the purity of the DNA sample to be subjected to PCR amplification need not be high. A single cell, a crude cell lysate, or even a small sample of degraded DNA template is usually adequate for successful amplification. The fundamental requirements of sample purity must be that the target contains at least one intact DNA strand encompassing the amplified region and that the impurities associated with the target be adequately dilute so as to not inhibit enzyme activity. However, for some amplifications, such as long PCR, it may be necessary to consider the quality and quantity of the DNA sample (23,24). For example,

1. When more template molecules are available, there is less occurrences of false positives caused by either cross-contamination between samples or “carryover” contamination from previous PCR amplifications;
2. When the PCR amplifications lacks specificity or efficiency, or when the target sequences are limited, there is a greater chance of inadequate product yield; and
3. When the fraction of starting DNA available to PCR is uncertain, it is increasingly difficult to determine the target DNA content (25).

1.10. Other Parameters for Successful PCR

1.10.1. Metal Ion Cofactors

Magnesium chloride is an essential cofactor for the DNA polymerase used in PCR, and its concentration must be optimized for every primer:template system. Many com-

ponents of the reaction bind magnesium ion, including primers, template, PCR products and dNTPs. The main 1:1 binding agent for magnesium ion is the high concentration of dNTPs in the reaction. Because it is necessary for free magnesium ion to serve as an enzyme cofactor in PCR, the total magnesium ion concentration must exceed the total dNTP concentration. Typically, to start the optimization process, 1.5 mM magnesium chloride is added to PCR in the presence of 0.8 mM total dNTPs. This leaves about 0.7 mM free magnesium for the DNA polymerase. In general, magnesium ion should be varied in a concentration series from 1.5–4.0 mM in 0.5 mM steps (1,25).

1.10.2. Substrates and Substrate Analogs

DNA polymerases incorporate dNTPs very efficiently, but can also incorporate modified substrates, when they are used as supplemental components in PCR. Digoxigenin-dUTP, biotin-11-dUTP, dUTP, c7deaza-dGTP, and fluorescently labeled dNTPs all serve as substrates for DNA polymerases. For conventional PCR, the concentration of dNTPs remains balanced in equimolar ratios, e.g., 200 μ M each dNTP (1). However, deviations (from these standard recommendations) may be beneficial in certain applications. For example, when random mutagenesis of a specific target is desired, unbalanced dNTP concentrations promote a higher degree of misincorporations by the DNA polymerase.

1.10.3. Buffers and Salts

The optimal PCR buffer concentration, salt concentration, and pH depend on the DNA polymerase in use. The PCR buffer for *Taq* DNA polymerase consists of 50 mM KCl and 10 mM Tris-HCl, pH 8.3, at room temperature. This buffer provides the ionic strength and buffering capacity needed during the reaction. It is important to note that the salt concentration affects the T_m of the primer:template duplex, and hence the annealing temperature.

1.10.4. Cosolvents

A variety of PCR cosolvents have been utilized to increase the yield, efficacy, and specificity of PCR amplifications. Although these cosolvents are advantageous in some amplifications, it is impossible to predict which additive will be useful for each primer:template duplex and therefore the cosolvent must be empirically tested for each combination. Some of the more popular cosolvents currently in use are listed in **Table 2** along with the recommended testing ranges (26).

1.10.5. Thermal Cycling Considerations

1.10.5.1. PCR VESSELS

PCR must be performed in vessels that are compatible with low amounts of enzyme and nucleic acids and that have good thermal transfer characteristics. Typically, polypropylene is used for PCR vessels and conventional, thick-walled microcentrifuge tubes are chosen for many thermal cycler systems. PCR is most often performed at a 10–100 μ L reaction scale and requires the prevention of the evaporation/condensation processes in the closed reaction tube during thermal cycling. A mineral oil overlay or

Table 2
PCR Cosolvents

Cosolvent	Recommended Testing Ranges	Comments
Betaine	Final concentration: 1.0–1.7 M	Reduces the formation of secondary structure caused by GC-rich regions (27)
Bovine serum albumin (BSA)	10–100 µg/mL	A nonspecific enzyme stabilizer which also binds certain DNA inhibitors (28)
7-deaza-2'-deoxyguanosine (dC ⁷ GTP)	Ratio 3:1 dC ⁷ GTP:dGTP	Facilitates amplification of templates with stable secondary structures when used in place of dGTP (1)
DMSO	2–10%	10% reduces <i>Taq</i> activity by 50%. Thought to reduce secondary structure. Useful for GC rich templates. Presumed to lower the <i>T_m</i> of the target nucleic acids.
Formamide	1–5%	Improve the specificity of PCR at lower denaturation temperatures (21,29)
Glycerol	1–10%	Improves the thermal stability of DNA Polymerases. Improves the amplification of high GC templates (30)
Nonionic detergents: Triton X-100, Tween 20 NP40	0.1–1%	Stabilizes <i>Taq</i> DNA polymerase. May suppress the formulation of secondary structure. May increase yield but may also increase non-specific amplification.
T4 gene 32 protein	20–150 µg/mL	Enhance PCR product yield and relieve inhibition (31)
Tetramethylammonium chloride (TMAC)	Final concentration: 15–100 mM	To eliminate nonspecific priming. Also used to reduce potential DNA-RNA mismatch. Improves the stringency of hybridization reactions.
TMA oxalate	2 mM	Decreases the formation of nonspecific DNA fragments and increases PCR product yield (32)

wax layer serves this purpose. More recently, 0.2-mL thin-walled vessels have been optimized for the PCR process and oil-free thermal cyclers have been designed that use a heated cover over the tubes held within the sample block.

1.10.5.2. TEMPERATURE AND TIME OPTIMIZATION

It is essential that the reaction mixtures reach the denaturation, annealing, and extension temperatures in each thermal cycle. If insufficient hold time is specified at any temperature, the temperature of the sample will not be equilibrated with that of the sample block. Some thermal cycler designs time the hold interval based on the block temperature, whereas others base the hold time on predicted sample temperature.

If a conventional thick-walled tube used in a cycler controlled by block temperature, a 60-s hold time is sufficient for equilibration. Extra time may be recommended at the (72°C) extension step for longer PCR products (23). Using a thin-walled 0.2-mL tube in a cycler controlled by predicted sample temperature, only 15 s is required. To use existing protocols or to development protocols for use at multiple laboratories, it is very important to choose hold times according to the cycler design and tube wall thickness.

1.10.6. PCR Amplification Cycles

The number of PCR amplification cycles should be optimized with respect to the starting concentration of the target DNA. Innis and Gelfand (1) recommend from 40–45 cycles to amplify 50 target molecules, and 25–30 cycles to amplify 3×10^5 molecules to the same concentration. This nonproportionality is caused by a so-called plateau effect, in which a decrease in the exponential rate of product accumulation occurs in late stages of a PCR. This may be caused by degradation of reactants (dNTPs, enzyme); reactant depletion (primers, dNTPs); end-product inhibition (pyrophosphate formation); competition for reactants by non-specific products; or competition for primer binding by reannealing of concentrated (10 nM) product. It is usually advisable to run the minimum number of cycles needed to see the desired specific product, because unwanted nonspecific products will interfere if the number of cycles is excessive.

1.10.7. Enzyme/Target

In a standard aliquot of *Taq* DNA polymerase used for a 100- μ L reaction, there are about 10^{10} molecules. Each PCR sample should be evaluated for the number of target copies it contains or may contain. For example, 1 ng of lambda DNA contains 1.8×10^7 copies. For low-input copy number PCR, the enzyme becomes limiting and it may be necessary to give the extension process incrementally more time. Thermal cyclers can reliably perform this automatic segment extension procedure in order to maximize PCR yield (1,25).

1.10.8. Hot Start

All of the above optimizations also apply to a PCR that is designed, from the beginning, with a hot start method. Often, a hot start can be incorporated successfully into a

previously optimized PCR without changing the reaction conditions. However, it usually pays to reoptimize after adding a hot start. Optimization is often a balance between producing as much product as possible and overproducing nonspecific, background amplifications. Because hot start greatly reduces background amplifications, the upper restraints are raised on conditions such as enzyme concentration, cycle number, and metal ion cofactor concentration. Sensitive PCRs that have been highly tuned without a hot start may fail when a hot start is added. This can be caused by slight delays in early cycles caused by mixing or enzyme activation. The PCR usually can be restored, often with substantial increase in specific product, by detuning—that is, simply increasing limiting parameters or reagents. In addition, there are optimizations specific to each hot start method. Mixing or enzyme activation can be affected by PCR volume, buffer composition and pH, cosolvents, cycling conditions, and so on. The specific product's literature, often a product insert, should be consulted for information on these considerations.

2. Materials

The protocol described later illustrates the basic principles and techniques of PCR and can be modified to suit other particular applications. The example chosen uses HIV Primer pair, SK145 and SK431 (Applied Biosystems), in conjunction with Applied Biosystem's GeneAmp 10X PCR Buffer II, MgCl₂ Solution, GeneAmp dNTPs, and PCR Carry-Over Prevention Kit, to amplify a 142-bp DNA fragment from the conserved gag region of HIV-1 using the AmpliTaq Gold Hot Start process.

1. 10X PCR Buffer II: 100 mM Tris-HCl, 500 mM KCl, pH 8.3 at room temperature.
2. 25 mM MgCl₂ solution.
3. dNTPs: 10 mM stocks of each of dATP, dCTP, dGTP; 20 mM stock of dUTP; all neutralized to pH 7.0 with NaOH.
4. Primer 1: SK145. 25 mM in 10 mM Tris-HCl, pH 8.3 at room temperature. Sequence: 5'-AGTGGGGGGACATCAAGCAGCCATGCAAAT-3'.
5. Primer 2: SK431. 25 mM in 10 mM Tris-HCl, pH 8.3 at room temperature. Sequence: 5'-TGCTATGTCAGTCCCCTTGTTCTCT-3'.
6. AmpErase[®] UNG: Uracil-N-glycosylase, 1.0 U/mL pH 8.3 at room temperature in 150 mM NaCl, 30 mM Tris-HCl, pH 7.5 at room temperature, 10 mM ethylenediaminetetraacetic acid (EDTA), 1.0 mM dithiothreitol (DTT), 0.05% Tween-20, 5% (v/v) glycerol.
7. HIV-1 Positive Control DNA: 10³ copies/mL in 10 mg/mL human placental DNA.
8. AmpliTaq Gold: 5 U/mL.
9. 0.5 mL microcentrifuge tubes (Applied Biosystems GeneAmp PCR microcentrifuge tubes).
10. Thermal Cycler (PE Applied Biosystems GeneAmp PCR System).

3. Methods

3.1. Hot Start Process

In the AmpliTaq Gold Hot Start process (33), a master mix is prepared at room temperature, aliquoted into individual tubes, and thermal cycled.

1. Assemble the reagent mix as shown here:

Reagent	Volume (1X mix, μL)	Final Concentration (per 100 μL Volume)
Sterile Water		N/A
10X PCR Buffer II	10.0	1X
25 mM MgCl ₂	10.0	2.5 mM
10 mM dATP	2.0	200 μM
10 mM dCTP	2.0	200 μM
10 mM dGTP	2.0	200 μM
20 mM dUTP	2.0	400 μM
25 μM Primer 1 (SK145)	2.0	50 pmol
25 μM Primer 2 (SK431)	2.0	50 pmol
1 U/ μL AmpErase UNG	1.0	1.0 U/reaction
10 ³ copies/ μL (+Control)	0.1–10.0	10 ² –10 ⁴ copies/ μL
5 U/ μL AmpliTaq Gold 0.5		2.5 U/reaction
Total Volume	100 μL	

2. Add 100 μL of the above reagent mix to the bottom of each GeneAmp PCR reaction tube. Avoid splashing liquid onto the tube walls. If any liquid is present on the tube walls, spin the tube briefly in a microcentrifuge.
3. Amplify the PCR amplifications within a programmed thermal cycler. For the Perkin Elmer DNA Thermal Cycler 9600, program and run the following linked files:
 - a. CYCL File: 95°C for 9 min, 1 cycle; link to file (b).
 - b. CYCL File: 94°C for 30 s, 60°C for 1 min, 43 cycles; link to file (c).
 - c. CYCL File: 60°C for 10 min; 1 cycle; link to file (d).
 - d. HOLD File: 10°C hold.

3.2. Analysis of PCR Products (see Note 2)

3.2.1. Agarose Gel Electrophoresis

PCR products can be easily and quickly analyzed and resolved using a 3% NuSieve GTG agarose (FMC Bioproducts, Rockland, ME) and 1% Seakem GTG agarose (FMC Bioproducts) gel run in either TBE (89 mM Tris-borate, 2 mM EDTA) or TAE (40 mM Tris-acetate, 2 mM EDTA, pH approx 8.5). The resolved DNA bands are detected by staining the gels with either approx 0.5 $\mu\text{g}/\text{mL}$ of ethidium bromide, followed by destaining with water or SYBR[®] Green 1 (Molecular Probes Inc., Eugene, OR) and finally photographed under UV illumination. Use a 123-basepair (bp) or 1-kilobasepair (kbp) ladder as a convenient marker for size estimates of the products (34).

3.2.2. Other Analytical Methods

A variety of other detection methods are available for PCR product analysis, such as ethidium bromide-stained 8–10% polyacrylamide gels run in TBE buffer, Southern gels or dot/blots, subcloning and direct sequencing, HPLC analysis, and the use of 96-well microplates, to name a few. The reverse dot-blot method combines PCR amplification with nonradioactive detection (35).

The introduction of fluorescent dyes to PCR, together with a suitable instrument for real-time, online quantification of PCR products during amplification has led to the development of kinetic PCR or quantitative PCR. Quantitative PCR (QPC) measures

PCR product accumulation during the exponential phase of the reaction and before amplification becomes vulnerable, i.e., when reagents become limited. The ABI Prism 7700 (Applied Biosystems) and the LightCycler (Roche Molecular Biochemicals, Mannheim, Germany) are integrated fluorescent detection devices that allow fluorescence monitoring either continuously or once per cycle. These instruments can also characterize PCR products by their melting characteristics, e.g., to discriminate single-base mutations from a wild-type sequence. The recently designed Mx4000™ Multiplex Quantitative PCR System (Stratagene, La Jolla, CA) can generate and analyze data for multiple fluorescent real-time QPCR assays.

4. Notes

1. Even though the PCR process has greatly enhanced scientific studies, a variety of problems with the process, easily revealed by ethidium-bromides-stained agarose gel electrophoresis, can and may need to be considered when encountered. For example, unexpected molecular weight size bands (nonspecific banding) or smears can be produced. These unexpected products accumulate from the enzymatic extension of primers that annealed to nonspecific target sites. Second, primer-dimer (approx 40–60 bp in length, the sum of the two primers) can be produced. Primer-dimer can arise during PCR amplification when the DNA template is left out of the reaction, too many amplification cycles are used, or the primers are designed with partial complementarity at their 3' ends. Note, an increase in primer-dimer formation will decrease the production of the desired product. Third, *Taq* DNA polymerase, which lacks the 3'-5' exonuclease “proofreading” activity, will occasionally incorporate the wrong base during PCR extension. The consequences of *Taq* misincorporations usually have little effect, but should be considered during PCR cloning and subsequent cycle sequencing.
2. PCR amplification for user-selected templates and primers are considered “failures” when 1) no product bands are observed; 2) the PCR product band is multibanded; or 3) the PCR product is smeared. These “failures” can be investigated and turned into successful PCR by manipulation of a number of variables, such as enzyme and salt concentrations, denaturation and anneal/extend times and temperatures, primer design, and hot start procedures (35).

When no desired PCR product band is observed, initially verify the enzyme addition and/or concentration by titrating the enzyme concentration. Second, the magnesium ion concentration is also critical, so care should be taken not to lower the magnesium ion molarity on addition of reagents (i.e., buffers containing EDTA will chelate out the magnesium ion). The denaturation and anneal/extend times and temperatures may be too high or too low, causing failures, and can be varied to increase reaction specificity. Finally, the chemical integrity of the primers should be considered. In cases where the PCR product band is multibanded, consider raising the anneal temperature in increments of 2°C and/or review the primer design and composition.

If a smear of the PCR product band is seen on an ethidium-bromide-stained agarose gel, consider the following options initially, individually, or in combination: decreasing the enzyme concentration, lowering the magnesium ion concentration, lengthening and/or raising the denaturation time and temperature, shortening the extension time, reducing the overall cycle number, and decreasing the possibility of carryover contamination. Finally, in PCR amplifications where the PCR product band was initially observed, and

on later trials a partial or complete loss of the product bands is observed, consider testing new aliquots of reagents and decreasing the possibility of carryover contamination.

For PCR amplifications using a modified DNA polymerase such as AmpliTaq Gold, poor product amplification can occur owing to inadequate activation of the Hot Start polymerase. Incubation time, temperature, and pH are critical for Hot Start polymerase activation. Contaminants added with the target, whether remnants from the sample's source or artifacts of the sample's preparation, can affect the PCR pH. Contaminants may also directly inhibit the polymerase. Hot Start polymerase activation begins during the pre-PCR activation step and continues through the PCR cycles' denaturation steps. The temperature and duration of these steps and the total number of PCR cycles should be optimized. Additional PCR cycles may increase specific product yield without increasing background in a Hot Start PCR. Raising the temperature above 95°C for any PCR step may irreversibly denature the polymerase.

References

- Innes, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J., eds. (1990) *PCR Protocols, A Guide to Methods and Application*, Academic, San Diego, CA.
- Mullis, K. B. and Faloona F. A. (1987) Specific synthesis of DNA in vitro via a polymerase chain reaction. *Meth. Enzymol.* **155**, 335–350.
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.
- Saiki, R. K., Scharf, S. J., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354.
- Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1087.
- Wang, A. M., Doyle, M. V., and Mark, D. F. (1989) Quantitation of mRNA by the polymerase chain reaction. *Proc. Nat. Acad. Sci. USA* **86**, 9717–9721. *Nature* 1.
- Kwok, S. and Higuchi, R. (1989) Avoiding false positives with PCR. *Nature* **339**, 237, 238.
- Orrego, C. (1990) Organizing a laboratory for PCR work, in *PCR Protocols. A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.), Academic, San Diego, CA, pp. 447–454.
- Kitchin, P. A., Szotyori, Z., Fromholz, C., and Almond, N. (1990) Avoiding false positives. *Nature* **344**, 201.
- Longo, N., Berninger, N.S., and Hartley, J. L. (1990) Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. *Gene* **93**, 125–128.
- Chou, Q., Russell, M., Birch D. E., Raymond, J., and Bloch, W. (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* **20**, 1717–1723.
- Birch, D. E., Kolmodin, L., Laird, W. J., McKinney, N., Wong, J., Young, K. K. Y., et al. (1996) Simplified Hot Start PCR. *Nature* **381**, 445,446.
- Ailenberg, M. and Silverman, M. (2000) Controlled hot start and improved specificity in carrying out PCR utilizing touch-up and loop incorporated primers (TULIPS). *Biotechniques* **29**, 1018–1020, 1022–1024.
- Kaboev, O. K., Luchkina, L. A., Tret'iakov, A. N., and Bahrmand, A. R. (2000) PCR hot start using primers with the structure of molecular beacons (hairpin-like structure). *Nucl. Acids Res.* **28**, E94.

15. Kainz, P., Schmiedlechner, A., and Strack, H. B. (2000) Specificity-enhanced hot-start PCR: addition of double-stranded DNA fragments adapted to the annealing temperature. *Biotechniques* **28**, 278–82.
16. Dang, C. and Jayasena, S. (1996) Oligonucleotide inhibitors of *Taq* DNA polymerase facilitate detection of low copy number targets by PCR. *J. Molec. Biol.* **264**, 268–278.
17. Innis, M. A., Myambo, K. B., Gelfand, D. H., and Brow, M. A. D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Nat. Acad. Sci. USA* **85**, 9436–9440.
18. Abramson, R. D. (1995) Thermostable DNA polymerases, in PCR Strategies (Innes, M. A., Gelfand, D. H., and Sninsky, J. J., eds.), Academic, San Diego, CA, pp. 39–57.
19. Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Nat. Acad. Sci. USA* **88**, 7276–7280.
20. Sobral, B. W. S. and Honeycutt, R. J. (1993) High output genetic mapping of polyploids using PCR generated markers. *Theor. Appl. Genet.* **86**, 105–112.
21. Myers, T. W. and Gelfand, D. H. (1991) Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* **30**, 7661–7666.
22. Myers, T. W. and Sigua, C. L. (1995) Amplification of RNA, in PCR Strategies (Innes, M. A., Gelfand, D. H., and Sninsky, J. J., eds.), Academic, San Diego, CA, pp. 58–58.
23. Cheng, S., Fockler, C., Barnes, W. M., and Higuchi, R. (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Nat Acad Sci USA* **91**, 5695–5699.
24. Cheng, S., Chen, Y., Monforte, J. A., Higuchi, R., and Van Houten, B. (1995) Template integrity is essential for PCR amplification of 20– to 30–kb sequences from genomic DNA. *PCR Meth. Amplificat.* **4**, 294–298.
25. Erlich, H. A., ed. (1989) *PCR Technology, Principles and Application for DNA Amplification*. Stockton, New York.
26. Landre, P. A., Gelfand, D. H., and Watson, R. H. (1995) The use of cosolvents to enhance amplification by the polymerase chain reaction, in PCR Strategies (Innes, M. A., Gelfand, D. H., and Sninsky, J. J., eds.), Academic, San Diego, CA, pp. 3–16.
27. Henke, W., Herdel, K., Jung, K. Schnorr, D., and Loening, S. (1997) Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucl. Acids Res.* **25** (19), 3957–3958.
28. Paabo, S., Gifford, J. A., and Wilson, A. C. (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucl. Acids Res.* **16**, 9775–9787.
29. Sarker G, Kapeiner, S., and Sommer, S. S. (1990) Formamide can drastically increase the specificity of PCR. *Nucl. Acid Res.* **18**, 7465.
30. Smith, K. T., Long, C. M., Bowman, B. and Manos, M. M. (1990) Using cosolvents to enhance PCR amplification. *Amplifications* **9/90** (5), 16,17.
31. Kreader, C. (1996) Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Appl. Environ. Microbiol.* **62**, 1102–1106.
32. Kovarova, M. and Draber, P. (2000) New specificity and yield enhancer of polymerase chain reactions. *Nucl. Acids Res.* **28**, E70.
33. AmpliTaq Gold. Package Insert. **BIO-142**, 54,670–3/96. Applied Biosystems, Foster City, CA.
34. Sambrook, J., Fritsch, E. F., and Maniatis, T. eds (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 6.20, 6.21, B.23, B.24.

35. Saiki, R. K., Walsh, P. S., Levenson, C. H., and Erlich, H. A. (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc. Nat. Acad. Sci. USA* **86**, 6230–6234.
36. Kolmodin, L., Cheng, S., and Akers, J. (1995) GeneAmp XL PCR Kit. Amplifications: A Forum for PCR Users (The Perkin-Elmer Corporation) **13**, 1–5.

Computer Programs for PCR Primer Design and Analysis

Bing-Yuan Chen, Harry W. Janes, and Steve Chen

1. Introduction

1.1. Core Parameters in Primer Design

1.1.1. T_m , Primer Length, and GC Content (GC %)

Heat will separate or “melt” double-stranded DNA into single-stranded DNA by disrupting its hydrogen bonds. T_m (melting temperature) is the temperature at which half the DNA strands are single-stranded and half are double-stranded. T_m characterizes the stability of the DNA hybrid formed between an oligonucleotide and its complementary strand and therefore is a core parameter in primer design. It is affected by primer length, primer sequence, salt concentration, primer concentration, and the presence of denaturants (such as formamide or DMSO).

All other conditions set, T_m is characteristic of the primer composition. Primer with higher G+C content (GC %) has a higher T_m because of more hydrogen bonds (three hydrogen bonds between G and C, but two between A and T). The T_m of a primer also increases with its length. A simple formula for calculation of the T_m (1,2) (see **Note 1**) is

$$T_m = 2 \times AT + 4 \times CG$$

where AT is the sum of A and T nucleotides, and CG is the sum of C and G nucleotides in the primer.

1.1.2. Primer Specificity

Primer specificity is another important parameter in PCR primer design. To amplify only the intended fragment, the primers should bind to the target sequence only but not somewhere else. In other words, the target sequence should occur only once in the template. Primer length not only affects the T_m , as discussed earlier, but also the uniqueness (specificity) of the sequence in the template (3). Suppose the DNA sequence is entirely random (which may not be true), the chance of finding an A, G, C,

or T in any given DNA sequence is one quarter ($1/4^1$), so a 16 base primer will statistically occur only once in every 4^{16} bases, or about 4 billion bases, which is about the size of the human genome. Therefore, the binding of a 16 base or longer primer with its target sequence is an extremely sequence-specific process. Of course, to be absolutely sure that the target sequence occurs only once, you would need to check the entire sequence of the template DNA, which is not possible in most cases. However, it is often useful to search the current DNA sequence databases to check if the chosen primer has gross homology with repetitive sequences or with other loci elsewhere in the genome. For genomic DNA amplification 17-mer or longer primers are routinely used.

1.1.3. Primer Sequence and Hairpin (Self-Complementarity) and Self-Dimer (Dimer Formation)

The hardest part in PCR primer design is to avoid primer complementarity, especially at the 3' ends. When part of a primer is complementary to another part of itself, the primer may fold in half and form a so-called hairpin structure, which is stabilized by the complementary base pairing. The hairpin structure is a problem for PCR because the primer is interacting with itself and is not available for the desired reaction. Furthermore, the primer molecule could be extended by DNA polymerase so that its sequence is changed and it is no longer capable of binding to the target site.

Similar to the hairpin structure, if not carefully designed, one primer molecule may hybridize to another primer molecule and acts as template for each other, resulting in primer-dimers. Primer-dimer formation causes the same problems to PCR reaction as the hairpin structure. It may also act as a competitor to amplification of the target DNA (4). Usually it is very hard and time-consuming to catch the hairpin structure or primer-dimer formation manually by a naked eye. However, they can be easily detected by primer analysis programs.

1.2. General Rules for PCR Primer Design

According to Innis and Gelfand (5) the rules for primer design is as follows:

1. Primers should be 17–28 bases in length;
2. Base composition should be 50–60% (G+C);
3. Primers should end (3') in a G or C, or CG or GC: this prevents “breathing” of ends and increases efficiency of priming;
4. Tms between 55–80°C are preferred;
5. Avoid primers with 3' complementarity (results in primer-dimers). 3'-ends of primers should not be complementary (i.e., basepair), as otherwise primer dimers will be synthesised preferentially to any other product;
6. Primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided;
7. Runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing), and should be avoided.

Because two different primers are needed for PCR reaction, primer-dimer formation between the two primers should also be checked and avoided if possible. It is desirable that primer $T_{m,s}$ should be similar (within 8°C or so). If they are too different, a suitable annealing temperature may be hard to find. At high annealing temperature,

the primer with the lower T_m may not work, whereas at low annealing temperature, amplification will be less efficient because the primer with the higher T_m will misprime.

In reality, primer selection is often empirical. It varies greatly from researcher to researcher in regard to the criteria they use.

1.3. Computer Programs for PCR Primer Design and Analysis

1.3.1. Computer Programs for Nondegenerate PCR Primer Design

For primer design, most researchers used to visually inspect target DNA sequence to find primer(s) with the characteristics they prefer, which are usually similar to the guidelines we mentioned earlier. As computers are widely used in molecular biology, a large number of computer programs have been specifically developed for nondegenerate primer selection, which makes the PCR primer design more efficient and reliable. Most sequencing analysis packages, such as Vector NTI (InforMax Inc.), usually contain a primer design module. In this chapter, we focus on free online (web) primer design programs (*see Note 2*). Selected computer programs for nondegenerate PCR primer design and their features are listed in **Table 1**.

From a computational point of view the design of nondegenerate PCR primers is relatively simple: find short substrings from DNA nucleotide string that meet certain criteria. Although the criteria vary between programs, the core parameters, such as the primer length, T_m , GC content, and self-complementarity, are shared by these programs.

1.3.2. Computer Programs for Degenerate PCR Primer Design

In the experiments to amplify the novel members of gene families or cognate sequences from different organisms by PCR, the exact sequence of the target gene is not known. We usually align all known sequences for this gene and find the most conserved regions, then design corresponding “degenerate” primers, which are a set of primers with nucleotide diversity at several positions in the sequence. Degeneracies obviously increase the chances of amplifying the target sequence but reduce the specificity of the primer(s) at the same time.

Designing degenerate primers has been considered more of an art than a science. There are much less computer programs for degenerate primer design (*see Table 2*) than for nondegenerate primer design.

1.3.3. Computer Programs for Primer Analysis

Even if you prefer to design primers by yourself, not by a computer program, it is advised that your primers should be analyzed by a computer program to determine T_m , possible hairpin structure, primer-dimers, and other properties before you place the order for them. **Table 3** lists two computer programs for this purpose.

2. Materials

1. Computer: A computer (PC or Macintosh) with high-speed internet access.
2. Programs: Web Browser, Netscape (5.0 or above) or Internet Explorer (4.0 or higher).
3. Input files for primer design: DNA sequence file DNA.txt (*see Table 4*) and protein sequence file Protein.txt (*see Table 5*) (*see Note 4*).

Computer Programs for PCR Primer Design and Analysis

Bing-Yuan Chen, Harry W. Janes, and Steve Chen

1. Introduction

1.1. Core Parameters in Primer Design

1.1.1. T_m , Primer Length, and GC Content (GC %)

Heat will separate or “melt” double-stranded DNA into single-stranded DNA by disrupting its hydrogen bonds. T_m (melting temperature) is the temperature at which half the DNA strands are single-stranded and half are double-stranded. T_m characterizes the stability of the DNA hybrid formed between an oligonucleotide and its complementary strand and therefore is a core parameter in primer design. It is affected by primer length, primer sequence, salt concentration, primer concentration, and the presence of denaturants (such as formamide or DMSO).

All other conditions set, T_m is characteristic of the primer composition. Primer with higher G+C content (GC %) has a higher T_m because of more hydrogen bonds (three hydrogen bonds between G and C, but two between A and T). The T_m of a primer also increases with its length. A simple formula for calculation of the T_m (1,2) (see **Note 1**) is

$$T_m = 2 \times AT + 4 \times CG$$

where AT is the sum of A and T nucleotides, and CG is the sum of C and G nucleotides in the primer.

1.1.2. Primer Specificity

Primer specificity is another important parameter in PCR primer design. To amplify only the intended fragment, the primers should bind to the target sequence only but not somewhere else. In other words, the target sequence should occur only once in the template. Primer length not only affects the T_m , as discussed earlier, but also the uniqueness (specificity) of the sequence in the template (3). Suppose the DNA sequence is entirely random (which may not be true), the chance of finding an A, G, C,

or T in any given DNA sequence is one quarter ($1/4^1$), so a 16 base primer will statistically occur only once in every 4^{16} bases, or about 4 billion bases, which is about the size of the human genome. Therefore, the binding of a 16 base or longer primer with its target sequence is an extremely sequence-specific process. Of course, to be absolutely sure that the target sequence occurs only once, you would need to check the entire sequence of the template DNA, which is not possible in most cases. However, it is often useful to search the current DNA sequence databases to check if the chosen primer has gross homology with repetitive sequences or with other loci elsewhere in the genome. For genomic DNA amplification 17-mer or longer primers are routinely used.

1.1.3. Primer Sequence and Hairpin (Self-Complementarity) and Self-Dimer (Dimer Formation)

The hardest part in PCR primer design is to avoid primer complementarity, especially at the 3' ends. When part of a primer is complementary to another part of itself, the primer may fold in half and form a so-called hairpin structure, which is stabilized by the complementary base pairing. The hairpin structure is a problem for PCR because the primer is interacting with itself and is not available for the desired reaction. Furthermore, the primer molecule could be extended by DNA polymerase so that its sequence is changed and it is no longer capable of binding to the target site.

Similar to the hairpin structure, if not carefully designed, one primer molecule may hybridize to another primer molecule and acts as template for each other, resulting in primer-dimers. Primer-dimer formation causes the same problems to PCR reaction as the hairpin structure. It may also act as a competitor to amplification of the target DNA (4). Usually it is very hard and time-consuming to catch the hairpin structure or primer-dimer formation manually by a naked eye. However, they can be easily detected by primer analysis programs.

1.2. General Rules for PCR Primer Design

According to Innis and Gelfand (5) the rules for primer design is as follows:

1. Primers should be 17–28 bases in length;
2. Base composition should be 50–60% (G+C);
3. Primers should end (3') in a G or C, or CG or GC: this prevents “breathing” of ends and increases efficiency of priming;
4. Tms between 55–80°C are preferred;
5. Avoid primers with 3' complementarity (results in primer-dimers). 3'-ends of primers should not be complementary (i.e., basepair), as otherwise primer dimers will be synthesised preferentially to any other product;
6. Primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided;
7. Runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing), and should be avoided.

Because two different primers are needed for PCR reaction, primer-dimer formation between the two primers should also be checked and avoided if possible. It is desirable that primer $T_{m,s}$ should be similar (within 8°C or so). If they are too different, a suitable annealing temperature may be hard to find. At high annealing temperature,

the primer with the lower T_m may not work, whereas at low annealing temperature, amplification will be less efficient because the primer with the higher T_m will misprime.

In reality, primer selection is often empirical. It varies greatly from researcher to researcher in regard to the criteria they use.

1.3. Computer Programs for PCR Primer Design and Analysis

1.3.1. Computer Programs for Nondegenerate PCR Primer Design

For primer design, most researchers used to visually inspect target DNA sequence to find primer(s) with the characteristics they prefer, which are usually similar to the guidelines we mentioned earlier. As computers are widely used in molecular biology, a large number of computer programs have been specifically developed for nondegenerate primer selection, which makes the PCR primer design more efficient and reliable. Most sequencing analysis packages, such as Vector NTI (InforMax Inc.), usually contain a primer design module. In this chapter, we focus on free online (web) primer design programs (*see Note 2*). Selected computer programs for nondegenerate PCR primer design and their features are listed in **Table 1**.

From a computational point of view the design of nondegenerate PCR primers is relatively simple: find short substrings from DNA nucleotide string that meet certain criteria. Although the criteria vary between programs, the core parameters, such as the primer length, T_m , GC content, and self-complementarity, are shared by these programs.

1.3.2. Computer Programs for Degenerate PCR Primer Design

In the experiments to amplify the novel members of gene families or cognate sequences from different organisms by PCR, the exact sequence of the target gene is not known. We usually align all known sequences for this gene and find the most conserved regions, then design corresponding “degenerate” primers, which are a set of primers with nucleotide diversity at several positions in the sequence. Degeneracies obviously increase the chances of amplifying the target sequence but reduce the specificity of the primer(s) at the same time.

Designing degenerate primers has been considered more of an art than a science. There are much less computer programs for degenerate primer design (*see Table 2*) than for nondegenerate primer design.

1.3.3. Computer Programs for Primer Analysis

Even if you prefer to design primers by yourself, not by a computer program, it is advised that your primers should be analyzed by a computer program to determine T_m , possible hairpin structure, primer-dimers, and other properties before you place the order for them. **Table 3** lists two computer programs for this purpose.

2. Materials

1. Computer: A computer (PC or Macintosh) with high-speed internet access.
2. Programs: Web Browser, Netscape (5.0 or above) or Internet Explorer (4.0 or higher).
3. Input files for primer design: DNA sequence file DNA.txt (*see Table 4*) and protein sequence file Protein.txt (*see Table 5*) (*see Note 4*).

Table 1
Selected Computer Programs for Nondegenerate PCR Primer Design

Program	Operating System	Features	URL (<i>see Note 3</i>)
Oligos	Windows 9X/NT	Free download The program includes several tools: make complement, reverse complement and inverted strand; search the sequence; extract from selected sites. (Reference Lowe T 1990)	http://www.biocenter.-helsinki.fi/bi/bare-1_html/oligos.htm
GCG Prime	Unix	Commercial Available within GCG This program selects primers according to a number of user-specified criteria including length, GC content, and annealing temperature. Potential primers can also be tested for self-complementarity and complementarity to each other to minimize the formation of primer dimers during the PCR.	http://www.gcg.com/products/wis-pkg-programs.html#Primer
Primer3	Internet Browser	Free Lots of user-configurable parameters Primer design for both PCR and hybridization Nice interface with useful help pages	http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
Web Primer	Internet Browser	Free Best for designing primers to clone yeast genes. Can use a standard yeast gene name or systematic yeast name as DNA source input	http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer
iOligo	Windows 9X/NT, Mac	Commercial Retrieval of Sequences from NCBI Sequence Editor Analysis of Oligonucleotide's Characteristics Submission of Oligo Orders by email	http://www.caesarsoftware.com/pages/products/ioligo/ioligo.shtml
xprimer	Internet Browser	Free The user can select repeat database and genome model. Nice graphical display of suggested primers	http://alces.med.umn.edu/webprimers.html
PCR Help!	Windows 9X/NT	Commercial Free Demo Download User-friendly "PCR Wizard" allows you to design primers to any given DNA template sequence as well as to generate a Techne Genius Thermal Cycler program file, which can be sent from a PC directly to multiple Genius thermal cyclers (up to 32)	http://www.techneuk.co.uk/CatMol/pcrhelp.htm
Oligo	Windows 9X/NT, Mac	Commercial Free Demo Download Nice graphical interface for searching, selecting, and analyzing primers from known sequences Cross-compatible Multiplex PCR Primer Search Priming Efficiency Calculations	http://www.oligo.net/
The Primer Generator	Internet Browser	Free Designs Site Directed Mutagenesis primers The program analyzes the original nucleotide sequence and desired amino acid sequence and designs a primer that either has a new restriction enzyme site or is missing an old one. This allows for faster sorting out of mutated and nonmutated sequences.	http://www.med.jhu.edu/medcenter/primer/primer.cgi

Table 2
Selected Computer Programs for Degenerate PCR Primer Design

Program	Operating System	Features	URL
GeneFisher	Internet Browser	Free Processes aligned or unaligned sequences of DNA or protein	http://bibiserv.techfak.uni-bielefeld.de/genefisher/
CODEHOP	Internet Browser	Free Design degenerate PCR primers from protein multiple sequence alignments. The multiple-sequence alignments should be of amino acid sequences of the proteins and be in the Blocks Database format	http://www.blocks.fhcr.org/codehop.html
Primer Premier 5	Mac and Windows 9X/NT	Commercial Free Demo Download Reverse translate a protein sequence and design primers in regions of low degeneracy	http://www.premierbiosoft.com/primerdesign/primerdesign.html

Table 3
Selected Computer Programs for PCR Primer Analysis

Program	Operating System	Features	URL
Oligo Analyzer	Internet Browser	Free Calculate T _m , find possible primer hairpin structure and primer dimer formation, Blast search databases for primer homologs	http://playground.idtdna.com/program/oligocalc/oligocalc.asp
NetPrimer	Internet Browser	Free Java Applet Analyze basic properties and second structures for an individual primer or primer pair. Also give a primer rating and a report of the analysis results	http://www.premierbiosoft.com/netprimer/netprimer.html

3. Methods

3.1. Designing Nondegenerate PCR Primers Using Primer3

Primer3 was developed at Whitehead Institute for Biomedical Research and Howard Hughes Medical Institute. It contains so many parameters that most people only need a subset of them to use as the criteria for primer selection.

3.1.1. Design Primers with the Default Settings

Primer3 provides default values for core parameters (*see* **Table 6** for a selected list. Go to Primer3 web page for a complete list and their meanings). If these default settings meet your needs, then use the following method to select your primers.

Table 4
Input File DNA.txt

```

1 GGGGAAGTGC AATCACACTC TACCACACAC TCTCTATAGT ATCTATAGTT GAGAGCAAGC
61 TTTGTTAACA ATGGCGGCTT CCATTGGAGC CTAAAATCT TCACCTTCTT CCCACAATTG
121 CATCAATGAG AGAAGAAATG ATTCTACACG TGCAATATCC AGCAGAAATC TCTCATTTC
181 GTCTTCTCAT CTCGCCGGAG ACAAGTTGAT GCCTGTATCG TCCTTACGTT CCCAAGGAGT
241 ACGATTCAAT GTGAGAAGAA GTCCATTGAT TGTGTCTCCT AAGGTGTTT CTGATTCGCA
301 GAATTCACAG ACATGTCTGG ATCCAGATGC TAGCAGGAGT GTTTTGGGAA TTATTCTTGG
361 AGGTGGAGCT GGGACCCGAC TTTATCCTCT AACTAAAAAA AGAGCAAAAC CTGCGGTTCC
421 ACTTGGAGCA AATTATCGTC TGATTGACAT TCCCCTAAGC AATTGCTTGA ACAGTAACAT
481 ATCCAAGATC TATGTTCTCA CACAATTCAA CTTCGCCTCT CTTAAATCGCC ACCTTTCACG
541 GGCATATGCT AGCAATATGG GAGAATACAA AAACGAGGGC TTTGTGGAAG TTCTTGCTGC
601 TCAACAAAGT CCGGAGAACC CCGATTGGTT CCAGGGCACT GCGGACGCTG TCAGACAATA
661 TCTGTGGTTG TTTGAGGAGC ATAATGTTCT TGAATACCTT ATACTTGCTG GAGATCATCT
721 GTATCGAATG GATTATGAAA AGTTTATTCA AGCCACAGG GAAACAGATG CTGATATTAC
781 TGTTGCCGCA CTGCCAATGG ACGAGAAGCG TGCCACTGCA TTCGGTCTCA TGAAGATTGA
841 CGAAGAAGGA CGCATTATTG AATTTGCAGA GAAACCGCAA GGAGAGCAAC TGCAAGCAAT
901 GAAAGTGGAT ACTACCATTT TAGGTCTTGA TGACAAGAGA GCTAAAAGAAA TGCCTTTTAT
961 CGCCAGTATG GGTATATAGT TCATTAGCAA AGACGTGATG TTAACCTAC TCGTGCACAA
1021 GTTCCCTGGG GCCAATGATT TTGGTAGTGA AGTTATTCTT GGTGCAACTT CACTTGGGAT
1081 GAGAGTGCAG GCTTATTTAT ATGATGGGTA CTGGGAAGAT ATTGGTACCA TTGAAGCTTT
1141 CTACAATGCC AATTTGGGCA TTACAAAAAA GCCGGTGCCA GATTTTAGCT TTTACGACCG
1201 ATCAGCCCCA ATCTACACC AACCTCGATA TTTGCCACCT TCAAAAATGC TTGATGCCGA
1261 TGTACAGGAT AGTGTCAATT GTGAAGGTTG TGTGATCAAG AACTGTAAGA TTCACCATT
1321 CGTGGTTGGG CTCAGATCAT GCATATCAGA GGGAGCAATT ATAGAAGACT CACTTTTGAT
1381 GGGGGCAGAT TACTACGAGA CTGATGCTGA GAGGAAGCTG CTGGCTGCAA AGGGCAGTGT
1441 CCCAATTGGC ATCGGCAAGA ATGTCTATA CAAAAGAGCC ATTATCGACA AGAATGCTCG
1501 TATAGGGGAC AATGTGAAGA TCATTAACAA AGACAATGTT CAAGAAGCGG CTAGGGAAAC
1561 AGATGGATAC TTCATCAAGA GTGGGATCGT CACTGTCTATC AAGGATGCTT TGATTCCAAG
1621 TGGAATCGTC ATTTAAAGGA ACGCATTATA ACTTGGTTGC CCTCCAAGAT TTTGGCTAAA
1681 CAGCATGAG TTACAAACGT GCCGAAGTTT TATTTTCTTA TGCTGTAGAA ATCTAGTGTA
1741 CATCTGCTT TTATGATACT TCTATTACC TGTTGCTGTG AAAAATTATT CGTCTAAAA
1801 AAAAAATAAT CTACCATTAC ACCA

```

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi and hit return. After connection Primer3 web page will appear in your browser.
3. Open the DNA sequence input file DNA.txt (*see Note 5*) using your favorite text editor, such as Notepad in Windows, then copy the sequence by going to **Edit/Select All, Edit/Copy** in the menubar. Close file DNA.txt.
4. In your browser click on the top sequence input box, then paste the above sequence by going to **Edit/Paste** in the menubar.
5. Click **Pick Primers** Button (there are six Pick Primers buttons on the page. Click any one of them will do the same). After a few s/min, Primer3 Output will be returned. The top part of the output is shown in **Table 7**.

The other parts of the output not shown are: whole input sequence and arrows, which nicely indicate the location of the primers above; additional four primer pairs; and statistics about the primer selection process.

Table 5
Input File Protein.txt

```

>Protein1
MKSTVHLGRVSTGGFNNGEKEIFGEKIRGSLNNLRINQLSKSL
KLEKKIKPGVAYSVITTENDTETVFVDMRLEERRANPKDVAAVILGGGEGTKLFLPLT
SRTATPAVPVGGCYRLIDIPMSNCINSAINKIFVLTQYNSAALNRHIARTYFGNGVSF
GDGFVEVLAATQTPGEAGKKWFQGTADAVRKFIVWFEDAKNKNIENILVLSGDHLYRM
DYMELVQNHIDRNADITLSCAPAEDSRASDFGLVKIDSRGRVQFAENQRFELKAMLV
DTSLVGLSPQDAKKSPYIASMGVYVFKTDVLLKLLKWSYPTSNDFGSEIIPAAIDDY
VQAYIFKDYWEDIGTIKSFYNASLALTOEFPEFQFYDPKTPFYTSRFLPPTKIDNCK
IKDAIISHGCFLRDCTVEHSIVGERSRLDCGVELKDTFMMGADYYQTESEIASLLAEG
KVPIGIGENTKIRKCIIDKNAKIGKNVSIINKDGVQEADRPEEGFYIRSGIIISEKA
TIRDGTVI
>Protein2
MDALCAGTAQSVAICNQESTFWGQKISGRRLINKGFGVRWCKSF
TTQQRGKNVTS AVLTRDINKEMPLFENS MFEEQPTAEPKAVASVILGGGVGTRFLPLT
SRRAKPAVPIGGCYRVIDVPMNSCINS GIRKIFILTQFNSFSLNRHLARTYNFGNGVG
FGDGFVEVLAATQTPGDAGKMWFQGTADAVRQFIWVFENQKNKNVEHIILSGDHLYR
MNYMDFVQKHIDANADITVSCVPMDDGRASDFGLMKIDETGRIHQFVEKPKGPALKAM
QVDTSILGLSEQEASNFPYIASMGVYVFKTDVLLNLLKSAYPSCNDFGSEIIPSAVKD
HNVQAYLFNDYWEDIGTVKSFFDANLALTKQPPKFDNDPKTPFYTSARFLPPTKVDK
SRIVDAIISHGCFLECNQHSIVGVRSLDYGVFEFKDTMMMADYYQTESEIASLLA
EGKVPIGVGPNTKIQKCIIDKNAKIGKDVVILNKQGVVEADRSAEGFYIRSGITVIMK
NATIKDGTVI
    
```

Table 6
Selected Default Settings for Primer3

Parameter	Minimum	Optimum	Maximum
Primer size (base pairs)	18	20	27
Primer T_m (°C)	57	60	63
Max T_m Difference (°C)			100
Primer GC%	20		80
Product size (basepairs)	100	200	1000

Table 7
Primer3 Output

```

WARNING: Numbers in input sequence were deleted.

No mispriming library specified
Using 1-based sequence positions
OLIGO  start len  tm  gc%  any  3' seq
LEFT PRIMER  890  20  59.99  45.00  4.00  0.00  ctgcaagcaatgaaagtgga
RIGHT PRIMER 1090  20  59.83  50.00  4.00  2.00  ttgcactctcatccaagtg
SEQUENCE SIZE: 1824
INCLUDED REGION SIZE: 1824

PRODUCT SIZE: 201, PAIR ANY COMPL: 7.00, PAIR 3' COMPL: 3.00
    
```

3.1.2. Design Primers with User-Defined Settings

Often the default values need to be altered because they do not meet a researcher's needs or Primer3 did not find an appropriate PCR primer pair. The following are helpful guidelines for adjusting these parameters if Primer3 failed to select a primer:

- a. Adjust location: pick a wider range to examine and allow for longer product size;
- b. Change primer size: usually easier to find compatible primers if they are shorter;
- c. Lower primer T_m .

Because there are so many configurable parameters in Primer3, it is impossible to explain their uses and try to change them here. Fortunately, the default values need not to be altered for most parameters. The readers should read the Primer3 help page and understand the uses of the parameters before trying to change them.

In the following method, we will try to design primers to clone the coding region in DNA.txt, which is from nucleotide 71 to 1636.

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi and hit return. After connection, Primer3 web page will appear in your browser.
3. Open the DNA sequence input file DNA.txt using your favorite text editor, such as Notepad in Windows, then copy the sequence by going to **Edit/Select All**, **Edit/Copy** in the menubar. Close file DNA.txt.
4. In your browser, click on the top sequence input box, then paste the above sequence by going to **Edit/Paste** in the menubar.
5. Type 71,1565 in the **Targets** input box. Change **Product Size/Max** from 1000 to 1824, then click **Pick Primers** button. (There are six Pick Primers buttons on the page. Click any one of them will do the same). After a few s/min, Primer3 Output will be returned. Are there any primers returned?

3.2. Designing Degenerate PCR Primers Using GeneFisher

GeneFisher is an interactive degenerate primer design software. The current version, GeneFisher 1.22, processes aligned or unaligned sequences of DNA or protein. In the following method, we will use two unaligned protein sequences as the sequence input and design degenerate primers which could amplify the cDNAs encoding these two proteins and their related family members (if any).

1. Start web browser (*see Note 6*).
2. Replace the default URL address with <http://bibiserv.techfak.unibielefeld.de/genefisher/> and hit return. After connection, **GeneFisher Interactive PCR Primer Design** home page will appear in your browser.
3. Click **Start** button on the page. After a few s/min, the Interactive **Primer Design** interface will be open.
4. In the **User Data** area of the page, type your **E-mail ID** and **Project** name. Click **this** button in the **Sequence Data** area to clear the sample sequence, then copy the two protein sequences from Protein.txt and paste to the **Sequence Data** area (*see Note 7*). Click **OK** button in the **Submit Query** area to accept your choice.

5. After a few s/min **GeneFisher Sequence Input** page will appear. Check that the protein lengths match with the input sequences. Click **OK** button to accept the two protein sequences.
6. After **GeneFisher Alignment Status** page returns, click **OK** button to use ClustalW as the alignment tool. ClustalW Multiple Sequence Alignment Setup page will appear. Click **OK** button to accept the default parameters.
7. Click **Progress** button on the **GeneFisher Clustal Alignment** page, which will open a new browser window. Click **Reload** button repeatedly in the new window to check the status of the alignment. If the last line on the page shows “GDE-Alignment file created” (The alignment time depends on your input. It takes several minutes for Protein.txt.) Then click **Alignment** button in the original window, which will show the alignment results in a new window.
8. We are satisfied with the alignment results, so go to the original window and click the **Consensus** button. **Sequence Consensus** page will return. Click **OK** button to accept the default consensus parameters, which will open the **GeneFisher Consensus** page.
9. Click **Progress** button to check the consensus calculation progress. If you are satisfied with the consensus calculation, click **Consensus** button on the **GeneFisher Consensus** page, which will show the alignment results in a new window. Click **Go!** button in the original window to generate primers.
10. After a few s/min, **Primer Design** page will appear. Click **OK** button to accept the default settings for primer design, which will open **GeneFisher Primer Calculation** page. Wait a few s/min, then click **Results** button, which will open the **Primer Calculation Results** page. Unfortunately, the results show that no primer pairs were generated. The rejection statistics underneath give some clues on why the primer selection fails. Click the **Redo** button to return to the **Primer Design** page.
11. Make the following changes to the primer parameters:
 - a. Set primer length from 15 to 22 bp.
 - b. Set GC content from 35 to 85%.
 - c. Set melting temperature T_m from 42 to 65°C.
 - d. Set product size from 100 to 1500 bp.
 - e. Set primer degeneracy 512-fold.
 - f. Set 3' GC content from 35 to 85%.

Repeat the primer design step above (**step 10**). This time, seven primer pairs were returned (*see Table 8*). If you click the primer sequence link (Forward Primer or Reverse Primer), the **GeneFisher Primers Profile - Data Sheet** about that primer pair will be returned in a new window. If you click the primer position link (FPPos. RPPos.), the **Textual Primer Pair Visualization** of that primer pair will be shown in a new window.

3.3. Analyze PCR Primers Using NetPrimer

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with <http://www.premierbiosoft.com/netprimer/netprimer.html> and hit return. After connection, NetPrimer web page will appear in your browser.
3. Click the **click here** link in the page to launch the NetPrimer applet.
4. After the applet is launched, type the following sequence in the **Oligo Sequence** input area: ctgcaagcaatgaaagtga, then click the **Analyze** button. The analysis results of the primer, such as T_m , molecular weight, GC%, rating, and stability, will be shown in the

Table 8
GeneFisher Output (IUB Code for Sequence)

7 best Pairs (of max. 7)							
ID	Forward Primer	Reverse Primer	Qual.	Prod.		FPPos.	RPPos.
				Len.	T _m Diff.		
1	NTAYMGNATGRAYTAYATGGA	GTyTGrTArTArTCnGCnCCCA	659	653	6	653	1306
2	NTAYMGNATGRAYTAYATGGA	TyTGrTArTArTCnGCnCCCAT	658	652	5	653	1305
3	NTAYMGNATGRAYTAYATGGA	AynGTnCCdATrTCyTCCCA	398	388	6	653	1041
4	NTTYAANGAYTAYTGGGA	GTyTGrTArTArTCnGCnCCCA	290	278	12	1028	1306
5	NTTYAANGAYTAYTGGGA	TyTGrTArTArTCnGCnCCCAT	289	277	11	1028	1305
6	NTAYMGNATGRAYTAYATGGA	GTyTTrAAnACrTAnACnCCCA	264	248	6	653	901
7	NTAYMGNATGRAYTAYATGGA	TyTTrAAnACrTAnACnCCCAT	262	247	5	653	900

Results area of the applet. You may also click the following buttons: **Hairpin**, **Dimer**, **Palindrome**, and **Repeat & Run**, to check the corresponding properties about the primer.

4. Notes

1. This formula only gives a very approximate T_m in the absence of denaturing agents such as formamide and DMSO, and it is only valid for primers < 20 nucleotides in length. For PCR purposes $T_m - 5^\circ\text{C}$ is a good annealing temperature to start with. However, optimal annealing temperatures can only be determined experimentally for a certain primer/template combination and there is no formula currently available to accurately define their relationships.

For longer primers, the nearest-neighbor method (6) offers a reliable estimation of the T_m and its formula is the following:

$$T_m = \Delta H / (A + \Delta S + R \times \ln[C/4]) - 273.15 + 16.6 \times \log[\text{salt}]$$

where:

ΔH (cal/mole) is the sum of the nearest-neighbor enthalpy changes for DNA helix formation (<0).

A (cal/degree Celsius/mole) is a constant for helix initiation, which is equal to -10.8 cal/degree Celsius/mole for nonself-complementary sequences and -12.4 for self-complementary sequences.

ΔS (cal/degree Celsius/mole) is the sum of the nearest-neighbor entropy changes for helix formation (<0).

R is the molar gas constant (1.987 cal/degree Celsius/mole).

C is the primer concentration.

[salt] is the salt concentration.

However, primer design programs may use different formula to calculate T_m . For example, the Primer3 program uses the following formula:

$$T_m = 81.5 + 16.6(\log_{10}([\text{Na}^+])) + 0.41 \times (\%GC) - 600/\text{length}$$

where $[\text{Na}^+]$ is the molar sodium concentration, (%GC) is the percent of Gs and Cs in the sequence, and length is the length of the sequence.

2. Keep in mind that internet is not a secure place to send your sequences. If you care about the privacy of your sequences or do not have internet access, then download a freeware or buy a commercial package and use them instead to design your primers.
3. URL stands for Universal Resource Locator, which is a unique address on the internet. However, URL is quite dynamic. Old web sites could be shut down and new sites could be set up, resulting in change of web address for a particular page or disappearance of a web page. Try to search for the new address of a web page by using a search engine, such as www.google.com.
4. You can also use your own in-house sequences for primer designs discussed in the **Methods** section. Of course, the results will vary.
5. For the input sequence of Primer3, numbers and blanks are ignored. Other letters are treated as N. FASTA format is acceptable. It is assumed that the strand direction is 5'→3'.
6. Although GeneFisher system has been optimized for Netscape Navigator version 4.x and above, our testing showed that it works fine with Internet Explorer 5.0 and above when protein sequences are used as inputs. Netscape Navigator should be used when DNA sequences are used as inputs.
7. If you use your own input sequences, make sure that the sequences have significant homology. Otherwise the primer pair which meets your parameters will be very hard to find, if not impossible. Do not use the **Browse...** button to load your sequences, as it appears that there is a bug in reading the sequences from a file by GeneFisher.

References

1. Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T., and Itakura, K. (1979) Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. *Nucl. Acids Res.* **6**, 3543–3557.
2. Rychlik, W. and Rhoads, R. E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucl. Acids Res.* **17**, 8543–8551.
3. Wu, D. Y., Ugozzoli, L., Pal, B. K., Qian, J., and Wallace, R. B. (1991) The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA Cell Biol.* **10**, 233–238.
4. Rychlik, W. (1993) Selection of primers for polymerase chain reaction, in *PCR Protocols. Current Methods and Applications* (White, B. A., ed.), Humana, Totowa, NJ, pp. 31–40.
5. Innis, M. A. and Gelfand, D. H. (1990) Optimization of PCRs, in *PCR Protocols* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.), Academic, New York, pp. 3–12.
6. Breslauer, K. J., Frank, R., Blocker, H., and Markey, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.

Single-Step PCR Optimization Using Touchdown and Stepdown PCR Programming

Kenneth H. Roux

1. Introduction

Polymerase chain reaction (PCR) optimization and troubleshooting can consume considerable energy and resources because of the finicky and often unpredictable nature of the reactions. Small variations in any of the many variables in a given reaction can have a pronounced effect on the resultant amplicon profile. Reactions that are too stringent yield negligible product, and reactions that are not stringent enough yield artifactual amplicons. Variables include concentrations of Mg^{2+} , H^+ , dNTPs, primers, and template, as well as cycling parameters. Regarding the latter, the value selected for the annealing temperature is most critical. Unfortunately, even with the most sophisticated algorithms (i.e., OLIGO) it is often difficult to predict the amplification optima *a priori* leaving no other choice but to employ empirical determination.

Touchdown (TD) PCR (1,2) and its sister technique, stepdown (SD) PCR (2) represent a markedly different approach that, in a single amplification regimen, inherently compensates for suboptimal reagent concentrations and less than perfect cycling parameters. Rather than guessing (or using imprecise calculations) to arrive at an appropriate temperature for the primer extension segment of the cycle, one can cast a wider net by using progressively lower annealing temperatures over consecutive cycles. The goal is to select a broad range of annealing temperatures that begins above the estimated melting temperature (T_m) and ends below it (see **Note 1**). Typically, one runs a TD PCR program at 2 cycles/°C declining over a 10–20°C range at 1°C intervals. In this way, the first primer–template hybridizations and primer extensions will be those with the highest specificity, i.e., presumably, the combination that gives the desired amplicon. Although the annealing temperature continues to drop in subsequent cycles to levels that normally would promote spurious amplification, the desired product, having already experienced several cycles of amplification, will be in a position to out-compete most lower T_m (spurious) amplicons. If, for example, there is only a 3°C

difference between the T_m of the target amplicon and the T_m of the first-primed spurious amplicon, the desired product will have undergone up to a 64-fold (2^6) amplification.

Our experience has been that TD PCR is applicable to a wide range of PCR situations (2,3). At one extreme, TD PCR generally yields a single strong amplicon from genomic DNA even when the primer–template combinations are grossly mismatched (see **Note 2**). Mismatching might occur when attempting to amplify specific members from a complex multigene family, when using nucleotide sequence information deduced from an amino acid sequence, or when amplifying across species lines (4–6). Primer–template basepair mismatches are permissible and can even be near (but probably not at) the 3' end of the primer (3). TD PCR can also compensate for suboptimal buffer composition (e.g., Mg^{2+} concentration) (2). On the other hand, reactions that are already optimal, as assessed by conventional PCR, will usually yield equally strong amplicons even when using a broad temperature range TD PCR protocol in which the annealing temperature dips well below the T_m . Stated another way, TD PCR appears to greatly aid marginal reactions while not imposing significant penalties on already robust reactions. Hence, TD PCR can be used routinely in lieu of conventional PCR and need not be viewed solely as an optimization procedure (7).

One potential drawback to TD PCR stems from the complexity of the programming on some thermal cyclers (see **Note 3**). Because of the numerous (10 to 20) annealing temperatures used, a large segment of the programming capacity of conventional thermal cyclers can be encumbered. Also, attempts to adjust the annealing temperature range can involve considerable reprogramming (see **Note 4**). Some newer thermal cyclers circumvent these problems by permitting the programming of automatic incremental temperature changes in progressive cycles. We have tested modified versions of TD PCR, which we term SD PCR, that utilize simplified programming (2). For SD PCR, one uses fewer, but larger, annealing temperature steps with proportionately more cycles per step. For example, a program might consist of three or four steps, at three to four cycles per step, with 3–5°C temperature differences between steps. SD PCR is not quite as universally applicable as TD PCR, but is adequate for many applications. The ease in programming may frequently be worth the trade-off.

2. Materials

2.1. Touchdown PCR with Mismatched Primer–Template Pairs

1. Template DNA (rabbit genomic liver 125 ng/ μ L).
2. Primers. Stock solutions are at 200 μ g/mL in H_2O . The following primer pairs yield a 445-bp amplicon. The sites of mismatches are in capital letters. The sequence of the corresponding genomic homologous strand (shown in brackets) are for comparative purposes (see **Note 5** for comments on degenerate primer design).
 - a. Primers: 5'cttgccaGtaatatAcgcctgcTaaCTtg3'; 5'ggatcttctgtgatgtctgactGttGgAg 3'
 - b. Homologous genomic sequences: [5'cttgccaAtaatatCgcctgcCaaTCTg 3']; 5'ggatcttctgtgatgtctgactAttTgTg 3'
3. 10X PCR buffer: 500 mM KCl, 100 mM Tris-HCl, pH 9.0, 1% Triton X-100, 2 mM of each dNTP (A, T, C, and G), and 15 mM $MgCl_2$ (see **Note 6**).
4. AmpliTaq DNA Polymerase (Perkin Elmer, Norwalk, CT).
5. Sterile mineral oil.

6. Standard wall 0.6-mL capped conical tubes.
7. Equipment and reagents for 1.5% agarose gel electrophoresis.

2.2. Stepdown PCR with a Mismatched Degenerate Primer

1. Template DNA (rabbit genomic liver 125 ng/ μ L).
2. Primers: Prepare stock solutions at 200 μ g/mL. The following primer pair yields a 703-bp amplicon. The sites of mismatches are in capital letters. Degeneracies are separated by a slash and are in parenthesis. The sequence of the genomic homologous strand corresponding to the primer with degeneracies is presented (shown in brackets) for comparative purposes.
 - a. Primers: 5'agggatcgggtgaaaggggtctcagc3'; 5'ttAtgagcattcat(a/G)aactctggagg3'.
 - b. Homologous genomic sequence: [5'agggatcgggtgaaaggggtctcagc3']; [5'ttGtgagcattcat-Aaactctggagg 3'].
3. 10X PCR buffer: 500 mM KCl, 100 mM Tris-HCl, pH 9.0, 1% Triton X-100, 2 mM of each dNTP (A, T, C, and G), and 15 mM MgCl₂.
4. AmpliTaq DNA Polymerase.
5. Equipment and reagents for 1.5% agarose gel electrophoresis.

3. Methods

Programming of the thermal cycler for TD and SD PCR are described first, followed by specific PCR conditions used for both reactions.

3.1. TD PCR Programming

1. Set thermal cycler to denature for 1 min at 94°C, anneal for 2 min, and primer extend for 3 min at 74°C.
2. Follow the cycling program with a 7-min primer extension step and a 4°C soak step (*see Note 7*).
3. Set the annealing stage for 2 cycles/°C beginning at 55°C and decreasing at 1°C increments to 41°C (i.e., 30 total cycles in 15 steps) to be followed by 10 additional cycles at 40°C.

3.2. SD PCR Programming

1. Set thermal cycler to denature for 1 min at 94°C, anneal for 2 min, and primer extend for 3 min at 74°C as detailed later. Follow the cycling program by a 7-min primer extension step and a 4°C soak step (*see Note 7*). Program the annealing stage for six cycles per temperature step beginning at 70°C and decreasing at 3°C increments to 58°C (i.e., 30 total cycles in five steps) to be followed by 10 additional cycles at 55°C (*see Note 8* for programming considerations).
2. Analyze and reamplify as described later.

3.3. PCR Setup

1. Set up master mix for 50- μ L reactions as indicated in **Table 1**.
2. Dispense master mix to 0.6-mL standard wall PCR tubes. Add 50 μ L mineral oil to each tube and place in thermal cycler (Perkin Elmer DNA Thermal Cycler). Begin initial cycle and add 2.5 μ L (1.25 U) of a 1:10 dilution of polymerase to each tube only after the temperature exceeds 80°C in the thermal cycler (i.e., the hot start protocol, *see Note 9*). Cap tubes and continue cycling.
3. Following amplification, monitor results by running 3 μ L on a 1.5% agarose/ethidium bromide gel and view by UV illumination. If a product is not evident or the desired

Table 1
PCR Master Mix

Components	Stock concentration	Amount per 50 μ L reaction, μ L	Final concentration
dNTP mix	2.0 mM ea.	5	0.2 mM
PCR buffer	10X	5	1X
Primer 1	200 ng/ μ L	1	4.0 ng/ μ L
Primer 2	200 ng/ μ L	1	4.0 ng/ μ L
Template	125 ng/ μ L	3	7.5 ng/ μ L
MgCl ₂	25 mM	3	1.5 mM
H ₂ O	—	29.5	—
		Subtotal: 47.5	

amplicon is of insufficient amount, amplify for an additional 5–10 cycles (with TD PCR, you are not sure exactly how many nonproductive and suboptimal cycles preceded the start of efficient amplification) at 5°C below the lowest annealing temperature previously used.

4. If the desired product is still not evident, consider conventional or TD nested PCR on a 1:100 to 1:1000 dilution of the initial TD PCR reaction.

4. Notes

1. The T_m for the primer–template combination can be roughly estimated using the formula:

$$T_m = 2(A + T) + 4(G + C) \quad (1)$$

For primer–template combinations with known or suspected mismatches, 5–20°C should be subtracted from the normal annealing temperature. Of course, more sophisticated programs, such as OLIGO Primer Analysis Software (National Biosciences, Inc., Plymouth, MN) (10,11), may also be used to calculate the T_m .

2. An estimation of the T_m is particularly difficult when using primers and templates containing mismatched base pairs. We have successfully amplified, to a single intensely staining band, primer–template pairs containing three to five mismatches with the template (3). In all cases, we used a TD PCR program in which the annealing temperatures dropped from 55–41°C at 2 cycles/°C.
3. If using a thermal cycler that has a programmable automatic temperature variation feature, set the annealing stages to decline by 0.5°C/cycle. For a standard thermal cycler, program 2 cycles/°C drop (for example, cycles 1 and 2, 65°C; cycles 3 and 4, 64°C, and so forth). In both instances, the TD portion of the program should be followed by 10 cycles at a fixed annealing temperature about 10°C below the estimated T_m . One should bear in mind that for situations in which the template is not fully complementary to the primers, once amplification commences, the amplicon will be fully complementary to the primers and thus will have a greater T_m than initially estimated. On the other hand, we have noted that final stage amplification at 10°C or more below the estimated T_m (rather than the 4–5°C usually recommended for standard PCR) can significantly increase the yield of otherwise marginal reactions (2). Avoid the temptation of adding too many cycles to this terminal fixed annealing temperature stage of the program. Excessive cycling can degrade the product and lead to spurious banding and high molecular weight smearing (14).

4. A convenient way to adjust the TD temperature range segment of a conventional thermal cycler (i.e., having linked or sequential file programming) is to program files covering a wide range of annealing temperature (20–25°C). The specific subset of files to be used in any given amplification protocol can be incorporated by simply linking the initial 5 min denaturation file to the file having the highest annealing temperature to be used and linking the file containing the lowest annealing temperature in the TD range to the terminal primer extension file. Be sure to keep note of these changes because they must be undone before the next alteration in the range. If this approach will tie up many files, you may wish to reprogram those files containing the segments to be deleted from the range (say 5°C from the bottom) and using the freed file capacity to add new files to the top (beginning) of the program. Again, the initial denaturation stage and terminal primer extension steps must be linked to the beginning and end of the new range, respectively. Thermal cyclers in which individual files cannot be linked but which rely on a single long multi-step program are even less versatile and may require complete reprogramming. Fortunately, TD and SD PCR are very forgiving and given temperature range protocol can be applied to a wide variety of situations.
5. Design the best primer set based on the information available. When designing primers to amplify genes of uncertain complementarity, try to cluster the sequence of greatest certainty near the 3' ends of the primers. Degeneracy derived from multiple nucleotides or inosine residues at positions of uncertainty are permissible (4,5,8) but not necessary. Note that some polymerases other than *Taq* cannot prime from inosine-containing primers (9).
6. Because of the minimal effort involved, it is generally advantageous to vary one of the buffer components (usually Mg²⁺) during the initial optimization.
7. Most time and temperature characteristics of the cycling program (denaturation, primer extension) will be the same as conventional PCR for your system if using primer–template combinations other than the examples described.
8. The temperature range of SD PCR may be divided into three to five more or less equal increments (steps) and be programmed accordingly. If you can afford to tie up the programming capacity of the thermal cycler, more steps are better than fewer. Proportion the total number of cycles to be used in the SD segment of the program equally among the steps. Add 10 cycles at a fixed temperature well below the T_m as described earlier. When background problems are expected to be minimal, a simple two step SD PCR protocol can still be advantageous. Here, the initial stage has perhaps six to 10 cycles at a fixed temperature 5°C above that which would normally be used in standard PCR (i.e., slightly above the calculated T_m). Even though a full 30 cycles at this elevated temperature would not be expected to yield a detectable amplicon, we have found (2) that sufficient, highly specific, amplification is occurring to allow the desired amplicon to dominate the amplification throughout the remainder of the standard temperature cycles, thus reducing the possibility that unwanted amplicons will be generated.
9. Because TD PCR is based on the use of high temperature to prevent spurious priming, it is imperative that hot start procedures be followed (12,13). If multiple samples are to be run, add an extended denaturation step to the beginning of the program or use the hold option.

References

1. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991) 'Touch-down' PCR to circumvent spurious priming during gene amplification. *Nucl. Acids Res.* **19**, 4008.

2. Hecker, K. H. and Roux, K. H. (1996) High and low annealing temperatures increase both specificity and yield in TD and SD PCR. *BioTechniques* **20**, 478–485.
3. Roux, K. H. (1994) Using mismatched primer–template pairs in TD PCR. *BioTechniques* **16**, 812–814.
4. Knoth, K., Roberds, S., Poteet, C., and Tamkun, M. (1988) Highly degenerate inosine-containing primers specifically amplify rare cDNA using the polymerase chain reaction. *Nucl. Acids Res.* **16**, 10,932.
5. Patil, R. V. and Dekker, E. E. (1990) PCR amplification of an *Escherichia coli* gene using mixed primers containing deoxyinosine at ambiguous positions in degenerate amino acid codons. *Nucl. Acids Res.* **18**, 3080.
6. Batzer, M. A., Carlton J. E., and Deininger, P. L. (1991) Enhanced evolutionary PCR using oligonucleotides with inosine at the 3'-terminus. *Nucleic Acids Res.* **19**, 5081.
7. Roux, K. H. (1995) Optimization and troubleshooting in PCR. *PCR Meth. Applicat.* **4**, S185–S194.
8. Peterson, M. G., Inostroza, J., Maxon, M. E., Flores, O., Adomon, A., Reinberg, D., and Tjian, R. (1991) Structure and functional properties of human general transcription factor II ϵ . *Nature* **354**, 369–373.
9. Knittel, T. and Picard, D. (1993) PCR with degenerate primers containing deoxyinosine fails with Pfu DNA polymerase. *PCR Meth. Applicat.* **2**, 346–347.
10. Rychlik, W. and Spencer, W. J. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucl. Acids Res.* **17**, 8543–8551.
11. Rychlik, W. (1994) New algorithm for determining primer efficiency in PCR and sequencing. *J. NIH Res.* **6**, 78.
12. D' Aquila, R. T., Bechtel, L. J., Viteler, J. A., Eron, J. J., Gorczyca, P., and Kaplin, J. C. (1991) Maximizing sensitivity and specificity of PCR by preamplification heating. *Nucl. Acids Res.* **19**, 3749.
13. Erlich, H. A., Gelfand, D., and Sninsky, J. J. (1991) Recent advances in the polymerase chain reaction. *Science* **252**, 1643–1651.
14. Bell, D. A. and DeMarini, D. (1991) Excessive cycling converts PCR products to random-length higher molecular weight fragments. *Nucl. Acids Res.* **19**, 5079.

XL PCR Amplification of Long Targets from Genomic DNA

Lori A. Kolmodin

1. Introduction

Long polymerase chain reaction (PCR) $>(1-5)$, specifically, XL PCR (Extra-Long Polymerase Chain Reaction), has enabled amplification of expanded trinucleotide repeats of the neuromuscular disease myotonic dystrophy (6) a 9-kb HIV-1 provirus from primary isolate DNA (7) , 24.2-kb fragments from nanogram quantities of genomic DNA for DNA damage repair (8) , and up to 42 kb of human genomic DNA (4) . The capability of the long PCR process stems through the use of:

1. Two thermophilic polymerases, one of which is highly processive (5' to 3' polymerase activity) and the other one possesses proofreading activity (3'-5' exonuclease activity) known to improve fidelity by removing mismatch bases that can cause pauses in polymerization or termination of strand synthesis (2) ;
2. Reduced denaturation time at moderately high temperatures aimed at protecting the DNA template against damage;
3. An alkaline tricine buffer suggested to protect DNA from being nicked at elevated temperatures as a result of acidic conditions;
4. Cosolvents such as dimethyl sulfoxide (DMSO) and glycerol known to stabilize enzymes (9) , lower secondary structure, and lower melting and strand separation temperatures (1) ;
5. Increased extension time to allow the polymerase to complete strand synthesis and to lower the frequency of recombination between partial extension products;
6. Consideration of the integrity of the template (1) .

Primer design and reaction stringency must also be considered for successful single-copy gene amplifications from complex genomic DNA. Optimal reaction conditions are system dependent, particularly as the target length increases. With increasing understanding of the key PCR parameters and subsequent improvements in PCR process, increasingly longer targets will be routinely amplified.

This ability to amplify longer targets by PCR has numerous applications. For example, larger steps may be taken in gene mapping and sequencing efforts. Amplification

of increasingly larger segments of any gene can also enable the simultaneous characterization of several regions of interest, for both research and diagnostic purposes. XL PCR protocols may facilitate studies of apparently unclonable regions, and of certain viral genomes or recombinant phage λ clones that are not readily cultured (3,5). XL PCR can also complement cloning approaches by providing larger quantities of longer inserts.

2. Materials

2.1. Preparation of Genomic DNA (see Notes 1 and 2)

1. Source of Template DNA, such as cultured cells available from the NIGMS Human Genetic Mutant Cell Repository (see Coriell Institute for Medical Research, Camden, NJ).
2. Phosphate-buffered saline or other balanced salt solution.
3. 10 mM Tris-HCl, pH 8.2, 0.4 M NaCl, 2 mM ethylenediamine tetracetic acid (EDTA).
4. 10% sodium dodecyl sulfate (SDS).
5. Proteinase K (2 mg/mL). Prepare fresh in 1% SDS, 2 mM EDTA.
6. RNase A (10 mg/mL). Heat inactivated at 95(–100°C for 20 min).
7. Saturated NaCl (approx 6 M).
8. 95% (v/v) ethanol.
9. 70% (v/v) ethanol.
10. TE buffer: 10 mM Tris-HCl, pH 7.5–8.0, 1 mM EDTA.

2.2. Analysis of Template Integrity

1. Agarose (see Subheading 2.3.).
2. 50 mM NaCl, 1 mM EDTA.
3. 10X Alkaline running buffer: 0.5 N NaOH, 10 mM EDTA.
4. 6X Ficoll/bromocresol green solution: 0.3 N NaOH, 6 mM EDTA, 18% Ficoll (type 400 from Pharmacia, Piscataway, NJ, or Sigma, St. Louis, MO), 0.15% bromocresol green, 0.25% xylene cyanol FF.
5. 0.1 M Tris-HCl, pH 8.0, 1 mM EDTA.
6. Ethidium bromide solution: 0.5 μ g/mL in TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH approx 8.5).

2.3. Primers

Use primers designed with high specificity (see Subheading 3.3. and Note 3). Primer sequences for the 7.5 kb and 17.7 kb of the human β -globin gene cluster and 16.3 kb of the mitochondrial genome are listed in Note 4. Nontemplate sequences can be added at the ends of primers to utilize differential display methodology for detecting changes in gene expression and for subsequent restriction digestion and cloning (see Notes 4 and 5). Primers with one to two phosphorothioate bonds at the 3'-terminal internucleotide linkage can be employed to minimize primer degradation (see Notes 6 and 7) (10,11).

2.4. PCR Amplification

1. GeneAmp XL PCR Kit (Applied Biosystems, Foster City, CA) consisting of: r*Tth* DNA polymerase, XL; XL PCR Buffer II; dNTP blend; Mg(OA)₂; and control template and primers for amplification of a 20.8-kb sequence from the phage λ genome (see Notes 8 and 9).

2. AmpliWax PCR Gem 100s (Applied Biosystems).
3. Perkin Elmer GeneAmp PCR Instrument System and PCR tubes with which the XL PCR reagents have been optimized and quality control tested (e.g., Perkin Elmer GeneAmp PCR System 9600 with 0.2 mL MicroAmp reaction tubes [Applied Biosystems]).

2.5. Product Analysis

1. SeaKem GTG agarose or SeaKem Gold agarose (FMC Bioproducts, Rockland, ME).
2. TAE buffer: 40 mM Tris-acetate, 2 mM EDTA, pH 8.5 or TBE buffer: 89 mM Tris borate, 2 mM EDTA.
3. Ethidium bromide (10 mg/mL). Store in the dark. Wear gloves when handling.
4. Restriction endonucleases.
5. Micron-100 spin column (Amicon, Beverly, MA) or Qiaquick PCR Purification Kit (QIAGEN, Chatsworth, CA).

3. Methods

3.1. Preparation of Total Genomic DNA

In general, the template DNA must be of good quality. The DNA should have minimal single-stranded nicks and double-stranded breaks within the long PCR template sample, particularly for targets longer than 15 kb. Successful amplification of long targets depends on the presence of intact copies of the single-stranded template, and the intactness of a DNA sample will depend on the method of preparation used. The method of choice may vary with the target and the source of template DNA (*see Notes 1 and 2*). The procedure of Miller et al. (*12*), based on a high-salt extraction of proteins, has been adapted for isolation of DNA from cultured cells (*1*) as described in the following steps:

1. Use a hemacytometer to determine the cell density. Freshly grown cells should be kept on ice until used. Cells may also be frozen in growth medium with 5% (v/v) DMSO for future use.
2. Rinse the cells once with phosphate-buffered saline (PBS) or other balanced salt solution.
3. Thoroughly resuspend a pellet of $1-3 \times 10^6$ cells with 0.6 mL of 10 mM Tris-HCl, pH 8.2, 0.4 M NaCl, 2 mM EDTA in a 1.5-mL microcentrifuge tube.
4. Add 40 μ L of 10% SDS, 100 μ L of 2 mg/mL proteinase K and 7.5 μ L of RNase A.
5. Mix thoroughly by inverting the tube several times or by vortexing briefly.
6. Incubate each sample for at least 1 h, but preferably overnight, in a 50°C shaking water bath to digest the proteins. Complete digestion is critical for efficient recovery of DNA in **Step 7**.
7. Add 0.2 mL saturated NaCl, shake the tube vigorously for 15 s, then microcentrifuge at 10,000 rpm (7000g) for 15 min. The precipitated proteins should form a tight, white pellet. If the pellet is not tight, repeat the microcentrifuge step. Incomplete separation of the proteins from the DNA will result in incomplete proteinase K/SDS treatment.
8. Carefully divide the clear supernatant containing the DNA into two new microcentrifuge tubes, then add a 2 \times volume of 95% (v/v) Ethanol at room temperature to each sample and gently invert each tube several times to mix well.
9. Centrifuge the tubes at 10,000g for 10–15 min to pellet the precipitated DNA.
10. Remove the supernatant and rinse the pellet with 0.5–1.0 mL of 70% (v/v) ethanol. Drain the excess fluid onto absorbent paper and allow the pellet to air dry for 5–10 min before

resuspension (*see step 11*). If the DNA becomes too dry, it will be difficult to resuspend in solution. If the amount of recovered DNA is sufficiently high, the precipitated DNA aggregate can be lifted out of the tube (using the tip of a Pasteur pipet that has been reshaped to form a hook with a Bunsen burner flame), rinsed in 70% ethanol, and air-dried for 5 min.

11. Based on expected yields (e.g., approx 15–20 μg total human genomic DNA from approx 3×10^6 cells), resuspend the genomic DNA in TE buffer to a final concentration of ≤ 300 ng/ μL . Avoid highly concentrated DNA solutions that can be heterogeneous and too viscous to handle easily. The precipitated DNA may require overnight incubation at 4°C to fully solubilize. Do not vortex the sample to aid in resuspending the DNA, because the shearing forces may introduce double-stranded breaks.
12. Determine the DNA concentration from the absorbance reading (between 0.2 and 0.8, for accuracy) at 260 nm. If accurate DNA concentrations are critical, and residual RNA is present, a method based on ethidium bromide fluorescence is described in **Note 10**.
13. Store genomic DNA stocks in TE buffer at 4°C to minimize the introduction of nicks through repeated freeze-thaw cycles. Alternatively, divide each stock across several tubes and freeze. Thaw aliquots as needed and store working stocks at 4°C.

3.2. Analysis of Template Integrity (Optional)

The single-stranded integrity of a template DNA preparation can be qualitatively assessed using alkaline agarose gel electrophoresis, essentially as in **ref. 13**. Use 0.3–0.5% agarose gels to visualize from 2 to > 30-kb single-stranded DNA (*see Note 11*).

1. Prepare a molten agarose solution in 50 mM NaCl, 1 mM EDTA.
2. When the solution has cooled to approx 50°C, add 0.1X-volume of 10X alkaline running buffer. Swirl to mix and then pour the gel.
3. Presoak the solidified gel in 1X alkaline running buffer for 30 min to ensure pH equilibration.
4. Load samples with a 6X Ficoll/Bromocresol green solution (*see Subheading 2.2.4*, or **ref. 13**).
5. Run the gel at 0.5–1.8 V/cm (e.g., 3.5–5 h) using a peristaltic pump to circulate the buffer. The buffer may become quite warm. Both the buffer level and gel position should be checked periodically during the run.
6. Neutralize the gel by gently shaking in 0.1 M Tris-HCl, pH 8.0, 1 mM EDTA for 30 min, and then stain with approx 0.5 $\mu\text{g}/\text{mL}$ ethidium bromide in TAE buffer.

3.3. Primer Design

As with standard PCR, successful primers for the XL PCR process need to be determined empirically. Certain guidelines are to be considered when designing the primer sequences for optimal reaction specificity.

1. Choose primers with higher, balanced melting temperatures (T_m) of approx 62–70°C to allow the use of relatively high annealing temperatures (65–70°C). Sequences of 20–24 bases can work well if the G+C-content is sufficiently high (50–60%), but longer primers (25–30 bases) may be needed if the A+T content is higher.
2. Primer sequences should not be complementary within themselves or to each other. Regions of complementarity, particularly at the 3' end, may result in “primer-dimer” or “smeared” products (*see Note 5*).

3. Primers with balanced T_m s within 1–2°C of each other are more likely to have the same optimal annealing temperature. If the difference in T_m is $\geq 3^\circ\text{C}$, the primer with the higher T_m may anneal to secondary primer sites during incubation at the lower temperature optimal for the second primer and contribute to nontarget products.
4. Primers used in standard PCR, generally designed with annealing temperatures of 5°C below the T_m (14), may also work at higher annealing temperatures, particularly when longer incubation times are used.
5. Primers that can be used for control XL PCR amplifications with human genomic DNA are listed in **Note 4**.
6. Software programs to calculate melting temperatures (include Oligo 5.0 (National Biosciences, Plymouth, MN) and Melt (in BASIC, from J. Wetmur, Mt. Sinai School of Medicine, NY), Primer Premier 5 (Premier Biosoft International), and PrimerSelect (from DNASTar)(see **Note 12**). Wu et al. (15) have developed an algorithm for an Oligonucleotide's "effective priming temperature" (T_p) based on its "effective length" (Ln):

$$T_p \text{ C} = 22 + 1.46 Ln \quad (1)$$

$Ln = 2$ (number of G and C bases) + (number of A and T bases) (5).

7. In any PCR, nontarget sites within the genome that have sufficient complementarity to the 3' end of a primer sequence can be secondary priming sites. As target length increases, however, secondary priming sites within the target are also likely. Shorter secondary products are likely to be more efficiently amplified than a long target, and thus may accumulate at the expense of the desired product. Consequently, specificity at the primer annealing step is critical for successful amplification of long targets.

Whenever possible, candidate primers for XL PCR should be screened against available sequence databases, particularly against any known sequences within the target and related loci (e.g., for gene families). Avoid primers within interspersed repetitive elements, such as Alu sequences (16). The program Oligo (see **Subheading 3.3.6**) can be used to scan a template sequence for potential secondary priming sites. Right Primer 1.01 (BioDisk Software, San Francisco, CA) can be used to screen sequences deposited in GeneBank (National Institutes of Health) for various target genomes to estimate the relative frequency of selected primer sequences. Web sites such as NCBI offers BLAST for doing sequence homology searches online (see **Note 12**).

3.4. PCR Amplification

The GeneAmp XL PCR Kit (see **Note 8**) is designed to use the AmpliMax PCR Gem-facilitated hot start process (17). In this process, a solid wax layer is formed over a subset of PCR reagents (e.g., the lower reagent mix containing 30–50% of the total reaction volume), with the remaining reagents (e.g., the upper reagent mix containing the remaining 50–70% of the total reaction volume) added above the wax layer. During the temperature ramp to the first denaturation step, the wax layer melts and is displaced by the upper reagent mix, which is more dense. Thermal convection suffices to completely mix the combined lower and upper reagent mixes, and the melted wax layer serves as a vapor barrier throughout the PCR amplification. Manual hot start processes can also be used (see **Note 13**), although reproducibility may be lower. The wax-mediated process also helps to minimize contamination between samples.

The protocol below describes XL PCR amplification of a human genomic DNA target (e.g., using the primers RH1024 and RH1053 as in **Note 4**) and a volume ratio of the lower:upper reagent mixes equal to 40:60. Each mix can be assembled as a master mix sufficient for multiple reactions which allows for volume losses during aliquoting (e.g., a 10X master mix for nine reactions).

1. Assemble the lower reaction mix as listed in **Table 1**. For a 100- μ L reaction, place 40 μ L of this lower reagent mix into the bottom of each MicroAmp reaction tube. Avoid splashing liquid onto the tube wall. If any liquid is present on the tube wall, spin the tube briefly in a microcentrifuge.
2. Carefully add a single AmpliWax PCR Gem 100 to each tube containing the lower reagent mix (*see Note 17*). Melt the wax by incubating the reaction tubes at 75–80°C for 3–5 min. Allow the tube to cool to room temperature (or on ice) in order to solidify the wax layer.
3. Assemble the upper reaction mix as listed in **Table 2**. For a 100- μ L reaction, aliquot 60 μ L of this upper reagent mix to each room temperature tube, above the solidified wax layer. Avoid splashing liquid onto the tube walls. If any liquid is present on the tube wall, tap the tube lightly to collect any droplets into the upper reagent layer. Do not spin the tubes in a microcentrifuge, as this will dislodge the wax layer.
4. Amplify the PCR amplifications in a programmable thermal cycler (*see Notes 24*). For the Perkin Elmer GeneAmp System 9600, program the following method:
 - a. HOLD: 94°C for 1 min (reagent mixing and initial template denaturation);
 - b. CYCL: 94°C for 15 s (denaturation, *see Note 25*) and 68°C for 12 min (annealing and extension, *see Notes 26–28* for 20 cycles);
 - c. AUTO: 94°C for 15 s and 68°C for 12 min, adding 15 s per cycle (*see Note 28*) for 17 cycles (*see Note 29*);
 - d. HOLD: 70–72°C for 10 min (final completion of strand synthesis);
 - e. HOLD: 4°C until tubes are removed (use the “Forever” software option).

3.5. Product Analysis

1. To withdraw a sample, gently insert a pipet tip through the center of the solid wax layer to form a small hole. If the tip becomes plugged during this procedure, use a fresh tip to withdraw the reaction sample.
2. The presence of PCR products can be quickly determined using a 0.6% SeaKem GTG agarose gel, run in either TAE or TBE buffer and stained in approx 0.5 μ g/mL ethidium bromide solution (*see Notes 30–32*).
3. High-molecular weight products can be more accurately sized with a 0.3% SeaKem Gold agarose gel (*see Note 11*) run in TAE buffer and stained in approx 0.5 μ g/mL ethidium bromide solution. Depending on the level of resolution needed, run the gel at 7 V/cm for 2 min and then either at 0.8 V/cm for up to 15 h or at 1.5 V/cm for up to 6 h or at 5 V/cm for 1–2 h. Products may also be analyzed by pulse field gel electrophoresis (*see Note 32*).
4. In general, XL PCR products may be analyzed directly by restriction digestion. If further manipulations, such as ligation and cloning, are planned, the reactions should be treated to remove the unincorporated dNTPs, primers and the *rTth* DNA polymerase, XL. If the polymerase and dNTPs are present during restriction digestions, recessed 3' termini may be filled in as they are created, eliminating such sites for ligation. Approaches to remove buffer components and unused primers is to use a Microcon Spin-100 column or a Qiaquick PCR Purification Kit (*see Note 33*).

Table 1
Guidelines for Preparation of the Lower Reagent Mix

Reagent	Volume, 1X Mix, μL	Final Concentration per 100 μL Volume
Sterile Water	14.0–15.2	N/A
3.3X XL Buffer II	12.0	1X
10 mM dNTP Blend	8.0	800 μM (200 μM each dNTP) (see Note 14)
25 mM MgCl_2	4.4–5.2	1.1–1.3 mM (see Note 15)
25 μM Primer RH1024	0.4–0.8	0.1–0.2 μM , 10–20 pmol/reaction (see Note 16)
25 μM Primer RH1053	0.4–0.8	0.1–0.2 μM , 10–20 pmol/reaction (see Note 16)
Total Volume	40.0 μL	

Table 2
Guidelines for Preparation of the Upper Reagent Mix

Reagent	Volume, 1X Mix, μL	Final Concentration per 100 μL Volume
Sterile Water	1.0–40.0	N/A
3.3X XL Buffer II	18.0	1X
2 U/ μL <i>rTth</i> DNA Polymerase, XL	1.0	2 U/reaction (see Note 18)
Human Genomic DNA	1.0–40.0	Up to (see Notes 19–23)
Total Volume	60.0 μL	

4. Notes

1. A number of other methods are available to prepare highly intact DNA. As discussed in **ref. 1**, these include the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN), QIAGEN Genomic-tips (Chatsworth, CA), phenol-extraction (as in **ref. 13**, with high quality phenol), and Megapore dialysis (**18**). The Puregene DNA Isolation Kit is based on high-salt extraction of the proteins. It is important to fully resuspend the cell pellet with the lysis solution for the highest yields of DNA. The method does call for vigorous vortexing to mix the high-salt solution with the cell lysate. The benefits of this 15–20 s vortexing step for efficient precipitation of proteins and subsequent recovery of DNA outweigh the potential for damaging the DNA. In general, however, template stocks for long PCR should be handled gently. The QIAGEN Genomic-tips, based on an anion-exchange resin, are particularly useful for isolation of cosmid DNA without the use of organic solvents. To avoid clogging the Genomic-tip, take care to not load too many cells

and to thoroughly resuspend the dilute sample with buffer QBT, as recommended, before loading the sample onto the column. The Megapore method (18) utilizes dialysis through type HA 0.45 μm membranes (Millipore, Bedford, MA) to remove denatured proteins and cellular debris. This approach is designed to generate very high-molecular weight DNA fragments for cloning.

2. A similar method that was not tested by the authors is the salt/chloroform extraction method by Mullenback et al. (19). The addition of chloroform facilitates the separation between the DNA and protein phases.
3. Purification of primers by polyacrylamide gel electrophoresis does not appear to be generally necessary, although significant levels of truncated sequences could contribute to priming at nonspecific, secondary sites.
4. Three positive control primer sets for XL PCR with human genomic DNA are listed in **Table 3 (20,21)**. These primers have been designed for use with a 67–68°C annealing temperature; at lower temperatures (e.g., 62°C), secondary products will also accumulate. The 16.3-kb human mitochondria genome primers can be used to confirm the presence of low-copy DNA because the mitochondrial genome is present in many copies per cell (22). Longer products are likely to be lower in yield because of lower reaction efficiencies. Certain sequences may also be difficult to amplify regardless of the target length, perhaps because of their base composition and potential to form secondary structures. Consequently, the best controls for a troublesome system may be a series of primer pairs (e.g., one constant, the others variable in position) that define increasingly larger subsets of the target.
5. In utilizing the Differential Display methodology, primers with 3' end modifications and 5' end arbitrary primers can be introduced into long PCR amplifications. This technique should enable isolation of cDNAs that contain both the 3' untranslated transcript region and parts of the 3' end coding region, thus generating highly reproducible primer-set specific fingerprints (23).
6. For subsequent cloning of PCR product, primers can be used to introduce recognition sites for restriction endonucleases. Such sites would be added to the 5'-end of the target-binding sequence, with an additional GC-rich "5'-clamp" of several bases for efficient binding of the restriction endonucleases and subsequent digestion. These nontemplate bases at the 5'-end of the primer should not noticeably affect reaction specificity, but should be accounted for in determining the annealing temperature for the first few cycles (before the entire sequence has been incorporated into the template population).
7. Primers synthesized with a single or double phosphorothioate (sulfur) linkage between 3' most bases has been shown to improve amplification by preventing digestion by the exonuclease activity of Vent DNA polymerase (10).
8. If designing primers with mismatches to the template within the primer sequence, and intend to use a proofreading DNA polymerase, one must consider the position of the mismatch carefully. Due to the tendency of the proofreading polymerase to degrade single-stranded DNA (primer) one base at a time from the 3' end during PCR amplification, some of the primer may be degraded past the positions containing the desired base changes before the primer anneals to the template. However, primers designed and synthesized with phosphorothioate linkages on the last two 3' bases will render the primer extendable but not degradable (11).
9. XL PCR buffer II is composed of tricine to maintain a protective pH during thermal cycling and the cosolvents glycerol and DMSO to effectively lower melting and strand separation temperatures. The 3'-5' exonuclease, "proofreading," activity of *rTth* DNA polymerase, XL facilitates the completion of strand synthesis in each cycle by removing

Table 3
Positive Control Primer Sets

Primer	Sequence	Positions	Amplicon Size
Right-hand primer for the human β -globin gene cluster:			
RH1053	5'-GCACTGGCTTAGGAGTTGGACT-3'	Complements 61986-62007	
Left-hand primer for the human β -globin gene cluster:			
RH1022	5'-CGAGTAAGAGACCATTGTGGCAG -3'	48528-48550	13.5 kb
RH1024	5'-TTGAGACGCATGAGACGTGCAG-3'	44348-44369	17.7 kb
Right-hand primer for the human mitochondrial DNA genome:			
RH1066	5'-TTTCATCATGCGGAGATGTTGGATGG-3'	Complements 14816-14841	
Left-hand primer for the human mitochondrial DNA genome:			
RH1065	5'-TGAGGCCAAATATCATTCTGAGGGGC-3'	15149-15174	16.3 kb

misincorporated nucleotides. The XL PCR amplification protocol uses relatively short denaturation times at moderately high temperatures to minimize template damage while ensuring complete denaturation. The XL PCR protocol also uses a wax-mediated hot-start method, relatively high-annealing temperatures, and reduced enzyme levels to enhance reaction specificity.

10. There are a variety of different enzyme and buffer systems for long PCR that are commercially available, as shown in **Table 4**.
11. Ethidium bromide fluorescence is highly specific for double-stranded DNA. Prepare a standard curve using solutions of 10–400 ng of λ /*Hind*III DNA in 1.2 mL volume of 0.5 μ g/mL ethidium bromide, 20 mM KH_2PO_4 , and 0.5 mM EDTA (pH 11.8–12.0). The fluorescence of aliquots of template DNA can then be compared against these standards. The high pH is critical to minimize any RNA contribution to the fluorescence. An A-4 Filter Fluorometer (Optical Technology Devices, Elmsford, NY) can be used with a bandpass filter (360 nm maximum) for excitation and an interference glass filter at 610 nm for emission spectra (B. Van Houten, University of Texas Medical Branch, Galveston, TX, communication with S. Cheng).
12. Higher molecular weight material will be better resolved on a 3% agarose gel than on a higher percentage gel. Because a 0.3% gel is fragile, use a high-strength agarose, such as Seakem Gold. Chill the gel at 4°C before removing the comb. Submerge the gel in chilled buffer before removing the combs, as the wells may collapse.
13. A few helpful Web addresses that contain tips on primer design, software to design primers, and databases of DNA sequences are:
<http://www.chemie.uni-marburg.de/~becker/prim-gen.html> (PrimerDesign); <http://www.hybsimulator.com/design.html> (HYBsimulator); <http://www.premierbiosoft.com/primerdesign> (Primer Premier); <http://alces.med.umn.edu/VGC.html> (Primer Selection).
<http://www.hgmp.mrc.ac.uk/GenomeWeb/nuc-primer.html> (a collection of sites).
 Note that these addresses were found during a routine search (March 2001) of the Web and have not been evaluated or verified by the author.

Table 4
Some Commercially Available Long PCR Products

Company	Product	Components	Enzyme Sources	For Use With
Roche www.roche.com	Expand High Fidelity PCR System	<i>Taq</i> DNA Polymerase <i>Pwo</i> DNA Polymerase 10X Buffer, MgCl ₂	<i>Thermus aquaticus</i> <i>Pyrococcus woesei</i>	≤5 kb genomic DNA
	Expand Long Template PCR Systems	<i>Taq</i> DNA Polymerase <i>Pwo</i> DNA Polymerase 10X Buffers	<i>Thermus aquaticus</i> <i>Pyrococcus woesei</i>	≤26 kb genomic DNA
	Expand 20 kb PCR System	<i>Taq</i> DNA Polymerase <i>Pwo</i> DNA Polymerase 10X Reaction Buffer MgCl ₂ Buffer Solution	<i>Thermus aquaticus</i> <i>Pyrococcus woesei</i>	≤35 kb genomic DNA
CLONTECH www.clontech.com	Advantage cDNA Polymerase Mix and PCR Kit	KlenTaq polymerase <i>Tth</i> Polymerase TaqStart antibody (Kit includes Buffer, dNTPs, template, primer)	<i>Thermus aquaticus</i> <i>Thermus thermophilus</i>	Up to 10 kb cDNA
	Advantage Genomic Polymerase Mix and PCR Kit	<i>Tth</i> polymerase Proofreading enzyme TaqStart antibody (Kit includes Buffer, dNTPs, template, primer)	<i>Thermus thermophilus</i>	Up to 30 kb genomic DNA
Stratagene www.stratagene.com	<i>TaqPlus</i> Long PCR System	<i>TaqPlus</i> Long Polymerase 10X <i>TaqPlus</i> Long Reaction Buffer	<i>Thermus aquaticus</i> <i>Pyrococcus furiosus</i>	Up to 35 kb
TaKaRa www.takara.com	TaKaRa Ex Taq DNA	Ex <i>Taq</i> polymerase 10X Taq Buffer dNTP Mixture	<i>Taq</i> polymerase Proofreading enzyme	Up to 17 kb genomic DNA
	LA PCR Kit	LA <i>Taq</i> polymerase 10X LA PCR Buffer II Control λ Template Primers, Markers	<i>Taq</i> polymerase Proofreading enzyme	Up to 27 kb genomic DNA

14. For a manual hot start, assemble a single master mix comprised of all but one key reaction component (usually DNA polymerase, Mg(OAc)₂, or dNTPs). Add a 75–80°C Hold (e.g., 5 min for up to 20–25 samples) before the initial denaturation step of the thermal cycling profile (see **Step 4** in **Section 3.4**). Bring all samples to this hold temperature for approx 1 min in the thermal cycler, then add the remaining component to each reaction mix. Note that the volume of this addition must be large enough to minimize pipeting variations, yet small enough to minimize the cooling effect on the reaction mixture already within the tube. If the *rTth* XL DNA polymerase or dNTP blend is the component that is withheld, dilute the components with 1X XL Buffer II to facilitate complete mixing. Remember, to change pipet tips after each component addition to the reaction mix.
15. XL PCR appears to be more sensitive to the integrity of the dNTP solutions than is standard PCR. Stock dNTP solutions should be at pH 7.0–7.5. Reproducibility may be best if these solutions are aliquoted and subjected to a minimal number of freeze-thaw cycles. Changing dNTP stock solutions in an optimized PCR may require slight adjustment of the Mg(OAc)₂ concentration.
16. XL PCR amplifications can be quite sensitive to Mg(OAc)₂ levels, and each new target or primer pair may have a different optimal range for the Mg(OAc)₂ concentration. In general, the acceptable Mg(OAc)₂ window empirically narrows as the starting DNA copy number is decreased and/or the amplicon length is increased. For optimal levels, titrate the Mg(OAc)₂ in increments of 0.1 mM.
17. Repeated freeze–thawing of primers may reduce the efficiency of XL PCR. Make small working stock aliquots of primers and dispose after two or three freeze–thaw cycles.
18. Tap or rotate the tube of AmpliWax PCR Gem 100s to empty a few beads onto either a clean sheet of weighing paper, into a clean weighing boat, or into the tube cap. Use a clean pipet tip to carefully direct a single bead into each tube containing the lower reagent mix.
19. The optimal amount of *rTth* DNA polymerase, XL depends on target length and starting copies. For example, with the Perkin Elmer GeneAmp PCR System 9600, amplification of genomic targets at 1X10⁴ starting copies requires two enzyme units of rTthXL/ per 100 μL reaction, while amplification of Lambda DNA at 1X10⁷ starting copies requires four enzyme units of rTthXL/100 μL reaction. Optimal enzyme concentration can be determined empirically by a titration of rTth DNA Polymerase, XL in increments of 1 U/100 μL reaction.
20. In general, 50–100 ng of total human genomic DNA (with high-average single-stranded molecular weight) will suffice for a 100-μL reaction. Excessive amounts of genomic DNA may contribute to the accumulation of nonspecific products.
21. If the template volume represents a large fraction of the final reaction volume, the DNA should be diluted in water or 10 mM Tris-HCl, 0.1 mM EDTA, to minimize chelation of the Mg(OAc)₂ in the final reaction mix (see **Note 15**).
22. Long PCR amplifications may be more sensitive to potential reaction inhibitors than shorter target amplifications. In such cases, the addition of 50–500 ng/μL nonacetylated BSA may enhance yields, possibly by binding nonspecific inhibitors.
23. Amplification of GC-rich nucleic acids can often be problematic. Reagents including DMSO, glycerol, TMAC, betaine, and 7-deaza GTP have routinely been used to disrupt base pairing or isostabilize the DNA allowing efficient amplification of difficult templates (see **Chapter 1**). A combination of Betaine and DMSO, in particular, has recently been shown to suggest improved processivity. The mixture may increase the resistance of the polymerase to denaturation (25).

24. Recently, *Escherichia coli* exonuclease III has been shown to be helpful for XL PCR amplifications using DNA samples induced with strand breaks and/or apurinic/aprimidinic sites via in vitro treatments such as high temperature (99°C), depurination at low pH and near-UV radiation. Exonuclease III also permitted amplification with DNA aged samples isolated by the phenol/chloroform method (26).
25. XL PCR amplifications are sensitive to the times and temperatures of denaturation and annealing, thus different types of thermal cyclers are likely to require adjustments to the recommended thermal cycling profile. Profiles for all the Perkin-Elmer GeneAmp PCR instruments are provided with the GeneAmp XL PCR Kit.
26. Complete denaturation of the template strands is critical for successful PCR amplifications. The presence of extended GC-rich regions may require use of 95–96°C denaturation temperatures, but the time should be kept short to minimize damage to the single-stranded template and loss of *rTth* DNA polymerase, XL activity over the course of the PCR run.
27. The choice of annealing temperature should be based on the actual primer pair being used (see **Subheading 3.3.**). In general, the highest possible temperature should be used to minimize annealing to secondary priming sites. For reactions in which only the desired product is obtained, a lower annealing temperature may improve product yields.
28. This thermal cycling profile uses two temperatures, compared to the three temperatures typically used in standard PCR. Strand synthesis by *rTth* DNA Polymerase, XL is efficient between 60–70°C. Consequently, when an annealing temperature of approx 62–70°C is used, the same temperature can be set for the extension phase of the cycle. If lower annealing temperatures are necessary, a third temperature step at 65–70°C should be added for efficient completion of strand synthesis, but sufficient time at the annealing temperature must be allowed for efficient priming before the reactions are raised to the extension temperature.
29. In general, use extension times sufficient for 30–60 s/kb of the target. In the two-temperature cycling profile, this applies to the total annealing-plus-extension time. As product accumulates, the ratio of template to polymerase molecules will increase, and the overall reaction efficiency may decrease. The AUTO feature of the Perkin Elmer GeneAmp PCR System 9600 allows the extension time to be incrementally increased during late cycles of the run, which helps maintain reaction efficiency. The potential disadvantage of using very long extension times initially is that during early cycles, excessively long extension times may permit nonspecific products to accumulate.
30. The optimal number of cycles will depend on the initial copy number of the template and the reaction efficiency. Reaction efficiency is generally higher for shorter (5–10 kb) vs longer (20–30 kb) targets. For example, from approx 104 copies of human genomic DNA (37 ng, in a 50- μ L reaction), the 16.3-kb multicopy mitochondrial genome target can be readily amplified with a total of 30–35 cycles, whereas the 17.7-kb single-copy β -globin target requires at least 35–37 cycles. The number of AUTO cycles (9600) will also depend on the initial copy number of the template and the length of the target.
31. Amplified target bands may be identified by size (gel mobility relative to standards), a Southern blot analysis (as in **ref. 13**), and/or analytical restriction digests. High-molecular-weight smears tend to reflect high levels of nonspecific synthesis, as in the cases of excess *rTth* DNA polymerase, XL or excess Mg(OAc)₂. Excessively long extension times or too many cycles of amplification can also result in the appearance of high-molecular weight bands or nonspecific smears. Low-molecular weight secondary bands may reflect insufficient specificity, and may be reduced by the use of a higher annealing temperature,

and/or lower concentrations of template, primers, or *rTth* DNA polymerase, XL. If accumulation of products other than the desired product is significant, the best solution may be to redesign one or both primers. Absence of any detectable product from a known template may indicate that the denaturation temperature was either too low for the template or too high for the DNA polymerase, XL; that the annealing temperature was too high for the primer pair; or that either the polymerase or the $\text{Mg}(\text{OAc})_2$ concentration was too low. If a Southern blot analysis or reamplification using primers located within the original target (nested PCR) reveals that the desired product is present at a very low level, the explanation may be that too few cycles of amplification were used for the starting target copy number (see **Notes 20** and **21**).

32. Optimization of the denaturation or annealing temperatures should be made in increments of 1–2°C. Adjustments to enzyme concentration can be made in increments of 0.5–1 U/100 μL reaction. Optimization of the $\text{Mg}(\text{OAc})_2$ concentration should be carried out in increments of 0.1 mM (see **Notes 27** and **28**).
33. Carryover contamination (see Chapter 1) and dNTP stock solutions of poor quality can both significantly reduce the apparent optimal range for the $\text{Mg}(\text{OAc})_2$ concentration. These problems may not be observed initially, but may become apparent during late amplifications with targets and primers previously observed to work well.
34. Resolution of high-molecular weight DNA (>50 kb) is best achieved using pulse field, for example, field inversion electrophoresis (27). One such system is made by Hoefer (San Francisco, CA).
35. The 3' ends of the PCR may have an additional one or two nontemplated nucleotides (28,29). Although the 3'–5' exonuclease activity present in *rTth* DNA Polymerase, XL would be expected to remove these 3'-additions, there is evidence that a certain fraction of XL PCR product molecules have an additional 3'-A (30). This fraction is likely to be less than that observed in standard PCR with *Taq* DNA polymerase, and using methods, such as the TA Cloning Kit (Invitrogen, San Diego, CA), which take advantage of the 3'A addition may therefore be inefficient. If necessary, the 3'-additions can be removed by incubation with Pfu DNA polymerase (31) or with Klenow fragment of *E. coli* DNA polymerase I (24,29).

References

1. Cheng, S., Chen, Y., Monforte, J. A., Higuchi, R., and Van Houten, B. (1995) Template integrity is essential for PCR amplification of 20- to 30-kb sequences from genomic DNA. *PCR Meth. Appl.* **4**, 294–298.
2. Cheng, S. (1995) Longer PCR amplifications, in *PCR Strategies* (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds.) Academic, San Diego, CA, pp. 313–324.
3. Cheng, S., Chang, S.-Y., Gravitt, P., and Respass, R. (1994) Long PCR. *Nature* **369**, 684–685.
4. Cheng, S., Fockler, C., Barnes, W. M., and Higuchi, R. (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. USA* **91**, 5695–5699.
5. Barnes, W. M. (1994) PCR amplification of up to 35 kb with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
6. Brugnoli, R., Morandi, L., Brambati, B., Briscioli, V., Cornelio, F., and Mantegazza, R. (1998) A new non-radioactive method for the screening and prenatal diagnosis of myotonic dystrophy patients. *J. Neurol.* **245**, 289–293.

7. Salminen, M. O., Koch, C., Sanders-Buell, E., Ehrenberg, P. K., Michael, N. L., Carr, J. K., Burke, D. S., and McCutchan, F. E. (1995) Recovery of virtually full length HIV-1 provirus of Diverse Subtypes from primary virus cultures using the polymerase chain reaction. *Virology* **213**, 80–86.
8. Van Houten, B., Cheng, S., and Chen, Y. (2000) Measuring gene-specific nucleotide excision repair in human cells using quantitative amplification of long targets from nonogram quantities of DNA. *Mutation Research* **460**, 81–94.
9. Landre, P. A., Gelfand, D. H., and Watson, R. M. (1995) The use of cosolvents to enhance amplification by the polymerase chain reaction, in *PCR Strategies* (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds.) Academic, San Diego, CA, pp. 3–16.
10. Skera, A. (1992) Phosphorothioate primers improve the amplification of DNA sequences by DNA polymerase with proofreading activity. *Nucl. Acids Res.* **20**, 3551–3554.
11. de Noronha, C. and Mullins, J. (1992) *PCR Meth. Appli.* **2**, 131–136.
12. Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl. Acids Res.* **16**, 1215.
13. Sambrook, J., Fitch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 6.20, 6.21, 9.16-9.19, 9.34-9.57, and B.23-B.24.
14. Innis, M. A. and Gelfand, D. H. (1990) Optimization of PCRs, in *PCR Protocols* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds) Academic, San Diego, CA, pp. 3–12.
15. Wu, D. Y., Ugozzoli, L., Pal, B. K., Qian, J., and Wallace, R. B. (1991) The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA Cell Biol.* **10**, 233–238.
16. Schmid, C. W. and Jelinek, W. R. (1982) The Alu family of dispersed repetitive sequences. *Sciences* **216**, 1065–1070.
17. Chou, Q., Russell, M., Birch, D. E., Raymond, J., and Block, W. (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* **20**, 1717–1723.
18. Monforte, J. A., Winegar, R. A., and Rudd, C. J. (1994) Megabase genomic DNA isolation procedure for use in transgenic mutagenesis assays. *Environ. Mol. Mutagen.* **23**, 46.
19. Mullenbach, R., Lagoda, P. J. L., and Welter, C. (1989) Technical Tips: an efficient salt-chloroform extraction of DNA from blood and tissues. *Trends in Gen.* **5**, 391.
20. Kolmodin, L., Cheng, S., and Akers, J. (1995) GeneAmp XL PCR Kit, in *Amplifications: A Forum for PCR Users* (The Perkin Elmer Corporation), Issue 13.
21. Cheng, S., Higuchi, R., and Stoneking, M. (1994) Complete mitochondrial genome amplification. *Nature Gen.* **7**, 350, 351.
22. Robin, E. D. and Wong, R. (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J. Cell Physiol.* **136**, 507–513.
23. Jurecic, R., Nachtman, R. G., Colicos, S. M., and Belmont, J. W. (1998) Identification and Cloning of Differentially Expressed Genes by Long-Distance Differential Display. *Anal. Biochem.* **259**, 235–244.
24. Scharf, S. J., (1990) Cloning with PCR, in *PCR Protocols* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.) Academic, San Diego, CA, pp. 84–91.
25. Fromenty, B., Demeilliers, Mansouri, A., and Pessayre, D. (2000) Escherichia coli exonuclease III enhances long PCR amplification of damaged DNA templates. *Nucl. Acids Res.* **28**, 50.

26. Carle, G. F., Frank, M., and Olson, M. V. (1989) Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. *Science* **232**, 65–68.
27. Hengen, P. N. (1997) Methods and reagents: optimizing multiplex and LA-PCR with betaine. *TIBS* **22**, 225,226.
28. Clark, J. M. (1988) Novel nontemplated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677–9686.
29. Hu, G. (1993) DNA Polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol.* **12**, 763–770.
30. Stewart, A. C., Gravitt, P. E., Cheng, S., and Wheeler, C. M. (1995) Generation of entire human papilloma virus genomes by long PCR: frequency of errors produced during amplification. *Genome Res.* **5**, 79–88.
31. Costa, G. L. and Weiner, M. P. (1994) Protocols for cloning and analysis of blunt-ended PCR-generated DNA fragments. *PCR Meth. Appl.* **3**, S95–S106.

Coupled One-Step Reverse Transcription and Polymerase Chain Reaction Procedure for Cloning Large cDNA Fragments

Jyrki T. Aatsinki

1. Introduction

Although *Thermus aquaticus* (*Taq*) and *Thermus thermophilus* (*Tth*) DNA polymerases have the ability to reverse transcribe RNA to complementary DNA (cDNA) and subsequently amplify the target cDNA, they are not usually the first choices for reverse transcription-polymerase chain reactions (RT-PCR) (1–4). Because they only synthesize short cDNA fragments, their use is not widespread. In general, avian myeloblastosis virus (AMV), or moloney murine leukemia virus (M-MLV) reverse transcriptases (RTs) are used to reverse transcribe RNA to cDNA templates for subsequent PCR. Previous coupled methods are also unable to amplify large cDNA fragments and, thus, they are suitable only for the detection of gene expression (5–8). The one-step RT-PCR procedure presented here was developed to amplify large cDNA fragments suitable for cloning full-length open reading frames (ORFs) encoding rat LH/CG receptor isoforms (9–12). As we all know, the construction of clones, including library screening and restriction mapping, by conventional cloning methods is very laborious and difficult.

The one-step RT-PCR procedure was first optimized for its specificity. Low concentrations of dNTP (0.2 mM of each), MgCl₂ (1.5 mM), and primer (0.1 μM of each) and a relatively high annealing temperature (55°C) were used, because these conditions have been found to enhance specific amplification. The commercially available PCR buffer (10 mM Tris-HCl, pH 8.4, and 50 mM KCl) was found to be suitable for primer extension by AMV-RT although it differed in its constituents from the recommendations of the manufacturer. To assure that primer extension was completed, long extension times were used, both in reverse transcription and PCR (60 min and 10 min + 59 s/cycle, respectively). Possible aggregates and secondary structures were eliminated by denaturing both primers and RNA at 65°C, for 15 min, before starting the amplification. Subsequent to a 1 h incubation at 42°C, the temperature was raised to

From: *Methods in Molecular Biology*, Vol. 192: *PCR Cloning Protocols*, 2nd Edition
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

95°C for 3 min to dissociate RNA-cDNA hybrids. Finally, RT-PCR products could be easily cloned for in vitro translation studies and for transfections in different cell lines, because suitable restriction enzyme sites were incorporated at both ends. Examples of other potential applications of the present coupled RT-PCR procedure, in addition to cDNA cloning and the detection of gene expression, include the quantitation of mRNA and clinical diagnostics (e.g., the detection of viral RNA, tumor cells, parasites, and genetic disorders). Our procedure has been used for over a decade and it still meets current demands. Although commercial applications have become available, the author thinks that the coupled one-step RT-PCR procedure was the first and still the best available procedure to produce large RT-PCR products sensitively and reliably.

2. Materials

2.1. Coupled One-Step RT-PCR

1. Total RNA isolated by the TRIZOL[®] following the instructions of manufacture (Gibco-BRL, Gaithersburg, MD).
2. Oligonucleotide primers, typically 32–35 nucleotides long and with a 40–60% G + C content, are designed to have internal unique restriction sites and an additional 8–9 nucleotide complementary sequence at the 5' end of the restriction sites. These added nucleotides are important for helping restriction enzymes to cleave the RT-PCR product. When creating a new restriction site, select a sequence that requires as few changes as possible. Primer-dimer formation during PCR is best avoided by using primers having noncomplementary 3' ends.
3. Ribonuclease inhibitor, Inhibit-ACE (5 prime-3 prime, Boulder, CO).
4. 10X PCR buffer: 500 mM KCl, 100 mM Tris-HCl, pH 8.4 (Gibco-BRL).
5. 50 mM MgCl₂ (Gibco-BRL).
6. Deoxynucleotide triphosphates (dNTPs) (Pharmacia, Uppsala, Sweden).
7. AMV-RT (Promega, Madison, MI).
8. *Taq* DNA polymerase, recombinant (Gibco-BRL).
9. DNA Engine[™] Peltier Thermal Cycler (MJ Research, Inc., Watertown, MA).
10. 200 µL-thin-wall PCR tubes.
11. 6X loading dye solution (MBI Fermentas, Vilnius, Lithuania).
12. DNA electrophoresis size markers: GeneRuler[™] Ladder Plus, ready-to-use (MBI Fermentas).
13. Reagents and supplies for agarose gel electrophoresis: *See* preparation of mixtures and use of equipments in the laboratory manual (13).

2.2. Cloning of RT-PCR Product

1. Wizard PCR product purification column (Promega).
2. Restriction enzymes (e.g., *Bam*HI and *Eco*RI) (Pharmacia).
3. Plasmid DNA, pUCBM20 (Boehringer Mannheim GmbH, Mannheim, Germany)
4. *Escherichia coli* (*E. Coli*) JM109 strain.
5. T4 DNA ligase (high concentration; 5 U/µL) and supplied 10X ligase buffer (Boehringer Mannheim GmbH).
6. Alkaline phosphatase and supplied 10X alkaline phosphatase buffer (Promega).
7. Reagents for DNA precipitation: 4 M NaCl; absolute ethanol; 70% ethanol.
8. Reagents and supplies for molecular cloning: 1 M CaCl₂; isopropyl-β-D-thio-galactoside (IPTG); 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal); Luria-Bertani medium; Luria-Bertani medium plates. *See* preparation of mixtures in the laboratory manual (13).

3. Methods

3.1. Coupled One-Step RT-PCR

1. Dilute primers to a concentration of 20 ng/ μ L in sterile distilled H₂O. Add 2.5 μ L of each primer to a reaction tube containing a volume of sterile distilled H₂O sufficient to bring the total reaction volume to 50 μ L after the addition of the rest of the reactants from **steps 3 and 4**. Add 0.5 U Inhibit-ACE and incubate for 20–30 min at room temperature.
2. Prepare stock mixtures of 10X PCR buffer (containing a final concentration of 1.5 mM MgCl₂ in 1X PCR buffer) and 10 mM dNTPs (2.5 mM of each dNTP) and incubate for 20–30 min at room temperature after adding 1 U Inhibit-ACE/100 μ L of stock mixture.
3. Add total RNA (100 pg–10 μ g) to the reaction tube (*see step 1*), denature at 65°C for 15 min and cool to 4°C using a programmed DNA thermal cycler (*see Note 1*).
4. Add 6.5 μ L of 10X PCR buffer and 4 μ L of 10 mM dNTPs from **step 2** to the reaction tube. Add 10 U AMV-RT and 2.5 U DNA polymerase, mix carefully, and collect by brief centrifugation.
5. Incubate at 42°C for 1 h to allow RT.
6. **Step 5** is linked to PCR cycles. Initial denaturation at 95°C for 3 min, followed by 30 cycles consisting of denaturation at 95°C for 1 min, primer annealing at 55°C for 2 min, and extension at 72°C for 10 min + 59 s/cycle.
7. Take 5–25 μ L of the RT-PCR product, add gel loading buffer, and size fractionate on a 1% ethidium bromide-stained agarose gel. Gel electrophoresis of the RT-PCR products are shown in Fig. 1 (*see Notes 2–6*).

3.2. Multiplex Coupled One-Step RT-PCR

1. Primers for low-copy number mRNA are added as aforementioned in **Subheading 3.1., step 1**, where 50 ng of each of the LH/CG receptor primer was added to a reaction tube. Primers for more abundant mRNAs are reduced such as 25 ng of both carbonic anhydrase II primers and 1 ng of both β -actin primers are sufficient to produce visible bands on the ethidium bromide-stained agarose gel (*see Fig. 1, lane 5*). Mix all the primers to a reaction tube containing a volume of sterile distilled H₂O sufficient to bring the total reaction volume to 50 μ L after the addition of the rest of the reactants from **steps 3 and 4**. Add 0.5 U Inhibit-ACE and incubate for 20–30 min at room temperature (*see Note 7*).
2. Follow **step 2** from **Subheading 3.1.**
3. Follow **step 3** from **Subheading 3.1.**
4. Add 6.5 μ L of 10X PCR buffer and 4 μ L of 10 mM dNTPs from **step 2** to the reaction tube. Add 20 U AMV-RT and 2.5 U DNA polymerase, mix carefully, and collect by brief centrifugation.
5. Follow **steps 5–7** from **Subheading 3.1.**

3.3. Cloning of RT-PCR Product

1. Purify the RT-PCR product using a Wizard PCR product purification column following the manufacturer's instructions.
2. Digest the purified RT-PCR product with restriction enzymes using a several-fold excess of enzyme and long incubation times. Usually, 50 U of restriction enzyme can be added at the beginning of digestion and a further 20 U during the incubation, which can be done overnight (*see Note 8*).
3. Purify the sample as in **step 1**.
4. Precipitate the sample by adding 1/20 vol of 4 M NaCl and 2 vol of cold absolute ethanol, allow to stand overnight at –20°C, or for 30 min at –80°C. Centrifuge at 12 000g for

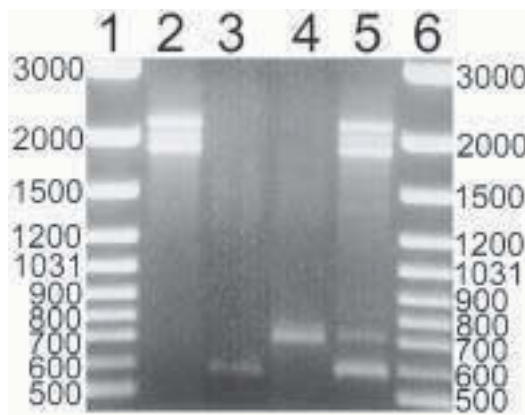


Fig. 1. Coupled one-step RT-PCR amplification of three different cDNA species from rat ovary total RNA. Lane 2 shows LH/CG receptor isoforms amplified using CAATT TTGGAATTCTAGTGAGTTAACGCTCTCG as the reverse primer, GGGAGCTCGA ATTCAGGCTGGCGGGCCATGGGGCGG as the forward primer (mismatched nucleotides, to create internal restriction sites, are underlined), and 5 μ g of total RNA as the template. Lane 3 shows the carbonic anhydrase II RT-PCR product amplified using GAGCACTATCCAG GTCACACATTCCAG as the reverse primer, ACTGGCACAAGGAGTTCCCCATTGCCA as a forward primer, and 100 ng of total RNA as the template. Lane 4 shows the β -actin RT-PCR product amplified using GATGCCACAGAATTCCATACCCAGGAAGGAAGGC as the reverse primer, GCCGCCCTAGGATCCAGGGTGTGATGGTGGGTAT as the forward primer (mismatched nucleotides, to create internal restriction sites, are underlined), and 100 pg of total RNA as the template. Lane 5 shows the multiplex RT-PCR, where three different cDNA species were simultaneously amplified in same reaction tube using the above mentioned primers, and 5 μ g of total RNA as the template. Other conditions for multiplex RT-PCR are as mentioned in **Subheading 3.2**. 25 μ L of RT-PCR products were size-fractionated on a 1% ethidium bromide-stained agarose gel. Lanes 1 and 6 show 10 μ L of the molecular weight standards (GeneRuler™ Ladder Plus).

20 min at +4°C and discard the supernatant. Add 1 mL of 70% ethanol to the pellet and centrifuge as before. Dissolve the dried pellet in 20 μ L of sterile distilled H₂O.

5. Prepare plasmid DNA by digesting with suitable restriction enzymes for 2–3 h using 10 U/1 μ g DNA. Dephosphorylate the DNA according to the manufacturer's instructions if using a single restriction enzyme. Purify the sample as described in **steps 1 and 4**.
6. Set up three ligation mixtures using 3:1, 1:1, and 1:3 molar ratios of insert and plasmid DNA. Add sterile distilled H₂O to final volume of 8 μ L, heat at 45°C for 5 min and cool to 16°C. Add 1 μ L of 10X ligation buffer, 1 μ L of high concentration T4 DNA ligase (5 U/ μ L), and incubate overnight at 16°C.
7. Transform competent cells using standard laboratory protocols (**13**). Pick up positive clones and analyze them by restriction digestion and agarose gel electrophoresis. Determine the nucleotide sequences of some clones by the double-stranded dideoxy sequencing method (**14**).

4. Notes

1. It is always necessary to test the amount of RNA for optimal amplification. If the total RNA contains high amounts of the target mRNA, efficient amplification is obtained with picogram amounts of RNA. On the other hand, if the total RNA contains only very small amounts of the target mRNA, up to 10 μg of the total RNA can be used to obtain an efficient amplification.
2. If no specific bands are visible on the ethidium bromide-stained agarose gel, use a gradually increasing amount of reverse transcriptase. Do not use excess amounts of *Taq* DNA polymerase, because this has been reported to lower the amount of specific RT-PCR product (10,12).
3. If no specific bands are visible on the ethidium bromide-stained agarose gel after optimization, prepare a Southern blot (15) of the gel. Hybridize with an appropriate probe that does not contain overlapping sequences with the primers used for the RT-PCR.
4. If, after Southern blotting, a specific hybridization signal is obtained, use nested PCR to produce visible bands on the ethidium bromide-stained agarose gel. Prepare a new pair of primers located further outside the region where the first primer set was designed. This new primer set does not need modifications (e.g., restriction sites) and can be shorter (about 22–25 nucleotides). Use these new primers in RT-PCR; take 1–5 μL of the RT-PCR product and use it as a template and the original modified oligonucleotides as primers for the second round of PCR.
5. If no specific bands are seen after procedures described in Notes 1–4, check the total RNA used in the experiments. Use primers of abundant mRNA (e.g., β -actin), instead of your primers, in the coupled one-step RT-PCR procedure. A positive signal in the control reaction leaves only two possibilities for explaining negative results: the sample RNA does not contain the target template RNA, or the primers anneal inefficiently to the template RNA. Try a new pair of primers located in a different region of the cDNA, because primers are sometimes chosen in a region of secondary structure, causing difficulties in priming.
6. Negative controls in RT-PCR should be done to eliminate the possibility of potential DNA contamination. Prepare two control samples following the above procedure, but omit the template RNA in one control and omit RT in the second.
7. Test empirically the amount of primers used in multiplex coupled one-step RT-PCR. Primers for low-copy number mRNAs are added as in basic procedure and primers for more abundant mRNAs are reduced until all the RT-PCR products are visible on the ethidium bromide-stained agarose gel. Excess amount of primers for abundant mRNAs compete efficiently for enzymes resulting no visible RT-PCR product of rare mRNAs.
8. The procedure for directional cloning of RT-PCR products is also included in this chapter because it has been problematic for many laboratories. In the present procedure, the use of a several-fold excess of restriction enzymes and long incubation times are critical for optimal results. Commercial methods for cloning PCR products (e.g., T/A cloning and blunt end DNA ligation kits) are also recommended, although they are expensive to use.

Acknowledgments

This work was performed at the Department of Anatomy and Cell Biology, and the Institute of Dentistry, University of Oulu, Finland. The support of grants from the Finnish Cultural Foundation and the Memorial Foundation of Maud Kuistila are gratefully acknowledged.

References

1. Jones, M. D. and Foulkes, N. S. (1989) Reverse transcription of mRNA by *Thermus aquaticus* DNA polymerase. *Nucl. Acids Res.* **17**, 8387–8388.
2. Shaffer, A. L., Wojnar, W., and Nelson, W. (1990) Amplification, detection, and automated sequencing of gibbon interleukin-2 mRNA by *Thermus aquaticus* DNA polymerase reverse transcription and polymerase chain reaction. *Analyt. Biochem.* **190**, 292–296.
3. Tse, W. T. and Forget, B. G. (1990) Reverse transcription and direct amplification of cellular RNA transcripts by *Taq* polymerase. *Gene* **88**, 293–296.
4. Myers, T. W. and Gelfand, D. H. (1991) Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* **30**, 7661–7666.
5. Goblet, C., Prost, E., and Whalen, R. G. (1989) One-step amplification of transcripts in total RNA using the polymerase chain reaction. *Nucl. Acids Res.* **17**, 2144.
6. Singer-Sam, J., Robinson, M. O., Bellvé, A. R., Simon, M. I., and Riggs, A. D. (1990) Measurement by quantitative PCR of changes in HPRT, PGK-1, PGK-2, APRT, MTase, and Zfy gene transcripts during mouse spermatogenesis. *Nucl. Acids Res.* **18**, 1255–1259.
7. Zafra, F., Hengerer, B., Leibrock, J., Thoenen, H., and Lindholm, D. (1990) Activity dependent regulation of BDNF and NGF mRNAs in the rat hippocampus is mediated by non-NMDA glutamate receptors. *EMBO J.* **9**, 3545–3550.
8. Wang, R.-F., Cao, W.-W., and Johnson M. G. (1992) A simplified, single tube, single buffer system for RNA-PCR. *BioTechniques* **12**, 702–704.
9. Aatsinki, J. T., Pietilä, E. M., Lakkakorpi, J. T., and Rajaniemi, H. J. (1992) Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an extensive alternative splicing of the primary transcript. *Mol. Cell. Endocrinol.* **84**, 127–135.
10. Aatsinki, J. T., Lakkakorpi, J. T., Pietilä, E. M., and Rajaniemi, H. J. (1994) A coupled one-step reverse transcription PCR procedure for generation of full-length open reading frames. *BioTechniques* **16**, 282–288.
11. Aatsinki, J. T. (1997) Coupled one-step reverse transcription and polymerase chain reaction procedure for cloning large cDNA fragments, in *Methods in Molecular Biology*, vol. 67, *PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering* (White, B.A., ed.) Humana, Totowa, NJ, pp. 55–60.
12. Aatsinki, J. T., Lakkakorpi, J. T., Pietilä, E. M., and Rajaniemi, H. J. (1998) A coupled one-step reverse transcription PCR procedure for generation of full-length open reading frames, in *BioTechniques® Update Series, The PCR Technique: RT-PCR* (Siebert, P., ed.), Eaton, Natick, MA, pp. 261–268.
13. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
15. Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.

Long Distance Reverse-Transcription PCR

Volker Thiel, Jens Herold, and Stuart G. Siddell

1. Introduction

Polymerase chain reaction (PCR) has been applied to the amplification of long DNA fragments from a variety of sources, including genomic, mitochondrial, and viral DNAs (1–5). We have adapted the concept of long PCR technology to reverse-transcription (RT) PCR (6). Here, we describe the parameters critical in producing RT-PCR products of up to 20 kbp. The nature of RT-PCR requires the synthesis of a cDNA by RT prior to its amplification in the PCR reaction. Thus, we focus on the three steps of RT-PCR: the preparation and requirements of the RNA template, the reverse transcription reaction, and the amplification of the cDNA by PCR.

To carry out these studies, we used the genomic RNA of the human coronavirus HCoV 229E as template (7). The HCoV 229E genomic RNA has a length of ca. 27,000 nucleotides and the homogeneity of the RNA can be readily assessed by electrophoresis and hybridization analysis (8). HCoV 229E genomic RNA has two major advantages for the studies reported here. First, as a viral RNA, it is relatively abundant in the infected cell. Second, coronaviruses are positive strand RNA viruses and the genomic RNA has a 3' polyadenylate tract that can be used for affinity chromatography (9).

In **Subheading 3.1.**, we describe a simple and fast technique to purify poly(A)-containing RNA from tissue culture cells. In fact, we believe that the integrity and purity of the RNA template is the most critical parameter when the RT-PCR amplification of sequences more than 5 kb in length is desired (6) (see **Note 4, Fig. 1**). Depending on the source of the RNA template, a method of preparation should be chosen that minimizes degradation of the RNA. In our hands oligo (dT)-based affinity chromatography with magnetic beads has proven to be reliable for the isolation of poly(A)-RNA that can be used to produce cDNA of more than 20 kb by RT.

The conditions of the reverse transcription reaction strongly influence the outcome of the subsequent PCR. Reverse transcription reactions have been performed using the RNase H-deficient reverse transcriptase, SuperScript II and the cDNA has been used for PCR amplification without further purification. In general, an RT-primer should

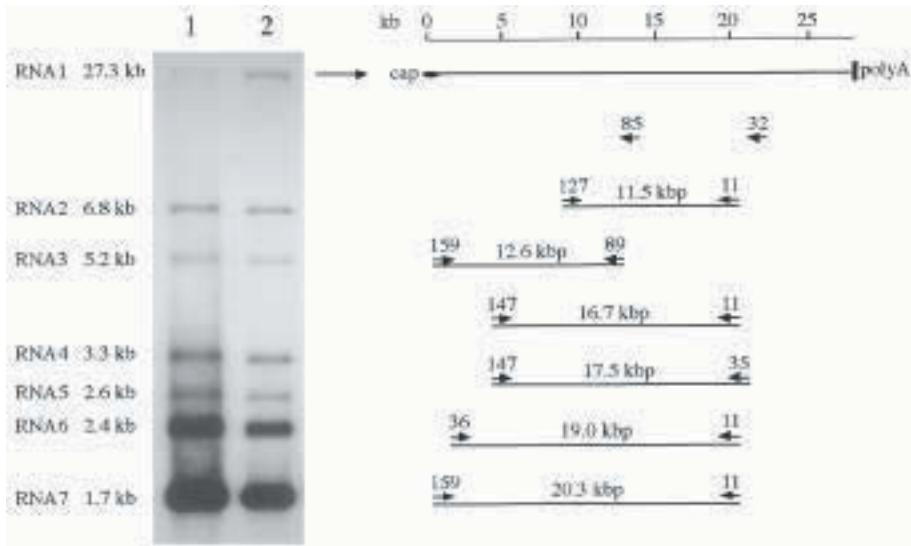


Fig. 1. Northern hybridization analysis of HCoV 229E poly(A)-containing RNA isolated from infected MRC-5 cells. The material shown in lane 1 was prepared using poly(U)-Sephrose (6). The material shown in lane 2 was prepared using oligo(dT)₂₅ magnetic beads as described in **Subheading 3.1**. The poly(A)-containing RNAs were separated by gel electrophoresis and the viral mRNAs were hybridized to a ³²P-(5'end)-labeled oligonucleotide (5'AGAAACTTCTCACGCACTGG 3'). Also shown is the relationship of the HCoV 229E genomic RNA template, oligonucleotide primers and RT-PCR products. The oligonucleotides are indicated as arrows with their orientation and position relative to the HCoV 229E genomic RNA. The expected sizes of the RT-PCR products are indicated.

be used that is highly specific and great care should be taken to adjust the optimal concentration of the RT-primer in the reverse transcription reaction. In our experience, the major problem that arises is “less stringent” priming during the RT reaction. The fortuitous cDNAs that are synthesized and the small amounts of RT-primer that are carried over into the PCR are responsible for most of the background amplification products observed (6).

Finally, the optimal conditions, regarding the amount of cDNA template, the choice of PCR primers and the cycle profile of the PCR, have to be determined. In this respect, the considerations that apply to the PCR amplification of dsDNA templates are equally applicable to amplification of cDNA produced by reverse transcription.

2. Materials

1. Phosphate-buffered saline (PBS).
2. Lysis buffer: 10 mM Tris-HCl, pH 7.5, 140 mM NaCl, 5 mM KCl, 1% Nonidet P-40 (NP40).
3. Oligo(dT)₂₅-Dynabeads (Dynal) (3.3×10^8 beads/mL).
4. Dynal Magnetic Particle Concentrator (Dynal).

5. 2X binding buffer: 20 mM Tris-HCl, pH 7.5, 1 M LiCl, 2 mM ethylenediamine tetraacetic acid (EDTA), 1% sodium dodecylsulfate (SDS).
6. Wash buffer: 10 mM Tris-HCl, pH 7.5, 150 mM LiCl, 1 mM EDTA, 1 mM EDTA pH 7.5.
7. SuperScript II reverse transcriptase (Life Technologies).
8. 5X first-strand buffer (Life Technologies).
9. 10 mM dNTPs (10 mM of each dNTP).
10. 0.1 M dithiothreitol (DTT).
11. RNasin (50 U/ μ L) (Pharmacia).
12. Thin-wall PCR tubes.
13. Elongase Enzyme Mix (Life Technologies).
14. PCR buffer B (Life Technologies).

3. Methods

3.1. The RNA Template: Preparation of Poly(A)-Containing RNA Using Oligo(dT)₂₅-Dynabeads

The protocol below describes the isolation of poly(A)-containing RNA from a confluent layer of adherent cells grown in a 175 cm² tissue culture flask (*see Note 1*).

1. Wash 200–500 μ L oligo(dT)₂₅-Dynabeads twice with 2X binding buffer using an appropriate Dynal Magnetic Particle Concentrator (*see Note 2*) and resuspend the beads in 1.5 mL 2X binding buffer.
2. Wash the cells twice with ice-cold PBS and then scrape in 10 mL ice-cold PBS.
3. Pellet the cells at 1000g for 2 min at 4°C.
4. Resuspend the cell pellet in 1.5 mL ice-cold lysis buffer and incubate for 30 s on ice.
5. Centrifuge the cell lysate at 1500g for 1 min at 4°C to remove nuclei.
6. Mix the supernatant with oligo(dT)₂₅-Dynabeads resuspended in 1.5 mL 2X binding buffer (**Step 1**) and incubate for 5 min at 23°C. Gently mix the sample every 1–2 min.
7. Wash the oligo(dT)₂₅ magnetic beads twice with wash buffer using an appropriate Dynal Magnetic Particle Concentrator (*see Note 3*).
8. Completely remove the wash buffer and add 50–100 μ L 2 mM EDTA, pH 7.5.
9. Transfer the solution into a 1.5-mL Eppendorf tube and incubate for 2 min at 65°C to elute the bound poly(A)-RNA.
10. Take the supernatant containing the poly(A)-RNA, add 5 μ L RNasin and store at –70°C in aliquots (*see Note 4*).
11. Regenerate the oligo(dT)₂₅-Dynabeads according to the manufacturer's instructions (optional; *see Note 5*).

3.2. The RT Reaction

1. Add the following components to a volume of 19 μ L (*see Note 6*):
 - RNase-free water
 - 4 μ L 5X first strand buffer;
 - 2 μ L 10 mM dNTPs (10 mM of each dNTP);
 - 2 μ L 0.1 M DTT;
 - 0.5 μ L RNasin (50 U/ μ L);
 - 5–15 pmol reverse transcription primer (*see Note 7*);
 - 10–500 ng poly(A)-RNA (0.5–3 μ L of poly(A)-containing RNA prepared as described earlier).

2. Incubate for 2 min at 42°C (*see Note 8*).
3. Add 1 μL (200 U) SuperScript II reverse transcriptase.
4. Incubate for 60–90 min at 42°C (*see Note 9*).
5. Incubate for 2 min at 94°C and chill on ice (*see Note 10*). Store at -20°C until use.

3.3. The PCR Reaction

To amplify long DNA fragments from cDNA templates the basic principles of long PCR technology are applicable (*see Note 11*). The protocol below describes a typical reaction using the Elongase Enzyme Mix (Life Technologies). In addition, we provide a list of PCR cycle profiles (**Fig. 2**) and corresponding oligonucleotide primers (**Fig. 1, Table 1**) that have been used to produce RT-PCR products of 11.5–20.3 kbp in length.

1. Set up the PCR reaction mix in thin-wall PCR tubes on ice [50 μL volume; final concentrations: 60 mM Tris-SO₄ (pH 9.1), 18 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.2 mM dNTPs, 0.2–0.4 μM PCR primer]:
 - 10 μL buffer B (Life Technologies) (*see Note 12*);
 - 1 μL 10 mM dNTPs (10 mM of each dNTP);
 - 1–2 μL forward PCR primer (10–40 pmol);
 - 1–2 μL reverse PCR primer (10–40 pmol);
 - 0.2–3 μL RT-reaction;
 - 1 μL Elongase Enzyme Mix (1 U/ μL);
 - water to a final volume of 50 μL ;
2. Place PCR tubes in an appropriate PCR cycler at 94°C (*see Note 13*).
 - Cycle conditions:
 - 1 min 94°C, followed by 30 cycles of 20 s denaturation at 94°C, 20 s annealing at 50°C and elongation for 1 min/kb expected product length at 68°C. Increase the elongation time during the last 18 cycles by 15–30 s in each successive cycle. Incubate additional 10 min at 72°C and terminate the reaction by decreasing the temperature to 4°C.

4. Notes

1. The number of cells will be about 10^7 – 10^8 depending on the tissue culture cell line used. We recommend preparing the RNA from a confluent cell layer using at least one 175 cm² tissue culture flask. This ensures that you will get enough poly(A)-containing RNA to perform several RT reactions using the same RNA preparation.
2. There are different Magnetic Particle Concentrators recommended, depending on the size of the tubes. To lyse the cells and bind the RNA to the oligo(dT)₂₅ magnetic beads, we use 15 mL tubes. Before we elute the poly(A)-containing RNA, we transfer the oligo(dT)₂₅ magnetic beads into a 1.5-mL Eppendorf tube.
3. If the oligo(dT)₂₅ magnetic beads appear to clump and do not resuspend well, you have cellular DNA in your sample. However, you can proceed without affecting the RNA quality or yield, if you perform additional washing steps (2–4) until the oligo(dT)₂₅ magnetic beads can be resuspended easily.
4. We strongly recommend analyzing the RNA template integrity before cDNA synthesis. We routinely perform a northern hybridization analysis. In **Fig. 1** poly(A)-containing RNA prepared from HCoV 229E infected MRC-5 cells using two different methods are shown. In both RNA preparations, it is possible to identify the HCoV 229E genomic RNA (27.3 kb) and the six subgenomic mRNAs (1.7 kb–6.8 kb) that are characteristic of

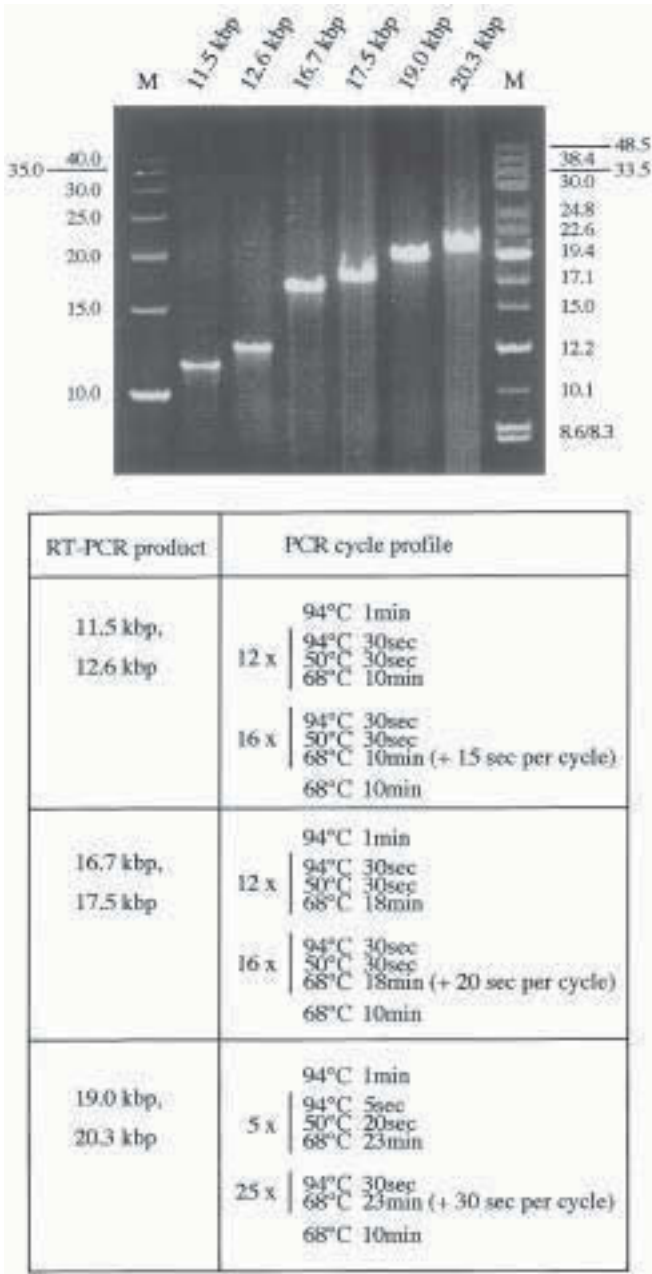


Fig. 2. Pulse-field gel electrophoresis of HCoV 229E RT-PCR products. Five microliters of each PCR reaction were separated by PFGE together with a 5-kbp DNA ladder and a high-molecular weight DNA marker (Life Technologies). Also shown are the cycle profiles that have been used to produce RT-PCR products ranging from 11.5–20.3 kbp in length.

Table 1
Oligonucleotide Primers

Oligonucleotide	Sequence (5' to 3') ^a	Position ^b	Orientation ^c	Application
85	ACACACGGTGTATGTCCTCATT	12979–13000	–	RT
32	TATAGGCATTGCGCAACCACCGG	21747–21769	–	RT
127	cgatcgcggccgctggccgaataggccatgGCTGATTACCGTTGCGCTTGT	9071–9091	+	PCR
11	gagaggatccGAAAACAAACATTTTATTAGTTGAGAC	20554–20582	–	PCR
159	cgatcgcggccgctggccgaataggccATGGCCTGCAACCGTGTGACACT	293–315	+	PCR
89	TCATGGTGTATTTAGTAAGAT	12830–12850	–	PCR
147	AAACCAGTCTGCTCATCA	3860–3877	+	PCR
35	tgggacgTCAAAGGACAACCTGGTCACATCTCAG	21353–21378	–	PCR
36	TTGGTCTGTTGGTGATTGGACCGGT	1048–1072	+	PCR

^aThe nucleotides corresponding to HCoV 229E sequences are shown in capitals.

The nucleotides shown in small case were added for cloning purposes.

^bThe position refers to the nucleotide sequence of HCoV 229E genomic RNA.

^cOligonucleotides with mRNA orientation are designated as +.

coronavirus infection. The hybridization analysis indicates that the material prepared using the oligo(dT)₂₅ magnetic beads (lane 2) is less degraded than the material prepared using poly(U)-Sepharose (lane 1). RT-PCR amplifications of more than 5 kbp were only possible when poly(A)-containing RNA shown in lane 2 was used as the template for reverse transcription.

5. The oligo(dT)₂₅ magnetic beads can be regenerated according to the manufacturer's instructions (Dynal). However, during the repeated washing steps, you will lose about 20% of the magnetic beads.
6. Add the reagents in the order listed in **Subheading 3.2**. To melt RNA structures before the reverse transcription, some protocols recommend to heat the RNA for 10 min at 70°C. However, in most cases this is not necessary.
7. Because the RT is performed at 42°C, cDNAs can be generated during the RT reaction by "less stringent" priming events. This is the reason for most of the background PCR products observed in our system. Therefore, it is absolutely necessary to use a highly specific RT-primer and to adjust the optimal primer concentration in the RT-reaction (6). Furthermore, we recommend the use of different primers for the RT and PCR reactions.
8. Before adding the SuperScript II enzyme, the RNA template and the RT-primer should be at 42°C to minimize unspecific binding.
9. We recommend incubation for 90 min when cDNA synthesis of more than 10 kb is desired. Increasing the incubation temperature above 42°C was not beneficial in our hands.
10. Some protocols for the amplification of long mRNAs have required digestion of the RNA with RNase H after cDNA synthesis (11). This step did not seem to be necessary for amplification with HCoV 229E genomic RNA. This may be owing to the fact that in our protocols the HCoV 229E RNA was relatively abundant. Experiments with dystrophin mRNA required treatment with RNase H prior to amplification (6).
11. The PCR cycle conditions have to be optimized according to the amount of template, the PCR primers and the cycle profile. We recommend that these parameters should be optimized with RT-PCRs of expected product sizes below 5 kbp before trying to synthesize longer RT-PCR products.
12. This will result in a Mg²⁺ concentration of 2 mM.
13. Optional: a "real" hot start can be performed by adding the enzyme mix after the PCR sample has reached 94°C.

Acknowledgments

The authors would like to thank A. Rashtchian for helpful discussions and providing SuperScript II reverse transcriptase and Elongase Enzyme Mix. This work was supported by a Grant (SFB 165/B1) from the German Research Council (DFG) to SGS.

References

1. Barnes, W. M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
2. Cheng, S., Fockler, C., Barnes, W. M., and Higuchi, R. (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. USA* **91**, 5695–5699.
3. Cheng, S., Chang, S.Y., Gravitt, P., and Respass, R. (1994) Long PCR. *Nature* **369**, 684–685.
4. Cheng, S., Higuchi, R., and Stoneking, M. (1994) Complete mitochondrial genome amplification. *Nature Genet.* **7**, 350–351.

5. Cheng, S., Chen, Y., Monforte, J. A., Higuchi, R., and Van Houten, B. (1995) Template integrity is essential for PCR amplification of 20- to 30-kb sequences from genomic DNA. *PCR Meth. Appl.* **4**, 294–298.
6. Thiel, V., Rashtchian, A., Herold, J., Schuster, D. M., Guan, N., and Siddell, S. G. (1997) Effective amplification of 20-kb DNA by reverse transcription PCR. *Analyt. Biochem.* **252**, 62–70.
7. Herold, J., Raabe, T., and Siddell, S. (1993) Molecular analysis of the human coronavirus (strain 229E) genome. *Arch.Virol. [Suppl]* **7**, 63–74.
8. Raabe, T., Schelle-Prinz, B., and Siddell, S. G. (1990) Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus HCV 229E. *J. Gen. Virol.* **71**, 1065–1073.
9. Siddell, S. (1983) Coronavirus JHM: coding assignments of subgenomic mRNAs. *J. Gen.Virol.* **64**, 113–125.
10. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. D., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, (Benson Chanda, V., ed.), Wiley, New York.
11. Nathan, M., Mertz, L. M., and Fox, D. K. (1995) Optimizing Long RT-PCR. *Focus* **17**, 78–80.

Increasing PCR Sensitivity for Amplification from Paraffin-Embedded Tissues

Abebe Akalu and Juergen K. V. Reichardt

1. Introduction

Many common molecular biology techniques including polymerase chain reaction (PCR) (1), Southern blotting (2), comparative genomic hybridization (3), and *in situ* hybridization (4) have been adapted for use with paraffin-embedded tissue (PET). PCR-amplified products from PET can be used for, among others, the analysis of loss of heterozygosity (1), gene amplification (5), direct sequencing, cloning, and characterization of genes (6).

Fixed and embedded material is suboptimal for PCR amplification because of the poor quality of extracted genomic DNA. The integrity of the DNA from PET is critically dependent on a multitude of factors, including the fixative used, fixation time, embedding process, and storage time. Fixation-induced DNA degradations occur as a result of extensive cross-linking of proteins to DNA and acid depurination of the DNA, especially for formalin-based fixatives (6,7). As a result, the DNA is often nicked, yielding relatively short PCR fragments. In addition, PCR inhibitors (histological stains and preservatives) that are coextracted from PET can either cause failure of PCR amplifications or greatly reduce the yield of PCR products (8).

Several approaches and methods have been reported to improve PCR amplification from PET. These include dilution of samples (50–100 cells per 1 L of extraction buffer) and boiling of extracts (8), phenol/chloroform extraction (9), optimizing the fixation and staining process (10), optimization of PCR conditions by prolonging the annealing and extension times, reamplification of DNA from certain samples, and stringent primer selection (11). Each of these approaches has limitations, and determining the optimum conditions for each sample can be laborious and challenging.

Recently, we have reported an improved method for PCR amplification from PET (12). This method uses DNA purified with the QIAquick™ gel extraction kit followed by amplification with AmpliTaq Gold®. The combination of these two approaches has allowed the routine amplification of PCR fragment up to 959 bp from PET, which

exceeds the previously expected upper limit of 800 bp (6). The advantages of these approaches over the conventional methods are as follows:

1. The use of QIAquick™ kit for purification is suitable for management of a large number of samples by eliminating time-consuming step and hazards of organic extractions. In addition, the yield of DNA recovery with the kit is substantially higher than that from phenol/chloroform extraction. This variation is possibly attributed to the loss of samples during organic extraction and subsequent ethanol precipitation. The alternative approach given here is particularly useful when the starting material is small for purification of microdissected PET from a single slide.
2. The gradual activation of AmpliTaq Gold® during thermal cycling allows for high-fidelity and higher-throughput PCR amplification from PET in which the quality and quantity of DNA template can be poor. The reduction of the pre-PCR activation step from 9 to 5 min and the increase in the number of PCR cycles further enhance the time release PCR reaction because polymerase activity builds as specific PCR product accumulates.

This chapter presents an improved strategy for the preparation of PET DNA for use in PCR amplifications including the deparaffinization of PET and extraction of DNA. Because the QIAquick™ kit is primarily designed for isolation of DNA from agarose gels, the adaptation of this kit for DNA purification from PET is also outlined. In addition, a simplified microdissection protocol is presented, as there is an increasing interest in using microdissected PET for the analyses of molecular events leading to malignant transformation step. Finally, the essential aspects of PCR parameters using AmpliTaq Gold® for amplification from PET are described.

2. Materials

2.1. Deparaffinization of PET Block

1. PET blocks (*see Note 1*).
2. Absolute ethanol.
3. Xylene (*see Note 2*).
4. Scalpel or razor blade.

2.2. Microdissection of PET Section from Slides

1. Stained or unstained slides containing tissue sections from PET (*see Note 1*).
2. 30-gauge needle.
3. Cover slip.
4. 50% glycerol (diluted with sterile water).
5. Microscope.

2.3. Digestion of PET with Proteinase K

1. Digestion buffer: 10 mM Tris-HCl (pH 8.0), 1 mM ethylenediaminetetraacetic acid (EDTA), and 1% Tween-20.
2. Proteinase K (20 mg/mL stock solution).
3. Water bath.

2.4. Purification of DNA from PET

1. TE buffer: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA.
2. 100% Isopropanol.
3. QIAquick™ Gel Extraction Kit (Qiagen, Chatsworth, CA).

2.5. PCR Amplification from PET

1. 100 mM dNTP stock solution (Gibco-BRL® Rockville, MD). Prepare a mixture of 10 mM of each dNTP (dGTP, dATP, dCTP, dTTP) in sterile water.
2. AmpliTaq Gold® Kit (Perkin-Elmer, NJ) containing 10X buffer II, 25 mM MgCl₂, and AmpliTaq Gold® (5 U/μL).
3. Purified DNA sample or crude extract from PET.

3. Methods

3.1. Microdissection of PET

Selective procurement of histopathologically defined cell populations from stained or unstained tumor sections on glass slides can be microdissected to study the molecular genetic events that drive the multistep transformation in tumors. Several kinds of microdissection techniques have been described with the development of multiplex molecular analysis using PCR technology (4,13,14). The following simplified protocol is routinely used in our laboratory for microdissection from deparaffinized and stained slides (see **Notes 3** and **4**; see **Fig. 1**):

1. Place a drop of 50% glycerol on the area to be selectively microdissected from the tissue section on a slide.
2. Place a glass cover slip over the drop of glycerol under a light pressure in order to moisten and ease detachment of the cells from the slide. After 5–10 min, carefully lift cover slip off with a scalpel.
3. Place the slide under a microscope and identify the selected area. Cells can be easily identified using a 100× magnification or wide-field microscopy such as an inverted microscope for cell culture.
4. Scratch selected cells, which were identified by microscopic visualization (see **Notes 5–7**), with a disposable sterile 30-gauge needle.
5. Transfer the tissue fragments adhering to the tip of the needle into a 1.5-mL microcentrifuge tube containing digestion buffer (see **Note 8**).
6. Repeat **steps 4** and **5** until the area is entirely microdissected and detached tissues are removed. Proceed with **Subheading 3.3**.

3.2. Deparaffinization of PET Blocks (see **Note 9**)

1. If a microtome is available cut paraffin blocks into single or multiple 10-μm sections (see **Note 10**). If no microtome is available, using a razor blade or disposable scalpel, cut a small section of PET blocks on a clean surface and then transfer the sample into a 1.5-mL microcentrifuge tube using a sterile toothpick or forceps. Sections may be minced into small pieces with scalpel for ease of transferring into a microcentrifuge tube.
2. Add 1 mL xylene and vortex until the paraffin wax sections dissolve.
3. Centrifuge at 10,000g for 3 min at room temperature to pellet the tissues.
4. Carefully remove wax/xylene supernatant into a waste bottle by pipeting.
5. Resuspend the tissue pellet in 1 mL of 100% ethanol by gently vortexing and then centrifuge at 10,000g for 3 min at room temperature. Remove the supernatant.
6. Repeat **step 5**. Air-dry the tissue pellet at room temperature until the ethanol evaporates completely.
7. Resuspend the pellet with digestion buffer and proceed with **Subheading 3.3**.

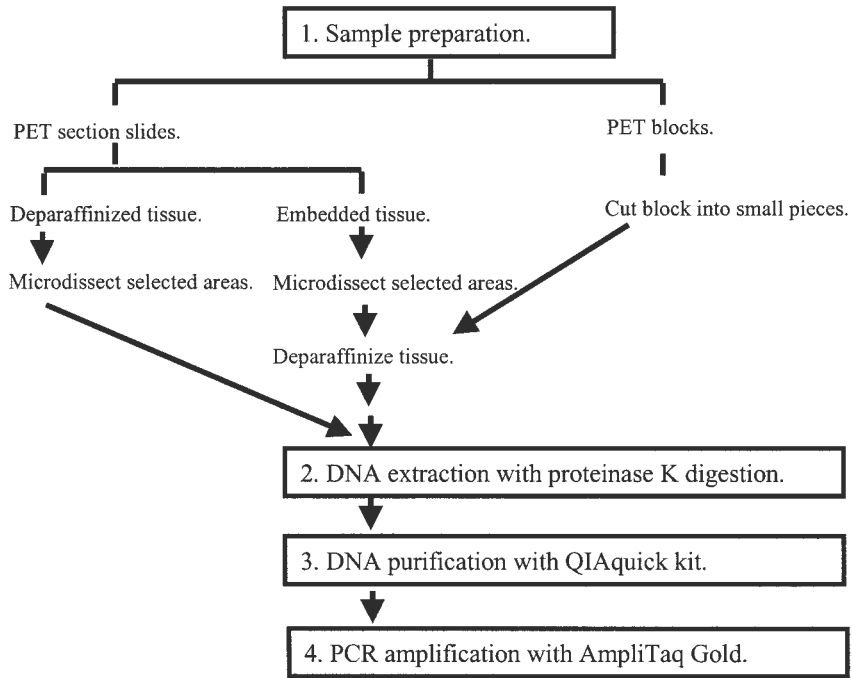


Fig. 1. General overview of processing PET for PCR amplification. After DNA extraction with proteinase K digestion, purification of DNA from PET may not be necessary if the target sequence is short (usually less than 300 bp) and optimal primers, PCR conditions, and polymerase are used. In this case, PCR amplification can be tried directly from crude tissue extracts.

3.3. Digestion of Tissue with Proteinase K

1. Add digestion buffer into 1.5-mL microcentrifuge tube containing deparaffinized tissue pellet from PET blocks or microdissected tissues. The volume of the buffer usually ranges 50–200 μL depending on the amount of tissues (*see* **Notes 7** and **10**). Boil the tube for 3 min and suspend tissues by gently vortexing. Add 100 $\mu\text{g}/\text{mL}$ proteinase K to the digestion buffer and incubate the tube at 52°C for 12–14 h.
2. Inactivate the proteinase K by boiling the tube for 10 min.
3. The extracted DNA can be stored at -20°C until it is used.

3.4. Purification of DNA

Purification is suggested to remove coextracted stains, residual fixation chemicals, and proteins. It is also possible to remove most of the degraded DNAs that compete with intact target sequences for dNTPs and primers during purification. The QIAquick™ Gel Extraction Kit is a silica-based technology used for isolation of DNA fragments (70 bp to 10 kb) from agarose gels as well as for DNA cleanup from enzymatic reactions (*see* **Note 11**). We adapted this kit also for the purification of DNA extracts from PET as follows (**12**):

1. Add 4–5 volumes of binding buffer QG to DNA extract from PET.
2. Add one volume of isopropanol to DNA extract and mix gently by pipeting.
3. Place the QIAquick™ spin column in a 2-mL tube. Load the sample from step 2 into the spin column. Centrifuge at 5000g for 1 min. Discard the flowthrough.
4. Place the spin column back into the same 2-mL tube. Add 0.75 mL of wash buffer PE and incubate the column for 5 min at room temperature. Centrifuge at 5000g for 1 min. Discard the flowthrough.
5. Place the spin column into the same tube. Centrifuge at 1000g for 1 min.
6. Transfer the spin column into a new 1.5-mL recovery tube. Add 30–100 µL of TE buffer directly to the center of the spin, incubate for 3 min at room temperature, centrifuge at 10,000g to elute the DNA, and then store at –20°C until it is used.

3.5. PCR Amplification from PET

Set up the PCR reaction (50 µL/reaction) by adding the AmpliTaq Gold® reagents into 0.5-mL microcentrifuge tube as follows:

1. 5 µL of 10X PCR buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl).
2. 4 µL of 25 mM MgCl₂ (final concentration 2 mM).
3. 4 µL of 10 mM dNTPs (final concentration 200 µM of each dNTP).
4. 0.1–0.2 µM of each forward and reverse primer (*see Note 12*).
5. 5–20 ng of genomic DNA (*see Note 13*).
6. 2.5 U AmpliTaq Gold®/reaction.
7. Bring to a total volume of 50 µL with sterile water.
8. Overlay the reaction mixture with 1 drop of sterile mineral oil.
9. Amplify the PCR reaction in a programmable thermal cycler. The following thermal cycling profile in RoboCycler® Gradient 40 (Stratagene, La Jolla, CA) is optimized for the amplification of fragments shown in Fig. 2:
 - a. Step 1: 1 cycle of initial denaturation at 95°C for 5 min (*see Note 14*).
 - b. Step 2: 60 cycles of denaturation at 95°C for 2 min, annealing at 60°C for 80 s, and elongation at 72°C for 80 s (*see Note 15*).
 - c. Step 3: 1 cycle of final extension at 72°C for 5 min.

4. Notes

1. For PCR amplifications, prepared specimen of PET on glass slides or as paraffin blocks are usually obtained from a pathology laboratory.
2. Xylene is a toxic organic solvent. Octane or commercially available solvents such as Hemo-De® Clearing Agent (Fisher Scientific) or AmeriClear® (Baxter Scientific) can be substituted.
3. Unstained sections on slide can also be used as long as the areas of interest are identified. The areas of tumors to be microdissected are selected and circled with a felt-tip pen on the inverted side of the slide usually by a pathologist.
4. A good morphological distinction between benign and malignant cells usually requires histologic staining of tissues for microscopic visualization. Usually sections are stained after deparaffinization. Sometimes, tissue sections on slides are stained without prior deparaffinization since paraffin holds the tissue fragments together during the microdissection process (10). In this case, sections are deparaffinized after microdissection. To deparaffinize microdissected tissues follow **steps 2–7** in **Subheading 3.2.**, by using less xylene (300 µL) and ethanol (300 µL).

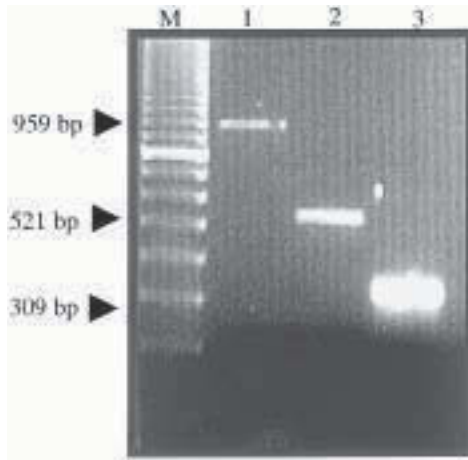


Fig. 2. An example of amplification from PET following the PCR parameters in **Subheading 3.5**. The tumor tissues were microdissected from deparaffinized and stained slide, digested with proteinase K, and purified with QIAquick™ kit. About 20 ng of DNA template was used for a PCR amplification with AmpliTaq Gold®. Lanes 1, 2, and 3 show amplification of 959 bp, 521 bp, and 309 bp fragments, respectively. M, DNA ladder. This figure is modified from **ref. 12** with permission from Elsevier Science.

5. Given the heterogeneity of biopsies, and the infiltrative nature of many tumors, an important consideration in microdissection is the minimization or possibly eliminating the contamination of normal cells with neoplastic cells or vice versa.
6. For the purpose of analyzing loss of heterozygosity from the same slide, normal tissues adjacent to the tumor cells should be microdissected and placed in a separate tube for analyses.
7. The size of tumor area is variable usually ranging 5–20 mm². If the size of the tumor is small, the same specific region from duplicate tissue slides can be microdissected, and scrapes can be pooled in one microcentrifuge tube to give sufficient DNA. Thus, the volume of the digestion buffer can range from 50–200 μL depending on the size of the tumor and the number of slides to be microdissected.
8. Microdissected samples from deparaffinized tissue slides are directly scraped off into 1.5-mL microcentrifuge tube containing digestion buffer. From undeparaffinized slides, tissues are scraped off into tube containing xylene, and then deparaffinized (*see Note 4*).
9. Deparaffinization of PET is a widely used procedure to promote digestion of tissues with proteinase K and PCR amplification. However, it is reported that deparaffinization may not be necessary for PCR analysis of RNA from PET (*15*).
10. For a number of analyses, tissue sections (5–10 μm) are prepared from PET blocks and mounted on slides. Sectioning of paraffin blocks using a microtome requires training and experience. The amount of sections required for extraction depends on the availability and the intended use of the tissue. If more tissue is required, a section of blocks 50–100 mg (or more) can be cut and processed. Do not exhaust all the blocks.
11. Recently, a DNeasy™ Tissue Kit for purification of DNA from PET has been introduced. In our laboratory, no differences in the quality and quantity of purified DNA for PCR

- amplification were observed between DNeasy™ Tissue Kit and QIAquick™ Gel Extraction Kit. Because the gel extraction kit is routinely used for isolation of DNA from agarose gels for cloning, the use of the kit for DNA purification from PET is an added benefit.
12. As always, primer design is one of the most important aspects of PCR. Because DNA template from PET is a complex mixture of degraded DNA fragments, careful attention is required to designing primers for PCR amplification from PET. Not all primers that amplify from fresh tissue can amplify fragments from PET. Although common guidelines are available, software programs such as Primer3 (from S. Rozen, H. and J. Skaletsky, <http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>) can be used for optimum primer selection. In the event of failure to amplify fragments, it usually helps to try different sets of primer pairs. It is also helpful to screen extracts from PET for the integrity and quantity of amplifiable DNA by running a control amplification reaction for a single-copy housekeeping gene such as β -actin primer (**16**) or another suitable target.
 13. The critical factor for successful PCR amplification from PET is the integrity of the target sequence. The amount of usable genomic DNA extracted from PET depends on multiple factors, including the amount of tissue, storage time, the chemical composition of the fixative, fixation time, and the embedding process. In addition, yields of genomic DNA will vary from tissue to tissue from which the DNA is extracted. The amount of DNA is greater in tissues containing concentrated nucleated cells than tissues with fewer nucleated cells. DNA extraction from PET in most cases yields approx 10 ng of DNA from 1000 cells. Determination of the concentration of DNA extracted from PET may not be necessary, if the amount of starting material is too small. This may also help in saving irreplaceable DNA sample. However, if the DNA is extracted from large blocks of PET, the concentration can be determined following a standard procedure. From our experience, 1–2 μ L of purified DNA from microdissected tissue with an area of 5–20 mm² gives a reasonable yield of PCR products. In some cases, if the DNA is either concentrated or crude cell extract is used for amplification, appropriate dilution of the template should be made.
 14. AmpliTaq Gold® is reversibly chemically inactivated. It gains about 40% of its activity with a pre-PCR heat step of 95°C for 9–12 min and its reactivation continues with subsequent thermal cycles in a time-release manner. The manufacturer's recommended time (at 95°C for 9–12 min) should be reduced to 2–5 min in order to increase the time-release effect of AmpliTaq Gold®. To compensate for the shortening of the pre-cycling heat step, increased thermal cycles up to 50 or more are required so that enough PCR products is generated for restriction analysis, sequencing, and cloning.
 15. The number of thermal cycles is determined by the amount of input DNA and intended use of the PCR product for analysis. For instance, when [γ -³²P]ATP radiolabeled primers are used for screening mutations and genotyping by single-strand conformation polymorphism analysis, the number of thermal cycles should not exceed 30 and the preincubation of 95°C for 9 min can be used. Because the method is so sensitive, radiolabeled PCR products from over 30 cycles usually show nonspecific amplifications and high background on autoradiography. This will largely compromise the results and lead to incorrect genotyping.

Acknowledgments

This work was supported by grants from the DoD for the USC Prostate Center (PC 992018; project A) and the TJ Martell Foundation to JKVR.

References

1. Akalu, A., Elmajian D. A., Highshaw, R. A., Nichols, P. W., and Reichardt, J. K.V. (1999) Somatic mutations at the SRD5A2 locus encoding prostatic steroid 5 α -reductase during prostatic cancer progression. *J. Urol.* **161**, 1355–1358.
2. Bramwell, N. H. and Burns, B. F. (1988) The effects of fixative type and fixation time on the quantity and quality of extractable DNA for hybridization studies on lymphoid tissue. *Exp. Hematol.* **116**, 730–732.
3. Speicher, M. R., Jauch, A., Walt, H., du Manoir, S., Schrock, E., Holtgreve-Grez, H., Schoell, C., Lengauer, C., Cremer, T., and Reid, T. (1993) Molecular cytogenic analysis of formalin-fixed, paraffin embedded solid tumors by comparative genomic hybridization. *Hum. Mol. Genet.* **2**, 1907–1914.
4. Nuovo, G. J. and Silverstein, S. J. (1988) Methods in laboratory investigation: comparison of formalin, buffered formalin, and Bouin's fixation on the detection of human papillomavirus deoxyribonucleic acid from genital lesions. *Lab. Invest.* **59**, 720–724.
5. Ribot, E. M., Quinn, F. D., Bai, X., and Murtagh, J. J., Jr. (1998) Comparative PCR: An improved method to detect gene amplification. *BioTechniques* **24**, 22–26.
6. Crisan, D. and Mattson, J. C. (1993) retrospective DNA analysis using fixed tissue specimens. *DNA Cell Biol.* **12**, 455–464.
7. Jacson, V. (1978) Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. *Cell* **15**, 945–954.
8. Zhuang, Z., Bertheau, P., Emmert-Buck, M. R., Liotta, L. A., Gnarra, J., Linehan, W. M., and Lubensky, I. A. (1995) A microdissection technique for archival DNA analysis of specific cell populations in lesions < 1 mm in size. *Am. J. Pathol.* **146**, 620–625.
9. Greer, C. E., Peterson, S. L., Kiviat, N. B., and Manos, M. M. (1991) PCR amplification from paraffin-embedded tissues. Effects of fixative and fixation time. *Am. J. Clin. Pathol.* **95**, 117–124.
10. Burton, M. P., Schneider, B. G., Brown, R., Escamilla-Ponce, N., and Gulley, M. L. (1997) Comparison of histologic stains for use in PCR analysis of microdissected, paraffin-embedded tissues. *BioTechniques* **24**, 86–92.
11. Wright, D. K. and Manos, M. M. (1990) Sample preparation from paraffin-embedded tissues, in *PCR Protocols, A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds.) Academic, San Diego, CA, pp. 32–38.
12. Akalu, A. and Reichardt, J. K. V. (1999) A reliable PCR amplification method for microdissected cells obtained from paraffin-embedded tissue. *Genet. Analyt.* **15**, 229–233.
13. Walch, A., Komminoth, P., Hutzler, P., Aubele, M., Hofler, H., and Werner, M. (2000) Microdissection of tissue section: application to the molecular genetic characterization of premalignant lesions. *Pathology* **68**, 9–17.
14. Baisse, B., Bian, Y.-S., and Benhattar, J. (2000) Microdissection by exclusion and DNA extraction for multiple PCR analyses from archival tissue sections. *BioTechniques* **28**, 856–862.
15. Shimizu, R. and Burns, J. C. (1995) Extraction of nucleic acids: sample preparation from paraffin-embedded tissues, in *PCR Strategies* (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., White T. J., eds.), Academic, San Diego, CA, pp. 153–136.
16. Benz-Ezra, J., Johnson, A., Rossi, J., Cook, N., and Wu, A. (1991) Effect of fixation on the amplification of nucleic acids from paraffin-embedded material by the polymerase chain reaction. *J. Histochem. Cytochem.* **39**, 351–354.

GC-Rich Template Amplification by Inverse PCR

DNA Polymerase and Solvent Effects

**Alain Moreau, Da Shen Wang, Steve Forget, Colette Duez,
and Jean Dusart**

1. Introduction

The amplification of GC-rich templates by any PCR method is usually a difficult task and despite the development of modified methods and conditions, this type of amplification still remains a specific case approach. Problems usually observed with GC-rich DNA are constraint of template amplification by stable secondary structures that stall or reduce the DNA polymerase progress, and the presence of secondary annealing sites giving rise to nonspecific amplified bands. This latter point is not exclusive to GC-rich templates but is frequently encountered in other types of templates. In order to design a more general method for GC-rich templates, different DNA polymerases were compared in combination with different organic solvents with the purpose of abolishing stable secondary structures (*1*). Our attention focused on the inverse polymerase chain reaction (iPCR) used to perform site-directed mutagenesis (*1,2*). This very attractive method requires a single pair of primers and involves the amplification of the whole recombinant plasmid, a difficult step with high GC-content DNA. Inverse PCR also proves useful in cloning missing parts of genes by using a self-ligated genomic DNA fragment as template.

A recent survey of the literature showed the absence of comparative studies regarding the use of different DNA polymerases in the amplification of GC-rich DNA with or without the addition of organic solvents, such as dimethyl-sulfoxide (DMSO) (*3,4*), formamide (*3,5,6*), and tetramethylammonium chloride (TEMAC) (*7*). Furthermore, little is known of the exact role of these chemicals in PCR. It was suggested that these compounds primarily affect the annealing kinetics as well as the efficiency of the DNA polymerase used. In order to identify critical parameters involved in iPCR with GC-rich templates, we analyzed the influence of DNA polymerases in combination

with the aforementioned solvents (**1**). The results obtained allowed us to improve iPCR for difficult template amplifications by either iPCR or standard PCR. Our iPCR method can be divided into two steps: amplification of the whole recombinant plasmid and iPCR product purification and ligation. This method was used to perform site-directed mutagenesis by amplification of a 4.8-kb plasmid derived from pUC18 and containing a 1980-bp insert, the gene encoding the extracellular DD-carboxypeptidase from *Actinomadura* R39, a 74% GC-content DNA (**14**). Different DNA polymerases were tested according to the manufacturer's specifications. However, correct amplification was not detected with any DNA polymerase tested (**1**). The efficiency of amplification by addition of DMSO, formamide, or TEMAC in the reaction mixture was evaluated according to the conditions described above. Analysis of PCR products in the presence of these organic solvents revealed that only Vent™ DNA polymerase (New England Biolabs, Beverly, MA) amplified the 4.8-kb plasmid (**1**). We focused our attention on the conditions to amplify GC-rich DNA templates with Vent DNA polymerase. In addition, the Mg²⁺ concentration was increased to obtain a 10-mM final concentration. This allowed us to amplify large DNA fragments in the PCR assay.

2. Materials

1. pBlueScript™ vector (Stratagene, La Jolla, CA) or any other pUC derivative plasmid used to clone the gene to be mutated (*see Note 1*).
2. 25–50 ng of DNA template (*see Note 2*).
3. 2 μM of each oligonucleotide, primers A and B corresponding respectively to the sense (coding) and antisense (noncoding) message (*see Note 3*).
4. 10X DNA polymerase buffer (provided by the manufacturer of the Vent DNA polymerase) (*see Note 4*) and Vent DNA polymerase (*see Note 5*).
5. dNTPs: Mix 10 mM each (Pharmacia LKB, Piscataway, NJ).
6. 100 mM MgSO₄.
7. Fresh, deionized formamide (Sigma, St. Louis, MO).
8. Sterile water.
9. Light mineral oil (Sigma).
10. Reagents for agarose gel electrophoresis (Life Technologies [Gibco-BRL], Gaithersburg, MD).
11. Sephadex G-50 fine (Pharmacia LKB), siliconized wool and 1-mL syringe.
12. T4 polynucleotide kinase (PNK) (New England Biolabs).
13. Reagents for ligation: 10X ligation buffer (Boehringer Mannheim, Indianapolis, IN); T4 DNA ligase (Boehringer Mannheim, Germany).
14. 400-μL and 1.5-mL sterile Eppendorf tube.
15. Thermocycler.

3. Methods

3.1. Inverse PCR

1. Perform iPCR by adding in a 400-μL sterile Eppendorf tube the following reagents (*see Note 6*): 5 μL DNA template (25–50 ng); 10 μL 10X Vent DNA polymerase buffer; 8 μL MgSO₄ (100 mM); 2 μL dNTPs (10 mM each); 4 μL primer A (50 pmol/μL); 4 μL primer B (50 pmol/μL); 10 μL formamide; 56 μL sterile H₂O; and 1 μL Vent DNA polymerase (2 U/μL); to a total volume of 100 μL.

2. Overlay the sample with 30 μL of light mineral oil.
3. Submit the samples to a standard three-step cycling protocol according to the following parameters (see Note 7): 95°C for 1 min (initial denaturation) (1 cycle); 94°C for 30 s (denaturation); XX°C, 1 min (annealing); 72°C for Y min (extension) 30 cycles; 72°C for 10 min (final extension) (1 cycle).

3.2. iPCR Product Purification and Ligation

1. Separate the iPCR reaction from the light mineral oil by simply pipeting only the reaction volume from the bottom of the tube into a new 1.5-mL sterile Eppendorf tube.
2. Take 10–20 μL aliquot of the iPCR reaction to visualize the product by agarose gel electrophoresis (0.7% agarose, see Note 8).
3. Purify the iPCR product by passing the remaining iPCR reaction through a Sephadex G-50 spun column of 1-mL and elute with 100 μL of TE buffer or H₂O.
4. Perform the phosphorylation and ligation reaction by adding the following reagents in a 1.5-mL sterile Eppendorf tube: 10 μL aliquot of purified iPCR reaction; 2 μL 10X ligation buffer; 0.5 μL T4 polynucleotide kinase (10 U/ μL); and 7.5 μL H₂O sterile; to a total volume of 20 μL .
5. Incubate the mixture 15 min at 37°C, then again add 0.5 μL of T4 polynucleotide kinase and incubate for another 15 min at 37°C.
6. Add 1 μL of T4 DNA ligase (1 U/ μL) to the reaction mixture and incubate overnight at 4°C, followed by 16°C for 3 h.
7. Transform competent *Escherichia coli* (*E. coli*) cells with 5- μL aliquot of the ligation mixture.

4. Notes

1. The targeted DNA is initially cloned in a vector, in general, a pUC-derived plasmid. Because the difficulty of amplification increases with plasmid length, it is better to avoid unnecessarily large plasmid and clone only a part of the gene flanking the targeted DNA. Once mutated by iPCR, this part will then be used to reconstruct the whole gene. The latter step reduces the subsequent sequencing necessary to check the integrity of the mutated DNA fragment.
2. The template concentration used to perform iPCR is about 25–50 ng. Higher concentrations will increase the background because of the wild-type plasmid that easily transforms *E. coli*. Lower template concentration reduces the final amount of amplified material, thus requiring more PCR cycles. These additional cycles contribute to the introduction of more errors by DNA polymerases. Under our conditions, 30 cycles are sufficient to amplify GC-rich templates.
3. The primer design is a crucial step. In iPCR, the primers to be used are oriented in inverted tail-to-tail direction, i.e., one primer corresponding to the coding sense (5'–3') whereas the other is antisense (3'–5'). Usually, one primer harbors the mutation, which can be a substitution, deletion, or insertion of one or more nucleotides. The selection of this pair of primers is sometimes difficult, but can be simplified by using one of several new computer programs for oligonucleotide selection (12,13). However, this initial step is frequently overlooked, resulting in great difficulties in template amplification (not just GC-rich ones). There are three basic rules to follow to avoid primer design problems:
 - a. Elimination of duplex formation at the 3' ends with either one or both primers, plus elimination of hairpin structure formation within primers.
 - b. Design of primers with T_m (°C) close to each other, i.e., less than a 10°C difference. The addition of any organic solvent to the PCR reaction will decrease the T_m for a

specific primer; and its partner primer bearing the mutation will also show reduced T_m depending on the introduced mutation.

- c. Location of the chosen mutation in the middle of the primer in order to maintain the internal stability of the oligonucleotide.
4. The buffer supplied by the manufacturer is 1X: 20 mM Tris-HCl, pH 8.8 (at 25°C), 10 mM KCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 2 mM MgSO_4 , and 0.1% Triton X-100.
5. The choice of DNA polymerase is a key point for the amplification of the whole recombinant plasmid. Among several DNA polymerases tested, our choices were Vent DNA polymerase (8–10) and *Pfu* DNA polymerase (Stratagene) (11). Indeed, these two enzymes produce almost exclusively blunt ends, whereas *Taq* DNA polymerase requires additional manipulations to obtain blunt ends. Furthermore, Vent and *Pfu* DNA polymerase are more accurate during iPCR than *Taq* DNA polymerase. We recommend testing the first iPCR reactions without the addition of organic solvents. It is difficult to choose the solvent and its optimal concentration without empirical assay. Furthermore, it has been noted that most DNA polymerases are sensitive to organic solvents, especially formamide. However, the use of 10% formamide is suggested in combination with Vent DNA polymerase (1). This DNA polymerase proved to be the most robust enzyme tested in the presence of formamide, because other DNA polymerases (*Taq* and **Pfu**) could not amplify DNA under similar conditions (Moreau, A., unpublished observations). Another possibility is the use of *Pfu* DNA polymerase without any solvent addition. In some cases, *Pfu* DNA polymerase gives rise to good amplification with GC-rich templates but does not tolerate formamide concentrations greater than 2.5%.
6. We observed that the Mg^{2+} concentration is very important in obtaining proper amplification with the Vent DNA polymerase, especially with large plasmids (5 kb). In the absence of amplification products, we recommend the modification of Mg^{2+} concentration by supplemental addition of increasing amounts of Mg^{2+} in the iPCR reaction. The optimum Mg^{2+} concentration usually occurs in a narrow range with the Vent DNA polymerase. Among a choice of several organic solvents, 10% formamide was the most useful addition to correctly amplify the 4.8 kb plasmid (see Fig. 1). Higher formamide concentrations appeared detrimental for the iPCR. It was also observed that addition of <8% DMSO failed to amplify the plasmid but higher concentrations did not significantly improve the iPCR, even with addition of T4 gene 32 protein, a single-stranded binding protein used to overcome secondary structures (15,16). The addition of TEMAC at concentrations from 10^{-2} – $10^{-5}M$ was not effective (see Fig. 1). The iPCR assay was repeated with different pairs of primers, in different locations and each time, the 4.8-kb plasmid was successfully amplified.
7. We choose to perform iPCR with the *Actinomadura* R39 DD-carboxypeptidase (74% GC-rich) with an annealing temperature of 55°C/1 min, and 5 min extension time at 72°C. This iPCR has been performed with the following two primers: 5'-GCCCTCGGGC GCGGTACCGCAT-3' ($T_m=70^\circ\text{C}$), which anneals perfectly with the target sequence and 5'-GTGGTCGAGGCCACACCGGGACGATG-3' ($T_m=61.6^\circ\text{C}$) introducing two non-contiguous mismatches in the sequence (because these substitutions are not contiguous, we have considered them as three mismatches). The annealing temperature has also been reduced, based on the guideline of T_m decrease of both primers by about 0.6°C/% of formamide added to the iPCR reaction. We used the extension time of 1 min/kb of plasmid for the polymerization at 72°C, and a final 10-min polymerization at 72°C. The guideline of 1 min/kb for the extension at 72°C is suitable when the plasmid size is 5 kb or less,

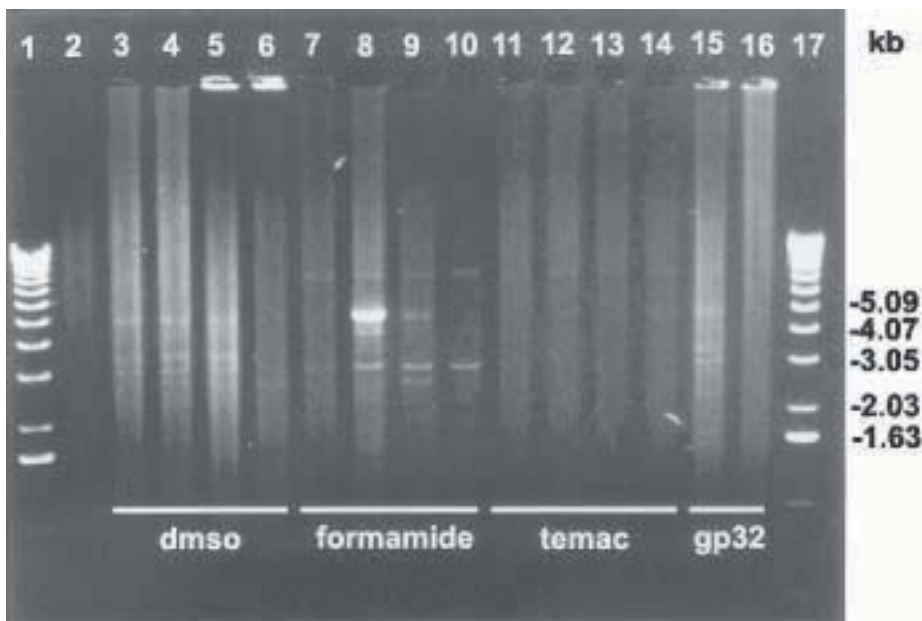


Fig.1. Solvent effects on iPCR. A 4.8-kb plasmid containing a GC-rich (74%) 1980-bp insert was amplified by iPCR using Vent™ DNA polymerase. Samples covered with mineral oil (30 μ L) were submitted to 30 amplification cycles on a Biometra Trio-Thermoblock™ :1 min denaturation at 95°C, 1 min annealing at 55°C, and 5 min extension at 72°C, followed by a final 10-min extension at 72°C. A 10- μ L aliquot of each sample was loaded onto a 0.7% agarose gel. Lanes 1 and 17, 1-kb DNA ladder size standard; lane 2, control iPCR under standard conditions without organic solvent; lanes 3–6, iPCR in presence of 8, 10, 15, and 20% DMSO respectively; lanes 7–10, iPCR in presence of 5, 10, 15, and 20% formamide respectively; lanes 11–14, iPCR in presence of 10^{-2} , 10^{-3} , 10^{-4} , and 10^{-5} M TEMAC, respectively; lane 15, iPCR in presence of 8% DMSO with 3 μ g of T4 gene 32 protein; and lane 16, iPCR under standard conditions with 3 μ g of T4 gene 32.

whereas with larger plasmid size, the extension time should be increased on an empirical basis. The initial denaturation is an important step, but prolonging it may be detrimental for longer templates.

8. An aliquot of the iPCR reaction is visualized by agarose gel electrophoresis. The presence of minor bands of smaller size than the wild-type plasmid is not a real problem. Most of them will generate partial plasmids missing key regions of replication or antibiotic resistance and will be easily eliminated after transformation. However, several methods are available to purify PCR products. We use a Sephadex G-50 spun-column to remove unincorporated dNTPs. A purified aliquot is then phosphorylated and ligated. It is not unusual to find frameshifts+1 or -1 at the junction of each primer after isolation of plasmid DNA from transformants obtained through iPCR product ligation. However, we observed that most of the transformants are the result of correct ligation.

References

1. Moreau, A., Duez, C., and Dusart, J. (1994) Improvement of GC-rich template amplification by inverse PCR. *BioTechniques* **17**, 233–234.
2. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–623.
3. Bookstein, R., Lai, C. C., To, H., and Lee, W. H. (1990) PCR-based detection of a polymorphic BamHI site in intron 1 of the human retinoblastoma (RB) gene. *Nucl. Acids Res.* **18**, 1666.
4. Winship, P. R. (1989) An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucl. Acids Res.* **17**, 1266.
5. Sarkar, G., Kapelner, S., and Sommer, S. S. (1990) Formamide can dramatically improve the specificity of PCR. *Nucl. Acids Res.* **18**, 7465.
6. Schuchard, M., Sarkar, G., Ruesink, T., and Spelsberg, T. C. (1993) Two-step “hot” PCR amplification of GC-rich avian c-myc sequences. *BioTechniques* **14**, 390–394.
7. Hung, T., Mak, K., and Fong, K. A. (1990) Specificity enhancer for polymerase chain reaction. *Nucl. Acids Res.* **18**, 4953.
8. Eckert, K. A. and Kunkel, T. A. (1991) DNA polymerase fidelity and the polymerase chain reaction. *PCR Meth. Applicat.* **1**, 17–24.
9. Kong, H., Kucera, R. B., and Jack, W. E. (1993) Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J. Biol. Chem.* **268**, 1965–1975.
10. Mattila, P., Korpela, J., Tenkanen, T., and Pitkanen, K. (1991) Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase— an extremely heat stable enzyme with proofreading activity. *Nucl. Acids Res.* **19**, 4967–4973.
11. Lundberg, K. S., Shoemaker, D. D., Adams, M. W., Short, J. M., Sorge, J. A., and Mathur, E. J. (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**, 1–6.
12. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
13. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
14. Granier, B., Duez, C., Lepage, S., Englebert, S., Dusart, J., Dideberg, O., et al (1992) Primary and predicted secondary structures of the *Actinomadura* R39 extracellular DD-peptidase, a penicillin-binding protein (PBP) related to the *Escherichia coli* PBP4. *Biochem. J.* **282**, 781–788.
15. Alberts, B. and Sternglanz, R. (1977) Recent excitement in the DNA replication problem. *Nature* **269**, 655–661.
16. Bittner, M., Burke, R. L., and Alberts, B. M. (1979) Purification of the *T4* gene 32 protein free from detectable deoxyribonuclease activities. *J. Biol. Chem.* **254**, 9565–9572.

PCR Procedure for the Isolation of Trinucleotide Repeats

Teruaki Tozaki

1. Introduction

Microsatellites, also referred to as short tandem repeats (STR) or simple sequence repeats (SSR), are highly polymorphic and abundant sequences dispersed throughout most eukaryotic nuclear genomes (1–3). In recent years, microsatellites have been used for linkage map construction, population genetics, molecular evolution studies, forensic sciences, and as parentage testing markers (4,5).

Trinucleotide repeats, such as $(CAG)_n$ and $(CGG)_n$ repeats, have drawn increasing attention in various aspects of both human gene mapping and clinical genetics. Genes with trinucleotide repeats, such as transcription-regulatory proteins and homeobox genes, are frequently found in mammalian genomes (6). Some of these genes are involved in neuropsychiatric disorders in human, such as Huntington's disease (7) and spinocerebellar ataxia type 1 (8). At present, the function of the trinucleotide repeats is not clear.

Polymorphism of trinucleotide repeats can be resolved more easily in alleles differing by one repeat than those of dinucleotide repeats. Thus, tri- and tetranucleotide repeats are useful for human forensic sciences. The abundance and polymorphic information of trinucleotide repeats in the human genome have been characterized by Gastier et al. (9). In the human genome, $(CAG)_n$ repeats occur, on average, once every 4400 kb, making them less abundant than dinucleotide repeats. Thus, it is considered that the construction of an enrichment library is useful for the isolation of trinucleotide repeats.

A novel, rapid, and convenient cloning method by the construction of an enrichment library has been developed for the isolation of trinucleotide repeats in human genome, and applied for the isolation of tri- and dinucleotide repeats in equine genome (10–14). The method includes the following procedures: adapter PCR of genomic DNA, enrichment procedure, adapter PCR for large preparation, followed by cloning. The enrichment procedure includes the following steps: capturing some other element from

genomic DNA by hybridization to biotin-labeled probes in solution, a nucleotide substrate-biased polymerase reaction, a subsequent extraction with magnetic beads coated with streptavidin. In particular, the nucleotide substrate-biased polymerase reaction is useful for the isolation of microsatellites containing long repeating units. The method is designed to select only the longer repeats. The advantage of the method is the efficient isolation of long repeat-containing highly polymorphic microsatellites, because microsatellites greater than 12-repeats in length are more likely to be polymorphic (15).

In addition, this method offers the major advantage that it can be repeated any number of times for enrichment, indicating that the method is useful for the cloning of less frequent microsatellites, such as trinucleotide repeats.

This chapter describes the construction of an enrichment library for the microsatellite-isolation. The method described here is a detailed version of a method that have been reported previously (10–14). The method is outlined in **Fig. 1**.

2. Materials

1. Autoclaved sterile 1.5-mL tube.
2. Autoclaved sterile 200 μ L-PCR tube.
3. Purified genomic DNA.
4. *Sau3AI* (Takara Shuzo).
5. 10X buffer for *Sau3AI*: 500 mM Tris-HCl, pH 7.5, at 25°C, 100 mM MgCl₂, 10 mM dithiothreitol, 1000 mM NaCl.
6. Mung bean nuclease (Takara Shuzo).
7. 1X buffer for Mung Bean nuclease: 30 mM acetate, pH 5.0, at 25°C, 100 mM NaCl, 1 mM zinc acetate, 5% glycerol.
8. Autoclaved, distilled, deionized water.
9. Oligonucleotide primers for adapter ligation.

For first enrichment:

Primer (1) 5'-GCACTCTCCAGCCTCTCAGTGCAG-3' (100 μ M and 25 μ M).

Anti primer (1) 5'-GATCCTGCACTG-3' (100 μ M).

For second enrichment:

Primer (2) 5'-AGCACTCTCCAGCCTCTCACCGAG-3' (100 μ M and 25 μ M).

Anti primer (2) 5'-GAT CCT CGG TGA-3' (100 μ M).

10. Biotin-labeled oligo (CAG)₈ (20 μ M).
11. *T4* DNA ligase (Takara Shuzo).
12. 10X ligase buffer: 660 mM Tris-HCl, pH 7.6, at 25°C, 66 mM MgCl₂, 100 mM dithiothreitol, 1 mM ATP.
13. *rTaq* polymerase (5 U/ μ L) (Takara Shuzo).
14. 10X *rTaq* buffer (MgCl₂ free): 100 mM Tris-HCl, pH 8.3, at 25°C, 500 mM KCl.
15. *ExTaq* polymerase (5 U/ μ L) (Takara Shuzo).
16. 10X *ExTaq* buffer (MgCl₂ plus) (Takara Shuzo).
17. *LATaq* polymerase (5 U/ μ L) (Takara Shuzo).
18. Perfect Match PCR enhancer (Stratagene, CA).
19. An enzyme mixture: 0.5 μ L of *rTaq* polymerase, 0.2 μ L of 10X diluted *LATaq* polymerase, and 2 μ L of Perfect Match PCR enhancer (see **Note 5**).
20. dNTP (each 2.5 mM of dATP, dCTP, dGTP, and dTTP).
21. d3NTP (each 2.5 mM of dATP, dCTP, and dGTP).

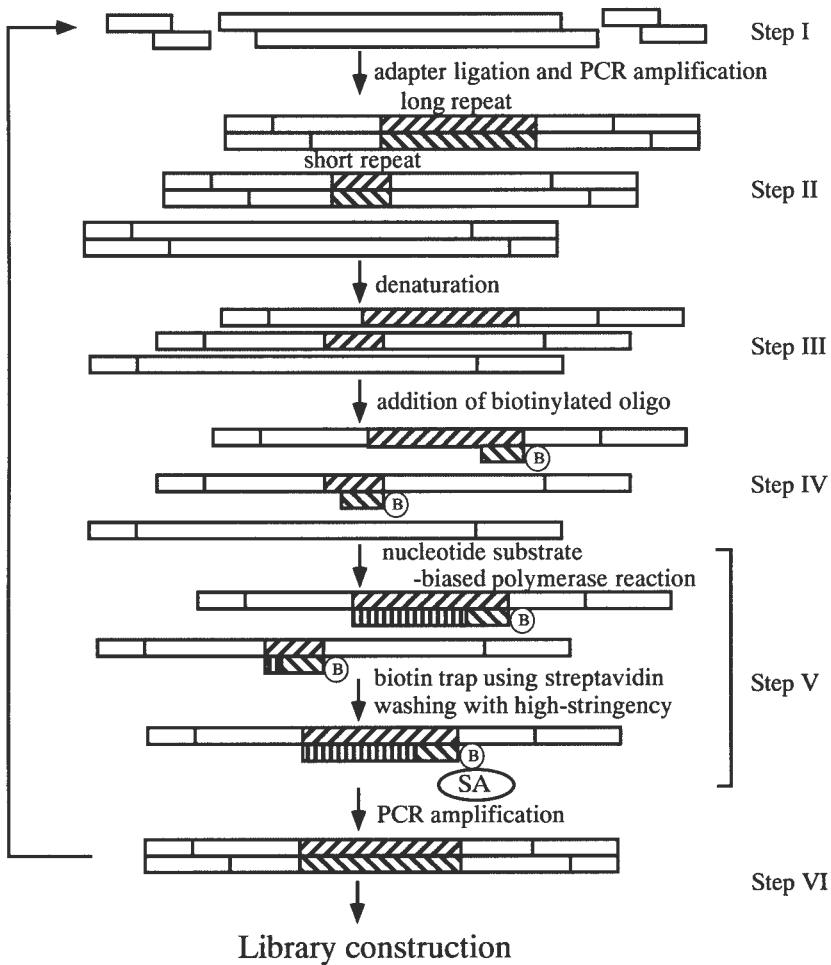


Fig. 1. A schematic diagram of the construction of a library enriched for microsatellites. The first line represents the *Sau3AI*-digested genomic DNA and the adapter sequences. The striped boxes are microsatellite target sequences, SA denotes the streptavidin-coated magnetic particles and B indicates biotin molecules.

22. 1X binding/washing buffer: 1 M NaCl, 10 mM Tris-HCl, pH 8.0, at 25°C, 1 mM EDTA.
23. 2X binding/washing buffer: 2 M NaCl, 10 mM Tris-HCl, pH 8.0, at 25°C, 1 mM EDTA.
24. Alkaline buffer: 1 mM NaOH, 10 mM Tris, 1 mM EDTA.
25. *ColI* DNA.
26. Streptavidin MagneSphereR Paramagnetic Particle (Promega, WI). Wash by adding and removing 0.5X SSC at room temperature twice, and resuspend in 100 μ L of 1X binding/washing buffer with *ColI* DNA (100 ng/ μ L).
27. Magna-Sep (a magnetic stand) (Gibco BRL, MD).

28. TE buffer: 10 mM Tris-HCl, pH 8.0, at 25°C, 1 mM ethylenediaminetetraacetic acid (EDTA).
29. Phenol equilibrated in TE buffer.
30. Chloroform:isoamyl alcohol: 24:1, v/v.
31. Glycogen (20 µg/µL) (Roche, Mannheim, Germany).
32. 3 M Sodium acetate.
33. 5 M Ammonium acetate.
34. 100% Ethanol.
35. 70% Ethanol.
36. G-50 Spin column (Roche, Mannheim, Germany).
37. T-vector (Promega, WI).
38. Competent *Escherichia coli* (*E. coli*) cells for electroporation.

3. Methods

3.1. Adapter PCR of Genomic DNA

The steps described in this method correspond to step I and step II of **Fig. 1**.

1. Add in the following order: approx 20 µg of genomic DNA, 10 µL of 10X buffer for *Sau3AI*, 5 µL of *Sau3AI*, and up to 100 µL of sterile water. Incubate at 37°C overnight. Add 100 µL of sterile water.
2. Extract by adding 100 µL of phenol that has been preequilibrated with TE buffer. Vortex. Add 100 µL of chloroform:isoamyl alcohol. Vortex. Centrifuge in a microfuge at room temperature for 30 s to separate the phases. Remove the upper (aqueous) phase (approx 200 µL) and place in a fresh sterile 1.5-mL tube.
3. Repeat **step 2** once.
4. Add 0.1 volume of 3 M sodium acetate. Mix well, and add 2 volumes of cold 100% ethanol. Incubate at -80°C for at least 10 min.
5. Recover the restricted genomic DNA by centrifugation in a microfuge (12 000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 400 µL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min. Resuspend the pellet in 50 µL of TE buffer (*see Note 1*).
6. In a sterile 1.5-mL tube, add in the following order: 500 ng of the restricted genomic DNA, each 2.5 µL of 100 µM primer (1) and antiprimer (1), 3 µL of 10X buffer for ligase, and up to 30 µL of sterile water.
7. Incubate at 53°C for 10 min in a heat block.
8. Incubate the tube with the heat block in refrigerator at 4°C for 60 min.
9. Add 1 µL of ligase, and incubate at 12°C overnight.
10. Add 470 µL of sterile water.
11. In a sterile 200-µL PCR tube, add in the following order: 10 µL of the ligation mixture, 5 µL of 20 µM primer (1), 10 µL of 10X buffer for *ExTaq* polymerase (MgCl₂ plus), 8 µL of each 2.5 mM dNTP, 0.5 µL of *ExTaq* polymerase, and 66.5 µL of sterile water.
12. Prepare four tubes for the sample.
13. Amplify by PCR using the following cycle profiles: initial denaturation at 72°C for 5 min; approx 20 cycles of 1 min each at 94°C and 72°C; and then 10 min at 72°C for final extension (*see Note 2*).
14. Extract by adding 200 µL of phenol that has been preequilibrated with TE buffer. Vortex it. Add 200 µL of chloroform:isoamyl alcohol. Vortex and centrifuge it in a microfuge at

- room temperature for 30 s to separate the phases. Remove the upper phase (approx 400 μL) and place in a fresh sterile 1.5-mL tube.
15. Repeat **step 14** once.
 16. Add 0.1 volume of 3 M sodium acetate. Mix well, and add 2 volumes of cold 100% ethanol. Incubate at -80°C for at least 10 min.
 17. Recover the PCR products by centrifugation in a microfuge (12 000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 600 μL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min.
 18. Resuspend the pellet in 50 μL of 1X mung bean nuclease buffer.
 19. Add 1 μL of mung bean nuclease. Incubate at 37°C for 45 min. Add 150 μL of sterile water.
 20. Extract by adding 100 μL of phenol that has been preequilibrated with TE buffer. Vortex. Add 100 μL of chloroform:isoamyl alcohol. Vortex and centrifuge it in a microfuge at room temperature for 30 s to separate the phases. Remove the upper phase (approx 200 μL) and place in a fresh sterile 1.5-mL tube.
 21. Repeat **step 20** once.
 22. Add 0.1 volume of 3 M sodium acetate. Mix well, and add 2 volumes of cold 100% ethanol. Incubate at -80°C for at least 10 min.
 23. Recover the PCR products by centrifugation in a microfuge (12 000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 600 μL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min. Resuspend the pellet in 50 μL of TE buffer.
 24. Purify the PCR product by passing the remaining excess primers through a G-50 spin column by centrifugation (*see Notes 1 and 3*).

3.2. Enrichment Procedure

The steps described in this method correspond to step III, step IV, and step V of **Fig. 1**.

1. In a sterile 200- μL PCR tube, add in the following order: 500 ng of the purified PCR products, 1 μL of 20 μM biotinylated oligo-(CAG)₈, 10 μL of 10X buffer for *rTaq* polymerase (MgCl₂ free), 10 μL of 15 mM MgCl₂, 8 μL of each 2.5 mM d3NTP, and up to 100 μL of sterile water. As negative controls, prepare two reaction mixtures without the purified PCR products or the biotinylated oligo-(CAG)₈, respectively.
2. Denature the sample at 94°C for 3 min to make target microsatellites accessible to the probe, and incubate it for each 5 min at 80°C and 75°C , respectively, to hybridize the biotinylated oligo-(CAG)₈ to the denatured PCR products.
3. Add 2.7 μL of enzyme mixture to the reaction solution kept at 75°C above (*see Note 4*).
4. Incubate it at 77°C for 10 min to carry out the nucleotide substrate-biased polymerase reaction.
5. Chill on ice, then move all the solution to a fresh 1.5-mL tube on ice which contains 200 μL of 2X binding/washing buffer to stop the polymerase reaction.
6. Add 100 μL of streptavidin-coated magnetic beads in the presence of 100 ng/ μL of *ColI* DNA as the competitor.
7. Incubate the sample at 37°C for 30 min (*see Note 5*).
8. Capture the magnetic bead-complexes with a magnet stand, and then remove the supernatant containing excess unbound oligos and noncomplementary sequences from the tube.
9. Wash the sample by adding 500 μL of 1X binding/washing buffer at room temperature. Gently pipet the sample. Capture the magnetic bead-complexes with the magnet stand. Remove the supernatant.

10. Repeat **step 9** twice.
11. Wash by adding 100 μL of 1X binding/washing buffer at 80°C. Gently pipet the sample. Capture the magnetic bead-complexes with the magnet stand. Remove the supernatant.
12. Repeat **step 11** once.
13. Elute ssDNAs, which were containing repeat sequences, by adding and incubating 50 μL of alkaline buffer at 80°C from the biotinylated oligo-(CAG)₈. Gently pipet the sample. Capture the magnetic beads with the magnet stand. Recover the supernatant to a fresh 1.5- μL tube.
14. Repeat **step 13** once.
15. Precipitate the eluted ssDNAs by adding 0.5 μL of glycogen (20 $\mu\text{g}/\mu\text{L}$), 40 μL of 5 M ammonium acetate, vortex briefly, and then 400 μL of cold 100% ethanol. Incubate at -80°C for at least 15 min.
16. Recover the ssDNAs by centrifugation in a microfuge (12 000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 600 μL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min. Resuspend the pellet in 20 μL of TE buffer.

3.3. Adapter PCR for Large Preparation of the Repeat Sequences

The step described in this method corresponds to step VI of **Fig. 1**.

1. In a sterile 200- μL PCR tube, add in the following order: 2 μL of the resuspended ssDNAs, 2 μL of 20 μM primer (1), 5 μL of 10X buffer for ExTaq polymerase (MgCl₂ plus), 4 μL of each 2.5 mM dNTP, 0.25 μL of ExTaq polymerase, and 36.75 μL of sterile water. Also prepare reaction mixtures of the two resuspended products without the purified PCR products or the biotinylated oligo-(CAG)₈, respectively (*see Note 6*).
2. Amplify by PCR using the following cycle profiles: initial denaturation at 94°C for 2 min; 15 cycles of 1 min each at 94°C and 72°C; and then 10 min at 72°C for final extension.
3. As a preliminary experiment, confirm the number of cycles of the second PCR procedure for large preparation. In three fresh sterile 200- μL PCR tubes, add in the following order: 10 μL of the first PCR products of sample, the negative control without the purified PCR products, or the negative control without the biotinylated oligo-(CAG)₈ as templates, 4 μL of 20 μM primer (1), 10 μL of 10X buffer for ExTaq polymerase (MgCl₂ plus), 8 μL of each 2.5 mM dNTP, 0.5 μL of ExTaq polymerase, and 67.5 μL of sterile water.
4. Amplify the sample DNA by PCR using the following cycle profiles: initial denaturation at 94°C for 2 min; approx 13–18 cycles of 1 min each at 94°C and 72°C; and then 10 min at 72°C for final extension.
5. Take 5–10 μL aliquot of the PCR to visualize the products by agarose gel electrophoresis (1.5% agarose). Confirm the number of cycles of the second PCR that does not amplify PCR products of the two negative controls.
6. After the confirmation, amplify the first PCR products by PCR using the same condition and the confirmed cycle profiles.
7. Prepare and amplify four tubes of the PCR products with the condition of **step 6**.
8. Extract by adding 200 μL of phenol that has been preequilibrated with TE buffer. Vortex. Add 200 μL of chloroform:isoamyl alcohol. Vortex. Centrifuge in a microfuge at room temperature for 30 s to separate the phases. Remove the upper phase (approx 400 μL) and place in a fresh sterile 1.5-mL tube.
9. Repeat **step 9** once.
10. Add 0.1 volume of 3 M sodium acetate. Mix well, and add 2 volumes of cold 100% ethanol. Incubate at -80°C for at least 10 min.

11. Recover the PCR products by centrifugation in a microfuge (12 000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 600 µL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min. Resuspend the pellet in 50 µL of TE buffer.
12. Purify the PCR product by passing the remaining PCR reaction through a G-50 spin column by centrifugation (*see Note 1*).

3.4. Second Cycle for Enrichment

The step described in this method corresponds to step I of **Fig. 1**.

1. When the second enrichment procedure was desired, add in the following order: 5 µg of the first enriched PCR products, 10 µL of 10X buffer for *Sau3AI*, 5 µL of *Sau3AI*, and up to 100 µL of sterile water. Incubate at 37°C overnight (*see Note 7*). Add 100 µL of sterile water.
2. Extract by adding 100 µL of phenol that has been preequilibrated with TE buffer. Vortex the sample. Add 100 µL of chloroform:isoamyl alcohol. Vortex and centrifuge it in a microfuge at room temperature for 30 s to separate the phases. Remove the upper phase (approx 200 µL) and place in a fresh sterile 1.5-mL tube.
3. Repeat **step 2** once.
4. Add 0.1 volume of 3 M sodium acetate. Mix well, and add 2 volumes of cold 100% ethanol. Incubate at -80°C for at least 10 min.
5. Recover the PCR products by centrifugation in a microfuge (12,000g) at 4°C for 20 min. Discard the supernatant and wash the pellet with 400 µL of cold 70% ethanol. Carefully discard the supernatant, remove excess liquid from the walls of the tube, and vacuum-dry the pellet in Speed-Vac for 5 min. Resuspend the pellet in 50 µL of TE buffer.
6. Purify the PCR product by passing the remaining PCR through a G-50 spin column by centrifugation.
7. In a sterile 1.5-mL tube, add in the following order: 250 ng of the purified PCR products, each 2.5 µL of 100 µM primer (2) and antiprimer (2), 3 µL of 10X buffer for ligase, and up to 30 µL of sterile water.
8. Incubate at 53°C for 10 min in the heat block.
9. Incubate the tube with the heat block at 4°C for 60 min.
10. Add 1 µL of ligase, and incubate at 12°C overnight.
11. Add 70 µL of sterile water.
12. Repeat **steps 3.1.3–3.1.5, 3.2, and 3.3** for second cycle enrichment.

3.5. Construction of Enrichment Library

The PCR products enriched for (CAG)_n repeats should be constructed by direct cloning into a T-vector using T4 DNA ligase, taking advantage of the 3'-A overhangs often produced by *Taq* polymerase. These recombinants should be transformed into competent *E. coli* cells by electroporation (*see Note 8*).

4. Notes

1. Take an aliquot containing 500 ng of genomic DNA or the PCR products to check the enrichment rate of microsatellites by Southern blot analysis (*see Note 8*).
2. With the adapter PCR procedure, PCR conditions are optimized to generate a smear of the PCR products without specific bands. If some specific bands appear, the adapter sequences (primer 1 or primer 2) should change to new designed adapter sequences. The

enrichment rate of microsatellites would be influenced by the generation of specific bands. By increasing amount of the ligation mixtures, the number of cycles for PCR could be reduced. It is possible that the reduced cycles of the PCR might get rid of the specific bands by the adapter primers.

3. Check the absence of excess primers by agarose gel electrophoresis. If the excess primers are visualized with the electrophoresis, the methods to purify them (*see Subheading 3.1.14–24*) should be repeated until the absence of the excess primers. The presence of the excess primers decreases the efficiency of the nucleotide substrate-biased polymerase reaction.
4. Prepare the fresh enzyme mixtures to prevent the inactivation of the enzymes.
5. Gently mix the reaction tubes, because the magnetic beads precipitate during the incubation.
6. The confirmation of the number of cycles of PCR for large preparation should be carried out for every lot of the experiments. The PCR with superfluous cycles might reduce the enrichment rate of microsatellites.
7. For this method to be repeated, the primer for the second enrichment must be changed from that for the first enrichment. The adapter change prevented amplification from the first primer. Amplification with alternative sets of primers would then produce different specific libraries of sequences complementary to a common target.
8. To analyze the enrichment rate, we performed Southern blot analysis of each product. The results reveal that the enrichment cloning method achieved 10^2 -fold enrichment by the first round procedure. Finally, 10^{2-3} -fold enrichment was achieved for $(CAG)_n$ repeat by the second round. A potential risk of the second round of PCR was considered, because a single DNA fragment was amplified to produce many copies. The results obtained in the previous study, however, indicated that only about 10% of the clones isolated were identical (*11*).

References

1. Weber, J. L. and May, P. E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**, 388–396.
2. Litt, M. and Luty, J. A. (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**, 397–401.
3. Tozaki, T., Sakagami, M., Mashima, S., Hirota, K., and Mukoyama, H. (1995) ECA-3: equine (CA) repeat polymorphism at chromosome 2p1.3–4. *Anim. Genet.* **26**, 283.
4. Marklund, S., Ellegren, H., Eriksson, S., Sandberg, K., and Andersson, L. (1994) Parentage testing and linkage analysis in the horse using a set of highly polymorphic horse microsatellites. *Anim. Genet.* **25**, 19–23.
5. Takahashi, H., Nirasawa, K., Nagamine, Y., Tsudzuki, M., and Yamamoto, Y. (1998) Genetic relationships among Japanese native breeds of chicken based on microsatellite DNA polymorphisms. *J. Hered.* **89**, 543–546.
6. Gerber, H. P., Seipel, K., Georgiev, O., Hofferer, M., Hug, M., Rusconi, S., and Schaffner, W. (1994) Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science* **263**, 808–811.
7. The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's Disease chromosomes. *Cell* **72**, 971–983.

8. Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J. Jr., Servadio, A., Beaudet, A. L., et al. (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* **4**, 221–225.
9. Gastier, J. M., Pulido, J. C., Sunden, S., Brody, T., Buetow, K. H., Murray, J. C., et al. (1995) Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. *Hum. Mol. Genet.* **4**, 1829–1836.
10. Inoue, S., Takahashi, T., and Ohta, M. (1999) Sequence analysis of genomic regions containing trinucleotide repeats isolated by a novel cloning method. *Genomics* **57**, 169–172.
11. Tozaki, T., Inoue, S., Mashima, S., Ohta, M., Miura, N., and Tomita, M. (2000) Sequence analysis of trinucleotide repeat microsatellite from an enrichment library of the equine genome. *Genome* **43**, 354–365.
12. Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., et al. (2000) The isolation and characterization of 18 equine microsatellite loci, TKY272–TKY289. *Anim. Genet.* **31**, 149–150.
13. Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., et al. (2000) The isolation and characterization of 36 equine microsatellite loci, TKY290–TKY323. *Anim. Genet.* **31**, 234–236.
14. Tozaki, T., Mashima, S., Hirota, K., Miura, N., Choi-Miura, N., and Tomita, M. (2001) Characterization of equine microsatellites and microsatellite-linked repetitive elements (eMLREs) by efficient cloning and genotyping method. *DNA Res.* **8**, 33–45.
15. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., et al. (1992) A second-generation linkage map of the human genome. *Nature* **359**, 794–801.

Methylation-Specific PCR

Haruhiko Ohashi

1. Introduction

1.1. Significance of DNA Methylation as an Epigenetic Phenomenon

Methylation of the DNA is an important epigenetic (i.e., not associated with alteration in the primary structure of the DNA) phenomenon, which plays important roles in regulation of gene expression, maintenance of genome integrity, and genomic imprinting. Although not only other nucleotides, but also proteins and lipids can be methylated, in the context of the present discussion, “methylation” designates only that of cytosine residues that are located 5' to guanines (CpG cytosines). A methyl residue is added to the 5 position of the pyrimidine ring of cytosine (5-methylcytosine) in the course of DNA replication, a process mediated by DNA(cytosine-5)-methyltransferases.

1.2. Other Methods for DNA Methylation Analysis

Investigation on DNA methylation had long been performed by Southern hybridization with methylation-sensitive endonucleases. Some endonucleases digest double-strand DNA at their cognitive sequences only when cytosines within are free of methylation, whereas others do so, regardless of the presence or absence of 5-methylcytosines. Genomic DNA is digested by a methylation-sensitive endonuclease and electrophoresed on agarose gel and probed with specific DNA probe. If the cytosines in the endonuclease recognition site is methylated, the DNA is left uncleaved at that site, thus appears band(s) of larger-than-expected size(s).

The main problems with Southern blotting with methylation-sensitive endonucleases is that only the cytosines in the context of available methylation-sensitive endonucleases [i.e., *HpaII* (CCGG), *HhaI* (GCGC)] could be examined; also, relatively large amount of DNA (5–10 μg) is required for each analysis. The latter problem could be circumvented by the use of polymerase chain reaction (PCR). Since cytosine methylation should not be conserved after PCR (5-methylcytosines should be converted to cytosines), digestion by methylation-sensitive endonucleases should be per-

formed *before* amplification by polymerase chain reaction (PCR). Primers should be positioned upstream and downstream of the recognition site of a methylation-sensitive endonuclease. DNA digested by a methylation-sensitive endonuclease is amplified with the primers: If the restriction site is methylated, the template DNA should be left uncleaved, thus the PCR product should be generated; and if the site is unmethylated, because the template is cut into fragments, there should be no product. Even if the template DNA is fully unmethylated at the restriction site, however, only slight incomplete digestion may lead to amplification because of high sensitivity of PCR. Thus, interpretation of the results should be made with caution for this method.

In 1992, Frommer et al. (1) described a method called bisulfite sequencing. This strategy of identifying the presence and absence of cytosine methylation has greatly contributed to the advancement of research on DNA methylation. Treatment of denatured DNA by sodium bisulfite deaminates cytosines at their 4 position, and converts cytosines to uracils. When the bisulfite-treated DNA is subjected to PCR, uracils are converted to thymines. 5-Methylcytosines are resistant to deamination by sodium bisulfite, and converted to cytosines after PCR. Thus, after PCR following bisulfite treatment, unmethylated cytosines are converted to thymines, while methylated cytosines are changed to cytosines (*see Fig. 1*). This means that cytosine methylation, which is an epigenetic modification of the DNA, can be translated into difference in primary structure (base composition) of the DNA. In bisulfite sequencing, the amplified products are cloned and sequenced. Because all the non-CpG cytosines are uniformly unmethylated, only the 5-methylcytosines in CpGs should remain cytosines on sequencing. Comparison with the original DNA sequence shows which cytosines are methylated.

1.3. Basic Concept of Methylation-Specific PCR

Methylation-specific PCR (MSP), which was first described by Herman et al. (2) in 1996, is an application of bisulfite sequencing method. For a sequence in a gene containing CpGs, the allele on which those CpGs are methylated and another on which those CpGs are unmethylated should give different sequences after bisulfite modification. When a primer set that are complementary to the sequence with methylated CpGs, but are not complementary to the originally same sequence with unmethylated CpGs, is used for PCR, only the sequence (allele) with methylated CpGs should be amplified. The same is true for the primer pair specific for sequence with unmethylated CpGs. The interpretation of the result is simple: If PCR product of the expected size is seen on agarose gel electrophoresis, the sample is considered to contain the methylated or unmethylated allele of the gene, depending on the primer pair used. Usually, primer pairs specific for methylated and unmethylated sequences, respectively, are used for the same gene, and the amplified products are run side-by-side on agarose gel for comparison.

1.4. Application of MSP

The greatest advantage of MSP over other methods of DNA methylation analysis is its simplicity. In contrast to Southern blotting with methylation-sensitive restriction enzymes, MSP requires much less amount of DNA, isotope use is usually unneces-

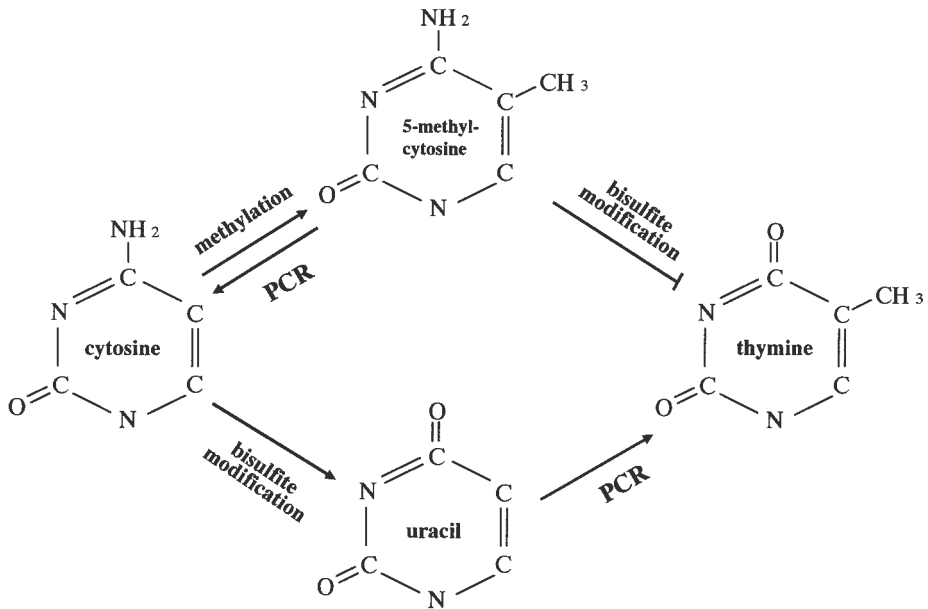


Fig. 1. Conversion of pyrimidines residues by chemical and biochemical reactions. Changes between pyrimidine residues by methylation, deamination, and PCR are being shown. Cytosines in DNA strands can be methylated to be 5-methylcytosines by DNA(cytosine-5)-methyltransferases in vivo, or by other methylases in vitro. Sodium bisulfite deaminates cytosines to uracils, whereas 5-methylcytosines are resistant to modification. PCR with dCTP as nucleotide source converts 5-methylcytosines to cytosines and uracils to thymines, respectively.

sary, and any CpGs, regardless of the sequence around, can be evaluated. Also, interpretation of the results is much more straightforward. MSP is superior to bisulfite sequencing in that it does not require cloning and sequencing, which usually take several days, and can be done in 1 or 2 d. All these characteristics fit well to researches that examine a large number of clinical samples at a time, and indeed it has been widely used for detection of aberrant hypermethylation and inactivation of tumor suppressor genes in cell lines and tumor samples (3). This method can also be used to evaluate methylation status of any DNA sequences, such as viral genes (4), and imprinted X-linked and autosomal genes (5 and 6).

2. Materials

2.1. Bisulfite Modification of the DNA

1. Hydroquinone (Sigma, St. Louis, MO).
2. Sodium bisulfite (Sigma).
3. Wizard DNA purification resin (Promega, Madison, WI), or similar product.
4. CpGenome DNA Modification Kit (Intergen, Purchase, NY).
5. 3 M Sodium hydroxide (NaOH).

6. 70 % Ethanol.
7. TE Buffer: 10 mM Tris-HCl, 0.1 mM ethylenediamine tetraacetic acid (EDTA), pH 7.5.
8. 5 M Ammonium acetate (NH₄OA_c).
9. 5 M Sodium chloride (NaCl).

2.2. PCR with Methylation-Specific Primers

1. 10X PCR reaction buffer: 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 15 mM MgCl₂, (TaKaRa, Osaka, Japan).
2. dNTP mixture: 2.5 mM each (TaKaRa).
3. *Taq* DNA polymerase (TaKaRa).

3. Methods

The procedure of the MSP consists of two steps: bisulfite modification of the sample DNA, and PCR with methylation-specific primers. Chemical modification of double-strand DNA by sodium bisulfite consists of four steps:

1. Denaturation of the double-strand DNA by NaOH (*see Note 1*);
2. Sulfonation of the 6 position of cytosine by sodium bisulfite (cytosine sulfonate);
3. Hydrolytic deamination at 3 position by hydroquinone (uracil sulfonate);
4. Alkali desulfonation by NaOH (uracil).

These reactions can be carried out with the protocol described below. Alternatively, a kit for bisulfite modification (CpGenome DNA Modification Kit, Intergen) is commercially available. It is important to note that each single-strand DNA (the sense strand and the antisense strand), generated as the result of initial denaturation, is modified independently by sodium bisulfite in the subsequent steps. This means that the sense and antisense strands are no longer complementary to each other after modification. Bisulfite-modified DNA can be stored at -20°C and used for repeated analyses.

PCR with methylation-specific primers is not different from ordinary PCR, but needs some precautions (*see Note 2*). There may be three sets of methylation-specific primers for each gene, or a region of a gene: primers that amplify sequences on which CpGs are methylated (M-primers); primers that amplify sequences on which CpGs are unmethylated (U-primers); and primers that are designed for the region without CpGs and should amplify the gene regardless of methylation status (C-primers, C for common). M- and U-primers are the ones used for MSP, and C-primers are for bisulfite sequencing. There could be another set of primers, W-primers (W for wild-type), which amplify DNA not modified by sodium bisulfite, or escaped modification for some technical defects. PCR with the forward M-primer and the reverse M-primer should detect, if present, the gene that are methylated at these CpGs. PCR with the forward M-primer and the reverse C-primer, or the other way around, should also work for the purpose. Because selecting regions fit for M- and U-primers is sometimes difficult, while the region without CpGs can usually be found, combination of either the forward M-primer or the forward U-primer, and the reverse C-primer may be a practical choice.

In fact, it is fairly easy to perform MSP on a gene for which workable methylation-specific primer sequences are already reported. If one wishes to examine a gene for which MSP has not been performed by others, the difficult part is selecting methylation-specific primers. The primers used in MSP should be designed so that they dis-

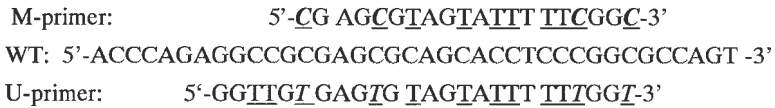


Fig. 2. Example of primer selection for MSP. MSP primers for the human androgen receptor gene (HUMARA) is being shown (5), as an example of primer designing. WT: Unmodified sense strand sequence rich in CpG and non-CpG cytosines. M-primer: Primer specific for the methylated allele. All the cytosines except for those preceding guanines are changed to thymines, while cytosines 5' to guanines remain unchanged. The *T*s are thymines converted from cytosines. The *C*s are CpG cytosines that are methylated and are not to be converted to thymines. U-primer: Primer specific for the unmethylated allele. All the cytosines are changed to thymines. The *T*s are thymines converted from cytosines. The *T*s are CpG cytosines that are unmethylated and are to be converted to thymines.

criminate between methylated and unmethylated sequences *and*, at the same time, between bisulfite-modified and bisulfite-unmodified sequences. Thus, the primers should be designed for the region rich in both CpGs and non-CpG cytosines. Both U- and M-primers must contain *T*s that are non-CpG *C*s in the unmodified (wild-type) sequence. U-primers must have *T*s located 5' to *G*s, at their 3' ends, and M-primers must contain *C*s in the CpG context at their 3' ends. **Fig. 2** shows, as an example, the genomic sequence and the methylation-specific primers designed for the human androgen receptor gene (5). If the purpose of the investigation is to know whether the expression of a gene is regulated by methylation, the part of the gene to be examined is the promoter region of the gene, where clustering of CpGs (CpG islands) is found about half of the genes. Sometimes, however, the CpG island in the 5' region of a gene may extend to the 5' untranslated region (5'-UTR) or even to the coding region.

3.1. Bisulfite Modification of the DNA

1. Take 1 μ g of DNA and add double-distilled water to 50 μ L in a 1.5-mL tube.
2. Add 3.5 μ L of 3 *M* NaOH (final concentration of 0.2 *M*), and incubate at 37°C or 10 min.
3. Freshly prepare 10 *mM* hydroquinone, and add 35 μ L of it to the tube.
4. Freshly prepare 3 *M* sodium bisulfite, adjust to pH 5.0 by adding 3 *M* NaOH, add 520 μ L of it to the tube, mix well, overlay with mineral oil, and incubate at 50°C for 16 h or longer.
5. Put the tube on ice, add 5 μ L of Wizard DNA purification resin (or similar DNA purification resin), mix, and incubate at room temperature for 10 min, spin at 5000*g* for 10 s at room temperature, discard the supernatant.
6. Put 1 mL of 70% ethanol, mix by Vortex, spin at 5000*g* for 10 s, and discard the supernatant. Do this two more times, and remove the supernatant completely.
7. Add 50 μ L of TE, mix well, incubate at 50°C for 5 min, spin at 12,000*g* for 1 min, and transfer the supernatant to a fresh 1.5-mL tube.
8. Add 5.5 μ L of 3 *M* NaOH, mix, and incubate at room temperature for 5 min.
9. Add 10 μ L of 5 *M* NH₄OAc and mix (for neutralization).
10. Add 1 μ L of 5 *M* NaCl and 200 μ L of 100% ethanol, mix well, keep the tube at -20°C for 1 h. Spin at 12 000*g*, at 4°C for 5 min, discard the supernatant, rinse the pellet with 70% ethanol, dry the pellet, and dissolve in 20–50 μ L of TE, and store at -20°C.

3.2. PCR with Methylation-Specific Primers

1. Make mixture of the following for the number of samples to be examined, and put the mixture into a 0.5-mL tube for PCR (*see Note 3*).

a. 10X PCR reaction buffer (15 mM Mg ²⁺)	2 μ L;
b. 2.5 mM dNTP mixture	2 μ L;
c. forward primer (10 μ M)	1 μ L;
d. reverse primer (10 μ M)	1 μ L;
e. ddH ₂ O	10 μ L.

 Overlay with mineral oil.
2. Add the bisulfite-modified DNA (2 μ L), or a control sample (*see Note 4*) into the tube.
3. Put the tubes on a thermal cycler and start denaturation at 95°C for 5 min.
4. Make mixture of the following for the number of tubes.

a. <i>Taq</i> DNA polymerase (5 U/ μ L)	0.2 μ L;
b. ddH ₂ O	1.8 μ L.

 Take a tube from the thermal cycler, add 2 μ L of the *Taq* mixture to each tube through mineral oil, and put the tube back to the heat block. This should be done as quick as possible.
5. Go through 30–40 cycles of the following amplification (*see Note 3*).
 - a. Denaturation 95°C for 30 s;
 - b. Annealing 55°C for 30 s;
 - c. Extension 72°C for 30 s.
6. Take 5–10 μ L of the reaction solution and run on 2–3% agarose gel.

4. Notes

1. Complete denaturation of the DNA in the first step of bisulfite modification is important, because unmethylated cytosines in undenatured DNA cannot be converted to uracils, thus leads to misinterpretation that they are methylated. For bisulfite sequencing of specific genes or plasmids, digestion of the template DNA by an endonuclease that cleaves outside the region to be amplified is recommended, in order to ensure complete denaturation (*I*). This pretreatment of DNA does not fit to examination for clinical samples, with which more than one genes may be examined for methylation. Another strategy for ensuring initial denaturation is to shear DNA through fine gage needles. Our experience with MSP for various genes *without* such pretreatment, however, did not show inconsistent results, and it seems that pretreatments may not be necessary. One can always check, if needed, the validity of bisulfite treatment of the given gene by PCR with W-primers or by sequencing MSP product.
2. By experience amplification efficiency with MSP is lower than with ordinary PCR, and one often needs to optimize PCR condition, in the direction of lowering stringency, to see clearer bands. The situation sometimes becomes tricky, because the M-primers and U-primers are designed to recognize sequences only partially different to each other. When one lowers the stringency of PCR and undergo a larger number of cycles to get clearer bands, they may notice that specificity of the reaction is lost: the same band can be seen with the control samples with the opposite specificity (e.g., M controls for U-primers). This may be the greatest pitfall for methylation analysis with MSP.
3. The factors that appear to influence PCR efficiency are reaction buffer [Mg²⁺ concentration, addition of demethyl sulfoxide (DMSO)], annealing temperature, and hot start regimens. We usually start with commercially available 10X PCR reaction buffer (10 mM

Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, final concentration), annealing temperature of 55°C, and hot start by adding *Taq* polymerase after incubating the reaction for 5 min at 95°C. When no product of the expected size is seen after agarose electrophoresis, elevation of MgCl₂ concentration, addition of DMSO (5% of the volume) should be tried. Elevating the annealing temperature may also help. Hot start appears to be critical for MSP, thus we usually do not even try nonhot start protocols. When ordinary hot start protocol of adding *Taq* after incubation at 95°C does not work, use of reagents designed for hot start, such as Platinum *Taq* (Gibco-BRL, Rockville, MD) may still help. Optimization of PCR conditions for MSP seems much more difficult than for ordinary PCR, and it may sometimes be wiser to redesign primer sequences before comparing all the possible conditions.

4. Not only for initial optimization of MSP but also for each experiment, as for any PCR, control templates are necessary. To obtain the perfect controls for MSP is sometimes difficult, since one needs both DNA that is methylated at the CpGs of the gene of interest, and DNA that is unmethylated in the same region. The ideal controls for U-primers are the PCR-amplified DNA segment, because PCR products are fully unmethylated. If one methylates the PCR product with methylases in vitro, that would be a good control for M-primers. There are several kinds of commercially available bacterial methylases with different sequence specificity. For example, Sss I methylase (New England Biolabs, Beverly, MA) methylates all the CpG cytosines. All these preparations are time consuming, but may be necessary for analysis of nonhuman (i.e., viral) genes. When DNA from some sources (e.g., cell lines) in which the gene in question is known to be either methylated or unmethylated is available, that could serve as controls. When one possesses DNA that is unmethylated for the gene, then in vitro treatment with methylases can provide controls for methylated gene.

References

1. Frommer, M., McDonald, L. E., Miller, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **89**, 1827–1831.
2. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D., and Baylin, S. B. (1996) Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. USA* **93**, 9821–9826.
3. Herman, J. G., Civin, C. I., Issa, J. P., Collector, M. I., Sharkis, S. J., and Baylin, S. B. (1997) Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies. *Cancer Res.* **57**, 837–841.
4. Tao, Q., Swinnen, L. J., Yang, J., Strivastava, G., Robertson, K. D., Ambinder, R. F. (1999) Methylation status of the Epstein-Barr virus major latent promoter C in iatrogenic B cell lymphoproliferative disease. Application of PCR-based analysis. *Am. J. Pathol.* **155**, 619–625.
5. Uchida, T., Ohashi, H., Aoki, E., Nakahara, Y., Hotta, T., Murate, T., Saito, H., and Kinoshita, T. (2000) Clonality analysis by methylation-specific PCR for the human androgen-receptor gene (HUMARA-MSP). *Leukemia* **14**, 207–212.
6. Kubota, T., Das, S., Christian, S. L., Baylin, S. B., Herman, J. G., Ledbetter, D. H. (1997) Methylation-specific PCR simplifies imprinting analysis. *Nature Genet.* **16**, 16–17.

Direct Cloning of Full-Length Cell Differentially Expressed Genes by Multiple Rounds of Subtractive Hybridization Based on Long-Distance PCR and Magnetic Beads

Xin Huang, Zhenglong Yuan, and Xuetao Cao

1. Introduction

Subtractive cloning is an ideal technique for identifying genes differentially expressed in two nuclear acids population (*1*). The polymerase chain reaction (PCR)-based subtraction is the method of choice when the starting samples are heterogeneous or difficult to obtain, which often occurs in the tissues to be compared. PCR amplification is the easiest method for generating adequate amount of nuclear acids for multiple-round hybridization. However, the bias in the relative representation of mRNA molecules in the starting materials and the accumulation of shorter fragments become the major deficiencies for this method and should be overcome. The bias caused by PCR amplification is because of the tendency of preferentially amplifying short fragments and certain templates with unique sequences in the sample. The thermophilic polymerase that is optimized to amplify multiple genes would be helpful and the adoption of gel filtration in preparation of templates for amplification could hinder the tendency of short fragment accumulation. In addition, increasing the amount of starting samples would represent much more molecules in tissues.

For cloning differentially expressed genes with full-length, the tracer (from which to isolate differentially expressed sequences) is expected to generate full-length cDNAs. A reverse transcription (RT) method based on switching mechanism at 5' end of RNA template (SMART) can generate higher percentage of full-length cDNAs than that produced by conventional method (*2,3*). Amplifying the cDNAs by long-distance PCR (LD-PCR) (*4*) could improve the size of the product span and thus increase the proportion of full-length sequences. The requirement for the driver (which is used to hybridize tracer to remove common sequences) is different to that of the tracer: short fragments are better than long fragments because they have high mobility and could

approach tracer DNA easily in hybridization. In such circumstances, the limitation of PCR that small fragments are preferentially amplified is no longer a limitation.

The hybridization step is very important to the success of the subtractive cloning and the tracer:driver ratio is an important parameter controlling the efficiency of hybridization. The ratio should be at least 1:10 to enable the driver to govern the hybridization and much higher ratio would be better. We recommend the tracer:driver ratio be at least 1:50 to 1:100. It is also important to remove driver-tracer hybrids and excess driver completely. Driver-tracer hybrids and excess driver could be removed efficiently by using streptavidin paramagnetic beads if the driver is biotinylated because streptavidin can bind biotin with high affinity ($K_d = 10^{-12}$) and biotin-streptavidin system had been found to be the most competent method for subtractive and positive selection (5,6).

Taken together, to improve the efficiency of subtractive cloning and make the technique less laborious, some preeminent methods were adopted to design a convenient, high-efficiency strategy for cloning of differentially expressed genes with full-length directly. First, a SMART method was employed, which not only can increase the population of full-length cDNAs, it can ligate the cDNAs with primers at the 5' and 3' termini as well when mRNA is being reverse transcribed. Second, the LD-PCR was adopted for the amplification of full-length cDNA. Third, to separate the tracer-driver hybrids and excess driver efficiently, the biotin-streptavidin system was used. Finally, to reduce the accumulation of shorter fragments, a Sepharcyl S-400 spin column chromatography was very convenient and suitable.

2. Materials

2.1. Isolation of mRNA from Cells

mRNA purification kit (Promega, Madison, WI, cat. no. MZ5400) including the following.

1. GTC extraction buffer: 4 M guanidine thiocyanate, 25 mM sodium citrate, pH 7.1.
2. 48.7 % β -mercaptoethanol.
3. Dilution buffer: 22.5 mM NaCl, 11.25 mM sodium citrate, 10 mM Tris-HCl, pH 7.4, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.25 % sodium dodecyl sulfate (SDS).
4. 50 μ M Biotinylated Oligo(dT) probe.
5. Streptavidin MagneSphere Paramagnetic Particles (SA-PMPs).
6. 0.5X SSC: 75 mM NaCl, 37.5 mM sodium citrate.
7. Nuclease-free water.
8. 3 M NaAc, nuclease-free.
9. Isopropanol.
10. 70% ethanol.

2.2. Preparation of Tracer

1. 5X first-strand buffer: 250 mM Tris-HCl, pH 8.3, 375 mM KCl, 30 mM MgCl₂, 50 mM dithiothreitol (DTT).
2. 10 mM dNTP.
3. rRNasin (40 U/ μ L) (Promega, Madison, WI, cat. no. N2511).
4. SUPERSCRIPT II RT (200 U/ μ L) (Gibco-BRL, Gaithersburg, MD, cat. no. 18064-014).

5. 3' Anchoring primer: 5' TACGGCTGCGAGAAGACGACAGAAT₍₃₀₎VN-3'.
6. CAP-oligo: 5' TACGGCTGCGAGAAGACGACAGAAGGG-3' (Clontech, Palo Alto, CA).
7. Tracer primer: 5' TACGGCTGCGAGAAGACGACAGAA-3'.
8. 50X Advantage cDNA polymerase mix (Clontech, cat. no. 8417-1).
9. 10X PCR buffer: 400 mM Tricine-KOH pH 9.2, 150 mM KAc, 35 mM Mg(Ac)₂, 37.5 µg/mL BSA.

2.3. Preparation of Driver

1. Avian myeloblastosis virus reverse transcriptase (AMV RT), 9 U/µL (Promega, cat. no. M5101).
2. 5X AMV RT buffer: 250 mM Tris-HCl, pH 8.3, 250 mM KCl, 50 mM MgCl₂, 50 mM DTT, 2.5 mM spermidine.
3. 40 mM sodium pyrophosphate.
4. Recombinant *Taq* polymerase (any brand).
5. 100 µM Oligo(dT)₁₅.
6. 3' anchoring primer: 5' T₍₁₅₎(A/G/C)-3'.
7. Biotin-21-dUTP (Clontech, cat. no. 5021).
8. 10X PCR buffer: 100 mM Tris-HCl, pH 9.0, 500 mM KCl, 15 mM MgCl₂, 1% Triton X-100.

2.4. Hybridization

1. 20X SSC: 3 M NaCl, 1.5 M sodium citrate.
2. 10% SDS.
3. Sephacryl S-400 HR (Amersham Pharmacia Biotech, Uppsala, Sweden, cat. no. 17-0609-10).
4. Spin column (Promega, cat. no. C128a).
5. TEN: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 25 mM NaCl.
6. SA-PMPs (Promega, cat. no. Z5482).
7. PCR primer: 5' TACGGCTGCGAGAAGACGACAGAA-3'.
8. 50X Advantage cDNA polymerase mix (Clontech, cat. no. 8417-1).
9. 10X PCR buffer: 400 mM Tricine-KOH, pH 9.2, 150 mM KAc, 35 mM Mg(Ac)₂, 37.5 µg/mL BSA.

2.5. Cloning of Differentially Expressed Genes

1. AmpliTaq Gold™ (Perkin-Elmer, Foster City, CA, cat. no. N808-0241).
2. GeneAmp 10X PCR buffer: 100 mM Tris-HCl, pH 8.3, 500 mM KCl.
3. 25 mM MgCl₂.
4. pGEM-T vector systems (Promega, cat. no. A3600).
5. Isopropylthio-β-D-galactoside (IPTG).
6. 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal).
7. DH5α *Escherichia coli* (*E. coli* strain).
8. 0.1 M CaCl₂.
9. Restriction endonucleases (New England Biolabs, Beverly, MA).
10. Sequencing primers: SP6, 5' TCAAGCTATGCATCCAAC-3'; T7, 5' TCACTATAGGGC GAATTG-3'.
11. BigDye terminator ready reaction mix (Perkin-Elmer, cat. no. 4303154).
12. 95% ethanol.
13. Sequencing loading buffer.

2.6. Evaluation of the Efficiency of Subtraction

1. Hybond-N™ Nylon membrane (Amersham, Little Chalfont, Buckinghamshire, UK, cat. no. RPN. 303N.)
2. BLUGENE® Nonradioactive Nucleic Acid Detection System, which contains:
 - a. SA-AP conjugate (1 mg/mL in 3 M NaCl, 1 mM MgCl₂, 0.1 mM ZnCl₂, 30 mM triethanolamine pH 7.6)
 - b. Nitroblue tetrazolium (NBT) (75 mg/mL in 70% dimethylformamide).
 - c. 5-bromo-4-chloro-3-indolylphosphate (BCIP) (50 mg/mL in dimethyl-formamide).
3. Formamide.
4. 20X SSC: 3 M NaCl, 1.5 M Sodium Citrate.
5. 50X Denhardt's solution: 1% Ficoll 400, 1% polyvinylpyrrolidone, 1% albumin bovine fraction V.
6. 0.5 M sodium phosphate (pH 6.5): 65.25 g NaH₂PO₄·2H₂O/L, 29.2 g Na₂HPO₄·12H₂O/L.
7. Shared herring sperm DNA.
8. Dextran sulfate.
9. Prehybridization solution: 50% formamide, 5X SSC, 5X Denhardt's solution, 20 mM sodium phosphate pH 6.5, 0.5 mg/mL freshly denatured sheared herring sperm DNA.
10. Hybridization solution: 45% formamide, 5X SSC, 1X Denhardt's solution, 20 mM sodium phosphate pH 6.5, 0.2 mg/mL freshly denatured sheared herring sperm DNA, 5% dextran sulfate.
11. 7.5 M NH₄Ac.
12. 5% SDS.
13. HYBRITUBE™ (Gibco-BRL, Cat. 20116-018, 10117-018).
14. Buffer 1: 0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl.
15. Buffer 2: 3% albumin bovine fraction V, 0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl.
16. Buffer 3: 0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, 50 mM MgCl₂.
17. Stopping solution: 20 mM Tris-HCl, pH 7.5, 0.5 mM Na₂EDTA.

3. Methods (see Note 1)

3.1. Isolation of mRNA from Cells

Messenger RNA was isolated directly from the crude cells with the PolyAtract® System 1000 (Promega). About 1×10^6 to 1×10^7 peripheral blood mononuclear cells (PBMCs) or PHA stimulated PBMCs were harvested for mRNA purification each time (see Note 2).

1. Wash cells with PBS twice and collect cells in a microcentrifuge tube.
2. Add 400 μ L of GTC extraction buffer into the microcentrifuge tube and 33 μ L of β -mercaptoethanol (48.7%). Homogenize the cells completely with a small homogenizer.
3. Add 800 μ L of dilution buffer and 1.2 μ L of Biotinylated Oligo(dT) Probe (50 μ M) and mix well. Incubate this mixture at 70°C for 5 min.
4. Centrifuge this mixture at 12,000g for 10 min at room temperature and collect the supernatant.
5. Wash 1 mL of SA-PMPs twice with equal volume of 0.5X SSC and suspend in 1 mL of 0.5X SSC. Mix the SA-PMPs with the supernatant from **step 4** and incubate at room temperature for 2 min.
6. Wash the SA-PMPs with 0.5X SSC for three times by magnetic capturing.

7. Add 200 μL of nuclease-free water to suspend the SA-PMPs completely and incubate at room temperature for 2 min.
8. Collect the supernatant and add 20 μL of 3 M NaAc, 220 μL of isopropanol. Mix the contents well and incubate at -70°C overnight.
9. Centrifuge the mixture at 15,000g for 15 min at 4°C .
10. Add 1 mL of 70% ethanol and centrifuge at 15,000g for 5 min at 4°C .
11. Dry the pellet briefly and elute it with nuclease-free water. Store at -70°C for later use.

3.2. Preparation of Tracer

3.2.1. First-Strand cDNA Synthesis

The first strand cDNA was synthesized in a volume of 20 μL (see **Note 3**).

1. Elute about 1 μg of mRNA in 11 μL of nuclease-free water and add 4 μL of 5X first-strand buffer, 1 μL of 10 mM dNTP, 1 μL of rRNasin (40 U/ μL), 1 μL of 3' Anchoring primer (10 μM), 1 μL of CAP-oligo (10 μM), 1 μL of SUPERSCRIPT II RT (200 U/ μL).
2. Mix the contents well and incubate at 42°C for 1 h.
3. Inactivate the reverse transcriptase by heating at 75°C for 15 min.

3.2.2. Amplification of Tracer

Advantage cDNA polymerase mix, which includes KlenTaq-1 polymerase and a minor quantity of DeepVent polymerase, was used to amplify tracer in order to improve the proportion of full-length sequences (see **Note 4**).

1. 5 μL of synthesized cDNA was used as template and mixed with 1 μL of Advantage cDNA polymerase mix, 5 μL of reaction buffer, 10 nmole of dNTP and 20 pmole of tracer primer (see **Subheading 2.2**). Add water to a total volume of 50 μL .
2. PCR condition: 95°C for 1 min; 95°C for 10 s, 68°C for 4 min, repeat for 15 cycles; 68°C for 10 min (see **Note 5**).

3.3. Preparation of Driver

3.3.1. First-Strand cDNA Synthesis (see **Note 6**)

The first-strand cDNA of driver was synthesized by using AMV reverse transcriptase according to the routine method (7) and the manufacturer's instruction manual.

1. Elute about 1 μg of mRNA in 10 μL of nuclease-free water and mix with 4 μL of 5X AMV RT reaction buffer, 1 μL of 10 mM dNTP, 2 μL of 40 mM sodium pyrophosphate, 1 μL of 100 μM oligo(dT), 2 μL of AMV RT. Mix well and incubate at 42°C for 1 h.
2. Inactivate AMV RT by boiling the reaction for 3 min.

3.3.2. Preparation of Driver Labeled with Biotin

Single-strand driver was amplified by priming with single primer and labeled with biotin-21-dUTP simultaneously. In 100 μL PCR reaction volume, add 40 nmole of dNTP, 2 nmole of Biotin-21-dUTP, 100 pmole of 3' anchoring primer (see **Subheading 2.3**), 2 μL of synthesized cDNA, 20 U of Taq polymerase, 10 μL of 10X PCR buffer. PCR condition: 95°C for 10 s, 42°C for 2 min, 72°C for 1 min and repeat for 40 cycles.

7. Add 200 μL of nuclease-free water to suspend the SA-PMPs completely and incubate at room temperature for 2 min.
8. Collect the supernatant and add 20 μL of 3 M NaAc, 220 μL of isopropanol. Mix the contents well and incubate at -70°C overnight.
9. Centrifuge the mixture at 15,000g for 15 min at 4°C .
10. Add 1 mL of 70% ethanol and centrifuge at 15,000g for 5 min at 4°C .
11. Dry the pellet briefly and elute it with nuclease-free water. Store at -70°C for later use.

3.2. Preparation of Tracer

3.2.1. First-Strand cDNA Synthesis

The first strand cDNA was synthesized in a volume of 20 μL (see **Note 3**).

1. Elute about 1 μg of mRNA in 11 μL of nuclease-free water and add 4 μL of 5X first-strand buffer, 1 μL of 10 mM dNTP, 1 μL of rRNasin (40 U/ μL), 1 μL of 3' Anchoring primer (10 μM), 1 μL of CAP-oligo (10 μM), 1 μL of SUPERSCRIPT II RT (200 U/ μL).
2. Mix the contents well and incubate at 42°C for 1 h.
3. Inactivate the reverse transcriptase by heating at 75°C for 15 min.

3.2.2. Amplification of Tracer

Advantage cDNA polymerase mix, which includes KlenTaq-1 polymerase and a minor quantity of DeepVent polymerase, was used to amplify tracer in order to improve the proportion of full-length sequences (see **Note 4**).

1. 5 μL of synthesized cDNA was used as template and mixed with 1 μL of Advantage cDNA polymerase mix, 5 μL of reaction buffer, 10 nmole of dNTP and 20 pmole of tracer primer (see **Subheading 2.2**). Add water to a total volume of 50 μL .
2. PCR condition: 95°C for 1 min; 95°C for 10 s, 68°C for 4 min, repeat for 15 cycles; 68°C for 10 min (see **Note 5**).

3.3. Preparation of Driver

3.3.1. First-Strand cDNA Synthesis (see **Note 6**)

The first-strand cDNA of driver was synthesized by using AMV reverse transcriptase according to the routine method (7) and the manufacturer's instruction manual.

1. Elute about 1 μg of mRNA in 10 μL of nuclease-free water and mix with 4 μL of 5X AMV RT reaction buffer, 1 μL of 10 mM dNTP, 2 μL of 40 mM sodium pyrophosphate, 1 μL of 100 μM oligo(dT), 2 μL of AMV RT. Mix well and incubate at 42°C for 1 h.
2. Inactivate AMV RT by boiling the reaction for 3 min.

3.3.2. Preparation of Driver Labeled with Biotin

Single-strand driver was amplified by priming with single primer and labeled with biotin-21-dUTP simultaneously. In 100 μL PCR reaction volume, add 40 nmole of dNTP, 2 nmole of Biotin-21-dUTP, 100 pmole of 3' anchoring primer (see **Subheading 2.3**), 2 μL of synthesized cDNA, 20 U of Taq polymerase, 10 μL of 10X PCR buffer. PCR condition: 95°C for 10 s, 42°C for 2 min, 72°C for 1 min and repeat for 40 cycles.

3.4. Subtractive Hybridization

3.4.1. First Round of Subtractive Hybridization

1. The first-round hybridization was performed with 5 μL of amplified tracer (*see Subheading 3.2.*) and 100 μL of amplified driver (*see Subheading 3.3.*). Add 20X SSC to a final concentration of 3X SSC and add SDS to a final concentration of 0.1%.
2. Incubate the reaction at 65°C for 12 h.
3. Fill the spin column with 1 mL of Sephacryl S-400, wash Sephacryl S-400 with 1 mL of TEN by centrifugation at 1600g for 4 min.
4. Repeat washing twice.
5. Apply 50 μL of reaction to the top of the gel and centrifuge at 1600g for 4 min. Collect the eluted sample.
6. Wash 1 mL of SA-PMPs with equal volume of 0.5X SSC and suspend in 200 μL of 0.5X SSC. Mix the SA-PMPs with the sample obtained from **step 5**. Incubate the mixture at room temperature for 5 min and then incubate at 65°C for 5 min.
7. Put the mixture in magnetic field instantly and collect the supernatant after the SA-PMPs congregated into pellet. This supernatant would be used as template in second-round hybridization.

3.4.2. Second- and Third- round Subtractive Hybridization

1. Amplify first round subtractive sample as follows: mix 5 μL of template with 1 μL of Advantage cDNA polymerase mix, 5 μL of reaction buffer, 10 nmole of dNTP and 20 pmole of PCR primer in 50 μL reaction volume; run the PCR at 95°C for 1 min, then 95°C for 10 s, 68°C for 4 min, repeat for 15 cycles, followed with extension at 68°C for 10 min.
2. Mix 5 μL of the amplified reaction with 100 μL of amplified driver (*see Subheading 3.3.*), and perform the subtractive hybridization as described in **Subheading 3.4.1**.
3. Repeat **steps 1** and **2** once (*see Note 7*)
4. The supernatant obtained from the final round subtractive hybridization would be used as the template for cloning differentially expressed genes.

3.5. Cloning of Differentially Expressed Genes

1. The subtracted template was amplified by PCR with AmpliTaq GOLD polymerase. Mix 20 μL of subtracted template with 10 nmole of dNTP, 20 pmole of tracer primer (*see Subheading 2.2.*), 2 U of AmpliTaq GOLD polymerase, 5 μL of GeneAmp 10X PCR buffer, 3 mM MgCl_2 , adjust the reaction volume to 50 μL with water. PCR condition: 94°C for 10 min; 95°C for 10 s and 68°C for 3 min, repeat for 20 cycles; 72°C extension for 10 min.
2. Run the PCR products on 1.2 % agarose and retrieve the DNA ranging from 1 to 3 kb.
3. Clone the retrieved DNA into pGEM-T vector and transform DH5 α competent cells. Select transformants by blue/white selection and restriction enzymes digestion.
4. Mix 100 ng of purified plasmid DNA with 1.6 pmole of sequencing primers and 4 μL of BigDye terminator ready reaction mix in 10 μL of reaction volume. The sequencing PCR condition is 96°C for 10 s, 50°C for 5 s, 60°C for 4 min, repeat for 25 cycles.
5. Add 1 μL of 3 M NaAc, 25 μL of 95% ethanol into the sequencing reaction. Mix them well and incubate at 4°C for 15 min. Centrifuge the mixture at 12,000g for 15 min at 4°C. Aspirate the supernatant carefully, add 150 μL of 70% ethanol and centrifuge at 12,000g for 10 min at 4°C. Aspirate the supernatant and dry the pellet briefly.
6. Dissolve the pellet in 2 μL of loading buffer, denature at 90°C for 2 min and load onto gel for sequencing.

3.6. Evaluation of the Efficiency of Subtraction

In our experiments, the efficiency of subtraction was evaluated by dot blotting with BLUGENE[®] nonradioactive nucleic acid detection system (Gibco-BRL) (*see Note 8*).

1. Adjust the concentration of DNA samples to be detected to 100 ng/ μ L. Denature the samples by boiling for 2 min and spot 1 μ L of each sample on the 1-cm² squares of nylon membranes.
2. Dry membranes for 2 h at 80°C.
3. Prepare biotin-labeled DNA probes by purifying driver with repeated ethanol precipitation.
4. Soak the membrane in 2X SSC.
5. Add freshly denatured sheared herring sperm DNA to prehybridization solution, mix, and place in a HYBRITUBE (Gibco-BRL) with the membranes. Incubate at 42°C water bath for 2 h.
6. Add heat-denatured probe and herring sperm DNA to hybridization solution at probe concentration of 100 ng/mL and mix. Remove prehybridization solution and add the hybridization solution. Hybridize at 42°C overnight.
7. Wash the membranes with 250 mL of 2X SSC/0.1% SDS for 3 min at room temperature twice, with 250 mL of 0.2X SSC/0.1% SDS for 3 min at room temperature twice, and then with 250 mL of 0.16X SSC/0.1% SDS for 15 min at 65°C twice. Finally rinse the membranes in 2X SSC briefly.
8. Rinse the membranes in Buffer 1.
9. Incubate the membranes at 65°C for 1 h in Buffer 2.
10. Dilute SA-AP conjugate to 1 μ g/mL with Buffer 1 and incubate the membranes in this diluted SA-AP conjugate at room temperature for 10 min with gentle agitation.
11. Wash the membranes with 250 mL of Buffer 1 for 15 min at room temperature, and with 100 mL of Buffer 3 for 10 min at room temperature.
12. Prepare the dye solution for each membrane by mixing 33 μ L of NBT solution and 25 μ L of BCIP solution in 7.5 mL of Buffer 3. Incubate the membranes in dye solution for 30 min in dark.
13. Wash the membranes with stopping buffer.

4. Notes

1. The method described here is based on our experiments of cloning differentially expressed genes from immunocytes, such as monocytes, dendritic cells (**8,9**), bone marrow derived stromal cells, etc. **Figure 1** shows the schematic diagram of the method. Several full-length genes have been cloned by this method and confirmed to be expressed differentially (**10–12**).
2. Other equivalent purification systems can be used. The volume of GTC extraction buffer relies on cell types as well as cell numbers. In our experiments, because lymphocytes are of small size and contain less plasma, the 400 μ L of GTC extraction buffer is sufficient for 1×10^7 PBMCs.
3. It is noticeable that the efficiency of first strand cDNA synthesis is about 10–30% according to the reverse transcriptase to be used, which indicates that higher amounts of starting mRNA should be used if displaying of multiple genes or the genes expressed in low abundance is desired.
4. All the PCRs in this experiment were performed on a Perkin-Elmer DNA thermal cycler model 9600. AmpliTaq GOLD polymerase also worked well in amplifying tracer in our experiment.

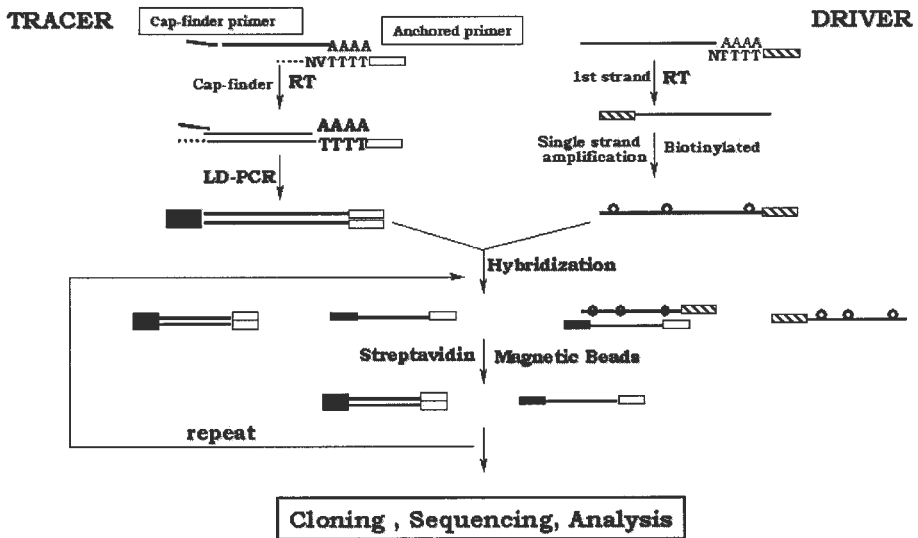


Fig. 1. The schematic diagram of the strategy for cloning of cell differentially expressed genes with full-length directly.

5. The extension time relies on not only the expected sequence size being cloned, but also the limitation of the techniques used in cloning and sequencing. The amplification cycles should be minimized in order to reduce PCR mutations.
6. In the preparation of driver, the accuracy of DNA polymerase is less important than the efficiency of DNA amplification and biotin labeling, and short fragments are favorable in hybridization, so the requirements for reverse transcriptase and polymerase are different from those for tracer preparation.
7. The template amount used in amplification and the PCR cycles should be adjusted according to the amount of products that obtained in subtractive hybridization, which should be monitored by electrophoresis in each step. **Figure 2** shows the subtracted cDNA amplification of PBMCs stimulated with PHA.
8. There are 60 positive clones among the 188 clones detected, so the proportion of specific clones is about 63.83%. Considering the frequency of Alu sequences (**13**), the efficiency of this subtractive cloning strategy is appropriate for cloning of differentially expressed genes.

References

1. Sagerstrom, C. G., Sun, B. I., and Sive, H. L. (1997) Subtractive cloning: past, present, and future. *Annu. Rev. Biochem.* **66**, 751–783.
2. Furuichi, Y. and Miura, K. (1975) A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. *Nature* **253**, 374,375.
3. Chenchik, A., Moqadam, F., and Siebert, P. (1996) A new method for full-length cDNA cloning by PCR, in *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*, (Krieg, P. A., ed.), Wiley-Liss, New York, pp. 273–321.

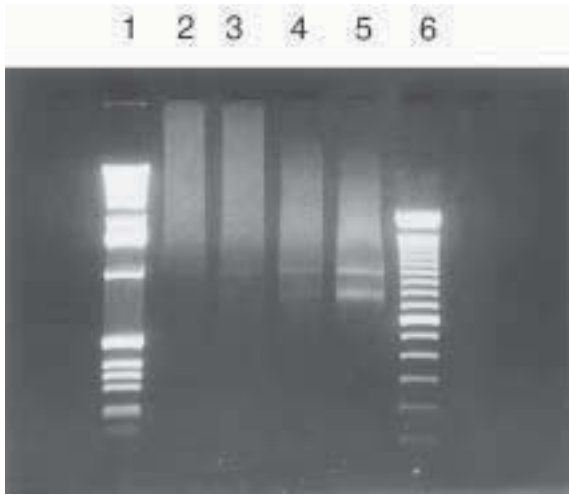


Fig. 2. The Subtracted cDNA Amplification of PBMCs Stimulated with PHA. 1.2% agarose gel shows amplified products of three rounds of subtractive hybridization. **Lane 1:** 1 kb DNA ladder (Gibco, cat. no. 15615); **Lane 2:** PCR products by using 0.02 μ L reverse-transcribed products as a template; **Lane 3, 4, 5:** the first, second and third subtractive PCR products by using 40 μ L subtracted cDNA as template; **Lane 6:** 100 bp DNA ladder (Gibco, cat. no. 15628).

4. Barnes, W. M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from (bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
5. Thompson, J., Solomon, R., Pellegrino, M., Sakai, K., Lewin, M., Field, M., et al. (1989) A noise-free molecular hybridization procedure for measuring RNA in cell lysates. *Analyt. Biochem.* **181**, 371–378.
6. Swaroop, A., Xu, J. Z., Agarwal, N., and Weissman, S. M. (1991) A simple and efficient cDNA library subtraction procedure: isolation of human retina-specific cDNA clones. *Nucl. Acids Res.* **19**, 1954.
7. Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) Reverse transcriptase (RNA-dependent DNA polymerase), in *Molecular Cloning: A Laboratory Manual* (2nd ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 5.52–5.55.
8. Zhao, Z., Huang, X., Li, N., Zhu, X., and Cao, X. (1999) Screening and cloning of new genes differentially expressed in dendritic cells in large scale. *Chin. J. Biochem. Mol. Biol.* **15**, 624–628.
9. Zhao, Z., Huang, X., Li, N., Zhu, X., Chen, S., and Cao, X. (1999) Direct cloning of cell differential expression genes with full-length by a new strategy based on the multiple rounds of “long distance” polymerase chain reaction and magnetic beads mediated subtraction. *J. Biotech.* **73**, 35–41.
10. Huang, X., Yuan, Z., Chen, G., Zhang, M., Zhang, W., Yu, Y., and Cao, X. (2001) Cloning and characterization of a novel ITIM containing Lectin-like immunoreceptor LLIR and its two transmembrane region deletion variants. *Biochem. Biophys. Res. Commun.* **281**, 131–140.
11. Huang, X., Zhao, Z., Yuan, Z., Zhang, M., Zhu, X., Chen, G., and Cao, X. (2000) Cloning and characterization of a novel deletion mutant of heterogeneous nuclear ribonucleoprotein M4 from human dendritic cells. *Science In China (Series C)* **43**, 648–654.

12. Li, N., Huang, X., Zhao, Z., Chen, G., Zhang, W., and Cao, X. (2000) Identification and characterization of a novel gene KE04 differentially expressed by activated human dendritic cells. *Biochem. Biophys. Res. Commun.* **279**, 487–493.
13. Claverie, J. M. and Makalowski, W. (1994) Alu alert. *Nature* **371**, 752.

II

CLONING PCR PRODUCTS

Cloning PCR Products

An Overview

Baotai Guo and Yuping Bi

1. Introduction

A successful cloning of polymerase chain reaction (PCR)-derived DNA fragment is a key step for further analysis of the amplified DNAs, though it is often a difficult task. Many cloning methods have been established; various commercial cloning kits are also available. These methods can be separated into two groups: ligation-dependent cloning and ligation-independent cloning. The former is used more widely than the latter. According to DNA ends, the existing ligation-dependent cloning methods for PCR products can be further divided into three types: blunt-end cloning, sticky-end cloning, and T-A cloning.

To choose a suitable ligation-dependent cloning method, several factors should be taken into account. First, consider the purpose of your cloning manipulation. For example, if the selected clones are to be expressed into proteins, then the directional cloning in an expression vector is desirable. Second, choose or prepare a suitable vector that matches with your unmodified or modified PCR products. Different DNA polymerases generate PCR products with different DNA ends, which are compatible to blunt-end or sticky-end cloning vectors. Third, for cloning extremely long PCR products (e.g. more than 10 kb), which are beyond the cloning capacity of the commonly used plasmid vectors, utilize cosmid or λ DNA vectors. To facilitate the selection of the clones with the inserts and improve the cloning efficiency, an efficient screening system for the positive clones is also very useful.

In this chapter, various cloning methods will be briefly reviewed; their advantages and disadvantages are pointed out. The strategies for positive selection are also discussed briefly.

2. Ligation-Dependent Cloning

2.1. Blunt-End Cloning

Blunt-end cloning is a conventional technique used in cloning PCR-amplified DNA fragments as well as any other blunt-end and double-stranded DNAs. In this method, PCR products with blunt ends and the blunt-cut (often dephosphorylated) vectors are ligated together. DNA fragments are inserted into the vectors in both orientations, therefore, directional cloning is not available.

Thermostable DNA polymerases, such as Vent, possess 3'-5' exonuclease activity, the so-called proofreading activity, which generates blunt-end PCR products (1). The resultant DNAs are relatively easy to be directly cloned through blunt-end ligation, because they are compatible with the blunt-end vectors. On the other hand, DNA polymerase like *Taq* exhibits nontemplate terminal transferase activity, which can add one deoxynucleotide to the 3'-hydroxyl termini of double-stranded PCR products and produce DNA fragments with single 3'-overhang nucleotide (2). The presence of 3'-overhangs prohibits the ligation of these fragments with blunt-end plasmid vectors. Therefore blunt-end cloning strategy is problematic with PCR products from DNA polymerases such as *Taq* and *Tth*. To overcome the cloning difficulty, PCR products may be treated with Klenow fragment of DNA polymerase I (3), T4 (4,5) or *Pfu* (4) polymerase to polish the ragged DNA ends, which can greatly increase the ligation and cloning efficiency.

Generally speaking, blunt-end cloning of PCR products is not as efficient as sticky-end cloning because of the inefficiency of blunt-end ligation and high tendency of self-ligation of vectors, which yields lower recombinant clones. Besides removal of 3'-overhangs, purification of PCR products, and addition of excessive T4 DNA ligase could also improve the blunt-end ligation efficiency. Adding restriction enzymes with a rare recognition sequence to ligation mixture can prevent the linearized vectors from recirculation during ligation reaction, resulting in lower self-ligation of vectors and higher recombinant clones (6). To facilitate the blunt-end cloning of PCR products, a blunt-end cloning procedure was established for cloning PCR products up to 10 kb (7). The cloning steps include an optimized polishing (if necessary) / kinase treatment of amplified DNAs, and ligation between the polished DNAs and the blunt dephosphorylated plasmid vectors (to inhibit self-ligation). PCR products from any DNA polymerase can be cloned using this method, and the cloning efficiency is very high. In addition, introduction of a positive selection system could further increase the cloning efficiency. Because only recombinants containing inserted DNA can survive the selection, the self-ligation of vectors is no longer a problem.

Blunt-end cloning takes advantages of ease in manipulation, low expenses, and lack of addition of a base to the amplified sequence owing to the 3'-overhangs produced by polymerase like *Taq*, which is especially useful in site-directed mutagenesis analysis (4). Because there is no need to incorporate restriction recognition sequences or other special sequences to the PCR primers, many researchers still think blunt-end cloning is an attractive cloning approach and prefer to use it.

2.2. Sticky-End Cloning

Sticky-end cloning is routinely used to clone PCR products in many laboratories. Restriction sites can be introduced into the 5' end of the PCR primers. As amplification reaction proceeds, the primers are integrated into the PCR products. After digestion by the corresponding restriction enzymes, sticky ends form at both ends of the PCR products. Digested PCR products and equivalently cut vectors can be ligated with very high efficiency, resulting in fairly efficient sticky-end cloning. Restriction sites introduced into the two PCR primers can be identical or different. If they are different, then the PCR products may be inserted into the vectors in only one orientation, enabling directional cloning of these products (8,9).

Some problems are frequently encountered with sticky-end cloning. Many restriction endonucleases fail to efficiently cleave the restriction recognition sites located at the 5' ends of double-stranded PCR products, especially when the sites are near or extremely near the 5' ends (10,11), which may be a result of the unstable binding of the restriction enzymes to the DNA termini. One solution is to add extra nucleotides to the 5' end outside the restriction site, but this increases the cost of primer synthesis and may also decrease the binding efficiency of the primers to the templates. For sticky-end cloning, another problem is the undesirable cleavage of internal restriction site(s) identical to those in primers. This problem is hard to predict, because the insert sequences are often unknown. The third problem is the generation of primer dimers. Restriction recognition sequences are commonly inverse repeat sequences (palindromes) that facilitate formation of primer dimers, which decrease amplification efficiency.

Another solution to overcome the restriction cleavage difficulty is by PCR product concatenation. The primary PCR products generated with 5'-phosphorylated primers can be first self-ligated by the T4 DNA ligase to form DNA concatamers, then the concatamerized DNAs can be readily cleaved, resulting in a higher efficiency sticky-end cloning (11). In another approach, only half of the palindromic restriction site is incorporated into one primer, and the other half into another primer. After concatamerization of the PCR-derived DNAs, the entire site is reconstituted, resulting in an easy sticky-end cloning as aforementioned (11). In both methods, the key is to convert the terminal sites into internal sites, nevertheless, directional cloning is not possible and both are intolerant to 3'-overhangs in PCR products. In addition, more steps are required after PCR, so the manipulation is laborious.

In 1999, an autosticky PCR (AS-PCR) cloning method was reported (9). This method uses two special primers that contain tetrahydrofuran derivative abasic sites, which makes DNA polymerase stall. The PCR products are therefore directly of different single-stranded DNA (5'-overhang) at both ends, corresponding to sticky ends of two different restriction sites. The intact PCR products can be directionally cloned into correspondingly cut plasmid vectors. While retaining the advantages of common sticky-end cloning, this method needs no post-PCR processing steps, avoiding the sensitivity of the internal sites to restriction enzymes and cleavage difficulty of sites at DNA ends.

Sticky-end cloning is really a personalized manipulation, and many flexible designs are employed to meet various cloning demands in different laboratories.

2.3. T-A Cloning

Taq is the most widely used thermostable DNA polymerase, and it has terminal transferase activity. In the presence of four dNTPs, deoxyadenosine (dA) is preferentially added to the 3' termini of PCR-amplified duplex DNAs, leaving a single 3' dA overhang (2). Without modification by restriction enzymes or any other enzymes, PCR products with 3' dA overhangs can be directly cloned into a linearized plasmid vector with complementary single 5' deoxythymidine (dT) overhangs at both ends. The dT overhang-containing vector is called T-vector, and the corresponding cloning is named T-A cloning. In fact, T-A cloning is also a sticky-end cloning, but a special one with only one base in the sticky end.

T-vectors can be prepared in laboratories through enzymatic reactions. By taking advantage of the terminal transferase activity of *Taq* polymerase, when only deoxythymidine triphosphate (dTTP) is present in the reaction, dT can be uniquely added to the 3' termini of a blunt-ended plasmid, resulting in a T-vector (12). A similar method is used to make T-vector in the presence of ddTTP, in which DNA terminal transferase introduces only one ddTTP to 3' termini of linear plasmids that are opened up with blunt-cutting restriction enzymes (13). The yielding T-vector does not possess 3'-hydroxyl group, so phosphodiester bond cannot form between 3' end of vector and 5' end of DNA fragment. The recombinant plasmid is of two nicks, but there is not much impact on transformation.

T-vector can also be produced directly by cutting plasmids with selected enzymes. Restriction enzymes, such as *XcmI*, *HphI*, and *AhdI*, generate single 3' dT overhang in the digested DNAs. *XcmI* recognition sequences are not present in the multiple cloning sites of commonly used plasmids, so it is convenient to add the *XcmI* sites to the desired plasmids in order to create the T-vector. The recognition sequence of *XcmI* is 5' CCANNNNNNNTGG 3', the internal bases (N) are alterable. Two adjacent *XcmI* sites with appropriate sequences are needed for the T-vector creation (14–18). T-vector is formed simply by *XcmI* digestion. An example of T-vector preparation by *XcmI* digestion is shown in **Fig. 1**.

T-A cloning is simple, reliable, and more efficient than blunt-end cloning. As long as DNA polymerases without 3'–5' exonuclease activity are used, T-A cloning is applicable and the PCR products may need no post-PCR processing steps. Therefore, T-A cloning is the most widely used cloning method for PCR products. It can be also used in cloning DNA fragments from other PCR-based techniques, such as RAPD (random amplified polymorphic DNA) and AFLP (amplified fragment length polymorphism). It has been used in establishing specific SCAR (sequence-characterized amplified region) markers (19,20) for molecular identification or recovery of DNA bands of interest in AFLP analyses. T-A cloning proves to be convenient and reliable in these applications. Although DNA polymerase with proofreading activity does not produce 3'-overhangs, their PCR products can still be cloned into T-vector after conducting a tailing procedure using *Taq* polymerase prior to ligation, which makes T-A cloning a universal method applicable to all DNA polymerases.

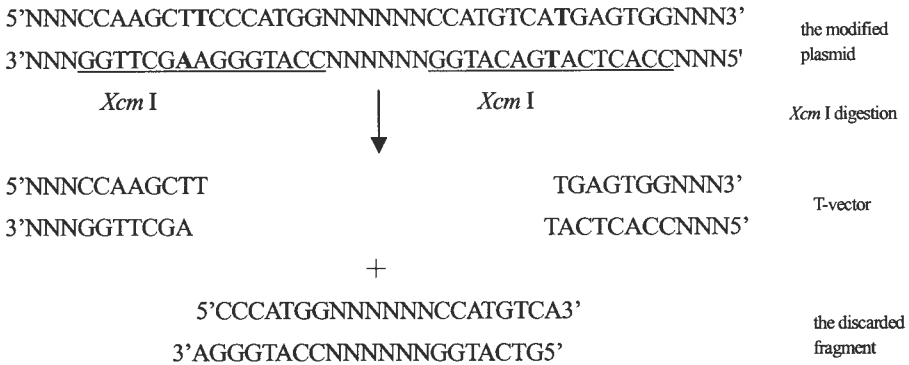


Fig. 1. Preparation of T-vector by *Xcm*I digestion

Although it is not required to do any post-PCR processing in T-A cloning, sometimes it is helpful to purify the PCR products before ligation to remove primer dimers, which compete favorably with the PCR products for the T-vector. In addition to the disadvantages of nondirectional cloning and the need for T-vector, T-A cloning also has the drawback of variable cloning efficiencies. This presumably is a consequence of variability in 3'-dA addition to PCR products mediated by *Taq* polymerase. It was found that not only dA but also dG, dC, or dT can be added to the 3' ends of PCR products (21), and the dA addition frequency is determined by the 3' end base and the adjacent bases in DNA fragments (22). It also appears that the dT overhangs in the T-vector and dA overhangs in the PCR products are unstable and may get lost during storage, especially at room temperature. The typical characteristics of blunt-end, sticky-end, and T-A cloning methods are summarized in **Table 1**.

3. Ligation-Independent Cloning

Conventional cloning means DNA recombination *in vitro*, that is, creation of phosphodiester bonds between DNA fragments and vectors by DNA ligase *in vitro*. Ligation-independent cloning (LIC), on the other hand, does not involve *in vitro* ligation, so it is ligase-free cloning. Various LIC methods exist, and different cloning strategies are adopted. PCR products and linearized vector containing identical terminal sequences to them can be cotransferred into *Escherichia coli* strain JC8679, then DNA recombination occurs *in vivo* by the help of the host homologous recombination system (23). This *in vivo* cloning (IVC) method is simple, but the host strains are very restricted. Like conventional sticky-end cloning, some LIC methods require the creation of sticky ends in both the inserts and the vectors. But the differences are that the sticky ends in the PCR amplification products are not created by the restriction enzymes, but by various other DNA enzymes, and the extra sequences at the 5' end of the primers are usually longer than common restriction recognition sequences. This type of cloning is called enzymatic modification-mediated LIC. In another LIC called PCR-induced LIC, the PCR products and the plasmids are fused together in a second PCR reaction, independent of any enzymatic modifications. In LIC the resulting

Table 1
Comparison of Three Ligation-Dependent Cloning Methods for PCR Products

Method	Vector	Primers	PCR Products	3'-overhang	Ligation	Advantages
Blunt-End	Common	Common	Intact	Intolerant	Less efficient	Simple and Universal
Sticky-End	Common	Additional Sequences	Additional Sequences	Tolerant	Efficient	Directional Cloning
T-A	T-vector	Common	Additional A	Tolerant	Efficient	Simple and High Efficiency

recombinant DNAs are open circular DNA molecules with nicks, but they can be directly applied to transform bacterial competent cells with high efficiency.

3.1. Enzymatic Modification-Mediated LIC

In enzymatic modification-mediated cloning, special sequences are introduced into 5' end of PCR primers. The PCR products are then treated with various DNA enzymes, such as T4 DNA polymerase (24), exonuclease III (25), or uracil DNA glycosylase (UDG) (26–28) to generate single-stranded tails (sticky ends), which will hybridize with the complementary sticky ends in vectors, forming a relatively stable recombinant plasmid. When using T4 DNA polymerase, additional 12 nucleotide sequences lacking dCMP are incorporated into the 5' end of primers, the resultant PCR products, lacking dGMP at their 3' ends, are treated by T4 DNA polymerase to form 5' overhangs in the controlled reaction conditions because of its 3'-5' exonuclease activity (24). Exonuclease III is also efficient to create 5'-overhangs in PCR products (25). With UDG-mediated LIC (27), 12-base dUMP-containing sequences are added to the 5' end of the PCR primers, resulting in the selective replacement of dUMP residues into 5' end of PCR products. After UDG treatment, which selectively degrades dU in the PCR products, 3'-overhangs are generated. Unlike many restriction enzymes, UDG functions effectively near DNA termini, thus enabling high efficiency degradation and subsequent cloning. Directional cloning can be accomplished by adding different extra sequences to the two PCR primers (25,27). These methods are highly efficient, but require long additional sequences in primers and post-PCR enzymatic treatments.

3.2. PCR-Induced LIC

By the skillful use of PCR techniques, it is possible to join any DNA fragments and vectors together. In PCR-induced LIC, cloning of PCR products is mediated by further PCR, in contrast to the conventional ligation mediated by DNA ligase. One strategy is that inserts and vectors are amplified separately and the two amplification products are then mixed, denatured, and annealed to form a recombinant DNA (29). Another strategy includes four pairs of primers, but inserts are first amplified with primers of additional sequences at the 5' ends that result in PCR products whose 3' ends are complementary to the 3' ends of the recipient linearized plasmid, then mixed with this linearized plasmid, followed by second PCR to produce nicked cyclic recombinant

DNAs (30,31). These methods are efficient and fast, but more than one primer pair is needed and two PCR amplifications are conducted. They are very useful in studies of site-directed mutagenesis and *in vitro* recombination (29), but researchers should be aware of the complexity of the design of additional sequences at the 5' end of the primers and the matching of these sequences with the corresponding vector sequences.

Still based on PCR-induced strategy, Chen et al. developed a new cloning method that enables inserting a PCR product into a vector flexibly and precisely at any desired location with high efficiency (32). For details, see the corresponding chapter in this section.

In LIC methods, PCR primers contain extra sequences for combination of PCR products with vectors or sticky-end creation, and the primers are usually longer than those used in conventional sticky-end cloning. Because of the need for special sequences in primers, LIC methods are hardly useful in cloning fragments from RAPD and AFLP.

4. Positive Selection System

As in general cloning, the cloning efficiency of PCR products is also dependent on three factors: ligation efficiency, transformation efficiency of competent cells, and selection efficiency of desired transformants. In most cases, transformation efficiency is not a limiting factor, but ligation and selection of positive transformants are. Improving the selection efficiency will improve the entire cloning efficiency. Based on the α -complementation of β -galactosidase, the blue/white screening system is widely used in DNA cloning, including PCR products cloning. This method allows visual color discrimination of recombinants from nonrecombinants in the presence of X-gal and IPTG in bacterial culture medium. One disadvantage is that it can lead to a very high background of nonrecombinant clones because all self-ligated vectors can be transformed into host and produce blue colonies. Another disadvantage is that in-frame cloning also leads to false-negative blue colonies.

To decrease the background and facilitate the selection of recombinants, many positive screening vectors or systems have been established (33–36). They use the strategy of insertional inactivity of lethal gene such as *ccd B*, which encodes potent poison protein to gyrase (33), and *barn*, which encodes barnase (35), or toxic sensitivity gene (36). Positive system creates powerful selection, because the bacteria with self-ligated vectors carrying lethal or toxin sensitive gene fail to form colonies on the plate. Because of the limitations in host range, requirements for special components in medium or limited cloning sites, positive selection systems are not used widely in PCR products cloning.

In a new positive-selection method applicable to clone PCR products up to 9 kb (37), the plasmid vector harbors the lethal mutant gene *crp^s* encoding an altered catabolite gene activator protein CAP^s. CAP is also known as cyclic AMP receptor protein CRP. CAP^s-cAMP complex is toxic to host cells, so wild-type cells with *crp^s* are killed. Only recombinant plasmids containing inserts in the unique restriction site within *crp^s* gene to make it disrupted can survive, leading to a very strict selection for recombinant transformants. The advantages of this method are of no need for special components in culture medium and suitable for a broad range of host bacterial strains, but the

vector must be maintained in *E. coli cyo⁻* strains, which are unable to produce adenylate cyclase. A commercial kit from Roche Biochemicals is also based on mutant *crp* gene, which allows efficient blunt-end cloning of PCR products.

Another positive screening system was specially developed to clone PCR products (38). In this method, the vector with a translation deficiency in *lacZ α* gene is created by deletion of the Shine-Dalgarno sequence and initiation codon. Instead, the Shine-Dalgarno sequence and initiation codon are incorporated into one of the PCR primers to allow complementation of inactive *lacZ α* gene by the PCR products, which result in blue transformed bacterial colonies.

References

1. Lohff, C. J. and Cease, K. B. (1992) PCR using a thermostable polymerase with 3' to 5' exonuclease activity generates blunt products suitable for direct cloning. *Nucl. Acids Res.* **20**, 144.
2. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677–9686.
3. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. (1989) A simple method for site directed mutagenesis using the polymerase chain reaction. *Nucl. Acids Res.* **17**, 6545–6551.
4. Costa, G. L. and Weiner, M. P. (1994) Polishing with T4 or Pfu polymerase increases the efficiency of cloning PCR fragments. *Nucl. Acids Res.* **22**, 2423.
5. Skryabin, B. and Vassilacopoulou, D. (1993) A simple and fast method for cloning and analyzing polymerase chain reaction products. *Genet. Analyt. Tech. Appl.* **10**, 113–115.
6. Costa, G. L., Graftsky, A., and Weiner, M. P. (1994) Cloning and analysis of PCR-generated DNA fragments. *PCR Meth. Appl.* **3**, 338–345.
7. Novy, R. E., Yaeger, K. Y., and Kolb, K. (1996) Perfectly blunt cloning: a superior method for cloning PCR products or any DNA. *Innovations* **6**, 711.
8. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1078.
9. Gal, J., Schnell, R., Szekeres, S., and Kalman, M. (1999) Directional cloning of native PCR products with preformed sticky ends (autosticky PCR). *Mol. Gen. Genet.* **260**, 569–573.
10. Kaufman, D. L. and Evans, G. A. (1990) Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* **9**, 304–306.
11. Jung, V., Pestka, S. B., and Pestka, S. (1990) Efficient cloning of PCR generated DNA containing terminal restriction endonuclease recognition sites. *Nucl. Acids Res.* **18**, 6156.
12. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**, 1154.
13. Holton, T. A. and Graham, M. W. (1991) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucl. Acids Res.* **19**, 1156.
14. Mead, D. A., Pey, N. K., Herrnstadt, C., Marcil, R. A., and Smith, L. M. (1991) A universal method for the direct cloning of PCR amplified nucleic acid. *Biotechnology* **9**, 657–663.
15. Cha, J., Bishai, W., and Chandrasegaran, S. (1993) New vectors for direct cloning of PCR products. *Gene* **136**, 369,370.
16. Harrison, J., Molloy, P. L., and Clark, S. J. (1994) Direct cloning of polymerase chain reaction products in an *Xcm* I T- vector. *Analyt. Biochem.* **216**, 235–236.
17. Kwak, J. H. and Kim, M. Y. (1995) Construction of T-vector for direct cloning and expression of cloned genes in *Escherichia coli*. *Analyt. Biochem.* **228**, 178–180.
18. Borovkov, A. Y. and Rivkin, M. I. (1997) *Xcm* I-containing vector for direct cloning of PCR products. *BioTechniques* **22**, 812,813.

19. Hermosa, M. R., Grondona, I., Diaz-Mingues, J. M., Iturriaga, E. A., and Monte, E. (2001) Development of a strain-specific SCAR marker for the detection of *Trichoderma atroviride* 11, a biological control agent against soilborne fungal plant pathogens. *Curr. Genet.* **38**, 343–350.
20. Xu, M., Huaracha, E., and Korban, S. S. (2001) Development of sequence-characterized amplified regions (SCARs) from amplified fragment length polymorphism (AFLP) markers tightly linked to the Vf gene in apple. *Genome* **44**, 63–70.
21. Hu, G. (1993) DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol.* **12**, 763–770.
22. Brownstein, J. M., Carpten, J. D., and Smith, J. R. (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques* **20**, 1004–1006.
23. Oliner, J. D., Kinzler, K. W., and Vogelstein, B. (1993) In vivo cloning of PCR products in *E. coli*. *Nucl. Acids Res.* **21**, 5192–5197.
24. Aslanidis, C. and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucl. Acids Res.* **18**, 6069–6074.
25. Hsiao, K.C. (1993) Exonuclease III induced ligase-free directional subcloning of PCR products. *Nucl. Acids Res.* **21**, 5528–5529.
26. Nisson, P. C., Rashtchian, A., and Watkian, P. C. (1991) Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase. *PCR Meth. Appl.* **1**, 120–123.
27. Rashtchian, A., Buchman, G. W., Schuster, D. M., and Berninger, M. S. (1992) Uracil DNA glycosylase-mediated cloning of polymerase chain reaction-amplified DNA: application to genomic and cDNA cloning. *Analyt. Biochem.* **206**, 91–97.
28. Smith, C., Day, P. J., and Walker, M. R. (1993) Generation of cohesive ends of PCR products by UDG-mediated excision of dU, and application for cloning into restriction digest-linearized vectors. *PCR Meth. Appl.* **2**, 328–332.
29. Jones, D. H., Sakamoto, K., Vorce, R. L., and Howard, B. H. (1990) DNA mutagenesis and recombination. *Nature* **344**, 793,794.
30. Shuldiner, A. R., Scott, L. A., and Roth, J. (1990) PCR-induced (ligase-free) subcloning: a rapid reliable method to subclone polymerase chain reaction (PCR) products. *Nucl. Acids Res.* **18**, 1920.
31. Shuldiner, A. R., Tanner, K., Scott, L. A., Moore, C. A., and Roth, J. (1991) Ligase-free subcloning: a versatile method to subclone polymerase chain reaction (PCR) products in a single day. *Analyt. Biochem.* **194**, 9–15.
32. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. (2000) Restriction site-free insertion of PCR products directionally into vectors. *BioTechniques* **28**, 498–500, 504,505.
33. Bernhard, P., Gabant, P., Bahassi, E. M., and Couturier, M. (1994) Positive-selection vectors using the F plasmid *ccd^B* killer gene. *Gene* **148**, 71–74.
34. Henrich, B. and Schmidtberger, B. (1995) Positive-selection vector with enhanced lytic potential based on a variant of (*X174* phage gene. *E. Gene* **154**, 51–54.
35. Yazynin, S. A., Deyev, S. M., Jucovic, M., and Hartley, R. W. (1996) A plasmid vector with positive selection and directional cloning based on a conditionally lethal gene. *Gene* **169**, 131,132.
36. Kast, P. (1994) pKSS-A second-generation general purpose cloning vector for efficient positive selection of recombinant clones. *Gene* **138**, 109–114.
37. Schlieper, D., von Wilcken-Bergmann, B., Schmidt, M., Sobek, H., and Muller-Hill, B. (1998) A positive selection vector for cloning of long polymerase chain reaction fragments based on a lethal mutant of the *crp* gene of *Escherichia coli*. *Analyt. Biochem.* **257**, 203–209.
38. Keese, P. and Graf, L. (1996) A positive screen for cloning PCR products. *Nucl. Acids Res.* **24**, 3474,3475.

Using T4 DNA Polymerase to Generate Clonable PCR Products

Kai Wang

1. Introduction

Polymerase chain reaction (PCR) mediated through *Taq* DNA polymerase has become a simple and routine method for cloning, sequencing, and analyzing genetic information from very small amount of materials (1). *Taq* DNA polymerase, like some other DNA polymerases, lacks 3' to 5' exonuclease activity and will add nontemplate-directed nucleotides to the ends of double-stranded DNA fragments (2). Because of the strong preference of the *Taq* polymerase for deoxyadenosine s'triphosphate (dATP), the nucleotide added is almost exclusively an adenosine. This results in generating "ragged" unclonable amplification products (3). Restriction endonuclease sites are often incorporated into the amplification primers so that clonable PCR products can be generated by restriction enzyme cleavage (4). However, the possible secondary sites located within the amplified products often complicate the cloning and interpretation of PCR results. A cloning system exploiting the template-independent terminal transferase activity of *Taq* polymerase has been developed (5–7). However, a special vector with thymidine (T) overhanging ends has to be used in the process.

T4 DNA polymerase has strong exonuclease and polymerase activities in a broad range of reaction conditions (8). By adapting its strong enzymatic activities, a simple and efficient method to generate clonable PCR fragments with T4 DNA polymerase has been developed. The T4 DNA polymerase not only repairs the ends of the PCR products, but also removes the remaining primers in the reaction with its strong single-stranded exonuclease activity. Therefore, this method does not require multiple sample handling, buffer changes, or gel purification steps. Instead, a simple alcohol precipitation step is used to purify the PCR products.

2. Materials

2.1. PCR

1. DNA template containing the sequence interested.
2. Oligonucleotide primers.
3. *Taq* DNA polymerase.
4. 10X PCR and enzymatic repair buffer: 500 mM Tris-HCl, pH 9.0, 25 mM MgCl₂, 500 mM NaCl, and 5 mM dithiothreitol (DTT). Commercial 10X PCR buffer also works well.
5. 1.5 mM 10X deoxynucleotides (dNTP) solution. Concentrated stock solution (100 mM) can be obtained from Pharmacia (Piscataway, NJ) or Boehringer Mannheim (Indianapolis, IN).
6. Gel electrophoresis and PCR equipment.

2.2. End Repair and Cloning

1. Enzymes: T4 DNA polymerase, T4 polynucleotide kinase, and T4 DNA ligase. Enzymes can be purchased from Boehringer Mannheim, Life Technologies (Gaithersburg, MD) or any other provider.
2. 1 mM of ATP solution. Concentrated stock solution (100 mM) can be obtained from Boehringer Mannheim.
3. Isopropyl alcohol.
4. Vector (blunt end and dephosphorylated).
5. 10X Ligase buffer: 660 mM Tris-HCl, pH 7.6, 66 mM MgCl₂, 10 mM adenosine (ATP), 1 mM Spermidine, and 10 mM DTT. Commercially available 10X ligase buffer also works well.
6. TE buffer: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA.
7. 5 M NaCl.

3. Methods

3.1. Primer Design

No special primer is needed, however, secondary structure and stretch of homopolymer should be avoided.

3.2. PCR Reaction

1. Prepare the following in a PCR reaction tube:
 - a. 5 μ L of 10X PCR buffer (500 mM Tris-HCl, pH 9.0, 25 mM MgCl₂, 500 mM NaCl, and 5 mM DTT).
 - b. 0.25 μ g of DNA template such as genomic DNA.
 - c. 1 μ M of each primer.
 - d. 0.15 mM of each deoxynucleotides (dNTP).
 - e. 1 U *Taq* polymerase (Perkin-Elmer, Cetus, Norwalk, CT).
 - f. add deionized H₂O to a final volume of 50 μ L.
2. Amplification conditions largely depend on the specific applications. However, a general cycling profile listed below can be used in most of the amplifications.
 - a. Initial denaturation 94°C for 5 min.
 - b. Amplification (30 cycles) 94°C for 30 s; 55°C for 45 s; and 72°C for 90 s.
 - c. Final extension 72°C for 10 min.
3. Examine the PCR amplification results with agarose gel electrophoresis (*see Note 1*).

3.3. End Repair for Cloning

1. Add the following to PCR reaction tubes directly to repair the PCR products (*see Note 2*):
 - a. 1 U of T4 DNA polymerase;
 - b. 1 μL of 4 mM dNTP solution (optional) (*see Note 3*);
 - c. 5 U of T4 polynucleotide kinase (*see Note 4*);
 - d. 1 μL of 1 mM ATP.
2. Incubate the reaction tubes at 25°C (room temperature) for 20 min (*see Notes 5 and 6*). Stop the reactions by adding 3 μL of 0.5 M EDTA (pH 8.0).
3. Incubate the reaction tubes at 70°C for 10 min to inactivate the enzymes.
4. Precipitate the PCR products at room temperature by adding 5 μL of 5 M NaCl and 60 μL of isopropyl alcohol (3) (*see Note 7*).
5. Resuspend the DNA fragments in 20 μL of TE or water.
6. Take 2 μL (containing approx 20–50 ng of PCR product) and mix with ligase and vector for ligation (*see Notes 4 and 8*): 1 μL of 10X ligase buffer; 1 μL of T4 DNA ligase; 2 μL of repaired PCR product; dephosphorylated vector DNA (60–150 ng); and add deionized H₂O to a final volume of 10 μL .
7. Incubate at 16°C overnight.
8. Dilute the ligation reaction fivefold in TE buffer. Use 2 μL of the diluted ligation reaction for transformation.

4. Notes

1. In case of multiple PCR products from a single reaction was observed, the specific products should be purified by gel electrophoresis based on its estimated size after repair reaction. Several different methods can be used to purify DNA fragments from agarose gel, such as phenol extraction from low-melting gel, “glassmilk” method, or by simple low-speed centrifugation.
2. This protocol utilizes a single buffer for all the enzymes that include *Taq* polymerase in PCR, T4 polymerase, and T4 polynucleotide kinase in end repairing. Therefore, slightly higher concentration of reagents and enzymes can be added in the reaction.
3. T4 DNA polymerase can be added directly into PCR tube without providing additional nucleotides. However, T4 DNA polymerase balances its exonuclease and polymerase activities based on the concentration of available deoxynucleotides. Depending on the length of amplification products, number of cycles, and nucleotide sequence composition of amplified region, the remaining nucleotide concentration after PCR amplification may be different from experiment to experiment. In order to avoid unnecessary confusion, supplemental nucleotides are routinely added for end-repair reaction.
4. T4 polynucleotide kinase is not needed when vector used has not been treated with phosphatase previously. However, dephosphorylated vector should be used to increase the cloning efficiency.
5. Room temperature (25°C) was chosen for the reaction, because T4 DNA polymerase has excessive exonuclease activity at 37°C.
6. The T4 polynucleotide kinase works well at room temperature.
7. Although the PCR products purified directly by alcohol precipitation after end repairing are sufficient for routine cloning, passing the repaired PCR product mixtures through a gel filtration column prior to the alcohol precipitation can greatly enhance the cloning efficiency.
8. In the ligation reaction, we routinely used 1:1 molar ratio between vector (dephosphorylated) and insert.

References

1. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677–9686.
2. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **230**, 1350–1354.
3. Scharf, S. (1990) *PCR Protocols: A Guide to Methods and Applications*. Academic, San Diego, CA.
4. Jung, V., Pestka, S. B., and Pestka, S. (1990) Efficient cloning of PCR generated DNA containing terminal restriction endonuclease recognition sites. *Nucl. Acids Res.* **18**, 6156.
5. Mead, D. A., Pey, N. K., Herrnstadt, C., Marcil, R. A., and Smith, L. M. (1991) A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/technology* **9**, 657.
6. Kovalic, D., Kwak, J., and Weisblum, B. (1991) General method for direct cloning of DNA fragments generated by the polymerase chain reaction. *Nucl. Acids Res.* **19**, 4560.
7. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**, 1154.
8. Sambrook, J., Fritsch, E. F., and Maniatis, T., eds. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Enzyme-Free Cloning of PCR Products and Fusion Protein Expression

Brett A. Neilan and Daniel Tillett

1. Introduction

Numerous techniques have been developed for the cloning of polymerase chain reaction (PCR) products. These include the incorporation of restriction enzyme sites into the PCR primers (*1*), blunt-end cloning (*2,3*), TA cloning (*4,5*), ligation-independent cloning (LIC) (*6–11*), and in vivo cloning (*12,13*). Because these methods are effective, they all require either extensive enzymatic treatment of the PCR product or vector (*1,4–11,14*), the use of PCR primers containing nonstandard bases (*9,10*), or specialized vectors or bacterial strains (*2,13,15*). In addition, cloning into specialized expression or reporter vectors requires multiple subcloning steps. Under most circumstances, these limitations pose little difficulty, however, under certain conditions the current techniques can be severely limiting. For example, the direct cloning of a PCR product into a particular site of an unusual expression vector is problematic. This can prove especially complex if the vector lacks suitable cloning sites, or a genetic basis for insert screening.

Recently, a novel technique to clone PCR products, termed “heterostagger PCR cloning,” was introduced (*16*). This technique involves the generation of two related PCR products by the design of two sets of primers. The two sets are identical except the second primer set contains a three-base guanosine five prime tail (e.g., first set 5'-TAT..., second set 5'-GGGTAT...). Two PCR reactions are performed using the first untailed-forward primer and the tailed-reverse primer, and then the tailed-forward primer with the untailed-reverse primer. This produces two PCR products that, when mixed, denatured, and allowed to reanneal, create fragments containing 3' CCC overhangs that are cloned by ligation to a vector containing 3' GGG overhangs. Although elegant, this technique is limited by the need for extensive vector preparation, and the requirement for a unique restriction enzyme site at the desired cloning location within the chosen vector.

We reasoned that if the required 3 bp tail was increased to 12–18 bp, then the resulting “heterostaggered” PCR products could be cloned using the LIC procedure (*6*). In addition, by PCR amplifying the vector using a compatible set of tailed and nontailed primers, the PCR products could be cloned without further enzymatic reaction (*see Fig. 1*).

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

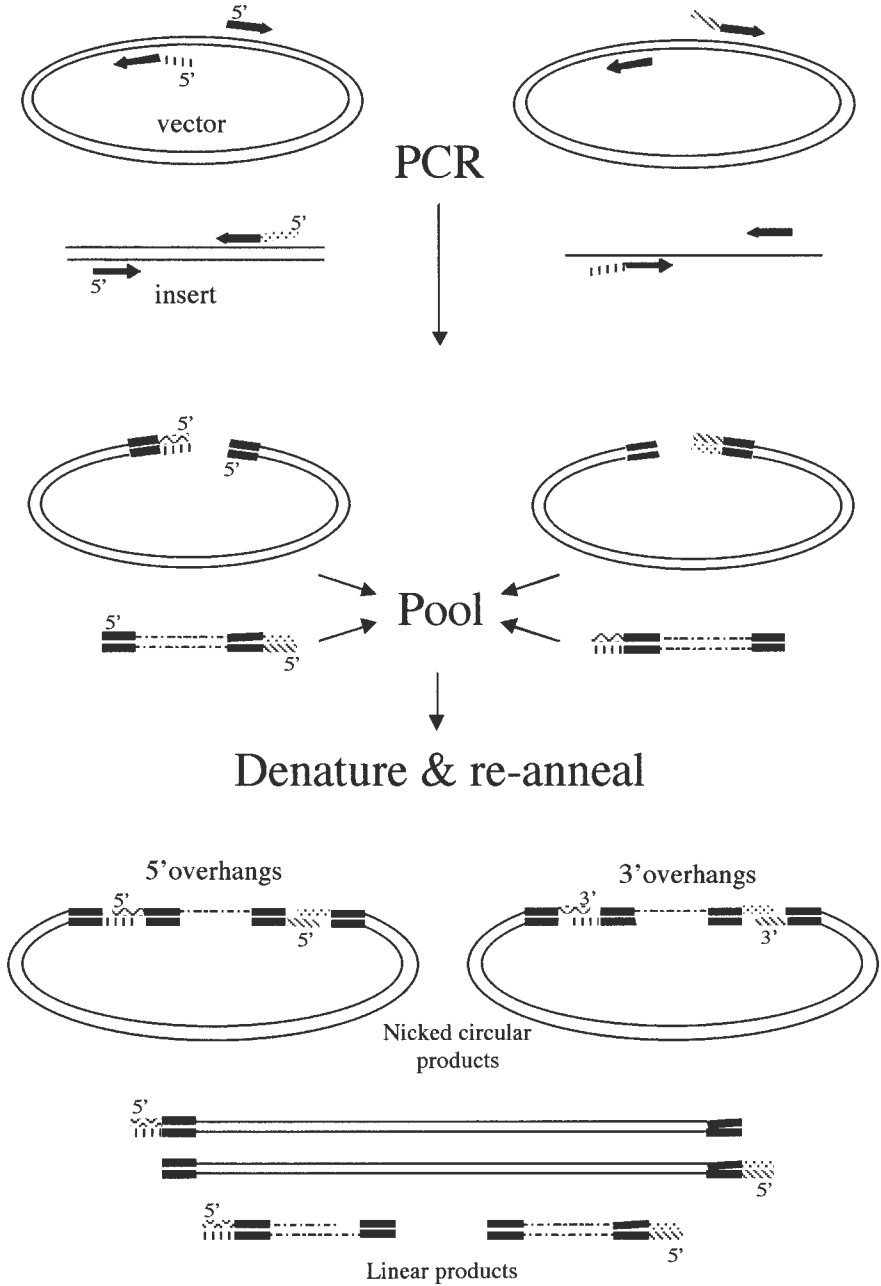


Fig. 1. The enzyme-free cloning strategy. Four PCR reactions are performed, two vector and two insert: the first vector PCR is performed with the 5' vector-short and the 3' vector-long (tailed) primers, whereas the second vector PCR uses the 5' vector-long and the 3' vector-short primers; the first insert PCR uses the 5' insert-short and 3' insert-long primers, whereas the second is performed using the 5' insert-long and the 3' insert-short primers. The long primers

The enzyme-free approach has a number of advantages over other cloning methods. First, no post-PCR enzyme reactions are required, such as restriction enzyme digestions, phosphatase treatments, partial DNase digestions, or ligations, and thus the potential for cloning failure is reduced. Second, the approach is simple and requires minimal sample handling. Third, the enzyme-free procedure enables the PCR insert to be cloned directly into a desired plasmid locus without the need for a suitable restriction enzyme site. This feature is particularly useful for ensuring a DNA fragment is directionally cloned into the correct reading frame for protein expression or for the creation of fusion proteins. In addition, the technique can be used to easily introduce a short amino acid tag sequence into either the amino or carboxy terminus of a protein.

2. Materials

2.1. DNA Amplification Reagents

It is advisable to store small aliquots of these solutions at 4°C for short periods (up to 3 wk). Stock solutions should be kept at -20°C for longer periods of storage. Preferably, stock solutions are kept as small aliquots to reduce deterioration of quality resulting from repeated freeze-thaw cycles. All solutions are made using deionized and autoclaved water. The pH values are for solutions at 25°C. Many suppliers of thermostable polymerases also supply the other PCR reagents.

1. 10X polymerase buffer: 200 mM Tris-HCl (pH 8.2), 100 mM KCl, 100 mM (NH₄)₂SO₄, 1.0% Triton X-100, and 1 mg/mL nuclease-free bovine serum albumin (BSA).
2. 2 mM deoxynucleotide 5'-triphosphate (dNTP) mix: combine equal volumes of 8 mM deoxyadenosine 5'-triphosphate (dATP), 8 mM deoxycytidine 5'-triphosphate (dCTP), 8 mM deoxyguanosine 5'-triphosphate (dGTP), and 8 mM deoxythymidine 5'-triphosphate (dTTP) (Boehringer).
3. 25 mM MgCl₂: solution is filter (0.22 μm) sterilized.
4. *Taq/Pfu* DNA polymerase mix: Two enzymes combined in 10:1 ratio. *Taq* supplied by Fischer Biotech (Perth, Australia) and *Pfu* from Promega (Madison, WI).
5. Oligonucleotide primers: available commercially and working solutions adjusted to 10 pmol/μL in ddH₂O (see Note 3).

2.2 Amplicon Purification, Hybridization, and Transformation Reagents

1. 3.0 M sodium acetate: solution is filter (0.22 μm) sterilized and the pH is not adjusted.
2. 80% ethanol: 80% (v/v) analytical-grade ethanol in ddH₂O.
3. Hybridization buffer: 100 mM NaCl, 10 mM Tris-HCl (pH 7.4), 1 mM ethylenediamine tetraacetic acid (EDTA) (pH 8.0).
4. Luria broth: 10 g Tryptone, 10 g NaCl, and 5 g Yeast Extract per liter of ddH₂O.

Fig. 1. (continued) differ from the short primers by possessing a 12 to 15 bp 5' tail, with each of the two nonhomologous insert primer tails complementary to only one of the two vector primer tails, thus ensuring the directional cloning of the desired fragment. Equimolar volumes of the four PCR products are mixed, heat-denatured, and allowed to reanneal. Eight annealing products of equal probability are created, of which four contain either 5' or 3' 12–18 bp overhangs. Low-temperature annealing (25°C) allows these four products to form two nicked circular products which can be transformed without ligation (see Note 8). Figure reproduced and modified from (19).

Table 1
Oligonucleotide Primers Described in this Chapter

Primer	Sequence (5' to 3')
NMLF	CATCACCATCACCATCAC-CAAGAAGGGGAACTCTTG
NMSF	CAAGAAGGGGAACTCTTG
NMLR	<u>TGTA</u> AAAACGACGGCCAGT-TTATTGGGGAATCAGTTTAC
NMSR	TTATTGGGGAATCAGTTTAC
PHL	<u>GTGATGGTGATGGTGATG</u> -CATAGCTGTTTCCTGTGTG
PHS	CATAGCTGTTTCCTGTGTG
PEL	<u>ACTGGCCGTCGTTTTACA</u> -ACGTCGTGACTGGGAAAAC
PES	ACGTCGTGACTGGGAAAAC

Dashes indicate the point of extension between short and long. Underlined and italicized sequences are the reverse complements for amplicon hybridization at the 3' and 5' ends, respectively, for the N-methyltransferase insert to be cloned (*see Note 6*). This table describes vector primers which encode a polyhistidine tag for fusion protein construction (PHL and NMSF). Alternatively, if such a tag is not required, an arbitrary sequence of 12–18 bp may be used to extend the short template specific primers sequences (as seen in the downstream primers, NMLR and PEL, used in this study).

3. Methods

3.1. Insert PCRs

Two insert reaction mixes are prepared and amplified. Each PCR mix contains a different insert short primer and insert long primer (*see Fig. 1* and **Table 1**).

1. Prepare genomic DNA as per standard protocols for the organism under investigation (*see Note 1*).
2. In two thin-walled 200- μ L microfuge tubes combine: 2 μ L of 10X polymerase buffer, 2 μ L of 2 mM dNTP mix, 2 μ L of 25 mM, 1 μ L of genomic DNA (approx 10–100 ng), 0.2 μ L of high fidelity *Taq/Pfu* polymerase mix (*see Note 2*), and 10.8 μ L of ddH₂O. Keep these insert reaction mixes on ice.
3. To one tube containing the insert reaction mix, add 1 μ L of insert short-forward primer and 1 μ L of insert long-reverse primer. To the other insert reaction mix, add 1 μ L of insert long-forward primer and 1 μ L of insert short-reverse primer.
4. Cover each reaction with mineral oil unless a thermocycler with a heated lid is available. Subject each tube to an initial denaturation of 2 min at 94°C in a thermal cycler. Incubate tubes with insert PCRs for 26 cycles of 94°C, 10 s; primer annealing temperature for 20 s; and 72°C for 90 s; followed by a final extension at 72°C for 5 min. PCR annealing temperature and extension times may vary depending on the primer design and the size and structure of the target gene (*see Notes 2* and **4**).
5. Determine that a single PCR amplicon of expected size has been generated for each insert PCR by agarose/TAE gel electrophoresis (**17**).

3.2. Vector PCRs

Two vector PCR mixes are prepared and amplified. Each PCR mix contains a different vector short primer and vector long primer (*see Fig. 1* and **Table 1**).

1. Prepare plasmid DNA as per standard protocols (17) (see **Note 5**).
2. In two thin-walled 200- μ L microfuge tubes combine: 2 μ L of 10X polymerase buffer, 2 μ L of 2 mM dNTP mix, 2 μ L of 25 mM, 1 μ L of plasmid DNA (approx 1–10 pg), 0.2 μ L of high-fidelity *Taq/Pfu* polymerase mix, and 10.8 μ L of ddH₂O. Keep these vector reaction mixes on ice.
3. To one tube containing the vector reaction mix, add 1 μ L of vector short-forward primer and 1 μ L of vector long-reverse primer. To the other vector reaction mix, add 1 μ L of vector long-forward primer and 1 μ L of vector short-reverse primer. For experiments requiring the heterologous expression of a fusion protein the vector primers are designed to incorporate the sequence encoding a 6 histidine residue tag (see **Note 6** and **Table 1**).
4. Subject each tube to an initial denaturation of 2 min at 94°C in a thermal cycler. Incubate tubes with vector PCRs for 26 cycles of 94°C, 10 s; primer annealing temperature for 20 s; and 72°C for 90 s; followed by a final extension at 72°C for 5 min. PCR annealing temperature and extension times may vary depending on the primer design and the size and structure of the cloning vector (see **Notes 2** and **4**).
5. Determine that a single PCR amplicon of expected size has been generated for each vector PCR by agarose/TAE gel electrophoresis (17).

3.3. Purification of Insert and Vector Amplicons

Although not essential, it is advisable for increased cloning efficiency to purify PCR products from other PCR components, including polymerases and oligonucleotides.

1. In one 1.5-mL centrifuge tube combine the two insert PCR products. In another tube, combine both vector PCRs. To each tube, add one-tenth volume of 3 M sodium acetate and 2 volumes of 80% ethanol. Alternatively, commercially available kits for the purification of PCR products may be employed, following the manufacturers instructions.
2. It is necessary at this stage to determine the concentration of each purified PCR product mix by either ultraviolet (UV) spectrometry (260 nm) or by ethidium bromide staining against known DNA standards.

3.4. Hybridization of Insert and Vector Amplicons

1. In a single 200- μ L microfuge tube combine purified insert and vector amplicons at an equimolar ratio. Approximately 50 ng of a 3 kb vector amplicon is required and the amount of insert added accordingly (e.g., 5 ng of a 300 bp insert PCR fragment).
2. This volume is lyophilized and resuspended in 50 μ L of hybridization buffer prior to denaturation and hybridization. A cloning control reaction, containing only 50 ng of vector PCRs in hybridization buffer should also be prepared in a microfuge tube.
3. Both test (containing insert and vector fragments) and control (vector only) hybridization reactions are then heated to 95°C for 3 min and thermal cycled at 65°C for 2 min and 25°C for 15 min for a total of 4 cycles. This cycling process is essential to the success of the cloning procedure.

3.5. Transformation of Hybridized PCR Amplicons

1. Approx 2 μ L of the hybridization reactions are added to 150 μ L of chemically (CaCl₂) prepared competent *Escherichia coli* (*E. coli*) DH5 α cells (17).
2. The PCR product/cell mix is allowed to transform on ice for 30 min followed by heat shock for 2 min at 42°C.

3. Cells are allowed to recover by incubation in Luria broth for 1 h, shaking at 37°C. An appropriate plasmid-borne antibiotic resistance marker is then challenged for transformant selection. Followed by screening for positive clones by colony PCR using both of the short-insert primers in an amplification reaction as detailed above (*see Subheading 3.1. and Notes 5 and 7*).

4. Notes

1. One benefit of this method is that the quality of genomic and plasmid DNA need not be as high as that required for successive restriction endonuclease and ligation reactions used in non-PCR cloning procedures.
2. The PCR annealing temperature is calculated for only the region that is target specific and would hybridize to template sequence in the first amplification cycle.
3. This approach requires that a number of primers (eight) are synthesized. However, the advantages offered by this technique outweigh the added cost, especially given the low cost of commercial oligonucleotides. It was found for many applications that a common set of vector primers can be used and thus only the two tailed insert primers are required over other regular PCR cloning procedures.
4. This approach requires that both the vector and insert are able to be amplified by PCR. In most circumstance this is not an important limitation. Zhang et al. (15) have recently introduced a PCR-based cloning procedure based on homologous recombination (termed ET cloning) able to overcome this limitation. However, this procedure requires two rounds of transformation and selection and either the use of specialized *E. coli* strains (*sbcA*), or the use of a second modifying plasmid (pBAD-ET λ). While the ET cloning procedure is superior for some applications (e.g., the engineering of very large BAC clones), the simplicity of the described enzyme-free procedure suggests it will prove generally advantageous.
5. As an alternative to purifying plasmid DNA a single bacterial colony containing the required cloning vector may be used in a colony PCR. If this is the case, the level of ddH₂O in the PCR should be adjusted to 11.8 μ L.
6. The NMT primers, NMLF, NMSF, NMLR and NMSR (*see Table 1*), were designed to amplify the N-methyltransferase (NMT) region of the microcystin synthetase gene *mcyA*. The forward long primer, NMLF, was designed to incorporate an amino terminal 6 histidine (6X His) purification tag into both expressed proteins. The pUC19 primers PHL, PHS, PEL, and PES (*see Table 1*), were designed to allow the regulated expression of the 6X His tagged NMT gene from the pUC19 *lac* promoter. Four parallel PCR reactions were performed with the corresponding long and short primer pairs to amplify the pUC19 vector, together with the NMT gene region from the cyanobacterium *Microcystis aeruginosa* PCC7806. The two NMT amplifications were performed using the two primer pairs NMLF/NMRS and NMSF/NMLR. The denaturation and hybridization reaction was performed with 40 ng of the pUC19 pooled PCR products and 15 ng of the NMT (1.2 kb) pooled PCR products in 50 μ L of hybridization buffer as described previously. Two microliters of the reaction was used to transform CaCl₂ competent cells of *E. coli* DH5 α . Transformation of *E. coli* DH5 α cells with the NMT PCR products resulted in 44 colonies. Eight colonies were selected from the NMT transformation and checked by PCR for the appropriate cloned insert DNA using the NMT short primers NMSF/NMSR. Five of the eight plasmids examined were found to contain the correct NMT fragment. A number of individual clones were sequenced to ensure mutations

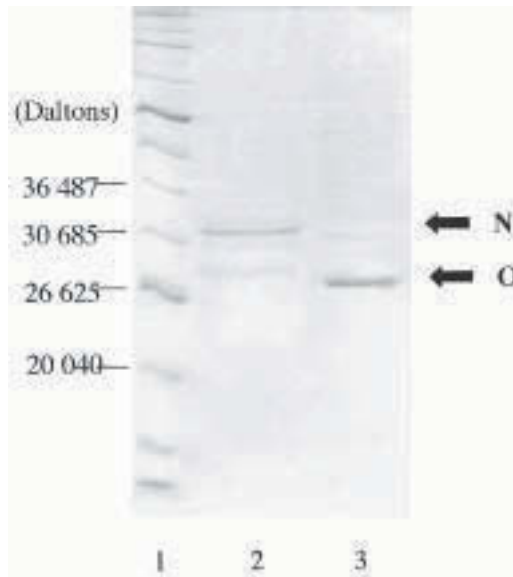


Fig. 2. Sodium dodecylsulfate-polyacrylamide gel electrophoreses (SDS-PAGE) analysis of purified recombinant N-methyltransferase (*see Note 6*) and O-methyltransferase protein fragments from *M. aeruginosa* PCC7806 overexpressed in *E. coli* BL21 using the described enzyme-free cloning method. The 8% polyacrylamide gel was stained with Coomassie blue. The masses of the markers are indicated at the left of the gel. Lane 1: molecular size marker (New England BioLabs), Lane 2: affinity column purified fraction of N-methyltransferase, Lane 3: affinity column purified O-methyltransferase from *M. aeruginosa* PCC7806.

had not been introduced into cloned sequences. Mutation-free clones of the NMT (pNMG) were selected and transformed into the *E. coli* expression strain BL21 (**18**). These clones were successfully used to over-express and purify these two proteins in *E. coli* BL21 (*see Fig. 2*).

7. The enzyme-free procedure provided a highly efficient means for cloning PCR products independent of vector restriction enzyme sites. Theoretically, 50% of the PCR products should be clonable, that is, all molecules with either 5' or 3' overhangs. In practice, transformation efficiencies of up to $8 \times 10^4/\mu\text{g}$ of insert DNA were obtained using standard CaCl_2 competent cells ($10^6/\mu\text{g}$ of pUC19 plasmid). The use of a high-fidelity PCR mix minimizes the introduction of PCR-derived mutations, both within the insert and vector. In addition, no colonies were obtained from the transformation of the control PCR amplified linear vector hybridization reaction.
8. PCR yield and efficiency is typically greater when applied to noncircularized templates. We have seen greater cloning efficiencies when the plasmid vector is linearized prior to vector amplifications. This step is not usually necessary given that only one correct clone is required in many experiments.

References

1. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) Direct cloning and sequencing analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1078.
2. Costa, G. L., Grafsky, A., and Weiner, M. P. (1994) Cloning and analysis of PCR-generated DNA fragments. *PCR Meth. Appl.* **3**, 338–345.
3. Scharf, S. J. (1990) Cloning with PCR, in *PCR Protocols*, Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds. Academic, New York, pp. 84–91.
4. Holton, T. A. and Graham, M. W. (1991) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucl. Acids Res.* **19**, 1156.
5. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**, 1154.
6. Aslanidis, C. and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucl. Acids Res.* **20**, 6069–6074.
7. Aslanidis, C., de Jong, P. J., and Schmitz, G. (1994) Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products. *PCR Meth. Appl.* **4**, 172–177.
8. Hsiao, K.-C. (1993) Exonuclease III induced ligase-free directional subcloning of PCR products. *Nucl. Acids Res.* **21**, 5528–5529.
9. Kaluz, S. and Flint, A. P. F. (1994) Ligation-independent cloning of PCR products with primers containing non-base residues. *Nucl. Acids Res.* **22**, 4845.
10. Nisson, P., Rashtchian, A., and Watkins, P. (1991) Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase. *PCR Meth. Appl.* **1**, 120–123.
11. Yang, Y.-S., Watson, W. J., Tucker, P. W., and Capra, J. D. (1993) Construction of recombinant DNA by exonuclease recession. *Nucl. Acids Res.* **21**, 1889–1993.
12. Bubeck, P., Winkler, M., and Bautsch, W. (1993) Rapid cloning by homologous recombination in vivo. *Nucl. Acids Res.* **21**, 3601–3602.
13. Oliner, J. D., Kinzer, K. W., and Vogelstein, B. (1993) In vivo cloning of PCR products in *E. coli*. *Nucl. Acids Res.* **21**, 5192–5197.
14. Shuldiner, A. R., Scott, L. A., and Roth, J. (1990) PCR-induced (ligase-free) subcloning: a rapid reliable method to subclone polymerase chain reaction (PCR) products. *Nucl. Acids Res.* **18**, 1920.
15. Zhang, Y., Buchholz, F., Muyrers, J. P. P., and Stewart, A. F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Gen.* **20**, 123–128.
16. Liu, Z. (1996) Hetero-stagger cloning: efficient and rapid cloning of PCR products. *Nucl. Acids Res.* **24**, 2458–2459.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, NY.
18. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* **185**, 60–89.
19. Tillett, D. and Neilan, B. A. (1999) Enzyme free cloning: a rapid method to clone PCR products independent of vector restriction enzyme sites. *Nucl. Acids Res.* **27**, e26.

Directional Restriction Site-Free Insertion of PCR Products into Vectors

Guo Jun Chen

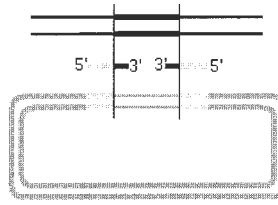
1. Introduction

The polymerase chain reaction (PCR) technique has proved to be a powerful tool for rapid amplification of DNA fragments of interest during cloning. Insertion of PCR products into suitable vectors in order to construct plasmids for protein expression, or to create chimeric genes or to study interactions between proteins and DNA requires a method that allows the precise insertion of a DNA fragment at a defined position in the vector without altering the surrounding DNA sequence. The commonly used cloning methods, such as blunt-end insertion (1,2), introduction of restriction sites on both ends of the PCR products (3,4), T vectors (5) and ligation-independent cloning (6,7) do not usually meet the needs because of the limited cloning sites available on the vector and/or the necessity of the nucleotide change. Hence, the cloning process may become very complicated to insert a PCR product into a defined location on the vector without the change of the nucleotides.

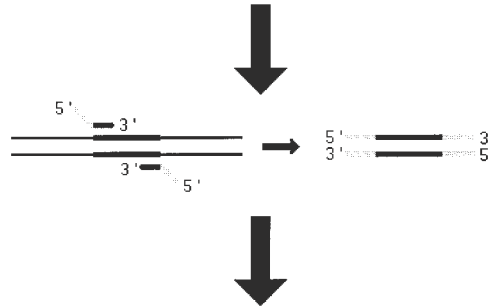
A high-efficiency restriction site-free cloning method (8) has been developed (*see Fig. 1*). It can precisely insert a DNA fragment of interest directionally into a vector at desired location without altering any nucleotide on either the DNA fragment or the vector. This cloning method uses a pair of primers to link two DNA molecules (insert and vector) at a precise junction to form a new double-stranded plasmid. The primer can functionally be divided into two parts: 3' portion and 5' portion. The 3' portion is used to amplify the DNA fragment of interest by PCR. The DNA sequence of this portion is thus homologous to the DNA insert at the junction. The 5' portion is used to generate a PCR product in which the DNA fragment is flanked on both sides by a piece of DNA that is homologous to the vector at the junction. In the thermal-cycling elongation (TCE) step, both the 3' end and the 5' end of each single-stranded DNA of the PCR product will anneal to their complementary strand of the vector. The 3' end will be extended by DNA polymerase using the vector as a template. This results in a fusion between the DNA fragment and the vector. At the end of the TCE reaction,

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

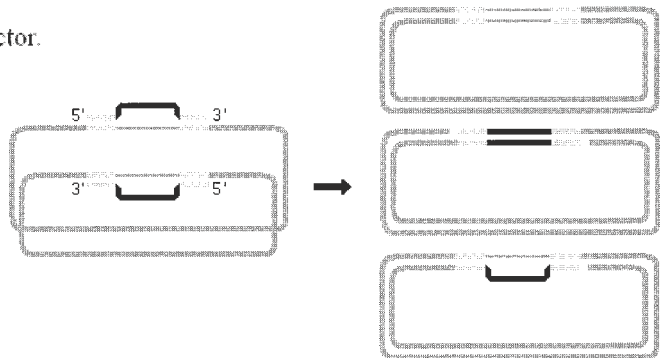
1. Design of primer.



2. Amplification of insert.



3. Fusion of insert and vector.



4. Selection of plasmid.

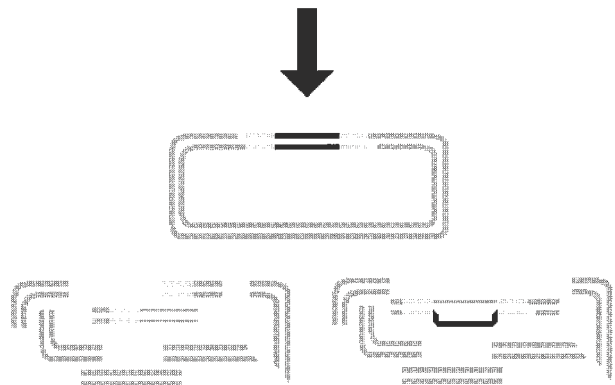


Fig. 1. Scheme of restriction site-free cloning method.

the newly synthesized single-stranded DNAs will anneal to their complementary strands to form stable nicked double-stranded plasmids. *DpnI* endonuclease, which specifically cuts double-stranded methylated and hemimethylated 5' -Gm₆ATC- 3' DNA sequences (9), is used to selectively digest the methylated parental DNA template. The enriched preparation of the plasmid is then used for transformation. The following is an application of this method to construct an IPTG inducible expression plasmid that produces wild type *Echerichia coli* (*E. coli*) peptidyl-tRNA hydrolase.

2. Materials

2.1. Enzymes and Reagents

1. Expand High Fidelity PCR System (Roche Molecular Biochemicals). Store at -20 °C.
2. *PfuTurbo*TM DNA polymerase (Stratagene). Store at -20°C.
3. Deoxynucleotide triphosphates (dNTPs) (Roche Molecular Biochemicals). Store at -20°C.
4. *E. coli* BL 21 chromosomal DNA [according to supplied protocol in QIAamp DNA Mini Kit (Qiagen)]. Store at -80°C.
5. 10X DNA loading buffer: 0.4% bromophenol blue, 0.4% xylene cyanol, 50% glycerol. Store at 4°C.
6. TAE agarose gels. Store at 4°C.
7. 1X TAE running buffer: 50X TAE: 242 g Tris base, 57.1 mL glacial acetic acid, 100 mL 0.5 M ethylenediamine tetraacetic acid (EDTA), H₂O to 1000 mL. Store at room temperature.
8. Ethidium bromide solution: dissolve 50 mg of ethidium bromide in 100 mL of H₂O. Use diluted 1:1000. Store at 4°C and protect from light.
9. QIAquick Gel Extraction Kit (Qiagen). Store at room temperature.
10. Vector pDS56/RBSII. (Caution: Vectors should be isolated from *dam*⁺ *E. coli* stains. Plasmid from *dam*⁻ strains is not suitable for this method.). Store at -80°C.
11. *DpnI* restriction enzyme (Roche Molecular Biochemicals). Store at -20°C.
12. Electroporation competent cells. (Wash cells in log phase (OD₆₀₀ = 0.6–0.8) with 100X cell pellet volume ice cold sterile distilled water or 5–10% glycerol twice by centrifugation at approx 5000 g. Resuspended the cell pellet in 5X volume 10% glycerol solution.) Store at -80°C.
13. SOC medium: 2% tryptone, 0.5% yeast extract, 0.05% NaCl and 20 mM glucose. Store at 4°C.
14. Luria Bertani (LB) ampicillin agar plates. Store at 4°C.
15. LB medium: 1% tryptone, 0.5% yeast extract, 0.5% NaCl. Store at 4°C.

2.2. Equipment

1. Centrifuge (up to 14,000g).
2. 1.5-mL Eppendorf tubes.
3. Thermal-cycler and 0.5-mL Eppendorf tubes.
4. DNA gel box and power supply.
5. Electroporation apparatus and 0.2-cm electroporation cuvetts.
6. 37°C Incubator.
7. Water bath (37°C and 65°C).
8. Ultraviolet (UV) light box.

Table 1
Sample Set for Cloning (see Subheading 3.1.)

Primer	5' Portion sequence	3' Portion sequence
forward	TCACACAGAATTCATTAAGAGGAGAAATTAAC	ATGACGATTAAATTGATTGTCGGCC
reverse	CCCCAGGCGTTTAAGGGCACCAATAACTGCC	TTATTGCGCTTTAAAGGCGTGCAAT

3. Methods

3.1. Primer Design

1. Draw a resulting plasmid map by inserting the DNA fragment into the desired position in the vector. At the junction of the two DNA molecules, select a sequence (at least 30 bases long) from the vector for the 5' portion sequence of the primer and a sequence (at least 20 bases long) from the DNA fragment for the 3' portion. The selected sequences for the 5' portion and the 3' portion should give a predicted melting temperature (T_m) around 60–70°C and 50–60°C, respectively. (The commonly used primer T_m estimation formula is $T_m(^{\circ}\text{C}) = 2 \times [\text{total number of A and T}] + 4 \times [\text{total number of G and C}]$.) The 5' portions or the 3' portions of forward primer and reverse primer should have a similar T_m . Ensure that both forward primers and reverse primers have the right orientation. They should always have the DNA sequence in the 5' to 3' direction from the vector to the DNA fragment (see Note 2).

Table 1 shows a sample set of primers being successfully used for cloning wild-type *E. coli* peptidyl-tRNA hydrolase to pDS56/RBSII (8):

3.2. Amplification of the Peptidyl-tRNA Hydrolase Encoding Region

1. Prepare a 100- μL PCR mixture in a 0.5-mL Eppendorf tube (on ice) as indicated as follows.
 - a. 1X PCR reaction buffer with Mg^{2+} (supplied with Expand High Fidelity PCR System).
 - b. 0.5 μg *E. coli* BL21 strain chromosomal DNA.
 - c. 0.5 μM forward primer.
 - d. 0.5 μM reverse primer.
 - e. 200 μM dNTPs each.
 - f. 2.5 U Expand High Fidelity DNA polymerase (add last, see Note 3).
 - g. Amount of fresh redistilled H_2O to bring the final volume up to 100 μL (add first).
2. Thoroughly mix and then centrifuge briefly to bring the sample to the bottom of the tube.
3. Set a 25-cycle two-step PCR program on a thermal-cycler. The denaturation is at 95°C for 30 s and the annealing/elongation is at 68°C for 1 min with an extra 10 min at 72°C at the end of the reaction. Then hold the reaction at 4°C (see Note 4).
4. Place the Eppendorf tube in the thermal-cycler and turn on the top heating cover. If the thermal-cycler does not have one, carefully add 50 μL mineral oil to overlay the reaction mixture to prevent evaporation. Then start the PCR program.
5. Mix the PCR product with 10 μL 10X DNA loading buffer. Load the sample on a 1% agarose TAE gel. Beware to connect the electrodes of the gel box in correctly to the power supply. Electrophoresis at 5 V/cm in 1X TAE buffer containing 0.5 $\mu\text{g}/\text{mL}$ ethidium bromide until the dye markers have migrated an appropriate distance. (Caution: Ethidium bromide is potent carcinogen. Wear glove to handle it.) (See Note 5).
6. Excise the expected band from the TAE gel on UV light box as quickly as possible. UV light will damage the DNA. (Caution: UV light induces DNA mutagenesis and is harmful to naked eyes. Wear protection glasses and/or a shield.) (See Note 6).

7. Extract the DNA fragment from the gel slice according to the protocol supplied in the QIAquick Gel Extraction Kit (Qiagen).

3.3. Insertion of PCR Product Into Vector

1. Quantify both the vector and the purified PCR product on an agarose gel. The amount of the DNA is estimated according to the known quantity of the marker. The vector should first be linearized by restriction enzyme digestion before applied to the agarose gel for electrophoresis. The running conditions are the same as that in **Subheading 3.2., step 5** (see **Note 7**).
2. Set up a 25- μ L TCE reaction mixture as following:
 - a. 1X PCR reaction buffer (supplied with PfuTurbo™ DNA polymerase).
 - b. 5–10 ng pDS56/RBSII.
 - c. 50–100 ng purified PCR products described above (see **Note 8**).
 - d. 200 μ M dNTPs each.
 - e. 1 U PfuTurbo™ DNA polymerase (add last, see **Note 9**).
 - f. Amount of fresh redistilled H₂O to the final volume of 25 μ L (add first).
3. The thermal-cycle program is 95°C 2 min for denaturation and 68°C 10 min for annealing/elongation with 20 cycles. Hold the reaction at 4°C (see **Note 4**).
4. Place sample in a thermal-cycler and start the TCE program.

3.4. DpnI Digestion and Transformation

1. After TCE reaction, add 5 U of *DpnI* restriction enzyme directly to the TCE reaction mixture. Mix thoroughly and digest for 1–2 h at 37°C (see **Note 10**).
2. Add 1–2 μ L *DpnI* treated TCE reaction mixture to ice-cold 50 μ L *E. coli* Sure 2 electroporation competent cells in a 1.5-mL Eppendorf tube (see **Note 11**).
3. Gently and thoroughly mix and transfer to the bottom of ice-cold Bio-Rad 0.2-cm electroporation cuvet. Be sure the cell droplet contacts both electrodes and is free from air bubbles.
4. Put the cuvet in a BioRad Gene Pulser™ chamber and perform electroporation with a setting of voltage at 2500 V, resistance at 200 Ω , and capacitance at 25 μ F. (Time constants of approx 5 ms are usually obtained.)
5. After electroporation, immediately add 450 μ L SOC medium (prewarmed to room temperature) to the electroporation cuvet and gently mix (see **Note 12**).
6. Transfer the above mixture to a 1.5-mL Eppendorf tube and put in a 37°C water bath for 30 min.
7. Plate 100 μ L of the transformation culture onto an LB agar plate containing 100 μ g/mL ampicillin (see **Note 13**).
8. Incubate the transformation plate upside down at 37°C overnight (12–16 h).
9. Check the plate and pick up single colonies for further analysis.

4. Notes

1. The method described above has also successfully been applied to mutagenesis (**8**). In such a case, the primer is divided into three functional parts: a 5' portion, a mutation-generating portion and a 3' portion. The mutation-generating portion is flanked by the 5' portion and the 3' portion. Changing the nucleotide(s) in this portion can introduce mutations such as substitution, insertion, or deletion in the plasmid. If both forward primer and reverse primer are used to introduce mutations, it can simultaneously mutate two positions that are far away from each other.

2. The length of the primer should be determined by the melting temperature (T_m) of the 5' portion and the 3' portion. The synthesis of long primers is prone to errors, therefore choose a high-quality primer supplier. Another solution is to design the primer as short as possible. In some cases, part of the DNA sequence will overlap between the DNA fragment and the vector (e.g., ATG codon exists in both DNA fragment and vector). This part can be used for both the 5' portion and the 3' portion of the primer. For substitution mutagenesis, the primer can even be reduced to the length of the 3' portion.
3. Expanding high-fidelity DNA polymerase gives high yield compared with *Pfu* DNA polymerase and low error compared with *Taq* DNA polymerase. Errors generated during early PCR steps can be amplified to a dramatic extent. It is suggested to use more templates, less PCR cycles, and high-yield and high-fidelity DNA polymerases.
4. A three-step program with a lower annealing temperature is also suitable for this task. High-annealing temperature usually yields a high-specificity product. Use the DNA extension rate recommended by DNA polymerase producer to calculate the elongation time. The elongation time should be long enough to obtain a full length PCR product or TCE product. It is especially important in the TCE step to ensure the synthesis of a full-length single-stranded plasmid.
5. Nonspecific PCR products will interfere with the TCE reaction and may generate unwanted plasmids because their ends have DNA sequences homologous to the insertion region of the vector. Removing these by-products through electrophoresis will increase the positive percentage.
6. To reduce the DNA damage, use long wavelength UV light or put a thin glass plate between the UV light box and the gel to block some UV light.
7. For more accurate estimation, apply two to three different amount of each sample and marker to the gel. Do not overexpose the film.
8. The ratio of the molarity of the plasmid to the insert is quite important to get good results. The optimum molar ratio is influenced by the number of TCE cycles. With the TCE program described in this protocol, the molar ratio (the plasmid to the insert) of 1 to 50–100 gives the best results for cloning of peptidyl-tRNA hydrolase. It is always a good start to fix the amount of the vector and vary the amount of the PCR product to find out the optimum condition.
9. Theoretically, the thermally stable DNA polymerase used for this task should have 3'–5' exonuclease activity but should not have strand displacement replication function and 5'–3' exonuclease activity. These will ensure that the DNA polymerase stops DNA synthesis right at the 5' end and no extra nucleotides are added to the 3' end of the PCR product.
10. The purpose of *DpnI* treatment is to reduce the background. As *DpnI* cuts fully methylated dsDNA and hemimethylated dsDNA, the plasmid used in this method should be isolated from *dam*⁺ *E. coli* strains. It may use other restriction systems such as the Kunkel method (10) for this task, in which the *E. coli dur*⁺ *ung*⁺ strain will in vivo eliminate the parental plasmid containing deoxyuracil after transformation.
11. At this stage, the *DpnI* treated DNA can also be transformed to the chemically treated competent cell following the protocol provided by the supplier of competent cells. DNA for electroporation must have a very low salt concentration. High ionic strength will cause arcing. Do not use more than 2 μ L *DpnI* treated TCE reaction mixture for a single electroporation. In case more DNA sample for transformation is needed, salt can be removed by DNA precipitation being described as follows.

- a. Add 2 μL Pellet Paint (Novagen) to the sample followed by 2.5 μL 3 M Na-Acetate pH 5.2. Mix the sample briefly. Use of Pellet Paint Co-Precipitant instead of tRNA or collagen as a DNA carrier in DNA precipitation has two advantages: one is that it is easy to visualize the DNA pellet; and two is that all the steps can be carried out at room temperature.
 - b. Add 50 μL ethanol and mix thoroughly.
 - c. Spin the sample in a microcentrifuge at 14,000–16,000g for 5 min.
 - d. Remove the supernatant with a pipet.
 - e. Wash the pink pellet by adding 500 μL of 70% ethanol. Centrifuge at 14,000–16,000g for 2 min. Remove the supernatant and air-dry the pellet.
 - f. Resuspend the pink pellet in deionized water.
12. Cells will die very quickly after electroporation if SOC medium is not added. Therefore, add SOC medium to the cuvet as soon as possible.
 13. Dry the agar plates (exposed upside down) at 37°C for 2–4 h just before use. The plate should be able to soak up to 0.5 mL of media when plating. Do not plate too many cells on a single plate. Sometimes it will mask small colonies on the plate. Try to use more plates if all the transformation samples need to be plated.

References

1. Bhat, G. J., Lodes, M. J., Myler, P. J., and Stuart, K. D. (1991) A simple method for cloning blunt ended DNA fragments. *Nucl. Acids Res.* **19**, 398.
2. Liu, Z. G. and Schwartz, L. M. (1992) An efficient method for blunt-end ligation of PCR products. *Biotechniques* **12**, 28.
3. Karn, J., Brenner, S., Barnett, L., and Cesareni, G. (1980) Novel bacteriophage lambda cloning vector. *Proc. Natl. Acad. Sci. USA* **77**, 5172–5176.
4. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1078.
5. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**, 1154.
6. Aslanidis, C. and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucl. Acids Res.* **18**, 6069–6074.
7. Haun, R. S., Serventi, I. M., and Moss, J. (1992) Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors. *Biotechniques* **13**, 515–518.
8. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. P. (2000) Restriction site-free insertion of PCR products directionally into vectors. *Biotechniques* **28**, 498–500, 504–505.
9. The NEB Transcript 1998, vol. 9 no. 1, page 7, GATC: DpnI, DpnII, MboI or Sau3A I.
10. Ling, M. M. and Robinson, B. H. (1997) Approaches to DNA mutagenesis: an overview. *Analyt. Biochem.* **254**, 157–178.

Autosticky PCR

Directional Cloning of PCR Products with Preformed 5' Overhangs

József Gál and Miklós Kálmán

1. Introduction

The polymerase chain reaction (PCR) is a method of central importance in molecular biology (1,2). The DNA fragment of interest is often amplified for cloning purposes. A frequently used experimental approach is to include extra restriction endonuclease cleavage sites in the amplification primers, digestion of the PCR product with the corresponding enzymes, and ligation of the product to a linearized cloning vector (3). However, the efficiency of cleavage by certain restriction endonucleases is rather low because of the cleavage site(s) being too close to the termini of a DNA fragment (4,5), and internal restriction sites of the fragment might also complicate the cloning.

During PCR, the DNA fragment of interest is amplified with primer oligonucleotides. Each strand of the amplified DNA fragment contains a built-in amplification primer in its 5' terminus. When a strand of the PCR product serves as a template for the amplifying DNA polymerase, the segment that is copied last is just the built-in primer.

In nucleic acids, abasic sites are positions where the base is missing from the sugar-phosphate backbone. Abasic sites in a DNA template are noninstructional for a DNA polymerase, making it stall during synthesis of the complementary strand (6–11). Autosticky PCR (AS-PCR) products are amplified with primers containing abasic sites (10,11). First, the primer is incorporated into a PCR product strand. When this strand acts as template during further cycles, the amplifying polymerase is stalled at the primer-borne abasic position, resulting in the formation of single-stranded 5' overhangs on the termini of the AS-PCR product. The overhangs enable ligation of the AS-PCR product to a vector cut by the corresponding enzymes.

A stable structural analog of 2'-deoxyribose, a tetrahydrofuran derivative, was chosen instead of naturally occurring abasic sites (7). It differs from the natural 2'-deox-

*Patent pending No. P9801320 Hungarian Patent Office.

From: *Methods in Molecular Biology*, Vol. 192: *PCR Cloning Protocols*, 2nd Edition
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

ribose abasic site by having a hydrogen instead of a hydroxyl group on the 1' carbon of the deoxyribose ring. This difference makes tetrahydrofuran abasic sites highly resistant to the chemical conditions applied during conventional oligonucleotide synthesis based on phosphoramidite chemistry, enabling the incorporation of these sites into primers. The protected derivative of the tetrahydrofuran abasic site, ready for primer synthesis, is commercially available.

The behavior of DNA polymerases at the terminus of a linear template (12,13) and at an abasic site proved to be similar. *Pfu* and *Vent*, which are proofreading polymerases (14,15), are stalled before abasic sites without extra 3' nucleotide addition (9,11). *Taq* polymerase does not possess proofreading 3'-5' exonuclease activity (16,17), and adds an extra nucleotide, predominantly a dAMP residue, opposite an abasic site (8,10,11). Our results showed that in the ligation junction of the AS-PCR product and the vector, the abasic site should be opposed either by a polymerase-added extra nucleotide, or the 5' terminal residue of the vector-borne overhang (11). According to the aforementioned, the nontemplated 3' extra nucleotide addition activity of the desired polymerase should be taken into consideration when designing AS-PCR primers.

In the resulting clones, the abasic position is either substituted by a nucleotide, frequently a dTMP residue, or is deleted. Rarely, the deletion of the abasic position is accompanied by a minor deletion in the vicinity of the position (10,11).

The AS-PCR method has several advantages. It allows directional cloning of PCR products. Modification of the amplification product is unnecessary before ligation, so the end sensitivity of restriction enzymes (4,5) does not hamper cloning. Theoretically, any desired 5' overhang can be generated, including overhangs that correspond to restriction sites present within the amplified sequence. A drawback of the method is that it is rather difficult to predict the result of the repair of the abasic positions.

2. Materials (see Note 1)

1. dSpacer, a protected tetrahydrofuran derivative phosphoramidite (7) and the so-called Chemical Phosphorylation Reagent (18) are commercially available (Glen Research, Sterling, VA).

Both chemicals should be dissolved in acetonitrile (chemical DNA synthesis grade) and kept moisture-free. After they are dissolved, they should be used within 2–3 d, 1 wk maximum. The reagents might be taken back off the machine. For best results, seal them under argon. They should then be stored at -80°C , where the acetonitrile is frozen. Be very careful to keep them protected from moisture when the solutions are warming up.

2. Vector DNA (preferably at 0.2–1 $\mu\text{g}/\mu\text{L}$ concentration).
3. Restriction endonucleases with the corresponding 10X buffers.
4. Sterile, deionized water, free from DNase and DNA.
5. 100% (or 96%) and 70% ethanol. Keep at -20°C .
6. 3 M Sodium acetate solution, pH 5.2.
7. Phenol:chloroform:isoamylalcohol 25:24:1 mixture (19).
8. Electrophoresis-grade agarose.
9. 1X TAE: 40 mM Tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA) (19).
10. 10 mg/mL ethidium bromide solution. Ethidium bromide is a powerful mutagen, so it should be used with caution.

11. 6X gel loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, 40% sucrose. Keep at 4°C.
12. The QIAexII Gel Extraction Kit (QIAGEN, Valencia, CA) is recommended for the recovery of DNA after preparative gel electrophoresis.
13. *Taq* DNA polymerase can be obtained from several providers, e.g., Invitrogen, Carlsbad, CA. *Pfu* DNA polymerase is available from Stratagene, La Jolla, CA. Vent DNA polymerase can be purchased from New England Biolabs, Beverly, CA. The commercially available DNA polymerases are usually provided with a 10X reaction buffer. Usually a solution of a Mg²⁺ salt is also supplied by the manufacturer, if it is not contained within the 10X reaction buffer.

In the lack of a manufacturer-provided buffer, the composition of 10X buffers are:

 - a. *Taq*: 200 mM Tris-HCl (pH 8.4 at 25°C), 500 mM KCl.
 - b. *Pfu*: 200 mM Tris-HCl (pH 8.75 at 25°C), 100 mM KCl, 100 mM (NH₄)₂SO₄, 1% Triton X-100, 1 mg/mL bovine serum albumin (BSA).
 - c. Vent: 200 mM Tris-HCl (pH 8.8 at 25°C), 100 mM KCl, 100 mM (NH₄)₂SO₄, 1% Triton X-100.
14. dNTP solutions are available from several commercial providers, e.g., Invitrogen. It is convenient for further work to prepare a stock solution containing 5 mM of each dNTP.
15. 10 μM (10 pmol/μL) solution of the AS-PCR primers.
16. Light mineral oil, if the PCR cycler operates without a heated lid.
17. The QIAquick PCR Purification Kit (QIAGEN), or a similar kit is recommended for the purification of the AS-PCR product.
18. T4 DNA ligase is available from several commercial providers, e.g., New England Biolabs. The enzyme is usually provided with a 10X concentrated buffer. If there is no manufacturer-provided 10X buffer, its recommended composition is 500 mM Tris-HCl (pH 7.5 at 25°C), 100 mM MgCl₂, 100 mM dithiothreitol (DTT), 10 mM rATP, 250 μg/mL BSA.
19. DH5α *Escherichia coli* (*E. coli*) competent cells, either commercially available (e.g., from Invitrogen), or homemade (20).
20. LB broth and LB/1.5% agar plates containing the corresponding antibiotic (19).
21. Sol-1: 50 mM Tris-HCl, 10 mM EDTA, pH 8.0 at 25°C.
22. Sol-2: 200 mM NaOH, 1% sodiumdodecyl sulfate (SDS). Best if freshly prepared before use.
23. Sol-3: 3 M potassium acetate, pH 5.5. Chill on ice before use.
24. Isopropanol. Keep at room temperature.
25. TE solution: 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 at 25°C.
26. RNaseA (19), 100 μg/mL solution (free from DNase).

3. Methods

3.1. Primer Design

The design of the primers is a very important step of AS-PCR. It was found that there should be a nucleotide opposite the abasic site in the AS-PCR product-vector ligation junction (11). Without this nucleotide, AS-PCR products were difficult or impossible to clone. The opposing nucleotide can either be a polymerase-added extra nucleotide, or the 5' terminal nucleotide of the vector-borne overhang. According to the above, the primer design should consider the 3' extra nucleotide addition capacity of the used polymerase.

Taq polymerase adds an extra nucleotide opposite abasic sites (8,10,11). It was also found that *Taq* polymerase can proceed a further nucleotide over the abasic position before it ceases synthesis (11), so it is advisable to contain a nucleotide between the restriction overhang and the abasic site, to provide “buffer” space for the read-through. This latter nucleotide can be exploited for the regeneration of the restriction site by means of the in vivo repair. Based on the above, AS-PCR primers designed for use with *Taq* polymerase should be arranged as: 5'-restriction overhang-"buffer" nucleotide-abasic position-template hybridizing segment-3'. For example, if the primer is designed for a *Bam*HI site, 5'-P-GATC-C-S-template hybridizing segment-3', where “P” stands for a phosphate group, “S” for a tetrahydrofuran abasic site (see Fig. 1).

Pfu and *Vent* polymerases do not add a nucleotide opposite the abasic position when stalled (9,11). In this case, AS-PCR primers should be designed so as in the ligation junction between the AS-PCR product and the linearized vector, the 5' terminal nucleotide of the vector-borne overhang is positioned opposite the abasic site. For example, if the primer is designed for a *Bam*HI site, 5'-P-GAT-S-template hybridizing segment-3' (see Fig. 1).

Double noninstructional sites in the template block DNA polymerases more efficiently than a single site (21). However, it is not recommended to use tandem abasic sites in AS-PCR primers, because it dramatically lowers the efficiency of cloning, probably because of the difficult repair situation (11).

It is also very important to phosphorylate AS-PCR primers on their 5' termini. Dephosphorylation of AS-PCR products significantly decreased the efficiency of cloning (11). It is very important that the 5' overhangs of the AS-PCR product could be ligated.

Only the template-hybridizing portion of AS-PCR primers should be taken into consideration when the annealing temperature is determined. For hybridizing segments around 20–25 nucleotides long, a well-working empirical formula is: $4X$ (number of G or C residues) + $2X$ (number of A or T residues) – 5 = recommended annealing temperature in °C. We suggest to design the hybridizing segment for about 55°C annealing temperature, based on the aforementioned formula. The general considerations of primer design (minimal secondary structure, minimal hybridization of the two primers with each other, no long homopolymeric segments, and so on) also apply to AS-PCR primer design.

3.2. Primer Synthesis (see Note 2)

Several companies that provide custom oligonucleotide synthesis (e.g., Sigma-Genosys, The Woodlands, TX, or New England Biolabs) offer the synthesis of tetrahydrofuran abasic site-containing oligonucleotides. The chemical 5' phosphorylation of AS-PCR primers is strongly recommended, because enzymatic phosphorylation is generally less efficient (18).

In case of having access to an oligonucleotide synthesizer, we recommend using dSpacer and the Chemical Phosphorylation Reagent of Glen Research, Sterling, VA. No changes of coupling conditions and deprotection are needed from the standard procedure of phosphoramidite chemistry-based oligonucleotide synthesis (see Note 3).

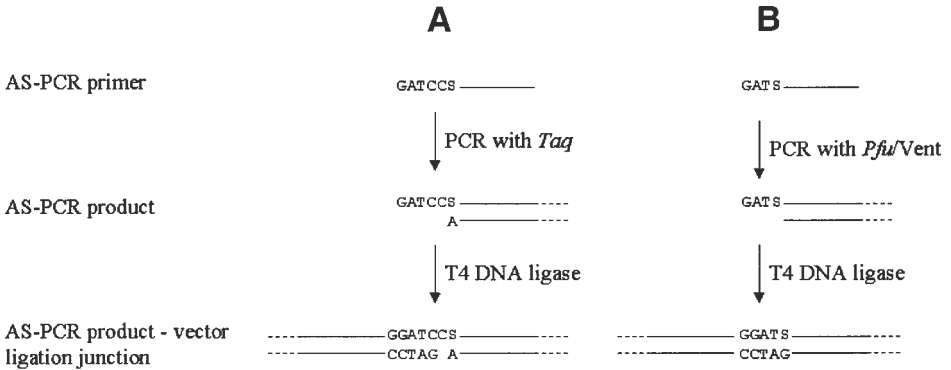


Fig. 1. The diagram of AS-PCR, shown in the case of primers designed for use with a *Bam*HI site. ‘S’, tetrahydrofuran abasic site. The dashed lines represent continuation of the DNA. *A.*, AS-PCR primer for use with *Taq* polymerase. *B.*, AS-PCR primer for use with *Pfu* or *Vent* polymerase.

3.3. Vector Preparation (see Notes 4 and 5)

The cloning vector should be cut with the two restriction endonucleases (generating noncompatible 5' overhangs), and the vector fragment isolated by preparative gel electrophoresis.

It is recommended to perform the two restriction cleavages one after the other, not together at the same time, because of the monitoring of the reactions. If possible, use the optimal, manufacturer-provided, usually 10X concentrated buffer for each digestion, to avert the adverse effects of the decreased specificity of restriction endonucleases (“star activity”). Consult the manufacturer’s catalog if there is a common buffer recommended for the desired double digestion, namely a buffer in which both enzymes work with high efficiency, without significant star activity. If no common buffer is available for the two restriction endonucleases, and at least one of the enzymes can be heat inactivated, start with that enzyme. Consult the product description of enzymes to find out if they can be heat inactivated (see Note 4).

1. Assemble the first restriction digestion in a 1.5-mL reaction tube:
 - a. Vector DNA 5 μ g
 - b. 10X buffer for first digestion 5 μ L
 - c. First restriction endonuclease 10–40 U (see Note 6)
 - d. Sterile deionized water up to 50 μ L

The restriction endonuclease should always be added last to the mixture. Mix components with a pipet, and incubate at the temperature optimum of the first restriction endonuclease for 1 h. Monitor the progress of the digestion by running a 3- μ L aliquot of the mixture next to uncut vector DNA on an agarose gel. If the digestion is only partial, add more enzyme (see Note 7), and prolong the incubation.

2. If there is a common buffer in which both enzymes perform well without star activity, go to step 2a. Otherwise, go to step 2b.

- a. If the first digestion is complete, add 10–40 U of the second enzyme to the mixture, mix with a pipet, and incubate at the temperature optimum of the second enzyme for at least 1 h (see **Notes 6** and **7**). If possible, monitor the second cleavage by agarose gel electrophoresis (see **Note 8**). Proceed to **step 4**.
 - b. If the first enzyme can be heat inactivated, heat inactivate the enzyme following the manufacturer's recommendations, then go to **step 3**, (see **Note 9**). Otherwise, extract the first digestion mixture with an equal volume of a 25:24:1 mixture of phenol:chloroform:isoamylalcohol, either by vortexing, or in the case of plasmids larger than 10 kb, by inverting or flicking the tube several times. Spin in a table-top centrifuge at 16,000g for 5 min. Carefully transfer the upper (aqueous) phase into a new 1.5-mL reaction tube with a pipet. Go to **step 3**.
3. Precipitate DNA by adding 0.1 volume of 3 M sodium acetate, pH 5.2 and 2.5 volume of cold 100% ethanol to 1 volume of DNA solution. Mix thoroughly by inverting the tube several times, and incubate at –20°C for 30 min. Pellet DNA by centrifugation at 16,000g in a table-top centrifuge for 10 min, discard supernatant, and wash pellet with 1 mL cold 70% ethanol. Air-dry pellet and dissolve it in a volume of sterile deionized water as determined below, then add the further components of the second restriction mixture.
 - a. 10X buffer for second enzyme 5 µL
 - b. Second restriction endonuclease 10–40 U
 - c. Sterile deionized water up to 50 µL

Mix components with a pipet and incubate at the temperature optimum for the second enzyme for at least 1 h. If possible, monitor the second cleavage with agarose gel electrophoresis (see **Note 8**). Proceed to **step 4**.
 4. The double-cut vector fragment should be isolated by preparative gel electrophoresis.

3.4. Amplification (see **Note 10**)

1. Assemble the following reaction mixture in a thin-wall PCR tube on ice:

a. 10X PCR buffer	10 µL
b. 25 mM Mg ²⁺ salt (see Note 11)	8 µL
c. 5 mM (each) dNTP	4 µL
d. Template DNA	Plasmid: 1–100 ng Genomic DNA: 100–250 ng
e. 10 µM upstream AS-PCR primer	10 µL
f. 10 µM downstream AS-PCR primer	10 µL
g. Thermophilic DNA Polymerase	<i>Taq</i> , <i>Pfu</i> : 2.5 U Vent: 2 U
h. Deionized, DNA-free water	up to 100 µL
2. If the thermal cycler operates without a heated lid, overlay the reaction mixture with 70 µL light mineral oil.
3. The recommended PCR program is:
 - a. Initial denaturation 94°C, 4 min (see **Note 12**)
 - b. 20–25 cycles of: 94°C, 1 min
50–55°C (see **Subheading 3.1.**), 1 min
72°C: *Taq*, Vent: 1 min per each kilobase, at least 1 min; *Pfu*:
2 min per each kilobase, at least 1 min
 - c. Final extension 72°C, 5 min

3.5. Analysis and Purification of the AS-PCR Product

Run 3 μL of the amplification mixture on an agarose gel next to a suitable molecular weight marker.

1. If the product with the expected molecular weight is absent, or is hardly visible, the reaction conditions should be optimized.
2. If only a single fragment with the expected length is apparently present, it is advised to purify the product from the amplifying polymerase and residual unused primers with the QIAquick PCR Purification Kit (QIAGEN), or a similar kit (*see Note 13*). Alternatively, preparative gel electrophoresis is also sufficient.
3. If unexpected bands are also present, it is strongly recommended to purify the fragment of interest by preparative gel electrophoresis.

3.6. Ligation

1. After purification, the concentrations of the AS-PCR product and the isolated, double-cut vector fragment should be estimated with agarose gel electrophoresis. An aliquot should be run next to a DNA standard of known concentration, e.g., a DNA size marker in which the amount of DNA running in a nearby band is known.
2. Assemble the ligation reaction as follows:
 - a. Vector DNA (double-cut, isolated) 50 ng
 - b. AS-PCR product (in fivefold *molar* excess over the vector; *see Note 14*)
 - c. 10X T4 DNA ligase buffer 1.5 μL
 - d. T4 DNA ligase 1.5 Weiss U
(= 100 cohesive end ligation U)
 - e. Sterile deionized water up to 15 μLIncubate overnight at 16 °C.
3. Also assemble a control reaction in which the AS-PCR product is replaced with the buffer that the AS-PCR product is dissolved in.

3.7. Transformation

To elevate the efficiency of transformation, T4 DNA ligase might be inactivated before transformation by incubation at 65°C for 10 min.

1. Thaw DH5 α *E. coli* competent cells on ice.
2. Add 5 μL of the ligation mixture to 100 μL of cells. Mix by flicking the tube a few times.
3. Incubate cells on ice for 20 min.
4. Heat shock cells at 42°C for 30 s.
5. Place cells back on ice for 2 min.
6. Add 900 μL of LB broth, and incubate cells at 37°C for at least 30 min, but no longer than 60 min.
7. Plate 100 and 900 μL of the transformation mixture onto LB/1.5% agar plates containing the appropriate antibiotic (*see Note 15*). Alternatively, plate 100 μL of the transformation mixture, and keep the remainder at 4°C. Usually, transformed cells can be kept at 4°C for a day without substantial loss of viability.
8. Incubate plates overnight at 37°C (*see Note 16*).

3.8. Plasmid Purification

1. Pick colonies with an inoculating loop or sterile toothpicks, and inoculate 3 mL LB broth containing the appropriate antibiotic (*see* **Note 15**). Shake overnight at 37°C. The bacterial culture should reach stationary phase.
2. Pellet cells from 1.5 mL of the bacterial culture in a 1.5 mL reaction tube by spinning in a table-top centrifuge at 16,000g for 1 min. Discard supernatant.
3. Repeat **step 2** with the rest of the bacterial culture.
4. Resuspend cells in 300 µL Sol-1 with vortexing.
5. Add 300 µL Sol-2, and invert the tube several times. Do not vortex or shake the tube at this point. Incubate at room temperature for 5 min.
6. Add 300 µL ice-cold Sol-3, and invert the tube several times. Do not vortex or shake the tube at this point. Place on ice for 10 min.
7. Pellet the precipitated material by spinning in a table-top centrifuge at 16,000g for 10 min.
8. Carefully transfer supernatant (approx 850 µL) with a pipet into a new 1.5-mL reaction tube.
9. Precipitate nucleic acids by adding 600 µL of room-temperature isopropanol. Mix thoroughly by inverting the tube several times, and spin at 16,000g for 10 min in a table-top centrifuge.
10. Discard supernatant. Wash pellet with 1 mL cold 70% ethanol.
11. Air-dry pellet.
12. Dissolve pellet in 180 µL of TE buffer.
13. Add 20 µL 100 µg/mL RNaseA solution. Incubate at 37°C for 1 h.
14. Extract the RNaseA mixture with 200 µL phenol:chloroform:isoamylalcohol (25:24:1) by vortexing, or in the case of plasmids larger than 10 kb, by inverting the tube several times.
15. Spin at 16,000g for 5 min in a table-top centrifuge. Transfer upper (aqueous) phase into a new 1.5-mL reaction tube. Add 20 µL 3 M sodium acetate, pH 5.2 and 500 µL cold 100% ethanol. Mix thoroughly by inverting the tube several times.
16. Incubate at -20°C for 30 min.
17. Spin at 16,000g for 10 min in a table-top centrifuge. Discard supernatant. Wash pellet with 1 mL cold 70% ethanol. Air-dry pellet.
18. Dissolve pellet in TE buffer or sterile deionized water. (In the latter case, it is important to store the sample at -20°C for prolonged periods of time.) The recommended volumes:
 - a. High copy number plasmids (e.g., pBluescript) 100 µL
 - b. Intermediate copy number plasmids (e.g., pBR322) 30 µL
 - c. For low copy number plasmids (e.g., pSC101), decrease the volume further.

3.9. Analysis of Clones

3 µL of the plasmid solution should be checked by restriction digestion with enzymes cleaving in the flanking regions of the insert (*see* **Note 17**). In the case of the *Taq* AS-PCR products, if the primers were designed to regenerate the cloning sites, the resulting clones can be checked by digestion with the restriction endonucleases used for the cloning. Positive clones should also be sequenced.

4. Notes

1. For useful hints in the preparation of stock solutions, turn to **ref. 19**.
2. Do not use tandem abasic sites in AS-PCR primers (**II**). Do not forget to phosphorylate the 5' termini of the primers (**II**).

3. Detailed information on dilution and coupling conditions for various synthesizers is available at <http://www.glenres.com/>.
4. Detailed information on restriction digestion conditions, recommended buffers for double digests, heat inactivation of restriction endonucleases, etc. is available at <http://www.neb.com/>.
5. For a detailed description of analytical and preparative agarose gel electrophoresis, turn to (19). For the recovery of the DNA from a cut-out gel slice, we recommend to use the QIAexII Gel Extraction Kit (QIAGEN) for its high yield, but several other methods are also available, e.g., phenol extraction of low melting temperature agarose gel (19), low-speed centrifugation (22), or electroelution (23).
6. In most of the cases, 2 μL of a manufacturer-provided restriction endonuclease stock, usually corresponding to 10-40 U of enzyme, is enough for the digestion of 5 μg vector DNA.
7. Restriction endonucleases are usually provided in a stock solution containing 50% glycerol. The final concentration of glycerol in the restriction mixture should not exceed 5%, otherwise it might elicit star activity of certain restriction endonucleases. According to the aforementioned, the volume of the enzyme added as a 50% glycerol stock should not exceed 10 % of the volume of the mixture.
8. If the second cleavage does not remove at least 5% of the single-cut vector, e.g., if both cleavages occur within a polycloning site of a vector, monitoring of the second cleavage by agarose gel electrophoresis might be difficult, or even impossible. In this case, do a separate, analytical scale digestion to make sure that the second enzyme does cut the vector.
9. Many restriction endonucleases can be inactivated with a 20-min incubation at 65°C. In some cases, an 80°C incubation might be necessary. However, certain restriction endonucleases survive even an 80°C heat treatment. In the latter case, a phenol:chloroform:isoamylalcohol extraction of the DNA is necessary. For the thermal stability of a given enzyme, consult the manufacturer's product description.
10. The presented conditions work well with most amplifications. However, optimization of certain conditions, e.g., concentration of the Mg^{2+} , annealing temperature, number of cycles, amount of enzyme, concentration of the primers, etc. might be necessary in certain cases.
11. For *Taq* polymerase, MgCl_2 is recommended; for *Pfu* and Vent, use MgSO_4 . Add the Mg^{2+} salt only if it is not contained within the 10X PCR buffer.
12. If an antibody-based "hot start" polymerase is used, a longer initial denaturation, or additional cycles might be necessary; follow the manufacturer's instructions.
13. It is strongly recommended to purify the AS-PCR product from the amplifying polymerase before ligation. Polymerase and dNTP carryover from the PCR mixture caused self-circularization of the vector because of filling-in of the vector-borne 5' restriction overhangs, even at 16°C (10). It is also recommended to purify the amplification product from AS-PCR primers. An AS-PCR primer, due to its structure, could also ligate to vector-borne 5' overhangs, decreasing the amount of the vector available for cloning. For the purification of the AS-PCR product from PCR mixture components, we used the QIAquick PCR Purification Kit (QIAGEN), but other methods should also work.
14. The highest cloning efficiency was obtained when AS-PCR products were cloned right after amplification and purification.
15. The recommended final concentrations of some frequently used antibiotics: ampicillin, 100 $\mu\text{g}/\text{mL}$; chloramphenicol, 30 $\mu\text{g}/\text{mL}$; Kanamycin, 50 $\mu\text{g}/\text{mL}$; streptomycin, 25 $\mu\text{g}/\text{mL}$; tetracycline, 12.5 $\mu\text{g}/\text{mL}$.

16. The plate with the vector control should contain a low number of colonies (background) as compared to the plate with the AS-PCR product-vector ligation. Otherwise, probably at least one of the restriction digestions of the vector was only partial, or some solution might have been contaminated with DNase, and removal of the vector-borne sticky ends might have led to recircularization of the vector by blunt-ligation.
17. If the cloned DNA fragment is short (below 500 basepairs), it might be necessary to digest more plasmid to generate visible amount of the insert on agarose gel.

References

1. Mullis, K. B. and Faloona, F. A. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Meth. Enzymol.* **155**, 335–350.
2. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.
3. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1078.
4. Jung, V., Pestka, S. B., and Pestka, S. (1990) Efficient cloning of PCR generated DNA containing terminal restriction endonuclease recognition sites. *Nucl. Acids Res.* **18**, 6156.
5. Kaufman, D. L. and Evans, G. A. (1990) Restriction endonuclease cleavage at the termini of PCR products. *Biotechniques* **9**, 304–306.
6. Schaaper, R. M., Kunkel, T. A., and Loeb, L. A. (1983) Infidelity of DNA synthesis associated with bypass of apurinic sites. *Proc. Natl. Acad. Sci. USA* **80**, 487–491.
7. Takeshita, M., Chang, C.-N., Johnson, F., Will, S., and Grollman, A. P. (1987) Oligodeoxynucleotides containing synthetic abasic sites. *J. Biol. Chem.* **262**, 10,171–10,179.
8. Paabo, S., Irwin, D. M., and Wilson, A. C. (1990) DNA damage promotes jumping between templates during enzymatic amplification. *J. Biol. Chem.* **265**, 4718–4721.
9. Greagg, M. A., Fogg, M. J., Panayotou, G., Evans, S. J., Connolly, B. A., and Pearl, L. H. (1999) A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc. Natl. Acad. Sci. USA* **96**, 9045–9050.
10. Gál, J., Schnell, R., Szekeres, S., and Kálmán, M. (1999) Directional cloning of native PCR products with preformed sticky ends (Autosticky PCR). *Mol. Gen. Genet.* **260**, 569–573.
11. Gál, J., Schnell, R., and Kálmán, M. (2000) Polymerase dependence of Autosticky polymerase chain reaction. *Analyt. Biochem.* **282**, 156–158.
12. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677–9686.
13. Hu, G. (1993) DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol.* **12**, 763–770.
14. Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A., and Mathur, E. J. (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**, 1–6.
15. Mattila, P., Korpela, J., Tenkanen, T., and Pitkanen, K. (1991) Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase - an extremely heat stable enzyme with proofreading activity. *Nucl. Acids Res.* **19**, 4967–4973.
16. Chien, A., Edgar, D. B., and Trela, J. M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* **127**, 1550–1557.
17. Tindall, K. R. and Kunkel, T. A. (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**, 6008–6013.

18. Horn, T. and Urdea, M. S. (1986) A chemical 5'-phosphorylation of oligodeoxyribonucleotides that can be monitored by trityl cation release. *Tetrahedron Lett.* **27**, 4705–4708.
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580.
21. Newton, C. R., Holland, D., Heptinstall, L. E., Hodgson, I., Edge, M. D., Markham, A. F., and McLean, M. J. (1993) The production of PCR products with 5' single-stranded tails using primers that incorporate novel phosphoramidite intermediates. *Nucl. Acids Res.* **21**, 1155–1162.
22. Heery, D. M., Gannon, F., and Powell, R. (1990) A simple method for subcloning DNA fragments from gel slices. *Trends Genet.* **6**, 173.
23. Zhen, L. and Swank, R. T. (1993) A simple and high yield method for recovering DNA from agarose gels. *Biotechniques* **14**, 894–898.

A Rapid and Simple Procedure for Direct Cloning of PCR Products into Baculoviruses

Tamara S. Gritsun, Michael V. Mikhailov, and Ernest A. Gould

1. Introduction

Recombinant baculovirus expression systems are among the most commonly used for expressing foreign genes in eukaryotic cells (1,2). This eukaryotic expression system became very popular for many reasons, including 1. potentially high-protein expression levels, 2. ease and speed of genetic engineering, 3. ability to accommodate large DNA inserts, 4. protein processing similar to higher eukaryotic cells (e.g., mammalian cells), and 5. ease of insect cell growth (3–5). Owing to their popularity, there have been considerable advances in the development of recombinant baculovirus systems to extend their range of application to include medicinally important pharmaceuticals (6), vaccine development (7,8) and rapid-action biological insecticides (9). As an expression system, the baculoviruses have been improved and modified to facilitate easy laboratory manipulation and high-level protein expression (10–12).

Baculoviruses are a diverse group of insect viruses. They have a double-stranded circular DNA genome about 130 kb in length. The most commonly used baculovirus for protein expression is *Autographa californica* nuclear polyhedrosis virus (AcMNPV). Conventionally, the foreign gene is inserted into the baculovirus expression vector through a bacterial transfer vector, which contains two “recombinant arms” (RA), i.e., long baculovirus sequences consisting of 300–3000 nucleotides. These RA provide classical homologous recombination between the transfer vector and the baculovirus during cotransfection (see Fig. 1A). In some cases this conventional approach is impossible because of incompatibility between the eukaryotic sequences and the bacterial systems used for their cloning. One method that avoids the bacterial cloning stage is based on the direct ligation of foreign DNA into a specific restriction site (see Fig. 1B) within the baculovirus (13,14). Another approach uses ligation between the transfer vector and the foreign DNA, followed by cotransfection of the purified ligation complex (see Fig. 1C) and the baculovirus DNA (15).

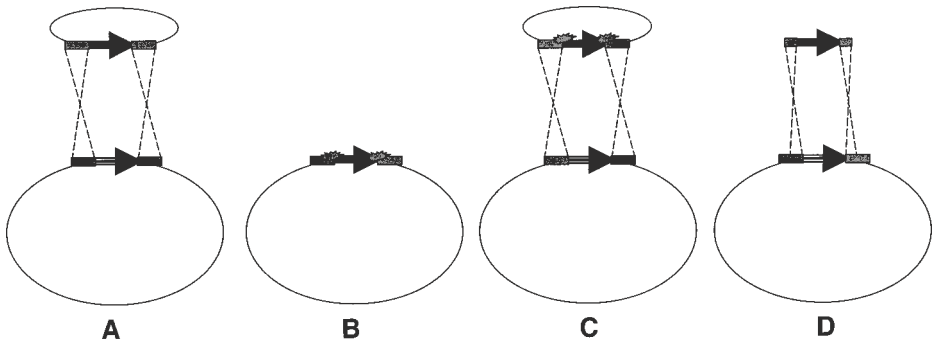


Fig. 1. Different cloning strategies of a protein gene. The baculovirus genome is shown as the larger ellipse with authentic gene designed for substitution (stripe-shafted arrow) and recombination arms (RA, gray boxes). The transfer vector is shown as the smaller ellipse with the gene destined to be cloned into the baculovirus (thick black arrow) flanked by the RA. Crossed dashed lines show homologous recombination resulting in genetic exchange between genes. Sites of ligation are depicted as shadowed stars. (A) The gene Y is cloned into transfer vector. (B) Direct ligation of gene to baculovirus DNA genome. (C) The gene is ligated to transfer vector. (D) PCR product of gene contains the truncated RA that enable homologous recombination.

Here we present a novel alternative approach, by introducing foreign genes into baculoviruses without an intermediate bacterial stage. The method, which is based on the principal of homologous recombination, utilizes polymerase chain reaction (PCR) primers that contain 50 nucleotides of RA representing the site where recombination is required (*see Fig. 1D*). The method is simple, rapid, and reliable and avoids the use of cumbersome techniques associated with enzymatic treatment and DNA purification. The method could be used not only for expression of genes, but also for cloning purposes when bacterial cloning of DNA with “difficult” sequences appears impossible.

This chapter does not provide the details for manipulation of baculoviruses and baculovirus DNA because these procedures have been described many times in specialized manuals and books (*see Note 1*). Moreover, many commercial companies offer ready-to-use kits supplying the full range of required chemicals, purified DNA, and manuals. Our method for direct PCR cloning could help scientists who are already familiar with the baculovirus expression system. For those who are not, we recommend the comprehensive laboratory guides (*1,2*) describing the details of virus growth, plaque assay, and virus genomic DNA purification.

1.1. Theory of the Method

The key elements of our PCR cloning method are to define the locus on the baculovirus genome for insertion of the PCR product and then to design appropriate primers with RA that will enable homologous recombination between the PCR product and the baculovirus. Here we present the sequence of the baculovirus polyhedrin locus with an inserted β -galactosidase gene (*lacZ*) [baculovirus Acrp23-*lacZ* (*16*)]

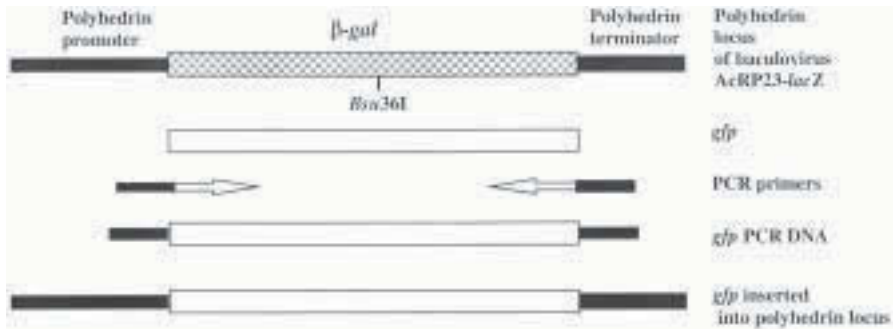


Fig. 2. Strategy for direct cloning of GFP gene into baculovirus genome. Restriction enzyme *Bsu36I* was used to linearize baculovirus DNA (AcRP23-*lacZ*) inside the β -galactosidase gene. Both upstream and downstream primers for PCR contained sequences of the baculovirus polyhedrin locus (black boxes) and GFP gene (white arrows). The viral DNA and GFP PCR DNA were mixed and cotransfected into insect cells. The resulting recombinant virus expressed the 238 amino acid GFP, replacing expression of the β -galactosidase gene.

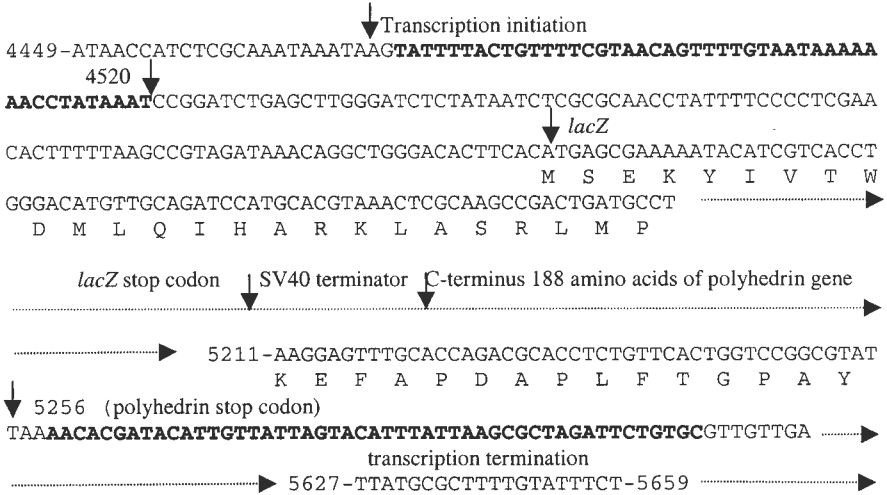
that in our experiments was replaced by the gene for green fluorescent protein (GFP describes the protein and *gfp* describes the gene) (see Figs. 2 and 3). We have chosen this example to demonstrate the principle of primer design. The same principles can be applied to different protein genes and to a broad variety of baculovirus expression vectors with different loci and different promoters.

Fig. 2 represents the schedule for the insertion of *gfp* in place of the *lacZ* of AcRP23-*lacZ* baculovirus (16). β -galactosidase is an enzyme that converts the white chromogenic substrate X-gal to a blue product. Parent virus AcRP23-*lacZ* expressing β -galactosidase, under the virus polyhedrin promoter, forms blue-coloured plaques in *Sf21* insect cells under agarose overlay containing X-gal. The recombinant baculovirus AcRP23-GFP acquires *gfp* and loses *lacZ* during the homologous recombination event and therefore produces white plaques in cell culture.

The GFP belongs to a family of fluorescent proteins from the jellyfish *Aequorea victoria* and it is an important reporter molecule for monitoring gene expression and protein localisation in vivo and *in situ*. The fluorescence excitation and emission spectra of *gfp* are similar to those of fluorescein, and the conditions used to visualise this fluorophore are also suitable for GFP (17).

Fig. 3A depicts the nucleotide and amino acid sequences comprising the polyhedrin locus, for expression of the *lacZ*. The locus commences with the polyhedrin promoter followed by the nucleotide A for transcription initiation. It also contains sequences, encoding the C-terminal 188 amino acids of the polyhedrin gene followed by sequences important for polyhedrin transcription termination. The C-terminal 188 amino acids of the polyhedrin protein gene are redundant and resulted from genetic manipulation during the construction of blue-plaque baculovirus AcRP23-*lacZ* (16). The sequences highlighted in bold at each terminus were selected as RA and are included in the primers for the PCR.

A



B

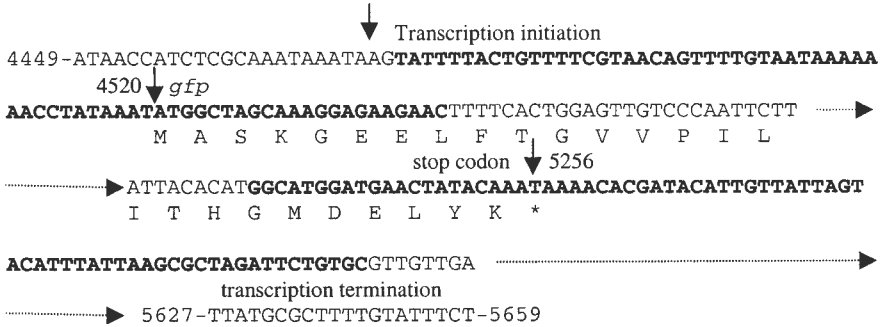


Fig. 3. Sequences of polyhedrin locus with cloned **A** *lacZ* and **B** *gfp*. Numbers correspond to the numeration of baculovirus AcMNPV (Accession number NC001623). Solid arrows indicate the main functional features of the locus—transcription initiation and stop codon of *lacZ*, SV40 terminator of *lacZ* transcription, the redundant sequences of C-terminal 188 amino acids of polyhedrin gene, including stop codon in position 5256 and the transcription termination signal of the authentic polyhedrin gene. Dashed arrowhead lines correspond to sequences not included in this picture. Translation of the beginning of the *lacZ* and *gfp* and the end of the polyhedrin gene is shown. The sequences highlighted in bold were used to design primers for cloning the *gfp* in this locus.

As the result of insertion of the *gfp* into the locus instead of the *lacZ* (**Fig. 3B**), the pretranslation region preceding the first initiation codon is shortened and transcription termination is provided by the transcription termination signal following the stop codon of the polyhedrin gene, not by the SV40 terminator as required for transcription

termination of *lacZ*. The sequences highlighted in bold from each terminus represent the final version of the upstream and downstream primers for the PCR including 50 nucleotides of baculovirus polyhedrin locus sequence followed by either 19 or 21 nucleotides from the 5' or 3' ends, respectively, of the *gfp* (see **Notes 3** and **4**).

Therefore the following principles to design primer sequences with RA must be established.

1. RA for the upstream primers should include the 5' untranslated region and initiation codon ATG. The RA should follow the promoter. The appropriate 5' untranslated regions provide high levels of protein expression. The best way to select the 5' untranslated region is to use the sequence that was already used for the expression of the authentic gene to be substituted. As illustrated earlier, in some cases the 5' untranslated region can be shortened without a significant reduction in protein expression. Each particular case has to be considered carefully.
2. RA for the downstream primer should be placed behind the stop codon and should include it. The sequences of the termination transcription signal must follow the RA. If the transcription termination signal is not known, it is safe to place the RA immediately after the stop codon of the authentic gene.
3. The length of "protein" sequence in the "recombinant" primers depends only on the temperature of primer annealing to template during PCR. For example, if the PCR product is less than 2 kb and nonspecific DNA products are not expected, the temperature of annealing could be in the range 45–60°C with the length of the "protein" part of the primers about 16–20 nucleotides. If the expected PCR product is longer than 2 kb and amplification of nonspecific PCR products is expected, then the calculated annealing temperature of the "protein" part of the primers must be 72–80°C with a length of 24–30 nucleotides (**18–20**).
4. In our experience, 50 nucleotides of RA are sufficient for recombination. Among the 84% of white plaques formed by the recombinant virus about 20% produced green fluorescence. We also tried 30 nucleotides of RA for the upstream primer only and found the same rate of recombination. We did not try shorter RA, but it is quite possible that shorter RA would work albeit with lower recombination efficiency.

2. Materials

2.1. Buffers

1. TE buffer: 10 mM Tris-HCl, pH 8.0, 2 mM ethylene diaminetetraacetic acid (EDTA).
2. 10XTBE buffer for (1 L): 54 g Tris-base, 27.5 g boric acid, 4.65 g EDTA (**21**).

2.2. Chemicals and Equipment for PCR

1. Mixture of deoxynucleotide 5-triphosphates (dNTPs) (Amersham): 10 mM each.
2. Liquid paraffin (Fisons Scientific Equipment, Inc. Griffin & George).
3. pGFP (plasmid containing *gfp*) (CLONTECH Laboratories, Inc.)(see **Note 4**).
4. Thermostable DNA polymerase (Taq) (Sigma-Aldrich) provided with appropriate 10X buffer (see **Note 5**).
5. Agarose (electrophoresis grade) (Gibco-BRL).
6. PCR equipment (thermal cycler) of any design.
7. Tank, tray, and power pack of any design for the electrophoresis of DNA in agarose gel.
8. Gel Extraction Kit (Qiagen).

2.3. Primers for the PCR (see Notes 3 and 4)

1. GFP-BAC-up (baculovirus sequence is emboldened, initiation codon is in small letters):
5'**TATTTTACTGTTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAAT**atg
GCTAGCAAAGGAGAAGAAC3'.
2. GFP-BAC-down (baculovirus sequence is emboldened, stop codon is in small letters):
5'**GCACAGAATCTAGCGCTTAATAAATGTACTAATAACAATGTATCGTGT**Tta
TTGTATAGTTCATCCATG3'.

2.4. Cells

All types of work associated with cells must be carried out aseptically using sterile plasticware and reagents.

Insect cells (*Sf21*) should be grown in suspension using TC100 medium (Gibco-BRL) with 10% foetal calf serum (FCS) (Gibco-BRL) (see **Note 1**) in 500 mL spinner bottles with a magnetic stirrer. For routine culture, cells should be grown to a density of 1×10^6 /mL and then subcultured to another bottles at a dilution rate 1:3–1:4 usually twice per wk.

2.5. Chemicals and Equipment for Baculovirus Transfection and Plaque Assay

1. Purified DNA of baculovirus AcRP23-*lacZ* reconstituted in TE buffer at a concentration of 500 ng/ μ L (see **Note 1**).
2. Restriction endonuclease *Bsu36* I (New England Biolabs Ltd.) provided with appropriate 10X buffer.
3. Fluorescence microscope.
4. Medium TC100 (Gibco-BRL) with 10% FCS (Gibco-BRL).
5. Lipofectin™ (Gibco-BRL).
6. Low melting agarose (Flowgen). Prepare 3% agarose in water using 100 mL glass bottles, sterilize it by autoclaving and keep at room temperature until use.
7. 24-well sterile plastic plates for cell culture (NUNC™).
8. 33-mm sterile plastic dishes for cell culture (CORNING™).
9. 96-well sterile plastic plates for cell culture (NUNC™).
10. Sterile glass Pasteur pipet.
11. Rubber teat.
12. Stain: sterile stock neutral red 3.3 g/mL (SIGMA).
13. Chromogenic substrate for β -galactosidase X-gal (Melford Laboratories Ltd.). Prepare 2% stock using dimethylformamide (BDH) and store at -20°C .

3. Methods

3.1. PCR Amplification of Protein Gene

1. Amplify *gfp* gene in PCR (see **Note 5**):

10X buffer	10 μ L
dNTPs (10 mM)	2 μ L
Primer GFP-BAC-up (10 pM)	5 μ L
Primer GFP-BAC-down (10 pM)	5 μ L

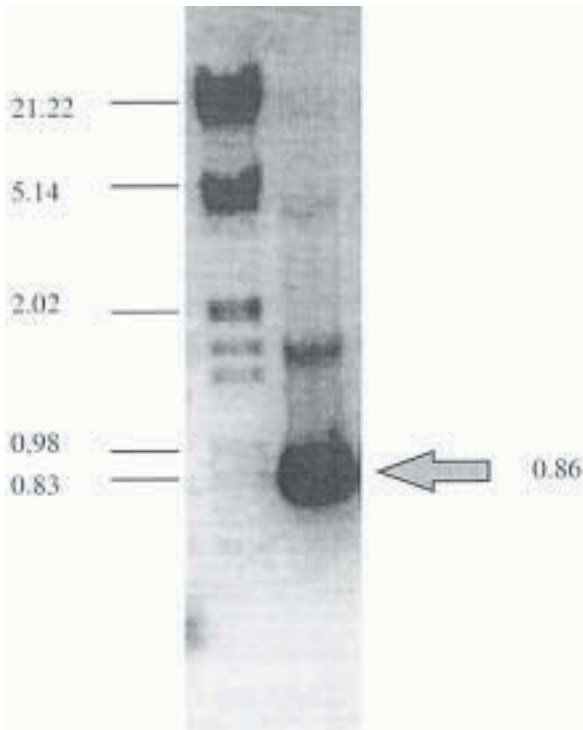


Fig. 4. Analysis of PCR product of *gfp* by electrophoresis in 1% Agarose gel. Molecular weight marker DNA (in kb) is shown on the left. The arrow indicates the positions of *gfp* PCR DNA.

Water	76.5 μ L
pGFP DNA (10 ng/(L))	1 μ L
Taq polymerase (5 U/(L))	0.5 μ L

- Use the following PCR program: 3 cycles at 95°C for 40 s, 52°C for 1 min and 72°C for 1 min, followed by 27 cycles at 95°C for 40 s and 72°C for 1 min (see **Note 6**).
- Test 5 μ L of PCR product in 1% agarose gel using TBE buffer. The derived PCR product in our case consisted of 860 base pairs (see **Fig. 4**). The 50 nucleotides of RA did not interfere with PCR (see **Note 3**). The additional DNA band represents nonspecific priming.
- Repeat PCR (if successful in the first place) in 3–5 tubes and pool the PCR products to obtain 10–50 mg of DNA.
- Separate the amplified DNA from nonspecific DNA by electrophoresis of the pooled PCR product in 1% agarose gel (see **Note 7**).
- Extract DNA from the excised agarose using Qiagen Gel Extraction Kit and adjust concentration to 500 ng/ μ L.

3.2. Transfection Experiments

All following procedures should be carried out aseptically and must be performed in a cabinet that provides HEPA-filtered air.

1. Linearize baculovirus AcRP23-*lacZ* DNA (**12**) (see **Notes 1** and **2**) using a unique restriction site (*Bsu36I* in our case) within the *lacZ*. The linearization of circular DNA significantly reduces the appearance of unwanted parent blue-plaque viruses after transfection. Carry out the following reaction in a volume of 100 μL :

DNA (500 ng/ μL)	10 μL
10X buffer	10 μL
Water	80 μL
<i>Bsu36I</i> (10 U/ μL)	2 μL

After incubation at 37°C for 2 h add the additional 2 μL of *Bsu36I* and incubate the mixture for another 2 h.

2. Seed dishes (35 mm) with $1\text{--}1.5 \times 10^6$ *Sf21* cells and incubate at 28°C for 2 h to allow cells to attach to the plastic.
3. Prepare the following mixture:

Linearized baculovirus DNA (from Step 1) (see Notes 1 and 2)	1 μL
PCR product (Step 6)	1 μL
Water	48 μL
4. Dilute the commercial Lipofectin™ 1:10 with water in a volume of 50 μL and add it dropwise to the mixture of baculovirus DNA and PCR product (**Step 3**). Incubate the mixture for 20 min at room temperature.
5. Wash the cells on 35 mm dishes (**Step 2**) with 1 mL TC100 medium without serum and add 1 mL of TC100 medium without serum.
6. Add the mixture of Lipofectin and DNA from **Step 4** dropwise to the dishes, gently swirling, and incubate for 5–20 h (as convenient) at 28°C.
7. Remove supernatant medium from cells, add 1 mL of TC100 medium with 10% FCS and incubate at 28°C for 3 d.
8. Harvest the supernatant medium (recombinant virus stock) in a sterile container and keep it at 4°C, before analysis for virus by plaque assay.

3.3. Plaque Assay

1. Seed 24 35-mm dishes with $1\text{--}1.5 \times 10^6$ *Sf21* cells in 2 mL of TC100 medium containing 10% FCS and incubate at 28°C for 2 h to allow cells to attach.
2. Make 10X dilutions of supernatant medium from transfection experiments (**Subheading 3.2., step 8**) as follows: place 90 μL of TC100 medium in 6 wells of a 96-well microtitre plate or suitable sterile vials; Add 10 μL of recombinant virus stock (**Subheading 3.2., step 8**) to the first well and then discard it; do not use the same tip to mix up liquid in the first well. Use a second tip to pipet the mixture into the first well, then using the same second tip transfer 10 μL from the first virus dilution to the second well and repeat the procedure until a final virus dilution of 1×10^{-6} is reached.
3. Aspirate the medium from the 35-mm dishes from **Step 1** and quickly apply four replicate aliquots of 0.1 mL of each dilution from supernatant medium (**Step 2**) starting with the 1×10^{-6} dilution. Try to place the diluted virus stock in the center of the dish and make sure that the liquid is spread thoroughly across the cell monolayers. Using the same pipet, continue to apply the next higher concentration of virus to appropriate dishes. To avoid the risk of the monolayer drying out during this process, remove the medium from

four dishes and apply the inoculum, then prepare four more dishes, etc. Incubate for 1 h at room temperature with gentle agitation every 15 min. Make sure that the monolayers are covered completely with the liquid all the time.

4. During this 1-h period of virus adsorption prepare the agarose overlay:
 - a. Melt 3% agarose stock in a microwave oven and cool in a water bath at 42°C.
 - b. Prewarm TC100 medium containing 10% FCS in water bath at 37°C.
5. Aspirate 0.1 mL of virus inoculum from 35 mm dishes from **Subheading 3.2., step 3**.
6. Mix melted 3% agarose with equal volume of prewarmed TC100 medium containing 10% FCS and gently apply 2 mL to the center of the 35 mm dish from **Step 5**. Let the agarose solidify for 15–20 min, add 1 mL TC100 medium containing 10% fetal calf serum (FCS) and place dishes in incubator at 28°C.
7. The plaques will have developed by day 3 or 4. To visualize plaques, gently remove 1 mL of liquid covering the agarose overlay in each dish and replace it with 1 mL of TC100 medium (without serum) containing 15 μ L of 2% X-gal solution and 50 μ L of neutral red solution (3.3 g/mL). Allow cells to stain for 16 h at 28°C. The parent virus will form blue plaques while recombinant virus will form white plaques on the red background.

3.4. Evidence of Protein Expression (see Note 7)

1. To reveal expression of GFP, seed 24-well culture plates with $2-3 \times 10^5$ Sf21 cells in 1 mL of TC100 medium containing 10% FCS and incubate at 28°C for 2 h to allow the cells to attach (the cells should form semiconfluent monolayers).
2. Remove the stain covering the agarose from dishes described in **Subheading 3.3., step 7**, leave the dishes open in the hood to dry, a little, for 30 min.
3. Using a wetted glass Pasteur pipet fitted with a rubber teat, stab the agarose immediately over a single well-defined white plaque. Try to pick-up the piece of agarose and cells that immediately cover a single plaque and then inoculate this directly into the medium above a cell monolayer of one well in a 24-well plate (from **Step 1**). Rinse the pipet 2–3 times with the medium in this well. Using a fresh sterile pipet each time, repeat this process for a large number of white plaques.
4. Incubate the plates at 28°C for 3–4 d.
5. To observe green fluorescence produced by recombinant baculoviruses remove nearly all the medium from 24-well plates (leave a small amount of medium to prevent drying of cells) and site the ultraviolet (UV) objective of the UV-microscope (with a filter for fluorescein) directly over each well. Look for fluorescence in the cells of the monolayer (see **Note 8**). A bright green color in the cells of the monolayer indicates that the recombinant baculovirus is present.

4. Notes

1. Different commercial companies (CLONTECH, Gibco-BRL, NOVAGEN,) offer a wide variety of baculovirus vectors and kits with purified DNA and provide comprehensive guidance. The purified and linearized DNA from kits might also be used as a cloning vector for PCR products of different genes. The choice of baculovirus vector is frequently defined by the screening system for the recombinant viruses and depends on the selective marker/reporter gene inserted in the baculovirus vector.
2. The choice of baculovirus vector and locus for the expression depends on the nature of the intended cloning procedure. The reporter proteins could be cloned in a baculovirus vector even without preliminary linearization. Previously we described a cloning procedure for the β -glucuronidase gene in another locus (P10) of PacBac6 baculovirus without

preliminary linearization of circular DNA (22). The resulting stock of recombinant viruses was heavily contaminated with parent virus forming white plaques in cell culture in the presence of chromogenic substrate X- β -D glc. Nevertheless, a single blue plaque expressing β -glucuronidase was easily visualized.

3. In our experience the precise length of the primers is not critically important for successful PCR. The presence of long 50 nucleotide "tails" representing the RA of baculoviruses does not interfere with amplification of even longer template (up to 10 kb) in PCR (unpublished results).
4. The sequences at either end of *gfp* vary depending on the sources of the information (accession numbers U19276, U36201).
5. The experimental protocols presented here for the PCR are suitable for the amplification of short templates. For the amplification of long PCR templates the conditions may be different (18–20). Many companies (Perkin Elmer, Promega, Advanced biotechnology, Bioline) offer extension PCR amplification kits that supply chemicals and manual guides for the production of long PCR molecules.
6. The programme for long PCR is different from those described above; it includes 72°C for annealing and a longer time for the extension depending on the length of the amplified region (18–20).
7. Confirmation of expression of proteins other than fluorescent proteins can be obtained by more conventional methods, for example, by analysis of infected cell lysates from each well of 24-well plates, by protein electrophoresis in polyacrylamide gels and/or immunoblotting. If the expression of protein is not required and the baculovirus vector is used only for cloning purposes, the evidence of successful cloning can be produced by extracting the DNA from the supernatant medium of 24-well plates as described elsewhere (1,2) and using it as a template for a PCR test.
8. We tested 48 white plaques and found 10 positive for green fluorescence, i.e., about 20% (22). For comparison, the use of a transfer vector containing genes designed for cloning in baculoviruses routinely gives 80–90% recombinant viruses (1,2). Nevertheless, when we tested recombinant virus DNA for the presence of cloned PCR product (by PCR) we found quite a high level of cloning (more than 90%) (unpublished). The reason for the lower efficiency of protein expression is not clear, but it does not detract from the usefulness of the principle. Indeed this method has been used by us when bacterial cloning has proved to be virtually impossible.

References

1. King, L. A. and Possee, R. D. (1992) *The Baculovirus Expression System: A Laboratory Guide*, Chapman and Hall, London, U.K.
2. O'Reilly, D. R., Miller, L. K. and A., L. V. (1994) *Baculovirus Expression Vectors: A Laboratory Manual*, Oxford University Press, New York.
3. Possee, R. D., Thomas, C. J., and King, L. A. (1999) The use of baculovirus vectors for the production of membrane proteins in insect cells. *Biochem. Soc. Trans.* **27**, 928–932.
4. Murhammer, D. W. (1991) Review and patents and literature. The use of insect cell cultures for recombinant protein synthesis: Engineering aspects. *Appl. Biochem. Biotechnol.* **31**, 283–310.
5. Kidd, I. M. and Emery, V. C. (1993) The use of baculoviruses as expression vectors. *Appl. Biochem. Biotechnol.* **42**, 137–159.
6. Kutchan, T. M. (1996) Heterologous expression of alkaloid biosynthetic genes—a review. *Gene* **179**, 73–81.

7. van Oirschot, J. T. (1999) Diva vaccines that reduce virus transmission. *J. Biotechnol.* **73**, 195–205.
8. Kelly, E. P., Greene, J. J., King, A. D., and Innis, B. L. (2000) Purified dengue 2 virus envelope glycoprotein aggregates produced by baculovirus are immunogenic in mice. *Vaccine* **18**, 2549–2559.
9. Bonning, B. C. and Hammock, B. D. (1996) Development of recombinant baculoviruses for insect control. *Annu. Rev. Entomol.* **41**, 191–210.
10. Belyaev, A. S., Hails, R. S., and Roy, P. (1995) High-level expression of five foreign genes by a single recombinant baculovirus. *Gene* **156**, 229–233.
11. Kitts, P. A. and Possee, R. D. (1993) A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* **14**, 810–817.
12. Kitts, P. A., Ayres, M. D., and Possee, R. D. (1990) Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucl. Acids Res.* **18**, 5667–5672.
13. Ernst, W. J., Grabherr, R. M., and Katinger, H. W. (1994) Direct cloning into the *Autographa californica* nuclear polyhedrosis virus for generation of recombinant baculoviruses. *Nucl. Acids Res.* **22**, 2855–2856.
14. Lu, A. and Miller, L. K. (1996) Generation of recombinant baculoviruses by direct cloning. *Biotechniques* **21**, 63–68.
15. Khromykh, A. A., Meka, H., and Westaway, E. G. (1995) Preparation of recombinant baculovirus by transfection of a ligated cDNA fragment without prior plasmid amplification in *E. coli*. *Biotechniques* **19**, 356–360.
16. Mann, S. G. and King, L. A. (1989) Efficient transfection of insect cells with baculovirus DNA using electroporation. *J. Gen. Virol.* **70**, 3501–3505.
17. Kain, S. R., Adams, M., Kondepudi, A., Yang, T. T., Ward, W. W., and Kitts, P. (1995) Green fluorescent protein as a reporter of gene expression and protein localization. *Biotechniques* **19**, 650–655.
18. Gritsun, T. S. and Gould, E. A. (1995) Infectious transcripts of tick-borne encephalitis virus, generated in days by RT-PCR. *Virology* **214**, 611–618.
19. Gritsun, T. S. and Gould, E. A. (1998) Development and analysis of a tick-borne encephalitis virus infectious clone using a novel and rapid strategy. *J. Virol. Meth.* **76**, 109–120.
20. Cheng, S. and Kolmodin, L. (1997) XL PCR amplification of long targets from genomic DNA, in *PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering*, Vol. 67 (White, B. A., ed.), Humana, Totowa, pp. 17–29.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, Vol. 3 (Nolan, C., ed.), Cold Spring Harbor Laboratory Press, New York, pp. B23.
22. Gritsun, T. S., Mikhailov, M. V., Roy, P., and Gould, E. A. (1997) A new, rapid and simple procedure for direct cloning of PCR products into baculoviruses. *Nucl. Acids Res.* **25**, 1864–1865.

III

MUTAGENESIS AND RECOMBINATION

PCR Approaches to DNA Mutagenesis and Recombination

An Overview

Binzhang Shen

1. Introduction

Current polymerase chain reaction (PCR) innovations provide powerful tools for the cloning of previously unknown genes as well as characterization of their functions. Examples of the latter used include PCR-mediated *in vitro* mutagenesis and recombination of the cloned genes. PCR approaches have become the method of choice to generate arrays of predefined mutations or recombination within the gene of interest. In the postgenomic era, these mutants or recombinants are highly desirable for the study of functional genomics, gene expression, protein structure-function relationships, protein-protein interactions, protein engineering, and *in vitro* evolution of enzymes.

There are numerous PCR-based approaches to DNA mutagenesis and recombination; therefore, a complete coverage in this review of methods and new developments for such a rapidly advancing field is impractical. Instead, a selected subset of basic and representative approaches will be presented. These approaches are the prototypes upon which many of the other designs in the current literature are based. Interested readers should also refer to a previous review (1).

PCR-mediated nucleotide changes, deletions, or insertions can be accomplished by performing PCR synthesis reactions with carefully chosen or modified PCR reaction components. One strategy for this kind of PCR mutagenesis is based on the principle of “mispriming” (2). Because mismatches between templates and primers are tolerated under certain PCR conditions, primers can be designed to include predefined changes (so-called mutagenic primers). Other strategies make use of either the built-in high error rate of *Taq* DNA polymerase or degenerate base analogs such as deoxyinosine triphosphate (dI) in the PCR reaction. In general, mutagenic primer PCR approaches are used for introducing site-directed mutagenesis (SDM) into the genes of interest, whereas *Taq* DNA polymerase or base analog PCR approaches are useful in creating random and extensive mutagenesis (REM) in the target gene.

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

A novel strategy to make recombinant molecules is to use PCR-mediated *in vitro* recombination. It involves formation of a heteroduplex intermediate between the two fragments to be recombined. There are two basic approaches to this PCR mediated *in vitro* recombination. In the first approach, PCR is employed to generate two overlapping primer-tailed genes capable of forming heteroduplex intermediates. Extension of the heteroduplex generates the construct of interest. In the second approach, new heteroduplex intermediates are formed and extended following each PCR cycle until the chimeric constructs become full-length. Approaches have also been developed which utilize the *in vivo* homologous recombination strategy to create recombinants from transformed PCR products.

In practice, which PCR strategy or approach to choose depends on the researcher's objective (SDM, REM or recombination), the template used (linear or circular), and the efficiency desired. In some instances, modification of the basic approaches is necessary, whereas in others, several approaches could be combined to achieve maximum performance. A comprehensive comparison of the basic PCR approaches discussed in the present review is summarized in **Table 1** and **Fig. 1**.

2. Mutagenic Primer

A perfect base match between templates and primers is favored under most PCR conditions. Mutagenic primers or mismatched primers, albeit less efficiently, can direct DNA polymerization under conditions where

1. A perfect match is not present;
2. Lower annealing temperature is applied;
3. A mismatched nucleotide is located at the 5' end or in the middle of a long mutagenic primer.

A mutagenic primer can be designed to include nucleotide substitutions, small deletions or insertions into its nucleotide sequence. As primers are incorporated into the final PCR products, so are these changes. In general, mutagenic primers are used either to create mutations in the target gene or to facilitate ^ocloning of the PCR products.

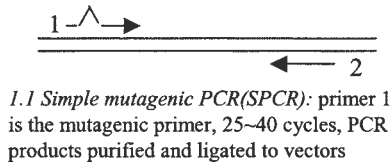
3. Location of Mutations

The mutagenic primer also defines the location of mutations. After a simple mutagenic PCR (SPCR) (**3**, **Table 1–1.1** and **Fig. 1–1.1**) using a mutagenic primer, mutations are introduced into the PCR products. Mutations generated in this manner are usually restricted to the termini of the final PCR products.

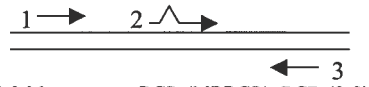
To overcome this limitation of SPCR, Ho et al. (**4**) devised overlap extension PCR (OEPCR). In OEPCR (**3**, **Table 1–1.2** and **Fig. 1–1.2**), two overlapping primers and two flanking primers are used. Two separate SPCR amplifications are done, with each using one overlapping primer and one flanking primer. The two PCR products are then annealed and extended. The resulting full-length DNA is either used directly or further PCR amplified with the two flanking primers. The use and design of two overlapping primers is key to OEPCR. These primers, when mutagenic, allow the introduction of mutations throughout the whole construct. One drawback of OEPCR is its low efficiency, which is primarily because of short sequence homology between the overlapping primers.

Table 1
Major Advantages and Limitation of Different PCR Approaches for DNA Mutagenesis and Recombination

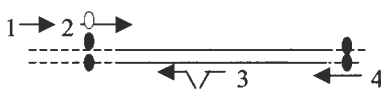
Types of PCR	Advantages	Limitations
<i>1. Basic mutagenic PCR</i>		
1.1 Simple mutagenic PCR	simple and useful	mutations at the end of the PCR product, 50% mutants
1.2 Overlap extension PCR	mutations anywhere in the PCR product,	Low efficiency
1.3 Megaprimer PCR	highly (100%) efficient, mutations anywhere desired	non-specific, non-template addition of bases at 3' end
1.4 Inverse PCR	for previously unknown genes	low efficiency
<i>2. Restriction site PCR</i>		
2.1 Modified restriction site PCR	highly efficient	for previously cloned gene
2.2 New restriction site PCR	highly efficient	mutation sites restricted
3.1 Marker couples PCR	simple process	low efficiency
<i>4. PCR-mediated recombination</i>		
4.1 Gene SOEing	highly efficient (100%)	costly primers
4.2 DNA shuffling and Staggered Extension Process (STep)	random and extensive recombination	highly specific, efficient selection systems needed
4.3 Recombinant circular PCR	good overall efficiency	more PCR reactions
4.4 Recombinant PCR	no post-PCR processing	overall efficiency 50%
<i>5. Nonmutagenic primer PCR</i>		
5.1 Erroneous PCR	Random and extensive mutations	sequence-dependent error rate, biased mutations
5.2 Degenerate base analogs mediated PCR	random and extensive mutations	highly specific, efficient selection systems needed,
5.3 Stepwise elongation of sequence PCR	Extensive and specific mutations	multiple primers and PCR



1.1 Simple mutagenic PCR (SPCR): primer 1 is the mutagenic primer, 25~40 cycles, PCR products purified and ligated to vectors



1.3 Megaprimer PCR (MPPCR): PCR (2-3) product purified, used as a megaprimer, second PCR (1-megaprimer)



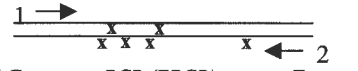
2.1 Modified restriction site PCR (MRPCR): polylinkers with restriction site at template ends, two PCR (1-3 and 2-4) combined, PCR (1-4), digested and ligated to vectors



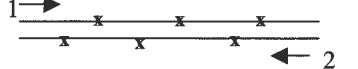
3.1 Marker coupled PCR (MCPCR): marker gene activation primer 1 and target gene mutagenic primer 2, PCR (1-2), as a megaprimer, extended and transformed



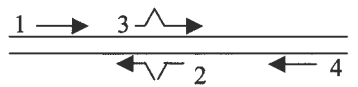
4.2 DNA shuffling and Staggered extension Process (StEP): various alleles mixed, fragmented by DNase I, random and briefly primed by repeated PCR cycles



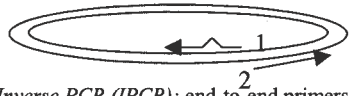
5.1 Erroneous PCR (EPCR): excess Taq DNA polymerase, high Mg⁺⁺, Mn⁺⁺, biased dNTP ratio



5.2 Degenerate base analogs mediated PCR (DBAPCR): dITP in PCR reaction, dI integrated products as templates for further PCR



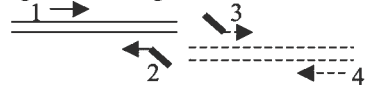
1.2 Overlap extension PCR (OEPCR): two PCR (primers 1-2 and 3-4), combined, melted, annealed, extended and final PCR (primers 1-4)



1.4 Inverse PCR (IPCR): end-to-end primers, PCR products re-circulated and transformed



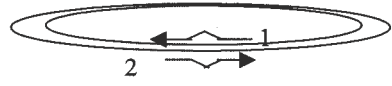
2.2 New restriction site PCR (NRPCR): new restriction site and mutation on internal primer 2 and 3, two PCR (1-2 and 3-4), digested and re-ligated



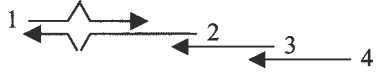
4.1 Gene SOEing: linkers on primer 2 and 3, two PCR (1-2 and 3-4), overlap extension of 1-2 and 3-4, finally PCR (1-4)



4.3 Recombinant circular PCR (RCPCR): overlapping mutagenic primer 1 and 4, two PCR (1-2 and 3-4), combined, annealed and transformed



4.4 Recombinant PCR (RPCR): overlapping mutagenic primers 1 and 2, PCR (1-2) combined, annealed, transformed, *in vivo* recombination



5.3 Stepwise elongation of sequence PCR (SESPCR): overlapping synthetic primers spans the known gene sequence, serial PCR

Fig. 1. Design of different PCR approaches to DNA mutagenesis and recombination. Note: ○ inactive restriction site, ● restriction site, □ inactive form, ■ active form, mutagenic primer.

Megaprimer PCR (MPCR) (5) has been developed to improve the efficiency of OEPCR. Instead of using a pair of overlapping primers, MPCR (3, **Table 1–1.3** and **Fig. 1–1.3**) starts with a SPCR using an internal mutagenic primer and a flanking primer. PCR products generated in this manner are purified and used as a primer (hence “megaprimer”) for the second round of PCR along with another flanking primer. Wild-type sequences are used as the template in both PCR reactions. Compared to OEPCR, MPCR uses fewer primers and a longer homology match between megaprimer and template, and this leads to a higher efficiency. A major limitation of MPCR is the nonspecific, nontemplate addition of nucleotides at 3' end of the megaprimer.

Both OEPCR and MPCR approaches discussed above have an advantage over SPCR for generating mutations throughout the whole construct, but they also have limitations. To overcome the SPCR limitation, a circular template can be used. The gene of interest can first be cloned into a vector and amplified using two end-to-end primers by an inverted PCR (IPCR) approach (6, **Table 1–1.4** and **Fig. 1–1.4**). The IPCR products are recirculated and amplified in *Escherichia coli* (*E.coli*). Depending on the location of the mutagenic primer-matched sequence within the template, mutations can also be introduced throughout the whole gene.

4. Cloning Efficiency

The low efficiency of ligation in the post-PCR cloning processing is a common problem experienced by most users of the above-mentioned approaches. Depending on whether a linear template or a circular template is available, either in vitro or in vivo recombination strategy could be used to improve post-PCR cloning efficiency. When linear templates have to be used, in vitro ligations of the PCR products are often necessary. One of the PCR based approaches for increasing ligation efficiency is to introduce new restriction sites into the PCR products through the use of mutagenic primers. In the modified restriction site PCR (MRPCR) approach (7, **Table 1–2.1** and **Fig. 1–2.1**), polylinkers with a restriction site are ligated to the template and are then PCR amplified using different primers. One of the primers pairs with the polylinker, but bears an inactive restriction site. Another internal primer serves to introduce mutation into the construct at the expected site. Using this approach, new restriction sites are introduced into the ends of the final PCR products whereas mutations can be targeted anywhere in the construct. Likewise, new restriction sites can also be created at any internal position of the gene using internal mutagenic primers by the new restriction site PCR (NRPCR) approach (8, **Table 1–2.2** and **Fig. 1–2.2**). These measures could improve post-PCR ligation, and therefore increase overall cloning efficiency. Because a restriction enzyme digestion is included, restriction sites on the primers are carefully chosen to avoid internal cleavage of the target sequence by the same restriction enzymes.

In the case of circular templates, in vivo homologous recombination can be employed. Two different approaches have been developed for use with circular templates. The so-called recombinant circular PCR (RCPCR)(9, **Table 1–4.3** and **Fig. 1–4.3**) approach represents the combined use of OEPCR and IPCR. The second approach, recombinant PCR (RPCR) (9, **Table 1–4.4** and **Fig. 1–4.4**), is a modified IPCR where

two overlapping primers are used, instead of end-to-end primers. The use of *in vivo* homologous recombination eliminates post-PCR processing and *in vitro* ligation, and therefore further improves the overall cloning efficiency.

5. Selection for Desired Products

Yet another advantage of the aforementioned MRPCR approach is that a restriction enzyme digestion adds a selection tool for the mutagenized products in the post-PCR process. By using carefully designed mutagenic primers, the only PCR products that can be digested and ligated into a vector are those amplified from primers bearing an activated restriction site. This added measure of selection helps eliminate the nonmutant background.

Marker coupled PCR (MCPCR) approach (*10*, **Table 1–3.1** and **Fig. 1–3.1**) is an example in which selection strategy is used for circular templates. An antibiotic resistance marker gene is targeted for selective activation in MCPCR instead of using restriction sites. For this approach to work, the template plasmid must have a point and null mutation in one of its otherwise functional antibiotic resistance gene. One of the PCR primers must be designed so that when incorporated into the new PCR product it restores the function of that marker gene. Following MCPCR, DNA products are transformed into *E. coli*. Transformants carrying desired mutagenized plasmids are then selected on media with the appropriate antibiotics.

6. Random and Extensive Mutagenesis

All of the approaches discussed so far are designed to create either single or a few site-directed mutations. Sometimes, it is desirable to obtain random and extensive mutations (REM) in the gene of interest or to generate a library of such molecules. There are two basic approaches to achieve these goals, and both are based on selected uses of PCR reaction components. The first approach (*11*) to generate REM makes use of error-prone DNA polymerases (**Table 1–5.1** and **Fig. 1–5.1**). Certain thermostable DNA polymerases, like *Taq* DNA polymerase, have an intrinsic error rate due to the lack of a 3'-5' exonuclease activity. Each pass of the polymerase during PCR allows the possibility of mutations; therefore, the cumulative error rate can be substantial. This tendency is further enhanced by other factors such as buffer composition (e.g., high-magnesium concentration, high pH, or addition of 0.5 mM MnCl₂) and other experimental conditions (e.g., a large amount of polymerase, a great number of cycles, a low-annealing temperature, a biased pool of the four dNTPs).

The second approach (**Table 1–5.2** and **Fig. 1–5.2**) to generate REM in the genes of interest is based on the base pairing property of degenerate base analog (*12*). For example, base analog dI can form base pairing with nucleotides A, C, G, and T under normal reaction conditions. In the presence of dI and a biased ratio of dNTPs, DNA polymerase tends to randomly incorporate a substantial amount of dI in the newly synthesized DNA strand. This dI-containing DNA can serve as a template in subsequent PCR amplifications and allows random base insertion at the dI-inserted sites. As a result, the final PCR products will have base substitutions at multiple and random sites.

Mutations at multiple but predetermined sites of a protein are sometimes desired. It is not difficult to see that the aforementioned approaches, when used singly, are not suitable for every need. Alternatively, stepwise elongation of sequence PCR (SESPCR) (13, Table 1–5.3 and Fig. 1–5.3) can be employed. In SESPCR, multiple mutagenic primers are utilized in a serial PCR to introduce mutations into the synthesizing gene. SESPCR is therefore practical only for relatively small genes.

7. Creation of Novel Genes

Our discussion has so far been focused on different PCR approaches for introducing mutations (single or multiple and site-specific or random) into cloned genes. In fact, some of these basic approaches have also been modified to create novel genes or chimeric constructs from pre-existing mutations or genes through a process known as *in vitro* recombination. One of the approaches is Gene-Splicing Overlap Extension (Gene SOEing) (14, Table 1–4.1 and Fig. 1–4.1), which is an OEPCR-based approach. In Gene SOEing, two pairs of primers are used. One primer from each pair is actually a bipartite or modular primer with an oligonucleotide linker attached at the 5' end of the primer. The linkers from the two primer pairs are designed to overlap. When these two modular primers are from the same gene, Gene SOEing becomes OEPCR. Otherwise, if the two modular primers are based on DNA sequences from different genes, Gene SOEing produces a recombinant molecule. Such modular primer design allows virtually any two DNA sequences to be recombined. Consequently, Gene SOEing PCR is highly useful for *in vitro* gene fusion and protein engineering when appropriate restriction sites are not available at the expected sites of recombination.

The second approach to *in vitro* DNA recombination is DNA shuffling or “sexual PCR” (15, Table 1–4.2 and Fig. 1–4.2). It is a modified megaprimer PCR. Basically, alleles of a target gene are mixed and fragmented by DNase I to different sizes (averaging 50–100 bp). These fragments serve as both primers and templates for PCR amplifications. Another slightly different modification is termed staggered extension primer (StEP) PCR (16, Table 1–4.2 and Fig. 1–4.2), which uses short random oligonucleotides to prime the synthesis of the new strand along the fragmented template.

It is interesting to note that recombinant molecules generated by these two PCR approaches of *in vitro* recombination are dramatically different. In Gene SOEing, PCR products are not recombinants *per se*. When combined, melted, reannealed, PCR products from different genes can form a heteroduplex intermediate as mediated by the overlapping linkers. Subsequent extension of the heteroduplex leads to the formation of recombinant molecules. Because linker primers define the recombination site, Gene SOEing-mediated DNA recombination is site-specific. A major application for this approach is in the study of gene fusion and protein domain swapping. In DNA shuffling and StEP, however, PCR products are already recombinant molecules. Recombination happens during each PCR cycle when templates and primers switch following each denaturalization and reannealing process. As a result, the site of recombination is randomized and the recombination frequency is a factor of template length and extension duration of each PCR cycle. PCR products generated in this manner represent a heterogenic mix of highly chimeric constructs. The resulting library of recombinant molecules is ideal for the study of *in vitro* evolution of enzymes.

We have briefly discussed some basic PCR approaches to DNA mutagenesis and recombination. Without going into much details of each approach, attempts have been made to present different approaches in a historical and logical prospective. In the postgenomic era, it is anticipated that PCR-mediated DNA mutagenesis and recombination will become a common tool of molecular biologists, and hence, more innovative PCR approaches will certainly emerge over time. It is hoped that this review offers some guidelines for the reader to choose a PCR approach and to create additional innovations.

References

1. Ling, M. M. and Robinson, B. H. (1997) Approaches to DNA mutagenesis: An overview. *Analyt. Biochem.* **254**, 157–178
2. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 263–273
3. Kammann, M., Laufs, J., Schel, J. and Gronenborn, B. (1989) Rapid insertional mutagenesis of DNA by PCR. *Nucl. Acids Res.* **17**, 5404.
4. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the PCR. *Gene* **77**, 51.
5. Sarkar, G. and Sommer, S. S. (1990) The ‘megaprimer’ method of site-directed mutagenesis. *BioTechniques* **8**, 404.
6. Ochman, H., Gerber, A. S. and Hartl, D. L. (1988) Genetic applications of inverse PCR. *Genetics* **120**, 621.
7. Ito, W., Ishinuro, H. and Kurosawa, Y. (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene* **102**, 67
8. Dulau, L., Cheyrou, A. and Aigle, M. (1989) Directed mutagenesis using PCR. *Nucl. Acids Res.* **17**, 2873.
9. Jones, D. H. and Winistorfer, S. C. (1991) Site-specific mutagenesis and DNA recombination by using PCR to generate recombinant circles in vitro or by recombination of linear PCR products in vivo. *Methods: Compan. Meth. Enzymol.* **2**, 2.
10. Shen, T. J., Zhu, L. Q. and Sun, X. (1991) A marker-coupled method for site-directed mutagenesis. *Gene* **103**, 73–77.
11. Cadwell, C. and Joyce, G. F. (1992) Randomization of genes by PCR mutagenesis. *PCR Meth. Appl.* **2**, 28–33.
12. Spee, J. H., de Vos W. M. and Kuipers, P. (1993) Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucl. Acids Res.* **21**, 777–778.
13. Majumder, K. (1992) Ligation-free gene synthesis by PCR: synthesis and mutagenesis at multiple loci of a chimeric gene encoding OmpA signal peptide and hirudin. *Gene* **110**, 89–94.
14. Horton, R. M. (1997) In vitro recombination and mutagenesis of DNA: SOEing together tailor-made genes, in *Methods in Molecular Biology*, vol. 67, *PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering* (White, B. A., ed.), Humana, Totowa, NJ, pp. 141–150.
15. Stemmer, W. P. C. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391
16. Zhao, H., Giver, L., Shao, Z., Affholter, J. A. and Arnold, F. H. (1998) Molecular evolutions by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258–261.

In-Frame Cloning of Synthetic Genes Using PCR Inserts

James C. Pierce

1. Introduction

Because many genes of biological interest are larger than the maximum size that current synthetic oligonucleotide synthesizers can produce (approx 110 bases), there is a need for methods that allow rapid production and expression of genes constructed from multiple synthetic DNA fragments. A number of synthetic genes have been generated using the recursive polymerase chain reaction (PCR) method, but problems concerning primer design and incorrect final gene sequence, especially with large synthetic genes, are a concern (1–3). The cloning method described here follows a series of steps in which multiple PCR products or synthetic duplex oligonucleotides are positionally cloned into a plasmid vector (4). A synthetic gene of practically any sequence or length can be built using the in-frame cloning method. Genes are assembled such that open reading frames are maintained by linking DNA fragments through the use of six basepair blunt-end restriction sites. Each cloning step uses an anchored sticky-end restriction site and a variable blunt-end restriction site that result in specific insert orientation and high cloning efficiencies. The overall strategy of in-frame cloning allows the researcher total control over nucleotide sequence, codon usage, promoter and other regulatory elements, and the placement of unique restriction sites throughout the recombinant construct. One advantage of the in-frame cloning method described here is that it allows for flexible yet precise construction of synthetic genes using standard recombinant techniques. Another advantage is that it employs inexpensive, readily available materials.

The in-frame cloning method is based on the observation that standard plasmid cloning vectors such as pUC or pGEM contain very few six basepair blunt-end restriction recognition sites (5). The amino acid sequence of a protein, whose gene is to be cloned, is scanned for those amino acids encoded by the blunt-end restriction sites listed in **Table 1**. Each of these “signpost” amino acids can then be used to fragment a protein sequence into sections that are easily encoded by synthetic oligonucleotides. When two contiguous DNA fragments are joined in the plasmid vector by blunt-end

Table 1
Blunt-End Restriction Sites
that Can be Used for In-Frame Cloning

Enzyme	Recognition Site	Amino Acids
<i>AfeI</i>	AGC-GCT	Ser-Ala
<i>BsaAI</i>	CAC-GTG	His-Val
<i>BsaAI</i>	TAC-GTA	Tyr-Val
<i>BstZ17I</i>	GTA-TAC	Val-Tyr
<i>BtrI</i>	CAC-GTC	His-Val
<i>EcoRV</i>	GAT-ATC	Asp-Ile
<i>FspI</i>	TGC-GCA	Cys-Ala
<i>HincII</i>	GTC-GAC	Val-Asp
<i>HincII</i>	GTT-AAC	Val-Asn
<i>HpaI</i>	GTT-AAC	Val-Asn
<i>MscI</i>	TGG-CCA	Trp-Pro
<i>NaeI</i>	GCC-GGC	Ala-Gly
<i>NruI</i>	TCG-CGA	Ser-Arg
<i>PmlI</i>	CAC-GTC	His-Val
<i>PvuII</i>	CAG-CTG	Gln-Leu
<i>Scal^a</i>	<u>AGT</u> -ACT	Ser-Thr
<i>SfoI^a</i>	GGC-GCC	Gly-Ala
<i>SmaI</i>	<u>CCC</u> -GGG	Pro-Gly
<i>SnaBI</i>	TAC-GTA	Tyr-Val
<i>SspI^a</i>	AAT-ATT	Asn-Ile
<i>StuI</i>	<u>AGG</u> -CCT	Arg-Pro

^aThese restriction enzyme recognition sites are present at least once outside of the multiple cloning regions of pUC and pGEM vectors. Underlined codons can be problematic in *E. coli* expression systems.

ligation the open reading frame of the synthetic gene is maintained. As described later, a combination of relatively simple cloning techniques and electroporation give high overall cloning efficiencies. The PCR is used for both synthesis of duplex DNA from oligonucleotides and for the rapid screening of intermediate and final synthetic plasmid constructs by direct amplification of plasmid DNA from transformed host bacteria (*Escherichia coli*) colonies. By monitoring the final plasmid-synthetic gene construct for the presence or absence of diagnostic restriction sites, one can have good confidence that the correct synthetic gene has indeed been cloned before confirmation by DNA sequencing.

The in-frame cloning method was initially developed to make synthetic genes from complementary synthetic oligonucleotides that were annealed to generate a duplex DNA molecule and then directly cloned into the plasmid vector (4). One advantage of using synthetic oligonucleotides as DNA fragments for direct cloning is the ability to use them without further purification. Also, because oligonucleotides do not contain a terminal phosphate group, no problems are encountered with multiple tandem copies

of the DNA insert following ligation reactions in which the insert is in significant molar excess relative to the plasmid scaffold. The protocol described here is based on the work of a number of researchers in which duplex insert DNA is first made by the PCR using overlapping oligonucleotide primers (1,6,7,8). This modification allows for larger blocks of insert DNA to be made per cloning event and will decrease the overall cost of oligonucleotide primers. Two problems associated with PCR-derived templates are poor DNA sequence fidelity and difficulty in obtaining flush blunt-ends. With the availability of thermostable DNA polymerase with high replication fidelity and 3'-5' exonuclease activity that removes terminal sequences, many of the problems associated with poor-quality PCR derived templates can be overcome.

2. Materials

1. Plasmid vector DNA (e.g., pUC or pGEM).
2. Synthetic oligonucleotides designed using the format in **Fig. 1**. Resuspend deprotected oligonucleotides in sterile water or Tris-ethylenediaminetetraacetic acid (EDTA) buffer, pH 8.0 at a final concentration of about 2 mg/mL. There is no need to purify oligonucleotides as long as the synthesis was relatively efficient. The quality of the oligonucleotides can be checked by simple agarose or acrylamide gel electrophoresis and staining with ethidium bromide.
3. *E. coli* host strain for plasmid transformation (e.g., strain DH10B, Life Technologies, Rockville, MD).
4. Restriction enzymes and buffers.
5. Reagents for agarose gel electrophoresis including low-melting point agarose for in-gel cloning (e.g., GTG SeaPlaque agarose, FMC BioProducts, Rockland, ME).
6. DNA ligase and buffer.
7. Microdialysis membrane such as Millipore VSP 0.025-mm filters (Bedford, MA).
8. A high efficiency method to transform *E. coli* cells with plasmid DNA (e.g., Bio-Rad Gene Pulser electroporator, Hercules, CA).
9. Media, Petri plates, and ampicillin for growth and selection of transformed bacteria.
10. PCR primers that border the cloning site of the plasmid vector (e.g., Sp6 and T7 primers are used for pGEM vectors).
11. Reagents and thermocycler for PCR.
12. Reagents needed for alkaline lysis, miniplasmid purification procedure (9).
13. DNA sequencing reagents and equipment for final confirmation of synthetic gene sequence.
14. TE buffer: 10 mM Tris-HCl, pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA).
15. TBE buffer: 90 mM Tris, 90 mM borate, pH 8.3, 2 mM EDTA.
16. TE-saturated phenol.
17. Chloroform:isoamyl alcohol (24:1).
18. 95% ethanol.
19. 70% ethanol.

3. Methods

3.1. Organizational Strategy for Oligonucleotide Design

1. The in-frame cloning method assumes that the amino acid sequence of the protein is known for the synthetic gene under construction. Plasmid vectors such as pUC and pGEM

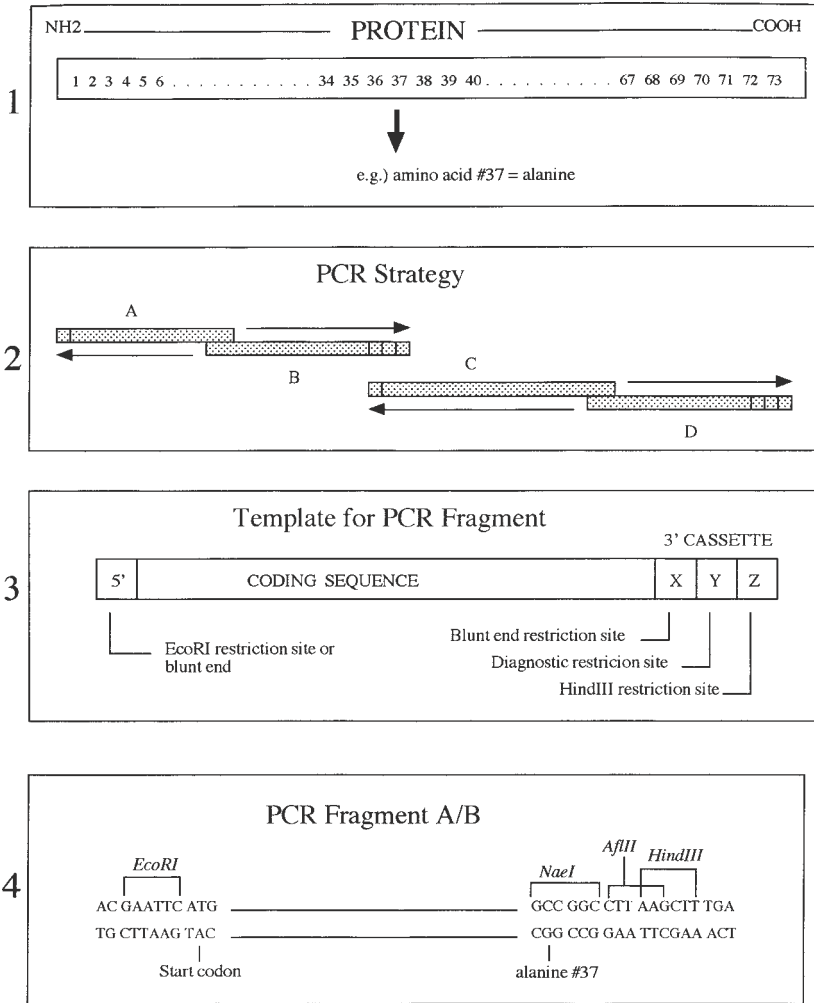


Fig. 1. Design strategy for the in-frame cloning method using PCR-generated DNA inserts. The design strategy is broken into four parts. Part 1 illustrates the analysis portion of the in-frame cloning procedure. The primary amino acid sequence is scanned for amino acids that are present in **Table 1**. The protein sequence is divided into segments that are encoded by overlapping synthetic oligonucleotides (Part 2). In this example, amino acid #37 is an alanine, which defines the end point between segment 1 (encoded by oligos A/B) and segment 2 (encoded by oligo C/D). The PCR is used to extend the 3' end of each oligonucleotide. Part 3 demonstrates the template from which to design the synthetic oligonucleotides. Each PCR-generated fragment contains three main components: the 5' end, the coding sequence and the 3' cassette. The 3' cassette is further divided into three sections. Section X represents the blunt-end restriction site picked from **Table 1**, section Y represents the diagnostic restriction site for segment identification, and section Z represents the 3' cloning site, usually a *HindIII* sticky-end. Part 4 shows the partial sequence of PCR fragment A/B.

are used as the scaffold to build the synthetic gene. These plasmids have relatively few six base-pair blunt-end restriction recognition sites (listed in **Table 1**). Seventeen of the 20 naturally occurring amino acids are available using the restriction sites listed in **Table 1**. Note that some of the codons illustrated in **Table 1** have been identified as being problematic when used in *E. coli*-based recombinant expression experiments (10). The in-frame cloning method uses this knowledge to design a cloning strategy that allows the forced orientation cloning of insert DNA such that the open reading frame is maintained by the assembly of DNA fragments at one endpoint by blunt-end ligation and at the other endpoint by sticky-end ligation (for overview, see **Figs. 1** and **2**).

2. A number of symmetrical six base-pair blunt-end restriction sites, listed in **Table 1**, can be used for in-frame cloning. The protein is broken into a number of fragments that contain one of the amino acids listed in **Table 1**. These selected amino acids define the last amino acid of a protein segment in each of the cloning steps. It is possible to build a gene starting from either the amino terminus or the carboxyl terminus. This gives more flexibility in the choice of amino acids that are encoded by the blunt-end restriction sites. Once the protein sequence has been segmented into specific insert blocks (see **Fig. 1**, part 1 and 2), a PCR strategy is utilized which will result in DNA fragments that contain all the elements needed for in-frame cloning.
3. **Fig. 1** illustrates the design strategy for generating PCR fragments that are used in the in-frame cloning method. The protein sequence is segmented into blocks that are easily covered by moderate sized oligonucleotides of about 50–90 bases. Part 1 of **Fig. 1** shows a hypothetical 73 amino acid protein that will be broken into two parts between amino acids #37 and #38. Note that amino acid #37 is alanine, which is encoded by one or more of the blunt-end restriction enzymes listed in **Table 1**. The oligonucleotides are designed such that a 20 base-pair region of complementary DNA sequence overlap will result when the two cognate oligonucleotides are annealed together (e.g., oligos A and B in part 2 of **Fig. 1**). As an example, if two 80 base oligos were annealed together with a 20 base overlap, a protein coding sequence of up to 46 amino acids could be generated. Once annealed, duplex DNA is made by extension of the 3' end of each oligo using PCR for a limited number of cycles (see **Note 1**).
4. The PCR fragment must contain a number of elements as illustrated in section 3 of **Fig. 1**. The 5' end contains either an *EcoRI* restriction site or a blunt-end. The *EcoRI* site is used during the first round of cloning. Subsequent rounds of cloning use a 5' blunt-end, which, upon ligation to the 3' blunt-end generated during the cloning process, will result in the correct translational reading frame. The 5' blunt end can be made by either of two (different) methods. If a DNA polymerase is used that generates flush blunt ends during the amplification process, then oligonucleotides can be designed such that the 5' end of the fragment (first three nucleotides) will contain the next codon in the amino acid sequence. If there is uncertainty concerning the sequence (e.g., terminal synthetase activity) at the 5' (left) end, then it would be necessary to generate an oligonucleotide that contains an internal blunt-end recognition site which upon restriction digestion of the PCR fragment with the appropriate enzyme will leave a suitable blunt-end. This second scenario is somewhat restrictive in that the amino acid must be encoded by a symmetrical six-base blunt-end restriction enzyme (not necessarily from **Table 1**) and makes the design of oligonucleotides more difficult. The internal part of the oligonucleotide contains the nucleotide sequence information that encodes the amino acids of the protein. The 3' end contains a number of restriction sites and is termed the 3' cassette. As shown in **Fig. 1**, part 3, the 3' cassette contains three sites labeled X, Y, and Z. Site X contains the blunt-end restriction site that will align the amino

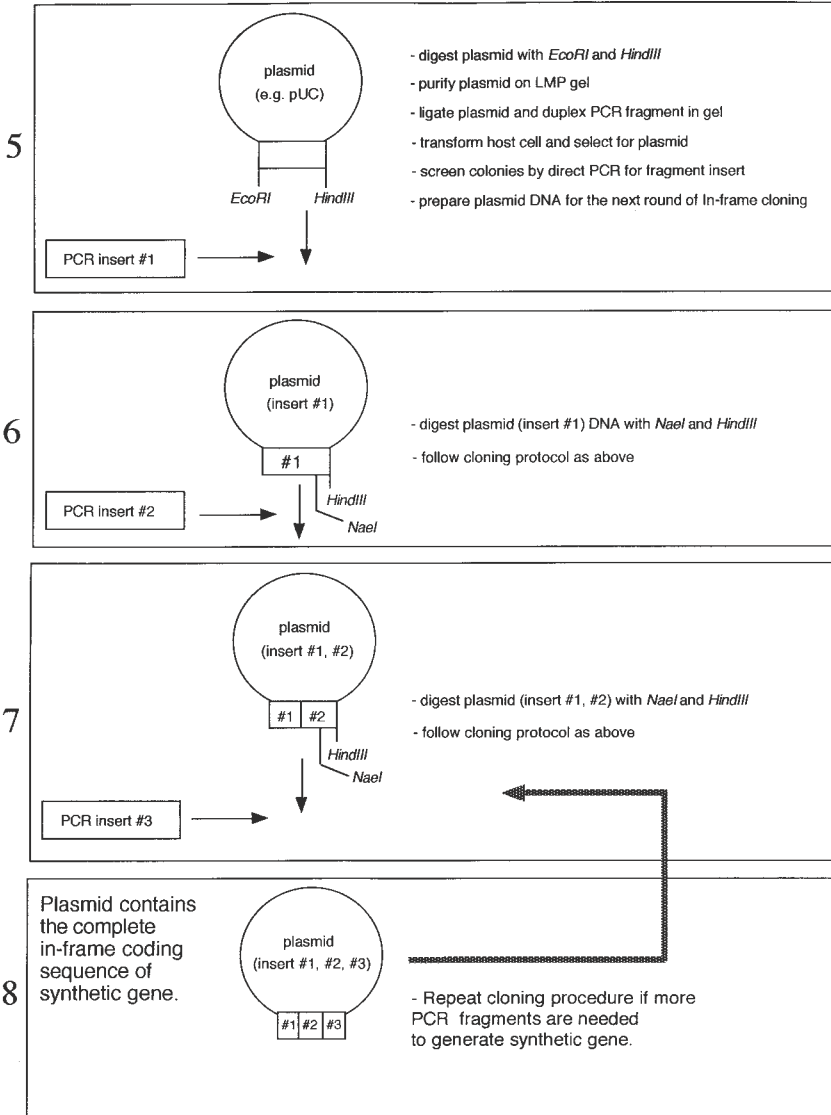


Fig. 2. Implementation of the in-frame cloning method for a three-insert synthetic gene. The execution of the cloning strategy is broken into four parts, 5–8. Part 5 shows the cloning of PCR insert #1 into a standard plasmid vector (e.g., pUC). This duplex PCR-generated segment is cloned as an *EcoRI*/*HindIII* fragment. The steps used in the in-frame cloning protocol are listed in sequential order next to the plasmid diagram. Part 6 illustrates the cloning of PCR insert #2 (blunt-end/*HindIII* fragment). Part 7 illustrates the cloning of PCR insert #3 (blunt-end/*HindIII* fragment). Part 8 shows the completed plasmid construct containing the synthetic gene. The bold arrow represents the ability to continue to clone as many DNA inserts as are needed to complete an open reading frame for any synthetic gene.

acids of the open reading frame upon blunt-end ligation. Site Y contains a diagnostic restriction site used to monitor placement of the insert when screening plasmid clones. Site Z contains the *Hind*III restriction site that is used to anchor the 3' DNA fragment to the plasmid during each ligation step of the cloning process.

5. **Figure 1**, part 4 illustrates the partial sequence of PCR fragment A/B that has been generated using the template from **Fig. 1**, part 3. This fragment contains a 5' *Eco*RI cloning site, a translational start codon, a *Nae*I blunt-end site, an *Afl*III diagnostic site, and a 3' *Hind*III cloning site. After restriction digestion with *Eco*RI and *Hind*III, this fragment can be cloned into the corresponding sites of the plasmid vector. Note that after cloning this fragment into the plasmid vector, cleavage with the restriction enzyme *Nae*I will generate a blunt-end which encodes the GCC codon of alanine (amino acid #37) at the 3' terminus. As shown below (*see Fig. 2*), subsequent rounds of in-frame cloning will insert the 5' blunt-end of the incoming fragment (C/D) next to the 3' *Nae*I-generated blunt-end of the first insert (A/B). Assuming the first three basepairs of fragment C/D encode amino acid #38, the translational reading frame of the synthetic gene will be maintained.
6. The design of primers for PCR fragment synthesis must be carefully monitored before oligonucleotide synthesis. Three points are worth noting: First, make sure that the inadvertent engineering of necessary unique restriction sites in other parts of the fragment does not occur. Screen the oligonucleotide sequence for all six-base restriction recognition sites before synthesis. It is important that the *Eco*RI, *Hind*III, blunt-end, and diagnostic restriction sites all remain unique to the engineered cloning vector. Second, make sure that the restriction sites needed for moving the synthetic gene from one vector to another have been correctly engineered. In most cases, the synthetic gene will be engineered in a standard plasmid vector (e.g., pGEM or pUC) and then subcloned into an expression vector (e.g., pMAL, baculovirus, and so on). Finally, remember to include all regulatory elements such as ribosome binding sites, transcription terminator sequences, and so on, if these elements are not present in the expression vector which the synthetic gene is to be cloned into.

3.2. Construction of a Synthetic Gene Using In-Frame Cloning

1. **Figure 2** illustrates the in-frame cloning process once an insert fragment has been generated. The protocol described in this review uses PCR generated fragments but synthetic duplex oligonucleotides or other cloned plasmid-derived fragments could also be used. The first round of in-frame cloning uses a PCR insert, which has *Eco*RI and *Hind*III sticky-ends at the 5' (left) and 3' (right) ends, respectively. All subsequent rounds use a 5' blunt-end and 3' *Hind*III sticky-end duplex DNA fragments as inserts. By a reiterative process of cutting the plasmid vector with *Eco*RI/blunt-end and *Hind*III restriction enzymes followed by agarose gel purification and ligation, a synthetic gene of almost any length can be constructed. The only limiting factor is the availability of appropriately spaced single amino acids in the protein sequence that are encoded in the blunt-end restriction sites listed in **Table 1**. Screening for plasmids with successful ligation of insert DNA is done by direct PCR analysis of antibiotic resistant bacterial colonies. This method greatly simplifies the isolation of positive clones and allows for the rapid construction of moderate to large-sized synthetic genes.
2. Assemble the following for a standard 40 μ L PCR: 0.5–1.5 μ g primer A (approx 30 pmol of a 60–80 base oligo), 0.5–1.5 μ g primer B (approx 30 pmol of a 60–80 base oligo), 4 μ L 10X buffer, 4.0 μ L 25 mM $MgCl_2$, 4.0 μ L of 2 mM dNTPs, 1 U *Taq* polymerase, and water to 40 μ L total volume.

3. Perform PCR using the following conditions (*see Note 1*): 94°C for 1 min (denaturation); 55°C for 2 min (annealing); 72°C for 2 min (extension) for five cycles; and 72°C for 10 min (final extension).
4. If a mineral oil overlay was used add an equal volume of chloroform, mix and transfer the aqueous layer to a sterile microcentrifuge tube. Incubate the tube with the lid open at 37°C for 15 min or until there is no trace odor of chloroform.
5. While the sample is drying, fill a 100-mm sterile Petri plate with 35 mL of 1X TE buffer (*see Note 2*). Float a Millipore VSM 0.025- μ m filter (shiny side up) on the TE buffer.
6. Place the PCR amplified DNA sample (up to 100 μ L) on the filter and allow 30–60 min for dialysis. Remove the sample and place it in a sterile microcentrifuge tube.
7. Digest an aliquot of the PCR amplified DNA sample by adding the following reagents to a microcentrifuge tube and incubating at 37°C for 1 h: 16 μ L DNA sample, 2 μ L 10X buffer, 1 μ L *Hind*III and 1 μ L *Eco*RI.
8. After the digestion is complete, purify the PCR DNA fragment away from the released cut ends (*see Note 3*). The protocol listed here is based on the Qiagen QIAquick-spin PCR purification columns. For more details, see manufacturer's product description.
9. Following restriction digestion add 0.5 mL of binding buffer to DNA sample. Apply sample to QIAquick-spin column and centrifuge column in a 2-mL centrifuge tube for 60 s at maximum speed. Remove flowthrough and wash column with 0.75 mL of wash buffer as above. Remove flowthrough and centrifuge again for 60 s to remove residual wash buffer. Elute DNA by adding 50 μ L of TE buffer to the spin column, placing the spin-column in a sterile 1.5-mL microcentrifuge tube and centrifuging for 60 s at maximum speed. The DNA fragment is now ready to be cloned into the plasmid cloning vector.
10. Plasmid vector DNA is prepared by cutting with restriction enzymes in the following manner: 1–2 μ g plasmid DNA (e.g., pGEM), 2 μ L 10X buffer, 1 μ L *Eco*RI, 1 μ L *Hind*III, and water to 20 μ L total volume. Incubate restriction digest at 37°C for 1 h.
11. The cut plasmid DNA is purified away from the small DNA fragment released during the double digest using low melting point (LMP) agarose gel electrophoresis (*see Note 4*). The purified DNA fragment is then ligated to the PCR insert fragment that was purified in **step 9**. Add to a microcentrifuge tube the following components: 5 μ L purified PCR fragment (approx 0.5 μ g), 10 μ L plasmid DNA fragment in agarose plug (50–100 ng), 4 μ L 5X ligation buffer, and 1 μ L DNA ligase. The ligation reaction is incubated overnight at room temperature (*see Note 5*).
12. Transform appropriate *E. coli* host cell (e.g., strain DH10B from Life Technologies) by electroporation or other high efficiency transformation protocol. Before electroporation, remove the salt from the ligation reaction by dialysis against 0.1X TE buffer for one hour at room temperature (*see Note 2*).
13. Electroporation using a Bio-Rad *E. coli* Gene Pulser (Hercules, CA) is performed as follows. The agarose plug containing the ligation reaction is melted at 70°C for 10 min and 5 μ L of the molten reaction is added to 80 μ L of electrocompetent host cells (*E. coli* strain DH10B). The DNA/host cell mixture is then placed in a 0.1-cm cuvet and electroporated at a setting of 1.80 kV. Add 1 mL of liquid media (e.g., L broth) and incubate with shaking at 37°C for 1 h. One-tenth of the transformed cells are then spread on a LB agar selector plate containing 50 mg/mL ampicillin. The plates are then incubated overnight at 37°C.
14. Screening for positive PCR fragment inserts is performed by direct PCR analysis of colonies from the ampicillin selector plates. An agar plug (*see Note 6*) of an individual ampicillin resistant colony is resuspended in 50 μ L of sterile water and mixed well. To

- each 40 μ L PCR sample tube add the following reagents: 5 μ L agar plug colony mixture (DNA template), 100 ng (ca. 15 pmol) each forward and reverse pUC/M13 primers (or T7 and Sp6 for pGEM vectors), 4 μ L 10X buffer, 4 μ L 25 mM MgCl₂, 4.0 μ L 2.0 mM dNTPs, 1 U *Taq* polymerase, and water to 40 μ L final volume.
15. Perform PCR using the following conditions: 94°C for 1 min (denaturation), 55°C for 2 min (annealing), and 72°C for 2 min (extension), 30 cycles.
 16. Analysis of PCR samples is performed by agarose gel electrophoresis using 0.5 X TBE buffer. To observe DNA fragment sizes in the range of 100 to 300 basepairs, a 3% agarose gel fractionates well. Bacteria that contain clones that are positive for the correctly ligated DNA insert will contain a PCR DNA fragment that is larger than those PCR DNA fragments amplified from bacteria that contain only vector DNA (*see Note 8*).
 17. Once positive colonies have been identified, plasmid DNA is prepared by the alkaline lysis or other DNA mini-prep procedures (**9,11**).
 18. To confirm that the PCR fragment has been cloned into the plasmid vector a restriction digest is performed. The choice of restriction enzyme depends upon which diagnostic site was engineered into the DNA insert (*see Fig. 1*, part 3). In addition to the insert diagnostic restriction enzyme, use restriction enzymes that will reveal correct insert ligation and the removal of the plasmid vector multiple cloning region (*see Note 9*). To a microcentrifuge tube add the following reagents: 5 μ L plasmid DNA (approx 0.2 μ g), 2 μ L 10X restriction buffer, 1 μ L restriction enzyme, (e.g., *Af*/III) and water to 20 μ L total volume. Incubate the restriction digest(s) at the appropriate temperature for 1 h.
 19. Fractionate the plasmid DNA from the restriction digests by agarose gel electrophoresis using a 0.8 % gel and 0.5X TBE buffer. A plasmid clone that is positive for the correct DNA insert will have a restriction pattern that indicates the presence of the diagnostic restriction site (e.g., *Af*/III). Once a positive clone is identified by PCR and then confirmed by restriction mapping, the next round of in-frame cloning is initiated. The next round of in-frame cloning (*see Note 10*) proceeds with a ligation reaction using the blunt-end and *Hind*III cut plasmid (insert 1) that was engineered in the first round of in-frame cloning (*see Fig. 2*, part 5), and the PCR fragment generated from oligonucleotides C and D (*see Fig. 1*, part 2).
 20. Using the plasmid DNA prepared in the mini-prep procedure (**step 17**) from one of the positive clones, perform a restriction enzyme double digest. One restriction enzyme will be *Hind*III. The second restriction enzyme will be a six basepair, blunt-end recognition site enzyme listed in **Table 1**. If the buffer and/or temperature conditions are incompatible for both enzymes together in one reaction, then perform the digest sequentially. To a microcentrifuge tube add the following: 1 μ g plasmid DNA (containing insert 1), 2 μ L 10X buffer, 1 μ L *Hind*III, 1 μ L blunt-end restriction enzyme from **Table 1**, and water to 20 μ L total volume. Incubate restriction enzyme at the appropriate temperature for 1 h.
 21. Purify the cut plasmid (insert 1) away from the small DNA fragment released during the double restriction digest using low-melting agarose gel electrophoresis exactly as described in **step 11** and **Note 4**.
 22. Prepare the second DNA insert (insert 2) by following the protocol outlined in **steps 1–9**. It is generally more efficient to prepare all of the DNA inserts (e.g., insert 1, insert 2, and so on) at the same time. After PCR extension to generate the double-stranded DNA fragment, the fragment must be digested with *Hind*III. As mentioned in **Subheading 3.1.**, part 4, if blunt-ends are difficult to achieve during the PCR, then a second restriction digest must be performed to generate the correct fragment ends needed for ligation. Perform the *Hind*III digest as follows: 17 μ L DNA sample, 2 μ L 10X buffer, and 1 μ L

*Hind*III. Incubate the reaction at 37°C for 1 h. Purify the DNA fragment away from small released fragment(s) by column chromatography as described in **steps 8 and 9**.

23. A ligation reaction is prepared exactly as described in **step 11**. After ligation and transformation (**steps 11–13**), positive clones are identified as described in **steps 14–19**. In-frame cloning can be reiterated as many times as is necessary to engineer a synthetic gene of the desired length (*see* **Fig. 2**, part 8).
24. When all of the necessary rounds of in-frame cloning have been completed, the insert nucleotide sequence must be confirmed by DNA sequence analysis. Candidate plasmid constructs should first be confirmed by analysis of the correct insert size (as determined by PCR amplification and comparison to vector-with-no-insert and intermediate plasmid constructs) and by restriction mapping to indicate that the correct diagnostic restriction sites are present and that certain vector restriction sites have been removed. Sequence analysis can be performed by double stranded PCR sequencing using M13/pUC universal primers (**II**). Single-stranded DNA can be made and sequenced using vectors such as the pGEM-f1 ori series.
25. A rough timeline is presented below to serve as a guide for experiment design. It is assumed that all oligonucleotide design and synthesis has been completed. It also assumes that a preliminary PCR experiment has shown that the oligonucleotide PCR fragments are of the correct size and of decent quality.

- Day 1: PCR amplification of oligonucleotide insert fragments
Gel electrophoretic analysis of PCR samples
- Day 2: Cut PCR fragments with restriction enzyme(s)
Cut plasmid vector with restriction enzymes
Purify PCR fragments by spin chromatography
Purify plasmid vector by LMP agarose gel electrophoresis
Set up ligation reactions and incubate overnight
- Day 3: Transform *E. coli* host cells
Grow colonies overnight on ampicillin selector plates
- Day 4: Screen selected colonies by direct PCR
Gel electrophoresis of PCR/colony samples
Overnight culture of tentatively positive colonies
- Day 5: Miniprep plasmid DNA isolation
Restriction mapping using diagnostic site analysis
Gel electrophoresis of plasmid restriction digests
Prepare vector for the next round of in-frame cloning

4. Notes

1. The generation of the duplex DNA fragments in this protocol is really just a DNA polymerase extension reaction. The PCR is used because it is simple and the reagents are generally available. A small number of amplification cycles (e.g., 5) should generate sufficient product. If priming artifact occurs try using fewer cycles. One cycle may be enough to generate sufficient duplex DNA for cloning. If PCR product artifacts are a problem, the use of 5' outside primers to amplify the correct size duplex fragment from the initial PCR amplification may be needed. The use of synthetic oligonucleotide duplex DNA for direct insert cloning is always an option.
2. The author uses microdialysis routinely to prepare DNA substrates that must undergo a number of different and separate enzymatic treatments. Although many DNA/enzyme reactions work adequately under “universal” buffer conditions, complex protocols that require a number of linked reactions often show low overall efficiency. It has been my

experience that if each DNA modification reaction is treated individually, the success of the entire experiment is greatly enhanced. Microdialysis is a technique that allows the efficient exchange of buffers for any DNA sample. Once mastered it is often more convenient and efficient than ethanol precipitation or spin chromatography. The following protocol works well for DNA fragments 100 bases or larger.

- a. Fill a 100-mm sterile Petri plate with 35 mL of 1X TE buffer. Float a Millipore VSM 0.025-mm filter (shiny side up) on the TE buffer. Make sure no buffer wets the top side of the filter membrane.
 - b. Place the DNA sample on the filter. Volumes from 10–100 μ L can be used but care must be taken with larger volumes. Multiple samples can be placed on one filter as long as they do not contact one another. As described later, sample volumes often increase or decrease, so beware!
 - c. Place the lid on the Petri plate and incubate for about one hour at room temperature. Make sure the Petri plate is in a safe place on the bench because if it is disturbed, the sample may be lost. Length of time will depend on how dramatic the change in buffer conditions will be. Simple desalting can usually be accomplished in less than 30 min.
 - d. Remove the sample using a pipeting device and place it in a sterile tube. The volume of the DNA sample can change significantly because of osmotic imbalances or evaporation, so it is often useful to record the pre-and post-microdialysis volumes. It has been my experience that very little DNA is lost by non-specific binding to the membrane in this procedure.
 - e. This protocol works well with agarose plugs when it is necessary to remove the electrophoresis buffer before enzymatic manipulation. The agarose plug is melted at 70°C and then placed directly on the filter that is floating on the TE buffer. The molten agarose will form a semi-solid plug on the filter but this does not interfere with buffer exchange. To remove the agarose plug containing the DNA sample, first lift the entire filter membrane off the buffer using forceps. Then scrape the agarose plug off the filter directly into a microcentrifuge tube using a sterile scalpel. The sample is then incubated at 70°C and the agarose-DNA sample can then be added to the next reaction mixture.
3. Ligation reactions are not always logical! Often the products recovered from a ligation reaction are a mixture of intended constructs and a collection of obscure and unlikely side-products. The source of unintended DNA constructs is often insert-insert ligation events. Therefore, it behooves one to think critically about just what DNA substrates are being placed in the ligation reaction. When synthetic duplex oligonucleotides are used as DNA inserts, there is not much concern about insert ligating to itself since the 5' ends are not phosphorylated. With the PCR generated DNA fragment approach described in this protocol, self-ligation is a potential problem. Although the forced orientation cloning approach minimizes ligation artifacts, there is still the possibility of multiple inserts being cloned into the vector during a ligation reaction.

The first round of in-frame cloning uses an *EcoRI/HindIII* sticky-ended fragment. There is significant possibility for self-ligation under the conditions used here (*see later*). Subsequent rounds of in-frame cloning using blunt-end/*HindIII* substrates pose less of a problem because of the lower efficiency of blunt-end ligation. A tripartite *EcoRI/HindIII* fragment will insert into the plasmid vector giving an incorrect construct. Because the in-frame cloning method screens for positive DNA inserts using direct PCR amplification of selected colonies followed by fragment size analysis using gel electrophoresis, this multiple insert problem should not interfere with the engineering experiment.

4. The double-digested plasmid is purified on a 0.6% LMP agarose gel using 0.5X TBE buffer. The gel is stained with ethidium bromide and the linearized plasmid cut out and then placed in 1 mL of sterile water to remove excess TBE buffer. After 30 min, remove the water and heat the agarose gel plug at 70°C for 10 min. It is important that the components of next reaction have already been prepared before the agarose gel fragment is melted. After heating, the molten agarose is added directly to the ligation reaction before it solidifies. Using the pipet tip, gently mix the molten agarose with the ligation reaction components and allow the tube to sit at room temperature. Do not centrifuge the sample once the agarose has been added, as this will separate the plug from the other reaction components. Depending on the reaction buffer and the amount of agarose added, the reaction mixture may or may not solidify at room temperature.
5. The molar ratio of insert to vector in the ligation reaction described in this protocol is approx 100:1. This is calculated by assuming the plasmid vector to be 3000 bp of which 0.1 µg is added to the ligation reaction. This would equal about 0.05 pmole of DNA (1 bp = 660 Daltons, 1 pmole of a basepair of DNA = 660 pg). If the average insert fragment is about 150 bp and 0.5 µg were used in the ligation reaction then about 5 pmoles would be present in the sample. Thus, a 5–0.05 pmole DNA ratio is achieved. This ratio is somewhat high for a standard sticky-end ligation reaction but is used to drive the ligation reaction in the agarose plug. If problems with multiple inserts occur, then decrease the concentration of fragment insert. It is difficult to modulate the concentration of vector DNA because an agarose plug strategy is being used and low cloning efficiencies are a problem if too little plasmid vector is used during transformation. In general, the conditions described here work well and the majority of plasmids contain the correct insert. Forced orientation and removal of restriction cleavage ends “push” the ligation reaction in the direction of correct insertion.
6. Screening for positive vector inserts is best accomplished by direct PCR amplification of bacterial colonies taken from the ampicillin selector plates. A sterile 50-mL capillary pipet (or 1-mL pipet) is used to isolate or “plug” well-separated colonies from the agar plate. The agar plug is then placed in a sterile 1.5-mL microcentrifuge tube containing 50 µL of sterile water and mixed well using a vortex mixer. A 5-µL sample of this bacterial colony/agar plug/water mixture is then added to a standard 40 µL PCR tube as described in the protocol. The author has found it convenient to work-up 18 samples at a time and then perform gel electrophoretic analysis using a 20-well comb with two lanes used as molecular weight marker and control PCR (vector no-insert), respectively. The tubes containing the bacterial colony/agar plug are stored at 4°C until the PCR and electrophoresis results are obtained. Once positive colonies are tentatively identified, an overnight culture is made using 3 mL of Luria broth, 50 mg per mL of ampicillin, and 25 µL of the bacterial colony/agar plug sample. This is grown overnight at 37°C with moderate shaking. The overnight culture can then be used to make an alkaline lysis miniprep DNA sample that is then used for restriction mapping and/or prepared for the next round of in-frame cloning. Once positive constructs are identified, the remaining overnight culture can be kept for long term storage by adding glycerol to 30% of total volume, mixing well and then storing at –80°C.
7. These primers flank the multiple cloning site of many pUC-based cloning vectors. If another vector is being used for in-frame cloning you may need a different set of primers. A control reaction should be performed from a cell-containing vector DNA that has not been manipulated (usually taken from the selector plate used to monitor transformation efficiency, e.g., uncut pGEM DNA).

8. The difference in fragment size between the control PCR sample (colonies with vector-no insert) and experimental samples (colonies with vector plus insert) will allow you to determine which host cells contained plasmid DNA with the PCR fragment insert cloned into the *EcoRI/HindIII* site. Remember to subtract the length of the multiple cloning region (approx 50 bp) of uncut pGEM or pUC-based vectors when comparing fragment sizes.
9. Analysis of the diagnostic restriction site (e.g., *AflIII*) as an identifier for the correct insert fragment, and loss of restriction sites (e.g., *PstI*) from the multiple cloning region (MCR) of the vector, ensures that this PCR fragment was cloned and the vector MCR removed. You can also check for the regeneration of the *EcoRI* and *HindIII* sites. This restriction digest screen will allow you to have good confidence that the correct fragment has indeed been cloned. During this restriction analysis the plasmid DNA can be prepared for the next step of in-frame cloning by cutting with *HindIII* and the blunt end recognition restriction enzyme. All of these samples can be run on a low-melting-point agarose gel and the *HindIII*/blunt end cut plasmid purified as described above (also see **Note 2**).
10. This process can be reiterated as many times as is necessary to clone a moderate to large-sized synthetic gene (200–1000 bp). If a very large gene is going to be made by the synthetic in-frame cloning method it may be more practical to use two or more plasmids as the templates for constructing different portions of the gene. Just duplicate each of the cloning and processing steps using different DNA inserts. The assembled gene fragments are then isolated by blunt-end/blunt-end or blunt-end/*HindIII* restriction digestion and then linked to the primary plasmid template by the in-frame cloning process.

References

1. Prodromou, C. and Pearl, L. H. (1992) Recursive PCR: a novel technique for total gene synthesis. *Protein Eng.* **5**, 827–829.
2. Jaffe, E. K., Volin, M., Bronson-Mullins, C., Dunbrack, R. L., Kervinen, J., Martins, J., et al. (2000) An artificial gene for human porphobilinogen synthase allows comparison of an allelic variation implicated in susceptibility to lead poisoning. *J. Biol. Chem.* **275**, 2619–2626.
3. Johnson, T. M., Quick, M. W., Sakai, T. T., and Krishna, N. R. (2000) Expression of functional recombinant beta-neurotoxin C_{ss} II in *E. coli*. *Peptides* **6**, 767–772.
4. Pierce, J. C. (1994) In-frame cloning of large synthetic genes using moderate-size oligonucleotides. *BioTechniques* **16**, 708–715.
5. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.
6. Barnett, R. W. and Erfle, H. (1990) Rapid generation of DNA fragments by PCR amplification of crude, synthetic oligonucleotides. *Nucl. Acids Res.* **18**, 3094.
7. Dillion, P. J. and Rosen, C. A. (1990) A rapid method for the construction of synthetic genes using the polymerase chain reaction. *BioTechniques* **9**, 298–300.
8. Sandhu, G. S., Aleff, R. A., and Kline, B. C. (1992) Dual asymmetric PCR: One-step construction of synthetic genes. *BioTechniques* **12**, 14–16.
9. Birnboim, H. C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513–1523.
10. Kane, J. F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotech.* **5**, 494–500.
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Megaprimer PCR

Sailen Barik

1. Introduction

A large variety of procedures of site-directed mutagenesis based on polymerase chain reaction (PCR) have been developed over the last decade. Among them, the “megaprimer” method, originally reported in 1990 (1), and its subsequent updates (8,10) still retain their popularity because they combine simplicity and versatility. Our most recent search in the PUBMED, using “megaprimer” as the keyword, generated 24 publications, many of which were improvements on the original theme. This is an impressive number, considering that “megaprimer” is essentially a specialized technique. In this chapter, we provide an updated protocol incorporating the variations and improvements of the basic technique published over the past decade. These include: a combination of megapriming and overlap extension, improvement of yield, use of single-stranded DNA, spiking with a proofreading polymerase (e.g., *Pfu*) to avoid unwanted mutations arising from nontemplated insertions by *Taq* polymerase, and the inclusion of various kinds of mutations, including multiple, nonadjacent ones (2–12,18,19,26–32).

The basic method, described in **Fig. 1**, still requires three oligonucleotide primers and two PCRs (termed PCR-1 and -2 here) employing the wild-type DNA as template (1,2,8,10). The “mutant” primer is represented by M, and the two “flanking” primers, by A and B. The M primer may encode a substitution, deletion, insertion, or a combination of these mutations, thus providing versatility while using the same fundamental strategy (10). The first PCR (PCR-1) is performed using the mutant primer M and one of the flanking primers, such as A (see **Fig. 1**). The double-stranded product A-M is purified and used as a primer (hence the name “megaprimer”; **ref. 1**) in the second PCR (PCR-2) together with the other flanking primer B. Note that both strands of the megaprimer have the potential to prime on the respective complementary strands of the template. However, the fundamental principles of PCR amplification ensure that only that strand of the megaprimer, which extends to the other primer (B in **Fig. 1**), will be *exponentially* amplified into the double-stranded product in PCR-2. As aforementioned, the wild-type DNA is used as template in both PCRs.

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

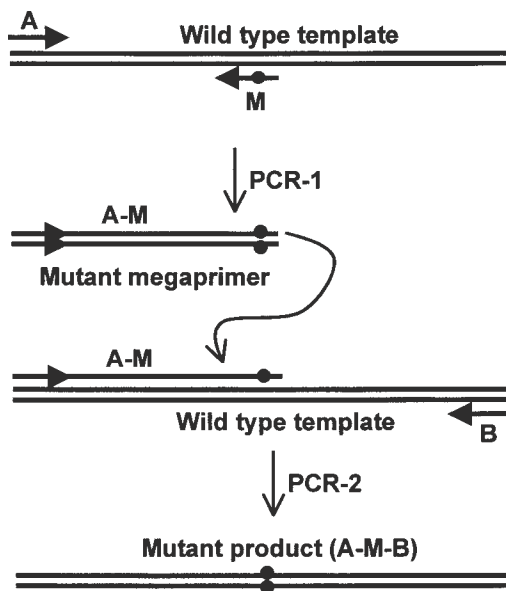


Fig. 1. The basic megaprimer method. Primers A, B, M, and the priming strand of the megaprimer AM are indicated by thinner lines with arrowhead, while the thicker double lines represent the wild type template (usually part of a plasmid clone, not shown). Primers A and B contain restriction sites (e.g., *NdeI* and *BamHI*) indicated as thicker regions, and extra “clamp” sequence at the 5' end indicated by double lines. The sequence to be inserted is shown as the dotted region in primer M and the subsequent PCR products. The final product containing the insertion is restricted and cloned.

Poor yields from PCR-2 have sometimes been reported, especially when the megaprimer is large (0.8 kb and above). Although the exact reason remains unclear, it must have something to do with the unique features of the megaprimer: its double-stranded nature and large size. Strand separation of the double-stranded megaprimer is essentially achieved in the denaturation steps of the PCR cycle. Under some conditions, however, self-annealing of the megaprimer apparently tends to reduce the yield of the product (4).

Here, we provide a brief overview of the various solutions suggested to overcome the low yield. In one approach, a biotin tag is added to the 5' end of primer A, which generates a biotin-labeled megaprimer in PCR-1. The megaprimer is denatured, and the biotinylated strand is purified on streptavidin-coupled magnetic beads (30). In another method (31), the use of two parallel templates allowed inclusion of two flanking primers as well as the megaprimer in PCR-2, resulting in a direct amplification of the final product. Use of a “one-tube” method (see Note 6), when properly optimized, may help eliminate the loss of megaprimer during the purification step. Other strategies for increasing the yield of PCR-2 rely on optimizing the concentrations of template and megaprimer (4,11,12,32). The use of higher amounts of template (in the microgram

range, as opposed to nanogram quantities used in standard PCR) in PCR-2 has been shown to dramatically increase the product yield for some sequences (4). A more general strategy, however, is to increase the amount of the megaprimer. A method that we have found useful is to carry out the first several cycles of PCR-2 with the megaprimer only. After this initial asymmetric PCR, the small primer is added (11), and PCR is continued. In an optimized method (31), the starting concentration of megaprimer is increased to 6 μg (from 25 ng) per 100 μL PCR-2. We have adopted a combination of the last two findings in this article.

2. Materials

1. Template: About 100 ng DNA template to be mutated (e.g., a gene cloned in a plasmid).
2. Primers: 100 pg of oligonucleotide primers A and B (see Fig. 1), and 50 ng of mutant primer M; one primer, say A, in the opposite sense, and other primer, B, in the same sense as the mutant primer M. If the plan is to digest the final product with restriction enzymes for the purposes of cloning, include restriction sites, preferably unique, in these primers. Realize that the mutant primer may contain a point mutation, or insertion, or deletion, as desired (see Notes 1 and 2).
3. PCR buffer: 10X PCR buffer for *Pfu* polymerase (Stratagene Cloning Systems, La Jolla, CA) is 200 mM Tris-HCl (pH 8.0–8.3), 100 mM KCl, 20 mM MgCl_2 , 60 mM ammonium sulfate, 1% Triton X-100, 100 $\mu\text{g}/\text{mL}$ nuclease-free BSA. The buffer is usually supplied with the enzyme.
4. Deoxyribonucleotides: Use a final dNTP concentration of 200 μM for each nucleotide. Make a stock dNTP mix containing 2 mM of each dNTP (dATP, dCTP, dGTP, dTTP). We make it by adding 50 μL of 10 mM stock solutions of each nucleotide, available commercially, into 50 μL H_2O , to produce 250 μL of the stock mix.
5. Analysis and purification of DNA: A system for purifying the PCR products, such as gel electrophoresis, followed by recovery of the appropriate DNA band in the excised agarose fragment.

Wherever needed in this procedure, use deionized (e.g., Millipore) autoclaved water.

3. Methods

3.1. PCR-1: Synthesis of the Megaprimer

1. We have assumed that the reader is familiar with the basic PCR. Use the following recipe for the first PCR. Make the following 100 μL reaction mix in an appropriate microcentrifuge tube (0.5 or 1.7 mL, dictated by the heating block of your thermal cycler):

H_2O	75 μL
10X PCR buffer	10 μL
dNTP mix (2 mM each)	10 μL
(final concentration of each nucleotide is 200 μM)	
Primer A	50 pmole
Primer M	50 pmole
DNA template	10–100 ng
<i>Pfu</i> polymerase (2.5 U)	0.5 μL
(or 2.5 U <i>Taq</i> plus 0.1 U <i>Pfu</i> polymerase, see Note 1)	
(Total = 100 μL)	

2. Vortex to mix, then spin briefly in a microfuge. If the thermal cycler has a heated lid, then proceed to perform PCR; otherwise, reopen the tube, overlay the reaction mixture with

enough mineral oil to cover the reaction (approx 100 μL for a 0.5-mL microfuge tube), then close cap. The tube is now ready for thermal cycling.

3. Perform PCR-1 using the following cycle profiles.
 - Initial denaturation 94°C, 3 min
 - 30–40 main cycles 94°C, 1 min (denaturation)
 - T°C (depending on the T_m of the primers), 2 min (annealing)
 - 72°C, appropriate time, depending on product length (extension)
 - Final extension 72°C, 1.5X n min

Following synthesis, the samples are maintained at 4°C (called “soak” file in older Perkin-Elmer programs) for a specified time. Some instruments lack an active cooling mechanism and keep samples at an ambient temperature of about 20°C by circulating tap water around the heat block, which appears to be adequate for overnight runs.
4. After PCR, proceed directly to the next step if there is no oil overlay. Otherwise, first remove the oil as follows. (If oil is not removed completely, the sample will float up when loaded in horizontal agarose gels!) Add 200 μL of chloroform to each tube. The mineral oil and chloroform will mix to form a single phase and sink to the bottom of the tube. Spin for 30 s in a microfuge. Carefully collect approx 80 μL of top aqueous layer and transfer to a fresh Eppendorf tube.
5. Purify the megaprimer using any standard procedure of your choice (as long as the nonmutagenic primer A is removed) and use it in PCR-2 below.

3.2. PCR-2: Synthesis of the Mutant Using the Megaprimer

1. Reconstitute 100 μL PCR as follows:

10XPCR buffer	10 μL
dNTP mix (2 mM each)	10 μL
(final concentration of each nucleotide is 200 μM)	
All of the recovered megaprimer (A-M)	
from the previous step (<i>see Note 3</i>)	20–50 μL
DNA template	0.2 μg
Make up volume to 100 μL with H_2O .	
Mix well.	
2. Start reaction essentially as described for PCR-1, except that a “hot-start” is preferred (*see Note 4*) and is performed as follows. When the reaction is in the annealing step of the first cycle, open the cap briefly, quickly add 0.5 μL *Pfu* polymerase (2.5 U) (or 2.5 U *Taq* plus 0.1 U *Pfu* polymerase), and mix by pipeting. Close the cap and let PCR continue.
3. After five cycles, when the reaction is again at an annealing step, promptly add 50 pmole of primer B, mix well, and let PCR continue another 30 cycles. (The small amounts of primer B and *Pfu* polymerase do not contribute significantly to the total reaction volume and therefore, have been ignored in the volume calculations.)
4. Do another PCR in parallel, using primers A and B (and no megaprimer) and the same wild-type template; use an aliquot (5 μL) of this PCR as a size marker when analyzing PCR-2 by gel electrophoresis. This will also help in identifying the real product (in PCR-2) among the wrong ones that sometimes result from mispriming.
5. Gel-purify the final mutant PCR product essentially as described earlier for the purification of the megaprimer (*see Notes 5 and 6*). Now it is ready for restriction, cloning, sequencing etc.

4. Notes

1. The problem of nontemplate nucleotides and its solution. Perhaps the most unique feature of the megaprimer method is that the product of one PCR becomes a primer in the next, which creates the following potential problem. *Taq* polymerase, due to its lack of proof-reading activity, tends to extend the product DNA beyond the template by adding one or two nontemplate residues, predominantly A's (14). When the product is used as a primer in the next round of PCR (PCR-2), these nontemplated A residues may not match with the template, and therefore, will either abrogate amplification (15–17) or produce an undesired A-substitution. A variety of solutions to this problem have been recommended (5,8,10,18). The first is to design the mutant primer such that there is at least one T residue beyond the 5' end of the primer sequence in the template. Thus, when the complementary strand incorporates a nontemplated A at the 3' end, it will still be complementary to the other strand. If the template sequence does not permit this, a second solution, which we have recommended in this chapter, is to use a mixture of *Taq* and *Pfu* DNA polymerases in 20:1 ratio in PCR-2 (3), or to use *Pfu* exclusively. The 3' exonuclease activity of *Pfu* should remove any mismatch at the 3' end of the megaprimer; however, this proof-reading ability also necessitates the addition of at least ten perfectly matched bases on both the 5' and 3' ends of the mutagenic primer (8,10,15,19). We have not actually tested other polymerases that are proficient in proofreading; but some of them might be used in lieu of *Pfu*.

In addition to these unique considerations, the general rules of primer design, some of which are described below, should be followed.

2. Length of the megaprimer. Try to avoid making megaprimers (A-M) that approach the size of the final, full-length product (gene) A-B (see Fig. 1). Briefly, if M is too close to B, it will make separation of AB and AM (unincorporated, left-over megaprimer) difficult after PCR. When the mutation is to be created near B, one should make an M primer of the opposite polarity, and synthesize BM megaprimer (rather than AM), and then do PCR-2 with BM megaprimer and the A primer. When the mutation is at or very near the 5' or 3' end of the gene (within 1–50 nucleotides), there is no need to use the megaprimer method; one can simply incorporate the mutation in either A or B primer and do a straightforward PCR using A and B primers! For borderline situations, such as when the mutation is, for example, 120 nucleotides away from the 5' end of the gene, incorporation of the mutation in primer A may make the primer too big to synthesize; or else, it will make the megaprimer AM too short to purify away from primer B. In such a case, simply back up primer A a few hundred bases further upstream in order to make the AM megaprimer longer. In general, realize that primers A and B can be located virtually anywhere on either side of the mutant primer M, and therefore, try to utilize this flexibility as an advantage when designing these primers.
3. Molar amount of megaprimer. Since the megaprimer is large, one needs to use a greater quantity of it to achieve the same number of moles as a smaller primer. Example: 50 pmoles of a 20 nt-long single-stranded primer will equal 0.3 μg ; however, 50 pmoles of a 500 nt-long double-stranded megaprimer will equal 6 μg . A good yield and recovery of megaprimer is, therefore, important. If needed, do 2X 100 μL PCRs to generate the megaprimer. There is no need to remove the template DNA after PCR-1, because the same DNA will be used as template in PCR-2.

4. "Hot start" PCR-2. The hot-start technique used in PCR-2 works just as well as the more expensive commercial methods. Hot start tends to reduce false and nonspecific priming in PCR in general (39) and is particularly useful in PCR-2 of the megaprimer method (our unpublished observation).
5. Poor yield of mutant. If the final yield is poor, the surest strategy is to amplify a portion of the gel-purified mutant product in a third PCR (PCR-3) using primers A and B and hot start. This may also be necessary if PCR-2 produces nonspecific products in addition to the specific one. Before PCR-3 is carried out, however, it is very important to ensure that the mutant product of PCR-2 is well separated from the wild-type template in the gel purification; otherwise, PCR-3 will amplify the wild-type DNA as well. The final gel-purified mutant DNA (from either PCR-2 or PCR-3) is ready for a variety of applications, as described later in brief.
6. Single-tube methods. Recently, various investigators have reported successful modifications of the megaprimer method in which the purification step is either simplified or not required (19,26–29). One involves cleavage of the template, coupled with enzymatic removal of PCR-1 primers, to ensure amplification of the correct product in PCR-2 (27). A second possibility is to exploit the unusually high T_m of the megaprimer by designing a short, low T_m flanking primer for PCR-1, and a long flanking primer for PCR-2. This enables the use of a higher T_m for PCR-2 such that it will only allow annealing of the appropriate flanking primer (28). A third method uses a limiting amount of the first flanking primer, such that when the second flanking primer is added, the principle product will be the mutant DNA (19). Although we have not tested any of these modifications, the interested reader is advised to consult the original papers.

Acknowledgment

Research in the author's laboratory was supported in part by NIH Grant AI 45803 and by a Burroughs Wellcome New Initiatives in Malaria Research Award (to S.B.).

References

1. Sarkar, G. and Sommer, S. S. (1990) The "megaprimer" method of site-directed mutagenesis. *BioTechniques* **8**, 404–407.
2. Landt, O., Grunart, H.-P., and Hahn, U. (1990) A general method for rapid mutagenesis using the polymerase chain reaction. *Gene* **96**, 125–128.
3. Perrin, S. and Gilliland, G. (1990) Site-specific mutagenesis using asymmetric polymerase chain reaction and a single mutant primer. *Nucl. Acids Res.* **18**, 7433–7438.
4. Barik, S. and Galinski, M. S. (1991) "Megaprimer" method of PCR: Increased template concentration improves yield. *BioTechniques* **10**, 489–490.
5. Kuipers, O. P., Boot, H. J., and de Vos, W. M. (1991) Improved site-directed mutagenesis method using PCR. *Nucl. Acids Res.* **19**, 4558.
6. Merino, E., Osuna, J., Bolivat, F., and Soberon, X. (1992) A general PCR-based method for single or combinatorial oligonucleotide-directed mutagenesis on pUC/ML3 vectors. *BioTechniques* **12**, 508–510.
7. Aiyar, A. and Leis, J. (1993) Modification of the megaprimer method of PCR mutagenesis: improved amplification of the final product. *BioTechniques* **14**, 366–369.
8. Barik, S. (1993) Site-directed mutagenesis by double polymerase chain reaction. *Meth. Mol. Biol.* **15**, 277–286.
9. Pont-Kingdon, G. (1994) Construction of chimeric molecules by a two-step recombinant PCR method. *Biotechniques* **16**, 1010–1011.

10. Barik, S. (1995) Site-directed mutagenesis by PCR: Substitution, insertion, deletion, and gene fusion. *Meth. Neurosci.* **26**, 309–323.
11. Datta, A. K. (1995) Efficient amplification using ‘megaprimer’ by asymmetric polymerase chain reaction. *Nucl. Acids Res.* **23**, 4530–4531.
12. Brons-Poulsen, J., Petersen, N. E., Horder, M., and Kristiansen, K. (1998) An improved PCR-based method for site-directed mutagenesis using megaprimers. *Mol. Cell. Probes* **12**, 345–348.
13. Colosimo, A., Xu, Z., Novelli, G., Dallapiccola, B., and Gruenert, D. C. (1999) Simple version of “megaprimer” PCR for site-directed mutagenesis. *Biotechniques* **26**, 870–873.
14. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677–9686.
15. Barnes, W. M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
16. Sarkar, G., Cassady, J., Bottema, C. D. K., and Sommer, S. S. (1990) Characterization of polymerase chain reaction amplification of specific alleles. *Analyt. Biochem.* **186**, 64–68.
17. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J. J. (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type I model studies. *Nucl. Acids Res.* **18**, 999–1005.
18. Chattopadhyay, D., Raha, T., and Chattopadhyay D. (1997) PCR mutagenesis: treatment of the megaprimer with mung bean nuclease improves yield. *BioTechniques* **6**, 1054–1056.
19. Picard, V., Ersdal-Badju, E., Lu, A., and Bock, S. C. (1994) A rapid and efficient one-tube PCR-based mutagenesis technique using Pfu DNA polymerase. *Nucl. Acids Res.* **22**, 2587–2591.
20. Rychlik, W. (1993) Selection of primers for polymerase chain reaction. *Meth. Mol. Biol.* **15**, 31–40.
21. *New England Biolabs Catalog* 1995, pp. 208,209.
22. Innis, M. A., Myambo, K. B., Gelfand, D. H., and Brow, M. A. D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. USA* **85**, 9436–9440.
23. Suggs, S. V., Hirose, T., Miyake, E. H., Kawashima, M. J., Johnson, K. I., and Wallace, R. B. (1981) Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences, in *ICN-UCLA Symposium on Developmental Biology Using Purified Genes* vol. 23, (Brown, D. D., ed.), Academic, New York, pp. 683–693.
24. Tautz, D. and Renz, M. (1983) An optimized freeze-squeeze method for recovering long DNA from agarose gels. *Analyt. Biochem.* **132**, 14–19.
25. Zintz, C. B. and Beebe, D. C. (1991) Rapid re-amplification of PCR products purified in low melting point agarose gels. *BioTechniques* **11**, 158–162.
26. Ling, M., and Robinson, B. H. (1995) A one-step polymerase chain reaction site-directed mutagenesis method for large gene-cassettes with high efficiency, yield, and fidelity. *Analyt. Biochem.* **230**, 167–172.
27. Seraphin, B. and Kandels-Lewis, S. (1996) An efficient PCR mutagenesis strategy without gel purification step that is amenable to automation. *Nucl. Acids Res.* **24**, 3276–3277.
28. Ke, S. H., and Madison, E. L. (1997) Rapid and efficient site-directed mutagenesis by single-tube ‘megaprimer’ PCR method. *Nucl. Acids Res.* **25**, 3371,3372.
29. Karreman, C. (1998) Fusion PCR, a one-step variant of the “megaprimer” method of mutagenesis. *Biotechniques* **24**, 736–742.
30. Fieschi, J., Niccoli, P., Camilla, C., Chames, P., Chartier, M., and Baty, D. (1996) Polymerase chain reaction-based site-directed mutagenesis using magnetic beads. *Analyt. Biochem.* **234**, 210–214.

31. Upender, M., Raj, L., and Weir, M. (1995) Megaprimer method for in vitro mutagenesis using parallel templates. *BioTechniques* **18**, 29–32.
32. Smith, A. M., and Klugman, K. P. (1997) “Megaprimer” method of PCR-based mutagenesis: the concentration of megaprimer is a critical factor. *BioTechniques* **22**, 438–442.
33. Meltzer, S. J. (1993) Direct sequencing of polymerase chain reaction products. *Meth. Mol. Biol.* **15**, 137–142.
34. Sarkar, G. and Bolander, M. E. (1995) Semi exponential cycle sequencing. *Nucl. Acids Res.* **23**, 1269–1270.
35. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 113–119.
36. Barik, S. (1993) Expression and biochemical properties of a protein Ser/Thr phosphatase encoded by bacteriophage __ *Proc. Natl. Acad. Sci. USA* **15**, 10,633–10,637.
37. Mazumder, B., Adhikary, G., and Barik, S. (1994) Bacterial expression of human respiratory syncytial viral phosphoprotein P and identification of Ser²³⁷ as the site of phosphorylation by cellular casein kinase II. *Virology* **205**, 93–103.
38. Krowczynska, A. M. and Henderson, M. B. (1992) Efficient purification of PCR products using ultrafiltration. *BioTechniques* **13**, 286–289.
39. Mullis, K. B. (1991) The polymerase chain reaction in an anemic mode: how to avoid cold oligodeoxyribonuclear fusion. *PCR Meth. Applicat.* **1**, 1–4.
40. Stoffel, E. S., Koeberl, D. D., Sarkar, G., and Sommer, S. S. (1988) Genomic amplification with transcript sequencing. *Science* **239**, 491–494.

PCR-Mediated Recombination

A General Method Applied to Construct Chimeric Molecular Clones

**Guowei Fang, Barbara Weiser, Aloise Visosky, Timothy Moran,
and Harold Burger**

1. Introduction

Molecular cloning has proven to be a powerful tool in biology, and chimeric clones are useful in a variety of fields including microbial pathogenesis and the development of vaccines. Chimeras can be created from DNA by using conventional cloning techniques, specifically restriction cleavage and DNA ligation. Such techniques, however, have limitations; most commonly, limitations result from the lack of restriction sites to provide points of entry for inserts in the desired regions or the multiplicity of restriction sites in other regions of the DNA. Because recombinant DNA molecules may be created during polymerase chain reaction (PCR) when two or more different DNA sequences are brought together (1,2), PCR-mediated recombination has been exploited to join DNA fragments of a few hundred bases (3–8). There are two drawbacks to these methods. First, they often involve multiple steps, and second, sequence errors frequently are introduced by certain thermostable polymerases during the PCR reaction (9,10).

We have developed a widely applicable, improved method to construct recombinant DNA molecules without reliance on restriction sites. The method differs from older PCR-mediated recombination procedures (3–8) in several ways: it is useful for a wide range of constructions, ranging from a few hundred bases to approximately 10 kb; it is based on asymmetric PCR that greatly increases the yield of desired products; it employs high-fidelity DNA polymerase; and results in a very low error rate. The technique utilizes PCR-amplified DNA, including DNA synthesized from RNA by reverse transcription.

To demonstrate the power of PCR-mediated recombination and provide an example of the general utility of the method, we constructed chimeric infectious molecular clones of HIV-1 derived from plasma viral RNA and proviral long-terminal repeats (LTR)s. To study pathogenesis and develop vaccines, it would be desirable to use infectious HIV and SIV clones derived from plasma viral RNA, which is more representative of the replicating virus pool than proviral DNA (11,12). The PCR-based method described here made construction of such clones possible without reliance on restriction sites. DNA sequences were first amplified by high-fidelity PCR using Pfu polymerase; they were then used both as megaprimers and templates in subsequent asymmetric long PCR amplifications to form chimeric clones. Biological characterization of these clones showed that most were infectious in tissue culture and sequence analysis demonstrated an error rate of only one base change in 20 kb of DNA sequence.

2. Materials

2.1. Clones and Strains

The sources of DNA are the following: plasmid HIV-1 pNL4-3 (13), which contains a full-length HIV-1 proviral genome; plasmid FG901-18, which contains the 5' half of an HIV-1 cDNA genome (5 kb); and plasmid FG902-12, which contains the 3' half of the same HIV-1 genome (5 kb). Both plasmids FG901-18 and FG902-12 are derived from plasma viral RNA from Wadsworth Center patient 001 (WC1-PR) and were constructed as previously described (14). The *Escherichia coli* strain used for transformation is X-blue ultra (Stratagene, La Jolla, CA).

2.2. PCR Primers

A set of oligonucleotide primers are designed and synthesized to amplify particular DNA fragments. Each primer contains the sequences necessary to amplify the fragment of interest from the source DNA; it also contains the 3' and 5' sequences complementary to the ends of the other DNA fragments chosen to be joined. Because the amplified fragments serve as megaprimers and templates in subsequent PCRs, the sequence and length of the overlap region is also critical. The primers are designed to achieve annealing at temperatures $\geq 60^{\circ}\text{C}$ and to avoid hairpins, self-priming, and primer-dimer formation. Highly conserved sequences are selected as primer sites based on available HIV-1 sequence data. Primers designed for amplifying the 5' end of the LTR region of HIV-1 pNL4-3 (GenBank M19921) are: NL5F (forward) 5'-TGG AAG GGC TAA TTT GGT CCC AAA AAA G-3'; NL5R (reverse) 5'-CAT CTC TCT CCT TCT AGC CTC C-3'. The primers for the 3' end of the LTR region of HIV-1 pNL4-3 are: NL3F (forward) 5'-CAC AAG TAG CAA TAC AGC AGC TAC CAA TGC-3'; and NL3R (reverse) 5'-TGC TAG AGA TTT TCC ACA CTG AC-3'.

Primers for amplifying the 5' half of the HIV-1 cDNA genome (5 kb) derived from the plasma viral RNA of patient WC1-PR are: FGF60 (forward) 5'-CAG ACC CTT TTA GTC AGT GTG GAA AAT C-3'; and FGR53 (reverse) 5'-GTC TAC TTG TGT GCT ATA TCT CTT TTT CCT CC-3'. The primers for amplifying the 3' half genome (5 kb) are: FGF46 (forward) 5'-GCA TTC CCT ACA ATC CCC AAA G-3'; and FGR95 (reverse) 5'-GGT CTA ACC AGA GAG ACC CAG TAC AG-3'. Primers were prescreened for optimal sensitivity and efficiency.

2.3. Polymerase Chain Reaction (PCR)

High-fidelity PCRs are performed using *Pfu* DNA polymerase (Stratagene). AmpliMax PCR[®] Gems (Perkin-Elmer, Foster City, CA) are used to minimize undesired primer interaction. *DpnI* endonuclease (BioLabs, Beverly, MA) is added to the reaction to digest the parental DNA template after PCR. The PCR products are analyzed by gel electrophoresis, and the desired fragments are isolated and purified using the QIAquick Gel Extraction kit (Qiagen, Chatsworth, CA).

Long asymmetric PCR amplification and construction use both Tth XL (Perkin-Elmer) and *Pfu* polymerases. All of the reactions are run in a Perkin-Elmer GeneAmp 9600 thermal cycler (Perkin Elmer). The PCR products are assayed on 0.8% SeaKem GTG agarose gel (FMC, Rockland, ME).

2.4. Cloning and Colony Screening

The purified DNA fragments are directly ligated into a phagemid TA vector (pCR II plasmid) by using the Original TA Cloning Kit (Invitrogen, Carlsbad, CA). The plasmid is used to transform X-blue ultra *E.coli* competent cells (Stratagene). After transformation, white colonies from X-Gal plates are verified by restriction digestion and partial DNA sequencing.

2.5. DNA Transfection and Virus Replication Assays

Plasmid DNA (10 µg of each) encoding HIV-1 genomes are transfected into normal human dermal fibroblast cell (NHDF 710, Clonetics, San Diego, CA) by the calcium phosphate precipitation method and Mammalian Transfection kit (Stratagene). Supernatant are harvested, clarified, and used to infect peripheral blood mononuclear cells (PBMCs) using standard methods.

HIV-1 reverse transcriptase assays are performed using the Reverse Transcriptase Assay, Nonradioactive kit (Boehringer-Mannheim, Indianapolis, IN) following the manufacturer's instructions; HIV-1 p24 antigen capture assays are performed using the Alliance[™] HIV-1 p24 ELISA kit (Dupont, Boston, MA).

2.6. Sequence Analysis

Plasmid DNA is extracted and purified using the Qiagen Max kit (Qiagen). The DNA templates are sequenced using fluorescent dye-labeled terminators and an Applied Biosystems DNA sequencer (Applied Biosystems, Foster City, CA). Sequence data are analyzed using University of Wisconsin Genetics Computer Group Software Packages on a Sun computer. The full sequences of two infectious HIV-1 clones, FG9012-38 and FG9012-40, have been contributed to GenBank under accession numbers AF003887 and AF003888.

3. Methods

3.1. Construction of Chimeric HIV-1 Clones

To construct full-length infectious clones of HIV-1, we first used a method we previously developed to clone complete HIV-1 genomes directly from plasma viral RNA using long reverse transcription and PCR (RT-PCR) (14,15). Long cDNAs were

derived from the HIV-1 RNA genome. Then, two 5-kb DNA fragments containing the 5' and 3' half-genomes were amplified and cloned. Because of the complex replication cycle of retroviruses, the sequence of the virus that appears in the RNA form, the form found in virions, differs in the LTR region from that found in the DNA or proviral form, which is found integrated in cells. For this reason, the HIV-1 genome cloned from virion-associated RNA has incomplete LTRs; it needs to have complete LTRs attached for clones to be infectious as DNA. Therefore, DNA fragments from proviral 5' and 3' LTR regions were linked to the 5' and 3' ends of the constructs, respectively. DNA fragments were obtained from two sources in the experiments described here: cloned cDNA reverse transcribed from plasma viral RNA (plasmids FG901-18 and FG902-12) and cloned proviral DNA (plasmid pNL4-3, which contains a full-length HIV-1 LTR). Construction of clones using conventional restriction and ligation methods was not feasible in this case because of the absence of convenient restriction sites in these HIV-1 genomes. We, therefore, developed a strategy to assemble the full-length HIV-1 clones by using PCR-mediated recombination (*see Fig. 1*).

3.2. High Fidelity Long PCR

The four DNA fragments (A, B, C, D) encompassing the full-length HIV-1 genome were first amplified individually in a high fidelity PCR by using *Pfu* DNA polymerase. Each of the fragments shared a segment of an overlapping sequence with one or two other fragments.

The optimal conditions for using *Pfu* polymerase in our high fidelity PCR were evaluated (*see Notes 1 and 2*). The final conditions for the amplification using *Pfu* polymerase were: 10 cycles of PCR at 94°C for 15 s, 55°C for 45 s, and 72°C for 2–10 min in a final volume of 100 μ L with 0.2 μ g plasmid DNA template, 2.5 U of *Pfu* polymerase, 20 pmol of each primer, 20 μ M each dNTP, and 0.5X *Pfu* buffer (containing 1.0 mM Mg²⁺).

Because of the high initial concentration of template used, to eliminate the carryover contamination of parental templates, the restriction endonuclease *DpnI* was added to the reaction after PCR. *DpnI* is specific for methylated and hemimethylated DNA; it was used to digest parental plasmid DNA and to select for amplified DNA.

Using the optimized PCR conditions, four fragments were amplified from corresponding plasmids. They are: fragment A, a 792-bp fragment amplified from HIV-1 pNL4-3, which contains the entire 5'LTR; fragment B, a 5-kb 5' half genome from plasmid clone FG901-18; fragment C, a 5 kb 3' half genome from plasmid clone FG902-12 and fragment D, a 750 bp fragment from 3' end of HIV-1 pNL4-3, which contains the entire 3' LTR (*see Fig. 2*). All four fragments to be assembled were gel purified and quantitated individually.

3.3. Asymmetric PCR Recombination

After the first PCR, an asymmetric PCR, which favors single-strand extension in the desired direction, was used in order to construct the full-length molecule (*see Fig. 1*). A combination of *Pfu* and *rTth* polymerase was used in the asymmetric PCR amplification and construction to take advantage of the *Pfu* polymerase's precision as well as to reduce PCR time, increase PCR yield, and enable us to use TA vector clon-

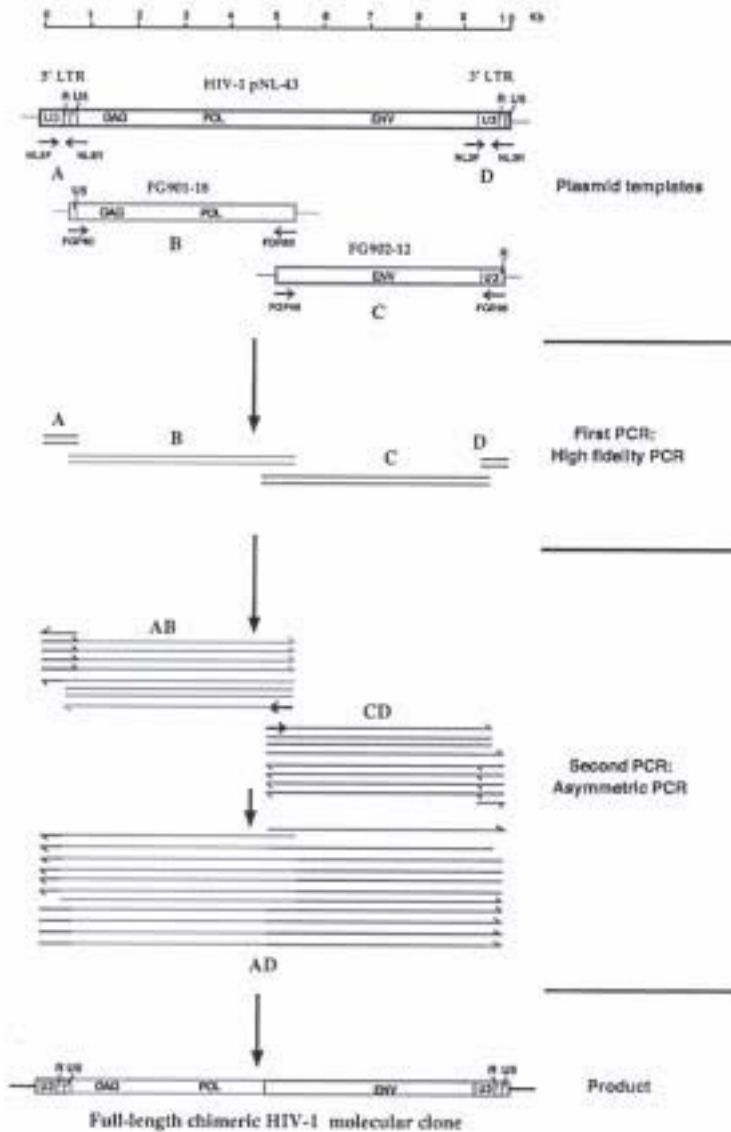


Fig. 1. Schematic diagram of the PCR-mediated recombination strategy used in constructing HIV-1 chimeric infectious clones. The segments of the HIV-1 genome in plasmid vectors are shown as rectangles with the plasmid name on the top; each plasmid is color coded. Synthetic oligonucleotide primers are shown as single strands, with full arrowheads indicating the direction. The PCR-amplified products are shown as two paired strands and are named by capital letters A–D and color coded according to the plasmid template. A and D are both derived from plasmid pNL4-3 (15) and, therefore, are depicted in one color. The intermediate PCR products are shown as single strands, with half arrowheads indicating the direction in which each strand can act as a primer for DNA polymerase (the 5'-3' direction).

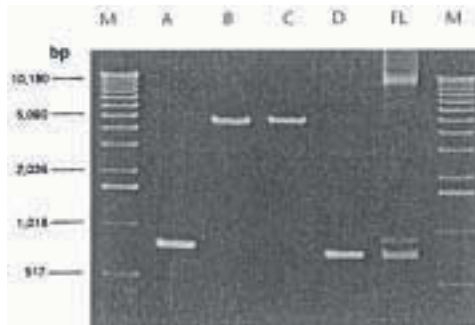


Fig. 2. Agarose gel analysis of DNA fragments amplified by PCR. Lane A shows a 792-bp fragment amplified from HIV-1 pNL4-3, which contains the entire 5' LTR; lane B, a 5 kb 5' half genome from plasmid clone FG901-12; fragment C, a 5-kb 3' half genome from plasmid clone FG902-12; and fragment D, a 750 bp fragment from the 3' end of HIV-1 pNL4-3, which contains the entire 3' LTR. Lane FL shows the full-length 10 kb HIV-1 fragment constructed by PCR-mediated recombination. Lane M depicts a set of molecular weight standards (1 kb ladder; GIBCO BRL/Life Technologies). Numbers on the left indicate the molecular weight (bp) of some of the bands in lane M. The source of the plasmids is given in the text.

ing. The four PCR-amplified fragments (A, B, C, and D) were mixed in an asymmetric ratio with two synthetic oligonucleotide primers, which serve as internal primers. Fragments A and D were added in excess and served as external flanking megaprimers, whereas the quantity of the two internal oligonucleotide primers was very limited. Conditions for asymmetric PCR were evaluated (*see* **Notes 2** and **3**). The optimal conditions were: 2 μg of each external flanking fragments (A and D, megaprimers), 0.1 μg of each 5-kb internal fragment (B and C) produced by high fidelity PCR were mixed with 0.2 pmol of synthetic oligonucleotide primers (FGR53 and FGF46). The mixture was subjected to 20 cycles of PCR in a volume of 100 μL containing 2.5 U of *rTth* polymerase, XL, 0.004 U of *Pfu* polymerase, 10 μM each of dNTPs, 1.2 mM $\text{Mg}(\text{OAc})_2$, and 1X XL PCR buffer II. The PCR cycling parameters were 94°C for 15 s, and 72°C for 8 min.

The resulting PCR produced an asymmetric single-strand amplification of two 5-kb halves (AB and CD) of the genome as the majority forms. In the subsequent PCR cycles, these intermediate single-stranded DNAs, each representing half of the genome (AB and CD) overlapped and annealed to complementary strands of each other. The annealed strands then served as megaprimers for one another and were extended by PCR to form full-length molecules (*see* **Fig. 2**). It was not necessary to purify the intermediate products. As soon as the first 10-kb full-length strand was synthesized, it too was PCR-amplified using the flanking primers A and D.

The 10-kb PCR-amplified fragment was detected by gel electrophoresis as a major band (*see* **Fig. 2**). The correctly assembled DNA was a major species and could be easily separated by electrophoresis from the portions of the genome serving as parental templates. The full-length product was then isolated and directly cloned into a TA

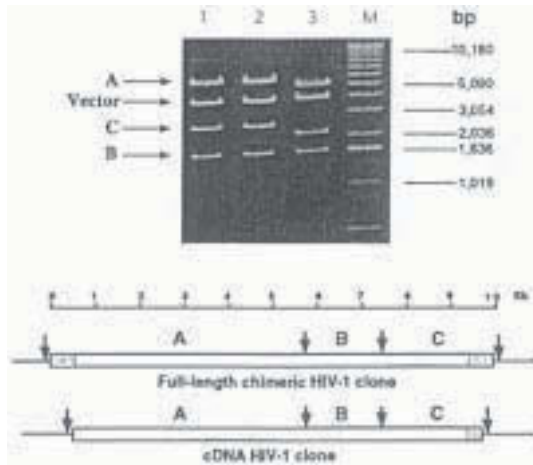


Fig. 3. Restriction digestion of constructed full-length HIV-1 molecular clones. Lanes 1 and 2 are two full-length chimeric HIV-1 molecules constructed by PCR-mediated recombination, with intact LTRs, and cloned into the TA vector (Invitrogen). Lane 3 is a full-length cDNA clone generated directly from plasma HIV-1 RNA as previously described (13) and cloned in the TA vector; it has partial LTRs. Plasmid DNA was digested with EcoRI. The bottom part of the figure is a schematic diagram of EcoRI restriction sites (shown as arrows) on cloned HIV-1 genomes (shown as rectangles) and vectors (shown as single strands); the shadowed parts are HIV-1 LTR regions. Lane M depicts a molecular weight standard (1 kb ladder; Gibco-BRL/Life Technologies).

vector. The structure of the chimeric DNA product was verified by restriction digestion patterns (see Fig. 3) and partial DNA sequencing. Most of the clones screened contained complete or nearly complete HIV-1 genomes with the proper length and organization; no unwanted deletions or insertions in overlap regions were found.

3.4. Sequence Fidelity of Clones

To determine the sequence fidelity of full-length clones constructed by PCR-mediated recombination, two of the constructed full-length infectious clones (clones FG9012-38 and FG9012-40, GenBank accession numbers AF003887 and AF003888) and their parental plasmids were entirely sequenced. Only one mutation was found in clone 38 and none in clone 40. The PCR-introduced error was a transition (A to G), and was located in the untranslated region. No other mutations including frame shifts, deletions, or insertions were found.

The cumulative error rate after PCR amplification of two complete genomes was 0.005%, or a total of one base change detected in greater than 20 kb of plasmid DNA. This error rate is much lower than the previously reported rates of 0.14-0.35% obtained by using thermostable enzymes (9,10,16). The high precision in the data reported here most likely resulted from using high fidelity polymerase and controlled PCR conditions aimed at producing a low error rate (see Notes 2 and 4).

3.5. Infectivity of Constructed Clones

The full-length chimeric molecules constructed from plasma-derived HIV-1 RNA and proviral DNA-derived LTRs were tested for infectivity by transfection into human dermal fibroblast cells. Controls included the transfection of pNL4-3 plasmid DNA and the TA vector plasmid DNA. Approximately half of the constructed clones tested (5 of 9 clones) produced virus particles as determined by both HIV-1 p24 antigen and reverse transcriptase assays. These particles were confirmed to be infectious virus by performing either cocultivation (**14**) of the transfected cells with PBMCs or by inoculation of donor PBMCs with cell-free supernatant from transfected cells. No virus was detected from the control experiments using vector plasmid DNA only.

4. Notes

1. The optimal concentration ranges of dNTPs (10–100 μM each dNTP), DNA templates (50–2000 ng per reaction), PCR extension times (0.5–3 min per kb DNA to be amplified) and numbers of PCR cycles (5–25 cycles) were evaluated (data not shown). The final conditions for the amplification using Pfu polymerase were: 10 cycles of PCR at 94°C for 15 s, 55°C for 45 s, and 72°C for 2–10 min in a final volume of 100 μL with 0.2 μg plasmid DNA template, 2.5 U of Pfu polymerase, 20 pmol of each primer, 20 μM each dNTP, and 0.5X Pfu buffer (containing 1.0 mM Mg^{2+}). To construct infectious molecular clones using PCR-mediated recombination, the PCR polymerase used needs to have the lowest error rate possible to minimize the chance of producing unwanted mutations. Pfu polymerase has the lowest error rate of all known thermophilic polymerases (**17,18**). Unlike some polymerases (e.g., *Taq*, *Tth*), Pfu produces perfectly blunt ends of PCR products. The blunt end avoids the introduction of undesired adenosine residues onto the 3' ends of the PCR products, which may lead to insertion of an extra nucleotide onto the full-length molecule during subsequent PCR.
2. As an additional consideration, the PCR error rate is related to the number of PCR cycles. To minimize the number of PCR cycles for DNA synthesis, the highest concentration of template plasmid consistent with amplification should be employed. Because PCR error rates increase with the Mg^{2+} and dNTP concentrations (**17**), the lowest Mg^{2+} and dNTP concentrations compatible with amplification should be used. Conditions for asymmetric PCR, minimum Mg^{2+} requirement, ratios of each fragment and primers, and the necessary length of overlap were investigated to develop the optimal conditions.
3. The length of the overlap sequence influences the efficiency of PCR and its yield. Therefore, it is important to define the required length of overlap, particularly when the overlap region between two fragments includes basepairs that are not identical. To determine the minimum length of overlap required for construction of 10-kb clones, we tested overlapping lengths ranging from 50 bp up to 690 bp, between fragments B and C, which have almost identical overlapping sequences (99.43% similarity). Because an overlap of 270 bp between B and C did produce 10-kb products, longer overlaps (560, 640, and 690 bp) resulted in a higher yield of full-length molecules. To increase the yield of complete products in PCR with shorter overlaps, more PCR cycles are needed, which may result in more errors being introduced. An overlap region of 50–100 bp per kb of DNA to be assembled appeared adequate for most long DNA constructions (data not shown). For overlapping regions with heterogeneous sequences, the optimal length of overlap needs to be determined and a longer overlapping sequence may be required.

In the overlap regions between fragments A and B, and C and D, there were 3.66 and 6.42% nucleotide differences respectively, yet fragments could still be assembled by using the relatively long overlapping (180 and 590 bp).

The effect of the concentration of megaprimers on production of complete clones was studied by using from 0.2 μg up to 2 μg of megaprimers (DNA fragments A and D) per 0.1 μg of templates (DNA fragments B and C) in PCRs. The production of 10-kb product was correlated with the concentration of megaprimers; high concentration of megaprimers greatly improved the yield of final products (data not shown). A high concentration of megaprimers may facilitate the formation of intermediate strands AB and CD in asymmetric PCR amplification and increase the yield of final products.

Because there were nucleotide differences between fragments A and B, and C and D, it was possible to determine which parental templates were included in the final full-length molecule. The sequence analysis indicated that the overlapping sequences between fragments A and B; and C and D in the full-length molecule were derived from HIV-1 pNL4-3 (A and D). This result confirmed our hypothesis that, in most cases, the two 5-kb halves (AB and CD) of the genome were first synthesized as single-stranded intermediates that then formed a full-length molecule in subsequent PCR cycles (see **Fig. 1**).

4. PCR-mediated recombination provides a powerful method of recombining DNA sequences from any source without reliance on restriction sites. What is required to perform PCR-mediated recombination is the sequence of the 3' and 5' overlapping regions of the desired PCR products. This method may be extended to include construction of chimeras between any DNA fragments lacking sequence homology. Such chimeras may be constructed by introducing overlapping sequences to one of the fragments (**19**). To ensure that unwanted mutations have not been introduced into the clones constructed by this method, each clone should be sequenced. Our results demonstrate that by using a high-fidelity polymerase and highly controlled PCR conditions, the PCR-introduced error rate can be greatly minimized.

This new procedure may be used to construct infectious chimeras of HIV or SIV for studies of vaccines and pathogenesis. Moreover, the method is designed to exchange viral genes at precise boundaries to study individual gene products from different HIV genomes. It can also be used to construct expression vectors for production of specific proteins or delivery vectors for gene transfer and gene therapy. Finally, the technique described here provides a versatile tool to transfer genes or gene fragments from different sources for genetic investigation and engineering.

Acknowledgments

This work was originally published by Fang et al. (1999) in *Nature Medicine* (**20**). The authors thank Ellen Shippey and Anne Klugo for help in preparing the manuscript and the Wadsworth Center Molecular Genetics Core Laboratory for oligonucleotide synthesis for PCR. This work was supported in part by grants from the National Institutes of Health (RO1A133334 and U01AI35004 from the National Institute of Allergy and Infectious Disease and the National Institute on Drug Abuse).

References

1. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.

2. Meyerhans, A., Vartanian, J. P., and Wain-Hobson, S. (1990) DNA recombination during PCR. *Nucl. Acids Res.* **18**, 1687–1691.
3. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J., and Pease, L. R. (1989) Engineering hybrid gene without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61–68.
4. Horton, R. M., Cai, Z. L., Ho, S. N., and Pease, L. R. (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* **8**, 528–535.
5. Yolov, A. A. and Shabarova, Z. A. (1990) Constructing DNA by polymerase recombination. *Nucl. Acids Res.* **18**, 3483–3486.
6. Klug, J., Wolf, M., and Beato, M. (1991) Creating chimeric molecules by PCR directed homologous DNA recombination. *Nucl. Acids Res.* **19**, 2793.
7. Gu, H., Planas, J., Gomez, R., and Wilson, D. J. (1991) Full length mouse glycophorin gene constructed using recombinant polymerase chain reaction. *Biochem. Biophys. Res. Commun.* **177**, 202–208.
8. Sandhu, G. S., Aleff, R. A., and Kline, B. C. (1992) Dual asymmetric PCR: one-step construction of synthetic genes. *Biotechniques* **12**, 14–16.
9. Eckert, K. A. and Kunkel, T. A. (1991) DNA polymerase fidelity and the polymerase chain reaction. *PCR Meth. Appl.* **1**, 17–24.
10. Mattila, P., Korpela, J., Tenkanen, T., and Pitkanen, K. (1991) Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase—an extremely heat stable enzyme with proofreading activity. *Nucl. Acids Res.* **19**, 4967–4973.
11. Coffin, J. M. (1995) HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **267**, 483–489.
12. Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., et al. (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183–188.
13. Myers, G. Korber, B., Foley, B., Jeang, K. T., Mellors, J. W., and Wain-Hobson, S. (1996) *Human Retroviruses and AIDS*. Los Alamos National Laboratory, Los Alamos, NM.
14. Fang, G., Weiser, B., Visosky, A., Townsend, L., and Burger, H. (1996) Molecular cloning of full-length HIV-1 genomes directly from plasma viral RNA. *J. AIDS* **12**, 352–357.
15. Fang, G., Burger, H., Grimson, R., Tropper, P., Nachman, S., Mayers, D., et al. (1995) Maternal plasma human immunodeficiency virus type 1 RNA level: a determinant and projected threshold for mother-to-child transmission. *Proc. Natl. Acad. Sci. USA* **92**, 12,100–12,104
16. Salminen, M. O., Koch, C., Sanders-Buell, E., Ehrenberg, P. K., Michael, N. L., Carr, J. K., et al. (1995) Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction. *Virology* **213**, 80–86.
17. Lundberg, K. S., Shoemaker, D. D., Adams, M. W., Short, J. M., Sorge, J. A., and Mathur, E. J. (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**, 1–6.
18. Flaman, J. M., Frebourg, T., Moreau, V., Charbonnier, F., Martin, C., Ishioka, C., et al. (1994) A rapid PCR fidelity assay. *Nucl. Acids Res.* **22**, 3259–3260.
19. Horton, R. M. (1995) SOEing together tailor-made genes. *Mol. Biotech.* **3**, 93–99.
20. Fang, G., Weiser, B., Visosky, A., Moran, T., and Burger, H. (1999) PCR-mediated recombination: a general method applied to construct chimeric infectious molecular clones of plasma-derived HIV-1 RNA. *Nat Med.* **5**, 239–242.

PCR Method for Generating Multiple Mutations at Adjacent Sites

Jiri Adamec

1. Introduction

Site-directed mutagenesis is a commonly used tool for identifying the role of specific amino acids in the structure and function of proteins. Various methods of *in vitro* mutagenesis have been described and are widely used for introducing modified coding sequences (1–7). In comparison, polymerase chain reaction (PCR)-based methods (1–6) are generally faster and more efficient than non-PCR-based methods (7). On the other hand, some PCR-based methods require two or more primers for each round of mutagenesis, whereas others need a single very long oligonucleotide or two round of PCR for introducing one mutation (5–7). These all can increase the cost of mutagenesis.

The PCR method for generating multiple mutations at adjacent sites is a two-step procedure. This can be very efficient and economical in those cases where a large number of nucleotide (amino acid) changes, deletions or insertions are to be programmed in a small region of the sequence. In the first step, a new, unique restriction site is introduced at the middle of or near to the part of DNA sequence to be changed without resulting in a change of the amino acid sequence. For this step, two PCR products cloned into a plasmid are used (*see Fig. 1*). The unique restriction site (URS3) is introduced into the plasmid by using overlapping oligonucleotide primers (P3 and P4) which contain sequences that complement the template DNA sequence at the 3'-end and each other at the 5'-end (*see Fig. 1A*). The sequences at the 5'-end of the primers contain the unique restriction site, which maintains the amino acid sequence of the expressed protein. Two PCR products are generated by using these internal overlapping primers (P3 and P4) paired with the corresponding external primers (P1 and P2). The external primers P1 and P2 cross the natural, unique restriction sites (URS1 and URS2) of the plasmid (*see Fig. 1A*). The product of PCR I or PCR II is then cleaved with restriction enzymes corresponding to URS1 and URS3 or URS2 and URS3, respectively. Digested fragments are separated on agarose gel electrophoresis,

A

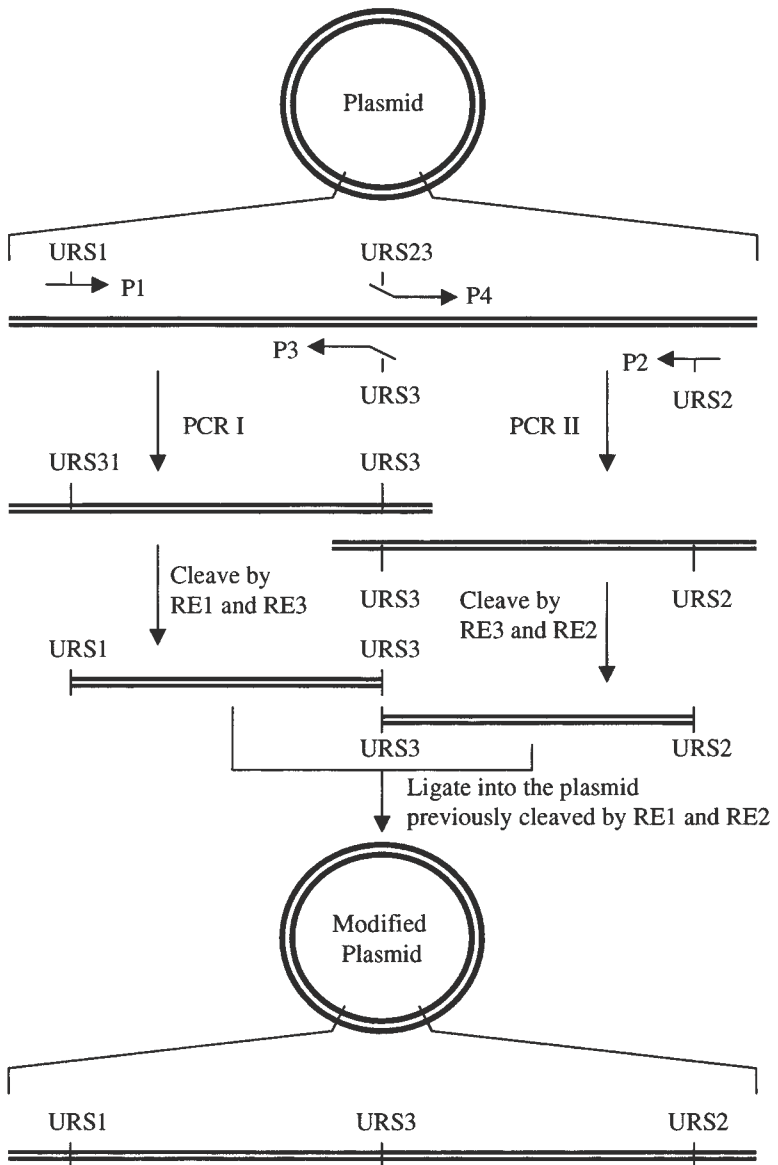


Fig. 1. Principle of PCR method for generating multiple mutations at adjacent sites. Oligonucleotide primers are represented by the arrows at their annealing sites in the target DNA sequences of plasmid. (A) The two internal primers P3 and P4 containing the unique restriction site URS3 anneal to opposite strands of the DNA template at the region to be modified. The external primers P1 and P2 cross the natural unique restriction sites URS1 and URS2, respectively. PCR I and PCR II are performed separately, their product cleaved by restriction enzymes (RE) corresponding to URS1 (RE1) and URS3 (RE3) or URS2 (RE2) and URS3

B

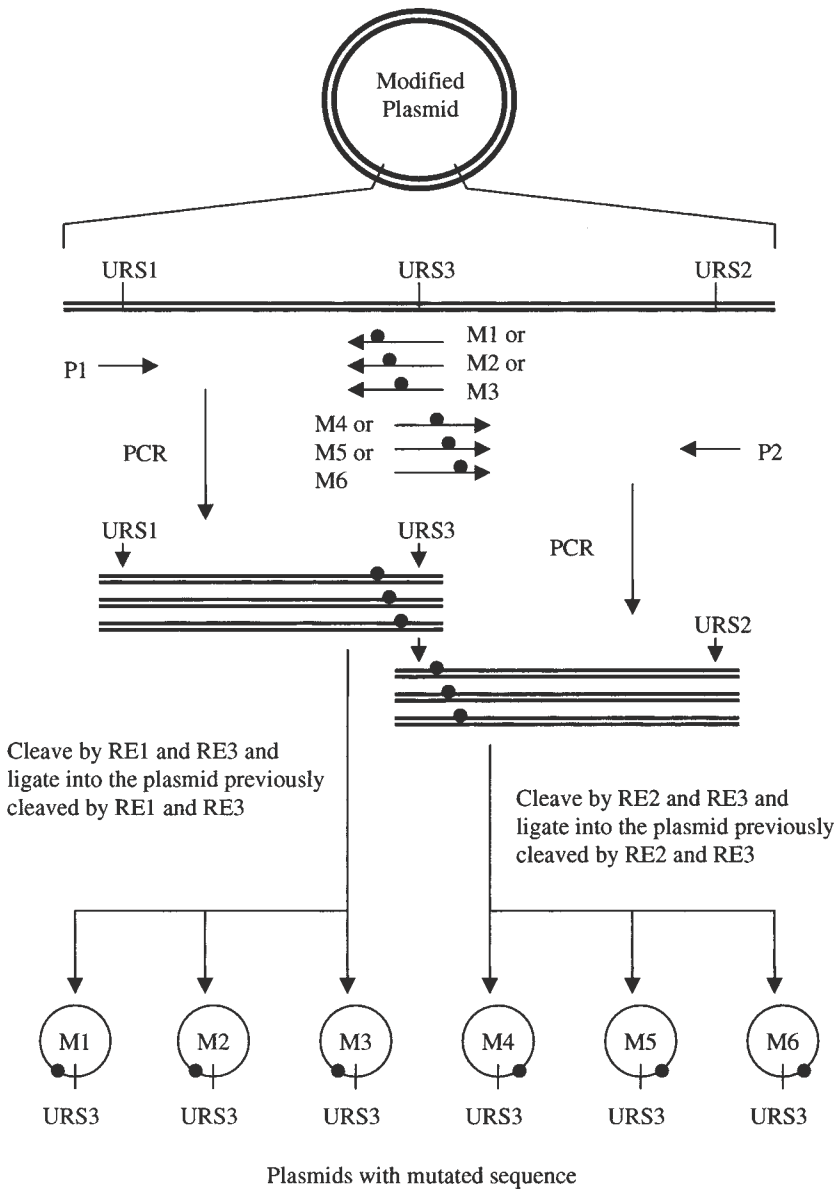


Fig. 1. (continued from opposite page) (RE3), respectively, and the fragments are ligated into the plasmid previously cleaved by RE1 and RE2 enzymes, to create new plasmid containing URS3. (B) To generate mutations, mutagenic primers containing URS3 are paired with external primers P1 (M1–M3, for generation of mutations at left side from URS3) or P2 (M4–M6, for generation of mutations at right side from URS3). The PCR fragments are then cleaved by appropriate restriction enzymes and used to reconstruct mutant sequences. Nucleotide changes are indicated as “●.”

extracted from gel slices, and ligated together into the original plasmid previously digested with restriction enzymes corresponding to URS1 and URS2 (see **Fig. 1A**). After transformation of *Escherichia coli* with ligated material, colonies are selected on LB plates (1% bacto tryptone, 0.5% bacto yeast extract, 1% NaCl and 1.2% agar) containing appropriate antibiotics. Plasmids are then isolated and tested for the presence of the URS3 site using a corresponding restriction enzyme, and those containing the URS3 site are used for the second step, which is the introduction of desired mutations.

In the second step, mutagenic primers are synthesized to contain a URS3 sequence at the 5'-end and a sequence introducing a mutation at the 3'-end (see **Fig. 1B**). In PCRs, mutagenic primers are paired with P1 primer (primers M1–M3) or with P2 primer (M4–M6), respectively (see **Fig. 1B**). Amplified fragments, designated here as F1–F6, are digested by restriction enzymes corresponding to URS1 and URS3 (F1–F3) or by the enzymes corresponding to URS2 and URS3 (F4–F6), respectively. They are subsequently separated on an agarose gel, isolated, and individually ligated into the plasmid from first step previously digested with enzymes corresponding URS1 and URS3 for fragments F1–F3 or URS2 and URS4 for F4–F6, respectively (see **Fig. 1B**). After transformation, isolated plasmids are tested for the presence of the correct mutation by sequencing.

Using this approach, we were able to generate a large number of mutations to investigate the function of the α -subunit of mitochondrial processing peptidase (8,9). Moreover, the very high efficiency of introducing mutations (90–95%) greatly reduced the number of colonies that needed to be screened, since we only had to test 2–3 colonies for each mutation.

2. Materials

1. PCR machine.
2. Plasmid containing DNA sequence to be mutated.
3. Oligonucleotide primers.
4. High-fidelity *Pfu* DNA polymerase and 10X Reaction Buffer (Stratagene, La Jolla, CA).
5. PCR Nucleotide Mix (Promega, Madison, WI).
6. QIAquick PCR Purification Kit (Qiagen, Valencia, CA).
7. Low-melting agarose (e.g., SeaPlaque GTG agarose, FMC BioProducts, Rockland, ME).
8. Reagents and apparatus for agarose gel electrophoresis.
9. AgarACE enzyme (Promega, Madison, WI).
10. Restriction enzymes and buffers (e.g., from New England Biolabs, Beverly, MA).
11. DNA ligase and buffer (e.g., from New England Biolabs).
12. *E. coli* strain for transformation (e.g., MAX Efficiency DH5 α Competent Cells, Life Technologies, Grand Island, NY).
13. LB medium and LB agar (BIO 101, Vista, CA).
14. Antibiotic(s) (e.g., from Life Technologies).
15. QIAprep Spin Miniprep Kit (Qiagen).

3. Methods

3.1. Introduction of Unique Restriction Site

3.1.1. Primer Design

Primers P1 and P2 should be approx 20 nt (nucleotides) in length with the sequence of the natural unique restriction site (URS) located in the middle. The choice of URS depends on the plasmid used. The ideal distance between primer P1 and P3 or P2 and P4 is about 150 bps (basepairs). The primers P3 and P4, which introduce the new URS into the sequence at 5'-end, must be followed by certain number of any nucleotides. The number of these nucleotides varies and depends on the URS used as the restriction enzyme, because recognizing this site requires nucleotide sequence on both sites of URS to be efficient (*see Note 1*). In this process a critical factor is the selection of a new URS. There are four main criteria, namely:

1. the site is not present in the original plasmid sequence;
2. the site is about the middle of region to be mutated;
3. nucleotide changes do not affect amino acid composition;
4. the restriction enzyme recognizing this site is commercially available.

Hence, in order to fulfill this strategy, we generate a silent mutation as previously described (9) and highlighted in **Fig. 2A**. To determine possible silent mutations that can be introduced into the region of interest, we recommend using computer programs available on the Internet (e.g., WebCutter at <http://www.firstmarket.com/cutter/cut2.html>).

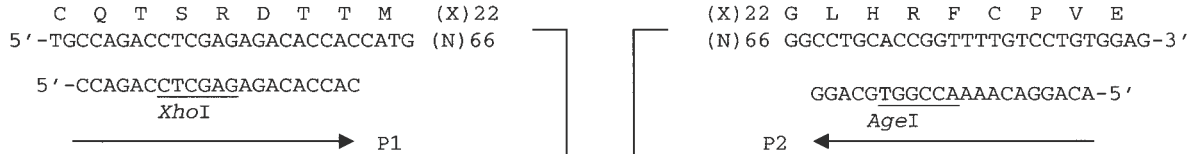
3.1.2. PCR, purification, and cloning

1. Assemble two 50 μ L PCRs (*see Note 2*):

	PCR I	PCR II
Template DNA (50 ng/ μ L)	2 μ L	2 μ L
10X Reaction Buffer	5 μ L	5 μ L
10X dNTP (2.5 mM each)	5 μ L	5 μ L
Primer P1 (20 μ M)	1 μ L	—
Primer P2 (20 μ M)	—	1 μ L
Primer P3 (20 μ M)	1 μ L	—
Primer P4 (20 μ M)	—	1 μ L
H ₂ O	35 μ L	35 μ L
<i>Pfu</i> DNA polymerase	1 μ L	1 μ L

2. Use following steps to perform PCR:
 - a. 94°C-5 min (initial denaturation);
 - b. 94°C-30 s (denaturation);
 - c. 50°C-30 s (annealing);
 - d. 72°C-1 kb/min (extension);
 - e. Repeat 30 times steps 2–4 (30 cycles);
 - f. 72°C-10 min (final extension step).

A



H P R L T D E E I E M T R M A V Q F E L E D L N M
 CACCCCGCCTGACAGATGAGGAAATTGAGATGACGAGGATGGCTGTTTCAGTTTGAACCTTGAGGACCTGAACATG

↓ Find restriction sites which can be introduced by silent mutation.
 Use Webcutter at "<http://www.firstmarket.com/cutter/cut2.html>".

CCGCGG
 Cfr42I

ACGCGT
 MluI

TGTACA
 BsrGI

GAGCTC
 SacI

↓ Select restriction site.

ACGCGT
 MluI

↓ Design primers.

TAACTCTACTGCGCAATACCGA-5'
 P3 ←

5'-GAGATGACGCGTATGGCTGTT
 → P4

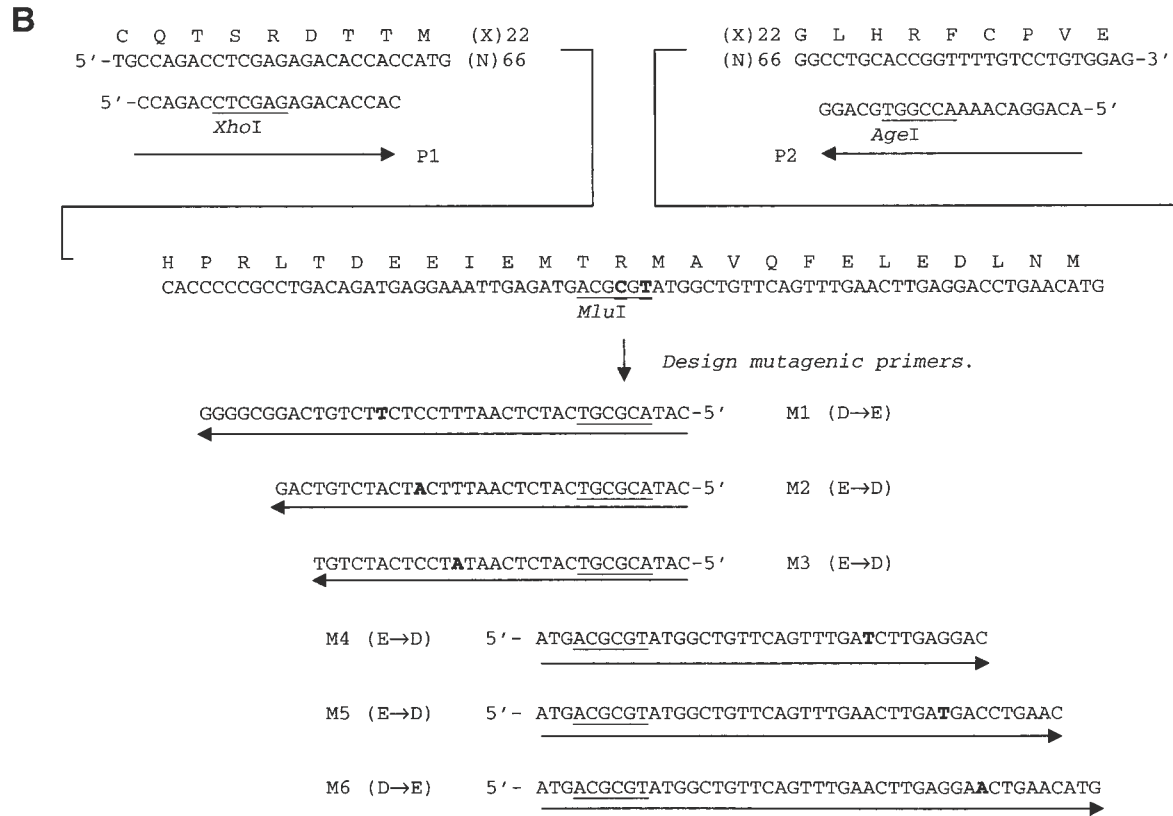


Fig. 2. Primer design strategy for (A) introducing new, unique restriction site and (B) generating mutations (modified from ref. 9). Underlined sequences indicate restriction enzyme sites. Bold letters indicate point mutations.

3. Clean both PCR products using QIAquick PCR Purification Kit and protocols supplied by vendor. Use 50 μL of H_2O to elute DNA from cartridges.
4. Cleave products of PCR I and PCR II, and plasmid with corresponding restriction enzymes (*see Note 3*).
5. Purify digested PCR products and plasmid (*see Note 4*).
6. Ligate purified PCR fragments and plasmid DNA from previous step. For ligation, use the molar ratio 3 : 3 : 1 of PCR I fragment : PCR II fragment : plasmid. Calculate DNA concentration of samples using: $C \text{ (pmol}/\mu\text{L}) = (75 \times A_{260} \times d) / (\text{bp} \times l)$, where A_{260} is absorbance at 260 nm, d is dilution factor, bp is length of DNA in basepair, and l is cuvet width in cm. For ligation, add the following components to a microcentrifuge tube: cleaved PCR I fragment, PCR II fragment, plasmid, 2 μL 10X ligation buffer, 1 μL DNA ligase and H_2O to 20 μL . Incubate reaction overnight at room temperature.
7. Transform appropriate *E. coli* strain using the following protocol (*see Note 5*). Mix ligation reaction from previous step with 100 μL of competent cells and keep on ice for 30 min. Incubate the mixture at 42°C for 1 min, transfer back on ice for 1 min and then add 0.5 mL LB medium. Place on shaker and incubate an additional 1 h at 37°C. Spin cells down (20 s at maximum speed), discard supernatant and resuspend cells gently in 200 μL LB medium. Plate on LB agar with appropriate antibiotic and let cells grow overnight at 37°C.
8. Place individual colonies in 4 mL LB medium containing appropriate antibiotic and grow overnight at 37°C. Isolate plasmids using a QIAprep Spin Miniprep Kit following manufacture's protocol and cleave them with restriction enzyme corresponding to introduced restriction site URS3. Sequence positive plasmids (colonies that are cleaved with enzyme) and use for next steps.

3.2. Mutations Generation

1. To design mutagenic primers follow the protocols outlined in **Subheading 3.1.1**.
2. For introduction of mutations on the left side from URS3, use the set up and conditions for PCR I described in **steps 1** and **2** except the mutagenic primer is used instead of primer P3. Similarly, to introduce mutations on the right side from URS3 use set up and conditions for PCR II, but use the mutagenic primer instead of primer P4 (*see Fig. 2B*).
3. Use the same protocol for purification, cleavage, and ligation as described in **steps 3–6**. Digested PCR fragments are then individually cloned into the cleaved plasmid, obtained from the previous step (for introduction of mutations on left side of URS3 use plasmid digested with restriction enzymes corresponding to URS1 and URS3, for mutations on right side, plasmid digested with restriction enzymes corresponding to URS2 and URS3). In the ligation reaction, combine the PCR fragment:plasmid in 3:1 molar ratio.
4. After transformation, plate cells on LB agar with appropriate antibiotic and let cells grow at 37°C overnight. Place two individual colonies from each transformation into 4 mL LB medium containing appropriate antibiotic and grow overnight at 37°C.
5. Isolate plasmids using QIAprep Spin Miniprep Kit and sequence cloned fragments to confirm the presence of the mutations and to verify that no undesired changes occurred in the sequence during the PCR process.

4. Notes

1. More detailed information about the cleavage close to the end of DNA fragments could be found in (**10**). Usually, it is enough to use 6 nt (e.g., AATAAT).

2. To decrease the number of undesired mutations generated by PCR process, we have used high-fidelity *Pfu* DNA polymerase and higher concentrations of template DNA (approx 100 ng). If the PCR machine is without a heated lid, we used 50 μ L of mineral oil on the top of each reaction mixture.
3. PCR products and plasmid DNA have to be cleaved by two different enzymes. Some enzymes are compatible (cut DNA with high efficiency using same buffer composition and conditions) and can be used simultaneously. If this is not the case or if one of the enzymes has low activity in the buffer used, we recommend that the DNA is cleaved separately with an additional purification step (use QIAquick PCR Purification Kit) between these two cleavages.
4. Several methods are available for purification of both the PCR products and plasmids including electroelution, Freeze-Squeeze (Bio-Rad, Hercules, CA), GeneClean (Bio 101, La Jolla, CA), etc. Below, we describe the method we used in our lab, because this was fast, versatile (e.g., does not depend on DNA size) and DNA recovery is more than 95%. Both cleaved PCR products and plasmid are loaded onto 1% low-melting temperature agarose gel and run. If the volume of the sample is too large, use SpeedVac to decrease volume. Excise DNA fragment from the gel with a clean scalpel and transfer to a 1.5-mL microcentrifuge tube. Incubate tube for 10 min at 72°C (use waterbath) with occasional mixing. Transfer tube to preheated 42°C water bath quickly, add 3.5 μ L AgarACE after 2 min, and mix well. Incubate for an additional 45 min. Purify DNAs by QIAquick PCR Purification Kit following the manufacture's instruction. Use 50 μ L of H₂O to elute DNAs.
5. For routine cloning we use MAX Efficiency DH5 α Competent Cells as they lack endonuclease activity, accept large plasmid and increase insert stability owing to mutations in genes *endA1*, *deoR*, and *recA1*, respectively.

Acknowledgment

The author would like to thank Stephen Naylor for discussions and critical reading of the manuscript.

References

1. Chen, B. and Przybyla, E. A. (1994) An efficient site-directed mutagenesis method based on PCR. *BioTechniques* **17**, 657–659.
2. Sang, N., Condorelli, G., De Luca, A., MacLachlan, T. K., and Giordano, A. (1996) Generation of site-directed mutagenesis by extra long, high-fidelity polymerase chain reaction. *Analyt. Biochem.* **233**, 142–144.
3. Tomic, M., Sunjevaric, I., Savtchenko, E. S., and Blumenberg, M. (1990) A rapid and simple method for introducing specific mutations into any position of DNA leaving all other position unaltered. *Nucl. Acids Res.* **18**, 1656.
4. Barik, S. (1993) Site-directed mutagenesis by double polymerase chain reaction: megaprimer method. *Meth. Mol. Biol.* **15**, 277–286.
5. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59.
6. Zhong, D. and Bajaj, S. P. (1993) A PCR-based method for site specific domain replacement that does not require restriction recognition sequences. *BioTechniques* **15**, 874–878.
7. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Nat. Acad. Sci. USA* **82**, 488–492.

8. Striebel, H.-M., Rysavy, P., Adamec, J., Spizek, J., and Kalousek, F. (1996) Mutational analysis of both subunits from rat mitochondrial processing peptidase. *Arch. Biochem. Biophys.* **335**, 211–218.
9. Adamec, J. and Kalousek, F. (1999) PCR method for generating multiple mutations at adjacent sites. *Folia Microbiol.* **44**, 11–14.
10. Moreira, R. F. and Noren, C. J. (1995) Minimum duplex requirements for restriction enzyme cleavage near the termini of linear DNA fragments. *BioTechniques* **19**, 58–59.

A Fast Polymerase Chain Reaction-Mediated Strategy for Introducing Repeat Expansions into CAG-Repeat Containing Genes

Franco Laccone

1. Introduction

Since their first description in 1991 (1), CAG-disease causing genes are increasing in number. Up to date, there are at least nine genetic diseases caused by CAG expansions. The creation of transgene and knock-in mice with CAG expansions is an useful tool for understanding the pathological mechanisms of the corresponding diseases, which could lead to therapeutic target(s) for the diseases. However, owing to the short life expectancy of the mice and to the low expression levels of transgenes, it is necessary to introduce CAG expansions larger than that in humans in order to elicit a pathological phenotype within the life of mice models. Naturally occurring “huge” CAG expansions (>100–150 CAGs), which could induce disease phenotype in the mice, are very seldom. In vitro synthesis of isolated CAG repeats have already been described (2,3). Most of these methods, however, require further cloning steps and often contain some flanking extraneous sequences. Here, the author describes a fast and simple way for expanding/introducing CAG repeats (or other repeats!) without altering the flanking 5' and 3' sequences of the gene of interest. This method was successfully employed for expanding the CAG repeat of the MJD/SCA3 gene (4). **Fig. 1** outlined the strategy of this method. Two independent polymerase chain reactions (PCRs) amplify the target gene from 5' to the CAG repeat region (PCR I) and from the CAG repeat to the 3' region (PCR II) of the gene. The amplicons of PCR I and PCR II will be then mixed, elongated and then a third PCR will be carried out with the two most “outsider” primers. We used this strategy for elongating the CAG of the MJD/SCA3 gene from 22 up to 138 CAG repeats (*see Figs. 2 and 3*). This method can be used for elongating different repeats in different genes or even to insert and elongate any simple or complex repeat into a DNA sequence. However, it is not possible in this chapter to give specific conditions for each applications. It is important to adapt

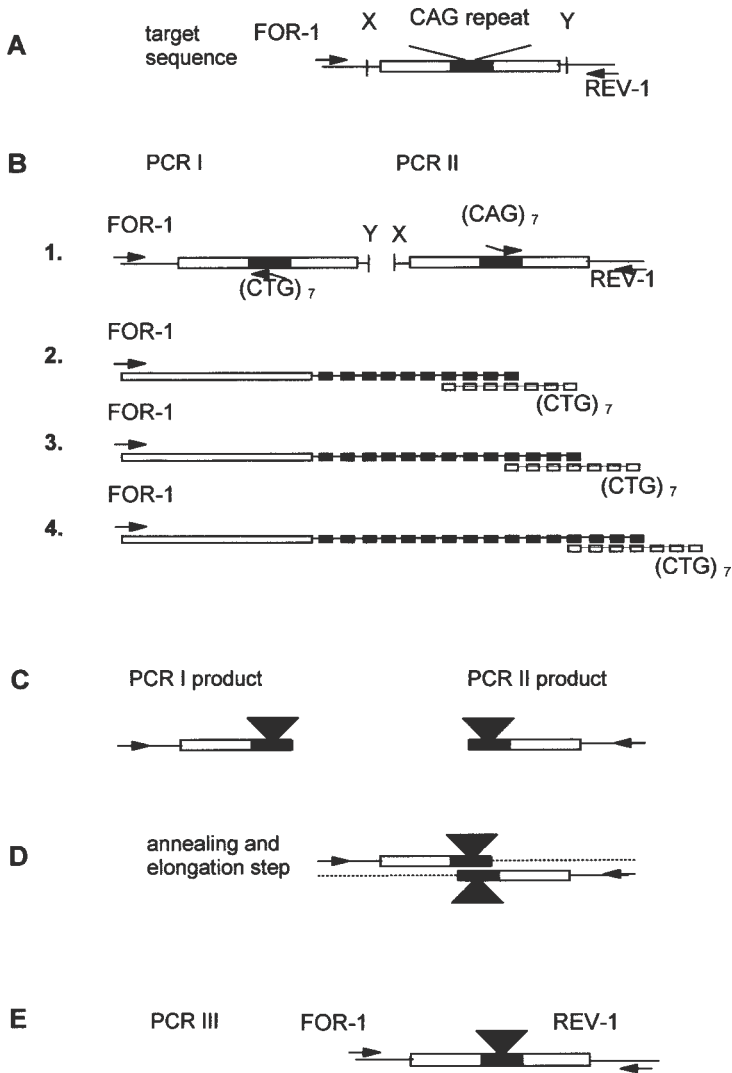


Fig. 1. Strategy for introducing CAG expansions into CAG-containing genes. (A) Diagrammatic representation of the target sequence including the repeat region (CAG repeat), restriction sites X and Y, where X is the enzyme for the "5'-digestion" and Y is the enzyme for the "3'-digestion" as explained in the text and the two outsider specific forward primer (FOR-1) and reverse primer (REV-1). (B) amplification of the target DNA in two distinct reactions after digestion with the X and Y restriction endonucleases in case of a circular template. As primers the pair FOR-1/(CTG)₇ for PCR I and the pair (CTG)₇/REV-1 for PCR II were used. (B2, B3, B4), subsequent cycles of PCR I leading to a progressive elongation of the repeat. (C) Amplicons of both PCRs. (D) Aliquots of both PCRs were mixed, denatured, annealed and extended with DNA polymerase. (E) Amplification of the elongated products of step D in PCR III with the specific (vector or gene specific) primers FOR-1 and REV-1. Arrowhead: variable CAG repeats. (Slightly modified from Laccone et al. (4), reprinted by permission of Wiley-Liss, Inc. Jossey-Bass Inc., a subsidiary of Wiley and Sons, Inc.)

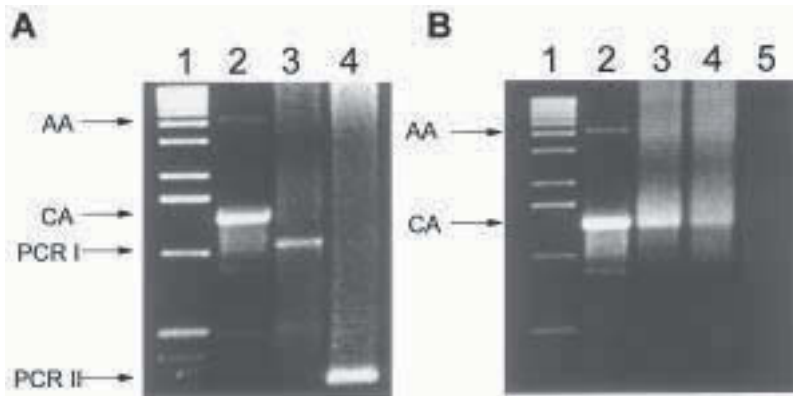


Fig. 2. (A) Electrophoresis of the PCR products (7 μ L each) on a 1% agarose gel of the MJD/SCA3 cDNA. Lane-1: 1-kb-ladder; Lane 2: amplification product of the complete 1.4-kb cDNA (CA, control amplification) obtained with the primers (FOR-1 and REV-1); Lane-3: PCR I product of about 1.1 kb (primers FOR-1/(CAG)₇); Lane-4: PCR II product obtained with primers primers (FOR-1 and REV-1); Lane-3: PCR I product of about 1.1 kb (primers FOR-1/(CAG)₇); Lane 4: PCR II product obtained with primers (CTG)₇/REV-1. The visible smear in both amplicons of PCR I and II should represent most probably a continuous expansion of the repeat region. (B) Lane 1: 1-kb ladder; lane 2: amplification of the target cDNA with the FOR-1 and REV-1 specific primers. Lane 3-5: results of PCR III obtained by mixing 1, 2, and 4 μ L of PCR I and PCR II amplicons with the two specific primers FOR-1 and REV-1. The visible smear represents most probably a series of products with elongated CAGs. In lane 2 the faint band of about 4.4 kb (AA = additional amplicons) may be caused by the linear amplification of the target circular clones which in this case have been not digested prior to the PCR. (From Laccone et al. (4), reprinted by permission of Wiley-Liss Inc. Jossey-Bass Inc., a subsidiary of Wiley and Sons, Inc.)

this method to your own application, optimizing in particular the PCR conditions to your target. The required chemicals and instruments are usually available in every molecular genetic laboratory.

2. Materials

1. Gene of interest.
2. Restriction endonucleases (In our case: *NotI*, *SalI*, *EcoRV*, New England Biolabs, Inc. Beverly, MA).
3. Gene or vector specific primers.
4. Repeat primers: in our case (CAG)₇ and (CTG)₇.
5. HotStarTaq Master Mix Kit (Qiagen, Bielefeld, Germany).
6. pBluescript vector (Stratagene, La Jolla, CA).
7. T4 DNA Ligase (Roche, Switzerland).
8. *Pfu* polymerase (Stratagene).
9. pSure *Escherichia Coli* competent cells (Stratagene).
10. Plasmid isolation Mini Kit (Qiagen).
11. Sequencing facilities.
12. Basic devices of cloning experiments: electrophoresis chambers, centrifuges, PCR thermo cyclers, shakers and so on.

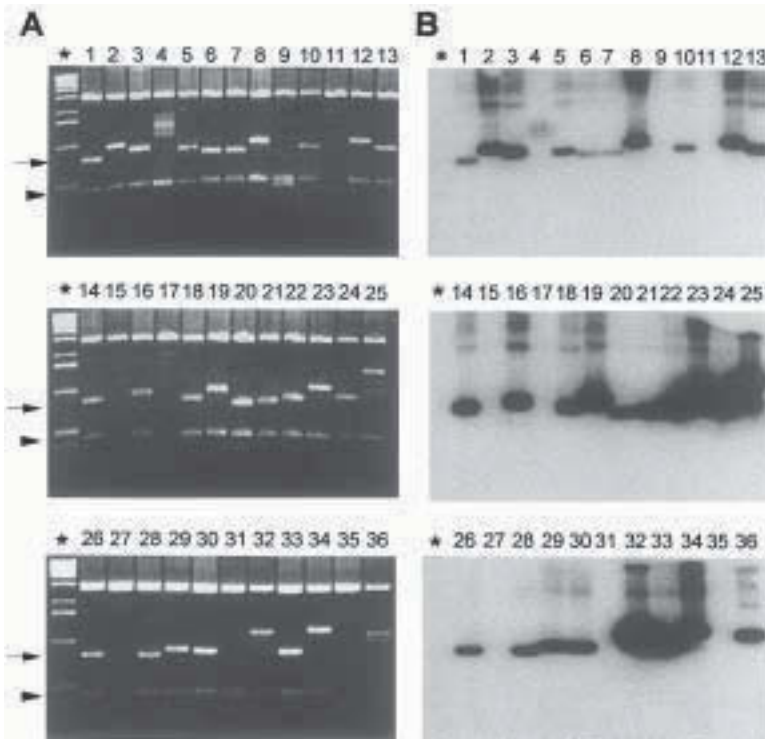


Fig. 3. (A) *EcoRI/HindIII* digestion of the cloned products of PCR III to release the plasmid insert. Because the *MJD/SCA3* gene contains an *EcoRI* restriction site at position 435, we expected a constant fragment of about 450 bp (arrowhead) and a fragment of $700 + (\text{CAG})_n$ bp (arrow) attributable to the presence of CAGs variable in size. (B) Hybridization of the blotted plasmid DNA with a probe containing 63 CAGs. (from Laccone et al. (4), reprinted by permission of Wiley-Liss Inc. Jossey-Bass Inc., a subsidiary of Wiley and Sons, Inc.).

3. Methods

1. Digestion of circular target DNA (see Note 1).

Two independent reactions containing each 1 μg of plasmid and a restriction endonuclease cutting at 3' and 5' end of the repeat region respectively in a volume of 50 μL for linearizing the plasmid will be carried out. The digested products will be called for our convenience 3'-digestion and 5'-digestion, respectively. If the target DNA is linear, this step can be skipped. 20 μL of each digested product are then diluted with 180 μL H_2O in clean tubes to a concentration of about 1 $\text{ng}/\mu\text{L}$. The rest of the digestion products should be stored at -20°C for further use. In case of linear DNA, the concentration of the DNA should be also of about 1 $\text{ng}/\mu\text{L}$ (see Fig. 1).

2. PCR I and PCR II.

In this step, the gene of interest will be amplified in two amplicons overlapping on the repeat region.

- a. PCR for 5' end (PCR I):
 - 1 μL 3'-digestion product from **step 1**
 - 1 μL specific forward primer (10 pmol/ μL)
 - 1 μL (CTG)₇ as reverse primer (10 pmol/ μL)
 - 25 μL HotstarTaq mix (Qiagen)
 - 22 μL H₂O
 - b. PCR for 3' end (PCR II):
 - 1 μL 5'-digestion product from **step 1**
 - 1 μL specific reverse primer (10 pmol/ μL)
 - 1 μL (CAG)₇ as forward primer (10 pmol/ μL)
 - 25 μL HotstarTaq mix (Qiagen)
 - 22 μL H₂O
 - c. Polymerase activation step:
 - 97°C for 15 min (polymerase activation step)
 - d. PCR conditions for 30 cycles (*see Note 2*):
 - 96°C for 30 s
 - 55°C for 30 s
 - 72°C for 2 min
3. Gel-electrophoresis of the amplicons of PCR I and PCR II.
A successful amplification should show primary products of the expected size with a smear of larger products. After that, the products must be purified either with cold ethanol precipitation or using commercial spin columns for removing unincorporated primers (*see Fig. 2*).
4. Heteroduplex formation and elongation.
Three reactions with different amount of amplicons from PCR I and II will be carried out:
- a. 1 μL of amplicons from PCR I
 - 1 μL of amplicons from PCR II
 - 25 μL HotstarTaqmix (Qiagen)
 - 21 μL H₂O
 - b. 2 μL of amplicons from PCR I
 - 2 μL of amplicons from PCR II
 - 25 μL HotstarTaqmix (Qiagen)
 - 19 μL H₂O
 - c. 4 μL of amplicons from PCR I
 - 4 μL of amplicons from PCR II
 - 25 μL HotstarTaqmix (Qiagen)
 - 15 μL H₂O
 - d. reaction conditions:
 - 97°C for 15 min
 - 60°C for 1 min
 - 72°C for 20 min
5. PCR III.
1 μL of the “outsiders” Forward and Reverse primer (10 pmol/ μL) will be added to each reactions of **step 4** above and the PCR will be carried out for 30 cycles at the following cycling conditions (*see Notes 2 and 3*): 96°C for 30 s, 55°C for 30 s, 72°C for 2 min.
The products of PCR III will be analyzed by agarose gel electrophoresis and the bands of interest will be cloned.

6. Cloning strategies.

In case the target gene was cloned into a plasmid and the outsiders primers were plasmid specific primers, it could be possible to cut the products with restriction endonucleases specific to the multiple cloning sites on the plasmid and subclone the digested product into a suitable vector. A more general method is to use a commercial T-vector according to the manufacturer's protocols or polishing the product with a proof-reading polymerase and cloning into blunt-ended vectors. This latter method will be described here because it is very convenient and economical.

- a. Digest 1 μg of pBluescript II with *EcoRV* restriction endonuclease in a total volume of 50 μL . After digestion 5 μL will be diluted to a final concentration of 1 $\text{ng}/\mu\text{L}$ by adding 95 μL H_2O .
- b. 20 μL of amplicons of PCR III will be incubated with 0.5 μL of *Pfu* polymerase (Stratagene) at 72°C for 30 min.

c. Ligation:

- 1 μL pBluescript/*EcoR V* digestion (1 $\text{ng}/\mu\text{L}$)
- 6 μL polished amplicons (from **step b**)
- 1 μL T4 DNA ligase (Roche, Germany)
- 1 μL 10X buffer
- 1 μL *EcoR V* restriction endonuclease (NEB)

The ligation reaction will be carried out overnight at 14°C (*see Note 4*).

d. "Killing" of religated vectors:

Add 0.5 μL of *EcoR V* restriction enzyme to the ligation reaction and incubate at 37°C for 60 min.

7. Transformation of the ligation reaction into suitable *E. coli* cells (*see Note 5*).
8. Picking and analysis of colonies for the presence of inserts. Single white colonies should be put into liquid culture with the corresponding antibiotic (in our case ampicillin). Isolate the plasmid using the Plasmid mini kit of Qiagen. Digestion of the positive plasmid and agarose gel electrophoresis analysis will reveal the size of the expansions (*see Fig. 3*).
9. Sequence analysis of the recombinant clones.

4. Notes

1. The digestion of the circular plasmid is very important to avoid a linear amplification of the original circular sequence. It is advisable to cut out the complete sequence of the gene of interest from the vector.
2. The PCR conditions should most probably be adapted depending on the different target genes. The annealing temperature and elongation time are the two important variables that should be changed if required. As a general rule, it would be advisable to identify the optimal cycling conditions for the amplification of the complete gene with the two "outsiders" primers. Furthermore it would be advisable to develop cloning strategies for the amplification of products that are not very large (up to 2-kb each). The restriction map of the gene of interest should be helpful in identifying the desired positive clones (*see Fig. 3A*).
3. It would be useful to amplify prior to PCR III the complete target with the "outsiders" primers for finding the optimal PCR conditions. In the PCR III the complete target should be amplified as a control of the PCR efficiency.
4. The addition of *EcoRV* restriction endonuclease is advisable for reducing the background due to the vector's self-religation, provided that no *EcoRV* recognition sequence is contained within the gene of interest. An alternative to the *EcoRV* might be the *SmaI* restric-

tion endonuclease. However, the ligation into the *Sma*I site was in our hands not so efficient as the *Eco*RV site. The ligation reaction will be carried out overnight at 14°C.

5. Dealing with expanded CAG repeats requires caution. The most important one is to inhibit the transcription of the gene. As a matter of fact the transcription of elongated repeat can result in deletions or expansion of the repeat and furthermore can have a toxic effect on the cells. The amount of extractable plasmids from induced cells in this latter case decreases consistently. In our hand the pSure containing a F episome (F' proAB lac^qZΔM15 Tn10 [Tet^r]) has been very reliable. When preparing the competent cells in one's own laboratory it is very important to grow the pSure cells on plates containing tetracyclin. By omitting tetracyclin it might be possible that the cells will lose the F episome and the repression of the transcription by the lac^q will be no longer efficient reducing the amount of extractable plasmids per liquid cultures.

The choice of the DNA polymerase for the amplification steps (PCR I, II and III) depends on the desired amplification fidelity. We used the Hotstar Taq DNA polymerase as a compromise between fidelity and amplification efficiency. Using this polymerase we have obtained clones up to 138 CAGs repeats without any errors and some clones with errors within the repeat. This random insertion of errors within the repeat (CAG to CCG or CAG to TAG) might be useful in some instances (e.g., effect of perfect versus imperfect repeats on cell culture or in mouse models). The use of a proof-reading polymerase, however, should reduce the error rate during the amplification.

References

1. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E., and Fischbeck, K. H. (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **4**, 77–79.
2. Ordway, J. M. and Detloff, P. J. (1996) In vitro synthesis and cloning of long CAG repeats. *Biotechniques* **21**, 609–612.
3. Merry, D. E., Kobayashi, Y., Bailey, C. K., Taye, A. A., and Fischbeck, K. H. (1998) Cleavage, aggregation and toxicity of the expanded androgen receptor in spinal and bulbar muscular atrophy. *Hum. Mol. Genet.* **7**, 693–701.
4. Laccone, F., Maiwald, R., and Bingemann, S. (1999) A fast polymerase chain reaction-mediated strategy for introducing repeat expansions into CAG-repeat containing genes. *Hum Mutat.* **13**, 497–502.

PCR Screening in Signature-Tagged Mutagenesis of Essential Genes

Dario E. Lehoux and Roger C. Levesque

1. Introduction

Signature tagged-mutagenesis (STM) is a functional genomics technique that identifies microbial genes required for infection within an animal host, or within host cell (1,2). As first described by Hensel et al., 1995 (3), transposon mutants are generated and each one tagged with a unique DNA sequence. Originally, STM used comparative hybridization to isolate mutants unable to survive in specified environmental conditions and to identify genes critical for survival in the host (3). The original STM has been modified to use defined oligonucleotides for tag construction into mini-Tn5 and to use polymerase chain reaction (PCR) instead of hybridization for rapid screening of bacterial mutants in vivo (4). The modified STM technique has been called PCR-based signature-tagged mutagenesis (PBSTM).

STM is divided into two steps: the construction of a library of tagged mutants and the in vivo screening of the library. First, PBSTM scheme involves designing pairs (12 in this case, but 24, 48, and 96 could be utilized) of 21-mers (see **Table 1**) synthesized as complementary DNA strands for cloning into the mini-Tn5 plasmid vector. The tagged minitransposons are used to mutagenize a microorganism. Each individual mutant can in theory be distinguished from every other mutant based on the different tag carried by the transposon in its genome (5). The set of 12 tags is repeatedly used to construct 12 libraries (see **Fig. 1**) and used for specific DNA amplifications easily detectable as signature tags (see **Fig. 2**).

A key step in PCR is the design of primers with specific DNA sequence. In this goal, primers should be between 18-mers to 24-mers in length (6). Moreover, higher free energy for duplex formation (ΔG) (7) caused by insertion of certain nucleotides at the 5'-end of a PCR primer stabilized primer-template duplex and optimized amplification reactions (8). On the other hand, the 3'-terminal position in primers was found essential for controlling mispriming (9). The insertion of a nucleotide mismatch at the 3'-terminus of a primer-template duplex is more detrimental to PCR amplification

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

Table 1
DNA sequences of oligonucleotides synthesized for STM

Tag number	Nucleotide sequence ^a
1	5'-GTACCGCGCTTAAAC CGTTCAG -3'
2	GTACCGCGCTTAAAT AGCCTG
3	GTACCGCGCTTAAA AGTCTCG
4	GTACCGCGCTTAATA ACGTGG
5	GTACCGCGCTTAA ACTGGTAG
6	GTACCGCGCTTA AGCATGTTG
7	GTACCGCGCTTA ATGTAACCG
8	GTACCGCGCTTAAA ATCTCGG
9	GTACCGCGCTTA ATAGGCAAG
10	GTACCGCGCTTAA CAATCGTG
11	GTACCGCGCTTA ATCAAGACG
12	GTACCGCGCTTAA CTAGTAGG

^aEach 21-mers has a T_m of 64°C and permits PCRs in one step when primer combinations are used for screening. The consensus 5'-ends comprising the first 13 nucleotides has higher ΔG s for optimizing PCRs. The variable 3'-ends indicated in bold define tag specificity and allow amplification of specific DNA fragments. Each tag is used as a primer in PCR with a primer synthesized within the Km resistant gene of mini-Tn5 Km 2. The set of 12 21-mers representing the complementary DNA strand in each tag are not represented and can be deduced from the sequences presented.

than internal mismatches (9). Specific oligonucleotides should be designed to optimize PCR and to have high specificity during screening by PCR. Twelve pairs of 21-mers were designed as tags following three basic rules: i) similar T_m of 64°C to simplify tag comparisons by using one step of PCRs; ii) invariable 5'-ends with higher ΔG than at the 3'-end to optimize PCR amplification reactions; iii) variable 3'-end for an optimized yield of specific amplification product from each tag. The 21-mers are double stranded, and are cloned into a minitransposon (mini-Tn5 Km2) which is used to mutagenize and tag bacteria. All STM tags showed specificity as a unique DNA amplification product by PCR when using primers 1 to 12 in combination with the Km primer (see Fig. 2).

A series of suicide plasmids carrying mini-Tn5s each with a specific tag are used to mutagenize targeted bacteria giving 12 libraries of mutants; 96 groups of 12 mutants are pooled and arrayed into 96-well plates (see Fig. 1). The 12 mutants from the same pool are grown separately overnight at 37°C. Aliquots of these cultures are pooled and a sample is removed for PCR analysis (the in vitro pool). A second sample from the same pool is used for the in vivo passage. After this passage, bacteria recovered from the animal organ (the in vivo pool) and the in vitro pool are used as templates in 12 distinct PCR. Amplicons obtained with the in vitro and in vivo pools are compared.

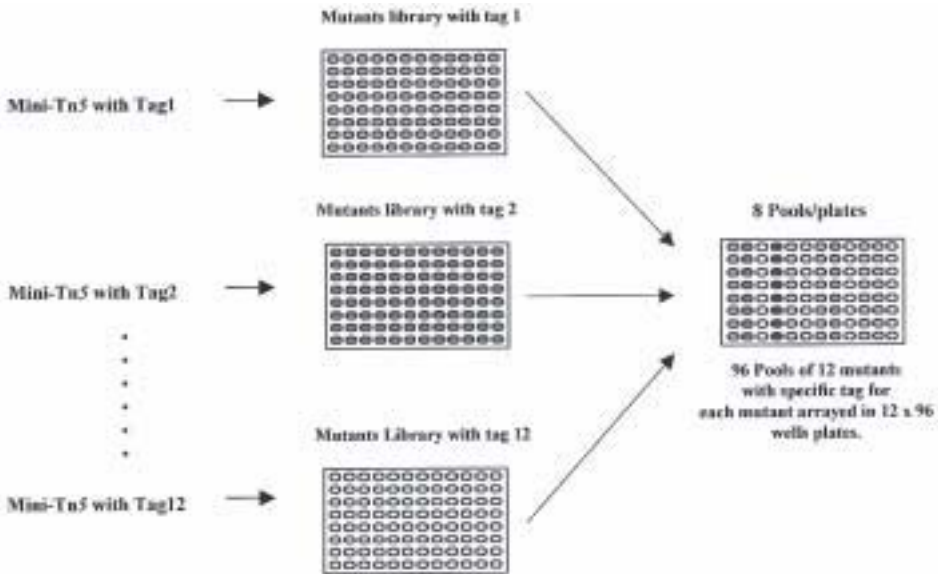


Fig. 1. Construction of 12 libraries of *P. aeruginosa* mutants tagged with mini-Tn5 *Km2*. Double-stranded DNA tags were cloned into the pUTmini-Tn5 *Km2* plasmid (see methods). *Km*-resistant exconjugants were arrayed as libraries of 96 clones. In a defined library, each mutant has the same tag but inserted at different locations in the bacterial chromosome. One mutant from each library is picked to form 96 pools of 12 mutants with a unique tag for each. The differences between tags are represented by colors. O and I represent the 19 bps inverted repeats at each extremity of the mini-Tn5.

Undetected mutants after the *in vivo* passage are *in vivo* attenuated. This simple STM method can be adapted to any bacterial system and used for genome scanning in various growth conditions.

2. Materials

- 10X medium salt buffer: 10 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 50 mM NaCl, 1 mM dithiothreitol (DTT).
- pUT mini-Tn5 *Km2* plasmid (10).
- KpnI* (New England Biolabs, Mississauga, Ont., Canada).
- 10X NEB #1 buffer: 10 mM bis-Tris Propane-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.0.
- 10X bovine serum albumin (BSA) (1 mg/mL) (NEB).
- T₄ DNA polymerase (Gibco-BRL, Burlington, Ont., Canada)
- dNTPs (dATP, dGTP, dCTP, dTTP from Amersham Pharmacia Biotech, Baie d'Urfé, QC, Canada).
- T₄ DNA ligase 10X buffer (NEB): 50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, 25 µg/mL BSA.
- T₄ DNA ligase (NEB).

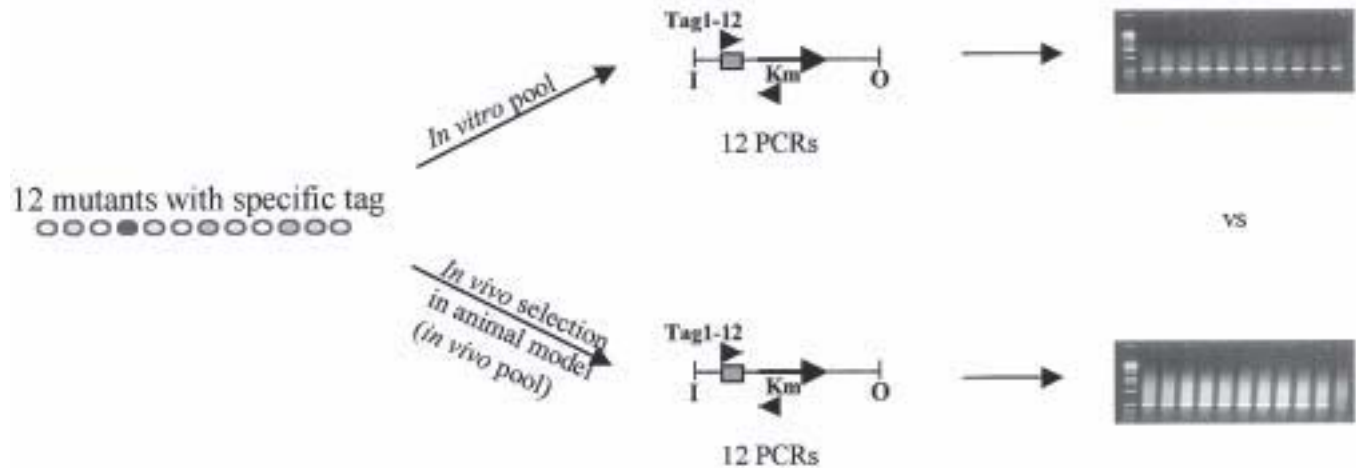


Fig. 2. STM scheme for comparisons between the in vitro and in vivo negative selection step. Mutants from the same pool were grown as separate cultures. An aliquot was kept as the in vitro pool, and a second aliquot was used for injection into an animal model for in vivo selection. After this passage, bacteria were recovered from animal organs and constitute the in vivo pool. The in vitro and in vivo pools of bacteria were used to prepare DNA templates in 12 PCRs using the 21-mers 1 to 12 in Table 1 and the Km primer. PCR products were analyzed by agarose gel electrophoresis. Lanes 1 to 12: PCR products obtained with the primers 1 to 12. In this example, the mutant with Tag11 was not recovered after the in vivo selection.

10. Micropure-EZ pure (Millipore, Montréal, QC, Canada).
11. Microcon 30 (Millipore).
12. Microcon PCR (Millipore).
13. Electrocompetent *Escherichia coli* S17-1 λ pir (washed in glycerol 10%).
14. 2-mm electroporation gap cuvetts (BTX, distributed by VWR Can Lab, Mississauga, Ont., Canada), SOC (Fisher, Montréal, QC, Canada),
15. SOC medium (Fisher): Formula per liter: 20 g Bacto tryptone, 5 g Bacto yeast extract, 0.5 g sodium chloride, 2.4 g magnesium sulfate anhydrous, 0.186 g potassium chloride, 20 mL of filter sterilized 20% glucose.
16. Tryptic soy broth (TSB; Fisher): Formula per liter: 17 g Bacto tryptone, 3 g Bacto soytone, 2.5 g Bacto dextrose, 5 g sodium chloride, 2.5 g dipotassium phosphate.
17. Ampicillin (Sigma Chemical Company, St. Louis, MO).
18. Kanamycin (Sigma).
19. TE buffer: 10 mM Tris-HCl, pH 7.4, 0.1 mM ethylenediaminetetraacetic acid (EDTA) 0.1 mM.
20. PCR premix: 10X *Taq* polymerase (Gibco-BRL) reaction buffer without Mg²⁺: 200 mM Tris HCl, pH 8.4, 500 mM KCl, 50 mM MgCl₂, 1.25 mM dNTPs, 10 pmoles oligonucleotide tag (see **Table 1**), 10 pmoles pUTKanaR1 (5'-GCGGCCTCGAGCAAGACGTTT-3'), *Taq* polymerase (Gibco-BRL).
21. Mineral oil.
22. Agarose.
23. 1X Tris-borate EDTA buffer: 5X concentrated stock solution per liter: 54 g Tris base, 27.5 g boric acid, 20 mL EDTA, pH 8.0.
24. 0.5 μ g/mL ethidium bromide solution.
25. Nylon membrane.
26. Brain heart infusion agar (BHIA; Fisher).
27. BHI (without agar).
28. Sterile 1X phosphate-buffered saline (PBS): 137 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.3 mM KH₂PO₄, pH 7.4.
29. 2-mL 96-wells plates (Qiagen).
30. 96-wells microtiter plates.
31. Animals.
32. Dissection kit.
33. Potter homogenizer.
34. QIAGEN genomic Tips (QUIAGEN).
35. Selected endonuclease.
36. pTZ18R (Amersham Pharmacia Biotech).
37. Electrocompetent *E. coli* DH5 α (washed in 10% glycerol).
38. X-gal (Sigma).
39. IPTG (Roche Diagnostics, Laval, QC, Canada).
40. QIAGEN midi preparation kit.
41. DNA sequencing service (Nucleic acids analysis and synthesis units, Laval University, <http://www.rsvs.ulaval.ca>).
42. PC computer.

3. Methods

3.1. Construction of Tagged Mini-Tn5 Km2

3.1.1. Double-stranded DNA tags

1. Twelve defined 21-mers oligonucleotides should be synthesized along with their complementary DNA strands as tags (*see Table 1*) (*see Note 1*) (Nucleic acids analysis and synthesis units, Laval University, <http://www.rsvs.ulaval.ca>.)
2. 50 pmoles of both complementary oligonucleotides are mixed in 100 μ L of medium salt buffer for the annealing reactions.
3. The annealing reaction mix is heated 5 min at 95°C, left to cool slowly at room temperature in the block heater, and kept on ice (*see Note 2*).

3.1.2. Minitransposon Tagging

1. 20 μ g of pUT mini-Tn5 Km2 plasmid are digested with 20 units of *KpnI* in 40 μ L of 1X NEB #1 buffer with 1X BSA (*see Note 3*).
2. Incubate 2 h at 37°C.
3. Inactivate 20 min at 65°C.
4. 4 nmoles of each dNTPs and 5 units of T₄ DNA polymerase are added to digested plasmid solution to blunt extremities.
5. Purify modified plasmid from enzymes with micropure-EZ and microcon 30 systems in a single step as described by the manufacturer's protocol.
6. 0.04 pmoles of plasmid are ligated to 1 pmole of double stranded DNA tags in a final volume of 10 μ L of T₄ DNA ligase 1X buffer containing 400 U of T₄ DNA ligase (*see Note 4*).
7. Ligated products are purified using microcon PCR as described by the manufacturer's instructions and resuspended in 5 μ L of H₂O.
8. All the 5 μ L containing ligated products are electroporated in *E. coli* S17- λ pir (*II*) (*see Note 5*) using a Bio-Rad apparatus at 2.5 KV, 200 Ohms, 25 μ F in a 2-mm electroporation gap cuvet. After electroporation, 0.8 mL of SOC is added to cells, which are transferred in culture tubes to be incubated 1 h at 37°C.
9. Transformants are selected on TSB supplemented with 50 μ g/mL of ampicillin and 50 μ g/mL of kanamycin by plating 100 μ L of transformed cells.
10. Single colonies are selected, purified and screened by colony PCR (*see Note 6*) in 50- μ L reaction volumes containing: 10 μ L of boiled bacterial colonies in 100 μ L of TE buffer; 5 μ L of 10X *Taq* polymerase (Gibco-BRL) reaction buffer; 1.5 mM MgCl₂; 200 μ M of each dNTPs; 10 pmoles of one of the oligonucleotide tag (*4*) used to construct the DNA tags as a 5' primer and 10 pmoles of the pUTKanaR1 (5'-GCGGCCTCGAGCA-AGACGTTT-3') as the 3' primer in the kanamycin resistance gene; 2.5 U of *Taq* polymerase (Gibco-BRL). Thermal cycling conditions were (touchdown PCR) (*see Note 7*): a hot start for 7 min at 95°C, 2 cycles at 95°C for 1 min, 70 to 60°C for 1 min, and at 72°C for 1 min, then followed by 10 cycles at 95°C for 1 min, 60°C for 1 min, 72°C for 1 min in a DNA Thermal Cycler (Perkin Elmer Cetus). 10 μ L of the amplified products were analyzed by electrophoresis in a 1% agarose gel, 1X Tris-borate EDTA buffer and stained for 10 min in 0.5 μ g/mL ethidium bromide solution (*12*). The amplicons should be around 500 basepairs (*see Fig. 2*).

3.2. Tagged Mutants Libraries Construction

3.2.1. Mutagenesis

1. *E. coli* S17-1 λ pir strain (containing tagged pUT mini-Tn5 plasmid) is used as a donor for conjugal transfer into the recipient strain. Separate conjugation must be done for each tagged minitransposon. The donor:recipient ratio should be established to obtain the maximum exconjugants by doing preliminary experiments. Cells are mixed and spotted as a 50- μ L drop on a nylon membrane placed on a nonselective BHIA plate. Plates are incubated at 30°C for 8 h (see **Note 8**).
2. Filters are washed with 10 mL of phosphate buffered saline (PBS) to recover bacteria.
3. Five 100 μ L aliquots of the PBS solution containing exconjugants are plated on five BHIA plates supplemented with the appropriate antibiotic to select for the strain. Kanamycin is used to select exconjugants with the mini-Tn5 inserted into their chromosomes (see **Note 9**). Plates are incubated overnight at 37°C.
4. Exconjugants are selected on BHIA supplemented with ampicillin. Mutants resistant to ampicillin are removed from the pool, since they carry the suicide vector inserted into the chromosome.
5. Kanamycin resistant and ampicillin sensitive exconjugants from a single conjugation are arrayed as a library of 96 clones in 2 mL 96-wells plate in 1.5 mL of BHI supplemented with kanamycin and appropriate antibiotic. The 2-mL 96-wells plates are incubated 18–22 h at 37°C (see **Note 10**). At this step, 12 differently tagged libraries are obtained.
6. As an STM working scheme, one mutant from each library is picked to form 96 pools of 12 unique tagged mutants (see **Fig. 1**) contained in the 2-mL 96-wells plates.

3.3. In Vivo Screening of Tagged Mutants

1. Each mutant from the same pool are inoculated individually in 200 μ L of TSB containing kanamycin and grown overnight at 37°C without agitation in microtiter plate.
2. Aliquots of these cultures are pooled.
3. A first sample is diluted from 10⁻¹ to 10⁻⁴, and plated on BHIA supplemented with kanamycin.
4. After overnight incubation at 37°C, 10⁴ colonies are recovered in 5 mL of PBS and a sample of 1 mL is removed for PCR and called the in vitro pool.
5. The 1 mL in vitro pool sample is spun down and the cell pellet is resuspended in 1 mL of TE buffer.
6. The in vitro pool is boiled 10 min, spun down, and 10 μ L of supernatant are used in PCR analysis as described above.
7. A second sample from the pooled cultures is used to inoculate animals.
8. After the appropriate in vivo incubation time, animals are sacrificed and bacteria are recovered from the targeted organs (see **Note 11**).
9. Tissues are recovered by dissection and homogenized with a Potter homogenizer in 10 mL of sterile phosphate buffered saline pH 7.0 contained in a 50-mL falcon tube (see **Note 12**).
10. 100 L of homogenized tissues are plated on BHIA supplemented with kanamycin. After the in vivo selection, 10⁴ colonies recovered from a single plate are pooled in 5 mL of PBS. From the 5 mL, 1 mL is spun down and resuspended in 1 mL of TE buffer (the in vitro pool).

11. The *in vivo* pool is boiled 10 min, spun down, and 10 μ L of supernatant is used in PCR analysis as described above. 10 μ L of PCR are used for 1% agarose gel electrophoresis separation.
12. PCR amplification products of tags present in the *in vivo* pool are compared with amplified products of tags present in the *in vitro* pool (*see* Fig. 2).
13. Mutants that give PCR amplicon from *in vitro* pool and not from *in vivo* pool are purified and kept for further analysis (*see* Note 13).

3.4. Cloning and Sequence Analysis of Transposon-Flanking DNA from Attenuated Mutants

1. Chromosomal DNA from attenuated mutants is prepared using the Qiagen genomic DNA extraction kit as described in the manufacturer's protocol.
2. Chromosomal DNA (1 μ g) is digested with endonuclease giving a large range of fragment sizes (*see* Note 14).
3. Digested chromosomal DNAs are cloned into pTZ18R predigested with the corresponding endonuclease. Ligation reactions are done as follows: 1 μ g of digested chromosomal DNA is mixed with 50 ng of digested pTZ18r in 20 μ L of 1X T₄ DNA ligase buffer with 40 units of T₄ DNA ligase.
4. Incubate overnight at 16°C.
5. Ligated products are purified using microcon PCR as described by the manufacturer's instructions and resuspended in 5 μ L of H₂O.
6. The 5- μ L recombinant plasmid solution is used for electroporation in *E. coli* DH5 α as described previously.
7. All the electroporation cells are spun down and resuspended in 100 μ L of BHI to be plated on a selective plate. Recombinant bacteria are selected as white versus blue colonies on X-gal/IPTG containing plates (0.005% and 0.1 mM, respectively) with ampicillin (100 μ g/mL) and kanamycin (50 μ g/mL) (*see* Note 15).
8. Clones are kept and purified for plasmid analysis.
9. Plasmid DNAs are prepared with QIAGEN midi preparation kit as described by the manufacturer.
10. These plasmids are sequenced using the complementary primer of the corresponding tagged mutant. Automated sequencing (ABI 373) is done as suggested by the manufacturer.
11. DNA sequences obtained are assembled and subjected to database searches using BLAST included in the GCG Wisconsin package (version 10.0). Complete open reading frames (ORF) of disrupted genes and similarity searches with complete genomes can be performed at NCBI using the microbial genome sequences (<http://www.ncbi.nlm.nih.gov>).

4. Notes

1. Here we present an example of a set of twelve 21-mers as DNA tags. However, it is possible to elaborate more or other DNA tags by following previously described rules. Specificity and quality of amplification should be tested prior to using them for tagged minitransposon in a mutagenesis experiment because it should have no cross amplification from different DNA tags.
2. The annealing oligonucleotide mixture should be made before each ligation.
3. It is possible to digest several small quantities of DNA preparation and after purification (step 3.1.2.5.) pool all digested plasmid preparations.

4. Using freshly made annealing oligonucleotide mixture raises the efficiency of ligation because tags might be degraded.
5. For replication and maintenance of the recombinant plasmid, it might be useful to use the well-known *E. coli* DH5 α λ *pir* strain. However, it will be necessary to transfer plasmids in the S17-1 λ *pir* strain to transfer DNA by conjugation.
6. It might be necessary to screen several colonies to find the good recombinant. It is possible to pool several colonies to reduce the number of PCRs (13). This ensures that you have the good recombinant among the selected colonies in very few PCRs. To bypass the necessity of doing plasmid preparations, PCRs can be done on bacterial cell lysates. One or several colonies are resuspended in 100 μ L of TE buffer, boiled 10 min, and spun down. 10 μ L of supernatant are used for the PCR template. After the pool PCR, the specific clone containing tagged plasmid should be identified within the pool.
7. Touchdown PCR was preferred to the standard PCR cycle because establishment of optimal PCR conditions for two primers is facilitated, and it increased specificity of amplification products obtained. It involves decreasing the annealing temperature by 1°C every second cycle to a “touchdown” annealing temperature, which is then used for 10 or so cycles. In this case, annealing temperature takes place at 6°C above the calculated T_m . During the following cycles, the annealing temperature is gradually reduced by 1°C until it has reached a level of approx 4°C below T_m .
8. Temperature and incubation time should be determined by preliminary experiments.
9. It is very important to use the good kanamycin concentration to eliminate background related to the inoculum effect. Minimal inhibitory concentration can be determined to evaluate the effective kanamycin concentration.
10. In a defined library, each mutant has the same tag but is assumed to be inserted at a different location in the bacterial chromosome. Southern blot hybridization is necessary to confirm the random integration of the mini-Tn5 (12).
11. Parameters concerning animal model should be particularly well defined. The inoculum size necessary to cause infection determines the complexity of mutants pooled. In fact, each mutant in a defined input pool has to be in a sufficient cell number to initiate infection. The inoculum size must not be too high, resulting in the growth of mutants which would otherwise have not been detected (2). Other important parameters in STM include the route of inoculation and the time-course of a particular infection. Also, certain gene products important directly or indirectly for initiation or maintenance of the infection may be niche-dependent or expressed specifically in certain animal or plant tissues only. If the duration of the infection in STM *in vivo* selection is short, genes important for establishment of the infection will be found, and if the duration is long, genes important for maintenance of infection will be identified (2). Several routes of inoculation and different animal models can be used.
12. Keep homogenates on ice.
13. Each STM attenuated mutant has to be confirmed by: a second round of STM screening (14), comparisons between *in vivo* bacterial growth rate of mutants versus growth of the wild-type in single (15) or competitive (16) infections, or estimation of LD₅₀ (3).
14. More than one endonuclease or partial digestion can be used to obtain more DNA fragments ranging from 1 to 4 Kb that are easier to clone in pTZ18r.
15. Only clones that contain plasmid with chromosomal fragments and the mini-Tn5 marker are obtained.

References

1. Shea, J. E., Santangelo, J. D., and Feldman, R. G. (2000) Signature-tagged mutagenesis in the identification of virulence genes in pathogens. *Curr. Opin. Microbiol.* **3**, 451–458.
2. Lehoux, D. E. and Levesque, R. C. (2000) Detection of genes essential in specific niches by signature-tagged mutagenesis. *Curr. Opin. Biotechnol.* **11**, 434–439.
3. Hensel, M., Shea, J. E., Gleeson, C., Jones, M. D., Dalton, E., and Holden, D. W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**, 400–403.
4. Lehoux, D. E., Sanschagrín, F., and Levesque, R. C. (1999) Defined oligonucleotide tag pools and PCR screening in signature-tagged mutagenesis of essential genes from bacteria. *Biotechniques* **26**, 473–478, 480.
5. Chiang, S. L., Mekalanos, J. J., and Holden, D. W. (1999) *In vivo* genetic analysis of bacterial virulence. *Annu. Rev. Microbiol.* **53**, 129–154.
6. Dieffenbach, W. C. and Dveksler, G. S., eds. (1995) *PCR Primer: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 714.
7. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
8. Rychlik, W. (1993) Selection of primer chain reaction, in *PCR protocols: current application*, W.H. Press, Editor.: Totowa NJ.
9. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J. J. (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. *Nucl. Acids Res.* **18**, 999–1005.
10. De Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative eubacteria. *J. Bacteriol.* **172**, 6568–6572.
11. Simon, R., Priefer, U., and Pühler, A. (1983) A broad range mobilization system for *in vitro* genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/Technology* **1**, 784–791.
12. Sambrook, J., Fritsch, E. F., and Maniatis, T., (eds) (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. Dewar, K., Sabbagh, L., Cardinal, G., Veilleux, F., Sanschagrín, F., Birren, B., and Levesque, R. C. (1998) *Pseudomonas aeruginosa* PAO1 bacterial artificial chromosomes: strategies for mapping, screening, and sequencing 100 kb loci of the 5.9 Mb genome. *Microb. Comp. Genom.* **3**, 105–117.
14. Darwin, A. J. and Miller, V. L. (1999) Identification of *Yersinia enterocolitica* genes affecting survival in an animal host using signature-tagged transposon mutagenesis. *Mol. Microbiol.* **32**, 51–62.
15. Camacho, L. R., Ensergueix, D., Perez, E., Gicquel, B., and Guilhot, C. (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol. Microbiol.* **34**, 257–267.
16. Chiang, S. L. and Mekalanos, J. J. (1998) Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. *Mol. Microbiol.* **27**, 797–805.

Staggered Extension Process (StEP) In Vitro Recombination

Anna Marie Aguinaldo and Frances Arnold

1. Introduction

In vitro polymerase chain reaction (PCR)-based recombination methods are used to shuffle segments from various homologous DNA sequences to produce highly mosaic chimeric sequences. Genetic variations created in the laboratory or existing in nature can be recombined to generate libraries of molecules containing novel combinations of sequence information from any or all of the parent template sequences. Evolutionary protein design approaches, in which libraries created by in vitro recombination methods are coupled with screening (or selection) strategies, have successfully produced variant proteins with a wide array of modified properties including increased drug resistance (1,2), stability (3–6), binding affinity (6), improved folding and solubility (7), altered or expanded substrate specificity (8,9), and new catalytic activity (10).

Stemmer reported the first in vitro recombination, or “DNA shuffling,” method for laboratory evolution (11). An alternative method called the staggered extension process (StEP) (12) is simpler and less labor intensive than DNA shuffling and other PCR-based recombination techniques that require fragmentation, isolation, and amplification steps (1,11,13,14). StEP recombination is based on cross hybridization of growing gene fragments during polymerase-catalyzed primer extension (12). Following denaturation, primers anneal and extend in a step whose brief duration and suboptimal extension temperature limit primer extension. The partially extended primers randomly reanneal to different parent sequences throughout the multiple cycles, thus creating novel recombinants. The procedure is illustrated in **Fig. 1**. The full-length recombinant products can be amplified in a second PCR, depending on the product yield of the StEP reaction. The StEP method has been used to recombine templates with sequence identity ranging from single base differences to natural homologous genes that are approx 80% identical.

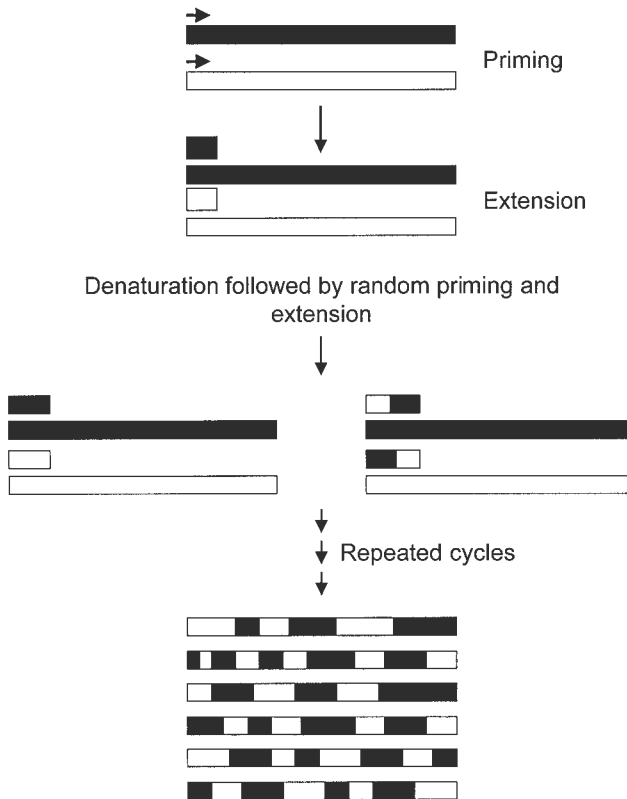


Fig. 1. StEP recombination, illustrated for two gene templates. Only one primer and single strands from the two genes (open and solid blocks) are shown for simplicity. During priming, oligonucleotide primers anneal to the denatured templates. Short fragments are produced by brief polymerase-catalyzed primer extension that is interrupted by denaturation. During subsequent random annealing-abbreviated extension cycles, fragments randomly prime the templates (template switching) and extend further, eventually producing full-length chimeric genes. The recombinant full-length gene products can be amplified in a standard PCR (optional).

2. Materials

1. DNA templates containing the target sequences to be recombined (*see Note 1*).
2. Oligonucleotide primers universal to all templates to be recombined (*see Note 2*).
3. *Taq* DNA polymerase (*see Note 3*).
4. 10X PCR buffer: 500 mM KCl, 100 mM Tris-HCl, pH 8.3.
5. 25 mM MgCl₂.
6. dNTP solution: 10 mM of each dNTP.
7. Agarose gel electrophoresis supplies and equipment.
8. *DpnI* restriction endonuclease (20 U/μL) and 10X supplied buffer (New England Biolabs, Beverly, MA).
9. QIAquick gel extraction kit (Qiagen, Valencia, CA) or your favorite method.

3. Methods

1. Combine 1–20 ng total template DNA, 0.15 μ M each primer, 1X PCR buffer, 200 μ M dNTP mix, 1.5 mM MgCl₂, 2.5 U *Taq* polymerase, and sterile dH₂O to 50 μ L. Set up a negative control reaction containing the same components but without primers (see **Note 4**).
2. Run the extension protocol for 80–100 cycles using the following parameters: 94°C for 30 s (denaturation) and 55°C for 5–15 s (annealing/extension) (see **Notes 5** and **6**).
3. Run a 5–10 μ L aliquot of the reactions on an agarose gel to check the quality of the reactions (see **Note 7**). If a discrete band with sufficient yield for subsequent cloning is observed after the StEP reaction, and the size of the full-length product is clearly distinguishable and easily separated from the original starting templates, proceed to **step 8**.
4. If parental templates were purified from a *dam+* *Escherichia coli* (*E. Coli*) strain (see **Note 1**), combine 2 μ L of the StEP reaction, 1X *DpnI* reaction buffer, 5–10 U *DpnI* restriction endonuclease, and sterile dH₂O to 10 μ L. Incubate at 37°C for 1 h (see **Note 8**).
5. Amplify the target recombinant sequences in a standard PCR using serial dilutions (1 μ L of undiluted, 1:10, 1:20, and 1:50 dilutions) of the *DpnI* reaction (or the StEP reaction if *DpnI* digestion was not done). Mix 1X PCR buffer, 1.5 mM MgCl₂, 200 μ M dNTP mix, 20 μ M of each primer, 2.5 U *Taq* DNA polymerase, and sterile dH₂O to 100 μ L.
6. Run the amplification reaction for 25 cycles using the following parameters: 94°C for 30 s, 55°C for 30 s, and 72 C for 60 s for each 1 kb in length.
7. Run a 10- μ L aliquot of the amplified products on an agarose gel to determine the yield and quality of amplification (see **Note 9**). Select the reaction with high yield and low amount of nonspecific products.
8. Gel purify the desired full-length reaction product following the manufacturer's protocol in the QIAquick gel purification kit. Digest the purified fragment with the appropriate restriction endonucleases for ligation into the preferred cloning vector.

4. Notes

1. Appropriate templates include plasmids carrying target sequences, sequences excised by restriction endonucleases and PCR amplified sequences. Reactions are more reproducible as template size decreases because this reduces the likelihood of nonspecific priming. For example, three 8.5-kb plasmid templates containing different 1.7-kb target sequences were less efficient for StEP recombination than 3-kb restriction fragments containing the target sequences. Unusually large plasmids and templates should be avoided.
Short template lengths may also pose a problem when the size is indistinguishable in length from the desired product. Conventional physical separation techniques, such as agarose gel electrophoresis, cannot be used to isolate the reaction product from the template, resulting in a high background of nonrecombinant clones. To minimize parental templates that may contribute to background nonrecombinant clones, plasmids used for template preparations (both intact plasmids and restriction fragments containing target sequences excised from plasmids) should be isolated from a methylation positive *E. coli* strain, e.g., DH5 α (BRL Life Technologies, Gaithersburg, MD) or XL1-Blue (Stratagene, La Jolla, CA). These *dam+* strains methylate DNA. *DpnI*, a restriction endonuclease that cleaves methylated GATC sites, can then be used to digest parental templates without affecting the PCR products.
2. Primer design should follow standard criteria including elimination of self-complementarity or complementarity of primers to each other, similar melting temperatures (within approx 2–4°C is best), and 40–60% G + C content. Primers of 21–24 bases in length work well.

3. Other investigators have used Vent DNA polymerase instead of *Taq* DNA polymerase in StEP recombination (15). Vent DNA polymerase is one of several thermostable DNA polymerases with proofreading activity leading to higher fidelity (16). Use of these alternative polymerases is recommended for DNA amplification when it is necessary to minimize point mutations. In addition, the proofreading activity of high-fidelity polymerases slows them down, offering an additional way to increase recombination frequency (17). Vent polymerase, for example, is reported to have an extension rate of 1000 nucleotides/min and processivity of 7 nucleotides/initiation event as compared to the higher 4000 nucleotides/min and 40 nucleotides/initiation event of *Taq* DNA polymerase (18). Slower rates lead to shorter extension fragments and greater crossover frequency.
4. The negative control reaction should be processed in the same manner as the sample reactions for all the steps of the procedure. No product should be visible for the no primer control throughout the procedure. If bands are present in the negative control similar to the sample reactions, products of the sample reactions may be the result of template contamination resulting in nonrecombinant clones.
5. The annealing/extension times chosen are based on the number of crossover events desired. Shorter extension times as well as lower annealing temperatures lead to increased numbers of crossovers due to the shorter extension fragments produced for each cycle. The size of the full-length product determines the number of reaction cycles. Longer genes require a greater number of reaction cycles to produce the full-length genes. The annealing temperature should be a few degrees lower than the melting temperature of the primers.
6. The progression of the fragment extensions can be monitored by taking 10 μ L aliquots of a duplicate StEP reaction at defined cycle numbers and separating the fragments on an agarose gel. For example, samples taken every 20 cycles from StEP recombination of two subtilisin genes showed reaction product smears with average sizes approaching 100 bp after 20 cycles, 400 bp after 40 cycles, 800 bp after 60 cycles, and a clear discrete band around 1 kb (the desired length) within a smear after 80 cycles (12). DNA polymerases currently used in DNA amplification are very fast. Even very brief cycles of denaturation and annealing provide time for these enzymes to extend primers for hundreds of nucleotides. For *Taq* DNA polymerase, extension rates at various temperatures are: 70°C, > 60 nt/s; 55°C, approx 24 nt/s; 37°C, approx 1.5 nt/s; 22°C, approx 0.25 nt/s (19). Therefore, it is not unusual for the full-length product to appear after only 10–15 cycles. The faster the full-length product appears in the extension reaction, the fewer the template switches that have occurred and the lower the crossover frequency. To increase the recombination frequency, everything possible should be done to minimize time spent in each cycle: selecting a faster thermocycler, using smaller test tubes with thinner walls, and, if necessary, reducing the reaction volume.
7. Possible reaction products are full-length amplified sequence, a smear, or a combination of both. Appearance of the extension products may depend on the specific sequences recombined or the template used. Using whole plasmids may result in nonspecific annealing of primers and their extension products throughout the vector sequence, which can appear as a smear on the gel. A similar effect may be observed for large templates. If no reaction products are visible, the annealing/extension times and the temperature of the StEP reaction will need to be determined empirically. Try reducing the annealing temperatures as well as modifying the primer and/or template concentrations.
8. The background from non-recombinant clones can be reduced following the StEP reaction by *Dpn*I endonuclease digestion to remove methylated parental DNA (see Note 1). At this point you want to get rid of the DNA template that is still in your reaction mixture before proceeding to amplification to prevent carryover contamination.

9. If the amplification reaction is not successful and you get a smear with a low yield of full-length sequence, reamplify these products using nested internal primers separated by 50–100 bp from the original primers.

References

1. Stemmer, W. P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **81**, 10,747–10,751.
2. Cramer, A., Raillard, S. A., Bermudez, E., and Stemmer, W. P. C. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291.
3. Giver, L., Gershenson, A., Freskgard, P. O., and Arnold, F. A. (1998) Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* **95**, 12,809–12,813.
4. Zhao, H. M. and Arnold, F. H. (1999) Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng.* **12**, 47–53.
5. Miyazaki, K., Wintrose, P. L., Grayling, R. A., Rubingh, D. N., and Arnold, F. H. (2000) Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J. Mol. Biol.* **297**, 1015–1026.
6. Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Pluckthun, A. (2001) Tailoring in vitro evolution for protein affinity or stability. *Proc. Natl. Acad. Sci. USA* **98**, 75–80.
7. Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotech.* **17**, 691–695.
8. Zhang, J. H., Dawes, G., and Stemmer, W. P. C. (1997) Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. USA* **94**, 4504–4509.
9. Kumamaru, T., Suenaga, H., Mitsuoaka, M., Watanabe, T., and Furukawa, K. (1998) Enhanced degradation of polychlorinated biphenyls by directed evolution. *Nat. Biotech.* **16**, 663–666.
10. Altamirano, M. M., Blackburn, J. M., Aguayo, C., and Fersht, A. R. (2000) Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold. *Nature* **403**, 617–622.
11. Stemmer, W. P. C. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391.
12. Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258–261.
13. Shao, Z., Zhao, H., Giver, L., and Arnold, F. H. (1997) Random-priming in vitro recombination: an effective tool for directed evolution. *Nucl. Acids Res.* **26**, 681–683.
14. Volkov, A. A. and Arnold, F. H. (2000) Methods for in vitro DNA recombination and random chimeragenesis. *Meth. Enzymol.* **28**, 447–456.
15. Ninkovic, M., Dietrich, R., Aral, G., and Schwienhorst, A. (2001) High-fidelity in vitro recombination using a proofreading polymerase. *BioTechniques* **30**, 530–536.
16. Cline, J., Braman, J. C., and Hogrefe, H. H. (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucl. Acids Res.* **24**, 3546–3551.
17. Judo, M. S. B., Wedel, A. B., and Wilson, C. (1998) Stimulation and suppression of PCR-mediated recombination. *Nucl. Acids Res.* **26**, 1819–1825.
18. Kong, H., Kucera, R. B., and Jack, W. E. (1993) Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J. Biol. Chem.* **268**, 1965–1975.
19. Innis, M. A., Myambo, K. B., Gelfand, D. H., and Brow, M. D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. USA* **85**, 9436–9440.

Random Mutagenesis by Whole-Plasmid PCR Amplification

Donghak Kim and F. Peter Guengerich

1. Introduction

1.1. General Introduction

Mutagenesis is a popular tool used in the analysis of protein structure and function. Polymerase chain reaction (PCR)-based mutagenesis can be used to introduce mutations with the use of the appropriate primer. Although the majority of attention has been given to site-directed mutagenesis, random mutagenesis is actually an older approach and has considerable potential because of its limited bias, if an appropriate screening method is available. This approach has successfully been used to obtain “gain-of-function” mutants (1). The ability to target mutants to individual proteins and parts of proteins with modern molecular tools has considerable applicability.

Generation of reliable random libraries for screening presents a particular challenge. Ideally, all potential clones should be represented in the library. Conventional methods include duplex oligonucleotide cassette synthesis followed by subcloning between two unique restriction sites (2), degenerate PCR-based methods using Mn^{2+} or splicing by overlap extension (3,4), single-stranded mutagenesis using bacteriophage M13-based vectors (5), and various chemical methods (6,7). Depending on the technique employed, introduction of multiple unique restriction sites for subcloning or single-stranded DNA rescue is often required. Other disadvantages of some of these methods include mutations outside the targeted region, intolerably high background from the native sequence, and mutational bias in terms of the types of nucleotide substitutions observed (8).

1.2. PCR-Based Whole Plasmid Amplification

We describe a less cumbersome method for random mutagenesis of up to five consecutive amino acids within a protein by PCR-based whole plasmid amplification using complementary degenerate primers. This method was originally introduced from

Table 1
Primer Codons for Random Mutagenesis
with No Wild-Type Background Remaining

amino acid	A	C	D	E	F	G	H	I	K	L
mutation primer	SHN	YNN	YNN	RNN	YNN	SDN	YNN	SBN	RNN	SBN
amino acid	M	N	P	Q	R	S	T	V	W	Y
mutation primer	DNN	YNN	SHN	RNN	RDN	SWN	SHN	SBN	DNN	YNN

The mutation primers represent the antisense primers from 5' to 3'. The sense primer can be designed as complementary. IUB Codes for Bases (N = G, A, T, or C in equal amount; S = G or C in equal amount; Y = C or T in equal amount; R = A or G in equal amount).

this laboratory (9), based on a modification of the Stratagene® QuikChange™ Site-Directed Mutagenesis approach. Also, a restriction enzyme-marker site is introduced to exclude the wild-type gene (10). The combination of reliability, convenience, low cost, and wide applicability of this technique renders it very practical for the construction of mutant libraries randomized within a limited target zone.

2. Materials

- Oligonucleotides: Sequences of oligonucleotide primer pairs for random mutagenesis can be chosen according to the following criteria:
 - Sense and antisense primers are of equal length, with the same start and end positions with respect to the coding strand of the template DNA (see Note 1).
 - Codons in the target region for random mutagenesis are encoded as in Table 1 (see Note 2).
 - Primers encompass the desired target region with a minimum of 12 bp of wild-type sequence on either side of the mismatches, preferentially terminating in multiple G or C residues to anchor the primer (see Note 3).
 - In cases where a unique restriction site is present within the template plasmids, the oligonucleotide is extended beyond the restriction site by the number of bases required for efficient (>90%) cleavage of the linear PCR product (see Fig. 1).
 - In cases where no unique site is present, a restriction site is incorporated into the primer by silent mutagenesis with the primer length increased correspondingly to ensure annealing (see Note 4).
 - Oligonucleotides for mutagenesis are synthesized on a 200 nmol scale and then purified with the appropriate methods.
- Plasmids: Small plasmids (containing the cDNA insert of interest) such as pBluescript® plasmid DNA (Stratagene, La Jolla, CA) are preferred for efficient PCR.
- DNA Polymerase and PCR Buffer: 2.5 U native *pfu* DNA polymerase (Stratagene); *pfu* DNA polymerase buffer: 20 mM Tris-HCl (pH 8.75), 10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.10% Triton X-100 (w/v), and 100 µg BSA/mL (see Note 5).

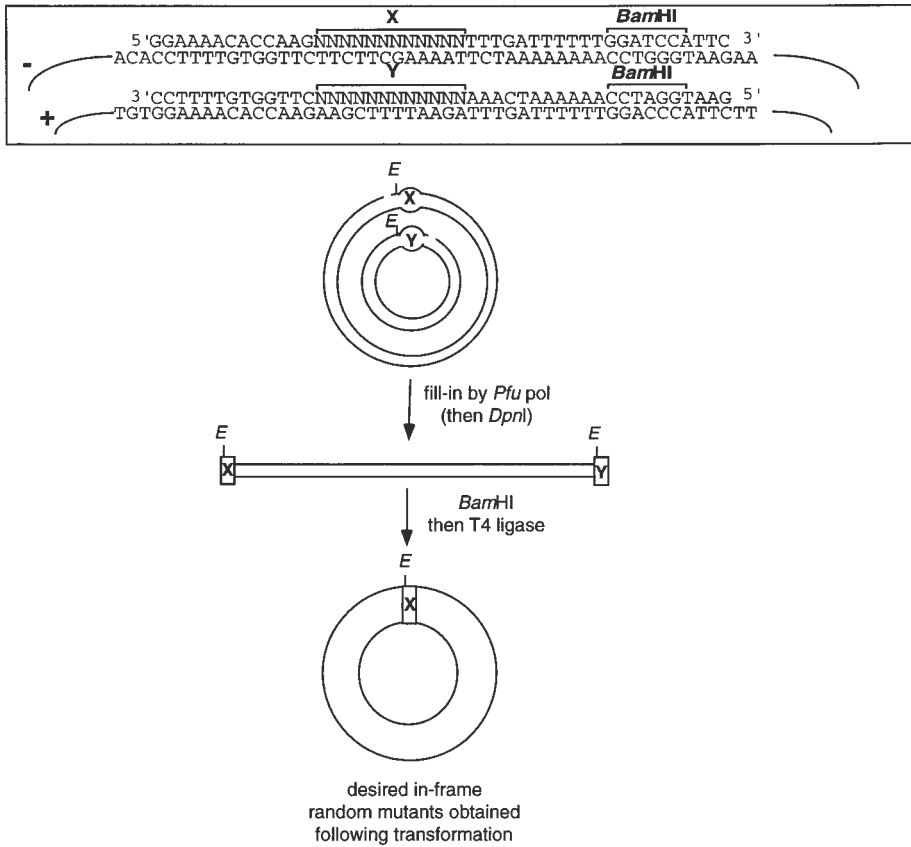


Fig. 1. Schematic representation of example primer design strategy for random mutagenesis of codons. Digestion at a unique restriction site followed by circularization with T4 DNA ligase excises a copy of the duplicated zone Y yielding desired in-frame mutants containing only X.

4. Bacterial strains: In order to increase the transformation efficiency of large and ligated DNA libraries, the ultrahigh competent cells are strongly recommended. *Escherichia coli* DH10B™ (Life Technologies, Gaithersburg, MD) or Epicurian Coli®XL10-Gold (Stratagene) is ideal for production of larger primary libraries.

3. Methods

1. Design degenerate primers spanning amino acids targeted for mutagenesis, with randomized codons encoded during primer synthesis.
2. Run PCR on plasmids DNA using purified primers (see Note 6):
 - 95°C initial denaturation, 5 min
 - (95°C, 1 min → 45°C, 1 min → 68°C, 2 min/kb) 30 cycles
 - 68°C extension, 10 min
 - 4°C hold

3. Clean and concentrate PCR by QIAquick[®] PCR purification kit (Qiagen, Valencia, CA) and resuspend DNA.
4. Digest cleaned DNA with the inserted restriction enzyme site (*see Note 7*).
5. Gel-purify the digested PCR product using available “gene-clean” methods, e.g., QIAquick[®] Gel Extraction kit (Qiagen). Elute DNA in 300 μ L distilled water or 10 mM Tris-HCl, pH 8.5.
6. Ligate DNA with T4 DNA ligase in a 350- μ L volume overnight at 16°C.
7. Precipitate reaction mixture by ethanol/NaOAc, wash pellet twice with 75% ethanol at room temperature, dry, resuspend DNA in 22 μ L water (*II*).
8. Perform transformation using ultra competent cells such as *E. coli* DH10b (Life Technologies) or Epicurian Coli XL10 Gold (Stratagene) and recover in 300 μ L SOC medium (*II*, *see Note 8*).
9. Plate pool on selective agar medium and incubate overnight at 37°C.
10. Add 2–5 mL Luria-Bertani medium; scrape and shake cells into a thick suspension.
11. Pool suspensions for each library and conduct alkaline lysis DNA preparation (*II*).

4. Notes

1. Only one of the primers (sense or antisense) needs to cover the target region in order to prevent self-hybridization, but both primers have to include any restriction enzyme sites.
2. Although all of the 20 amino acids can not be introduced, PCR primer design according to **Table 1** can eliminate the wild-type plasmid efficiently. In order to recover all 20 amino acids codons in PCR, the codons in the target region can be encoded as 5'NNS3' (S = G or C in equal amounts) in the sense primer and 5'SNN3' in the corresponding region of the antisense primer.
3. The total length of primers can be about 40–50 bp and can be extended up to 70 bp.
4. The restriction enzyme site is incorporated into the primer by silent mutagenesis using the instruction provided in the New England Biolab catalog (Beverly, MA).
5. The native *pfu* DNA polymerase is used for high PCR performance, but the conventional thermostable enzymes, such as *Taq* DNA polymerase, can be substituted when PCR conditions are optimized.
6. The primer concentrations for PCR must be optimized by serial dilution (**9**). Because the primers are self-complementary, there is a strong tendency toward self-hybridization. Either too little or too much primer may result in reduced PCR yield.
7. Additional digestion with the *DpnI*, which cuts the dam-methylated parental DNA, may help to remove the contamination of the wild-type plasmids originating from the templates in PCRs.
8. High efficiency transformation depends highly on the use of ultrahigh-competent *E. coli*. Twenty successive transformations are conducted for the full production of the library.

Acknowledgment

The authors are grateful to Dr. A. Parikh for valuable suggestions and discussions.

References

1. Botstein, D. and Shortle, D. (1985) Strategies and application of in vitro mutagenesis. *Science* **229**, 1193–1201.
2. Hill, D. E., Oliphant, A. R., and Struhl, K. (1997) Mutagenesis with degenerate oligonucleotides: an efficient method for saturating a defined DNA region with base pair substitutions. *Meth. Enzymol.* **155**, 558–568.

3. Fromant, M., Blanquet, S., and Plateau, P. (1995) Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Analyt. Biochem.* **224**, 347–353.
4. Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnol.* **16**, 258–261.
5. Nakamaye, K. L. and Eckstein, F. (1986) Inhibition of restriction endonuclease NciI cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis. *Nucl. Acids Res.* **14**, 9679–9698.
6. Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996) An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleotide analogues. *J. Mol. Biol.* **255**, 589–603.
7. Tange, T., Taguchi, S., Kojima, S., Miura, K., and Momose, H. (1997) Improvement of a useful enzyme (substilisin BPN') by an experimental evolution system. *Appl. Microbiol. Biotechnol.* **41**, 239–244.
8. Zoller, M. J. (1992) New recombinant DNA technology for protein engineering. *Curr. Opin. Biotechnol.* **3**, 348–354.
9. Parikh, A. and Guengerich, F. P. (1998) Random mutagenesis by whole-plasmid PCR amplification. *BioTechniques* **24**, 428–431.
10. Lanio, T. and Jeltsch, A. (1998) PCR-Based random mutagenesis method using spiked oligonucleotides to randomize selected parts of a gene without any wild-type background. *BioTechniques* **25**, 958–965.
11. Sambrook, J., Fritsch, E. F., and Maniatis, T., eds. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

IV

CLONING UNKNOWN NEIGHBORING DNA

PCR-Based Strategies to Clone Unknown DNA Regions from Known Foreign Integrants

An Overview

Eric Ka-Wai Hui, Po-Ching Wang, and Szecheng J. Lo

1. Introduction

Many foreign DNAs, such as some virus DNAs and almost all transposable elements (transposons), are capable of integrating host genomes, and the effects of integration can be pleiotropic. To investigate the mechanism and biological effect of foreign DNA insertions, characterization of the integration site, called integrant-host junction (IHJ), in the host genome becomes important. Traditional genomic library construction and screening for the cloning and analysis of IHJ are time-consuming, labor-intensive, and tedious. Therefore, a variety of efficient and reliable polymerase chain reaction (PCR)-based techniques have been developed. Application of the PCR to yield enough amounts of DNA for cloning and analysis is highly recommended especially for those specimens that are in a minute amount. Because the amplification process of PCR requires a pair of primers that can anneal to known sites at two end of the target DNA template, it seems that PCR is not applicable to IHJ searching because only one side of the fragment sequence in the integrant is known. A number of PCR-based techniques, however, have been developed to amplify the unknown cellular DNA flanking sequence from the foreign DNA. In this chapter, we introduce the PCR-based methodologies for the rapid acquisition of unknown DNA sequences. Based on the underlying principles, we classified these techniques into five categories: 1) PCR after intramolecular circularization; 2) interspersed repetitive sequence PCR (IRS-PCR); 3) ligation-anchored PCR (LA-PCR); 4) arbitrarily primed PCR (AP-PCR); and 5) reverse transcription PCR (RT-PCR). These techniques include inverse PCR (IPCR), partial IPCR (PI-PCR), long IPCR (LR-iPCR or LI-PCR), novel *Alu*-PCR, long interspersed repetitive element PCR (LINE-PCR), B1-PCR, vectorette-PCR, multiplestep-touchdown vectorette-PCR (MTV-PCR), long-distance vectorette-PCR

(LDV-PCR), splinkerette-PCR, thermal asymmetric interlaced PCR (TAIL-PCR), and retroviral LTR arbitrarily primed PCR (RELAP-PCR); and the capture PCR (C-PCR), which can improve the PCR amplification, is also discussed.

2. PCR After Intramolecular Circularization

A PCR technique, which is used to amplify an unknown DNA region adjacent to an integrated sequence after its intramolecular circularization, is called IPCR (inverse or inverted or inside-out PCR).

2.1. IPCR

The concept of IPCR first came in the 1980s (1–5) and has been in use for many years. The principle of this technique is illustrated in detailed in **Fig. 1**. IPCR begins with the digestion of genomic DNA with a restriction enzyme. Intramolecular self-ligation of restriction enzyme digested DNA fragments created a small monomeric circle. Within this circularized form of DNA, a conventional PCR technique is applied to amplify the IHJ region by using two opposite direction primers on the known integrant sequences called integrant specific primers (ISP). Hence, the primers are designed to anneal to the region of known sequence in IPCR. Generally, this technique has been used to characterize fragments up to 5 kb (6). However, a new DNA sequence or restriction sites could be used to start a new round of IPCR to obtain additional information. This strategy may be repeated to “walk” both upstream and downstream of a known DNA sequence (7). IPCR has been already applied to identify the integration sites of hepatitis B virus (HBV) (8,9), human T-lymphotrophic virus type-1 (HTLV-1) (10–15), human immunodeficiency virus type-1 (HIV-1) (16), and some transposable elements such as IS30 (17), T-DNA, *Ds* element (18), *dTph1* (19), Tn5 (20), Tn55 (21), and P element (22). The insertion of reticuloendotheliosis virus (REV) on Marek’s disease virus (MDV) genome is also identified by this method (23).

2.2. PI-PCR and Long-Distance IPCR

Some alternative IPCR methodologies have been published in recent years.

Partial inverse PCR (PI-PCR) (see **Fig. 2**) employs the genomic DNA partially digested by using 4-base recognition restriction enzyme (such as *Sau3A1*). After self-ligation, the circular DNA fragments are used as templates for IPCR (24). This is based on the preference of PCR for amplifying relatively smaller fragments. A DNA fragment that is less than 1 kb facilitates amplification via self-ligation and the IPCR process (9). A wide range of partial digests should be generated to find one that gives an optimal PCR amplification. Moreover, this approach eliminates the need to have any prior knowledge of restriction enzyme sites surrounding the integrant.

Long-distance IPCR method, designated long-range IPCR (LR-iPCR) (25) or long inverse PCR (LI-PCR) (26), enables the direct amplification of relatively large size flanking from circularized DNA fragments. The central to the long-distance IPCR is the use of a thermostable polymerase. This technique has been adapted to amplify relatively large-size flanking fragments up to 10 kb by using a highly thermal stable polymerase.

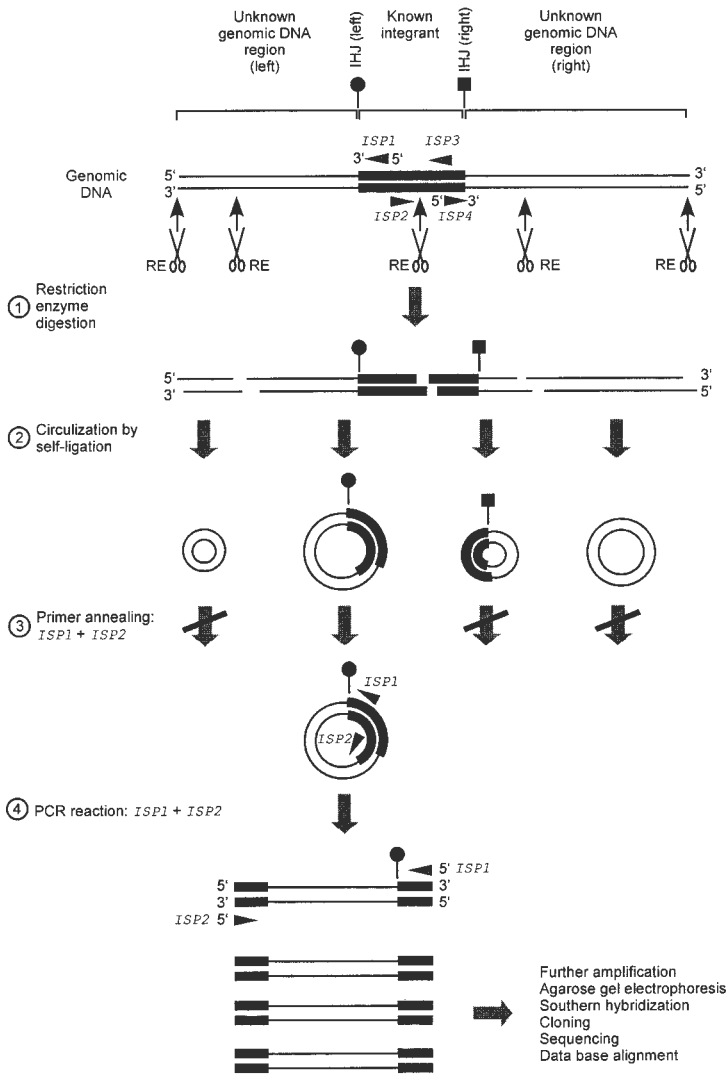


Fig. 1. Schematic flow diagram of IPCR protocol. Two complementary strands of genomic DNA have been shown at the top. The heavy and thin line regions represent the integrant fragment and cellular genomic unknown sequence, respectively. The positions of both left and right IHJs are indicated as closed circle and square signs, respectively. Integrant specific primers *ISP1*, 2, 3, and 4 for I-PCR are shown as arrowheads. In this particular example, the restriction enzyme cutting site (RE, scissors shape) is present in the integrant. *ISP1* and *ISP2* have been applied to amplify the left IHJ. The right IHJ has been amplified by using the other pair of IHJ using the primers *ISP3* and *ISP4*, under the same principle. If the RE is not present in the integrant, then any set of primer can be used to amplify both left and right IHJ (71). The slant lines on thick arrows indicate that no primer annealing will occur and no further amplification products. For detail manipulations see (9,17,97).

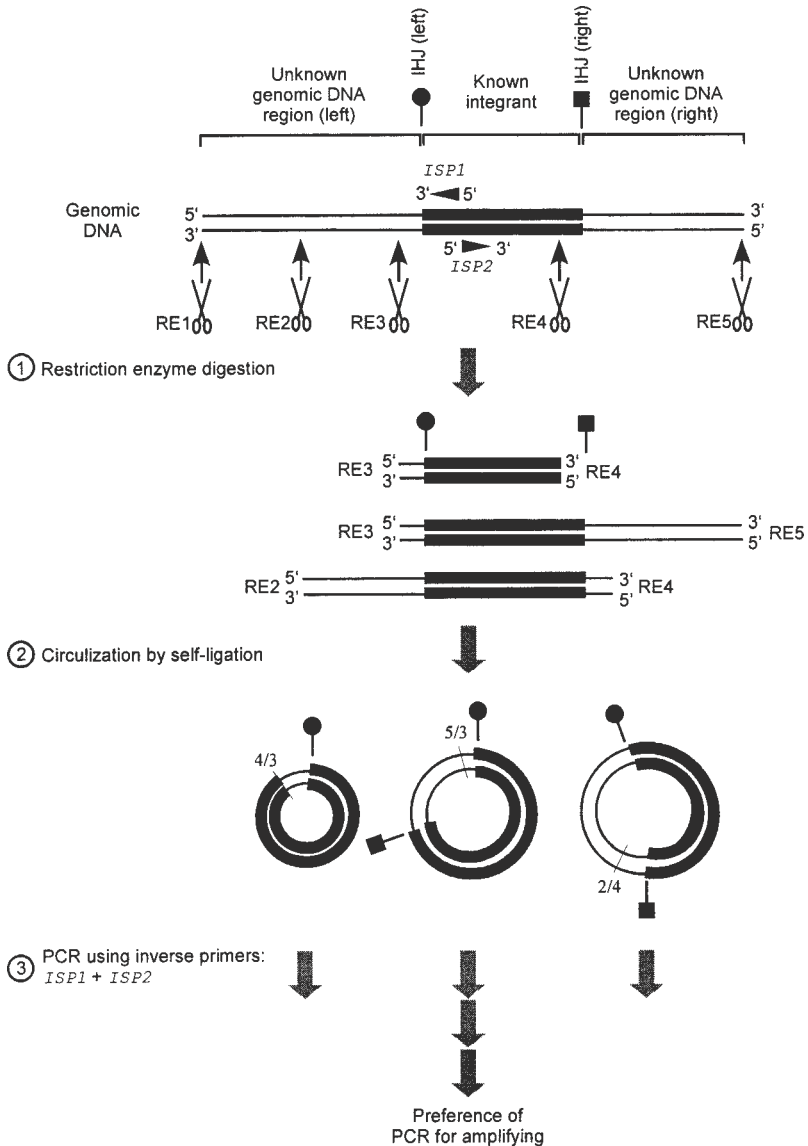


Fig. 2. Description of PI-PCR. Two complementary strands of genomic DNA have been shown at the top. The heavy and thin line regions represent the integrant fragment and cellular genomic unknown sequence, respectively. The positions of both left and right IHJs are indicated as closed circle and square signs, respectively. The different 4-base restriction enzyme recognition sites (RE, scissors shape) have been marked from 1 to 5. The partial digested DNA fragments are self-ligated and can be amplified from a known ISP, *ISP1* and *ISP2*. PCR is preference to amplify smaller fragments even if there are comparable amounts of large and small DNAs. For detail manipulations see (24).

2.3. Remarks for IPCR

The major advantage of IPCR is to amplify the flanking unknown sequence by using two known specific primers on the integrant, so-called integrant specific primer (ISP). The intramolecular circularization (self-ligation) of template is a key step for IPCR (*I*). This technique, occasionally, does not produce any product in a particular reaction, presumably because of the ineffective intramolecular ligation. This illustrates that the circularization step in IPCR is a fastidious procedure and not easy to optimize. Wang et al. (*9*) have demonstrated that by choosing the nearest restriction site, which can be determined by conventional PCR, gives a higher successful rate for cloning the IHJ by IPCR. Moreover, the noncircularized, intermolecular ligation, and free viral or transposon DNA fragment may interfere with the PCR. If intermolecular ligation had occurred, multiple PCR products would have been generated. To avoid intermolecular ligation (ligation between two digested fragments), the concentration of DNA has to be decreased and this results in a large volume for ligation. In addition, IPCR has been proved to be less sensitive than the other PCR-based methods (*II*). Therefore, IPCR requires a relatively large size of sample for compensating the low efficiency of ligation.

3. IRS-PCR-Based Techniques

The principle of IRS-PCR (interspersed repetitive sequence PCR) is based on the fact that the IRS elements are interspersed in the human genome. In this technique, amplification proceeds with one ISP to the known integrant sequence and the other primer specific to the known cellular interspersed repetitive sequence (IRS), which is distributed among the genome. IRSs are present in a high copy number in most multicellular organisms (*see Table 1*) (reviewed in *ref. 27*). The extension products from these specific primers include a segment containing the region of IHJ.

3.1. Novel *Alu*-PCR

Novel *Alu*-PCR (novel *Alu* element-mediated PCR) is an IRS-PCR, which uses a primer to *Alu* element and ISP for the PCR amplification. The overall strategy of novel *Alu*-PCR is outlined in *Fig. 3*. *Alu* elements are the largest family of short interspersed repetitive elements (SINEs) (*see Table 1*). The average density of *Alu* repeats on human genome is at the mean interval of about 3–6 kb, although *Alu* is not uniformly distributed (reviewed in *refs. 28–29*). This technique actually was first applied to amplify human genomes in the background of nonhuman genome and called “*Alu*-PCR” (*30–32*). Hence, extending the applicability of *Alu*-PCR, the inserted foreign sequence can be directly amplified between the known inserted sequence and the *Alu* consensus sequence, and to identify the IHJ (*33*).

Two specific primers are needed in novel *Alu*-PCR: ISP annealing to the known integrant sequence and the other to human *Alu* repeat sequences. In order to avoid illegitimate products, which are amplified from *Alu* sequences itself (*Alu*-*Alu* or inter-*Alu* amplification), two technical skills have been suggested (*33*). First, the primers should be synthesized by deoxyuridine triphosphates (dUTPs). This chemically modified primer can then be destroyed by uracil DNA glycosylase (UDG) after the first

Table 1
Organization of the Human Genome

DNA organization	% of total genome	Size of repeat unit	Copy #
Human genome			
I. Mitochondrial genome	0.0005 %		
II. Nuclear genome	99.9995 %		
A. Genes and gene-related sequences	approx 25.0 %		
1. Coding DNA	approx 2.5%		
2. Noncoding	approx 22.5%		
a. Introns, untranslated region, etc.			
b. Pseudogenes			
c. Gene fragments			
B. Extragenic DNA	approx 75.0 %		
1. Unique or low copy no.	approx 45.0 %		
2. Moderate to highly repetitive	approx 30.0 %		
a. Tandemly repeated/clustered repeats	approx 3.7 %		
i. Megasatellite DNA			
ii. Satellite DNA			
iii. Minisatellite DNA			
iv. Microsatellite DNA			
b. IRS approx 26.3 %			
i. SINE class	approx 8.7 %		
- <i>Alu</i> family	approx 7.0 %	Full: 280 bp	approx $0.5-1 \times 10^6$
- MIR families	approx 1.7 %	Average: 130 bp	approx 4×10^5
ii. LINE class	approx 10.6 %		
- LINE-1 (L1H or <i>Kpn</i>) family	approx 8.5 %	Full: 6 kb; Average: 0.8 kb	approx $1-5 \times 10^5$
- LINE-2 family	approx 2.1 %	Average: 250 bp	approx 2.7×10^5
iii. LTR class	approx 4.6 %		
- HERV/RTL V family	approx 1.3%	Average: 1.3 kb	approx 5×10^4
- THE-1, MER, and other families	approx 3.3%		approx 2×10^5
iv. DNA transposon	approx 1.6 %		approx 2×10^5
- <i>mariner</i> family		Average: 250 bp	
- Others			
v. Others	approx 0.8 %		approx 6×10^4

HERV: human endogenous retroviruses; IRS: interspersed repetitive sequence; LINE: long interspersed nuclear element; LTR: long terminal repeat; MER: medium reiteration frequency; MIR: mammalian-wide interspersed repeat; RTL V: retrovirus-like elements; SINE: short interspersed nuclear element; THE-1: transposable human element.

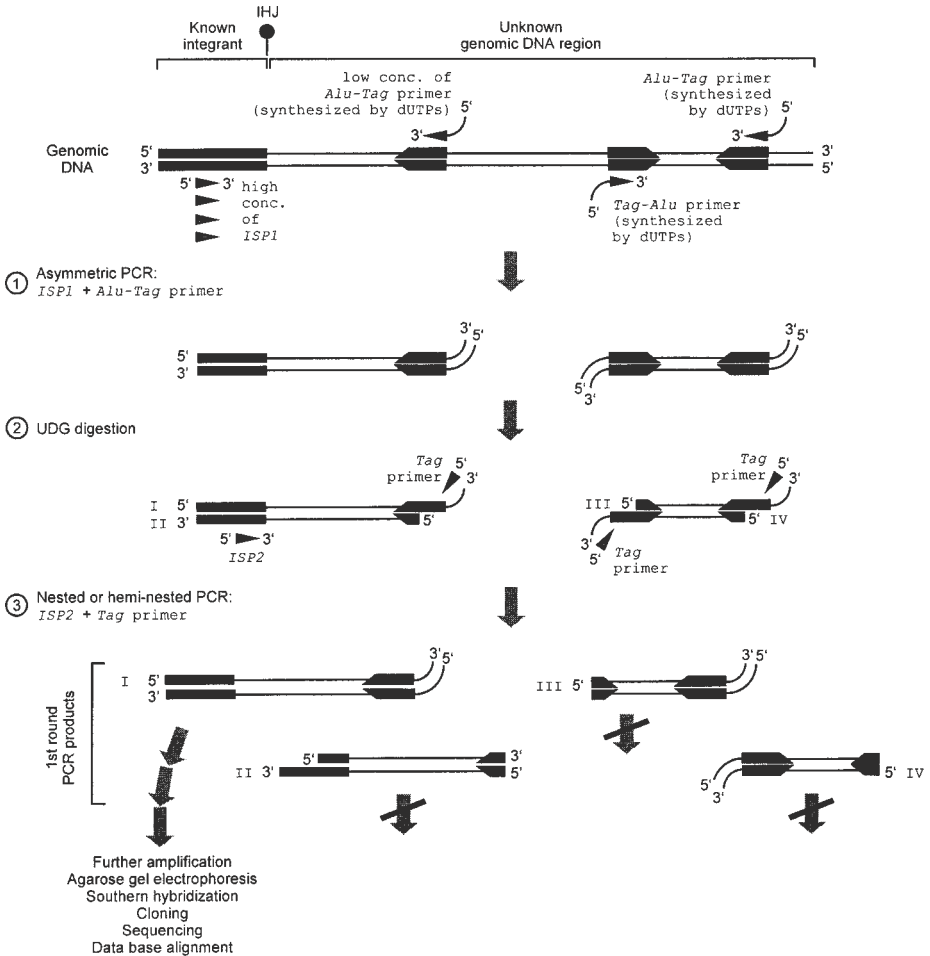


Fig. 3. Schematic diagram of the novel *Alu*-PCR. Two complementary strands of genomic DNA have been shown at the top. The heavy and thin line regions represent the foreign integrant fragment and cellular genomic unknown sequence, respectively. The position of IHJ is indicated by closed circle. Closed arrow boxes represent *Alu* elements on the human genome in different orientations. Primers for PCR, *ISP1*, *ISP2*, *Alu-Tag*, and *Tag* primers, are shown as arrowheads. The non-annealing *Tag* region on *Alu-Tag* primer is shown by curved thin line. Human genomic DNA was amplified by using the first set of primers: *ISP1* and *Alu-Tag* primer (step 1). After an initial 10 cycles of PCR, the *Alu-Tag* primers on new synthesized DNA is destroyed by UDG (step 2). These digest PCR products are further amplified by using an internal primer: *ISP2* and *Tag* primers (step 3; nested PCR). Only DNA product I, which still contains both *ISP2* and *Tag* nested primer target sites, will be amplified further. The cross marks on thick arrows indicate that no primer annealing will occur and no further amplification products. For detail manipulations see (34).

10–15 cycles of amplification. Such modification can break the *Alu-Alu* specific amplification (see **Fig. 3, step 2**). Second, an asymmetric amplification (unequal ratio of two primers) is performed before UDG treatment (see **Fig. 3, step 1**). The primer on the known integrant sequence is added at least 10-fold higher concentration than the primer for the *Alu* sequence (34). In general, during the first 10–20 cycles, dsDNA products are generated. But when the limiting primer is exhausted, ssDNA is produced for the next cycles by primer extension (35,36). No matter the accumulation of dsDNA and ssDNA, the products of integrant-*Alu* amplification are higher than the *Alu-Alu* products, and thus this asymmetric PCR does not favor *Alu-Alu* amplification.

In addition, the design of the primer contains a tag sequence, which can be applied to the other standard PCR protocols such as the nested or hemi-nested PCR (see **Fig. 3, step 3**), to decrease the nonspecific amplification of PCR. Moreover, a single primer control can exclude the false-positive amplification and Southern hybridization has been suggested to facilitate cloning. Some investigators, by using novel *Alu*-PCR, have successfully identified cellular sequences flanking from integrated HBV (34), HIV-1 (37,38), and human papillomavirus type-16 (HPV-16) (39) DNA. The adeno-associated virus (AAV) vector insertion for gene therapy, in addition, was also detected by this way (40).

3.2. B1-PCR and LINE-PCR

The *Alu* repeat is primate-specific but other mammals have similar types of sequence such as the B1 family in mouse. Thus, the novel *Alu*-PCR equivalent, B1-PCR has been applied to find the AAV vector integration site from the rat tissues (40).

Based on the same principles, the *Alu* primer can be replaced by the others primers, which can anneal to the other genomic repetitive sequences. In addition to the SINE (such as *Alu* sequence) applied in novel-*Alu* PCR, other IRS has also been used under the same approach such as long interspersed nuclear element (LINE), and so-called LINE-PCR. LINE-PCR has been applied to identify the HPV integration site (39).

3.3. Remarks for IRS-PCR

The IRS-PCR offers at least four advantages over IPCR. First, less amount of DNA is required. Second, in contrast to IPCR, an intramolecular ligation reaction is not required in the IRS-PCR. This can overcome the low efficiency of the self-ligation reaction. Third, IRS-PCR is based on only two steps: UDG digestion and conventional PCR procedures, thereby saving most time. Fourth, IRS-PCR avoids the attendant problems in interpretation resulting from episomal contamination.

However, this technique is not suitable for the case of IRS elements within a short distance to the integrant. Unfortunately, many virus genomes tend to insert adjacent to or into repetitive sequence, such as HBV (41–46), simian virus 40 (SV40) (47), murine leukemia virus (MuLV) (48), hamster endogenous retrovirus (49), HIV-1 (16,50–51), HPV-16 (52–54), woodchuck hepatitis virus (WHV) (55), and duck hepatitis B virus (DHBV) (56). Furthermore, the use of IRS-PCR is also limited by the requirement for the adjacent repeat sequences to be in the correct orientation.

4. LA-PCR-Based Techniques

LA-PCR (ligation-anchored PCR or cassette ligation-anchored PCR) is based on the ligation of an oligonucleotide cassette unit, called adapter or linker, to the cleaved

genomic fragments (57). The amplification proceeds with ISP to the known integrant sequence and the other primer specific to the known ligated adapter. Indeed, the principle of single-specific-primer PCR (SSP-PCR) (58), rapid amplification of cDNA ends (RACEs) (reviewed in ref. 59), and rapid amplification of genomic DNA ends (RAGE) (60) have been developed using the similar concept. In these PCR methods, the ligated unit enables PCR to amplify the DNA fragment between itself and a known primer from known integrant sequence.

4.1. LM-PCR

In ligation mediated-PCR (LM-PCR), genomic DNA is digested by a restriction enzyme and ligated with a primer using T4 DNA ligase. Then, ISP and the ligated primer are used in a classical PCR amplification. This protocol has been applied to the samples from HTLV-I integration genome (11,61–62).

4.2. Vectorette-PCR

The unique feature of vectorette-PCR method is the special secondary structure of the cassette, which termed vectorette unit (see Fig. 4A). The vectorette unit contains a central non-complementarity mismatched region resulting in a bubble-shape (see Fig. 4A), therefore, the vectorette-PCR is also termed “bubble PCR.” VectorettePCR was first used for rapid isolation of terminal sequences from yeast artificial chromosome (YAC) clones (63), and then applied to the intronic DNA sequence characterization (64). The procedure of vectorette-PCR begins with the digestion of genomic DNA with a restriction enzyme to generate a 5'-overhang, and then ligation with a vectorette unit. The flanking sequences are then amplified by using an ISP of the integrant and the universal vectorette specific primer. The amplification strategy is summarized in Fig. 4B. The vectorette primer, which applied to the PCR is, actually, identity sequence to, but not complementary to, the noncomplementarity mismatched region of vectorette unit and therefore it only process PCR extension from the second round of the reaction. This enhances the PCR amplification specific to the IHJ containing genomic fragment. The HIV integration site has been identified by this method (16).

4.3. MTV- and LDV-PCRs

Vectorette-PCR, in addition, has been modified to be a multistep-touchdown vectorette-PCR (MTV-PCR) (65), which is suitable for analysis of the high CG content region. MTV-PCR starts at a hot-start technique and proceeds at touchdown PCR cycle profile. Because the high GC content DNA results in the PCR conflicting secondary structures, the application of touchdown cycling parameters prevent significantly the formation of unspecific DNA fragments.

A vectorette-based long-distance PCR has been developed to amplify the fragment up to 5 kb (66), and so-called long-distance vectorette-PCR (LDV-PCR). The use of a mixture of thermostable DNA polymerases is central to this approach.

4.4. Splinkerette-PCR

However, undesirable amplifications of nonspecific “end-repair priming” may involve the free cohesive ends of unligated free vectorettes and 5'-overhangs of unknown cellular region (see Fig. 4C). These ends are filled during the first cycle of PCR. After the denaturing step, these ends are able to anneal together (as shown in step 4 of

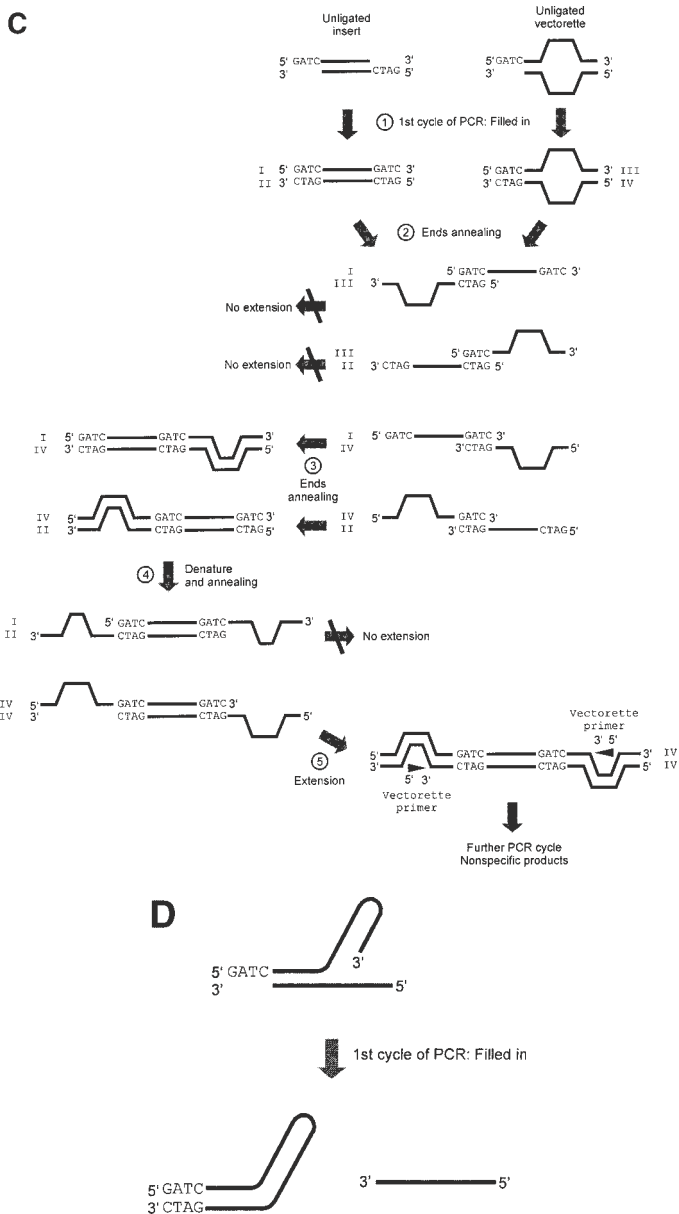


Fig. 4. (A, facing page) Structure of the vectorette and splinkerette units. The primer for vectorette and splinkerette are also shown. (B, facing page) Schematic representation of the principle of vectorette-PCR. Two complementary strands of DNA are shown at the top. The heavy and thin line regions represent the foreign integrant fragment and cellular genomic unknown sequence, respectively. The position of IHJ is indicated by closed circle. *ISP* and vectorette primers for PCR are shown as arrowheads. The slant lines on thick arrows indicate that no primer annealing will occur and no further amplification products. For detail manipulations see (63). (C) A diagram showing the effect of “end-repair priming” in vectorette-PCR. (D) The splinkerette unit do not has “end-repair priming” effect.

Fig. 4C). In this procedure, the complementary strand of vectorette primer is generated from unwanted fragments, and this decrease the specificity of vectorettes-PCR. The splinkerette is therefore designed as a hairpin structure on one strand rather than a central DNA mismatch (67), as compared in **Fig. 4A**. The advantage of splinkerette-PCR over vectorette-PCR is the elimination of the end-repair priming phenomena (see **Fig. 4D**). Some researchers have successfully identified flanking regions of transposon *Sleeping Beauty* (*Tc1/mariner* superfamily) by using this method (68).

4.5. Remarks for LA-PCR

LM-PCR has been proved to be more sensitive than IPCR in the detection of integrant (11). Some commercial products (Invitrogen and TaKaRa), which use PCR technology to quickly identify the unknown sequence, are also based on the principle for LA-PCR. However, it requires a proper ligation between oligomer linker to genomic DNA fragments. The ratio of linker DNA and genomic DNA has to be serially diluted to obtain a maximum intermolecular ligation (58).

5. AP-PCR-Based Techniques

The principle of AP-PCR (arbitrarily primed PCR) is using nonspecific arbitrary primers for PCR amplification (69). Following this “hemispecific” concept, the targeted gene walking PCR (70; reviewed in 71), single primer reaction (72), differential display PCR (DD-PCR) (73; reviewed in 74,75), and restriction site PCR (RS-PCR) (76; reviewed in 71) have been developed for the amplification of DNA sequences by nonspecific arbitrary primers.

5.1. TAIL-PCR

TAIL-PCR utilizes three nested specific primers on known integrant in successive three rounds of PCRs together with a shorter arbitrary degenerate (AD) primer. The basis for this strategy is thermal asymmetric PCR. Arbitrary priming creates nontarget molecules, because degenerate primers that hybridize randomly in genomic DNA and constitute the bulk of the final unwanted products. The interspersing asymmetric and symmetric PCR cycles are used geometrically to favor amplification of target molecules over nonspecific products.

A schematic diagram of targeted TAIL-PCR is shown in **Fig. 5**. During the high-stringency cycle at first round of PCR (high-stringency PCR program at step 1 in **Fig. 5**) only the long integrant specific primer *ISR1* can efficiently anneal to the DNA template, therefore only specific product (product I in **Fig. 5**) is amplified and little or no nontarget sequence product (which is primed at both ends by *AD* primers; product II in **Fig. 5**) has been formed. In the following single reduced-stringency cycle (low-stringency PCR program at step 1 in **Fig. 5**), however, both *ISR1* and *AD* primers can anneal to the template DNA. The single-stranded target DNA, which is produced during last high-stringency cycles is replicated to dsDNA and hence providing a several-fold increase of target template for the next round of amplification. In following TAIL-cycling (TAIL programme at step 1 in **Fig. 5**), the specific product (product I) is

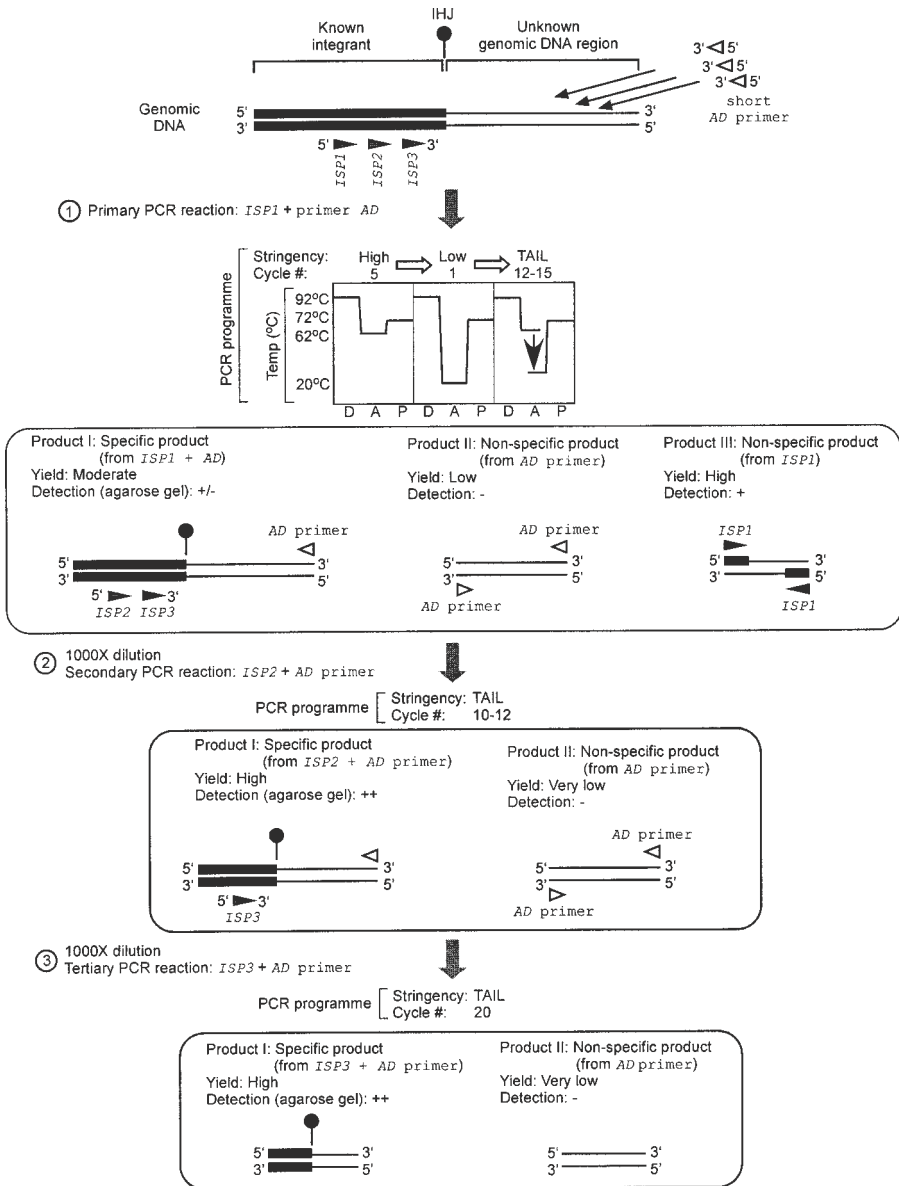


Fig. 5. Summary of the TAIL-PCR procedure. Two complementary strands of DNA have been shown at the top. The heavy and thin line region represent the integrant fragment and cellular genomic unknown sequence, respectively. The position of IHJ is indicated by closed circle. *ISP* and *AD* (arbitrary degenerate) primers for PCR are shown as close and open arrow-heads, respectively. The PCR program with different stringencies and cycle number is shown on the top of the box. D, A, and P represents denature, annealing, and polymerization step, respectively, in PCR cycle. For detail manipulation see (84).

possible to be amplified preferentially over nontarget sequence product. But at the same time, nonspecific product (at both ends by the long *ISP1* primers; product III in **Fig. 5**) can also arise efficiently through mispriming. Such undesired products are diluted out, however, in subsequent secondary (**step 2**) and tertiary (**step 3**) round PCR, which using internally nested specific primers *ISP2* and *ISP3*, respectively. The TAIL cycling in both secondary and tertiary is performed in lower background further. In fact, the T-DNA insertion (**77,78**), Ds elements (**79,80**), and *Tto1* introduced (**81**) in *Arabidopsis* have been identified by this technique.

5.2. RELAP-PCR

The combination of a long ISP designed to detect retroviral long terminal repeat (LTR) and a short arbitrary primer (AD primer) from the AD primer set in different lengths binding in a random fashion under a low-stringency condition are used (*see Fig. 6*, upper panel). Therefore, AP-PCR had been adapted to allow the amplification of LTR-containing retrovirus integration site. This is called RELAP-PCR (retroviral LTR-arbitrarily primed PCR) and has been applied to identify the integration site of mouse mammary tumor virus (MMTV) (**82**).

Hot spot-combined PCR (HS-cPCR), modified AP-PCR for retrovirus, is based on previous finding regarding the host spot of retroviral LTR integration sites. In this technique, the primers have been designed to target on both known retroviral LTR and nonintegrant region in different combination (**83**). It is possible to design primers, which border the “suspected” fragment.

5.3. Remarks for AP-PCR

These AP-PCR methods allow rapid detection without any DNA manipulation before PCR, such as restriction enzyme digestion or ligation. Amplification occurs either upstream or downstream from a known sequence. In TAIL-PCR, a set of nested long primer and a short arbitrary primer are important (**77,84**). Besides the primer design, the stringency in the primer-template interaction is an important parameter of this class of PCR (**85**). Specificity of the amplification reaction has been further confirmed by Southern blotting of the PCR products. The single primer control is always necessary and important to exclude the false positive results.

In the case of RELAP-PCR, a series of walking primers have been designed to increase the incidence of positive results. In addition, a series of walking reaction are usually done in parallel. This can be laborious and time-consuming. Moreover, this strategy is only suitable for the LTR-containing integrant.

6. RT-PCR-Based Technique

In the case of retrovirus, the promoter activation within 3' untranslated LTR initiates chimeric mRNA transcripts, which consist the viral LTR and the cellular gene fragment in the same transcriptional orientation (**86,87**). Based on this property of retrovirus, the poly(A)-tail containing mRNA is purified and cDNA is synthesized by reverse transcription using an oligo(dT)-adaptor primer which primers on the poly(A)

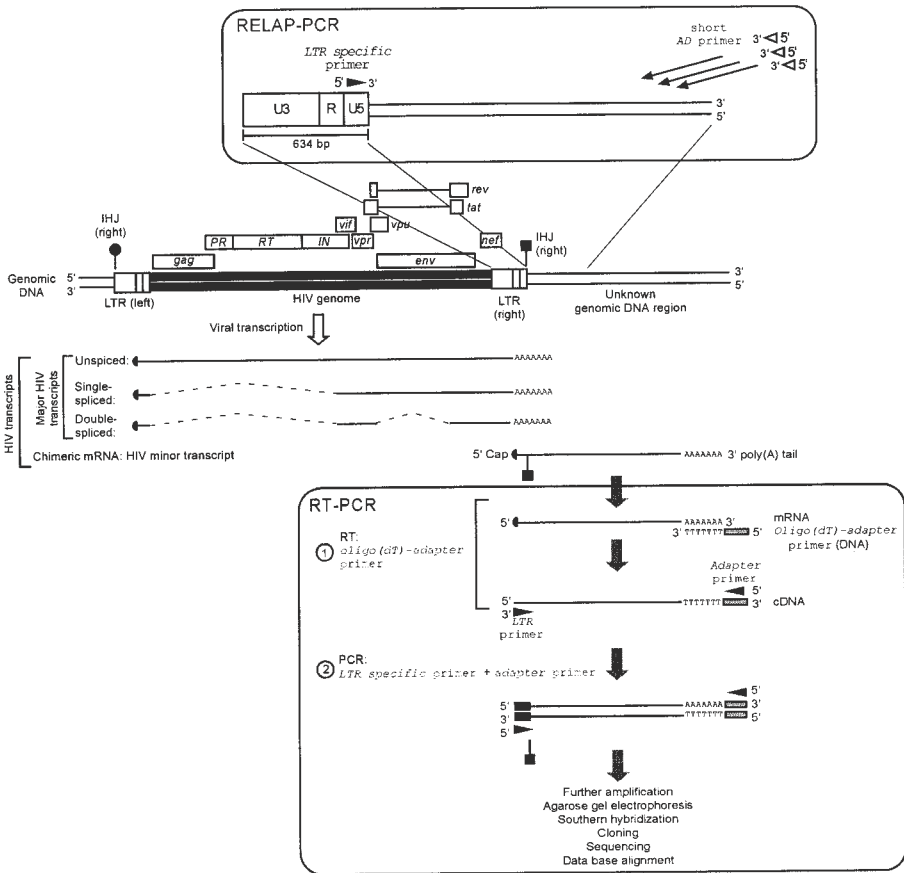


Fig. 6. Summary of the procedure for RELAP- and RT-PCR in the studying on retrovirus insertion. Two complementary strands of HIV DNA have been shown at the middle. The heavy and thin line region represent the integrated HIV fragment and cellular genomic unknown sequence, respectively. The position of left and right IHJ is indicated by closed circle and square, respectively. Boxes with HIV open reading frames (ORFs) are shown. The right and left LTR region of HIV are shown as an open box (the size is not to scale). Primers for PCR are shown as arrowheads. Upper panel is the schematic diagram for RELAY-PCR. For detail manipulation for *see* (82). Lower panel is the schematic flow diagram for RT-PCR based protocol. The major and minor viral transcripts from HIV viral transcription are shown. Only the cDNA from chimeric mRNA contains LTR primer target site. For detail manipulation see (90).

tails. Using the adapter primer and an LTR-specific primer, the chimeric mRNA containing the retroviral insertion site is amplified by PCR. The overall RT-PCR based method to isolate these chimeric cDNAs is schematically shown in the lower panel in **Fig. 6**. This principle is similar as anchored PCR (A-PCR) (88), one-sided PCR (89), and RACE. This RT-PCR based method is rapid and simple, and successful to identify the retrovirus integration site (90). However, this method only works on the virus,

cellular genes that are involved in cell proliferation and differentiation (95,96). However, genomic cloning method requires the establishment of genomic DNA libraries, which is time-consuming and laborious. Therefore, many PCR-based techniques have been developed for the elucidation of unknown flanking DNA sequence adjacent to a region of known integrant sequence. The genomic sequences flanking foreign integrant can then be determined rapidly with these techniques (within 1 wk). PCR-based techniques offer an inexpensive and flexible alternative to IHJ searching, and can be performed in any laboratory equipped with basic molecular biology. The limitation of PCR is the need for the sequence of two target specific primers that flank the region that is intended for amplification. The problem here is how to allow the direct amplification of DNA without a prior knowledge of sequence information. Several strategies have resolved this limitation. IPCR can amplify flanking region directed away from the core region of known integrant sequence after DNA self-ligation (circularization). IRS-PCR depends on the distribution of IRS on the genome. Many techniques for amplifying flanking unknown regions of DNA are based on the creation of new primer binding sites on the potential PCR template by ligating oligonucleotide linkers or cassette unit of known sequences to the ends of DNA fragments, such as LA-PCR. And some other techniques allow primer binding in a random fashion under low-stringency condition, such as AP-PCR. All methods described in this chapter are compared in **Table 2**.

Beside the IHJ searching, these techniques can also be applied to the determination for YAC end points (31,63), cDNA ends (97), genomic breakpoints (deletion or translocation) (66,98), intron-exon junctions (99), gene rearrangements (6), promoter sequence sequences (100,101), mating-type gene switching (102), and gene-targeting vector construction (103). Although, some PCR-based techniques such as RAGE (104), RS-PCR (76; reviewed in 71), panhandle PCR (105,106; reviewed in 107), multiplex RS-PCR (108), gene walking PCR (109), homo-oligomeric tailing based PCR (110), novel step-down PCR (111), and nonspecifically primed suppression PCR (NSPS-PCR) (112) have not yet been used to IHJ study, these principles are also applicable in the integration site searching.

Among all techniques introduced in this chapter, the IRS-PCR- and TAIL-PCR-based methods are highly recommended to apply in the integrant seeking. One reason is the high sensitivity of these two methods. Second, genomic DNA manipulation, such as DNA digestion and ligation reaction, is not required. Third, only the simple straightforward technique of conventional PCR protocol is needed. In fact, each of the approaches to identify IHJ is useful, and the experimental context is the critical feature that determines success. If an experiment is poorly designed (especially the primer set sequence) or the sample is contaminated, the result is a large number of bands after PCR, that are difficult and time-consuming to analyze. Any inefficiencies, mispriming, or incomplete reaction in the PCR or restriction enzyme digestion steps can result in artifacts that are misleading. Some improvements, therefore, are applied in the PCR reaction to increase the specificity of PCR amplification. These are hot start PCR (in any PCR techniques), nested PCR (in IRS-PCR, TAIL-PCR), touch-

Table 2
Methods Described in this Chapter

Principle Base on	Methods	DNA digestion	DNA ligation	Amount of DNA used	Primers on	Step to optimize	Source of nonspecific products	Sensitivity
Inverse	IPCR (Fig. 1) PI-PCR (Fig. 2) LR-iPCR LI-PCR	+	+(self-ligation)	D: 0.2–10 µg L: 0.2-5 ng/µL	Integrand	1. Relative high DNA sample amount 2. DNA concentration of self-ligation	1. Unintegrated DNA 2. Unligated DNA	Low
IRS-PCR	Novel <i>Alu</i> -PCR (Fig. 3) LINE-PCR B1-PCR	–	–	P: 10–106 ng	1. Integrand 2. IRS (eg. <i>Alu</i> , LINE, B1)	1. IRS orientation 2. Distance to IRS	Inter-IRS amplification	High
LA-PCR	LM-PCR Vectorette-PCR (Fig. 4B) MTV-PCR LDV-PCR Splinkerette-PCR Capture (Fig. 7B)	+	+	D: 0.5-2 µg L: 0.5–2 µg	1. Integrand 2. Adapter unit	1. Ligation efficiency 2. Adapter unit design (Fig. 4C)	End-repair priming	Medium
AP-PCR	TAIL-PCR (Fig. 5) RELAP-PCR (Fig. 6) Capture (Fig. 7A)	–	–	P: 20–150 ng	1. Integrand 2. Arbitrary primer	Primer set design	Arbitrary priming	High
RT-PCR	RT-PCR (Fig. 6)	–	+	R: 3 µg poly(A) ⁺ RNA	1. LTR (integrand) 2. Adapter	cDNA synthesis		Medium

D: DNA digestion reaction; L: ligation reaction; P: PCR reaction; R: RT reaction.

down PCR (in IRS-PCR, MTV-PCR, and TAIL-PCR), specific primer cassette structure (in vectorette- and splinkerette-PCR), asymmetric (or unequal) ratio of the two amplification primers is used (in IRS-PCR and TAIL-PCR), primer synthesis by dUTP and then digest by UDG (in IRS-PCR), isolated biotinylated products (in C-PCR), and -NH₂ and -PO₄ groups modification on the adaptors to prevent nonspecific 3' end elongation during PCR reaction (in novel step down PCR). These modifications are pivotal to successful amplification. It might also be helpful to optimize the PCR conditions with respect to magnesium and dimethyl-sulfoxide (DMSO) concentrations according to standard protocols. Although false positive results could happen, both single primer control and Southern blotting would minimize this problem and confirm the specificity of the amplified products.

References

1. Collins, F. S. and Weissman, S. M. (1984) Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. *Proc. Natl. Acad. Sci. USA* **81**, 6812–6816.
2. Ochman H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–623.
3. Ochman, H., Ajioka, J. W., Garza, D., and Hartl, D. L. (1989) Amplification of flanking sequences by inverse PCR, in *PCR Technology* Ehrlich, H. A., ed., Stockton, New York, pp. 105–111.
4. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequence. *Nucl. Acid Res.* **16**, 8186.
5. Silver, J. and Keerikatte, V. (1989) Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. *J. Virol.* **63**, 1924–1928.
6. Willis, T. G., Jadayel, D. M., Coignet, L. J. A., Abdul-Rauf, M., Treleaven, T. J., Catorsky, D., and Byer, M. J. S. (1997) Rapid molecular cloning of rearrangements of the IGHJ locus using long-distance inverse polymerase chain reaction. *Blood* **90**, 2456–2462.
7. Garces, J. A. and Gavin, R. H. (2001) Using an inverse PCR strategy to clone large contiguous genomic DNA fragments. *Meth. Mol. Biol.* **161**, 3–8.
8. Tsuei, D.-J., Chen, P.-J., Lai, M.-Y., Chen, D.S., Yang, C.-S., Chen, J.-Y., and Hsu, T.-Y. (1994) Inverse polymerase chain reaction for cloning cellular sequences adjacent to integrated hepatitis B virus DNA in hepatocellular carcinomas. *J. Virol. Meth.* **49**, 269–284.
9. Wang, P.-C., Hui, E. K.-W., Chiu, J.-H., and Lo, S. J. (2001) Analysis of integrated hepatitis B virus DNA and flanking cellular sequence by inverse polymerase chain reaction. *J. Virol. Meth.* **92**, 83–90.
10. Takemoto, S., Matsuoaka, M., Yamaguchi, K., and Takatsuki, K. (1994) A novel diagnostic method of adult T-cell leukemia: monoclonal integration of human T-cell lymphotropic virus type I provirus DNA detected by inverse polymerase chain reaction. *Blood* **84**, 3080–3085.
11. Cavois, M., Wain-Hobson, S., and Wattel, E. (1995) Stochastic events in the amplification of HTLV-I integration sites by linker-mediated PCR. *Res. Virol.* **146**, 179–184.
12. Cavois, M., Wain-Hobson, S., Gessain, A., Plumelle, Y., and Wattel, E. (1996) Adult T-cell leukemia/lymphoma on a background of clonally expanding HTLV-1 positive cells. *Blood* **88**, 4646–4650.
13. Ohshima, K., Suzumiya, J., Kato, A., Tashiro, K., and Kikuchi, M. (1997) Clonal HTLV-I-infected CD4⁺ T-lymphocytes and non-clonal non-HTLV-infected giant cells in incipient ATLL with Hodgkin-like histologic features. *Int. J. Cancer* **72**, 592–598.

14. Ohshima, K., Mukai, Y., Shiraki, H., Suzumiya, J., Tashiro, K., and Kikuchi, M. (1997) Clonal integration and expression of human T-cell lymphotropic virus type I in carriers detected by polymerase chain reaction and inverse PCR. *Am. J. Hematol.* **54**, 306–312.
15. Leclercq, I., Cavois, M., Mortreux, F., Hermine, O., and Gessain, A. (1998) Oligoclonal proliferation of human T-cell leukemia virus type 1 bearing T cells in adult T-cell leukemia/lymphoma without deletion of the 3' provirus integration sites. *J. Haematol.* **101**, 500–506.
16. Carteau, S., Hoffmann, C., and Bushman, F. (1998) Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. *J. Virol.* **72**, 4005–4014.
17. Ochman, H., Ayala, F. J., and Hartl, D. L. (1993) Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. *Meth. Enzymol.* **218**, 309–321.
18. Knapp, S., Larondelle, Y., Robberg, M., Furtek, D., and Theres, K. (1994) Transgenic tomato lines containing Ds elements at defined genomic positions as tools for targeted transposon tagging. *Mol. Gen. Genet.* **243**, 666–673.
19. Souer, E., Quattrocchio, F., de Vetten, N., Mol, J., and Koes, R. (1995) A general method to isolate genes tagged by a high copy number transposable element. *Plant J.* **7**, 677–685.
20. Martin, V. J. and Mohn, W. W. (1999) An alternative inverse PCR (IPCR) method to amplify DNA sequence flanking Tn5 transposon insertions. *J. Microbiol. Meth.* **35**, 163–166.
21. De Lencastre, H., Wu, S. W., Pinho, M. G., Ludovice, A. M., Filipe, S., Gardete, S., et al. (1999) Antibiotic resistance as a stress response: complete sequencing of a large number of chromosomal loci in *Staphylococcus aureus* strain COL that impact on the expression of resistance to methicillin. *Microbiol. Drug Res.* **5**, 163–175.
22. Liao, G.-C., Rehm, E. J., and Rubin, G. M. (2000) Insertion site preferences of the P transposable element in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **97**, 3347–3351.
23. Isfort, R., Jones, D., Kost, R., Witter, R., and Kung, H.-J. (1992) Retrovirus insertion into herpesvirus in vitro and in vivo. *Proc. Natl. Acad. Sci. USA* **89**, 991–995.
24. Pang, K. M. and Knecht, D. A. (1997) Partial inverse PCR: a technique for cloning flanking sequences. *BioTechniques* **22**, 1046–1048.
25. Mathur, J., Szabados, L., Schaefer, S., Grunenberg, B., Lossow, A., Jonas-Straube, E., et al. (1998) Gene identification with sequenced T-DNA tags generated by transformation of *Arabidopsis* cell suspension. *Plant J.* **13**, 707–716.
26. Raponi, M., Dawes, I. W., and Arndt, G. M. (2000) Characterization of flanking sequences using long inverse PCR. *BioTechniques* **28**, 839–843.
27. Smit, A. F. (1996) The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.* **6**, 743–748.
28. Mighell, A. J., Markham, A. F., and Robinson, P. A. (1997) Alu sequence. *FEBS Lett.* **417**, 1–5.
29. Szmulewicz, M. N., Novick, G. E., and Herrera, R. J. (1998) Effects of Alu insertions on gene function. *Electrophoresis* **19**, 1260–1264.
30. Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirez-Solis, R., Webster, T. D., et al. (1989) Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci. USA* **86**, 6686–6690.
31. Nelson, D. L., Ballabio, A., Victoria, M. F., Pieretti, M., Bies, R. D., Gibbs, R. A., et al. (1991) Alu-primer polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: identification of clones associated with a disease locus. *Proc. Natl. Acad. Sci. USA* **88**, 6157–6161.
32. Ledbetter, S. A., Nelson, D. L., Warren, S. T., and Ledbetter, D. H. (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics* **6**, 475–481.

33. Puskas, L. G., Fartmann B., and Bottka S. (1994) Restricted PCR: amplification of an individual sequence flanked by a highly repetitive element from total human DNA. *Nucl. Acids Res.* **22**, 3251–3252.
34. Minami M., Poussin K., Brechot C., and Paterlini P. (1995) A novel PCR technique using Alu-specific primers to identify unknown flanking sequences from the human genome. *Genomics* **29**, 403–408.
35. Gyllensten, U. B. and Erlich, H. A. (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* **85**, 7652–7656.
36. Gyllensten, U. B. and Allen, M. (1993) Sequencing of in vitro amplified DNA. *Meth. Enzymol.* **218**, 1–16.
37. Courcoul, M., Patience, C., Rey F., Blanc, D., Harmache, A., Sire, J., et al. (1995) Peripheral blood mononuclear cells produce normal amounts of detectable Vif- human immunodeficiency virus type 1 particles which are restricted for the preretrotranscription steps. *J. Virol.* **69**, 2068–2074.
38. Sonza, S., Maerz, A., Deacon, N., Meanger, J., Mills, J., and Crowe, S. (1996) Human immunodeficiency virus type 1 replication is blocked prior to reverse transcription and integration in freshly isolated peripheral blood monocytes. *J. Virol.* **70**, 3863–3869.
39. Carmody, M. W., Jones, M., Tarraza H., and Vary, C.P. (1996) Use of the polymerase chain reaction to specifically amplify integrated HPV-16 DNA by virtue of its linkage to interspersed repetitive DNA. *Mol. Cell. Probes* **10**, 107–116.
40. Wu, P., Phillips, M.I., Bui, J., and Terwilliger, E.F. (1998) Adeno-associated virus vector-mediated transgene integration into neurons and other nondividing cell targets. *J. Virol.* **72**, 5919–5926.
41. Choo, K. B., Liu, M. S., Chang, P. C., Wu, S. M., Su, M. W., Pan C. C., and Han, S. H. (1986) Analysis of six distinct integrated hepatitis B virus sequences cloned from the cellular DNA of a human hepatocellular carcinoma. *Virology* **154**, 405–408.
42. Shaul, Y., Garcia, P. D., Schonberg, S., and Rutter, W. J. (1986) Integration of hepatitis B virus DNA in chromosome-specific satellite sequences. *J. Virol.* **59**, 731–734.
43. Nagaya, T., Nakamura, T., Tokino, T., Tsurimoto, T., Imai, M., Mayumi, T., et al. (1987) The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma. *Genes Deve.* **1**, 773–782.
44. Matsumoto H., Yoneyama T., Mitamura K., Osuga T., Shimojo H., and Miyamura, T. (1988) Analysis of integrated hepatitis B virus DNA and cellular flanking sequences cloned from a hepatocellular carcinoma. *Int. J. Cancer* **42**, 1–6.
45. Quada, K., Saldanha, J., Thomas, H., and Monjardino, J. (1992) Integration of hepatitis B virus DNA through a mutational hot spot within the cohesive region in a case of hepatocellular carcinoma. *J. Gen. Virol.* **73**, 179–183.
46. Chen J. Y., Harrison, T. J., Tsuei, D. J., Hsu, T. Z., Zuckerman, A. J., Chan, T. S., and Yang, C. S. (1994) Analysis of integrated hepatitis B virus DNA and flanking cellular sequences in the hepatocellular carcinoma cell line HCC36. *Intervirology* **37**, 41–46.
47. Dhruva, B. R., Shenk, T., and Subramanian, K. W. (1980) Integration in vivo into simian virus 40 DNA of a sequence that resembles a certain family of genomic interspersed repeated sequences. *Proc. Natl. Acad. Sci. USA* **77**, 4514–4518.
48. Ou, C.-Y., Boone, L. R., and Yang, W. K. (1983) A novel sequence segment and other nucleotide structural features in the long terminal repeat of a BALB/c mouse genomic leukemia virus-related DNA clone. *Nucl. Acids Res.* **11**, 5603–5620.

49. Taruscio, D. and Maneulidis, L. (1991) Integration site preferences of endogenous retroviruses. *Chromosoma* **101**, 141–156.
50. Stevens, S. W. and Griffith, J. D. (1994) Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1Hs repetitive elements. *Proc. Natl. Acad. Sci. USA* **91**, 5557–5561.
51. Stevens, S. W. and Griffith, J. D. (1996) Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. *J. Virol.* **70**, 6459–6462.
52. Awady, M. K., Kaplan, J. B., O'Brien, S. T., and Burd, R. D. (1987) Molecular analysis of integrated human HPV16 sequences in the cervical cancer cell line SiHa. *Virology* **159**, 389–398.
53. Baker, C. C., Phelps, W. C., Lindgren, V., Braun, M. J., Gonda, M. A., and Howley, P. M. (1987) Structural and transcriptional analysis of HPV16 sequences in cervical carcinoma cell lines. *J. Virol.* **61**, 962–971.
54. Wagatsuma, M., Hashimoto, K., and Matsukura, T. (1990) Analysis of integrated human HPV16 DNA in cervical cancers, amplification of viral sequences together with cellular flanking sequences. *J. Virol.* **64**, 813–821.
55. Bruni, R., Argentini, C., D'Ugo, E., Giuseppetti, R., and Rapicetta, M. (1997) Woodchuck hepatitis virus DNA integration in a common chromosomal region of the woodchuck genome in two independent hepatocellular carcinomas. *Arch. Virol.* **142**, 499–509.
56. Gong, S. S., Jensen, A. D., Chang, C. J., and Rogler, C. E. (1999) Double-stranded linear duck hepatitis virus virus (DHBV) stably integrates at higher frequency than wild-type DHBV in LMH chicken hepatoma cells. *J. Virol.* **73**, 1492–1502.
57. Mueller, P. R. and Wold, B. (1989) In vivo footprinting of a muscle specific enhancer by ligated mediate PCR. *Science* **246**, 780–786.
58. Shyamala, V. and Ames, G. F.-L. (1989) Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene* **84**, 1–8.
59. Zhang, Y. and Frohman, M. A. (1997) Using rapid amplification of cDNA ends (RACE) to obtain full-length cDNAs, in *cDNA Library Protocols*, Cowell, I. G. and Austin, C. A., eds., Humana, Totowa, NJ, pp.61–87.
60. Mizobuchi, M. and Frohman, L.A. (1993) Rapid amplification of genomic DNA ends. *BioTechniques* **15**, 214–216.
61. Wattel, E., Vartanian, J. P., and Wain-Hobson, S. (1995) Clonal expansion of HTLV-I infected cells in asymptomatic and symptomatic carries without malignancy. *J. Virol.* **69**, 2863–2868.
62. Cavrois, M., Gessain, A., Wain-Hobson, S., and Wattel, E. (1996) Proliferation of HTLV-I infected circulating cells in vivo in all asymptomatic carries and patients with TSP/HAM. *Oncogene* **12**, 2419–2423.
63. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell S., et al. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucl. Acids Res.* **18**, 2887–2890.
64. Arnold, C. and Hodgson, I. J. (1991) Vectorette PCR: a novel approach to genomic walking. *PCR Meth. Appl.* **1**, 39–42.
65. Rubie, C., Schulze-Bahr, E., Wedekind, H., Borggreffe, M., Haverkamp, W., and Breithardt, G. (1999) Multistep-touchdown vectorette-PC-a rapid technique for the identification of IVS in genes. *BioTechniques* **27**, 414–418.
66. Proffitt, J., Fenton, J., Pratt, G., Yates, Z., and Morgan, G. (1999) Isolation and characterization of recombination events involving immunoglobulin heavy chain switch regions in

- multiple myeloma using long distance vectorette PCR (LDV-PCR) *Leukemia* **13**, 1100–1107.
67. Devon, R. S., Porteous, D.J., and Brookes, A. J. (1995) Splinkerettes - improved vectorettes for greater efficiency in PCR walking. *Nucl. Acids Res.* **23**, 1644–1645.
 68. Ivics, Z., Hackett, P. B., Plasterk, R. H., and Izsvak, Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510.
 69. Williams, J. G. K., Kubelik, A. R., Livak, K. J., and Rofalski, J. A. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.* **18**, 6531–6535.
 70. Parker, J. D., Rabinovitch, P. S., and Burmer, G. C. (1991) Targeted gene walking polymerase chain reaction. *Nucl. Acid Res.* **19**, 3055–3060.
 71. Hui, E. K.-W., Wang, P.-C., and Lo, S. J. (1998) Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell. Mol. Life Sci.* **54**, 1403–1411.
 72. Parks, C. L., Chang, L.-S., and Sheuk, T. (1991) A polymerase chain reaction mediated by a single primer: cloning of genomic sequences adjacent to a serotonin receptor protein coding region. *Nucl. Acids Res.* **19**, 7155–7160.
 73. Liang, P. and Pardee, A. B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.
 74. Carulli, J. P., Artinger, M., Swain, P. M., Root, C. D., Chee, L., Tulig, C., et al. (1998) High throughput analysis of differential gene expression. *J. Cell. Biochem. Suppl.* **30/31**, 286–296.
 75. Matz, M. V. and Lukyanov, S. A. (1998) Different strategies of differential display: area of application. *Nucl. Acids Res.* **26**, 5537–5543.
 76. Sarkar, G., Turner, R. T., and Bolander, M. E. (1993) Restriction-site PCR: a direct method of unknown sequence retrieval adjacent to a known locus by using universal primers. *PCR Meth. Appl.* **2**, 318–322.
 77. Liu, Y.-G., Mitsukawa, N., Oosumi, T., and Whittier, R. F. (1995) Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric inter-laced PCR. *Plant J.* **8**, 457–463.
 78. Campisi, L., Yang, Y., Heiling, E., Herman, B., Carsista, A. J., Allen, D. W., et al. (1999) Generation of enhancer trap lines in Arabidopsis and characterization of expression patterns in the inflorescence. *Plant J.* **17**, 699–707.
 79. Smith, D., Yanai, Y., Liu, Y.-G., Ishiguro, S., Okada, K., Shibata, D., et al. (1996) Characterisation and mapping of Ds-Gus-T-DNA lines for targeted insertional mutagenesis. *Plant J.* **10**, 721–732.
 80. Parinov, S., Sevugan, M., Ye, D., Yang, W.-C., Kumaran, M., and Sundaresan, V. (2000) Analysis of flanking sequences from Dissociation insertion lines: a database for reverse genetics in Arabidopsis. *Plant Cell* **11**, 2263–2270.
 81. Okamoto, H. and Hirochika, H. (2000) Efficient insertion mutagenesis of Arabidopsis by tissue culture-induced activation of the tobacco retrotransposon Tto1. *Plant J.* **23**, 291–304.
 82. Casper, C., Leib-Mosch, C., Salmons, B., Gunzburg, W. H., Baumann, G., Hofler, H., et al. (1998) Mapping of mouse mammary tumor virus integration site by retroviral LTR-arbitrary polymerase chain reaction. *Virus Res.* **54**, 207–215.
 83. Borenshtein, R. and Davidson, I. (1999) Development of the hot spot-combined PCR assay for detection of retroviral insertions into Marek's disease virus. *J. Virol. Meth.* **82**, 119–127.

84. Liu, Y.-G. and Whittier, R. F. (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* **25**, 674–681.
85. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.
86. Morishita, K., Parker, D. S., Mucenski, M. L., Jenkins, N. A., Copeland, N. G., and Ihle, J. N. (1988) Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. *Cell* **54**, 831–840.
87. Askew, D. S., Bartholomew, C., Buchbeg, A. M., Valentine, M. B., Jenkins, N. A., Copeland, N. G., and Ihle, J. N. (1991) His-1 and His-2: identification and chromosomal mapping of two commonly rearranged sites of viral integration in a myeloid leukemia. *Oncogene* **6**, 2041–2047.
88. Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989) Polymerase chain reaction with single-sided specificity: analysis of T cell receptor d chain. *Science* **243**, 217–220.
89. Ohara, O., Dorit, R. L., and Gilbert, W. (1989) One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci. USA* **86**, 5673–5677.
90. Valk, P. J. M., Joosten, M., Vankan, Y., Lowenberg, B., and Delwel, R. (1997) A rapid RT-PCR based method to isolate complementary DNA fragments flanking retrovirus integration sites. *Nucl. Acids Res.* **25**, 4419–4421.
91. Lagerstrom, M., Parik, J., Malmgren, H., Stewart, J., Pettersson, U., and Landegren, U. (1991) Capture PCR: efficient amplification of DNA fragments adjacent to a known sequence in human and YAC DNA. *PCR Meth. Appl.* **1**, 111–119.
92. Sorensen, A. B., Duch, M., Jorgensen, P., and Petersen, F. S. (1993) Amplification and sequence analysis of DNA flanking integrated proviruses by a simple two-step polymerase chain reaction method. *J. Virol.* **67**, 7118–7124.
93. Frey, M., Settner, C., and Gierl, A. (1998) A general method for gene isolation in tagging approaches: Amplification of insertion mutagenised sites (AIMS) *Plant J.* **13**, 717–721.
94. Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmaier, A., et al. (1997) Analysis of a chemical defense mechanism in grasses. *Science* **277**, 696–699.
95. Peters, G. (1990) Oncogenes at viral integration sites. *Cell Growth Diff.* **1**, 503–510.
96. Jonkers, J. and Berns, A. (1996) Retroviral insertional mutagenesis as a strategy to identify cancer genes. *Biochim. Biophys. Acta* **1287**, 29–57.
97. Huang, S.-H. (1997) Inverse PCR approach to cloning cDNA ends, in *cDNA Library Protocols* (Cowill, J. G. and Austin, C. A., eds., Humana, Totowa, NJ, pp.89–96.
98. Hodzic, D., Frey, B., Marechal, D., Scarcez, T., Grootclaes, M., and Winkler, R. (1999) Cloning of breakpoints in and downstream the IGF2 gene that are associated with overexpression of IGF2 transcripts in colorectal tumours. *Oncogene* **18**, 4710–4717.
99. Moynihan, T. P., Markham, A. F., and Robinson, P. A. (1996) Genomic analysis of human multigene families using chromosome-specific vectorette PCR. *Nucl. Acids Res.* **24**, 4094–4095.
100. Triglia, T. (2000) Inverse PCR (IPCR) for obtaining promoter sequence. *Meth. Mol. Biol.* **130**, 79–83.
101. Terauchi, R. and Kahl, G. (2000) Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of Pal and Pgi genes from yams (*Dioscorea*) *Mol. Gen. Genet.* **263**, 554–560.

102. Witthuhn, R. C., Harrington, T. C., Wingfield, B. D., Steimel, J. P., and Wingfield, M. J. (2000) Deletion of the MAT-2 mating-type gene during uni-directional mating-type switching in *Ceratocystis*. *Curr. Genet.* **38**, 48–52.
103. Akiyama, K., Watanabe, H., Tsukada, S., and Sasai, H. (2000) A novel method for constructing gene-targeting vectors. *Nucl. Acids Res.* **28**, E77.
104. Cormack, R. S. and Somssich, I. E. (1997) Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. *Gene* **194**, 273–276.
105. Jones, D. H. and Winistorfer, S. C. (1992) Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. *Nucl. Acids Res.* **20**, 595–600
106. Jones, D. H. and Winistorfer, S. C. (1993) Genome walking with 2– to 4–kb steps using panhandle PCR. *PCR Meth. Appl.* **2**, 197–203.
107. Jones, D. H. (1995) Panhandle PCR, in *PCR Primer: A Laboratory Manual*, Dieffenbach, C. W. and Dveksler, G. S., eds., Cold Spring Harbor Laboratory Press, NY, pp.411–419.
108. Weber, K. L., Bolander, M. E., and Sarkab, G. (1998) Rapid acquisition of unknown DNA sequence adjacent to a known segment by multiplex restriction site PCR. *BioTechniques* **25**, 415–419.
109. Baury, B., Masson, D., Lustenberge, P., and Denis, M. G. (1999) Gene walking by PCR amplification of short fragments from Taq DNA polymerase - modified P1 plasmid DNA and TA cloning. *BioTechniques* **27**, 1118–1122.
110. Rudi, K., Fossheim, T., and Jakobsen, K. S. (1999) Restriction cutting independent method for cloning genomic DNA segments outside the boundaries of known sequences. *BioTechniques* **27**, 1170–1177.
111. Zhang, Z. and Gurr, S. J. (2000) Walking into the unknown: a “step down” PCR-based technique leading to the direct sequence analysis of flanking genomic DNA. *Gene* **253**, 145–150.
112. Tamme, R., Camp, E., Kortschak, R. D., and Lardelli, M. (2000) Nonspecific, nested suppression PCR method for isolation of unknown flanking DNA. *BioTechniques* **28**, 895–902.

Long Distance Vectorette PCR (LDV PCR)

James A. L. Fenton, Guy Pratt, and Gareth J. Morgan

1. Introduction

Vectorette polymerase chain reaction (PCR) is a method designed to amplify DNA when the sequence of one end of the target DNA is unknown (1,2). This technique, therefore, gives a handle on unknown sequence, which flanks DNA that has already been characterized, or sequenced. The vectorette method was conceived and patented in 1988 when it was used to sequence the termini of YAC clone inserts (1), as well as to undertake genomic walking (2). Other applications have been developed, including sequencing of cosmid insert termini, mapping of promoters, and/or introns in genomic DNA from cDNA subclones, sequencing of large clones without subcloning, mapping of regions containing deletions, insertions, and translocations. Vectorette PCR has also been adapted to clone full-length cDNA and determine the 5' and 3' ends of mRNAs (3).

Vectorette PCR has been utilized for the amplification and sequencing of genomic breakpoints in translocations in chronic myeloid leukemia (CML) (4,5). These breakpoints were isolated using relatively small PCR fragments across the breakpoint of a known translocation, t(9;22) (q34;11), which produces the chimaera gene *BCR-ABL*. More recently, we have developed and applied a robust long-distance vectorette (LDV PCR) PCR strategy, which was initially developed to isolate the specific sites of DNA breakpoints in chromosomal translocations and recombination events in the hematological malignancy multiple myeloma (6). Unlike the previous studies where a known translocation was being analyzed, the aim was to effectively screen patient DNA for unknown translocation and recombination events. The most obvious requirement for such a method is that longer PCR products would have to be obtained from vectorette PCR. To this end, the protocol for LDV PCR was developed.

The vectorette unit (*see Fig. 1*) is an oligonucleotide linker of synthetic double-stranded DNA, which possesses a restriction fragment-compatible end that can be ligated to. In fact, a vectorette unit is only partially double stranded because there is a central mismatched region (*see Fig. 1*). This mismatch region was incorporated as part

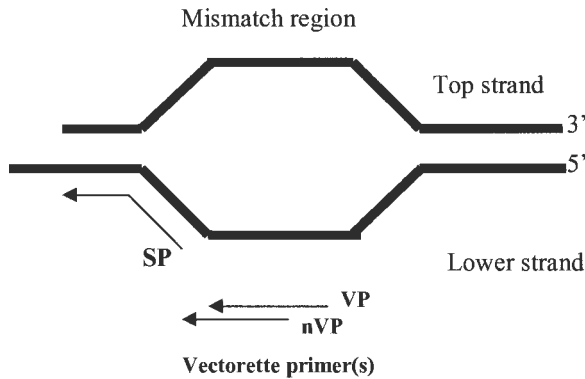


Fig. 1. Schematic representation of a vectorette unit. The vectorette unit is partially double stranded and contains a central mismatched region to avoid first strand synthesis by the vectorette primers. Two primers can be utilized for two rounds of PCR, vectorette primer (VP) and the nested vectorette primer (nVP). There are also sequencing primers (SP), which bind internally to the PCR primers, SP can be used to directly sequence any PCR products obtained.

of a strategy to avoid non-specific amplification, which can commonly occur in one-sided PCR techniques. Vectorette PCR consists of three basic stages:

1. Restriction digest of the sample DNA, this usually generates a 5' overhang (*see Fig. 2*)
2. Ligation of a compatible vectorette unit to the restriction enzyme-digested end (*see Fig. 2*)
3. PCR using primers from the known DNA sequence (called the initiating primer, or IP) and primers for the vectorette unit (vectorette primer, or VP) (*see Fig. 3*)

The importance of the central mismatch region in vectorette units is that it is this part of the design of the vectorette unit, which negates any nonspecific priming occurred. The vectorette primer has the same sequence as the bottom strand of the mismatched region and thus is unable to bind to this sequence, i.e., anneal to the vectorette DNA during the first cycle of PCR. Therefore, only the IP, complimentary to the known part of the sequence of interest, will anneal in the first cycle of PCR and therefore prime DNA synthesis (*see Fig. 3*). During the first cycle of PCR, the priming from the IP will eventually produce a sequence complimentary to the bottom strand of the vectorette unit. This now provides a template to which the VP can bind and, thus, prime, which means that consequent rounds of PCR (from cycle 2 onward) proceed as conventional PCR, so that there is only amplification of fragments containing the sequence of interest and ligated vectorette units. Our experience has shown that in order for a successful LDV PCR protocol to work, a nested PCR is also required.

The type of sequence to be investigated will determine what is observed when LDV PCR products are run out on an agarose gel. If one is simply trying to characterize unknown sequence adjacent to known DNA, then any band obtained will be potentially interesting. However, if studying recombination events (including translocations) in a specific sequence region where the germline configuration is already

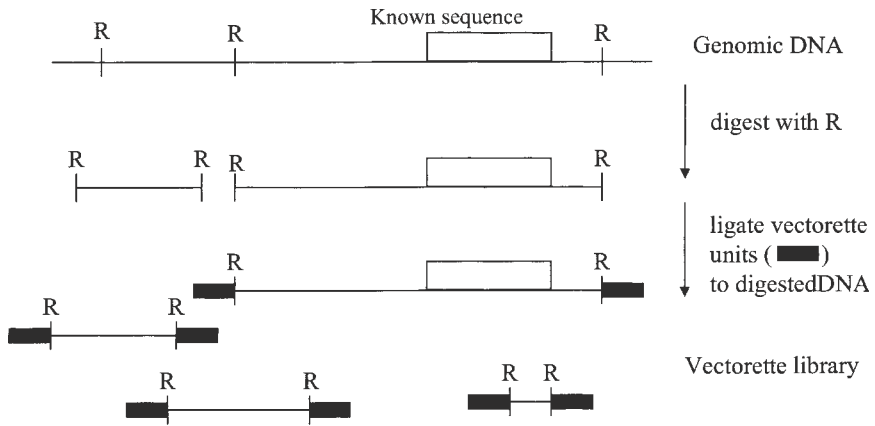


Fig. 2. Formation of vectorette library. Genomic DNA is digested with a Restriction enzyme (R) for 1 h. Vectorette units (dark rectangles) are ligated to the now compatible ends of the restriction digested DNA, giving the vectorette library, which will include fragments with the known DNA of interest.

known and germline DNA is known to present within a sample, then more caution is required. When an appropriate restriction site is within range, LDV PCR will amplify a band of predictable size from any germline DNA present, and in informative cases, a band of different size from DNA template that has been subject to recombination (*see Fig. 4*). Of course, amplifying germline bands does provide suitable positive controls for the technique.

Reviewers of inverse PCR methodology have reported that vectorette PCR can give spurious amplification of nontarget DNA (7,8). Such undesirable amplification of non-specific "end-repair priming" may involve the free cohesive ends of unligated free vectorette units and 5' overhangs of unknown cellular regions (8). In both cases, the ends are filled during the first cycle of PCR and are thus able to anneal together after the subsequent denaturation step. A complimentary strand for vectorette primer is now in position for PCR to be initiated from this nonspecific fragment. As a rule, we have not encountered major problems with such nonspecific priming, this may be a result of the hot start initiation steps, which has been incorporated into the protocol.

One of the advantages of vectorette PCR, in our hands, is that we have found it to be a simpler and more rapid assay to perform than Southern blotting, which may otherwise have been undertaken. LDV PCR allows the simultaneous detection and isolation of recombination breakpoint regions, i.e., the boundary between known sequences of DNA and unknown sequences. Additionally, the technique can be applied to quantities of DNA that are too low to permit realistic Southern blotting to take place. The amplification nature of the protocol described also allows one to isolate DNA recombination events such as a translocation in a small subpopulation of cells, against a background of nonrearranged (germline) DNA from other (normal) cells.

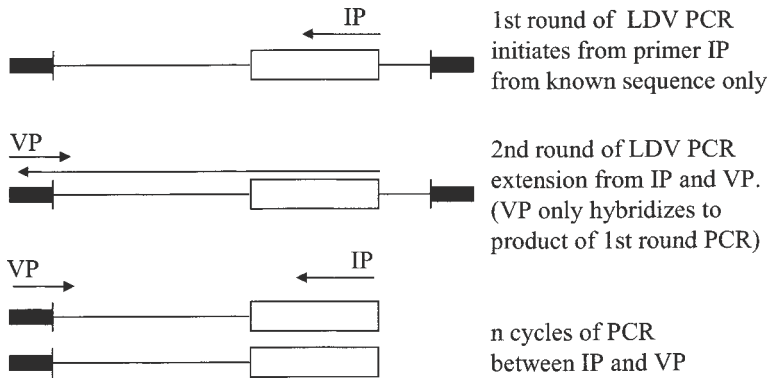


Fig. 3. LDV PCR. The IP binds to the known sequence (white rectangle) and primes DNA synthesis. If a restriction site with a ligated vectorette unit is within range, a sequence complementary to the bottom strand of that vectorette unit will be produced. This provides a template to which VP can bind and thus prime allowing conventional PCR to occur from cycle 2 onward.

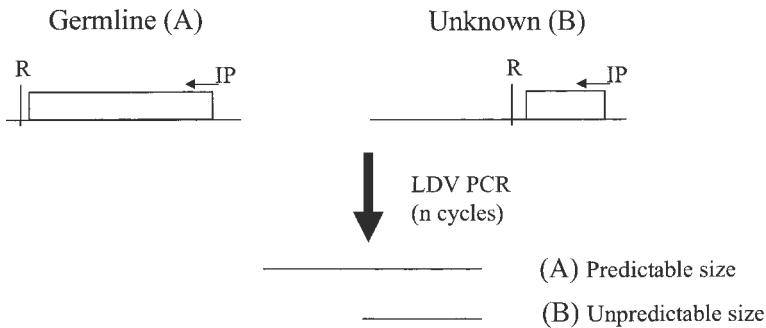


Fig. 4. Schematic of LDV PCR products on agarose gel. When investigating a recombination event, bands of predictable size (A) may be amplified from germline DNA in addition to a band of different size from DNA template, which has undergone a recombination event (B). The band (A) of predictable size acts as a useful control in this case.

2. Materials

2.1. DNA Extraction

Genomic DNA extracted from the sample of interest. The DNA to be investigated must be digestible with an appropriate restriction enzyme, yielding a general population of DNA fragments (*see Note 1*).

2.2. Construction of Vectorette Libraries

1. Restriction enzymes. A range of restriction enzymes is recommended for LDV PCR (*see Note 2*).
2. Vectorette units. These are supplied commercially. There is a starter kit available, namely Vectorette II (Sigma-Genosys).

3. 100 mM Dithiothreitol (DTT).
4. 100 mM Adenosine triphosphate (ATP).
5. T4 Ligase.

2.3. Primers

1. Vectorette and nested vectorette primers (Sigma-Genosys).
2. Initiating primers (*see Subheading 3.3.*).

2.4. LDV PCR

1. *Taq* polymerase, a suitable “long accurate” polymerase enzyme should be employed, e.g., LA *Taq* (TaKaRa) (*see Note 3*).
2. Thin-walled/ultrathin-walled PCR tubes.
3. Agarose.
4. Ethidium bromide.
5. TBE buffer.

2.5. Sequencing

1. QIAquick Gel Extraction kit (QIAGEN).
2. Vectorette sequencing primer, a 15mer and 20mer are both available (Sigma-Genosys).
3. ABI Prism Big dye terminator (Applied Biosystems).

2.6. Cloning

1. pT7Blue Vector Perfectly Blunt Cloning Kit (Novagen)

3. Methods

3.1. DNA Extraction

“Good quality” DNA is required as a template. The DNA should be accurately quantified so exactly 1 μg is used in the construction of the Vectorette libraries (*see Note 1*).

3.2. Construction of Vectorette Libraries

3.2.1. Restriction Digest

For each vectorette library, 1 μg of DNA is required, and the number of vectorette libraries to be constructed should be decided (*see Note 2*).

The appropriate amount of DNA solution is added to a 0.5 ml microfuge tube, along with the following reagents for a restriction digest:

1. 1 μg of genomic DNA.
2. 5 μL of 10X restriction buffer.
3. 20 Units of restriction enzyme.
4. Sterilized/deionized water to 50 μL final volume.

The sample is incubated in a heat block/water bath at 37°C for 1 h, and then placed on ice for at least 2 min. Once chilled, the restriction digested DNA is ready for the ligation step.

3.2.2. Ligation

Each vectorette unit is supplied in a vial as 15 pmol of annealed, lyophilized DNA. The contents of the vial should be resuspended in 25 μL of sterilized/dionized water.

1. 5 μL of the corresponding vectorette unit is ligated to the digested DNA sample with the following cofactors:
 - a. 5 μL Vectorette unit, (now in solution).
 - b. 1 μL 100 mM ATP.
 - c. 1 μL 100 mM DTT.
 - d. 1 μL (1 Unit) T4 DNA ligase.
2. The sample is incubated at 20°C for 60 min, followed by 37°C for 30 min, both steps are repeated twice on a thermocycler (i.e., 4.5 h in total). The purpose of this cycling is to increase the efficiency of the ligation step (*see Note 4*).
3. Add 200 μL of sterilized/dionized water to each tube and store the vectorette libraries in aliquots at -20°C. This provides the DNA template for the PCR.

3.3. Primers

Two lyophilized PCR primers, vectorette II primer and nested vectorette II primer, are commercially available (Sigma-Genosys). These should be resuspended in 1 mL of sterilized/dionized water to give a working concentration of 10 μM . It is recommended that aliquots of these solutions are made up and stored at -20°C.

IPs need to be designed for the known (anchor) end of the LDV PCR. This can easily be achieved by using a primer design program, such as Primer Express (Applied Biosystems) if there is enough known sequence available. Obviously, such programs will only design primer pairs, but this can be useful because it allows the opportunity to test out beforehand primers, which are going to be applied to LDV PCR. The primers used must be totally specific for the known target sequences and be based on the “standard” design parameters of length (no. of basepairs), the GC/AT ratio and melting temperature (T_m). Primers and nested primers are required to bind to the known sequence close to the boundary between the known and unknown DNA sequences. Seminested primers are just as acceptable, if there are constraints on space. Ideally, we prefer to use primers that have a high T_m value so that a 2-step PCR protocol can be used (*see Note 5*).

3.4. LDV PCR

The following protocol describes the use of LA *taq* (TaKaRa, Japan), but this is not the only suitable DNA polymerase enzyme commercially available (*see Notes 3*).

3.4.1. First-Round PCR (30 cycles)

1. For the 50- μL volume reaction, use thin-walled/ultrathin-walled PCR tubes appropriate to the thermal cycler, which will be employed. Set up each reaction by adding 8 μL dNTPs (2.5 mM each), 5 μL 10X LA PCR buffer (Mg plus), 0.5 μL Vectorette primer (10 μM or 10 pmol/ μL), 0.5 μL IP (10 μM or 10 pmol/ μL), 14 μL sterilized/deionized water.
Always ensure everything is kept on ice at all times. These reagents can be made up together in master mix form and added 28 μL (*see Note 6*) to 2 μL DNA (vectorette library), giving a total volume of 30 μL .
2. If necessary, overlay the reaction with mineral oil and briefly spin down the tubes.
3. Place tubes on thermocycler and start a program with following cycling profile: Initial denaturation at 95°C for 3 min, followed by 30 cycles consisting of denaturation at 95°C for 45 s, annealing/extension at 68°C for 4 min. The reaction is completed with a hold for final extension at 72°C for 10 min and a final hold at 4°C, until the tubes are removed.

4. Employ a manual hot start (*see Note 7*). Ensure that there is a hold/pause during the initial 3 min 95°C denaturation step. After at least a minute at this temperature, the contents of the tubes will effectively have equilibrated. Add 20 μL of sterilized/dionized water per tube, containing 0.5 μL of LA *Taq* (TaKaRa), through the mineral oil layer, giving a final total volume of 50 μL . It is easier to make up enough of this latter solution for the required number of tubes as a master mix beforehand. When all the enzyme has been added to every tube, restart the cycle program.

3.4.2. Second-Round (Nested) PCR (35 cycles)

A nested PCR is then undertaken.

1. 1 μL of each first-round reaction is diluted in 1 mL of sterilized/dionized water (1:1000 dilution) to make the DNA template (*see Note 8*).
2. A further PCR is set up as above (**Subheading 3.4.1., steps 1–4**) using the nested (or seminested) primers and 2 μL of the diluted template made above.
3. When the reaction is completed, run out at least 10 μL of each sample on a 1% agarose gel in the presence of ethidium bromide.
4. Any band(s) of interest should be remade from the original vectorette library through the full-nested LDV PCR protocol, i.e., 65 cycles as before. If the same PCR product is observed again it should be cloned and/or directly sequenced. Suitable positive controls should also be employed (*see Note 9*). If nonspecific bands keep appearing then steps should be taken to eradicate these (*see Note 10*).

3.5. Sequencing LDV PCR Product

Specific bands from LDV PCR can be directly sequenced if prepared correctly. A protocol is described to purify and to directly sequence LDV PCR products using the Big Dye Terminator Cycle sequencing ready reaction kit (Applied Biosystems). For longer products of several kilobases in length, it may be preferable to first clone and then sequence the bands of interest.

1. Run out 25 μL of the product of interest on an ethidium bromide stained 0.9% agarose gel. When the band of interest is viewed under UV light and is visibly separated from primers (and any other possible bands, such as known germline products), use a clean scalpel to cut out the smallest possible slice of agarose containing the band.
2. Purify the band using the QIAquick Gel Extraction kit, microcentrifuge protocol (Qiagen) and elute into 30 μL of elution buffer. This gives the template for the sequencing reaction (*see Note 11*).
3. For direct sequencing from the Vectorette end of a LDV PCR product internal Vectorette II sequencing primers are available (Sigma-Genosys, The Woodlands, TX). Vectorette sequencing primers are supplied lyophilized at 500-pmol concentration. These should be resuspended in sterilized/deionized water at a suitable working concentration for a cycle sequencing reaction, in this case 1.6 pmol/ μL .
4. It is also possible to use the nested (nIP) primer used in the second round of PCR to sequence from the known end of the LDV PCR product to check that the LDV PCR did initiate from the correct location.
5. Set up a Big Dye sequencing reaction, depending on the type of thermal cycler being used, the following reaction should be made up in an appropriate tube: 1 μL of sequencing primer (1.6 pmol/ μL), 5 μL of the LDV PCR product, and 4 μL of Terminator Ready Reaction Mix to give a 10- μL total volume (*see Note 12*).
6. The standard automated sequencing protocol should be followed (*see Note 12*).

3.6. Cloning LDV PCR Products

Cloning methods need not be described in great detail, but we have successfully used the pT7Blue Perfectly Blunt Cloning System (Novagen) to clone LDV PCR products. Candidate LDV PCR products can be either gel purified, e.g., QIAquick Gel Extraction kit (Qiagen) or column purified, e.g., Wizard PCR Preps DNA purification system (Promega). Obviously, once cloned into a plasmid, the insert can be sequenced from either side with primers flanking the cloning site or using internal primers that bind to the known sequence of the PCR product, which includes the vectorette sequencing primer since the vectorette unit is still ligated to the insert.

4. Notes

1. High-molecular-weight genomic DNA can be prepared from samples by proteinase K digestion, phenol-chloroform extraction, and ethanol precipitation (8). To quantify the genomic DNA to be used in LDV PCR, measure it on a spectrophotometer and check the quality via a 260/280 nm ratio reading (8). If the integrity of the DNA is doubted, then a standard long PCR with control primers can be attempted first to prove that large PCR products can be made from this template.
2. The greater the number of vectorette libraries made per sample, the greater the chance of there being a restriction site (with a Vectorette unit ligated to it) within the range of which the polymerase enzyme can reach from the anchoring (IP) primers designed. Anything between 5–8 vectorette libraries is recommended depending on the amount of DNA available. The obvious restriction enzymes to use are *HindIII*, *BamHI*, *EcoRI*, *ClaI* (for which compatible Vectorette units are available). For the blunt Vectorette unit, which is also available, we have successfully employed blunt cutter enzymes such as *PvuII*, *RsaI*, and *HincII*. It is also possible to use enzymes which recognize the same restriction sites as the other enzymes, e.g., *BglII* with a *BamHI* Vectorette unit.
3. The DNA polymerase enzyme used will be extremely important, since it is required to generate as long a PCR product as possible. The further one can amplify along a PCR fragment from the initial first round of the PCR the more chance there is of reaching an appropriate restriction site with a vectorette unit ligated to it. For the PCR, we have successfully used both *LA Taq* (Takara, Japan) and *Elongase* (Gibco-BRL) enzyme systems. Other workers have reported the successful use of a Boehringer hi-fidelity enzyme system in vectorette PCR.
4. The temperature cycling for the vectorette ligation step makes the process more efficient. Restriction enzyme binding sites are reformed when target DNA fragments ligate to each other but not when they ligate to the appropriate vectorette unit, therefore this temperature cycling increases the relative proportion of target DNA correctly linked to vectorette units. Further digestion therefore makes more compatible DNA available to ligate to vectorette units.
5. When possible, it is recommended that a two-step PCR protocol where annealing and denaturation are undertaken at the same temperature after a denaturation step. The T_m of the Vectorette PCR primers are relatively high (both are given as $>70^\circ\text{C}$), however, a separate annealing step can be employed. We have successfully utilized the more conventional thermal cycling profile of three separate temperature steps typically used in standard PCR, with primers known to have lower annealing temperatures, e.g., 60°C or 62°C . In such a case, the extension step is still undertaken at 68°C , giving a typical cycling profile of 95°C for 3 min, followed by cycles consisting of denaturation at 95°C for 45 s,

annealing at 60 or 62°C for 45 s, and extension at 68°C for 4 min. The reaction is completed with a hold for final extension at 72°C for 10 min and a final hold at 4°C until the tubes are removed. Shorter or longer extension times can be employed as desired.

6. It is easier to make up master mix (in excess by at least one tube), which contains all the common component reagents of a reaction, dNTP, buffer, magnesium solution, primers and water and add 28 μL of this master mix to each reaction tube. 2 μL of each DNA library is then added to the appropriate tube and then the mineral oil placed on top.
7. A hot start protocol is recommended for all the PCRs. This can be undertaken in one of several ways. First, for a manual hot start, as described in **Subheading 3.4.1.**, namely adding the polymerase predissolved in water to the rest of the reaction components. This method can only be followed if a mineral oil overlay is being used so that the enzyme is loaded through the oil layer when the tube and its contents are held at the initial 95°C. The 20- μL volume of this addition is used solely to minimize pipeting variations. Second, a variation of this protocol can be utilized, especially when employing thermal cyclers such as the Gene AMP PCR System 9700 (Applied Biosystems) which incorporate heated lids so that mineral oil is not required. All the contents of the PCR are mixed together, including the *Taq* polymerase, with the reaction tube being kept on ice. The thermal cycler is then programmed to have an initial hold step at 4°C (for say 30 s) before the initial, 3 min 95°C denaturation step. The tubes are then transferred from the ice to the cycler when it is held at 4°C. The profile then proceeds to heat up to 95°C for the first denaturation step, thus also giving an effective hot start protocol.
8. We have found that a 1:1000 dilution was the optimal dilution to produce a substrate for the second round (nested) PCR, but this can be varied, e.g., a 1:10,000 or 1:100 dilution.
9. When an appropriate known restriction enzyme recognition site is within range LDV PCR will amplify a band of predictable size from germline DNA, which can give a good positive control band and prove that the initiating primers designed are good and working. When looking at recombination events it may be advisable to perform the LDV PCR procedure with libraries made from germline DNA (e.g., placenta), as well as the sample of interest. Any germline bands will be observed in both lanes when the products are run out side by side on a gel. However, a band of different size will result from DNA that has undergone a recombination event. In the case of a translocation, this will provide the actual translocation breakpoint that will obviously not be observed with a germline template.
10. As was discussed in **Subheading 1.** spurious amplification can occur. If this is observed then the hot start protocol used should be evaluated and if necessary revised. Spurious bands can very infrequently be observed in vectorette libraries made with a blunt cutting restriction enzyme after incorrect ligation has apparently occurred (JF personal experience!), although this has never presented itself as a major problem.
11. The concentration of the PCR product is critical to the sequence reaction. An easy way to determine that there is enough LDV PCR product for sequencing is to do a quick “eyeball” method check. Simply run a very small volume of the purified product, say 4 μL , out on a 1 % agarose minigel and check that the correct band is clearly visible to the naked eye when illuminated with UV light in the presence of ethidium bromide.
12. The reaction described is half the standard volume recommended by Applied Biosystems. The relatively high concentration of primer allows a greater volume to be added for the template. Good, clean sequence can be obtained from a direct sequencing protocol. However, if a bigger band (several kb in length) is required to be sequenced it may be necessary to clone this product into a suitable vector (*see Subheading 3.5.2.*) and then sequence it. This is because up to 100 times more of a 2000 bp product (a typical LDV

PCR product size), is required in a sequencing reaction than say a smaller 200 bp product. Therefore, it may be difficult to get a high enough concentration of this band in the sequencing reaction.

Acknowledgment

Vectorette is a trade mark of Sigma-Genosys. The development of the LDV PCR technique in our lab was supported by a grant from Yorkshire Cancer Research.

References

1. Riley, J., Butler, R., Ogilvie, D., Finnear, R., Jenner, D., Powell, S., et al. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucl. Acid Res.* **18**, 2887–2890.
2. Arnold, C. and Hodgson, I. J. (1991) Vectorette PCR: A novel approach to genomic walking. *PCR Meth. Appl.* **1**, 39–42.
3. Chenchik, A., Diachenko, L., Moqadam, F., Tarabykin, V., Lukyanov, S., and Siebert, P. D. (1996) Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *Biotechniques* **21**, 526–534.
4. Mills, K. I., Sproul, A. M., Ogilvie, D., Elvin, P., Leibowitz, D., and Burnett, A. K. (1992) Amplification and sequencing of genomic breakpoints located within the M-bcr region by Vectorette-mediated polymerase chain reaction. *Leukemia* **6**, 481–483.
5. Zhang, J. G., Goldman, J. M., and Cross, N. C. P. (1995) Characterization of genomic BCR-ABL breakpoints in chronic myeloid leukemia by PCR. *Br. J. Haematol.* **90**, 138–146.
6. Proffitt, J., Fenton, J., Pratt, G., Yates, Z., and Morgan, G. (1999) Isolation and characterisation of recombination events involving immunoglobulin heavy chain switch regions in multiple myeloma using long distance vectorette PCR (LDV-PCR). *Leukemia* **13**, 1100–1107.
7. Hengen, P. N. (1995) Methods and reagents—Vectorette, splinkerette and boomerang DNA amplification. *TIBS* **20**, 372–373.
8. Hui, E. K., Wang, P., and Lo, S. J. (1998) Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell. Mol. Life Sci.* **54**, 1403–1411.
9. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 9.16–9.19.

Nonspecific, Nested Suppression PCR Method for Isolation of Unknown Flanking DNA (“Cold-Start Method”)

Michael Lardelli

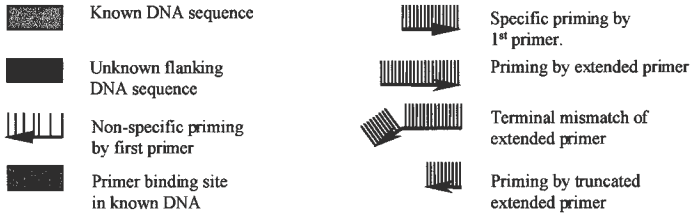
1. Introduction

The ability of the polymerase chain reaction (PCR) to amplify DNA depends upon the existence of defined primer binding sites. Thus, to amplify a region of flanking unknown DNA sequence, a defined primer binding site must be created. Numerous strategies have been found to do this, such as addition of nucleotides or ligation of oligonucleotides to DNA ends, restriction of the flanking DNA followed by ligation of known DNA sequences and, (for cDNA amplification), ligation of RNA oligonucleotides to the RNA molecule followed by reverse transcription PCR (reviewed in *ref. 1*). All of these methods require considerable molecular biological processing of the source nucleic acid.

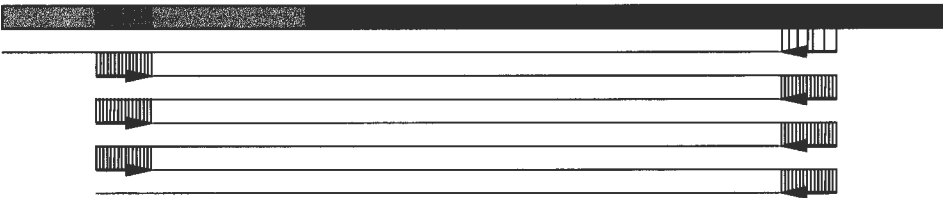
A far simpler strategy is to use nonspecific binding by oligonucleotides to generate a primer binding site in unknown flanking DNA (*see Fig. 1*). Numerous variations on this strategy have been described (2–5). A general problem with these strategies is detection and purification of the desired product from among numerous spurious amplification products. However, a method has now been developed that is very simple to perform, can generate long PCR fragments, requires no processing of the source DNA other than PCR and produces few spurious products. Because of its simplicity and sensitivity it can be used as the method of first choice before resorting to more complex procedures.

The “cold-start method” consists of two sequential PCRs. In the first PCR (“nonspecifically primed PCR,” NSPPCR), a single primer binding to known sequence and priming toward unknown flanking DNA is used under conditions of low specificity, e.g., low annealing temperature. The intention is to generate DNA strands primed from within the unknown region and extending past the primer binding site in the known DNA region. In the following PCR cycle, the primer can now bind to the spe-

Symbol Key



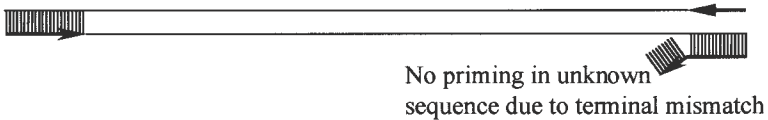
Reaction 1 - NSPPCR



Reaction 2 – Reamplification PCR



Before exonuclease truncation of extended primer:



After partial exonuclease truncation of extended primer:

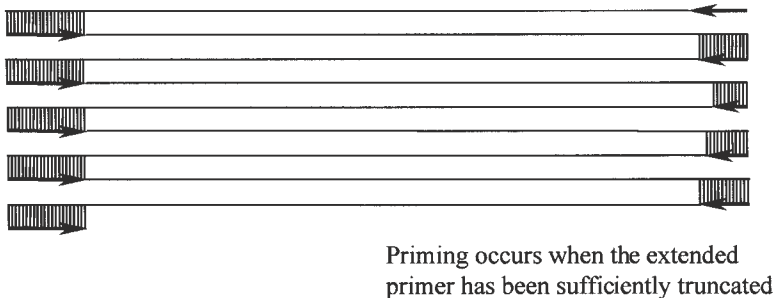


Fig. 1. The “cold-start method” for cloning flanking unknown sequences. The method consists of two sequential PCRs separated by a dilution step.

cific site in the known sequence and, by generating the complement of the first DNA strand, will produce a perfect primer binding site at the other end of the DNA molecule (*see Fig. 1*). This DNA fragment will now amplify exponentially during subsequent PCR cycles. Numerous spurious DNA fragments can be generated during this first PCR, but the desired flanking DNA fragments begin with the advantage that they already possess one perfect primer annealing site and so they are amplified at relatively high frequency.

Some methods that rely on non-specific priming use a primer binding specifically in the known sequence region and a second primer binding nonspecifically in the unknown flanking region. The advantage of using a single primer for both purposes is that inverted repeats are generated at both termini of the PCR products and so the reaction becomes a suppression PCR (6). Normally, amplification by conventional PCR favors short fragments over longer ones. In a nonspecific PCR, this leads to a predominance of shorter fragments among the products. However, suppression PCR counters this effect because, in shorter fragments, the terminal inverted repeats anneal at high frequency, thus blocking primer annealing and DNA synthesis. Low primer concentrations can enhance this suppression effect.

The second PCR of the Cold-start method (“reamplification PCR”) specifically amplifies the desired product(s) from the NSPPCR. To do this, the products of the NSPPCR are diluted and then a novel form of seminested PCR is performed. Normally, seminested PCR could not specifically amplify the desired products because the only primer known to bind in the unknown flanking region is the initial primer. The initial primer binds at both ends of all the products of the NSPPCR. Any subsequent PCR using this primer amplifies both desired and spurious products. However, we discovered that by extending the initial primer by six nucleotides and using a proofreading thermostable DNA polymerase, amplification of the desired products is greatly favored. Presumably, during the reamplification PCR, the proofreading polymerase amplifies the desired fragment in a linear manner until the primer is truncated sufficiently by the polymerase to bind at both ends of the desired fragment after which exponential amplification occurs (*see Fig. 1*). Meanwhile, partial binding of the extended primer at the ends of spurious products inhibits their reamplification. Reamplification PCR is conducted under stringent conditions and is also a form of suppression PCR so that it favors the amplification of longer fragments.

In its simplest form, the cold-start method consists of two PCRs separated by a dilution step. The number of final products is small and the desired products can be identified by Southern blotting against the known DNA sequence or simply by cloning and sequencing. There are a number of optional enhancements of the method that can further increase its success rate. The basic method is described here and references are given for the enhancements.

2. Materials

1. Oligonucleotide primers for NSPPCR and reamplification PCR (*see Subheading 3.1. Primer Design* below).
2. A DNA solution including the known DNA region with desired flanking sequences.
3. Double-distilled water.

4. Thermostable DNA polymerase *lacking* 3'-5' exonuclease activity and recommended 10X concentrated reaction buffer (e.g., *Taq* DNA polymerase and associated buffer from Stratagene Cloning Systems, La Jolla, CA).
5. Thermostable DNA polymerase *possessing* 3'-5' exonuclease activity and recommended 10X concentrated reaction buffer (e.g., *Pfu Turbo*[®] DNA polymerase from Stratagene Cloning Systems).
6. 10 mM dNTPs (2.5 mM each of dATP, dCTP, dGTP, and dTTP).
7. Paraffin oil (if performing reactions under oil, *see Note 1*).
8. PCR thermal cycler and accessories (e.g., a thermal cycler that accepts 96-well microassay plates with paraffin oil such as the PTC-200 Peltier Thermal Cycler with Multiplate[™] 96 Polypropylene V-Bottom Microplates from MJ Research Inc., Watertown, MA).
9. Micropipeters and pipet tips (e.g., Gilson, France) for liquid.
10. 1.5-mL capped, polypropylene microfuge tubes.
11. 6X gel loading buffer: 15% w/v Ficoll (Type 400; Amersham Pharmacia Biotech AB, Uppsala, Sweden), 0.35% w/v Orange G (Sigma, St. Louis, MO), 60 mM ethylene diaminetetraacetic acid (EDTA), pH 8.0.
12. Equipment and reagents for agarose gel electrophoresis.
13. DNA electrophoresis size markers (e.g., 1 Kb DNA Ladder; Life Technologies, Rockville, MD).
14. Reagents for Southern blotting, radioactive probe synthesis and hybridization, and posthybridization washing.
15. A scalpel for excision of DNA-containing bands from agarose gels.
16. Reagents for purification of DNA from agarose gel (e.g., QIAquick PCR Purification Kit, QIAGEN GmbH, Hilden, Germany).
17. Reagents for cloning of blunt-ended DNA fragments (e.g., the Zero Blunt[™] TOPO[®] PCR Cloning Kit from Invitrogen Corporation, Carlsbad, CA).

3. Methods

3.1. Primer Design

Two primers are required for this procedure, “first primer” and “extended primer”. The first primer should have a melting temperature (T_m) of around 60°C and be around 20 nucleotides in length. Ideally, it should be designed to bind 100–200 bp from the unknown flanking DNA and it must prime towards this DNA. The extended primer is identical to the first primer but is extended at its 3' end by six nucleotides corresponding to the known sequence (*see Fig. 1* and **Note 2**).

3.2. Nonspecifically Primed PCR

Details are given below for one reaction. Scale up the premixes described if multiple reactions are to be performed, e.g. to test different annealing temperatures, primer concentrations or Mg^{2+} concentrations for NSPPCR (*see Note 3*). As a guide to when NSPPCR is occurring, lower the annealing temperature or raise the Mg^{2+} concentration until bands of DNA can be observed under UV light on an agarose gel stained with ethidium bromide.

1. Assemble premixes as follows: a) PCR premix: 8 μ L of water, 2 μ L of the DNA source (e.g., 20 ng of genomic DNA, *see Note 4*), 2 μ L of 10X *Taq* DNA polymerase buffer (e.g., 500 mM KCl, 100 mM Tris-HCl pH 8.3), 5 μ L of 2.5–20 mM $MgCl_2$ (*see Note 3*), 1 μ L of first primer @ 2.5–40 μ M (*see Note 5*), 2 μ L dNTPs (10 mM). b) Polymerase

premix: 0.5 μL of 10X *Taq* DNA polymerase buffer, 4.1 μL of water, 0.4 μL of *Taq* DNA polymerase (5 U/ μL) (see **Note 6**).

2. Place the premix in a well of a PCR microassay plate or in a PCR tube. Cover with one drop of oil (see **Note 7**).
3. Place the plate or tube in the PCR thermal cycler and heat to 94°C (e.g., “pause” the cycler during the first denaturation step of the PCR cycling protocol given in **step 4** below). Wait 30 s and then eject the polymerase premix into the plate well / tube from just above the oil (e.g., onto the wall of the plate well or tube). The polymerase mix will fall through the oil and mix with the other PCR components (see **Note 1**).
4. Perform PCR cycling as follows: 35 cycles: Denaturation at 94°C for 30 s, annealing at 5 to 30°C below the oligonucleotide’s calculated annealing temperature for 1 min (see **Note 8**), temperature ramp of 0.5°C/s to 72°C then 72°C for 3 min.

3.3. Reamplification PCR

From the NSPPCR, it is possible to proceed directly to reamplification PCR or, if multiple annealing temperatures and/or Mg^{2+} concentrations are being tested, one can examine these first for the presence of desired sequences by Southern hybridization (see **Note 3**). Even if one cannot detect desired sequences by this method (because there is too little known sequence in the desired products against which to probe or because the concentration of desired products is too low) it is worthwhile to proceed with a reamplification PCR because this may, nevertheless, reveal desired sequences.

1. Dilute a sample of the NSPPCR products 1 : 1000 in water.
2. Assemble premixes as follows: a) PCR premix: 14 μL of water, 1 μL of the 1:1000 diluted NSPPCR products, 2 μL of 10X *Pfu Turbo*[®] DNA polymerase buffer (200 mM Tris-HCl pH 8.8, 20 mM MgSO_4 , 100 mM KCl, 100 mM $[\text{NH}_4]_2\text{SO}_4$, 1% Triton X-100, 0.1% nuclease-free BSA), 1 μL of extended primer @ 10 μM (see **Note 9**), 2 μL dNTPs (10 mM). b) Polymerase premix: 4.3 μL of water, 0.5 μL of 10X *Pfu Turbo*[®] DNA polymerase buffer, 0.2 μL of *Pfu Turbo* DNA polymerase (2.5 U/ μL) (see **Note 10**).
3. Place the premix in a well of a PCR microassay plate or in a PCR tube. Cover with one drop of oil (see **Note 7**).
4. Place the plate or tube in the PCR cycler and heat to 95°C. (e.g., “pause” the PCR cycler during the first denaturation step of the PCR cycling protocol given in **step 5** below). Wait 30 s, and then eject the polymerase premix into the plate well / tube from just above the oil (e.g., onto the wall of the plate well or tube). The polymerase mix will fall through the oil and mix with the other PCR components (see **Note 1**).
5. Perform PCR cycling as follows: 35 cycles: Denaturation at 95°C for 30 s, annealing at 5 to 10°C below the oligonucleotide’s calculated annealing temperature for 1 min, temperature ramp of 0.5°C/s to 72°C then 72°C for 4 min.

3.4. Identification of Desired Products from the Reamplification PCR

1. Remove a 10- μL sample from the reamplification PCR, add 2.5 μL of 5X loading buffer and conduct electrophoresis on a 1.5 % agarose gel beside size markers (e.g., 1 kb DNA ladder). If one or more bands are seen there are two alternative ways to proceed:
2. (optional) The gel can be Southern blotted. Any bands containing the desired sequences can be detected by probing the blot using DNA corresponding to the area of known sequence between the primer binding site and the unknown flanking region (see **ref. 7** for methods). When a band containing the desired sequences is identified, the remaining reaction products can be processed as in **step 3** below (see **Note 11**).

3. The bands on the gel are excised with a scalpel and the DNA they contain is purified (e.g., using the QIAquick PCR Purification Kit from Qiagen GmbH), cloned using a system allowing cloning of blunt-ended PCR fragments (e.g., the Zero Blunt™ TOPO® PCR Cloning Kit from Invitrogen Corporation) and then sequenced. Desired fragments can be identified since they contain the area of known sequence between the primer binding site and the unknown flanking region.

4. Notes

1. The method for “hot-starting” PCR that is described in this protocol is the optimal one for starting a PCR that is being conducted under paraffin oil. If you wish to hot-start the reaction using thermostable DNA polymerase to which a blocking antibody is initially bound, then the entire reaction can be made up as one solution (rather than being divided into reaction and enzyme premixes).
2. The extended primer may be shortened by a number of bases at the 5' end to reduce the melting temperature. However, note that the primer is still required to bind at the chosen annealing temperature even when the additional six nucleotides at its 3' end are removed by the exonuclease activity of the polymerase.
3. Multiple NSPPCRs can be performed using different annealing temperatures and Mg²⁺ concentrations to find those conditions that give sufficient nonspecific binding by the first primer in the unknown flanking DNA to amplify the desired sequence. The NSPPCRs are then electrophoresed on an agarose gel, Southern blotted and hybridized with a probe containing the known sequence between the known first primer binding site and the flanking, unknown DNA (see ref. 7 for methods). When desired DNA fragments are identified in a particular NSPPCR, this reaction can then be used for reamplification PCR. This procedure has the advantage that different conditions produce NSPPCR products of different sizes and the NSPPCR producing the largest fragment of flanking DNA can be selected (see also Note 5). However, note that reamplification PCR can amplify NSPPCR products that are present at levels below the level of detection of Southern hybridization.
4. The efficiency of amplification of flanking cDNA sequences can be boosted by initial purification of the desired cDNA on magnetic beads before NSPPCR. See ref. 8 for details.
5. Varying the concentration of the primer can cause variation in the length of the desired NSPPCR products. Lower primer concentrations will tend to select for longer desired fragment lengths.
6. This protocol should give final reagent concentrations of: 20 ng of genomic DNA (or other DNA source), 50 mM KCl, 10 mM Tris-HCl pH 8.3, 0.5–4 mM MgCl₂, 0.1–1.6 μM first primer, 0.4 mM dNTPs, and 2U of *Taq* DNA polymerase.
7. It may not be necessary to use oil and/or “pausing of the PCR cycling” if the containment of the PCR and/or the method of “hot starting” obviate this. See also Note 1.
8. If the annealing temperature initially used for NSPPCR does not result in amplification of desired products after the reamplification PCR, then lower the annealing temperature further. See also Note 3.
9. Despite that the reamplification PCR is a form of suppression PCR, lowering the extended primer concentration has not been shown to have a great effect on the size of the reamplified products. Lower primer concentrations apparently simply produce less DNA.
10. This protocol should give final reagent concentrations (not including template DNA) of: 20 mM Tris-HCl pH 8.8, 2 mM MgSO₄, 10 mM KCl, 10 mM [NH₄]₂SO₄, 0.1% Triton X-100, 0.01% nuclease-free BSA, 0.4 μM extended primer, 0.4 mM dNTPs, 0.5 U *Pfu Turbo* DNA polymerase.

11. If the probe sequence includes the primer binding site, then a low level of hybridization of the probe with all PCR products is possible.

References

1. Hui, E. K., Wang, P. C., and Lo, S. J. (1998) Strategies for cloning unknown flanking DNA from sequences flanking foreign integrants. *Cell. Mol. Life Sci.* **54**, 1403–1411.
2. Malo, M. S., Srivastava, K., Andresen, J. M., Chen, X.-N., Korenberg, J. R., and Ingram, V. M. (1994) Targeted gene walking by low stringency polymerase chain reaction: Assignment of a putative human brain sodium channel gene (*SCN3A*) to chromosome 2q24-31. *Proc. Natl. Acad. Sci. USA* **91**, 2975–2979.
3. Parker, J. D., Rabinovitch, P. S., and Burner, G. C. (1991) Targeted gene walking polymerase chain reaction. *Nucleic Acids Res.* **19**, 3055–3060.
4. Parks, C. L., Chang, L.-S., and Shenk, T. (1991) A polymerase chain reaction mediated by a single primer: cloning of genomic sequences adjacent to a serotonin receptor protein coding region. *Nucl. Acids Res.* **19**, 7155–7160.
5. Trueba, G. A. and Johnson, R. C. (1996) Random primed gene walking PCR: a simple procedure to retrieve nucleotide fragments adjacent to known DNA sequences. *BioTechniques* **21**, 20.
6. Lukyanov, K. A., Launer, G. A., Tarabykin, V. S., Zaraisky, A. G., and Lukyanov, S. A. (1995) Inverted terminal repeats permit the average length of amplified DNA fragments to be regulated during preparation of cDNA libraries by polymerase chain reaction. *Analyt. Biochem.* **229**, 198–202.
7. Sambrook, J. and Russel, D. *Molecular Cloning: A Laboratory Manual*, 3rd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
8. Tamme, R., Camp, E., Kortschak, R. D., and Lardelli, M. (2000) Non-specific, nested, suppression PCR method for isolation of unknown flanking DNA. *BioTechniques* **28**, 895–902.

Inverse PCR

cDNA Cloning

Sheng-He Huang

1. Introduction

Since the first report on cDNA cloning in 1972 (1), this technology has been developed into a powerful and universal tool in isolation, characterization, and analysis of both eukaryotic and prokaryotic genes. But the conventional methods of cDNA cloning require much effort to generate a library that is packaged in phage or plasmid and then survey a large number of recombinant phages or plasmids. There are three major limitations in those methods. First, substantial amount (at least 1 μg) of purified mRNA is needed as starting material to generate libraries of sufficient diversity (2). Second, the intrinsic difficulty of multiple sequential enzymatic reactions required for cDNA cloning often leads to low yields and truncated clones (3). Finally, screening of a library with hybridization technique is time-consuming.

Polymerase chain reaction (PCR) technology can simplify and improve cDNA cloning. Using PCR with two gene-specific primers, a piece of known sequence cDNA can be specifically and efficiently amplified and isolated from very small numbers ($<10^4$) of cells (4). However, it is often difficult to isolate full-length cDNA copies of mRNA on the basis of very limited sequence information. The unknown sequence flanking a small stretch of the known sequence of DNA cannot be amplified by the conventional PCR. Recently, anchored-PCR (5–7) and inverse PCR (8–10) have been developed to resolve this problem. Anchored-PCR techniques have the common point: DNA cloning goes from a small stretch of known DNA sequence to the flanking unknown sequence region with the aid of a gene-specific primer at one end and a universal primer at other end. Because of only one gene specific primer in the anchored-PCR it is easier to get a high level of nonspecific amplification by PCR than with two gene-specific primers (10,11). The major advantage of inverse PCR (IPCR) is to amplify the flanking unknown sequence by using two gene-specific primers.

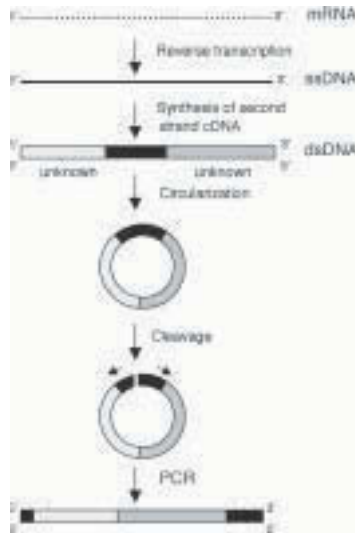


Fig. 1. Diagram of IPCR for cDNA cloning. The procedure consists of five steps: reverse transcription, synthesis of second strand cDNA, circularization of double-strand cDNA, reopen the circle DNA, and amplification of reverse DNA fragment. The black and open bars represent the known and unknown sequence regions of double-stranded cDNA, respectively.

At first, IPCR was successfully used in the amplification of genomic DNA segments that lie outside the boundaries of known sequence (8,9). I have a new procedure which extends this technique to the cloning of unknown cDNA sequence from total RNA (10). Double-stranded cDNA is synthesized from RNA and ligated end to end (see Fig. 1). Circularized cDNA is nicked by selected restriction enzyme or denatured by NaOH treatment (12,13). The reopened or denatured circular cDNA is then amplified by two gene-specific primers. Recently, this technique has been efficiently used in cloning full-length cDNAs (14–16). The following protocol was used to amplify cDNA ends for the human stress-related protein ERp72 (10) (see Fig. 2).

2. Materials

2.1. First-Strand cDNA Synthesis

1. Total RNA prepared from human CCRF/CEM leukemic lymphoblast cells (17,18).
2. dNTP mix: 10 mM of each dNTP.
3. Random primers (Boehringer Mannheim, Indianapolis, IN). Prepare in sterile water at 1 $\mu\text{g}/\mu\text{L}$. Store at -20°C .
4. RNasin (Promega, Madison, WI).
5. Actinomycin D (1 mg/mL). Actinomycin D is light sensitive and toxic. It should be stored in a foil-wrapped tube at -20°C .
6. MMLV reverse transcriptase.
7. 5X First-strand buffer: 0.25 M Tris-HCl (pH 8.3), 0.375 M KCl, 50 mM MgCl_2 , 50 mM dithiothreitol (DTT), and 2.5 mM spermidine. The solution is stable at -20°C for more than 6 mo.

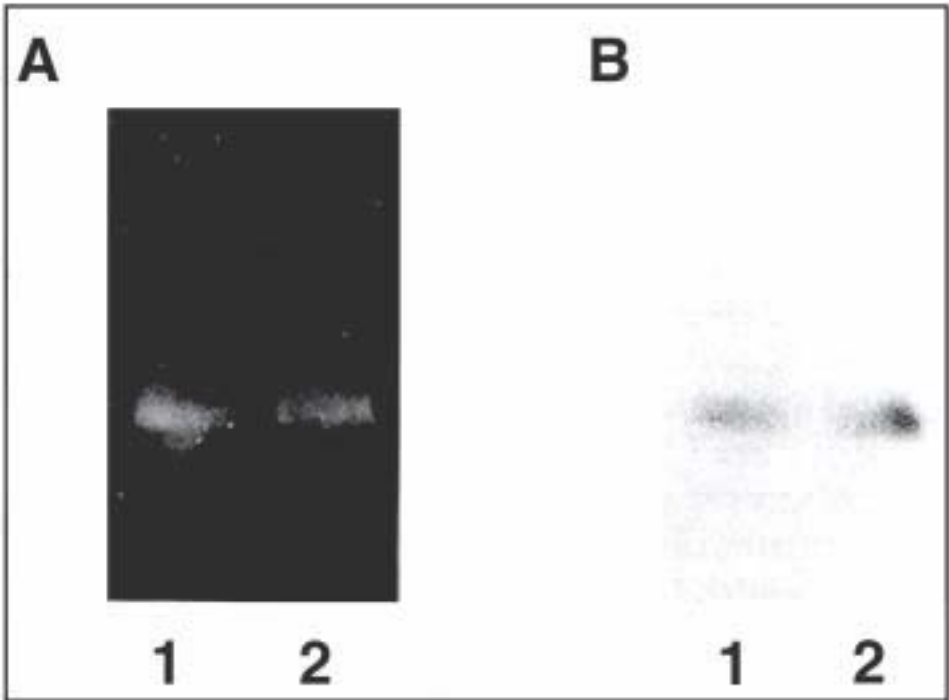


Fig. 2. Application of IPCR to amplifying the joining region (280 bp) from 5' (160 bp) and 3' (120 bp) sequences of human ERp72 cDNA. Amplified DNAs from CCRF/CEM cells sensitive (lane 1) and resistant (lane 2) to cytosine arabinoside stained by ethidium bromide (A) or hybridized with ^{32}P -labeled ERp72 cDNA (B). See text for the sequences of the primers and the parameters of IPCR.

2.2. Second-Strand Synthesis

- 10X second strand buffer: 400 mM Tris-HCl (pH 7.6), 750 mM KCl, 30 mM MgCl_2 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, 30 mM DTT, and 0.5 mg/mL of bovine serum albumen (BSA). The solution is stable at -20°C for at least 6 mo.
- 1 mM nicotinamide adenine dinucleotide (NAD).
- RNase H (2 U/ μL).
- Escherichia coli* (*E. coli*) DNA polymerase I (5 U/ μL).
- E. coli* DNA ligase (1 U/ μL).
- Nuclease-free H_2O .
- T4 DNA polymerase.
- 200 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0.
- GeneClean (Bio 101 Inc., La Jolla, CA).
- TE buffer: 10 mM Tris-HCl, pH 7.6, 1 mM EDTA.
- DNA standards. Prepare 1-mL aliquots of a purified DNA sample at 1, 2.5, 5, 10, and 20 $\mu\text{g}/\text{mL}$ in TE buffer. Store at -20°C for up to 6 mo.
- TE/ethidium bromide: 2 $\mu\text{g}/\text{mL}$ of ethidium bromide in TE buffer. Store at 4°C for up to 6 mo in a dark container.

2.3. Circularization and Cleavage or Denaturation

1. 5X ligation buffer (supplied with T4 DNA ligase).
2. T4 DNA ligase (1 U/ μ L).
3. T4 RNA ligase (4 μ g/ μ L).
4. 15 μ M Hexaminecobalt chloride.
5. Phenol:CHCl₃:isoamyl alcohol (25:24:1).
6. 3 M sodium acetate, pH 7.0.
7. Absolute ethanol.
8. 70% ethanol.

2.4. Inverse PCR

1. 10X PCR buffer: 100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM MgCl₂, 0.01%(w/v) gelatin.
2. 15 mM MgCl₂.
3. Deoxyoligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) 380B DNA synthesizer and purified by OPEC column from the same company. The primer pairs were selected from the 5' and 3' sequence of the cDNA coding for human ERp72 stress-related protein (5'-primer: 5'-TTCCTCCTCCTCCTCCTCTT-3'; 3'-primer: 5'-ATCTAAATGTCTAGT-3') (10).
4. Light mineral oil.
5. *Taq* DNA polymerase.

3. Methods

3.1. First-Strand cDNA Synthesis (19)

1. Perform reverse transcription in a 25- μ L reaction mixture, adding the following components:

5X first-strand buffer	5.0 μ L
dNTP mix	2.5 μ L
random primers	2.5 μ L
Rnasin	1.0 U
actinomycin D	1.25 μ L
MMLV reverse transcriptase	250 U
RNA	15–25 μ g of total RNA (Heat denature RNA at 65°C for 3 min prior to adding to reaction)
Nuclease-free H ₂ O	to 25 μ L final vol.

3.2. Second-Strand Synthesis (20)

1. Add components to the first-strand tube on ice in the following order:

10X second-strand buffer	12.5 μ L
1 mM NAD	12.5 μ L
RNase H (2 μ / μ L)	0.5 μ L
<i>E. coli</i> DNA polymerase I (5 (μ L)	5.75 μ L
<i>E. coli</i> ligase (1 μ / μ L)	1.25 μ L
nuclease-free water	92.5 μ L
2. Incubation at 14°C for 2 h.
3. Heat the reaction mix to 70°C for 10 min, spin for few seconds, and then put in ice.
4. Add 4 U of T4 DNA polymerase and incubate at 37°C for 10 min to blunt the ends of double-stranded cDNA.

5. Stop the reaction with 12.5 μL of 0.2 M EDTA and 200 μL sterile H_2O .
6. Concentrate and purify the sample with GeneClean. Resuspend the DNA in 100–200 μL of sterile H_2O .
7. Estimate the DNA concentration by comparing the ethidium bromide fluorescent intensity of the sample with that of a series of DNA standards on a sheet of plastic wrap (**21**). Dot 1–5 μL of sample onto plastic wrap on a UV transilluminator. Also dot with 5 μL of DNA standards. Add an equal volume of TE buffer containing 2 $\mu\text{g}/\text{mL}$ of ethidium bromide, mix by repipeting up and down. Use proper UV shielding for exposed skin and eyes.

3.3. Circularization and Cleavage (see Notes 1–4)

1. Set up the circularization reaction mixture (150 μL) containing the following components: 100 μL (100 ng DNA) of the purified sample, 30 μL of 5X ligation buffer, and 6 μL of T4 DNA ligase. Finally, add 2 μL of T4 RNA ligase or 15 μL of 15 μM hexaminecobalt chloride (see **Note 5**).
2. Incubate at 18°C for 16 h.
3. Boil the ligated circular DNA for 2–3 min in distilled water or digest with an appropriate restriction enzyme to re-open circularized DNA.
4. Purify the DNA sample with GeneClean as described in step 6 in **Subheading 3.2**, or extract with water-saturated phenol/ CHCl_3 and then precipitate with ethanol (**20**).

3.4. Inverse PCR (see Note 6)

1. Add 1/10 of the purified cDNA to 100 μL of amplification mixture (**22**):

10X PCR buffer	10 μL
15 mM MgCl_2	10 μL
dNTP mix (2.5 mM of each)	10 μL
5'-primer (10 pmole/ μL)	10 μL
3'-primer (10 pmole/ μL)	10 μL
cDNA	10 μL
Nuclease-free H_2O	39.5 μL
Taq DNA polymerase (2.5 $\mu\text{g}/\mu\text{L}$)	0.5 μL
2. Cap and vortex the tubes to mix. Spin briefly in a microfuge. Cover each reaction with a few drops of light mineral oil to prevent evaporation.
3. Put a drop of mineral oil into each well of the thermal cycler block that will hold a tube. Load the reaction tubes.
4. Amplify by PCR using the following cycle profile:

25 cycles	94°C 1 min (denaturation)
	65°C 2 min (annealing)
	72°C 4 min (elongation)

4. Notes

1. For maximum efficiency of intra-molecular ligation, low concentration of cDNA should be used in the ligation mix. High density of cDNA may enhance the level of heterogeneous ligation, which creates nonspecific amplification.
2. Cleavage or denaturation of circularized double-strand cDNA is important because circular double-strand DNA tends to form supercoil and is poor template for PCR (**23**). Circularized double-strand DNA is only good for amplification of a short DNA fragment.
3. The following three ways can be considered to introduce nicks in circularized DNA. Boiling is a simple and common way. Owing to the unusual secondary structure of some

circular double-strand DNA, sometimes this method is not sufficient in nicking and denaturing circular double-strand DNA. A second method is selected restriction enzyme digestion. The ideal restriction site is located in the known sequence region of cDNA. In most cases, it is difficult to make the right choice of a restriction enzyme because the restriction pattern in unidentified region of cDNA is unknown. If an appropriate enzyme is not available, EDTA-oligonucleotide-directed specific cleavage may be tried (24,25). Oligonucleotide linked to EDTA-Fe at T can bind specifically to double-stranded DNA by triple-helix formation and produce double-stranded cleavage at the binding site.

4. Alkali denaturation has been successfully used to prepare plasmid DNA templates for PCR and DNA sequencing (12,13,26). This method should be feasible in denaturing circularized double-strand cDNA.
5. Inclusion of T4 RNA ligase or hexaminecobalt chloride can enhance the efficiency of blunt-end ligation of double-strand DNA catalyzed by T4 DNA ligase (27).
6. IPCR can be used to efficiently and rapidly amplify regions of unknown sequence flanking any identified segment of cDNA or genomic DNA. This technique does not need construction and screening of DNA libraries to obtain additional unidentified DNA sequence information. Some recombinant phage or plasmid may be unstable in bacteria and amplified libraries tend to lose them (23). IPCR eliminates this problem.

Acknowledgments

The author would like to acknowledge Dr. J. Holcenberg for his invaluable comments and generous support. The author especially thanks C.-H. Wu and B. Cai for their technical assistance.

References

1. Verma, I. M., Temple, G. F., Fan, H., and Baltimore, D. (1972) In vitro synthesis of double-stranded DNA complimentary to rabbit reticulocyte 10S RNA. *Nature* **235**, 163–169.
2. Akowitz, A. and Mamuelidis, L. (1989) A novel cDNA/PCR strategy for efficient cloning of small amounts of undefined RNA. *Gene* **81**, 295–306.
3. Okayama, H., Kawaichi, M., Brownstein, M., Lee, F., Yokota, T., and Arai, K. (1987) High-efficiency cloning of full-length cDNA; Construction and screening of cDNA expression libraries for mammalian cells. *Meth. Enzymol.* **154**, 3–28.
4. Brenner, C. A., Tam, A. W., Nelson, P. A., Engleman, E. G., Suzuki, N., Fry, K. E., and Larrick, J. W. (1989) Message amplification Phenotyping (MAPPING): a technique to simultaneously measure multiple mRNAs from small numbers of cells. *BioTechniques* **7**, 1096–1103.
5. Frohman, M. A. (1990) RACE: Rapid amplification of cDNA ends, in: *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.), Academic, San Diego, CA, pp 28–38.
6. Shyamala, V. and Ames, G. F.-L. (1989) Genome walking by single-specific-primer polymerase chain reaction : SSP-PCR. *Gene* **84**, 1–8.
7. Huang, S.-H., Jong, A. Y., Yang, W., and Holcenberg, J. (1993) Amplification of gene ends from gene libraries by PCR with single-sided specificity. *Meth. Mol. Biol.* **15**, 357–363.
8. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–625.

9. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res.* **16**, 8186.
10. Huang, S.-H., Hu, Y. Y., Wu, C.-H. and Holcenberg, J. (1990) A simple method for direct cloning cDNA sequence that flanks a region of known sequence from total RNA by applying the inverse polymerase chain reaction. *Nucl. Acids Res.* **18**, 1922.
11. Delort, J., Dumas, J. B., Darmon, M. C., and Mallet, J. (1989) An efficient strategy for cloning 5' extremities of rare transcripts permits isolation of multiple 5'-untranslated regions of rat tryptophan hydroxylase mRNA. *Nucl. Acids Res.* **17**, 6439–6448.
12. Cusi, M. G., Cioe', L., and Rovera, G. (1992) PCR amplification of GC-rich templates containing palindromic sequences using initial alkali denaturation. *BioTechniques* **12**, 502–504.
13. Lau, E. C., Li, Z.-Q., and Slavkin, S. C. (1993) Preparation of denatured plasmid templates for PCR amplification. *BioTechniques* **14**, 378.
14. Green, I. R. and Sargan, D. R. (1991) Sequence of the cDNA encoding ovine tumor necrosis factor- α : problems with cloning by inverse PCR. *Gene* **109**, 203–210.
15. Zilberberg, N. and Gurevitz, M. (1993) Rapid Isolation of full length cDNA clones by "Inverse PCR:" purification of a scorpion cDNA family encoding α -neurotoxins. *Analyt. Biochem.* **209**, 203–205.
16. Austin, C. A., Sng, J.-H., Patel, S., and Fisher, L. M. (1993) Novel HeLa topoisomerase II is the II β isoform: complete coding sequence and homology with other type II topoisomerases. *Biochim. Biophys. Acta* **1172**, 283–291.
17. Delidow, B. C., Lynch, J. P., Peluso, J. J., and White, B. A. (1993) Polymerase Chain Reaction: Basic Protocols. *Meth. Mol. Biol.* **15**, 1–29.
18. Davis, L. G., Dibner, M. D., and Battey, J. F. (1986) *Basic Methods in Molecular Biology*, Elsevier Science, New York.
19. Kru, M. S. and Berger, S. L. (1987) First strand cDNA synthesis primed by oligo(dT). *Meth. Enzymol.* **152**, 316–325.
20. Promega (1996) *Protocols and Applications* 3rd ed., pp. 179–190.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, 2nd ed., Cold Spring Harbor Laboratory Press, New York.
22. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.
23. Moon, I. S. and Krause, M. O. (1991) Common RNA polymerase I, II, and III upstream elements in mouse 7SK gene locus revealed by the inverse polymerase chain reaction. *DNA Cell Biol.* **10**, 23–32.
24. Strobel, S. A. and Dervan, P. B. (1990) Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triple-helix formation. *Science* **249**, 73–75.
25. Dreyer, G. B. and Dervan, P. B. (1985) Sequence-specific cleavage of single-stranded DNA: Oligodeoxynucleotide-EDTA.Fe(II). *Proc. Natl. Acad. Sci. USA* **82**, 968–972.
26. Zhang, H., Scholl, R., Browse, J., and Somerville, C. (1988) Double strand DNA sequencing as a choice for DNA sequencing. *Nucl. Acids Res.* **16**, 1220.
27. Sugino, A., Goodman, H. M., Heynecker, H. L., Shine, J., Boyer, H. W., and Cozzarelli, N. R. (1977) Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J. Biol. Chem.* **252**, 3987.

Inverse PCR

Genomic DNA Cloning

Ambrose Y. Jong, Anna T'ang, De-Pei Liu, and Sheng-He Huang

1. Introduction

Inverse PCR (IPCR) was designed for amplifying anonymous flanking genomic DNA regions (1,2). The technique involves the digestion of source DNA, circulation of restriction fragments, and amplification using oligonucleotides that prime the DNA synthesis directed away from the core region of a known sequence, i.e., opposite of the direction of primers used in normal or standard PCR (Fig. 1). Prior to the invention of the polymerase chain reaction (PCR), the acquisition of a specific DNA fragment usually entailed the construction and screening of DNA libraries, and the traditional “walking” into flanking DNA fragments involved the successive probing of libraries with clones obtained in the prior screening. These time-consuming procedures could be replaced by IPCR. Because IPCR can be used to efficiently and rapidly amplify regions of unknown sequence flanking any identified segment of cDNA or genomic DNA, researchers do not need to construct and screen DNA libraries to obtain additional unidentified DNA sequence information using this technique. Some recombinant phage or plasmids may be unstable in bacteria and amplified libraries tend to lose them. IPCR eliminates this problem.

The IPCR is an effective method for identifying flanking cDNA and genomic DNA segments that lie outside the primers without any information (1–3). Therefore, it has been widely used for the identification of flanking DNA sequences from simple or complex genomes (4), or for obtaining promoter sequences (5,6). For example, IPCR can be used for the detection of adjacent DNA fragments in bacterial chromosomes (7), or for amplifying Tn transposon insertions (8). Several modified applications of IPCR have been created. For example, inverse PCR mutagenesis can generate epitope-tagged proteins (9), create an internal deletion (10), make a deletion library (11), or construct a regional peptide library (12).

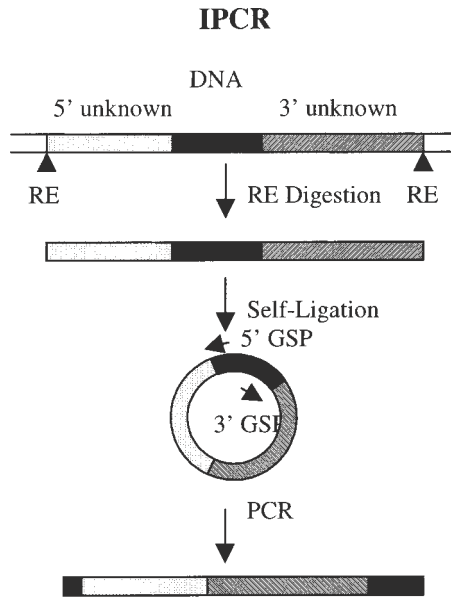


Fig. 1. Diagram of IPCR for genomic DNA cloning. The procedure consists of four steps: genomic DNA isolation, circularization of double-stranded DNA, reopening of the circular DNA, and amplification of reverse DNA fragment. The black and open bars represent the known and unknown sequence regions of double-stranded cDNA, respectively. RE: restriction enzymesite; 65p: gene-specific primer.

IPCR is relatively simple because intramolecular ligation can be achieved at low DNA concentration. One feature of IPCR is the involvement of ligation of separated regions at the ends of a sequence (a restriction fragment or a PCR product). This technique can be applied to closely linked polymorphisms. Haplotypes consisting of heterozygous polymorphisms have been determined by pedigree analysis. However, using inverse PCR, the linked polymorphisms can be analyzed by a single procedure using allele-specific primers (13). A modified protocol, long inverse PCR (LI-PCR), has been used to generate up to 10 kb flanking DNA segments (14). Another modified protocol, partial inverse PCR (PI-PCR), has been performed using a *Sau3A1* or other 4-base cutter enzymes to do partial digestion of genomic DNA before the religation step (15).

Recent genomic projects have made significant progress on the generation of physical markers for the refinement of genetic maps, determination of the linear sequences of chromosomes, and identification of short portions of cDNA clones from mRNA (expressed sequence tags, or ESTs) from eukaryotic to prokaryotic organisms. PCR has been a useful tool for amplifying specific single-copy DNA sequences from total genomic DNA. In the post-genomic era, to find appropriate methods for sequencing the two DNA fragments adjacent to an already sequenced gene or DNA fragment in

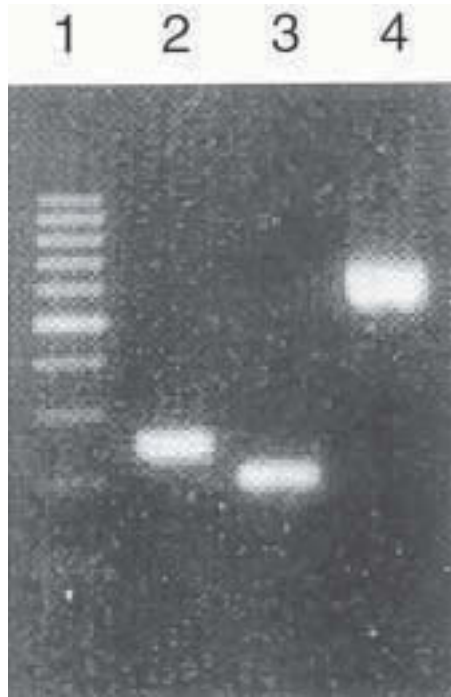


Fig. 2. Electrophoresis of the amplified products of the IPCR on 2% agarose gel. Lane 1: DNA marker, 100-bp DNA ladder, the most intense band is 500 bp. Lanes 2 and 3: the DNA from human KB-466 cells was digested by *TaqI*, ligated, and used as IPCR template. Lane 4: the KB-466 DNA was digested by *SspI*, ligated, and used as IPCR template.

eukaryotic or prokaryotic genomes is of significant importance. Needless to say, IPCR becomes useful in identifying flanking regions of a known DNA fragments (**Fig. 2**). Even if the complete sequence of a particular genome is known, IPCR is still a powerful tool; for example, IPCR can be used to analyze the viral DNA integration sites, such as hepatitis B virus DNA integrating into human hepatoma cell line (**16**).

In this chapter, we present a general protocol to describe how to use IPCR method for selectively amplifying specific genomic DNA fragments without resorting to conventional cloning procedures.

2. Materials

2.1. Genomic DNA Isolation

1. Cell lines or tissues.
2. Premix I: 1 mL 2 M Tris-HCl, pH 7.4, 1 mL 0.5 M ethylenediaminetetraacetic acid (EDTA), pH 8.0, 1 mL 20% sodium dodecyl sulfate (SDS), 7 mL phosphate buffered saline (PBS). Prepare freshly.
3. 10 mg/mL Proteinase K.

4. Phenol/chloroform/isoamyl alcohol 25:24:1.
5. $N_{250}T_{10}E_1$: 250 mM NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA.
6. 10 M Ammonium acetate.
7. Absolute ethanol.
8. 70% Ethanol.

2.2. Restriction Enzyme Digestion

1. *TaqI* (10 U/ μ L).
2. *SspI* (10 U/ μ L).
3. Other restriction enzymes.

2.3. Circularization and Cleavage or Denaturation

1. 5X ligation buffer (supplied with T4 DNA ligase).
2. T4 DNA ligase (1 U/ μ L).
3. T4 RNA ligase (4 units/ μ L).
4. 15 μ M Hexaminecobalt chloride.
5. Phenol: $CHCl_3$:isoamyl alcohol (25:24:1).
6. 3 M Sodium acetate, pH 7.0.
7. Absolute ethanol.
8. 70% Ethanol.

2.4. Inverse PCR

1. 10X PCR buffer: 100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.01%(w/v) gelatin.
2. 15 mM $MgCl_2$.
3. Deoxyoligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) 392B DNA synthesizer and purified by OPC column from the same company.
4. *Taq* DNA polymerase.

3. Methods

3.1. Genomic DNA Isolation

1. For culture cells, trypsinize, wash twice with PBS and resuspend in cold PBS. Use approx 1 mL per 2×10^7 cells. For tissues, freeze quickly in liquid nitrogen and mince into small pieces with scalpel and place in approx 5 volumes of PBS. Work quickly to prevent degradation by nucleases.
2. Add 1 mL of proteinase K (10 mg/mL) per 10 mL PBS.
3. Immediately add equal volume of Premix I. Mix gently by inversion.
4. Incubate at 50°C overnight.
5. Extract with equal volume of phenol/chloroform/isoamyl alcohol 25:24:1.
6. Transfer aqueous phase to a new tube. Re-extract the organic phase with an equal volume of $N_{250}T_{10}E_1$. Pool the aqueous phases.
7. Ethanol precipitate with one quarter volume 10 M ammonium acetate and 2 volumes ethanol. The DNA should precipitate immediately. Spool the DNA out with a Pasteur pipet, or spin for 2 min.
8. Wash DNA pellet with 70% ethanol, then dry briefly (do not overdry or DNA will be difficult to dissolve) and resuspend in 25 μ L distilled water.

3.2. Restriction Enzyme Digestion

1. Add components to the genomic DNA on ice in the following order:

10X reaction buffer	12.5 μL
Nuclear-free water	95 μL
Restriction enzymes	5 μL
2. Incubation at 37°C for 2 h.
3. Stop the reaction with 12.5 μL of 0.2 M EDTA and 200 μL sterile H₂O.
4. Concentrate and purify the sample with GeneClean. Resuspend the DNA in 100–200 μL of sterile H₂O.
5. Estimate the DNA concentration by comparing the ethidium bromide fluorescent intensity of the sample with that of a series of DNA standards on a sheet of plastic wrap. Dot 1–5 μL of sample onto plastic wrap on a UV transilluminator. Also dot with 5 μL of DNA standards. Add an equal volume of TE buffer containing 2 $\mu\text{g}/\text{mL}$ of ethidium bromide, mix by repipetting up and down. Use proper UV shielding for exposed skin and eyes.

3.3. Circularization and Cleavage (see Notes 1–4)

1. Set up the circularization reaction mix containing the following components: 100 μL (100 ng DNA, *see Note 1*) of the purified sample, 25 μL of 5X ligation buffer, and 6 μL of T4 DNA ligase. Finally, add 2 μL of T4 RNA ligase or 15 μL of 15 μM hexaminecobalt chloride (*see Note 5*).
2. Incubate at 18°C for 16 h.
3. Boil the ligated circular DNA for 2–3 min in distilled water or digest with an appropriate restriction enzyme to reopen circularized DNA (*see Notes 2–4*).
4. Purify the DNA sample with GeneClean as described in **step 6** in **Subheading 3.2.** or extract with water-saturated phenol/CHCl₃ and then precipitate with ethanol.

3.4. Inverse PCR

1. Add 1/10 of the purified cDNA to 100 μL of amplification mix:

10X PCR buffer	10 μL
15 mM MgCl ₂	10 μL
dNTP mix (2.5 mM of each)	10 μL
5'-primer (10 pmole/ μL)	10 μL
3'-primer (10 pmole/ μL)	10 μL
cDNA	10 μL
Nuclease-free H ₂ O	39.5 μL
<i>Taq</i> DNA polymerase 2.5 (μL)	0.5 μL
2. Cap and vortex the tubes to mix. Spin briefly in a microfuge. Cover each reaction with a few drops of light mineral oil to prevent evaporation.
3. Put a drop of mineral oil into each well of the thermal cycler block that will hold a tube. Load the reaction tubes.
4. Amplify by PCR using the following cycle profile:

25 cycles	94°C, 1 min (denaturation)
	65°C, 2 min (annealing)
	72°C, 4 min (elongation)

4. Notes

1. The principle of IPCR is simple and the protocol is straightforward. The trickiest part is to obtain the targeted DNA after ligation, because of the high randomness of the ligation reaction. For maximum efficiency of intramolecular ligation, a low DNA concentration should be used in the ligation mix. High density of DNA may enhance the level of heterogeneous ligation, which creates nonspecific amplification.
2. Cleavage or denaturation of circularized double-stranded DNA is important because circular double-stranded DNA has a tendency to form supercoil which is a poor template for PCR (17). Circularized double-stranded DNA is only good for amplification of a short DNA fragment.
3. The following three ways can be considered to introduce nicks in circularized DNA. Boiling is a simple and common way. Because of the unusual secondary structure of some circular double-stranded DNA, sometimes this method is not sufficient in nicking and denaturing circular double-stranded DNA. A second method is selected restriction enzyme digestion. The ideal restriction site is located within the known sequence region of DNA. In most cases, however, it is difficult to make the right choice of a restriction enzyme because the restriction pattern in the unidentified region of DNA is unknown. If an appropriate enzyme is not available, EDTA-oligonucleotide-directed specific cleavage may be tried (18). Oligonucleotide linked to EDTA-Fe at T can bind specifically to double-stranded DNA by triple-helix formation and produce double-stranded cleavage at the binding site.
4. Alkaline denaturation has been successfully used to prepare plasmid DNA templates for PCR and DNA sequencing (19). This method should be feasible in denaturing circularized double-stranded cDNA.
5. Inclusion of T4 RNA ligase or hexaminecobalt chloride can enhance the efficiency of blunt-end ligation of double-strand DNA catalyzed by T4 DNA ligase (20).

References

1. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–625.
2. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res.* **16**, 8186.
3. Huang, S.-H., Hu, Y. Y., Wu, C.-H., and Holcenberg, J. (1990) A simple method for direct cloning cDNA sequence that flanks a region of known sequence from total RNA by applying the inverse polymerase chain reaction. *Nucl. Acids Res.* **18**, 1922.
4. Li, Z. H., Liu, D. P., and Liang, C. C. (1999) Modified inverse PCR method for cloning the flanking sequences from human cell pools. *Biotechniques* **27**, 660–662.
5. Triglia, T. (2000) Inverse PCR (IPCR) for obtaining promoter sequence. *Meth. Mol. Biol.* **130**, 79–83.
6. Yoshitomo-Nakagawa, K., Muramatsu, M., and Sugano, S. (1997) Cloning of the promoter regions of mouse TGF-beta receptor genes by inverse PCR with highly overlapped primers. *DNA Res.* **4**, 73–75.
7. Pham, H. S., Kiuchi, A., and Tabuchi, K. (1999) Methods for rapid cloning and detection for sequencing of cloned inverse PCR-generated DNA fragments adjacent to known sequences in bacterial chromosomes. *Microbiol. Immunol.* **43**, 829–836.
8. Martin, V. J. and Mohn, W. W. (1999) An alternative inverse PCR (IPCR) method to amplify DNA sequences flanking Tn5 transposon insertions. *J. Microbiol. Meth.* **35**, 163–166.

9. Gama, L. and Breitwieser, G. E. (1999) Generation of epitope-tagged proteins by inverse PCR mutagenesis. *Biotechniques* **26**, 814–816.
10. Xu, Y. and Gong, Z. (1999) Adaptation of inverse PCR to generate an internal deletion. *Biotechniques* **26**, 639–641.
11. Pues, H., Holz, B., and Weinhold, E. (1997) Construction of a deletion library using a mixture of 5'-truncated primers for inverse PCR (IPCR). *Nucl. Acids Res.* **25**, 1303–1304.
12. Eisinger, D. P. and Trumpower, B. L. (1997) Long-inverse PCR to generate regional peptide libraries by codon mutagenesis. *Biotechniques* **22**, 250–254.
13. Li, Z. X., Yoshimura, S., Kobayashi, T., and Akane, A. (1998) Allele-specific, inverse-PCR amplification for genotyping MN blood group. *Biotechniques* **25**, 358–360, 362.
14. Raponi, M., Dawes, I. W., and Arndt, G. M. (2000) Characterization of flanking sequences using long inverse PCR. *Biotechniques* **28**, 838–844.
15. Pang, K. M. and Knecht, D. A. (1997) Partial inverse PCR: a technique for cloning flanking sequences. *Biotechniques* **22**, 1046–1048.
16. Wang, P., Ka-Wai, H. E., Chiu, J., and Lo, S. J. (2001) Analysis of integrated hepatitis B virus DNA and flanking cellular sequence by inverse polymerase chain reaction. *J. Virol. Meth.* **92**, 83–90.
17. Moon, I. S. and Krause, M. O. (1991) Common RNA polymerase I, II, and III upstream elements in mouse 7SK gene locus revealed by the inverse polymerase chain reaction. *DNA Cell Biol.* **10**, 23–32.
18. Strobel, A. S and Dervan, P. B. (1990) Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triplex formation. *Science* **249**, 73–75.
19. Cusi, M. G., Cioe, L., and Rovera, G. (1992) PCR amplification of GC-rich templates containing palindromic sequences using initial alkali denaturation. *Biotechniques* **12**, 502–504.
20. Sugino, A., Goodman, H. M., Heynecker, H. L., Shine, J., Boyer, H. W., and Cozzarelli, N. R. (1977) Interaction of bacteriophage T4 RNA and DNA ligase in joining of duplex DNA at base-paired ends. *J. Biol. Chem.* **252**, 3987.

Gene Cloning and Expression Profiling by Rapid Amplification of Gene Inserts with Universal Vector Primers

Sheng-He Huang, Hua-Yang Wu, and Ambrose Y. Jong

1. Introduction

Isolation of a full-length gene and analysis of expression profiling are fundamental and challenging in the current molecular biology. A great deal of effort is needed to detect unknown gene sequences by screening cDNA or genomic libraries by nucleic acid or protein probes. As the complete genome sequences of many organisms have been reported, this has raised the most challenging issue that how global gene expression patterns can be detected. Recently, various PCR methods have been developed for cloning of unknown genes and analysis of expression profiling (1–10). Several PCR-based strategies such as iAFLP (introduced amplified fragment length polymorphism) and TOGA (total gene-expression analysis) are currently available for global gene-expression analysis (8–10). An outline of gene cloning and expression profiling by anchored-PCR with vector primers is illustrated in **Fig. 1B**. In this chapter, we focus on how to use a gene library and anchored-PCR for cloning unknown gene sequences (see **Fig. 1A**). Friedmann et al. (1) first used PCR to screen λ gt11 library with two gene-specific primers. This protocol can be effectively used to isolate a particular DNA fragment between two specific primers or to generate nucleic acid probe from cDNA libraries. The unknown sequences flanking the fragment between the two specific primers can not be amplified by this method.

Anchored-PCR or single-specific-primer PCR (2,3) and inverse PCR (4) have been adapted to cloning of full length cDNAs with the knowledge of a small stretch of sequence within the gene. Both methods start from mRNA and are good for cDNA cloning when a cDNA library is not available. Cloning of full-length cDNA is usually far more difficult than any other recombinant DNA work because the multiple sequential enzymatic reactions often result in low yield and truncated clones (11). Shyamala and Ames (3) extended the use of anchored PCR to amplify unknown DNA sequences

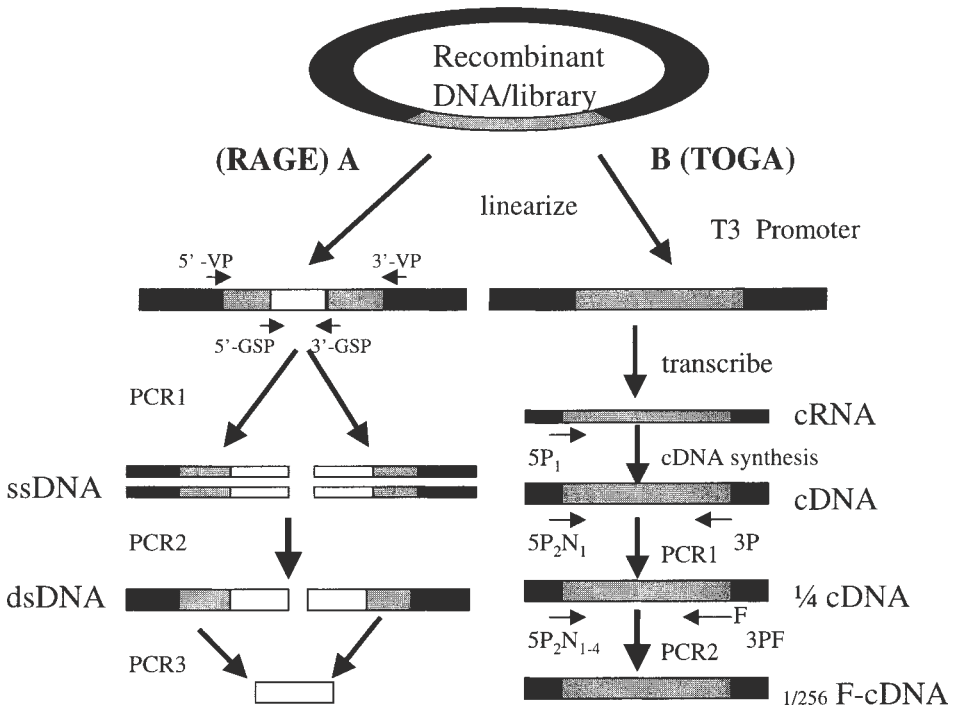


Fig. 1. The scheme for Gene Cloning and Expression Profiling by amplification of gene inserts from a library. **(A)** Gene cloning by rapid amplification of gene ends (RAGE). If a plasmid library is used, a restriction enzyme rarely within gene inserts such as *NotI* may be selected to linearize the recombinant DNA. A phage library can be directly used for PCR. Here is an example for cloning cDNA ends from a phage library (λ gt11). 5'-VP is λ gt11 forward primer (5'-GACTCCTGGAGCCCG-3'). 3'-VP: λ gt11 reverse primer (5'-GGTAGCGACCGGCGC-3'). 5'-GSP (gene specific primer): 5ASPR with *EcoRI* restriction site (5'-AGACTGAATTCGGTACCGGCGGTACTATCGCTTCC-3'). 3'-GSP: 3ASPB containing *BamHI* site (5'-CTGATGGATCCTGGCAGTGGCTGGACGC-3'). **(B)** Expression profiling by TOGA. Poly(A) RNA isolation, cDNA synthesis and library construction were carried out as described previously (10). The 3' *MspI-NotI* fragments were directionally cloned into *ClaI-NotI* cleaved plasmid pBC SK+ (Stratagene) in an orientation antisense to its T3 promoter. After cleavage with *MspI* to linearize insert-containing plasmids, antisense transcripts were produced with T3 RNA polymerase. The cRNAs were reverse transcribed into cDNA by using 5P₁ [a 5' vector primer including a 4-nt restriction endonuclease cleavage site (4-nt REC) derived from the ligation of the vector with the 5' insert]. PCR1 was performed with primers 5P₂N₁ [a 5' vector primer plus the 4-nt REC and the first adjacent nucleotide (N₁=A,G,T,C) of the insert at the 3' end] and 3P (a 3' vector primer). This step subdivides the cDNA species into four pools. PCR2 was carried out with 3PF (a fluorescent 3' vector primer) and 5P₂N₁₋₄ [(a 5' vector primer containing the 4-nt REC and four 3' degenerate nucleotides (N₁₋₄, N=A,C,G,T), which are the adjacent nucleotides of the insert)]. This amplification subdivides the input species into 256 subpools for electrophoretic resolution.

from genome on the basis of using a short stretch of known sequence for designing a gene-specific primer. We developed a much simpler method to isolate full-length cDNA and flanking genomic DNA from gene libraries by anchored-PCR. The principle of this technique is schematically depicted in **Fig. 1A**. Briefly, two ends of a gene were amplified by two vector primers (VP) complementary to the vector sequences flanking the polylinker region and two gene-specific primers (GSP) complementary to 5' and 3' parts of the known gene sequence.

We used the yeast gene coding for asparaginase II (**12**) as a model to verify this method. Two ends of the yeast asparaginase II gene (**12**) were amplified from λ gt11 yeast genomic library by this method. The size of the full-length gene is 1.7 kb and there is 353 bp overlapping sequence between the two GSPs (5ASPR and 3ASPB). The 3' gene end (from 5ASPR to the extreme 3' end of the gene) and the 5' gene end (from 3ASPB to the very beginning of 5' part of the gene) are 1.15 and 0.9 kbs, respectively. Because there may be two different orientations for each insert in the gene library, four reactions are performed for the amplification of two gene ends in the first round of PCR. **Fig. 2** shows that there was only one orientation for the asparaginase II gene clone and the expected sizes of the two amplified gene ends were obtained. The same size fragment was amplified from the gene library and the two purified gene ends by using the two GSPs as primers. The gel-purified PCR fragments were sequenced and the sequence data were consistent with the literature (**12**).

We amplified and isolated 5'-cDNA fragment of 5-hydroxytryptamine 2 (5-HT₂) receptor from lambda SWAJ-2 mouse brain cDNA library by this technique. 5'-cDNA region of 5-HT₂ receptor was successfully amplified from λ SWAJ-2 mouse brain cDNA library with a GSP (5'-TTCTGCCTGAGACTAAAAAGGGTTAAGCCCTTAGATGGCA-3') and a VP (Xba I-T15 adaptor primer: 5'-GTCGACTCTTAGAT-3') (**5**).

2. Materials

1. *Taq* DNA polymerase (Perkin/Elmer, Foster City, CA).
2. dNTP: 10 mM of each (Perkin/Elmer).
3. Yeast genomic library (ClonTech, Palo Alto, CA). See **Note 1**.
4. λ SWAJ-2 mouse brain cDNA library (ClonTech).
5. 5' and 3' λ gt11 primers (ClonTech).
6. GeneClean™ kits (Bio101, La Jolla, CA).
7. Sequenase kits (United States Biochemical Corporation, Cleveland, OH).
8. 10X *Taq* DNA polymerase buffer: 67 mM Tris-HCl, pH 8.8, 6.7 mM MgCl₂, 170 μ g/mL bovine serum albumin (BSA), 16.6 mM ammonium sulfate (New England Biolab, Beverly, MA) was prepared as described (2; see **Note 4**).
9. Oligodeoxyribonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer and purified by OPEC column from the same company.
10. Programmable Thermal Controller.

3. Methods

The following conditions were used to amplify the ends of the yeast asparaginase II gene from a λ gt11 yeast genomic library.

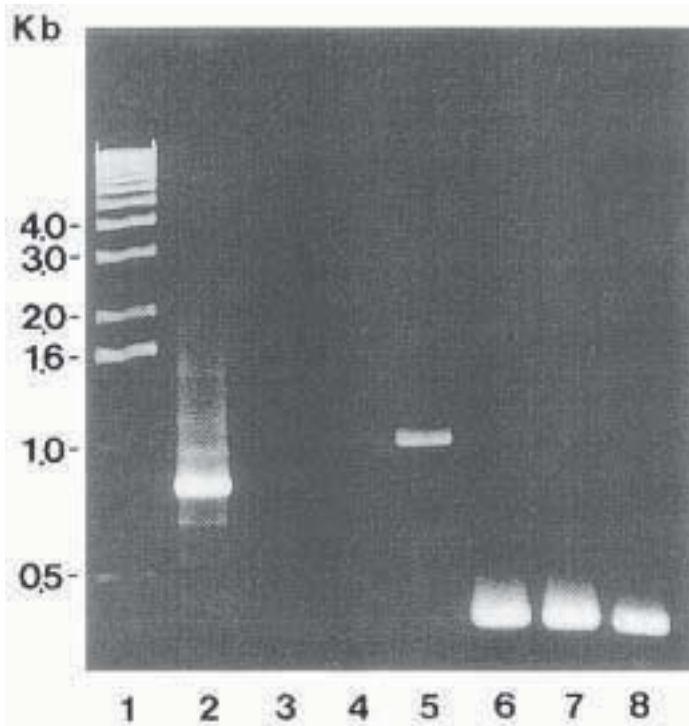


Fig. 2. Gel pattern of PCR products. Fragments amplified from the purified 5' (lane 6) and 3' (lane 7) gene ends and λ gt11 yeast genomic library (lane 2, 3, 4, 5, and 8) were resolved in 1.5% agarose gel and stained with ethidium bromide. Lane 1 was 1 Kb ladder DNA marker (BRL). There were five pairs of primers used in PCR: 1) 5'-VP and 3'-GSP (lane 2); 2) 3'-VP and 3'-GSP (lane 3); 3) 5'-VP and 5'-GSP (lane 4); 4) 3'-VP and 5'-GSP (lane 5), and 5) 5'-GSP and 3'-GSP (lane 6,7 and 8).

3.1. PCR#1 (see Notes 2 and 3)

1. In order to enhance the specificity of amplification, asymmetric PCR was carried out in 50 μ L of reaction mixture containing 1 μ L aliquot of λ gt11 yeast genomic library (about 1×10^7 PFU), 2.5 pmol vector primer (VP), 50 pmol GSP, 5 μ L of dimethylsulfoxide, 5 μ L of 10X *Taq* DNA polymerase buffer (2), 1.5 mM of each dNTP, and 2.5 U of *Taq* DNA polymerase. Before adding the enzyme, the PCR cocktail was heated at 94°C for 3 min to disrupt the phage particles.
2. PCR parameters were: 35 cycles at 94°C for 1 min, 48°C for 30 s, and 72°C for 8 min. The major product, single-strand DNA, can be used for sequencing and Southern blotting analyses.

3.2. PCR#2

1. Dilute the first PCR product at 1:10 in H₂O.
2. Amplify 1- μ L aliquots as described in PCR#1, with the exception of using equal amounts (50 pmol) of the VP and GSP. The selected dsDNA fragments are available for cloning, which can be facilitated by incorporating restriction sites in the primers.

3.3. PCR#3

PCR 3 is used to test the products of PCR 2 because amplification of the two GSPs should produce the same fragment from a gene library and the two purified gene ends.

1. Assemble a PCR with 1 μ L of 1:10 diluted PCR2 product or 1- μ L aliquot of gene library were used in the third round of PCR using two GSPs (50 pmol each).
2. Amplify for 35 cycles at 94°C for 30 s, 48°C for 30 s and 72°C for 4 min.
3. Purify DNA fragments from PCR#1, 2, and 3 with GeneClean™ (BIO 101 Inc.) from agarose gel and then sequence by the dideoxy chain termination method (13) with Sequenase (USB) with the aid of NP40 (14) (see Notes 4 and 5).

4. Notes

1. The successful isolation and detection of gene inserts from cDNA or genomic libraries is dependent on the quality of the library. There is a big difference between the primary and the amplified libraries because different recombinant clones may grow at very different rates, resulting in unequal distribution of the recombinants in the amplified libraries (19). It may be better to use the primary library for PCR amplification. If a short stretch of known DNA sequence (more than 100 bp) is available, it is easy to test the quality of the library by PCR with two GSPs. In case the library is not good, an alternative strategy is to use inverse or anchored PCR to isolate or detect gene inserts from genomic DNA or self-made cDNA.
2. The most obvious and common problem for PCR with single-sided specificity is nonspecific amplification of DNA fragments without significant homology with the gene of interest (15,17). In this chapter we found two modifications that improve specific amplification. At first, we used PCR buffer from Perkin Elmer Cetus and failed to amplify specific products for yeast asparaginase II gene. It appears that the *Taq* DNA polymerase buffer from New England Biolab is much better for PCR with single-specific primer. Second, asymmetric PCR with a relatively large amount of GSP was performed in the first round of PCR in order to enhance the specificity of amplification.
3. There is a limitation for DNA amplification with degenerate primers based on the highly conserved regions of a protein from other species or limited amino acid sequence data because degeneracy of primers can create more problems with nonspecific amplification. This limit may be reduced by incorporation of deoxyinosine into wobble positions of degenerate oligonucleotides (15).
4. This method can be used for directional genome walking from known into unknown flanking regions of the chromosomal DNA with genomic libraries. PCR amplification of large DNA fragments (up to 35 kb) has been achieved by the combination of a high level of an exonuclease-free, N-terminal deletion mutant of *Taq* DNA polymerase, *Klentaq1*, with a very low concentration of a thermostable DNA polymerase exhibiting a 3'-exonuclease activity (*Pfu*, *Vent*, or *Deep Vent*) (18).
5. The two gene ends with overlapping sequence can be simply linked by two ways. First, ssDNAs of the two gene ends from PCR1 can be annealed and end filled by Klenow in the presence of random primers. Second, when the sequence information is obtained from the two gene ends, a full length cDNA or gene can be amplified by two specific primers that represent the sequences at the extreme 3' and 5'-ends of the gene or cDNA.

Acknowledgments

The authors would like to acknowledge Dr. J. Holcenberg for his generous support. We especially thank C.-H. Wu and B. Cai for their technical assistance. This work was supported by the T.J. Martell Foundation.

References

1. Friedmann, K. D., Rosen, N. L., Newman, P. J., and Montgomery, R. R. (1988) Enzymatic amplification of specific cDNA inserts from λ gt11 libraries. *Nucl. Acids Res.* **16**, 8718.
2. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 998–9002.
3. Shyamala, V. and Ames, G. F.-L. (1989) Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene* **84**, 1–8.
4. Huang, S. H., Hu, Y. Y., Wu, C. H., and Holcenberg, J. (1990) A simple method for direct cloning cDNA sequence that flanks a region of known sequence from total RNA by applying the inverse polymerase chain reaction. *Nucl. Acids Res.* **18**, 922.
5. Huang, S.-H., Jong, A. Y., Yang, W., and Holcenberg, J. (1993) Amplification of gene ends from gene libraries by polymerase chain reaction with single-sided specificity. *Meth. Mol. Biol.* **15**, 357–363.
6. Tsurui, H., Hara, E., Oda, K., Suyama, A., Nakada, S., and Wada, A. (1990) A rapid and efficient cloning method with a solid-phase DNA probe: application for cloning the 5'-flanking region of the gene encoding human fibronectin. *Gene* **88**, 233–239.
7. Cormack, R. S. and Somssich, I. E. (1997) Rapid amplification genomic ends (RAGE) as a simple method to clone flanking genomic DNA. *Gene* **194**, 273–276.
8. Kawamoto, S., Ohnishi, T., Kita, H., Chisaka, O., and Okubo, K. (1999) Expression profiling by iAFLP: A PCR-based method for genome-wide gene expression profiling. *Genome Res.* **9**, 1305–1312.
9. Wang, A., Pierce, A., Judson-Kremer, K., Gaddis, S., Aldaz, C. M., Johnson, D. G., and MacLeod, M. C. (1999) Rapid analysis of gene expression (RAGE) facilitates universal expression profiling. *Nucl. Acids Res.* **27**, 4609–4618.
10. Sutcliffe, J. G., Foye, P. E., Erlander, M. G., Hilbush, B. S., Bodzin, L. J., Durham, J. T., and Hasel, K. W. (2000) TOGA: an automated parsing technology for analyzing expression of nearly all genes. *Proc Natl Acad Sci USA* **97**, 1976–1981.
11. Okayama, H., Kawaichi, M., Brownstein, M., Lee, F., Yokota, T., and Arai, K. (1987) High-efficiency cloning of full-length cDNA: construction and screening of cDNA expression libraries for mammalian cells. *Meth. Enzymol.* **154**, 3–28.
12. Kim, K. W., Kamerud, J. Q., Livingston, D. M., and Roon, R. J. (1988) Asparaginase II of *Saccharomyces cerevisiae*: characterization of the ASP3 gene. *J. Biol. Chem.* **263**, 11,948–11,953.
13. Sanger, F., Nickler, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Bachmann, B., Lucke, W., and Hunsmann, G. (1990) Improvement of PCR amplified DNA sequencing with the aid of detergents. *Nucl. Acids Res.* **18**, 1309.
15. Patil, R. V. and Dekker, E. E. (1990) PCR amplification of an *Escherichia coli* gene using mixed primers containing deoxyinosine at ambiguous positions in degenerate amino acid codons. *Nucl. Acids Res.* **18**, 3080.
16. Frohman, M. A. (1990) RACE: rapid amplification of cDNA Ends, in *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.), Academic, San Diego, CA, pp. 28–38
17. Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989) Polymerase chain reaction with single-sided specificity: analysis of T cell receptor δ chain. *Science* **243**, 217–220.
18. Barnes, W. M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
19. Frischauf, A.-M. (1987) Construction and characterization of a genomic library in λ . *Meth. Enzymol.* **152**, 190–199.

The Isolation of DNA Sequences Flanking Tn5 Transposon Insertions by Inverse PCR

Vincent J. J. Martin and William W. Mohn

1. Introduction

Because of its versatility, the Tn5 transposon has become a powerful tool in the classical genetic studies of Gram-negative bacteria. The Tn5 transposon is functional in a broad range of Gram-negative bacteria and transposes at a high frequency with low specificity of insertion (1,2), allowing it to insert in a large number of locations in bacterial genomes. The initial use of the Tn5 transposon was in the identification of disrupted genes by insertion mutagenesis and phenotypic screening, however, derivatives of this mobile element were later developed for use in reporter gene fusion (3) and promoter probing experiments (4).

The characterization of the region of Tn5 insertion in a genome requires the isolation of DNA flanking the insertion. This was conventionally accomplished by cloning the mutagenized DNA fragment containing both upstream and downstream regions relative to the insertion, along with the transposon (5). To isolate the flanking DNA, clone libraries were constructed from the Tn5 mutated DNA and screened on selective medium containing kanamycin. This approach for isolating the DNA of interest is time-consuming and labor-intensive because it requires the production and screening of genomic clone libraries and subcloning of the positive clones prior to sequencing.

Several methods were developed to reduce the time required to isolate and sequence DNA regions flanking transposon insertions. These methods include vector-mediated rescue of mutations (6,7) and polymerase chain reaction (PCR)-based methods such as linker-mediated, single-primer, anchored, repetitive extragenic palindrome (REP) and inverse PCR (8-14). Because of their ease, these methods also allow for the simultaneous characterization and isolation of large numbers of clones from a mutation library.

In this chapter, we describe a simple and reliable experimental procedure to clone DNA flanking a Tn5 transposon by inverse PCR (IPCR). The IPCR method is based on the initial report from Rich and Willis (12) and incorporates the improvements from Martin and Mohn (13) and Huang et al. (14). The basic strategy for the Tn5 IPCR

is shown schematically in **Fig. 1**. The technique requires three basic steps: 1) the restriction enzyme digestion of Tn5-tagged genomic DNA, 2) circularization of the fragments by intra-molecular ligation, and 3) amplification with PCR primers designed to anneal to the Tn5 transposon. In a first approach, the choice of restriction endonuclease to use in IPCR template preparation is established by Southern blot analysis using an asymmetric probe, which hybridizes to a small region of both the left and right side IS50 regions as well as the genes encoding antibiotic resistance (*see Fig. 2*). For enzymes that do not cut the transposon, a single band will appear in the Southern blot corresponding to 5.8 kb + the size of flanking DNA from both ends. In using enzymes that cut once within the transposon, two bands will appear from the Southern blot. The thick band corresponds to the left side of the Tn5, which is complementary to the entire probe and the light band to the right side of the transposon, which is complementary to only one end of the probe (*see Fig. 2*). The size of the IPCR products from templates prepared from each of the restriction endonucleases tested can be inferred from the size of the bands in the Southern blot analysis. Although knowing the expected size of the IPCR product allows a certain level of confidence in the results, it is not always necessary. Huang et al. (**14**) reported on a method using a nested PCR primer (IR2, **Table 1**), which allowed discrimination of target from nonspecific products without resorting to Southern blot analysis.

2. Materials

2.1. Southern Blot Analysis

A list of general materials and methods used for genomic DNA preparations, Southern blot analysis of genomic DNA and radioisotope labeling of DNA probes is found in Sambrook and Russell (**15**). Additional requirements follow.

1. Genomic DNA (2.5–5 µg per restriction endonuclease tested) prepared from Tn5 transposon mutant strain(s) of interest.
2. A selection of restriction endonucleases that do not cut the Tn5 transposon or that cut the transposon only once (*see Note 1*). Restriction endonucleases are normally supplied with their appropriate buffers.

Examples of enzymes that do not cut the Tn5 transposon are: *AatII*, *AflIII*, *ApaI**, *ApaLI*, *AseI*, *AvrII*, *BsrGI*, *Clal**, *DraI***, *EcoRI*, *EcoRV***, *KpnI*, *MluI*, *NdeI*, *NsiI*, *PmlI***, *PvuI*, *SacI**, *ScaI***, *SpeI*, *StuI* *** and *XbaI*(for * and **, see Note 1)*.

Enzymes that cut the Tn5 transposon only once are: *BamHI*, *SalI*, *SmaI**/XmaI*, and *SphI*.

3. *HindIII-BamHI* 1861-bp probe from the Tn5 transposon (*see Fig. 2*) labeled with [$\alpha^{32}\text{P}$]-dCTP by nick translation (labeling kit sold by Life Technologies) (*see Notes 2 and 3*).

2.2. Preparation of Inverse PCR Template

1. Sterile, redistilled water.
2. Genomic DNA (1 µg) isolated from Tn5 mutant strain(s) of interest.
3. Restriction endonuclease(s) selected from the results of the Southern blot.
4. TE saturated phenol:chloroform:isoamyl alcohol (25:24:1).
5. Absolute ethanol.
6. 70 % ethanol.

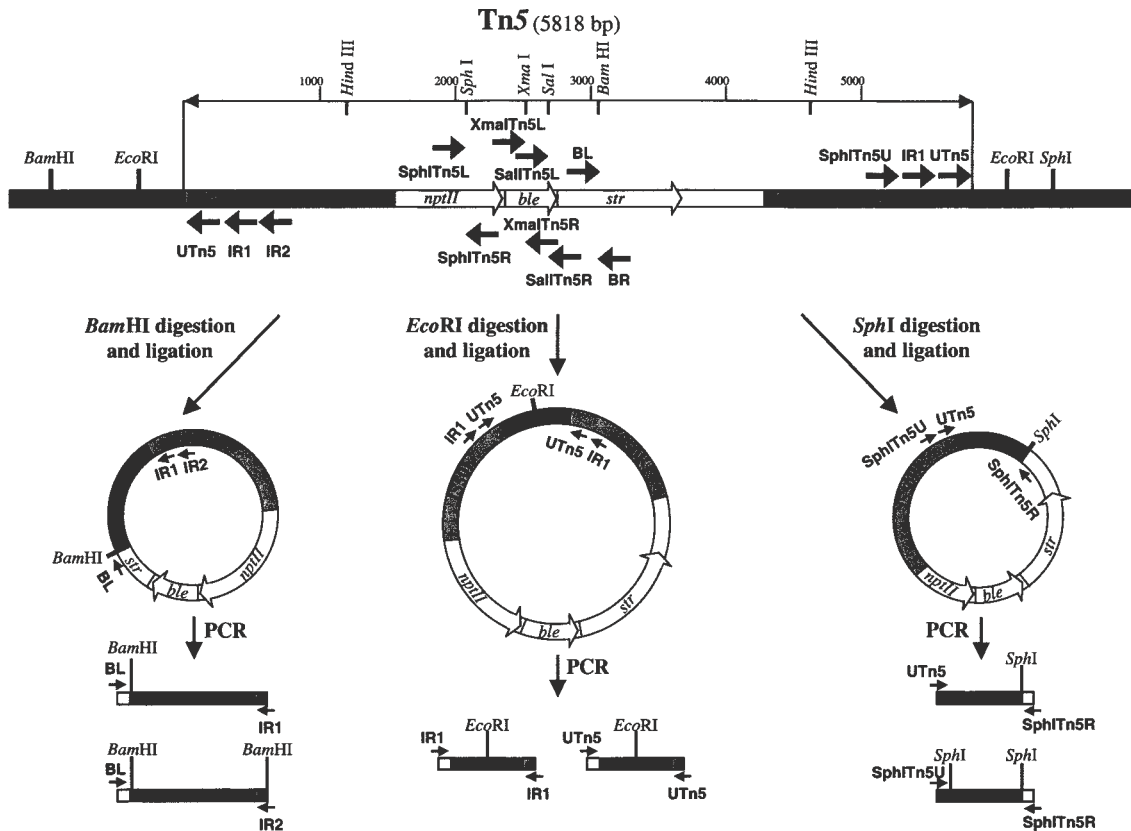


Fig. 1. Schematic diagram showing the strategy to amplify DNA sequences flanking the Tn5 transposon by IPCR. Black bars represent unknown DNA sequences flanking a Tn5 insertion; gray bars represent the left (IS50_L) and right (IS50_R) insertion sequences. Small arrows symbolize the location of the PCR primers described in Table 1. Antibiotic resistance genes are *nptII*, neomycin/kanamycin; *ble*, bleomycin ; *str*, streptomycin.

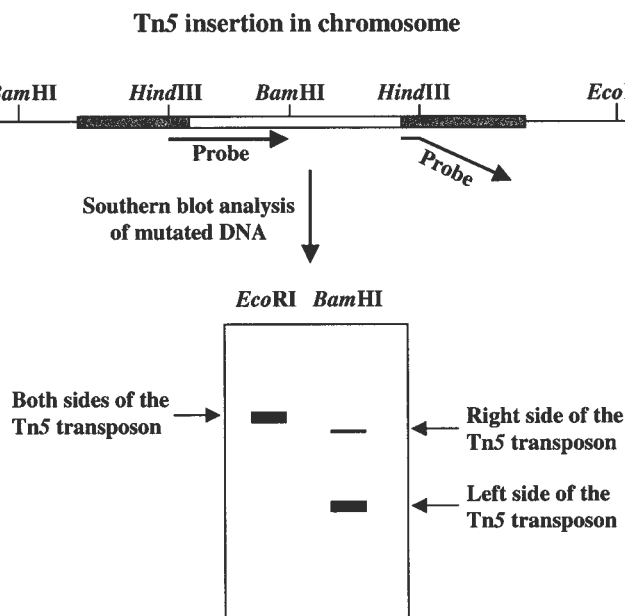


Fig. 2. Schematic diagram of the Southern blot analysis strategy used to predict the size of the IPCR products [based on Mitchell and Smit, 1990 (17)]. The gray bars represent the left and right insertion sequences IS50, which are inverted terminal repeats in the Tn5 transposon. The light band in the blot represents hybridization of the *HindIII*-*BamHI* DNA probe to the right side of the transposon and the thick band represents hybridization to the left side of the transposon.

7. 3 M cold sodium acetate pH 7.0.
8. T4 DNA ligase (400 U/ μ L) (New England Biolabs).
10. 10X ligation buffer containing 10 mM ATP (supplied with T4 DNA ligase).

2.3. Inverse PCR

1. *Taq* and *Pwo* DNA polymerase enzyme mix (Expand High Fidelity PCR System, Roche Molecular Biochemical) (see Note 4).
2. 10X PCR buffer (Expand HF buffer supplied with the PCR system).
3. Sterile 25 mM $MgCl_2$ stock solution (supplied with the PCR system).
4. PCR grade nucleotide mix containing 10 mM of each dNTP.
5. Deoxyoligonucleotide primers (Table 1).
6. Light mineral oil (may not be required if thermocycler uses a heated lid).
7. pRZ705, pSUP2021, or any other plasmid containing a Tn5 insertion to be used as positive control template for the IPCR reaction (optional).

3. Methods

3.1. Preparation of Inverse PCR Template

1. Based on the results of the Southern blot analysis, select an endonuclease enzyme(s) to use for preparing the IPCR template(s). The enzymes should produce an IPCR template that will yield IPCR products of ≤ 5 kb (see Note 4). The predicted size of the IPCR

Table 1
List of Primers for Tn5 IPCR

Name	DNA Sequence	Length (bp)	Restriction enzyme	Region of Tn5 transposon ^a	Reference
UTn5	5'-GGTTCGTTTCAGGACGCTAC-3'	20	All enzymes	37–18 and 5782–5801	(12)
IR1	5'-GAGCAGAAGTTATCATGAACG-3'	21	All enzymes	107–87 and 5712–5732	(14)
XmaITn5R	5'-AGGCAGCAGCTGAACCAA-3'	18	<i>XmaI</i>	2556–2539	(13)
XmaITn5L	5'-GCCGGCTGGATGATCCTC-3'	18	<i>XmaI</i>	2462–2479	This study
XmaITn5U ^b	5'-CGCCCGGGTCACATGGAAGTCAGATCCTG-3'	29	<i>XmaI</i>	70–50 and 5749–5769	This study
SalITn5R	5'-ATGCCTGCAAGCAATGC-3'	18	<i>SalI</i>	2739–2722	(13)
SalITn5L	5'-AACCAGCAGCGGCTATCC-3'	18	<i>SalI</i>	2614–2031	This study
SalITn5U ^b	5'-TGGGTCGACTCACATGGAAGTCAGATCCTG-3'	30	<i>SalI</i>	70–50 and 5749–5769	This study
SphITn5R	5'-ATTCGGCAAGCAGGCATC-3'	18	<i>SphI</i>	2135–2118	This study
SphITn5L	5'- TGGACGAAGAGCATCAGG-3'	18	<i>SphI</i>	2023–2040	This study
SphITn5U ^b	5'-CCGCATGCTCACATGGAAGTCAGATCCTG-3'	29	<i>SphI</i>	70–50 and 5749–5769	This study
BR	5'- CATTCTGTAGCGGATGGAGATC -3'	23	<i>BamHI</i>	3134–3112	(14)
BL	5'- GGGGACCTTGACAGATAGC -3'	20	<i>BamHI</i>	3007–3026	(14)
IR2	5'-CGGGATCCTCACATGGAAGTCAGATCCTG-3'	29	<i>BamHI</i>	70–50 and 5749–5769	(14)

^aBased on the complete sequence of *Escherichia coli* transposon Tn5 from GenBank accession# U00004

^bThe annealing regions of these primers are based on the IR2 primer of Huang et al. (14).

product is calculated by subtracting the size of the Tn5 DNA fragment flanking the target sequence from the size of the band that hybridized to the probe (see Fig. 2).

2. Digest 1 μg of the genomic DNA.
3. Heat inactivate the endonuclease(s) at 65°C for 20 min and centrifuge briefly to collect condensate.
4. Extract the digested DNA with an equal volume of TE saturated phenol:chloroform:isoamyl alcohol (25:24:1).
5. Precipitate the DNA with 1/10 vol cold sodium acetate and 2 vol of absolute ethanol. Centrifuge and wash the DNA pellet with 70% ethanol (15). Suspend the DNA in 180 μL of sterile H₂O.
6. Set up the template circularization reaction with the 180 μL of cut genomic DNA solution, 20 μL of 10X T4 ligase buffer and 0.5 μL (200 U) of T4 ligase (see Note 5).
7. Incubate at 16°C for 12 h (overnight).
8. Precipitate the circularized template DNA with ethanol, wash with 70% ethanol, and suspend DNA in 20 μL of sterile H₂O.

3.2. Inverse PCR Reaction

1. In a PCR tube, combine the following components for the inverse PCR amplification reaction: 10 μL of 10X PCR buffer, 6 μL of 25 mM MgCl₂ (final concentration of 1.5 mM), 2 μL 10 mM dNTP mix (final concentration of 200 μM each dNTP), 0.5 μM of each primer, 10 μL (500 ng) of circularized template and sterile redistilled H₂O to 99.25 μL . Mix the components and centrifuge briefly in a microfuge.
2. In a thermocycler, heat the reaction mix to 95°C for two minutes before adding 0.75 μL (2.6 U) of the polymerase enzyme mix ("hot start" PCR is used to improve the specificity, sensitivity, and yield of the reaction).
3. Amplify the target DNA sequence using the following cycle profile: first five cycles: 94°C for 30 s (denaturation), 60°C for 30 s (annealing) and 72°C (\leq 3kb) or 68°C (>3 kb) for 1 min/1.5 kb (elongation). For the last 25–30 cycles reduce the annealing temperature to 55 °C.
4. Analyze samples on a 0.7% agarose gel.

3.3. Alternative Method

Huang et al. (14) described the use of a nested PCR primer from the IS50 regions to distinguish the target product from nonspecific IPCR products. Using a nested primer set and two PCRs, it is possible to forgo the Southern blot analysis to determine the size of the PCR product (see Fig. 1).

1. Prepare the IPCR template as described in the materials and methods.
2. Set up a pair of PCR reactions for each template. Primer pairs to amplify either both, the left or right side of Tn5 are listed in Table 2.
3. The extension step of the temperature cycle profile should be lengthened to 10 min to ensure that long target sequences will be amplified.
4. Analyze the PCR products on a 0.7% agarose gel. Bands of the same size in both PCR and nested PCR indicate which is the correct product for the target sequence.

3.4. Cloning and Sequencing of IPCR Products

Because of the nature of the inverted terminal repeats of Tn5, it does not contain unique sites for sequencing primers to anneal at its ends. Therefore, IPCR products from templates generated by enzymes not cutting the transposon must be cloned first in order to provide priming sites for sequencing. This is achieved by blunt-end or TA cloning.

Table 2
List of Primer Pairs for Nested Tn5 IPCR

Enzyme used in template preparation		<i>Xma</i> I			<i>Sal</i> I		<i>Sph</i> I		<i>Bam</i> HI	
Side of Tn5 flanking DNA amplified	Both	Left	Right	Left	Right	Left	Right	Left	Right	
Primer pair	IR1 ^b	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c	UTn5 ^c
		XmaITn5L	XmaITn5R	SalITn5L	SalITn5R	SphITn5L	SphITn5R	BL	BR	
Nested primer pair	UTn5 ^b	XmaITn5U	XmaITn5U	SalITn5U	SalITn5U	SphITn5U	SphITn5U	IR2	IR2	
		XmaITn5L	XmaITn5R	SalITn5L	SalITn5R	SphITn5L	SphITn5R	BL	BR	

^aThe PCR product from these reactions can be cloned with cohesive ends introduced with the nested primer.

^bA single oligonucleotide primer is used for these PCRs.

^cThis primer can be substituted with IR1, if desired.

However, sequencing of the IPCR products without cloning is possible from amplicons of the left or right sides of Tn5 using the UR1 or UTn5 primers. Furthermore, the addition of a restriction endonuclease recognition site in the nested primers makes possible cohesive-end cloning of the PCR product (*see Fig. 1*)

4. Notes

1. * These enzymes can be blocked by overlapping DNA methylation sites. ** These enzymes leave a blunt-end on the cut DNA and may require the addition of PEG (**16**) or increased enzyme concentration in the reaction to increase ligation efficiency.
2. To simplify the preparation of the DNA probe, sub-clone the *HindIII-BamHI* fragment into a high-copy vector such as pUC19 and select on LB-kanamycin. Sufficient amounts of probe can be made by PCR using the M13 primers or by isolating the fragment from the high copy vector.
3. Finding the proper condition for washing the membrane is important because the asymmetric DNA probe partially hybridizes to the right IS50 target sequence and is easily washed off if the conditions are too stringent. Slowly increasing the wash stringency and using a phosphorimager with 1 h exposures to look at the blot between washes will ensure that you do not overwash the blot.
4. Using *Taq* DNA polymerase, the maximum size of the IPCR products we were able to amplify from a genomic template was approx 3 kb. The length of the IPCR product amplified may be increased by using a proofreading polymerase or a blend of the two enzymes. Huang et al. (**14**) were able to amplify, at a low yield, a 6.2 kb IPCR product using a *Taq/GB-D* polymerase enzyme blend. However, an additional limitation to the size of the IPCR fragment amplified is the low probability of intramolecular ligation of long DNA fragments.
5. Ligation reactions must be performed with dilute DNA samples (<5µg/mL) to favor intramolecular ligation.

References

1. Goryshin, I. Y., Miller, J. A., Kil, Y. V., Lanzov, V. A., and Reznikoff, W. S. (1998) Tn5/IS50 target recognition. *Proc. Natl. Acad. Sci. USA* **95**, 10,716–10,721.
2. Berg, D. E. (1989) Transposon Tn5 in Mobile DNA, (Berg, D. E. and Howe, M. M., eds.), American Society for Microbiology, WA, pp. 163–184.
3. De Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.* **172**, 6568–6572.
4. Simon, R., Oquandt, J., and Klipp, W. (1989) New derivatives of transposon Tn5 suitable for mobilization of replicons, generation of operon fusions and induction of genes in Gram-negative bacteria. *Gene* **80**, 161–170.
5. DeBruijn, F. J. and Rossbach, S. (1993) Transposon mutagenesis in *Methods for General and Molecular Bacteriology*, (Gerhardt, P., ed.), American Society for Microbiology, WA, pp. 387–405.
6. Dennis, J. J. and Zylstra, G. J. (1998) Plasmids: Modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl. Environ. Microbiol.* **64**, 2710–2715.
7. Hozbor, D. F., Pich Otero, A. J. L., Wynne, M. E., Petrucci, S., and Lagares, A. (1998) Recovery of Tn5-flanking bacterial DNA by vector-mediated walking from the transposon to the host genome. *Analyt. Biochem.* **259**, 286–288.

8. Ribot, E. M., Quinn, F. D., Bai, X., and Murtagh, J. J. (1998) Rapid amplification of transposon ends for the isolation, cloning and sequencing of transposon-disrupted chromosomal genes. *Biotechniques* **24**, 16–22.
9. Prod'hom, G., Lagier, B., Pelicic, V., Hance, A. J., Gicquel, B., and Guilhot, C. (1998) A reliable amplification technique for the characterization of genomic DNA sequences flanking insertion sequences. *FEMS Microbiol. Lett.* **158**, 75–81.
10. Karlyshev, A. V., Pallen, M. J., and Wren, B. W. (2000) Single-primer PCR procedure for rapid identification of transposon insertion sites. *Biotechniques* **28**, 1078–1081.
11. Subramanian, P. S., Versalovic, J., McCabe, E. R., and Lupski, J. R. (1992) Rapid mapping of *Escherichia coli*:Tn5 insertion mutations by REP-Tn5 PCR. *PCR Meth. Appl.* **1**, 187–192.
12. Rich, J. J. and Willis, D. K. (1990) A single oligonucleotide can be used to rapidly isolate DNA sequences flanking a transposon Tn5 insertion by the polymerase chain reaction. *Nucl. Acid Res.* **18**, 6673–6676.
13. Martin, V. J. J. and Mohn, W. W. (1999) An alternative inverse PCR (IPCR) method to amplify DNA sequences flanking Tn5 transposon insertions. *J. Microbiol. Meth.* **35**, 163–166.
14. Huang, G., Zhang, L., and Birch, R. G. (2000) Rapid amplification and cloning of Tn5 flanking fragments by inverse PCR. *Lett. Appl. Microbiol.* **31**, 149–153.
15. Sambrook, J. and Russell, D. W. (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press.
16. Upcroft, P. and Healey, A. (1987) Rapid and efficient method for cloning of blunt-ended DNA fragments. *Gene* **51**, 69–76.
17. Mitchell, D. and Smit, J. (1990) Identification of genes affecting production of the adhesion organelle of *Caulobacter crescentus* CB2. *J. Bacteriol.* **172**, 5425–5431.

Rapid Amplification of Genomic DNA Sequences Tagged by Insertional Mutagenesis

Martina Celerin and Kristin T. Chun

1. Introduction

Current and recent efforts to determine the genomic DNA sequence for numerous organisms (e.g., *Saccharomyces cerevisiae*, *Candida albicans*, *Neurospora crassa*, *Arabidopsis thaliana*, *Zea mays*, *Caenorhabditis elegans*, *Mus musculus*, *Homo sapiens*, *Schizosaccharomyces pombe*, *Danio rerio*, *Drosophila melanogaster*, *Oryza sativa*, and various archaea, and eubacteria) have revealed novel genes with unknown functions, and transposon mutagenesis provides a powerful method for assigning functions to these genes (e.g., 1,2–5). In addition, restriction enzyme-mediated integration (REMI) has been used widely for mutagenesis (e.g., 6,7, reviewed in 8). For both of these approaches, a unique DNA sequence is inserted by nonhomologous integration into genomic DNA to generate a disruption mutation that is physically tagged. Genetic analysis of the resulting mutants yields those with defects in the function of interest. Because each mutation is physically marked by a unique DNA sequence, it is possible to use this tag to identify each disrupted gene.

The large-scale characterization of these mutations requires a reliable, rapid, and simple method to identify the site of insertion. The polymerase chain reaction (PCR) enables the quick and easy identification and analysis of specific DNA sequences, but in its original form, this process requires knowledge of the DNA sequences at each end of the DNA being amplified (9). For collections of transposon or REMI-generated disruption mutations, the DNA sequence of the inserted DNA is known, but the sequence of the surrounding gene is not. Involving only two successive PCRs and two pairs of PCR primers, semi-random, two-step PCR (ST-PCR) provides a simple, rapid method for identifying these disrupted genes.

ST-PCR for disruption mutations in *S. cerevisiae* consists of two successive PCRs utilizing two different pairs of PCR primers. When used to amplify disruptions from *Coprinus cinereus* (whose genome is roughly three times larger than that of *S. cerevisiae*), a third PCR is sometimes used. For the first reaction, the template is

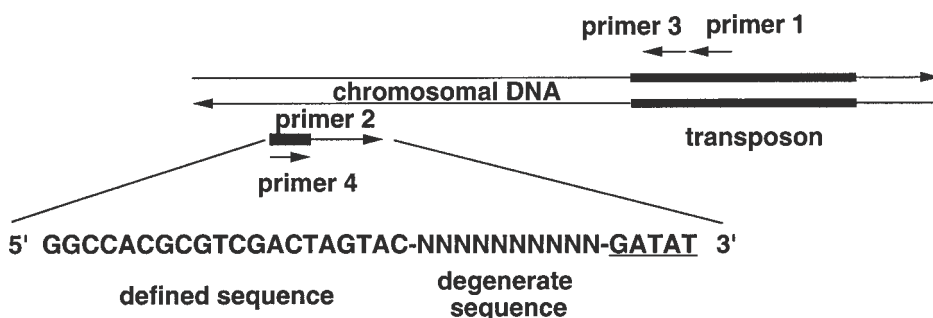


Fig. 1. Summary of ST-PCR scheme. (Reproduced from ref. 10 with permission)

genomic DNA from the mutant of interest, and primers 1 and 2 direct amplification (see Fig. 1). Primer 1 anneals specifically near one known end of the inserted DNA (e.g., transposon or REMI DNA element) and directs DNA polymerization away from that end and into the disrupted genomic DNA. Primer 2 contains a specific 20-nucleotide sequence followed by 10 bases of degenerate sequence and a specific five-nucleotide sequence (see Fig. 1). The 20 defined bases provide the sequence to which a primer in the second PCR will anneal. For the 10 random bases, during the synthesis of primer 2, a mixture of all four nucleosides is incorporated at each of the degenerate base positions, so in this region, the resulting collection of primers should include every possible permutation. The specific five-nucleotide sequence at the 3' end is predicted to occur every 500 to 1500 bp of the genome being analyzed. In the *S. cerevisiae* genome, which consists of approx 40% G+C, the sequence GATAT should occur approx every 600 bp. Assuming the occurrence of this sequence follows a Poisson distribution, there is an 80% chance that at least one "GATAT" occurs within 1 kb of any position in the genome and a 90% chance that one occurs within 1.5 kb (see Fig. 2). Therefore, at least one "GATAT" should be found near each transposon insertion. Consequently, at least one of the primer 2's is likely to anneal to a GATAT site near the insertion and, along with primer 1, promote the amplification of the DNA sequence in and next to the mutation.

The annealing temperatures for the first six rounds of the first amplification are relatively low to allow primer 2's with low melting temperatures as well as those with some mismatches to anneal (see Table 1). If the 15 3'-most bases of primer 2 contribute the most to its melting temperature (T_m), a primer with 10 A's and/or T's in the degenerate region will have a predicted T_m of about 32°C. Alternatively, a primer with 10 G's and/or C's will have a T_m of about 52°C. To favor primers that anneal under relatively stringent conditions, the first annealing temperature is 42°C, and for each of the next five PCR cycles, the annealing temperature is one degree cooler (41°C and so on).

For the second PCR, PCR products from the first reaction serve as the template, and a second pair of primers direct amplification. Primer 3 anneals to the inserted DNA element, nested relative to primer 1, and primer 4 anneals to the complement of

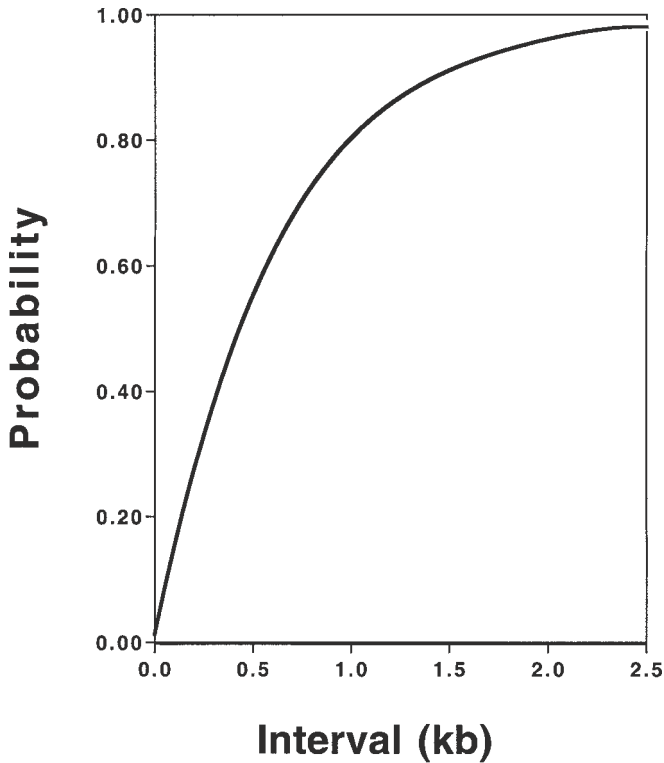


Fig. 2. Probability that the sequence “GATAT” will occur within a given interval of *S. cerevisiae* genomic DNA. This graph assumes that the occurrence of “GATAT” occurs randomly and conforms to a Poisson distribution and that the *S. cerevisiae* genome consists of 40% G+C basepairs.

Table 1
DNA Primers Used in ST-PCR

Primer	DNA sequence (5' to 3')
1	TAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
2/OL-106	GGCCACGCGTCGACTAGTAC (N) ₁₀ GATAT
3	(CAU) ₄ TGATAATCTCATGACCAAAATCCC
4/OL-107	(CUA) ₄ GGCCACGCGTCGACTAGTAC
OL-103	(CAU) ₄ GTCGCGGTGAGTTCAGGCTT
OL-104	ATCCACGCCCTCCTACATCG
OL-105	(CAU) ₄ CAGCTTCATTTTCCGTGTGG
OL-108	(CUA) ₄ TAGAATACTCAAGCTATGCA
OL-109	CCGGTGCGCAGCTGATCATG(N) ₁₀ CTATA
OL-110	(CAU) ₄ CCGGTGCGCAGCTGATCATG

the 20 bases of defined sequence at the 5' end of primer 2 (see **Fig. 1**). These primers are intended to amplify the PCR products from the first reaction that were specifically generated from primers 1 and 2, and not to amplify PCR products generated from pairs of primer 2's. The resulting PCR product can be used as the template for direct DNA sequence analysis to determine the sequence of the disrupted genomic DNA. Using this method, we rapidly and easily identified the essential genes in a collection of mini-Tn3 transposon mutated *S. cerevisiae* mutants (**2,10,11**).

With a few modifications, we also used this method to identify the mutated genes in a collection of *C. cinereus* sporulation-defective mutants (**7**). These mutations were generated by REMI, whereby *C. cinereus* was cotransformed with a restriction endonuclease and a linearized plasmid, pPHT1 (**7** and **Fig. 3**). This plasmid contains PHT1, a dominant selectable marker which consists of the open reading frame of the *E. coli* hygromycin B phosphotransferase gene (*hph*, **12**) fused to the promoter and terminator regions of the *C. cinereus* β -tubulin gene. This cotransformation results in the reasonably random, nonhomologous integration of the selectable marker and the generation of a disruption mutation at the insertion site (**6,7**, reviewed in **8**). We determined that during REMI mutagenesis of *C. cinereus*, roughly 50% of the insertion events are nonperfect integrations (**7**). These are likely to result from degradation of the ends of the transforming DNA and processing of the restriction enzyme-generated double-strand breaks in the target DNA prior to insertion, as well as integration of the transforming DNA independent of the restriction enzyme cleavage site. If substantial truncation of the transforming DNA occurs, terminal-proximal priming sites may be lost. Therefore, unlike ST-PCR in *S. cerevisiae* for transposon-disruption mutants for which only two pairs of primers were used, several different combinations of primers were designed to amplify REMI-generated disruptions for *C. cinereus*. By using this approach, we were able to analyze mutants in which the pPHT1 was truncated during integration. Three different primer pairs anneal to pPHT1, either near its *hph* end or to a proximal region (e.g., to the β -tubulin promoter or to *hph*). In addition, a fourth pair of primers was designed to anneal to the opposite end of pPHT1.

Numerous other modifications of PCR that allow the amplification of unknown nucleotide sequences next to known sequences have been described. The concept of using a known cDNA sequence and its 3' polyA tail to amplify the intervening sequence was introduced with RACE (**13**), one-sided PCR (**14**), and anchored PCR (**15**). When seeking DNA sequences that are not at the 3' end of a cDNA (and lack the polyA tail), these procedures all require terminal deoxynucleotidyl-transferase to attach to the end of the unknown sequence a polynucleotide tail. Inverse PCR (**16**), inverted PCR (**17**), ligation-mediated single-sided PCR (**18**), single-specific primer PCR (SSP-PCR) (**19**), and panhandle PCR (**20**), all require the generation of a DNA fragment containing the known DNA sequence as well as the adjacent unknown sequences of interest (for example, restriction enzyme digestion). In addition, these five methods require the ligation of a known DNA sequence to the unknown end. Rapid amplification of transposon ends (RATE), ligation-mediated PCR (LMPCR), and vectorette-mediated PCR also require a ligation step (**21–23**). Transposon inser-

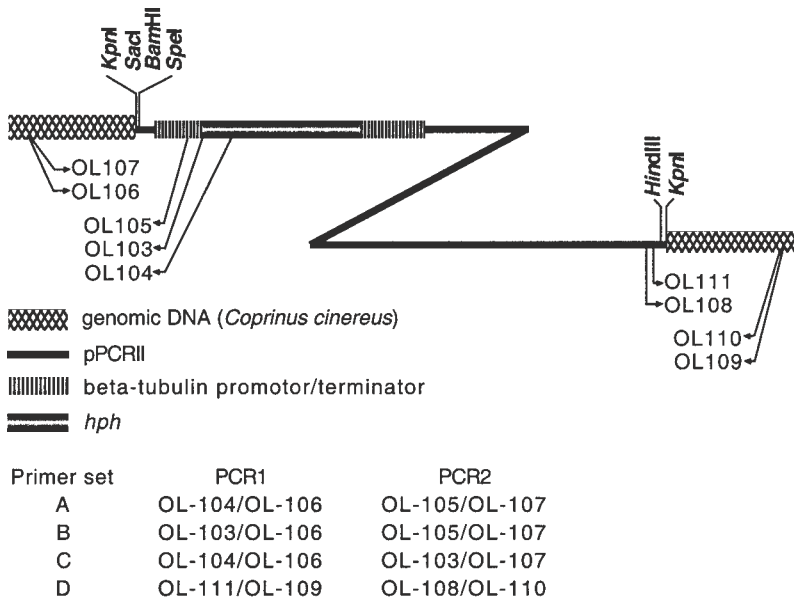


Fig. 3. ST-PCR strategy used for cloning the genomic DNA sequences adjacent to the pPHT1 insert in *C. cinereus* mutants.

tion display (TID) requires both restriction endonuclease digestion and ligation (24). In these types of modified PCR, one primer anneals to the DNA of known sequence, a second anneals to the sequences added to the other end, and the intervening, unknown sequences are amplified.

Several other PCR protocols (e.g., targeted gene walking, 25) employ a primer that anneals specifically at or near the end of a region of known sequence and a collection of primers that anneal imperfectly to a nearby, unknown sequence. Nonstringent annealing conditions allow the second primers to anneal. A similar method, PCR mediated by a single primer (26), uses only one primer and relies on conditions where it anneals specifically to its complementary sequence as well as imperfectly to a nearby sequence. These methods require none of the enzymatic steps described earlier, but they often yield multiple PCR products, because in addition to the expected DNA fragment, they yield products generated from two primers annealing imperfectly to genomic target sequences. The multiple PCR products generated must be screened by Southern analysis to identify the desired fragment.

Finally, thermal asymmetric interlaced (TAIL-) PCR has been used to amplify the genomic sequences flanking insertions in *A. thaliana* (27) and *C. cinereus* (P. Medina and P. Pukkila, personal communication). Like ST-PCR, it does not require additional steps before PCR, and it does not require screening afterward. This method involves three successive PCRs.

ST-PCR is a simple, efficient, and rapid method to identify novel DNA sequences next to previously known sequences. In addition to analyzing single sequences of interest, it should also be applicable to large-scale analyses. Although we have used this method to analyze mutations in fungi, it should be possible to modify it for examining more complex genomes.

2. Materials

1. Thermal cycler (PCR) machine. Ideally, it should have a program function that allows a progressive decrease of temperature each time a step in a cycle subroutine is executed.
2. *Taq* polymerase. We successfully used enzyme from Roche (Indianapolis, IN) for ST-PCR from *S. cerevisiae* and from Fisher Scientific (Pittsburgh, PA) for amplification from *C. cinereus*. It is likely that enzyme from other sources will also be suitable.
3. PCR buffer. 10X stock (500 mM KCl, 100 mM Tris-HCl pH 8.3, 2 mg/mL gelatin, 30 mM MgCl₂, and 2 mM of each dNTP) was used for *Taq* from Roche. The 10X buffer supplied by Fisher Scientific was used for *Taq* from Fisher Scientific, and this reaction was supplemented with MgCl₂ to a final concentration of 4.5 mM. All reagents should be molecular biology grade. Solutions should be made with double-distilled or ultrapure water and should be autoclave sterilized (except the TAE or TBE buffer for agarose gel electrophoresis, which does not need to be sterile).
4. PCR primers. The sequences for the PCR primers we use are listed in **Table 1** and shown in **Figs. 1** and **3**. One set of primers was designed to amplify mini-Tn3 transposon disrupted sequences in *S. cerevisiae* and another set was for amplifying REMI-generated pPHT1-disrupted sequences in *C. cinereus*. Therefore, for other applications, four or more primers should be designed to suit the characteristics of the inserted DNA and the genome being studied.

Primer 1 should anneal near one end of the inserted DNA and direct DNA polymerization through that end and into the disrupted genomic DNA.

Primer 2 consists of 20 bases of defined sequence, followed by 10 random bases, followed by five defined bases. The 20 defined bases provide the sequence to which a primer in the second PCR will anneal. For *S. cerevisiae*, we chose the 20 base sequence used in the 5' RACE primers sold by Life Technologies, Gibco-BRL, but any other unique sequence should work well. For the ten random bases, during the synthesis of primer 2, a mixture of all four nucleosides is incorporated at each of the degenerate base positions. Alternatively, nucleosides containing 2'-deoxyinosine or a universal base, such as 3-nitropyrrole or 5-nitroindole, could be used as degenerate bases (28–30). The 3'-most five base sequence should correspond to a sequence expected to occur in the genome close enough to the insertion site to allow efficient amplification of the intervening sequence. Consequently, assuming that the 15 3'-most bases of primer 2 are the most important for annealing to the template, at least one of these primers should be likely to anneal to the complement of the five base sequence occurring near the insertion site and, along with primer 1, promote the amplification of the DNA sequence in and next to the mutation.

Primer 3 anneals to the inserted DNA element (in our case mini-Tn3 or pPHT1), nested with respect to primer 1, and primer 4 anneals to the complement of the 20 bases of defined sequence at the 5' end of primer 2. These primers are intended to amplify the PCR products from the first reaction that were specifically generated from primers 1 and 2.

5. To analyze PCR products, standard agarose gel electrophoresis equipment and molecular biology grade agarose is used. To purify PCR products after agarose gel electrophoresis, standard methods [e.g. GeneClean (Bio 101) or QiaexII (Qiagen) kits] work well.

6. 50X TAE or 10X TBE is used as the stock for agarose gels and running buffer. 50X TAE: 2.0 M Tris-acetate, 50 mM EDTA, adjust pH to 7.9 with glacial acetic acid. 10X TBE: 0.9 M Tris-borate, and 20 mM EDTA adjusted pH to 8.0 with NaOH.
7. To determine the DNA sequence of the amplified DNA, the PCR product may be sequenced directly using a thermal cycle sequencing protocol. Alternatively, if primer 3 is designed to contain (CAU)₄ at its 5' end and primer 4 contains (CUA)₄ at its 5' end (in addition to the sequences for these primers described earlier), the Clone-Amp pAMP1 System for Rapid Cloning of Amplification Products (Life Technologies, Gibco-BRL) may be used to subclone the PCR fragment directionally into the pAMP1 plasmid, and the resulting recombinant can then be sequenced using standard methods. This DNA sequence can then be compared to the complete sequence of the *S. cerevisiae* genome to identify the complete sequence of the disrupted gene.
8. TE buffer: 10 mM Tris-HCl pH 7.4–8.0, 1.0 mM EDTA.

3. Method

3.1. ST-PCR from Mini-Tn3 Disruption Mutations in *S. cerevisiae*

1. Purify genomic DNA containing the disruption mutation using standard methods. Any method that yields DNA of sufficient purity for restriction endonuclease digestion or genomic blot analysis should work well. Resuspend the DNA in TE buffer or ultrapure water.
2. For the first PCR (PCR1, 20 μ L total volume) mix approximately 100 ng of template DNA, 1X PCR buffer, 20 pmol each of PCR primers 1 and 2, and 2.5 units of *Taq* polymerase. Reagents and each reaction should be stored on ice during preparation. Carry out this first amplification reaction using the following program:
 - a. 94°C, 2 min
 - b. 94°C, 30 s
 - c. Initial temp. 42°C, 30 s. Decrease 1.0°C for each subsequent cycle
 - d. 72°C, 3 min
 - e. Return to **step b** and repeat five more times
 - f. 94°C, 30 s
 - g. 65°C, 30 s
 - h. 72°C, 3 min
 - i. return to **step f** and repeat this subroutine 24 more times
 - j. 4°C, hold
 - k. End
3. Dilute the resulting PCR with 80 μ L of sterile water, and use 1 μ L of this as the template for the second PCR (PCR2), which is essentially the same as the first, except primers 3 and 4 and the following PCR program are used:
 - a. 94°C, 30 s
 - b. 65°C, 30 s
 - c. 72°C, 3 min
 - d. return to step a and repeat this subroutine 29 more times
 - e. 4°C, hold
 - f. End
4. Load 5 μ L of the 20 μ L reaction onto a 1.0% agarose gel in 0.5X TAE or 1.0X TBE and separate the PCR products using a standard electrophoresis gel box and power supply. 0.5 μ g/mL ethidium bromide may be included in the gel or the gel may be stained after electrophoresis.

5. Excise the PCR fragment from the gel using a razor blade or scalpel and purify the DNA from the agarose using standard methods (e.g., GeneClean (Bio101) or QiaexII (Qiagen) kit) into 20 μ L of TE buffer.
6. The DNA sequence of the purified fragment can then be determined directly using a thermal cycle sequencing protocol, 5 μ L of the 20 μ L of PCR2 fragment, and primer 3 as the sequencing primer. Alternatively, one can subclone the PCR2 fragment into a plasmid vector. One approach to accomplish this is to use the Clone-Amp pAMP1 System for Rapid Cloning of Amplification Products (Gibco-BRL). To do so, one would design primer 3 with (CAU)₄ at its 5' end and primer 4 with (CUA)₄ at its 5' end. PCR2 fragments amplified with these primers will contain four uridines at the 5' end of each strand. Treatment of this DNA with uracil DNA glycosylase removes uracil, disrupts base pairing, and exposes the complementary single-strand 3' ends. These ends can then anneal to complementary single-strand pAMP1 vector ends. Ligation is not required, so the annealed fragments can be transformed directly into competent *E. coli*. The resulting recombinant can then be sequenced using standard methods (see **Note 1**).

3.2. ST-PCR from REMI-Generated Disruption Mutations in *C. cinereus*

The ST-PCR protocol can be modified to amplify disrupted genomic sequences from genomes more complex than that of *S. cerevisiae*. Here, we describe modifications used to analyze REMI-generated mutations in *C. cinereus*, whose genome is 37.5 Mb, roughly three times larger than that of *S. cerevisiae*.

1. Just as for *S. cerevisiae*, DNA from *C. cinereus* REMI-generated mutants that is sufficient for genomic blot analysis should work well for ST-PCR.
2. The first PCR is the same as that described in **Subheading 3.1., step 2**, except for the following modifications.
 - a. A set of primers (A, B, C, or D) is used instead of *S. cerevisiae* primers 1 and 2 (see **Table 1** and **Fig. 3**). Primer sets A, B, and C were designed to amplify from the β -tubulin promoter end of pPHT1 ("head" end), and primer set D was designed to amplify from the other end ("tail" end) of pPHT1. Of course, additional primer sets could be designed (see **Note 2**).
3. PCR1 is diluted as described in **Subheading 3.1., step 3** for use as the template in the PCR2. The same PCR2 conditions are used as those described in **Subheading 3.1., step 3**, except for the following modifications.
 - a. 5% (w/v) acetamide is added to the reaction (see **Note 3**).
 - b. Only 20 cycles of amplification are used (see **Note 3**).
 - c. Depending upon which primer pair was used in PCR1, the corresponding pair for PCR2 is used in this subsequent PCR. (see **Table 1**, e.g., if OL-104 and OL-106 are used in PCR1, OL-105 and OL-107 are used in PCR2)
4. PCR2 fragments are resolved by agarose gel electrophoresis (see **Note 4**) and purified as described in **Subheading 3.1., steps 4 and 5**.
5. Several of the *C. cinereus* PCR2 amplifications generated only a small quantity of product. To increase the likelihood of successfully cloning these products, a third cycle of amplification can be used. In this third PCR, 1 μ L of the 20- μ L eluate from the excised PCR2 fragment is used as the template, and otherwise the reaction mix and conditions are as described for the PCR2 (see **step 3** above). The products of the amplification are separated by agarose gel electrophoresis, and the DNA fragments are excised, cloned into pAMP1 (as described above for *S. cerevisiae*), and sequenced (see **Notes 5 and 6**).

6. For PCR2 amplifications involving primer sets A, B, or C (see **Fig. 3**), the resulting products include a portion of the *C. cinereus* β -tubulin promoter from pPHT1. Each product from these amplifications should be screened by Southern analysis, and only those with this expected portion of pPHT1 should be analyzed further (see **Note 5**).
7. Depending upon the primer set used, each of the ST-PCR products should contain a predicted portion of amplified pPHT1. For ST-PCR amplifications involving primer sets A or B (see **Fig. 3**), the product should contain 181 bp of pPHT1. Ideally, the size of the genomic DNA amplified should be greater than 0.3 kb, but 100 bp can be sufficient and therefore any ST-PCR products greater than 0.3 kb are pursued. For similar reasons, only products greater than 0.6 kb and 0.2 kb from ST-PCR amplifications involving primer sets C or D, respectively, are pursued (see **Fig. 3, Note 4**).
8. The DNA fragment corresponding to the region of disrupted genomic DNA can be used to probe a blot of separated chromosomes to map the mutated gene to a chromosome. This probe can subsequently be used to probe either a corresponding chromosome-specific or a genomic cosmid library to isolate a complete genomic clone.

4. Notes

1. In theory, PCR2 might generate, in addition to a single PCR product amplified with primers 3 and 4, one or more PCR products amplified with a pair of primer 4's (that had annealed to a product from the first PCR generated from two primer 2's). Surprisingly, we did not observe such PCR products. It is possible that this is not a common occurrence and would only be observed after analysis of many more mutants. However, if the products from the second PCR are cloned into a vector in a way that requires that the ends of the insert contain primers 3 and 4, these extraneous PCR products are eliminated. In our case, specific DNA sequences were incorporated into the 5' ends of primers 3 and 4, and these were used to insert the PCR product asymmetrically into a compatible plasmid (see **Subheadings 2.7. and 3.1.6.**).
2. It is useful to know the number of PHT1 integration sites in the genome of the REMI transformant being analyzed. We determined this by probing with *hph* Southern blots of restriction-digested genomic DNA and intact, resolved chromosomes. REMI transformants with multiple sites of insertion were backcrossed to obtain isolates that contain a single site of integration. Also, depending on the relative orientation of multiple inserts at a single locus, some primer pairs were not useful. For example, in a tail-to-tail orientation, primer combination D (see **Fig. 3**) would not be expected to amplify flanking genomic DNA, whereas primer combinations A, B, or C might. Therefore, to predict which primer pairs should be used, the number of pPHT1 constructs per site of integration, as well as their relative orientation should be determined. We used restriction digests and Southern blot analysis to determine the orientation of inserts within a site (e.g., head-to-tail, head-to-head, tail-to-tail).
3. We found that including 5% acetamide in the PCR2, as well as decreasing the number of amplification cycles to only 20 (instead of the 30 used for *S. cerevisiae*) decreased the amount of smearing when these products were analyzed on agarose gels. This, in turn, allowed us to see less abundant PCR products, which could be excised, purified, and used in a third PCR. We also tried using less of the first PCR as the template for the second PCR (instead of only 1:5, diluting the first PCR 1:10, 1:100, or 1:1000, before taking 1 μ L for the second PCR). This modification decreased smearing, but unfortunately it also decreased product yield.
4. For *C. cinereus* REMI-generated mutants, the ST-PCR products obtained ranged in size from 0.3 to 1.2 kb.

5. Some of the products obtained in amplifications using primer sets A, B, or C did not contain the expected portion of the β -tubulin promoter from pPHT1. The specific cause of these PCR artifacts is unknown, but these were observed infrequently.
6. We found that in some of the PCR2 and PCR3 amplifications multiple products were generated. By cloning and sequencing these products, we determined that many were simply shorter or longer amplifications of the same genomic DNA. Additionally, some of the products turned out to be amplifications of the native β -tubulin gene and were identified as such during the sequencing step.

References

1. Cooley, L., Kelley, R., and Spradling, A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* **239**, 1121–1128.
2. Chun, K. T. and Goebel, M. G. (1996) The identification of transposon-tagged mutations in essential genes that affect cell morphology in *Saccharomyces cerevisiae*. *Genetics* **142**, 39–50.
3. Burns, N., Grimwade, B., Ross-Macdonald, P. B., Choi, E.-Y., Finberg, K., Roeder, G. S., and Snyder, M. (1994) Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* **8**, 187–1105.
4. Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J. D., Dean, C., Ma, H., and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **9**, 1797–1810.
5. Parinov, S., Sevugan, M., Ye, D., Yang, W.-C., Kumaran, M., and Sundaresan, V. (1999) Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in *Arabidopsis*. *The Plant Cell* **11**, 2263–2270.
6. Schiestl, R. H. and Petes, T. D. (1991) Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **88**, 7585–7589.
7. Cummings, W. J., Celerin, M., Crodian, J., Brunick, L. K., and Zolan, M. E. (1999) Insertional mutagenesis in *Coprinus cinereus*: use of a dominant selectable marker to generate tagged, sporulation-defective mutants. *Curr. Genet.* **36**, 371–382.
8. Riggle, P. J. and Kumamoto, C. A. (1998) Genetic analysis in fungi using restriction enzyme-mediated integration. *Curr. Opin. Microbiol.* **1**, 395–399.
9. Mullis, K. B., Faloona, F. A., Scharf, S. J., Sakai, R. K., Horn, G. T., and Erlich, H. A. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 263–273.
10. Chun, K. T., Edenberg, H. J., Kelley, M. R., and Goebel, M. G. (1997) Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. *Yeast* **13**, 233–240.
11. Chun, K. T. and Goebel, M. G. (1997) Mutational analysis of Cak1p, an essential protein kinase that regulates cell cycle progression. *Mol. Gen. Genet.* **256**, 365–375.
12. Gritz, L. and Davies, J. (1983) Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. *Gene* **25**, 179–188.
13. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
14. Ohara, O., Dorit, R. L., and Gilbert, W. (1989) One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci. USA* **86**, 5673–5677.

15. Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989) Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science* **243**, 217–220.
16. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–623.
17. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res.* **16**, 8186.
18. Mueller, P. R. and Wold, B. (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246**, 780–786.
19. Shyamala, V. and Ames, G. F. (1989) Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene* **84**, 1–8.
20. Jones, D. H. and Winistorfer, S. C. (1992) Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. *Nucl. Acids Res.* **20**, 595–600.
21. Ribot, E. M. (1998) Rapid amplification of transposon ends for the isolation, cloning and sequencing of transposon-disrupted chromosomal genes. *Biotechniques* **24**, 16–22.
22. Prod'hom, G., Lagier, B., Pelicic, V., Hance, A. J., Gicquel, B., and Guilhot, C. (1998) A reliable amplification technique for the characterization of genomic DNA sequences flanking insertion sequences. *FEMS Microbiol. Let.* **158**, 75–81.
23. Eggert, H., Bergemann, K., and Saumweber, H. (1998) Molecular screening of P-element insertions in a large genomic region of *Drosophila melanogaster* using polymerase chain reaction mediated by the Vectorsette. *Genetics* **149**, 1427–1434.
24. Yephremov, A. and Saedler, H. (2000) Display and isolation of transposon-flanking sequences starting from genomic DNA and RNA. *Plant J.* **21**, 495–505.
25. Parker, J. D., Rabinovitch, P. S., and Burmer, G. C. (1991) Targeted gene walking polymerase chain reaction. *Nucl. Acids Res.* **19**, 3055–3060.
26. Parks, C. L., Chang, L. S., and Shenk, T. (1991) A polymerase chain reaction mediated by a single primer: cloning of genomic sequences adjacent to a serotonin receptor protein coding region. *Nucl. Acids Res.* **19**, 7155–7160.
27. Liu, Y.-G., Mitsukawa, M., Oosumi, T., and Whittier, R. F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**, 457–463.
28. Knoth, K., Roberds, S., Poteet, C., and Tamkun, M. (1988) Highly degenerate, inosine-containing primers specifically amplify rare cDNA using the polymerase chain reaction. *Nucl. Acids Res.* **16**, 10,932.
29. Nichols, R., Andrews, P. C., Zhang, P., and Bergstrom, D. E. (1994) A universal nucleoside for use at ambiguous sites in DNA primers. *Nature* **369**, 492–493.
30. Loakes, D., Brown, D. M., Linde, S., and Hill, F. (1995) 3-Nitropyrrole and 5-nitroindole as universal bases in primers for DNA sequencing and PCR. *Nucl. Acids Res.* **23**, 2361–2366.

Isolation of Large-Terminal Sequences of BAC Inserts Based on Double-Restriction-Enzyme Digestion Followed by Anchored PCR

Zhong-Nan Yang and T. Erik Mirkov

1. Introduction

Large insert libraries are critical in genome-related research. Bacterial artificial chromosome (BAC) libraries are widely used in plant, animal, and human research; the ends of BAC clones are used as probes for chromosome walking and to confirm overlapping of contigs, as well as RFLP markers for mapping. Several methods have been developed to isolate BAC ends, including subcloning methods, plasmid rescue, and polymerase chain reaction (PCR)-based methods such as inverse PCR, thermal asymmetric interlaced PCR (TAIL-PCR), and Vectorette PCR (1, and references therein). Here, we described BAC end isolation based on double-restriction-enzyme digestion followed by anchored PCR. The BAC vector in the examples used here is pBeloBAC11 (2). Other vectors can be used, but the specific designation of left or right ends as used here may be different depending on the particular vector used.

The strategy to isolate BAC ends is outlined in **Fig. 1**. The first step is the double-restriction-enzyme digestion of the BAC DNA. The two enzymes used are *NotI* and any one of the following four enzymes: *EcoRV*, *HpaI*, *StuI*, and *XmnI*. There are only two *NotI* sites in the BAC vector flanking the insert, and *NotI* would be expected to cut infrequently in the inserts. For the second enzyme of the double digestion, any enzyme with a 6-basepair recognition sequence that is not present in the left or right arms of the BAC vector can be used. Here the four enzymes *EcoRV*, *HpaI*, *StuI*, and *XmnI* are chosen as they leave blunt termini after digestion. *EcoRV* is used as representative of the second enzyme in an example of the isolation strategy in **Fig. 1**. After double digestion, the DNA fragments containing the left arm (T7 side) with the left end of the BAC insert and the right arm (SP6 side) with the right end of the BAC insert will have one blunt termini and a *NotI* termini (L and R in **Fig. 1**). The second step is dephosphorylation of digested BAC DNA, which prevents the digested BAC DNA

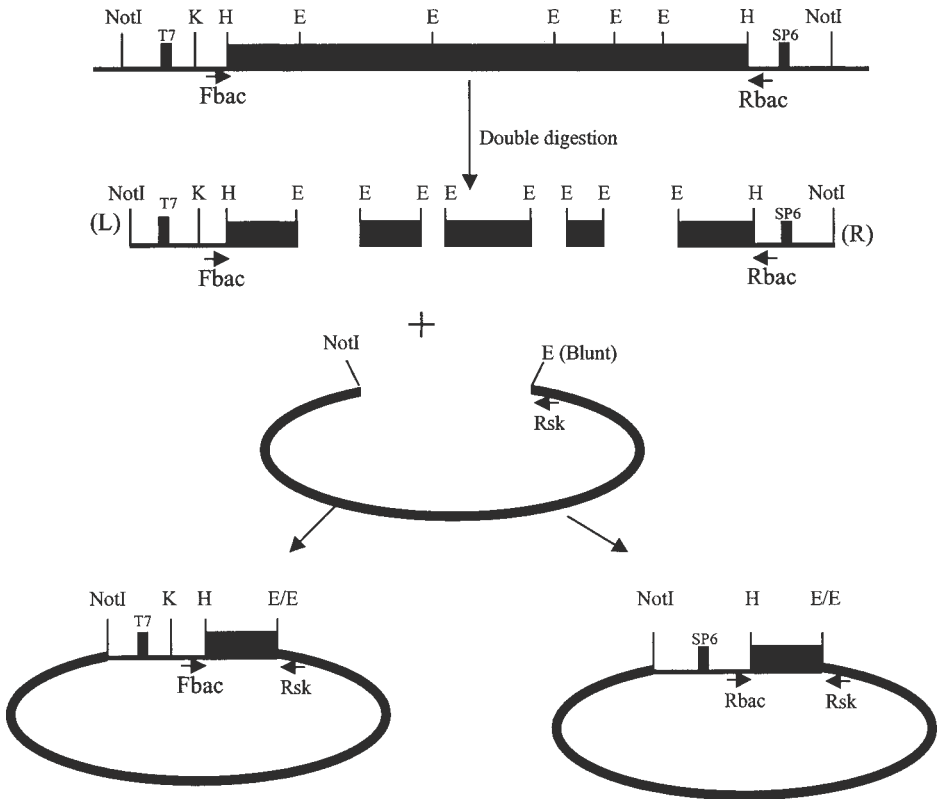


Fig. 1. Schematic diagram of the strategy used for specific amplification of terminal sequences of BAC clones. BAC ends containing the left arm (L) or right arm (R) of the BAC vector will be released after double-restriction-enzyme digestion (*NotI* and *EcoRV*). These DNA fragments are ligated to the *NotI* and *EcoRV* sites of pMSK. PCR products representing the left and right ends can be amplified specifically with primers Rsk and Fbac (left end) or Rsk and Rbac (right end). K, *KpnI*; H, *HindIII*; E, *EcoRV*; Fbac, left arm specific primer; Rbac, right arm specific primer; Rsk, pMSK specific primer adjacent to the *EcoRV* site. T7, T7 promoter; SP6, SP6 promoter.

from being religated. The third step is the ligation of digested BAC DNA with the *NotI* and *EcoRV*-digested anchor vector pMSK. The fourth step is PCR amplification with one primer annealing to pMSK and another primer annealing to the left arm (Fbac) or right arm (Rbac) to specifically amplify the left end or right end correspondingly.

With the optimized conditions described in this protocol, the PCR product is a single strong band that can be used directly to generate probes for hybridization. In most cases, PCR products that represent BAC ends can be obtained from two or three of the four double digestions and ligations. This allows reconfirmation of isolated BAC ends simply by restriction enzyme digestion (1). For each BAC clone, four double-restric-

tion enzyme digestions were used. The actual probability to obtain PCR products that represent both the left and right ends would be expected to be greater than 90%. This method is simple, fast and reliable and it has become a routine method for BAC end isolation in our laboratory.

2. Materials

1. BAC DNA prepared from a 5-mL overnight culture and resuspended in 35 μL H_2O (3).
2. Plasmid pBluescript II SK(+) (Stratagene, LA Jolla, CA) that contains the restriction sites *NotI* and *EcoRV*.
3. GENECLEAN II Kit from Bio 101, Inc., (Carlsbad, CA).
4. Restriction enzymes (see **Note 1**): *NotI*, *EcoRV*, *HpaI*, *KpnI*, *PvuII*, *SacI*, *StuI*, *XmnI* (20 U/ μL) and their appropriate reaction buffer from New England Biolabs (Beverly, MA).
5. Shrimp alkaline phosphatase (Amersham Pharmacia Biotech, Piscataway, NJ).
6. T4 DNA ligase and its reaction buffer (Life Technologies, Rockville, MD)
7. Primers: Rsk (5'-TCGAGGTCGACGGTATCG-3'); Fbac (5'-AGTCGACCTGCAG-GCATG-3'); and Rbac (5'-CGCCAAGCTATTTAGGTGA-3'), each at 1 $\mu\text{g}/\mu\text{L}$.
8. 5 U/ μL *Taq* DNA polymerase and 10X *Taq* reaction buffer.
9. 2 mM deoxynucleotide 5'-triphosphate (dNTP) mix (final concentration of 2 mM of each of the four dNTPs).
10. 1X TAE buffer: 0.04 M Tris-acetate, 0.001 M ethylenediaminetetraacetic acid (EDTA), pH 8.0.
11. Mineral oil.
12. Thin-walled PCR tubes.
13. Automated thermal cycler.

3. Methods

3.1. Preparation of pMSK from pBluescript II SK(+) (see Note 2)

After digestion of pBluescript II SK(+) with *SacI* and *KpnI*, a 101-bp fragment containing the multiple cloning site is released. This fragment is made blunt with T4 DNA polymerase, gel purified, and ligated to the gel purified 2.5-kb *PvuII* fragment of pBluescript II SK(+) containing the *f1(+)* origin, *ColE1* origin, and the ampicillin-resistance gene, to obtain the anchor vector pMSK (1).

3.2. Anchor Vector Preparation

1. Plasmid pMSK is used as an anchor vector to ligate with digested BAC DNA. It is prepared from pBluescript II SK(+) as described above.
2. Prepare a mix by adding the following to a tube on ice:
 - a. 4 μg pMSK DNA;
 - b. 8 μL NEB Reaction Buffer 3;
 - c. 4 μL (10 U/ μL) *NotI*
 - d. 4 μL (20 U/ μL) *EcoRV*
 - e. Add H_2O to a final volume of 80 μL .Incubate at 37°C for 2 h.
3. Electrophorese the digestion in a 1% agarose gel in 1X TAE buffer.
4. Excise the band of 2.6 kb and purify with the GenClean II kit following the manufacturer's instructions. Elute the DNA in a final volume of 80 μL .

3.3. BAC DNA Double Digestion and Dephosphorylation

1. Prepare a master mixture by adding the following to a tube on ice:
 - a. 8 μL 10X NEB Reaction Buffer 4;
 - b. 4 μL Miniprep BAC DNA (see **Note 3**);
 - c. 4 μL *NotI* (10 U/ μL);
 - d. 60 μL H_2O .

Aliquot 19 μL of the mixture into each of four tubes. Add 1 μL of one of the second restriction enzyme (*EcoRV*, *HpaI*, *StuI*, or *XmnI*) to each tube and incubate at 37°C for 1 h.
2. Add 0.5 μL (0.5 U) shrimp alkaline phosphatase to each tube and incubate at 37°C for 30 min.
3. Incubate at 65°C for 20 min to inactivate the shrimp alkaline phosphatase.

3.4. Ligation (PCR Template Preparation)

1. Prepare a master mixture by adding the following to a tube on ice:
 - a. 16 μL *EcoRV* and *NotI* digested pMSK vector;
 - b. 16 μL 5X T4 DNA ligase buffer;
 - c. 24 μL H_2O .

Aliquot 14 μL of the master mixture into each of four tubes.
2. Add 5 μL of one of the dephosphorylated digestions (see **Subheading 3.3.**) to one of the above tubes and then add 1 μL of T4 DNA ligase (1 U/ μL). Mix and incubate at room temperature for 1 h (see **Note 3**). One microliter of each ligation is then diluted to 50 μL .

3.5. PCR Amplification (BAC End Amplification)

1. Prepare two reaction master mixtures. Add the following to a microcentrifuge tube on ice:
 - a. 40 μL 10X *Taq* Reaction buffer;
 - b. 40 μL 2 mM dNTP mix;
 - c. 2 μL primer Rsk;
 - d. 3.2 μL *Taq* polymerase;
 - e. 312.8 μL H_2O .

Aliquot 199 μL of the mixture into each of two microfuge tubes labeled F and R that will be called master mixture F and R, respectively.
2. Add 1 μL of primer Fbac or Rbac to master mixture F and R, respectively. Aliquot 50 μL of the master mixture F into each of four empty tubes.
3. Add 1 μL of one of the four diluted ligations (*EcoRV*, *HpaI*, *StuI* or *XmnI*) to the tubes prepared earlier. The corresponding reactions are labeled EF, HF, SF, and XF. The PCR products from these tubes represent the left end of the BAC inserts.
4. Correspondingly, aliquot 50 μL of the master mixture R into each of four empty tubes and add 1 μL of one of the four diluted ligations. The PCR reactions are labeled ER, HR, SR, and XR. The PCR products from these tubes represent the right end of the BAC inserts.
5. Overlay the reaction mixture with 50 μL mineral oil.
6. Using the following cycles, program the automated thermal cycler according to the manufacturers's instructions (see **Note 4**):

94°C (2 min),	1 cycle
94°C (1 min), 62°C (1 min), 72°C (4 min)	30 cycles
72°C (5 min)	1 cycle.
7. Electrophorese 10 μL from each reaction in an agarose gel in 1X TAE buffer. An example of a typical result is shown in **Fig. 2**.
8. Excise the amplification products from the gel and purify with the GenClean II kit.

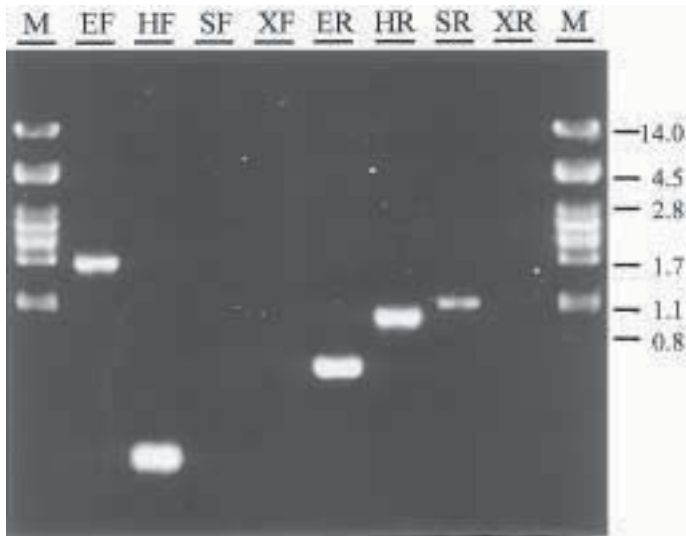


Fig. 2. Agarose gel analysis of PCR products representing the terminal sequences of a BAC clone. EF, HF, SF, and XF represent PCR products obtained using ligations E, H, S, and X with primers Fbac and Rsk. ER, HR, SR, and XR represent PCR products obtained using ligations E, H, S, and X with primers Rbac and Rsk. For BAC clone A, a product representing the left end of the insert (T7 side) was obtained from PCR performed on ligations E and H, and a product representing the right end (SP6 side) was obtained from PCR performed on ligations E, H, and S. M, DNA size marker (λ -Pst I (kb)).

4. Notes

1. In this protocol, the four enzymes *EcoRV*, *HpaI*, *StuI*, or *XmnI* are chosen as the second enzyme of the double digestion, as these enzymes leave blunt termini after digestion. Actually, any enzyme that cuts frequently in the genomic DNA, but not in the left or right arm of the BAC vector can be used. However, the digested DNA should be ligated to the corresponding site of the anchor vector. Use of an enzyme that leaves a blunt terminus allows the user to prepare only one anchor vector that is compatible with all four enzymes. In very few cases, BAC ends cannot be obtained using the four enzymes recommended earlier. Other restriction enzymes that do not cut in the left arm or right arm, and frequently cut in the insert can be chosen as a second enzyme for the double digestion. The digested DNA should be ligated to the corresponding sites of the anchor vector.
2. The plasmid pMSK that is prepared from pBluescript II SK(+) is used as an anchor vector to ligate with double digested BAC DNA. The majority of the *LacZ* gene along with the T7 promoter sequence has been removed as these sequences will cause nonspecific PCR amplification. Actually, any plasmid with *NotI* and *EcoRV* (or any other enzymes leaving blunt termini) sites can be used as an anchor vector as long as the primer sequences in the anchor vector are not in common with the primers used to anneal to the BAC arms.
3. In our lab, we usually prepare miniprep BAC DNA in the morning, and prepare PCR templates from the BAC DNA in the afternoon. The PCR can be run overnight, and BAC ends can be purified and labeled the second day. Usually, 1 μ L of miniprep BAC DNA is

sufficient as a source of template for the PCR amplification. If double-digested BAC DNA is ligated with the vector from 4 h to overnight, a higher quantity of the PCR product is obtained. As the DNA concentration from minipreps can vary, 2–5 μL of the diluted ligation can be used as the PCR template.

4. When using 1 μL of the undiluted ligation as a PCR template, nonspecific amplification is frequently observed after 30 cycles of PCR, although a dominant amplification product can be observed that usually represents the BAC end. When ligations are diluted 1:100, and 1 μL of the diluted ligation is used as the template for PCR, DNA fragments representing the BAC end can also be obtained, except that the quantity of the amplification product is less abundant. In our protocol, we dilute the ligation 1:50. If DNA yield from the minipreps is low, ligations can be diluted 1:10 and 1 μL of the diluted ligation can be used as the PCR template. Although this method was developed for BAC end isolation, it can also be used for YAC and PAC end isolation with some modifications. For PAC clones, the right arm (T7 side) has a *NotI* site (4). Therefore, the right end can be isolated using the same protocol. The left arm of the PAC vector has a *SfiI* site which is an enzyme with 8-bp recognition sequence, and would be expected to cut infrequently in the genomic sequence. Thus, the left end can be isolated with double digestion of *SfiI* and one of the *EcoRV*, *HpaI*, *StuI*, and *XmnI* digestions followed by PCR amplification. Primers should be designed accordingly. The primers specific for left ends or right ends of PAC clone should be designed based on the sequences of the left arm or right arm of the PAC vector. As for YAC clones, YAC2 has *NotI* and *SfiI* sites flanking the inserts (5). Thus, YAC ends can be isolated with this protocol except that *NotI* or *SfiI* is chosen based on the vector sequence.

References

1. Yang, Z. N. and Mirkov, T. E. (2000) Isolation of large terminal sequences of BAC inserts based on double-restriction-enzyme digestion followed by anchored PCR. *Genome* **43**, 412–415.
2. Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. USA* **89**, 8794–8797.
3. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
4. Pierce J. C., Sauer, B., and Sternberg, N. (1992) A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: improved cloning efficiency. *Proc. Natl. Acad. Sci. USA* **89**, 2056–2060.
5. Burke, D. T., Carle, G. F., and Olson, M. V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806–811.

A “Step Down” PCR-Based Technique for Walking Into and the Subsequent Direct-Sequence Analysis of Flanking Genomic DNA

Ziguo Zhang and Sarah Jane Gurr

1. Introduction

The hunt for missing sequence data, whether it be in pursuit of full-length clones or for promoter sequences, can be laborious and expensive. Indeed, protracted efforts to find missing sequences by library screening can be fraught with the frustration of foraging through libraries that may lack the relevant clones. Such problems have led several researchers to describe alternative methods to clone and analyze DNA adjacent to known sequences.

Several polymerase chain reaction (PCR)-based methodologies are available for walking from a known region into cloned or uncloned genomic DNA, including inverse PCR (1,2), randomly primed PCR (3) and adaptor ligation (4–7). An improvement on the adaptor ligation method, combining “vectorette PCR” (6) with “suppression PCR” (9) was made (8) and, furthermore, an adaptor-ligation PCR protocol, using blocked adaptors and exonuclease digestion to remove unligated products, was described (10). More protocols followed, with a technique for the rapid acquisition of unknown DNA sequences by multiplex restriction site PCR (11), a method for the rapid amplification of genomic ends (12), and a technique to amplify and clone genomic DNA without restriction digestion (13).

However, some of these PCR-based methods have proved to be complicated, inefficient, or to demand genomic DNA of 50 kb in length as the starting material. In certain instances, not only can the collection of adequate amounts of biological tissues be troublesome, but so too can the preparation of high-molecular weight DNA. Indeed, this is our experience, as we work with a plant pathogenic fungus (barley powdery mildew) that is a true obligate parasite and that cannot, therefore, be grown axenically.

We describe a novel and efficient PCR-based technique for walking into unknown flanking genomic DNA adjacent to a known sequence such as cDNA, without recourse

to protracted laborious library screening for overlapping sequences. This two component "hot start" and "step down" PCR method uses 6 X 1 µg of genomic DNA (about 20 kb in length) restricted with six different endonucleases and ligated to adaptors, with the inclusion of two further restriction enzymes to prevent self-ligation (**14**). It allows us to walk, in a single experiment, up to 6 kb into flanking DNA and gives us sufficient PCR products for up to 200 different walking experiments. The protocol is summarized diagrammatically in **Fig. 1**.

2. Materials

1. Genomic DNA: Many different methods describe the preparation of intact high-molecular weight DNA. However, genomic DNA of quality adequate for Southern blot analysis is sufficient for this protocol. We routinely use DNA extracted from barley powdery mildew spores, prepared essentially as described for plant DNA extraction (**15**) using hexadecyltrimethylammonium bromide (CTAB).
2. Restriction endonuclease enzymes: *SpeI*, *AvrII*, *NheI*, *XbaI*, *AgeI*, *XmaI*, *NgoMIV*, *BspEI*, and their respective buffers as supplied by New England BioLabs (NEB), Herts, UK. (see **Notes 1** and **2**).

Group I restriction enzyme set the residual "ends":

<i>SpeI</i> :	A/CTAGT	—N
* <i>AvrII</i> :	C/CTAGG	
<i>NheI</i> :	G/CTAGC	—NGATC
<i>XbaI</i> :	T/CTAGA	

Group II

* <i>AgeI</i> :	A/CCGGT	—N
<i>XmaI</i> :	C/CCGGG	
<i>NgoMIV</i> :	G/CCGGC	—NGGCC
<i>BspEI</i> :	T/CCGGA	

3. T4 DNA ligase by NEB.
4. Primers and Adaptors, synthesized in MWG Biotech, AG Ebersberg, Germany.
 - a. PP1: GTAATACGACTCACTATAGGGC
 - b. PP2: ACTATAGGGCACGCGTGGT
 - c. PPR1: PO₄-CTAGGGCCACCACG-NH₂
 - d. PPR2: PO₄-CCGGTGCCACCACG-NH₂ (see **Note 3**)
 - e. Pad1: GTAATACGACTCACTATAGGGCACGCGTGG TGGCC
 - f. Pad2: GTAATACGACTCACTATAGGGCACGCGTGGTGGCA

The relationship between the synthesized primers and adaptors is:

<u>5'-GTAAT ACGAC TCACT ATAGG GCACG CGTGG TGGCC</u>	
<u>*PP1</u>	<u>*PP2</u>
	H ₂ N-GCACC ACCGGGATC-PO ₄ -5'

and named adaptor 1 (ADP1)

<u>5'-GTAAT ACGAC TCACT ATAGG GCACG CGTGG TGGCA</u>	
<u>PP1</u>	<u>PP2</u>
	H ₂ N-GCACC ACCGTGGCC-PO ₄ -5'

and named adaptor 2 (ADP2)

The nested primer sequences, PP1 and PP2, are underlined and italicized (see **Note 4**). The end of adaptor 1 (ADP1) shares sequence commonality with the group I restriction enzyme set. The end of adaptor 2 (ADP2) shares sequence commonality with the group II restriction enzyme set. End self-ligation is prevented by the addition of *AvrII* or *AgeI* to the ligation buffer mix.

5. 10X PCR buffer: 400 mM Tricine-potassium hydroxide, 150 mM potassium acetate, 35 mM magnesium acetate, 37.5 µg/mL bovine serum albumin, pH 9.2.
6. 10X deoxynucleotide 5'-triphosphates (dNTPs): an aqueous solution of 2 mM of each dNTP deoxyadenosine 5'-triphosphate (dATP), deoxycytidine 5'-triphosphate (dCTP), deoxyguanosine 5'-triphosphate (dGTP), and deoxythymidine 5'-triphosphate (dTTP).
7. Advantage cDNA polymerase mix, supplied by Clontech Laboratories, CA (see **Note 4**).
8. Thermal cycler: Genius 96 wells cycler, supplied by Techne Ltd, Cambridge UK.
9. PCR tubes: 0.2 mL thin wall 8 strip tubes, supplied by Anachem Ltd, Luton, UK.
10. TE buffer: 10 mM Tris-HCl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.6.

3. Methods

3.1. Template Preparation

3.1.1. Genomic DNA Restriction Enzyme Digestion

1. Digest six 1 µg aliquots of barley powdery mildew DNA (about 20 kb in length) individually with one of 6 restriction enzymes (*SpeI*, *NheI*, *XbaI*, *XmaI*, *NgoMIV* or *BspEI*, as described in **Subheading 2.2.**, (see also **Note 1**) in a 50 µL volume of their appropriate restriction enzyme buffer (**Fig. 1**).
2. After 12 h digestion at 37°C extract the six separate DNA digestion samples twice in equal volumes of phenol/chloroform (v/v) and precipitate the DNA in sodium acetate/ethanol at -20°C (**16**).

3.1.2. Adaptor Ligation: The Protocol is Summarized in **Fig. 1**

1. Ligate the *SpeI*, *NheI*, and *XbaI* digested genomic DNA samples to adaptor 1 (PPR1 + Pad1) and the *XmaI*, *NgoMIV*, and *BspEI* digested samples to adaptor 2 (PPR2 + Pad2). The six 20 µL volume ligation mixes are made up of ligation buffer (50 mM NaCl, either 10 mM Tris-HCl (ligations 1–3) or 10 mM Bis.Tris-HCl (ligations 4–6), 10 mM MgCl₂, 1 mM dithiothreitol (DTT), 1 mM adenosine triphosphate (ATP), pH 7.9 for ligations 1–3; pH 7.0 for ligations 4–6), 1 µg DNA, 100 pmole adaptor (ADP 1 or ADP 2), 6 U of T4 ligase, 6 U restriction enzyme *AvrII* (ligation set 1, 1–3), and 6 U initial restriction enzyme (e.g., *SpeI*); or *AgeI* (ligation set 2, 4–6) and 6 U initial restriction enzyme (e.g., *XmaI*) (see **Notes 1** and **2**).

In summary, the 6 ligations set up are:

Ligation 1: *SpeI* digested genomic DNA, *AvrII* and *SpeI*, adaptor 1

Ligation 2: *NheI* digested genomic DNA, *AvrII* and *NheI*, adaptor 1

Ligation 3: *XbaI* digested genomic DNA, *AvrII* and *XbaI*, adaptor 1

Ligation 4: *XmaI* digested genomic DNA, *AgeI* and *XmaI*, adaptor 2

Ligation 5: *NgoMIV* digested genomic DNA, *AgeI* and *NgoMIV*, adaptor 2

Ligation 6: *BspEI* digested genomic DNA, *AgeI* and *BspEI*, adaptor 2

2. Place the ligation mixes in a Techne Genius PCR machine set to cycle between 12°C (60 min) and 25°C (20 min) for six cycles, followed by an 8°C for 120 min.
3. Stop the ligation reaction by placing the tubes at 70°C for 6 min.
4. Dilute the ligation mixes by the addition of 180 µL TE buffer.
5. Use 1 µL of each diluted ligation mix to set the first round PCR amplification reaction.

3.2. Primer Design

Design and synthesize two nested primers (e.g., P1 and P2) from the known sequence (see **Note 3**).

Any DNA sequence that is longer than 50 bp provides a suitable anchor from which to walk into adjacent unknown sequence. Wherever possible, each of the primers should be about 30 bases in length, with ideally, 10 to 15 GC residues. Unless unavoidable, the primer pairs should not overlap with each other's sequence.

3.3. PCR Amplification

3.3.1. First-Round PCR Amplification

1. Set up a 20 μ L PCR mixture with 40 mM Tricine-potassium hydroxide, 15 mM potassium acetate, 3.5 mM magnesium acetate, 3.75 μ g/mL bovine serum albumin, pH 9.2, 200 μ M dNTP mix, 1X Advantage cDNA polymerase mix, 10 pmole of each primer and 1 μ L of the diluted ligation mix (5 ng DNA).
2. Initiate the amplification reaction with a "hot start", which is automatically performed with the inclusion of the TaqStart antibody (Advantage cDNA polymerase mix, Clontech). Subsequently, utilize our "step down" PCR protocol with the primers PP1/P1 and PP2/P2.
3. Set the PCR machine to cycle:
 - 94°C for 2 s and 72°C for 3 min, 3 cycles;
 - 94°C for 2 s and 70°C for 3 min, 3 cycles;
 - 94°C for 2 s and 68°C for 3 min, 3 cycles;
 - 94°C for 2 s, 66°C for 20 s, 68°C for 3 min, 26 cycles;
 - 68°C for 8 min.
4. Dilute 5 μ L of the PCR products from first-round PCR amplification with 200 μ L sterile distilled water.
5. Use 1 μ L of the diluted first-round PCR products in second-round of PCR amplification.

3.3.2. Second-Round PCR Amplification

Amplicons may or may not be visible following gel electrophoresis of the products of first-round PCR amplification. If no bands are present after round 1 do not discard the tubes, but proceed to second-round PCR amplification. If bands are present after round 1, do not be tempted to stop and sequence these products: it is essential to progress to second-round PCR amplification, which progresses essentially as the first PCR, but it exploits the primers PP2 (nested for PP1) and P2 (nested for P1). However, if bands are present after round 1, but absent after second-round PCR amplification, this may be because of an intron boundary / primer problem (*see Note 7*).

1. Visualize the PCR products from first- and second-round reactions following electrophoresis through a 1.0% agarose gel (*see Note 5*).
2. Excise the dominant bands from the gel and sequence the products directly.

Fig. 2 shows a typical picture of the gel electrophoresis profile of the amplicons after second-round PCR amplification.

4. Notes

1. The choice of the initial six restriction enzymes used to cut the powdery mildew DNA was determined by:
 - a. the resulting average length of the restricted DNA (about 4 kb), which is a suitable length for PCR amplification;
 - b. the commonality of sequence at the cuts ends, i.e., N/NGATC with *SpeI*, *NheI*, and *XbaI* and N/NGGCC with *XmaI*, *NgoMIV*, and *BspEI*.

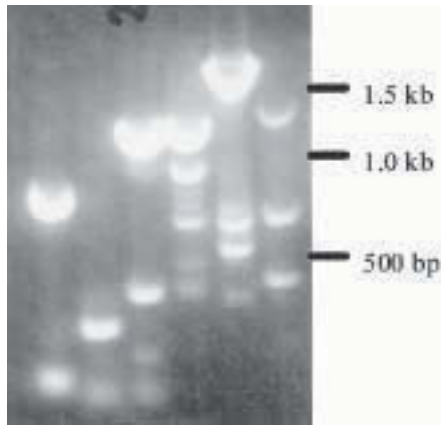


Fig. 2. Gel electrophoresis profile of the amplicons from second-round PCR amplification (with primers PP2 and P2) of the six adaptor-ligated enzyme digests (*SpeI*, *NheI*, *XbaI*, *XmaI*, *NgoMIV*, *BspEI*, loaded from left to right across the gel).

2. The restriction endonucleases *AvrII* and *AgeI* were reserved for the ligation step as both enzymes tolerate the conditions optimal for the ligation step. They were included to prevent self-ligation of the adaptors (*AvrII* for adaptor 1 and *AgeI* for adaptor 2). The second series of restriction enzyme digestions with *SpeI*, *NheI*, *XbaI*, *XmaI*, *NgoMIV*, and *BspEI* is included to obviate problems with genomic DNA fragment self-ligation. The reactions were therefore performed at 12°C (to favor ligation) and 25°C (to favor restriction digestion) and the resultant ligation products are / were restricted genomic DNA with adaptors 1 or 2 at both ends.
3. The -PO₄ and -NH₂ groups were added to the ends of PPR1 and PPR2 to prevent nonspecific 3' end elongation during the PCR. Without the -NH₂ modification, the 3' end of the PPR1 (or PPR2) would elongate using Pad1 (or Pad2) as template at the beginning of the cycling protocol. As a result, no target fragment would amplify—as Pad1 (or Pad2) carries PP1 and PP2 primer sequence, and as both ends of all the genomic fragments in the DNA “pool” from the template preparation carry the Pad1 (Pad2) sequence. Similarly, the -PO₄ group modification is essential to ensure that PPR1(2) ligates to the 3' end of the genomic DNA. Indeed, without the —O₄ modification the PPR1(2) would be lost in the first denaturation step and the 3' ends of both strands of the genomic DNA fragments would elongate using Pad1(2) as template. Both the -NH₂ and -PO₄ modifications are therefore pivotal to successful gene-specific amplification. Without them, there would be spurious and nonspecific amplification.
4. Three additional steps in the protocol are designed to further restrict spurious annealing and sequence amplification. They are as follows.
 - a. The “hot start” technique that is automatically performed by the Advantage cDNA polymerase mix.
 - b. A series of “step down” temperature cycles (72°C to 66°C, ramped by 2°C), which allows the preferential annealing of gene specific primer (P1 or P2). Notably, in first-round PCR amplification, the PP1 primer anneal sequence remains unavailable until the gene specific primer (P1) anneals to the target sequence (i.e., the initial DNA

synthesis). The “step down” protocol ensures that this initial synthesis is as specific as is possible.

- c. The inclusion of the nested primer pair that is pivotal to the success of second-round PCR amplification.
5. The visualization of the products of second-round PCR amplification, following gel electrophoresis, often reveals several amplicons to be present in different samples (**Fig. 2**). Select one amplicon of the approximate length of sequence you require; this will usually be adequate. However, if a sequence longer than 1 kb is needed, select two bands (about 500 to 1000 bp) to be sequenced by the gene-specific primer (e.g., P2) and select all bands of different length to be sequenced by PP2 primer. This may facilitate the rapid collection of all overlapping sequences and so avoid the necessity for further experimentation.
6. The DNA isolated from each gel band is sufficient for up to eight sequencing reactions. Furthermore, each sequencing reaction can give clear readable sequence data up to about 800 bp. However, some of the longer amplicons can sometimes yield relatively poor data, with only 300–400 bp of readable sequence. Here, it is useful to design new primer pairs based on this preliminary sequence data and to use these either to resequence second-round PCR amplicon or to use them to reamplify the products of first-round PCR amplification.
7. Amplification problems could be encountered using a gene-specific primer designed from cDNA sequence data, where the primer spans an intron-exon boundary. In this instance, we have successfully sequenced a DNA fragment, where the P2 primer sequence spans such a boundary, by sequencing an amplicon from first-round PCR amplification. Alternatively, a further primer could be synthesized to avoid this problem area.
8. Regions of genomic DNA that are rich in repetitive sequences may pose a problem for this protocol. We have had no experience of this problem.
9. Theoretically, if a particular gene is distant from the restriction enzyme recognition sites, then amplification will not occur. How likely is this? Given that the average distance between recognition sites is 4096 bp and add this to our ability to walk some 6000 bp then the likelihood of one 6 bp recognition site being absent per 6000 bp is 23%. However, as six different restriction enzymes are used and six amplifications are performed per walk, then the likelihood of all restriction sites being absent is about 0.01%. On a more practical level, in our lab we have performed 80 different walking procedures and have covered some 200 kb sequence (**14,17**) without a single failed attempt!
10. Use a proofreading DNA polymerase to reduce errors in sequence amplification. Furthermore, the direct-sequence analysis of the amplicons, rather than the adoption of cloning procedures, further limits sequence error.

Acknowledgment

This work was supported by BBSRC grant ASD 43/A06377 and BBSRC/COGEME grant 43/IGF 12456.

References

1. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–623.
2. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res.* **16**, 8186.

3. Parker, J. D., Rabinovitch, P.S., and Burmer, G. C. (1991) Targeted gene walking polymerase chain reaction. *Nucl. Acids Res.* **19**, 3055–3060.
4. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., and Markham, A. F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clone. *Nucl. Acids Res.* **18**, 2887–2890.
5. Rosenthal, A. and Jones, D. S. (1990) Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucl. Acids Res.* **18**, 3095–3096.
6. Lagerstrom, M., Parik, J., Malmgren, H., Stewart, J., Pettersson, U., and Landegren, U. (1991) Capture PCR: efficient amplification of DAN fragments adjacent to a known sequence in human and Yac DNA. *PCR Meth. Appl.* **1**, 111–119.
7. Jones, D. H. and Winistorfer, S. C. (1993) A method for the amplification of unknown flanking DNA: targeted inverted repeat amplification. *Biotechniques* **15**, 894–904.
8. Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A., and Lukyanov, S. A. (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucl. Acids Res.* **23**, 1087–1088.
9. Lukyanov, S. A., Gurskaya, N. G., Lukyanov, K. A., Tarabykin, V. S., and Sverdlov, E. D. (1994) Highly efficient subtractive hybridization of cDNA. *Bioorgan. Chem.* **20**, 701–704.
10. Padegimas, L. S. and Reichert, N. A. (1998) Adaptor ligation-based polymerase chain reaction-mediated walking. *Analyt. Biochem.* **260**, 149–153.
11. Weber, K. L., Bolander, M. E., and Sarkab, G. (1998) Rapid acquisition of unknown DNA sequence adjacent to a known segment by multiplex restriction site PCR. *Biotechniques* **25**, 415–419.
12. Cormack, R. S. and Somssich, I. E. (1997) Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. *Gene* **194**, 273–276.
13. Rudi, K., Fossheim, T., and Jakobsen, K. S. (1999) Restriction cutting independent method for cloning genomic DNA segments outside the boundaries of known sequences. *Biotechniques* **27**, 1170–1172, 1176–1177.
14. Zhang, Z. and Gurr, S. J. (2000) Walking into the unknown: a “step down” PCR-based technique leading to the direct sequence analysis of flanking genomic DNA. *Gene* **253**, 145–150.
15. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Deidman, J. G., and Strahl, K. (Eds.), *Current Protocols on Molecular Biology*, 1987. Wiley, NY, p.233.
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, NY.
17. Zhang, Z. and Gurr, S. J. (2001) Expression and sequence analysis of the *Blumeria graminis* mitogen-activated protein kinase genes, *mpk1* and *mpk2*. *Gene* **266**, 57–65.

V

LIBRARY CONSTRUCTION AND SCREENING

Use of PCR in Library Screening

An Overview

Jinbao Zhu

1. Introduction

Traditionally, libraries are screened with different probes to isolate target genes or sequences. These probes can be a particular sequence such as a cDNA, a polymerase chain reaction (PCR) product, or a genomic fragment (**1**). An oligonucleotide can be a probe if no closely related cDNAs or gene clones are available to use, or if only the amino acid sequence of a protein from a gene is available and the gene sequences can be derived by reverse-translating (**2,3**). Sequences derived from mRNA can also be used as a probe to identify gene sequences according to the differences in their concentrations from different biological sources (**1**). In addition, antibody probes can also be used to isolate cDNA clones for target proteins (**4**). However, using cDNA, PCR product and gene probes is limited in library screening depending on the availability of a sequence of sufficient similarity to cross-hybridize with a clone containing the target gene (**1**). In screening a library with a nucleic acid probe, plates are prepared with either bacterial colonies or with plaques from a phage library (**1**). The general screening procedure, for example, in screening cDNA in phage, includes titering the library to determine the phage concentration; plating and lifting the phage particles; hybridizing the immobilized cDNA with appropriate radiolabeled DNA probes; and selecting positive plaques by autoradiography. The procedure is usually cycled in order to isolate a plaque with a single cDNA. This procedure, however, can be time-consuming, labor-intensive, and difficult with requirements such as radiolabeling. Today, the PCR has become a significant technique in identifying and isolating positive clones from libraries. With the sensitivity and specificity of PCR technique, the advantages of PCR-based library screening can be recognized: 1. False positive clones can be avoided since positives are identified by DNA bands of appropriate size; 2. Overall screening time can be shortened; 3. Screening of multiple genes with appropriate prim-

ers can be performed in the same PCR (5). As a result, PCR has been extensively used in the screening of cDNA or genomic libraries (2,6–10) or large-insert DNA libraries of bacterial artificial chromosome (BAC) (11–13) and yeast artificial chromosome (YAC) (14–16).

There are many applications of using PCR techniques in library screening and some applications are reviewed as follows.

2. cDNA and Genomic Library Screening

Narayana R. Isola et al. (8) described a rapid and less expensive method to screen cDNA libraries utilizing PCR to amplify target sequences from the libraries. In studying the expression of H-CAM gene family, they designed oligonucleotide primers to amplify a region of the apparently conserved, putative extracellular domain to simplify and more efficiently identify the presence of CD44 transcripts (17) within cDNA libraries from a variety of tissues and cells. With the screening technique, a novel CD44 transcript was isolated from a reticulocyte cDNA library. And the total procedure could be accomplished in three days (8). This method also can be used to isolate multiple members of a protein family as well as homologous genes in different species by designing appropriate primers to amplify the most conserved regions (8).

Bloem and Yu (6) introduced a time-saving method to screen cDNA or genomic libraries using PCR, which eliminates the time consuming steps of filter hybridization. Briefly, plating was made at the density of 50,000 pfu/150 mm plate. Then 20 μ L filter rinse was used as the template for PCR and followed by agarose gel electrophoresis to determine if the plate had phage plaques with the target sequence by comparing to the DNA band of appropriate size in the gel. Filtering was then made on the target plate and the filter along with the agar was cut into different sections. PCR was then performed on each section and the PCR products were checked by agarose gel electrophoresis. The target phage could be isolated further by more plating and PCR or plating and filter hybridization. The method was simplified with only one round of plating and two orders of PCR before any filter hybridization. With this procedure, multiple pairs of primers can be used in the same PCR reaction to isolate several genes/fragments at the same time.

Similar to filter hybridization, plaque hybridization is labor-intensive, and often creates false positive recombinant clones in screening genomic or cDNA fragments wrapped within phage libraries (9). As an alternative, David I. Israel (9) described a PCR-based screening method that improved the specificity of the screening by requiring that three oligonucleotides (two PCR primers and one hybridization probe) correctly annealed to the desired gene fragment and decreased the odds of acquiring false positive clones. In this protocol, the initial genomic library in λ phage was subdivided into 64 wells. PCR was then performed on the pooled amplified phage using specific oligonucleotide primers. A single well with the target gene sequence was identified through a PCR product of the right size that hybridized with an internal oligonucleotide probe. The positive phage was then used to infect bacteria at a lower number of phage particles per well for a secondary screening, and followed by a subsequent tertiary screening using the same protocol. This protocol can be applied to screen phage

and plasmid libraries of all types and for all genes as long as sequence information is available for designing the appropriate primers and an internal hybridization probe.

According to James W. Larrick (*10*), using quantitative reverse transcription (RT)-PCR to screen cDNA libraries is an extremely sensitive technique, which will find wide application in biomedical research and diagnostics for cancer, metabolic diseases, and autoimmune diseases. In neurobiology, the technique can be used to dissect complex immunological processes and it may also be used to examine functional aspects of the nervous system; in oncology, it can be used to detect the expression of abnormal genes using very small quantities of tumor tissue and studying biopsy material of limited availability; in virology and infectious diseases, the technology provides a useful tool in the study of the pathogenesis of viral diseases, particularly to study the expression and kinetics of viral proteins as well as the host cell's response to proteins such as cytokines, interferons, and stress proteins. RT-PCR also can be used to create a cDNA probe that can be used for library screening. For example, Caruso et al. (*18*) obtained a specific cDNA probe by RT-PCR and screened a cDNA library.

Chiang et al. (*7*) recently studied transcription patterns of sequences on human chromosome 21 using PCR to screen embryonic, fetal, and adult human cDNA libraries. Seventy-three primer pairs were used to screen 41 different cDNA libraries using primers previously used to screen a fetal brain cDNA library. They concluded that their data indicated significant overlap between the genes expressed in different tissues, which is consistent with the supposition that there are as few as 60,000–70,000 human genes. They indicated that PCR-based cDNA library screening technique was easier to perform than Northern blots or RNase protection assays although large numbers of primers might be required to be designed and synthesized.

3. YAC Library Screening

In screening of yeast artificial-chromosome (YAC) clones, colony hybridization was commonly used in identifying target sequences of genes (*19*). However, screening of YAC libraries by colony hybridization causes certain problems according to Heard et al. (*15*), and becomes cumbersome with a large scale of screening (*14*). The problems are: 1. The requirement of primary yeast transformants must be implanted in agar; 2. The number of YACs per cell and the number of cells per yeast colony are much lower than in cloning in bacteria; 3. The spheroplasting of the colonies on the filters prior to lysis can be inefficient leading to under-representation of some colonies (*14*). However, YAC library screening utilizing PCR has been proved to be rapid, sensitive and efficient. For example, Heard et al. (*14*) described a method in which YAC clones were screened using PCR techniques. In brief, 96 colonies were obtained from primary yeast transformants by growing on 96 well microtitre dishes and then followed by several steps to obtain the DNA for screening. For screening, 1 μL of DNA from each of the pools was used in a single PCR using a pair of oligonucleotide primers derived from a known target sequence. A positive colony could be identified with a band of appropriate DNA size.

Green and Olson (*14*) developed an approach for screening large, ordered libraries of YAC clones based on PCR techniques with sequence tag site (STS) content map-

ping. The method starts with individual clones growing in arrays of 384 colonies per nylon filter. The general strategy for screening human YAC libraries includes two consecutive PCR after pooling single-filter pools containing 384 human-YAC clones. The first PCR is performed on multifilter pools to identify the positive target pools. The second PCR is made on each of the constituent single-filter pools from the positive multifilter pools. If a single-filter pool is detected with the appropriate PCR product as separated in polyacrylamide gels and detected by ethidium bromide staining, the positive clone in the 384-clone array can be located by colony hybridization using the probe of the radiolabeled PCR product. With this strategy, a 23,000-clone YAC library for several single-copy human genes has been successfully screened. The PCR-based approach offers several advantages over colony hybridization methods. For instance, the high sensitivity and specificity of this technique can allow a single positive clone to be screened using only 100 ng of DNA from a pool representing thousands of YAC clones. With this protocol, it is possible to determine if a particular DNA sequence is present in a YAC library in less than a day. However, the limitation of this method might be the requirement of DNA sequence and extensive specific primer synthesis. As suggested by the authors, this protocol could be further improved by eliminating the colony hybridization in the second stage for better overall efficiency. For example, Kwiatkowski et al. (16) described a PCR protocol for screening single filters of matrix pools eliminating the step of colony hybridization, and introduced a method for PCR of crude yeast lysates.

4. BAC Library Screening

Asakawa et al. (11) reported a protocol to screen a human genomic BAC with a two-step PCR. The first PCR screening was applied against ten superpools of 9600 BAC clones. The second screening was employed to DNA samples achieved by four-dimensional PCR (4D-PCR) for an identified superpool. The two-step PCR screening could allow isolating a desired BAC clone(s) within a day or so. However, according to the authors, only one positive clone should be present in a given superpool on the 4D or multidimensional two-step PCR screening procedure. If more than one positive clone ($n \geq 1$) were detected in a single superpool, third screening step with at most n^4 PCR assays would be required. In comparison to a 3D-PCR screening procedure (13), the 4D-system might be simpler with less PCR reactions. However, the 3D-screening procedure (13) had the advantage of avoiding a third screening (11). Two-dimensional PCR was also used to screen BAC library. For example, Crooijmans et al. (12) illustrated a protocol using two-dimensional PCR with 125 microsatellite markers and successfully screened a chicken BAC library.

Xu et al. (20) described a protocol using arbitrary primers (AP-PCR) to screen overlapping BAC libraries. They used a rice BAC library to prepare the pools of BAC DNA and chose 22 arbitrary primers to examine the pooled BAC DNAs and individual BAC DNAs. Each primer identified 1–10 loci with the average of 4.4 loci. A total number of 245 overlapping BAC clones was identified and confirmed by DNA-DNA hybridization. This method takes the advantages of Green and Olson's method (14) but eliminates the need of synthesizing specific primers.

The PCR technique has dramatically improved library screening compared to filter or plaque hybridization. In addition to simplifying the procedure, PCR has increased the specificity and efficiency of library screening. The disadvantages for screening using PCR technique might be the availability of sequence information and preparation of oligonucleotide primers. This drawback, however, has been diminished with the development of DNA sequencing techniques and availability of DNA sequences. The relevant detailed PCR-based screening protocols will be presented in the following chapters of this section.

References

1. Halford, N. G. (1996) Heterologous and homologous gene probes. in *Plant Gene Isolation: Principles and Practice*, Foster, G. D. and Twell, D. Eds., Wiley, New York, pp. 103–123.
2. Sugs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H., and Itakura, K. (1981) Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences from human microglobulin. *Proc. Natl. Acad. Sci. USA* **78**, 6613.
3. Lathe, R. (1985) Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* **183**, 1–12.
4. Young, R. A. and Davis, R. W. (1983) Efficient isolation of genes using antibody probes. *Proc. Natl. Acad. Sci. USA* **80**, 1194–1198.
5. Yu, L. and Bloem, L. J. (1993) Use of polymerase chain reaction to screen phage libraries. in *PCR Protocols: Current Methods and Applications*, Vol. 15, White, B. A., Humana Press, Totowa, N J, pp. 211–214.
6. Bloem, L. J. and Yu, L. (1990) A time-saving method for screening cDNA or genomic libraries. *Nucl. Acids Res.* **18**, 2830.
7. Chiang, P. W., Kurnit, D. N., Osemlak-Hanzlik, M., and Trbusek, M. (1995) Rapid PCR-based method to directionally pull out longer cDNA fragments from cDNA libraries. *Biotechniques* **18**, 36–40.
8. Isola, N. R., Harn, H. J., and Cooper, D. L. (1991) Screening recombinant DNA libraries: a rapid and efficient method for isolating cDNA clones utilizing the PCR. *Biotechniques* **11**, 580–582.
9. Israel, D. I. (1993) A PCR-based method for high stringency screening of DNA libraries. *Nucl. Acids Res.* **21**, 2627–2631.
10. Larrick, J. W. (1992) Message amplification phenotyping (MAPPING)-principles, practice and potential. *Trends Biotechnol.* **10**, 146–152.
11. Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Ryo, K., Kudoh, J., Kawasaki, K., Minoshima, S., and Shimizu, N. (1997) Human BAC library: construction and rapid screening. *Gene* **191**, 69–79.
12. Crooijmans, R. P., Vrebalov, J., Dijkhof, R. J., van der Poel, J. J., and Groenen, M. A. (2000) Two-dimensional screening of the Wageningen chicken BAC library. *Mamm. Genome.* **11**, 360–363.
13. Kim, U. J., Birren, B. W., Slepak, T., Mancino, V., Boysen, C., Kang, H. L., Simon, M. I., and Shizuya, H. (1996) Construction and characterization of a human bacterial chromosome library. *Genomics* **34**, 213–218.
14. Green, E. D. and Olson M. V. (1990) Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **87**, 1313–1317.

15. Heard, E., Davis, B., Feo, S., and Fried, M. (1989) An improved method for the screening of YAC libraries. *Nucl. Acids Res.* **17**, 5861.
16. Kwiatkowski, T. J., Zoghbi, H. Y., Ledbetter, S. A., Ellison, K. A., and Chinault, A. C. (1990) Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR. *Nucl. Acids Res.* **18**, 7191–7192.
17. Harn, H. J., Isola, N. R., and Cooper, D. L. (1991) The multispecific cell adhesion molecule CD44 is represented in reticulocyte cDNA. *Biochem. Biophys. Res. Comm.* **178**, 1127–1134.
18. Caruso, C., Bertini, L., Tucci, M., Caporale, C., Leonardi, L., Saccardo, F., Bressan, R. A., Veronese, P., and Buonocore, V. V. (2000) Isolation and characterization of wheat cDNA clones encoding PR4 proteins. *DNA Seq.* **10**, 301–307.
19. Brownstein, B. H., Silverman, G. A., Little, R. D., Burke, D. T., Korsmeyer, S. J., Schlessinger, D., and Olson, M. V. (1989) Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244**, 1348–1351.
20. Xu, J., Yang, D., Domingo, J., Ni, J., and Huang, N. (1998) Screening for overlapping bacterial artificial chromosome clones by PCR analysis with an arbitrary primer. *Proc. Natl. Acad. Sci. USA* **95**, 5661–5666.

Cloning of Homologous Genes by Gene-Capture PCR

Renato Mastrangeli and Silvia Donini

1. Introduction

Conventional procedures to isolate a gene belonging to an ortholog family usually imply the use or the construction of double-stranded cDNA libraries derived from a specific mRNA source of interest (cells or tissues) (1). The double-stranded DNA library plated on various membranes is screened by filter hybridization with a radiolabeled probe (1) derived from the known homologous gene. Clones or plaques hybridizing with the probe are then isolated and sequenced to find the gene of interest. This approach is time-consuming and a large number of false positive clones might be obtained, given that no homology is *a priori* available, especially when the homologous probe contains a short and stable stretch of a sequence sharing clustered homology with both strands of the cDNA library. Alternatively, homologous genes may be isolated by PCR, but only when specific primers are available.

Genomic sequence data are now available from a limited number of eukaryote model organisms (2–6). However, this number will increase in the near future (7,8) and bioinformatics will provide the easiest and fast way to search homologous genes. Meanwhile, data derived from the sequenced genomes and from differential gene expression analysis are making available a huge number of genes that have to be functionally characterized. To assess the gene function and to study phylogenetic relationships in various model organisms, screening procedures to search for gene families (orthologs and paralogs) are still necessary when the sequencing data are not available for the organism of interest.

Gene-Capture polymerase chain reaction (PCR) or (GC-PCR) is a cloning strategy successfully used to isolate the full-length coding cDNA sequence of the mouse ortholog of human lymphocyte activation gene 3 (*LAG-3*) from total RNA of mouse activated thymocytes (9). The method is based on PCR amplification of single-stranded cDNA molecules selectively isolated from a starting cDNA mixture by multiple and sequential capture steps driven by magnetic beads bound homologous probes (capture probes). Two capture steps with biotinylated human homologous probes and

streptavidin-coupled magnetic beads were sufficient to isolate the mouse *LAG-3* cDNA, which was found to be 79% homologous to the human cDNA (9).

The strategy of GC-PCR is shown in panels A-D of **Fig. 1**. The starting single-stranded cDNA library must contain suitable sequences at the 5' and 3' ends, to further allow PCR amplification (*see* panel A and D). This mixture contains all the reverse cDNA molecules of the library with 3' modified by homopolymeric tailing (*see* panel A).

The capture probe is a biotinylated forward single-stranded DNA molecule (a synthetic oligonucleotide or a cDNA fragment prepared by PCR) linked to streptavidin-coupled magnetic beads. Sequences of the various homologous capture probes are derived from nonoverlapping encoding regions of the gene of interest (*see* panel B).

A capture step consists of a fast liquid-phase hybridization reaction (first-order kinetic) between a single-stranded cDNA mixture and a high concentration of single-stranded homologous capture probe followed by magnetic separation of hybridized (captured) single-stranded cDNA molecules. All cDNA molecules forming stable complexes with the capture probe will be retrieved from the solution containing the high abundance of nonhybridizing cDNA molecules. The selection of full-length encoding cDNA molecules is performed by repeated capture steps with different homologous capture probes on the single stranded cDNA molecules released in the previous capture step. Only long cDNA molecules forming stable complexes with the new capture probe will be retrieved from the solution containing the nonspecific single-stranded cDNA molecules that hybridized in the previous capture step. Shorter cDNA molecules of the target gene derived from incomplete retrotranscription and/or partially degraded RNA, and lacking the hybridizing region will be left in solution (*see* panel C).

The capture steps allow the easy and fast magnetic separation of DNA complexes bound to magnetic beads from a complex cDNA mixture of the target source, when a magnet is applied (*see* panel C). The high sensitivity of PCR allows then the amplification of the minute amounts of captured and released cDNA molecules. The use of one base anchored oligo dT mix in both RT reaction and in the first PCR cycle of captured molecules (*see* panel A and D) allows the synthesis of specific cDNA fragments with homogeneous terminal ends (10). Southern blot analysis of the amplified cDNA molecules will help to identify and isolate the fragments corresponding to the full-length target gene. Sequencing of the isolated cDNA will confirm the success of the applied procedure. Two independent PCR amplifications followed by DNA sequencing must be performed to unequivocally determine the DNA sequence of interest (*see* panel D).

Compared to conventional library screening GC-PCR represents a sensitive, fast, simple, specific, and safe method to easily isolate full-length ortholog genes. The starting material is a small amount of total RNA extracted from the cell or tissue of interest. This turns out to be useful when a small RNA source is available. The starting amplifiable single-stranded cDNA library has less complexity when compared to double-stranded cDNA libraries. The sequential hybridization with different probes performed on single-stranded cDNA molecule allows a further high-complexity

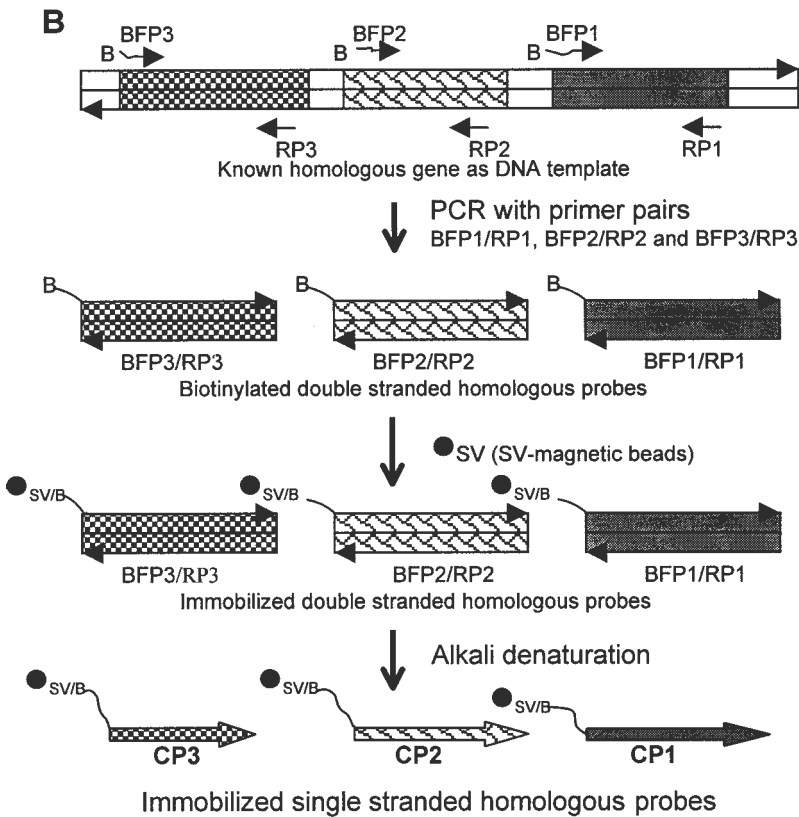
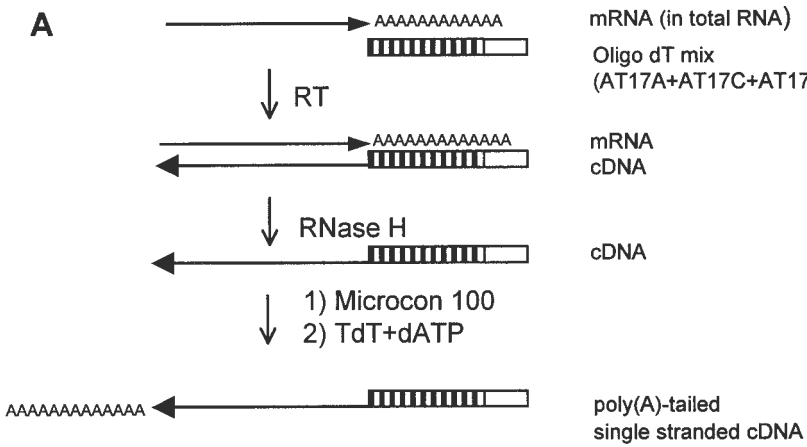


Fig. 1. (A) Synthesis of amplifiable single-stranded cDNA. (B) Synthesis and immobilization of single-stranded homologous probes CP1, CP2, and CP3 on magnetic beads.

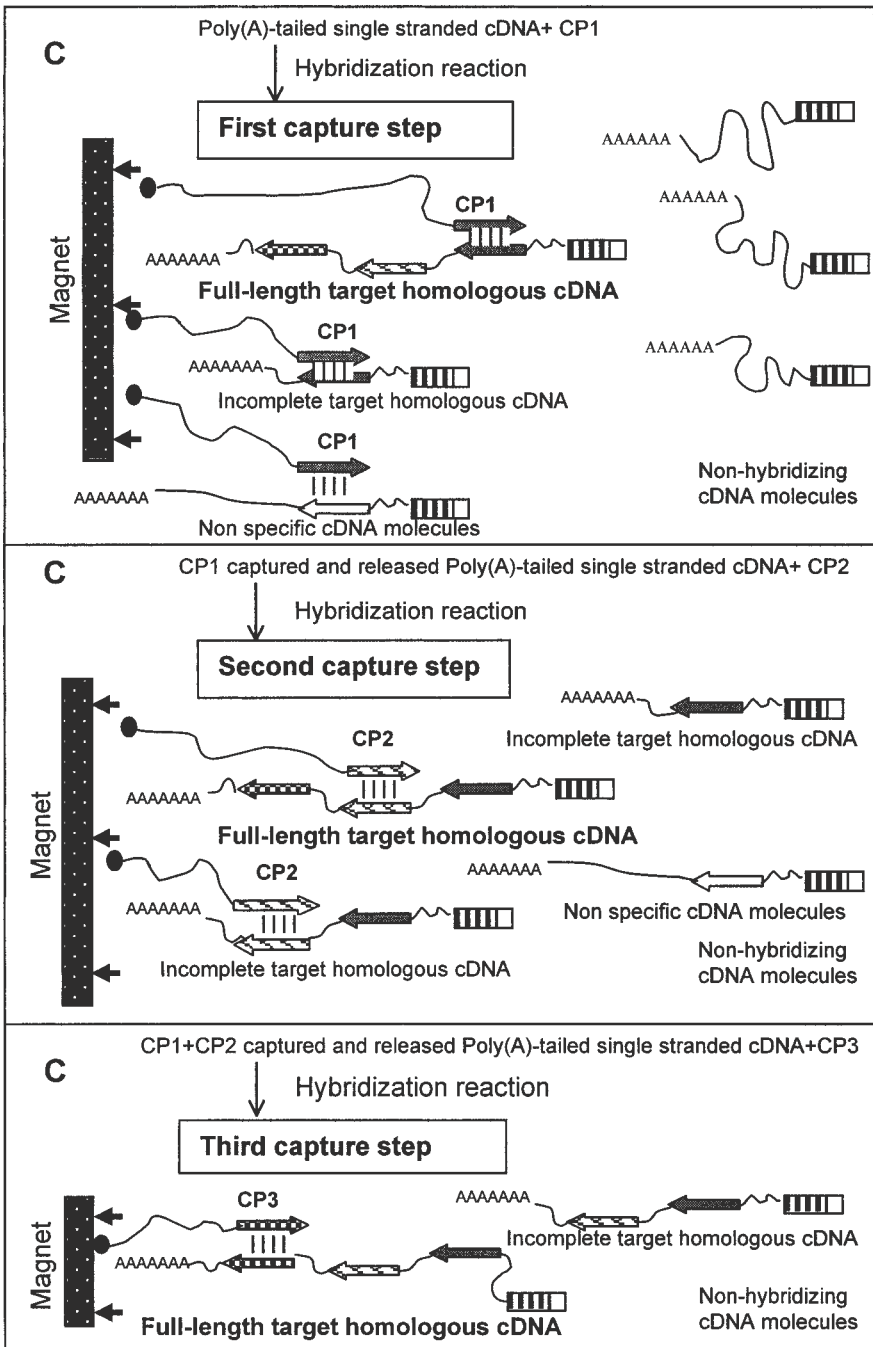


Fig. 1. (C) Selection by capture steps of full-length target homologous cDNA.

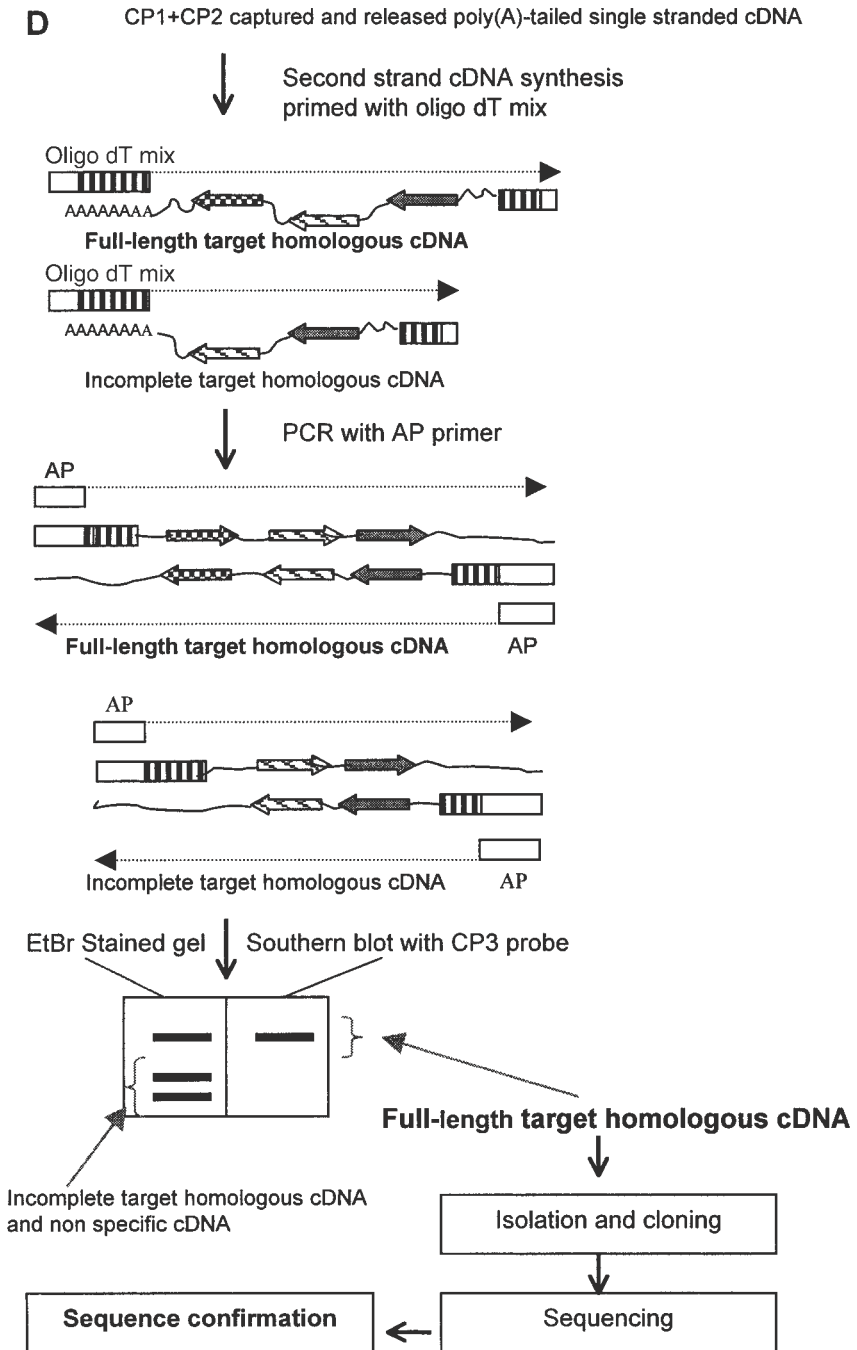


Fig. 1. (D) PCR, Southern analysis, and sequencing of selected single-stranded cDNA.

reduction in each capture step, resulting in a strong reduction of the hybridization background. In addition, the hybridization reaction is performed in a microcentrifuge tube, thus allowing the simultaneous processing of many samples. cDNAs from different libraries (source and/or species) may be therefore easily analyzed.

The method has the potential to isolate genes sharing conserved regions of suitable length, gene variants, and gene-encoding proteins with only limited knowledge of their amino acid sequence. Finally, GC-PCR may also work on DNA libraries (**II**).

2. Materials

Use molecular biology grade reagents. All described oligonucleotides may be purchased from any DNA Synthesis vendors or synthesized in house as described (**II**).

2.1. Extraction of Total RNA and Synthesis of Amplifiable cDNA

1. TriZol reagent (Life Technologies, Gaithersburg, MD).
2. RNase-free water.
3. Agarose gel electrophoresis reagents and equipment for RNA analysis.
4. 30 μ M Oligonucleotide Anchored T₁₇ G (AT17G) in RNase-free water: 5'GGCCCTG-GATCCGGACCTAATTTTTTTTTTTTTTTTTTTG3'.
5. 30 μ M Oligonucleotide Anchored T₁₇ A (AT17A) in RNase-free water: 5'GGCCCTG-GATCCGGACCTAATTTTTTTTTTTTTTTTTTTA3'.
6. 30 μ M Oligonucleotide Anchored T₁₇ C (AT17C) in RNase-free water: 5'GGCCCTG-GATCCGGACCTAATTTTTTTTTTTTTTTTTTTC 3'.
7. 10 μ M Oligo dT mix : 30 μ M AT17G + 30 μ M AT17A + 30 μ M AT17C (1:1:1).
8. RNase H (Roche Diagnostics).
9. RNase inhibitor (Roche Diagnostics).
10. SuperScript™ II RNase H⁻ Reverse Transcriptase (Life Technologies).
11. 0.1 M Dithiothreitol (DTT).
12. 10 mM Deoxynucleotide 5'-triphosphate (dNTP).
13. 0.2X TE buffer: 2 mM Tris-HCl, 0.2 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0.
14. Microcon 100 (Millipore, Beverly, MA).
15. TdT buffer: 200 mM Tris-HCl pH 8.4, 500 mM KCl, 25 mM MgCl₂, 1 mg/mL BSA.
16. Terminal deoxynucleotide transferase (TdT) (Life Technologies).
17. 2.5 mM Deoxyadenosine 5'-triphosphate (dATP) (*see Note 1*).

2.2. Synthesis and Immobilization of Biotinylated Capture Probes to Magnetic Beads

1. DNA source containing the known cDNA of interest (user).
2. Oligonucleotides designed by the user:
 - a. 5'-Biotinylated forward primer 1 (BFP1).
 - b. 5'-Biotinylated forward primer 2 (BFP2).
 - c. 5'-Biotinylated forward primer 3 (BFP3).
 - d. Reverse primer 1 (RP1).
 - e. Reverse primer 2 (RP2).
 - f. Reverse primer 3 (RP3).
3. *Pfu* DNA polymerase (Stratagene, La Jolla, CA).
4. 10X *Pfu* buffer provided with the enzyme.
5. Glycerol.
6. 2.5 mM dNTP.

7. Agarose gel electrophoresis reagents and equipment.
8. TEN buffer: 10 mM Tris-HCl, 0.1 mM EDTA, 1M NaCl, pH 8.0.
9. Dynabeads M-280 Streptavidin 10 mg/mL (Dynal, Oslo, Norway).
10. Magnetic particle concentrator for Eppendorf microtubes (Dynal MPC-E)(Dynal, Oslo, Norway).
11. 0.15 M NaOH.
12. OLIGO Primer Analysis Software (National Bioscience, Inc., Plymouth, MN).

2.3. Capture Step

1. 20X SSC: 3 M NaCl, 0.3 M sodium citrate.
2. 10% (w/v) sodium dodecyl sulfate (SDS).
3. Hybridization and washing buffer: 6X SSC, 0.1% (w/v) SDS.
4. Washing buffer: 1X SSC, 0.1% (w/v) SDS.
5. 1X TE buffer: 10 mM Tris-HCl, 1 mM EDTA, pH 8.0.
6. DNase-free BSA 20 mg/mL (Roche Diagnostics).
7. Hybridization oven and water bath.

2.4. DNA Amplification

1. Reagents and equipment for PCR.
2. Thermostable DNA polymerases:
 - a. *Taq* DNA polymerase (Advanced Biotechnologies Ltd., Leatherhead, Surrey, UK).
 - b. *Pfu* DNA polymerase (Stratagene, La Jolla, CA).
3. 10X *Taq* buffer (without MgCl₂) (buffer IV, vial 1 provided with *Taq* enzyme).
4. Dimethyl sulfoxide (DMSO).
5. 25 mM MgCl₂ (buffer IV, vial 2 provided with *Taq* enzyme).
6. 6.5 mM dNTP.
7. Anchor primer (AP): 5' GGCCCTGGATCCGGACCTAA 3'.

2.5. Reagents for Southern Blot Analysis

1. PCR DIG Labeling Mix (Roche Diagnostics).
2. Blocking Reagent (Roche Diagnostics).
3. Positively charged nylon membrane (Roche Diagnostics).
4. DNA Molecular Weight Markers and digoxigenin-labeled DNA Molecular Weight Markers (Roche Diagnostics).
5. Hybridization buffer: 5X SSC, 0.1% (w/v) SDS, 1% N-lauryl sarcosine, and 0.5% (w/v) Blocking reagent.
6. 5X SSC; 2X SSC; and 1X SSC, 0.1% SDS.
7. Buffer A: 100 mM maleic acid, 150 mM NaCl, pH 7.5.
8. Buffer B: 1% Blocking reagent in Buffer A.
9. Anti-Digoxigenin-POD, Fab fragments (Roche Diagnostics).
10. 100 mM Tris-Cl, 150 mM NaCl, pH 7.5.
11. Washing buffer: 0.3% Tween 20 in buffer A.
12. ECL reagents (Amersham Pharmacia Biotech).
13. Photogene development folder (Life Technologies) or Saran Wrap.
14. Hyperfilm ECL (Amersham Pharmacia Biotech).

2.6. DNA Cloning and Sequencing

1. Conventional cloning reagents and equipment.
2. Reagents and equipment for DNA sequencing.

3. Method

3.1. Synthesis of Amplifiable Single-Stranded cDNA

Starting from Total RNA

3.1.1. Total RNA Extraction

1. Extraction of total RNA from cells or tissues is performed with TriZol reagent, according to the manufacturer's instructions; alternatively, other conventional procedures might be used. The RNA preparation should be of good quality.
2. At the end of the procedure, dissolve total RNA in RNase-free water and store at -20°C .

3.1.2. Reverse Transcriptase (RT) Reaction

Total RNA (5 μg) is reverse transcribed in 0.5-mL Eppendorf tube (*see Note 2*).

1. Add a volume corresponding to 5 μg total RNA, then 2 μL oligo dT mix (10 μM each primer), and RNase-free water up to 10 μL total volume.
2. Incubate at 65°C for 10 min, then chill on ice.
3. Spin down the liquid and add the following components: 10 μL of 5X SuperScript II buffer, 5 μL 0.1 M DTT, 5 μL 10 mM dNTP, 1 μL RNase inhibitor (10–50 U), 2 μL SuperScript II (400 U), 17 μL RNase-free water. Total volume is 50 μL .
4. Incubate samples at 37°C for 90–120 min.
5. Stop the reaction by incubating at 90°C for 5 min, then chill on ice for 10 min.
6. Add 1 μL RNase H (1 U) and incubate samples for 30 min at 37°C to degrade RNA.
7. Add 0.2 mL of 0.2X TE buffer and transfer the resulting solution to the sample reservoir of a Microcon 100. Insert into a microcentrifuge, ultrafiltrate at $500 \times g$ for 5 min following Microcon instructions. Four more ultrafiltration steps are then performed with 0.5 mL of 0.2X TE buffer at $500g$ for 12 min. Final volume is adjusted to 50 μL with 0.2X TE buffer (*see Note 3*).
8. Store single-stranded cDNA at -20°C .

3.1.3. Synthesis of Amplifiable Single-Stranded cDNA by Poly (A) Tailing with TdT

In this step the synthesized reverse single-stranded cDNA is modified to allow PCR amplification. Many procedures may be used, each procedure has its own advantages in terms of simplicity, cost, rapidity (*see Note 4*). Here, the simple TdT homopolymeric tailing with dATP is used. A poly (A) tail is added to the 3' end of the reverse strand of single-stranded cDNA. This allows the same primer mixture (oligo dT mix) to be used in the RT reaction and for the second-strand cDNA synthesis (*see Note 5*).

1. Transfer 16 μL of single-stranded cDNA mixture (corresponding to 1.6 μg total RNA) in a 0.5-mL Eppendorf tube. Incubate at 70°C for 5 min, then chill on ice.
2. Spin down the liquid and add the following components: 1 μL TdT buffer, 2 μL dATP (2.5 mM), 1 μL TdT (10 U). Total volume is 20 μL . Incubate at 37°C for 10 min (*see Note 1*).
3. Inactivate TdT by incubating at 70°C for 10 min.
4. Store amplifiable single-stranded cDNA at -20°C .

3.2. Synthesis and Immobilization of Single-Stranded Biotinylated Homologous Probes to Dynabeads M-280 Streptavidin

5'-biotinylated probes may be easily prepared by oligonucleotide synthesis (*see Note 6*) or by PCR. Here, the preparation by PCR from a DNA source containing the

gene of interest, i.e., plasmid or PCR fragment, is described. The design of oligonucleotides used as primer pairs should be performed using a suitable primer design software, i.e., OLIGO.

3.2.1. PCR Amplification of Specific cDNA Fragments to be Used as Homologous Probes

Three homologous probes have to be generally designed and synthesized on translated exons. The first probe will be designed on terminal exons, the second on middle exons, and the third toward the 5' exons. This design will allow the isolation of molecules containing the full-length cDNA sequence. For each homologous probe to be amplified use 5'-biotinylated forward and nonbiotinylated reverse primer sets: BFP1/RP1, BFP2/RP2, and BFP3/RP3 (see Fig. 1, panel B).

1. Add the following to the 0.5-mL PCR tube kept on ice: 150 fmoles of DNA template (i.e., plasmid, cDNA fragment), 9 μ L 10X *Pfu* buffer, 3 μ L 10 μ M biotinylated forward primer (i.e., BFP1, BFP2, and BFP3), 3 μ L 10 μ M nonbiotinylated reverse primer (RP1, RP2, and RP3), 10 μ L 2.5 mM dNTP, 5 μ L glycerol, and H₂O up to 90 μ L volume.
2. Add 100 μ L mineral oil and transfer reaction tubes to a thermal cycler.
3. Denature PCR mixture at 96°C for 5 min then maintain the temperature at 80°C.
4. Add 10 μ L of prewarmed *Pfu* DNA polymerase (2.5 U in 1X *Pfu* buffer) at 80°C (see Note 7).
5. Amplify using the following conditions: 96°C for 30 s, optimal annealing temperature (derived from OLIGO software) for 30 s, 75°C for 3 min, 25 cycles with a final extension step of 5 min.
6. Store at 4°C up to purification.
7. 5 μ L of biotinylated PCR product (BFP1/RP1, BFP2/RP2, and BFP3/RP3) is analyzed by agarose gel electrophoresis following standard procedure.
8. If the expected product is unique, the PCR mixture may be directly ultrafiltered by Microcon 100 and four ultrafiltration steps with 500 μ L TEN buffer (500 \times g for 12 min) (see Note 8). Otherwise, purify the expected fragment by conventional procedure. The final volume is adjusted to 100 μ L with TEN buffer.

3.2.2. Immobilization of Biotinylated PCR Product (Homologous Probe) to Dynabeads M-280 Streptavidin

1. For each homologous probe to be prepared add 100 μ L of SV-magnetic beads suspension (1 mg) (Dynabeads M-280 Streptavidin 10 mg/mL) in an Eppendorf tube. Use the Dynabeads following the Dynal instructions.
2. Apply the magnet using the magnetic particle concentrator for Eppendorf microtubes (Dynal MPC-E). When the liquid phase is completely separated from the solid phase, discard the liquid phase.
3. Add 100 μ L TEN, mix beads, apply the magnet and wait separation, discard the liquid phase.
4. Repeat twice the previous washing step.
5. The biotinylated PCR product (BFP1/RP1, BFP2/RP2, and BFP3/RP3) is added (90 μ L) to the beads, mix beads, and incubated for 30 min at room temperature with gentle agitation (immobilization step).
6. Apply the magnet, wait separation and transfer the liquid phase to an Eppendorf tube. This liquid phase will be used to check the efficiency of the immobilization step.

7. Wash the beads containing the immobilized probe three times with 100 μL TEN as described in **step 3**.
8. Add 100 μL TEN; mix beads and store at 4°C up to denaturation.
9. Check the immobilization efficiency by loading on a suitable agarose gel 5 μL of biotinylated probe before binding and 5 μL of the supernatant after the binding reaction. A decrease of ethidium bromide signal between the pre- and postbinding indicates that most of the probe is immobilized.

3.2.3. Single-Stranded Homologous Probe Immobilized to Dynabeads M-280 Streptavidin

1. Resuspend by mixing the magnetic beads linked to double-stranded homologous probe. Apply the magnet and discard the TEN liquid phase.
2. Add 100 μL of 0.15 M NaOH and incubate for 10 min at room temperature with gentle agitation. The double-stranded probe is denatured.
3. Apply the magnet and discard the liquid phase containing the single-stranded reverse sequence not linked to magnetic beads.
4. Add 100 μL of 0.15 M NaOH, mix beads, apply the magnet and discard the liquid phase (fast wash with denaturing solution).
5. Wash the beads three times with 100 μL TEN as described in **Subheading 3.2.2., step 3**. The magnetic beads contain now the forward sequence of the single-stranded homologous probe.
6. Add 200 μL TEN and store beads at 4°C. For long-term storage add 0.02% sodium azide.

PCR fragments BFP1/RP1, BFP2/RP2, and BFP3/RP3 generate the homologous capture probes CP1, CP2, and CP3, respectively (*see Fig. 1, panel B*).

3.3. Selection by Capture Steps

Start the selection procedure using the homologous capture probe CP1 that is designed more closely to the 3' end of the encoding DNA sequence. The second capture steps will be performed with nonoverlapping homologous capture probe (capture probe 2, CP2) that is upstream to CP1. If necessary, a third capture step with a third capture probe (CP3) further upstream to CP2 will be performed.

3.3.1. First Capture Step

1. In an Eppendorf tube add 50 μL suspension of SV-magnetic beads linked to homologous capture probe 1 (CP1), mix beads, apply the magnet, and wait separation, discard the liquid phase (*see Note 9*).
2. Add 100 μL hybridization buffer (6X SSC, 0.1% SDS), mix beads, apply the magnet, and wait separation, discard the liquid phase.
3. Repeat twice the previous step, then add 100 μL hybridization buffer (prewarmed at 70°C).
4. Denature the amplifiable poly (A)-tailed single-stranded cDNA by incubating at 90°C for 10 min. Chill on ice.
5. Add 10 μL of denatured amplifiable poly(A)-tailed single-stranded cDNA (equivalent to 0.8 μg total RNA) to beads. Mix gently.
6. Incubate the mixture for 10 min at 70°C with gentle agitation.
7. Incubate the hybridization mixture for 4 h in hybridization oven at 45°C under gentle rotation to maintain the beads in suspension.

8. Mix beads, apply the magnet, and wait separation, discard the liquid phase. Hybrids between the homologous capture probe 1 (CP1) and poly(A)-tailed single-stranded cDNA are retrieved from the solution by the MPC-E magnet.
9. Add 100 μ L hybridization buffer, mix beads, apply the magnet, and wait separation, discard the liquid phase.
10. Repeat twice. Resuspend in 300 μ L of the same buffer.
11. The beads suspension is divided into three 100 μ L aliquots (each containing hybrids deriving from 0.27 μ g starting total RNA).
12. Using prewarmed solutions, wash each aliquot five times as described in **step 2** at different temperatures to determine the optimal selection conditions, i.e., 45°C, 55°C, and 65°C, respectively (*see Note 10*).
13. Equilibrate the beads containing hybridization complexes at room temperature by two washing steps with 100 μ L of TEN. The suspension can be stored at 4°C up to the next elution step (*see step 14*).
14. Mix beads, apply the magnet, wait separation, and carefully discard all the liquid phase. Resuspend the beads containing hybridization complexes with 100 μ L 1X TE buffer. Incubate at 90°C for 5 min. The captured poly(A)-tailed single-stranded cDNA molecules are released into solution.
15. Apply the magnet to the warm solution, wait separation, and carefully collect the liquid phase, which contains released single-stranded cDNA molecules captured by CP1 homologous probe.
16. Add bovine serum albumin (BSA) (50 μ g/mL final concentration) to the recovered cDNA molecule and store at -20°C.

3.3.2. PCR of Single-Stranded cDNA Molecules

Derived from the Capture Step

1. In a 0.5-mL PCR tube, add 30 μ L of amplifiable single-stranded cDNA, 0.5 μ L oligo dT mix (10 μ M each primer), 7.5 μ L 10X *Taq* buffer (without $MgCl_2$), 8 μ L DMSO, 21 μ L 25 mM $MgCl_2$, 8 μ L 6.25 mM dNTP. The final volume is 75 μ L.
2. Add 80 μ L mineral oil and transfer reaction tubes to a thermal cycler.
3. Denature PCR mixture at 95°C for 5 min, and then maintain the temperature at 80°C.
4. Add 5 μ L of prewarmed *Taq* DNA polymerase (2.5 U in 1X *Taq* buffer) at 80°C (*see Note 7*).
5. For the first PCR cycle, use the following conditions: 96°C for 30 s, 48°C for 2 min, and 72°C for 30 min (*see Note 11*).
6. The first cycle is followed by a denaturation step at 96°C 4 min, 80°C 5 min.
7. Add 6 μ L of 10 μ M AP primer at 80°C (*see Note 12*).
8. Amplify for 35 cycles using the following conditions: 96°C for 30 s, 60°C for 30 s, 72°C for 4 min. Final extension time 10 min.
9. Store at 4°C.

3.3.3. Southern Blot of Captured and Amplified cDNA

The first capture steps may be easily followed by Southern blot analysis with the homologous probe 2 (*see Note 13*).

3.3.3.1. SYNTHESIS OF THE PROBE FOR SOUTHERN BLOT

This step provides a DIG-labeled single-stranded probe for the chemiluminescent detection of hybridizing molecules.

1. Add to 0.5 mL PCR tube: 15 ng of DNA template (i.e., PCR fragments BFP2/RP2, or BFP3/RP3 obtained in **Subheading 3.2.1.**), 20 pmoles of reverse primer (RP2 or RP3), 9 μ L 10X *Taq* buffer, 6 μ L 25 mM MgCl₂, 10 μ L of PCR DIG-labeling dNTP mix, and H₂O up to 90 μ L volume.
2. Add 100 μ L mineral oil and transfer the reaction tube to a thermal cycler.
3. Denature the mixture at 95°C for 5 min then maintain the temperature at 80°C.
4. Add 10 μ L of prewarmed *Taq* DNA polymerase (2.5 U in 1X *Taq* buffer) at 80°C.
5. Perform 25 cycles using the following conditions: 96°C for 30 s, annealing temperature (derived from OLIGO software) for 30 s, 72°C for 2 min.
6. Store the reaction mixture at -20°C until required.

3.3.3.2. SOUTHERN BLOT

1. Ten microliters of amplified cDNA from **Subheading 3.3.2.**, 10 μ L of DIG-labeled and nonlabeled molecular weight markers are subject to agarose-gel electrophoresis and blotted onto a positively charged nylon membrane by conventional procedure.
2. The blot is prehybridized for 60 min at 45°C with hybridization buffer (0.25 mL/cm² filter).
3. Remove the prehybridization solution and add 0.2 mL/cm² filter of hybridization solution containing 50 μ L of DIG-labeled probe that has been previously treated for 10 min at 70°C. Incubate the blot overnight at 45°C with gentle agitation.
4. Remove the filter and rinse with 5X SSC for 2 min at room temperature.
5. Wash the filter four times with 1X SSC, 0.1% SDS (0.7 mL/cm² filter) for 30 min with agitation at a suitable temperature depending on the requested stringency (*see Note 10*).
6. Remove the filter and equilibrate in buffer A for 2 min at room temperature.
7. Filter is incubated in buffer B (blocking buffer) (0.2 mL/cm² filter) for 30 min at room temperature.
8. Remove the blocking solution and add 0.2 mL/cm² filter of Anti-Digoxigenin-POD solution. The Anti-Digoxigenin-POD solution is prepared by diluting in Buffer B (1:10,000) the Anti-Digoxigenin-POD stock solution (150 U/mL in 100 mM Tris-Cl, 150 mM NaCl, pH 7.5). Incubate for 30 min at room temperature with gentle agitation.
9. Remove the filter and rinse with buffer A for 2 min at room temperature.
10. Wash the blot twice with 0.7 mL/cm² filter of washing buffer for 15 min at room temperature and vigorous agitation.
11. Remove the filter and rinse with 2X SSC for 2 min at room temperature.
12. Place the filter on filter paper to drain the liquid, then on a flat container.
13. Hybrid detection is performed with ECL reagents: mix an equal volume of detection solutions 1 and 2 and add (0.15 mL/cm²) directly to the blot on the side containing the DNA. Incubate for 1 min at room temperature. Drain off excess detection buffer and place the filter between two sheets of Photogene development folder or alternatively wrap filter in plastic wrap. Gently smooth out air pockets and in a darkroom (red safelight on) place the filter in a cassette (DNA side up). Cover with a sheet of Hyperfilm-ECL and expose for 1, 10, and 60 min, depending on the required exposure time. Develop the film to detect hybrids between the probe and the DNA fragments as well as the digoxigenin-labeled MW markers.

3.3.4. Second Capture Step

In this step, the magnetic beads linked to the homologous single-stranded probe 2 (Capture Probe 2, CP2) are used to bind cDNA released from the first capture step to select the cDNA molecules sharing homology with both CP1 and CP2 probes (*see Fig. 1*, panel C).

1. In a 0.5-mL tube, pool the Southern positive fractions of released single-stranded cDNAs from **Subheading 3.3.1., step 16**, total volume ranges from 70 μL up to 210 μL .
2. Incubate at 90°C for 10 min. Chill on ice.
3. In a 1.5-mL Eppendorf tube, add 138 μL of the denatured-pooled fractions (if only one fraction is available, add 70 μL of this fraction and 68 μL of 10 mM Tris-HCl, 1 mM EDTA, pH 8.0, BSA 50 $\mu\text{g}/\text{mL}$ instead), 60 μL 20X SSC and 2 μL 10% SDS. Final volume is 200 μL .
4. In a 1.5-mL Eppendorf tube, add 50 μL suspension of magnetic beads linked to homologous probe 2 (CP2) (0.25 mg beads), mix beads, apply the magnet, and wait separation; discard the liquid phase.
5. Add 200 μL hybridization buffer, mix beads, apply the magnet, and wait separation, discard the liquid phase.
6. Add 200 μL solution from **step 3** to CP2 beads from **step 5**.
7. Incubate the mixture for 10 min at 70°C with gentle agitation.
8. Incubate the hybridization mixture with gentle agitation for 4 h at a single stringent temperature according to results of the previous Southern blot.
9. Mix beads, apply the magnet and wait separation, discard the liquid phase. Hybrids between the homologous capture probe CP2 and amplifiable single-stranded cDNA are retrieved from the solution by the magnet.
10. Add 200 μL hybridization buffer, mix beads, apply the magnet, and wait separation; discard the liquid phase. Repeat twice.
11. The suspension in 200 μL hybridization buffer is divided into two 100 μL aliquots.
12. Wash each aliquot 5X at the hybridization temperature with either 6X SSC, 0.1% SDS or at more stringent conditions with 1X SSC, 0.1% SDS.
13. Equilibrate the beads containing hybridization complexes by two washing steps with 100 μL TEN. The suspension can be stored at 4°C up to the next elution step.
14. Apply the magnet, wait separation, and discard the liquid phase. Resuspend the beads containing hybridization complexes with 1X TE. Incubates at 90°C for 5 min.
15. Apply the magnet to the warm solution, wait separation, and carefully collect the liquid phase, which contains released single-stranded cDNA molecules selected by CP1 and CP2 homologous probes.
16. Add BSA (50 $\mu\text{g}/\text{mL}$ final concentration) to the recovered cDNA and store at -20°C.

3.3.5. Southern Blot Analysis After the Second Capture Step

Thirty microliters of each recovered fraction is amplified by PCR and analyzed by Southern blot with a DIG-labelled probe as described in **Subheadings 3.3.2.** and **3.3.3.** (**Fig. 1**, panel D). This analysis should now confirm the isolation of the putative homologous gene and indicate whether an additional capture step is required (*see Note 14*).

To evaluate by PCR and Southern blotting the capture step efficiency, the controls should include 1. starting amplifiable single-stranded cDNA, 2. cDNA released from the first capture step, and 3. cDNA released from the second capture step (*see Note 14*).

3.3.6. Third Capture Step

Perform this additional capture step only if necessary. In this case follow **Subheading 3.3.4.** by using the homologous capture probe 3 (CP3) instead of CP2 on the released single-stranded cDNA from the second capture step (**Fig. 1**, panel C). Specific amplified fragments corresponding to the full-length homologous gene should be

obtained by PCR. A further Southern blot analysis with CP3 probe might be performed to get a general view of the applied GC-PCR procedure by adding the necessary controls and following the steps in **Subheadings 3.3.2.** and **3.3.3.**

Please *see* **Note 15** if GC-PCR fails.

3.4. Cloning of Captured cDNA

The cloning step is necessary for DNA sequencing of the captured DNA (*see* **Note 16**). The user will choose his own preferred cloning and screening strategy.

It is advisable to use *Escherichia coli* (*E. coli*) strain that allows the cloning of unstable DNA structures such as inverted repeats, which are present at 5' and 3' ends of the obtained fragment (*see* **Note 17**). The following are very general instructions:

1. Gel purify the selected PCR fragment.
2. Ligate the PCR fragment to the chosen plasmid vector.
3. Transform a suitable *E. coli* strain.
4. Screen recombinant plasmids.
5. Select a positive plasmid for DNA sequencing.

3.5. DNA Sequencing and Analysis

Plasmid DNA is sequenced to finally confirm the cloned cDNA is the gene of interest.

3.5.1. DNA Sequencing of Captured cDNA

1. To start DNA sequencing, use universal primers present on the plasmid vector both upstream and downstream to the cloned insert.
2. Use the most suitable sequencing chemistry for your equipment.
3. For long DNA fragments, DNA sequence obtained by the universal primers may be used to design additional sequencing primers to complete the sequencing of the internal DNA. Available primers used to obtain the homologous probes (BFP1, BFP2 BFP3, RP1, RP2, and RP3) can also be used as sequencing primers if the homology is very high and the 3' end perfectly matches with the homologous gene; the annealing temperature should be decreased accordingly as the homology decrease.
4. Complete DNA sequence on both strands of the captured DNA fragment following the most suitable sequencing strategy.

3.5.2. Confirmation of the DNA Sequence

Taq DNA polymerase and a modified PCR buffer were used to obtain the preliminary sequencing data on isolated fragment (PCR1). These conditions increase the yield and also the amplification error rate. Therefore, the obtained sequence has to be confirmed by two additional independent RT-PCR amplifications (PCR2 and PCR3).

1. Design PCR primers on the extremities of isolated homologous gene (*see* **Note 18**).
2. With the designed primer pairs, perform two RT-PCR reactions on starting RNA (PCR2 or PCR3). At least one PCR amplification has to be performed with a different enzyme (with increased fidelity, i.e., *Pfu*) and in selected conditions that minimize misincorporations.
3. Perform DNA sequencing on both strands of purified PCR2 and PCR3 fragments.
4. Align the sequencing data of the isolated fragment (PCR1) and those derived from PCR2 and PCR3.
5. Determine the correct DNA sequence of the isolated gene. The resulting consensus sequence will be unequivocally the correct DNA sequence of the target homologous gene (*see* **Note 19**).

4. Notes

1. The 5'RACE system kit (Life Technologies) contains most of the reagents used in this procedure, i.e., terminal deoxynucleotide transferase (TdT), SuperScript, *E. coli* RNase H, TdT buffer, 0.1 M DTT and 10 mM dNTP. In addition, if it is preferred to perform a dC tailing to facilitate the cloning, the kit contains reagents to perform dC homopolymeric tailing and the 5'RACE anchor primer for the further PCR amplification.
2. The RT reaction is a critical step for long retrotranscripts (12). In our RT-PCR conditions, a DNA smear larger than 300 bp is expected. Indeed, a smear between 500 and 2700 bp with higher intensity in the 700–1150 bp region was observed in ethidium bromide stained gel (9). The reaction conditions of reverse transcription should favour the synthesis of full-length cDNA molecules. Commercial kits are now available for this purpose.
3. This step is critical and is necessary to remove residual primers, dNTP, degraded RNA and single-stranded cDNA molecules shorter than 300 nucleotides. Unwanted TdT reactions such as different nucleotide incorporation or primer tailing are therefore avoided. Store all Microcon 100 filtrates, which in case of leakage should be concentrated and filtered with a new device.
4. Different procedures or commercial kits may be used to obtain amplifiable single-stranded cDNA: terminal deoxynucleotide transferase (TdT) tailing (12), single-strand ligation of cDNA or "SLIC strategy" (13), Gene Racer kit from Invitrogen, and the Switch Mechanism at the 5' end of RNA Template (SMART technology) by using the SMART oligonucleotide during RT reaction (see SMART™ kit, Clontech).
5. The disadvantage in our method is that the PCR product has to be cloned to obtain the DNA sequence. Possible problems with the inverted repeat at the 5' and 3' ends such as the suppression PCR effect (14) were not observed (9).
6. When capture probe sequence is shorter than 100 nucleotides, the 5' biotinylated forward homologous probe might be easily synthesized. Biotin may be directly introduced to 5' terminus of the capture probes during automated synthesis by using LC biotin-ON from Clontech. For a longer sequence, use the described PCR procedure.
7. Hot-start procedure (15) should be used to prevent mispriming and primer dimerization. When adding further reagents, i.e., the prewarmed enzyme, be careful to add the solution under the mineral oil layer.
8. The purification step is necessary to remove the residual biotinylated primer that could affect the next binding step to SV-magnetic beads.
9. From Dynal instructions, 1 mg beads can bind 10–30 pmoles of DNA molecules with 50–3000 nucleotides in length, at inverted ratio with respect to the DNA length. 0.25 mg SV-magnetic beads can bind 2.5–7.5 pmoles of homologous capture probe. For sequences 100–1000 nucleotides in length, the bounded amount might range from 0.25 to 0.825 µg of probe in 100–200 µL hybridization volume. A high-probe concentration ranging from 1.25 to 8.25 µg/mL is, therefore, obtained allowing fast hybridization reactions compared to conventional library screening.
10. Incubation temperature should be adjusted depending on the required stringency. The salt concentration may also be decreased to increase the stringency. OLIGO program provides melting temperatures at several salt concentrations for a given probe and for various homology conditions. As the homology is not *a priori* known, the first step is performed in conditions that allow hybridization of molecules sharing a homology higher than a minimal assumed value, i.e., 60%–70%.
11. PCR must be performed in conditions suitable to obtain full-length molecules, especially in the critical first cycle when amplification fails if the full-length double stranded cDNA

is not obtained. Captured cDNA is amplified by PCR with *Taq* DNA polymerase according to the method of Frohman (12), with minor modifications. A buffer composition (6.5 mM MgCl₂, 10% DMSO, 1.5 mM dNTP) was found by Frohman to give high amplification yield of cDNA.

12. AP primer sequence is present on 5' terminus of both forward and reverse sequence. If a different tailing is performed two different anchor primers have to be used. Anchor sequences containing suitable cloning site might also be designed to facilitate the downstream cloning step.
13. Although Southern blot after the first capture step may be omitted to speed-up the procedure, it provides useful information on both CP1-captured cDNA and CP2 capture probe. The ethidium bromide stained gel shows the complexity of CP1-captured material at various washing temperatures. The Southern positive signal with the homologous probe 2 used at a given washing stringency will give indications on the presence of the target gene, its expected size and specificity when compared to ethidium bromide signals. Furthermore, it will give indications that the second homologous probe may be successfully hybridized and washed at a given temperature in the further capture step. The use of a probe corresponding to the sequence of the first capture probe has to be avoided as it will not be specific and almost all ethidium bromide signals will be positive. This is a serious problem in the first capture step.
14. In our study, the homology of capture probes 1 and 2 (713 and 290 nt long, respectively) with the gene of interest was found to be 79% and 82%, respectively. After the first and nonselective capture step on activated and nonactivated thymocytes, a smear was still observed by ethidium bromide staining, but the homologous gene was clearly visible by Southern blotting in activated thymocytes only. After the second capture step, ethidium bromide stained gel clearly showed only the bands of interest (9).
15. GC-PCR may fail when: 1. the target gene is present in the mRNA source at a very low level (use an activated cell source and/or increase the PCR cycle number); 2. starting mRNA is degraded (use a new RNA preparation and avoid RNase contamination); 3. homopolymeric tailing fails or tailed cDNA is lost in the Microcon purification step (check ethidium bromide staining of the amplified noncaptured cDNA mixture, the absence of an intense and smeared signal indicates that cDNA is not amplifiable; repeat the tailing reaction with new reagents); 4. first strand cDNA synthesis of long mRNA molecules is uncompleted (shorter fragment than expected are observed by ethidium bromide stained gel after the second capture step, if the 5' encoding region is not present, Southern blot will be negative by using a probe corresponding to CP3 region and positive with a probe corresponding to CP2. If this is the case, isolation of the fragment followed by 5'RACE will provide the full-length gene. Alternatively, design additional capture probe closest to 3'encoding sequence or optimise the RT reaction to obtain longer retrotranscripts.); 5. Second-strand cDNA synthesis of long cDNA is uncompleted in the first PCR cycle (increase extension time, use conditions and enzymes which allow long distance PCR).
16. A different TdT tailing should avoid the cloning step and a direct sequencing of the isolated fragment may be performed with the anchor primers if an acceptable homogeneity of the PCR product is achieved (*see Note 1*).
17. SURE strains are available from Stratagene which allow the cloning of unstable DNA structures.
18. Anchored specific primers containing suitable cloning sites might be used to facilitate the gene cloning into expression vector for recombinant protein production.

19. The sequencing data are derived from three independent RT-PCR reactions, one performed with a different enzyme chosen between high fidelity proofreading enzymes such as *Pfu*, ThermalAce (Invitrogen), *Pwo* (Roche), etc. If present in a given position, an error will be present only in one sequence whereas the other two sequences will show the correct nucleotide.

In PCR1 conditions used to maximize the yield (12), several misincorporations were observed (9). Two additional PCRs were performed with *Taq* and *Pfu* DNA polymerase, respectively. Both PCR2 and PCR3 sequences were found identical, suggesting identity with the natural sequence (9).

Acknowledgments

The authors would like to thank Carlo Serafini for supporting this work and Emilia Micangeli for the expert and valuable technical assistance.

References

1. Sambrook, J., Fritsch, E. F., Maniatis, T., and Irwin, N. (1989) Construction and analysis of cDNA libraries, Preparation of radiolabeled DNA and RNA probes, Synthetic oligonucleotides probes, Screening expression libraries with antibodies and oligonucleotides, in *Molecular Cloning, A Laboratory Manual*, 2nd edition, (Nolan, C. ed.), Book 2. Chapters 8, 10, 11, and 12. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., et al. (1996) Life with 6000 genes. *Science* **274**, 563–567.
3. The *C. elegans* Sequencing Consortium. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018.
4. Adams, M. D., et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195.
5. The Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
6. Venter, J.C., et al. (2001) The Sequence of the Human Genome. *Science* **291**, 1304–1351.
7. Pennisi, E. (2000) Genomics. Rat genome off to an early start. *Science* **289**, 1267–1269.
8. Gibbons, A. (2000) Genomics. Building a case for sequencing the chimp. *Science* **289**, 1267.
9. Mastrangeli, R., Micangeli, E., and Donini, S. (1996) Cloning of murine LAG-3 by magnetic bead bound homologous probes and PCR (gene-capture PCR). *Analyt. Biochem.* **241**, 93–102.
10. Liang, P., Zhu, W., Zhang, X., Guo, Z., O'Connell, R. P., Averboukh, L., Wang, F., and Pardee, A. B. (1994). Differential display using one-base anchored oligo-dT primers. *Nucl. Acids Res.* **22**, 5763–5764.
11. Mastrangeli, R. and Donini, S. (1998) *Gene-Capture PCR in BioTechniques Books: Gene Cloning and Analysis by RT-PCR*, (Siebert, P.D., and Larrick, J. eds.), Eaton, Natick, MA. pp 271–288.
12. Frohman, M. A. (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Meth. Enzymol.* **218**, 340–356.
13. Edwards, J. B., Delort, J., and Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucl. Acids Res.* **19**, 5227–5232.
14. Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., et al. (1996) Suppression subtractive hybridization: a method for generating differentially

regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. USA* **93**, 6025–6030.

15. Chou, Q., Russell, M., Birch, D. E., Raymond, J., and Bloch, W. (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* **20**, 1717–1723.

Rapid and Nonradioactive Screening of Recombinant Libraries by PCR

Michael W. King

1. Introduction

The advent of the polymerase chain reaction (PCR) has greatly facilitated the isolation and characterization of clones from both cDNA and genomic libraries (1–3). Given the complexity of the genome of a particular organism or the relative abundance of a particular mRNA, within the cell type from which a cDNA library was constructed, the ability with which one can isolate a gene or cDNA clone of interest is quite variable. With respect to genomic libraries, the number of clones needed to be screened to isolate a single-copy sequence is a function of the complexity of the genome and the average size of the cloned fragments in the library. In the case of cDNA libraries, the frequency of a given clone of interest depends on the abundance of the messenger RNA. Highly abundant messages can represent 10% or more of total mRNA, whereas, very rare messages can be as low as one in 10^6 . In addition, the representation of some sequences in a cDNA library, particularly the 5' ends of large mRNAs, will be less than expected owing to the technical difficulties in converting the mRNA into full-length cDNA copies. In some cases, a particular sequence of interest can be depleted or lost at various steps of screening owing to its inefficiency to be replicated relative to other clones in the library.

Prior to the advent of the PCR, the principal technique for screening bacteriophage λ -based libraries involved screening nitrocellulose filters replicas with radioactively labeled probes. Recently, more highly sensitive methods for screening that utilize the PCR have been described (1–3). This chapter describes a PCR-based selection method for the isolation of clones from recombinant DNA libraries prepared in λ -based vectors. The technique uses no radioisotopes and can be completed in as few as 7 d. The technique is amenable to the use of highly specific PCR primers as well as degenerate primers designed to isolate families of related clones. In this method, a cDNA or genomic library is initially plated on as few as ten 100 mm plates at a density near 3000–5000 plaques per plate. The phage from each plate are soaked in SM buffer (4)

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

to generate the starting aliquots for PCR. Positive aliquots from the primary plating are identified and replated at lower and lower densities to generate subaliquots for secondary and tertiary screens until a positive clone is identified by PCR of phage soaked from a single plaque.

PCR-based library screening can be performed on both cDNA libraries and genomic libraries that are cloned in any variation of bacteriophage λ -based vectors. Screening can be performed from aliquots of unamplified newly packaged libraries or from plate lysate amplified libraries. It is helpful, yet not necessary, to know the titer (plaque forming units per milliliter; pfu/mL) of the library being screened. In general, good primary libraries have a titer of anywhere from 1,000,000–5,000,000 pfu in the entire initial packaging reaction (usually a 500- μ L library). Libraries that have been amplified generally have a titer of 10^{10} – 10^{11} pfu/mL. Unless it is known that the titer of a given library is uncharacteristically low, determining the precise titer of the starting aliquot is not necessary (*see Subheading 3.5.*).

2. Materials

1. Recombinant DNA library.
2. Oligonucleotide primers for PCR (*see Subheading 3.1.*).
3. Bacterial culture media, agar plates and top agar media.
4. SM Buffer (4): 100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl, pH 7.5.
5. Reagents for PCR.
6. Reagents for agarose gel electrophoresis.
7. CHCl₃.

3. Methods

3.1. Design of Primers

1. In designing primers for PCR one needs to consider several important factors such as making the primers with a near 50% GC content, a high degree of specificity with respect to nucleotide sequences and the absence of primer self-complementarity. The GenomeWeb site (<http://www.hgmp.mrc.ac.uk/GenomeWeb/nuc-primer.html>) maintained by the UK. HGMP Resource Center has links for several sites supplying primer design software tools to aid in the proper design of primers for PCR-based techniques. Primers that are from 18–21 nucleotides in length are optimal for this (and most) PCR-based technique because they allow high annealing temperatures that result in greater specificity in the reaction. In most cases, where possible, the primers should be designed to be maximally useful in a PCR with 2 mM MgCl₂ and an annealing temperature above 62°C. It has been shown that primer annealing is very specific over a wide range of MgCl₂ concentrations (from 0.5 mM to 2 mM). In addition, a higher degree of fidelity can be obtained, in some cases, by substitution of MgSO₄ for MgCl₂.
2. With the use of degenerate primer pairs, as for the isolation of families of related cDNAs or genes, it is extremely important that the 3'-nucleotide position of each primer set be nondegenerate in order to prevent an increase in nonspecific templating. In most cases, it is also optimal if the annealing temperature used with degenerate primers is higher than 55°C, although examples are available where successful isolation of related cDNAs has been carried out using degenerate primer pairs with annealing temperatures of 52°C.

3. For increased specificity it is possible to add a nested primer that resides within the sequences to be amplified with the primary primer pair. This nested primer can then be used in a second PCR, in combination with one of the primary primers, to test positive PCRs for the presence of correct internal sequences.

3.2. Basics of Phage Growth

1. Each step in this PCR-based screening protocol begins with the overnight culture of the appropriate host *Escherichia coli* (*E. coli*) (see **Table 1**). The cells should be grown at 37°C with agitation in NZCYM media supplemented with 1 mM MgSO₄ and 0.3% maltose to ensure optimal phage infection and growth (**4**). The minimum volume of the overnight culture of cells depends upon the number of plates that will be used. The standard volume of cells is 100 μL of bacteria for each 100 mm plate.
2. The next morning an aliquot of top agar is melted and held at 45°C. The appropriate number of agar plates are prewarmed at 37°C for 15–20 min.
3. The library or plate lysate aliquots are diluted appropriately in SM buffer, then added to an aliquot of the fresh overnight host *E. coli* that corresponds to the total volume needed for plating. Adding the phage to the entire volume of cells, instead of to individual aliquots, ensures an equal distribution of phage on each plate. The cells and phage are then incubated at 37°C for 10 min to allow the infection cycle to initiate.
4. The infected cells are then separated into individual aliquots for plating. The plating is carried out by adding melted 45°C top agar to each tube (3 mL for each 100 mm plate), pouring the solution onto individual plates and ensuring even spreading of the top agar.
5. Let the plates sit for 10 min at room temperature to harden the top agar. The plates are then inverted and incubated at 37°C for 6–8 h to allow plaques to form. (See **Subheadings 3.5–3.8**, for the screening protocol.)

3.3. Preparation for Screening

1. It is critically important that, prior to screening any library with this technique, the library be tested for the presence of clones that contain sequences that will amplify with a given primer pair.
2. Test a 1-μL aliquot (undiluted) of the library to be screened using the standard PCR protocol described in **Subheading 3.4**. As described in the Introduction, an amplified library will have approx 10⁷–10⁸ phage in a 1-μL aliquot. Therefore, a 1-μL aliquot will contain 100- to 1000-fold more phage than is statistically necessary to screen to find a clone of a given sequence. If the primers are unable to amplify the correct fragment from this amount of phage then the library is either devoid of clones or the primer pair is not functional as expected.

3.4. Standard PCR

1. All PCRs are performed in a volume of 25 μL containing 1X polymerase buffer, 2 mM MgCl₂ (or a concentration appropriate for a given primer pair), 10 pmole of each primer (approx 60 ng; assuming primers of 18–21 nucleotides) and 200 μM deoxynucleotide 5'-triphosphates (dNTPs). It is optimal to prepare a master mix (of sufficient volume for all the reactions) containing all components of the PCR except the template. This ensures equivalence of components in each individual reaction. It is important to note that an excess of any component of the PCR, in particular, an excess of primers, can lead to failure of the reaction or false positive results.

Table 1
Outline of Screening Protocol

Day	Activity	Identity
1	Start culture of host <i>E. coli</i>	
2	Plate 5–10 100-mm plates (approx 3000–5000 pfu/plate), grow 6–8 h, soak phage in SM overnight at 4°C	Primary screen
3	Process phage lysates, PCR screen	Primary screen
3	Start culture of host <i>E. coli</i>	
4	Plate 5 100 mm plates, of 1 or all 1° positives (approx 1000 pfu/plate), grow 6–8 h, soak phage in SM overnight at 4°C	Secondary screen
5	Process phage lysates, PCR screen	Secondary screen
5	Start culture of host <i>E. coli</i>	
6	Plate 5 100-mm plates, (approx 100–500 pfu/plate), grow 6–8 h, soak phage in SM overnight at 4°C, is possible to pick single plaques at this stage	Tertiary screen
7	Process phage lysates, PCR screen	Tertiary screen
7	Start culture of host <i>E. coli</i>	
8	Plate 1 100 mm plate (approx 50 pfu/plate), grow 6–8 h, pick single plaques into SM and elute overnight at 4°C	Final screen
9	Screen single plaque lysates by PCR	Final screen

- To ensure that the phage particles in the plate lysates are disrupted, the PCRs are “hot started.” This is accomplished by an initial denaturation at 95°C for 10 min. The reaction should then be held at 80°C whereaas adding polymerase. Many vendors sell *Taq* polymerase prebound by antibody. This prevents the enzyme from acting on any template until the antibody is denatured by the “hot start” conditions. Using these forms of polymerase, it is possible to add the enzyme directly to the master mix.
- The following cycle profiles have been shown to be optimal for the majority of specific primers (*see* **Notes** for degenerate primers):
30 cycles: 95°C, 0.5–1 min (denaturation); *n*°C, 0.5–2 min (annealing temperature defined by primer sequences, *see* **Note 1**); 72°C, *n* min (extension time depends on product length, for most reactions 0.5 min is sufficient; *see* **Note 2**).

3.5. Primary Screen

- For the primary screen it is usually necessary to plate 10 aliquots of the library at a density in the range of 3000–5000 pfu/plate. This density will nearly destroy the entire bacterial lawn. If the titer of the library is unknown, most amplified libraries can be plated using 1 µL of a 10³ dilution dispersed onto the 10 plates (i.e., the equivalent of 0.1 µL/plate). In some cases, it is possible to screen as few as 5 plates.
- Incubate the plates upside down at 37°C. Stop the incubation when the plaques begin to merge with one another. This usually takes 6–8 h at 37°C.

3. Overlay the plate with 4 mL of SM buffer and let stand at 4°C overnight. It is possible to incubate the plates with SM buffer for 2 h at 37°C or room temperature for 4–5 h. However, the PCR results are frequently smeary because of bacterial growth in the SM buffer. Also, the titer of the resultant lysates can be 100- to 1000-fold lower than the lysates prepared by 4°C overnight incubation. This latter fact is important to remember for subsequent screens.
4. Collect the SM buffer as a separate aliquot from each of the plates and remove agar and bacterial debris by centrifugation at 3000g for 10 min in a JA17 rotor (Beckman high-speed centrifuge) or an SS34 rotor (Sorval centrifuge).
5. Save 1 mL from each aliquot. Add CHCl₃ to 0.3% to prevent bacterial growth in these aliquots and allow for their longer term storage at 4 °C.
6. Use a 1-μL aliquot for the PCR assay (*see Subheading 3.4.*).
7. Analyze a 15-μL aliquot of each PCR on an agarose gel to determine which aliquot(s) have amplified the target of interest (*see Note 3*).
8. In most screens, there should be at least 1 plate lysate exhibiting a positive signal by PCR (*see Fig. 1*).
9. Provided the library was tested for the presence of DNA that can be amplified with the primer pair being used (*see Subheading 3.3.*), it will be possible to find primary lysate that contain positive signals. It may be necessary to continue to screen more primary plates until a positive is found.

3.6. Secondary Screen (*see Note 4*)

1. It is not necessary to titer the primary plate lysates although this can be done if accurate plaque numbers in the secondary screen are desired. However, based upon the fact that the primary plates should have experienced near complete lysis of the bacterial lawn, the secondary screens (using five 100 mm plates) are plated using 1 μL of a 10⁴ dilution (i.e., 0.2 μL/plate) of each positive primary lysate. Plating at this density (approx 500–1000 plaques/plate) is to ensure that the secondary lysates are dense enough to allow for enrichment of the clones of interest.
2. Allow plaques to grow as for the primary screen.
3. Prepare the phage lysates as for the primary screen using 3 mL of SM buffer.
4. Process the plate lysates and save a 1-mL aliquot of each with 0.3% CHCl₃ for long-term storage.
5. Screen a 1-μL aliquot of each in a 25-μL PCR as for the primary screen.
6. Analyze a 10-μL aliquot of each PCR by agarose gel electrophoresis (*see Fig. 1*).

3.7. Tertiary Screen (*see Note 5*)

1. The tertiary screen is the last plate lysate screen. It is possible, in some instances, to proceed directly to the screening of single plaques from the secondary screen. In some screens it may be necessary to perform an additional quaternary plate lysate screen prior to assay of single plaques (*see Notes 6 and 7*).
2. Plate 5 plates (100 mm plates) using 1 μL of a 10⁴ dilution of a positive secondary lysate (i.e., 0.2 μL/plate). In some cases it may be necessary to use a 10⁵ dilution, depending upon the density of plaques in the secondary screen.
3. Allow plaques to grow as for the primary screen.
4. Prepare the phage lysates as for the primary screen using 3 mL of SM buffer.
5. Process the plate lysates and save a 1-mL aliquot of each with 0.3% CHCl₃ for long-term storage.

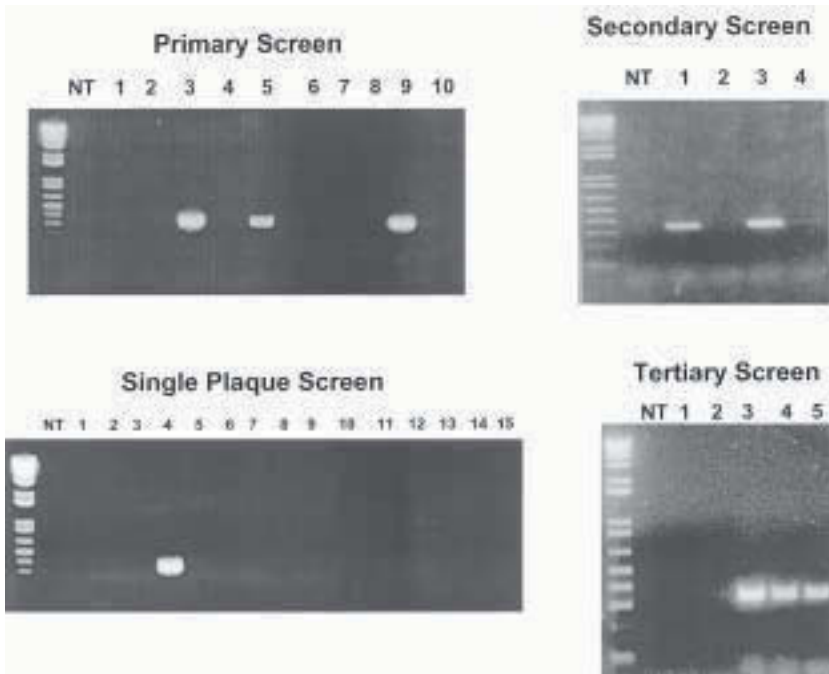


Fig. 1. Typical PCR-based library screen results. Ten plate lysates were assayed for the primary screen. Lysate #3 was used for the secondary screen. Lysate #3 from the secondary screen was used for the tertiary screen. Lysate #3 of the tertiary screen was used for the screening of single plaques, a total of 15 were assayed. A single plaque was found to be positive and used for phagemid rescue.

6. Screen a 1- μ L aliquot of each in a 25- μ L PCR as for the primary screen.
7. Analyze a 10- μ L aliquot of each PCR by agarose gel electrophoresis (see Fig. 1).

3.8. Single Plaque Screen

1. Plate a single 100 mm plate from a tertiary (or a secondary) positive using 1 μ L of 10^3 or 10^4 dilution.
2. The level of dilution used for the single plaque screening plate is less than for a tertiary screen and the amount of the dilution plated is 1 μ L. This is because the density of plaques in the tertiary screen should have been in the range of less than 250–500 pfu/plate such that the lysates from those plates will have a lower titer.
3. Pick individual plaques into 75–100 μ L of SM buffer using sterile glass Pasteur pipets to “scoop” the plaques out of the top agar. The volume of SM buffer used to soak single plaques depends on the size of the plaques themselves. Generally, if the plaques are similar in diameter to the diameter of a glass Pasteur pipet then one should use 100 μ L of SM buffer, less for smaller plaques.
4. Elute the phage particle from the plaque overnight at 4°C (see Notes 8 and 9).
5. At this step it is best to screen a 3- μ L aliquot in the standard 25 μ L PCR. Volumes less than 3 μ L can be used but have a tendency to give variable amplification (see Fig. 1).

4. Notes

1. The annealing temperature (T_m) used during the PCR, for any given primer, is determined from the base composition of the primer. To calculate the annealing temperature use the following formula: $T_m = 2(A+T) + 4(G+C)$. As an example for a 21mer primer with 12 G+C and 9 A+T the $T_m = 2(9) + 4(12) = 66^\circ\text{C}$. For the majority of PCRs it is optimal to use an annealing temperature that is 2–4°C below the calculated T_m . However, increased specificity is obtained by annealing at the T_m .
2. The elongation time used in the PCR is determined by the length of the resultant product. Given the rate of the majority of DNA polymerases at approx 1000 bases per second it would, in theory, be possible to use extremely short elongation times. However, in practice it has been observed that an elongation time of approx 1 min per 1000 bases is optimal.
3. To analyze PCR products by agarose gel electrophoresis one should use different percentages of agarose (dependent upon product size) in order to obtain good resolution of products. In general use 1.5–2% gels for products of less than 500 bp, 1–1.5% for products of 500–1000 bp and 0.7–1% for products greater than 1000 bp.
4. Following the primary screen, each positive plate lysate will most probably represent a different type of cDNA or gene clone unless degenerate primers were utilized. Therefore, to maximize the possibility of obtaining full-length cDNA clones or overlapping genomic clones each primary positive should be carried through to the secondary stage. However, to reduce the screening “load” one single primary positive at a time can be carried through to single plaque isolation.
5. When using specific primers, each positive secondary lysate represents the same single type of clone that was present in the primary lysate. Therefore, only one of the positive, secondary lysates is carried through to the tertiary screen. Also, only a single tertiary positive is carried through to the single plaque screen.
6. When using degenerate primer pairs, the complexity of possible clone types in any given primary positive can be large. For this reason it is necessary to plate at least 10 plates for the secondary screens. Because of this complexity of clones in a primary positive, it is best to carry only one primary positive at a time through to single plaque isolations. In addition, because each of the secondary positives will likely represent different types of clones, each of them needs to be screened in the tertiary screen. The latter fact is also likely in many cases at the level of the tertiary (and beyond) screens.
7. The use of degenerate primers in this screening technique will require at least four and possibly as many as six rounds of plate lysate screening prior to the screening of single plaques.
8. The single plaque lysate is used to prepare phage DNA as well as a permanent stock of the clone by small scale liquid lysis. Start an overnight culture of the appropriate host *E. coli* in NZCYM plus maltose and MgSO_4 (see **Subheading 3.2.**). The next morning add 25 μL of the 100 μL of SM (into which the positive plaque was eluted) to 100 μL of overnight cells and 100 μL NZCYM with maltose and MgSO_4 . Incubate with agitation at 37°C for 10 min. Transfer this culture to 50 mL of NZCYM without maltose or MgSO_4 . Incubate with agitation at 37°C until the cells in the culture begin to lyse. This takes approx 6 h and is visible as debris in the normally silky appearance of the growing *E. coli*. At this time add CHCl_3 to 0.5% and incubate an additional 10 min to accelerate the cell lysis as well as to prevent further growth of the cells. Centrifuge the cells and debris at 7K rpm for 10 min. Save an aliquot of the supernatant as a stock of the clone, either at 4°C or by adding dimethyl sulfoxide (DMSO) to 7% and storing at -80°C. To the remainder of the phage supernatant add RNase A and Dnase I to 1.5 $\mu\text{g}/\text{mL}$ and incubate at 37°C for

30 min. Precipitate the phage particles by addition of solid PEG 6000 to 10% (w/v) and solid NaCl to 0.5 *M*. Place at 4°C overnight. Collect the precipitate by centrifugation at 9k rpm for 25 min. Resuspend the precipitate in 0.5–1 mL of TE buffer (4). Add Proteinase K to 150 µg/mL and incubate at 45°C for 45 min. Extract the released phage DNA with an equal volume of phenol then phenol/CHCl₃ and again with CHCl₃. Precipitate the DNA by addition of 0.1 vol of 2 *M* ammonium acetate pH 5 and 2 vol of ethanol. The DNA should form a stringy precipitate immediately. It is best to remove the precipitating DNA by collecting it on a swirling glass rod. This reduced RNA and protein contamination that may affect restriction enzyme digestion. Centrifuge the precipitate at top speed in a microfuge, remove any supernatant and resuspend the pellet in 50–100 µL of TER buffer (4). Use from 3–10 µL for restriction enzyme digestion.

9. The single plaque lysate is used for the rescue of phagemid DNA, containing the cDNA clone, from λZap[®] (Stratagene, Inc.) or (ZipLox[®] (Life Technologies, Inc.) if the cDNA library was constructed with these modified forms of bacteriophage λ. The protocol for phagemid rescue from λZap is described. Start a culture of XL1-Blue[®] overnight in NZCYM plus 0.3% maltose and 1 mM MgSO₂. Also start a culture of SOLR[®] in LB media (4). The next morning dilute 40 µL of the overnight XL1-Blue cells into 1 mL of NZCYM plus 0.3% maltose and 1 mM MgSO₄. Incubate with agitation at 37°C for 60 min. Transfer 200 µL of these cells to a new tube, add 25 µL of the 100 µL SM (into which the positive plaque was eluted) and 1 µL of ExAssist[®] helper phage. Incubate with agitation at 37°C for 15 min. Add 3 mL of LB media (4) and incubate with agitation at 37°C for 2.5 h. Centrifuge the solution at 2000 rpm for 15 min. Transfer the supernatant to a new tube and heat at 70°C for 15 min. Centrifuge at 6000g for 15 min. Save the supernatant in a sterile tube as this is a stock of the rescued single-stranded phagemid. This solution can be stored at 4°C for up to 2 mo. To obtain colonies with the double-stranded phagemid, add 10–50 µL of the phagemid stock solution to 100 µL of the fresh overnight SOLR cells. Generally, 10 µL of the phagemid stock is more than enough to yield several hundred colonies when passaged through SOLR. Incubate with agitation at 37°C for 15 min. Spread 10–50 µL onto a single LB plus ampicillin plate and incubate overnight at 37°C. In most cases a single 10 µL aliquot is sufficient for plating since only a single colony is necessary. The double-stranded phagemid DNA can then be isolated from colonies by standard miniprep techniques (4).

References

1. King, M. W. (2000) Screening recombinant libraries by polymerase chain reaction. in *PCR Cloning Protocols*, in the series: *Methods in Molecular Biology*, 2nd ed., Humana Press, Totowa, NJ.
2. Amaravadi, L. and King, M. W. (1994) A rapid and efficient, non-radioactive method for screening recombinant DNA libraries. *BioTechniques* **16**, 98–103.
3. Isola, N. R., Harn, H. J., and Cooper, D. L. (1991) Screening recombinant DNA libraries: A rapid and efficient method for isolating cDNA clones utilizing the PCR. *BioTechniques* **11**, 580–582.
4. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1992) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 443.

Rapid cDNA Cloning by PCR Screening (RC-PCR)

Toru Takumi

1. Introduction

Now that the draft sequence of the human genome is available (1,2), cDNA cloning based on its sequence from a library is no longer an experimental goal, but a starting point and a routine laboratory practice. Hybridization screening with either radio-labeled or nonradio-labeled probes had been commonly used for cDNA cloning before, but it is laborious and time-consuming in the post-genome era. Application of the polymerase chain reaction (PCR) is surprisingly widespread since its discovery (3). Here an application to cDNA library screening is reported: rapid cloning of full-length cDNAs by screening pools of cDNAs by PCR (RC-PCR) (4). This PCR-based cDNA screening technique is applicable to both bacteria and phage libraries.

Providing that one has identified unique DNA sequences not only in a PCR clone or a partial-length cDNA clone, but also in ESTs or genomic sequences in databases, this RC-PCR screening technique is very simple and extremely rapid for isolation of a full-length cDNA from a library. This RC-PCR technique involves no radioisotopes, and avoids labor-intensive and expensive procedures such as transferring bacterial colonies to filters and hybridizing them to radiolabeled DNA probes. RC-PCR enables us to isolate a single clone from the library in a period of a few days.

2. Materials

1. cDNA library (*see Note 1*).
2. Reagents for PCR (store at -20°C):
 - a. PCR buffer: 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂.
 - b. dNTP mixture.
 - c. *Taq* DNA polymerase.
 - d. Sterile water.
3. Sequence specific primers.
4. 96-well dishes.
5. LB medium.
6. Agarose gel for analysis of PCR products.

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

3. Methods

3.1. Primer Design

1. The sequence of primers should be specific for the target sequence. In the case of cloning members of a gene family, choose the region that is less conserved among the family.
2. The parameters for designing primers are the length (basepairs), the GC/AT ratio and melting temperature (5).
3. As to the length of primers, a 17–18-mer is thought to be sufficient for RC-PCR if a positive control gives a single, specific band in an agarose gel. Longer primers are appropriate to amplify larger PCR products.
4. Amplification of 200–400 bp DNA is the most efficient (5).

3.2. PCR Screening

1. Divide the cDNA library into ten fractions, each of which contains, for example, approx 100,000 clones (*see Note 1*). Spread each of the library over 150 mm LB agar plates containing the appropriate antibiotics. After overnight growth, collect the colonies by a scraper and put them into microfuge tubes containing LB solution.
2. For each of the ten fractions, assemble a 50- μ L PCR in a 500- μ L microfuge tube containing the following:

10X PCR buffer	5.0 μ L
dNTP mix	200 μ M each
Primer	0.15 μ M each
DNA template	2.5 μ L of broth containing the colonies (<i>see Notes 2–4</i>)
sterile H ₂ O	to 49.5 μ L final volume
<i>Taq</i> DNA polymerase	0.5 μ L (2.5 u)
3. After a 5-min denaturation at 94°C, the cycling parameters are the following (*see Note 4*):

25 cycles	94°C	45 s
	60–65°C	1 min
	72°C	2 min
4. Run an agarose gel to detect the positive fractions.
5. For the second screening, dilute the positive pool of cells in LB broth to a concentration of about 30,000 clones per mL and distribute 100 μ L of this suspension (about 3000 clones) into each of the 96 wells of a microplate. Combine 10 μ L from each well in a column (giving a total of 120 μ L from each of the plate's 12 columns); similarly 10 μ L from each well in a row (giving 80 μ L from each of the 8 rows). From each of the resulting 20 mixtures, directly use 2.5 μ L as the DNA template of the next PCR.
6. Perform PCR as described in **steps 2 and 3** and analyze the results by gel electrophoresis.
7. Repeat screening and subdividing of the positive pools until a single clone is obtained (*see Note 5*).

4. Notes

1. As described earlier, this protocol is applicable to both bacteria and phage libraries. In the case of screening of phage libraries, a high phage titer is necessary for RC-PCR in the first screening. Once the titer of the positive pools drops to less than the order of 10³ in a serial screening, it is recommended to amplify the phages on a small plate (10 cm) before proceeding. For other protocols of the screening of the phage libraries by PCR, *see refs. (6–8)*.
2. Extraction of plasmid DNA is not necessary for PCR; intact bacterial colonies are sufficient for use as PCR templates (9,10). LB solution containing bacterial colonies can be

used directly as a template without any culture. In theory, it is not necessary to incubate the medium if it includes a positive clone. Incubation of the broth, however, may help to detect the positive signal especially for the initial screening. The number of colonies may determine whether the medium should be incubated or not.

3. The quality of cDNA library determines whether or not it is possible to isolate a full-length cDNA or a partial one. Use of primers that correspond to the relatively upper stream region of the cDNA may help increase the chances of isolating a full clone.
4. The PCR conditions should be optimized in each case by use of a positive control. Primer design, concentration of magnesium (Mg^{+}) and annealing conditions, including temperature, are important parameters. In the case of phage libraries, note that phage dilution buffer already contains Mg^{+} . In some cases, 10% of DMSO in the PCR buffer may help to produce a clean, specific PCR product. For a small fragment, it may be necessary to amplify more than 30 cycles (7). Careful precautions to avoid contamination are essential because of the extremely high sensitivity of PCR (11).
5. Here, an example of this rapid cDNA cloning by PCR screening (RC-PCR) from a bacteria library is illustrated. In an attempt to clone gene family members of transforming growth factor- β (TGF- β) receptors, a novel clone, termed clone B1, was isolated from GH₃ rat pituitary tumor cells by PCR using degenerate oligonucleotide primers (4,12). The method of RC-PCR was used to isolate a full-length B1 cDNA from the GH3 cell library (12).

First, a bacteria library of GH3 cells was divided into 16 pools, each of which included approx 100,000 clones. As primers for PCR, two oligonucleotides that were expected to be highly specific for the B1 PCR product and that were derived from the sequence of the reverse transcription-polymerase chain reaction (RT-PCR) product were utilized: ATC-GTGGTTCCGGGAGGCAGAGATC (25-mer, the sequence corresponding to nucleotides 726–750 of B1, *see Fig. 1A in ref. 12*) as a 5'-primer and CTGATTTGGAGCAATGTTCTATGGTG (25-mer, nucleotides 1095–1119 of B1, *see Fig. 1A in ref. 12*) as a 3'-primer. These primers correspond to a region of the kinase domain that is poorly conserved among members of the TGF- β receptor family.

The screening by PCR was done as follows: After a 5-min denaturation step at 94°C, the cycling parameters were 94°C for 45 s, 65°C for 1 min, and 72°C for 2 min, for a total of 25 cycles. Three pools of clones generated a PCR product of the expected size. One pool generated a slightly larger PCR product; subsequently this fraction was shown to contain a B1 clone with an insertion of 34 nucleotides in the kinase domain.

For the second screening, this pool of cells was diluted in LB broth to a concentration of about 30,000 clones per mL and 100 μ L aliquots of this suspension (about 3000 clones per aliquot) were distributed into 96-wells of a microplate. Ten microliters from each well in a column were combined (giving a total of 120 μ L from each of the plate's 12 columns); similarly 10 μ L from each well in a row were combined (giving 80 μ L from each of the 8 rows) (*see Fig. 1*). From each of the resulting 20 mixtures, 2.5 μ L were used directly in a PCR, and analyzed by gel electrophoresis. **Fig. 2** shows the results of the second screening; the PCR products were analyzed directly by electrophoresis through 2% agarose gel and ethidium bromide staining. Two of the twenty pools, that corresponding to column "9" and that corresponding to row "E," generated PCR products of the same size (394 bp) obtained from the original cloned B1 PCR product (lane P).

With a total of only 20 PCR reactions (12 columns and 8 rows) it was possible to identify a positive column and a positive row and thus a single positive pool (*see Fig. 1*). The positive pool was again diluted, this time to about 300 clones per mL, and subdivided

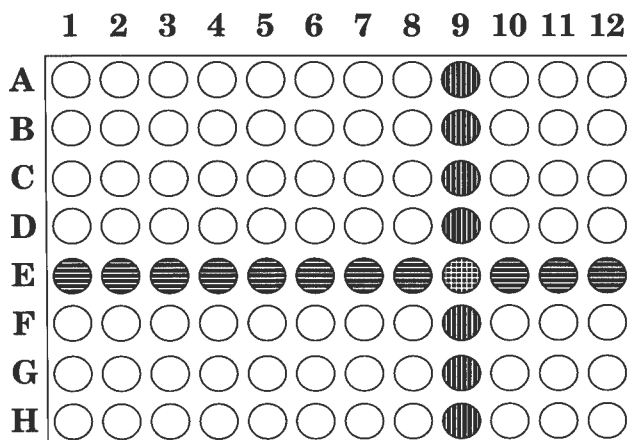


Fig. 1. Schematic representation of a 96-well microplate. The row and column were identified as described in the text.



Fig. 2. Analysis of the products of second screening by gel electrophoresis. An aliquot of each PCR reaction, one for each column (1-12) and one for each row (A-H) was analyzed directly by electrophoresis through 2% agarose gel with ethidium bromide staining. In the reaction in lanes P, the same PCR analysis was done on a plasmid containing the original B1 PCR product obtained as described in the text. Lanes M contain a 1-kb DNA ladder (Life Technologies, MD).

into 96-well plates (each well having 100 μ L or approx 30 clones). Again, column and row aliquots were pooled and PCR performed using each of the resulting 20 pools. The one positive pool identified in this subsequent screening was finally streaked onto several LB plates and, after overnight growth, 96 individual colonies were inoculated into 100 μ L of broth in wells of a microplate. As above, 20 PCR reactions were sufficient to identify a single positive clone. The cDNA isolated by this procedure contained 2276 bp and encodes a novel 505-amino-acid protein belonging to the family of type I receptor serine/threonine kinases (12,13).

Acknowledgment

The author would like to thank Paul Burke for his comments on this manuscript.

References

1. International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–919.
2. Venter, J. C. et al. (2001) The sequence of the human genome. *Science* **291**, 1304–1351.
3. Dieffenbach, C. W. and Dveksler, G. S. eds. (1995) *PCR PRIMER*, Cold Spring Harbor Laboratory Press, Plainview, NY.
4. Takumi, T. and Lodish, H. F. (1994) Rapid cDNA cloning by PCR screening. *BioTechniques* **17**, 443–444.
5. Rychlik, W. (1993) Selection of primers for polymerase chain reaction, in *PCR Protocols: Current Methods and Applications* (White, B. A., ed.), Methods Mol. Biol., vol. 15, Humana, Totowa, NJ, pp. 31–40.
6. Friedman, K. D., Rosen, N. L., Newman, P. J., and Montgomery, R. R. (1990) Screening of (gt11 libraries, in *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White T. J. eds.), Academic, San Diego, CA, pp. 253–258.
7. Yu, L. and Bloem, L. J. (1993) Use of polymerase chain reaction to screen phage libraries, in *PCR Protocols: Current Methods and Applications* (White, B. A., ed.), Methods Mol. Biol., vol. 15, Humana, Totowa, NJ, pp. 211–215.
8. Munroe, D. J., Loebbert, R., Bric, E., Whitton, T., Prawitt, D., Vu, D., Buckler, A., Winterpacht, A., Zabel, B., and Housman, D. E. (1995) Systematic screening of an arrayed cDNA library by PCR. *Proc. Natl. Acad. Sci. USA* **92**, 2209–2213.
9. Güssow, D. and Clackson, T. (1989) Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucl. Acids Res.* **17**, 4000.
10. Sandhu, G. S., Precup, J. W., and Kline, B. C. (1989) Rapid one-step characterization of recombinant vectors by direct analysis of transformed *Escherichia coli* colonies, *BioTechniques* **7**, 689–690.
11. Kwok, S. and Higuchi, R. (1989) Avoiding false positives with PCR. *Nature* **339**, 237–238.
12. Takumi, T., Moustakas, A., Lin, H. Y., and Lodish, H. F. (1995) Molecular characterization of a type I serine-threonine kinase receptor for TGF- β and activin in the rat pituitary tumor cell line GH3. *Exp. Cell Res.* **216**, 208–214.
13. Moustakas, A., Takumi, T., Lin, H. Y., and Lodish, H. F. (1995) GH3 pituitary tumor cells contain heteromeric type I and type II receptor complexes for transforming growth factor β and activin-A. *J. Biol. Chem.* **270**, 765–769.

Generation and PCR Screening of Bacteriophage λ Sublibraries Enriched for Rare Clones (the “Sublibrary Method”)

Michael Lardelli

1. Introduction

The simple sublibrary method described in this chapter allows the detection and rapid isolation of rare clones from bacteriophage λ libraries. The method is based on the ability of polymerase chain reaction (PCR) to detect clones present in a library at very low frequencies. Clones present at frequencies as low as one in 10,000,000, which would normally be impractical to isolate by conventional probe hybridization, can be rapidly isolated in this way (*1*). The method can also be used to isolate clones that are initially undetectable by PCR in λ libraries.

The sublibrary method can be divided into three steps (*see Fig. 1*):

1. 10^6 or 10^7 clones from a λ library are divided into pools of 25,000 or 250,000 clones, respectively, which are then amplified to form sublibraries.
2. The sublibraries are screened for the presence of the desired clone using PCR.
3. The desired clone is isolated from a positive sublibrary by conventional techniques (i.e., hybridization of a labeled probe to blots from a sublibrary that has been plated out).

The sublibrary method is useful in many different cloning strategies. For example, when a novel DNA sequence has been identified by PCR it is often desirable to isolate a corresponding λ clone for sequence analysis since the thermostable polymerases used in PCR have considerably greater in vitro error rates than those of bacteriophage DNA polymerases operating within their bacterial hosts (*2*). PCR screening of λ sublibraries not only allows rapid isolation of rare bacteriophage clones but also ensures that an isolated clone contains the entire region identified by the PCR. This avoids the necessity of isolating (by probe hybridization) multiple bacteriophage clones in order to be certain of covering the entire identified region. Other advantages of the sublibrary method are the savings in time and expensive materials (nylon/nitrocellulose disks and radioactive isotopes) that it permits. It is also possible to extend the

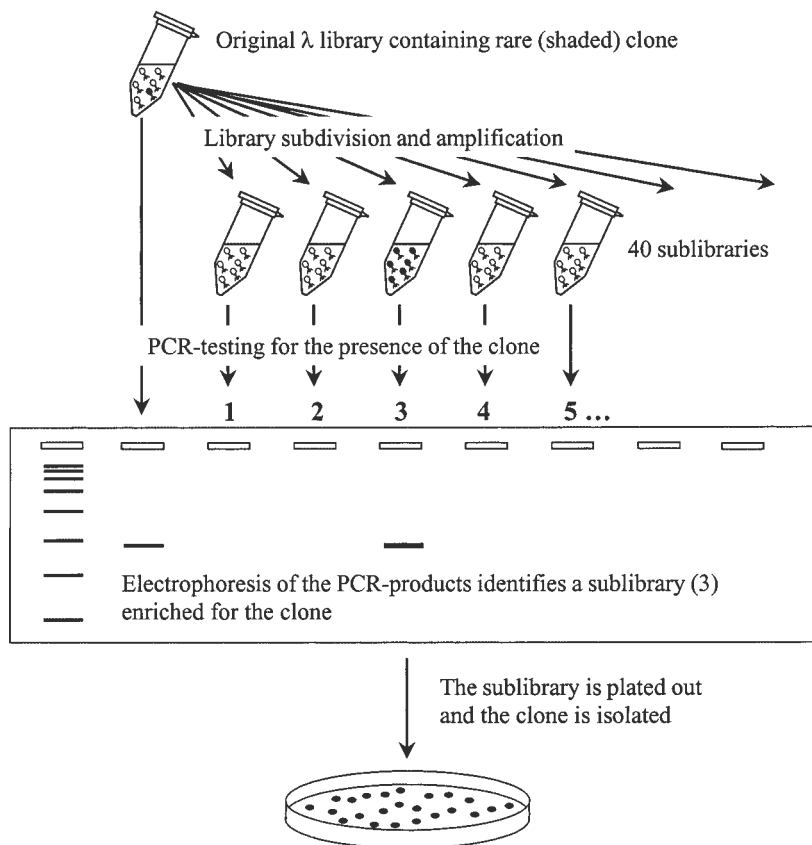


Fig. 1. The sublibrary method for detection and enrichment of rare clones in λ libraries.

method to the isolation of novel DNA sequences directly from λ libraries, for example, by use of degenerate PCR primers.

2. Materials

1. DyNAzyme™ II DNA Polymerase and supplied buffer (from Finnzymes Oy, Espoo, Finland).
2. Double-distilled water.
3. Oligonucleotide primers for PCR that amplify the DNA region of interest.
4. 10 mM dNTPs (2.5 mM each of dATP, dCTP, dGTP, and dTTP).
5. Paraffin oil.
6. PCR machine and accessories (recommended: a PCR machine that accepts 96-well microassay plates with paraffin oil, e.g., a MJ Research Inc. (Watertown, MA) PTC-200 Peltier Thermal Cycler with Multiplate™ 96 Polypropylene V-Bottom Microplates).
7. Micropipeters and pipet tips (e.g., Gilson, France) for liquid.
8. A DNA sample containing the sequence of interest (*see Note 2*).

9. A cDNA or genomic DNA library in a λ vector.
10. Centrifuge(s) capable of accepting microtiter plates, 50-mL tubes and 1.5-mL microfuge tubes.
11. 6X gel loading buffer: 15% w/v Ficoll (Type 400; Amersham Pharmacia Biotech AB, Uppsala, Sweden), 0.35% w/v Orange G (Sigma, St. Louis, MO), 60 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0.
12. Equipment and reagents for agarose gel electrophoresis.
13. DNA electrophoresis size markers (e.g., 1 Kb DNA Ladder; Life Technologies, Rockville, MD).
14. A sterile 100-mL conical flask for bacterial culture.
15. An *Escherichia coli* strain suitable for infection by, and growth of, bacteriophage λ , e.g., NM538 (3).
16. NZYCM liquid medium, NZYCM agar, and top agar. Prepare according to **ref. 4** by adding to 950 mL deionized water: 10 g of NZ amine, 5 g of NaCl, 5 g of bacto-yeast extract, 1 g of casamino acids, and 2 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$. Shake until solutes have dissolved. Adjust pH to 7.0 with 5N NaOH. Adjust the volume to 1 L with deionized water. For NZYCM agar add 15 g bacto-agar/L. For NZYCM top agar add 7 g bacto-agar/L. Autoclave for 20 min at 15 lb/in² on liquid cycle. Remelt the NZYCM agar and soft agar as required in a microwave oven on low power.
17. 20% maltose in distilled water (sterilized by filtration).
18. An oscillator incubator at 37°C.
19. A spectrophotometer and cuvetts with a 1-cm light path to measure light absorbance at 600 nm.
20. A 50-mL centrifuge tube.
21. 10 mM MgSO_4 sterilized by autoclaving.
22. SM solution: Prepare according to **ref. 4** by combining 5.8 g of NaCl, 2 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 50 mL of 1 M Tris-HCl, pH 7.5, and 5 mL of 2% gelatin solution. Adjust volume to 1 L with double-distilled water. Sterilize by autoclaving.
23. 10-mL pipets.
24. 40 capped glass test tubes.
25. Chloroform.
26. 1.5-mL capped, polypropylene microfuge tubes.
27. 10- and 15-cm diameter plastic Petri dishes (to be filled with NZYCM agar for plating out of bacteriophage).
28. An oven for incubation at 37°C.
29. 13-cm Nylon disks for bacteriophage blotting (e.g., from NEN(r) Life Science Products, Inc. Boston, MA).
30. Reagents for bacteriophage blotting, radioactive probe hybridization, and posthybridization washing (e.g., see **ref. 4**).
31. Needle (1–2-mm diameter).
32. DNA for probe manufacture (e.g., see **Note 15**).
33. Rediprime II DNA Labelling System from Amersham Pharmacia Biotech AB, Uppsala, Sweden and ($\alpha^{32}\text{P}$)-dCTP.
34. Adhesive tape.
35. Thick card (the same size as the X-ray film used for autoradiography).
36. X-ray film suitable for autoradiography with ^{32}P .
37. X-ray film developing solutions.

3. Methods

The sublibrary method described below is divided into five subheadings (3.1.–3.5.). First, the PCR for detection of the desired DNA region is tested. The PCR is then used to test λ libraries for the presence of clones containing the desired DNA region. A positive λ library is used to generate sublibraries. PCR is then used to test for sublibraries enriched for the desired clone. Finally, the desired clone is isolated from a positive sublibrary.

3.1. Testing the PCR Screen

1. Assemble a PCR premix and a polymerase premix as follows: (a) PCR premix: 2 μL of 10X Dynazyme II™ buffer (see **Note 1**), 13 μL of water, 1 μL of primer A (20 μM), 1 μL of primer B (20 μM), 2 μL of dNTPs (10 mM). (b) Polymerase premix: 0.5 μL of 10X Dynazyme™ buffer, 4.0 μL of water, and 0.5 μL of Dynazyme II (2 U/ μL) (see **Note 1**).
2. Prepare a DNA sample for PCR:
 - a. Place one drop of mineral oil in the well of a microassay plate (or PCR tube). Then pipet 1–2 μL of a suitable DNA sample (see **Note 2**) under the oil. Note that preparation of a negative (no DNA) control is essential (see **Note 3**).
 - b. (Necessary only when the DNA sample is a bacteriophage library.) Place the microassay plate (or PCR tube) in the PCR cyclor and heat to 94°C for 2 min to denature bacteriophage capsids and inactivate any DNases.
3. Add the PCR mix and centrifuge the microassay plate (or PCR tube) for a couple of seconds at 1000 rpm (approx 170g).
4. Place the microassay plate (or PCR tube) in the PCR cyclor and heat to 94°C (e.g., “pause” the PCR cyclor during the first denaturation step of the PCR cycling protocol given in **step 5**). Wait 30 s and then eject the polymerase premix into the well from just above the top of the oil (e.g., onto the wall of the plate well or tube). The polymerase mix will fall through the oil and mix with the other PCR components.
5. Perform PCR cycling as follows: 35 cycles: denaturation at 94°C for 1 min, annealing at a suitable temperature for 1 min (see **Note 4**), temperature ramp of 0.5°C/s to 72°C, and elongation at 72°C for a suitable time (see **Note 5**). Then hold the PCRs at 4°C (or freeze them) until analysis.
6. Add 5 μL of 6X gel loading buffer to the top of the oil and centrifuge the plate or tube as in **step 3**.
7. Electrophorese 20 μL of the PCR on an agarose gel of suitable concentration and using suitable size markers (see **Note 6**). When loading the PCR onto the gel, draw the solution up and down in the pipet a couple of times to mix it thoroughly with the loading buffer. Avoid drawing oil into the tip.
8. If a specific PCR fragment of the expected size is revealed by the electrophoresis, then the PCR screening method probably works.

3.2. PCR-Testing of λ Libraries

1. Test the λ library for the presence of a clone containing the desired DNA sequence by performing PCR on 1–2 μL of undiluted λ library as described in **steps 1–7** of **Subheading 3.1**. Always perform a negative (no DNA) control and a positive control using the known DNA (from **Subheading 3.1**) to give a PCR signal (see **Note 7**).
2. If the PCR detects a desired clone in the λ library then proceed to the manufacture of sublibraries as described under **Subheading 3.3**. (see **Note 8**).

3.3. Generation of λ Sublibraries

1. Grow a 50-mL culture of *E. coli* strain NM538 overnight at 37°C in NZYCM medium plus 0.2% maltose with vigorous shaking (*see Note 9*).
2. Measure the optical density (OD) of the culture relative to sterile NZYCM medium at 600 nm with a light path of 1 cm. The OD of an overnight culture of NM538 in NZYCM plus 0.2% maltose is usually between 3 and 5.
3. Centrifuge the culture in a 50-mL centrifuge tube at approx 6000g for 10 min. Pour off the supernatant and resuspend the cells in 10 mL of 10 mM MgSO₄. The cells can be stored for up to 2 d at 4°C (*see Note 10*).
4. Dilute the cells in NZYCM medium (prewarmed to 37°C) to a final volume of 40 mL so that the OD is calculated to equal 0.2.
5. Immediately infect the diluted culture with 10⁶ or 10⁷ plaque forming units (pfu) from the λ library (dilute the λ library as required in SM solution; *see Note 11*). Mix the bacteriophage briefly and gently by inversion.
6. Incubate the infected culture at 37°C (without agitation) for 10 min to allow bacteriophage absorption by the cells and then rapidly divide the culture into 40X 1-mL aliquots in numbered, capped glass test tubes. Incubate the tubes at 37°C with vigorous shaking.
7. When cell lysis occurs (usually after 3–5 h of incubation, *see Note 12*) add three drops of chloroform to each tube and continue vigorous shaking for 5 min. The tubes now contain λ sublibraries. Store them (for a limited period; *see Note 13*) capped at 4°C. (Alternatively, store an aliquot of each sublibrary at –70°C and prepare DNA for PCR testing using the remaining volume; *see Note 13*).

3.4. PCR Screening of the Sublibraries

Test the λ sublibraries for the presence of a clone containing the desired DNA sequence by performing PCR on 1–2 μ L of each λ sublibrary as described in **steps 1–7 of Subheading 3.1**. (Alternatively, test DNA prepared from each sublibrary; *see Note 13*). A positive-control PCR using 1–2 μ L of the original λ library and a negative (no DNA) control should also be performed for comparison with the sublibrary PCRs (*see Note 14*).

3.5. Isolation of the Desired Bacteriophage Clone from a Sublibrary

1. Determine the titer of a positive sublibrary by infecting 100- μ L aliquots of NM538 at OD = 2.0 in 10 mM MgSO₄ with 10 μ L aliquots of 10⁻⁵, 10⁻⁷, and 10⁻⁹ dilutions of the positive sublibrary in SM solution. After absorption of the bacteriophage onto the bacteria for 10 min at 37°C, add 3 mL of NZYCM top agar plus 0.2% maltose (melt the NZYCM top agar, add the maltose, and then hold at 42°C), mix gently but rapidly, and pour onto a 10-cm Petri dish of NZYCM agar that has been slightly dried and prewarmed to 37°C. Rock the plate briefly to ensure even spreading of the top agar. When the top agar has set, invert the plate and incubate overnight at 37°C. Count the bacteriophage plaques produced and calculate the sublibrary titer using the following formula:

$$\text{Titer in pfu/mL} = (100 \times \text{No. of plaques}) / \text{Sublibrary dilution} \quad (1)$$

Proceed immediately to **step 2** (*see Note 13*).

2. If the positive sublibrary was generated from 25,000 pfu (i.e., the initial 40-mL culture was infected with 10⁶ pfu) then plate out 100,000 pfu onto two 15-cm diameter plates of NZYCM agar (i.e., 50,000 pfu/plate). If the positive sublibrary was generated from

250,000 pfu (i.e., the initial 40-mL culture was infected with 10^7 pfu) then plate out 10^6 pfu on 20 15-cm plates or start a new round of sublibrary generation using 10^6 pfu from the positive sublibrary (i.e., generate “subsublibraries”). To plate out 50,000 pfu on a 15-cm plate, dilute an aliquot of the positive sublibrary to a concentration of 5×10^6 pfu/mL in SM solution and then add 10 μ L of this to 300 μ L of NM538 at OD = 2.0 in 10 mM MgSO_4 . Incubate 10 min at 37°C and then add 7 mL of NZYCM top agar plus 0.2% maltose (melted, then held at 42°C). Mix gently but rapidly and pour onto a 15-cm NZYCM agar plate slightly dried and prewarmed to 42°C. Ensure that the top agar covers the entire surface of the plate evenly by tilting it as necessary. When the top agar has set, invert the plate and incubate at 37°C until the plaques are the size of small pin-pricks. Cool the plates to 4°C.

3. Blot the plated-out bacteriophage onto 13-cm nylon or nitrocellulose disks. Remember to stab the disks with a needle when they are lying on the plates so that the disks can be reoriented to the plate after hybridization and autoradiography. Hybridize the disks with a radiolabeled probe that detects the desired clone (see **Note 15**). Tape the washed and dried disks to card and tape the card to X-ray film for autoradiography. Remember to stab through the card and film with a pin so that the autoradiograph can be reoriented to the blots after development.
4. After autoradiography of the blots, pick out areas of the plates containing positive clones using a 1-mL micropipet with a cutoff tip. Place these agar samples in individual 1.5-mL microfuge tubes containing 0.5 mL of SM solution and three drops of chloroform. Use a fresh pipet tip to break up the agar piece and invert the tube a number of times. Place tube overnight at 4°C. Repeat this procedure for an area of a plate lacking a positive signal (for use as a negative control in the following PCR).
5. Invert the tube again a couple of times and then centrifuge for 10 s at top speed in a microcentrifuge. Remove a 1- μ L sample and perform PCR on this according to **steps 1–7 of Subheading 3.1**. A positive signal (properly controlled) indicates that the picked plate area contains a clone of the desired type (see **Note 16**).
6. Titrate the bacteriophage solution from the picked area, as in **step 1**, using undiluted, 10^{-2} and 10^{-4} dilutions of the solution. Plate out 1000 pfu on two 15-cm NZYCM plates (see **Note 17**) and allow plaques to grow to a diameter of 2 mm before repeating **steps 3 and 4**. Positive clones should give a strong signal on the autoradiograph and do not need to be checked by PCR. If individual positive clones (plaques) cannot be picked, then pick an area containing a positive clone and repeat this step after plating out 100 pfu.
7. Amplify the isolated clone by infecting 10 mL of NM538 cells at OD = 0.2 in NZYCM medium (prepared as described in **steps 1–4 of Subheading 3.3.**) with 250,000 pfu followed by incubation at 37°C until lysis or by elution of plated-out bacteriophage in SM solution (see **Note 18**). Store at 4°C with added chloroform.

4. Notes

1. PCR conducted on samples from λ libraries can be problematic because media used for bacteriophage λ growth contain significant concentrations of Mg^{2+} ions (that are required to prevent disassociation of the capsids). For example, standard NZYCM medium contains 8 mM MgSO_4 (4). The author has found that use of DynazymeTM (a thermostable polymerase purified from *Thermus brockianus* by Finnzymes Oy) with the buffer supplied by the manufacturer allows PCR to be performed successfully without consideration for variations in Mg^{2+} concentration in most cases. Other thermostable polymerases can be used in this procedure, but these may be sensitive to small variations in Mg^{2+}

concentration. One solution to the problem is to adjust the Mg^{2+} concentration in the PCR buffer so that addition of an aliquot from a λ library gives the desired concentration. If the titer of the library is sufficiently high, it is also possible to make a 1:10 dilution of the library, thus reducing the Mg^{2+} concentration while still providing enough bacteriophage for detection by PCR.

2. Before beginning the sublibrary method, the PCR must be tested for its ability to detect the desired sequence. The test substrate will depend on the type of library to be screened. Primers for screening cDNA libraries can be tested on single-stranded cDNA produced from whole cell mRNA or on a suitable dilution of a preexisting plasmid clone of the cDNA sequence (for example, a clone of a PCR fragment). When screening genomic λ libraries, the PCR primers should be tested on diluted genomic DNA. During PCR-testing of the λ library, 1–2 μ L of the library is used in a reaction. If the library has a titer of 10^{10} pfu/mL (10^7 pfu/ μ L) and contains the desired clone at a frequency of between 10^{-6} and 10^{-7} (i.e., between 10 copies and 1 copy/ μ L), then this corresponds to a concentration of the desired clone in the library of between 1.7×10^{-18} and 1.7×10^{-19} M, respectively, and in the PCR reaction of between 6.8×10^{-20} and 6.8×10^{-21} M. Ideally, the PCR should detect the desired DNA sequence at this concentration. Higher titers in the original library allow the use of less-efficient PCR. When ejecting samples under oil, check that the sample volume has actually been placed in the oil and has not clung to the outside of the micropipet tip.
3. The negative control is prepared and handled in exactly the same way as the other PCRs except that no DNA sample is added. To avoid cross-contamination with the other PCRs when testing a PCR screen, add PCR premix and polymerase premix to the negative control first.
4. The annealing temperature for the PCR will depend on the primers used. For standard 18-mer PCR primers with a 50% GC content, try annealing first at 50°C. If this is not successful, then test annealing temperatures in the range 40–60°C.
5. The elongation time depends on the progressivity of the polymerase. The author usually assumes a progressivity of 1 kb/min and sets the elongation time at twice the time estimated for synthesis of the desired DNA fragment. Thus, elongation time in minutes = $2 \times$ fragment length in kb.
6. Size markers should allow identification of the desired PCR product. DNA Ladder markers from Life Technologies allow identification of PCR fragments with a size of 200 bp–12 kb.
7. A 4- μ L library sample may aid in the detection of very rare clones (**I**), but in this case the reaction volume should be increased to 50–100 μ L. Purification of DNA from each sublibrary for use in PCR may also assist in detection of very rare clones—see **Note 13**.
8. If PCR on a sample of the λ library does not produce an identifiable PCR product it may still be worth proceeding with the sublibrary procedure starting with 10^7 clones. A suitable clone present in the original library at very low concentration may be sufficiently amplified in one of the sublibraries to be detected by the PCR.
9. The author has only used *E. coli* strain NM538 for generation of sublibraries. Other strains suitable for λ infection should also work but the initial OD of the culture that is infected for sublibrary generation may need to be adjusted slightly. Addition of maltose to the NZYCM medium induces synthesis of the maltose transporter protein to which bacteriophage λ binds when infecting a cell.
10. The sublibrary procedure as originally described (**I**) used maltose in both the overnight bacterial culture and after bacteriophage infection. However, the author has since found

that higher sublibrary titers are obtained if the infected bacteria are cultured in medium lacking maltose. This may be because of the presence of fewer maltose receptors in the debris resulting from cell lysis that can absorb the bacteriophage and/or because a slower rate of infection allows higher cell numbers to form before mass lysis of cells occurs. Dead cells in the initial cell culture will also needlessly absorb bacteriophage so only 1–2-d-old cultures should be used. NM538 cells resuspended in 10 mM MgSO₄ and stored at 4°C can be used for up to a week for plating-out of bacteriophage.

11. In order to infect the bacteria with the correct amount of bacteriophage, the library should be titrated a short time before use using the method described in **step 1** of **Subheading 3.5**.
12. If cell lysis fails to occur, then the infection can be modified by decreasing the OD of the culture to be infected (e.g., from 0.2 to 0.1) and/or by adding maltose to the culture medium (*see Note 10*). In general, lysis occurs later with infections from λ libraries of genomic DNA relative to those from cDNA libraries.
13. Unfortunately, despite their high titers, the sublibraries may be unstable. It is therefore desirable to proceed with PCR screening and bacteriophage isolation within a few days of their manufacture (except see below). Although PCR signals do not appear to be greatly affected by a decrease in a sublibrary's infectivity, the titer of a sublibrary should always be measured immediately before plating it out for clone isolation. Instability of sublibraries is not a great problem because generation of sublibraries is so simple and rapid that it is easily performed each time one wishes to isolate a new clone. If the sublibraries are stored in microfuge tubes, then ensure that three drops of chloroform are always present in each tube. To remove some cell debris, the sublibraries may also be centrifuged (10 min at 16,000 rpm) and transferred to new microfuge tubes with chloroform. Increased longevity of the sublibraries can be obtained by addition of 40 μ L of dimethyl sulfoxide (DMSO, Sigma) to 500 μ L of each sublibrary and then freezing these aliquots at -70°C . The remaining 500 μ L of each sublibrary can be used for DNA purification, i.e., add 10 μ L of 20 mg/mL proteinase K to each sublibrary, incubate at 37°C overnight, then add 10 μ L of 5 M NaCl, extract with 200 μ L of a 1:1 mixture of phenol and chloroform (that has been equilibrated against 1 M Tris-HCl pH 8.0 and then stored under 10 mM Tris-HCl pH 7.5, 1 mM EDTA) before precipitation with ethanol, drying and redissolution in 50–200 μ L of 10 mM Tris-HCl pH 7.5, 1 mM EDTA. The DNA preparation is then stored at -20°C and can be used for screening of the sublibraries by PCR with increased sensitivity. The titers of the sublibraries at -70°C decrease more slowly than otherwise but should, nevertheless, be redetermined before plating out of a sublibrary for clone isolation.
14. To control for cross contamination resulting from the pipeting technique when setting up the sublibrary PCRs, add PCR premix to the sublibrary samples first, then to the positive control, and finally to the negative control. Use the same order of addition for the polymerase mix.
15. Detailed descriptions of bacteriophage blotting and hybridization techniques are provided by **ref. 4**. A suitable probe can be synthesized by excising the PCR product of a positive control from a low gelling temperature agarose gel after electrophoresis in Tris-acetate buffer (*see ref. 4*) and then using this in a randomly-primed DNA synthesis reaction with the Rediprime II DNA Labelling System from Amersham Pharmacia Biotech AB; see manufacturer's instructions.
16. This step ensures that clones detected by probe hybridization contain the entire region detected by the PCR. Clones containing partially overlapping regions of DNA and false positives resulting from artifactual spots on the autoradiograph are thus excluded.

17. To save time, the titration can be performed in duplicate on 15-cm plates. A pair of plates with approx 500 pfu each can then be used directly for blotting and clone isolation.
18. Detailed descriptions of all the techniques for handling bacteriophage clones can be found in **ref. 4**.

References

1. Lardelli, M. and Lendahl, U. (1994) Generating bacteriophage λ sublibraries enriched for rare clones. *BioTechniques* **16**, 420–422.
2. Mattila, P., Gelfand, D. H., Sninsky, J. J., and White, T. J. (1991) Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase—an extremely heat stable enzyme with proofreading activity. *Nucl. Acids Res.* **19**, 4967–4973.
3. Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. (1983) Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170**, 827–842.
4. Sambrook, J. and Russel, D. (2001) *Molecular Cloning: A Laboratory Manual* (Third Edition) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

PCR-Based Screening for Bacterial Artificial Chromosome Libraries

Yuji Yasukochi

1. Introduction

The bacterial artificial chromosome (BAC) (1) is now the most widely utilized vector system to construct large-insert genomic libraries for genome analysis. It has the following advantages over another vector system, the yeast artificial chromosome (YAC) (2): ease of DNA purification, a lower rate of chimerism, stable maintenance of inserts, and compatibility with templates for sequence determination. Regarding screening, it is very important that BAC-DNAs can be separated from genomic DNA of the host, as the host DNA often competes with BAC-DNAs for the amplification of expected PCR fragments.

Screening for BAC libraries is essentially different from the conventional screening, in which only positive clones are picked up and the rest are discarded. In contrast, the same set of BAC clones (BACs) must be used for multiple times. It is very difficult to gain a large number of BACs compared with commonly used vector systems, because preparation of high-molecular-weight insert DNA is complicated and transformation efficiency of BAC is quite low. Therefore, it is necessary to develop screening techniques which enable sustainable utilization of limited number of BACs.

Two methods are generally used to isolate BACs, PCR-based screening and colony hybridization using high-density replica (HDR) filters, a kind of colony blot on which BACs are dotted at high-density (i.e., 3072 clones per the area of a standard microtiter plate). Colony hybridization is effective when genome structure of a target is unknown and unlikely to include multi-copy sequences, i.e., cDNA and EST (expressed sequence tag) clones. It is also a powerful tool to analyze organisms of relatively small genome size or restricted chromosomal region of specific interest. However, there are several difficulties to perform colony hybridization in a large scale.

First, colony hybridization requires skilled technicians, which limits the scale of the experiment. In contrast, a less experienced person can perform PCR screening and it is not so difficult to operate multiple thermal cyclers by one person. Second, HDR

filters must be prepared repeatedly because they can be used for 20–30 hybridizations at most. Thousands of PCR screening can be performed without additional maintenance after BAC-DNAs are once prepared and pooled. Third, sequences containing multicopy or highly conserved sequences cannot be used as a probe, which can be used in PCR screening as far as such regions are carefully excluded from primer sequences. Finally, the area of HDR filters required for screening increases in proportion to the genome size of analyzed organisms, which makes this technique more laborious. In PCR-based screening, larger genome size only affects the first-step screening, which is just a small portion of the procedure. As described above, PCR-based screening is suitable for a wide range of experiments, especially for extensive analysis such as contig construction around the whole genome of higher eukaryotes.

PCR-based screening is composed of two steps (3). The first-step screen is performed on primary DNA pools, which represents BACs of several 96-well microtiter plates (superpool), and the second-step “three-dimensional” screen is performed on secondary DNA pools representing each row, column and plate prepared for each superpool. Preparation of DNA pools is in reverse order, that is, primary DNA pools are prepared from the mixture of secondary DNA pools. Pooling bacterial cell cultures and DNA solutions is the most critical process for accuracy of this method. Therefore, pooling process must be performed in 96-well format using 96-well plates and 8- or 12-channel pipets, if manually performed. Utilization of 384-well plates is not a good idea because 16- or 24-channel pipets fully adapted for 384-well plates are not available at present.

The first thing to determine is the number of plates in one superpool. Total of rows, columns and plates is preferable to be a divisor of 96 for efficient utilization of 96-well plates in the second-step screen. Since the sum of rows and columns is fixed to 20, candidates of the number of plates are 4, 12, and 28, representing 384, 1152, and 2688 clones, respectively.

It is not desirable that more than two positive clones are frequently found in one superpool because an additional step (*see Subheading 3.5.*) is needed to identify candidate clones. The probability of the existence of the positive clones in one superpool is calculated by the following formula: $p = xCn \times (a/g)^n \times (1-a/g)^{x-n}$, where a is the average insert size; g is the haploid genome size; n is the number of the positive clones; p is the probability; x is the number of clones in one superpool. **Table 1** shows the estimated probabilities when the average insert size of a BAC library is 150 kb. In general, the number of plates should be determined as the probability of detecting more than two positive clones in one superpool is less than 20%. Protocols described below are based on the case of 12 plates.

However, in case of organisms whose genome size is very small (<100 Mb), it is an alternative choice that n is set to be larger and the first screening is omitted. On the other hand, in case of organisms of larger genome size (>2000 Mb), too many primary DNA pools must be screened in the first screening. To reduce the number of superpools, the number of plates in one superpools must be increased. For this purpose, it is effective to determine positive plates “two-dimensionally” (4). For example, if 196 plates are arrayed in 14×14 grid and DNA pools from 14 columns and 14 rows

Table 1
The Probabilities (%) of the Existence of the Positive Clones

Haploid Genome Size (Mb)	4 Plates / superpool				12 Plates / superpool No. of positive clones				28 Plates / superpool			
	0	1	2	3 \leq	0	1	2	3 \leq	0	1	2	3 \leq
50	31.5	36.5	21.0	11.0	3.1	10.9	25.3	60.7	0.0	0.3	1.0	98.7
100	56.2	32.4	9.3	2.1	17.7	30.7	30.8	20.8	1.8	7.1	14.4	76.7
200	75.0	21.6	3.1	0.3	42.1	36.4	16.9	4.6	13.3	26.8	27.0	32.8
300	82.5	15.9	1.5	0.1	56.2	32.4	9.8	0.6	26.1	35.0	23.5	15.4
500	89.1	10.3	0.5	0.1	70.8	24.5	4.4	0.3	44.6	36.0	14.5	4.9
750	92.6	7.1	0.0	0.0	79.4	18.3	2.1	0.2	58.4	31.4	8.4	1.8
1000	94.4	5.4	0.0	0.0	84.1	14.5	1.3	0.1	66.8	26.9	5.4	0.9
1500	96.2	3.7	0.0	0.0	89.1	10.3	0.6	0.0	76.4	20.5	2.8	0.3
2000	97.2	2.8	0.0	0.0	91.7	7.9	0.3	0.0	81.7	16.5	1.6	0.2
3000	98.1	1.9	0.0	0.0	94.4	5.4	0.2	0.0	87.4	11.7	0.7	0.2
4000	98.6	1.4	0.0	0.0	95.8	4.1	0.0	0.0	90.4	9.1	0.4	0.1
5000	98.9	1.1	0.0	0.0	96.6	3.3	0.0	0.0	92.3	7.4	0.3	0.0

On the assumption that average insert size is 150 kb.

are used instead of 28 individual plates, total of insert size ($8 \times 12 \times 28$ format) increases by seven times.

Secondary DNA pools can be directly prepared from mixtures of cell cultures pooling for each row, column, and plate of the superpool. However, it is strongly recommended to prepare row and column pools for each plate. It enables to identify positive clones by PCR of candidate pools without picking up all potential positive clones (i.e., 8 clones for 2 positives and 27 clones for 3 positives in one superpool) from the stock (*see Subheading 3.5.*). It is also effective to identify positive clones when no signal is detected from any of row, column and plate pools (*see Subheading 3.5.*). DNA preparation systems for a large amount of samples have drastically been improved owing to the large-scale genome sequencing. Also, the increase in labor and cost of BAC-DNA preparation is now acceptable to consider the benefit.

2. Materials

1. 8- and 12-channel pipets.
2. 96-Well plates (standard and deepwell type).
3. 96-Well plates for PCR.
4. Plasmid isolator or 96-well format plasmid purification kits (i.e., QIAwell 96 Ultra Plasmid Kit, Qiagen).
5. Thermal cycler adapted for 96-well plates.
6. Gel electrophoresis equipment.
7. LB freezing buffer (4): 36 mM K_2HPO_4 , 13.2 mM KH_2PO_4 , 1.7 mM citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)_2SO_4$, 7.5% v/v glycerol, an antibiotic for selection, in LB medium.
8. TE buffer: 10 mM Tris-HCl, pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA).
9. Binary gel (*see Note 1*): 0.7% agarose, standard grade, 0.7% Synergel (Diversified Biotech, Boston, MA), in 0.5X TBE buffer.
10. Loading dye: 30% v/v glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol.

3. Methods

This procedure is a modified version of that described by Wu et al. (6).

3.1. Preparation of DNA Pools for Each Plate

1. Pick up white colonies harboring BACs into 96-deep well plates, each well containing 0.6 mL of LB freezing buffer, and incubate at 37°C for more than 16 h with vigorous shaking (*see Note 2*).
2. Transfer 100 μ L each overnight culture into triplicated 96-well plates with a 8- or 12-channel pipet and store at -80°C . The rest of the culture is used to prepare DNA pools.
3. Assemble 100 μ L each overnight culture in the same row of the same plate into a column of 96-deep well plates with a 8-channel pipet (*see Fig. 1*).
4. Transfer the resulting mixture (800 μ L) into tubes for an automatic plasmid isolator with a Pasteur pipet (*see Note 3*). Isolate BAC DNAs as recommended by the manufacture and dissolved BAC-DNA in 200 μ L TE buffer (*see Note 4*). Transfer into 96-well plates (basic row DNA pools, **Fig. 1**).
5. Assemble 100 μ L each overnight culture in the same column of the same plate into a row of 96-deepwell plates with a 12-channel pipet (*see Fig. 1*). Isolate BACs in the same way as described above and transfer into 96-well plates (basic column DNA pools, **Fig. 1**).

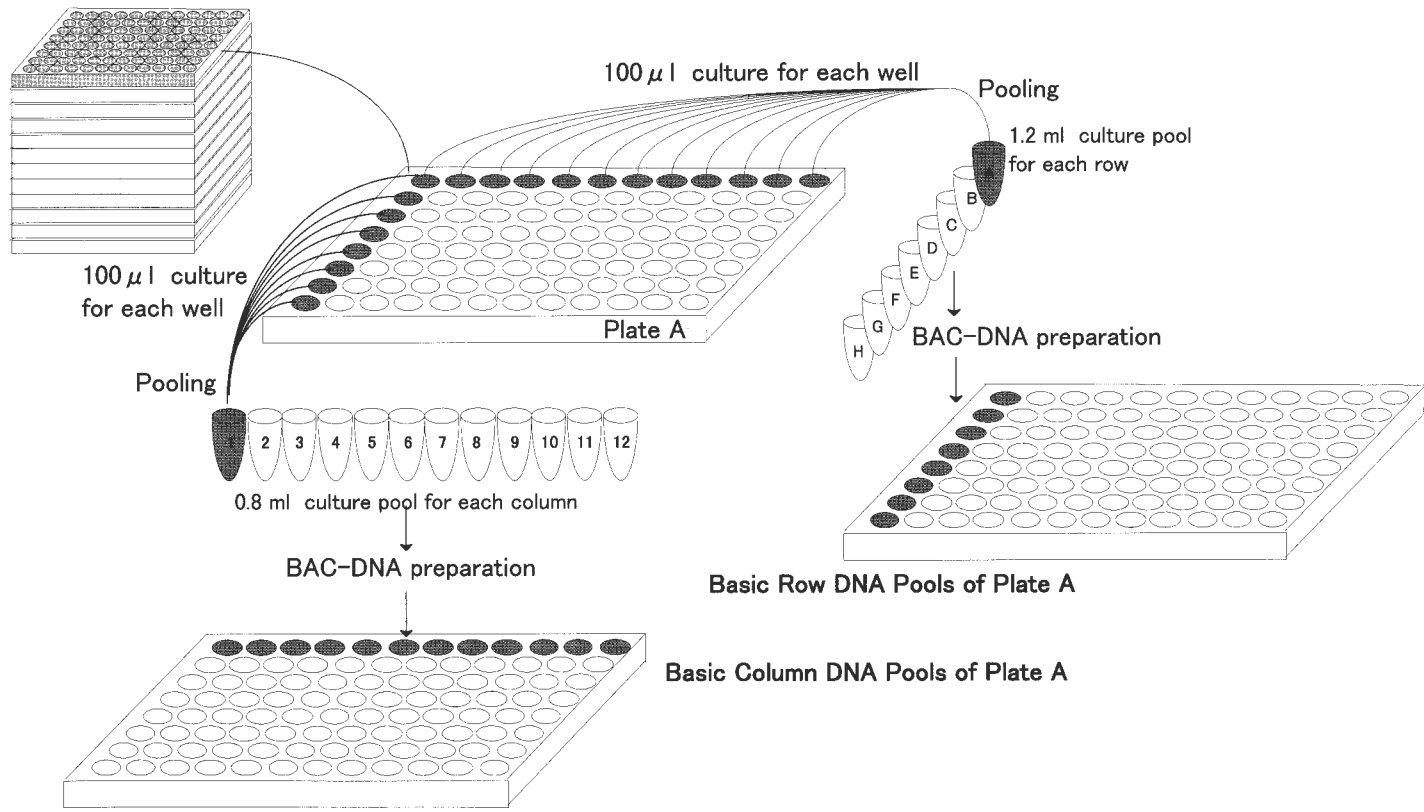


Fig. 1. Schematic representation of preparation of basic row and column DNA pools.

3.2. Preparation of DNA Pools for PCR-based Screening

1. Assemble 40- μ L aliquots of each basic row pool of the same position into a row of 96-deep well plates with a 12-channel pipet (*see Note 5*) to prepare row pools of the superpool (secondary row DNA pools, **Fig. 2**).
2. Assemble 40- μ L aliquots of each basic column pool of the same position into a column of 96-deep well plates with a 12-channel pipet to prepare column pools of the superpool (secondary column DNA pools, **Fig. 2**).
3. Assemble 25- μ L aliquots of row pools of the same plate into a column of 96-deep well plates with a 8-channel pipet. Similarly, assemble 25- μ L aliquots of each column pool of the same plate into a row of 96-deep well plates with a 12-channel pipet. Mix plate pools prepared from row pools and column pools to prepare plate pools of the superpool (*see Note 6*) (secondary plate DNA pools, **Fig. 2**).
4. Assemble 40- μ L aliquots of plate pools of the same superpool to prepare primary DNA pools (*see Fig. 2*).
5. Dilute primary and secondary DNA pools with nine-fold volume TE buffer to prepare working solutions. Dilute working solutions of primary and secondary DNA pools and original solutions of basic DNA pools with 19-fold volume 1.1X PCR buffer (*see Note 7*) to prepare solutions ready for PCR-based screening. Store all DNA pools at -20°C .

3.3. First-step Screening

1. Add 150 μ L of 1.1X PCR buffer and 12 μ L of primer (*see Note 8*) solutions (5 μ M) to a 1.5-mL microtube for each reaction. Add 2.4 μ L of *Taq* polymerase (5 U/ μ L) just prior to starting reaction and mix well.
2. Transfer 56 μ L of reactions into three wells in row H of a 96-well plate designed for PCR. Divide the reactions equally into eight wells in the same column with a 12-channel pipet.
3. Add 5 μ L of primary DNA solution to each well. Overlay the reaction mixture with 20 μ L of mineral oil (*see Note 9*).
4. Place the plate in a thermal cycler and perform amplification reaction.
5. Add 7 μ L of loading dye to each well, and mix well by pipeting. Load 7 μ L of the mixture on a binary gel and run electrophoresis until bromophenol blue dye has arrived at the end of the gel (*see Note 10*).
6. Stain the gel with ethidium bromide and record PCR banding pattern with UV illumination. Identify positive superpools to perform the second-step screening (*see Note 11*).

3.4. Second-step Screening

1. Add 225 μ L of 1.1X PCR buffer and 12 μ L of primer solutions (5 μ M) to a 1.5-mL microtube for each reaction. Add 2.8 μ L of *Taq* polymerase (5 U/ μ L) and mix well.
2. Transfer 56 μ L of reactions into four wells in row H of a 96-well plates and repeat procedures from **Subheading 3.3., steps 2 to 6** except for adding secondary DNA solutions instead of primary DNA solutions. Score positive signals for row, column, and plate pools to identify positive clones.
3. If single signal is detected from all of row, column and plate pools (*see Fig. 3A*), pick up the candidate clone and isolate BAC-DNA. Confirm the individual BAC DNA as harboring target sequences by PCR (*see Note 12*).
4. If multiple signals are detected from any of row, column and plate pools (*see Note 13*) or no signal is detected for any of row, column, and plate pools, proceed to **Subheading 3.5**.

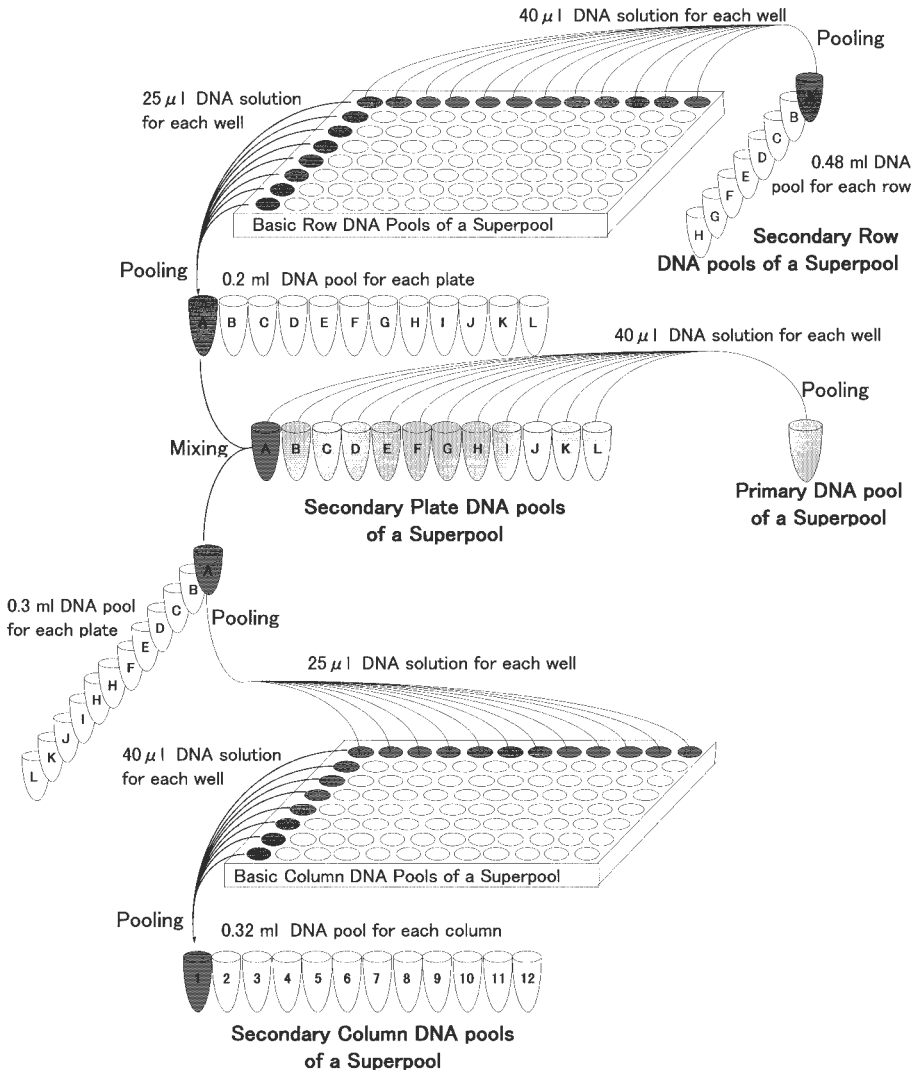


Fig. 2. Schematic representation of preparation of primary and secondary DNA pools.

3.5. Supplementary Experiments to Identify Positive Clones

1. If multiple signals are detected from any of row, column and plate pools (see Fig. 3B), pick up 5 µL of all candidate basic DNA pools of the positive plates, and perform PCR to narrow candidate clones.
2. If no signal is detected from row or column pools (see Fig. 3C), pick up 5 µL of basic DNA pools of the positive plate, and perform PCR to identify a candidate clone.
3. If no signal is detected from plate pools (see Fig. 3D), pick up 5 µL of basic DNA pools of the positive row or column, and perform PCR to identify a candidate clones. Pick up the candidate clone and isolate BAC-DNA for PCR confirmation.

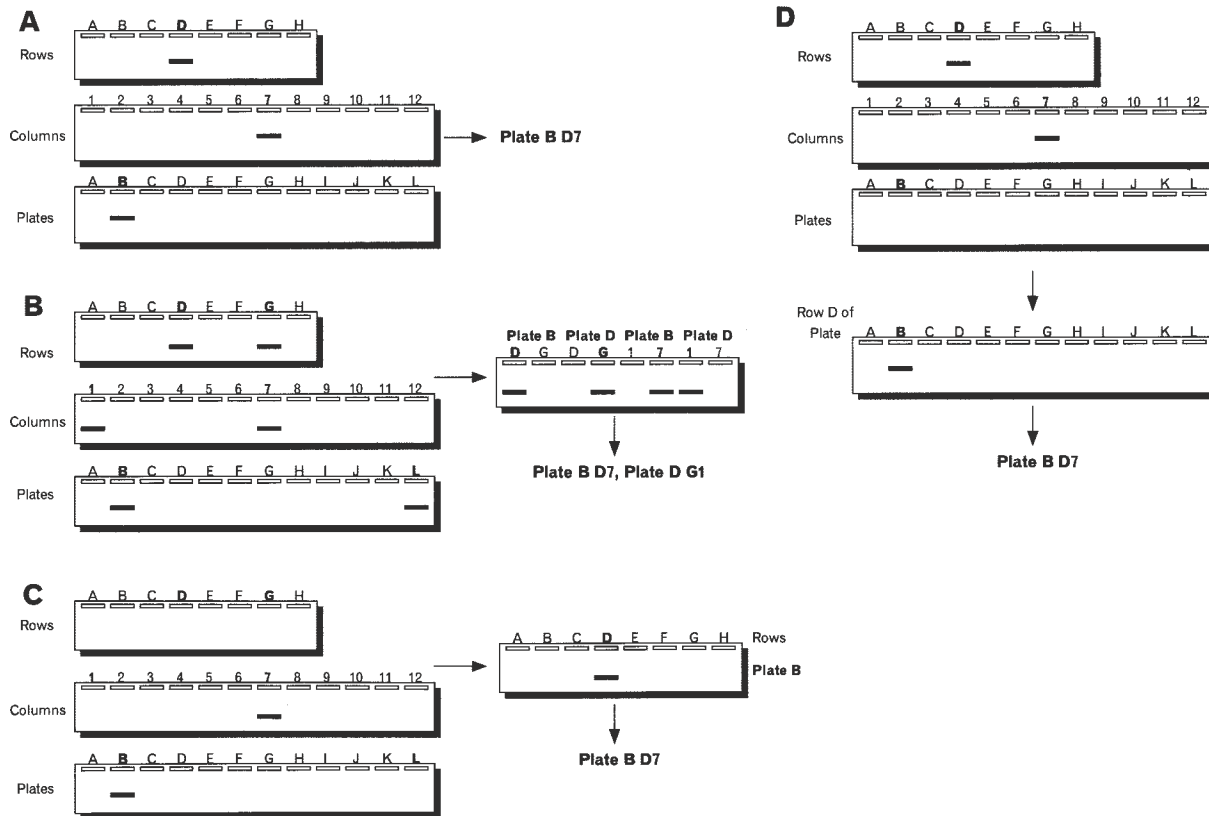


Fig. 3. Summary of results in the second-step screening and supplementary experiments required for each case. (A) Single signal is detected from all of row, column and plate pools, and supplementary experiments are not necessary. (B) More than two signals are detected (*see Subheading 3.5.1*). (C) No signal is detected from row or column pools (*see Subheading 3.5.2*). (D) No signal is detected from plate pools (*see Subheading 3.5.3*).

4. Notes

1. Addition of Synergel improves resolution and transparency of gels compared with a higher concentration (>2%) standard-grade agarose gel. Because Synergel powder tends to foam, mix agarose and Synergel powder thoroughly, add 0.1 volume of 0.5X TBE buffer, and wait until the powder is completely saturated. Then, add 0.5X TBE buffer up to final volume, and autoclave at 120°C for 5 min. Pour 250 mL autoclaved gel into 500-mL conical beakers and store them with plastic wrap at room temperature until gel casting. Melt the stored gel with a microwave, and pour to a casting tray. Gels should be cooled in a refrigerator for at least 30 min prior to loading samples.
2. Preparation of DNA pools should preferably be performed simultaneously with picking up and preservation of BACs in order to save labor and minimize the damage of frozen cells. In case DNA pools must be prepared from already preserved BAC library, pick up clones with a sterile 96-pin gridding tool and seed them to 0.4 mL-LB medium.
3. Alternatively, centrifuge the deepwell-plates at the maximum speed for your centrifuge until bacterial cells are completely precipitated. Remove the supernatant with a Pasteur pipet, and apply the cells to a 96-well format plasmid purification kit (i.e., QIAwell 96 Ultra Plasmid Kit, Qiagen) according to the provider's manual. Utilization of a 96-well format kit or plasmid isolator is preferable because it eliminates human errors of misidentification of samples.
4. DNase-free RNase (0.5 µg/mL) should be added to TE buffer because excessive RNA might disturb PCR-based screening.
5. Picking up clones from multiple frozen plates is troublesome and has a bad influence to plates which in fact have no positive clones. Therefore, the situation that positive plates are unknown should be avoided. Independent preparation of secondary plate DNA pools is very important to improve the efficiency of determining positive plates.
6. Take great care not to transfer basic DNA pools to a wrong row. Place the upper end of the plate lid on the lower end of the correct row to prevent a careless mistake.
7. It is very difficult to keep the pipeted volume constant for all channels of a multichannel pipette. If DNA pools are diluted with TE buffer or water, the final concentration of the reaction varies for each well because the ratio between the buffer diluting DNA pools and the buffer containing primers and *Taq* polymerase is not constant. Therefore, the composition of them should be the same to improve reliability of the analysis. However, the salt concentration of DNA pools diluted with PCR buffer is apt to increase due to evaporation of water, which leads to a smearing banding pattern in PCR. Thus, DNA pools should be stored in the way to minimize evaporation of water.
8. Primers should be designed to generate clear expected fragments, which can be directly amplified from genomic DNA as a template. It is strongly recommended that similarity search should be performed for sequences used for primer design to find any repetitive sequences. Generally speaking, too short (<100 bp) or long (>1 kb) fragments are not desirable to obtain stable results. Arbitrary primers, typically used in RAPD analysis, and degenerate primers can be used as far as they can generate clear bands directly from genomic DNA.
9. Mineral oil should be overlaid even if a thermal cycler with a hot lid is used, because mineral oil prevents evaporation of the finished reaction and keeps its volume constant.
10. It is recommended to utilize combs designed for loading with a multichannel pipet, whose interval between wells is adjusted to that of 96-well plate.
11. If expected fragments are invisible for all lanes, increase of PCR cycles (up to 50 cycles) might be effective. Nested PCR is also useful to improve sensitivity of the screening. Use

- 5 μL of the finished reaction of the first reaction instead of DNA pools as templates for the second reaction with internal primers.
12. If the isolated BAC-DNA does not contain target sequence, confirm if scoring of positive signals is correct first of all. When there is no problem about scoring, pick up 5 μL of the positive basic row and column DNA pools of the positive plates, and perform PCR to confirm whether DNA pooling was correctly performed. When no signal is detected from them, proceed to **Subheading 3.5.2**. You may find a lot of false-positive clones from specific plates, indicating BAC-DNAs were not properly prepared or pooled. In such cases, prepare basic DNA pools of such plates again, or otherwise exclude them from isolating clones.
 13. If positive signals are detected from nearly all lanes, a multicopy sequence motif might be involved in primer sequences. In many cases, such fragments are amplified very well, and sometimes smearing appears because of polymorphisms among them.

References

1. Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. USA* **89**, 8794–8797.
2. Burke, D. T., Carle, G. F., and Olson, M. V. (1987) Cloning of large-segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806–812.
3. Chumakov, I., et al. (1992) Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* **359**, 380–387.
4. Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Kubota, R., et al. (1997) Human BAC library: construction and rapid screening. *Gene* **191**, 69–79.
5. Nakamura, S., Asakawa, S., Ohmido, N., Fukui, K., Shimizu, N., and Kawasaki, S. (1997) Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene *Pi-ta* using a highly representative rice BAC library. *Mol. Gen. Genet.* **254**, 611–620.
6. Wu, C., Asakawa, S., Shimizu, N., Kawasaki, S., and Yasukochi, Y. (1999) Construction and characterization of bacterial artificial chromosome libraries from the silkworm, *Bombyx mori*. *Mol. Gen. Genet.* **261**, 698–706.

A 384-Well Microtiter-Plate-Based Template Preparation and Sequencing Method

Lei He and Kai Wang

1. Introduction

As the genome research continues to grow and genomic information avalanches at the speed of light, biological research labs are facing continuous effort to redefine the scale and scope of their research. Large-scale research applications such as high-throughput sequencing and cDNA microarrays are becoming standard tools in most labs, demanding an increasing ability to generate and handle a large amount of highly purified DNA templates (**1**). As a result, developing protocols that address the rising needs in sample tracking, robotic compatibility, and efficient integration has become a priority.

Sample tracking is essential for validating results associated with particular samples. It is especially challenging to labs that use both 96- and 384-well format and need to transfer samples back and forth between the two. We (**2,3**) and others (**4,5**) have developed 96-well microtiter plate-based DNA preparation and purification protocols. During the past few years, we have optimized our protocol to accommodate 384-well microtiter plate-based DNA sample handling. This protocol involved a single-step same-well PCR purification process that eliminates the need for tedious liquid transfer and sample tracking (*see Fig. 1*). The purification process is sufficient to support subsequent sequencing reactions, array fabrications, and other molecular techniques that require purified DNA samples.

2. Materials

2.1. DNA Preparation

1. LB media with 50 µg/mL ampicillin.
LB media composition (1 L): 10 g of bacto-tryptone, 5 g of bacto-yeast extract, 5 g of NaCl, adjust to pH 7.0 with 5 M NaOH, and bring to volume with Milli-Q water. Sterilize by autoclaving.
2. Sterile toothpicks.

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

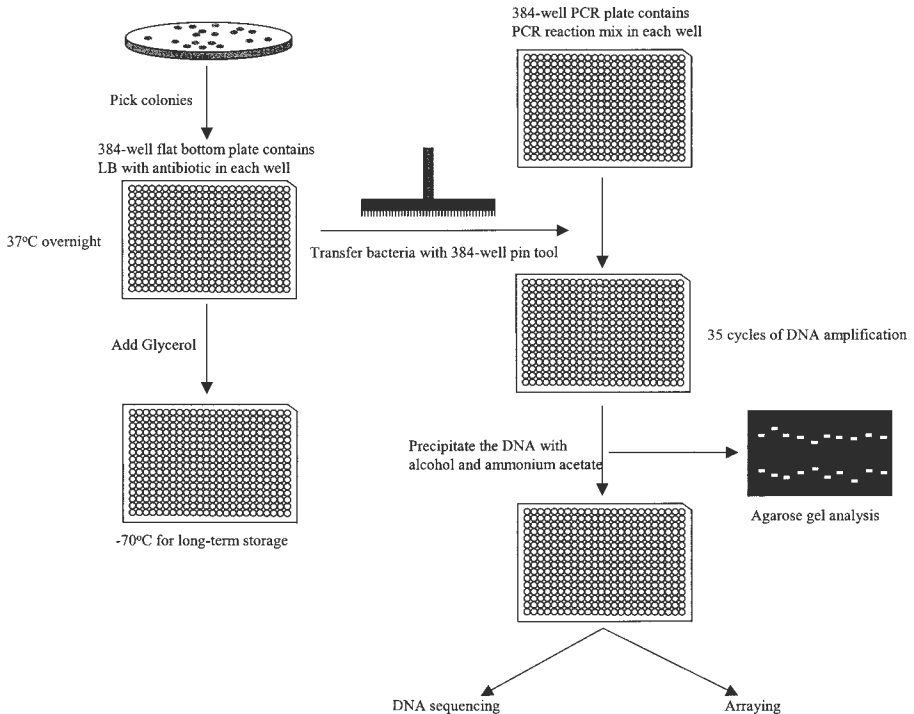


Fig. 1. Schematic drawing of the DNA purification process.

- 384-well culture plate (cat. no. 242757, Nunc, IL or similar). These are flat-bottom plates with 250 μL total volume per well.
- Repeat pipetor (cat. no. 53106-220, VWR Scientific Products).

2.2. PCR-Based DNA Isolation

- 384 pin tool (Catalog number 250393, V & P Scientific, CA).
- 384-well V bottom PCR plate (cat. no. TF-0384, Marsh, NY) with 200 μL total volume per well.
- Micro-Seal A (cat. no. MSA-5001, MJ Research, MA).
- 384-well thermocycler (PTC-200, MJ Research, MA) or a tetrad (cat. no. PTC-225, MJ Research, MA) with a 384-well alpha unit (cat. no. ALS-1238, MJ Research, MA).
- PCR premix: 50 mM NaCl, 20 mM Tris-HCl (pH 9.0), 2.5 mM MgCl_2 , 0.5 mM dithiothreitol (DTT) 50 μM of dATP, dCTP, dGTP and dTTP, and 0.5 U of *Taq* DNA polymerase (Roche, IN).
- Vector primers for pUC and pUC derived vector (0.4 μM)
 Forward: 5'-CGCCAGGGTTTTCCAGTCACGACGTTG-3'
 Reverse: 5'-GTATGTTGTGGGAATTGTGAGCGGA-3'
- 1X agarose gel loading dye made from 10X loading dye stock. 10X loading dye composition (5 mL total volume):
 2.5 mL of glycerol, 50 μL of 10% sodium dodecyl sulfate (SDS), 21 mg of bromophenol blue, 21 mg of xylene cyanole, 1 mL of 0.5 M ethylenediaminetetraacetic acid (EDTA) (pH 8.0) and finally, 1.45 mL of Milli Q water.

2.3. PCR Product Purification

1. 95% ethanol with 2 M ammonium acetate (NH_4OAc) (cat. no. number A639-500, Fisher Scientific, Inc. Houston, TX).
2. Aluminum tape (cat. no. Scotch 60597, 3M, MN).
3. 1X TE: 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA.
4. Agarose gel electrophoresis apparatus.
Horizontal minigel apparatus (cat. no. HMG 34, Ellard instrumentation Ltd., Seattle, WA) and power supply (such as Bio-Rad power pac 300).

2.4. DNA Sequencing

1. 384-well V bottom PCR plate (cat. no. TF-0384, Marsh, NY) with 200- μL total volume per well.
2. 4 μM stock of sequencing oligos. We routinely use the M13 forward and reverse primers indicated above for sequencing.
3. Big Dye Terminator sequencing premix (cat. no. 4303135, Applied Biosystems, Inc., Foster City, CA).
4. 2.5X sequencing dilution buffer: 5 mL of 1 M Tris, pH 10.0, 125 μL of 1 M MgCl_2 , Adjust pH to 9.0 with HCl.
5. Fluorescent DNA sequencer (ABI 377, Applied Biosystems, Inc.).
6. Long ranger single packs (cat. no. 50693, BMA, Rockland, ME).
7. 95% ethanol with 2 M ammonium acetate (NH_4OAc).
8. Aluminum tape (cat. no. Scotch 60597, 3M, MN).
9. Loading buffer: deionized formamide: 25 mM EDTA with blue dextran (5:1, v:V).

3. Methods

3.1. DNA Preparation

1. Dispense 50 μL of LB media with 50 ng/ μL ampicillin into each well of a flat bottom 384-well culture plate using a repeat pipetor from Eppendorf (cat. no. 022-26-900-3, Brinkmann Instruments, Inc., Westbury, NY). Set the pipet tip against the side wall of each well to avoid splashing.
2. Pick one bacteria colony into each individual well of the 384-well culture plate using a sterile toothpick and leave it in the well.
3. After the entire plate is filled, take out the toothpicks and discard in biohazard container.
4. Place the culture plate onto a plate shaker in a 37°C incubator. The shaker speed should be set at 0 first and gradually turned up to where the culture is gently shaken, but not spilling into other wells.
5. Incubate the culture overnight (*see Note 1*).
6. For long term storage of these clones, 25 μL of LB with 30% glycerol (v/v) is dispensed into each of the wells in the original culture plate (final glycerol concentration is 10%). The plate then can be stored in a -70°C freezer.

3.2. PCR Amplification

1. Prepare PCR premix for one 384-well plate by mixing the following:
 - a. 40 μL Forward primer (10 μM)
 - b. 40 μL Reverse primer (10 μM)
 - c. 100 μL *Taq* DNA polymerase
 - d. 600 μL 10X buffer
 - e. 120 μL 10 mM dNTPs
 - f. 5100 μL Milli-Q water

Be sure to keep *Taq* polymerase on ice. Always add *Taq* last and mix by gently pipeting up and down. Do not vortex the premix after *Taq* has been added.

2. Dispense 15 μL of premix into each individual well of a V bottom 384-well PCR plate. Set the pipet tip against the side wall of each well to avoid splashing.
3. Dip the 384-pin tool into the overnight culture to transfer a small amount of the bacterial culture into the PCR plate containing the amplification premix (see **Notes 2** and **3**).
4. Seal the plate with Micro-Seal A (cat. no. MSA-5001, MJ Research, MA) and put it into a 384-well thermocycler for DNA amplification.
5. The thermocycler program should be set as the following:
 - a. 94°C 10 s denaturation
 - b. 60°C 30 s primer annealing
 - c. 72°C 90 s extension of product35 cycles
6. To check the PCR amplification, we take 3 μL of the amplification mixture from the PCR tube and mix it with 1 μL of gel loading dye. Load the entire volume onto a 1% agarose gel for analysis.
7. After electrophoresis, the gel is examined under an UV light and a picture is taken for the record.

3.3. PCR Product Purification

Excess primers and salts from PCR need to be removed before further application such as DNA sequencing.

1. Add 30 μL of 95% ethanol with 2 M ammonium acetate to the remaining 12 μL of amplification product in each well. This may be done either using a multichannel pipettor or a robotic instrument such as the Hydra.
2. Seal the entire plate with aluminum tape. Vortex briefly, then incubate the plate at room temperature for 20 min (see **Note 4**).
3. Spin down the plate at room temperature using a microplate centrifuge (cat. no. CT422, Jouan, VA). The centrifugation time is 20 min at 2500g.
4. Wash each pellet twice with 30 μL of 70% ethanol. Spin down the pellets afterwards using the same centrifuge for 5 min at 2500g.
5. Discard the supernatant and turn the plate upside down on a stack of paper towels.
6. Spin the plate briefly with it being upside down on the paper towel in the same centrifuge. The centrifugation time is 30 s and the speed should not exceed 800g.
7. Turn the plate right side up and air-dry the pellets at room temperature. Average drying time is about 30 min.
8. Resuspend the pellets in 15 μL of 1X TE. The yield and quality of the purified PCR product can be verified by agarose-gel electrophoresis. We can routinely recover about 80% of DNA without noticeable primer contamination.

3.4. DNA Sequencing

1. Transfer 3 μL of purified PCR products into each well of a new 384-well V bottom plate. This is done using a Hydra 96 pipet, but also could be done using a multichannel pipettor (see **Note 5**).
2. Make sequencing premix for 384-well sequencing reactions by mixing the following reagents:

- a. 200 μL 4 μM sequencing oligo
 - b. 1000 μL Milli-Q water
 - c. 1600 μL Big Dye Terminator sequencing mix
3. Using a Rainin Multidispensing pipetor, dispense 7 μL of the premix into each individual well of the sequencing plate. We do not use the Hydra for this purpose to conserve expensive sequencing reagents. Bring entire volume down to the bottom of the wells by a quick spin with microplate centrifuge.
 4. Seal the plate with Micro-Seal A and put it in a MJ 384-well thermocycler or a 384-well alpha unit of a MJ Tetrad. The thermocycling conditions are set following ABI's Big Dye Terminator manual:
 - a. 96°C 10 s
 - b. 50°C 5 s
 - c. 60°C 4 min25 cycles
 5. The reaction takes approx 2.5 h. After amplification, add 20 μL of 95% ethanol with 2 M ammonium acetate to each well to purify the sequencing product (*see Note 6*).
 6. Seal the entire plate with aluminum tape and vortex the plate briefly.
 7. Spin down the plate at room temperature using microplate centrifuge (cat. no. CT422, Jouan, VA) at 2500g for 30 min.
 8. Wash the pellets twice with 20 μL of 70% ethanol, spin it down briefly using the same centrifuge at 2500g for 5 min.
 9. Discard the supernatant and turn the plate upside down on a stack of paper towels.
 10. Spin the plate briefly with it being upside down on the paper towel in the same centrifuge. The centrifugation time is 30 s and the speed should not exceed 800g.
 11. Turn the plate right side up and air-dry the pellets at room temperature. Average drying time is about 30 min.
 12. Resuspend the pellets in 3 μL of loading buffer, 1 μL of which is loaded onto a polyacrylamide gel subjected to 9 h of electrophoresis on a 377 fluorescent sequencer. The electrophoresis parameters are set up according to ABI's standard, which runs at 2850 V, 40 mA, and the gel temperature is kept at 50°C.

4. Notes

1. When preparing the growth culture for amplification, the incubation period may need to be adjusted depending on the vector used. It is best not to overgrow cells to their apoptosis stage because PCR will amplify genomic DNA contamination and result in by-products interfering with sequencing. In our lab, we generally work with cultures that have grown for 4–6 h.
2. To avoid cross-contamination between templates, the 384-pin tool needs to be cleaned thoroughly. We submerge the pin tool in a plate containing 70% ethanol and flame it quickly in between every transfer.
3. PCR amplification requires only a very small amount of growth culture as starting material. Therefore, when inoculating the PCR premix with growth culture, a light touching of the culture surface using the 384-pin tool will carry sufficient amount of DNA for the transfer.
4. During the PCR product purification process, it is critical to thoroughly mix the precipitant with unpurified PCR products within each well. The 20-min incubation time is also important for the proper interaction to occur between salt and small nucleotides.

5. Our experience has been that many brands of 384-well PCR plates show various degrees of warpage after thermocycling. The deformed plates can not be used with any robotic system including the Hydra station (Robbins Scientific, CA). The 384-well plates from Robbins Scientific (cat. no. 1047-90-0, Robbins Scientific, CA) proves to have the least warping, especially when used in conjunction with their plate stretcher.
6. At the time this method was developed, sequencing dye removal was done via ethanol precipitation. Since then, a few 384-well column based dye removal systems became available, offering other options. These systems, however, still rely heavily on centrifugation. Automating this step continues to be a challenge.
7. We routinely obtain over 700 bp of reliable sequence reads with no significant quality difference when compared to DNA purified through alkaline lysis method.
8. For DNA sequencing application, the turnaround time from picking bacterial colonies to achieving final DNA sequences based on this protocol is only a bit over one day in our lab.
9. All the reagents were delivered using an 8-channel pipet at the time this protocol was developed. We have since adapted most of the procedures using Hydra 96 for increased throughput.

References

1. DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680–686.
2. Huang, G. M., Wang, K., Kuo, C., Paepfer, B., and Hood, L. (1994) A high-throughput plasmid DNA preparation method. *Analyt. Biochem.* **223**, 35–38.
3. Wang, K., Gan, L., Boysen, C., and Hood, L. (1995) A microtiter plate based high throughput DNA purification method. *Analyt. Biochem.* **226**, 85–90.
4. Itoh, M., Carninci, P., Nagaoka, S., Sasaki, N., Okazaki, Y., Ohsumi, T., et al. (1997) Simple and rapid preparation of plasmid template by a filtration method using microtiter filter plates. *Nucl. Acids Res.* **25**, 1315–1316.
5. Engelstein, M., Aldredge, T. J., Madan, D., Smith, J. H., Mao, J. I., Smith, D. R., and Rice, P.W. (1998) An efficient, automatable template preparation for high throughput sequencing. *Microb. Comp. Genomics* **3**, 237–241.
6. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

A Microtiter-Plate-Based High Throughput PCR Product Purification Method

Ryan Smith and Kai Wang

1. Introduction

The construction and sequencing of DNA libraries, such as genomic libraries and cDNA libraries, is an important and powerful approach to understand the biology at the most fundamental level, that of the DNA. The high-throughput DNA sequencing approach has also been used in novel gene discovery and gene expression analysis. By sequencing such libraries, one can gain valuable information about the genome and the identity, as well as the prevalence of individual messages between the cell's nucleus and transcriptional machinery. Such knowledge can be invaluable in the quest to elucidate complex biological systems. Unfortunately, the preparation of a quality template DNA can be a time-consuming and costly hurdle on the path to efficient library sequencing. Here, we present a quick and reliable high-throughput approach to inexpensive DNA purification. This method is also suitable for the expeditious processing of whole colony or plaque polymerase chain reaction (PCR) products (1).

A number of different methods can be used to prepare either single- or double-stranded DNA sequencing template. However, most necessitate the investment in semiautomated equipment or require expensive consumables such as commercially prepared microcentrifugation or ion-exchange columns, whereas other methods can be low-throughput owing to multiple and tedious liquid handling steps (2–7). Using the protocol described here, one can quickly obtain purified, double-stranded DNA ready for use directly in fluorescent sequencing reactions. We regularly achieve 500 bp or more of quality sequence data from DNA purified using this protocol. In addition to sequencing, the purified product is ready for use in hybridizations and many other enzymatic manipulations.

Purification is achieved by passing the cDNA through a gel filtration matrix prepared as minispin columns. The minispin columns are prepared by pipeting the gel filtration matrix into a 96-well membrane bottomed assay plate and then packing the

column by vacuum or centrifugation. After rinsing the columns, the unpurified DNA is then loaded on top of the filtration matrix. The matrix retains the salts, oligonucleotide primers, and unincorporated nucleotides, whereas the large DNA fragments pass through the columns in the subsequent centrifugation step (*see Fig. 1*). The entire procedure, including the preparation of the minispin columns and verification of product recovery by agarose gel electrophoresis, can be accomplished in roughly 1 h.

2. Materials

1. Unpurified PCR products (5-40 μ L).
2. 96-well membrane-bottomed assay plate, (cat. no. 256081, Nalge Nunc International, Rochester, NY) or MultiScreen96 PCR plate (cat. no. MANU03010, Millipore, Bedford, MA).
3. Sephacryl 500-HR (cat. no. 17-0613-10, Amersham Pharmacia Biotech, Piscataway, NJ).
4. 96-well collection plate (Multi 0.2 mL Rigid PCR Ultraplate, cat. no. 53550-036, VWR, West Chester, PA, or similar products).
5. 96-well PCR tray and base (cat. no. 403081 and number N801-0531, Applied Biosystems, Foster City, CA, or similar products).
6. Centrifuge with microplate rotor (Sorvall RC-3B Plus, DuPont, Wilmington, DE, or similar products).
7. 1X TE buffer: 10 mM Tris-HCl, pH 7.4, 0.1 mM ethylenediaminetetraacetic acid (EDTA).
8. Agarose (Low EEO Agarose, cat. no. 300013, Stratagene, La Jolla, CA, or similar products).
9. 1X TAE buffer: 40 mM Tris acetate, 1 mM EDTA, pH 8.0.
10. Ethidium bromide (cat. no. 300083, Stratagene, or similar products).
11. Microwave oven or hot plate.
12. Horizontal electrophoresis apparatus (cat. no. 400450, Stratagene, or similar products).
13. 40-well comb (cat. no. 400458, Stratagene, or similar products).
14. Power supply (cat. no. 400690, Stratagene, or similar products).
15. Gel loading buffer (10X BlueJuice, cat. no. 10816-015, Life Technologies, Rockville, MD, or similar products).
16. DNA size standard (123 bp DNA Ladder, cat. no. 15613-011, Life Technologies, or similar products).
17. Transilluminating UV light box (cat. no. 401150, Stratagene, or similar products).

3. Methods

3.1. Mini Spin Column Preparation in 96-Well Plates (see Notes 1 and 2)

1. Load approx 200 μ L of Sephacryl 500-HR into each well.
2. Place membrane-bottomed plate with Sephacryl 500-HR on top of a 96-well collection plate and spin at 800g for 2 min in centrifuge with microplate rotor.
3. Discard flowthrough and add an additional 100 μ L of Sephacryl 500-HR to each well.
4. Spin at 800g for 2 min in centrifuge with microplate rotor (*see Note 3*).
5. Discard flowthrough. The final column height should now be roughly three-quarters of the well depth. Add more Sephacryl 500-HR and repeat centrifugation as needed.
6. Rinse the columns with 100 μ L 1X TE and spin at 800g for 2 min in centrifuge with microplate rotor. Discard flowthrough.
7. Repeat **step 6** twice.

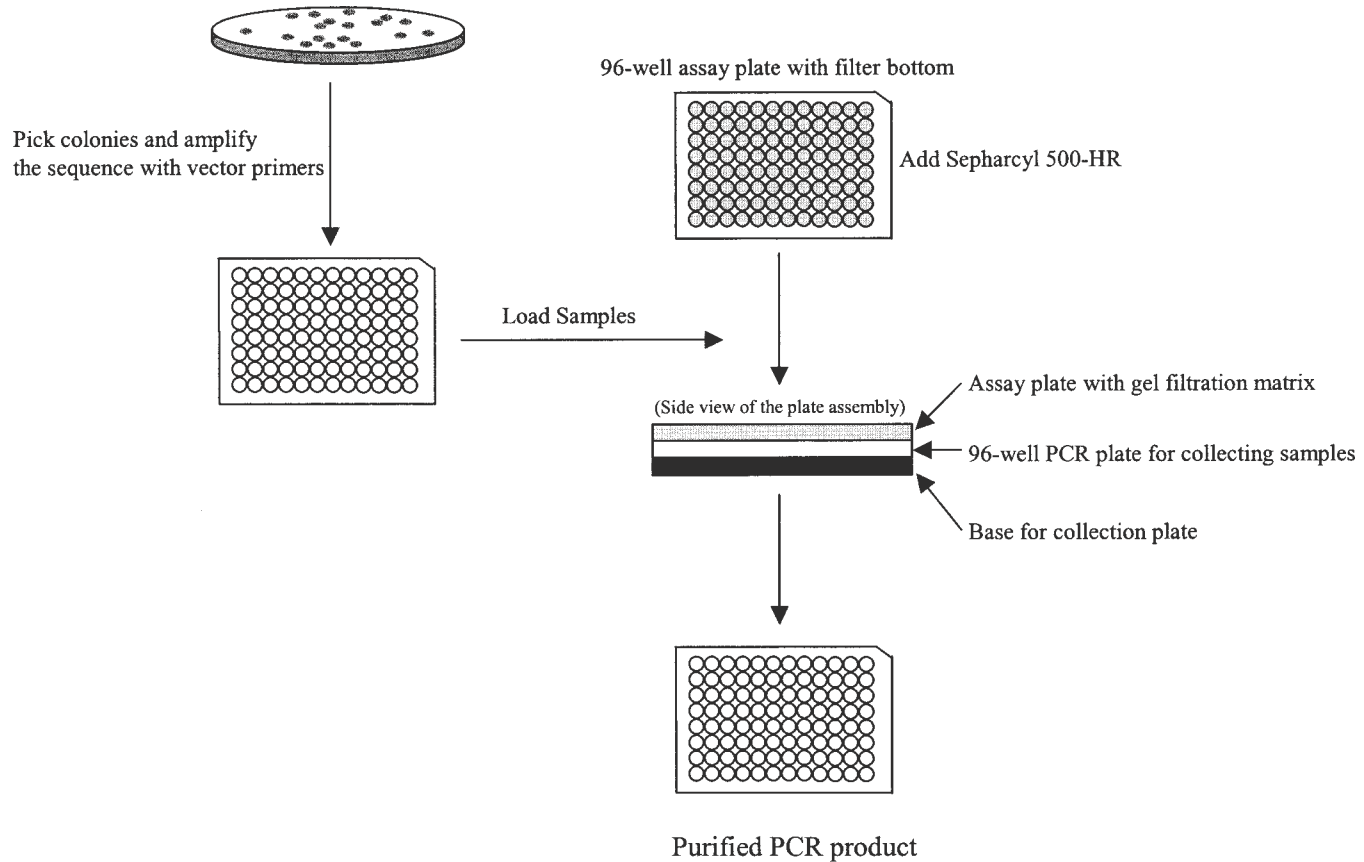


Fig. 1. A brief outline of the DNA purification method.

3.2 Purification

1. Apply PCR products to center of the columns (*see Note 4*).
2. Place membrane-bottomed plate with Sephacryl 500-HR on top of a 96-well collection plate/base assembly and spin at 800g for 2 min in centrifuge with microplate rotor. Collect the flowthrough.
3. Using the same 96-well collection plate, add an additional 20 μ L of 1X TE buffer on top of each column and spin at 800g for 2 min in centrifuge with microplate rotor (*see Note 5*).
4. Cover 96-well collection plate and place at 4°C for short-term storage or at -20°C for long-term storage.

3.3. Agarose Gel Electrophoresis

1. Prepare a 2% agarose gel by melting agarose in 1X TAE buffer using a microwave oven or hot plate.
2. Cool the molten agarose to 65°C in a water bath. Add ethidium bromide and pour agarose solution into the casting tray with a comb, having enough teeth to create wells for all your samples plus a DNA size standard, in place.
3. Allow the gel to cool and set.
4. Place the gel in electrophoresis apparatus and submerge in 1X TAE. Remove the comb.
5. Mix the purified PCR product from **Subheading 3.2.** with an appropriate amount of loading dye and then load the samples into the wells. Also load an aliquot of DNA size standard.
6. Apply current to the apparatus setting the voltage at approx 100 V. Run the gel until the dye has migrated 30 mm.
7. Examine the gel on an UV light box and confirm PCR product recovery and removal of primers and unincorporated nucleotides.

4. Notes

1. We have observed some variability in the performance of the matrix (and, to a lesser degree, the membrane-bottomed plates) from different lots. We suggest performing a test purification of a small number of samples (*see Note 2*) to assess the performance of the materials before processing large numbers of samples.
2. This protocol may be used with fewer than 96 samples. Seal the unused wells of the membrane-bottomed plate with tape before loading the matrix (*see Subheading 3.1.1.*). The plate and the unused wells may be used at a later date. Do not use the same wells more than once.
3. The protocol for column preparation can be performed using a vacuum manifold (cat. no. MAVM0960R, Millipore, Bedford, MA) in place of a centrifuge. When using a vacuum manifold, apply vacuum until column stops dripping.
4. While preparing the mini spin columns, take care to avoid sloped columns. If unpurified PCR products are applied to sloped columns, the products may not pass through enough of the matrix to achieve purification. Sloped columns may be repaired by applying 1X TE buffer to the column and stirring the well with the pipet tip, followed by another round of centrifugation.
5. Depending on the concentration of the PCR products and their intended use, **step 3.2.3.** may be omitted. The majority of the DNA is recovered in the first centrifugation step. A second centrifugation will increase total DNA recovery but will decrease the concentration of DNA.

References

1. Wang, K., Gan, L., Boysen, C., and Hood, L. (1995) A microtiter plate-based high-throughput DNA purification method. *Analyt. Biochem.* **226**, 85–90.
2. Eperon, I. C. (1986) Rapid preparation of bacteriophage DNA for sequence analysis in sets of 96 clones, using filtration. *Analyt. Biochem.* **156**, 406–412.
3. Mardis, E. R. and Roe, B. A. (1989) Automated methods for single-stranded DNA isolation and dideoxynucleotide DNA sequencing reactions on a robotic workstation. *Biotechniques* **7**, 840–850.
4. Smith, V., Brown, C. M., Bankier, A. T., and Barrell, B. G. (1990) Semiautomated preparation of DNA templates for large-scale sequencing projects. *DNA Seq.* **1**, 73–78.
5. Cao, X., and Brosius, F. C. (1993) Direct sequencing of double-stranded PCR products isolated from conventional agarose gels. *Biotechniques* **15**, 384–386.
6. Anderson, B., Lu, J., Edwards, K. E., Muzny, D. M., and Gibbs, R. A. (1996) Method for 96-well M13 DNA template preparations for large-scale sequencing. *Biotechniques* **20**, 1022–1027.
7. Leonard, J. T., Grace, M. B., Buzard, G. S., Mullen, M. J., and Barbagallo, C. B. (1998) Preparation of PCR products for DNA sequencing. *Biotechniques* **24**, 314–317.

Index

A

- ABI Prism 7700, 15
- AcMNPV, 153
- Adaptor ligation, 346
- Agarose gel analysis
 - DNA fragments, 202
- Agarose gel electrophoresis, 14
 - iPCR, 79
 - microtiter-plate-based high throughput PCR product purification, 420
- AmeriClear, 71
- Amplifiable cDNA
 - synthesis, 364
- Amplifiable single-stranded cDNA
 - synthesis, 366
- Amplification cycles, 12
- AmpliTaq DNA Polymerase, 6–7
 - associated activities, 7
 - biochemical reactions, 6–7
 - biophysical properties, 6
- AmpliTaq Gold, 5–6, 7, 73
- AmpliTaq Gold Hot Start process, 13
- AmpliWax PCR gems, 5
- Anchored-PCR, 309
- Anchor vectors
 - preparation, 339
- Annealing stages
 - thermal cycler, 34
- Annealing temperature
 - bacteriophage λ sublibraries, 397
 - recombinant libraries rapid and nonradioactive screening, 383
- TCE, 138
- XL PCR, 48–49
- Antibiotics
 - AS-PCR, 149
- AP-PCR, 249, 260
- Arbitrarily primed PCR (AP-PCR), 249, 260
- AS-PCR. *see* Autosticky PCR (AS-PCR)
- Asymmetric PCR recombination
 - PCR-mediated recombination methods, 200–203
- Autographa californica* nuclear polyhedrosis virus (AcMNPV), 153
- Autosticky PCR (AS-PCR), 113, 141–151
 - amplification, 146
 - antibiotics, 149
 - clone analysis, 148
 - ligation, 147
 - materials, 142–143
 - methods, 142–148
 - plasmid purification, 148
 - primer design, 143–144
 - primer synthesis, 144–145
 - product
 - analysis and purification, 147
 - restriction digestion
 - web site, 149
 - restriction endonucleases, 149
 - synthesizers
 - web site, 149
 - transformation, 147
 - vector preparation, 145–146

B

- Bacterial artificial chromosome (BAC)
 - amplification, 338
 - DNA
 - double digestion, 340
 - end
 - amplification, 340
 - large-terminal sequence isolation, 337–341
 - materials, 339
 - methods, 339–340
- Bacterial artificial chromosome (BAC) libraries, 337, 401–410
- DNA pools, 404–406, 409
- DNase-free RNase, 409
- materials, 404
- methods, 404–408
- mineral oil, 409
- positive clones, 407
- screening, 356–357
- Synergel, 409
- Bacteriophage λ sublibraries, 391–399
 - annealing temperature, 397
 - cross contamination, 398
 - elongation time, 397
 - materials, 392–393
 - methods, 394–396
 - negative control, 397
- Baculovirus cloning, 153–162
 - buffers, 157
 - cells, 158
 - chemicals, 157
 - materials, 157–158
 - methods, 158–161
 - method theory, 154–157
 - PCR amplification
 - protein gene, 158–159
 - PCR equipment, 157

- PCR primers, 158
- plaque assay chemicals and equipment, 158
- transfection experiments, 160
- Baculovirus genome
 - GFP gene direct cloning, 155
- Baculovirus vectors
 - commercial sources, 161
- Betaine, 11
- Bio-Rad *E. coli* Gene Pulser, 182
- Biotinylated capture probes
 - immobilization, 364–365
 - synthesis, 364–365
- Bisulfite
 - DNA denaturation, 96
- Bisulfite sequencing, 92
- BLAST
 - web site, 8
- Blunt-end cloning, 112
- Blunt-end restriction sites
 - in-frame cloning, 176
- Bovine serum albumin (BSA), 11
- B1-PCR, 256
- Browsers
 - GeneFisher, 29
- BSA, 11
- Buffers, 10
- C**
- CAG-repeat containing genes
 - cloning strategies, 222
 - introducing CAG expansions, 218
 - materials, 219
 - methods, 220–222
 - restriction endonuclease, 222–223
- Captured cDNA
 - cloning, 372
 - sequencing, 372
- Capture PCR (C-PCR), 264–265

- Capture steps, 360–362, 365, 368–372
 - Southern blot, 369–370
 - cDNA
 - amplifiable
 - synthesis, 364
 - amplifiable single-stranded
 - synthesis, 366
 - cloning
 - iPCR, 249–250
 - libraries, 354–355
 - Chimeric HIV-1 clones
 - PCR-mediated recombination
 - methods, 199–200
 - Circular plasmid
 - digestion, 222
 - Circular target DNA
 - digestion, 220
 - CLONTECH
 - web site, 46
 - CODEHOP, 23
 - Cold-start method, 285–291
 - annealing temperatures, 290
 - extended primer, 290
 - materials, 287–288
 - methods, 288–290
 - nonspecifically primed PCR, 288–289
 - primer design, 288
 - reamplification PCR, 289–290
 - Colony hybridization
 - skills, 401
 - Computer programs. *see* Software packages
 - Cosolvents, 10, 11
 - Coupled one-step RT-PCR
 - materials, 54
 - methods, 55
 - C-PCR, 264–265
 - Crime labs
 - validation, 3
 - Cross contamination
 - bacteriophage λ sublibraries, 398
 - 384-well microtiter-plate-based
 - template, 415
 - D**
 - Degenerate PCR primer
 - primer design software
 - packages, 21
 - GeneFisher, 26–27
 - Deoxynucleoside triphosphate (dNTP), 3–4
 - XL PCR, 47
 - Design software packages
 - primer, 8
 - Differential Display methodology, 44
 - Differentially expressed genes
 - subtractive cloning, 104
 - Directional restriction site-free
 - insertion, 133–139
 - DpnI digestion and
 - transformation, 137
 - enzymes, 135
 - equipment, 135
 - materials, 135
 - methods, 136–137
 - peptidyl-tRNA hydrolase
 - encoding region
 - amplification, 136–137
 - primer design, 136
 - reagents, 135
- Disruption mutations
 - mini-Tn3
 - ST-PCR, 331–332
- DMSO, 11
- DNA amplification
 - GC-PCR, 365

- DNA cloning and sequencing
 - GC-PCR, 365
- DNA denaturation
 - bisulfite, 96
- DNA fragments
 - agarose gel analysis, 202
- DNA methylation
 - analysis, 91–92
- DNA mutagenesis and
 - recombination, 167–174
 - advantages and limitation, 169
 - cloning efficiency, 171–172
 - design, 170
 - desired products selection, 172
 - mutations location, 168–171
 - novel genes, 173–174
 - primer, 168
 - random and extensive, 172–173
- DNA polymerase, 121–123
 - commercially available, 7
 - enzyme
 - LDV-PCR, 282
 - oligonucleotide inhibitors
 - hot start, 6
- DNA pools
 - BAC libraries, 404–406, 409
- DNA purification
 - PET
 - materials, 68
 - methods, 70–71
- DNA recombination, 167–174
- DNase-free RNase
 - BAC libraries, 409
- DNA sequence databases
 - web sites, 45
- DNA sequencing
 - GC-PCR, 365
 - oligonucleotides
 - STM, 226
- DNA shuffling, 173, 235–239
- DNA transfection
 - PCR-mediated recombination
 - materials, 199
- DNA transposon-flanking
 - cloning and sequence analysis, 232
- DNeasy Tissue Kit, 72
- dNTP, 3–4, 47
- Double-digested plasmid
 - synthetic gene in-frame
 - cloning, 186
- Double-stranded DNA tags, 230
- DpnI digestion and transformation
 - directional restriction site-free
 - insertion, 137
 - primer
 - length, 138
- E**
- Electroelution, 215
- Electroporation, 182
- Elongation time
 - bacteriophage λ sublibraries, 397
 - recombinant libraries rapid and nonradioactive screening, 383
- End-repair priming, 277
- Environmental pathogens
 - identification, 3
- Enzymatic modification-mediated
 - LIC, 116
- Enzyme, 6–7, 12
- Enzyme-free cloning, 125–131
 - amplicon
 - hybridization, 129–130
 - purification, 129
 - amplification purification, hybridization, and transformation reagents
 - materials, 127

- DNA amplification reagents
 - materials, 127
 - insert PCRs
 - methods, 128–129
 - materials, 127
 - methods, 128–130
 - vector PCRs
 - methods, 128
 - Epigenetic phenomenon
 - methylation-specific PCR, 91
 - Ethidium bromide fluorescence, 45
 - Ethidium-bromides-stained agarose gel electrophoresis, 15
 - Expression profiling
 - scheme, 310
 - Extension PCR amplification kits
 - sources, 162
 - Extra-long polymerase chain reaction (XL PCR), 37–48
 - amplification
 - materials, 38–39
 - methods, 41–42
 - annealing temperature, 48–49
 - carryover contamination, 49
 - control primer sets, 44
 - dTNP, 47
 - materials, 38–39
 - methods, 39–42
 - positive control primer sets, 45
 - primer design
 - materials, 38
 - methods, 40–41
 - product analysis, 42
 - reagent mix
 - preparation guidelines, 43
 - template integrity analysis
 - materials, 38
 - methods, 40
 - thermal cyclers, 48
 - total genomic DNA, 39–40
- F**
- Food pathogens
 - identification, 3
 - Foreign DNAs, 249
 - Forensics, 3
 - Formamide, 11
 - Freeze-Squeeze, 215
 - Fusion protein expression, 125–131
- G**
- GCG Prime, 22
 - GC-PCR. *see* Gene-capture PCR (GC-PCR)
 - GC rich template amplification
 - inverse PCR, 75–79
 - Gel electrophoresis
 - RC-PCR, 388
 - Gene-capture (GC) rich template amplification
 - inverse PCR, 75–79
 - materials, 76
 - methods, 76–77
 - Gene-capture PCR (GC-PCR), 359–374
 - DNA amplification, 365
 - DNA cloning and sequencing, 365
 - failure, 374
 - hot start, 373
 - incubation temperature, 373
 - materials, 364–365
 - methods, 366–372
 - southern blot
 - reagents, 365
 - GeneClean, 215
 - Gene cloning
 - scheme, 310
 - GeneFisher, 23
 - browsers, 29
 - degenerate PCR primer
 - primer design software
 - packages, 26–27
 - output, 28

- Gene Racer kit, 373
- Gene-Splicing Overlap Extension (Gene SOEing), 173
- Genome web site, 378
- Genomic DNA cloning
 - iPCR, 301–306
- Genomic DNA restriction enzyme digestion, 346
- Genomic library screening, 354–355
- GFP gene direct cloning
 - baculovirus genome, 155
- Glycerol, 11
- Google, 29
- H**
- Hairpin
 - primer sequence, 20
- HDR filters, 401
- Hemo-De Clearing Agent, 71
- Hexaminecobalt chloride RNA ligase
 - iPCR, 298
- High-density replica (HDR)
 - filters, 401
- High fidelity long PCR
 - PCR-mediated recombination methods, 200
- High-molecular-weight genomic DNA, 282
- Hind*III digest, 183
- HIV-1 clones
 - chimeric
 - PCR mediated recombination, 199–200
 - molecular
 - restriction digestion, 203
- Hot start, 5–6, 12–13
 - DNA polymerases
 - oligonucleotide inhibitors, 6
 - GC-PCR, 373
 - LDV-PCR, 283
 - modified DNA polymerases, 5–6
 - PCR-2, 194
 - process, 13–14
 - step down PCR, 348
- Human genome
 - organization, 254
- I**
- iAFLP, 309
- IHJ, 249
- Incubation temperature
 - GC-PCR, 373
- In-frame cloning
 - blunt-end restriction sites, 176
- Input file DNA.txt, 24
- Input file protein.txt, 25
- Insertional mutagenesis tags
 - rapid amplification of genomic DNA sequences, 325–334
- Integrand-host junction (IHJ), 249
- Internet
 - security, 29
- Interspersed repetitive sequence
 - PCR (IRS-PCR), 249, 253, 256, 266
- Introduced amplified fragment length polymorphism (iAFLP), 309
- Inverse PCR (iPCR), 253
 - agarose gel electrophoresis, 79
 - cDNA cloning, 249–250, 293–298
 - alkali denaturation, 298
 - circularization, 296, 297
 - diagram, 294
 - first-strand cDNA synthesis, 294, 296
 - materials, 294–296
 - methods, 296–297

- second-strand cDNA synthesis,
 - 295, 296–297
- T4 RNA ligase, 298
- DNA polymerase, 78
- GC-rich template amplification,
 - 75–78
- genomic DNA cloning, 301–306
 - alkaline denaturation, 306
 - circularization, 304–305
 - isolation, 303–304
 - materials, 303–304
 - methods, 304–305
 - restriction enzyme digestion,
 - 304–305
 - magnesium, 78
 - primer design, 77
 - product purification and ligation, 77
 - schematic flow diagram, 251
 - Tn5, 315–322
- In vitro combination
 - StEP, 235–239
- In vivo cloning (IVC), 115
- iOligo, 22
- iPCR. *see* Inverse PCR (iPCR)
- IRS-PCR, 249, 253, 256, 266
- IVC, 115
- L**
- LA-PCR, 249, 256–260
- LD-PCR, 99
- LDV-PCR. *see* Long distance vectorette-PCR (LDV-PCR)
- Library screening, 353–357
- LIC, 115–118
- Ligation-anchored PCR (LA-PCR),
 - 249, 256–260
- Ligation-dependent cloning method,
 - 112–115
 - blunt-end cloning, 112
 - selection, 111
 - sticky-end cloning, 113–114
 - T-A cloning, 114–115
- Ligation-independent cloning (LIC),
 - 115–118
 - enzymatic modification-mediated, 116
 - induced PCR, 115–116
 - positive selection system, 117–118
- Ligation-mediated-PCR (LM-OCR), 257
- Ligation-mediated single-sided PCR, 328
- LightCycler, 15
- LINE-PCR, 249, 256
- LM-OCR, 257
- Long-distance iPCR, 250
- Long distance PCR (LD-PCR), 99
- Long distance reverse-transcription PCR, 59–65
 - materials, 60–61
 - methods, 61–62
 - PCR reaction, 62
 - RNA template, 61
 - RT reaction, 61–62
- Long distance vectorette-PCR (LDV-PCR), 249, 257, 275–284
 - cloning
 - materials, 279
 - methods, 282
 - DNA extraction
 - materials, 278–279
 - methods, 279–282
 - primers
 - materials, 279
 - methods, 280
 - sequencing
 - materials, 279
 - methods, 281
 - vectorette libraries
 - materials, 278–279
 - methods, 279–280

- Long interspersed repetitive element PCR (LINE-PCR), 249, 256
 - Long IPCR (LR-iPCR), 249
 - LR-iPCR, 249
 - M**
 - Magnesium chloride, 9–10
 - Manual hot start, 5, 47
 - Marker coupled PCR (MCPCR), 172
 - Master mix, 34
 - LDV-PCR, 283
 - MAX Efficiency DH5 α Competent Cells, 215–223
 - MCPCR, 172
 - Megapore dialysis, 43
 - Megaprimer
 - length, 193
 - molar amounts, 193
 - publications, 189
 - Megaprimer PCR (MPCR), 171, 189–195
 - basic method, 190
 - materials, 191
 - methods, 191–192
 - mutant synthesis, 192
 - nontemplate nucleotides, 193
 - PCR-1, 191–192
 - PCR-2, 192
 - single-tube methods, 194
 - synthesis, 191–192
 - Melting temperature. *see* T_m (melting temperature)
 - Metal ion cofactors, 9–13
 - Methylation-sensitive endonucleases, 91
 - Methylation-specific PCR (MSP), 91–97
 - amplification efficiency, 96–97
 - application, 92–93
 - bisulfite DNA modification materials, 93–94
 - methods, 95–96
 - concept, 92–93
 - control templates, 97
 - epigenetic phenomenon, 91
 - primers
 - materials, 94
 - methods, 95
- Microdialysis, 184–185
 - Microtiter-plate-based high throughput PCR product purification, 417–420
 - agarose gel electrophoresis, 420
 - materials, 418
 - methods, 418–420
 - mini spin column in 96-well plates, 418–420
 - purification, 420
- Microtiter-plate-based template, 411–416
 - cross contamination, 415
 - DNA isolation materials, 412
 - DNA preparation materials, 411–412
 - methods, 413
 - DNA sequencing materials, 412
 - methods, 414–415
 - materials, 411–412
 - methods, 413–415
 - PCR amplification methods, 413–414
 - PCR product purification materials, 412
 - methods, 414
- Mineral oil
 - BAC libraries, 409

- Mini spin column in 96-well plates, 418–420
- Mini-Tn3 disruption mutations
 - ST-PCR, 331–332
- Minitransposon tagging, 230
- Mismatched degenerate primer stepdown PCR, 33
- Mismatched primer-template pairs
 - touchdown PCR, 32–33
- Modified DNA polymerases
 - hot start, 5
- Modified restriction site PCR (MRPCR), 171
- Monoclonal antibodies to DNA polymerases hot start, 5
- MPCR. *see* Megaprimer PCR (MPCR)
- MRPCR, 171
- MSP. *see* Methylation-specific PCR (MSP)
- MTV-PCR, 249, 257
- Multiple mutation generation, 207–215
 - materials, 210
 - methods, 211–214
 - mutations generation, 214
 - PCR purification and cloning, 211–214, 215
 - primer design, 211
 - principle, 208–209
 - unique restriction site, 211–212
- Multiplex coupled one-step RT-PCR methods, 55
- Multistep-touchdown vectorette-PCR (MTV-PCR), 249, 257
- Mutagenesis, 231. *see also* DNA mutagenesis and recombination
- Mutant synthesis
 - megaprimer, 192
- Mutations
 - DNA mutagenesis, 168–171
- Mutations generation
 - multiple mutation generation, 214
- Mx4000 Multiplex quantitative PCR System, 15
- N
- Negative control
 - bacteriophage λ sublibraries, 397
- NetPrimer, 23
 - primer analysis
 - primer design software packages, 27–28
- Nondegenerate PCR primer
 - primer design software packages, 21
 - Primer3, 23–26
- Nonionic detergents, 11
- Nonradioactive screening
 - recombinant libraries, 377–384
 - and recombinant libraries rapid, 378–383
- Nontemplate nucleotides
 - MPCR, 193
- Novel *Alu* element mediated PCR, 253–256
 - schematic diagram, 255
- Novel genes
 - DNA mutagenesis and recombination, 173–174
- NuSieve GTG agarose, 14
- O
- Octane, 71
- OEPCR, 168
- Oligo, 8, 22
- Oligo Analyzer, 23
- Oligonucleotide
 - DNA sequences
 - STM, 226

- Oligonucleotide design
 - synthetic gene in-frame cloning, 177–181
- Oligonucleotide primers, 64
 - enzyme-free cloning, 128
 - PCR-mediated recombination materials, 198
- Open reading frames (ORFs), 53
- Overlap extension PCR (OEPCR), 168
- Overlap sequence length
 - PCR-mediated recombination, 204–205
- P**
- Paraffin-embedded tissue (PET), 67–73
 - deparaffinization, 72
 - materials, 68–70
 - DNA purification, 68
 - methods, 69–71
 - DNA purification, 70–71
 - microdissection
 - materials, 68
 - methods, 69
 - PCR amplification
 - materials, 69
 - methods, 71
 - proteinase K digestion
 - materials, 68
 - methods, 70
 - target sequence, 73
 - thermal cycler, 73
- Paraffin-embedded tissue (PET) block
 - deparaffinization
 - materials, 68
 - methods, 69
- Partial inverse PCR (PI-PCR), 249–250, 252
- pBluescript II SK(+), 339
- pBluescript II SK9(+), 341
- PCR, 3–16. *see also* individual types
 - achievements, 6
 - after intramolecular circularization, 249, 250–253
 - amplification
 - step down PCR, 347
 - applications, 3
 - carryover prevention, 4–5
 - defined, 3
 - failures, 15–16
 - hot start, 5–6
 - induced
 - LIC, 115–116
 - inserts, 175–187
 - materials, 13
 - methods, 13–15
 - process, 4–5
 - products
 - analysis, 14–15
 - commercially available, 46
 - web site, 46
 - reaction
 - long distance reverse-transcription PCR, 62
 - samples, 9
 - amount, 9
 - purity, 9
 - types, 9
 - software program, 22
 - vessels
 - thermal cycling, 10–12
- PCR-1, 191–192
 - amplicon
 - gel electrophoresis, 221
- PCR-2, 192
 - amplicon
 - gel electrophoresis, 221
 - hot start, 194

- PCR-3, 221
 - cloned products
 - digestion, 220
- PCR Help!, 22
- PCR mediated recombination,
 - 197–205
 - asymmetric PCR recombination,
 - 200–203
 - chimeric HIV-1 clones, 199–200
 - clones
 - infectivity, 204
 - materials, 198, 199
 - sequence fidelity, 203
 - colony screening
 - materials, 199
 - DNA transfection
 - materials, 199
 - high fidelity long PCR, 200
 - materials, 198–199
 - methods, 199–204
 - overlap sequence length, 204–205
 - primers
 - materials, 198
 - schematic diagram, 201
 - sequence analysis
 - materials, 199
 - strains
 - materials, 198
 - virus replication assays
 - materials, 199
- Peptidyl-tRNA hydrolase encoding region amplification
 - directional restriction site-free insertion, 136–137
- PET. *see* Paraffin-embedded tissue (PET)
- PfuTurbo Hotstart DNA polymerase, 5
- Phage growth
 - recombinant libraries rapid and nonradioactive screening, 379
- Phenol-extraction, 43
- Physical barrier hot start, 5
- PI-PCR, 249–250, 252
- Plague hybridization, 354
- Plaque assay
 - transfection experiments
 - methods, 160–161
- Plasmid purification
 - AS-PCR, 148
- Plasmids
 - StEP, 237
- Polymerase chain reaction (PCR).
 - see* PCR
- Positive clones
 - BAC libraries, 407
- Positive control
 - RC-PCR, 387
- Primer, 8–9
 - DNA mutagenesis, 168
 - LDV-PCR, 279–280
 - temperature, 282–283
 - PCR-mediated recombination
 - materials, 198
 - Tn5 iPCR, 319, 321
- Primer3, 22, 23–26
 - web site, 73
 - default settings, 23, 25
 - output, 25
 - user-defined settings, 26
- Primer analysis
 - primer design software packages
 - NetPrimer, 27–28
 - software packages, 21
- Primer analysis software, 23
- Primer design, 35
 - AS-PCR, 143–144
 - cold-start method, 288
 - core parameters, 19–20
 - directional restriction site-free insertion, 136

- general rules, 20–21
- iPCR, 77
- multiple mutation generation, 211
- RC-PCR, 386
- recombinant libraries rapid and nonradioactive screening, 378–379
- StEP, 237
- stepdown PCR, 346–347
- T4 DNA polymerase methods, 122
- web sites, 45
- Primer design software packages, 19–29
 - degenerate PCR primer, 21
 - GeneFisher, 26–27
 - materials, 21–23
 - methods, 23–28
 - nondegenerate PCR primer, 21
 - Primer3, 23–26
 - primer analysis
 - NetPrimer, 27–28
 - web site, 8, 45
- Primer Generator, 22
- Primer Premier 5, 23
- Primer sequence
 - hairpin, 20
 - primer design, 20
 - self-dimer, 20
- Primer specificity
 - primer design, 19–20
- Primer synthesis
 - AS-PCR, 144–145
- Privacy
 - Internet, 29
- Proofreading DNA polymerase
 - step down PCR, 349
- Proteinase K digestion
 - PET
 - materials, 68
 - methods, 70
- Protein gene
 - cloning strategies, 154
- Pseudomonas aeruginosa* mutants
 - library construction, 227
- Publications
 - megaprimer, 189
- Pulse-field gel electrophoresis, 63
- Puregene DNA Isolation Kit, 43
- Pyrimidine residues
 - conversion, 93
- Q**
 - QiAexII Gel Extraction (QIAGEN)
 - genomic-tips, 43
 - kit, 149
 - QIAquick kit, 68
 - Quantitative PCR (QPC), 14–15
- R**
 - RACE system kit, 373
 - RAGE. *see* Rapid amplification of gene inserts (RAGE)
 - Random and extensive mutations (REM), 172–173
 - Rapid amplification of gene inserts (RAGE)
 - universal vector primers, 309–313
 - materials, 311
 - methods, 311–313
 - Rapid amplification of genomic DNA sequences
 - insertional mutagenesis tags, 325–334
 - materials, 330–331
 - methods, 331–333
 - Rapid and nonradioactive screening
 - recombinant libraries, 377–384
 - Rapid cDNA cloning by PCR
 - screening (RC-PCR), 385–388
 - gel electrophoresis, 388

- materials, 385
- methods, 386–388
- positive control, 387
- primer design, 386
- RC-PCR. *see* Rapid cDNA cloning by PCR screening (RC-PCR)
- Reamplification PCR, 287
- Recombinant arms, 153
- Recombinant libraries rapid and nonradioactive screening
 - agarose gel electrophoresis, 383
 - annealing temperature, 383
 - elongation time, 383
 - materials, 378
 - methods, 378–382
 - phage growth, 379
 - primer design, 378–379
 - single plaque screen, 382
- Recombinant PCR (RPCR), 171–172
- Recombination. *see* DNA mutagenesis and recombination
- RELAP-PCR, 262
- REM, 172–173
- REP, 315
- Repeat expansions
 - introduction, 217–223
- Repetitive extragenic palindrome (REP), 315
- Restriction digestion
 - AS-PCR
 - web site, 149
 - HIV-1 molecular clones, 203
 - iPCR
 - genomic DNA cloning, 304–305
- Restriction endonucleases
 - AS-PCR, 149
 - CAG-repeat containing genes, 222–223
 - step down PCR, 348
- Restriction enzymes
 - iPCR
 - genomic DNA cloning, 304–305
 - step down PCR, 347
- Restriction site
 - multiple mutation generation, 211–214
- Restriction site-free cloning scheme, 134
- Retroviral LTR-arbitrarily primed PCR (RELAP-PCR), 262
- Reverse transcription-polymerase chain reaction (RT-PCR), 53–58, 249, 262–264, 355
 - directional cloning, 57
 - negative controls, 57
 - product cloning
 - materials, 54
 - methods, 55–56
- Reverse transcription (RT) reaction
 - long distance reverse-transcription PCR, 61–62
- RNA
 - ligase
 - iPCR, 298
 - template
 - LD-RT-PCR, 61
 - total
 - extraction, 364
- Roche
 - web site, 46
- RPCR, 171–172
- RT-PCR. *see* Reverse transcription-polymerase chain reaction (RT-PCR)
- RT reaction
 - long distance reverse-transcription PCR, 61–62

S

- Salt-chloroform extraction, 44
- Salts, 10
- SD-PCR. *see* Stepdown PCR (SD-PCR)
- SDS-PAGE, 131
- Seakem GTG agarose, 14
- Search engines, 29
- Security
 - Internet, 29
- Self-dimer
 - primer sequence, 20
- Semi-random two-step PCR (ST-PCR), 325–327
 - mini-Tn3 disruption mutations, 331–332
 - REMI-generated disruption mutations, 332–333
- Sexual PCR, 173
- Short tandem repeats (STR), 81
- Signature-tagged mutagenesis (STM), 225–234
 - attenuated mutant
 - confirmation, 233
 - materials, 227–229
 - methods, 230–232
 - oligonucleotides
 - DNA sequences, 226
 - scheme, 228
 - steps, 225–226
 - tagged mini-Tn5 Km2, 230
 - tagged mutants libraries
 - construction, 231–232
 - transposon-flanking DNA
 - cloning and sequence analysis, 232
- Simple sequence repeats (SSR), 81
- Single plaque screen
 - recombinant libraries rapid and nonradioactive screening, 382
- Single-specific primer
 - Taq DNA polymerase buffer, 313
- Single-specific primer PCR (SSP-PCR), 309, 328–330
- Single-step PCR optimization³, 31–35
- Single-stranded biotinylated homologous probes
 - synthesis and immobilization, 366–368
- Single-strand ligation of cDNA (SLIC strategy), 373
- Site-directed mutagenesis, 207
- SLIC strategy, 373
- Sodium dodecylsulfate-polyacrylamide gel electrophoreses (SDS-PAGE), 131
- Software packages
 - primer, 8
 - primer analysis, 21, 23, 27–28
 - primer design, 19–29
- Southern blot, 57
 - capture steps, 369–370
 - iPCR
 - Tn5 transposon insertions, 316, 318
 - reagents
 - GC-PCR, 365
- Southern hybridization, 91
- Splinkerette-PCR, 257–260
- SSP-PCR, 309, 328–330
- SSR, 81
- Staggered extension process (StEP), 235–240
 - annealing/extension times, 238
 - illustrated, 236

- materials, 236
- methods, 237
- negative control reaction, 238
- plasmids, 237
- primer design, 237
- StEP. *see* Staggered extension process (StEP)
- Stepdown PCR (SD-PCR), 31–35, 343–349
 - hot start, 348
 - materials, 344–346
 - methods, 33, 346–347
 - mismatched degenerate primer, 33
 - PCR amplification, 347
 - primer design, 346–347
 - proofreading DNA polymerase, 349
 - restriction endonucleases, 348
 - restriction enzymes, 347
 - temperature, 35, 348–349
- Sticky-end cloning, 113–114
- STM. *see* Signature-tagged mutagenesis (STM)
- ST-PCR, 325–333
- STR, 81
- Stratagene
 - web site, 46
- Sublibrary method, 391–399
- Substrate analogs, 10
- Substrates, 10
- Subtractive cloning, 99
 - differentially expressed genes, 104
 - driver preparation, 103
 - efficiency, 102, 105
 - methods, 102–105
 - mRNA isolation, 102–103
 - subtractive hybridization, 104
 - tracer preparation, 103
- Suicide plasmids, 226
- Switch Mechanism, 373
- Synthetic gene in-frame cloning, 175–187
 - double-digested plasmid, 186
 - ligation, 185
 - materials, 177
 - methods, 177–184
 - oligonucleotide design, 177–181
 - vector insert screening, 186
- T**
- T-A cloning, 114–115
- Tagged mini-Tn5 Km2, 230
 - STM, 230
- Tagged mutants
 - in vivo screening, 231–232
- Tagged mutants libraries
 - construction
 - STM, 231–232
- TAIL-PCR. *see* Thermal asymmetric interlaced PCR (TAIL-PCR)
- TaKaRa
 - web site, 46
- Taq DNA polymerase buffer
 - single-specific primer, 313
- TaqStart, 5, 114–115
- Targeted gene, 329
- Targets, 12
- TCE, 133–134, 138
- T4 DNA polymerase, 121–123
 - end repair and cloning
 - materials, 122
 - methods, 123
 - materials, 122
 - methods, 122–123
 - PCR
 - materials, 122
 - PCR reaction
 - methods, 122
 - primer design
 - methods, 122

- TD-PCR. *see* Touchdown PCR (TD-PCR)
- TdT tailing, 373
- TEMAC, 75
- Temperature, 12
- SD PCR, 35
 - step down PCR, 348–349
- Temperature melting. *see* T_m (melting temperature)
- Terminal deoxynucleotide transferase (TdT) tailing, 373
- Tetramethylammonium chloride (TMAC), 11
- Tetramethylammonium chloride (TEMAC), 75
- T4 gene 32 protein, 11
- Thermal asymmetric interlaced PCR (TAIL-PCR), 250, 260–262, 329–330, 337
- Thermal cycler
- annealing stages, 34
 - PET, 73
 - TD PCR, 35
- Thermal cyclers
- XL PCR, 48
- Thermal cycling, 10–13
- Thermal -cycling elongation (TCE), 133–134
- annealing temperature, 138
- 384-well microtiter-plate-based template, 411–416
- Time optimization, 12
- TMAC, 11
- TMA oxalate, 11
- T_m (melting temperature)
- formula, 34
 - predictions, 8–9
 - primer design, 19, 28
- Tn5 inverse PCR, 315–322
- cloning and sequencing, 320–322
 - materials, 316–318
 - methods, 318–322
 - primers, 319, 321
 - southern blot, 316, 318
- TOGA, 309
- Total gene-expression analysis (TOGA), 309
- Total RNA
- extraction, 364
- Touchdown PCR (TD-PCR), 31–35
- materials, 32–33
 - methods, 33
 - mismatched primer-template pairs, 32–33
 - STM, 233–234
 - thermal cycler, 35
- Transfection experiments
- baculovirus cloning, 160
 - plaque assay
 - methods, 160–161
 - protein expression, 161
- Transposon-flanking DNA
- cloning and sequence analysis, 232
- Trinucleotide repeats isolation, 81–88
- adapter PCR
 - genomic DNA, 84–85
 - repeat sequences, 86–87
- enrichment library construction, 87
 - enrichment procedure, 85–87
 - materials, 82–84
 - methods, 84–97
- Triton X-100, 11
- T4 RNA ligase
- iPCR, 298
- Tween 20 NP40, 11

U

- Universal Resource Locator (URL)
 - stability, 29
- Universal vector primers
 - RAGE, 309–313
- Uracil-N glycosylase (UNG), 4–5

V

- Vectorette library, 278–280, 282
 - formation, 277
- Vectorette PCR, 257, 343
 - structure, 259
- Vectorette primer, 276
- Vectorette units, 275–276
 - central mismatch region, 276
- Vector preparation
 - AS-PCR, 145–146
- Vent, 112, 238
- Viral targets
 - identification, 3
- Virus replication assays
 - PCR-mediated recombination materials, 199

W

- War
 - human identification, 3
 - Web Primer, 22
 - Web sites
 - AS-PCR, 149
 - BLAST, 8
 - CLONTECH, 46
 - DNA sequence databases, 45
 - Genome, 378
 - long PCR products
 - commercially available, 46
 - Primer3, 73
 - default settings, 23–26
 - primer design, 45
 - primer design software packages,
 - 8, 45
 - Roche, 46
 - Stratagene, 46
 - TaKaRa, 46
 - 384-well microtiter-plate-based
 - template, 411–416
- X–Y**
- XL-PCR. *see* Extra-Long Polymerase Chain Reaction (XL PCR)
 - xprimer, 22
 - Xylene, 71
 - Yeast artificial-chromosome (YAC)
 - clones, 355

Series Editor: *John M. Walker*

PCR Cloning Protocols

Second Edition

Edited by

Bing-Yuan Chen and Harry W. Janes*Rutgers University, New Brunswick, NJ***From Reviews of the First Edition:***"...a balanced and easy to use guide to the principal contemporary options." —Microbiology Today**"...eliminat[es] many laborious and expensive techniques associated with traditional isolation and analysis." —Biochemical Education*

In the post-genomic era, PCR has become the method of choice not only for cloning existing genes, but also for generating a wide array of novel genes by mutagenesis and/or recombination within the genes of interest. *PCR Cloning Protocols, Second Edition*, updates and expands Bruce White's best-selling *PCR Cloning Protocols* (1997) with the newest procedures for DNA cloning and mutagenesis. Here the researcher will find readily reproducible methods for all the major aspects of PCR use, including PCR optimization, computer programs for PCR primer design and analysis, and novel variations for cloning genes of special characteristics or origin, with emphasis on long-distance PCR and GC-rich template amplification. Also included are both conventional and novel enzyme-free and restriction site-free procedures to clone PCR products into a range of vectors, as well as state-of-the-art protocols to facilitate DNA mutagenesis and recombination and to clone the challenging uncharacterized DNA flanking a known DNA fragment. Powerful applications of PCR in library construction and sublibrary generation and screening are also presented.

Authoritative and up-to-date, *PCR Cloning Protocols, Second Edition*, constitutes a gold-standard collection of the fastest, simplest, and most popular methods for isolating genes from all biological samples and creating novel genes from them by mutagenesis/recombination—essential methods for today's study of functional genomics, gene expression, protein structure–function relationships, protein engineering, and molecular evolution.

FEATURES

- Readily reproducible methods for isolating genes from all biological samples
- Cutting-edge techniques described by hands-on masters
- Proven methods for cloning uncharacterized DNA flanking a known DNA fragment
- Tested techniques for in vitro DNA recombination

SECTIONAL CONTENTS

Part I. Performing and Optimizing PCR. Part II. Cloning PCR Products. Part III. Mutagenesis and Recombination.

Part IV. Cloning Unknown Neighboring DNA. Part V. Library Construction and Screening.

Methods in Molecular Biology™ • 192
PCR CLONING PROTOCOLS
SECOND EDITION
 ISBN: 0-89603-969-2

humanapress.com

ISBN 0-89603-969-2



9 780896 039698

9 0000

