

MX

PCR

METHODS EXPRESS

The **METHODS EXPRESS** series

Series editor: B. David Hames

Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

Bioinformatics

Biosensors

Cell Imaging

DNA Microarrays

Expression Systems

Genomics

Immunohistochemistry

PCR

Protein Arrays

Proteomics

Whole Genome Amplification

METHODS EXPRESS

The editors would like to dedicate this book to the late Nat Bumstead,
a friend and mentor to us both.

*'Most of the fundamental ideas of science are essentially simple, and may, as
a rule, be expressed in language comprehensible to everyone.'* Albert Einstein

MX

MX

PCR

METHODS EXPRESS

edited by **S. Hughes**

*Centre for Tumour Biology,
Institute of Cancer and CR-UK Clinical Centre,
London, UK*

and **A. Moody**

AstraZeneca, Macclesfield, UK



Scion



© Scion Publishing Ltd, 2007

First published 2007

All rights reserved. No part of this book may be reproduced or transmitted, in any form or by any means, without permission.

A CIP catalogue record for this book is available from the British Library.

ISBN: 978 1 904842 28 6 (paperback)

ISBN: 978 1 904842 29 3 (hardback)

Scion Publishing Limited

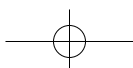
Bloxham Mill, Barford Road, Bloxham, Oxfordshire OX15 4FF

www.scionpublishing.com

Important Note from the Publisher

The information contained within this book was obtained by Scion Publishing Limited from sources believed by us to be reliable. However, while every effort has been made to ensure its accuracy, no responsibility for loss or injury whatsoever occasioned to any person acting or refraining from action as a result of information contained herein can be accepted by the authors or publishers.

Typeset by Phoenix Photosetting, Chatham, Kent, UK
Printed by Ajanta Offset and Packagings Ltd, Delhi, India



Contents

Contributors	xiii
Foreword	xvii
Preface	xviii
Abbreviations	xix
Color section	xxi
Chapter 1.	
An introduction to the polymerase chain reaction	
<i>Adrian Moody</i>	
1. Introduction	1
1.1 History of PCR	1
1.2 Utility of PCR – modifications and applications	3
2. Methods and approaches	4
2.1 Components of PCR	4
2.2 The PCR cycle	5
2.3 Thermal cyclers	6
2.4 Oligonucleotide primers	6
2.5 Standard PCR	8
2.6 Optimizing a PCR	11
2.7 Optimizing MgCl ₂ concentration	13
2.8 General hints and tips	17
3. Troubleshooting	18
4. References	20
Chapter 2.	
Polymerases for PCR	
<i>Meg Martel, Simon Baker, Ian Kavanagh, and Simon May</i>	
1. Introduction	21
1.1 Function of DNA polymerases	21
1.2 Origins of DNA polymerases	21

vi ■ CONTENTS

2. Methods and approaches	22
2.1 Unit definition of DNA polymerase	22
2.2 DNA polymerase activity	22
2.3 Choice of enzyme	22
2.4 Hot-start PCR	24
2.5 Buffer composition	26
2.6 Inhibitors	27
2.7 Detection methods for assessing enzyme activity	27
3. Troubleshooting	33
4. References	33

Chapter 3.**A detailed guide to quantitative RT-PCR***Pete Kaiser*

1. Introduction	35
2. Methods and approaches	36
2.1 Technologies available	36
2.2 Choice of reference gene	38
2.3 Analysis of real-time qRT-PCR data	38
2.4 Designing TaqMan primers and probes	41
2.5 Recommended protocols	41
2.6 Analyzing primer optimization results	43
3. Troubleshooting	46
4. References	47

Chapter 4.**Use of quantitative PCR for the detection of genomic microdeletions or microduplications***Simon Hughes, Rosanna Weksberg, Laura Moldovan, and Jeremy A. Squire*

1. Introduction	49
1.1 Test model	50
2. Methods and approaches	51
2.1 SYBR Green	51
2.2 Measuring amplification	51
2.3 Primer design	53
2.4 Housekeeping/reference genes and data normalization	54
2.5 Primer optimization	54
2.6 Generating a standard curve	54
2.7 Recommended protocols	56
2.8 DNA quantification and data analysis	58
2.9 Case study	58
2.10 Concluding remarks	60
3. Troubleshooting	61
4. References	62

Chapter 5.	
Robust and unique PCR for single-nucleotide polymorphism genotyping applications	
<i>Xiangning Chen</i>	
1. Introduction	63
1.1 Variable number of tandem repeats	63
1.2 Single-nucleotide polymorphisms	64
2. Methods and approaches	65
2.1 General guidelines for genotyping applications	65
2.2 SNP genotyping methods	68
2.3 General considerations	70
2.4 Recommended protocols	70
2.5 Typical results	75
2.6 Summary	76
3. Troubleshooting	77
4. References	79
Chapter 6.	
Using PCR and linkage mapping to identify single genes and quantitative trait loci for livestock traits	
<i>Jillian F. Maddox, Imke Tammen, and Sonja Dominik</i>	
1. Introduction	81
1.1 Genetic traits	81
1.2 Genetic markers	82
2. Methods and approaches	83
2.1 Genotyping	83
2.2 Data analysis: linkage analysis and trait mapping	100
2.3 Definition of terms	102
3. Troubleshooting	103
4. References	104
Chapter 7.	
PCR restriction fragment length polymorphism analysis for genotyping of single-nucleotide polymorphisms	
<i>Simon Hughes</i>	
1. Introduction	107
1.1 Restriction enzymes	107
1.2 RFLP	108
1.3 PCR-RFLP	108
2. Methods and approaches	109
2.1 Restriction enzyme identification	110
2.2 DNA extraction and PCR amplification	116
2.3 Results and data interpretation	119
3. Troubleshooting	119
4. References	121

viii ■ CONTENTS

Chapter 8.**Forensic genetic DNA typing with PCR-based methods***Claus Børsting, Juan J. Sanchez, and Niels Morling*

1. Introduction	123
1.1 PCR and forensic genetics	123
1.2 Mitochondrial DNA analysis	124
1.3 Single nucleotide polymorphisms	125
2. Methods and approaches	125
2.1 PCR fragment analysis	125
2.2 PCR optimization	125
2.3 PCR multiplexing	126
2.4 Recommended protocols	127
3. Troubleshooting	140
4. References	141

Chapter 9.**Large PCR multiplexes with special reference to forensic single-nucleotide polymorphism typing***Juan J. Sanchez, Claus Børsting, and Niels Morling*

1. Introduction	143
1.1 Genetic markers	143
1.2 SNP typing	144
1.3 SNP multiplexing	145
2. Methods and approaches	146
2.1 Multiplex design	146
2.2 Recommended protocols	148
2.3 Multiple-injection protocol	153
3. Troubleshooting	156
4. References	157

Chapter 10.**Pre-implantation genetic diagnosis of monogenic disease: PCR-based methods for the identification of mutations in single cells***Dagan Wells*

1. Introduction	159
1.1 Principles of PGD	159
1.2 Contamination	160
1.3 Amplification efficiency	162
1.4 Allele dropout	163
1.5 Whole genome amplification	163
2. Methods and approaches	165
2.1 Expected results from multiplex PCR	169
3. Troubleshooting	170
4. References	171

Chapter 11.**Rapid generation of gene-targeting constructs***Trevor J. Wilson, Dirk Truman, Antonietta Giudice, and Paul Hertzog*

1. Introduction	173
1.1 The function (or phenotype)-driven approach	173
1.2 The gene-driven approach	174
2. Methods and approaches	176
2.1 Obtaining gene sequences and identification of BACs	178
2.2 ET cloning procedure	182
2.3 Cloning of floxed exon	185
2.4 Preparation and electroporation of DNA	186
2.5 Screening of ES cell clones	195
3. Troubleshooting	195
4. References	196

Chapter 12.**Construction of long DNA molecules from multiple fragments using PCR***Nikolai A. Shevchuk and Anton V. Bryksin*

1. Introduction	197
2. Methods and approaches	197
2.1 Principles of long multiple fusion	197
2.2 Limitations of long multiple fusion	199
2.3 Factors critical for successful long multiple fusion	200
2.4 Recommended protocols	203
3. Troubleshooting	214
4. References	214

Chapter 13.**Efficient PCR-based mutagenesis method applicable to diverse mutagenesis strategies using type IIs restriction enzymes***Jae-Kyun Ko and Jianjie Ma*

1. Introduction	217
2. Methods and approaches	218
2.1 Principles of mutagenesis	218
2.2 Design of mutagenic primers and choice of type IIs restriction enzyme for mutagenesis	219
2.3 Application to diverse mutagenesis	224
2.4 Summary	225
3. Troubleshooting	225
4. References	228

Chapter 14.**Inverse PCR-based restriction fragment length polymorphism for identifying low-level mutations in tumors***G. Mike Makrigiorgos*

1. Introduction	229
-----------------	-----

x ■ CONTENTS

2. Methods and approaches	230
2.1 Principles of inverse PCR-based amplified restriction fragment length polymorphism	230
2.2 Size-separation analysis of PCR products and estimation of mutation frequencies	236
2.3 Sample results	236
3. Troubleshooting	240
4. References	240

Chapter 15.**PCR methods for infectious disease diagnosis**

Padmini Ramachandran, Andrew Hardick, Charlotte Gaydos, Samuel Yang, and Richard Rothman

1. Introduction	243
1.1 PCR for infectious disease diagnosis in a clinical setting	244
1.2 Antimicrobial resistance profiling	245
1.3 PCR and biodefense	245
2. Methods and approaches	245
2.1 Cost of PCR	245
2.2 False-negative and -positive results	246
2.3 Sample processing for PCR	247
2.4 Sample preparation from various human specimens	247
2.5 Recommended protocols	249
3. Troubleshooting	261
4. References	263

Chapter 16.**Use of PCR for DNA methylation analyses**

Mario F. Fraga and Manel Esteller

1. Introduction	265
2. Methods and approaches	266
2.1 PCR-based techniques	266
2.2 Summary	275
3. Troubleshooting	275
4. References	277

Chapter 17.**PCR-based methods to determine DNA methylation status at specific CpG sites using methylation-sensitive restriction enzymes**

Helmtrud I. Roach and Ko Hashimoto

1. Introduction	279
2. Methods and approaches	280
2.1 Methylation-sensitive restriction enzymes	280
2.2 Principle of the MSRE PCR method	281
2.3 Identifying CpG sites and suitable MSREs	282
2.4 Extraction of nucleic acids	283

2.5	Detection of methylation status using MSREs	286
2.6	Applications	289
3.	Troubleshooting	290
4.	References	291

Chapter 18.**PCR-based whole genome amplification***Nona Arneson, Simon Hughes, Richard Houlston, and Susan Done*

1.	Introduction	293
2.	Methods and approaches	294
2.1	DOP-PCR	296
2.2	PEP-PCR	299
2.3	Ligation-mediated PCR	301
2.4	Downstream applications	313
3.	Troubleshooting	314
3.1	General troubleshooting	314
3.2	I-PEP-PCR	316
3.3	PRSG	316
4.	References	316

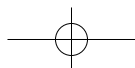
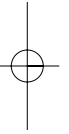
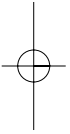
Chapter 19.**PCR sequencing of human genes for the discovery of DNA sequence variants***Abizar Lakdawalla*

1.	Introduction	319
1.1	DNA sequence variations	319
2.	Methods and approaches	320
2.1	Resequencing methods	320
2.2	Analysis of results	334
2.3	Mutation nomenclature	339
3.	Troubleshooting	340
4.	References	342

Appendix 1

List of suppliers	343
-------------------	-----

Index	345
--------------	------------



Contributors

Nona Arneson, Division of Applied Molecular Oncology, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario, M5G 2M9, Canada. E-mail: arneson@uhnres.utoronto.ca

Simon Baker, ABgene, Blenheim Road, Epsom Surrey KT19 9AP, UK; and School of Biological & Molecular Sciences, Oxford Brookes University, Gypsy Lane, Oxford OX3 0BP, UK. E-mail: simon.baker@brookes.ac.uk

Claus Børsting, Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, DK-2100 Copenhagen, Denmark.
E-mail: claus.boersting@forensic.ku.dk

Anton V. Bryksin, Department of Microbiology and Immunology, New York Medical College, Valhalla, NY, USA; and Institute of Chemical Biology and Fundamental Medicine, 8 Lavrent'eva St, Novosibirsk, 630090, Russia.
E-mail: anton_bryksin@gorodok.net

Xiangning Chen, Virginia Institute for Psychiatric and Behavioral Genetics and the Department of Psychiatry, Virginia Commonwealth University, Richmond, VA 23298, USA. E-mail: xchen@vcu.edu

Sonja Dominik, CSIRO Livestock Industries, Locked Bag 1, Armidale, New South Wales 2350, Australia. E-mail: sonja.dominik@csiro.au

Susan Done, Division of Applied Molecular Oncology, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario, M5G 2M9, Canada. E-mail: sdone@uhnres.utoronto.ca

Manel Esteller, Cancer Epigenetics Laboratory, Molecular Pathology Programme, Spanish National Cancer Centre (CNIO), Melchor Fernandez Almagro 3, 28029 Madrid, Spain. E-mail: mesteller@cnio.es

Mario F. Fraga, Cancer Epigenetics Laboratory, Molecular Pathology Programme, Spanish National Cancer Centre (CNIO), Melchor Fernandez Almagro 3, 28029 Madrid, Spain. E-mail: mffraga@cnio.es

xiv ■ CONTRIBUTORS

Charlotte Gaydos, Johns Hopkins University, Division of Infectious Diseases, Baltimore, MD, USA. E-mail: cgaydos@jhmi.edu

Antonietta Giudice, Monash Immunology and Stem Cell Laboratories, STRIP Building 75, Monash University, Wellington Rd, Clayton 3800, Victoria, Australia. E-mail: antonietta.giudice@med.monash.edu.au

Andrew Hardick, Johns Hopkins University, Division of Infectious Diseases, Baltimore, MD, USA. E-mail: ahardic1@jhmi.edu

Ko Hashimoto, Department of Orthopaedic Surgery, Tohoku University School of Medicine, Sendai, Japan. E-mail: hasshie@mail.tains.tohoku.ac.jp

Paul Hertzog, CRC for Chronic Inflammatory Diseases and Centre for Functional Genomics and Human Disease, Monash Institute of Medical Research, Monash University, 27–31 Wright St, Clayton 3168, Victoria, Australia. E-mail: paul.hertzog@med.monash.edu.au

Richard Houlston, Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Surrey, SM2 5NG, UK.

Simon Hughes, Centre for Tumour Biology, Institute of Cancer and CR-UK Clinical Centre, Bart's and The London, Queen Mary's School of Medicine and Dentistry, Ground Floor, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ, UK. E-mail: simon.hughes@cancer.org.uk

Pete Kaiser, Institute for Animal Health, Compton, Berkshire RG20 7NN, UK. E-mail: pete.kaiser@bbsrc.ac.uk

Ian Kavanagh, ABgene, Blenheim Road, Epsom, Surrey KT19 9AP, UK. E-mail: ian.kavanagh@thermofisher.com

Jae-Kyun Ko, Department of Physiology and Biophysics, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA. E-mail: koja@umdnj.edu

Abizar Lakdawalla, Applied Biosystems, 45 West Gude Drive, Rockville, MD 20850, USA. E-mail: abizar.a.lakdawalla@appliedbiosystems.com

Jianjie Ma, Department of Physiology and Biophysics, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA. E-mail: maj2@umdnj.edu

Jillian F. Maddox, Department of Veterinary Science, University of Melbourne, Victoria, Australia. E-mail: jillm@rubens.its.unimelb.edu.au

G. Mike Makrigiorgos, Dana Farber/Brigham and Women's Cancer Center, Harvard Medical School, Boston, MA 02115, USA. E-mail: mmakrigiorgos@lroc.harvard.edu

Meg Martel, ABgene, Blenheim Road, Epsom, Surrey KT19 9AP, UK.

E-mail: megm@abgene.com

Simon May, ABgene, Blenheim Road, Epsom, Surrey KT19 9AP, UK.

E-mail: simon.may@thermofisher.com

Laura Moldovan, Program in Genetics and Genomic Biology, The Research Institute, The Hospital for Sick Children, Toronto, Canada.

Adrian Moody, AstraZeneca, 19B13, Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK. E-mail: adrian.moody@astrazeneca.com

Niels Morling, Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, DK-2100 Copenhagen, Denmark.

E-mail: niels.morling@forensic.ku.dk

Padmini Ramachandran, Johns Hopkins University, Department of Emergency Medicine, Baltimore, MD, USA. E-mail: pramach1@jhmi.edu

Helmtrud I. Roach, Bone and Joint Research Group, University of Southampton, Southampton SO16 6YD, UK. E-mail: hr@soton.ac.uk

Richard Rothman, Johns Hopkins University, Division of Infectious Diseases and Department of Emergency Medicine, Baltimore, MD, USA.

E-mail: rrothman@jhmi.edu

Juan J. Sanchez, Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, DK-2100 Copenhagen, Denmark.

E-mail: jjsanchz@ull.es

Nikolai A. Shevchuk, Center for Cancer and Immunology Research, Children's Research Institute, Washington, DC, USA; and Institute for Biomedical Sciences/Program in Molecular and Cellular Oncology, The George Washington University, Washington, DC, USA. E-mail: shevchuk@hotmail.com

Jeremy A. Squire, Ontario Cancer Institute and Department of Laboratory Medicine, Pathology and Medical Biophysics, University of Toronto, Toronto, Canada. E-mail: jeremy.squire@utoronto.ca

Imke Tammen, Reprogen: Centre for Advanced Technologies in Animal Genetics and Reproduction, Faculty of Veterinary Science, The University of Sydney, 425 Werombi Rd, PMB 3, Camden, New South Wales 2570, Australia.

E-mail: itammen@camden.usyd.edu.au

Dirk Truman, Centre for Functional Genomics and Human Disease, Monash Institute of Medical Research, Monash University, 27-31 Wright St, Clayton 3168, Victoria, Australia. E-mail: dirk.truman@med.monash.edu.au



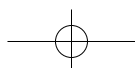
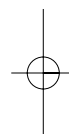
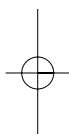
xvi ■ CONTRIBUTORS

Rosanna Weksberg, Program in Genetics and Genomic Biology, The Research Institute, The Hospital for Sick Children, Toronto, Canada; and Division of Clinical and Metabolic Genetics, Department of Paediatrics, The Hospital for Sick Children, Toronto, Canada. E-mail: rosanna.weksberg@sickkids.ca

Dagan Wells, Yale University Medical School, Department of Obstetrics and Gynecology, 333 Cedar Street, New Haven, Connecticut 06520, USA. E-mail: dagan.wells@yale.edu

Trevor J. Wilson, CRC for Chronic Inflammatory Diseases and Centre for Functional Genomics and Human Disease, Monash Institute of Medical Research, Monash University, 27–31 Wright St, Clayton 3168, Victoria, Australia. E-mail: trevor.wilson@med.monash.edu.au

Samuel Yang, Johns Hopkins University, Department of Emergency Medicine, Baltimore, MD, USA. E-mail: syang10@jhmi.edu



Foreword

Polymerase chain reaction (PCR) methods began to appear in the literature in the early 1980s, with a variety of adaptations of the basic technique emerging over the subsequent 20 years. These adaptations have relevance to many fields including basic research, clinical investigations, and forensic science. In this volume of the Scion book series *Methods Express*, we present a collection of some of the most interesting adaptations and applications of PCR. The chapters are supported by references covering the development, testing, and validation of each method, along with examples of research applications. References of particular interest are also indicated with asterisks to guide the reader to papers of importance in understanding the uses, capabilities, and expected performance of each method.

In any method-based book, it is important to realise that information is not knowledge. We once assumed the role of books was the transfer of knowledge, but in actual fact, what is being transferred is information, which requires a context for the reader to maximize learning. In the laboratory, if an experienced researcher sees you struggle, they can offer a solution or a technique they learned after being in a similar situation. The work of Michael Polanyi on learning (Polanyi was a chemist in early life) established a two-way look at knowledge: explicit knowledge is written down, as in a book or manual, whilst tacit knowledge is the result of experience and typically resides in the expert's mind and is difficult to collect and distil. The key lies in the ability to transfer expert, or tacit, knowledge, and in finding ways of capturing the lessons learned by experts after years of working in a scientific field, or with a particular technique or tool, such as PCR, and convey this as a written document. As PCR continues to grow in use and value across scientific disciplines, this book will provide an approach to pass the knowledge of most value on to readers working in the field.

PCR: Methods Express is an up-to-date compendium of techniques and approaches. More than a reference manual for PCR methods, it is also a deliberate attempt to capture and share tricks-of-the-trade, lessons learned, and 'simple solutions to common problems'. Tell me how to do this, but also tell me which of the possible approaches offered is best for my circumstances, and help me troubleshoot when things happen differently than expected. This volume provides just that: it provides the tacit knowledge of PCR.

Jeff Witherly
Author of *An A to Z of DNA Science: What Scientists Mean When They Talk About Genes and Genomes*

Preface

This book is intended to supply fundamental practical information for basic, clinical, and student researchers interested in using PCR methods in their research. It is structured not only to impart protocols, but also to illustrate the great variety of applications in which PCR plays a fundamental role.

The motivation for preparing this book came from the realization that an up-to-date, affordable book covering a wide array of the practical aspects of PCR techniques does not exist. The offerings of information in the text are intended to cater to the broad range of abilities among students, clinicians, and technologists, and hopefully will permit more exploratory experiments using this amazingly versatile tool.

We wish to thank all those who kindly gave of their time and skill to prepare the exceptional chapters herein. We would also like to acknowledge all the rest of the people who have made this book possible. We hope you find this text valuable and we welcome comments and ideas for future editions.

Simon Hughes & Adrian Moody
February 2007

Abbreviations

22q11DS	chromosome 22q11 deletion syndrome
ADO	allele dropout
BAC	bacterial artificial chromosome
BS	bisulfite sequencing
BSA	bovine serum albumin
CCD	charge-coupled device
CE	capillary electrophoresis
CGH	comparative genomic hybridization
CODIS	Combined DNA Indexing System
CSF	cerebrospinal fluid
C _T	cycle threshold
DMSO	dimethyl sulphoxide
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxynucleoside triphosphate
DOP-PCR	degenerate-oligonucleotide-primed PCR
EDNAP	European DNA Profiling Group
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EMPOP	EDNAP mitochondrial DNA population database
ePCR-RFLP	engineered PCR-RFLP
ES	embryonic stem
FISH	fluorescent <i>in situ</i> hybridization
FP-TDI	fluorescence polarization template-directed dye-terminator incorporation
FRET	fluorescence resonance energy transfer
FRT	Flp recognition target
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GBS	group B streptococci
HIV	human immunodeficiency virus
HPCE	high-performance capillary electrophoresis
HPLC	high-performance liquid chromatography
iFLP	inverse PCR-based amplified RFLP
I-PEP	improved PEP
LB	Luria-Bertani

xx ■ ABBREVIATIONS

LOD	logarithm of odds
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
MSP	methylation-specific PCR
MSRE	methylation-sensitive restriction enzyme
mtDNA	mitochondrial DNA
NTC	no-template control
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEP-PCR	primer-extension pre-amplification PCR
PGD	pre-implantation genetic diagnosis
PRSG	PCR of randomly sheared genomic DNA
PVA	polyvinyl alcohol
qPCR	quantitative PCR
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
RT	reverse transcriptase
RT-PCR	reverse transcriptase PCR
SAP	shrimp alkaline phosphatase
SBE	single-base extension
SCOMP	single-cell comparative genomic hybridization
SDS	sodium dodecyl sulfate
SNP	single-nucleotide polymorphism
STR	short tandem repeat
TEMED	tetramethylethylenediamine
VCFS	velocardiofacial syndrome
VNTRs	variable number of tandem repeats
WGA	whole genome amplification

CHAPTER 12

Construction of long DNA molecules from multiple fragments using PCR

Nikolai A. Shevchuk and Anton V. Bryksin

1. INTRODUCTION

When a DNA molecule has to be assembled from three or more unrelated fragments, it can often be more convenient to use a PCR-based method called long multiple fusion (1) instead of the traditional restriction enzyme/ligation cloning. If an error rate of about 1 per 7000 bp is acceptable, long multiple fusion can be employed to assemble a linear recombinant DNA molecule of up to 10 kb from five fragments, or of up to 20 kb long from three fragments, precisely and quickly. This linear recombinant DNA molecule can then be cloned into a plasmid vector, if necessary.

The method described in this chapter could prove useful for such applications as the development of multi-domain vaccines and the construction of plasmid vectors and gene-targeting vectors, as well as for the assembly of viral genomes for basic and vaccine research.

The technique offers more flexibility and is less time- and labor-consuming than traditional cloning, which can often lead to introduction of unwanted sequence at fusion points. The long multiple fusion method has proved itself over a wide range of DNA sources, sizes of fused fragments, and final application of the obtained recombinant molecules (2–8).

2. METHODS AND APPROACHES

2.1 Principles of long multiple fusion

The method consists of three major steps (see *Fig. 1*).

- Step 1. During the first step, the fragments are amplified from the original source using chimeric primers carrying overlapping sequence at their 5' ends.

PCR: *Methods Express* (S. Hughes and A. Moody, eds.)
© Scion Publishing Limited, 2007

198 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

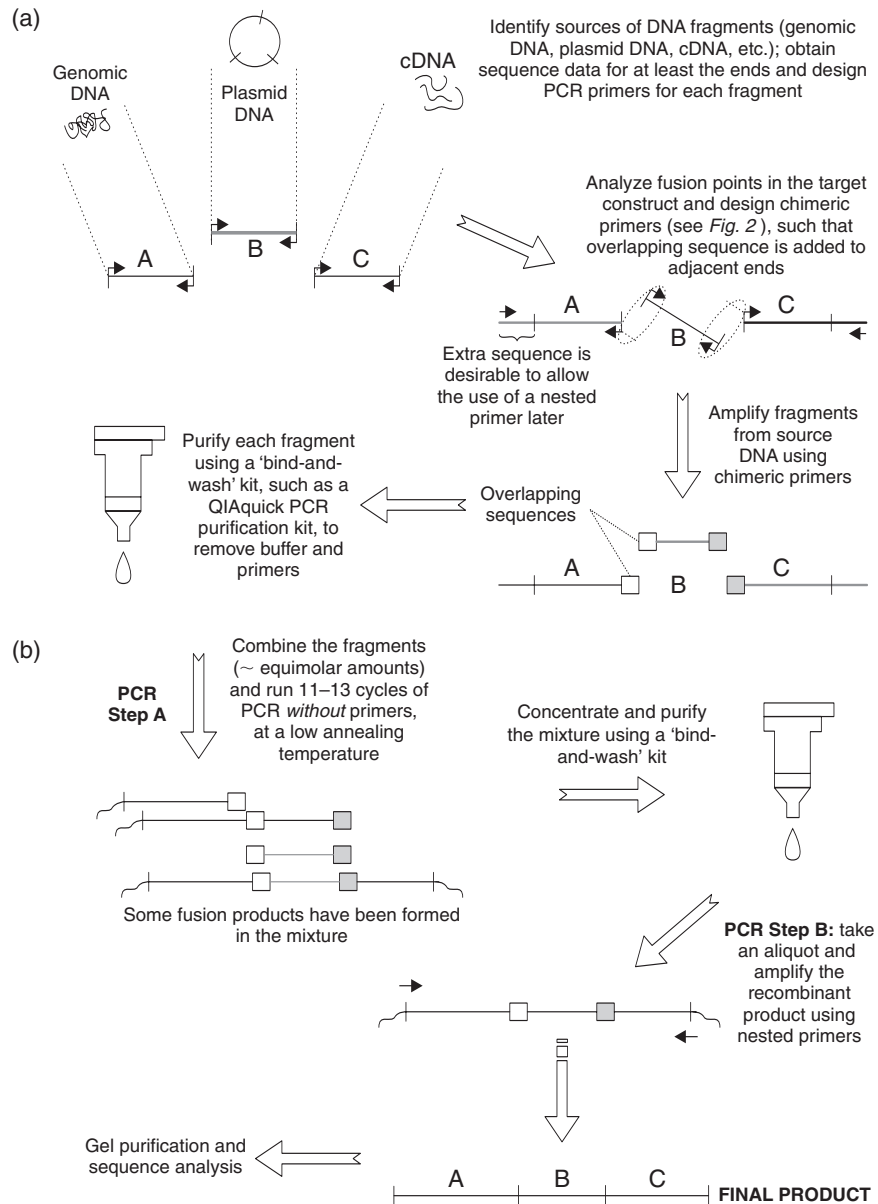


Figure 1. Outline of long triple fusion.

(a) Fragments are amplified from the original source using chimeric primers carrying overlapping sequence at their 5' ends. (b) Fragments are allowed to anneal and form intermediary products (PCR Step A) and, finally, the recombinant end product is amplified (PCR Step B). The correctness of the assembly is then verified using agarose gel electrophoresis and/or sequencing of critical regions.

- Step 2. Fragments are allowed to anneal and form intermediary products (PCR Step A in *Fig. 1*).
- Step 3. The recombinant end product is amplified (PCR Step B in *Fig. 1*)

The accuracy of the assembly can then be verified using agarose gel electrophoresis and/or sequencing of critical regions. The method has several pitfalls that need to be avoided in order to assemble the recombinant product successfully. All of the key points that are often overlooked by researchers who follow the method described in our original paper (1) are discussed in the section below and in section 3.

The long multiple fusion technique was born from the synthesis of long PCR with overlap-extension PCR (9, 10). Existing protocols of overlap-extension PCR are limited to regular (short) PCR, i.e. with a limit of about 3–4 kb (11, 12), and have been limited to the fusion of only two DNA fragments (13). Recently, others have described methods for the assembly of up to ten short fragments using modifications of overlap-extension PCR (14), but the length of the end product was limited to about 5.5 kb. Our long multiple fusion method allows the creation of recombinant products as long as 20 kb from three fragments (1), and has been used successfully to assemble 10 kb products from four (1) or five (A.V. Bryksin, unpublished data) fragments. Long multiple fusion can facilitate the construction of highly complex recombinant DNA molecules for various applications (2–8, 14). For example, potentially it can allow a vaccine researcher to create multiple custom-made viral genomes within a short time frame or develop complex multi-domain vaccines and other recombinant proteins. It can also be used for the assembly of sophisticated gene-targeting constructs (7, 12).

2.2 Limitations of long multiple fusion

PCR, even when performed using a proofreading enzyme, will lead to a small number of base changes (PCR errors). In a typical long multiple fusion set-up, we found the error rate to be below 1 base change per 6.6 kb (1). To ensure fidelity of the assembly and the absence of PCR errors in critical regions of a recombinant construct, sequencing of critical regions prior to assembly is advisable. Resequencing of critical regions in the final product is recommended. The expected error occurrence in the final product can be calculated from the total number of PCR cycles used throughout the whole procedure and the error rates of the PCR kits used. If the error rate (expressed in errors per megabase (Mb) per cycle) is multiplied by the total number of PCR cycles, this will yield the expected number of errors in the final product per Mb, e.g. 150 errors/Mb. If this value is then divided by 1000, this will give the error rate per kilobase (kb), i.e. 0.15 errors/kb. Finally, if this value is then multiplied by the length of the final product, it will give the number of expected errors in the final fusion product. A typical error rate of a long triple fusion procedure is approximately one error per 6.6 kb (1). If this error rate is unacceptable, then traditional restriction enzyme/ligation cloning should be used. The planning of a complex cloning project using the

200 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

traditional approach can be simplified with the software application VECTOR NTI ADVANCE (Invitrogen), available free of charge to academic users at the time of writing.

It should also be noted that a PCR product, even when gel purified, can contain a very small amount of nonspecific PCR products. To obtain an extremely pure product (which may not be necessary for many applications), the PCR fusion product may have to be cloned into a plasmid vector and verified by partial sequencing. For this purpose, TOPO vectors (Invitrogen) or T-vectors can be used.

2.3 Factors critical for successful long multiple fusion

- The component fragments shown at the top of *Fig. 1* have to be amplified with a polymerase that has a 3'→5' proofreading activity, for example *Pfu* polymerase, rather than *Taq* polymerase. A combination of *Taq* with *Pfu*, as found in long-PCR kits, is also suitable. The problem with *Taq* polymerase is that it leaves single A nucleotide overhangs at the 3' ends, which will disrupt priming of the overlapping regions during the overlap extension reaction (15), shown in *Fig. 1* (PCR Step A). An enzyme with proofreading activity will generate PCR products that have predominantly blunt ends (without extra A nucleotides added at the end) allowing successful overlap extension among DNA fragments (see *Protocol 2*). The following is a list of PCR kits that we have used successfully for amplification of component fragments:

- TripleMaster polymerase (Eppendorf)
- Advantage-HF 2 PCR kit (Clontech)
- Herculase HotStart DNA polymerase (Stratagene)
- EXL polymerase PCR kit (Stratagene)
- Long Template Expand polymerase (Roche)

Avoid using the DeepVent polymerase kit, even though it is mentioned in the original publication (1); this kit requires extensive optimization and can cause many problems.

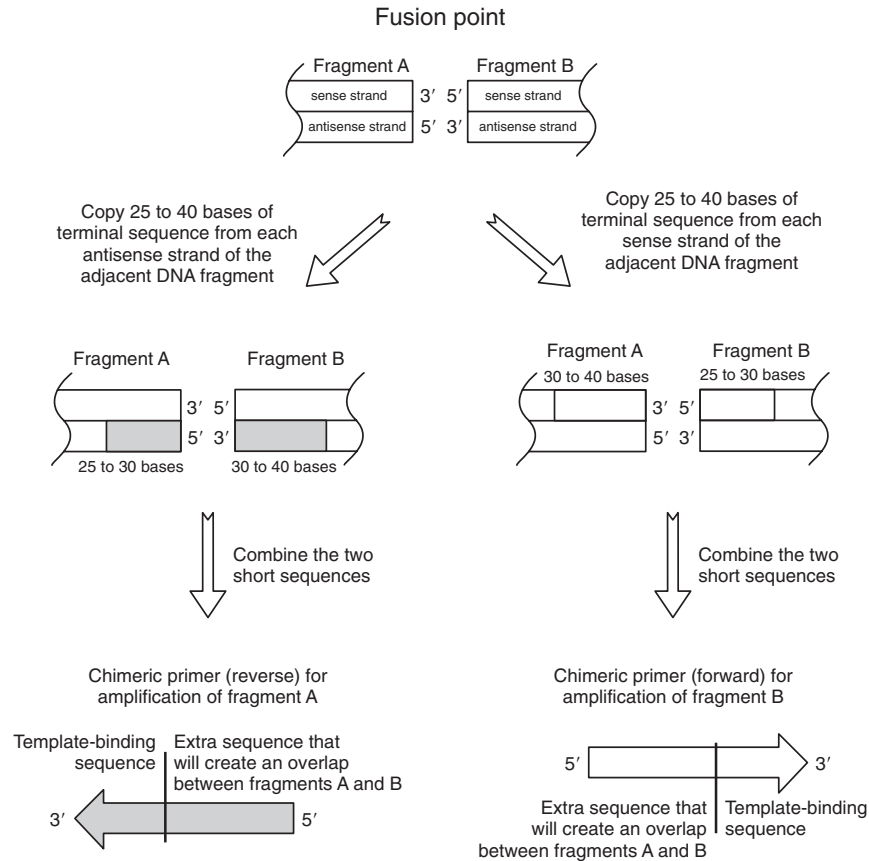
- Exposure of the PCR products to UV light should be avoided and all purification steps should be limited to 'bind-and-wash' or desalting methods, such as a QIAquick PCR purification kit (Qiagen). We found that UV light and handling of agarose slices containing DNA bands will make the DNA template unusable for long multiple fusion, most likely due to damage and nicking of the DNA. An exception can be made for gel purification when *both* of the following conditions are true:

1. The length of the final product is below 10 kb.
2. Crystal violet (Sigma) is used in the agarose gel instead of ethidium bromide.

Crystal violet allows visualization of DNA bands without UV light, thus minimizing damage to the DNA template. A 10 kb final product was successfully assembled from five fragments that were gel purified using agarose gels containing crystal violet (A.V. Bryksin, unpublished data).

- No primers should be used in PCR Step A in *Fig. 1*. This will allow the overlaps to anneal and extend forming various fusion products, including the target product. Although omission of this step can sometimes successfully produce a fusion product (12, 14), in our experience, inclusion of this step produces reliable results. The 'bind-and-wash' purification steps before and after PCR Step A partially remove residual primers from the previous amplification reactions and thus are also critical for success. Gel purification with crystal violet visualization (instead of UV and ethidium bromide) can be used for purification of fragments before PCR Step A, but *only if* the length of the final recombinant product is less than or equal to 10 kb. We have not tested whether this works for 20 kb fusions.
- The length of overlapping sequence (see *Fig. 2*) for each fusion point should be at least 20–30 nt when the final product is under 3 kb, 35–40 nt when the final product is 3–10 kb, and 50–70 nt when the final recombinant product is greater than 10 kb. *Fig. 2* shows the design of two chimeric primers for one fusion point. The length of overlapping sequence equals the sum of the extra sequences included in the two chimeric primers shown in *Fig. 2*.
- It is possible and often desirable to design only one chimeric primer per fusion point and to use a regular primer on the other side. In our experience, chimeric primers do not work well when the fragment to be amplified is 7 kb or longer. Also, chimeric primers may not work when the fragment has to be amplified from genomic DNA. In this case, it is preferable to use regular (nonchimeric) primers to amplify this problem fragment and to use chimeric primers for amplification of the adjacent fragments. In this case, in *Fig. 2*, a chimeric primer would be used for amplification of fragment B and a regular primer would be used for amplification of fragment A. If only one chimeric primer will be used for a fusion point, it should contain twice the length of the extra sequence (overlap) that would normally be required for a fusion point assembled using two chimeric primers.
- It is strongly recommended that nested primers (20–50 nt away from the terminus) are used during amplification of the final product (PCR Step B in *Fig. 1*). This will improve the purity of the final product and also decrease the chance of failure of the project. Proofreading polymerases are believed to degrade PCR products at the ends due to their exonuclease activity. This can prevent binding of the original terminal primers (top of *Fig. 1*) to the final product (bottom of *Fig. 1*) because the homologous sequence at the ends has been degraded. This is why it is desirable to use nested primers in the final amplification step.
- Long PCR in general is much more sensitive to the quality of the template (16) than regular short PCR. Repeated freezing and thawing and/or storage in distilled water, rather than in a buffer with a pH above 7, will also render a template unusable for long PCR (16). It is recommended that DNA templates for long PCR are stored at 4°C in 5 mM Tris/HCl (pH 8 or 8.5). This rule is less critical for projects in which all fragments are shorter than 5 kb.

202 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

**Figure 2. Designing chimeric primers.**

Only adjacent ends of two DNA fragments are shown in the figure. One needs to know the terminal DNA sequence of the adjacent fragments in a prospective fusion point in order to design chimeric primers. Typically, two chimeric primers are designed for each fusion point. One chimeric primer will serve as a reverse primer for amplification of fragment A and this chimeric primer is created by combining short terminal sequences from *complementary* DNA strands of fragments A and B, as shown in the flowchart directed to the left. The forward primer for amplification of fragment A is not shown in the figure. The second chimeric primer will serve as a forward primer for amplification of fragment B and this chimeric primer is created by combining short terminal sequences from *direct* DNA strands of fragments A and B as shown in the flowchart directed to the right. The reverse primer for amplification of fragment B is not shown in the figure. A DNA sequence editor and/or primer design software would be helpful for this task. The length of overlapping sequence for each fusion point should be at least 40 nt when the final product is under 10 kb, and 50–70 nt when the final recombinant product is longer than 10 kb. The total length of overlapping sequence equals the sum of the lengths of extra sequences included in the two chimeric primers. For example, if the reverse chimeric primer has 40 nt of extra sequence in it, whilst the direct chimeric primer has 30 nt of extra sequence, then, after fragments A and B are amplified using those chimeric primers (and two other primers not shown in the figure), fragments A and B will have an overlap of $30 + 40 = 70$ bp.

2.4 Recommended protocols

2.4.1 Long triple fusion

The protocols described below can be used for the creation of linear DNA molecules of 3–20 kb from three fragments.

Primer design

To design chimeric primers, it is essential to know the sequence of each DNA fragment, or at least 40–50 bp of sequence at each end of every fragment if the complete sequence is unavailable.

For amplification of fragment A, identify the desired fusion point between the fragments of interest (fragments A and B in *Fig. 2*) and select 25–30 nt from the antisense strand of fragment A (as if you are designing a normal reverse primer for amplification of this DNA fragment). The exact number of terminal nucleotides will depend on the melting temperature of the resulting oligonucleotide, which should be between 62 and 68°C. Remember that the 5' end of your oligonucleotide is fixed at the end of fragment A (fusion point) and you can only vary the 3' terminus of the oligonucleotide.

Now add 30–40 nt of sequence from the terminus of fragment B (also from the complementary strand, as shown in the downward flow chart in *Fig. 2*) to the 5' end of this reverse primer. The combined oligonucleotide is a reverse chimeric primer for amplification of fragment A. Design a forward primer for fragment A (not shown in *Fig. 2*). This should be a regular primer if fragment A is the first fragment in the construct. The extra sequence included in the reverse chimeric primer will ensure that after you have amplified fragment A from its source DNA, the amplicon will contain additional sequence that overlaps fragment B.

Use a similar approach to design a forward chimeric primer for amplification of fragments B and C. Make sure that the melting temperature of the template-binding moiety of each chimeric primer (see *Fig. 2*) is around 62–68°C. During amplification of the component fragments (see *Protocol 1* and top of *Fig. 1*), a high annealing temperature should be used (62–68°C). This will ensure a high specificity of PCR and a low level of side-products, as a high annealing temperature minimizes mispriming and the formation of secondary structures (15).

For each fusion point, make sure that the total length of extra 5' sequence included in the chimeric primer(s) is at least:

- 20–30 nt when the final product is under 3 kb
- 35–40 nt when the final recombinant product is less than or equal to 10 kb
- 50–70 nt when the final recombinant product is between 10 and 20 kb long

When designing primers for generating amplification fragments, it is important to include an additional 20–50 nt at the extreme 5' end of fragment A and at the 3' end of fragment C (or the final fragment) in order to allow for nested primers at the final amplification step (see top and bottom of *Fig. 1*).

The steps outlined above should provide two primers for each DNA fragment

204 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

and at least one chimeric primer for each fusion point. When ordering primers, request polyacrylamide gel electrophoresis (PAGE) purification for chimeric primers. If chimeric primers (long oligonucleotides) are not PAGE purified, they will contain erroneous oligonucleotides (resulting from the shortcomings of automated oligonucleotide synthesis), which can introduce base changes and deletions in the final product.

Protocol 1

Amplification of fragments from source DNA

Equipment and Reagents

- One of the following PCR kits:
 - TripleMaster polymerase (Eppendorf)
 - Long Template Expand polymerase (Roche)
 - Herculase HotStart polymerase (Stratagene)
 - EXL polymerase (Stratagene)
 - Advantage-HF 2 PCR kit (Clontech)^a
- Nuclease-free water (Promega)
- DMSO (Sigma)
- Thermal cycler
- 10 mM Ultrapure dNTP mix (Sigma)
- GM3 synthase gene BAC clone DNA (50 ng/μl)
- GM3 synthase forward primer: 5'-TCTGAGAGTAACTGCCCTCTTGACATC-3' (50 μM)
- GM3 synthase reverse primer: 5'-CATCTTGCTTTGAGCTCGGGTG-3' (50 μM)
- p3XFLAG-CMV-9 vector (1 μg/μl; Sigma)
- p3XFLAG-CMV-9 vector forward primer (chimeric): 5'-GTGATTGCTCGAGGCCTCCCTGCAATGGTACACCCGAGCTCAAAGCAAGATGATTGAACAAGATGATTGCACGCAGGTTC-3' (50 μM)
- p3XFLAG-CMV-9 vector reverse primer (chimeric): 5'-ATGCATTTTTTTCATGTCACATTCTCAGTAGTATAATTAACCTTGAGGATATAAAGGATCCCACTCCA GGGAAATTGATCCAGACATGATAAGATACA-3' (50 μM)
- Human genomic DNA (0.1 μg/μl)
- Example forward primer: 5'-TGGAGTGTGGATCCTTATATCC-3' (50 μM)
- Example reverse primer: 5'-AGACCTTCTTCTGCCCATATACATC-3' (50 μM)

Method

1. Amplify each of your three DNA fragments using a typical protocol recommended by the manufacturer of your PCR kit(s)^{b,c}.
2. For the generation of a 20 kb fusion, use the following protocols.
3. For fragment A (11.2 kb), combine:
 - 33 μl of nuclease-free water
 - 5 μl of reaction buffer (Herculase HotStart DNA polymerase kit)
 - 2.5 μl of dNTPs (10 mM)
 - 5 μl of human GM3 synthase gene BAC clone DNA
 - 1 μl of GM3 synthase forward primer
 - 1 μl of GM3 synthase reverse primer
 - 1 μl of Herculase HotStart polymerase
 - 1.5 μl of DMSO

4. Place the tube in a thermal cycler and run the following program:
 - Initial denaturation step at 92°C for 1 min
 - 27 cycles of denaturation at 92°C for 10 s, annealing at 65°C for 30 s, and extension at 68°C for 11 min 30 s (plus automatic extension of the extension time by 5 s per cycle)
 - Final additional extension step at 68°C for 13 min
 - Hold at 4°C
5. For fragment B (1.7 kb), combine:
 - 36.7 µl of nuclease-free water
 - 5 µl of HF buffer (Advantage-HF 2 PCR kit)
 - 5 µl of dNTPs (10 mM)
 - 0.3 µl of p3XFLAG-CMV-9 vector
 - 1 µl of p3XFLAG-CMV-9 vector forward primer
 - 1 µl of p3XFLAG-CMV-9 vector reverse primer
 - 1 µl of HF polymerase
6. Place the tube in a thermal cycler and run the following program:
 - Initial denaturation step at 94°C for 30 s
 - 26 cycles of denaturation at 94°C for 15 s, annealing at 65°C for 40 s, and extension at 68°C for 1 min 50 s
 - Final additional extension step at 68°C for 3 min
 - Hold at 4°C
7. For fragment C (7.5 kb), combine:
 - 35 µl of nuclease-free water
 - 5 µl of reaction buffer (Herculase HotStart DNA polymerase kit)
 - 2.5 µl of dNTPs (10 mM)
 - 3 µl of human genomic DNA
 - 1 µl of example forward primer
 - 1 µl of example reverse primer
 - 1 µl of Herculase HotStart polymerase
 - 1.5 µl of DMSO
8. Place the tube in a thermal cycler and run the following program:
 - Initial denaturation step at 92°C for 1 min
 - 26 cycles of denaturation at 92°C for 10 s, annealing at 60°C for 30 s, and extension at 68°C for 7 min 30 s (plus automatic extension of extension time by 4 s per cycle)
 - Final additional extension step at 68°C for 9 min
 - Hold at 4°C

Notes

^aThis kit will only amplify fragments shorter than 4 kb, but it offers the best PCR error rate.

^bIt is OK to use different kits for different fragments as long as they are from the list of recommended PCR kits above.

^cBe sure to use a high annealing temperature (62–68°C) in your PCRs, especially the ones that include chimeric primers, in order to minimize the level of nonspecific PCR products.

206 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

Protocol 2 serves three purposes:

1. To remove primers and buffer from a PCR product.
2. To analyze the length of the PCR product.
3. To quantify the PCR product.

We usually use crystal violet instead of ethidium bromide for staining of agarose gels as it is safer and does not require UV exposure, but this protocol will also work if ethidium bromide is used throughout instead of crystal violet. If one of your fragments is shorter than 0.5 kb, which is rare for a long fusion project, then please see note e (*Protocol 2*).

Protocol 2

Analysis and purification of each fragment

Equipment and Reagents

- Crystal violet (2 mg/ml; Sigma)
- 1% (or 0.5% for long PCR products) agarose gel, containing 30 μ l of the crystal violet stock solution per 150 ml of agarose.
- 5 \times Loading buffer (10% Ficoll 400, 0.1 M EDTA, pH 8.0, 120 μ g/ml crystal violet)
- White light box and camera
- Apparatus required for electrophoresis
- 1 \times TAE running buffer (4.84 g/l Tris base, 1.142 ml/l glacial acetic acid, 0.372 g/l EDTA in distilled water)
- QIAquick PCR purification kit (Qiagen)
- QIAquick gel extraction kit (Qiagen) (optional)
- DNA molecular weight markers, such as a 1 kb DNA step ladder (Promega)
- Spectrophotometer

Method

1. Purify each PCR product using a 'bind-and-wash' DNA purification kit, such as a QIAquick PCR purification kit^a, following the manufacturer's instructions.
2. Elute the purified PCR product with 30 μ l of 5 mM Tris/HCl (pH 8.5) (this is a twofold dilution of the Qiagen elution buffer supplied with the kit)^b.
3. Load and run 5 μ l of each PCR product (see *Protocol 1*) on your agarose gel. Also include a lane with 0.5 μ g of DNA marker in 10 μ l of the same loading buffer as the PCR samples^c.
4. Use a white light table (instead of UV light)^d to view the results of the electrophoresis^e.
5. Assess whether your PCR products are the correct length.
6. Quantify your PCR products either visually (by comparing the intensity of PCR bands with the intensity of DNA marker bands) or by using spectrophotometry to calculate approximate concentrations of the PCR products.

Notes

^aThis purification is necessary in order to remove (at least partially) the primers and buffer from the previous PCR.

^bThis is optional: if the length of your final recombinant product is less than or equal to 10 kb, you may use gel purification using a crystal violet-containing agarose gel and a gel extraction kit such as a QIAquick gel extraction kit (Qiagen). In this case you can skip the 'bind-and-wash' purification step in step 1, but may have to repeat electrophoresis in order to quantify your DNA samples. The gel purification procedure will completely remove previous primers and primer dimers from your PCR products.

^cThis is necessary for correct visual quantification of DNA amounts in PCR samples.

^dYou may use the regular ethidium bromide-stained agarose gels and UV light throughout this protocol for analysis of PCR results. However, if you decide to use gel purification of your PCR products (see note b), ethidium bromide and UV light should be avoided as they will damage the DNA and may lead to failure of the multiple fusion procedure.

^eYou may have a problem viewing fragments smaller than 0.9 kb using crystal violet. In this case, increase the amount loaded on the gel two- to threefold and use a dark room to view the results. It is almost impossible to see DNA shorter than 0.5 kb using crystal violet. If this is the case, use 0.75 μ l of ethidium bromide (10 mg/ml stock) per 150 ml of agarose throughout this protocol instead of crystal violet.

Protocol 3

Overlap extension reaction (see Fig. 1, PCR Step A)

Equipment and Reagents

- One of the PCR kits mentioned in *Protocol 1*
- Nuclease-free water (Promega)
- DMSO
- Thermal cycler
- 10 mM Ultrapure dNTP mix (Sigma)

Method

1. Prepare a 50 μ l PCR^a. Combine equimolar^b amounts of all three purified fragments (from *Protocols 1* and *2*) in a final volume of 10 μ l, but *do not* add primers. An example for the generation of a 20 kb reaction is shown below. Combine:
 - 29 μ l of water
 - 5 μ l of reaction buffer (Herculase HotStart DNA polymerase kit)
 - 2.5 μ l of dNTPs (10 mM)
 - 3 μ l of 11 kb fragment A (total ~0.45 μ g) in 5 mM Tris/HCl (pH 8.5)
 - 1.5 μ l of 1.7 kb fragment B (total ~0.07 μ g) in 5 mM Tris/HCl (pH 8.5)
 - 5.5 μ l of 7.5 kb fragment C (total ~0.31 μ g) in 5 mM Tris/HCl (pH 8.5)
 - 1 μ l of Herculase HotStart polymerase
 - 2.5 μ l of DMSO
2. Run the reaction for 11–15 cycles using an annealing temperature of 60°C^c and the following program:
 - Initial denaturation step at 92°C for 1 min
 - 13 cycles of denaturation at 92°C for 10 s, annealing at 60°C for 1 min, and extension at 68°C for 21 min^d
 - Final additional extension step at 68°C for 21 min
 - Hold at 4°C

208 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

Notes

^aIf a larger final volume is required, i.e. 200 μl , set up four separate 50 μl reactions and combine following cycling.

^bEquimolar concentrations of fragments ensure that the overlapping sequence is as likely to anneal to an adjacent fragment as to its complementary strand in the same fragment. If one of the fragments has a higher molar concentration than the others, the overlapping sequence that it contains will be more likely to anneal with the complementary strand of the same fragment than with the other fragments. In the example, we have three purified DNA fragments of 11, 1.7, and 7.5 kb. The DNA concentration of fragments A, B, and C are 0.150, 0.047, and 0.056 $\mu\text{g}/\mu\text{l}$, respectively. The fragments should be in the following proportions by weight to achieve equimolar concentrations: 11 : 1.7 : 7.5 (molecular weight of DNA is roughly proportional to its length). From these weight proportions, we have to calculate the volumes of fragments given their known concentrations. We need to solve the equation:

$$V = V_A + V_B + V_C = (11 * M/C_A) + (1.7 * M/C_B) + (7.5 * M/C_C)$$

where V is the sum of volumes of all DNA fragments (in μl ; 10 μl in our case); V_A , V_B , and V_C are the volume of fragments A, B, C in μl , respectively; C_A , C_B , and C_C are the concentrations of fragments A, B, and C in $\mu\text{g}/\mu\text{l}$, respectively; and M is the $1/P$ fraction of the total amount of DNA in the reaction (in μg).

The value of P is calculated by adding up the values in the weight proportions above (or just the lengths of the DNA fragments); therefore, $P = (11 + 1.7 + 7.5) = 20.2$. Hence, M is defined as approximately $1/20$ of the total amount of DNA. It is not difficult to calculate M , which equals 0.041 μg . Now the formula for calculation of the final volume of each DNA fragment (corresponding to equimolar amounts of DNA fragments in the reaction) is:

$$V_x = L_x * M/C_x$$

where L_x is the length of a fragment (A, B, or C) in kilobases, M is the value we calculated above, and C_x is the concentration of the fragment (A, B, or C) in $\mu\text{g}/\mu\text{l}$.

For example, for fragment A:

$$\begin{aligned} V_A &= L_A * M/C_A \\ V_A &= 11 * 0.041/0.150 \\ V_A &= 3 \end{aligned}$$

Therefore, we have volumes of 3, 1.5, and 5.5 μl for fragments A, B, and C, respectively. The higher the total amount of DNA fragments that you use in this overlap-extension step, the better the outcome of the fusion; the recommended minimum total amount of DNA of all fragments in a 50 μl reaction is around 0.4 μg . Therefore, you should try to obtain the highest concentration of each fragment that is possible in *Protocol 2*. You could use one QIAquick column to process two 50 μl reactions of each fragment and elute the DNA with 30 μl of 5 mM Tris buffer.

^cThis should work for the annealing of 40 nt or longer overlaps. For troubleshooting, you can lower the annealing temperature to 55°C.

^dUse an extension time that corresponds to the total length of the final recombinant product, i.e. 1 min per kb for most PCR kits. For example, if the length of your target product is 10 kb, use a 10 min extension time in your cycling program. Theoretically, overlap extensions cover only part of the total length of the final product, but we find that using an extra extension time leads to better results.

Protocol 4

Purification of intermediary product^a

Equipment and Reagents

- QIAquick PCR purification kit (Qiagen)

Method

1. Purify each PCR product using a 'bind-and-wash' DNA purification kit, such as a QIAquick PCR purification kit^a, following the manufacturer's instructions^b. For the elution step, use 30 μ l of 5 mM Tris/HCl (pH 8.5)^c.

Notes

^aThis purification step is optional if you used crystal violet gel purification in *Protocol 2* as gel purification removes all primers completely. You may still want to do this step to concentrate your sample.

^bIf a reaction volume larger than 50 μ l was generated from *Protocol 3*, load the separate 50 μ l reactions sequentially onto a single QIAquick column to achieve the highest concentration. It is not necessary to run agarose gel electrophoresis at this point, as you will not see your fusion product here due to its very low concentration.

^cThis is a twofold dilution of the Qiagen elution buffer supplied with the kit.

Protocol 5

Amplification of the final recombinant product^a

Equipment and Reagents

- One of the PCR kits mentioned in *Protocol 1*
- Nuclease-free water
- Thermal cycler
- 10 mM Ultrapure dNTP mix (Sigma)
- Example forward primer: 5'-AAAGCAGGCAATTGAATGACAGTAATGATG-3' (50 μ M)
- Example reverse primer: 5'-GTGTAGCATTCAAGGCCTTTGCTATCTGG-3' (50 μ M)

Method

1. Amplify the final product using a typical protocol recommended by the manufacturer of your PCR kit(s). For example, for a 50 μ l reaction^b (see *Fig. 1*, PCR Step B), combine^c:
 - 28.2 μ l of water
 - 5 μ l of reaction buffer (Herculase HotStart DNA polymerase kit)
 - 2.5 μ l of dNTPs (10 mM)
 - 10 μ l of purified DNA from PCR Step A (*Protocol 4*)
 - 0.4 μ l of example forward primer
 - 0.4 μ l of example reverse primer
 - 1 μ l of Herculase HotStart polymerase
 - 2.5 μ l of DMSO
2. Place the tube in a thermal cycler and run the following program:
 - Initial denaturation step at 92°C for 1 min
 - 31 cycles of denaturation at 92°C for 10 s and combined annealing and extension at 68°C^d for 20 min 40 s (plus automatic extension of extension time by 10 s per cycle)
 - Final additional extension step at 68°C for 30 min
 - Hold at 4°C

Notes

^aThis step is shown as PCR Step B in *Fig. 1*. The use of nested primers here (if you planned for them in *Protocol 1*) will improve purity and the chances of success of your fusion.

^bA 50 μ l reaction volume per tube offers fast enough exchange of heat between the thermal cycler and the reaction mixture and at the same time is convenient for mixing all of the components in the right amounts.

^cAn aliquot of your template (from *Protocol 5*) can comprise up to 1/5 of the volume of the PCR.

^dMake sure that you use a high annealing temperature (65–68°C) to achieve the lowest level of nonspecific PCR products in your reaction. This assumes that you have designed primers that have a high melting temperature as described in section 2.4.1.

Protocol 6

Analysis and purification of final product

Equipment and Reagents

- Crystal violet (2 mg/ml; Sigma)
- 1% (or 0.5% for long PCR products) agarose gel, containing 30 μ l of the crystal violet stock solution per 150 ml of agarose
- 5 \times Loading buffer (10% Ficoll 400, 0.1 M EDTA, pH 8.0, 120 μ g/ml of crystal violet)
- White light box and camera
- Apparatus required for electrophoresis
- 1 \times TAE running buffer (4.84 g/l Tris base, 1.142 ml/l glacial acetic acid, 0.372 g/l EDTA in distilled water)
- QIAquick PCR purification kit or QIAquick gel extraction kit (Qiagen)
- DNA molecular weight markers, e.g. 1 kb DNA step ladder (Promega)

Method

1. Load and run 5 μ l of each PCR on your agarose gel^a. Also include a lane with 0.5 μ g of DNA marker in 10 μ l of the same loading buffer as the PCR samples^b.
2. Use a white light table (instead of UV light)^c to view the results of the electrophoresis to determine whether the PCR products are of the correct length^d.
3. Purify each PCR product using a 'bind-and-wash' DNA purification kit, such as a QIAquick PCR purification kit, following the manufacturer's instructions.
4. Quantify your PCR products visually (by comparing the intensity of PCR bands with the intensity of DNA marker bands) or using spectrophotometry.

Notes

^aLoad 20–40 μ l if gel purification is required.

^bThis is necessary for correct visual quantification of DNA amounts in PCR samples.

^cYou can use ethidium bromide-stained agarose gels and UV light throughout this protocol for both analysis and purification of PCR products. However, be advised that ethidium bromide and UV light can damage DNA and may decrease the quality of your final product.

^dIf you did not get the expected PCR product, refer to section 3 below.

Additional considerations

It is recommended that you sequence across the fusion points to verify whether the assembly has worked correctly. You can also sequence other critical regions of your construct to make sure they have no PCR errors.

Prepare a sufficient quantity of your final product for the downstream application. If the desired end product is linear, then it is not necessary to clone it into a plasmid vector to obtain a large quantity. Instead, one can set up a large-scale final amplification reaction (see *Protocol 5*), gel purify the PCR product and use it for a desired downstream application such as cell transfection. It is possible to obtain 5–10 μ g of pure final product using this approach.

212 ■ CHAPTER 12: CONSTRUCTION OF LONG DNA MOLECULES FROM MULTIPLE FRAGMENTS

Fig. 3 shows a 20 kb product that we obtained successfully by fusing three fragments of 10.7, 1.7, and 7.5 kb (1). If the size of the final product is 12–20 kb, you may need to use pulsed-field electrophoresis to measure the size of your PCR product. It is possible to use regular electrophoresis with 0.5% agarose for proper assessment of DNA fragments up to 12 kb with appropriate DNA size markers.

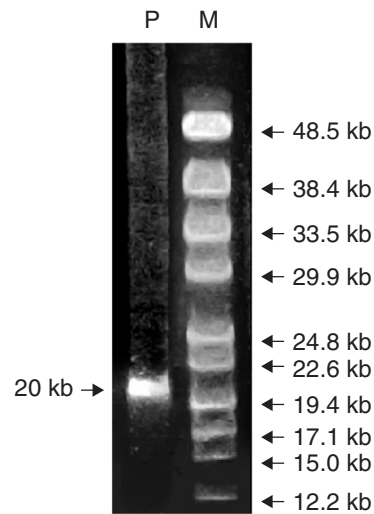


Figure 3. Pulsed-field gel electrophoresis of the 20 kb product.

Pulsed-field gel electrophoresis of the 20 kb product of long triple fusion of fragments of 10.7, 1.7, and 7.5 kb (adapted from 1). P, 20 kb product; M, DNA molecular weight markers.

2.4.2 Long multiple fusions

The process described below can be used for the assembly of four or five fragments into products up to 10 kb. The basic procedure is outlined briefly in *Fig. 4*. The detailed procedure is outlined in *Fig. 1*. The majority of the steps involved are identical to those outlined in the protocols above:

- *Analysis and purification of each fragment.* This protocol is identical to *Protocol 2*, except that, for a quintuple fusion, it is preferable to use gel purification using a crystal violet-based agarose gel.
- *Overlap extension reaction.* This step is depicted in *Fig. 1* as PCR Step A and in *Fig. 4* as PCR Steps A.1 and A.2. One should use two separate reactions for adjacent fragments: A+B and C+D as shown in *Fig. 4*. The protocol is largely identical to *Protocol 3*, except that extension times in PCRs should correspond to the length of products A+B or C+D, not the final product A+B+C+D (see *Fig. 4*).
- *Purification of intermediary product.* This protocol is identical to *Protocol 4* and is applied to products A+B and C+D separately.

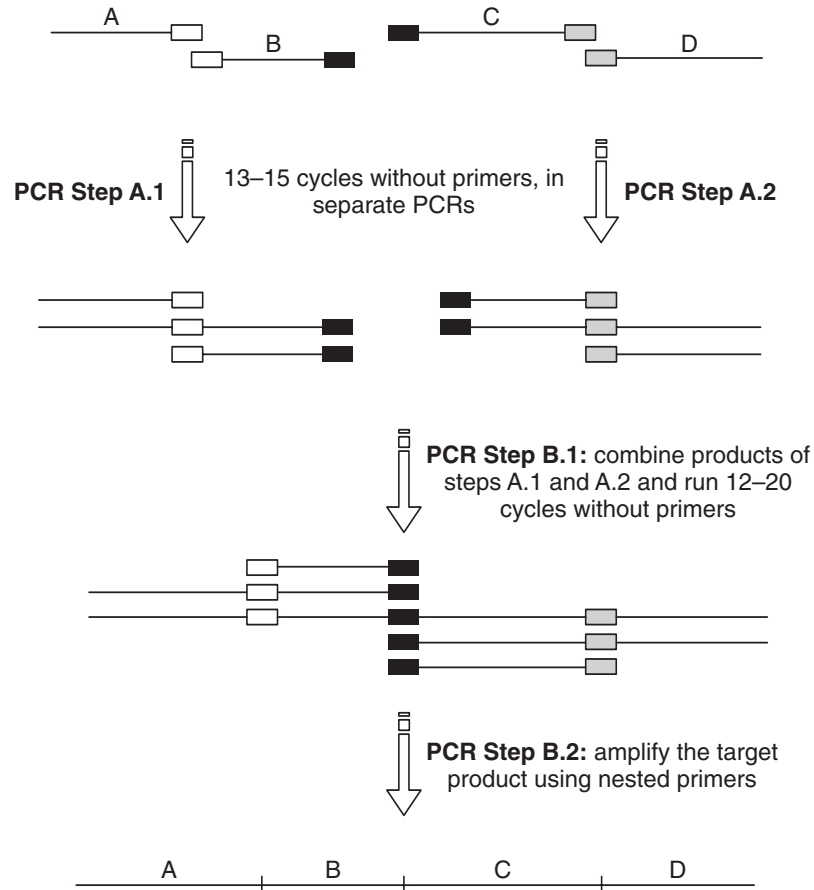


Figure 4. A brief outline of long quadruple fusion.

This technique was used successfully to assemble a 10 kb product. The procedure can also be used for the assembly of five fragments into products up to 10 kb. A more detailed procedure is outlined in *Fig. 1*, with the exception of the steps between PCR Step A and PCR Step B. For the quintuple we used pairwise fusions of A+B and D+E, and then mixed those two products with fragment C during PCR Step B.1.

- *Additional overlap extension reaction.* This step is depicted in *Fig. 4* as PCR Step B.1. The protocol is largely identical to *Protocol 3*, except that the calculation of equimolar amounts is omitted, and equal aliquots of products A+B and C+D are used for a total of 1/5 of the final volume of the PCR. For the quintuple fusion, products A+B, D+E, and fragment C are mixed in equal aliquots and fused.
- *Amplification of final recombinant product.* This step is shown as PCR Step B/B.2 in *Figs 1* and *4*. The use of nested primers here will improve purity and the chances of success of your fusion. This protocol is identical to *Protocol 5*.

3. TROUBLESHOOTING

- **There are problems with amplifying a fragment with chimeric primers**
Amplify the fragment with regular primers, purify it with a bind-and-wash kit, and then reamplify the fragment with chimeric primers. During the reamplification, use 25% or more of the total amount of DNA obtained in the first PCR and run the reamplification for only 10–12 cycles.
- **No product is obtained in the final amplification**
Check the sequences of all chimeric primers and the method shown in *Fig. 2*. Ensure that you have adhered to the principles outlined in section 2.3.
- **An individual fragment cannot be amplified**
 - Check whether the template contains extended GC-rich sequences.
 - Make sure that you are using one of the five recommended kits.
 - If you are using chimeric primers with genomic DNA or with long fragments (7 kb or more), try to redesign new primers such that chimeric primers are only used for shorter fragments or for fragments amplified from plasmid DNA.
 - Try using the Herculase HotStart DNA polymerase kit (Stratagene) for the problem template.
 - Some problems with amplification of individual fragments can arise from poor quality of the source DNA. If it is a genomic DNA, phenol extraction is recommended. A substantial amount of template is needed for long PCR, in addition to good quality of template. For a successful long PCR, 0.3–1 µg of DNA template per 50 µl reaction is often needed. This applies to fusion steps A and B, as well as to amplification of individual fragments.
- If you have read section 2.3 and followed the protocols but still cannot solve your problems with the method, feel free to contact me at shevchuk@hotmail.com.

ACKNOWLEDGMENTS

This work was supported by NIH grants R01 CA61010 (to Stephan Ladisch) and R01 AI48856 (to Felipe C. Cabello).

4. REFERENCES

- ★★ 1. Shevchuk NA, Bryksin AV, Nusinovich YA, Cabello FC, Sutherland M & Ladisch S (2004) *Nucleic Acids Res.* **32**, e19. – *Original publication describing long multiple fusion.*
2. Ternes P, Sperling P, Albrecht S, *et al.* (2006) *J. Biol. Chem.* **281**, 5582–5592.
3. DUBYTSKA L, Godfrey HP & Cabello FC (2006) *J. Bacteriol.* **188**, 1969–1978.
4. Strahilevitz J, Robicsek A & Hooper DC (2006) *Antimicrob. Agents Chemother.* **50**, 600–606.
5. Kato T, Muraski J, Chen Y, *et al.* (2005) *J. Clin. Invest.* **115**, 2716–2730.
6. Ojaimi C, Mulay V, Liveris D, Iyer R & Schwartz I (2005) *Infect. Immun.* **73**, 6791–6802.
7. Bugrysheva JV, Bryksin AV, Godfrey HP & Cabello FC (2005) *Infect. Immun.* **73**, 4972–4981.
8. Morozova OV, DUBYTSKA LP, Ivanova LB, *et al.* (2005) *Gene*, **357**, 63–72.

REFERENCES ■ 215

- ★★★ 9. **Yon J & Fried M** (1989) *Nucleic Acids Res.* **17**, 4895. – *One of the first publications describing overlap-extension PCR.*
- ★★ 10. **Yolov AA & Shabarova ZA** (1990) *Nucleic Acids Res.* **18**, 3983–3986. – *One of the first publications describing overlap extension PCR.*
- ★ 11. **Pont-Kingdon G** (1997) *Methods Mol. Biol.* **67**, 167–172. – *A useful review of the application of PCR to the creation of recombinant molecules and to mutagenesis.*
12. **Kuwayama H, Obara S, Morio T, Katoh M, Urushihara H & Tanaka Y** (2002) *Nucleic Acids Res.* **30**, e2.
13. **Horton RM** (1995) *Mol. Biotechnol.* **3**, 93–99.
- ★★ 14. **Xiong AS, Yao QH, Peng RH, et al.** (2004) *Nucleic Acids Res.* **32**, e98. – *Successful fusion of ten DNA fragments using PCR.*
15. **Sambrook J, Fritsch, E. & Maniatis, T.** (2001) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
16. **RedAccuTaq LA DNA polymerase mix** (1999) Technical bulletin no. MB-690. Sigma-Aldrich, Saint Louis, MO.

