1

Molecular Cloning

A LABORATORY MANUAL

SECOND EDITION

J. Sambrook
UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER

E.F. Fritsch
GENETICS INSTITUTE

T. Maniatis
HARVARD UNIVERSITY

Cold Spring Harbor Laboratory Press
1989

Molecular Cloning

A LABORATORY MANUAL

SECOND EDITION

Printed in the United States of America

9 8 7

Book and cover design by Emily Harste

Cover: The electron micrograph of bacteriophage λ particles stained with uranyl acetate was digitized and assigned false color by computer. (Thomas R. Broker, Louise T. Chow, and James I. Garrels)

Cataloging in Publications data

Sambrook, Joseph
Molecular cloning : a laboratory manual / E.F. Fritsch, T. Maniatis—2nd ed.
p. cm.
Bibliography: p.
Includes index.
ISBN 0-87969-309-6
1. Molecular cloning—Laboratory manuals. 2. Eukaryotic cells--Laboratory manuals. I. Fritsch, Edward F. II. Maniatis, Thomas III. Title.
QH442.2.M26 1987
574.87'3224—dc19 87-35464

All Cold Spring Harbor Laboratory Press publications may be ordered directly from Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803. Phone: 1-800-843-4388. In New York (516) 367-8423. FAX: (516) 367-8432.

Preface to First Edition

This manual began as a collection of laboratory protocols that were used during the 1980 Cold Spring Harbor course on the Molecular Cloning of Eukaryotic Genes. These procedures had been in use in our laboratories at that time but were scattered throughout the notebooks of many different people. In 1981 we decided to produce a more complete and up-to-date manual not only for use in the next Cold Spring Harbor course, but also for eventual publication. Out of the many permutations of the methods being used, we assembled a set of "consensus protocols," which were photocopied and widely distributed to many laboratories even as the 1981 course was underway. Then in the winter of 1981–1982, the manual was substantially rewritten, and new or revised protocols and figures, as well as entirely new chapters, were added.

Even since this last rewriting, however, the field has progressed: New methods are constantly being invented and existing techniques are altered in response to changing needs. Although we have included in this manual only those protocols that have been thoroughly tested and used successfully in our laboratories, we make no claim that they are inviolable or perfect. We would welcome suggestions for improvements, and we would be grateful to be told about any new procedures that are devised.

The evolution of protocols poses the difficult problem of attribution. We have tried to give credit at appropriate places in the text to the people who originally developed the procedures presented here, but in many cases tracing a particular method to its undisputed roots has proved to be impossible. We therefore wish to apologize—and to express gratitude—to those we have been unable to acknowledge for an idea, procedure, or recipe. Our major function has been to compile, to verify, and, we hope, to clarify; less frequently we have introduced modifications, and only in rare instances have we devised new protocols. In large part, then, the manual is based on procedures developed by others, and it is to them that any credit belongs.

Because the manual was originally written to serve as a guide to those who had little experience in molecular cloning, it contains much basic material. However, the current version also deals in detail with almost every laboratory task currently used in molecular cloning. We therefore hope that newcomers to cloning and veterans alike will find material of value in this book.

Although molecular cloning seems straightforward on paper, it is more difficult to put into practice. Most protocols involve a large number of

individual steps and a problem with any one of them can lead the experimenter into difficulty. It is a good idea to verify the products of each step and to include controls to check the efficiency of each reaction. To deal with these problems, a well-founded understanding of the principles underlying each procedure is essential. We have therefore provided background information and references that may be useful should trouble occur.

This manual could not have been written without the help and advice of members of our laboratories and contributions from many others. We therefore wish to thank Joan Brooks, John Fiddes, Mary-Jane Gething, Tom Gingeras, David Goldberg, Steve Hughes, David Ish-Horowicz, Mike Mathews, Patty Reichel, Joe Sorge, Jim Stringer, Richard Treisman, and Nigel Whittle. We wish particularly to thank Arg Efstratiadis for his helpful discussions and criticisms of Chapter 7; Brian Seed for permission to include a description of his unpublished procedure for screening libraries by recombination (Chapter 10) and many other useful suggestions; Doug Hanahan for advice on transformation (Chapter 8); Bryan Roberts for suggestions on methods of hybrid-selection and cDNA cloning; Doug Melton for providing a protocol for injection of *Xenopus* oocytes; Ronni Greene for suggesting improvements to many protocols; Nina Irwin for providing a critical anthology of methods available for expressing eukaryotic proteins in bacteria (Chapter 12); Rich Roberts for supplying the computer analysis of the sequence of pBR322; Barbara Bachmann and Ahmad Bukhari for reviewing and correcting the list of *E. coli* strains; and Tom Broker, Louise Chow, Jeff Engler, and Jim Garrels for producing the elegant photographs used for the front and back covers.

We also thank all those who participated in the Cold Spring Harbor Molecular Cloning courses of 1980 and 1981. They were an excellent group of students, who struggled through the first two drafts of the manual and made many useful suggestions. We also thank Nancy Hopkins, who helped us to teach the course the first year and convinced us that producing a manual would be a worthwhile task. In 1981 Doug Engel helped teach the course and suggested many improvements to the manual. Contributing to the success of both courses were the efforts of the teaching assistants, who were Catherine O'Connell and Helen Doris Keller in the summer of 1980 and Susan VandeWoude, Paul Bates, and Michael Weiss in 1981.

We wish to thank Patti Barkley and Marilyn Goodwin for their cheerfulness and forbearance during the typing of successive revisions of the manuscript. Our artists, Fran Cefalu and Mike Ockler, worked with great dedication and perseverance to produce the drawings for the manual. Joan Ebert kept track of the many references added to and deleted from the text and assembled the reference list. We are also grateful to Nancy Ford, Director of Publications, Cold Spring Harbor Laboratory, for her encouragement and support. Finally, without the patience, skill, and diplomacy of Doug Owen, who prepared the manuscript for the printer and helped us in many other ways, this book would not exist.

Tom Maniatis
Ed Fritsch
Joe Sambrook

Preface to Second Edition

Since the publication of the first edition of this Laboratory Manual in 1982, there has been a vast increase both in the number of people who use molecular cloning and in the range and power of the techniques used to handle recombinant DNA. This remarkable proliferation of cloning methods is reflected in the number of gene sequences in the GenBank DNA sequence data base. In 1982 there were fewer than 350 gene sequences on file, but in less than four years this number grew to almost 5000. Today over 15,000 sequences are listed. These figures, impressive as they are, do not do justice to the increased sophistication with which cloned genes are now analyzed. In 1982, respectable journals would still accept manuscripts containing little more than the partial sequence of a cDNA clone; today, the publication of a complete sequence is taken for granted, and papers describing the initial cloning of a cDNA will often also contain elegant accounts of the expression of the gene product in prokaryotic or eukaryotic hosts. In most cases, this first paper is rapidly followed by others that use site-directed mutagenesis to explore the relationship between the structure and function of the relevant protein. The cloning and analysis of elements that regulate the expression of eukaryotic genes has moved at a similar, exciting pace.

This increase in the range and speed of molecular cloning is reflected in the tripling of size of this manual and its consequent division into three volumes. Techniques that were mentioned only in passing in the first edition, such as mutagenesis, expression of cloned genes in mammalian cells, and dideoxy-mediated sequencing, are now described in depth; new sections have been added that deal with recently invented techniques, such as amplification of DNA by the polymerase chain reaction; and we have included modern variations and embellishments of many of the basic methods that were the mainstay of the first edition. We hope that any inconvenience caused by the expansion in size of the manual will be compensated by an increase in the richness of its content. We hope that this second edition will be a resource for the experienced cloner, a starting point for the student, and a guidebook for the next generation of investigators in molecular cloning.

The burgeoning of molecular cloning has also led to its commercialization, and we are now able to purchase a wide variety of high quality and reasonably priced reagents and enzymes. Although this is a very positive development, it has had a few unfortunate side effects. One is the proliferation of preassembled kits to carry out particular cloning tasks. Although these kits reduce the possibility of trivial errors, they also tend to discourage experimenters from thinking about what they are doing. It becomes all too

easy to follow blindly instructions to add 2 μl of solution A without knowing what the particular reagent is, why it is necessary, or why it is added at a particular point in the protocol. Kits therefore reward the scientifically illiterate and inhibit the development of improvements. In an attempt to counteract this trend, we have greatly increased the amount of background material at the beginning of each chapter of the manual and we have provided full references. Users of the manual who read this material should have no problem understanding either the general design or the specific details of the experimental protocols. We also strongly recommend reading the relevant protocols in their entirety before commencing work. This allows reagents to be prepared ahead of time and permits the investigator to carry out the protocol efficiently.

This manual could not have been produced without the help and encouragement of a large number of people. We are extremely grateful to colleagues from the University of Texas Southwestern Medical Center at Dallas, from Genetics Institute, from Harvard University, and from many other institutions who have read individual chapters, have submitted protocols, and have made many invaluable suggestions that have immeasurably improved the manual. We have listed all these individuals with their affiliations in a special section at the end of this preface.

We owe a special debt of thanks to Rick Myers and Alison Cowie of the University of California at San Francisco, who read all of the chapters in draft, eliminated many embarrassing errors, and made many valuable suggestions for improvements; to Winship Herr of Cold Spring Harbor Laboratory and Mary-Jane Gething of Southwestern Medical Center, who drafted the chapters on DNA sequencing and expression of cloned genes in mammalian cells, respectively; and to Judy Campbell, who contributed significantly to the preparation of the chapter on enzymes. We are also grateful to Mike Ockler, who provided all of the conceptual illustrations in this edition, and to Carolyn Doyle, who assembled the index.

Nina Irwin has played a special role in the writing of this manual. Among her many contributions are the *E. coli* strain list and the maps of the many strains of bacteriophage λ, plasmids, cosmids, and bacteriophage M13 used in this manual. These maps are based on many months of painstaking work to reconstruct and verify the genealogy of these host and vector strains. In addition, Nina also drafted the chapter on expression of cloned genes in prokaryotic hosts, read and reviewed all of the chapters in manuscript and galleys, and searched the scientific literature with great intelligence and skill to validate many of the facts that appear.

We also wish to thank Jim Watson for his continued interest in this manual. The support he has provided made this edition possible.

Finally, we would like to express our deep appreciation of the work of our editor Nancy Ford. For over two years, she has worked untiringly to bring order and sense to the entire project and to improve and clarify our writing. Nancy and her assistant Michele Ferguson have both served as a source of encouragement, sympathy, and friendship. We are greatly in their debt.

J. Sambrook
E.F. Fritsch
T. Maniatis

Acknowledgments

Ed Alderman, Genetics Institute
Patricia Ashley, University of Texas Southwestern Medical Center
Margaret Baron, Harvard University
Frank Baas, Harvard University
Phil Bird, University of Texas Southwestern Medical Center
Tom Bittick, University of Texas Southwestern Medical Center
Colleen Brewer, University of Texas Southwestern Medical Center
Gene Brown, Genetics Institute
H. Franklin Bunn, Harvard Medical School
Steve Clark, Genetics Institute
Preston Dunnmon, University of Texas Southwestern Medical Center
Anne Ephrussi, Harvard University
Henry Erlich, Perkin Elmer Cetus
Chen-Ming Fan, Harvard University
Ken Ferguson, Cold Spring Harbor Laboratory
Bill Garrard, University of Texas Southwestern Medical Center
Bob Gerard, University of Texas Southwestern Medical Center
Gary Gilliland, Harvard Medical School
Doug Hanahan, University of California, San Francisco
Ed Harlow, Cold Spring Harbor Laboratory
Jean Henneberry, University of Texas Southwestern Medical Center
Bill Huse, Stratagene
David Ish-Horowicz, Imperial Cancer Research Fund
David Israel, Genetics Institute
Randy Kaufman, Genetics Institute
Andrew Keller, Harvard University
John Knopf, Genetics Institute
Steve Lacey, University of Texas Southwestern Medical Center
Ray MacDonald, University of Texas Southwestern Medical Center
Ed Madison, University of Texas Southwestern Medical Center
Mike Mathews, Cold Spring Harbor Laboratory
Michael McClelland, University of Chicago
John McCoy, Genetics Institute
Steve McKnight, Carnegie Institution of Washington
Doug Melton, Harvard University
Alan Michelson, Harvard University
George Morris, Genetics Institute
Michael Nelson, University of Chicago
Robin Reed, Harvard University
Rich Roberts, Cold Spring Harbor Laboratory
Laura Roman, University of Texas Southwestern Medical Center
Susan Rosenberg, University of Utah
David Russell, University of Texas Southwestern Medical Center
Brian Seed, Harvard Medical School
Conrad Seghers, University of Texas Southwestern Medical Center
Mark Segal, University of Texas Southwestern Medical Center
Chuck Shoemaker, Harvard School of Public Health
Huda Shubeita, University of Texas Southwestern Medical Center
Harinder Singh, University of Chicago
Beth Smith, Information Center, Genetics Institute
Joe Sorge, Stratagene
Lisa Sultzman, Genetics Institute
Galvin Swift, University of Texas Southwestern Medical Center
Patty Temple, Genetics Institute
Jeff Vieira, Rutgers University
Geoff Wahl, Salk Institute
Steve Wasserman, University of Texas Southwestern Medical Center
Gordon Wong, Genetics Institute
Rick Young, Massachusetts Institute of Technology
Jian-Hua Zhang, University of Texas Southwestern Medical Center
Mark Zoller, Genentech

Contents

BOOK 1

1

Plasmid Vectors

2

Bacteriophage λ Vectors

3

Cosmid Vectors

1
Plasmid Vectors

Bacterial plasmids are double-stranded closed circular DNA molecules that range in size from 1 kb to more than 200 kb. They are found in a variety of bacterial species, where they behave as accessory genetic units that replicate and are inherited independently of the bacterial chromosome. Nevertheless, they rely on enzymes and proteins encoded by the host for their replication and transcription. Frequently, plasmids contain genes coding for enzymes that under certain circumstances in nature are advantageous to the bacterial host. Among the phenotypes conferred by plasmids are resistance to antibiotics; production of antibiotics; degradation of complex organic compounds; and production of colicins, enterotoxins, and restriction and modification enzymes.

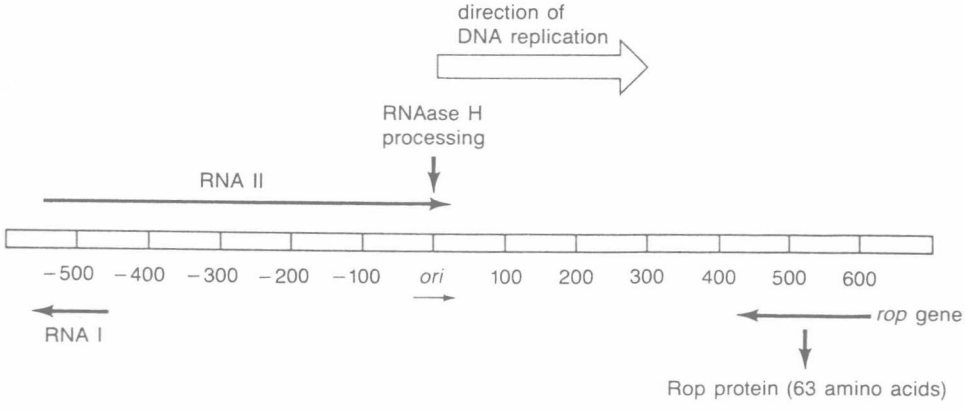

FIGURE 1.1
The figure shows the direction and approximate size of transcripts involved in the initiation of replication of plasmids carrying pMB1 (or ColE1) origins.