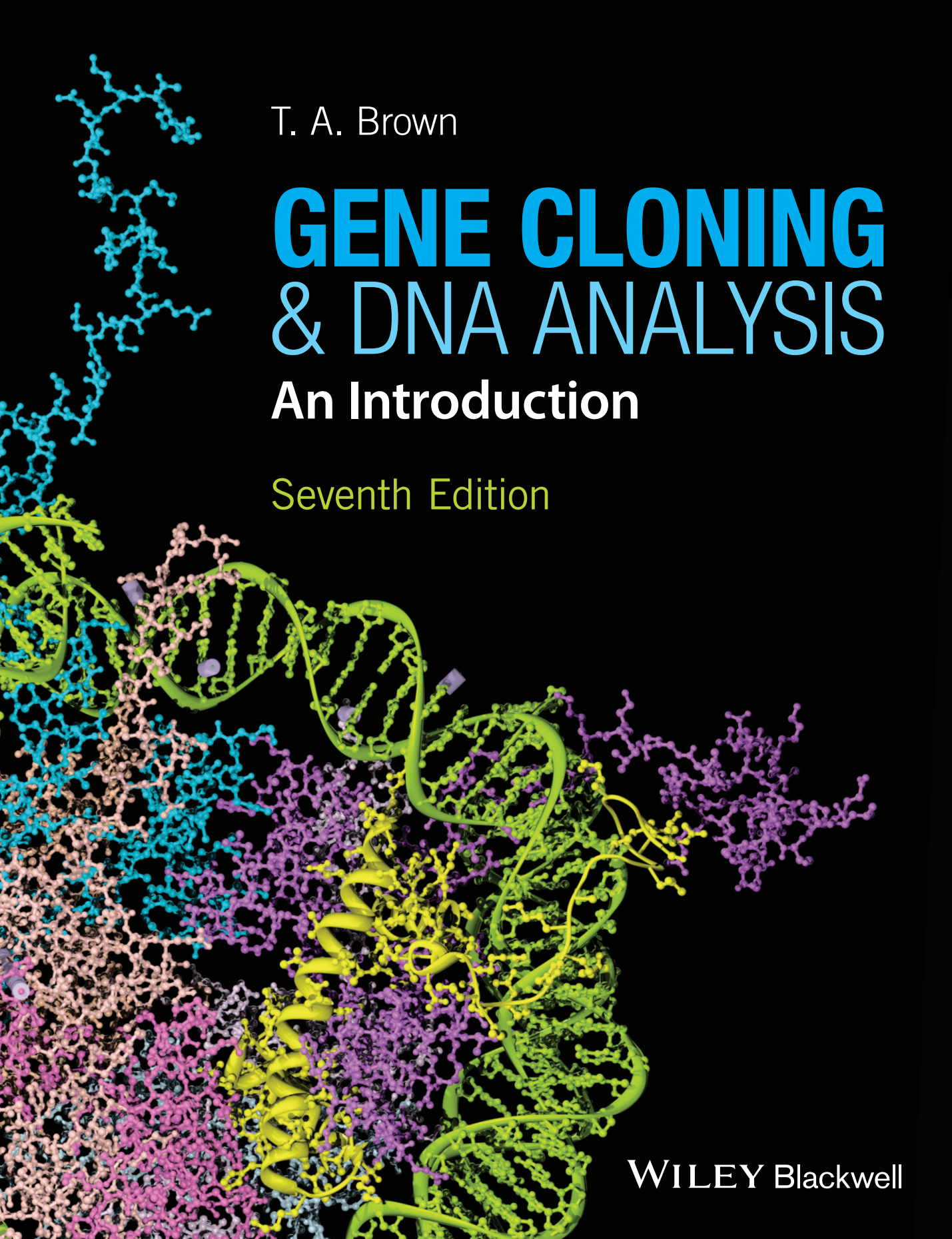

T. A. Brown
GENE CLONING
& DNA ANALYSIS
An Introduction
Seventh Edition
WILEY Blackwell

GENE CLONING AND DNA ANALYSIS

GENE CLONING AND DNA ANALYSIS

An Introduction

T.A. BROWN
University of Manchester
Manchester

Seventh Edition

WILEY Blackwell

This edition first published 2016 © 2016 by John Wiley & Sons, Ltd
Sixth edition published 2010 by John Wiley & Sons, Ltd
Fourth and fifth editions published 2001, 2006 by Blackwell Publishing Ltd
First, second and third editions published 1986, 1990, 1995 by Chapman & Hall

Registered office: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Brown, T.A. (Terence A.), author.
Gene cloning and DNA analysis : an introduction / T.A. Brown. – Seventh edition.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-119-07257-7 (cloth) – ISBN 978-1-119-07256-0 (pbk.)
I. Title.
[DNLM: 1. Cloning, Molecular. 2. DNA, Recombinant–analysis. 3. Sequence Analysis, DNA. QU 450]
QH442.2
572.8′633–dc23 2015015462

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

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Set in 10/12pt Sabon by Aptara Inc., New Delhi, India

1 2016

Contents in Brief

CONTENTS IN BRIEF

Contents

CONTENTS

Preface to the Seventh Edition

PREFACE TO THE SEVENTH EDITION

Anyone who works with DNA is well aware of the dramatic changes that have taken place during the past few years in DNA sequencing methodology. To reflect these advances, in this new edition *of Gene Cloning and DNA Analysis: An Introduction* I have completely remodelled the chapter on DNA sequencing to give the new 'next-generation' methods equal prominence alongside the traditional approaches to DNA sequencing, and also to modernize the description of the ways in which genome sequences are generated. Elsewhere, I have stressed the importance of RNA-seq as a means of studying transcriptomes, and ChIP-seq for locating protein-binding sites. These changes correct the major weakness of the Sixth Edition, which was written just before these methods came into mainstream use.

Elsewhere, I have made the usual updates, especially in Part III where I have tried to keep pace with the increasingly rapid developments in the applications of gene cloning and DNA analysis in industry, medicine and agriculture. I have also rewritten the last part of the final chapter, on archaeogenetics, in order to present some of the new information on the human past that has been revealed by the Neanderthal and Denisovan genome sequences. As always, my primary aim is to ensure that *Gene Cloning* remains an introductory text that begins at the beginning and does not assume that the reader has any prior knowledge of the techniques used to study genes and genomes.

For the n-th time I must thank my wife Keri for the unending support that she has given to me in my decision to use up evenings and weekends writing this and other books.

T.A. Brown
University of Manchester

About the companion website

ABOUT THE COMPANION WEBSITE

This book is accompanied by a companion website:

www.wiley.com/go/brown/cloning

The website includes:

- Powerpoints of all figures from the book for downloading
- PDFs of tables from the book

PART I

The Basic Principles of Gene Cloning and DNA Analysis

Chapter 1

Why Gene Cloning and DNA Analysis are Important

Chapter contents

CHAPTER CONTENTS

In the middle of the 19th century, Gregor Mendel formulated a set of rules to explain the inheritance of biological characteristics. The basic assumption of these rules is that each heritable property of an organism is controlled by a factor, called a **gene**, that is a physical particle present somewhere in the cell. The rediscovery of Mendel's laws in 1900 marks the birth of **genetics**, the science aimed at understanding what these genes are and exactly how they work.

1.1 The early development of genetics

For the first 30 years of its life this new science grew at an astonishing rate. The idea that genes reside on **chromosomes** was proposed by W. Sutton in 1903, and received experimental backing from T.H. Morgan in 1910. Morgan and his colleagues then developed the techniques for **gene mapping**, and by 1922 had produced a comprehensive analysis of the relative positions of over 2000 genes on the four chromosomes of the fruit fly, *Drosophila melanogaster*.

Despite the brilliance of these classical genetic studies, there was no real understanding of the molecular nature of the gene until the 1940s. Indeed, it was not until the

Gene Cloning and DNA Analysis: An Introduction, Seventh Edition. T.A. Brown.

Companion Website: www.wiley.com/go/brown/cloning

experiments of Avery, MacLeod, and McCarty in 1944, and of Hershey and Chase in 1952, that anyone believed that deoxyribonucleic acid (DNA) is the genetic material. Up until then it was widely thought that genes were made of protein. The discovery of the role of DNA was a tremendous stimulus to genetic research, and many famous biologists (Delbrück, Chargaff, Crick, and Monod were among the most influential) contributed to the second great age of genetics. During the 14 years between 1952 and 1966, the structure of DNA was elucidated, the genetic code cracked, and the processes of transcription and translation described.

1.2 The advent of gene cloning and the polymerase chain reaction

These years of activity and discovery were followed by a lull, a period of anticlimax when it seemed to some molecular biologists (as the new generation of geneticists styled themselves) that there was little of fundamental importance that was not understood. In truth, there was a frustration that the experimental techniques of the late 1960s were not sophisticated enough to allow the gene to be studied in any greater detail.

Then, in the years 1971–1973 genetic research was thrown back into gear by what at the time was described as a revolution in experimental biology. A whole new methodology was developed, enabling previously impossible experiments to be planned and carried out, if not with ease, then at least with success. These methods, referred to as **recombinant DNA technology** or **genetic engineering**, and having at their core the process of **gene cloning**, sparked another great age of genetics. They led to rapid and efficient **DNA sequencing** techniques that enabled the structures of individual genes to be determined, reaching a culmination at the turn of the century with the massive genome sequencing projects, including the human project which was completed in 2000. They led to procedures for studying the regulation of individual genes, which have allowed molecular biologists to understand how aberrations in gene activity can result in human diseases such as cancer. The techniques spawned modern **biotechnology**, which puts genes to work in production of proteins and other compounds needed in medicine and industrial processes.

During the 1980s, when the excitement engendered by the gene cloning revolution was at its height, it hardly seemed possible that another, equally novel and equally revolutionary process was just around the corner. According to DNA folklore, Kary Mullis invented the **polymerase chain reaction (PCR)** during a drive along the coast of California one evening in 1985. His brainwave was an exquisitely simple technique that acts as a perfect complement to gene cloning. PCR has made easier many of the techniques that were possible but difficult to carry out when gene cloning was used on its own. It has extended the range of DNA analysis and enabled molecular biology to find new applications in areas of endeavour outside of its traditional range of medicine, agriculture, and biotechnology. Archaeogenetics, molecular ecology, and DNA forensics are just three of the new disciplines that have become possible as a direct consequence of the invention of PCR, enabling molecular biologists to ask questions about human evolution and the impact of environmental change on the biosphere, and to bring their powerful tools to bear in the fight against crime. Forty years have passed since the dawning of the age of gene cloning, but we are still riding the rollercoaster and there is no end to the excitement in sight.

1.3 What is gene cloning?

What exactly is gene cloning? The easiest way to answer this question is to follow through the steps in a gene cloning experiment (Figure 1.1):

1. A fragment of DNA, containing the gene to be cloned, is inserted into a circular DNA molecule called a **vector**, to produce a **recombinant DNA molecule**.
2. The vector transports the gene into a host cell, which is usually a bacterium, although other types of living cell can be used.
3. Within the host cell the vector multiplies, producing numerous identical copies, not only of itself but also of the gene that it carries.
4. When the host cell divides, copies of the recombinant DNA molecule are passed to the progeny and further vector replication takes place.
5. After a large number of cell divisions, a colony, or **clone**, of identical host cells is produced. Each cell in the clone contains one or more copies of the recombinant DNA molecule. The gene carried by the recombinant molecule is now said to be cloned.

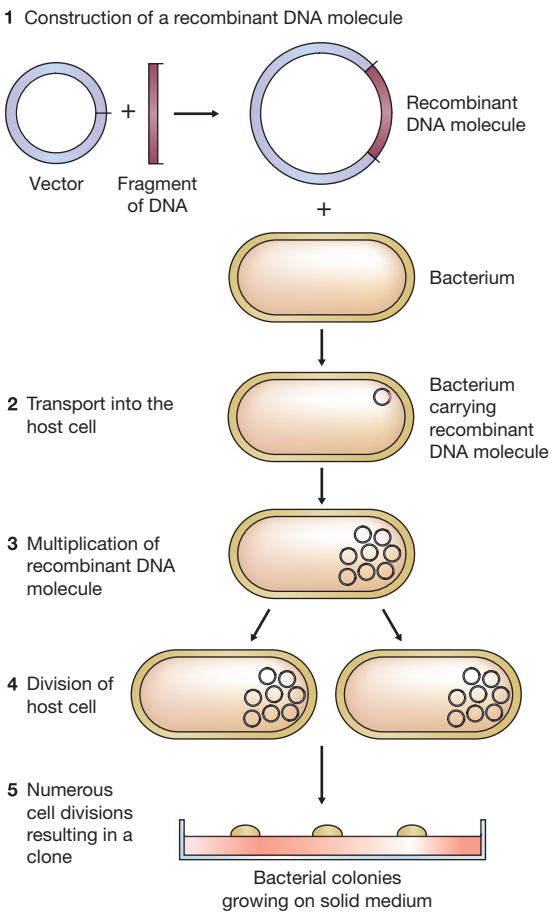


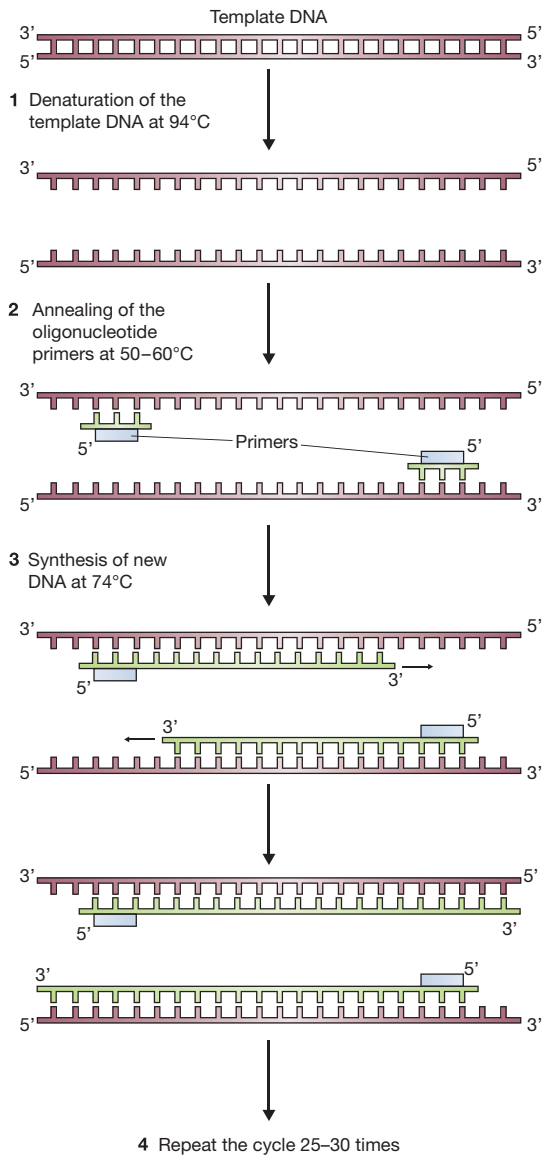
Figure 1.1
The basic steps in gene cloning.

1.4 What is PCR?

The polymerase chain reaction is very different from gene cloning. Rather than a series of manipulations involving living cells, PCR is carried out in a single test tube simply by mixing DNA with a set of reagents and placing the tube in a thermal cycler, a piece of equipment that enables the mixture to be incubated at a series of temperatures that are varied in a preprogrammed manner. The basic steps in a PCR experiment are as follows (Figure 1.2):

1 The mixture is heated to 94 °C, at which temperature the hydrogen bonds that hold together the two strands of the double-stranded DNA molecule are broken, causing the molecule to **denature**.

Figure 1.2
The basic steps in the polymerase chain reaction.



2 The mixture is cooled down to 50–60 °C. The two strands of each molecule could join back together at this temperature, but most do not because the mixture contains a large excess of short DNA molecules, called **oligonucleotides** or **primers**, which **anneal** to the DNA molecules at specific positions.

3 The temperature is raised to 74 °C. This is a good working temperature for the ***Taq* DNA polymerase** that is present in the mixture. We will learn more about **DNA polymerases** on p. 51. All we need to understand at this stage is that the *Taq* DNA polymerase attaches to one end of each primer and synthesizes new strands of DNA, complementary to the **template** DNA molecules, during this step of the PCR. Now we have four stands of DNA instead of the two that there were to start with.

4 The temperature is increased back to 94 °C. The double-stranded DNA molecules, each of which consists of one strand of the original molecule and one new strand of DNA, denature into single strands. This begins a second cycle of denaturation–annealing–synthesis, at the end of which there are eight DNA strands. By repeating the cycle 30 times the double-stranded molecule that we began with is converted into over 130 million new double-stranded molecules, each one a copy of the region of the starting molecule delineated by the annealing sites of the two primers.

1.5 Why gene cloning and PCR are so important

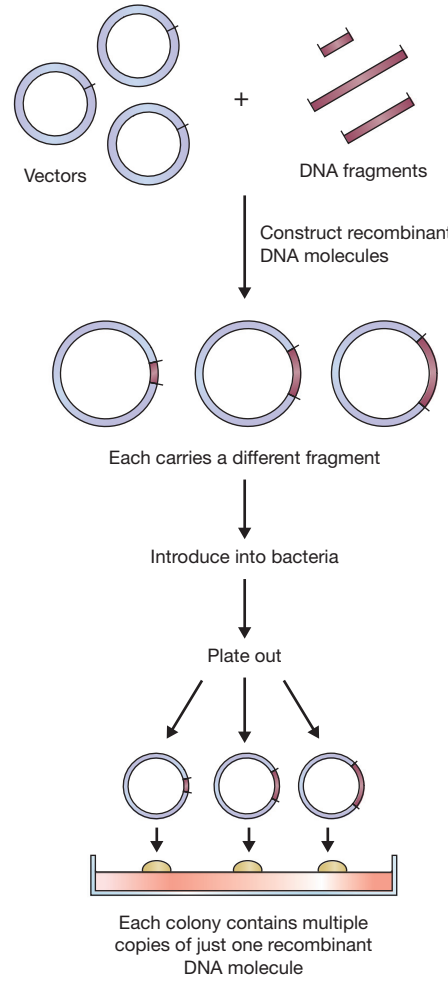
As can be seen from Figures 1.1 and 1.2, gene cloning and PCR are relatively straightforward procedures. Why, then, have they assumed such importance in biology? The answer is largely because both techniques can provide a pure sample of an individual gene, separated from all the other genes in the cell.

1.5.1 Obtaining a pure sample of a gene by cloning

To understand exactly how cloning can provide a pure sample of a gene, consider the basic experiment from Figure 1.1, but drawn in a slightly different way (Figure 1.3). In this example, the DNA fragment to be cloned is one member of a mixture of many different fragments, each carrying a different gene or part of a gene. This mixture could indeed be the entire genetic complement of an organism – a human, for instance. Each of these fragments becomes inserted into a different vector molecule to produce a family of recombinant DNA molecules, one of which carries the gene of interest. Usually, only one recombinant DNA molecule is transported into any single host cell, so that although the final set of clones may contain many different recombinant DNA molecules, each individual clone contains multiple copies of just one molecule. The gene is now separated away from all the other genes in the original mixture, and its specific features can be studied in detail.

In practice, the key to the success or failure of a gene cloning experiment is the ability to identify the particular clone of interest from the many different ones that are obtained. If we consider the **genome** of the bacterium *Escherichia coli*, which contains just over 4000 different genes, we might at first despair of being able to find just one gene among all the possible clones (Figure 1.4). The problem becomes even more overwhelming when we remember that bacteria are relatively simple organisms and that the human genome contains about five times as many genes. As explained in Chapter 8, a variety

Figure 1.3
Cloning allows individual fragments of DNA to be purified.



of different strategies can be used to ensure that the correct gene can be obtained at the end of the cloning experiment. Some of these strategies involve modifications to the basic cloning procedure, so that only cells containing the desired recombinant DNA molecule can divide and the clone of interest is automatically **selected**. Other methods involve techniques that enable the desired clone to be identified from a mixture of lots of different clones.

Once a gene has been cloned there is almost no limit to the information that can be obtained about its structure and expression. The availability of cloned material has stimulated the development of many different analytical methods for studying genes, with new techniques being introduced all the time. Methods for studying the structure and expression of a cloned gene are described in Chapters 10 and 11, respectively.

1.5.2 PCR can also be used to purify a gene

The polymerase chain reaction can also be used to obtain a pure sample of a gene. This is because the region of the starting DNA molecule that is copied during PCR

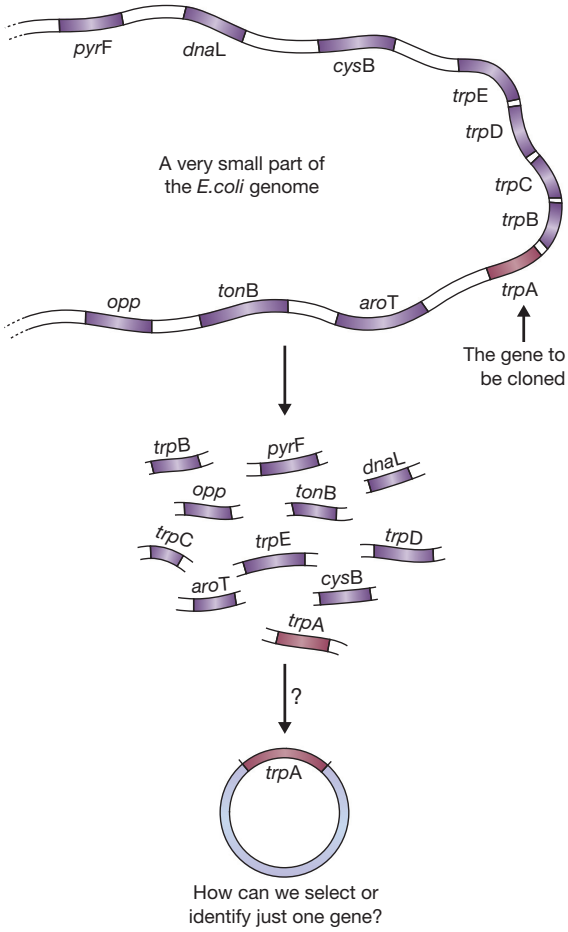


Figure 1.4
The problem of selection.

is the segment whose boundaries are marked by the annealing positions of the two oligonucleotide primers. If the primers anneal either side of the gene of interest, many copies of that gene will be synthesized (Figure 1.5). The outcome is the same as with a gene cloning experiment, although the problem of selection does not arise because the desired gene is automatically 'selected' as a result of the positions at which the primers anneal.

A PCR experiment can be completed in a few hours, whereas it takes weeks – if not months – to obtain a gene by cloning. Why then is gene cloning still used? This is because PCR has two limitations:

- In order for the primers to anneal to the correct positions, on either side of the gene of interest, the sequences of these annealing sites must be known. It is easy to synthesize a primer with a predetermined sequence (see p. 149), but if the sequences of the annealing sites are unknown then the appropriate primers cannot be made. This means that PCR cannot be used to isolate genes that have not been studied before – that has to be done by cloning.