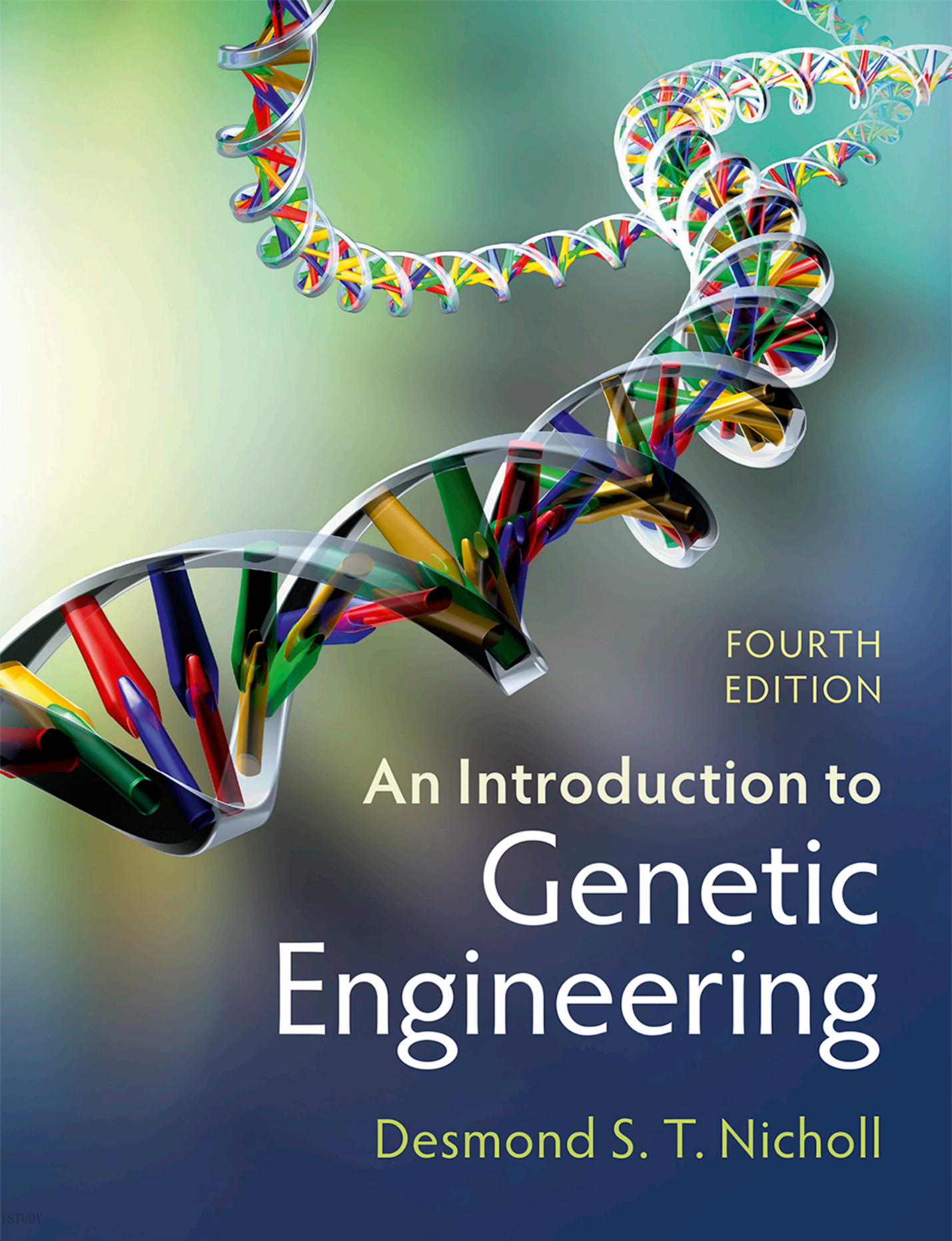

FOURTH EDITION

An Introduction to Genetic Engineering

Desmond S. T. Nicholl

ISTUDY

ISTUDY

An Introduction to Genetic Engineering
Fourth Edition

The fourth edition of this popular textbook retains its focus on the fundamental principles of gene manipulation, providing an accessible and broad-based introduction to the subject for beginning undergraduate students. It has been brought thoroughly up to date with new chapters on the story of DNA and genome editing, and new sections on bioethics, significant developments in sequencing technology and structural, functional and comparative genomics and proteomics, and the impact of transgenic plants. In addition to chapter summaries, learning objectives, concept maps, glossary and key word lists, the book now also features new concluding sections, further reading lists and websearch activities for each chapter to provide a comprehensive suite of learning resources to help students develop a flexible and critical approach to the study of genetic engineering.

Desmond S. T. Nicholl was Senior Lecturer in Biological Sciences, Head of Bioscience, Head of Quality Enhancement and Assistant Dean for Education at the University of the West of Scotland. As well as three previous editions of *An Introduction to Genetic Engineering*, he also authored *Cell and Molecular Biology* (Learning & Teaching Scotland, 2000).

ISTUDY

‘Genetic engineering represents a toolbox that all students within the basic and applied biology fields must get acquainted with. The fourth edition of *An Introduction to Genetic Engineering* is an excellent up-to-date version of a classic textbook. This ambitious book excellently balances the molecular biology knowledge required to grasp the comprehensive gene technology toolbox with a discussion of its impact on society.’

Per Amstrup Pedersen, University of Copenhagen

‘As a biomedical engineering professor teaching an undergraduate Genetic Engineering course for close to 10 years, I use Dr Nicholl’s *An Introduction to Genetic Engineering* as my go-to textbook. It is not one of those overly thick textbooks that overwhelm students. Its comprehensiveness captures readers’ attention with succinct fundamental concepts that truly promote one’s interest in exploring the wonder of many genetic engineering techniques and applications. To facilitate that further, the material provided at the end of each chapter encourages readers to expand their learning with relevant resources … Many of my students become so interested that they pursue graduate degrees and have a career in this field. Dr Nicholl’s textbook has a long-term influence on its readers.’

M. Ete Chan, State University of New York at Stony Brook

‘Dr Nicholl’s book covers all the basic material that one would expect from its title, but what particularly impressed me was how it isn’t afraid to move into political and socio-economic arenas. In Chapter 16, for example, balanced arguments are presented for and against the development of transgenic organisms, and these don’t always come out in favour of the science.’

Neil Crickmore, University of Sussex

ISTUDY

An Introduction to Genetic Engineering

Fourth Edition

Desmond S. T. Nicholl

CAMBRIDGE
UNIVERSITY PRESS

ISTUDY

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/highereducation/isbn/9781009180597

DOI: 10.1017/9781009180610

First and second editions © Cambridge University Press 1994, 2002

Third and fourth editions © Desmond S. T. Nicholl 2008, 2023

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 1994
Second edition 2002
Third edition 2008
Fourth edition 2023

Printed in the United Kingdom by TJ Books Limited, Padstow, Cornwall, 2023

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

A Cataloging-in-Publication data record for this book is available from the Library of Congress

ISBN 978-1-009-18059-7 Hardback
ISBN 978-1-009-18060-3 Paperback

Additional resources for this publication at www.cambridge.org/nicholl4

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

ISTUDY

Contents

ISTUDY

ISTUDY

Detailed Contents

ISTUDY

ISTUDY

ISTUDY

ISTUDY

ISTUDY

ISTUDY

ISTUDY

ISTUDY

Preface

Advances in genetics continue to be made at an ever increasing rate, which presents something of a dilemma when writing an introductory text on the subject. In the years since the third edition was published, many new applications of gene manipulation technology have been developed; genome sequencing has become available at bench-top scale and cost, and gene editing can be achieved using very modest laboratory infrastructure. Personal genome profiling is available from a range of companies, and genetic technology has played a major role in managing many aspects of the COVID-19 pandemic, from diagnostic testing to rapid development of safe and effective vaccines.

Information technology resources, coupled with the internet and World Wide Web, have been critical parts of all these developments, providing tools for the analysis of DNA sequences and instant sharing of data across the globe. At the same time, a level of mistrust has developed among some sections of society, largely driven by misinformation on social media channels, which has illustrated the power of the internet in a less positive way. It is against this background that some themes began to emerge for the fourth edition, reflecting the aim of encouraging students to use the excellent resources on the web, whilst retaining a level of critical assessment of the information. Aspects around the ethics of using genetic technology are perhaps now even more important than before, so these are discussed early in the text to enable the applications to be placed within an appreciation of the ethical framework.

Whilst aiming for a slight broadening in scope, I remain convinced that a basic technical introduction to the subject should be the major focus of the text. Thus, some of the original methods used in gene manipulation have been kept as examples of how the technology developed, even though some of these have become little used or even obsolete. From the educational point of view, this should help the reader cope with more advanced information about the subject, as a sound grasp of the basic principles is an important part of any introduction to genetic engineering. I have been gratified by the many positive comments about the third edition of the text, and I hope that this new edition continues to serve a useful purpose as part of the introductory literature on this fascinating subject.

This book is organised as four parts. *Part 1* (*Genetic Engineering in Context*; Chapters 1–3) sets the scene and brings the discussion of the ethical issues around DNA technology to the start of the book. *Part 2* (*The Basis of Genetic Engineering*; Chapters 4–6) provides an introduction to molecular biology and outlines the tools available to the genetic engineer, and *Part 3* (*The Methodology of Gene Manipulation*; Chapters 7–12) extends this theme further by examining how these tools enable

ISTUDY

sophisticated experiments and procedures to be carried out. Finally, in *Part 4* (*Genetic Engineering in Action*; Chapters 13–17), we look at the impact of DNA technology across a range of key areas.

In the fourth edition, I have expanded the range of features that should be useful as study aids where the text is used to support a particular academic course. In the book, there are *text boxes* sprinkled throughout the chapters. These highlight key points on the way through the text, and can be used as a means of summarising the content. At the start of each chapter, the *aims* of the chapter are presented, along with a *chapter summary* in the form of *learning objectives*. These have been written quite generally, so that an instructor can modify them to suit the level of detail required. A list of the *key words* in each chapter is also provided for reference. These are shown as bold in the text; terms in blue can also be found in the Glossary. A new addition to the end of each chapter is a *websearch* page that provides some structured web-based search exercises that help to set the chapter in context and act as a start point for further study using the resources available online. As in previous editions, a *concept map* has been generated for each chapter, showing how the main topics are linked. The concept maps provided here are essentially summaries of the chapters, and may be examined either before or after reading the chapter.

As this remains an introductory text, no in-text reference has been made to the primary (research) literature, but some suggestions for *further reading* are given at the end of each chapter. Most of these are available in open-access format or may be available through an institution's library subscription service. A *glossary* of terms used has also been provided.

A new development for the fourth edition is a set of *online resources* at www.cambridge.org/nicholl4. This provides access to a range of materials from the book (and additional information) that I hope will be useful in building a learning system to suit your preferred learning style. The resources have been provided in electronic format as a *study guide* to enable collation into a set of student-generated notes.

My thanks go to the anonymous (but appreciated) reviewers of the proposal and the early versions of the manuscript. Their comments and suggestions have made the book better; any errors of fact or interpretation of course remain my own responsibility. Special thanks to Megan Keirnan, Susan Francis, Helen Shannon and Rachel Norridge at Cambridge University Press, and to Joyce Cheung, for their cheerful advice, support, encouragement and patience, which helped bring the project to its conclusion.

My final and biggest thank you goes as ever to my wife Linda and to Charlotte, Thomas and Anna, who have grown up along with the various editions of 'IGE'. I dedicate this new edition to them.

ISTUDY

Part I

Genetic Engineering in Context

ISTUDY

Chapter 1 Summary

Learning Objectives

When you have completed this chapter, you will be able to:

- Define genetic engineering as it will be described in this book
- Outline the basic features of genetic engineering
- Describe the emergence of gene manipulation technology
- Explain the steps required to clone a gene
- Appreciate elements of the ethical debate surrounding genetic engineering
- Identify a range of internet-based resources related to DNA technology

Key Words

Genetic engineering, bioinformatics, gene manipulation, gene cloning, recombinant DNA (rDNA) technology, genetic modification, new genetics, DNA technology, molecular agriculture, genethics, Gregor Mendel, James Watson, Francis Crick, DNA ligase, type II restriction enzyme, plasmid, extrachromosomal element, replicon, clone, genetically modified organism (GMO), internet, World Wide Web, Tim Berners-Lee, uniform resource locator (URL), domain (*re*. URL), search engine, valid, reliable, peer review, Encyclopedia Britannica, Wikipedia, social media, misinformation, disinformation, suggested search term (SST), digital object identifier (DOI).

ISTUDY

Chapter 1

Introduction

1.1 What Is Genetic Engineering?

Making progress in any scientific discipline depends on continually developing techniques and methods to extend the range and sophistication of experiments that may be performed. Within the biosciences, this has been demonstrated in a spectacular way by the emergence and development of **genetic engineering**. In 2022, we marked the fiftieth anniversary of the creation of the first recombinant DNA molecules, an event that is often used to note the start of the recombinant DNA era of genetics. The five decades since 1972 have seen astonishing progress in the breadth and scope of the technology, and it is now routine practice to identify a specific DNA fragment from the genome of an organism, determine its base sequence and assess its function. The sequence might then be altered and replaced into the organism it came from, or a different organism, to achieve a particular goal. We have seen the expansion of the technology into the domain of 'big science' in the era of the Human Genome Project, and its return to the small-scale laboratory as new developments have appeared. Whole genomes can now be sequenced using a benchtop machine, and genome editing enables researchers to alter the genome of an organism with a high level of precision. All of this is now underpinned by the astonishing developments in **bioinformatics**, with sophisticated computational tools available to analyse almost unimaginable amounts of data that are generated on a daily basis.

Several terms may be used to describe the technologies involved in manipulating genes.

The term genetic engineering is often thought to be rather emotive or even trivial, yet it is probably the label that most people would recognise. However, there are several other terms that can be used to describe the technology, including **gene manipulation**, **gene cloning**, **recombinant DNA (rDNA) technology** and **genetic modification**. You may also come across the term the '**new genetics**', although we are at a point where this is perhaps less useful than was the case previously. A more useful generic term that covers a wide range of techniques and applications is simply **DNA technology**. There are also legal definitions used in administering regulatory mechanisms in countries where genetic engineering is practised.

ISTUDY

The genetic material provides a rich resource in the form of information encoded by the sequence of bases in the DNA.

Although there are many diverse and complex techniques involved, the basic principles of genetic manipulation are reasonably simple. The premise on which the technology is based is that genetic information, encoded by DNA and arranged in the form of genes, is a *resource* that can be manipulated in various ways to achieve certain goals in both pure and applied science, medicine, biotechnology and agriculture. There are many areas in which genetic manipulation has made a significant impact, including:

- Basic research on gene structure and function
- Production of useful proteins by novel methods
- Generation of transgenic plants and animals
- Medical diagnosis and treatment
- Forensic analysis of crime scene samples
- Molecular anthropology and the study of evolution
- Genome analysis and genome editing

In later chapters, we will look at how DNA technology has contributed to these areas.

Gene cloning enables isolation and identification of individual genes.

The mainstay of genetic manipulation is the ability to isolate a single DNA sequence from the genome. This is the essence of gene cloning and can be considered as a series of four steps (Fig. 1.1). Successful completion of these steps provides the genetic engineer with a specific DNA sequence, which may then be used for a variety of purposes. A useful analogy is to consider gene cloning as a form of **molecular agriculture**, enabling the production of large amounts (in genetic engineering, this means nanograms or micrograms) of a particular DNA sequence. Although the basic cloning methodology has been extended (and in many cases replaced) by technologies such as the polymerase chain reaction, large-scale DNA sequencing and genome editing, this ability to isolate a particular gene sequence is

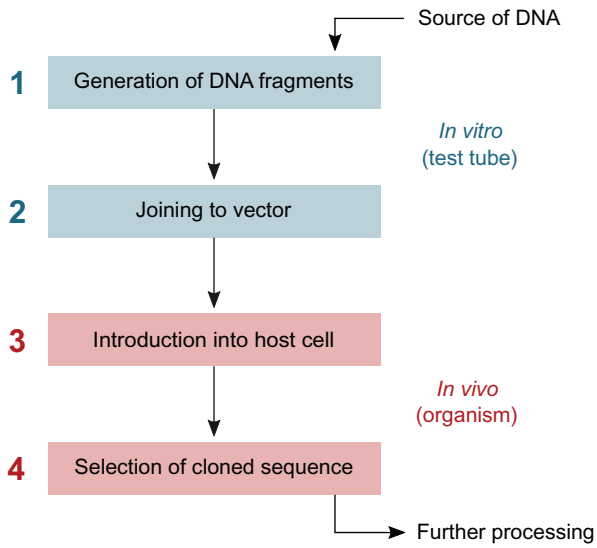

Fig. 1.1 The four steps in cloning a DNA sequence. Steps 1 and 2 are carried out *in vitro* and generate the recombinant DNA molecules. A host organism, such as a bacterium, is used for steps 3 and 4 (*in vivo*). The term *clone* refers to the colonies of identical host cells produced during amplification of the cloned fragments. The cloned sequence can then be isolated and processed further. Gene cloning is sometimes referred to as molecular cloning, to distinguish the process from the cloning of whole organisms.

ISTUDY

still a major part of gene manipulation as carried out on a day-to-day basis in research laboratories worldwide.

As well as technical and scientific challenges, modern genetics poses many moral and ethical questions.

One aspect of genetic engineering that has given cause for concern is the debate surrounding the potential applications of the technology. The term **genethics** is sometimes used to describe the ethical problems that exist in modern genetics, which are likely to increase in both number and complexity as genetic engineering technology becomes more sophisticated and implemented more widely. The use of transgenic plants and animals, investigation of the human genome, gene therapy, genome editing and many other topics are of concern not just to the scientist, but also to the population as a whole. Developments in genetically modified foods have provoked a well-documented public backlash against the technology in many countries. Additional developments in the cloning of organisms, and in areas such as *in vitro* fertilisation and xenotransplantation, raise further questions. Although not strictly part of gene manipulation technology, organismal cloning will be considered later in this book, as this is an area of much concern and can be considered as genetic engineering in its broadest sense. Research on embryonic stem cells, and the potential therapeutic benefits that this may bring, is another area of concern that is part of the general advance of genetics. We will look at some of these ethical aspects in more detail in Chapter 3.

1.2 Laying the Foundations

Gregor Mendel is often considered the 'father' of genetics.

Although the techniques used in gene manipulation began to appear in the 1970s, we should remember that development of these techniques depended on the knowledge and expertise provided by chemists, biochemists and microbial geneticists working in the earlier decades of the twentieth century. We can consider the development of genetics as falling into three main eras (Fig. 1.2). The science of genetics really began with the rediscovery of **Gregor Mendel**'s work at the start of the century, and the next 40 years or so saw the elucidation of the principles of inheritance and genetic mapping. Microbial genetics became established in the mid-1940s, and the role of DNA as the genetic material was confirmed. During this period, great advances were made in understanding the mechanisms of gene transfer between bacteria, and a broad knowledge base was established, from which later developments would emerge.

Watson and Crick's double helix is perhaps the most 'famous' and most easily recognised molecule in the world.

Determination of the structure of DNA by **James Watson** and **Francis Crick** in 1953 provided the stimulus for the development of genetics at the molecular level, and the next few years saw a period of intense activity and excitement as the main features of the gene and its expression were determined. This work culminated in the deciphering of the complete genetic code in 1966, and the stage was now set for the appearance of the new discoveries that would lead to the development of the early techniques in recombinant DNA technology.

ISTUDY

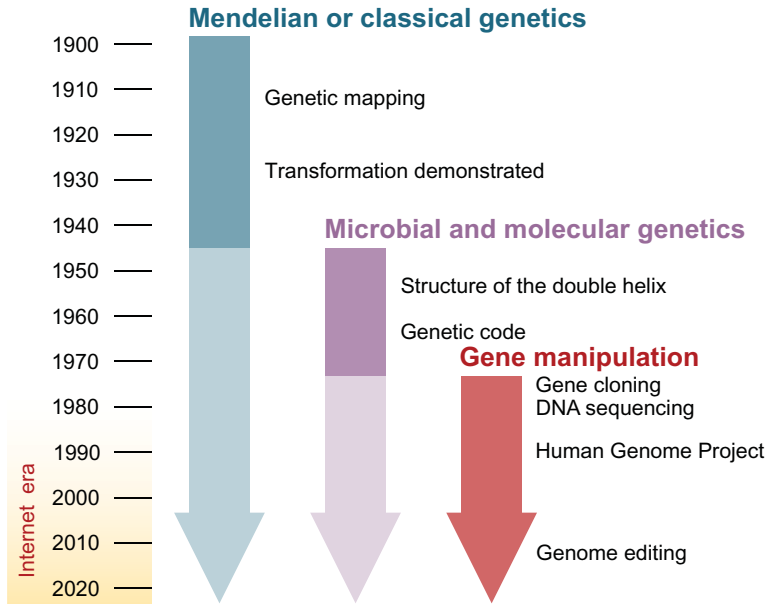

Fig. 1.2 The history of genetics since 1900. Three eras can be identified. Darker shaded areas represent the periods of major development in each branch of the subject, although advances continue to be made in all of these areas.

1.3 First Steps in DNA Cloning

By the end of the 1960s, most of the essential requirements for the emergence of gene technology were in place.

In the late-1960s, there was a sense of frustration among scientists working in the field of molecular biology. Research had developed to the point where progress was being hampered by technical constraints, as the elegant experiments that had helped to decipher the genetic code could not be extended to investigate the gene in more detail. However, a number of developments provided the necessary stimulus for gene manipulation to become a reality. In 1967, the enzyme **DNA ligase** was isolated. This enzyme can join two strands of DNA together, a prerequisite for the construction of recombinant molecules, and can be regarded as a sort of molecular glue. This was followed by the isolation of the first **type II restriction enzyme** in 1970, a major milestone in the development of genetic engineering. Restriction enzymes are essentially molecular scissors that cut DNA at precisely defined sequences. Such enzymes can be used to produce fragments of DNA that are suitable for joining to other fragments. Thus, by 1970, the basic tools required for the construction of recombinant DNA were available.

The first recombinant DNA molecules were generated at Stanford University in 1972, utilising the cleavage properties of restriction enzymes (scissors) and the ability of DNA ligase to join DNA strands together (glue). The importance of these first tentative experiments cannot be overstated. Scientists could now join different DNA molecules together and could link the DNA of one organism to that of a completely different organism. The methodology was extended in

ISTUDY

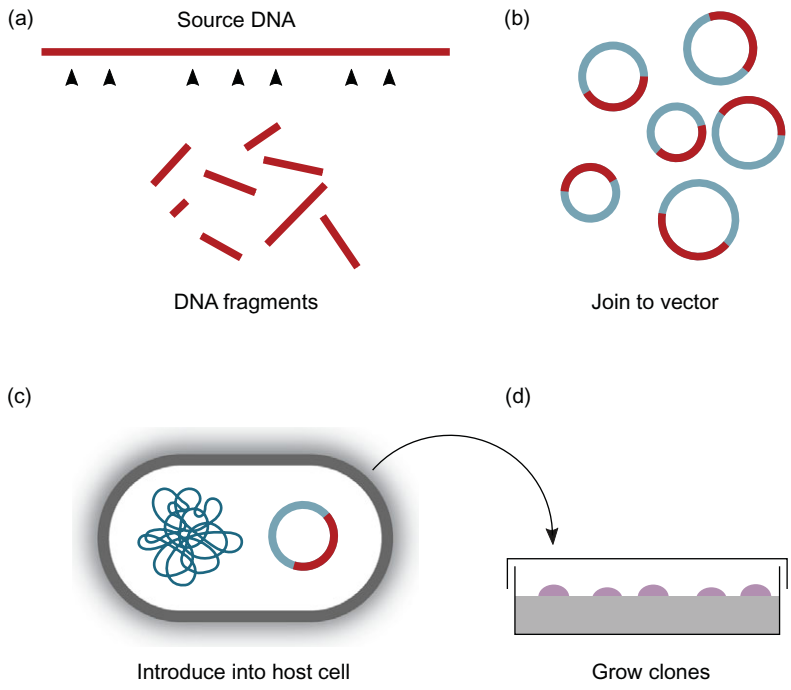

Fig. 1.3 Cloning DNA fragments. (a) The source DNA is isolated and fragmented into suitably sized pieces. This may be carried out mechanically or by using restriction enzymes. (b) The fragments are then joined to a carrier molecule or vector to produce recombinant DNA molecules. In this case, a plasmid vector is shown. (c) The recombinant DNA molecules are then introduced into a host cell (a bacterial cell in this example) for propagation as clones (d). These can be stored as a stable resource that can be used for further analysis.

1973 by joining DNA fragments to the **plasmid** pSC101, which is an **extrachromosomal element** derived from an antibiotic resistance plasmid originally isolated from the bacterium *Salmonella typhimurium*. These recombinant molecules behaved as **replicons**, *i.e.* they could replicate when introduced into *Escherichia coli* cells. Thus, by creating recombinant molecules *in vitro*, and placing the construct in a bacterial cell where it could replicate *in vivo*, specific fragments of DNA could be isolated from bacterial colonies that formed **clones** (colonies formed from a single cell, in which all cells are identical) when grown on agar plates. This development marked the emergence of the technology that became known as gene cloning (Fig. 1.3).

The key to gene cloning is ensuring that the target sequence is replicated in a suitable host cell.

The discoveries of 1972 and 1973 triggered what is perhaps the biggest scientific revolution of all – the 'new genetics' era had arrived. The use of the new technology spread very quickly, and a sense of urgency and excitement prevailed. This was dampened somewhat by the realisation that the new technology could give rise to potentially harmful organisms with undesirable characteristics. It is to the credit of the biological community that measures were adopted to regulate the use of gene manipulation, and that progress in contentious areas was restricted until more information became available about the possible consequences of the inadvertent release of organisms containing recombinant DNA. However, the development of **genetically modified organisms (GMOs)**, particularly crop plants, has reopened the debate about the safety of these organisms and the consequences

The development and use of GMOs pose some difficult ethical questions that may not arise in other areas such as gene cloning.

ISTUDY

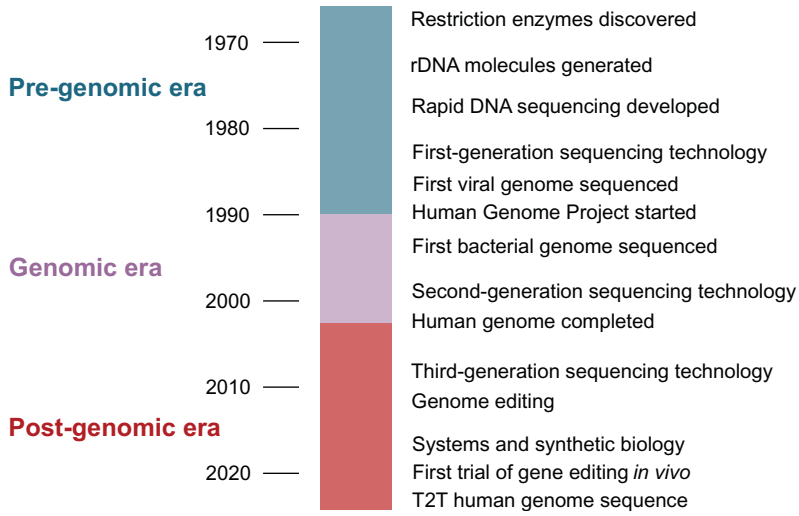

Fig. 1.4 Eras of genomics. Technological development in DNA science and computation in the 1970s and 1980s led to the project to sequence the human genome, which largely defined the genomic era. The knowledge gained has continued to drive the development of techniques and applications in the post-genomic era.

of releasing GMOs into the environment. In addition, many of the potential medical benefits of gene manipulation, genetics and cell biology pose ethical questions that may not be easy to answer. We will come across some of these issues later in this book.

DNA technology continued (and continues) to expand at pace, with a number of key techniques developed in the late-1970s and early-1980s that would enable a step-change in scale to be achieved, with the ambitious project to sequence the human genome being completed in 2003. Further developments have changed DNA sequencing significantly, to the extent that we can now usefully describe DNA technology as itself falling into three eras (Fig. 1.4). As we explore the topics in this book, it may be useful to keep this diagram in mind as a 'roadmap' to help us place the technology in context.

1.4 Using the Web to Support Your Studies

Since the first edition of this book was published in 1994, the growth and development of DNA technology has been impressive, and often astonishing. In many ways, the parallel development of the **internet** and **World Wide Web (www)** is equally impressive, and the generation of students who may be using this book has grown up in a world that is completely immersed in the technology associated with 'the web'. Although information on how to access the internet is no longer needed, some of the guidance for using the web that was included in earlier editions remains appropriate.

The internet was conceived and developed by **Tim Berners-Lee** (now Sir Tim) in 1989 whilst working at CERN in Geneva. If you are not familiar with CERN and Berners-Lee, look these up now using your computer, laptop, tablet or smartphone! The term *internet* is used to

ISTUDY

describe the network of computers that together provide the means to publish and share information, whilst the *World Wide Web* is a more general description of the information that is published using the internet. However, the two terms are often used interchangeably, and the phrases 'surfing the net', 'surfing the web' and 'look it up online' have become part of modern-day language. We will use 'the web' in this book when referring to the World Wide Web.

A really positive feature of the web is the range of material that can be found with a few clicks of a mouse or swipes with a finger. The fluidity and dynamic nature of the web mean that a printed book (like this one) cannot possibly compete; so we will not try to. A subtle shift in emphasis is therefore required – this book is a resource that can help steer you through the confusion of the web, as well as providing a structured look at the subject matter. I have therefore assumed that you will be reading this text with pretty much instant and continuous access to the internet, and that you will be able to use the web to help gather, collate and interpret additional information that you find.

This book can be used as a guide to explore topics more broadly, and/or in more depth, using the web.

Finding websites is generally straightforward if you know where you are going. Each website has an 'address' known as a **uniform resource locator (URL)**. A URL generally begins with http://www. followed by the specific **domain** address. This may end with a country identifier such as .uk, .ca, .cn, .us, .au, *etc*. The most common domain ending is .com (where the term 'dot com' comes from), which is used by around 50 per cent of all websites. Many student textbooks have associated websites, as do research groups, university departments, companies, *etc*.

If you don't know the URL, one of the many **search engines** enables you to look for information using a range of terms, and it is astonishing what you can find. However, caution is needed: there is an awful lot of information on the web, and there is a lot of awful information there as well! It is very easy to get sidetracked and end up wasting a lot of time searching through sites that are of no value (but may be interesting nonetheless). As I write this, I have just typed in the search terms that I used to illustrate this section in the third edition of this book. The number of 'hits' for each of the various terms, in 2007 and 2021, is shown in Table 1.1.

A number of points can be made when we consider the data. Firstly, the number of retrieved items is often far too large to be useful, and thus more specific, defined or restricted search terms will generally produce fewer results. However, even with something like 'sheep cloning' or 'plasmid vector', there is still too much information to look through, so learning to use the search and filter facilities provided by various search engines and websites is well worth the effort. Secondly, it seems that there has been an orders-of-magnitude increase in the number of hits generated by all of these terms. At first glance, this may not seem implausible, given the increase in research and development that will have occurred in these areas since 2007. However, we do need to be a little cautious about inferring too much,

ISTUDY