

# GENE CLONING AND DNA ANALYSIS

An Introduction

**T.A. BROWN**

Faculty of Life Sciences  
University of Manchester  
Manchester

Sixth Edition

 **WILEY-BLACKWELL**

A John Wiley & Sons, Ltd., Publication

# Brief Contents

*Preface to the Sixth Edition* xvi

## **Part I The Basic Principles of Gene Cloning and DNA Analysis** 1

- 1 Why Gene Cloning and DNA Analysis are Important 3
- 2 Vectors for Gene Cloning: Plasmids and Bacteriophages 13
- 3 Purification of DNA from Living Cells 25
- 4 Manipulation of Purified DNA 45
- 5 Introduction of DNA into Living Cells 72
- 6 Cloning Vectors for *E. coli* 88
- 7 Cloning Vectors for Eukaryotes 105
- 8 How to Obtain a Clone of a Specific Gene 126
- 9 The Polymerase Chain Reaction 147

## **Part II The Applications of Gene Cloning and DNA Analysis in Research** 163

- 10 Sequencing Genes and Genomes 165
- 11 Studying Gene Expression and Function 185
- 12 Studying Genomes 207

## **Part III The Applications of Gene Cloning and DNA Analysis in Biotechnology** 223

- 13 Production of Protein from Cloned Genes 225
- 14 Gene Cloning and DNA Analysis in Medicine 245
- 15 Gene Cloning and DNA Analysis in Agriculture 264
- 16 Gene Cloning and DNA Analysis in Forensic Science and Archaeology 282

*Glossary* 298

*Index* 312

Companion website available at [www.wiley.com/go/brown/cloning](http://www.wiley.com/go/brown/cloning)



# Contents

*Preface to the Sixth Edition* xvi

## Part I The Basic Principles of Gene Cloning and DNA Analysis 1

---

- 1 Why Gene Cloning and DNA Analysis are Important 3**
  - 1.1 The early development of genetics 3**
  - 1.2 The advent of gene cloning and the polymerase chain reaction 4**
  - 1.3 What is gene cloning? 5**
  - 1.4 What is PCR? 6**
  - 1.5 Why gene cloning and PCR are so important 7**
    - 1.5.1 Obtaining a pure sample of a gene by cloning 7
    - 1.5.2 PCR can also be used to purify a gene 9
  - 1.6 How to find your way through this book 11**
  
- 2 Vectors for Gene Cloning: Plasmids and Bacteriophages 13**
  - 2.1 Plasmids 13**
    - 2.1.1 Size and copy number 15
    - 2.1.2 Conjugation and compatibility 16
    - 2.1.3 Plasmid classification 16
    - 2.1.4 Plasmids in organisms other than bacteria 17
  - 2.2 Bacteriophages 17**
    - 2.2.1 The phage infection cycle 18
    - 2.2.2 Lysogenic phages 19
      - Gene organization in the  $\lambda$  DNA molecule 19
      - The linear and circular forms of  $\lambda$  DNA 19
      - M13—a filamentous phage 22
    - 2.2.3 Viruses as cloning vectors for other organisms 24

### 3 Purification of DNA from Living Cells 25

#### 3.1 Preparation of total cell DNA 25

3.1.1 Growing and harvesting a bacterial culture 26

3.1.2 Preparation of a cell extract 28

3.1.3 Purification of DNA from a cell extract 29

Removing contaminants by organic extraction and enzyme digestion 29

Using ion-exchange chromatography to purify DNA from a cell extract 30

3.1.4 Concentration of DNA samples 30

3.1.5 Measurement of DNA concentration 31

3.1.6 Other methods for the preparation of total cell DNA 32

#### 3.2 Preparation of plasmid DNA 33

3.2.1 Separation on the basis of size 35

3.2.2 Separation on the basis of conformation 36

Alkaline denaturation 36

Ethidium bromide–caesium chloride density gradient centrifugation 36

3.2.3 Plasmid amplification 39

#### 3.3 Preparation of bacteriophage DNA 39

3.3.1 Growth of cultures to obtain a high  $\lambda$  titer 40

3.3.2 Preparation of non-lysogenic  $\lambda$  phages 40

3.3.3 Collection of phages from an infected culture 42

3.3.4 Purification of DNA from  $\lambda$  phage particles 42

3.3.5 Purification of M13 DNA causes few problems 43

### 4 Manipulation of Purified DNA 45

#### 4.1 The range of DNA manipulative enzymes 46

4.1.1 Nucleases 46

4.1.2 Ligases 47

4.1.3 Polymerases 48

4.1.4 DNA modifying enzymes 49

#### 4.2 Enzymes for cutting DNA—restriction endonucleases 50

4.2.1 The discovery and function of restriction endonucleases 51

4.2.2 Type II restriction endonucleases cut DNA at specific nucleotide sequences 52

4.2.3 Blunt ends and sticky ends 53

4.2.4 The frequency of recognition sequences in a DNA molecule 53

4.2.5 Performing a restriction digest in the laboratory 54

4.2.6 Analysing the result of restriction endonuclease cleavage 56

Separation of molecules by gel electrophoresis 57

Visualizing DNA molecules in an agarose gel 58

4.2.7 Estimation of the sizes of DNA molecules 58

4.2.8 Mapping the positions of different restriction sites in a DNA molecule 59

4.2.9 Special gel electrophoresis methods for separating larger molecules 60

- 4.3 Ligation—joining DNA molecules together** 63
  - 4.3.1 The mode of action of DNA ligase 63
  - 4.3.2 Sticky ends increase the efficiency of ligation 64
  - 4.3.3 Putting sticky ends onto a blunt-ended molecule 64
    - Linkers 64
    - Adaptors 65
    - Producing sticky ends by homopolymer tailing 67
  - 4.3.4 Blunt end ligation with a DNA topoisomerase 69

## 5 Introduction of DNA into Living Cells 72

- 5.1 Transformation—the uptake of DNA by bacterial cells** 74
  - 5.1.1 Not all species of bacteria are equally efficient at DNA uptake 74
  - 5.1.2 Preparation of competent *E. coli* cells 75
  - 5.1.3 Selection for transformed cells 75
- 5.2 Identification of recombinants** 76
  - 5.2.1 Recombinant selection with pBR322—insertional inactivation of an antibiotic resistance gene 77
  - 5.2.2 Insertional inactivation does not always involve antibiotic resistance 79
- 5.3 Introduction of phage DNA into bacterial cells** 81
  - 5.3.1 Transfection 81
  - 5.3.2 *In vitro* packaging of  $\lambda$  cloning vectors 81
  - 5.3.3 Phage infection is visualized as plaques on an agar medium 81
- 5.4 Identification of recombinant phages** 83
  - 5.4.1 Insertional inactivation of a *lacZ'* gene carried by the phage vector 83
  - 5.4.2 Insertional inactivation of the  $\lambda$  *cl* gene 83
  - 5.4.3 Selection using the Spi phenotype 83
  - 5.4.4 Selection on the basis of  $\lambda$  genome size 84
- 5.5 Introduction of DNA into non-bacterial cells** 85
  - 5.5.1 Transformation of individual cells 85
  - 5.5.2 Transformation of whole organisms 85

## 6 Cloning Vectors for *E. coli* 88

- 6.1 Cloning vectors based on *E. coli* plasmids** 89
  - 6.1.1 The nomenclature of plasmid cloning vectors 89
  - 6.1.2 The useful properties of pBR322 89
  - 6.1.3 The pedigree of pBR322 90
  - 6.1.4 More sophisticated *E. coli* plasmid cloning vectors 90
    - pUC8—a Lac selection plasmid 92
    - pGEM3Z—*in vitro* transcription of cloned DNA 93
- 6.2 Cloning vectors based on M13 bacteriophage** 94
  - 6.2.1 How to construct a phage cloning vector 94
  - 6.2.2 Hybrid plasmid–M13 vectors 96

- 6.3 Cloning vectors based on  $\lambda$  bacteriophage** 97
  - 6.3.1 Segments of the  $\lambda$  genome can be deleted without impairing viability 98
  - 6.3.2 Natural selection can be used to isolate modified  $\lambda$  that lack certain restriction sites 98
  - 6.3.3 Insertion and replacement vectors 98
    - Insertion vectors 99
    - Replacement vectors 100
  - 6.3.4 Cloning experiments with  $\lambda$  insertion or replacement vectors 100
  - 6.3.5 Long DNA fragments can be cloned using a cosmid 101
- 6.4  $\lambda$  and other high-capacity vectors enable genomic libraries to be constructed** 102
- 6.5 Vectors for other bacteria** 104

## **7 Cloning Vectors for Eukaryotes** 105

- 7.1 Vectors for yeast and other fungi** 105
  - 7.1.1 Selectable markers for the 2  $\mu$ m plasmid 106
  - 7.1.2 Vectors based on the 2  $\mu$ m plasmid—yeast episomal plasmids 106
  - 7.1.3 A YEp may insert into yeast chromosomal DNA 107
  - 7.1.4 Other types of yeast cloning vector 108
  - 7.1.5 Artificial chromosomes can be used to clone long pieces of DNA in yeast 110
    - The structure and use of a YAC vector 110
    - Applications for YAC vectors 111
  - 7.1.6 Vectors for other yeasts and fungi 112
- 7.2 Cloning vectors for higher plants** 112
  - 7.2.1 *Agrobacterium tumefaciens*—nature's smallest genetic engineer 113
    - Using the Ti plasmid to introduce new genes into a plant cell 113
    - Production of transformed plants with the Ti plasmid 115
    - The Ri plasmid 117
    - Limitations of cloning with *Agrobacterium* plasmids 117
  - 7.2.2 Cloning genes in plants by direct gene transfer 118
    - Direct gene transfer into the nucleus 118
    - Transfer of genes into the chloroplast genome 119
  - 7.2.3 Attempts to use plant viruses as cloning vectors 119
    - Caulimovirus vectors 120
    - Geminivirus vectors 120
- 7.3 Cloning vectors for animals** 120
  - 7.3.1 Cloning vectors for insects 121
    - P elements as cloning vectors for *Drosophila* 121
    - Cloning vectors based on insect viruses 122
  - 7.3.2 Cloning in mammals 122
    - Viruses as cloning vectors for mammals 123
    - Gene cloning without a vector 124

## 8 How to Obtain a Clone of a Specific Gene 126

### 8.1 The problem of selection 126

8.1.1 There are two basic strategies for obtaining the clone you want 127

### 8.2 Direct selection 128

8.2.1 Marker rescue extends the scope of direct selection 129

8.2.2 The scope and limitations of marker rescue 130

### 8.3 Identification of a clone from a gene library 131

8.3.1 Gene libraries 131

8.3.2 Not all genes are expressed at the same time 131

8.3.3 mRNA can be cloned as complementary DNA 133

### 8.4 Methods for clone identification 133

8.4.1 Complementary nucleic acid strands hybridize to each other 133

8.4.2 Colony and plaque hybridization probing 133

Labeling with a radioactive marker 136

Non-radioactive labeling 137

8.4.3 Examples of the practical use of hybridization probing 137

Abundance probing to analyse a cDNA library 137

Oligonucleotide probes for genes whose translation products have been characterized 138

Heterologous probing allows related genes to be identified 141

Southern hybridization enables a specific restriction fragment containing a gene to be identified 142

8.4.4 Identification methods based on detection of the translation product of the cloned gene 144

Antibodies are required for immunological detection methods 144

Using a purified antibody to detect protein in recombinant colonies 145

The problem of gene expression 146

## 9 The Polymerase Chain Reaction 147

### 9.1 The polymerase chain reaction in outline 147

### 9.2 PCR in more detail 149

9.2.1 Designing the oligonucleotide primers for a PCR 149

9.2.2 Working out the correct temperatures to use 152

### 9.3 After the PCR: studying PCR products 153

9.3.1 Gel electrophoresis of PCR products 154

9.3.2 Cloning PCR products 154

9.3.3 Problems with the error rate of *Taq* polymerase 157

### 9.4 Real-time PCR enables the amount of starting material to be quantified 158

9.4.1 Carrying out a quantitative PCR experiment 159

9.4.2 Real-time PCR can also quantify RNA 160

## Part II The Applications of Gene Cloning and DNA Analysis in Research 163

---

### 10 Sequencing Genes and Genomes 165

#### 10.1 The methodology for DNA sequencing 165

##### 10.1.1 Chain termination DNA sequencing 166

Chain termination sequencing in outline 166

Not all DNA polymerases can be used for sequencing 168

Chain termination sequencing requires a single-stranded DNA template 169

The primer determines the region of the template DNA that will be sequenced 169

##### 10.1.2 Pyrosequencing 171

Pyrosequencing involves detection of pulses of chemiluminescence 171

Massively parallel pyrosequencing 171

#### 10.2 How to sequence a genome 173

##### 10.2.1 The shotgun approach to genome sequencing 174

The *Haemophilus influenzae* genome sequencing project 174

Problems with shotgun sequencing 176

##### 10.2.2 The clone contig approach 177

Clone contig assembly by chromosome walking 177

Rapid methods for clone contig assembly 178

Clone contig assembly by sequence tagged site content analysis 179

##### 10.2.3 Using a map to aid sequence assembly 180

Genetic maps 180

Physical maps 181

The importance of a map in sequence assembly 183

### 11 Studying Gene Expression and Function 185

#### 11.1 Studying the RNA transcript of a gene 186

11.1.1 Detecting the presence of a transcript and determining its nucleotide sequence 186

11.1.2 Transcript mapping by hybridization between gene and RNA 188

11.1.3 Transcript analysis by primer extension 190

11.1.4 Transcript analysis by PCR 191

#### 11.2 Studying the regulation of gene expression 192

11.2.1 Identifying protein binding sites on a DNA molecule 193

Gel retardation of DNA-protein complexes 193

Footprinting with DNase I 194

Modification interference assays 194

11.2.2 Identifying control sequences by deletion analysis 197

Reporter genes 197

Carrying out a deletion analysis 198



- 11.3 Identifying and studying the translation product of a cloned gene** 199
  - 11.3.1 HRT and HART can identify the translation product of a cloned gene 199
  - 11.3.2 Analysis of proteins by *in vitro* mutagenesis 200
    - Different types of *in vitro* mutagenesis techniques 202
    - Using an oligonucleotide to create a point mutation in a cloned gene 203
    - Other methods of creating a point mutation in a cloned gene 204
    - The potential of *in vitro* mutagenesis 205

## 12 Studying Genomes 207

- 12.1 Genome annotation** 207
  - 12.1.1 Identifying the genes in a genome sequence 208
    - Searching for open reading frames 208
    - Simple ORF scans are less effective at locating genes in eukaryotic genomes 209
    - Gene location is aided by homology searching 210
    - Comparing the sequences of related genomes 211
  - 12.1.2 Determining the function of an unknown gene 212
    - Assigning gene function by experimental analysis requires a reverse approach to genetics 212
    - Specific genes can be inactivated by homologous recombination 213
- 12.2 Studies of the transcriptome and proteome** 214
  - 12.2.1 Studying the transcriptome 215
    - Studying a transcriptome by sequence analysis 215
    - Studying transcriptomes by microarray or chip analysis 215
  - 12.2.2 Studying the proteome 217
    - Separating the proteins in a proteome 217
    - Identifying the individual proteins after separation 218
  - 12.2.3 Studying protein–protein interactions 220
    - Phage display 220
    - The yeast two hybrid system 220

# Part III The Applications of Gene Cloning and DNA Analysis in Biotechnology 223

## 13 Production of Protein from Cloned Genes 225

- 13.1 Special vectors for expression of foreign genes in *E. coli*** 227
  - 13.1.1 The promoter is the critical component of an expression vector 228
    - The promoter must be chosen with care 228
    - Examples of promoters used in expression vectors 231
  - 13.1.2 Cassettes and gene fusions 232

- 13.2 General problems with the production of recombinant protein in *E. coli*** 234
  - 13.2.1 Problems resulting from the sequence of the foreign gene 235
  - 13.2.2 Problems caused by *E. coli* 236
- 13.3 Production of recombinant protein by eukaryotic cells** 237
  - 13.3.1 Recombinant protein from yeast and filamentous fungi 237
    - Saccharomyces cerevisiae* as the host for recombinant protein synthesis 237
    - Other yeasts and fungi 238
  - 13.3.2 Using animal cells for recombinant protein production 239
    - Protein production in mammalian cells 239
    - Protein production in insect cells 240
  - 13.3.3 Pharming—recombinant protein from live animals and plants 241
    - Pharming in animals 241
    - Recombinant proteins from plants 242
    - Ethical concerns raised by pharming 243

## **14 Gene Cloning and DNA Analysis in Medicine** 245

- 14.1 Production of recombinant pharmaceuticals** 245
  - 14.1.1 Recombinant insulin 246
    - Synthesis and expression of artificial insulin genes 247
  - 14.1.2 Synthesis of human growth hormones in *E. coli* 247
  - 14.1.3 Recombinant factor VIII 249
  - 14.1.4 Synthesis of other recombinant human proteins 251
  - 14.1.5 Recombinant vaccines 252
    - Producing vaccines as recombinant proteins 252
    - Recombinant vaccines in transgenic plants 253
    - Live recombinant virus vaccines 253
- 14.2 Identification of genes responsible for human diseases** 255
  - 14.2.1 How to identify a gene for a genetic disease 256
    - Locating the approximate position of the gene in the human genome 256
    - Identification of candidates for the disease gene 258
- 14.3 Gene therapy** 259
  - 14.3.1 Gene therapy for inherited diseases 259
  - 14.3.2 Gene therapy and cancer 260
  - 14.3.3 The ethical issues raised by gene therapy 262

## **15 Gene Cloning and DNA Analysis in Agriculture** 264

- 15.1 The gene addition approach to plant genetic engineering** 265
  - 15.1.1 Plants that make their own insecticides 265
    - The  $\delta$ -endotoxins of *Bacillus thuringiensis* 265
    - Cloning a  $\delta$ -endotoxin gene in maize 266
    - Cloning  $\delta$ -endotoxin genes in chloroplasts 268
    - Countering insect resistance to  $\delta$ -endotoxin crops 269

- 15.1.2 **Herbicide resistant crops** 270
  - “Roundup Ready” crops 271
  - A new generation of glyphosate resistant crops 272
- 15.1.3 **Other gene addition projects** 273
- 15.2 Gene subtraction** 274
  - 15.2.1 **Antisense RNA and the engineering of fruit ripening in tomato** 274
    - Using antisense RNA to inactivate the polygalacturonase gene 274
    - Using antisense RNA to inactivate ethylene synthesis 276
  - 15.2.2 **Other examples of the use of antisense RNA in plant genetic engineering** 276
- 15.3 Problems with genetically modified plants** 277
  - 15.3.1 **Safety concerns with selectable markers** 277
  - 15.3.2 **The terminator technology** 278
  - 15.3.3 **The possibility of harmful effects on the environment** 279

## **16 Gene Cloning and DNA Analysis in Forensic Science and Archaeology** 282

- 16.1 DNA analysis in the identification of crime suspects** 283
  - 16.1.1 **Genetic fingerprinting by hybridization probing** 283
  - 16.1.2 **DNA profiling by PCR of short tandem repeats** 283
- 16.2 Studying kinship by DNA profiling** 286
  - 16.2.1 **Related individuals have similar DNA profiles** 286
  - 16.2.2 **DNA profiling and the remains of the Romanovs** 286
    - STR analysis of the Romanov bones 286
    - Mitochondrial DNA was used to link the Romanov skeletons with living relatives 287
    - The missing children 289
- 16.3 Sex identification by DNA analysis** 289
  - 16.3.1 **PCRs directed at Y chromosome-specific sequences** 289
  - 16.3.2 **PCR of the amelogenin gene** 290
- 16.4 Archaeogenetics—using DNA to study human prehistory** 291
  - 16.4.1 **The origins of modern humans** 291
    - DNA analysis has challenged the multiregional hypothesis 291
    - DNA analysis shows that Neanderthals are not the ancestors of modern Europeans 293
  - 16.4.2 **DNA can also be used to study prehistoric human migrations** 294
    - The spread of agriculture into Europe 294
    - Using mitochondrial DNA to study past human migrations into Europe 294

*Glossary* 298

*Index* 312