

Contents

CONTENTS

Preface to the Seventh Edition xvii
About the Companion Website xix

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 8
- 1.6 How to find your way through this book 11
 - Further reading 12

2 Vectors for Gene Cloning: Plasmids and Bacteriophages 13

- 2.1 Plasmids 13
 - 2.1.1 Size and copy number 14
 - 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 λ DNA molecule 19
 - The linear and circular forms of λ DNA 19
 - M13 – a filamentous phage 22
 - 2.2.3 Viruses as cloning vectors for other organisms 24
 - Further reading 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
 - Using silica to purify DNA from a cell extract 30
- 3.1.4 Concentration of DNA samples 32
- 3.1.5 Measurement of DNA concentration 33
- 3.1.6 Other methods for the preparation of total cell DNA 34

3.2 Preparation of plasmid DNA 35

- 3.2.1 Separation on the basis of size 35
- 3.2.2 Separation on the basis of conformation 37
 - Alkaline denaturation 37
 - Ethidium bromide–caesium chloride density gradient centrifugation 38
- 3.2.3 Plasmid amplification 39

3.3 Preparation of bacteriophage DNA 40

- 3.3.1 Growth of cultures to obtain a high λ titre 41
- 3.3.2 Preparation of non-lysogenic λ phages 41
- 3.3.3 Collection of phages from an infected culture 43
- 3.3.4 Purification of DNA from λ phage particles 43
- 3.3.5 Purification of M13 DNA causes few problems 43

Further reading 45

4 Manipulation of Purified DNA 47

4.1 The range of DNA manipulative enzymes 48

- 4.1.1 Nucleases 48
- 4.1.2 Ligases 50
- 4.1.3 Polymerases 51
- 4.1.4 DNA-modifying enzymes 52

4.2 Enzymes for cutting DNA: Restriction endonucleases 53

- 4.2.1 The discovery and function of restriction endonucleases 54
- 4.2.2 Type II restriction endonucleases cut DNA at specific nucleotide sequences 55
- 4.2.3 Blunt ends and sticky ends 55
- 4.2.4 The frequency of recognition sequences in a DNA molecule 57
- 4.2.5 Performing a restriction digest in the laboratory 58
- 4.2.6 Analysing the result of restriction endonuclease cleavage 59
 - Separation of molecules by gel electrophoresis 59
 - Visualizing DNA molecules in an agarose gel 60
- 4.2.7 Estimation of the sizes of DNA molecules 61
- 4.2.8 Mapping the positions of different restriction sites in a DNA molecule 62

4.2.9 Special gel electrophoresis methods for separating larger molecules 63

4.3 Ligation: Joining DNA molecules together 66

4.3.1 The mode of action of DNA ligase 66

4.3.2 Sticky ends increase the efficiency of ligation 67

4.3.3 Putting sticky ends on to a blunt-ended molecule 67

Linkers 68

Adaptors 68

Homopolymer tailing 70

4.3.4 Blunt end ligation with a DNA topoisomerase 71

Further reading 74

5 Introduction of DNA into Living Cells 75

5.1 Transformation: The uptake of DNA by bacterial cells 76

5.1.1 Not all species of bacteria are equally efficient at DNA uptake 77

5.1.2 Preparation of competent *E. coli* cells 78

5.1.3 Selection for transformed cells 78

5.2 Identification of recombinants 79

5.2.1 Recombinant selection with pBR322: Insertional inactivation of an antibiotic resistance gene 80

5.2.2 Insertional inactivation does not always involve antibiotic resistance 81

5.3 Introduction of phage DNA into bacterial cells 83

5.3.1 Transfection 83

5.3.2 *In vitro* packaging of λ cloning vectors 83

5.3.3 Phage infection is visualized as plaques on an agar medium 86

5.4 Identification of recombinant phages 86

5.4.1 Insertional inactivation of a *lacZ'* gene carried by the phage vector 87

5.4.2 Insertional inactivation of the λ *cl* gene 87

5.4.3 Selection using the *Spi* phenotype 88

5.4.4 Selection on the basis of λ genome size 88

5.5 Introduction of DNA into non-bacterial cells 88

5.5.1 Transformation of individual cells 88

5.5.2 Transformation of whole organisms 90

Further reading 90

6 Cloning Vectors for *Escherichia coli* 93

6.1 Cloning vectors based on *E. coli* plasmids 94

6.1.1 The nomenclature of plasmid cloning vectors 94

6.1.2 The useful properties of pBR322 94

6.1.3 The pedigree of pBR322 95

6.1.4 More sophisticated *E. coli* plasmid cloning vectors 95

pUC8: A Lac selection plasmid 97

pGEM3Z: *In vitro* transcription of cloned DNA 98

- 6.2 Cloning vectors based on λ bacteriophage** 99
 - 6.2.1 Segments of the λ genome can be deleted without impairing viability 99
 - 6.2.2 Natural selection was used to isolate modified λ that lack certain restriction sites 100
 - 6.2.3 Insertion and replacement vectors 102
 - Insertion vectors 102
 - Replacement vectors 102
 - 6.2.4 Cloning experiments with λ insertion or replacement vectors 103
 - 6.2.5 Long DNA fragments can be cloned using a cosmid 103
 - 6.2.6 λ and other high-capacity vectors enable genomic libraries to be constructed 104
- 6.3 Cloning vectors for the synthesis of single-stranded DNA** 106
 - 6.3.1 Vectors based on M13 bacteriophage 107
 - 6.3.2 Hybrid plasmid-M13 vectors 108
- 6.4 Vectors for other bacteria** 109
 - Further reading 110

7 Cloning Vectors for Eukaryotes 111

- 7.1 Vectors for yeast and other fungi** 111
 - 7.1.1 Selectable markers for the 2 μ m plasmid 112
 - 7.1.2 Vectors based on the 2 μ m plasmid: Yeast episomal plasmids 112
 - 7.1.3 A YE_p may insert into yeast chromosomal DNA 113
 - 7.1.4 Other types of yeast cloning vector 115
 - 7.1.5 Artificial chromosomes can be used to clone long pieces of DNA in yeast 116
 - The structure and use of a YAC vector 116
 - Applications for YAC vectors 118
 - 7.1.6 Vectors for other yeasts and fungi 118
- 7.2 Cloning vectors for higher plants** 119
 - 7.2.1 *Agrobacterium tumefaciens*: nature's smallest genetic engineer 119
 - Using the Ti plasmid to introduce new genes into a plant cell 120
 - Production of transformed plants with the Ti plasmid 122
 - The Ri plasmid 123
 - Limitations of cloning with *Agrobacterium* plasmids 123
 - 7.2.2 Cloning genes in plants by direct gene transfer 124
 - Direct gene transfer into the nucleus 125
 - Transfer of genes into the chloroplast genome 125
 - 7.2.3 Attempts to use plant viruses as cloning vectors 126
 - Caulimovirus vectors 127
 - Geminivirus vectors 127
- 7.3 Cloning vectors for animals** 127
 - 7.3.1 Cloning vectors for insects 128
 - P elements as cloning vectors for *Drosophila* 128
 - Cloning vectors based on insect viruses 129

- 7.3.2 Cloning in mammals 130
 - Viruses as cloning vectors for mammals 130
 - Gene cloning without a vector 131
- Further reading 132

8 How to Obtain a Clone of a Specific Gene 135

- 8.1 The problem of selection 135
 - 8.1.1 There are two basic strategies for obtaining the clone you want 136
- 8.2 Direct selection 137
 - 8.2.1 Marker rescue extends the scope of direct selection 138
 - 8.2.2 The scope and limitations of marker rescue 139
- 8.3 Identification of a clone from a gene library 140
 - 8.3.1 Gene libraries 140
 - Not all genes are expressed at the same time 140
 - mRNA can be cloned as complementary DNA 142
- 8.4 Methods for clone identification 143
 - 8.4.1 Complementary nucleic acid strands hybridize to each other 143
 - 8.4.2 Colony and plaque hybridization probing 144
 - Labelling with a radioactive marker 145
 - Non-radioactive labelling 146
 - 8.4.3 Examples of the practical use of hybridization probing 146
 - Abundancy probing to analyse a cDNA library 147
 - Oligonucleotide probes for genes whose translation products have been characterized 148
 - Heterologous probing allows related genes to be identified 150
 - Southern hybridization enables a specific restriction fragment containing a gene to be identified 151
 - 8.4.4 Identification methods based on detection of the translation product of the cloned gene 153
 - Antibodies are required for immunological detection methods 153
 - Using a purified antibody to detect protein in recombinant colonies 154
 - The problem of gene expression 155
- Further reading 155

9 The Polymerase Chain Reaction 157

- 9.1 PCR in outline 157
- 9.2 PCR in more detail 159
 - 9.2.1 Designing the oligonucleotide primers for a PCR 159
 - 9.2.2 Working out the correct temperatures to use 162
- 9.3 After the PCR: Studying PCR products 164
 - 9.3.1 Gel electrophoresis of PCR products 164
 - 9.3.2 Cloning PCR products 166
 - 9.3.3 Problems with the error rate of *Taq* polymerase 167

- 9.4 Real-time PCR enables the amount of starting material to be quantified 169
 - 9.4.1 Carrying out a quantitative PCR experiment 169
 - 9.4.2 Real-time PCR can also quantify RNA 171
- Further reading 171

Part II The Applications of Gene Cloning and DNA Analysis in Research 173

10 Sequencing Genes and Genomes 175

- 10.1 Chain-termination DNA sequencing 176
 - 10.1.1 Chain-termination sequencing in outline 176
 - 10.1.2 Not all DNA polymerases can be used for sequencing 178
 - 10.1.3 Chain-termination sequencing with *Taq* polymerase 179
 - 10.1.4 Limitations of chain-termination sequencing 180
- 10.2 Next-generation sequencing 182
 - 10.2.1 Preparation of a next-generation sequencing library 182
 - DNA fragmentation 183
 - Immobilization of the library 184
 - Amplification of the library 184
 - 10.2.2 Next-generation sequencing methods 185
 - Reversible terminator sequencing 186
 - Pyrosequencing 187
 - 10.2.3 Third-generation sequencing 188
 - 10.2.4 Directing next-generation sequencing at specific sets of genes 188
- 10.3 How to sequence a genome 189
 - 10.3.1 Shotgun sequencing of prokaryotic genomes 190
 - Shotgun sequencing of the *Haemophilus influenzae* genome 190
 - Shotgun sequencing of other prokaryotic genomes 193
 - 10.3.2 Sequencing of eukaryotic genomes 194
 - The hierarchical shotgun approach 194
 - Shotgun sequencing of eukaryotic genomes 196
 - What do we mean by 'genome sequence'? 198
- Further reading 198

11 Studying Gene Expression and Function 201

- 11.1 Studying the RNA transcript of a gene 202
 - 11.1.1 Detecting the presence of a transcript and determining its nucleotide sequence 203
 - 11.1.2 Transcript mapping by hybridization between gene and RNA 204
 - 11.1.3 Transcript analysis by primer extension 205
 - 11.1.4 Transcript analysis by PCR 206

- 11.2 Studying the regulation of gene expression** 207
 - 11.2.1 Identifying protein binding sites on a DNA molecule 209
 - Gel retardation of DNA–protein complexes 209
 - Footprinting with DNase I 210
 - Modification interference assays 212
 - 11.2.2 Identifying control sequences by deletion analysis 212
 - Reporter genes 213
 - Carrying out a deletion analysis 215
- 11.3 Identifying and studying the translation product of a cloned gene** 216
 - 11.3.1 HRT and HART can identify the translation product of a cloned gene 216
 - 11.3.2 Analysis of proteins by *in vitro* mutagenesis 216
 - Different types of *in vitro* mutagenesis techniques 218
 - Using an oligonucleotide to create a point mutation in a cloned gene 220
 - Other methods of creating a point mutation in a cloned gene 220
 - The potential of *in vitro* mutagenesis 223
- Further reading** 223

12 Studying Genomes 225

- 12.1 Genome annotation** 225
 - 12.1.1 Identifying the genes in a genome sequence 226
 - Searching for open reading frames 226
 - Simple ORF scans are less effective at locating genes in eukaryotic genomes 227
 - Gene location is aided by homology searching 228
 - Comparing the sequences of related genomes 229
 - Identifying the binding sites for regulatory proteins in a genome sequence 230
 - 12.1.2 Determining the function of an unknown gene 231
 - Assigning gene function by experimental analysis requires a reverse approach to genetics 231
 - Specific genes can be inactivated by homologous recombination 232
- 12.2 Studies of the transcriptome and proteome** 233
 - 12.2.1 Studying the transcriptome 234
 - Studying transcriptomes by microarray or chip analysis 234
 - Studying a transcriptome by SAGE 235
 - Sequencing a transcriptome by RNA-seq 236
 - Advantages of the different methods for transcriptome analysis 237
 - 12.2.2 Studying the proteome 237
 - Separating the proteins in a proteome 238
 - Identifying the individual proteins after separation 239
 - 12.2.3 Studying protein–protein interactions 240
 - Phage display 241
 - The yeast two-hybrid system 242
- Further reading** 243

Part III The Applications of Gene Cloning and DNA Analysis in Biotechnology 245

13 Production of Protein from Cloned Genes 247

13.1 Special vectors for the expression of foreign genes in *E. coli* 249

13.1.1 The promoter is the critical component of an expression vector 251

The promoter must be chosen with care 251

Examples of promoters used in expression vectors 253

13.1.2 Cassettes and gene fusions 254

13.2 General problems with the production of recombinant protein in *E. coli* 257

13.2.1 Problems resulting from the sequence of the foreign gene 257

13.2.2 Problems caused by *E. coli* 258

13.3 Production of recombinant protein by eukaryotic cells 259

13.3.1 Recombinant protein from yeasts and filamentous fungi 260
Saccharomyces cerevisiae as the host for recombinant protein synthesis 260

Other yeasts and fungi 261

13.3.2 Using animal cells for recombinant protein production 262

Protein production in mammalian cells 262

Protein production in insect cells 262

13.3.3 Pharming: Recombinant protein from live animals and plants 263

Pharming in animals 263

Recombinant proteins from plants 265

Ethical concerns raised by pharming 265

Further reading 266

14 Gene Cloning and DNA Analysis in Medicine 269

14.1 Production of recombinant pharmaceuticals 269

14.1.1 Recombinant insulin 270

Synthesis and expression of artificial insulin genes 270

14.1.2 Synthesis of human growth hormones in *E. coli* 271

14.1.3 Recombinant factor VIII 274

14.1.4 Synthesis of other recombinant human proteins 275

14.1.5 Recombinant vaccines 275

Producing vaccines as recombinant proteins 276

Recombinant vaccines in transgenic plants 277

Live recombinant virus vaccines 279

14.2 Identification of genes responsible for human diseases 280

14.2.1 How to identify a gene for a genetic disease 282

Locating the approximate position of the gene in the human genome 282

Linkage analysis of the human *BRCA1* gene 283

Identification of candidates for the disease gene 284

14.3 Gene therapy 286

14.3.1 Gene therapy for inherited diseases 286

14.3.2 Gene therapy and cancer 288

14.3.3 The ethical issues raised by gene therapy 288

Further reading 290

15 Gene Cloning and DNA Analysis in Agriculture 291**15.1 The gene addition approach to plant genetic engineering 292**

15.1.1 Plants that make their own insecticides 292

The δ -endotoxins of *Bacillus thuringiensis* 292Cloning a δ -endotoxin gene in maize 293Cloning δ -endotoxin genes in chloroplasts 295Countering insect resistance to δ -endotoxin crops 296

15.1.2 Herbicide-resistant crops 298

'Roundup Ready' crops 298

A new generation of glyphosate-resistant crops 299

15.1.3 Other gene addition projects 300

15.2 Gene subtraction 302

15.2.1 Antisense RNA and the engineering of fruit ripening in tomato 302

Using antisense RNA to inactivate the polygalacturonase gene 302

Using antisense RNA to inactivate ethylene synthesis 304

15.2.2 Other examples of the use of antisense RNA in plant genetic engineering 304

15.3 Problems with genetically modified plants 305

15.3.1 Safety concerns with selectable markers 305

15.3.2 The terminator technology 306

15.3.3 The possibility of harmful effects on the environment 307

Further reading 308

16 Gene Cloning and DNA Analysis in Forensic Science and Archaeology 311**16.1 DNA analysis in the identification of crime suspects 312**

16.1.1 Genetic fingerprinting by hybridization probing 312

16.1.2 DNA profiling by PCR of short tandem repeats 312

16.2 Studying kinship by DNA profiling 315

16.2.1 Related individuals have similar DNA profiles 315

16.2.2 DNA profiling and the remains of the Romanovs 315

STR analysis of the Romanov bones 315

Mitochondrial DNA was used to link the Romanov skeletons with living relatives 317

The missing children 318

16.3 Sex identification by DNA analysis 318

16.3.1 PCRs directed at Y chromosome-specific sequences 318

16.3.2 PCR of the amelogenin gene 319

- 16.4 Archaeogenetics: Using DNA to study human prehistory 320**
- 16.4.1 The origins of modern humans 320
 - DNA analysis has challenged the multiregional hypothesis 321
 - DNA analysis shows that Neanderthals are not the direct ancestors of modern Europeans 322
 - The Neanderthal genome sequence suggests there was interbreeding with *H. sapiens* 323
 - 16.4.2 DNA can also be used to study prehistoric human migrations 324
 - Modern humans may have migrated from Ethiopia to Arabia 324
 - Colonization of the New World 325
 - Further reading 328**

Glossary 329

Index 345