

ISBN 9780470683866

Table of Contents

Preface xiii

1 From Genes to Genomes 1

- 1.1 Introduction 1
- 1.2 Basic molecular biology 4
 - 1.2.1 The DNA backbone 4
 - 1.2.2 The base pairs 6
 - 1.2.3 RNA structure 10
 - 1.2.4 Nucleic acid synthesis 11
 - 1.2.5 Coiling and supercoiling 11
- 1.3 What is a gene? 13
- 1.4 Information flow: gene expression 15
 - 1.4.1 Transcription 16
 - 1.4.2 Translation 19
- 1.5 Gene structure and organisation 20
 - 1.5.1 Operons 20
 - 1.5.2 Exons and introns 21
- 1.6 Refinements of the model 22

2 How to Clone a Gene 25

- 2.1 What is cloning? 25
- 2.2 Overview of the procedures 26
- 2.3 Extraction and purification of nucleic acids 29
 - 2.3.1 Breaking up cells and tissues 29
 - 2.3.2 Alkaline denaturation 31
 - 2.3.3 Column purification 31
- 2.4 Detection and quantitation of nucleic acids 32
- 2.5 Gel electrophoresis 33
 - 2.5.1 Analytical gel electrophoresis 33
 - 2.5.2 Preparative gel electrophoresis 36
- 2.6 Restriction endonucleases 36
 - 2.6.1 Specificity 37
 - 2.6.2 Sticky and blunt ends 40
- 2.7 Ligation 42
 - 2.7.1 Optimising ligation conditions 44
 - 2.7.2 Preventing unwanted ligation: alkaline phosphatase and double digests 46
 - 2.7.3 Other ways of joining DNA fragments 48
- 2.8 Modification of restriction fragment ends 49
 - 2.8.1 Linkers and adaptors 50
 - 2.8.2 Homopolymer tailing 52
- 2.9 Plasmid vectors 53
 - 2.9.1 Plasmid replication 54
 - 2.9.2 Cloning sites 55
 - 2.9.3 Selectable markers 57
 - 2.9.4 Insertional inactivation 58
 - 2.9.5 Transformation 59
- 2.10 Vectors based on the lambda bacteriophage 61
 - 2.10.1 Lambda biology 61
 - 2.10.2 In vitro packaging 65
 - 2.10.3 Insertion vectors 66
 - 2.10.4 Replacement vectors 68
- 2.11 Cosmids 71

2.12 Supervectors: YACs and BACs 72

2.13 Summary 73

3 Genomic and cDNA Libraries 75

3.1 Genomic libraries 77

3.1.1 Partial digests 77

3.1.2 Choice of vectors 80

3.1.3 Construction and evaluation of a genomic library 83

3.2 Growing and storing libraries 86

3.3 cDNA libraries 87

3.3.1 Isolation of mRNA 88

3.3.2 cDNA synthesis 89

3.3.3 Bacterial cDNA 93

3.4 Screening libraries with gene probes 94

3.4.1 Hybridization 94

3.4.2 Labelling probes 98

3.4.3 Steps in a hybridization experiment 99

3.4.4 Screening procedure 100

3.4.5 Probe selection and generation 101

3.5 Screening expression libraries with antibodies 103

3.6 Characterization of plasmid clones 106

3.6.1 Southern blots 107

3.6.2 PCR and sequence analysis 108

4 Polymerase Chain Reaction (PCR) 109

4.1 The PCR reaction 110

4.2 PCR in practice 114

4.2.1 Optimisation of the PCR reaction 114

4.2.2 Primer design 115

4.2.3 Analysis of PCR products 117

4.2.4 Contamination 118

4.3 Cloning PCR products 119

4.4 Long-range PCR 121

4.5 Reverse-transcription PCR 123

4.6 Quantitative and real-time PCR 123

4.6.1 SYBR Green 123

4.6.2 TaqMan 125

4.6.3 Molecular beacons 125

4.7 Applications of PCR 127

4.7.1 Probes and other modified products 127

4.7.2 PCR cloning strategies 128

4.7.3 Analysis of recombinant clones and rare events 129

4.7.4 Diagnostic applications 130

5 Sequencing a Cloned Gene 131

5.1 DNA sequencing 131

5.1.1 Principles of DNA sequencing 131

5.1.2 Automated sequencing 136

5.1.3 Extending the sequence 137

5.1.4 Shotgun sequencing; contig assembly 138

5.2 Databank entries and annotation 140

5.3 Sequence analysis 146

5.3.1 Identification of coding region 146

5.3.2 Expression signals 147

5.4 Sequence comparisons 148

5.4.1 DNA sequences 148

5.4.2 Protein sequence comparisons 151

5.4.3 Sequence alignments: Clustal 157

5.5 Protein structure 160

5.5.1	Structure predictions	160
5.5.2	Protein motifs and domains	162
5.6	Confirming gene function	165
5.6.1	Allelic replacement and gene knockouts	166
5.6.2	Complementation	168
6	Analysis of Gene Expression	169
6.1	Analysing transcription	169
6.1.1	Northern blots	170
6.1.2	Reverse transcription-PCR	171
6.1.3	In situ hybridization	174
6.2	Methods for studying the promoter	174
6.2.1	Locating the promoter	175
6.2.2	Reporter genes	177
6.3	Regulatory elements and DNA-binding proteins	179
6.3.1	Yeast one-hybrid assays	179
6.3.2	DNase I footprinting	181
6.3.3	Gel retardation assays	181
6.3.4	Chromatin immunoprecipitation (ChIP)	183
6.4	Translational analysis	185
6.4.1	Western blots	185
6.4.2	Immunocytochemistry and immunohistochemistry	187
7	Products from Native and Manipulated Cloned Genes	189
7.1	Factors affecting expression of cloned genes	190
7.1.1	Transcription	190
7.1.2	Translation initiation	192
7.1.3	Codon usage	193
7.1.4	Nature of the protein product	194
7.2	Expression of cloned genes in bacteria	195
7.2.1	Transcriptional fusions	195
7.2.2	Stability: conditional expression	198
7.2.3	Expression of lethal genes	201
7.2.4	Translational fusions	201
7.3	Yeast systems	204
7.3.1	Cloning vectors for yeasts	204
7.3.2	Yeast expression systems	206
7.4	Expression in insect cells: baculovirus systems	208
7.5	Mammalian cells	209
7.5.1	Cloning vectors for mammalian cells	210
7.5.2	Expression in mammalian cells	213
7.6	Adding tags and signals	215
7.6.1	Tagged proteins	215
7.6.2	Secretion signals	217
7.7	In vitro mutagenesis	218
7.7.1	Site-directed mutagenesis	218
7.7.2	Synthetic genes	223
7.7.3	Assembly PCR	223
7.7.4	Synthetic genomes	224
7.7.5	Protein engineering	224
7.8	Vaccines	225
7.8.1	Subunit vaccines	225
7.8.2	DNA vaccines	226
8	Genomic Analysis	229
8.1	Overview of genome sequencing	229
8.1.1	Strategies	230
8.2	Next generation sequencing (NGS)	231
8.2.1	Pyrosequencing (454)	232

8.2.2 SOLiD sequencing (Applied Biosystems)	235
8.2.3 Bridge amplification sequencing (Solexa/Illumina)	237
8.2.4 Other technologies	239
8.3 De novo sequence assembly	239
8.3.1 Repetitive elements and gaps	240
8.4 Analysis and annotation	242
8.4.1 Identification of ORFs	243
8.4.2 Identification of the function of genes and their products	250
8.4.3 Other features of nucleic acid sequences	251
8.5 Comparing genomes	256
8.5.1 BLAST	256
8.5.2 Synteny	257
8.6 Genome browsers	258
8.7 Relating genes and functions: genetic and physical maps	260
8.7.1 Linkage analysis	261
8.7.2 Ordered libraries and chromosome walking	262
8.8 Transposon mutagenesis and other screening techniques	263
8.8.1 Transposition in bacteria	263
8.8.2 Transposition in Drosophila	266
8.8.3 Transposition in other organisms	268
8.8.4 Signature-tagged mutagenesis	269
8.9 Gene knockouts, gene knockdowns and gene silencing	271
8.10 Metagenomics	273
8.11 Conclusion	274

9 Analysis of Genetic Variation 275

9.1 Single nucleotide polymorphisms	276
9.1.1 Direct sequencing	278
9.1.2 SNP arrays	279
9.2 Larger scale variations	280
9.2.1 Microarrays and indels	281
9.3 Other methods for studying variation	282
9.3.1 Genomic Southern blot analysis: restriction fragment length polymorphisms (RFLPs)	282
9.3.2 VNTR and microsatellites	285
9.3.3 Pulsed-field gel electrophoresis	287
9.4 Human genetic variation: relating phenotype to genotype	289
9.4.1 Linkage analysis	289
9.4.2 Genome-wide association studies (GWAS)	292
9.4.3 Database resources	294
9.4.4 Genetic diagnosis	294
9.5 Molecular phylogeny	295
9.5.1 Methods for constructing trees	298

10 Post-Genomic Analysis 305

10.1 Analysing transcription: transcriptomes	305
10.1.1 Differential screening	306
10.1.2 Other methods: transposons and reporters	308
10.2 Array-based methods	308
10.2.1 Expressed sequence tag (EST) arrays	309
10.2.2 PCR product arrays	310
10.2.3 Synthetic oligonucleotide arrays	312
10.2.4 Important factors in array hybridization	313
10.3 Transcriptome sequencing	315
10.4 Translational analysis: proteomics	316
10.4.1 Two-dimensional electrophoresis	317
10.4.2 Mass spectrometry	318
10.5 Post-translational analysis: protein interactions	320
10.5.1 Two-hybrid screening	320
10.5.2 Phage display libraries	321
10.6 Epigenetics	323

- 10.7 Integrative studies: systems biology 324
- 10.7.1 Metabolomic analysis 324
- 10.7.2 Pathway analysis and systems biology 325

11 Modifying Organisms: Transgenics 327

- 11.1 Transgenesis and cloning 327
 - 11.1.1 Common species used for transgenesis 328
 - 11.1.2 Control of transgene expression 330
- 11.2 Animal transgenesis 333
 - 11.2.1 Basic methods 333
 - 11.2.2 Direct injection 333
 - 11.2.3 Retroviral vectors 335
 - 11.2.4 Embryonic stem cell technology 336
 - 11.2.5 Gene knockouts 339
 - 11.2.6 Gene knock-down technology: RNA interference 340
 - 11.2.7 Gene knock-in technology 341
- 11.3 Applications of transgenic animals 342
- 11.4 Disease prevention and treatment 343
 - 11.4.1 Live vaccine production: modification of bacteria and viruses 343
 - 11.4.2 Gene therapy 346
 - 11.4.3 Viral vectors for gene therapy 347
- 11.5 Transgenic plants and their applications 349
 - 11.5.1 Introducing foreign genes 349
 - 11.5.2 Gene subtraction 351
 - 11.5.3 Applications 352
- 11.6 Transgenics: a coda 353

Glossary 355

Bibliography 375

Index 379