

# Contents

<b>Preface</b>	<b>xi</b>
<b>About the Author</b>	<b>xiii</b>
<b>Acknowledgements</b>	<b>xv</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Review of Molecular Genetics	1
1.1.1 Nucleic acids	1
1.1.2 Constituents of the genome	3
1.1.3 Organelle genomes	4
1.1.4 The structure of genes	5
1.1.5 Transcription and translation	6
1.1.6 Genes, multigene families and conserved sequences	8
1.1.7 Coding and non-coding fractions of the genome	9
1.2 The Tools of Molecular Biology	9
1.2.1 Restriction enzymes and electrophoresis	9
1.2.2 Cloning and construction of libraries	11
1.2.3 Hybridization techniques	11
1.2.4 Enzymatic amplification (PCR)	14
1.2.5 Sequencing DNA and sequence analysis	16
1.2.6 Chromosomal assignment	16
1.2.7 Techniques specific to genomic analysis	18
1.3 Specifics of the Genome Programmes	20
1.3.1 Creation of the genome programmes	20
1.3.2 Genome research centres	22
1.3.3 The growing importance of bioinformatics	22
1.3.4 Automation and robotics	25
1.4 The Species Analysed	25
1.4.1 The main branches of life	25
1.4.2 Prokaryotes	26
1.4.3 Eukaryotes	28

<b>2</b>	<b>Linkage Maps</b>	<b>33</b>
2.1	Tools and Methods in Genetic Mapping	33
2.1.1	The genetic maps of <i>Drosophila</i> and humans	33
2.1.2	Tools for genetic mapping in humans	35
2.1.3	RFLP markers	36
2.1.4	Microsatellites	38
2.1.5	Mapping methodology	38
2.2	The Development of Genetic Maps	40
2.2.1	RFLPs and the first genetic maps	40
2.2.2	Microsatellites and modern maps	40
2.2.3	Progress of the human genetic map	43
2.2.4	SNP maps	43
2.2.5	Genetic maps of model and domestic species	46
2.2.6	Size of genetic maps	48
2.3	Radiation Hybrid Maps	50
2.3.1	Principles of radiation hybrid mapping	50
2.3.2	The establishment of radiation hybrid maps	50
2.3.3	The first maps	51
2.3.4	Size of maps	51
2.4	Conclusion	52
<b>3</b>	<b>Physical Maps</b>	<b>53</b>
3.1	Local Maps or Small Genomes	54
3.1.1	Mapping tools	54
3.1.2	Physical mapping of model species	54
3.1.3	Local maps in humans	57
3.2	Strategies for Physical Mapping of the Human Genome	58
3.2.1	The chromosome-specific approach	58
3.2.2	The whole-genome approach	58
3.2.3	STS mapping	59
3.2.4	Restriction mapping	61
3.2.5	Mapping by hybridization	61
3.3	Maps of the Human Genome	63
3.3.1	The CEPH/Généthon map	63
3.3.2	The WI/MIT map	64
3.3.3	Chromosome-specific mapping	66
3.3.4	New generation human physical maps	66
3.4	Conclusion	69
<b>4</b>	<b>Genome Sequencing</b>	<b>71</b>
4.1	Strategic Choices	71
4.1.1	Approaches used for large-scale sequencing	71
4.1.2	Sequencing strategies	72

4.1.3	Organisms sequenced	75
4.1.4	Identifying genes	78
4.2	Prokaryotic Genomes	80
4.2.1	Chromosome structure	80
4.2.2	Gene organization	81
4.2.3	Remarkable genes	84
4.2.4	Virulence	85
4.2.5	Non-coding sequences	87
4.2.6	Comparative genomics	90
4.3	Genomes of Model Eukaryotes	92
4.3.1	Chromosome structure	92
4.3.2	Identification of genes	94
4.3.3	Functions of recognized or predicted genes	95
4.3.4	Genes specific to metazoans	98
4.3.5	Plant genomics	99
4.3.6	Homologues of genes responsible for human disease	101
4.3.7	Non-coding regions	102
4.3.8	Evolutionary genomics	103
4.4	The Human Genome	104
4.4.1	Human chromosomes	104
4.4.2	Identification of genes	106
4.4.3	Repeated sequences	110
4.4.4	Evolution	111
4.5	Other Genomes Sequenced	113
4.5.1	<i>Plasmodium falciparum</i>	113
4.5.2	<i>Anopheles</i>	116
4.5.3	Microsporidium	118
4.5.4	<i>Leishmania major</i>	120
4.6	Conclusion	121
<b>5</b>	<b>Sequencing cDNA and the Transcriptome</b>	<b>123</b>
5.1	Strategies of cDNA Sequencing	123
5.1.1	Posing the problem	123
5.1.2	Production and sequencing of cDNA	124
5.1.3	Choice of tissue of origin	125
5.1.4	Large-scale sequencing of cDNA	126
5.2	The Economic Stakes	127
5.2.1	Data protection	127
5.2.2	The involvement of large pharmaceutical companies	127
5.2.3	The 'coining' of partial sequences	128
5.2.4	The cDNA race	129
5.3	The Analysis of cDNA Sequences	129
5.3.1	Assembly of partial sequences	129

5.3.2	Identification of new genes	130
5.3.3	The transcriptional map of the human genome	131
5.3.4	Identification of genes responsible for genetic diseases	131
5.3.5	cDNA programmes in animal and plant species	132
5.4	The Transcriptome	133
5.4.1	Local analyses of transcription	133
5.4.2	Massive sequencing and the global vision of cellular physiology	133
5.4.3	Use of nucleic acid chips	136
5.4.4	The SAGE technique	140
5.5	Conclusion	145
<b>6</b>	<b>The Proteome</b>	<b>147</b>
6.1	Basic Techniques	148
6.1.1	Electrophoresis	148
6.1.2	Chromatography	150
6.1.3	Ultracentrifugation	150
6.1.4	Use of antibodies	152
6.1.5	Study of protein interactions	154
6.1.6	Informatics	155
6.2	Transgenesis	156
6.2.1	Mouse transgenics	157
6.2.2	Examples of mouse transgenics	159
6.2.3	Transgenesis in other species	160
6.3	Mutagenesis	162
6.3.1	Directed mutagenesis	163
6.3.2	Directed mutagenesis in mice ( <i>knock-out</i> )	165
6.3.3	Examples of knock-out in mice	167
6.3.4	Random mutagenesis	168
6.4	Two-dimensional Electrophoresis and Identification of Proteins	169
6.4.1	Separation of proteins by two-dimensional electrophoresis	170
6.4.2	Protein identification by classic techniques	171
6.4.3	Identification of proteins by mass spectrometry	173
6.4.4	Examples of applications	177
6.4.5	Protein interactions	178
6.5	Identification of Protein Interactions by Two-hybrid System	179
6.5.1	Strategy	179
6.5.2	Protein–protein interactions	183
6.5.3	RNA–protein interactions	185
6.5.4	Global approaches	186

6.6	Chip Technology	188
6.6.1	Construction of protein chips	188
6.6.2	Strategies employed	189
6.7	Analysis of Three-dimensional Structure	191
6.7.1	Crystallography	191
6.7.2	Nuclear magnetic resonance	192
6.7.3	Protein structure	192
6.7.4	Example: histocompatibility antigens	193
6.8	Conclusion	194
<b>7</b>	<b>Identification of Genes responsible for Disease</b>	<b>197</b>
7.1	Genetic Diseases	197
7.1.1	Monogenic and multifactorial diseases	197
7.1.2	Mutations	199
7.1.3	Cloning genes responsible for diseases	201
7.2	Functional Cloning and Chromosomal Anomalies	202
7.2.1	Examples of functional cloning	202
7.2.2	Chromosomal anomalies	204
7.2.3	Mitochondrial diseases	205
7.3	Strategy for Positional Cloning	205
7.3.1	Recruitment of affected families	205
7.3.2	Genetic mapping and primary localization	205
7.3.3	Physical mapping	207
7.3.4	Identification of genes present in a localization interval, and of the disease-causing gene	207
7.3.5	The first success of positional cloning	208
7.4	What is the Future for the Cloning of Disease-causing Genes?	209
7.4.1	Monogenic diseases	209
7.4.2	Multifactorial diseases	211
7.4.3	Sequencing	212
7.4.4	Towards a redefinition of some genetic diseases	212
7.4.5	New mutational mechanisms	213
7.4.6	Genetic diseases and therapies	215
7.5	Conclusion	218
	<b>General Conclusion</b>	<b>219</b>
	<b>Further Reading</b>	<b>221</b>
	<b>Index</b>	<b>223</b>