
CONTENTS IN BRIEF

CHAPTER 1	GENOMES, TRANSCRIPTOMES, AND PROTEOMES	1
CHAPTER 2	STUDYING DNA	27
CHAPTER 3	MAPPING GENOMES	55
CHAPTER 4	SEQUENCING GENOMES	87
CHAPTER 5	GENOME ANNOTATION	119
CHAPTER 6	IDENTIFYING GENE FUNCTIONS	135
CHAPTER 7	EUKARYOTIC NUCLEAR GENOMES	155
CHAPTER 8	GENOMES OF PROKARYOTES AND EUKARYOTIC ORGANELLES	181
CHAPTER 9	VIRAL GENOMES AND MOBILE GENETIC ELEMENTS	203
CHAPTER 10	ACCESSING THE GENOME	219
CHAPTER 11	THE ROLE OF DNA-BINDING PROTEINS IN GENOME EXPRESSION	241
CHAPTER 12	TRANSCRIPTOMES	257
CHAPTER 13	PROTEOMES	293
CHAPTER 14	GENOME EXPRESSION IN THE CONTEXT OF CELL AND ORGANISM	329
CHAPTER 15	GENOME REPLICATION	357
CHAPTER 16	MUTATIONS AND DNA REPAIR	389
CHAPTER 17	RECOMBINATION AND TRANSPOSITION	411
CHAPTER 18	HOW GENOMES EVOLVE	429
GLOSSARY		463
INDEX		491

CONTENTS

CHAPTER 1 GENOMES, TRANSCRIPTOMES, AND PROTEOMES

1.1 DNA

Genes are made of DNA

DNA is a polymer of nucleotides

The double helix is stabilized by base pairing and base stacking

The double helix has structural flexibility

1.2 RNA AND THE TRANSCRIPTOME

RNA is a second type of polynucleotide

The RNA content of the cell

Many RNAs are synthesized as precursor molecules

There are different definitions of the transcriptome

1.3 PROTEINS AND THE PROTEOME

There are four hierarchical levels of protein structure

Amino acid diversity underlies protein diversity

The link between the transcriptome and the proteome

The genetic code is not universal

The link between the proteome and the biochemistry of the cell

SUMMARY

SHORT ANSWER QUESTIONS

IN-DEPTH PROBLEMS

FURTHER READING

CHAPTER 2 STUDYING DNA

2.1 ENZYMES FOR DNA MANIPULATION

The mode of action of a template-dependent DNA polymerase

The types of DNA polymerase used in research

Restriction endonucleases enable DNA molecules to be cut at defined positions

Gel electrophoresis is used to examine the results of a restriction digest

Interesting DNA fragments can be identified by Southern hybridization

Ligases join DNA fragments together

End-modification enzymes

2.2 THE POLYMERASE CHAIN REACTION

Carrying out a PCR

The rate of product formation can be followed during a PCR

PCR has many and diverse applications

2.3 DNA CLONING

Why is gene cloning important?

The simplest cloning vectors are based on *E. coli* plasmids

Bacteriophages can also be used as cloning vectors

Vectors for longer pieces of DNA

DNA can be cloned in organisms other than *E. coli*

SUMMARY

SHORT ANSWER QUESTIONS

IN-DEPTH PROBLEMS

FURTHER READING

CHAPTER 3 MAPPING GENOMES

3.1 WHY A GENOME MAP IS IMPORTANT

Genome maps are needed in order to sequence the more complex genomes

Genome maps are not just sequencing aids

27 3.2 MARKERS FOR GENETIC MAPPING

Genes were the first markers to be used

RFLPs and SSLPs are examples of DNA markers

Single-nucleotide polymorphisms are the most useful type of DNA marker

3.3 THE BASIS TO GENETIC MAPPING

The principles of inheritance and the discovery of linkage

Partial linkage is explained by the behavior of chromosomes during meiosis

From partial linkage to genetic mapping

37

38

38

39

40

41

41

41

43

44

44

47

48

48

50

51

51

51

52

55

55

57

58

58

59

61

63

63

65

65

57

58

58

59

61

63

63

65

65

68

68

68

68

68

68

3.4 LINKAGE ANALYSIS WITH DIFFERENT TYPES OF ORGANISMS

Linkage analysis when planned breeding experiments are possible	69
Gene mapping by human pedigree analysis	71
Genetic mapping in bacteria	73
The limitations of linkage analysis	74

3.5 PHYSICAL MAPPING BY DIRECT EXAMINATION OF DNA MOLECULES

Conventional restriction mapping is applicable only to small DNA molecules	75
Optical mapping can locate restriction sites in longer DNA molecules	77
Optical mapping can be used to map other features in a DNA molecule	79

3.6 PHYSICAL MAPPING BY ASSIGNING MARKERS TO DNA FRAGMENTS

Any unique sequence can be used as an STS	81
DNA fragments for STS mapping can be obtained as radiation hybrids	82
A clone library can be used as the mapping reagent	83

SUMMARY	84
----------------	----

SHORT ANSWER QUESTIONS	85
-------------------------------	----

IN-DEPTH PROBLEMS	85
--------------------------	----

FURTHER READING	86
------------------------	----

CHAPTER 4 SEQUENCING GENOMES 87**4.1 CHAIN-TERMINATION SEQUENCING**

Chain-termination sequencing in outline	87
Not all DNA polymerases can be used for sequencing	89
Chain-termination sequencing with <i>Taq</i> polymerase	90
Strengths and limitations of chain-termination sequencing	91

4.2 NEXT-GENERATION SEQUENCING

Preparation of a sequencing library is the common feature of next-generation methods	92
Various next-generation sequencing methods have been devised	95
Third- and fourth-generation methods enable sequencing in real time	97

4.3 HOW TO SEQUENCE A GENOME

The potential of the shotgun method was proven by the <i>Haemophilus influenzae</i> sequence	98
Many prokaryotic genomes have been sequenced by the shotgun method	99
	100

Shotgun sequencing of eukaryotic genomes requires sophisticated assembly programs	102
More complex genomes can be sequenced by a hierarchical shotgun approach	104
What is a genome sequence and do we always need one?	107

4.4 A SURVEY OF EUKARYOTIC GENOME SEQUENCING PROJECTS

The Human Genome Project: genome sequencing in the heroic age	109
The Neanderthal genome: assembly of an extinct genome by use of the human sequence as a reference	110
The giant panda genome: shotgun sequencing based entirely on next-generation data	111
The barley genome: the concept of gene space	113

SUMMARY	115
----------------	-----

SHORT ANSWER QUESTIONS	115
-------------------------------	-----

IN-DEPTH PROBLEMS	116
--------------------------	-----

FURTHER READING	117
------------------------	-----

CHAPTER 5 GENOME ANNOTATION 119**5.1 GENOME ANNOTATION BY COMPUTER ANALYSIS OF THE DNA SEQUENCE**

The coding regions of genes are open reading frames	119
Simple ORF scans are less effective with genomes of higher eukaryotes	120
Locating genes for noncoding RNA	122
Homology searches and comparative genomics give an extra dimension to gene prediction	123

5.2 GENOME ANNOTATION BY ANALYSIS OF GENE TRANSCRIPTS

Hybridization tests can determine if a fragment contains transcribed sequences	125
Methods are available for precise mapping of the ends of transcripts	126
Exon–intron boundaries can also be located with precision	126

5.3 ANNOTATION BY GENOMEWIDE RNA MAPPING

Tiling arrays enable transcripts to be mapped onto chromosomes or entire genomes	128
Transcript sequences can be directly mapped onto a genome	129

5.4 GENOME BROWSERS	131
----------------------------	-----

SUMMARY	132	CHAPTER 7	
SHORT ANSWER QUESTIONS	132	EUKARYOTIC NUCLEAR	
IN-DEPTH PROBLEMS	133	GENOMES	155
FURTHER READING	133	7.1 NUCLEAR GENOMES ARE CONTAINED	
		IN CHROMOSOMES	155
CHAPTER 6		Chromosomes are much shorter than the DNA	155
IDENTIFYING GENE FUNCTIONS	135	molecules they contain	157
6.1 COMPUTER ANALYSIS OF GENE		Special features of metaphase chromosomes	159
FUNCTION	135	DNA–protein interactions in centromeres	159
Homology reflects evolutionary relationships	135	and telomeres	
Homology analysis can provide information on	136	7.2 HOW ARE THE GENES ARRANGED IN A	
the function of a gene	136	NUCLEAR GENOME?	161
Identification of protein domains can help to	137	Genes are not evenly distributed within a genome	161
assign function to an unknown gene	137	A segment of the human genome	162
Annotation of gene function requires a common	138	The yeast genome is very compact	164
terminology	138	Gene organization in other eukaryotes	165
6.2 ASSIGNING FUNCTION BY		7.3 HOW MANY GENES ARE THERE AND	
GENE INACTIVATION AND		WHAT ARE THEIR FUNCTIONS?	167
OVEREXPRESSION	139	Gene numbers can be misleading	168
Functional analysis by gene inactivation	140	Gene catalogs reveal the distinctive features of	169
Individual genes can be inactivated by	140	different organisms	172
homologous recombination	140	Families of genes	174
Gene inactivation without homologous	142	Pseudogenes and other evolutionary relics	174
recombination	142	7.4 THE REPETITIVE DNA CONTENT OF	
Gene overexpression can also be used to	144	EUKARYOTIC NUCLEAR GENOMES	176
assess function	144	Tandemly repeated DNA is found at centromeres	176
The phenotypic effect of gene inactivation or	145	and elsewhere in eukaryotic chromosomes	176
overexpression may be difficult to discern	145	Minisatellites and microsatellites	177
6.3 UNDERSTANDING GENE FUNCTION		Interspersed repeats	177
BY STUDIES OF EXPRESSION PATTERN		SUMMARY	178
AND PROTEIN PRODUCT	146	SHORT ANSWER QUESTIONS	178
Reporter genes and immunocytochemistry	146	IN-DEPTH PROBLEMS	179
can be used to locate where and when genes	146	FURTHER READING	179
are expressed	146		
Directed mutagenesis can be used to probe gene	147		
function in detail	147		
6.4 USING CONVENTIONAL GENETIC		CHAPTER 8	
ANALYSIS TO IDENTIFY GENE		GENOMES OF PROKARYOTES	
FUNCTION	149	AND EUKARYOTIC ORGANELLES	181
Identification of human genes responsible for	150	8.1 PHYSICAL FEATURES OF PROKARYOTIC	
inherited diseases	150	GENOMES	181
Genomewide association studies can also	151	The traditional view of the prokaryotic	181
identify genes for diseases and other traits	151	chromosome	183
SUMMARY	152	Some bacteria have linear or multipartite genomes	
SHORT ANSWER QUESTIONS	153	8.2 GENETIC FEATURES OF PROKARYOTIC	
IN-DEPTH PROBLEMS	153	GENOMES	186
FURTHER READING	154	Gene organization in the <i>E. coli</i> K12 genome	186

Operons are characteristic features of prokaryotic genomes	188		
Prokaryotic genome sizes and numbers of genes vary according to biological complexity	189		
Genome sizes and numbers of genes vary within individual species	190		
Distinctions between prokaryotic species are further blurred by lateral gene transfer	192		
Metagenomes describe the members of a community	194		
8.3 EUKARYOTIC ORGANELLAR GENOMES	195		
The endosymbiont theory explains the origin of organellar genomes	195		
Most organellar genomes are circular	196		
The gene catalogs of organellar genomes	197		
SUMMARY	198		
SHORT ANSWER QUESTIONS	200		
IN-DEPTH PROBLEMS	201		
FURTHER READING	201		
CHAPTER 9			
VIRAL GENOMES AND MOBILE GENETIC ELEMENTS	203		
9.1 THE GENOMES OF BACTERIOPHAGES AND EUKARYOTIC VIRUSES	203		
Bacteriophage genomes have diverse structures and organizations	203		
Replication strategies for bacteriophage genomes	205		
Structures and replication strategies for eukaryotic viral genomes	206		
Some retroviruses cause cancer	207		
Genomes at the edge of life	209		
9.2 MOBILE GENETIC ELEMENTS	210		
RNA transposons with long terminal repeats are related to viral retroelements	210		
Some RNA transposons lack long terminal repeats	212		
DNA transposons are common in prokaryotic genomes	213		
DNA transposons are less common in eukaryotic genomes	214		
SUMMARY	216		
SHORT ANSWER QUESTIONS	216		
IN-DEPTH PROBLEMS	217		
FURTHER READING	217		
CHAPTER 10			
ACCESSING THE GENOME			219
10.1 INSIDE THE NUCLEUS			219
The nucleus has an ordered internal structure			220
The DNA content of a nondividing nucleus displays different degrees of packaging			221
The nuclear matrix is thought to provide attachment points for chromosomal DNA			222
Each chromosome has its own territory within the nucleus			223
Each chromosome comprises a series of topologically associated domains			224
Insulators mark the boundaries of topologically associated domains			226
10.2 NUCLEOSOME MODIFICATIONS AND GENOME EXPRESSION			228
Acetylation of histones influences many nuclear activities including genome expression			228
Histone deacetylation represses active regions of the genome			229
Acetylation is not the only type of histone modification			230
Nucleosome repositioning also influences gene expression			231
10.3 DNA MODIFICATION AND GENOME EXPRESSION			234
Genome silencing by DNA methylation			234
Methylation is involved in genomic imprinting and X inactivation			235
SUMMARY			236
SHORT ANSWER QUESTIONS			237
IN-DEPTH PROBLEMS			238
FURTHER READING			238
CHAPTER 11			
THE ROLE OF DNA-BINDING PROTEINS IN GENOME EXPRESSION			241
11.1 METHODS FOR STUDYING DNA-BINDING PROTEINS AND THEIR ATTACHMENT SITES			241
X-ray crystallography provides structural data for any protein that can be crystallized			241
NMR spectroscopy is used to study the structures of small proteins			243
Gel retardation identifies DNA fragments that bind to proteins			244

Protection assays pinpoint binding sites with greater accuracy	244	RNA silencing was first identified as a means of destroying invading viral RNA	276
Modification interference identifies nucleotides central to protein binding	246	MicroRNAs regulate genome expression by causing specific target mRNAs to be degraded	278
Genomewide scans for protein attachment sites	247		
11.2 THE SPECIAL FEATURES OF DNA-BINDING PROTEINS	249	12.4 INFLUENCE OF RNA PROCESSING ON THE COMPOSITION OF A TRANSCRIPTOME	278
The helix–turn–helix motif is present in prokaryotic and eukaryotic proteins	249	The splicing pathway for eukaryotic pre-mRNA introns	279
Zinc fingers are common in eukaryotic proteins	250	The splicing process must have a high degree of precision	280
Other nucleic acid-binding motifs	251	Enhancer and silencer elements specify alternative splicing pathways	282
11.3 INTERACTION BETWEEN DNA AND ITS BINDING PROTEINS	252	12.5 TRANSCRIPTOMES IN RESEARCH	284
Direct readout of the nucleotide sequence	252	Transcriptome analysis as an aid to genome annotation	284
The nucleotide sequence has a number of indirect effects on helix structure	253	Cancer transcriptomes	286
Contacts between DNA and proteins	253	Transcriptomes and the responses of plants to stress	287
SUMMARY	254	SUMMARY	289
SHORT ANSWER QUESTIONS	255	SHORT ANSWER QUESTIONS	289
IN-DEPTH PROBLEMS	256	IN-DEPTH PROBLEMS	290
FURTHER READING	256	FURTHER READING	290
CHAPTER 12	257	CHAPTER 13	293
TRANSCRIPTOMES		PROTEOMES	
12.1 COMPONENTS OF THE TRANSCRIPTOME	257	13.1 STUDYING THE COMPOSITION OF A PROTEOME	293
The mRNA fraction of a transcriptome is small but complex	257	The separation stage of a protein profiling project	294
Short noncoding RNAs have diverse functions	259	The identification stage of a protein profiling project	297
Long noncoding RNAs are enigmatic transcripts	260	Comparing the compositions of two proteomes	299
Microarray analysis and RNA sequencing are used to study the contents of transcriptomes	262	Analytical protein arrays offer an alternative approach to protein profiling	300
12.2 SYNTHESIS OF THE COMPONENTS OF THE TRANSCRIPTOME	263	13.2 IDENTIFYING PROTEINS THAT INTERACT WITH ONE ANOTHER	301
RNA polymerases are molecular machines for making RNA	264	Identifying pairs of interacting proteins	301
Transcription start points are indicated by promoter sequences	266	Identifying the components of multiprotein complexes	304
Synthesis of bacterial RNA is regulated by repressor and activator proteins	268	Identifying proteins with functional interactions	305
Synthesis of bacterial RNA is also regulated by control over transcription termination	271	Protein interaction maps display the interactions within a proteome	306
Synthesis of eukaryotic RNA is regulated primarily by activator proteins	272	13.3 SYNTHESIS AND DEGRADATION OF THE COMPONENTS OF THE PROTEOME	308
12.3 DEGRADATION OF THE COMPONENTS OF THE TRANSCRIPTOME	275	Ribosomes are molecular machines for making proteins	308
Several processes are known for nonspecific RNA turnover	275		

During stress, bacteria inactivate their ribosomes in order to downsize the proteome	311	Yeast mating types are determined by gene conversion events	338
Initiation factors mediate large-scale remodeling of eukaryotic proteomes	312	Genome rearrangements are responsible for immunoglobulin and T-cell receptor diversity	339
The translation of individual mRNAs can also be regulated	313		
Degradation of the components of the proteome	314		
13.4 INFLUENCE OF PROTEIN PROCESSING ON THE COMPOSITION OF THE PROTEOME		14.3 CHANGES IN GENOME ACTIVITY UNDERLYING DEVELOPMENT	341
The amino acid sequence contains instructions for protein folding	315	Bacteriophage λ : a genetic switch enables a choice to be made between alternative developmental pathways	342
Some proteins are activated by proteolytic cleavage	318	<i>Bacillus</i> sporulation: coordination of activities in two distinct cell types	343
Important changes in protein activity can be brought about by chemical modification	320	<i>Caenorhabditis elegans</i> : the genetic basis of positional information and the determination of cell fate	346
		Fruit flies: conversion of positional information into a segmented body plan	348
13.5 BEYOND THE PROTEOME	322	Homeotic selector genes are universal features of higher eukaryotic development	350
The metabolome is the complete set of metabolites present in a cell	322	Homeotic genes also underlie plant development	352
Systems biology provides an integrated description of cellular activity	323	SUMMARY	352
SUMMARY	326	SHORT ANSWER QUESTIONS	353
SHORT ANSWER QUESTIONS	326	IN-DEPTH PROBLEMS	354
IN-DEPTH PROBLEMS	327	FURTHER READING	354
FURTHER READING	327		
		CHAPTER 15	
CHAPTER 14		GENOME REPLICATION	357
GENOME EXPRESSION IN THE CONTEXT OF CELL AND ORGANISM	329		
14.1 THE RESPONSE OF THE GENOME TO EXTERNAL SIGNALS	330	15.1 THE TOPOLOGY OF GENOME REPLICATION	357
Signal transmission by import of the extracellular signaling compound	330	The double-helical structure complicates the replication process	358
Receptor proteins transmit signals across cell membranes	332	The Meselson–Stahl experiment proved that replication is semiconservative	359
Some signal transduction pathways have few steps between receptor and genome	333	DNA topoisomerases provide a solution to the topological problem	361
Some signal transduction pathways have many steps between receptor and genome	334	Variations on the semiconservative theme	363
Some signal transduction pathways operate via second messengers	336	15.2 THE INITIATION PHASE OF GENOME REPLICATION	364
		Initiation at the <i>E. coli</i> origin of replication	364
14.2 CHANGES IN GENOME ACTIVITY RESULTING IN CELLULAR DIFFERENTIATION	336	Origins of replication have been clearly defined in yeast	365
Some differentiation processes involve changes to chromatin structure	336	Origins in higher eukaryotes have been less easy to identify	366
		15.3 EVENTS AT THE REPLICATION FORK	367
		DNA polymerases are molecular machines for making (and degrading) DNA	367
		DNA polymerases have limitations that complicate genome replication	369

Okazaki fragments must be joined together to complete lagging-strand replication	370	Defects in DNA repair underlie human diseases, including cancers	406
15.4 TERMINATION OF GENOME REPLICATION	372	SUMMARY	406
Replication of the <i>E. coli</i> genome terminates within a defined region	373	SHORT ANSWER QUESTIONS	407
Little is known about termination of replication in eukaryotes	374	IN-DEPTH PROBLEMS	407
Telomerase completes replication of chromosomal DNA molecules, at least in some cells	375	FURTHER READING	408
Telomere length is implicated in cell senescence and cancer	378	CHAPTER 17	
<i>Drosophila</i> has a unique solution to the end-shortening problem	379	RECOMBINATION AND TRANSPOSITION	411
15.5 REGULATION OF EUKARYOTIC GENOME REPLICATION	380	17.1 HOMOLOGOUS RECOMBINATION	412
Genome replication must be synchronized with the cell cycle	380	The Holliday and Meselson–Radding models for homologous recombination	412
Origin licensing is the prerequisite for passing the G1–S checkpoint	380	The double-strand break model for homologous recombination	414
Replication origins do not all fire at the same time	382	RecBCD is the most important pathway for homologous recombination in bacteria	415
The cell has various options if the genome is damaged	383	<i>E. coli</i> can also carry out homologous recombination by the RecFOR pathway	417
SUMMARY	384	Homologous recombination pathways in eukaryotes	417
SHORT ANSWER QUESTIONS	385	The primary role of homologous recombination is thought to be DNA repair	418
IN-DEPTH PROBLEMS	385	17.2 SITE-SPECIFIC RECOMBINATION	419
FURTHER READING	386	Bacteriophage λ uses site-specific recombination during the lysogenic infection cycle	419
CHAPTER 16		Site-specific recombination is an aid in construction of genetically modified plants	421
MUTATIONS AND DNA REPAIR	389	17.3 TRANSPOSITION	421
16.1 THE CAUSES OF MUTATIONS	389	Replicative and conservative transposition of DNA transposons	422
Errors in replication are a source of point mutations	390	Retroelements transpose replicatively via an RNA intermediate	423
Replication errors can also lead to insertion and deletion mutations	391	SUMMARY	425
Mutations are also caused by chemical and physical mutagens	394	SHORT ANSWER QUESTIONS	426
16.2 REPAIR OF MUTATIONS AND OTHER TYPES OF DNA DAMAGE	398	IN-DEPTH PROBLEMS	427
Direct repair systems fill in nicks and correct some types of nucleotide modification	398	FURTHER READING	427
Base excision repairs many types of damaged nucleotide	399	CHAPTER 18	
Nucleotide excision repair is used to correct more extensive types of damage	401	HOW GENOMES EVOLVE	429
Mismatch repair corrects replication errors	402	18.1 GENOMES: THE FIRST 10 BILLION YEARS	429
Single- and double-strand breaks can be repaired	403	The first biochemical systems were centered on RNA	429
If necessary, DNA damage can be bypassed during genome replication	405	The first DNA genomes	432
		How unique is life?	433

18.2 EVOLUTION OF INCREASINGLY COMPLEX GENOMES	434
Genome sequences provide extensive evidence of past gene duplications	434
A variety of processes could result in gene duplication	438
Whole-genome duplication is also possible	439
Smaller duplications can also be identified in the human genome and other genomes	442
Both prokaryotes and eukaryotes acquire genes from other species	444
Genome evolution also involves rearrangement of existing genes	445
There are competing hypotheses for the origins of introns	448
The evolution of the epigenome	449
18.3 GENOMES: THE LAST 6 MILLION YEARS	450
The human genome is very similar to that of the chimpanzee	451
Paleogenomics is helping us understand the recent evolution of the human genome	452
18.4 GENOMES TODAY: DIVERSITY IN POPULATIONS	453
The origins of HIV/AIDS	454
The first migrations of humans out of Africa	455
The diversity of plant genomes is an aid in crop breeding	457
SUMMARY	458
SHORT ANSWER QUESTIONS	459
IN-DEPTH PROBLEMS	460
FURTHER READING	460
GLOSSARY	463
INDEX	491