

SECOND EDITION

GENETICS AND GENOMICS IN MEDICINE

TOM STRACHAN AND
ANNEKE LUCASSEN



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Contents

<i>Preface</i>	<i>xv</i>
<i>Acknowledgements</i>	<i>xviii</i>

1 FUNDAMENTALS OF DNA, CHROMOSOMES, AND CELLS..... 1

1.1 THE STRUCTURE AND FUNCTION OF NUCLEIC ACIDS.....2

General concepts: the genetic material, genomes, and genes	2
The underlying chemistry of nucleic acids	2
Base pairing and the double helix	4
DNA replication and DNA polymerases.....	5
Genes, transcription, and the central dogma of molecular biology.....	7

1.2 THE STRUCTURE AND FUNCTION OF CHROMOSOMES.....8

Why we need highly structured chromosomes, and how they are organized	8
Chromosome function: replication origins, centromeres, and telomeres.....	9

1.3 DNA AND CHROMOSOMES IN CELL DIVISION AND THE CELL CYCLE 10

Differences in DNA copy number between cells.....	10
The cell cycle and segregation of replicated chromosomes and DNA molecules.....	11

Mitosis: the usual form of cell division.....	13
Meiosis: a specialized reductive cell division giving rise to sperm and egg cells	13
Why each of our gametes is unique.....	16

SUMMARY	18
QUESTIONS	19
FURTHER READING	19

2 FUNDAMENTALS OF GENE STRUCTURE, GENE EXPRESSION, AND HUMAN GENOME ORGANIZATION.....21

2.1 PROTEIN-CODING GENES: STRUCTURE AND EXPRESSION22

Gene organization: exons and introns	22
RNA splicing: stitching together the genetic information in exons	23
Translation: decoding messenger RNA to make a polypeptide.....	24
From newly synthesized polypeptide to mature protein	29

2.2 RNA GENES AND NONCODING RNA.....32

The extraordinary secondary structure and versatility of RNA.....	33
RNAs that act as specific regulators: from quirky exceptions to the mainstream	34

2.3 WORKING OUT THE DETAILS OF OUR GENOME AND WHAT THEY MEAN	35	3 PRINCIPLES UNDERLYING CORE DNA TECHNOLOGIES.....	57
The Human Genome Project: working out the details of the nuclear genome.....	35	3.1 AMPLIFYING DNA BY DNA CLONING	58
What the sequence didn't tell us and the goal of identifying all functional human DNA sequences.....	37	Amplifying desired DNA within bacterial cells.....	59
2.4 A QUICK TOUR OF SOME ELECTRONIC RESOURCES USED TO INTERROGATE THE HUMAN GENOME SEQUENCE AND GENE PRODUCTS.....	39	The need for vector DNA molecules.....	59
Gene nomenclature and the HGNC gateway.....	40	Physical clone separation	60
Databases storing nucleotide and protein sequences	40	The need for restriction nucleases.....	60
Finding related nucleotide and protein sequences	40	DNA libraries and the uses and limitations of DNA cloning.....	61
Links to clinical databases	42	3.2 AMPLIFYING DNA USING THE POLYMERASE CHAIN REACTION (PCR).....	62
2.5 THE ORGANIZATION AND EVOLUTION OF THE HUMAN GENOME.....	42	Basics of the polymerase chain reaction (PCR)	62
A brief overview of the evolutionary mechanisms that shaped our genome.....	42	Quantitative PCR and real-time PCR.....	63
How much of our genome is functionally significant?	43	3.3 PRINCIPLES OF NUCLEIC ACID HYBRIDIZATION.....	63
The mitochondrial genome: economical usage but limited autonomy.....	44	Formation of artificial heteroduplexes.....	66
Gene distribution in the human genome.....	45	Hybridization assays: using known nucleic acids to find related sequences in a test nucleic acid population	66
The extent of repetitive DNA in the human genome	46	Microarray hybridization: large- scale parallel hybridization to immobilized probes.....	70
The organization of gene families.....	47	3.4 PRINCIPLES OF DNA SEQUENCING.....	71
The significance of gene duplication and repetitive coding DNA.....	50	Dideoxy DNA sequencing	72
Highly repetitive noncoding DNA in the human genome	51	Massively parallel DNA sequencing (next-generation sequencing)	74
SUMMARY.....	53	SUMMARY.....	75
QUESTIONS	54	QUESTIONS	76
FURTHER READING.....	54	FURTHER READING.....	76
		4 PRINCIPLES OF GENETIC VARIATION	77
		4.1 DNA SEQUENCE VARIATION ORIGINS AND DNA REPAIR	79
		Genetic variation arising from errors in chromosome and DNA function....	79

Various endogenous and exogenous sources can cause damage to DNA by altering its chemical structure	81	Programmed and random post-zygotic genetic variation.....	100
The wide range of DNA repair mechanisms.....	82	Somatic mechanisms allow cell-specific production of immunoglobulins and T-cell receptors.....	100
Repair of DNA damage or altered sequence on a single DNA strand.....	82	MHC (HLA) proteins: functions and polymorphism.....	102
Repair of DNA lesions that affect both DNA strands.....	83	The medical importance of the HLA system.....	104
Undetected DNA damage, DNA damage tolerance, and translesion synthesis	84	SUMMARY.....	107
4.2 POPULATION GENOMICS AND THE SCALE OF HUMAN GENETIC VARIATION	87	QUESTIONS	108
DNA variants, polymorphisms, and human population genomics.....	87	FURTHER READING.....	108
Small-scale variation: single nucleotide variants and small insertions and deletions	89	5 SINGLE-GENE DISORDERS: INHERITANCE PATTERNS, PHENOTYPE VARIABILITY, AND ALLELE FREQUENCIES	109
Microsatellites and other variable number of tandem repeat (VNTR) polymorphisms.....	90	5.1 INTRODUCTION: TERMINOLOGY, ELECTRONIC RESOURCES, AND PEDIGREES	110
Structural variation and low copy number variation.....	91	Background terminology and electronic resources with information on single-gene disorders	110
Taking stock of human genetic variation.....	92	Investigating family history of disease and recording pedigrees.....	111
4.3 FUNCTIONAL GENETIC VARIATION AND PROTEIN POLYMORPHISM.....	93	5.2 THE BASICS OF MENDELIAN AND MITOCHONDRIAL DNA INHERITANCE PATTERNS	112
The vast majority of genetic variation has a neutral effect on the phenotype, but a small fraction is harmful	93	Autosomal dominant inheritance.....	112
Different types of Darwinian natural selection operate in human lineages.....	94	Autosomal recessive inheritance.....	113
Generating protein diversity by gene duplication: the example of olfactory receptor genes.....	98	Sex-linked inheritance	116
4.4 EXTRAORDINARY GENETIC VARIATION IN THE IMMUNE SYSTEM.....	99	Matrilineal inheritance for mitochondrial DNA disorders.....	121
Pronounced genetic variation in four classes of immune system proteins....	99	5.3 UNCERTAINTY, HETEROGENEITY, AND VARIABLE EXPRESSION OF MENDELIAN PHENOTYPES.....	122
		Difficulties in defining the mode of inheritance in small pedigrees.....	122
		Heterogeneity in the correspondence between phenotypes and the underlying genes and mutations	124

Nonpenetrance and age-related penetrance	126
--	-----

5.4 ALLELE FREQUENCIES IN POPULATIONS..... 129

Allele frequencies and the Hardy-Weinberg law	130
Applications and limitations of the Hardy-Weinberg law	131
Ways in which allele frequencies change in populations	132
Population bottlenecks and founder effects.....	133
Mutation versus selection in determining allele frequencies.....	135
Heterozygote advantage: when natural selection favors carriers of recessive disease	136

SUMMARY..... 137

QUESTIONS 138

FURTHER READING..... 138

6 PRINCIPLES OF GENE REGULATION AND EPIGENETICS 139

The two fundamental types of gene regulation.....	139
<i>Cis</i> -acting and <i>trans</i> -acting effects in gene regulation	140

6.1 GENETIC REGULATION OF GENE EXPRESSION 141

Promoters: the major on–off switches in genes.....	141
Modulating transcription and tissue-specific regulation	142
Transcription factor binding and specificity	143
Genetic regulation during RNA processing: RNA splicing and RNA editing	144
Translational regulation by <i>trans</i> -acting regulatory proteins.....	147
Post-transcriptional gene silencing by microRNAs.....	148
Repressing the repressors: competing endogenous RNAs sequester miRNA	148

6.2 CHROMATIN MODIFICATION AND EPIGENETIC FACTORS IN GENE REGULATION 150

An overview of the molecular basis of epigenetic mechanisms.....	150
How changes in chromatin structure produce altered gene expression.....	151
Histone modification and histone substitution in nucleosomes	152
Modified histones and histone variants affect chromatin structure	154
The function of DNA methylation in mammalian cells	155
DNA methylation: mechanisms, heritability, and global roles during early development and gametogenesis	156
Long noncoding RNAs in mammalian epigenetic regulation.....	158
Genomic imprinting: differential expression of maternally and paternally inherited alleles.....	160
X-chromosome inactivation: compensating for sex differences in gene dosage	163

6.3 ABNORMAL EPIGENETIC REGULATION IN MENDELIAN DISORDERS AND UNIPARENTAL DISOMY 165

Principles of epigenetic dysregulation	165
“Chromatin diseases” due to mutations in genes specifying chromatin modifiers	167
Disease resulting from dysregulation of heterochromatin	168
Uniparental disomy and disorders of imprinting	171
Abnormal gene regulation at imprinted loci.....	172

SUMMARY..... 176

QUESTIONS 176

FURTHER READING..... 177

7 HOW GENETIC VARIATION IN DNA AND CHROMOSOMES CAUSES DISEASE 179

7.1 AN OVERVIEW OF HOW GENETIC VARIATION RESULTS IN DISEASE..... 180

The importance of repeat sequences in triggering pathogenesis..... 182

7.2 PATHOGENIC NUCLEOTIDE SUBSTITUTIONS AND TINY INSERTIONS AND DELETIONS183

Pathogenic single nucleotide substitutions within coding sequences..... 183

Mutations that result in premature termination codons..... 185

Genesis and frequency of pathogenic point mutations..... 188

Surveying and curating point mutations that cause disease 191

7.3 PATHOGENESIS DUE TO VARIATION IN SHORT TANDEM REPEAT COPY NUMBER.....192

The two main classes of pathogenic variation in short tandem repeat copy-number..... 192

Dynamic disease-causing mutations due to unstable expansion of short tandem repeats 194

Unstable expansion of short tandem repeats can cause disease in different ways 197

7.4 PATHOGENESIS TRIGGERED BY LONG TANDEM REPEATS AND INTERSPERSED REPEATS198

Pathogenic exchanges between repeats occurs in both nuclear DNA and mtDNA 198

Nonallelic homologous recombination and transposition 199

Pathogenic sequence exchanges between chromatids at mispaired tandem repeats 199

Disease arising from sequence exchanges between distantly located repeats in nuclear DNA202

7.5 CHROMOSOME ABNORMALITIES 204

Structural chromosomal abnormalities206

Chromosomal abnormalities involving gain or loss of complete chromosomes209

7.6 MOLECULAR PATHOLOGY OF MITOCHONDRIAL DISORDERS212

Mitochondrial disorders due to mtDNA mutation show maternal inheritance and variable proportions of mutant genotypes..... 213

The two major classes of pathogenic DNA variant in mtDNA: large deletions and point mutations..... 215

7.7 EFFECTS ON THE PHENOTYPE OF PATHOGENIC VARIANTS IN NUCLEAR DNA217

Mutations affecting how a single gene works: an overview of loss of function and gain of function..... 218

The effect of pathogenic variants depends on how the products of alleles interact: dominance and recessiveness revisited220

Gain-of-function and loss-of-function mutations in the same gene can produce different phenotypes223

Multiple gene dysregulation resulting from aneuploidies and mutations in regulatory genes224

7.8 A PROTEIN STRUCTURE PERSPECTIVE OF MOLECULAR PATHOLOGY 225

Pathogenesis arising from protein misfolding.....226

The many different ways in which protein aggregation can result in disease.....	226
---	-----

7.9 GENOTYPE–PHENOTYPE CORRELATIONS AND WHY MONOGENIC DISORDERS ARE OFTEN NOT SIMPLE231

The difficulty in getting reliable genotype–phenotype correlations.....	231
Modifier genes and environmental factors: common explanations for poor genotype–phenotype correlations.....	232

SUMMARY.....	236
QUESTIONS	237
FURTHER READING.....	237

8 IDENTIFYING DISEASE GENES AND GENETIC SUSCEPTIBILITY TO COMPLEX DISEASE..... 239

8.1 IDENTIFYING GENES IN MONOGENIC DISORDERS 240

A historical overview of identifying genes in monogenic disorders	240
Linkage analysis to map genes for monogenic disorders to defined subchromosomal regions.....	241
Chromosome abnormalities and other large-scale mutations as routes to identifying disease genes..	248
Exome sequencing: let’s not bother getting a position for disease genes!.....	248

8.2 APPROACHES TO MAPPING AND IDENTIFYING GENETIC SUSCEPTIBILITY TO COMPLEX DISEASE.....251

The polygenic and multifactorial nature of common genetic disorders	252
Difficulties with lack of penetrance and phenotype classification in complex disease.....	255

Estimating heritability: the contribution made by genetic factors to the variance of complex diseases	256
---	-----

The very limited success of linkage analyses in identifying genes underlying complex genetic diseases	259
---	-----

The fundamentals of allelic association and the importance of HLA-disease associations	262
--	-----

Linkage disequilibrium as the basis of allelic associations.....	266
--	-----

How genomewide association studies are carried out.....	270
---	-----

Moving from candidate subchromosomal region to identify causal genetic variants in complex disease can be challenging	273
---	-----

The limitations of GWA studies and the issue of missing heritability.....	274
---	-----

Alternative genome-wide studies and the role of rare variants and copy number variants in complex disease.....	276
--	-----

The assessment and prediction of risk for common genetic diseases and the development of polygenic risk scores.....	278
---	-----

8.3 ASPECTS OF THE GENETIC ARCHITECTURE OF COMPLEX DISEASE AND THE CONTRIBUTIONS OF ENVIRONMENTAL AND EPIGENETIC FACTORS 280

Common neurodegenerative disease: from monogenic to polygenic disease	283
---	-----

The importance of immune system pathways in common genetic disease.....	287
---	-----

The importance of protective factors and how a susceptibility factor for one complex disease may be a protective factor for another disease.....	289
--	-----

Gene–environment interactions in complex disease.....	290	Translating genomic advances and developing generic drugs as a way of overcoming the problem of too few drug targets.....	325
Epigenetics in complex disease and aging: significance and experimental approaches.....	294	Developing biological drugs: therapeutic proteins produced by genetic engineering	325
SUMMARY.....	297	Genetically engineered therapeutic antibodies with improved therapeutic potential.....	326
QUESTIONS	298		
FURTHER READING.....	298		
9 GENETIC APPROACHES TO TREATING DISEASE.....	301	9.3 PRINCIPLES OF GENE AND CELL THERAPY	329
9.1 AN OVERVIEW OF TREATING GENETIC DISEASE AND OF GENETIC TREATMENT OF DISEASE.....	303	Two broad strategies in somatic gene therapy.....	329
Three different broad approaches to treating genetic disorders.....	303	The delivery problem: designing optimal and safe strategies for getting genetic constructs into the cells of patients	330
Very different treatment options for different inborn errors of metabolism	305	Different ways of delivering therapeutic genetic constructs, and the advantages of <i>ex vivo</i> gene therapy.....	334
Genetic treatment of disease may be conducted at many different levels	309	Viral delivery of therapeutic gene constructs: relatively high efficiency but safety concerns.....	336
9.2 GENETIC INPUTS INTO TREATING DISEASE WITH SMALL MOLECULE DRUGS AND THERAPEUTIC PROTEINS	310	Virus vectors used in gene therapy	336
An overview of how genetic differences affect the metabolism and performance of small molecule drugs.....	311	The importance of disease models for testing potential therapies in humans.....	337
Phenotype differences arising from genetic variation in drug metabolism.....	313	9.4 GENE THERAPY FOR INHERITED DISORDERS: PRACTICE AND FUTURE DIRECTIONS.....	340
Genetic variation in enzymes that work in phase II drug metabolism.....	317	Multiple successes for <i>ex vivo</i> gene supplementation therapy targeted at hematopoietic stem cells	340
Altered drug responses resulting from genetic variation in drug targets.....	318	<i>In vivo</i> gene therapy: approaches, barriers, and recent successes.....	342
When genotypes at multiple loci in patients are important in drug treatment: the example of warfarin	321	An overview of RNA and oligonucleotide therapeutics.....	344
Translating genetic advances: from identifying novel disease genes to therapeutic small molecule drugs.....	322	RNA interference therapy	347
		Future therapeutic prospects using CRISPR-Cas gene editing.....	349
		Therapeutic applications of stem cells and cell reprogramming.....	353
		Obstacles to overcome in cell therapy.....	353

A special case: preventing transmission of severe mitochondrial DNA disorders by mitochondrial replacement..... 355

SUMMARY..... 356

QUESTIONS 358

FURTHER READING..... 358

10 CANCER GENETICS AND GENOMICS..... 361

10.1 FUNDAMENTAL CHARACTERISTICS AND EVOLUTION OF CANCER..... 362

The defining features of unregulated cell growth and cancer..... 362

Why cancers are different from other diseases: the contest between natural selection operating at the level of the cell and the level of the organism..... 364

Cancer cells acquire several distinguishing biological characteristics during their evolution..... 366

The initiation and multistage nature of cancer evolution and why most human cancers develop over many decades..... 369

Intratumor heterogeneity arises through cell infiltration, clonal evolution, and differentiation of cancer stem cells..... 372

10.2 ONCOGENES AND TUMOR SUPPRESSOR GENES 375

Two fundamental classes of cancer gene 375

Viral oncogenes and the natural roles of cellular oncogenes..... 376

How normal cellular proto-oncogenes are activated to become cancer genes..... 376

Tumor suppressor genes: normal functions, the two-hit paradigm, and loss of heterozygosity in linked markers 380

The key roles of gatekeeper tumor suppressor genes in suppressing G₁-S transition in the cell cycle..... 383

The additional role of p53 in activating different apoptosis pathways to ensure that rogue cells are destroyed 384

Tumor suppressor involvement in rare familial cancers and non-classical tumor suppressors..... 384

The significance of miRNAs and long noncoding RNAs in cancer..... 388

10.3 GENOMIC INSTABILITY AND EPIGENETIC DYSREGULATION IN CANCER 389

Different types of chromosomal instability in cancer..... 390

Deficiency in mismatch repair results in unrepaired replication errors and global DNA instability..... 392

Different classes of cancer susceptibility gene according to epigenetic function, epigenetic dysregulation, and epigenome-genome interaction 395

10.4 NEW INSIGHTS FROM GENOME-WIDE STUDIES OF CANCERS 397

Genome sequencing has revealed extraordinary mutational diversity in tumors and insights into cancer evolution..... 398

Defining the landscape of driver mutations in cancer and establishing a complete inventory of cancer-susceptibility genes..... 401

Tracing the mutational history of cancers: just one of the diverse applications of single-cell genomics and transcriptomics in cancer..... 404

Genome-wide RNA sequencing enables insights into the link

between cancer genomes and cancer biology and aids tumor classification..... 405

10.5 GENETIC INROADS INTO CANCER THERAPY..... 407

Targeted anticancer therapies are directed against key cancer cell proteins involved in oncogenesis or in escaping immunosurveillance..... 408

CAR-T Cell therapy and the use of genetically engineered T cells to treat cancer..... 410

The molecular basis of tumor recurrence and the evolution of drug resistance in cancers..... 411

The promise of combinatorial drug therapies 413

SUMMARY.....413

QUESTIONS415

FURTHER READING.....415

11 GENETIC AND GENOMIC TESTING IN HEALTHCARE: PRACTICAL AND ETHICAL ASPECTS 417

11.1 AN OVERVIEW OF GENETIC TESTING..... 418

The different source materials and different levels of genetic testing.....419

11.2 GENETIC TESTING FOR CHROMOSOME ABNORMALITIES AND PATHOGENIC STRUCTURAL VARIATION 423

Screening for aneuploidies using quantitative fluorescence PCR..... 424

Detecting large-scale copy number variants using chromosome SNP microarray analysis..... 425

Detecting and scanning for oncogenic fusion genes using, respectively, chromosome FISH and targeted RNA sequencing..... 428

Detecting pathogenic moderate- to small-scale deletions and duplications at defined loci is often achieved using the MLPA or ddPCR methods..... 430

Two very different routes towards universal genome-wide screens for structural variation: genome-wide sequencing and optical genome mapping..... 433

11.3 GENETIC AND GENOMIC TESTING FOR PATHOGENIC POINT MUTATIONS AND DNA METHYLATION TESTING 433

Diverse methods permit rapid genotyping of specific point mutations 436

The advantages of multiplex genotyping..... 438

Mutation scanning: from genes and gene panels to whole exome and whole genome sequencing..... 440

Interpreting and validating sequence variants can be aided by extensive online resources..... 442

Detecting aberrant DNA methylation profiles associated with disease..... 448

11.4 GENETIC AND GENOMIC TESTING: ORGANIZATION OF SERVICES AND PRACTICAL APPLICATIONS 450

The developing transformation of genetic services into mainstream genomic medicine 450

An overview of diagnostic and pre-symptomatic or predictive genetic testing 453

The different ways in which diagnosis of genetic conditions is carried out in the prenatal period... 456

Preimplantation genetic testing is carried out to prevent

the transmission of a harmful genetic defect using <i>in vitro</i> fertilization	460
Noninvasive prenatal testing (NIPT) and whole genome testing of the fetus.....	461
An overview of the different types of genetic screening	463
Pregnancy screening for fetal abnormalities.....	463
Newborn screening allows the possibility of early medical intervention.....	464
Different types of carrier screening can be carried out for autosomal recessive conditions.....	466
New genomic technologies are being exploited in cancer diagnostics.....	468
Bypassing healthcare services: the rise of direct-to-consumer (DTC) genetic testing.....	470
The downsides of improved sensitivity through whole genome sequencing: increased uncertainty about what variants mean.....	472
11.5 ETHICAL, LEGAL, AND SOCIETAL ISSUES (ELSI) IN GENETIC TESTING.....	473
Genetic information as family information.....	473
Consent issues in genetic testing.....	474
The generation of genetic data is outstripping the ability to provide clinical interpretation	477
New disease gene discovery and changing concepts of diagnosis	479
Complications in diagnosing mitochondrial disease	479
Complications arising from incidental, additional, secondary, or unexpected information.....	480
Consent issues in testing children	482
Ethical and societal issues in prenatal diagnosis and testing.....	483
Ethical and social issues in some emerging treatments for genetic disorders	485
The ethics of germline gene modification for gene therapy and genetic enhancement.....	487
SUMMARY	489
QUESTIONS	491
FURTHER READING.....	491
Glossary	493
Index.....	509