MOLECULAR GENETIC PATHOLOGY
MOLECULAR GENETIC PATHOLOGY

Edited by

LIANG CHENG, MD
Professor of Pathology and Urology
Director of Molecular Pathology Laboratory
Chief of Genitourinary Pathology Division
Department of Pathology and Laboratory Medicine
and Clarian Pathology Laboratory
Indiana University School of Medicine, Indianapolis, IN

DAVID Y. ZHANG, MD, PhD, MPH
Associate Professor of Pathology
Director, Molecular Pathology Laboratory
Department of Pathology
Mount Sinai School of Medicine
and the Mount Sinai Medical Center
New York, NY

Humana Press
Preface

We have had the opportunity to witness both the beginning and the subsequent growth of an exciting specialty that combines both pathology and medical genetics, a field commonly referred to as molecular genetic pathology. The birth of this specialty took place in 1988 when Kari Mullis developed a new DNA amplification technology called the polymerase chain reaction (PCR). Within a few years, this technology was no longer being used exclusively in research laboratories. The technique had found numerous new applications in clinical medicine as a tool for diagnosis and diseases monitoring. The use of PCR technology has greatly expanded the specialties of anatomic and clinical pathology and has increased the availability of genetic testing in the clinical setting. We expect that such advances as the completion of the Human Genome Project, the maturation of pharmacogenomics, the growth of proteomics, and the rapidly growing field of molecular genetic pathology will lead to a new era of personalized and customized patient care.

More recently, the American Board of Pathology (ABP), in conjunction with the American Board of Medical Genetics (ABMG), established a new subspecialty, molecular genetic pathology. Fellowship training for molecular genetic pathology is approved by the Accreditation Council for Graduation Medical Education (ACGME). Many pathologists and medical geneticists are applying for advanced training in this growing subspecialty. Training in molecular pathology is also becoming a required element in pathology residency curricula. To meet these demands, a team of more than 50 leading experts has compiled this quick reference book for medical students, general practitioners, medical technologists, pathologists, and medical geneticists. We also hope that residents or fellows who are training in pathology and medical genetics will find this book helpful in their preparation for board examinations.

Molecular Genetic Pathology contains two parts. Part I covers general molecular genetic pathology and technology, including principles of clinical molecular biology, principles of clinical cytogenetics, diagnostic methodology and technology, tissue microarrays and biomarker validation, laser capture microdissection, clinical flow cytometry, conceptual evolution in cancer biology, clinical genomics in oncology, clinical proteomics, clinical pharmacogenomics, clonality analysis in surgical pathology, fluorescence in situ hybridization (FISH), conventional cytogenetics for hematology and oncology diagnosis, instrumentation, genetic inheritance and population genetics, and genetic counseling. Part II provides disease-based information, including prenatal diagnosis, familial cancer syndromes, molecular testing for solid tumors, molecular pathology of the central nervous system, molecular virology, molecular bacteriology, mycology and parasitology, molecular testing for coagulopathies, molecular hemoglobinopathies, molecular diagnostics of lymphoid malignancies, molecular diagnostics of myeloid leukemias, HLA system and transfusion medicine (molecular approach), molecular forensic pathology, gene therapy, ethical and legal issues in molecular testing, and quality assurance and laboratory inspection. Each chapter begins with a detailed Table of Contents for easy reference.

Assembling this diverse guidebook has truly been a team effort, cutting across many traditional specialty boundaries. We are most grateful for all the contributors who made this project possible. Our special thanks go to Mr. Ryan P. Christy from the Multimedia Education Division of the Department of Pathology at Indiana University, who has edited the illustrations for the book. We would like to thank the staff at Humana Press/Springer, including Ms. Mary Jo Casey, Mr. Paul Dolgert, Mr. Richard Hruska, and Mr. David Casey for their assistance in the development and editing of this text, and in particular Ms. Amy Thau, without whose outstanding work this book would have been an impossible achievement.

Liang Cheng, MD
David Y. Zhang, MD, PhD, MPH
## Contents

Preface ................................................................. v  
Contributors ......................................................... ix

### Part I General Sections

1 Principles of Clinical Molecular Biology  
*Shaobo Zhang, Darrell D. Davidson, David Y. Zhang, Jodi A. Parks, and Liang Cheng* ........................................ 1

2 Principles of Clinical Cytogenetics  
*Stuart Schwartz* ................................................. 33

3 Diagnostic Methodology and Technology  
*Josephine Wu, Tao Feng, Ruliang Xu, Fei Ye, Bruce E. Petersen, Liang Cheng, and David Y. Zhang* ..................... 65

4 Tissue Microarrays  
and Biomarker Validation  
*Martina Storz and Holger Moch* .............................. 133

5 Laser Capture Microdissection  
*Mattthew Kuhar and Liang Cheng* ............................ 141

6 Clinical Flow Cytometry  
*Magdalena Czader* ............................................... 155

7 Conceptual Evolution in Cancer Biology  
*Shaobo Zhang, Darrell D. Davidson, and Liang Cheng* ....... 185

8 Clinical Genomics in Oncology  
*Hugo M. Horlings and Marc van de Vijver* .................... 209

9 Clinical Proteomics  
*David H. Geho, Virginia Espina, Lance A. Liotta, Emanuel F. Petricoin, and Julia D. Wulfkuhle* ......................... 231

10 Clinical Pharmacogenomics  
*Catalina Lopez-Correa and Lawrence M. Gelbert* ............. 241

11 Clonality Analysis in Modern Oncology and Surgical Pathology  
*Liang Cheng, Shaobo Zhang, Timothy D. Jones, and Deborah E. Blue* .................................................. 261

12 Fluorescence *In Situ* Hybridization (FISH) and Conventional Cytogenetics for Hematology and Oncology Diagnosis  
*Vesna Nafjeld* ...................................................... 303

13 Instrumentation  
*Bruce E. Petersen, Josephine Wu, Liang Cheng, and David Y. Zhang* .................................................. 365

14 Genetic Inheritance and Population Genetics  
*Tatiana Foroud and Daniel L. Koller* ............................ 393

15 Genetic Counseling  
*Kimberly A. Quaid and Lisa J. Cushman* ......................... 405

### Part II Disease-Based Sections

16 Molecular Medical Genetics  
*Lisa Edelmann, Stuart Scott, and Ruth Kornreich* ............. 415

17 Prenatal Diagnosis  
*Nataline Kardon and Lisa Edelmann* ............................. 441

18 Familial Cancer Syndromes  
*Michele P. Elieff, Antonio Lopez-Beltran, Rodolfo Montironi, and Liang Cheng* ........................................ 449
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Testing for Solid Tumors</td>
<td>467</td>
</tr>
<tr>
<td>Neal I. Lindeman and Paola Dal Cin</td>
<td></td>
</tr>
<tr>
<td>Molecular Pathology of the Central Nervous System</td>
<td>497</td>
</tr>
<tr>
<td>Eyas M. Hattab and Brent T. Harris</td>
<td></td>
</tr>
<tr>
<td>Molecular Virology</td>
<td>533</td>
</tr>
<tr>
<td>Josephine Wu, Mona Sharaan, and David Y. Zhang</td>
<td></td>
</tr>
<tr>
<td>Molecular Bacteriology, Mycology and Parasitology</td>
<td>581</td>
</tr>
<tr>
<td>Mona Sharaan, Josephine Wu, Bruce E. Petersen, and David Y. Zhang</td>
<td></td>
</tr>
<tr>
<td>Molecular Testing for Coagulopathies</td>
<td>623</td>
</tr>
<tr>
<td>Veshana Ramiah and Thomas L. Ortel</td>
<td></td>
</tr>
<tr>
<td>Molecular Hemoglobinopathies</td>
<td>637</td>
</tr>
<tr>
<td>Jodi A. Parks, Tina Y. Fodrie, Shaobo Zhang, and Liang Cheng</td>
<td></td>
</tr>
<tr>
<td>Molecular Diagnostics of Lymphoid Malignancies</td>
<td>655</td>
</tr>
<tr>
<td>Francisco Vega and Dan M. Jones</td>
<td></td>
</tr>
<tr>
<td>Molecular Diagnostics of Myeloid Leukemias</td>
<td>675</td>
</tr>
<tr>
<td>C. Cameron Yin and Dan M. Jones</td>
<td></td>
</tr>
<tr>
<td>The HLA System and Transfusion Medicine: Molecular Approach</td>
<td>689</td>
</tr>
<tr>
<td>S. Yoon Choo</td>
<td></td>
</tr>
<tr>
<td>Molecular Forensic Pathology</td>
<td>703</td>
</tr>
<tr>
<td>P. Michael Conneally and Stephen R. Dlouhy</td>
<td></td>
</tr>
<tr>
<td>Gene Therapy: Vector Technology and Clinical Applications</td>
<td>717</td>
</tr>
<tr>
<td>Kenneth Cornetta</td>
<td></td>
</tr>
<tr>
<td>Ethical and Legal Issues in Molecular Testing</td>
<td>731</td>
</tr>
<tr>
<td>Kimberly A. Quaid</td>
<td></td>
</tr>
<tr>
<td>Quality Assurance and Laboratory Inspection</td>
<td>737</td>
</tr>
<tr>
<td>Carol L. Johns and Liang Cheng</td>
<td></td>
</tr>
<tr>
<td>Molecular Diagnostics of Myeloid Leukemias: Molecular Approach</td>
<td></td>
</tr>
<tr>
<td>Liang Cheng and Shaobo Zhang</td>
<td>751</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>767</td>
</tr>
</tbody>
</table>
Contributors

DEBORAH E. BLUE, MD
Assistant Professor of Pathology
Associate Director, Molecular Pathology Laboratory
Department of Pathology and Laboratory Medicine
and Clarian Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN

LIANG CHENG, MD
Professor of Pathology and Urology
Director of Molecular Pathology Laboratory
Chief, Genitourinary Pathology Division
Department of Pathology and Laboratory Medicine
and Clarian Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN

S. YOON CHOO, MD
Associate Professor of Pathology and Medicine
Director of HLA Laboratory, Associate Medical Director
of Blood Bank
Departments of Pathology and Medicine
Mount Sinai School of Medicine and the Mount Sinai
Medical Center
New York, NY

P. MICHAEL CONNEALLY, PhD
Distinguished Professor Emeritus, Medical and
Molecular Genetics and Neurology
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

KENNETH CORNETTA, MD
Joe C. Christian Professor and Chairman
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

LISA J. CUSHMAN, PhD
Certified Genetic Counselor
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

MAGDALENA CZADER, MD, PhD
Assistant Professor of Pathology
Director of Clinical Flow Cytometry Laboratory
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
Indianapolis, IN

PAOLA DAL CIN, PhD
Associate Professor of Pathology
Cytogenetics Laboratory
Department of Pathology
Brigham and Women's Hospital
and Harvard Medical School
Boston, MA

DARRELL D. DAVIDSON, MD, PhD
Assistant Professor of Pathology
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
Indianapolis, IN

STEPHEN R. DLOUHY, PhD
Associate Research Professor
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

LISA EDELMANN, PhD
Assistant Professor
Department of Genetics and Genomic Sciences
Director, Molecular Cytogenetics
Co-Director, Genetic Testing Laboratory
Mount Sinai School of Medicine and the Mount Sinai
Medical Center
New York, NY

MICHELLE P. ELIEFF, MD
Staff Pathologist
Diagnostic Pathology Services, Inc.
Nampa, ID
Contributors

VIRGINIA ESPINA, MS
Research Professor
The Center for Applied Proteomics
and Molecular Medicine
George Mason University
Manassas, VA

TAO FENG, MS, MP (ASCP)
Research Assistant
Department of Pathology
Mount Sinai School of Medicine
and the Mount Sinai Medical Center
New York, NY

TINA Y. FODRIE, BS, MT, MP (ASCP)
Supervisor, Department of Molecular Pathology
Indiana University School of Medicine
and VA Medical Center
Indianapolis, IN

TATIANA FOROUD, PhD
Professor of Medical and Molecular Genetics
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

DAVID H. GEHO, MD, PhD
Associate Director of Imaging
Merck and Company, Inc.
West Point, PA

LAWRENCE M. GELBERT, PhD
Research Advisor
Eli Lilly and Company
Indianapolis, IN

BRENT T. HARRIS, MD, PhD
Assistant Professor of Pathology
Department of Pathology
Dartmouth Medical School
Lebanon, NH

EYAS M. HATTAB, MD
Associate Professor of Pathology
Department of Pathology and Laboratory Medicine
and Clarian Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN

HUGO M. HORLINGS, MD
Department of Pathology
The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital
Amsterdam, The Netherlands

CAROL L. JOHNS, PhD
Supervisor and Technical Coordinator
Clarian Molecular Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN

DAN M. JONES, MD, PhD
Professor of Pathology
Department of Hematopathology
Medical Director, Molecular Diagnostics Laboratory
The University of Texas M. D. Anderson Cancer Center
Houston, TX

TIMOTHY D. JONES, MD
Staff Pathologist
Department of Pathology
Floyd Memorial Hospital and Health Services
New Albany, IN

NATALINE KARDON, MD
Associate Professor
Director of Prenatal Cytogenetics Laboratory
Department of Genetics and Genomic Sciences
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

DANIEL L. KOLLER, PhD
Research Assistant Professor
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

RUTH KORNREICH, PhD
Research Assistant Professor of Human Genetics
Co-Director Genetic Testing Laboratory
Department of Human Genetics
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

MATTHEW KUHAR, MD
Resident
Department of Pathology and Laboratory Medicine and Clarian Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN
Contributors

NEAL I. LINDEMAN, MD
Associate Pathologist, Clinical Chemistry
Associate Pathologist, Molecular Diagnostics
Brigham and Women’s Hospital
Assistant Professor of Pathology
Harvard Medical School
Boston, MA

LANCE A. LIOTTA, MD, PhD
Director, The Center for Applied Proteomics and Molecular Medicine
Professor of Life Sciences
George Mason University
Manassas, VA

ANTONIO LOPEZ-BELTRAN, MD, PhD
Professor of Pathology
Department of Pathology and Surgery
Cordoba University Medical School
Cordoba, Spain

CATALINA LOPEZ-CORREA, MD, PhD
Principal Research Scientist
Eli Lilly and Company
Indianapolis, IN

HOLGER MOCH, MD
Professor and Chairman
Institute of Surgical Pathology
Department of Pathology
University Hospital Zürich
Zürich, Switzerland

RODOLFO MONTIRONI, MD, FRCPath
Professor of Pathology
Institute of Pathological Anatomy and Histopathology
Polytechnic University of the Marche Region (Ancona)
School of Medicine
United Hospitals
Ancona, Italy

VESNA NAJFELD, PhD
Professor of Pathology and Medicine
Director, Tumor Cytogenetics, and Oncology, Molecular and Cellular Tumor Markers
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

THOMAS L. ORTEL, MD, PhD
Associate Professor of Medicine and Pathology
Hemostasis and Thrombosis Center
Duke University Medical Center
Durham, NC

JODI A. PARKS, MD
Visiting Lecturer, Clinical Chemistry
Department of Pathology and Laboratory Medicine
and Clarian Pathology Laboratory
Indiana University School of Medicine
Indianapolis, IN

BRUCE E. PETERSEN, MD
Molecular Genetic Pathology Fellow
Department of Pathology
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

EMANUEL F. PETRICIONI, PhD
Professor of Life Sciences
Director, The Center for Applied Proteomics and Molecular Medicine
Chair, Department of Molecular and Microbiology
George Mason University
Manassas, VA

KIMBERLY A. QUAILD, PhD
Professor of Medical and Molecular Genetics
Department of Medical and Molecular Genetics
Indiana University School of Medicine
Indianapolis, IN

VESHANA RAMIAH, MD
Hematology/Oncology Fellow
Duke University Medical Center
Durham, NC

STUART SCHWARTZ, PhD, FACMG
Professor of Human Genetics
Department of Human Genetics
University of Chicago
Chicago, IL

STUART SCOTT, PhD
Clinical Molecular Genetics Fellow
Department of Human Genetics
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

MONA SHARAAN, MD
Molecular Genetic Pathology Fellow
Department of Pathology
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY
Contributors

MARTINA STORZ, BS
Director of TMA Core Facility
Institute of Surgical Pathology
Department of Pathology
University Hospital Zürich
Zürich, Switzerland

MARC VAN DE VIJVER, MD, PhD
Professor of Pathology
Department of Pathology
The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital
Amsterdam, The Netherlands

FRANCISCO VEGA, MD, PhD
Assistant Professor of Pathology
Department of Hematopathology
The University of Texas M. D. Anderson Cancer Center
Houston, TX

JOSEPHINE WU, DDS, CLSp(MB), CLDir
Assistant Professor
Department of Pathology
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

JULIA D. WULFKUHLE, PhD
Research Professor
The Center for Applied Proteomics and Molecular Medicine
George Mason University
Manassas, VA

RULIANG XU, MD, PhD
Associate Professor of Pathology
Department of Pathology
New York University School of Medicine
New York, NY

FEI YE, PhD
Assistant Professor
Department of Pathology
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

C. CAMERON YIN, MD, PhD
Assistant Professor of Pathology
Department of Hematopathology
The University of Texas M. D. Anderson Cancer Center
Houston, TX

DAVID Y. ZHANG, MD, PhD, MPH
Associate Professor of Pathology
Director, Molecular Pathology Laboratory
Department of Pathology
Mount Sinai School of Medicine and the Mount Sinai Medical Center
New York, NY

SHAOBO ZHANG, MD
Associate Research Professor
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
Indianapolis, IN
Part I

General Sections
Principles of Clinical Molecular Biology

Shaobo Zhang, MD, Darrell D. Davidson, MD, PhD, David Y. Zhang, MD, PhD, MPH, Jodi A. Parks, MD, and Liang Cheng, MD

CONTENTS

I. Deoxyribonucleic Acid (DNA) ..........1-4
   Overview ...........................................1-4
   Types of DNAs ......................................1-5
   DNA Replication ....................................1-7
   DNA Mutation .......................................1-8
   DNA Mutation and Disease ......................1-9
   Factors Related to DNA Aberrations ..........1-9
   DNA Repair Mechanisms ..........................1-12

II. Genes .............................................1-13
    Overview .........................................1-13
    Gene Components .................................1-13
    Functional Categories of Genes ..............1-13
    Cancer-Related Genes ............................1-14
    Regulation of Gene Expression ...............1-14
    Signal Transduction ..............................1-19

III. Chromosomes .................................1-19
     Overview .........................................1-19
     Chromatin ........................................1-19

IV. RNA and Proteins .........................1-22
    Overview .........................................1-22
    Types of RNA ......................................1-23
    Ribosome and Ribozyme ...........................1-23
    mRNA Processing ..................................1-24
    Protein Translation ..............................1-27

V. Mitochondrial DNA .........................1-28
    Overview .........................................1-28
    mtDNA Inheritance ................................1-28
    Characteristics of mtDNA ......................1-28
    Mitochondrial Genes and Gene Expression ....1-28
    mtDNA Replication ...............................1-29
    mtDNA Damage, Mutations, and Repair ........1-30
    Mitochondrial Disease ...........................1-31

VI. Suggested Reading .........................1-32
DEOXYRIBONUCLEIC ACID (DNA)

Overview

- **Definition**
  - DNA is a large nucleic acid polymer arranged in chromosomes for storage, expression and transmission of genetic information
  - The genetic information is encoded by a sequence of nucleotides

- **Components of DNA**
  - Bases are molecules containing carbon-nitrogen rings in DNA
    - Purines: adenine (A) and guanine (G) have two joined carbon-nitrogen rings
    - Pyrimidines: thymine (T) and cytosine (C) have one carbon-nitrogen ring
  - Nucleoside is made up of a five-carbon sugar (deoxyribose) and a base
  - Deoxyribose is the same sugar found in RNA, but with oxygen removed from the 2' carbon position

- Nucleotide is made up of a phosphate group, a pentose sugar (deoxyribose), and a nitrogenous base

  - **Phosphodiester bond**
    - The phosphate group is bond to the nucleoside at the hydroxyl group of the 5' carbon atom of deoxyribose
    - Phosphodiester bonds are strong covalent bonds between phosphate groups connecting the 5' carbon of one deoxyribose to the 3' carbon of the next deoxyribose of the adjacent nucleotide nucleotide
  - The phosphodiester bond determines DNA chain polarity (ends designated as either 5' or 3')

- **DNA sequence refers to the order of the nucleotides in a DNA strand, which code for unique sets of genetic information, both proteins and regulatory segments**

![Fig. 1](image_url)

Fig. 1. There are four bases in DNA: Adenine (A), guanine (G), thymine (T), and cytosine (C). Adenine and guanine are purines and thymine and cytosine are pyrimidines. Deoxyribose is the sugar in DNA. The carbon atoms are numbered as indicated. Note there is no oxygen on site 2 of deoxyribose. A nucleoside molecule is composed of a base and deoxyribose. When a phosphate group is added to nucleoside, the complex becomes a nucleotide. Nucleotides are the basic building blocks of DNA.
- The deoxyribonucleotides in DNA differ only in the bases they carry, so the DNA sequence is denoted by a base sequence (e.g., -ATTGCAT-)
- Base sequence is presented from 5' to 3'
- DNA strands are pairs of complementary molecules, which entwine each other in an antiparallel direction
- Two strands of DNA wind around each other to form a double helix (Figure 3)
  - Deoxyribose-phosphate backbone is on the exterior of the DNA double helix
  - The interior of the DNA is formed by paired bases attached to each other by hydrogen bonds. G (Guanine) pairs with C (Cytosine) via three hydrogen bonds, and A (Adenine) pairs with T (Thymine) via two hydrogen bonds inside the double helix. Note that the three hydrogen bonds joining G to C (GC bond) are stronger than the two hydrogen bonds joining A to T (AT bond) (Figure 4)
- DNA has two DNA chains; one is oriented 5'→3' while the other strand is oriented 3'→5' direction (antiparallel)
  - Sense is a DNA strand that could be transcribed. Sense strand has a sequence similar to its RNA transcript
  - Antisense is the complimentary strand of sense. Antisense works as template for the RNA transcript
  - A DNA fragment appears to have a unique function, either structural, regulatory, or coding

**Types of DNAs**

- Single copy DNA is a specific DNA sequence that is present only once in the genome
- Repetitive DNA is a DNA segment with a specific DNA sequence that is repeated multiple times in the genome
- Moderately repetitive DNA refers to 10–10⁵ copies of the sequence per genome
  - Moderate repeated DNA is found primarily in non-coding sequences
- Highly repetitive DNA describes DNA sequence present in greater than 10⁵ copies per genome
  - Highly repeated DNA is found primarily in centromere and telomere regions as tandem repeats
- Tandem repeat DNA contains a variable number of short DNA sequences repeated many times in series. The number of repeats is unique to each individual, and can be used for relationship testing
  - The tandem repeat pattern may vary from one base repeats (mononucleotide repeat) to several 1000-bp repeat sequences
Fig. 3. Human genomic DNA contains two polynucleotide chains wound around each other to form a double-stranded helix. The two chains are “antiparallel,” one running 5’–3’ and the other running 3’–5’ direction. The DNA strands are synthesized and read out by RNA polymerase in the 5’–3’ direction. The purine or pyrimidine attached to each deoxyribose projects into the center of the helix. Base A pairs with T and a G pairs with C through hydrogen bonds in the central axis.

- These segments of DNA are satellite DNA because of the experimental observation that they often form a minor satellite band near the major centrifugation fraction when DNA is separated by density gradient.
- Clusters of such repeats are scattered on many chromosomes. Each variant is an allele that is inherited co-dominantly.
- Megasatellite DNAs are tandem repeat DNA segments with a length greater than 1000 bp (1 kbp) repeated 50–400 times.
- Satellite DNAs comprise about 15% of human DNA. The repeated sequence ranges from 5 to 170 bp and the complex is about 100 kbp in length.
- Minisatellite DNAs are repeated sequences ranging from 14 to 500 bp in length. The repeat complex is 0.1–20 kbp in length. Minisatellite DNA is present in telomere region.
- Microsatellite DNAs are sequences <15 bp in length that repeat 10–100 times without interruption. There are approx 200,000 microsatellite loci in the human genome (Table 1).
  - Loss of heterozygosity (LOH) in a cell represents the loss of one parent’s contribution of DNA to a cell’s genome, often in microsatellite regions.
    - It often indicates the presence of tumor-suppressor gene loss around the microsatellite locus.
    - LOH can arise through deletion, nonreciprocal DNA transfer, mitotic recombination, or chromosome loss.
    - LOH is often used to analyze the clonal origin of cancer-associated loci.
    - Microsatellite DNA loci are useful markers for the detection of LOH.
The double helical structure of DNA is largely due to hydrogen bonding between the base pairs linking one complementary strand to the other. Hydrogen bonds are non-covalent, weak bonds between electron donors and recipients. There are two hydrogen bonds between A and T and three hydrogen bonds between G and C, thus the bonds between G and C are stronger than between A and T.

Table 1. Major Characteristics of Repetitive DNA

<table>
<thead>
<tr>
<th>Form of DNA</th>
<th>Length (bp)</th>
<th>Number of repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single copy</td>
<td>Vary</td>
<td>Single copy</td>
</tr>
<tr>
<td>Moderately repetitive</td>
<td>Vary</td>
<td>10–10^5</td>
</tr>
<tr>
<td>Highly repetitive</td>
<td>Vary</td>
<td>&gt;10^5</td>
</tr>
<tr>
<td>Tandem repeat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Megasatellite</td>
<td>&gt;1000</td>
<td>50–400</td>
</tr>
<tr>
<td>Satellite</td>
<td>5–170</td>
<td>500–2000</td>
</tr>
<tr>
<td>Minisatellite</td>
<td>14–500</td>
<td>7–40</td>
</tr>
<tr>
<td>Microsatellite</td>
<td>&lt;15</td>
<td>10–100</td>
</tr>
</tbody>
</table>

- When parents’ contributions of certain microsatellite loci are of different size, these microsatellite loci are informative
  - Microsatellite instability at critical loci is a marker for malignancy or premalignant genetic change (see details in Chapter 7)
- Mitochondrial DNA (mtDNA) (see Mitochondrial DNA)

DNA Replication
- Double-stranded DNA (dsDNA) is exactly duplicated prior to cell division so that each daughter cell is endowed with an exact replica of the parent cell DNA
- DNA replication occurs during S (synthesis) phase of the cell cycle
  - Cell cycle refers to a cycle of events in a eukaryotic cell from one cell division to the next; it consists of G_0, G_1, S, G_2, and M phases (Figure 5)
  - Semi-conservative replication means that each DNA molecule consists of one original and one newly synthesized chain (Figure 6)
- DNA Polymerases
  - DNA polymerases are enzymes involved in DNA replication. Eukaryotic cells have five different DNA polymerases
    - DNA polymerase α and δ replicate nuclear DNA
    - DNA polymerase β and ε are involved in DNA repair
    - DNA polymerase γ replicates mitochondria DNA (mtDNA)
  - Multiple replications means that replication begins at multiple sites within a DNA strand and proceeds bidirectionally from each origin (Figure 7)
    - The replication apparatus at each origin forms a bubble and extends toward both ends of the DNA molecule until it meets another bubble
    - The leading strand is the DNA chain that is synthesized continuously in the 5′–3′ direction
      - Synthesis of the leading strand is catalyzed by DNA polymerase δ
    - The lagging strand is the DNA chain that is synthesized as a series of short fragments, known as Okazaki fragments, polymerized in the 5′→3′ direction also (Figure 8)
      - The newly synthesized DNA fragments will eventually meet and ligate to create an intact strand
      - The lagging strand is synthesized by DNA polymerase α
      - The lagging strand polymerizes from 5′ to 3′ at the nucleotide level but overall growth by ligation of Okazaki fragments is in the 3′→5′ direction
Fig. 5. The cell cycle, or cell-division cycle, is the series of events in a eukaryotic cell between one cell division and the next. The cell cycle consists of four phases, G₁, S, G₂, and M phase. G₀ is a period in which cells exist in a quiescent state. G₀, G₁, G₂, and S phase are collectively known as interphase. Cells in G₀ phase are resting cells unable to divide without a signal to re-enter the cell cycle. DNA synthesis occurs during S phase, which is followed by a short G₂ phase. Mitosis and cytokinesis together are defined as the M (mitotic) phase, during which the mother cell divides into two daughter cells.

DNA Mutation

- DNA mutation is a permanent change in the genetic material sequence
- Most mutations are found in non-coding sequences
- Single base pair substitution (point mutation) involves a single nucleotide, which is replaced with another nucleotide
  - Point mutation is the most common form of mutation
  - It happens most commonly in non-coding sequences
  - It is also the most frequent type of mutation associated with tumor suppressor gene mutation
    - Transitions are the mutations that substitute a different purine for a purine or a pyrimidine for a pyrimidine
    - Transversions are mutations that substitute a different purine for a pyrimidine or a pyrimidine for a purine
    - Synonymous (silent) mutation is a single base pair substitution yielding a different codon that still codes for the same amino acid
    - Missense mutation is a single base pair substitution that results in a different codon and a different amino acid
    - Nonsense mutation is a single base pair substitution that converts a codon specifying an amino acid into a stop codon
    - Deletion is an irreversible mutation in which one or more nucleotides are removed from the DNA sequence
      - The deletion will cause a shift of the reading frame
      - A one base deletion, for example, will shift all codons left, altering the amino acids for which they code
    - Insertion is a mutation that adds one or more nucleotides to the DNA sequence
      - An insertion in the coding region of a gene may cause a shift in the reading frame
      - An insertion alters splicing of messenger RNA (mRNA) (splice site mutation)
    - Amplification increases the dosage of genes located within a locus by inserting multiple copies of the chromosomal region or by promulgating fragments of DNA containing the locus outside the chromosomes. Proteins produced from amplified genes are generally increased
    - Loss of heterozygosity (LOH) is a DNA alteration in which one allele from one parent’s contribution is lost, either by deletion or a recombination event
    - The most frequently observed gene associated with LOH in sporadic cancer is p53

- Deletion is an irreversible mutation in which one or more nucleotides are removed from the DNA sequence
  - The deletion will cause a shift of the reading frame
  - A one base deletion, for example, will shift all codons left, altering the amino acids for which they code
- Insertion is a mutation that adds one or more nucleotides to the DNA sequence
  - An insertion in the coding region of a gene may cause a shift in the reading frame
  - An insertion alters splicing of messenger RNA (mRNA) (splice site mutation)
- Amplification increases the dosage of genes located within a locus by inserting multiple copies of the chromosomal region or by promulgating fragments of DNA containing the locus outside the chromosomes. Proteins produced from amplified genes are generally increased
- Loss of heterozygosity (LOH) is a DNA alteration in which one allele from one parent’s contribution is lost, either by deletion or a recombination event
- The most frequently observed gene associated with LOH in sporadic cancer is p53
DNA Mutation and Disease

- DNA mutations cause errors in protein sequences, creating partially or completely non-functional proteins.
- DNA mutations give rise to offspring that carry the mutation in all their cells.
  - Human beings with a single allele mutation will transmit the mutation to half of the progeny.
  - If both alleles are mutated, all progeny will inherit the mutation.

Factors Related to DNA Aberrations

- Spontaneous chemical reactions
  - Spontaneous chemical reactions cause single base pair substitutions or single base pair deletion through the following processes:
    - Tautomerism—base change by repositioning a hydrogen atom
    - Depurination—loss of a purine base (A or G)
    - Deamination—changes a normal base to an atypical base
      - Change from C→U
      - Spontaneous deamination of 5-methycytosine (irreparable)
      - Change from A→HX (hypoxanthine)
    - Transition—a purine changes to another purine, or a pyrimidine to a pyrimidine
    - Transversion—a purine becomes a pyrimidine, or a pyrimidine becomes a purine
- Induced mutations
  - Chemical mutagenesis can modify bases and cause either interstrand or intrastrand cross-linking.
  - Common chemical mutagens include:
    - Nitrosoguanidine (N-methyl-N'-nitrosoguanidine)
    - Hydroxylamine (NH₂OH)
Multiple DNA replication sites

DNA fragment extension

New strands synthesised

Fig. 7. Multiple replications are an efficient way to synthesize chromosomal DNA. DNA synthesis begins at many locations and proceeds bidirectionally from each location. Eventually the replication bubbles merge and the DNA fragments are ligated to form two daughter DNA strands. Each of the newly synthesized double strands consists of one parental and one newly synthesized chain (semiconservative).

Leading strand

Lagging strand

Fig. 8. The DNA double helix is unwound by the enzyme helicase before synthesis of a new DNA chain begins. A DNA polymerase (shown in green) binds to the strand and moves along the strand assembling the leading strand (fragment inside the left fork). The lagging strand is synthesized in discontinuous polynucleotide segments called Okazaki fragments. A series of Okazaki fragments are linked by DNA ligase to form the lagging strand. In eukaryotic cells the leading and lagging strands are synthesized by DNA polymerase δ and α, respectively.

- Base analogs (e.g., bromodeoxyuridine) only mutate DNA when the analog is incorporated during S phase in replicating DNA
- Alkylating agents (e.g., cyclophosphamide) mutate both replicating and non-replicating DNA. The alkylating agent transfers an alkyl group, often to the N7 position of guanine
- Polycyclic hydrocarbons are converted within cells to highly reactive epoxy compounds that react with DNA (e.g., benzpyrenes found in internal combustion engine exhaust and cigarette smoke)
- DNA intercalating agents insert themselves between the stacked bases at the center of the DNA strand (e.g., ethidium bromide)
• DNA cross-linkers cause both interchain covalent bonds and stable bonds between the DNA strands and nuclear proteins (e.g., platinum)
• Oxidative damage caused by oxygen radicals accelerating hydroxylation of guanine to 8-hydroxyguanine, causing a G:C to A:T transversion

Radiation
• Ionizing radiation can cause individual base lesions, cross-linking, or strand breakage, sometimes mediated by oxygen radicals
• Ultra violet radiation causes covalent bonding between adjacent cytosine and thymine bases creating pyrimidine dimers

Viral mutagenesis involves a DNA virus, RNA or retrovirus integrating all or part of its sequence into the human genome
• An episome is a DNA molecule separate from the chromosomal DNA and capable of autonomous replication. It is the common status of viral particles during viral infection
• Integration means that DNA fragments of viral origin have become inserted into chromosomal DNA
• Epstein-Barr virus has been associated with lymphoproliferative disorder in immunocompromised patients (such as post-transplantation)
• Human herpesvirus 8 has been associated with Kaposi’s sarcoma, Castleman’s disease, body cavity lymphoma (primary effusion lymphoma), and multiple myeloma
• Human papillomavirus has been associated with premalignant and malignant transformation of the uterine cervix

Inborn errors of metabolism
• Inborn errors comprise a large class of genetic diseases involving metabolic disorders
• The majority of inborn errors are due to single gene defects that code for enzymes to convert intermediary metabolites

DNA Repair Mechanisms
• DNA repair includes a collection of processes through which a cell identifies and corrects damage to its DNA
• DNA repair is essential to cell survival
• If DNA damage is irreparable then programmed cell death (apoptosis) should ensue
• Failure to correct molecular lesions in gamete-forming cells leads to progeny with congenital mutations
• Single-strand damage repair
  - Direct repair is an enzyme-catalyzed reaction that directly reverses DNA damage. No template is needed to correct an altered base back to its natural state
  - Base excision repair (BER) removes a damaged base and replaces it with a normal base (Figure 9)

Fig. 9. A single base in DNA may be chemically mutated, for example, by deamination or alkylation, causing incorrect base pairing (X), and consequently, incorrect codons in the DNA. BER is initiated by DNA glycosylases linking particular types of chemically altered bases to the deoxyribose-phosphate backbone. This mutated base is excised as a free base, generating sites of base loss called apurinic or apyrimidinic (AP) sites. The AP sites are substrates for AP endonucleases. The ribose-phosphate backbone is then removed from the DNA by an exonuclease called deoxyribophosphodiesterase (dp-diesterase). Then the DNA polymerase and a ligase catalyze incorporation of a specific deoxyribonucleotide into the repaired site, enabling correct base pairing.

Mismatch repair removes nucleotides, which are mispaired with the corresponding base on the complementary chain, usually by DNA polymerase errors. This repair process can also remove up to 30 base insertions (Figure 10)

Translation synthesis temporarily replaces the conventional polymerase at a DNA lesion by one of a group of specialized polymerases that can replicate damaged DNA. It is better for the daughter cell to inherit a point mutation than to have a significant part of a chromosome deleted, causing framshift
Fig. 10. Mismatch repair is a cellular procedure for recognizing and repairing insertion, deletion or misincorporation. The DNA damage is repaired by excising the mis-incorporated base or segment and synthesizing a new stretch of DNA is synthesized to replace the excised segment. This process involves more than just the mismatched nucleotide itself and can lead to the removal and synthesis of a significant piece of DNA. Also there are multiple excision-repair systems in a single cell type.

Fig. 11. Homologous recombination repairs a double-stranded break. It allows the precise replacement of a sequence from one allele with a sequence from the homologous allele. The breaks in dsDNA use a homologous dsDNA molecule as a template. Homologous recombination requires a homologous sequence to guide the repair.

Fig. 12. NHEJ can also repair double-strand breaks in DNA. NHEJ directly ligates the breaks without a homologous template. NHEJ typically utilizes short homologous DNA sequences, termed microhomologous, to guide the repair but still results in some DNA sequence information being lost or "spliced out."

**GENES**

**Overview**
- Genes are DNA sequences that encode heritable biologic characteristics
  - The human genome is divided into two categories, nuclear and mitochondrial genome
- DNA in the human nuclear genome encodes about 30,000–40,000 different genes, much lower than previous estimates of around 100,000 genes before completion of the genome map
- DNA in the human mitochondrial genome encodes 37 genes

**Gene Components (Figure 13)**
- Promoter is a DNA fragment to which RNA polymerase binds to initiate transcription
  - Core promoter directs the basal transcription complex to initiate the transcription of the gene
- TATA box is a short sequence located within the promoter of most genes
  - TATA box has a core 5'-TATAAA-3' sequence
  - The TATA box is usually found as the binding site of RNA polymerase II
Fig. 13. A gene consists of both coding and non-coding sequences. The coding sequence (open reading frame, ORF) extends from a start codon to a stop codon. Introns are non-coding sequences that will be spliced out after transcription. 5' untranslated region (UTR) is a part of mRNA located between cap site and start codon. 3' untranslated region is also a part of mRNA following coding sequence. A promoter and different regulatory motifs are located up stream or downstream of a gene. Enhancer or silencer may be located upstream or downstream of the gene it regulates.

- CCAAT Box (CAT box) is located at -75 and serves as a modulator for the basal transcription
  - CCAAT box has a core 5'-CCAACTC-3' sequence
  - CCAAT box is the binding site of nuclear factor 1 (NF-1) and CCAAT box binding factor (CBF)
- GC box is also called Sp1 box. It has consensus sequences GGGCGG and is found within 100 bp from the transcription initiation site
  - GC boxes serve as a modulators to the basal transcription of the core promoter
- CpG sites are regions of DNA with a high frequency of phosphodiester-linked cytosine-guanine pairs. The “p” in CpG indicates that a normal phosphodiester bond between nucleosides gives the the CG sequence direction
  - CpG islands are located near or within 40% of mammalian gene promoters
  - The genes with CpG islands are expressed if the CpG islands are not methylated
- Enhancers are DNA sequences that when bound by certain factors increase transcription levels of genes
  - Unlike promoters, enhancers do not have to be within or near the genes they act on, or even located on the same chromosome
- Silencer is a DNA sequence that can bind regulators of transcription called repressors. The binding of repressor prevents RNA polymerase from initiating transcription
  - When repressor is bound to target DNA, RNA synthesis is decreased or fully suppressed
- An exon is any region of DNA within a gene that encodes a protein. Exons of many eukaryotic genes interleave with segments of non-coding DNA (introns). Mature mRNA contains only sequentially linked exons
- Introns are sections of non-coding DNA located between exons, which are transcribed into RNA but are spliced out to form mRNA
- Open reading frame (ORF) is the sequence of DNA or mRNA molecule from the start codon (ATG) to a stop codon (TAA, TAG, or TGA). An open reading frame codes for amino acid codons that can be translated into a protein
- Boundary elements (insulator elements) are regions of DNA that mark the 5' and 3' ends of a gene
- Gene expression is the process by which a gene’s DNA sequence is converted into the structures and functions of a cell
  - Protein-coding genes are translated into proteins
  - Non-protein coding genes code for RNAs, (e.g., ribosomal RNA [rRNA] genes, and transfer RNA [tRNA] genes), which usually have a structural, regulatory or catalytic role

Functional Categories of Genes
- Housekeeping genes are genes that are transcribed at a relatively constant level and remain unaffected by environmental conditions
  - Housekeeping gene products are necessary for cell maintenance
    - Since their expression is typically unaffected by experimental conditions, they may be used for normalization of other gene expression levels in the cell
    - Housekeeping genes often lack the CCAAT and TATA boxes
  - Actin and glyceraldehyde 3-phosphate dehydrogenase are examples of housekeeping genes commonly used as controls for mRNA quantitation
• Facultative genes are transcribed only when needed
• Inducible gene expression is either responsive to environmental changes or dependent on the stage of the cell cycle
• Pseudogenes are multiple copy genes characterized by defective copies, mostly truncated, of a functional gene
  – Pseudogenes arise from gene duplication or retrotransposition
• RNA genes transcribe mRNA as their end products without protein translation

Cancer-Related Genes (Table 2)
• Tumor-suppressor genes prevent cell overgrowth (neoplasia)
  – They are involved in cell cycle control, cell differentiation, and apoptosis
  – Tumor-suppressor gene products generally promote genomic stability
  – Tumor-suppressor gene inactivation mechanisms include point mutation, deletion, and epigenetic inactivation of the gene
• Proto-oncogenes are normal genes that cause a malignant phenotype either because of mutation or increased expression. Proto-oncogenes code for proteins to regulate cell growth and differentiation
  – A proto-oncogene becomes an oncogene when mutated, inappropriately expressed or over expressed, transforming the cell by unregulated growth and differentiation
• Oncogene
  – Oncogene products include
    • Growth factors bind receptors on the cell surface to stimulate cell proliferation or to control differentiation
    • Receptors are proteins on the cell surface, within the cytoplasm or in the cell nucleus for binding specific molecules (ligands) to initiate a cellular response
    • Protein kinases chemically add phosphate groups to specific amino acids of substrate proteins. This process usually results in a functional change of the target protein resulting in changed enzyme activity, altered cellular location, or modified association with other proteins
    • Receptor tyrosine kinases are membrane-bound enzymes to transfer a phosphate group from ATP to a tyrosine residue in a protein. The tyrosine kinase-binding hormones and growth factors are generally growth-promoting and mitogenic agents, such as epidermal growth factor receptor (EGFR)
    • Cytoplasmic tyrosine kinases are non-receptor tyrosine kinases (TK) to regulate many cellular processes, such as inducing gene of Rous sarcoma virus (src-family), SYK-Zeta-chain associated protein kinase 70 (Syk-ZAP-70 family), and Bruton’s tyrosine kinase (BTK family) oncogenes belongs to the cytoplasmic tyrosine kinases group. The bcr-abl transcript (fusion gene of the Philadelphia chromosome in chronic myelocytic leukemia (CML) is also a tyrosine kinase. The bcr-abl fusion gene kinase activates mediators of the cell cycle regulation system, leading to a clonal myeloproliferative disorder
  • Regulatory GTPases are a large family of enzymes that bind and hydrolyze GTP (guanosine triphosphate) existing in GTP-bound and -unbound states. They play important roles in the following cellular processes:
    • Signal transduction at the intracellular domain of transmembrane receptors
    • Protein biosynthesis at the ribosome
    • Control of differentiation during cell division
    • Translocation of proteins through membranes
    • Transport of vesicles within the cell
  • Ras oncogene produces a small regulatory GTPase important as a molecular switch for a variety of signal pathways. Ras controls such processes as cytoskeletal integrity, cell proliferation, adhesion, apoptosis, and migration
  • Cytoplasmic serine/threonine kinase phosphorylates the hydroxyl group of serine or threonine. The Raj kinase, and cyclin-dependent kinases belong to the serine/threonine kinase family
  • Adaptor proteins are small accessory proteins, which lack intrinsic enzymatic activity but bind signal transduction pathway components, driving the formation of active protein complexes
  • Transcription factors mediate the binding of RNA polymerase to DNA and initiation of transcription. A transcription factor may work to either stimulate or repress transcription of a gene

Regulation of Gene Expression
• Regulation of gene expression controls the amount and appearance agenda of a gene’s functional product
  – All steps of gene expression can be modulated
  – Regulation of gene expression is the basis for cell differentiation, diversity, and adaptation
• Cis-action factors are short regulatory sequences located within the promoter or in the vicinity of a gene’s structural portion. Cis-sequences facilitate the transcription of adjacent polypeptide-encoding sequences
• Trans-action factors bind to the cis-acting sequences to control gene expression
• Enhancer is a short region of DNA that upregulates transcription levels of genes. Enhancer sequences are active when bound to trans-action factors
Principles of Clinical Molecular Biology

- Response element is a short sequence of DNA within the promoter of a gene that can bind to a specific hormone receptor complex and regulate transcription of genes subject to that hormone

**Signal Transduction**
- A signal transduction pathway is a sequence of enzymes and second messengers by which a receptor communicates with the cell nucleus
- The signal transduction pathway “translates” the receptor ligand message at the surface into a cellular response in the nucleus
  - abl and ras are signal transducers
- Transcription factors

**Overview**
- A chromosome is an enormous macromolecule into which somatic DNA is packaged in eukaryotic cells. Three billion base pairs of nucleotides (a complete set of DNA) are divided among 46 chromosomes, each containing many genes, regulatory elements, and intervening nucleotide sequences (Figure 14)
- Chromosomes are found only in the eukaryotic nucleus and can be seen only during nuclear division
- During most of the life cycle, the genetic material occupies areas of nuclei in the form of chromatin, and individual chromosomes cannot be distinguished
- In eukaryotes, the basic function of the chromosome is to package and compress the DNA, exposing specific genes for transcription during certain phases of the cell life span

**Chromatin**
- Chromatin is the form of genetic material existing during interphase of eukaryotic cells and is made up of DNA and protein
- Chromatin can be seen with the light microscope after staining with nuclear stains
- Chromatin is a packaged state of DNA in a small volume to strengthen the DNA, to allow mitosis and meiosis, and to serve as a mechanism for expression control
- Chromatin functions as a gene regulator
  - The changes in chromatin structure are effected mainly by methylation (DNA and proteins) and acetylation (proteins)
  - Euchromatin is a loosely packed form of chromatin that is involved in active transcription or regulation and is lightly stained by nuclear stains

- A transcription factor is a molecule that initiates transcription of DNA in the eukaryotic nucleus
- Transcription factors interact with promoter or enhancer sequences either by binding directly to DNA or by interacting with other DNA-bound proteins
- myc is an example of a transcription factor that activates expression of many genes by binding to consensus sequences
- Programmed cell death regulators are molecules that prevent apoptosis. Activation of these regulators leads to overgrowth of abnormal cells
  - bcl-2 is an example of a programmed cell death regulator that governs mitochondrial outer membrane permeability and suppresses apoptosis

---

**CHROMOSOMES**

- Heterochromatin is a darkly stained and tightly packed form of DNA. Its major biologic characteristic is that it is not transcribed
- Chromatin composition
  - Histones are the major chromatin binding proteins. They act as spools around which DNA winds
    - Histones play a role in gene regulation
    - Histones H2A, H2B, H3, and H4 form octamers (two of each) with a cylindrical shape (Figure 15)
Fig. 15. Nucleosomes are the fundamental repeating subunits of all eukaryotic chromatin. They package DNA into chromosomes inside the cell nucleus and control gene expression. The DNA winding around the nucleosome core particle consists of about 146 bp of dsDNA wrapped in 1.65-left-handed superhelical turns around complexes of the four histone proteins known as the histone octamers. The DNA hanging between two nucleosome cores is typically 55-bp long and is known as linker DNA.

- When DNA winds 1.65 times around a histone octamer a nucleosome results
  - Each nucleosome contains 146 bp
- The nucleosomes are stacked and further coiled into a 30-nm fiber, which makes up the chromosome residing in the cell nucleus
- Linker DNA is the DNA hanging between two nucleosomes, typically 55-bp long
- Histone–DNA interaction regulates gene expression. Acetylation of histone modulates gene expression and leads to transcription activation. The extent of interaction between histone and DNA is affected by the degree of histone acetylation
- Histone acetylation is the process in which charged lysine side chains are acetylated, leading to reduced affinity between histone and DNA. After acetylation, RNA polymerase and transcription factors have better access to the promoter

### Chromosomes

- Chromosome structure
  - Each chromosome has two short arms (p), two long arms (q), one centromere, and four telomeres
  - The centromere is the constricted region of a chromosome, which has a special sequence and structure for attachment to the spindle filament during M-phase and for separation of chromosomes during mitosis (Figure 14)

  - The centromere divides the chromosome into four arms
    - The two equal short arms are designated “p” (petite)
    - The two equal long arms are designated “q” (follows p in the Latin alphabet)
  - Telomere: (see Chapter 7, Telomere section)

- Chromosome grouping
  - Chromosomes are numbered and grouped according to their morphologic characteristics
  - Chromosomes are numbered according to their relative sizes from largest to smallest. The position of the centromere determines chromosome grouping
    - Group A have nearly equal p and q arms whereas group E have the centromere almost at the telomere
  - Chromosome identification is confirmed by the banding pattern unique to each chromosome (Figure 16)

- Chromosome number
  - Human cells contain 46 chromosomes (23 from each parent) including 22 pairs of autosomes and one pair of sex chromosomes
  - The number of chromosomes doubles during cell division
    - Meiosis is a process allowing one diploid cell to divide in a special way into four haploid cells in eukaryotes
    - Mitosis is the process by which a cell separates its duplicated chromosomes into two identical sets of chromosomes
  - Ploidy is the number of homologous sets of chromosomes in a cell
    - Haploidy (monoploidy) is the number of chromosomes in the gamete of an individual (23 in a human). Haploid chromosomes have only one short and one long arm
    - Diploidy is the normal state of chromosomes in a cell, with two copies of each chromosome, one from each parent. The two chromosomes in a pair are said to be homologous
    - Polyploidy is the state of cells with extra chromosomes beyond the basic set of paired chromosomes
    - Aneuploidy is a condition in which the number of chromosomes is abnormal owing either to extra or missing chromosomes. The number of chromosomes in an aneuploide cell cannot be a multiple of the haploid set (Figure 17)
      - Monosomy is a type of aneuploidy with at least one missing parental chromosome (Figure 18)
      - Trysomy is a type of aneuploidy with one extra chromosome added to a pair of homologous chromosomes
Human Genome and Chromosomes

Genome size:
$3.2 \times 10^9$ nucleotides

22 autosomes and 2 sex chromosomes

Diploid (2N):
$(1-22) \times 2 + XY/XX = 46$

$\sim 35,000$ genes

Fig. 16. Chromosomes are grouped according to their relative size, the position of their centromere, and banding patterns.

Fig. 17. Aneuploidy is a condition in which the number of chromosomes is not a multiple of the haploid set due to gaining or losing chromosomes. The figure shows a fluorescent probe chromosome painting of tumor cell chromosomes featuring a series of chromosome gains and loses.

Fig. 18. Monosomy is a type of aneuploidy with loss of one chromosome from a pair in the cell’s diploid chromosome set. Fluorescence in situ hybridization shows the loss of one chromosome in these cells indicated by having only one signal present in each nucleus.
- Structural alterations of chromosomes (see details in Chapter 2)
  - Inversion occurs when a chromosome segment is flipped end to end. Inversion is designated by the symbol inv
  - Reciprocal translocation is a chromosomal rearrangement caused by the interchange of chromosome segments between non-homologous chromosomes. Reciprocal translocations are denoted by the symbol t followed by parentheses showing the exchanged chromosome breakpoints separated by a semicolon
  - Isochromosome is a chromosome which has lost one set of its arms, either p or q, and replaced them with an exact copy of the other arms. Isochromosomes thus have four identical arms, either p or q. Isochromosome is denoted by the symbol i
- Ring chromosome is a chromosome that is formed when the telomeres have been lost, and the ends of arms fuse together to form a ring. A ring chromosome is denoted by the symbol r
- Fragile sites are chromosome regions that are poorly connected to the rest of the chromosome
  - Fragile sites are often rich in CGG or CGC repeats and are inherited like a gene and break away frequently
  - Double chain breaks in fragile sites lead to the loss of genetic material
  - Fragile sites are especially prone to breakage when cells are cultured under conditions that inhibit DNA replication or repair

Selected tumors with chromosomal anomalies (Table 2)

Table 2. Selected Tumors with Commonly Found Chromosomal Anomalies

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Common chromosome anomalies</th>
<th>Genes involved</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epithelial tumors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>9q22.3</td>
<td>PTCH</td>
</tr>
<tr>
<td>Clear cell renal carcinoma</td>
<td>3p25-26</td>
<td>VHL</td>
</tr>
<tr>
<td>Translocation renal cell carcinoma</td>
<td>t(X;1)(p11.2q21)</td>
<td>PRCC-TFE3</td>
</tr>
<tr>
<td></td>
<td>t(X;17)(p11.2-q25)</td>
<td>ASPL-TFE3</td>
</tr>
<tr>
<td></td>
<td>t(X;1)(p11.2;q34)</td>
<td>PSF-TFE3</td>
</tr>
<tr>
<td>Papillary renal cell carcinoma</td>
<td>Gain 7, 17, loss Y, and 4</td>
<td></td>
</tr>
<tr>
<td>Hereditary papillary renal cell cancer</td>
<td>7q31</td>
<td>c-MET</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>1q</td>
<td>BRCA1, Her-2/neu</td>
</tr>
<tr>
<td></td>
<td>17q21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13q12</td>
<td>BRCA2</td>
</tr>
<tr>
<td></td>
<td>del(16q)</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>del(17p)</td>
<td>TP53</td>
</tr>
<tr>
<td></td>
<td>12p</td>
<td>Ras</td>
</tr>
<tr>
<td></td>
<td>3p14</td>
<td>FHIT</td>
</tr>
<tr>
<td></td>
<td>5q21-22</td>
<td>APC</td>
</tr>
<tr>
<td></td>
<td>18q21</td>
<td>DCC, SMAD4</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>del(3p)</td>
<td>FHIT</td>
</tr>
<tr>
<td></td>
<td>13q</td>
<td>RB</td>
</tr>
<tr>
<td></td>
<td>9p21</td>
<td>P16</td>
</tr>
<tr>
<td></td>
<td>17p</td>
<td>TP53</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Tumor</th>
<th>Common chromosome anomalies</th>
<th>Genes involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate cancer</td>
<td>t(21;21)(p22.2;q22.3)</td>
<td><strong>TMPRSS2-EGR</strong></td>
</tr>
<tr>
<td></td>
<td>del(8p12-21)</td>
<td><strong>NKK31</strong></td>
</tr>
<tr>
<td></td>
<td>1q24</td>
<td><strong>HPC1</strong></td>
</tr>
<tr>
<td></td>
<td>Xq27-28</td>
<td><strong>HPCX</strong></td>
</tr>
<tr>
<td></td>
<td>Xq11</td>
<td><strong>AR</strong></td>
</tr>
<tr>
<td></td>
<td>del(10q24)</td>
<td><strong>PTEN</strong></td>
</tr>
<tr>
<td></td>
<td>Trisomy 7</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Loss of Y</td>
<td>–</td>
</tr>
<tr>
<td>Bladder transitional cell carcinoma</td>
<td>gain 3, 7, 17, del(9p21)</td>
<td><strong>P53, P16</strong></td>
</tr>
<tr>
<td>(UroVysion panel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medullary thyroid carcinoma</td>
<td>10q11.2</td>
<td><strong>RET</strong></td>
</tr>
<tr>
<td>Papillary thyroid carcinoma</td>
<td>10q11-q13</td>
<td><strong>RET</strong></td>
</tr>
<tr>
<td>Mesothelioma</td>
<td>del(3p21)</td>
<td><strong>CTNNB1</strong></td>
</tr>
<tr>
<td>Ovarian papillary cystadenocarcinoma</td>
<td>t(6;14)</td>
<td>–</td>
</tr>
<tr>
<td>Granulosa cell tumor and Brenner tumor</td>
<td>trisomy 12</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>i(12p)</td>
<td>–</td>
</tr>
<tr>
<td>Testicular germ cell tumors</td>
<td>i(12p)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>12p overrepresentation</td>
<td>–</td>
</tr>
<tr>
<td>Wilm’s tumor</td>
<td>del(11p13)</td>
<td><strong>WT1</strong></td>
</tr>
</tbody>
</table>

**Soft tissue tumors**

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Common chromosome anomalies</th>
<th>Genes involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolar soft-part sarcoma</td>
<td>t(X;17)(p11;q25)</td>
<td><strong>TFE3-ASPL</strong></td>
</tr>
<tr>
<td>Alveolar rhabdomyosarcoma</td>
<td>t(2;13)(q35;q14),</td>
<td><strong>PAX3-FKHR,</strong></td>
</tr>
<tr>
<td></td>
<td>t(1;13)(p36;q14)</td>
<td><strong>PAX7-FKHR</strong></td>
</tr>
<tr>
<td>Clear cell sarcoma (melanoma of soft part)</td>
<td>t(12;22)(q13;q12)</td>
<td><strong>EWS-ATFI</strong></td>
</tr>
<tr>
<td>Dermatofibrosarcoma protubers and giant cell fibroblastoma</td>
<td>t(17;22)(q22;q13)</td>
<td><strong>COL1A1-PDGFB</strong></td>
</tr>
<tr>
<td>Myxoid chondrosarcoma</td>
<td>t(9;22)(q22;q12)</td>
<td><strong>EWS-CHN</strong></td>
</tr>
<tr>
<td>Lipoma</td>
<td>t(3;12)(q27;q13)</td>
<td><strong>HMGIC-LPP</strong></td>
</tr>
<tr>
<td>Lipoblastoma</td>
<td>8q rearrangement</td>
<td>–</td>
</tr>
<tr>
<td>Myxoid liposarcoma</td>
<td>t(12;16)(q13;p11)</td>
<td><strong>CHOP-FUS</strong></td>
</tr>
<tr>
<td></td>
<td>t(12;22)(q13;q12)</td>
<td><strong>EWS-CHOP</strong></td>
</tr>
<tr>
<td>Well-differentiated liposarcoma</td>
<td>Ring chromosome 12</td>
<td>–</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Tumor</th>
<th>Common chromosome anomalies</th>
<th>Genes involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ewing’s sarcoma/primitive neuroectodermal tumor</td>
<td>t(11;22)(q24;q12), t(21;22)(q22;q12)</td>
<td>EWS-FLI1, EWS-ERG</td>
</tr>
<tr>
<td>Desmoplastic small round cell tumor</td>
<td>t(11;22)(p13;q12)</td>
<td>EWS-WT1</td>
</tr>
<tr>
<td>Synovial sarcoma</td>
<td>t(X;18)(p11;q11), t(X;20)</td>
<td>SYT-SSX1</td>
</tr>
<tr>
<td>Infantile fibrosarcoma and congenital mesoblastic nephroma</td>
<td>t(12;15)(p13;q25)</td>
<td>ETV6-NTRK3</td>
</tr>
<tr>
<td>Inflammatory myofibroblastic tumor</td>
<td>t(1;2)(q22;p23) t(2;19)(p23;p13)</td>
<td>TPM3-ALK, TPM4-ALK</td>
</tr>
<tr>
<td>Gastrointestinal stromal tumor</td>
<td>4q11-21 c-kit exon 11</td>
<td></td>
</tr>
<tr>
<td>Hemangiopericytoma</td>
<td>t(12;19)</td>
<td></td>
</tr>
<tr>
<td>Uterine leiomyoma</td>
<td>t(12;14)(q13-15;q24.1)</td>
<td>HMGIC</td>
</tr>
<tr>
<td>Endometrial stromal sarcoma</td>
<td>t(7;17)(p15-p21;q12-q21)</td>
<td>JAZF1-JJAzi</td>
</tr>
<tr>
<td>Leiomyosarcoma</td>
<td>del(1p)</td>
<td></td>
</tr>
<tr>
<td>Pleomorphic adenoma</td>
<td>t(3;8)(p12;q12)</td>
<td>FGFR1-FIM</td>
</tr>
<tr>
<td>Aneurysmal bone cyst</td>
<td>17p rearrangement</td>
<td></td>
</tr>
<tr>
<td>Desmoplastic fibroblastoma and fibroma of tendon sheath</td>
<td>t(2;11)</td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td>del(9p21), 12q14, del(22q)</td>
<td>CDKN2, CDK4,</td>
</tr>
<tr>
<td>Neural/neuroendocrine tumors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acoustic neuroma</td>
<td>22q12.2</td>
<td>NF2</td>
</tr>
<tr>
<td>Schwannoma</td>
<td>del(22q13)</td>
<td>NF2</td>
</tr>
<tr>
<td>Meningioma</td>
<td>del(22q11-q13), Monosomy 22</td>
<td>NF2</td>
</tr>
<tr>
<td>Glioblastoma multiforme</td>
<td>Trisomy 7</td>
<td>EGFR gene amplification (7p11)</td>
</tr>
<tr>
<td>Medulloblastoma</td>
<td>del(10q), i(17q), del(17p13.2), Trisomy 8</td>
<td>PTEN and DMBTI (10q)</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>2p del(1p31-32)</td>
<td>N-Myc amplification (homogenous staining region and double minutes)</td>
</tr>
<tr>
<td>Tumor</td>
<td>Common chromosome anomalies</td>
<td>Genes involved</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Oligodendroglioma</td>
<td>del(1p36)</td>
<td>TP73</td>
</tr>
<tr>
<td></td>
<td>del(19q13)</td>
<td>PEG3</td>
</tr>
<tr>
<td>Retinoblastoma</td>
<td>del(13q14)</td>
<td>RB</td>
</tr>
<tr>
<td>Pheochromocytoma</td>
<td>del(22q13)</td>
<td>SLC1</td>
</tr>
<tr>
<td></td>
<td>del(1p11-36)</td>
<td>RIZ1</td>
</tr>
<tr>
<td><strong>Lymphomas/leukemia</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burkett’s lymphoma</td>
<td>t(8;14)(q24;q32),</td>
<td>c-myc-IgH</td>
</tr>
<tr>
<td></td>
<td>t(8;22)(q24;q11)</td>
<td>c-myc-IgL</td>
</tr>
<tr>
<td></td>
<td>t(2;8)(p12q24)</td>
<td>c-myc-IgK</td>
</tr>
<tr>
<td>Follicular lymphoma</td>
<td>t(14;18)(q32;q21)</td>
<td>IgH-BCL2</td>
</tr>
<tr>
<td>Mantle cell lymphoma</td>
<td>t(11;14)(q13;q32) 11q</td>
<td>IgH-BCL1 (cyclin D1)</td>
</tr>
<tr>
<td></td>
<td>del(13q14)</td>
<td>ATM mutation or deletion</td>
</tr>
<tr>
<td></td>
<td>Trisomy 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>del(17p)</td>
<td></td>
</tr>
<tr>
<td>Mucosa-associated lymphoid</td>
<td>t(11;18)(q21;q21)</td>
<td>API2-MALT1</td>
</tr>
<tr>
<td>tissue (MALT) lymphoma</td>
<td>Trisomy 3</td>
<td></td>
</tr>
<tr>
<td>Diffuse large cell lymphoma</td>
<td>t(14;18)(q32;q21)</td>
<td>IgH-BCL2</td>
</tr>
<tr>
<td></td>
<td>t(3;14)(q27;q32)</td>
<td>BCL6-IgH</td>
</tr>
<tr>
<td>Anaplastic large cell lymphoma</td>
<td>t(2;5)(p23;q35)</td>
<td>NPM-ALK</td>
</tr>
<tr>
<td>Lymphoplastomacytic lymphoma</td>
<td>t(9;14)(p13;q32)</td>
<td>PAX5</td>
</tr>
<tr>
<td>Myelodysplastic disorder</td>
<td>del(5q)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trisomy 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monosomy 7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>del(7q)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>del(17p)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>del(20q)</td>
<td></td>
</tr>
<tr>
<td>Chronic myelogenous leukemia</td>
<td>t(9;22)(q34;q11)</td>
<td>BCR-ABL</td>
</tr>
<tr>
<td>Acute myelogenous leukemia (AML)-M2</td>
<td>t(8;21)(q22;q22)</td>
<td>AML1-ETO</td>
</tr>
<tr>
<td>Acute promyelocytic leukemia –M3</td>
<td>t(15;17)(q22;q21)</td>
<td>PML-RARα</td>
</tr>
<tr>
<td>AML-M4eo</td>
<td>inv(16)(p13;p22)</td>
<td>MYH11-CBFB</td>
</tr>
<tr>
<td>AML-M4, M5</td>
<td>t(11;19) 11q23</td>
<td>MLL-ENL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLL</td>
</tr>
</tbody>
</table>

(Continued)