
2 An Overview of Nucleic Acid Chemistry, Structure, and Function

The Foundations of Molecular Biology

WILLIAM B. COLEMAN

1. INTRODUCTION

Chemists and early biochemists determined the essential building blocks of living cells and characterized their chemical nature. Among these building blocks were nucleic acids, long-chain polymers composed of nucleotides. Nucleic acids were named based partly on their chemical properties and partly on the observation that they represent a major constituent of the cell nucleus. That nucleic acids form the chemical basis for the transmission of genetic traits was not realized until about 60 years ago (1,2). Prior to that time, there was considerable disagreement among scientists as to whether genetic information was contained in and transmitted by proteins or nucleic acids. It was recognized that chromosomes contained deoxyribonucleic acid as a primary constituent, but it was not known if this DNA carried genetic information or merely served as a scaffold for some undiscovered class of proteins that carried genetic information. However, the demonstration that genetic traits could be transmitted through DNA formed the basis for numerous investigations focused on elucidation of the nature of the genetic code. During the last half-century, numerous investigators have participated in the scientific revolution leading to modern molecular biology. Of particular significance were the elucidation of the structure of DNA (3–9), determination of structure–function relationships between DNA and RNA (10,11), and acquisition of basic insights into the processes of DNA replication, RNA transcription, and protein synthesis (12–19). Molecular pathology represents the application of the principles of basic molecular biology to the investigation of human disease processes. Our ever broadening insights into the molecular basis of disease processes continues to provide an opportunity for the clinical laboratory to develop and implement new and novel approaches for diagnosis and prognostic assessment of human disease.

2. THE CENTRAL DOGMA OF MOLECULAR BIOLOGY

Molecular biology has developed into a broad field of scientific pursuit and, at the same time, has come to represent a basic

component of most other basic research sciences. This has come about through the rapid expansion of our insights into numerous basic aspects of molecular biology and the development of an understanding of the fundamental interaction among the several major processes that comprise the larger field of investigation. A theory, referred to as the “central dogma,” describes the interrelationships among these major processes (20,21). The central dogma defines the paradigm of molecular biology that genetic information is perpetuated as sequences of nucleic acid, but that genes function by being expressed in the form of protein molecules (20). The flow of genetic information among DNA, RNA, and protein that is described by the central dogma is illustrated in Fig. 1. Individual DNA molecules serve as templates for either complementary DNA strands during the process of replication or complementary RNA molecules during the process of transcription. In turn, RNA molecules serve as blueprints for the ordering of amino acids by ribosomes during protein synthesis or translation. This simple representation of the complex interactions and interrelationships among DNA, RNA, and protein was proposed and commonly accepted shortly after the discovery of the structure of DNA. Nonetheless, this paradigm still holds more than 45 years later and continues to represent a guiding principle for molecular biologists involved in all areas of basic biological, biomedical, and genetic research.

3. CHEMICAL NATURE OF DNA

Deoxyribonucleic acid is a polymeric molecule that is composed of repeating nucleotide subunits. The order of nucleotide subunits contained in the linear sequence or primary structure of these polymers represents all of the genetic information carried by a cell. Each nucleotide is composed of (1) a phosphate group, (2) a pentose (5 carbon) sugar, and (3) a cyclic nitrogen-containing compound called a base. In DNA, the sugar moiety is 2-deoxyribose. Eukaryotic DNA is composed of four different bases: adenine, guanine, thymine, and cytosine. These bases are classified based on their chemical structure into two groups:

From: *Molecular Diagnostics: For the Clinical Laboratorian, Second Edition*
Edited by: W. B. Coleman and G. J. Tsongalis © Humana Press Inc., Totowa, NJ

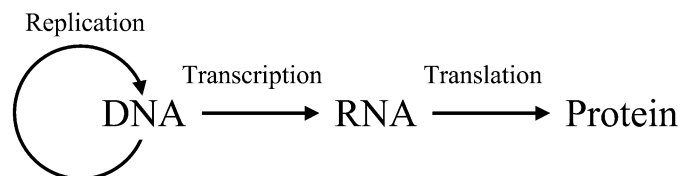


Fig. 1. The central dogma of molecular biology. The central dogma defines the paradigm of molecular biology that genetic information is perpetuated as sequences of nucleic acid, but that genes function by being expressed in the form of protein molecules. (From ref. 20.)

Adenine and guanine are double-ring structures termed purines, and thymine and cytosine are single-ring structures termed pyrimidines (Fig. 2). Within the overall composition of DNA, the concentration of thymine is always equal to the concentration of adenine, and the concentration of cytosine is always equal to guanine (22,23). Thus, the total concentration of pyrimidines always equals the total concentration of purines. These monomeric units are linked together into the polymeric structure by 3', 5'-phosphodiester bonds (Fig. 3). Natural DNAs display widely varying sizes depending on the source. Relative molecular weights range from 1.6×10^6 Daltons for bacteriophage DNA to 1×10^{11} Daltons for a human chromosome.

4. STRUCTURE OF DNA

The structure of DNA is a double helix, composed of two polynucleotide strands that are coiled about one another in a spiral (3,4). Each polynucleotide strand is held together by phosphodiester bonds linking adjacent deoxyribose moieties. The two polynucleotide strands are held together by a variety of noncovalent interactions, including lipophilic interactions between adjacent bases and hydrogen-bonding between the bases on opposite strands. The sugar-phosphate backbones of the two complementary strands are antiparallel; that is, they possess opposite chemical polarity. As one moves along the DNA double helix in one direction, the phosphodiester bonds in one strand will be oriented 5'–3', whereas in the complementary strand, the phosphodiester bonds will be oriented 3'–5'. This configuration results in basepairs being stacked between the two chains perpendicular to the axis of the molecule. The base-pairing is always specific: Adenine is always paired to thymidine, and guanine is always paired to cytosine. This specificity results from the hydrogen-bonding capacities of the bases themselves. Adenine and thymine form two hydrogen bonds, and guanine and cytosine form three hydrogen bonds. The specificity of molecular interactions within the DNA molecule allows one to predict the sequence of nucleotides in one polynucleotide strand if the sequence of nucleotides in the complementary strand is known (24). Although the hydrogen bonds themselves are relatively weak, the number of hydrogen bonds within a DNA molecule results in a very stable molecule that would never spontaneously separate under physiological conditions. There are many possibilities for hydrogen-bonding between pairs of heterocyclic bases. Most important are the hydrogen-bonded basepairs A:T and G:C that were proposed by Watson and Crick in their double-helix structure of DNA (3,24). However, other forms of base-pairing have been described (25,26). In addition, hydrophobic interactions between the stacked bases in the double helix lend additional stability to the DNA molecule.

Three helical forms of DNA are recognized to exist: A, B, and Z (27). The B conformation is the dominate form under physiological conditions. In B DNA, the basepairs are stacked 0.34 nm apart, with 10 basepairs per turn of the right-handed double helix and a diameter of approx 2 nm. Like B DNA, the A conformer is also a right-handed helix. However, A DNA exhibits a larger diameter (2.6 nm), with 11 bases per turn of the helix, and the bases are stacked closer together in the helix (0.25 nm apart). Careful examination of space-filling models of A and B DNA conformers reveals the presence of a major groove and a minor groove (27). These grooves (particularly the minor groove) contain many water molecules that interact favorably with the amino and keto groups of the bases. In these grooves, DNA-binding proteins can interact with specific DNA sequences without disrupting the base-pairing of the molecule. In contrast to the A and B conformers of DNA, Z DNA is a left-handed helix. This form of DNA has been observed primarily in synthetic double-stranded oligonucleotides, especially those with purine and pyrimidines alternating in the polynucleotide strands. In addition, high salt concentrations are required for the maintenance of the Z DNA conformer. Z DNA possesses a minor groove but no major groove, and the minor groove is sufficiently deep that it reaches the axis of the DNA helix. The natural occurrence and potential physiological significance of Z DNA in living cells has been the subject of much speculation. However, these issues with respect to Z DNA have not yet been fully resolved.

5. SEQUENCE OF THE HUMAN GENOME

The diploid genome of the typical human cell contains approx 3×10^9 basepairs of DNA that is subdivided into 23 pairs of chromosomes (22 autosomes and sex chromosomes X and Y). It has long been suggested that discerning the complete sequence of the human genome would enable the genetic causes of human disease to be investigated (28–30). Practical methods for DNA sequencing appeared in the mid to late 1970s (31–33), and numerous reports of DNA sequences corresponding to segments of the human genome began to appear. In the mid-1980s, a project to sequence the complete human genome was proposed, and this project began in the later years of that decade. The development of automated methods for DNA sequencing (34,35) made the ambitious goals of the Human Genome Project attainable (36,37). Subsequently, detailed genetic and physical maps of the human genome appeared (38–43), expressed sequences were identified and characterized (44–46), and gene maps of the human genome were constructed (47,48). Efforts by several consortiums using differing approaches (49–51) to large-scale sequencing of human DNA and sequence contig assembly culminated in 2001 with the publication of a draft sequence of the human genome (52,53). The actual number of genes contained in the human genome is not yet known. Early estimates suggested that the human genome might contain 70,000 to 100,000 genes (54–56). However, more recently, the number of genes contained in the human genome has been estimated to be approximately 30,000–40,000 (52,53,57,58). Early analysis of the draft sequences of the human genome revealed considerable variability between individuals, including in excess of 1.1–1.4 million single-nucleotide polymorphisms (SNPs) distributed throughout

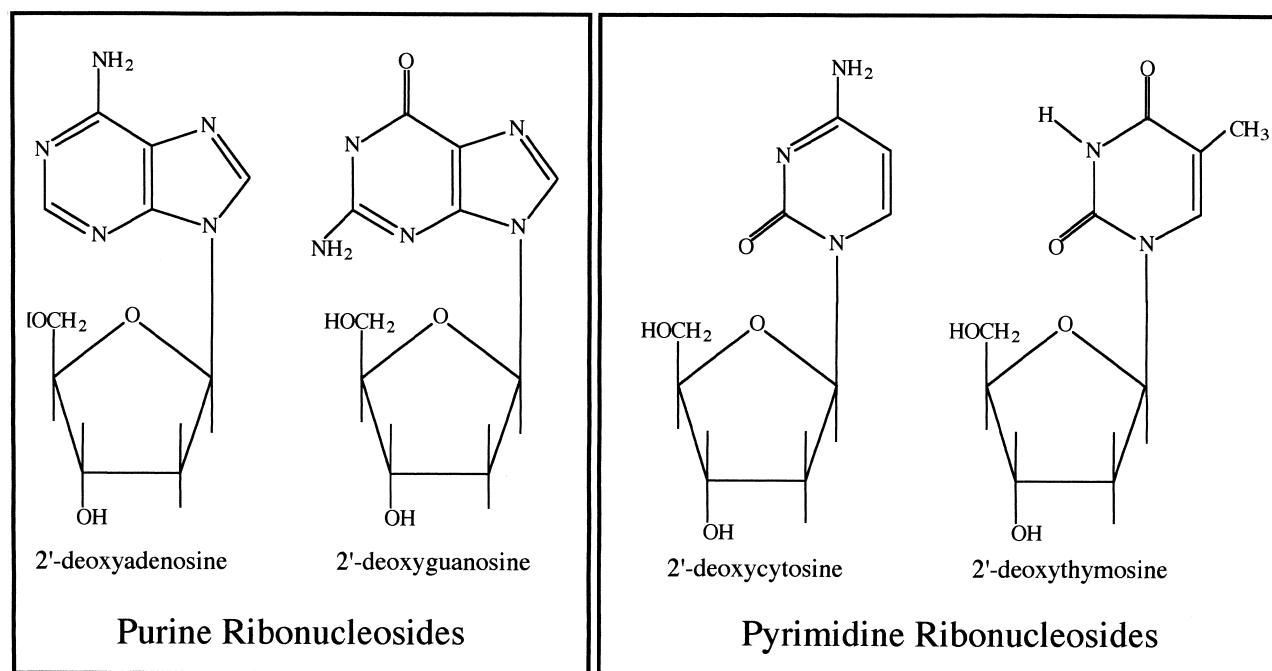


Fig. 2. The chemical structure of purine and pyrimidine deoxyribonucleosides.

the genome (53,59). Refinement of the human genome sequence, identification and characterization of the genes contained, and description of the features of the genome (SNPs and other variations) continues at a rapid pace (<http://www.nhgri.nih.gov/>). In early 2003, it was announced that the Human Genome Project had completely sequenced 99% of the gene-containing portion of the human genome, and new plans and goals for genomics research were described (60,61).

6. ORGANIZATION OF GENOMIC DNA

The human genomic DNA is packaged into discrete structural units that vary in size and genetic composition. The structural unit of DNA is the chromosome, which is a large continuous segment of DNA (62). A chromosome represents a single genetically specific DNA molecule to which are attached a large number of protein molecules that are involved in the maintenance of chromosome structure and regulation of gene expression (63). Genomic DNA contains both “coding” and “noncoding” sequences. Noncoding sequences contain information that does not lead to the synthesis of an active RNA molecule or protein (54,64). This is not to suggest that noncoding DNA serves no function within the genome. On the contrary, noncoding DNA sequences have been suggested to function in DNA packaging, chromosome structure, chromatin organization within the nucleus, or in the regulation of gene expression (65,66). A portion of the noncoding sequences represent intervening sequences that split the coding regions of structural genes. However, the majority of noncoding DNA falls into several families of repetitive DNA whose exact functions have not been entirely elucidated (67,68).

Coding DNA sequences give rise to all of the transcribed RNAs of the cell, including mRNA. The organization of transcribed structural genes consists of coding regions that are interrupted by intervening noncoding regions of DNA (Fig. 4). Thus,

the primary RNA transcripts contain both coding and noncoding sequences. The noncoding sequences must be removed from the primary RNA transcript during processing to produce a functional mRNA molecule appropriate for translation.

7. DNA FUNCTION

DNA serves two important functions with respect to cellular homeostasis: the storage of genetic information and the transmission of genetic information. In order to fulfill both of these functions, the DNA molecule must serve as a template. The cellular DNA provides the source of information for the synthesis of all the proteins in the cell. In this respect, DNA serves as a template for the synthesis of RNA. In cell division, DNA serves as the source of information inherited by progeny cells. In this case, DNA serves as a template for the faithful replication of the genetic information that is ultimately passed into daughter cells.

7.1. TRANSCRIPTION OF RNA Contained within the linear nucleotide sequence of the cellular DNA is the information necessary for the synthesis of all the protein constituents of a cell (Table 1). Transcription is the process in which mRNA is synthesized with a sequence complementary to the DNA of a gene to be expressed. The correct start and end points for transcription of a specific gene are identified in the DNA by a promoter sequence upstream of the gene and a termination signal downstream (Fig. 4). In the case of RNA transcription, only one strand of the DNA molecule serves as a template. This strand is referred to as the “sense” strand. Transcription of the sense strand ultimately yields a mRNA molecule that encodes the proper amino acid sequence for a specific protein.

7.2. REPLICATION OF DNA The double-stranded model of the structure of DNA strongly suggests that replication of the DNA can be achieved in a semiconservative manner (69–72). In semiconservative replication, each strand of the

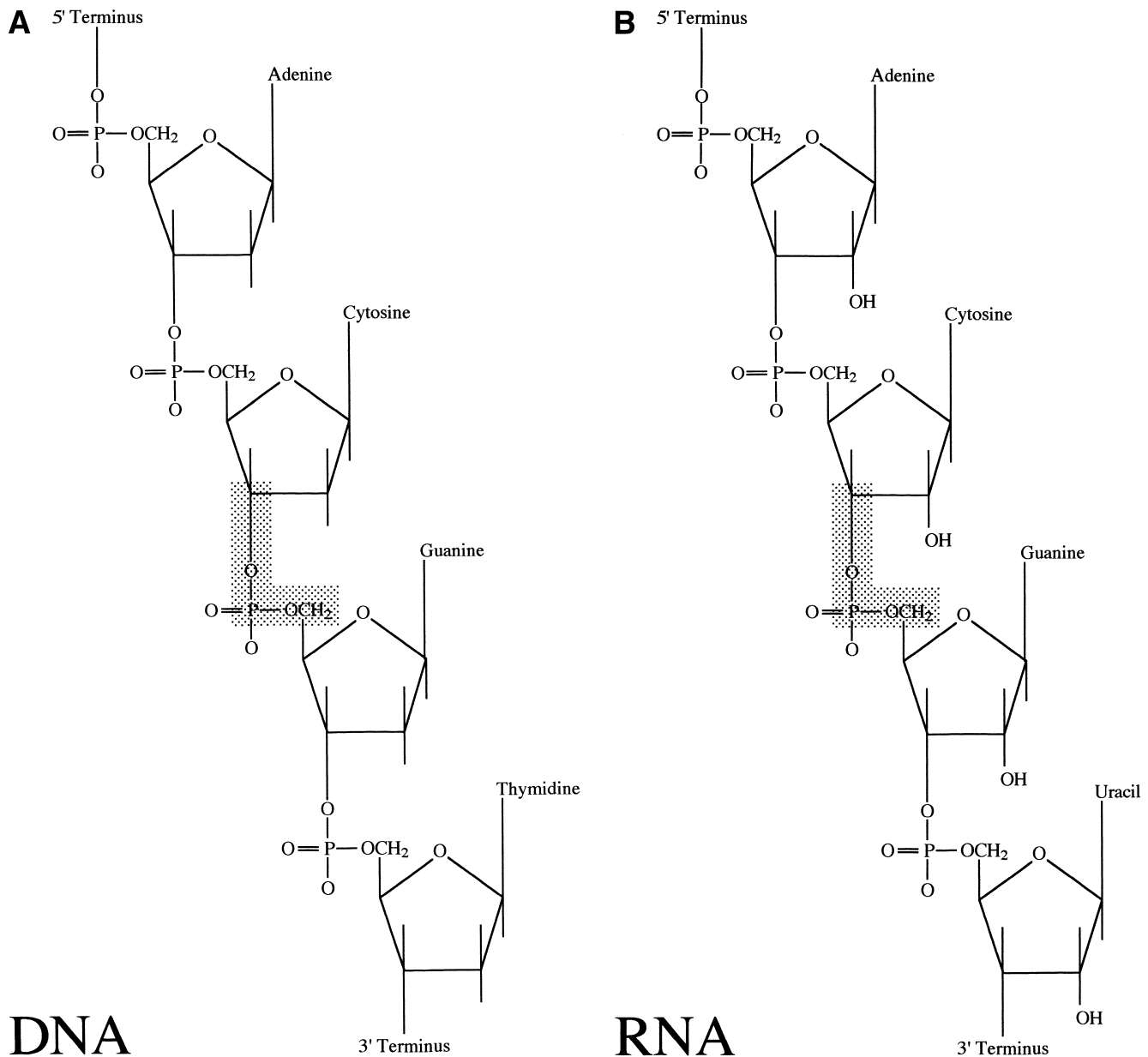


Fig. 3. The chemical structure of repeating nucleotide subunits in DNA and RNA. Each panel shows the sugar-phosphate backbone of a single polynucleotide strand of nucleic acid composed of four nucleotide subunits. The stippled area highlights a 3'-5' phosphodiester bond.

DNA helix serves as a template for the synthesis of complementary DNA strands. The result is the formation of two complete copies of the DNA molecule, each consisting of one strand derived from the parent DNA molecule and one newly synthesized complementary strand. The utilization of the DNA strands as the template for the synthesis of new DNA ensures the faithful reproduction of the genetic material for transmission into daughter cells (19).

7.3. GENETIC RECOMBINATION Genetic recombination represents one mechanism for the generation of genetic diversity through the exchange of genetic material between two homologous nucleotide sequences (73,74). Such an exchange of genetic material often results in alterations of the primary structure (nucleotide sequence) of a gene and, subsequently, alteration of the primary structure of the encoded protein product. In

organisms that reproduce sexually, recombination is initiated by formation of a junction between similar nucleotide sequences carried on the same chromosome from the two different parents. The junction is able to move along the DNA helix through branch migration, resulting in an exchange of the DNA strands.

7.4. DNA REPAIR Maintenance of the integrity of the informational content of the cellular DNA is absolutely required for cellular and organismal homeostasis (75,76). The cellular DNA is continuously subjected to structural damage through the action of endogenous or environmental mutagens. In the absence of efficient repair mechanisms, stable mutations can be introduced into DNA during the process of replication at damaged sites within the DNA. Mammalian cells possess several distinct DNA repair mechanisms and pathways that serve to maintain DNA integrity, including enzymatic reversal repair,

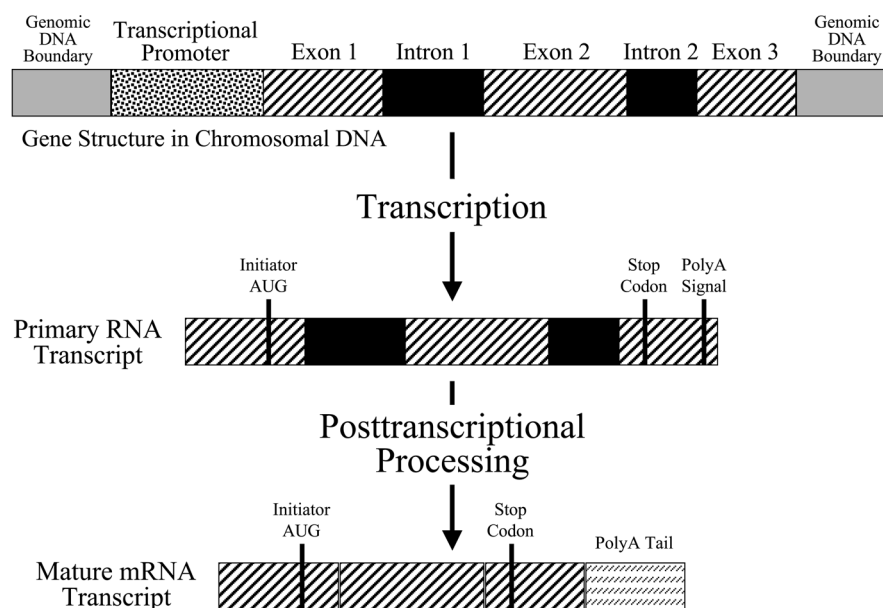


Fig. 4. Basic organization of a structural gene in DNA and biogenesis of mature mRNAs.

Table 1
The Universal Genetic Code

First position (5' end)	Second position				Third position (3' end)
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
U	Phe	Ser	Tyr	Cys	C
U	Leu	Ser	Stop	Stop	A
U	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
C	Leu	Pro	His	Arg	C
C	Leu	Pro	Gln	Arg	A
C	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
A	Ile	Thr	Asn	Ser	C
A	Ile	Thr	Lys	Arg	A
A	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
G	Val	Ala	Asp	Gly	C
G	Val	Ala	Glu	Gly	A
G	Val	Ala	Glu	Gly	G

nucleotide excision repair, and postreplication repair (77,78). Several steps in the process of DNA repair are shared by these multiple pathways, including (1) recognition of sites of damage, (2) removal of damaged nucleotides, and (3) restoration of the normal DNA sequence. Each of these steps in DNA repair are accomplished by specific proteins and enzymes. Surveillance of the cellular DNA is a continual process involving specific aspects of the transcription and replication machinery. However, in each case where the restoration of the normal DNA sequence is accomplished through the replacement of damaged nucleotides, the undamaged DNA strand serves as a template in the repair process. This ensures the faithful reproduction of the primary structure of the DNA at the damaged site.

8. CHEMICAL NATURE OF RNA

Like DNA, RNA is composed of repeating purine and pyrimidine nucleotide subunits. However, several distinctions can be made with respect to the chemical nature of RNA and DNA. Unlike the 2'-deoxyribose sugar moiety of DNA, the sugar moiety in RNA is ribose. Like DNA, RNA usually contains adenine, guanine, and cytosine, but does not contain thymidine. In place of thymidine, RNA contains uracil. The concentration of purines and pyrimidine bases do not necessarily equal one another in RNA because of the single-stranded nature of the molecule. The monomeric units of RNA are linked together by 3',5'-phosphodiester bonds analogous to those in DNA (Fig. 3). RNAs have molecular weights between 1×10^4 Daltons for transfer RNA (tRNA) and 1×10^7 Daltons for ribosomal RNA (rRNA).

9. STRUCTURE AND FUNCTION OF RNA

RNA exists as a long, regular, unbranched polynucleotide strand. The informational content of the RNA molecule is contained in its primary structure or nucleotide sequence. In spite of the fact that RNA exists primarily as a single-stranded molecule, significant higher-order structures are often formed in individual RNA molecules. In some cases, this higher-order structure is related to the actual function of the molecule. Three major classes of RNA are found in eukaryotic organisms: messenger RNA (mRNA), transfer RNA (tRNA), and ribosomal RNA (rRNA). Each class differs from the others in the size, function, and general stability of the RNA molecules (79). Minor classes of RNA include heterogeneous nuclear RNA (hnRNA), small nuclear RNA (snRNA), and small cytoplasmic RNA (scRNA).

9.1. STRUCTURE AND FUNCTION OF MESSENGER RNA The ability of DNA to serve directly as a template for the synthesis of protein is precluded by the observations that protein synthesis takes place in the cytoplasm, whereas almost all of the cellular DNA resides in the nucleus. Thus, genetic information contained in the DNA must be transferred to an

intermediate molecule that is translocated into the cell cytoplasm, where it directs the ordering of amino acids in protein synthesis. RNA fulfills this role as the intermediate molecule for the transport and translation of genetic information (11). Messenger RNA molecules represent transcripts of structural genes that encode all of the information necessary for the synthesis of a single-type polypeptide of protein. Thus, mRNAs serve two important functions with respect to protein synthesis: (1) mRNAs deliver genetic information to the cytoplasm where protein synthesis takes place and (2) mRNAs serve as a template (or blueprint) for translation by ribosomes during protein synthesis.

Mammalian cells (among others) express “interrupted” genes; that is, genes with coding sequences are not contiguous (continuous) in the DNA, and that require a posttranscriptional modification prior to translation of protein products (Fig. 4). The majority of structural genes in the higher eukaryotic organisms are interrupted. The average gene contains 7–10 exons, spread over 10–20 kb of DNA. For instance, the *p53* tumor suppressor gene is composed of 11 exons, occupies approx 16 kb in the genomic DNA, and produces a 2-kb mRNA (80–82). However, other genes are much larger. For example, the *Rb1* tumor suppressor gene occupies 200 kb in the genomic DNA, contains 27 exons, and gives rise to a 4.7-kb mRNA (83–85). The primary RNA transcript exhibits the same overall structure and organization as the structural gene and is often referred to as the pre-mRNA. Removal of intronic sequences yields a mature mRNA that is considerably smaller with an average size of 1–3 kb. The process of removing the intronic sequences is called RNA splicing (86–88).

The primary products of RNA transcription in the nucleus compose a special class of RNAs that are characterized by their large size and heterogeneity (79). These RNA molecules are referred to as heterogeneous nuclear RNAs (hnRNAs). Heterogeneous nuclear RNAs contain both intronic and exonic sequences encoded in the template DNA of structural genes. These hnRNAs are processed in the nucleus (87,89,90) to give mature mRNAs that are transported into the cytoplasm, where they participate in protein synthesis. Nuclear processing of RNA involves (1) chemical modification reactions (addition of the 5' CAP), (2) splicing reactions (removal of intronic sequences), and (3) polyadenylation [addition of the 3' poly(A) tail]. Additional processing of some specific mRNAs occurs in the cell cytoplasm, including RNA editing reactions (91–93). It has been suggested that some snRNAs function in the processing of hnRNAs (94). Mature mRNAs are transported into the cytoplasm of the cell, where they participate in the translational processes of protein synthesis.

In RNA splicing, intronic sequences are specifically removed from the primary RNA transcript and the remaining exonic sequences are rejoined into one molecule. There is no extensive homology or complementarity between the two ends of an intron precluding the general possibility that intronic sequences form extensive secondary structures (such as a hairpin loop) as a preliminary step in the splicing reaction. The splice junctions represent short, well-conserved consensus sequences. The generic intron contains a GT sequence at the 5' boundary and an AG sequence at the 3' boundary (Fig. 5). The 5' and 3' splice junctions are often referred to as the splice donor and splice acceptor sites, respectively. Splice sites are generic in that they do not

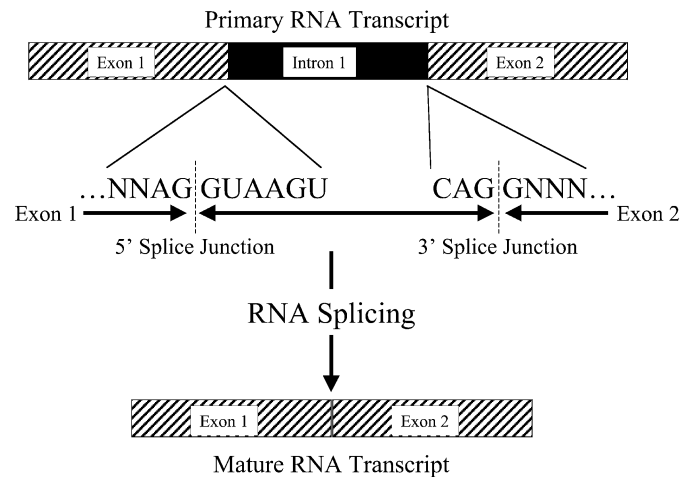


Fig. 5. Fundamental aspects of RNA splicing.

exhibit specificity for individual RNA precursors, and individual precursors do not convey specific information that is required for splicing. In principle, any splice donor site can react with any splice acceptor site. However, under normal conditions, these reactions are restricted to the donor and acceptor sites of the same intron. Analysis of molecular intermediates formed during the splicing of large precursor RNAs suggests that the introns are removed in a definitive pattern or through a preferred pathway that dictates the general order of intron removal. This suggestion might imply a mechanism in which the conformation of precursor RNA molecules limits the accessibility of splice junctions, such that as specific introns are removed, the conformation of the molecule changes and new splice sites become available. However, several other plausible mechanisms exist that could account for the observed patterns of splicing in RNAs that have been examined in detail (86,89,90). The RNA sequences that are required for successful splicing include (1) the 5' splice donor and 3' splice acceptor consensus sequences and (2) a consensus sequence known as the branch site. The branch site is located approx 20–40 bases upstream of the 3'-terminus of the intronic sequence and conforms to the following consensus sequence: Py-N-Py-Py-Pu-A-Py. The role of the branch site is to identify the nearest splice acceptor site for connection to the splice donor site. The first step in splicing involves a cleavage of the RNA molecule at the 5' end of the intron. The resulting intron–exon molecule forms a structure known as a lariat, through the formation of a 5'–2' bond between the G residue located at the 5' end of the intron and the adenine residue of the branch site. The second step involves cleavage of the RNA molecule at the 3' splice site. This cleavage releases the intron in lariat form, which is subsequently degraded.

The 5'-termini of all eukaryotic mRNAs are modified posttranscriptionally, through enzymatic reactions occurring in the nucleus and in the cytoplasm of the cell. Initially, a guanine residue is added to the 5'-termini of primary mRNA transcripts through the action of an enzyme called guanylyl transferase. This reaction occurs in the nucleus soon after the initiation of transcription. This guanine residue is linked to the first coded

nucleoside triphosphate through a 5′–5′ triphosphate linkage, rather than a 3′–5′ phosphodiester bond. Thus, this guanine residue occurs in the structure of the mRNA in reverse orientation from all the other nucleotides. The modified 5′-terminus is referred to as a “CAP” and is the site of additional modification reactions involving the addition of methyl groups. These additional modifications are catalyzed by enzymes located in the cytoplasm of the cell. The first methylation occurs in all mRNAs and consists of the addition of a methyl group to the 7-position of the terminal guanine through the action of guanine-7-methyltransferase. Additional methylation reactions can occur involving the additional bases at the 5′-terminus of the mRNA transcript, and less frequently, internal methylation of bases within an mRNA molecule takes place.

The poly(A) tail possessed by most eukaryotic mRNAs is not encoded in the DNA; rather, it is added to the RNA in the nucleus after transcription of the structural gene is complete. The addition of the poly(A) tail is catalyzed by the enzyme poly(A) polymerase, which adds approx 200 adenosine residues to the free 3′-OH terminus of the primary RNA transcript. The precise function of the poly(A) tail is unknown, but has been speculated to be involved in mRNA stability or in control of mRNA utilization (95). Although the removal of the poly(A) tail does precede degradation of certain mRNAs, a systematic correlation between mRNA stability or survival, and the length or presence of the poly(A) tail has not been established. Removal of the poly(A) tail can inhibit the initiation of translation *in vitro*, suggesting a potential role for this structure in the control of mRNA translation. However, it is not clear whether this effect is related to a direct influence of poly(A) structures on the initiation reaction or the result of some indirect cause.

Some other forms of posttranscriptional RNA processing occur with respect to a small subset of eukaryotic mRNAs. The process of RNA editing involves a posttranscriptional alteration of the informational content of a specific mRNA (91). The editing of RNA is revealed when the linear sequence of the mRNA molecule differs from the coding sequence carried in the DNA. In mammalian cells, there are examples where substitution of a single base occurs in the mRNA, resulting in an alteration of an amino acid in the final protein product. Because no known template source mediates the RNA editing reaction, the most likely mechanism for mRNA editing would involve a specific enzyme that can recognize the sequence or secondary structure of the specific target mRNA and catalyze the specific base substitution. However, there are examples of RNA editing in some lower eukaryotes that utilize “guide RNA,” which directs the RNA editing reaction (96). The final result of RNA editing is the generation protein products representing more than one polypeptide sequence from a single coding gene. The different protein products might possess different biological activities, suggesting that RNA editing might represent a mechanism for controlling the functional expression of genes through a post-transcriptional process that does not impact the normal mechanisms for controlling levels of gene expression.

9.2. STRUCTURE AND FUNCTION OF TRANSFER RNA Transfer RNAs are small molecules consisting of approx 75–80 nucleotides. Like mRNAs, tRNA is generated

through nuclear processing of precursor RNA transcripts. The structure of the tRNA molecule reflects its function as an adapter between the mRNA and amino acids during protein synthesis. Specific tRNAs correspond to each of the amino acids utilized in protein synthesis in any particular cell type. Although the specific tRNAs differ from each other with respect to their actual nucleotide sequence, tRNAs as a class of RNA molecules share several common structural features. Each tRNA contains information in the primary structure or nucleotide sequence that dictates the higher-order structure of the molecule. The secondary structure of the tRNA resembles a cloverleaf (97). The folding of the cloverleaf structure is maintained through intrastrand sequence complementarity and base-pairing interactions between nucleotides. In addition, each tRNA contains an ACC sequence at the 3′-terminus and an anticodon loop (97). Amino acids are attached to their specific tRNA through an ester bond to the 3′-hydroxyl group of the terminal adenine of the ACC sequence. The anticodon loop recognizes the triplet codon of the template mRNA during the process of translation. With the exception of the codons encoding methionine and tryptophan, there are at least two possible codons for each amino acid (Table 1). Nonetheless, each amino acid has only one corresponding tRNA. Thus, the hydrogen-bonding between nucleotides of the codon and anticodon often involve “wobble” pairing (98). This form of base-pairing allows mismatches in the third base of a codon triplet. In the overall structure–function relationship in tRNA, the nucleotide sequence of the anticodon loop dictates which amino acid will be attached to the ACC sequence of the tRNA.

Transfer RNAs serve as adapters between the mRNA template and the amino acids of growing polypeptide chains during the process of protein translation (97); that is, the tRNA serves to ferry the appropriate amino acid into the active site of the ribosome, where it becomes incorporated into the growing polypeptide being synthesized. Amino acids are coupled to their specific tRNA through the action of enzymes called aminoacyl tRNA synthetases. The specificity of the “charging” reaction is critical to the integrity of the translation process because the incorporation of amino acids at the level of the ribosome depends wholly on the sequence of the anticodon portion of the tRNA molecule. The charged tRNA the mRNA through the transient hybridization of the codon and anti-codon RNA sequences in the ribosome complex as it moves along the mRNA. The entry of the charged tRNA into the active site of the ribosome brings its associated amino acid into juxtaposition with the nascent polypeptide, facilitating the formation of a peptide bond. In this manner, the tRNA provides a link between the genetic information contained in the mRNA and the linear sequence of amino acids represented in the resulting polypeptide product.

9.3. STRUCTURE AND FUNCTION OF RIBOSOMAL RNA The ribosome is a nucleoprotein that serves as the primary component of a cell’s protein synthesis machinery. The ribosome is a complex structure consisting of two subunit particle types (60S and 40S). The overall composition of the fully assembled ribosome includes at least 4 distinct rRNA molecules and nearly 100 specific protein subunits. The major rRNAs in mammalian cells were named for their molecular size as determined by their sedimentation rates. Three of these rRNAs (5S,

Table 2
Secondary and Tertiary Structural Motifs in RNA

<i>Structure</i>	<i>Description</i>	<i>Functions</i>
Hairpin loops	Single-stranded loop that bridges one end of a double-stranded stem	Component of more complex RNA structures; May serve as a nucleation site for RNA folding, or recognition sites for protein–RNA interaction
Internal loops	Interruptions in double-stranded RNA caused by the presence of nucleotides on both strands that cannot participate in Watson–Crick base-pairing	Protein-binding sites and ribozyme cleavage sites
Bulges	Double-stranded RNA molecules with unpaired nucleotides on only one strand	Contribute to the formation of more complex RNA structures; recognition sites for protein–RNA interactions.
Nucleotide triples	Triple helical structures that form through hydrogen bonding between nucleotides of a single-stranded RNA molecule and nucleotides within a double-stranded RNA molecule; hydrogen bonds can involve nucleotide bases, sugars, or phosphate groups	Orient regions of secondary structure in large RNA molecules and stabilize three-dimensional RNA structures.
Pseudoknots	Results from base-pairing between nucleotide sequences within an RNA loop structure and a complementary nucleotide sequence outside the RNA loop	RNA self-splicing, autoregulation of translational processes, and ribosome frameshifting

5.8S, and 28S) are components of 60S ribosomal particle. The smaller 40S ribosomal particle contains a single 18S rRNA. The 5.8S, 18S, and 28S rRNAs are the products of the processing of a single 45S precursor RNA molecule. The 5S rRNA is independently transcribed and processed. The rRNAs assemble with ribosomal protein subunits in the nucleus. The precise role of rRNA in the function of the ribosome is not completely understood (99). However, it is recognized that interactions between the rRNAs of the ribosomal subunits might be important in the overall structure of the functioning ribosomal particle. In addition, rRNA sequences can interact with ribosome-binding sequences of mRNA during the initiation of translation. Likewise, it is likely that rRNAs bind to invariant tRNA sequences when these molecules enter the active site of the ribosome (99).

9.4. SPECIAL RNA STRUCTURES Higher-order RNA structures exhibit hydrogen-bonding between A:U and G:C basepairs. Several specific higher-order RNA structures have been recognized (Table 2) and characterized in detail (100). Hairpin loops consist of a double-stranded stem and a single-stranded loop that bridges one end of the stem. These structures are essential components of more complex RNA structures and probably serve as nucleation sites for RNA folding. Loops can also function as recognition sites for protein–RNA interactions. Internal loops represent interruptions in double-stranded RNA caused by the presence of nucleotides on both strands that cannot participate in Watson–Crick base-pairing. Several important functions are associated with internal loops, including protein-binding sites and ribozyme cleavage sites. In many cases, internal loops have been shown to represent highly ordered structures maintained by the formation of non-Watson–Crick basepairs. Bulges are structural motifs contained within double-stranded RNA molecules with unpaired nucleotides on only one strand. RNA bulges contribute to the formation of more complex, higher-order RNA structures and

can also serve as recognition sites for protein–RNA interaction. Nucleotide triples occur when single-stranded RNA sequences form hydrogen bonds with nucleotides that are already base-paired. These interactions serve to stabilize three-dimensional RNA structures and to orient regions of RNA secondary structure in large RNA molecules. Pseudoknots are tertiary structural elements that result from base-pairing between nucleotide sequences contained within a loop structure and sequences outside the loop structure. These are important in RNA self-splicing, translational autoregulation, and ribosomal frameshifting.

10. DNA DAMAGE, MUTAGENESIS, AND THE CONSEQUENCES OF MUTATION

DNA damage can result from spontaneous alteration of the DNA molecule or from the interaction of numerous chemical and physical agents with the structural DNA molecule (101). Spontaneous lesions can occur during normal cellular processes, such as DNA replication, DNA repair, or gene rearrangement (78), or through chemical alteration of the DNA molecule itself as a result of hydrolysis, oxidation, or methylation (102,103). In most cases, DNA lesions create nucleotide mismatches that lead to point mutations. Nucleotide mismatches can result from the formation of apurinic or apyrimidinic sites following depurination or depyrimidination reactions (103), nucleotide conversions involving deamination reactions (78), or, in rare instances, from the presence of a tautomeric form of an individual nucleotide in replicating DNA. Deamination reactions can result in the conversion of cytosine to uracil, adenine to hypoxanthine, and guanine to xanthine (78). However, the most common nucleotide deamination reaction involves methylated cytosines, which can replace cytosine in the linear sequence of a DNA molecule in the form of 5-methylcytosine. The 5-methylcytosine residues are always located next to guanine residues on the same chain, a motif

referred to as a CpG island. The deamination of 5'-methylcytosine results in the formation of thymine. This particular deamination reaction accounts for a large percentage of spontaneous mutations in human disease (104–106). Interaction of DNA with physical agents, such as ionizing radiation, can lead to single- or double-strand breaks as a result of scission of phosphodiester bonds on one or both polynucleotide strands of the DNA molecule (78). Ultraviolet (UV) light can produce different forms of photoproducts, including pyrimidine dimers between adjacent pyrimidine bases on the same DNA strand. Other minor forms of DNA damage caused by UV light include strand breaks and crosslinks (78). Nucleotide base modifications can result from exposure of the DNA to various chemical agents, including N-nitroso compounds and polycyclic aromatic hydrocarbons (78). DNA damage can also be caused by chemicals that intercalate the DNA molecule and/or crosslink the DNA strands (78). Bifunctional alkylating agents can cause both intrastrand and interstrand crosslinks in the DNA molecule.

The various forms of spontaneous and induced DNA damage can give rise to a plethora of different types of molecular mutation (107). These various types of mutation include both gross alteration of chromosomes and more subtle alterations to specific gene sequences in otherwise normal chromosomes. Gross chromosomal aberrations include (1) large deletions, (2) additions (reflecting amplification of DNA sequences), and (3) translocations (reciprocal and nonreciprocal). All of these forms of chromosomal abnormality can be distinguished through standard karyotype analyses of G-banded chromosomes (Fig. 6). The major consequence of chromosomal deletion is the loss of specific genes that are located in the deleted chromosomal region, resulting in changes in the copy number of the affected genes. The deletion of certain classes of genes such as tumor suppressor genes or genes encoding the proteins involved in DNA repair can predispose cells to neoplastic transformation (108,109). Likewise, amplification of chromosomal regions results in an increase in gene copy numbers, which can lead to the same type of circumstance if the affected region contains genes for dominant proto-oncogenes or other positive mediators of cell cycle progression and proliferation (108–110). The direct result of chromosomal translocation is the movement of some segment of DNA from its natural location into a new location within the genome, which can result in altered expression of the genes that are contained within the translocated region. If the chromosomal breakpoints utilized in a translocation are located within structural genes, then new hybrid genes can be generated.

The most common forms of mutation involve single-nucleotide alterations, small deletions, or small insertions into specific gene sequences. These microscopic alterations very often can only be detected through DNA sequencing. Single-nucleotide alterations that involve a change in the normal coding sequence of the gene are referred to as point mutations. The consequence of most point mutations is an alteration in the amino acid sequence of the encoded protein. However, some point mutations are “silent” and do not affect the structure of the gene product (Table 3). Silent mutations are possible because most amino acids can be encoded by more than one triplet codon (Table 1). Point mutations fall into two classes:

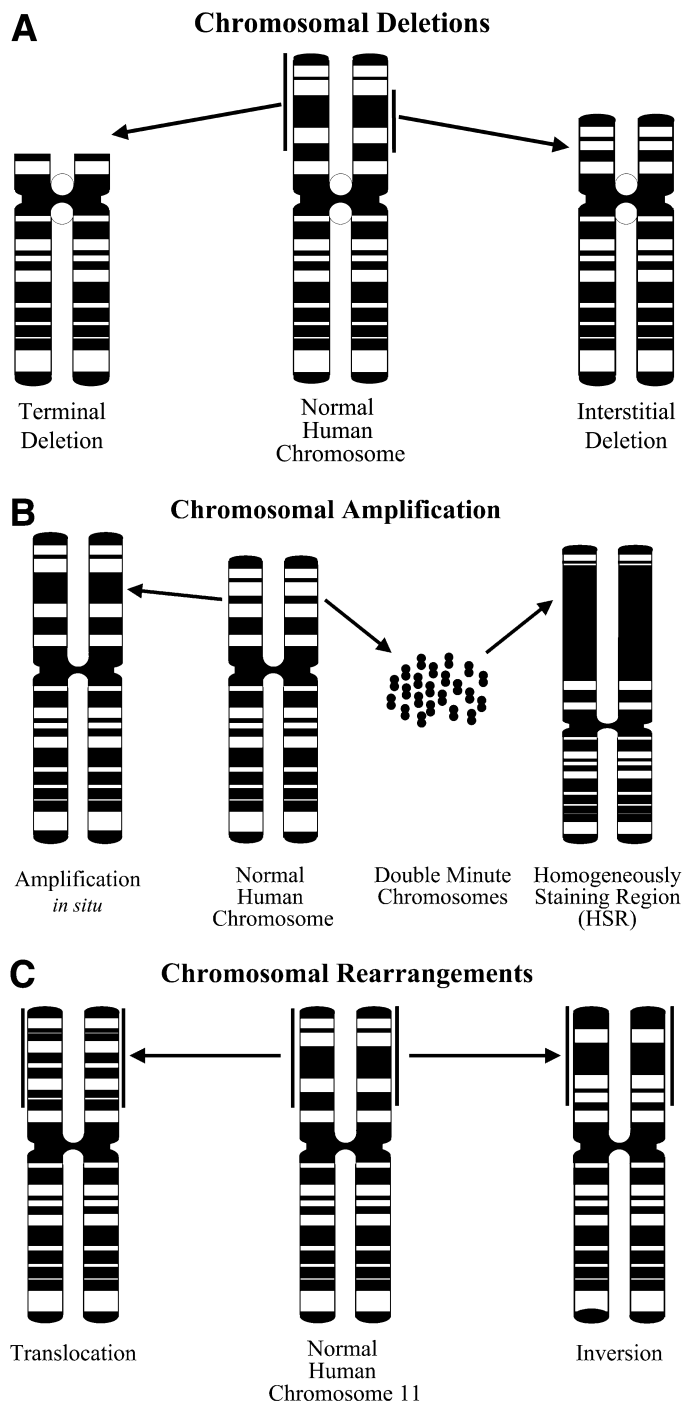


Fig. 6. Forms of gross chromosomal aberrations. Three examples of chromosomal aberrations are demonstrated using a standard G-band ideogram of human chromosome 11. The asterisks denote the chromosomal band that has expanded as a result of DNA amplification. The stippled area highlights a region from another chromosome that has been translocated to the long arm of the starting chromosome, displacing the normal qter region without altering the overall size of the chromosome (note the differences in G-banding in this region).

missense mutations and nonsense mutations. Missense mutations involve nucleotide base substitutions that alter the translation of the affected codon triplet. In contrast, nonsense mutations involve nucleotide base substitutions that modify a triplet codon that normally encodes for an amino acid into a

Table 3
Forms and Consequences of Molecular Mutation

Normal coding sequence and amino acid translation									
... Phe	Phe	Glu	Pro	Gly	Ser	Asn	Val	Tyr	...
... UUC	UUU	GAA	CCG	GGA	AGC	AAU	GUC	UAC	A...
Missense point mutation resulting in amino acid change									
... Phe	Phe	Glu	Pro	Val	Ser	Asn	Val	Tyr	...
... UUC	UUU	GAA	CCG	GCA	AGC	AAU	GUC	UAC	A...
Missense point mutation without amino acid change (silent mutation)									
... Phe	Phe	Glu	Pro	Gly	Ser	Asn	Val	Tyr	...
... UUC	UUU	GAA	CCG	GGC	AGC	AAU	GUC	UAC	A...
Frameshift mutation resulting from a single base insertion									
... Phe	Phe	Glu	Pro	Arg	Lys	Gln	Cys	Leu	...
... UUC	UUU	GAA	CCG	AGG	AAG	CAA	UGU	CUA	CA...
Frameshift mutation resulting from a single base deletion									
... Phe	Phe	Glu	Pro	Glu	Asp	Met	Ser	Asn	...
... UUC	UUU	GAA	CCG	GAA	GCA	AUG	UCU	ACA	...
Nonsense mutation resulting in a premature stop codon									
... Phe	Phe	Glu	Pro	Stop					...
... UUC	UUU	GAA	CCG	UGA	AGC	AAU	GUC	UAC	...

translational stop codon. This results in the premature termination of translation and the production of a truncated protein product. Small deletions and insertions can usually be classified as frameshift mutations because the deletion or insertion of a single nucleotide (for instance) alters the reading frame of the gene on the 3' side of the affected site. This results in the synthesis of a polypeptide product that might bear no resemblance to the normal gene product (Table 3). In addition, small insertions or deletions can result in the premature termination of translation resulting from the presence of a stop codon in the new reading frame of the mutated gene. Deletions or insertions that occur involving multiples of three nucleotides will not result in a frameshift mutation, but will alter the resulting polypeptide gene product, which will exhibit either loss of specific amino acids or the presence of additional amino acids within its primary structure. These types of alteration can also lead to a loss of protein function.

REFERENCES

- Avery, O. T., MacLeod, C. M., and McCarty, M. Studies on the chemical nature of the substance inducing transformation of *Pneumococcus* types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* Type III. *J. Exp. Med.* **79**:137–158, 1944.
- McCarty, M. and Avery, O. T. Studies of the chemical nature of the substance inducing transformation of pneumococcal types II. Effect of desoxyribonuclease on the biological activity of the transforming substance. *J. Exp. Med.* **83**:89–96, 1946.
- Watson, J. D. and Crick F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**:737–738, 1953.
- Wilkins, M. H., Stokes, A. R., and Wilson, H. R. Molecular structure of deoxypentose nucleic acids. *Nature* **171**:738–740, 1953.
- Franklin, R. E. and Gosling, R. G. Molecular structure of nucleic acids. Molecular configuration in sodium thymonucleate. 1953. *Ann. NY Acad. Sci.* **758**:16–17, 1995.
- Watson, J. D. and Crick, F. H. The structure of DNA. *Cold Spring Harbor Symp. Quant. Biol.* **18**:123–131, 1953.
- Franklin, R. E. and Gosling, R. G. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172**:156–157, 1953.
- Jacobson, B. Hydration structure of deoxyribonucleic acid and its physicochemical properties. *Nature* **172**:666–667, 1953.
- Wilkins, M. H., Seeds, W. E., Stokes, A. R., and Wilson, H. R. Helical structure of crystalline deoxypentose nucleic acid. *Nature* **172**:759–762, 1953.
- Watson, J. D. and Crick, F. H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**:964–967, 1953.
- Brenner, S., Jacob, F., and Meselson, M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**:576–581, 1961.
- Dounce, A. L. Duplicating mechanism for peptide chain and nucleic acid synthesis. *Enzymologia* **15**:251–258, 1952.
- Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J. Biol. Chem.* **233**:163–170, 1958.
- Kornberg, A. Biologic synthesis of deoxyribonucleic acid. *Science* **131**:1503–1508, 1960.
- Nirenberg, M. W. and Matthaei, J. H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. USA* **47**:1588–1602, 1961.
- Matthaei, H. and Nirenberg, M. W. The dependence of cell-free protein synthesis in *E. coli* upon RNA prepared from ribosomes. *Biochem. Biophys. Res. Commun.* **4**:404–408, 1961.
- Kornberg, R. D. and Lorch, Y. Chromatin structure and transcription. *Annu. Rev. Cell Biol.* **8**:563–587, 1992.
- Kozak, M. Regulation of translation in eukaryotic systems. *Annu. Rev. Cell Biol.* **8**:197–225, 1992.
- Coverley, D. and Laskey, R. A. Regulation of eukaryotic DNA replication. *Annu. Rev. Biochem.* **63**:745–776, 1994.
- Crick, F. H. C. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**:548–555, 1958.
- Crick, F. Central dogma of molecular biology. *Nature* **227**:561–563, 1970.
- Chargaff, E., Vischer, E., Doniger, R., Green, C., and Misani, F. The composition of the deoxypentose nucleic acids of thymus and spleen. *J. Biol. Chem.* **177**:405–416, 1949.

23. Chargaff, E. Structure and function of nucleic acids as cell constituents. *Fed. Proc.* **10**:654–659, 1951.
24. Crick, F. H. C. and Watson, J. D. The complementary structure of deoxyribonucleic acid. *Proc. R. Soc. A* **223**:80–96, 1954.
25. Hunter, W. N., Brown, T., Anand, N. N., and Kennard, O. Structure of an adenine–cytosine base pair in DNA and its implications for mismatch repair. *Nature* **320**:552–555, 1986.
26. Liu, K., Miles, H. T., Frazier, J., and Sasisekharan, V. A novel DNA duplex. A parallel-stranded DNA helix with Hoogsteen base pairing. *Biochemistry* **32**:11,802–11,809, 1983.
27. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., and Kopka, M. L. The anatomy of A-, B-, and Z-DNA. *Science* **216**:475–485, 1982.
28. Collins, F. S. The human genome project and the future of medicine. *Ann. NY Acad. Sci.* **882**:42–55, 1999; discussion 56–65.
29. Futreal, P. A., Kaspzyk, A., Birney, E., Mullikin, J. C., Wooster, R., and Stratton, M. R. Cancer and genomics. *Nature* **409**:850–852, 2001.
30. Jimenez-Sanchez, G., Childs, B., and Valle, D. Human disease genes. *Nature* **409**:853–855, 2001.
31. Sanger, F. and Coulson, A. R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**:441–448, 1975.
32. Sanger, F., Nicklen, S., and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467, 1977.
33. Maxam, A. M. and Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560–564, 1977.
34. Smith, L. M., Sanders, J. Z., Kaiser, R. J., et al. Fluorescence detection in automated DNA sequence analysis. *Nature* **321**:674–679, 1986.
35. Ansorge, W., Sproat, B. S., Stegemann, J., and Schwager, C. A non-radioactive automated method for DNA sequence determination. *J. Biochem. Biophys. Methods* **13**:315–323, 1986.
36. Collins, F. and Galas, D. A new five-year plan for the U.S. Human Genome Project. *Science* **262**:43–46, 1993.
37. Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., and Walters, L. New goals for the U.S. Human Genome Project: 1998–2003. *Science* **282**:682–689, 1998.
38. Hudson, T. J., Stein, L. D., Gerety, S. S., et al. An STS-based map of the human genome. *Science* **270**:1945–1954, 1995.
39. Berry, R., Stevens, T. J., Walter, N. A., et al. Gene-based sequence-tagged-sites (STSs) as the basis for a human gene map. *Nat. Genet.* **10**:415–423, 1995.
40. Gyapay, G., Schmitt, K., Fizames, C., et al. A radiation hybrid map of the human genome. *Hum. Mol. Genet.* **5**:339–346, 1996.
41. Cheung, V. G., Nowak, N., Jang, W., et al. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409**:953–958, 2001.
42. Olivier, M., Aggarwal, A., Allen, J., et al. A high-resolution radiation hybrid map of the human genome draft sequence. *Science* **291**:1298–1302, 2001.
43. McPherson, J. D., Marra, M., Hillier, L., et al. A physical map of the human genome. *Nature* **409**:934–941, 2001.
44. Adams, M. D., Dubnick, M., Kerlavage, A. R., et al. Sequence identification of 2,375 human brain genes. *Nature* **355**:632–634, 1992.
45. Adams, M. D., Soares, M. B., Kerlavage, A. R., Fields, C., and Venter, J. C. Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat. Genet.* **4**:373–380, 1993.
46. Adams, M. D., Kerlavage, A. R., Fields, C., and Venter, J. C. 3,400 new expressed sequence tags identify diversity of transcripts in human brain. *Nat. Genet.* **4**:256–267, 1993.
47. Schuler, G. D., Boguski, M. S., Stewart, E. A., et al. A gene map of the human genome. *Science* **274**:540–546, 1996.
48. Caron, H., van Schaik, B., van der Mee, M., et al. The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291**:1289–1292, 2001.
49. Venter, J. C., Adams, M. D., Martin-Gallardo, A., McCombie, W. R., and Fields, C. Genome sequence analysis: scientific objectives and practical strategies. *Trends Biotechnol.* **10**:8–11, 1992.
50. Venter, J. C., Smith, H. O., and Hood, L. A new strategy for genome sequencing. *Nature* **381**:364–366, 1996.
51. Venter, J. C., Adams, M. D., Sutton, G. G., et al. Shotgun sequencing of the human genome. *Science* **280**:1540–1542, 1998.
52. Lander, E. S., Linton, L. M., Birren, B., et al. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921, 2001.
53. Venter, J. C., Adams, M. D., Myers, E. W., et al. The sequence of the human genome. *Science* **291**:1304–1351, 2001.
54. Nowak, R. Mining treasures from ‘junk DNA.’ *Science* **263**:608–610, 1994.
55. Antequera, F. and Bird, A. Predicting the total number of human genes. *Nat. Genet.* **8**:114, 1994.
56. Fields, C., Adams, M. D., White, O., and Venter, J. C. How many genes in the human genome? *Nat. Genet.* **7**:345–346, 1994.
57. Ewing, B. and Green, P. Analysis of expressed sequence tags indicates 35,000 human genes. *Nat. Genet.* **25**:232–234, 2000.
58. Roest Crollius, H., Jaillon, O., Bernot, A. Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence. *Nat. Genet.* **25**:235–238, 2000.
59. Sachidanandam, R., Weissman, D., Schmidt, S. C., et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**:928–933, 2001.
60. Collins, F. S., Green, E. D., Guttacher, A. E., and Guyer, M. S. A vision for the future of genomics research. *Nature* **422**:835–847, 2003.
61. Collins, F. S., Morgan, M. and Patrinos, A. The Human Genome Project: lessons from large-scale biology. *Science* **300**:286–290, 2003.
62. Kavenoff, R., Klotz, L. C., and Zimm, B. H. One the nature of chromosome-sized DNA molecules. *Cold Spring Harbor Symp. Quant. Biol.* **38**:1–8, 1974.
63. Spector, D. L. The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* **72**:573–608, 2003.
64. Mattick, J. S. Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep.* **2**:986–991, 2001.
65. Small, D., Nelkin, B., and Vogelstein, B. Nonrandom distribution of repeated DNA sequences with respect to supercoiled loops and the nuclear matrix. *Proc. Natl. Acad. Sci. USA* **79**:5911–5915, 1982.
66. Tsongalis, G. J., Coleman, W. B., Smith, G. J., and Kaufman, D. G. Partial characterization of nuclear matrix attachment regions from human fibroblast DNA using Alu-polymerase chain reaction. *Cancer Res.* **52**:3807–3810, 1992.
67. Jelinek, W. R. and Schmid, C. W. Repetitive sequences in eukaryotic DNA and their expression. *Annu. Rev. Biochem.* **51**:813–844, 1982.
68. Schmid, C. W. and Jelinek, W. R. The Alu family of dispersed repetitive sequences. *Science* **216**:1065–1070, 1958.
69. Meselson, M. and Stahl, F. W. The replication of DNA. *Cold Spring Harbor Symp. Quant. Biol.* **23**:9–12, 1958.
70. Meselson, M. and Stahl, F. W. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **44**:671–682, 1958.
71. Franceschini, P. Semiconservative DNA duplication in human chromosomes treated with BUDR and stained with acridine orange. *Exp. Cell Res.* **89**:420–421, 1974.
72. Rude, J. M. and Friedberg, E. C. Semi-conservative deoxyribonucleic acid synthesis in unirradiated and ultraviolet-irradiated xeroderma pigmentosum and normal human skin fibroblasts. *Mutat. Res.* **42**:433–442, 1977.
73. West, S. C. Enzymes and molecular mechanisms of genetic recombination. *Annu. Rev. Biochem.* **61**:603–640, 1992.
74. Alberts, B. DNA replication and recombination. *Nature* **421**:431–435, 2003.
75. Friedberg, E. C. Biological responses to DNA damage: a perspective in the new millennium. *Cold Spring Harb. Symp. Quant. Biol.* **65**:593–602, 2000.
76. Friedberg, E. C. DNA damage and repair. *Nature* **421**:436–440, 2003.
77. Sancar, A. and Sancar G. B. (1988) DNA repair enzymes. *Annu. Rev. Biochem.* **57**:29–67.

78. Friedberg, E. C., Walker, G. C., and Siede, W. *DNA Repair and Mutagenesis*. ASM, Washington, DC, pp. xvii, 1995.
79. Lewin, B. Units of transcription and translation: sequence components of heterogeneous nuclear RNA and messenger RNA. *Cell* **4**:77–93, 1975.
80. Oren, M. The p53 cellular tumor antigen: gene structure, expression and protein properties. *Biochim. Biophys. Acta* **823**:67–78, 1985.
81. Hollstein, M., Hergenhahn, M., Yang, Q., Bartsch, H., Wang, Z. Q., and Hainaut, P. New approaches to understanding p53 gene tumor mutation spectra. *Mutat. Res.* **431**:199–209, 1999.
82. Bennett, W. P., Hussain, S. P., Vahakangas, K. H., Khan, M. A., Shields, P. G., and Harris, C. C. Molecular epidemiology of human cancer risk: gene–environment interactions and p53 mutation spectrum in human lung cancer. *J. Pathol.* **187**:8–18, 1999.
83. Bookstein, R., Lee, E. Y., To, H., et al. Human retinoblastoma susceptibility gene: genomic organization and analysis of heterozygous intragenic deletion mutants. *Proc. Natl. Acad. Sci. USA* **85**:2210–2214, 1988.
84. Lee, W. H., Bookstein, R., Hong, F., Young, L. J., Shew, J. Y., and Lee, E. Y. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science* **235**:1394–1399, 1987.
85. Hong, F. D., Huang, H. J., To, H., et al. Structure of the human retinoblastoma gene. *Proc. Natl. Acad. Sci. USA* **86**:5502–5506, 1989.
86. Cech, T. R. Self-splicing of group I introns. *Annu. Rev. Biochem.* **59**:543–568, 1990.
87. Jurica, M. S. and Moore, M. J. Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell* **12**:5–14, 2003.
88. Black, D. L. Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72**:291–336, 2003.
89. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55**:1119–1150, 1986.
90. Sharp, P. A. Splicing of messenger RNA precursors. *Science* **235**:766–771, 1987.
91. Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J., and Scott, J. A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* **50**:831–840, 1987.
92. Sollner-Webb, B. RNA editing. *Curr. Opin. Cell Biol.* **3**:1056–1061, 1991.
93. Landweber, L. F. and Gilbert, W. RNA editing as a source of genetic variation. *Nature* **363**:179–182, 1993.
94. Mattaj, I. W., Tollervey, D., and Seraphin, B. Small nuclear RNAs in messenger RNA and ribosomal RNA processing. *FASEB J.* **7**:47–53, 1993.
95. Darnell, J. E., Philipson, L., Wall, R., and Adesnik, M. Polyadenylic acid sequences: role in conversion of nuclear RNA into messenger RNA. *Science* **174**:507–510, 1971.
96. Blum, B., Bakalara, N., and Simpson, L. A model for RNA editing in kinetoplastid mitochondria: “guide” RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**:189–198, 1990.
97. Schimmel, P. and Ribas de Pouplana, L. Transfer RNA: from mini-helix to genetic code. *Cell* **81**:983–986, 1995.
98. Crick, F. H. Codon–anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* **19**:548–555, 1966.
99. Noller, H. F. Structure of ribosomal RNA. *Annu. Rev. Biochem.* **53**:119–162, 1984.
100. Shen, L. X., Cai, Z., and Tinoco, I., Jr. RNA structure at high resolution. *FASEB J.* **9**:1023–1033, 1995.
101. Drake, J. W. and Baltz, R. H. The biochemistry of mutagenesis. *Annu. Rev. Biochem.* **45**:11–37, 1976.
102. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**:709–715, 1993.
103. Ames, B. N., Shigenaga, M. K., and Gold, L. S. DNA lesions, inducible DNA repair, and cell division: three key factors in mutagenesis and carcinogenesis. *Environ. Health Perspect.* **101**(Suppl 5):35–44, 1993.
104. Cooper, D. N. and Youssoufian, H. The CpG dinucleotide and human genetic disease. *Hum. Genet.* **78**:151–155, 1988.
105. Rideout, W. M., 3rd, Coetzee, G. A., Olumi, A. F., and Jones, P. A. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* **249**:1288–1290, 1990.
106. Jones, P. A., Buckley, J. D., Henderson, B. E., Ross, R. K., and Pike, M. C. From gene to carcinogen: a rapidly evolving field in molecular epidemiology. *Cancer Res.* **51**:3617–3620, 1991.
107. Bishop, J. M. Molecular themes in oncogenesis. *Cell* **64**:235–248, 1991.
108. Coleman, W. B. and Tsongalis, G. J. Multiple mechanisms account for genomic instability and molecular mutation in neoplastic transformation. *Clin. Chem.* **41**:644–657, 1995.
109. Coleman, W. B. and Tsongalis, G. J. The role of genomic instability in human carcinogenesis. *Anticancer Res.* **19**:4645–4664, 1999.
110. Coleman, W. B. and Tsongalis, G. J. The role of genomic instability in the development of human cancer, in *The Molecular Basis of Human Cancer*, Coleman, W. B. and Tsongalis, G. J., eds., Humana, Totowa, NJ, pp. 115–142, 2002.



<http://www.springer.com/978-1-58829-356-5>

Molecular Diagnostics

For the Clinical Laboratorian

Coleman, W.B.; Tsongalis, G.J. (Eds.)

2005, XVIII, 567 p., Hardcover

ISBN: 978-1-58829-356-5

A product of Humana Press