GENETICS

Mendel and the Gene Idea
(Chapter 14)

YOU MUST KNOW…

- Terms associated with genetics problems: P, F_1, F_2, dominant, recessive, homozygous, heterozygous, phenotype, and genotype.
- How to derive the proper gametes when working a genetics problem.
- The difference between an allele and a gene.
- How to read a pedigree.
- How to use data sets to determine Mendelian patterns of inheritance.

Mendel used the scientific approach to identify two laws of inheritance (14.1)

- True-breeding parents in a genetic cross are called **the P (parental) generation**; their offspring are called the **F_1 (first filial) generation**. If the F_1 population is crossed, their offspring are called the **F_2 (second filial) generation**.
- The following are four related concepts that make up Mendel's model explaining the 3:1 inheritance pattern that he observed among F2 offspring. Use Figure 14.5 to find each of the four basic concepts of Mendel's model.
    - **Alternative versions of genes account for variations in inherited characteristics among offspring**. For example, consider flower color in peas. The gene for flower color in pea plants comes in two versions: white and purple. These alternative versions of the gene, called **alleles**, are the result of slightly different DNA sequences.
    - **For each character, every organism inherits one allele from each parent**.
    - **If the two alleles are different, then the dominant allele will be fully expressed in the offspring, whereas the recessive allele will have no noticeable effect on the offspring**.
    - **The two alleles for each character separate during gamete production**. If the parent has two of the same alleles, then the offspring will all get that version of the gene, but if the parent has two different alleles for a gene, each offspring has a 50% chance of getting one of the two alleles. This is Mendel's **law of segregation**.
- **Law of independent assortment** was Mendel's second law. It states that each pair of alleles will segregate (separate) independently during gamete formation. This occurs during anaphase I of meiosis.
- **Homozygous** organisms have two of the same alleles for a particular trait. If the dominant allele for a trait is designated *R* (dominant traits are generally capitalized), and the recessive allele is designated *r* (recessive traits are generally not capitalized), then an individual could be homozygous for the dominant trait (*RR*) or homozygous for the recessive trait (*rr*).

- A **heterozygous** organism has two different alleles for a trait (*Rr*).
- **Phenotype** refers to an organism's expressed physical traits, and **genotype** refers to an organism's genetic makeup. For example, the phenotype of a seed might be round, and its genotype could be *RR* or *Rr*.
- A **testcross** is done to determine if an individual showing a dominant trait is homozygous or heterozygous. A test cross is always done between the unknown genotype and a *homozygous recessive* individual. If the unknown parent is homozygous dominant, (*RR* x *rr*) all of the offspring will show the dominant trait, but if the unknown parent is heterozygous (*Rr* x *rr*) some of the offspring will show the recessive trait.
- A **monohybrid cross** is a cross involving the study of only one character (for example, flower color), whereas a **dihybrid cross** is a cross intended to study two characters (for example, flower color and seed shape).
- The diagram that follows shows the results of a dihybrid cross. In this case, in the parental generation two homozygous plants are crossed: one homozygous dominant for yellow and round seeds (*YYRR*) and one homozygous recessive for green and wrinkled seeds (*yyrr*). The only gamete type the first parent can produce is *YR*, and the only gamete the second parent can produce is *yr*. The F_1 generation, therefore, is composed of individuals with genotype *YyRr*. Crossing *YyRr* with a second *YyRr* gives an F_2 generation that completes the cross and looks like Figure 14.8.

The laws of probability govern Mendelian inheritance (14.2)

- Understanding how to predict offspring of genetic crosses involves familiarity with the basic laws of probability. There are two laws that you will use directly in solving genetics problems:
    - **The rule of multiplication**: When calculating the probability that two or more independent events will occur together in a specific combination, multiply the probabilities of each of the two events. Thus, the probability of a coin landing face up two times in two flips is ½ x ½ = ¼. If you cross two organisms with the genotypes *AABbCc* and *AaBbCc*, the probability of an offspring having the genotype *AaBbcc* is ½ x ½ x ¼ = 1/16.
    - **The rule of addition**: When calculating the probability that any of two or more mutually exclusive events will occur, you need to add together their individual probabilities. For example, if you are tossing a die, what is the probability that it will land on either side with 4 spots or the side with 5 spots? (1/6 + 1/6 = 1/3)

> **STUDY TIP**
>
> What are the chances of event 1 ***and*** event 2? Multiply them. What are the chances of event 1 ***or*** event 2? Add them.

Inheritance patterns are often more complex than predicted by simple Mendelian genetics (14.3)

- **Complete dominance** is dominance in which the heterozygote and the homozygote for the dominant allele are indistinguishable. A *Yy* yellow seed is just as yellow as a *YY* yellow seed.
- **Codominance** occurs when two alleles are dominant and affect the phenotype in two different but equal ways. The traditional example for this type of dominance is human blood types. The alleles for A and B blood are dominant to the allele for type O blood, but A and B are codominant to each other. A person who has alleles for both A and B blood will be blood type AB because these alleles are each completely expressed.
- **Incomplete dominance** is a type of dominance in which the F_1 hybrids have an appearance that is in between that of the two parents. For example, if two plants, one with white flowers and one with red flowers, were crossed and all of the offspring had pink flowers, you could conclude that the trait for flower color exhibits incomplete dominance. Breeding two of the hybrids with incomplete dominance gives a flower ratio of 1 red: 2 pink: 1 white.
- **Multiple alleles** occur when a gene has more than two alleles. Again, a good example of this is seen in human blood types. There are three alleles for human blood types: I^A (or A), I^B (or B), and *i* (or O) but one person only receives any combination of two alleles.
- **Pleiotropy** is the property of a gene that causes it to have multiple phenotypic effects. For example, sickle-cell disease has multiple symptoms all caused by a single defective gene.
- In **epistasis**, a gene at one locus alters the effects of a gene at another locus. For example, an individual may have genes for heavy skin pigmentation, but if a separate gene that produces the pigment is defective, the genes for pigment deposition will not be expressed. This would lead to a condition known as albinism.
- In **polygenic inheritance**, two or more genes have an additive effect on a single character in the phenotype (such as height or skin color in humans). When several genes are involved, the phenotype usually is described by a bell-shaped curve, with fewer individuals at each extreme and most individuals clustered in the middle.

Many human traits follow Mendelian patterns of inheritance (14.4)

- A **pedigree** is a diagram that shows the relationship between parents and offspring across two or more generations. (See Figure 14.15.) In a typical pedigree, circles represent females, and squares represent males. White open circles or squares indicate that the individual did not or does not express a particular trait, whereas the shaded ones indicate that the individual expresses or expressed the trait. Through the patterns they reveal, pedigrees can help determine the genome of individuals that comprise them; pedigrees can also help predict the genome of future offspring.
- **Recessively inherited disorders** require two copies of the defective gene for the disorder to be expressed. Examples include the following:
    - **Cystic fibrosis** is caused by a mutation in an allele that codes for a cell membrane protein that functions in the transport of chloride ions into and out of cells. The

resulting high extracellular levels of chloride cause mucus to be thicker and stickier, leading to organ malfunction and recurrent bacterial infections.
    - **Tay-Sachs** disease is caused by an allele that codes for a dysfunctional enzyme, which is unable to break down certain lipids in the brain. As these lipids accumulate in the brain cells, the child suffers from blindness, seizures, and degeneration of brain function, leading to death.
    - **Sickle-cell disease** is caused by an allele that codes for a mutant hemoglobin molecule that forms long rods when the oxygen levels in the blood are low. These long rods cause the red blood cell to sickle, clogging small blood vessels and leading to pain, organ damage, and even paralysis.
- **Lethal dominant alleles** require only one copy of the allele in order for the disorder to be expressed. Usually, only lethal alleles that act late in life are passed on.
- **Huntington's disease** is caused by a lethal dominant allele. It is a degenerative disease of the nervous system, which usually doesn't affect the individual until he or she is over 40 years old.
- Genetic testing may be used on a fetus to detect certain genetic disorders. Two common tests are amniocentesis and chorionic villus sampling (CVS).
    - **Amniocentesis** occurs when the physician removes amniotic fluid from around the fetus. The amniotic fluid can be utilized to detect some genetic disorders, and the cells in the fluid can be cultured for a karyotype.
    - **Chorionic villus sampling** involves using a narrow tube inserted through the cervix to suction out a tiny sample of the placenta that contains only fetal cells. A karyotype can immediately be developed from these cells.

The Chromosomal Basis of Inheritance
(Chapter 15)

YOU MUST KNOW…

- How the chromosome theory of inheritance connects the physical movement of chromosomes in meiosis to Mendel's laws of inheritance.
- The unique pattern of inheritance in sex-linked genes.
- How alteration of chromosome number or structurally altered chromosomes (deletions, duplications, etc.) can cause genetic disorders.
- How genomic imprinting and inheritance of mitochondrial DNA are exceptions to standard Mendelian inheritance.

Mendelian inheritance has its physical basis in the behavior of chromosomes (15.1)

- The **chromosome theory of inheritance** states that genes have specific locations (called *loci*) on chromosomes and that it is chromosomes that segregate and assort independently. It is important to connect this physical movement of chromosomes in meiosis to Mendel's laws of inheritance. If you are having trouble visualizing this, carefully work through Figure 15.2 in your book.
- A **sex-linked gene** is one located on a sex chromosome (X or Y in humans). After the chromosome theory of inheritance was formed, *Thomas Hunt Morgan* discovered the existence of sex-linked genes.

Sex-linked genes exhibit unique patterns of inheritance (15.2)

- In humans, there are two types of sex chromosomes, X and Y. A person who inherits two X chromosomes usually develops as a female, whereas a person who inherits one X and one Y normally develops as a male.
- Sex-linked genes may be either X-linked or Y-linked. Figure 15.7 shows the unique pattern of inheritance in X-linked genes. In addition to tracking the gene from one generation to the next, it is also necessary to track the sex of the offspring. Figure 15.7 is based on work with *Drosophila* (fruit flies) performed by Thomas Hunt Morgan, but the pattern is the same in humans.
- Each egg or ovum contains an X chromosome; there are two types of sperm: those with an X chromosome and those with a Y chromosome. In fertilization, there is a 50% chance that a sperm carrying an X or a Y will reach and penetrate the egg first. Thus, gender is determined by chance and by the male sperm cell in humans.
- Fathers pass X-linked genes on to their daughters but not to their sons; fathers pass the Y chromosome to all of their sons.

- Females will express an X-linked trait exactly like any other trait, but males, with only one X chromosome, will express the allele on the X chromosome they inherited from their mother. The terms *homozygous* and *heterozygous* do not apply to a male pattern of sex-linked genes.
- The vast majority of genes on the X chromosome are not related to sex.
- Several X-linked disorders have medical significance:
    - **Duchenne muscular dystrophy** is an X-linked disorder characterized by a progressive weakening of the muscles and loss of coordination. Affected individuals rarely live past their early 20s.
    - **Hemophilia** is an X-linked disorder characterized by having blood with an inability to clot normally, caused by the absence of proteins required for blood clotting.
- **X-inactivation** regulates gene dosage in females. Although female mammals inherit two X chromosomes, one of the X chromosomes (randomly chosen) in each cell of the body becomes inactivated during embryonic development by **methylation**. As a result, males and females have the same effective dose of genes with loci on the X chromosome.
- The inactive chromosome condenses into a **Barr body**, which lies along the inside of the nuclear envelope. Still, females are not usually affected as heterozygote carriers of problematic alleles, because half of their sex chromosomes are normal and produce the necessary proteins.

Linked genes tend to be inherited together because they are located near each other on the same chromosome (15.3)

- **Linked genes** are located on the same chromosome and therefore tend to be inherited together during cell division.
- **Genetic recombination** is the production of offspring with a new combination of genes inherited from the parents. This is due to crossing over during prophase I of meiosis. Many genetic crosses yield some offspring with the same phenotype as one of the parents (these offspring are referred to as **parental types**) and some offspring with phenotypes different from either parent (these offspring are referred to as **recombinants**).
- **Crossing over** can explain why some linked genes get separated during meiosis. During meiosis, unlinked genes follow independent assortment because they are located on different chromosomes. Linked genes are located on the same chromosome and would not be predicted to follow independent assortment. However, sometimes genetic crosses give results that seem to indicate that some independent assortment has occurred, even when genes are on the same chromosome. These results are not caused by independent assortment but can be explained by crossing over. Research further indicates that the farther apart two genes are on a chromosome, the higher the probability that crossing over will occur between them. The likelihood of crossing over between different genes on the same chromosome is expressed as a percent.
- A **linkage map** is a genetic map that is based on the percentage of crossover events.
- A **map unit** is equal to a 1% recombination frequency. Map units are used to express relative distance along the chromosome.

Alterations of chromosome number or structure cause some genetic disorders (15.4)

- **Nondisjunction** occurs when the members of a pair of homologous chromosomes do not separate properly during meiosis I, or sister chromatids don't separate properly during meiosis II.
- As a result of nondisjunction, one gamete receives two copies of the chromosome, whereas the other gamete receives none. If the faulty gametes engage in fertilization, the offspring will have an incorrect chromosome number. This is known as **aneuploidy**.
- Fertilized eggs that have received three copies of the chromosome in question are said to be **trisomic**; those that have received just one copy of a chromosome are said to be **monosomic** for the chromosome.
- **Polyploidy** is the condition of having more than two complete sets of chromosomes, forming a 3*n* or 4*n* individual. Rare in animals, this condition is fairly frequent in plants.
- Errors in meiosis or damaging agents like radiation can cause portions of a chromosome to be lost or rearranged, resulting in the following mutations:
    - A **deletion** occurs when a chromosomal fragment is lost, resulting in a chromosome with missing genes.
    - A **duplication** occurs when a chromosomal segment is repeated.
    - An **inversion** occurs when a chromosomal fragment breaks off and reattaches to its original position but backward, so that the part of the fragment that was originally at the attachment point is now at the end of the chromosome.
    - A **translocation** occurs when the deleted chromosome fragment joins a *nonhomologous* chromosome.
- Human disorders caused by chromosomal alterations include the following:
    - **Down syndrome**: An aneuploid condition that is the result of having an extra chromosome 21 (trisomy 21). Down syndrome includes characteristic facial features, short stature, heart defects, and developmental delays.
    - **Klinefelter syndrome**: An aneuploidy condition in which a male possesses the sex chromosome XXY (an extra X). Klinefelter males have male sex organs but are sterile.
    - **Turner syndrome**: A monosomic condition in which the female has just one sex chromosome, often designated as XO. Turner syndrome females are sterile because the reproductive organs do not mature. Turner syndrome is the only known viable monosomy in humans.

Some inheritance patterns are exceptions to standard Mendelian inheritance (15.5)

- In mammals, geneticists have identified traits that differ, depending on which parent passed along the allele for those traits. This phenomenon is called genomic imprinting. The phenotypic effect of a gene may depend on which allele is inherited from which parent.
    - Genomic imprinting occurs during gamete formation and results in the silencing of a particular allele of certain genes. The offspring expresses only one allele of an imprinted

gene, hence the exception to Mendelian inheritance. Over 60 imprinted genes have been identified, with hundreds more suspected.

- Genes that are present in mitochondria and plastids are inherited only from the mother because the zygote's cytoplasm comes only from the egg. You inherited your mitochondrial DNA only from your mother; your mother inherited her mitochondrial DNA only from her mother. Your mitochondrial DNA is your maternal grandmother's!

The Molecular Basis of Inheritance
(Chapter 16)

YOU MUST KNOW…

- The structure of DNA.
- The knowledge about DNA gained from the work of Griffith; Avery, MacLeod, and McCarty; Hershey and Chase; Wilkins and Franklin; and Watson and Crick. (LO3.2)
- Replication is semiconservative and occurs 5' to 3'.
- The roles of DNA polymerase, ligase, helicase, and topoisomerase in replication. (LO 3.3)
- The general differences between bacterial chromosomes and eukaryotic chromosomes.
- How DNA packaging can affect gene expression.

DNA is the genetic material (16.1)

- Once chromosomes were known to carry genes, the next question became which of the two organic compounds that make chromosomes, DNA or protein, was the genetic material?
    - Frederick Griffith studied two strains of the bacterium *Streptococcus pneumoniae*. In mice, bacteria of the smooth strain caused pneumonia, whereas the rough strain was nonpathogenic. However, when Griffith mixed heat-killed smooth strain (nonpathogenic) and rough strains (nonpathogenic) and injected mice with the mixture, the mice developed pneumonia and died! Griffith concluded that the living rough strain had been transformed into the pathogenic smooth form by a heritable agent. The question remained: What was the heritable agent, protein or DNA?
    - O.T. Avery, along with his colleagues McCarty and MacLeod, dedicated 14 years to identifying the "transforming agent" from Griffith's work. In 1944 they concluded that DNA was the transforming factor. Their results were greeted with interest but skepticism. Scientists had difficulty believing that a macromolecule as simple as DNA could be the genetic material and not the much more complex protein molecule.
- In 1952 **Alfred Hershey and Martha Chase** provided an unambiguous answer to the DNA or protein question utilizing *bacteriophages* – viruses that infect bacteria. Bacteriophages were excellent organisms for this study, in part because they are made of only two organic compounds, DNA and protein. Hershey and Chase used a radioactive isotope of phosphorus to tag the DNA in one culture of bacteriophages and radioactive sulfur to tag the protein in a second culture. Their results clearly showed that only the DNA entered bacteria infected by the

virus; the radioactive protein never entered the cell. This research convinced scientists that DNA must be the genetic material.

- The next big question centered on the structure of DNA. Would the structure of DNA give any clues as to how it functioned as the genetic material?
    - **James Watson and Francis Crick** were the first to solve the puzzle of the structure of DNA. Critical to their success was the work of Rosalind Franklin and Maurice Wilkins, both working in the field of X-ray crystallography.
    - **X-ray crystallography** is a process used to visualize molecules three-dimensionally. X-rays are diffracted as they pass through the molecule, and they bounce back to produce patterns that can be interpreted through mathematical equations. Through this technique, a rough blueprint of the molecule was formed and the helical structure was deduced.
- Watson and Crick's model determined four major features of DNA. Find each major point by following Figure 16.7 as the model is explained.
    - DNA is a **double helix**, which can be described as a twisted ladder with rigid rungs. The side, or backbone, is made up of sugar-phosphate components, whereas the rungs are made up of pairs of nitrogenous bases.
    - Notice that a single nucleotide is circled in Figure 16.7. It is composed of a sugar (deoxyribose) attached to a phosphate and a nitrogen base.
    - The nitrogenous bases of DNA are adenine (A), thymine (T), guanine (G), and cytosine (C). In DNA, adenine pairs only with thymine, and guanine pairs only with cytosine.
    - Notice that the chain on the right side of the model runs in one direction, while the left side of the chain runs in the opposite, upside-down direction. The strands are termed **antiparallel**. The left side runs 5'to 3' and the opposite strand runs 3' to 5'. (Recall that the carbons are numbered, and you will see that the number 5 carbon and the number 3 carbon and the resultant nucleotides are flipped relative to each other.) Nucleic acid strands are always antiparallel, whether they are DNA/DNA or DNA/RNA or RNA/RNA interactions.

Many proteins work together in DNA replication and repair (16.2)

- **Replication** is the making of DNA from an existing DNA strand. DNA replication is *semiconservative*. This means that at the end of replication, each of the daughter molecules has one old strand, derived from the parent strand of DNA, and one strand that is newly synthesized. Study Figure 16.10 to see the pattern of semiconservative replication.
- The replication of DNA includes six major points:
    - The replication of DNA begins at sites called the *origins of replication*.
    - Initiation proteins bind to the origin of replication and separate the two strands, forming a *replication bubble*. DNA replication then proceeds in both directions along the DNA strand until the molecule is copied.
    - A group of enzymes called **DNA polymerases** catalyzes the elongation of new DNA at the replication fork.

    - DNA polymerase adds nucleotides to the growing chain one by one, working in a 5' to 3' direction, matching adenine with thymine and guanine with cytosine.
    - Recall that the strands of DNA are antiparallel. This means that DNA replication occurs continuously along the 5' to 3' strand, which is called the **leading strand**. The strand that runs 3' to 5' is copied in a series of segments and termed the **lagging strand**. Read steps 1-3 in Figure5.3 to visualize this process.
    - The lagging strand is synthesized in separate pieces called **Okazaki fragments**, which are then sealed together by **DNA ligase** (steps 4-7, Figure 16.17), forming a continuous DNA strand.
- There are several factors contributing to the accuracy of DNA replication:
    - The specificity of base pairing (A = T, G = C)
    - **Mismatch repair**, in which special repair enzymes fix incorrectly paired nucleotides
    - **Nucleotide excision repair**, in which incorrectly placed nucleotides are excised or removed by enzymes termed **nucleases**, and the gap left over is filled with the correct nucleotides
- The fact that DNA polymerase can add nucleotides only to the 3' end of a molecule means that it would have no way to complete the 5' end of the DNA molecule at the end of the chromosome. Every time the chromosome is replicated for mitosis, a small portion of the tip of the chromosome is removed. To avoid losing the terminal genes, the linear ends of eukaryotic chromosomes are "capped" with **telomeres**, short, repetitive nucleotide sequences that do not contain genes. This means that any given cell has a finite number of times it can divide before essential information is lost. In tumor cells, such as those in HeLa cells, a mutation activates an enzyme called *telomerase*, which prevents this degradation and renders the cells "immortal."

A chromosome consists of a DNA molecule packed together with proteins (16.3)

- A bacterial chromosome is one double-stranded, circular DNA molecule associated with a small amount of protein.
- Eukaryotic chromosomes are linear DNA molecules associated with large amounts of protein.
- In eukaryotic cells, DNA and proteins are packed together as **chromatin**.
- As DNA becomes more highly packaged, it becomes less accessible to transcription enzymes, which reduces gene expression. In interphase cells, most chromatin is in the highly extended form (**euchromatin**) and is available for transcription. When the euchromatin condenses to chromosomes during mitotic division, the more condensed chromatin (**heterochromatin**) is no longer available for transcription. Heterochromatin is largely inaccessible to transcription enzymes and, thus, generally is not transcribed. Barr bodies are another example of heterochromatin.

Gene Expression: From Gene to Protein
(Chapter 17)

YOU MUST KNOW…

- How RNA and DNA are similar and different, and how this defines their roles.
- The differences between *replication*, *transcription*, and *translation* and the role of DNA and RNA in each process.
- How eukaryotic cells modify RNA after transcription.
- How genetic material is translated into polypeptides. (LO 3.4)
- How mutations can change the amino acid sequence of a protein and be able to predict how a mutation can result in changes in gene expression. (LO 3.6)

TIP FROM THE READERS

This is the central chapter for molecular genetics. It is one of the top five chapters you must know to perform well on the AP exam!

Be sure you know the difference between *replication* (DNA to DNA), *transcription* (DNA to RNA), and *translation* (RNA to protein). In essay questions that use these terms, often 20% of the students confuse the processes!

Genes specify proteins via transcription and translation (17.1)

- **Gene expression** is the process by which DNA directs the synthesis of proteins (or, in some cases, RNAs).
- The **one gene-one polypeptide hypothesis** states that each gene codes for a polypeptide, which can be – or can constitute a part of – a protein.
- **Transcription** is the synthesis of RNA using DNA as a template. It takes place in the nucleus of eukaryotic cells.
- **Messenger RNA**, or **mRNA**, is produced during transcription. It carries the genetic message of DNA to the protein-making machinery of the cell in the cytoplasm, the *ribosome*.
- In eukaryotes, transcription results in pre-mRNA, which undergoes **RNA processing** to yield the final mRNA.
- In prokaryotes, transcription results directly in mRNA, which is not processed. Transcription and translation can occur simultaneously.
- **Translation** is the production of a polypeptide chain using the mRNA transcript and occurs at the ribosomes.

- The instructions for building a polypeptide chain are written as a series of three-nucleotide groups; this is called a *triplet code*.
- During transcription, only one strand of the DNA is transcribed, and it is called the **template strand**. The mRNA that is produced is said to be *complementary* to the original DNA strand. The mRNA base triplets are called **codons**. *They are written in the 5' to 3' direction*.
- The genetic code is *redundant*, meaning that more than one codon codes for each of the 20 amino acids. The codons are read based on a consistent reading frame – the groups of three must be read in the correct groupings in order for translation to be successful.

Transcription is the DNA-directed synthesis of RNA: a closer look (17.2)

- **RNA polymerase** is an enzyme that separates the two DNA strands and connects the RNA nucleotides as they base-pair along the DNA template strand.
- The RNA polymerases can add RNA nucleotides only to the 3' end of the strand, so RNA elongates in the 5' to 3' direction. As RNA nucleotides are added, remember that uracil replaces thymine when base pairing to adenine.
- The DNA sequence at which RNA polymerase attaches is called the **promoter**, whereas the DNA sequence that signals the end of transcription is called the **terminator**.
- A **transcription unit** is the entire stretch of DNA that is transcribed into an RNA molecule. A transcription unit may code for a polypeptide or an RNA, like transfer RNA or ribosomal RNA.
- There are three main stages of transcription:
    - **Initiation**: In bacteria, RNA polymerase recognizes and binds to the promoter. In eukaryotes, RNA polymerase II, the specific RNA polymerase that transcribes mRNA, cannot bind to the promoter without supporting help from proteins known as transcription factors. **Transcription factors** assist the binding of RNA polymerase to the promoter and, thus, the initiation of transcription. The whole complex of RNA polymerase II and transcription factors is called a **transcription initiation complex** (Figure 17.8).
    - **Elongation**: RNA polymerase moves along the DNA, continuing to untwist the double helix. RNA nucleotides are continually added to the 3' end of the growing chain. As the complex moves down the DNA strand, the double helix re-forms, with the new RNA molecule straggling away from the DNA template. Find these key steps in Figure 17.9.
    - **Termination**: After RNA polymerase transcribes a terminator sequence in the DNA, the RNA transcript is released, and the polymerase detaches.

Eukaryotic cells modify RNA after transcription (17.3)

- In eukaryotes, there are a couple of key post-transcriptional modifications to RNA – the addition of a **5' cap** and the addition of a **poly-A tail**.
- The 5' cap and the poly-A tail facilitate the export of mRNA from the nucleus, help protect the mRNA from degradation by enzymes, and facilitate the attachment of the mRNA to the ribosome.

- **RNA splicing** also takes place in eukaryotic cells. In RNA splicing, large portions of the newly synthesized RNA strand are removed. The sections of the mRNA that are spliced out are called **introns**, and the sections that remain – and subsequently spliced together by a *spliceosome* – are called **exons**. Use Figure 17.12 to help you visualize exons and introns.
- One amazing thing about how spliceosomes work is the role of a special kind of RNA, termed **small nuclear RNA** (snRNA). The snRNA plays a major role in catalyzing the excision of the introns and joining of exons. When RNA serves a catalytic role, the molecule is termed a **ribozyme**. For many years it was thought that only proteins could be catalytic, but the discovery of ribozymes totally changed that idea!
- Another rethinking that has taken place came with the realization that we have only about 20,000 genes to make approximately 100,000 polypeptides. One gene can often make more than one polypeptide. An intron removed in the production of one polypeptide can be an exon in a second polypeptide made from the same gene! Alternative RNA splicing allows for different combinations of exons, resulting in more than one polypeptide per gene.

Translation is the RNA-directed synthesis of a polypeptide: a closer look (17.4)

- In addition to mRNA, two additional types of RNA play important roles in translation: transfer RNA (tRNA) and ribosomal RNA (rRNA).
- The **tRNA** functions in transferring amino acids from a pool of amino acids in the cell's cytoplasm to a ribosome. The ribosome accepts the amino acid from the tRNA and incorporatesthe amino acid into a growing polypeptide chain.
- Each type of tRNA is specific for a particular amino acid; at one end it loosely binds the amino acid, and at the other end it has a nucleotide triplet called an **anticodon**, which allows it to pair specifically with a complementary codon on the mRNA.
- A **codon** is an mRNA triplet. Because there are four different nucleotides (A, T, C, and G), taking them three at a time results in 64 different codons. The 64 different codons include 3 stop codons and 61 codons that code for the 20 different amino acids. Most amino acids can therefore be designated by more than one codon.
- A ribosome is composed of **rRNA** and protein and has 2 subunits. The large subunit of a ribosome has three binding sites for tRNA molecules (locate each tRNA binding site in Figure 17.18).
    - A **P site**, which holds the tRNA that carries the growing polypeptide chain.
    - An **A site**, which holds the tRNA that carries the amino acid that will be added to the chain next.
    - An **E site**, which is the exit site for each tRNA.
- Translation, like transcription, can be divided into three stages:
    - **Initiation**: Organize initiation into these three steps. Use Figure 17.18 to find each step.
        - A small ribosomal subunit binds to mRNA in such a way that the first codon of the mRNA strand, which is always AUG, is placed in the proper position.

        - The tRNA with anticodon UAC, which carries the amino acid methionine, hydrogen bonds to the first codon (initiation factors are proteins that assist in holding all this together).
        - Large subunit of ribosome attaches, allowing the tRNA with methionine to attach to the P site. Notice that the A site is now available to the tRNA that will bring the second amino acid.
    - **Elongation**: Read about the three steps of elongation – codon recognition, peptide bond formation, and translocation – using Figure 17.19. This is an important idea; follow the explanations and the diagrams carefully.
    - **Termination**: A stop codon in the mRNA is reached and translation stops. A protein called release factor binds to the stop codon, and the polypeptide is freed from the ribosome (Figure 17.20).
- Polypeptides then fold to assume their specific conformation, and they are almost always modified further to render them functional. The destination of a protein is often determined by the sequence of about 20 amino acids at the leading end of the polypeptide chain. The **signal peptide**, the sequence of the leading 20 or so amino acids, serves as a sort of cellular zip code, directing proteins to their final destination.

Mutations of one or a few nucleotides can affect protein structure and function (17.5)

- Mutations are alterations in the genetic material of the cell; **point mutations** are alterations of just one nucleotide base pair of a gene. They come in two basic types:
    - A **nucleotide-pair substitution** is the replacement of one nucleotide and its partner with another pair of nucleotides.
    - **Missense mutations** are those substitutions that enable the codon to still code for an amino acid, although it might not be the correct one.
    - **Nonsense mutations** are those substitutions that change a regular amino acid codon into a stop codon, ceasing translation.
- **Insertions** and **deletions** refer to the additions and losses of nucleotide pairs in a gene. If they interfere with the codon groupings, they can cause a **frameshift mutation**, which causes the mRNA to be read incorrectly on each remaining codon.
- **Mutagens** are substances or forces that interact with DNA in ways that cause mutations. X-rays and other forms of radiation are known mutagens, as are certain chemicals.