Chapter 16

The Molecular Basis of Inheritance

Lecture Outline

Overview: Life's Operating Instructions

- In April 1953, James Watson and Francis Crick shook the scientific world with an elegant double-helical model for the structure of deoxyribonucleic acid, or DNA.

- Your genetic endowment is your DNA, contained in the 46 chromosomes you inherited from your parents and the mitochondria you inherited from your mother.

- Nucleic acids are unique in their ability to direct their own replication.

- The resemblance of offspring to their parents depends on the precise replication of DNA and its transmission from one generation to the next.

- It is this DNA program that directs the development of your biochemical, anatomical, physiological, and (to some extent) behavioral traits.

Concept 16.1 DNA is the genetic material

- After T. H. Morgan's group showed that genes exist as parts of chromosomes, the two chemical constituents of chromosomes— DNA and proteins—became the candidates for the genetic material.

- Until the 1940s, the great heterogeneity and specificity of function of proteins seemed to indicate that proteins were the genetic material.

  - However, this view was challenged by experiments with bacteria and viruses.

Experimental evidence established that DNA was the genetic material.

- The discovery of the genetic role of DNA began with research by Frederick Griffith in 1928.

- Griffith studied *Streptococcus pneumoniae*, a bacterium that causes pneumonia in mammals.

  - One strain was harmless but the other strain was pathogenic (disease-causing).

- Griffith mixed a heat-killed, pathogenic strain of bacteria with a live, harmless strain and injected this into a mouse.

  - The mouse died, and Griffith recovered the pathogenic strain from the mouse's blood.

- Griffith called this phenomenon **transformation**, a phenomenon now defined as a change in genotype and phenotype due to the assimilation of external DNA by a cell.

- American bacteriologist Oswald Avery focused on the three main candidates as the transforming substance: DNA, RNA, and protein.

  - Avery broke open the heat-killed, pathogenic bacteria and extracted the cellular contents.

- He used specific treatments that inactivated each of the three types of molecules and then tested the ability of each sample to transform harmless bacteria.
- Only DNA was able to bring about transformation.
- In 1944, Oswald Avery, Maclyn McCarty, and Colin MacLeod announced that the transforming substance was DNA.
  - Many biologists were skeptical because proteins seemed to be better candidates.
  - It was also thought that the genes of bacteria could not be similar in composition and function to those of more complex organisms.
- Further evidence that DNA was the genetic material came from studies that tracked the infection of bacteria by viruses called **bacteriophages** (or **phages** for short).
- Viruses consist of DNA (or sometimes RNA) enclosed in a protective coat of protein.
  - To replicate, a virus infects a host cell and takes over the cell's metabolic machinery.
- In 1952, Alfred Hershey and Martha Chase showed that DNA was the genetic material of the phage T2.
  - The T2 phage, consisting almost entirely of DNA and protein, attacks *Escherichia coli* (*E. coli*), a common intestinal bacteria of mammals.
  - This phage can quickly turn an *E. coli* cell into a T2-producing factory that releases phages when the cell ruptures.
- To determine the source of genetic material in the phage, Hershey and Chase designed an experiment in which they labeled protein and DNA and then tracked which component entered the *E. coli* cell during infection.
- One batch of T2 phage was grown in the presence of radioactive sulfur, marking the proteins but not DNA.
  - Another batch was grown in the presence of radioactive phosphorus, marking the DNA but not proteins.
  - Each batch was allowed to infect separate *E. coli* cultures.
- Shortly after the onset of infection, Hershey and Chase spun the cultured infected cells in a blender, shaking loose any parts of the phage that remained outside the bacteria.
  - The mixtures were spun in a centrifuge, which separated the heavier bacterial cells in the pellet from the lighter free phages and parts of phages in the liquid supernatant.
  - The pellet and supernatant of the separate treatments were tested for the presence of radioactivity.
- When bacteria were infected with T2 phages containing radiolabeled proteins, most of the radioactivity was in the supernatant containing phage particles, not in the pellet with the bacteria.
  - When they examined the bacterial cultures with T2 phage that had radiolabeled DNA, most of the radioactivity was in the pellet with the bacteria.
- Hershey and Chase concluded that the injected DNA of the phage provides the genetic information that makes the infected cells produce new viral DNA and proteins to assemble into new viruses.
- By 1947, Erwin Chargaff had developed a series of rules based on a survey of DNA composition in organisms.
  - He knew that DNA was a polymer of nucleotides consisting of a nitrogenous base, deoxyribose, and a phosphate group.

- The base can be adenine (A), thymine (T), guanine (G), or cytosine (C).
- Chargaff noted that the base composition of DNA varies from species to species.
- Chargaff found a peculiar regularity in the ratios of nucleotide bases, known as *Chargaff's rules.*
  - In any one species, the four bases are found in characteristic, but not necessarily equal, ratios.
  - In all organisms, the number of adenines is approximately equal to the number of thymines and the number of guanines is approximately equal to the number of cytosines.
  - Human DNA is 30.3% adenine, 30.3% thymine, 19.5% guanine, and 19.9% cytosine.
- The basis for these rules remained unexplained until the discovery of the double helix.

Watson and Crick discovered the double helix by building models to conform to X-ray data.

- By the early 1950s, the challenge was to determine how the structure of DNA accounted for its role in inheritance.
- Among the scientists working on the problem were Linus Pauling in California and Maurice Wilkins and Rosalind Franklin in London.
- Wilkins and Franklin used X-ray crystallography to study the structure of DNA.
  - In this technique, X-rays are diffracted as they pass through aligned fibers of purified DNA. The diffraction pattern can be used to deduce the three-dimensional shape of molecules.
- James Watson learned from this research that DNA is helical.
  - Earlier data obtained by Franklin and others suggested the width of the helix and the spacing of nitrogenous bases.
- The width of the helix suggested that it is made up of two strands, contrary to a three-stranded model that Linus Pauling had recently proposed.
- Watson and his colleague Francis Crick began to work on a model of DNA with two strands, the **double helix**.
- From an unpublished annual report summarizing Franklin's work, Watson and Crick knew Franklin had concluded that sugar-phosphate backbones were on the outside of the double helix.
- Using molecular models made of wire, they placed the sugar-phosphate chains on the outside and the nitrogenous bases on the inside of the double helix.
  - This arrangement put the relatively hydrophobic nitrogenous bases in the molecule's interior, away from the surrounding aqueous solution.
- In their model, the two sugar-phosphate backbones are **antiparallel**, with the subunits running in opposite directions.
- The sugar-phosphate chains of each strand are like the side ropes of a rope ladder.
  - Pairs of nitrogenous bases, one from each strand, form rungs.
  - The ladder forms a full turn of the helix every ten bases.
- The nitrogenous bases are paired in specific combinations: adenine with thymine and guanine with cytosine.
- Pairing like nucleotides did not fit the uniform diameter indicated by the X-ray data.

- A purine-purine pair is too wide, and a pyrimidine-pyrimidine pair is too short.
- Only a pyrimidine-purine pair produces the 2-nm diameter indicated by the X-ray data.

- In addition, Watson and Crick determined that chemical side groups of the nitrogenous bases form hydrogen bonds, connecting the two strands.
  - Based on details of their structure, adenine forms two hydrogen bonds only with thymine, and guanine forms three hydrogen bonds only with cytosine.
  - This finding explained Chargaff's rules.

- The base-pairing rules dictate the combinations of nitrogenous bases that form the "rungs" of DNA.
  - The rules do not restrict the sequence of nucleotides *along* each DNA strand.

- The linear sequence of the four bases can be varied in countless ways, and each gene has a unique order of nitrogenous bases.

- In April 1953, Watson and Crick published a succinct, one-page paper in *Nature* reporting their double helix model of DNA.
  - The beauty of the double helix model was that the structure of DNA suggested the basic mechanism of its replication.

Concept 16.2 Many proteins work together in DNA replication and repair

- The specific pairing of nitrogenous bases in DNA was the flash of inspiration that led Watson and Crick to the correct double helix.

- The possible mechanism for the next step, the accurate replication of DNA, was clear to Watson and Crick from their double helix model.

During DNA replication, base pairing enables existing DNA strands to serve as templates for new complementary strands.

- In a second paper, Watson and Crick published their hypothesis for how DNA replicates.
  - Because each strand is complementary to the other, each can form a template when separated.
  - The order of bases on one strand guides the addition of complementary bases and therefore duplicates the pairs of bases exactly.

- When a cell copies a DNA molecule, each strand serves as a template for ordering nucleotides into a new complementary strand.
  - One at a time, nucleotides line up along the template strand according to the base-pairing rules. The nucleotides are linked to form new strands.

- Watson and Crick's **semiconservative model** of replication predicts that when a double helix replicates, each of the daughter molecules has one old strand and one newly made strand.

- Two competing models—a conservative model and a dispersive model—were also considered.
  - In the conservative model of replication, the two parental strands come back together after the process.
  - In the dispersive model, all four strands of DNA following replication have a mixture of old and new DNA.

- Experiments in the late 1950s by Matthew Meselson and Franklin Stahl supported the semiconservative model proposed by Watson and Crick over the other two models.
- Meselson and Stahl labeled the nucleotides of the old strands with a heavy isotope of nitrogen (^{15}N) and the new nucleotides with a lighter isotope (^{14}N).
  - Replicated strands could be separated by density in a centrifuge.
  - Each model—the semiconservative model, the conservative model, and the dispersive model—made a specific prediction about the density of replicated DNA strands.
  - The first replication in the ^{14}N medium produced a band of hybrid (^{15}N-^{14}N) DNA, eliminating the conservative model.
  - A second replication produced both light and hybrid DNA, eliminating the dispersive model and supporting the semiconservative model.

A large team of enzymes and other proteins carries out DNA replication.

- It takes *E. coli* less than an hour to copy each of the 4.6 million nucleotide pairs in its single chromosome and divide to form two identical daughter cells.
  - A human cell can copy its 6 billion nucleotide pairs and divide into daughter cells in only a few hours.
- This process is remarkably accurate, with only one error per 10 billion nucleotides.
- More than a dozen enzymes and other proteins participate in DNA replication.
- Much more is known about replication in bacteria than in eukaryotes.
  - However, the process is fundamentally similar for prokaryotes and eukaryotes.
- The replication of a DNA molecule begins at special sites called **origins of replication**.
- In bacteria, this site is a specific sequence of nucleotides that is recognized by the replication enzymes.
  - These enzymes separate the strands, forming a replication "bubble."
  - Replication proceeds in both directions until the entire molecule is copied.
- In eukaryotes, there may be hundreds or thousands of origin sites per chromosome.
  - At the origin sites, the DNA strands separate, forming a replication "bubble" with **replication forks** at each end, where the parental strands of DNA are being unwound.
  - The replication bubbles elongate as the DNA is replicated, and eventually fuse.
- Several kinds of proteins participate in the unwinding of parental strands of DNA.
  - **Helicases** untwist the double helix and separate the template DNA strands at the replication fork.
  - **Single-strand binding proteins** bind to unpaired DNA strands, stabilizing them.
  - This untwisting causes tighter twisting ahead of the replication fork; **topoisomerase** helps relieve this strain.
- Unwound sections of parental DNA strands are now available to serve as templates for the synthesis of new complementary DNA strands.
- However, the enzymes that synthesize DNA cannot *initiate* synthesis of a polynucleotide.
  - They can only add nucleotides to the end of an existing chain that is base-paired with the template strand.

- The initial nucleotide chain is a short stretch of RNA called a **primer**, synthesized by the enzyme **primase**.
  - Primase starts a complementary RNA chain from a single RNA nucleotide, adding RNA nucleotides one at a time, using the parental DNA strand as a template.
  - The completed primer is generally five to ten nucleotides long.
- The new DNA strand starts from the 3′ end of the RNA primer.
- As nucleotides align with complementary bases along the template strand, they are added to the growing end of the new strand by the polymerase.
- Enzymes called **DNA polymerases** catalyze the synthesis of new DNA by adding nucleotides to a preexisting chain.
  - In *E. coli*, two different DNA polymerases play major roles in replication: DNA polymerase III and DNA polymerase I.
  - In eukaryotes, at least 11 different DNA polymerases have been identified so far.
- Most DNA polymerases require a primer and a DNA template strand, along which complementary DNA nucleotides line up.
  - The rate of elongation is about 500 nucleotides per second in bacteria and 50 per second in human cells.
- Each nucleotide added to a growing DNA strand comes from a nucleoside triphosphate, a nucleoside (a base and a sugar) with three phosphate groups.
  - ATP is a nucleoside triphosphate with ribose instead of deoxyribose.
- Like ATP, the triphosphate monomers used for DNA synthesis are chemically reactive, partly because their triphosphate tails have an unstable cluster of negative charge.
  - As each nucleotide is added to the growing end of a DNA strand, the last two phosphate groups are hydrolyzed to form pyrophosphate: $P—P_i$.
  - The exergonic hydrolysis of pyrophosphate to two inorganic phosphate molecules drives the polymerization of the nucleotide to the new strand.

The two strands of DNA in the double helix are antiparallel.

- The two strands of DNA in a double helix are **antiparallel**, oriented in opposite directions to each other.
  - Each DNA strand has a 3′ end with a free hydroxyl group attached to deoxyribose and a 5′ end with a free phosphate group attached to deoxyribose.
  - The 5′→3′ direction of one strand runs opposite the 3′→5′ direction of the other strand.
- The two new strands formed during DNA replication must also be antiparallel to their template strands.
  - DNA polymerases can add nucleotides only to the free 3′ end of a growing DNA strand.
  - A new DNA strand can elongate only in the 5′→3′ direction.
- Let's examine one of two replication forks in a bubble.
- Along one template strand, DNA polymerase III can synthesize a complementary strand continuously by elongating the new DNA in the mandatory 5′→3′ direction.
  - DNA pol III remains in the replication fork on that template strand and continuously adds nucleotides to the new complementary strand as the fork progresses.

- ○ The DNA strand made by this mechanism, called the **leading strand**, requires a single primer.
- The other parental strand (5′→3′ into the fork), the **lagging strand**, is copied *away from* the fork.
  - ○ Unlike the leading strand, which elongates continuously, the lagging stand is synthesized as a series of short segments called **Okazaki fragments**.
  - o Okazaki fragments are about 1,000 to 2,000 nucleotides long in *E. coli* and 100 to 200 nucleotides long in eukaryotes.
  - o Although only one primer is required on the leading strand, each Okazaki fragment on the lagging strand must be primed separately.
- Another enzyme, **DNA ligase**, joins the sugar-phosphate backbones of the Okazaki fragments to form a single DNA strand.
- Another DNA polymerase, DNA polymerase I, replaces the RNA nucleotides of the primers with DNA versions, adding them one by one to the 3′ end of the adjacent Okazaki fragment.
  - o DNA polymerase I cannot join the final nucleotide of the replacement DNA segment to the first DNA nucleotide of the Okazaki fragment whose primer was just replaced.
- **DNA ligase** joins the sugar-phosphate backbones of all the Okazaki fragments into a continuous DNA strand.
- To summarize: At the replication fork, the leading strand is copied continuously into the fork from a single primer. The lagging strand is copied away from the fork in short segments, each requiring a new primer.

The various proteins that participate in DNA replication form a single large complex, a "DNA replication machine."

- It is conventional and convenient to think of the DNA polymerase molecules as moving along a stationary DNA template.
- In reality, the various proteins involved in DNA replication form a single large complex, a DNA replication "machine."
- Many protein-protein interactions facilitate the efficiency of this machine.
  - ○ For example, by interacting with other proteins at the fork, primase slows the progress of the replication fork and coordinates the rate of replication on the leading and lagging strands.
- The DNA replication machine is probably stationary during the replication process.
- In eukaryotic cells, multiple copies of the machine may anchor to the nuclear matrix, a framework of fibers extending through the interior of the nucleus.
- Two DNA polymerase molecules, one on each template strand, "reel in" the parental DNA and extrude newly made daughter DNA molecules.
- The lagging strand is looped back through the complex.

Enzymes proofread DNA during its replication and repair damage in existing DNA.

- Mistakes during the initial pairing of template nucleotides and complementary nucleotides occur at a rate of one error per 100,000 base pairs.
- DNA polymerase proofreads each new nucleotide against the template nucleotide as soon as it is added.

- ○ If there is an incorrect pairing, the enzyme removes the wrong nucleotide and then resumes synthesis.
- ○ The final error rate is only one mistake per 10 billion nucleotides.
- Mismatched nucleotides that are missed by DNA polymerase or mutations that occur after DNA synthesis is completed can often be repaired.
- In **mismatch repair**, enzymes fix incorrectly paired nucleotides.
- ○ A hereditary defect in one of these enzymes is associated with a form of colon cancer.
- Incorrectly paired or altered nucleotides can also arise after replication.
- DNA molecules are constantly subjected to potentially harmful chemical and physical agents, such as X-rays and cigarette smoke.
- DNA bases also undergo spontaneous chemical changes under normal cellular conditions.
- ○ These changes are usually corrected before they become permanent changes or *mutations*.
- Each cell continuously monitors and repairs its genetic material, with 100 repair enzymes known in *E. coli* and more than 130 repair enzymes identified in humans.
- Many cellular systems for repairing incorrectly paired nucleotides use a mechanism that relies on the base-paired structure of DNA.
- o A **nuclease** cuts out a segment of a damaged strand, and the resulting gap is filled in with nucleotides, using the undamaged strand as a template.
- o One such DNA repair system is called **nucleotide excision repair**.
- DNA repair enzymes in skin cells repair genetic damage caused by ultraviolet light, which can produce *thymine dimers* between adjacent thymine nucleotides.
- ○ This buckles the DNA double helix and interferes with DNA replication.
- The importance of the proper functioning of the enzymes that repair damage to DNA is clear from the inherited disorder xeroderma pigmentosum. Individuals with this disorder are hypersensitive to sunlight.
- ○ Mutations in their skin cells are left uncorrected and cause skin cancer.

Altered DNA nucleotides have evolutionary significance.

- Accurate replication and repair of DNA damage are important for the functioning of the organism and for passing on a complete, accurate genome to the next generation.
- Although the error rate is extremely low, rare mistakes – mutations – do slip through.
- Mutations can change the phenotype of an organism. If they occur in germ cells, which give rise to gametes, they can be passed on from generation to generation.
- ○ The vast majority of such changes are harmful, but a very small percentage can be beneficial.
- Mutations are the source of the variation on which natural selection operates during evolution and are ultimately responsible for the appearance of new species.
- ○ The balance between accuracy of DNA replication or repair and a low mutation rate has allowed the evolution of the rich diversity of species we see on Earth today.

The ends of DNA molecules are replicated by a special mechanism.

- Limitations of DNA polymerase create problems for the linear DNA of eukaryotic chromosomes and provide no way to complete the 5′ ends of daughter DNA strands.
  - Repeated rounds of replication produce shorter and shorter DNA molecules.
- Prokaryotes, with their circular DNA molecules, do not have this problem.
- The ends of eukaryotic chromosomal DNA molecules have special nucleotide sequences called **telomeres**.
  - Telomeres do not contain genes. Instead, the DNA typically consists of multiple repetitions of one short nucleotide sequence.
  - In human telomeres, this sequence is typically TTAGGG, repeated between 100 and 1,000 times.
- Telomeric DNA acts as a kind of buffer zone that protects the organism's genes.
- Telomeric DNA and specific proteins associated with it prevent the staggered ends of the daughter molecule from activating the cell's system for monitoring DNA damage.
- Telomeres become shorter during every round of replication.
  - Telomeric DNA tends to be shorter in the dividing somatic cells of older individuals and in cultured cells that have divided many times.
- It has been suggested that the shortening of telomeres is connected with the aging process of certain tissues and even to aging in general.
- Eukaryotic cells have evolved a mechanism to restore shortened telomeres in germ cells.
  - If the chromosomes of germ cells became shorter with every cell cycle, essential genes would eventually be lost.
- An enzyme called **telomerase** catalyzes the lengthening of telomeres in eukaryotic germ cells, restoring their original length.
- Telomerase is not present in the somatic cells of multicellular organisms.
- Normal shortening of telomeres may protect organisms from cancer by limiting the number of divisions that somatic cells can undergo.
  - Cells from large tumors often have unusually short telomeres because they have gone through many cell divisions.
- Active telomerase has been found in some cancerous somatic cells.
  - Active telomerase may overcome the progressive shortening that would lead to self-destruction of the cancer.
  - Immortal strains of cultured cells are capable of unlimited cell division.
- Telomerase may provide a useful target for cancer diagnosis and chemotherapy.

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

- Most bacteria have a single, circular, double-stranded DNA molecule associated with a small amount of protein.
  - The *bacterial chromosome* is very different from a eukaryote chromosome, which is a linear DNA molecule associated with a large amount of protein.

- The chromosomal DNA of *E. coli* consists of about 4.6 million nucleotide pairs, representing about 4,400 genes.
  - Stretched out, this DNA would be about a millimeter long, 500 times longer than an *E. coli* cell.
- A bacterium has a dense region of DNA called the **nucleoid**.
  - Within the nucleoid, proteins cause the chromosome to tightly coil and "supercoil," densely packing it so that it fills only part of the cell.
- Each eukaryotic chromosome contains a single linear DNA double helix that, in humans, has an average of about 1.5×10^8 nucleotide pairs.
  - If stretched out, such a DNA molecule would be about 4 cm long, thousands of times the diameter of a cell nucleus.
- In the cell, eukaryotic DNA is packaged with protein to form **chromatin**.
- The packing of chromatin varies during the course of the cell cycle.
  - It is highly extended during interphase.
  - As a cell enters mitosis, the chromatin coils and condenses, forming short, thick metaphase chromosomes.
- Although interphase chromatin is generally much less condensed than the chromatin of mitotic chromosomes, it shows several of the same levels of higher-order packing.
  - Some of the chromatin is present as a 10-nm fiber, but much is compacted into a 30-nm fiber, which may be further folded into looped domains.
- The looped domains of an interphase chromosome are attached to the nuclear lamina, on the inside of the nuclear envelope, and perhaps also to fibers of the nuclear matrix.
  - These attachments may help organize regions where genes are active.
- The chromatin of each chromosome occupies a specific restricted area within the interphase nucleus, preventing tangling of the chromatin fibers of different chromosomes.
- Even during interphase, the centromeres and telomeres of chromosomes exist in a highly condensed state similar to that seen in a metaphase chromosome.
  - Condensed chromatin is called **heterochromatin**, while more dispersed chromatin is called **euchromatin**.
- Because of its compaction, heterochromatic DNA is largely inaccessible to the machinery in the cell responsible for expressing the genetic information coded in the DNA.
  - More loosely packed euchromatin makes its DNA accessible to this machinery, so the genes present in euchromatin can be expressed.
- The chromosome is a dynamic structure that is condensed, loosened, modified, and remodeled as necessary for various cell processes, including mitosis, meiosis, and gene expression.
  - Chemical modifications of histones affect the state of chromatin condensation and affect gene activity.