Chapter 9
DNA

- First evidence that genes were not made of proteins was made by *Fred Griffith* in 1928.
- He worked with 2 strains of bacterial pneumonia (with the use of mice) that contained different genetic information.
- One strain of bacteria had a genetic code for which an external capsule was produced to protect the bacteria from attack by the animal's immune system (virulent strain). When injected into the mice, the animal died. The other strain could not produce a capsule and when injected in mice their immune system attacked the bacteria and did not kill the mice (nonvirulent strain).
- Griffith found that nonvirulent bacteria could be changed to the virulent form. Virulent bacteria were killed (by heat) and when injected into mice the mice did not die. – but if the bacteria was killed and injected into mice with the nonvirulent bacteria some mice died. It was found that the dead mice did contain the living virulent bacteria even though none had been injected. So Griffith drew the conclusion that some of the genetic material from the dead virulent bacteria must have entered the living nonvirulent bacteria, making them virulent.
- This phenomena is called *bacterial transformation*, which is involved in the transfer of genetic information from one bacterium to another. Refer to Fig 9-1 in text.
- Oswald Avery and colleagues worked for many years to determine what caused bacterial transformation. He demonstrated that the bacterial transformation Griffith studied could be produced by DNA from the virulent bacteria and not by their proteins or other substances. The transforming DNA gets incorporated into recipient cell's DNA giving the cell the genetic traits carried by the transforming DNA.
- Further evidence that *genetic material is DNA*, came from studies of phages (bacterial viruses) (Hershey and Chase). A phage is made of a DNA molecule inside a protein coat.
- Hershey and Chased tested a hypothesis that phages do not enter bacteria intact; instead the protein coat attaches to a cell wall and injects its DNA into the bacterium.
- These researchers used radioactive sulfur and phosphorus to distinguish between the phages protein and DNA. Their research found that *DNA was the hereditary material of viruses*.

Some background information

Chromosomes are paired structures that are made up of strands of *chromatin*, which contains DNA and protein, in roughly equal amounts (and occurs in cytoplasm of prokaryotic cells and nucleus of eukaryotic cells). In humans, there are 46 chromosomes (23 pairs) in the cell nucleus of regular cells of the body – called somatic cells – as opposed to the gametes, which are sperm and egg cells that contain only 23 unpaired chromosomes. A chromosome has a short arm and a long arm. The arms are held together by a *centromere*, which is a specialized area that provides a handle by which the chromosome can be moved around and consists of short DNA sequences repeated hundreds of times.

DNA – deoxyribonucleic acid – is the chemical molecule that serves as genetic material. A strand of DNA is a long chain (a polymer) of nucleotides (genes). Each nucleotide of DNA contains:

- a nitrogenous base
- a sugar with 5 carbon molecules (deoxyribose)
- a phosphate group

There are 4 different kinds of nitrogenous bases:

- adenine
- thymine
- cytosine
- guanine

The nitrogenous bases (and therefore the nucleotides) can be and are different throughout the long chain of DNA. DNA exists inside the chromosomes.

Genes lie along the chain of DNA. They are mad up of sections of nucleotides. Some genes can have many nucleotides; others just a few. Humans have thousands of different genes, which reside on different chromosomes, but on the same chromosome in all people. For example, the gene for cystic fibrosis is always found in the same location on gene #7 in all humans. The gene for Huntington's disease is on chromosome #4 in all humans. The gene for sickle cell anemia is on chromosome #11 in all humans. However, not all humans have all diseases. Some genes are expressed (they show the effect) whereas, other genes are repressed (they do not show an effect).

Alleles are the different forms of a trait. For example, the gene for hair color resides in a certain location on a certain gene in all humans. However, humans can have many different shades of hair color; the different shades are represented by different alleles. Say that the gene for hair color is on gene #2 (I do not know that it is, I am just giving you an example). You have a pair of #2 genes – 1 from mom and 1 from dad. Suppose that your gather has dark brown hair, and mom is a blonde. Brown hair color is more dominant that blond hair, so I will call it *H* – (*H* for hair, capital *H* for dominance). On one of your #2 chromosomes, you have a pair of *H* alleles at the location of the gene for hair color. But you also have your mom's genes. I will call her hair color *h* – (*h* for hair, but lowercase to represent a less dominant – recessive – color). On your other #2 chromosome you have a pair of h alleles at the location where the gene for hair color lies. Both *H* and *h* are alleles of a gene that controls a certain trait.

Nucleic Acids: large molecules that carry tons of small details: all the genetic information. Nucleic acids are found in every living thing – plants, animals, bacteria, viruses, fungi – that uses and converts energy. There are 2 types of nucleic acids: DNA and RNA. Nucleic acids are made up of strands of nucleotides, which are made up of a base containing nitrogen (called nitrogenous base), a sugar that contains a 5-carbon molecule and a phosphoric acid. That is it. Your entire genetic composition, personality, and maybe even intelligence hinges on molecules contain a nitrogen compound, some sugar, and an acid. However, DNA does not encode things you learn or experience during your life.

The nitrogenous bases are molecules either called purines or pyrimidines.

Purines include:

- adenine
- guanine

Pyrimidines include:

- Cytosine
- Thymine (in DNA)
- Uracil (in RNA)

DNA

DNA contains 2 strands of nucleotides arranged in a way that makes it look like a twisted ladder (double helix). (X-ray diffraction photographs produced in 1952 by Rosalind Franklin showed this double helix.)

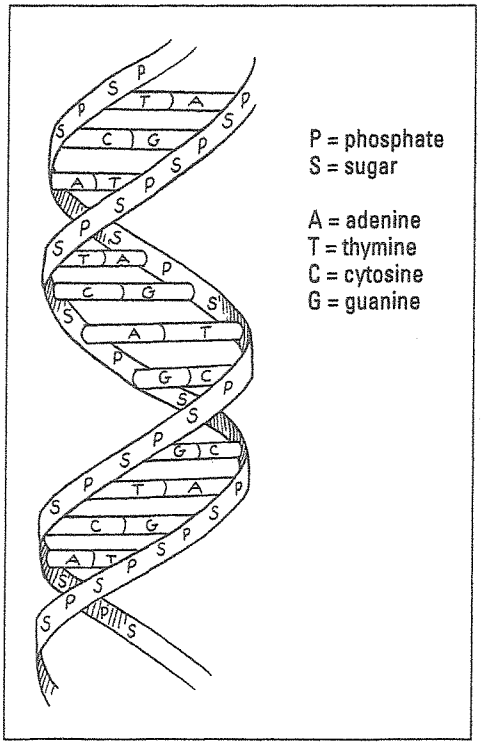

Fig. 1

The nitrogenous bases that DNA builds its double-helix upon are adenine (A), guanine(G), cytosine(C), and thymine (T). The sugar that is in the composition of DNA is 2-deoxyribose.

Refer to the above figure – ALL NITROGENOUS BASES ARE PAIRED - adenine is always paired with thymine (A-T), and guanine is always paired with cytosine (G-C). These bases are held together by hydrogen bonds, which form the "rungs" of the twisted ladder. The sides of the ladder are made up of the sugar (deoxyribose) and phosphate molecule held together by a covalent bond.

Certain sections of nitrogenous bases along the strand of DNA form a gene. A gene is a unit that contains the genetic information of codes for a particular product and transmits hereditary information to the next generation. But genes are not found only in reproductive cells. Every cell in an organism contains DNA (and therefore genes) because DNA also codes for the proteins that the organism produces. And proteins control cell function and provide structure. So, the basis of life happens in each and every cell. Whenever a new cell is made in an organism, the genetic material is reproduced and put into the new cell. The new cell can then create proteins within itself and also pass on the genetic information.

The order of the nitrogenous bases on a strand of DNA (or in a section of the DNA that comprises a gene) determines which amino acid is produced. And the order that amino acids are strung together determines which protein is produced. Which protein is produced determines what structural element is produced within your body (such as muscle tissue, skin, or hair) or what function can be performed (such as if hemoglobin is being produced to transport oxygen to all the cells.)

Every cellular process and every aspect of metabolism is based on genetic information and the production of the proper proteins. If the wrong protein is produced (such as in sickle cell disease), then disease occurs.

Watson-Crick model of DNA

- Structure consists of 2 strands of DNA.
- The 2 strands area arranged like a twisted "rope ladder", with the ladder's sides being the sugar-phosphate backbones of the 2 strands and the rungs being the bases.
- The "rungs" are made of complementary paired nitrogen-containing bases:
    - C joins G
    - A joins T
- If the nucleotide base sequence if one strand of a DNA molecule reads:

  TTG/CCA/CTG/ACC it will have a **complementary** strand which reads
  AAC/GGT/GAC/TGG

- A-T and C-G from the most stable combinations of hydrogen bonds (known as the *phenomenon of complementarity*.
- In order for the hydrogen bonds to from properly between the base pairs in DNA, the 2 nucleotide strands of the DNA molecule had to run in opposite directions. This is known as the *antiparallel phenomenon.*

DNA Replication

Your DNA does not copy itself – it replicates. Every cell in your body needs to be replaced periodically. Cells never stop working and eventually wear out. Cell turnover, as it is called, happens constantly. Blood cells need to be replaced every 120 days. But not every blood cell is replaced every 4 months. On any given day, your body can be replacing some blood cells, some skin cells, some hair cells, and some mucous cells. Whatever is on the body's to-do list, the process of DNA replication is the same.

Remember DNA looks like a twisted ladder, with the nucleotide bases forming the "rungs" of the ladder. During replication (see figure below) the DNA strand must "unzip" so that the rungs are split apart with 1 nucleotide on one side and 1 nucleotide on the other. Each side of the original DNA strand becomes a *template strand* upon which the new *complementary strand* forms (called *semiconservative DNA replication because each new double helix contains one old (template) and one newly formed strand).*

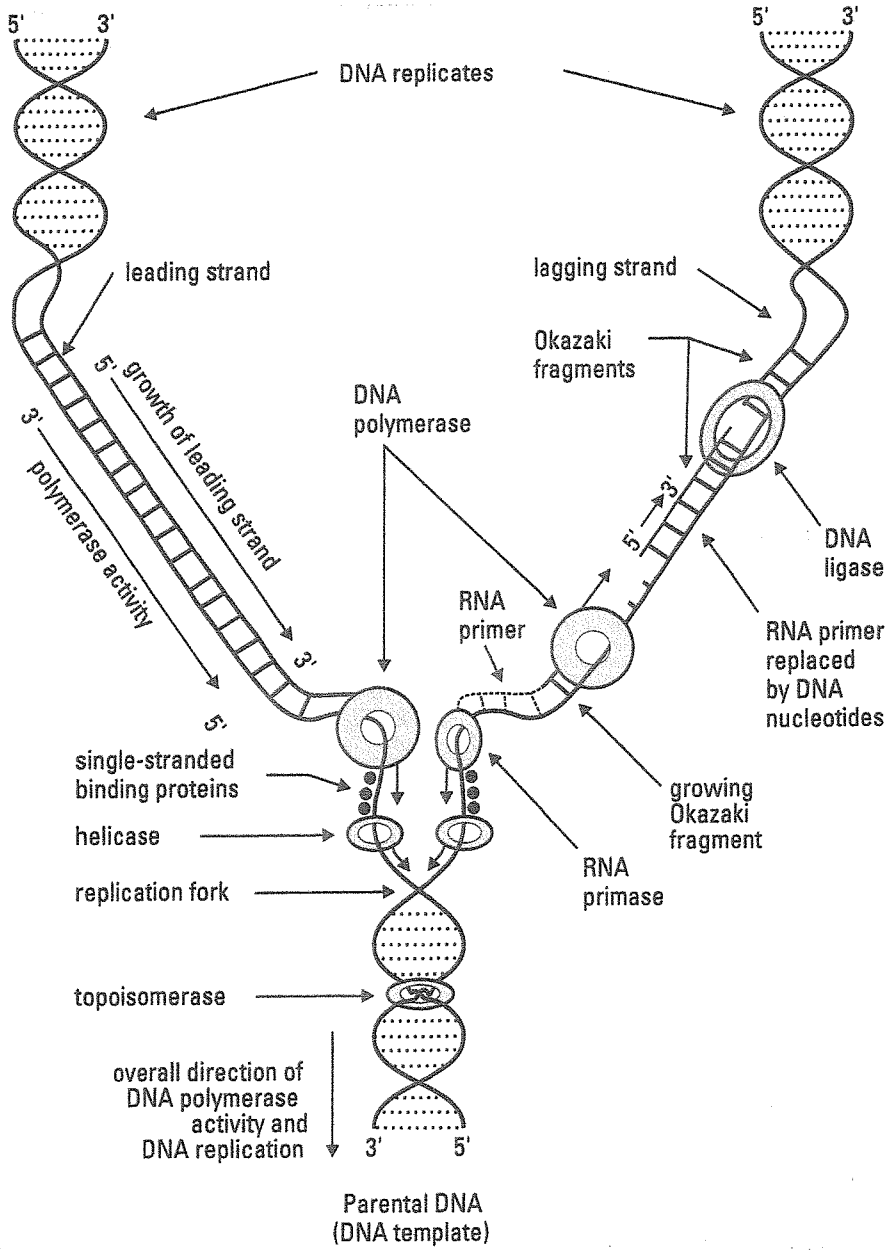

Fig. 2

The unzipping of the DNA helix is initiated by the *enzyme helicase*. The entire DNA strand does not unzip all at one time, however. Only part of the original DNA strand opens up at one time. When the top part of the helix is opened, the original DNA strand looks like a Y. This partly open/partly closed area where replication is going on is called the *replication fork.*

Note in the above figure the numbers 5' and 3' (read "5 prime" and "3 prime"). These numbers indicate the direction in which replication is occurring. Each template strand of DNA is "read" in the 3'5' direction. And, because the bases that are complementary (opposite) to the template strand are added, the complementary strand "grows" in the 5'3' direction – (also note in the figure above that the enzyme DNA polymerase is involved in this replication.)

The template strand tells some great stories, so you will want to know how to read it. Remember the nitrogen bases that make up each nucleotide along the strand of DNA include adenine, guanine, cytosine, and thymine, which are abbreviated as A, G, C, and T. In a molecule of DNA, A pairs with T and C pairs with G.

As the enzyme DNA polymerase moves along the template strand, if a base says A, than a T is added to the growing complementary strand. If a base on the template strand says G, then DNA polymerase adds a C to the growing complementary strand.

Remember the order of the bases is important because the order of bases delineate the genes, and the genes dictate what amino acids are produced, and the amino acids determine which proteins are produced, and proteins are needed in every cell of you body. Proteins make up cell structures themselves, as well as enzymes that initiate cellular processes that keep you alive. It all stars with the DNA, though.

Again refer to the above figure. See how the left side is opening and growing smoothly? The DNA polymerase works continuously on that side, and that side is called the *leading strand.* The other side (5'3') looks a little messier because the process does not occur smoothly. On that strand, called the *lagging strand*, the DNA polymerase reads the template strand and assembles the new bases in fragments. These fragments (separate and shorter pieces) are called *Okazaki fragments*, and they are them joined together by the *enzyme DNA ligase* to from the new complementary strand.

Now that you know what the nitrogenous bases of DNA do, I will tell you what the phosphate and sugar molecules of DNA do. The replicating DNA strand needs energy to go through the steps of reading the template, producing the complementary base, and joining the base to the growing strand. The molecules of the sugar deoxyribose provide that energy. The phosphate bonds that are broken apart when the original strand of DNA "unzips" provides the chemical energy needed to get the whole process started. Nature certainly is well organized, isn't it?

DNA replication and eukaryotes/prokaryotes – just remember this:

In DNA replication, prokaryotes have only 1 replication origin and 1 replication fork. Some bacteria may contain one or more circular DNA molecules called *plasmids*. (Plasmids are used in genetic engineering.) A bacterium replicates any plasmids it contains at the same time as its main DNA. Also prokaryotes have 2 "proofreading" mechanism. This allows few errors - approximately 1 copy error for every 250 Escherichia coli cells replicating their DNA.

Eukaryotes may have many replication origins and many replication forks. Plasmids may occasionally be found in yeast. Each chromosome contains a single DNA molecule extending from one end of the chromosome to the other, but coiled and folded many times. The DNA is associated with various proteins, forming a substance called *chromatin*. The chromosomal proteins can be divided into histones. Histone proteins (structural proteins with a high content of the amino acids arginine and lysine)

- are responsible, in part, for packing eukaryotic DNA into short, tight packages.
- have DNA wound around clusters of this material (histones)
- have a high content of (+) charged amino acids which are attracted to the net (-) charges of the phosphate groups of the DNA molecule
- are associated with the condensed state of DNA (chromatin packed as tightly as possible)

DNA Repair

Changes in the DNA molecules of your cells actually occur on a daily basis and with high frequency. DNA repair depends on the existence of 2 copies of the genetic information, one in each strand of the double helix. As long as one strand remains undamaged, the repair enzymes, of which there are more than 20 kinds, can use it as a template to replace a damaged segment in its partner. Most of the damage is repaired unless both strands are damaged beyond recognition at the same time. The following are examples of DNA repair:

- You are exposed to background radiation from your surroundings.
- You are exposed to mutagenic chemicals from a polluted environment.
- The damage is caused by single-stranded viral DNA.
- Both repair enzymes molecules are damaged and inoperable.

Mistakes (Mutations)

If a mistake in a new strand of DNA goes undetected or unrepaired, the mistake becomes a mutation. A *mutation* is a deviation from the original DNA strand. The nucleotides are not in the same sequence. Although mutations can and do cause serious defects, all mutations are not bad. A mutation may result from

- Environmental cause
- The addition or loss of one or more nucleotides (deletions or insertions)
- The substitution of one nucleotide for another during replication (substitutions)

FYI:
Substitution: These types of mutations occur when the wrong nucleotide is put in for another nucleotide. For example, if the code for a particular gene read 5'-A-T-C-G-T-C-A-G-3', the correct complementary sequence for the code on the new strand of DNA would be 3'-T-A-G-C-A-G-T-C-5'.

Genetic code is written in a specific direction. Because DNA is a double helix in which 2 strands intertwine, confusion can easily be created when trying to keep track of the ends of the strands. To avoid confusion, one strand of DNA is labeled 3' (3 prime), and the other is labeled 5' (5 prime). The convention is to read the strand in the 5' to 3' direction.

A substitution mutation occurs if the newly created sequence reads something like this: 3'-T-A-C-C-T-C-A-G-5'. The third base over should be guanine (G) instead of cytosine(C). That base could have been passed over during the "reading" of the strand of DNA, or a new C could have been put in instead of the G. In either case, it's wrong, so it's a mutation. Because it is just one base, it's called a *point mutation.* Chances are that the protein that gene creates would not be affected. If so, the mistake is called a *silent mutation.*

Deletions: If, during the creation of a new strand of complementary DNA, a nucleotide is read by the complementary base is not inserted, the complementary strand is missing a nucleotide. This type of mutation is called a deletion. Deletions can cause serious diseases. Cystic fibrosis is a disease that causes the lungs to continually fill with thick mucus, which can harbor bacteria and cause serious cases of pneumonia, as well as other problems. People with cystic fibrosis often do not live past their 20's or 30's. Cystic fibrosis is caused by a teeny little deletion on chromosome #7. Duchenne muscular dystrophy, the most common type of the disease (that the Jerry Lewis Telethon raises money for every Labor Day weekend), also is caused by a deletion on a chromosome. Muscular dystrophies are genetic defects that lead to muscle deterioration, which can be quite serious. People lose their ability not only to walk, but Duchenne muscular dystrophy eventually lead to sever weakening of the muscles that allow breathing and of the heart muscle.

Insertions: If an extra nucleotide is slipped into a newly developing complementary strand, the rest of the strand is read wrong. This type of mutation is called a *frameshift mutation* because the reading of the "frames" of genetic code is shifted (think of each nucleotide as a frame on a piece of film). One well-known disease that is caused by the addition of nucleotides is Huntington's disease, in which the sequence C-A-G is inserted up to 100 times into a normal gene. Although the sequence is a multiple of 3 (so technically not a frameshift mutation), the abundance of these insertions screw up the reading of the normal genetic code, causing abnormal protein production or a lack of protein production. In people with Huntington's disease, the nervous system degenerates starting when a person is in his or her 30's or 40's. Another disease caused by an insertion mutation became well know when a movies was made about a person affected with the disease. The movie "Elephant Man," gave an account of a man affected with neurofibromatosis, which causes deformities. The cause of this disease is the insertion of DNA sequences that do not code for anything right into the middle of DNA sequences making up a gene that does code for certain proteins. When these noncoding sequences are stuck into the gene, the code for the normal gene cannot be read, and errors in protein production occur.

Mutagenic agents (i.e. chemicals, certain kinds of radiation) that cause mutations can cause

- Somatic mutations (mutation in a body cell)
- Germ cell mutations (mutations in cells destined to form eggs or sperm)
- Cancer

Possible biological results of a mutation that involve a change at a single nucleotide base pair are

- It could cause death
- Its effect could be so slight as to be undetectable
- It could cause an effect in the next generation

Genomes

Genomes: total collection of DNA (genetic material) in any individual. They

- carry instructions for making all of an individual's proteins.
- dictate the sequence of nucleotides in the RNA of the ribosomes.
- regulate the activity of parts of the genome.

Mobile Genetic Elements

Barbara McClintock first cast doubt on the picture of a static genome. She discovered that corn plants have *transposable elements* (jumping genes), segments of DNA that can move from one part of the genome to another. These transposable elements contain "programs" for operating parts of the genome.

Some bacteria have small mobile elements called *transposons*, which are responsible for the passing of the gene ampicillin resistance from one bacterium to another. They can move from the cell's main DNA molecule into a smaller plasmid and back again. (Remember plasmids are circular DNA molecule, found in some bacteria and fungi in addition to the organism's own genome.)

Meselson and Stahl Experiment

- In the experiment detailed in the "How Do We Know" essay Meseleson and Stahl found that the hybrid DNA molecules of the first generation consisted of half containing only ^{14}N and half containing only ^{15}N. The best interpretation of their first generation results were that they supported both the semiconservative and *dispsersive replication* hypothesis. (Dispersive replication is when parental *DNA is broken up into short segments used as templates for the formation of new segments, which* are *then somehow joined together*. At cell division, each new sell inherits a DNA molecule with some old and some new nucleotides in each strand.).
- In the Meselson and Stahl experiment it was necessary to grow bacteria for a second generation in the normal nitrogen(^{14}N) medium in order to eliminates the dispersive replication hypotheses
- Please note it is now clear that the DNA of all organisms is replicated semiconservatively.