THE MOLECULAR BASIS OF INHERITANCE

## Introduction

* 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 the DNA you inherited from your parents.
* 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.

# A. DNA as the Genetic Material

**1. The search for genetic material led to DNA**

* Once T.H. Morgan’s group showed that genes are located on chromosomes, the two constituents of chromosomes - proteins and DNA - were 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 was not consistent with experiments with microorganisms, like bacteria and viruses.
* The discovery of the genetic role of DNA began with research by Frederick Griffith in 1928.
* He studied *Streptococcus pneumoniae*, a bacterium that causes pneumonia in mammals.
* One strain, the R strain, was harmless.
* The other strain, the S strain, was pathogenic.
* In an experiment Griffith mixed heat-killed S strain with live R strain bacteria and injected this into a mouse.
* The mouse died and he recovered the pathogenic strain from the mouse’s blood.
* Griffith called this phenomenon **transformation**, a change in genotype and phenotype due to the assimilation of a foreign substance (now known to be DNA) by a cell.
* For the next 14 years scientists tried to identify the transforming substance.
* Finally in 1944, Oswald Avery, Maclyn McCarty and Colin MacLeod announced that the transforming substance was DNA.
* Still, many biologists were skeptical.
* In part, this reflected a belief 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 was derived from studies that tracked the infection of bacteria by viruses.
* Viruses consist of a DNA (sometimes RNA) enclosed by a protective coat of protein.
* To replicate, a virus infects a host cell and takes over the cell’s metabolic machinery.
* Viruses that specifically attack bacteria are called **bacteriophages** or just **phages**.
* 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 where they could label protein or DNA and then track which entered the *E. coli* cell during infection.
* They grew one batch of T2 phage in the presence of radioactive sulfur, marking the proteins but not DNA.
* They grew another batch in the presence of radioactive phosphorus, marking the DNA but not proteins.
* They allowed each batch to infect separate *E. coli* cultures.
* Shortly after the onset of infection, they 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 lighter free phages and parts of phage in the liquid supernatant.
* They then tested the pellet and supernatant of the separate treatments for the presence of radioactivity.
* Hershey and Chase found that when the bacteria had been infected with T2 phages that contained radio-labeled proteins, most of the radioactivity was in the supernatant, not in the pellet.
* When they examined the bacterial cultures with T2 phage that had radio-labeled 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, which assemble into new viruses.
* The fact that cells double the amount of DNA in a cell prior to mitosis and then distribute the DNA equally to each daughter cell provided some circumstantial evidence that DNA was the genetic material in eukaryotes.
* Similar circumstantial evidence came from the observation that diploid sets of chromosomes have twice as much DNA as the haploid sets in gametes of the same organism.
* By 1947, Erwin Chargaff had developed a series of rules based on a survey of DNA composition in organisms.
* He already knew that DNA was a polymer of nucleotides consisting of a nitrogenous base, deoxyribose, and a phosphate group.
* The bases could be adenine (A), thymine (T), guanine (G), or cytosine (C).
* Chargaff noted that the DNA composition varies from species to species.
* In any one species, the four bases are found in characteristic, but not necessarily equal, ratios.
* He also found a peculiar regularity in the ratios of nucleotide bases which are known as *Chargaff’s rules*.
* The number of adenines was approximately equal to the number of thymines (%T = %A).
* The number of guanines was approximately equal to the number of cytosines (%G = %C).
* Human DNA is 30.9% adenine, 29.4% thymine, 19.9% guanine and 19.8% cytosine.

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

* By the beginnings of the 1950s, the race was on to move from the structure of a single DNA strand to the three-dimensional structure of DNA.
* Among the scientists working on the problem were Linus Pauling, in California, and Maurice Wilkins and Rosalind Franklin, in London.
	+ The phosphate group of one nucleotide is attached to the sugar of the next nucleotide in line.
	+ The result is a “backbone” of alternating phosphates and sugars, from which the bases project.
* Maurice Wilkins and Rosalind Franklin used X-ray crystallography to study the structure of DNA.
* In this technique, X-rays are diffracted as they passed through aligned fibers of purified DNA.
* The diffraction pattern can be used to deduce the three-dimensional shape of molecules.
* James Watson learned from their research that DNA was helical in shape and he deduced the width of the helix and the spacing of bases.
* Watson and his colleague Francis Crick began to work on a model of DNA with two strands, the **double helix**.
* Using molecular models made of wire, they first tried to place the sugar-phosphate chains on the inside.
* However, this did not fit the X-ray measurements and other information on the chemistry of DNA.
* The key breakthrough came when Watson put the sugar-phosphate chain on the outside and the nitrogen bases on the inside of the double helix.
* The sugar-phosphate chains of each strand are like the side ropes of a rope ladder.
* Pairs of nitrogen bases, one from each strand, form rungs.
* The ladder forms a twist 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 would be too wide and a pyrimidine-pyrimidine pairing would be too short.
* Only a pyrimidine-purine pairing would produce the 2-nm diameter indicated by the X-ray data.
* In addition, Watson and Crick determined that chemical side groups of the nitrogen bases would form hydrogen bonds, connecting the two strands.
* Based on details of their structure, adenine would form two hydrogen bonds only with thymine and guanine would form 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.
* However, this does not restrict the sequence of nucleotides along *each* DNA strand.
* The linear sequence of the four bases can be varied in countless ways.
* Each gene has a unique order of nitrogen bases.
* In April 1953, Watson and Crick published a succinct, one-page paper in *Nature* reporting their double helix model of DNA.

# B. 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.

1. 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.
* Essentially, because each strand is complementary to the other, each can form a template when separated.
* The order of bases on one strand can be used to add in complementary bases and therefore duplicate 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 model, semiconservative replication, predicts that when a double helix replicates, each of the daughter molecules will have one old strand and one newly made strand.
* Other competing models, the conservative model and the dispersive model, were also proposed.
* 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.
* In their experiments, they labeled the nucleotides of the old strands with a heavy isotope of nitrogen (15N), while any new nucleotides were indicated by a lighter isotope (14N).
* Replicated strands could be separated by density in a centrifuge.
* Each model— the semi-conservative model, the conservative model, and the dispersive model— made specific predictions on the density of replicated DNA strands.
* The first replication in the 14N medium produced a band of hybrid (15N-14N) DNA, eliminating the conservative model.
* A second replication produced both light and hybrid DNA, eliminating the dispersive model and supporting the semiconservative model.

2. 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 5 million base pairs in its single chromosome and divide to form two identical daughter cells.
* A human cell can copy its 6 billion base pairs and divide into daughter cells in only a few hours.
* This process is remarkably accurate, with only one error per billion nucleotides.
* More than a dozen enzymes and other proteins participate in DNA replication.
* The replication of a DNA molecule begins at special sites, **origins of replication**.
* In bacteria, this is a single 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.
* The replication bubbles elongate as the DNA is replicated and eventually fuse.
* **DNA** **polymerases** catalyze the elongation of new DNA at a replication fork.
* As nucleotides align with complementary bases along the template strand, they are added to the growing end of the new strand by the polymerase.
* The rate of elongation is about 500 nucleotides per second in bacteria and 50 per second in human cells.
* The raw nucleotides are nucleoside triphosphates.
* Each has a nitrogen base, deoxyribose, and a triphosphate tail.
* As each nucleotide is added, the last two phosphate groups are hydrolyzed to form pyrophosphate.
* The exergonic hydrolysis of pyrophosphate to two inorganic phosphate molecules drives the polymerization of the nucleotide to the new strand.
* The strands in the double helix are *antiparallel*.
* The sugar-phosphate backbones run in opposite directions.
* 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 counter to the 3’ -> 5’ direction of the other strand.
* DNA polymerases can only add nucleotides to the free 3’ end of a growing DNA strand.
* A new DNA strand can only elongate in the 5’->3’ direction.
* This creates a problem at the replication fork because one parental strand is oriented 3’->5’ into the fork, while the other antiparallel parental strand is oriented 5’->3’ into the fork.
* At the replication fork, one parental strand (3’-> 5’ into the fork), the **leading strand**, can be used by polymerases as a template for a continuous complementary strand.
* The other parental strand (5’->3’ into the fork), the **lagging strand**, is copied away from the fork in short segments (Okazaki fragments).
* Okazaki fragments, each about 100-200 nucleotides, are joined by **DNA ligase** to form the sugar-phosphate backbone of a single DNA strand.
* DNA polymerases cannot *initiate* synthesis of a polynucleotide because they can only add nucleotides to the end of an existing chain that is base-paired with the template strand.
* To start a new chain requires a **primer**, a short segment of RNA.
* The primer is about 10 nucleotides long in eukaryotes.
* **Primase**, an RNA polymerase, links ribonucleotides that are complementary to the DNA template into the primer.
* RNA polymerases can start an RNA chain from a single template strand.
* After formation of the primer, DNA polymerases can add deoxyribonucleotides to the 3’ end of the ribonucleotide chain.
* Another DNA polymerase later replaces the primer ribonucleotides with deoxyribonucleotides complementary to the template.
* Returning to the original problem at the replication fork, the leading strand requires the formation of only a single primer as the replication fork continues to separate.
* The lagging strand requires formation of a new primer as the replication fork progresses.
* After the primer is formed, DNA polymerase can add new nucleotides away from the fork until it runs into the previous Okazaki fragment.
* The primers are converted to DNA before DNA ligase joins the fragments together.
* In addition to primase, DNA polymerases, and DNA ligases, several other proteins have prominent roles in DNA synthesis.
* A **helicase** untwists and separates the template DNA strands at the replication fork.
* **Single-strand binding proteins** keep the unpaired template strands apart during replication.
* 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**.**
* 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 that may be anchored to the nuclear matrix.
* The DNA polymerase molecules “reel in” the parental DNA and “extrude” newly made daughter DNA molecules.

3. 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 10,000 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 per billion nucleotides.
* DNA molecules are constantly subject to potentially harmful chemical and physical agents.
* Reactive chemicals, radioactive emissions, X-rays, and ultraviolet light can change nucleotides in ways that can affect encoded genetic information.
* DNA bases often undergo spontaneous chemical changes under normal cellular conditions.
* Mismatched nucleotides that are missed by DNA polymerase or mutations that occur after DNA synthesis is completed can often be repaired.
* Each cell continually monitors and repairs its genetic material, with over 130 repair enzymes identified in humans.
* In **mismatch repair**, special enzymes fix incorrectly paired nucleotides.
* A hereditary defect in one of these enzymes is associated with a form of colon cancer.
* In **nucleotide excision repair**, a **nuclease** cuts out a segment of a damaged strand.
* The gap is filled in by DNA polymerase and ligase.
* The importance of the proper functioning of repair enzymes is clear from the inherited disorder xeroderma pigmentosum.
* These individuals are hypersensitive to sunlight.
* In particular, ultraviolet light can produce thymine dimers between adjacent thymine nucleotides.
* This buckles the DNA double helix and interferes with DNA replication.
* In individuals with this disorder, mutations in their skin cells are left uncorrected and cause skin cancer.

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

* Limitations in the DNA polymerase create problems for the linear DNA of eukaryotic chromosomes.
* The usual replication machinery provides no way to complete the 5’ ends of daughter DNA strands.
* Repeated rounds of replication produce shorter and shorter DNA molecules.
* The ends of eukaryotic chromosomal DNA molecules, the **telomeres**, have special nucleotide sequences.
* In human telomeres, this sequence is typically TTAGGG, repeated between 100 and 1,000 times.
* Telomeres protect genes from being eroded through multiple rounds of DNA replication.
* Eukaryotic cells have evolved a mechanism to restore shortened telomeres.
* **Telomerase** uses a short molecule of RNA as a template to extend the 3’ end of the telomere.
* There is now room for primase and DNA polymerase to extend the 5’ end.
* It does not repair the 3’-end “overhang,”but it does lengthen the telomere.
* Telomerase is *not* present in most cells of multicellular organisms.
* Therefore, the DNA of dividing somatic cells and cultured cells does tend to become shorter.
* Thus, telomere length may be a limiting factor in the life span of certain tissues and of the organism.
* Telomerase is present in germ-line cells, ensuring that zygotes have long telomeres.
* Active telomerase is also found in cancerous somatic cells.
* This overcomes the progressive shortening that would eventually lead to self-destruction of the cancer.