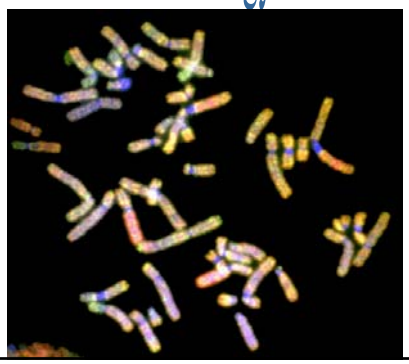
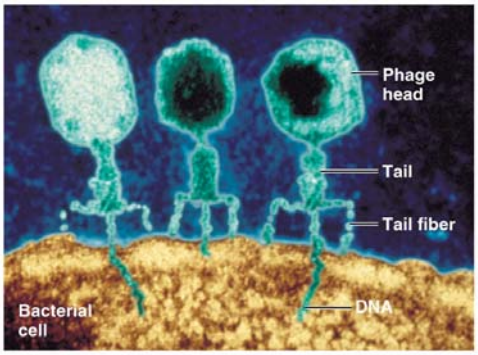
The Molecular Biology of the Gene



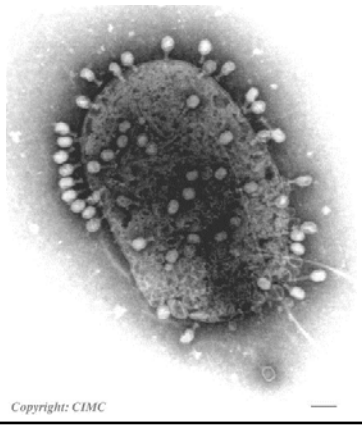
Introduction

- The chromosome theory of inheritance set the stage for the development of a molecular understanding of the gene.
- Many of the basics of molecular biology began to be understood by studying viruses that infect bacteria.
- Viruses are not considered alive because they lack cellular structure and metabolism.
- Viruses are composed of a protein coat and internal DNA (or RNA), and they depend on the metabolism of their host to make more viral particles.

- All living things are infected by viruses. Bacterial viruses are known as **bacteriophages**.
- Experimental systems using phages were a logical choice for early experiments on the molecular biology of the gene. Phages are simple, with simple genes infecting relatively simple and easily manipulated bacteria.

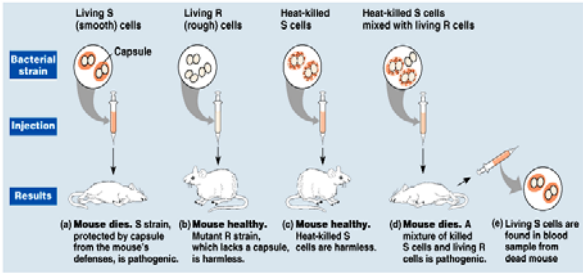


(a) T2 and related phages use their tail pieces to attach to the host cell and inject their genetic material (TEM).



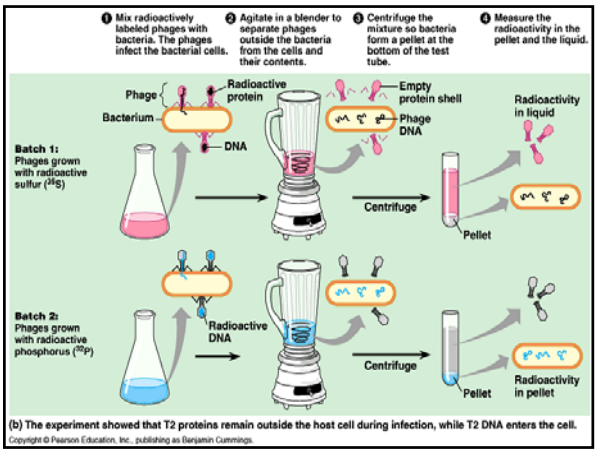
Experiments showed that DNA is the genetic material

- In 1928, **Griffith** showed that some substance conveyed traits from heat-killed bacteria to living bacteria without the trait.



•Evidence gathered in the 1930s and 1940s showed it was DNA rather than protein that was the genetic material.

•In 1952, **Hershey & Chase**, using T2 phage, showed the radioactive isotope of sulfur (found only in proteins) was not transferred into new viral particles, whereas the radioactive isotope of phosphorous (only found in DNA) was.



DNA and RNA are polymers of nucleotides

•Each monomer nucleotide is composed of a **5-caron sugar**, a **phosphate group** (acidic), and a **nitrogenous base**.

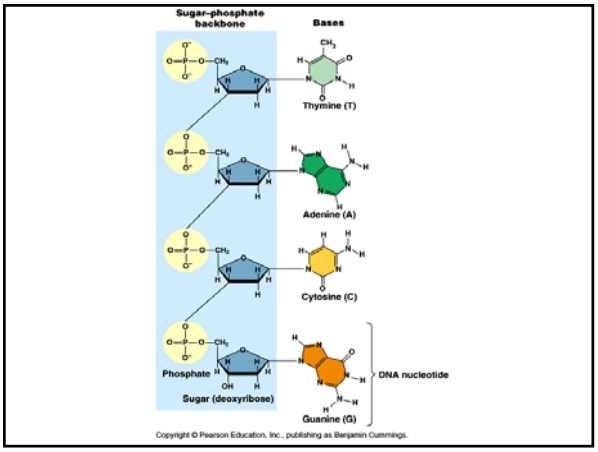
•Ribose rather than deoxyribose sugar is found in RNA.

•The four nitrogenous bases of DNA:

Pyrimidines: Thymine (**T**) and Cytosine (**C**)

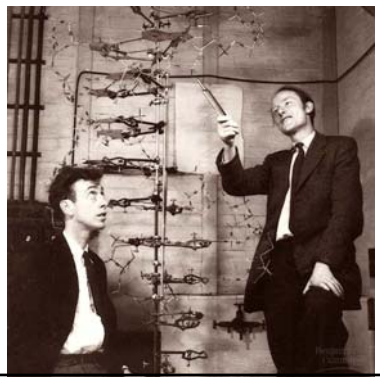
Purines: Adenine (**A**) and Guanine (**G**)

•**Uracil** is used instead of thymine in RNA.

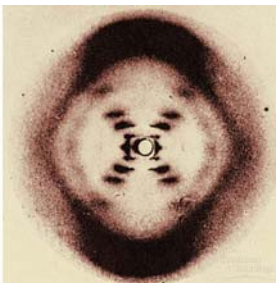


DNA is a double-stranded helix

•Many scientists contributed to the **Watson-Crick model**.



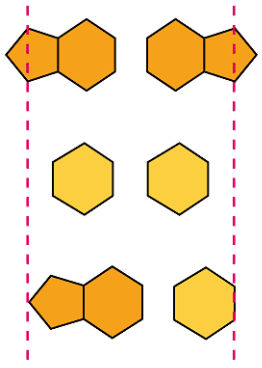
•**Wilkins and Franklin's** X-ray crysatllographs.

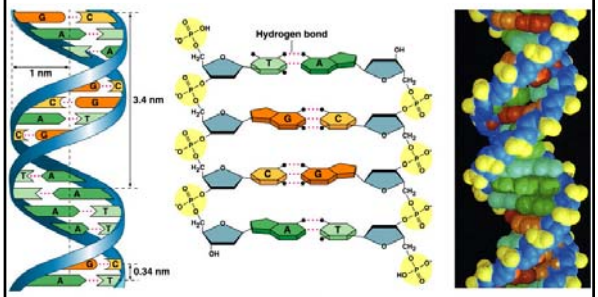


•**Chargaff's** chemical analysis showing the amounts of **A** and **T**, and **G** and **C**, were always equal

•The model that fit all the observations was a **double helix** (a twisted rope ladder) with sugar backbones on the outside and hydrogen-bounded nitrogenous bases on the inside.

•**A** always bonds with **T**, and **G** always bonds with **C**, but there are no restrictions on linear sequence.





•Two strands of the double helix run in opposite directions (**antiparallel**).

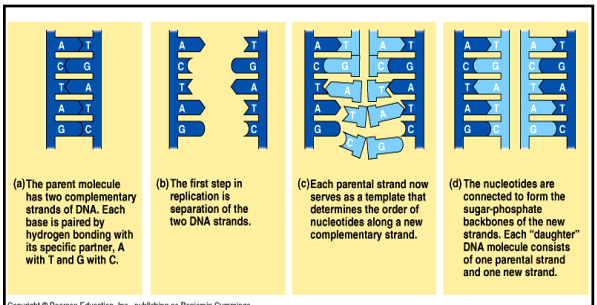
•The Watson-Crick model was proposed in a short paper in 1953 and almost immediately led to proposed mechanisms about DNA function.

DNA replication depends on specific base pairing

•The nature of the reproductive process, and of the cell cycle involved in it, requires that complete and faithful copies of DNA be reproduced (**replicated**).

•Watson & Crick stated that their model suggested a copying mechanism.

•The mechanism proposed and confirmed by the end of the 1950s involved each half of the double helix functioning as a template upon which a new, missing half is built.



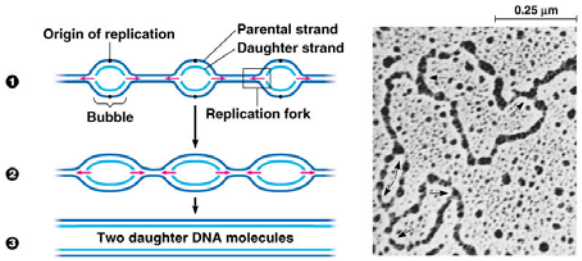
NOTE: Each new double helix consists of one old strand and one new strand; thus the mechanism of replication is ***semiconservative***.

•The actual mechanism involves a complex arrangement of molecular players, the help of enzymes, particularly **DNA polymerases**, and some geometric contortions including untwisting of the parent helix and retwisting the daughter helices.

•Despite its speed (50-500 pairs per second), replication is very accurate, with approximately one mistaken nucleotide pair in a billion.

DNA replication: A closer look

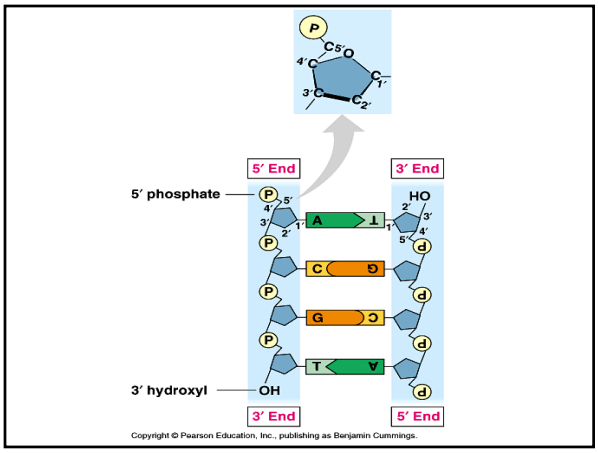
•Replication occurs simultaneously at many sites (**replication bubbles**) on a double helix. This allows DNA replication to occur in a shorter period of time than replication at a single origin would allow.



(a) In eukaryotes, DNA replication begins at many sites along the giant DNA molecule of each chromosome.

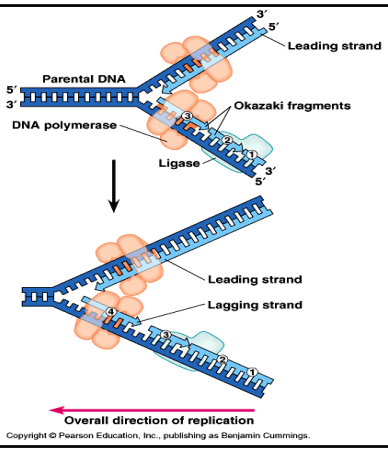
(b) In this micrograph, three replication bubbles are visible along the DNA of cultured Chinese hamster cells. The arrows indicate the direction of DNA replication at the two ends of each bubble (TEM).

•DNA polymerases can only attach nucleotides to the 3' end of a growing daughter strand. Thus replication always proceeds in the 5' to 3' direction.



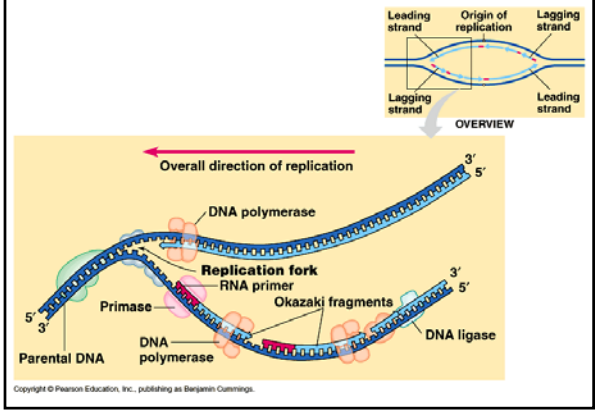
•Within the replication bubbles, one daughter strand is synthesized continuously while the other daughter strand must be synthesized in short pieces which are joined together by **DNA ligase**.

•These short pieces of DNA are called ***Okazaki fragments***.



•DNA polymerases proofread the new daughter strands

•This replication process assures that daughter cells will carry the same genetic information as each other and the parental cells.



The Flow of Genetic Information from DNA to RNA to Protein

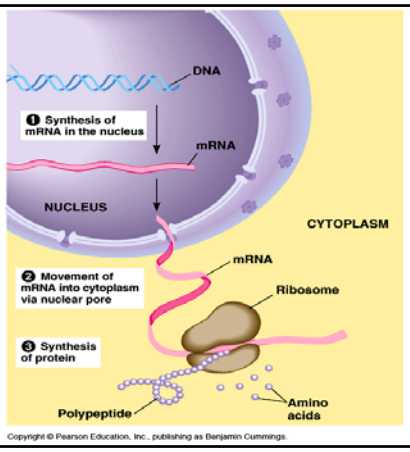
The DNA genotype is expressed as proteins, which provides the molecular basis for phenotypic traits

•The molecular basis of phenotypic traits are the proteins an organism can make.

•The **one-gene one enzyme hypothesis** was formulated in the 1940s by Beadle & Tatum, who were studying nutritional mutants of the mold *Neurospora*. They found that genetic mutants lacked single enzymes needed to complete metabolic pathways.

•Now stated as **one-gene one polypeptide**.

•The flow is known to occur in two stages: **transcription** of the genetic code in the nucleus to a messenger RNA (mRNA) molecule, and **translation** of the mRNA message in the cytoplasm.

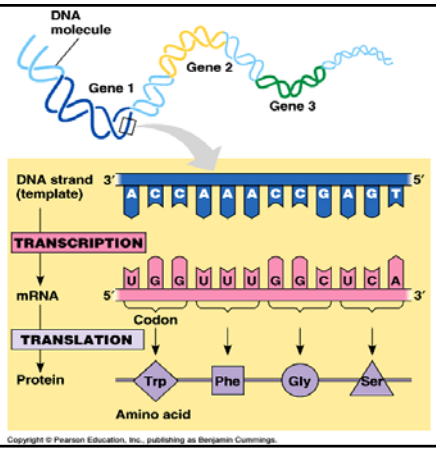


Genetic information is written as codons and translated into amino acids sequences

•The nucleotide monomers represent letters in an alphabet that can form words in a language. Each word codes for one amino acid in a polypeptide.

•There are four letters (A, T, G, and C) and 20 amino acids. One-letter words would create four distinct words. Two-letter words would create a vocabulary of 16 words (4 x 4). Three-letter words would create a vocabulary of 64 words (4 x 4 x 4).

•**Triplets of bases** are the <u>smallest words</u> of uniform length that can specify all the amino acids. These triplets are known as **codons**.

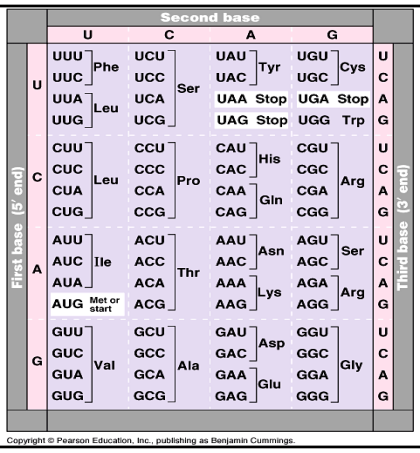


The Genetic Code is the Rosetta Stone of life

•The first codon was deciphered by Nirenberg in 1961.

•Nirenberg added polyuracil (an artificially made RNA polynucleotide) to a mixture containing ribosomes and other cell fractions required for translation. The polypeptide polyphenylalanine was produced indicating UUU was the codon for phenylalanine.

•The code was completely known by the end of the 1960s. It shows redundancy but no ambiguity.



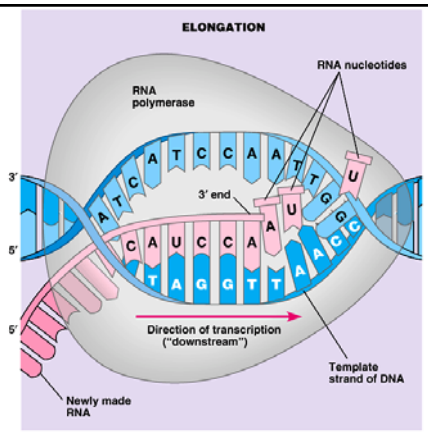
•The code is virtually the same for all organisms. Thus bacterial cells can translate the genetic messages of human cells, and vice versa.

•This gives evidence of the relatedness of all life and suggests that the genetic code was established early in the history of life.

Transcription produces genetic messages in the form of RNA

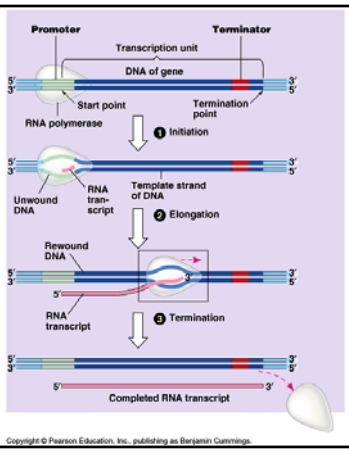
•In **transcription** one strand of DNA serves as a template for the new RNA strand.

•**RNA polymerases** construct the RNA strand.



•Transcription is initiated from one strand of DNA as indicated by a **promoter region** (the site at which RNA polymerase attaches), the DNA unwinds, and RNA polymerization and elongation occur.

•Finally, the mRNA sequence is terminated when the process reaches a special **terminator region** of the DNA.



NOTE: Transcription means copying a message to a new medium

•Two other types of RNA (rRNA and tRNA) play a role in translation and are transcribed by this process.

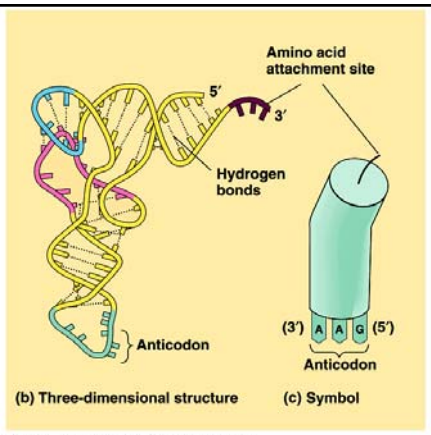
Genetic messages are translated in the cytoplasm

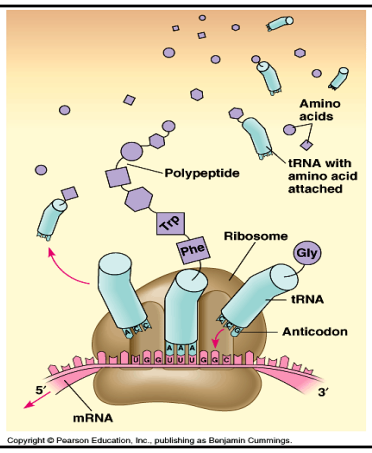
- RNA that encodes a messenger sequence is called messenger RNA (mRNA)
- In prokaryotes, transcription and translation both occur in the cytoplasm.
- In eukaryotes, a completed mRNA molecule leaves the nucleus and the message is translated in the cytoplasm.
- The players in the translation process include ribosomes, tRNA, enzymes and protein factors, and sources of cellular energy.

NOTE: Translation means rewording a message into a new language. This new language in this case is the linear sequence of amino acids in polypeptides.

Transfer of RNA molecules serve as interpreters during translation

- Amino acids that are to be joined in correct sequences cannot recognize the codons on the mRNA.
- Transfer RNA (tRNA) molecules, one or more for each type of amino acid to the right codon.
- Each tRNA contains a region (the **anticodon**) that recognizes and binds to the correct codon on the mRNA.





Ribosomes build polypeptides

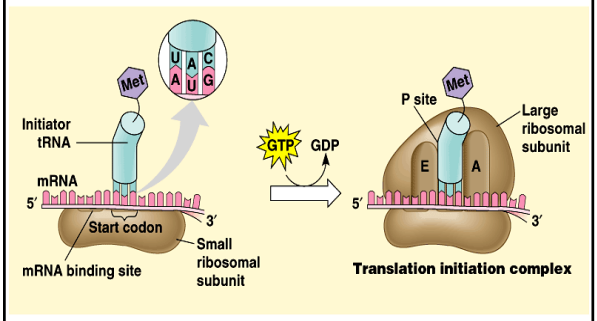
- Ribosomes are composed of ribosomal RNA (rRNA) and protein, arranged in two subunits.
- The shape of ribosomes provides a platform on which protein synthesis can take place. There are locations for the mRNA, and two tRNA-amino acid complex binding sites (an **A site** and a **P site**).

An initiation codon marks the start of an mRNA message

•Translation can be divided into the same three phases as transcription: initiation, elongation, and termination.

•An mRNA molecule is longer than the genetic message it contains. It contains a starting nucleotide sequence that helps the **initiation phase** and an ending sequence that helps the **termination phase**.

•During initiation, the initial sequence helps bind the mRNA to the ribosomal subunit, a specific start codon binds with an initiator tRNA anticodon carrying the amino acid methionine, and the large ribosome binds to the small subunit as the initiator tRNA fits into the P site on the large subunit.



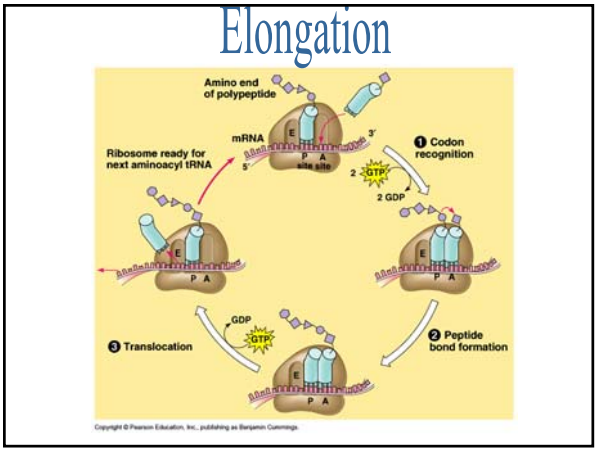
Elongation adds amino acids to the polypeptide until a stop codon terminates translation

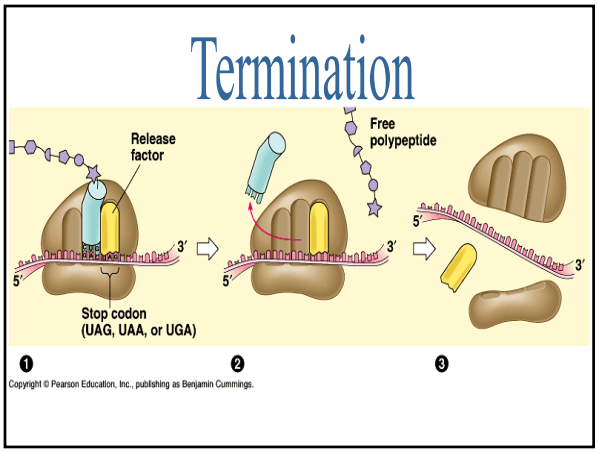
•Elongation involves three steps

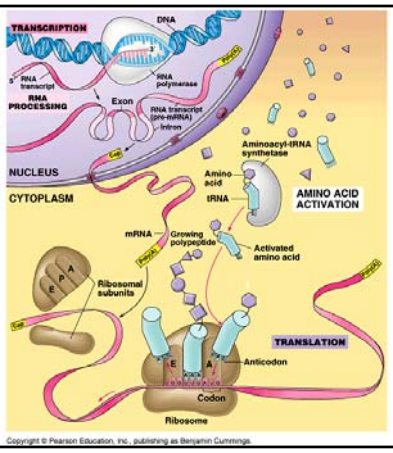
(a) codon recognition
(b) peptide bond formation
(c) translocation

•The formation of the polypeptide bond is catalyzed by an enzyme within the ribosome structure.

•Elongation continues until a special stop codon (UAA, UAG, or UGA) causes termination of the process. The finished polypeptide is freed, and the ribosome splits into its two subunits.







Mutations can change the meaning of genes

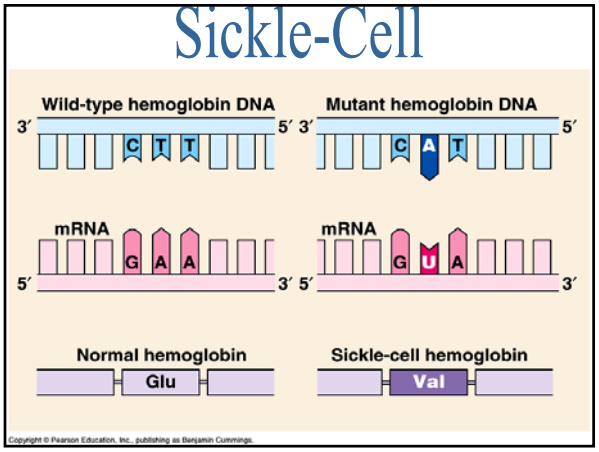
•Many differences in inherited traits in humans have been traced to their molecular causes.

•A change in the nucleotide sequence of DNA is known as a **mutation**.

•Certain substitutions of one nucleotide base for another will lead to mutations, resulting in the replacement of one amino acid for another in a polypeptide sequence.

•Base substitutions usually cause a gene to produce an abnormal product, or they result in no change if the new codon still codes for the same amino acid.

•A base substitution is known to account for the type of hemoglobin produced by the sickle-cell allele.



•Rarely, base substitutions lead to improved or changed genes that may enhance the success of the individual in which they occur.

•The addition or subtraction of nucleotides may result in a shift of the three-base reading frame; all codons past the affected one are likely to code for different amino acids. The differences that result will almost always result in a nonfunctional polypeptide.