

Chapter 12: From DNA to Protein: Genotype to Phenotype

I One Gene, One Polypeptide

- A gene is a DNA sequence.
- However, not all DNA sequences are genes.
- To be a gene the DNA sequence must be able to be transcribed.
- In the 1940's, Beadle and Tatum found that a phenotypic mutant had an altered enzyme.
 - They used *Neurospora crassa*.
 - This is an organism with a haploid vegetative life-cycle.
 - This makes recessive mutations easy to detect.
 - Neurospora were grown on minimal medium consisting of just sucrose, minerals and a few vitamins.
 - Wild-type neurospora were treated with a mutagen, which causes changes in the DNA.
 - After treatment, they were grown in a complete medium.
 - When testing some of the treated strains, some were found that could no longer grow on minimal medium, but instead needed certain supplements.
 - These nutrient requiring *auxotrophs* were assumed to have mutated.
 - For each auxotrophic strain, Beadle and Tatum were able to find a single compound that could support its growth.
 - One group of mutants needed arginine to grow.
 - Mapping studies established that some these *arg* mutations are at different loci, and therefore are in different genes.
 - Beadle and Tatum demonstrated that these different mutants had defective genes for the same biochemical pathway, the pathway leading to arginine synthesis.
 - See Figure 12.1.
 - If the gene defect affected earlier steps in the pathway, several different substances could substitute for arginine.
 - If the defect was for the step just before arginine synthesis, only arginine could substitute.
 - Beadle and Tatum postulated the *one gene, one enzyme* definition of a gene.
 - Later it was learned that some enzymes are composed of different subunits coded for by separate genes. The one gene, one enzyme hypothesis was changed to the *one-gene, one-polypeptide* definition.
 - Even this hypothesis requires modification because some genes code for RNA molecules that are never translated into polypeptides.

II DNA, RNA and the Flow of Information

- Two steps are used to express a gene:
 - Transcription makes a single-stranded RNA copy of a segment of the DNA.
 - Translation uses information encoded in a portion of the RNA to make a polypeptide.
- In eukaryotes, these two steps are physically separated.

A. RNA differs from DNA

- RNA is single-stranded because only one strand of the DNA is used as a template.
- The sugar in RNA is ribose, not deoxyribose.
- Wherever thymine is found in DNA, RNA has uracil.

B. Information flows in one direction when genes are expressed.

- The old central dogma stated that DNA codes for RNA, and RNA codes for protein.
- Once information passes into protein, it cannot get out again.
- Messenger RNA or mRNA moves from the nucleus of eukaryotic cells into the cytoplasm.
- Transfer RNA or tRNA is the link between the code of the mRNA and the amino acids of the polypeptide.
- The tRNA molecules are charged with the correct amino acid.
- *See Figure 12.3.*

C. RNA viruses modify the central dogma

- RNA viruses are viruses that use RNA as their information molecule during transmission.
- Examples are the influenza virus and poliovirus.
- HIV and certain tumor viruses have RNA as their infectious information molecule but convert it to DNA inside the host cell.
- *See Figure 12.2.*

III Transcription: DNA – Directed RNA Synthesis

- RNA polymerase is the enzyme that used DNA as a template to make RNA.
- Just one of the strands of a gene's DNA is used to make the RNA.
- This strand is called the template strand. The other strand is called the coding strand, and has a sequence that directly corresponds to the mRNA.
- Each chromosome, which is a continuous, single double-helix of DNA, has many regions that are read by RNA polymerase.
- One gene might use one of the strands of the double-helix on the template for RNA, and another on the same chromosome might use the other strand as the template.

A. Initiation of transcription requires a promoter and an RNA polymerase

- Transcription of a gene begins at a promoter, which is a certain sequence of nucleotides.
- The RNA polymerase binds to the promoter region, when conditions allow.
- The promoter sequence directs the RNA polymerase as to which of the double strands is the template and what direction the RNA polymerase should move.
- RNA is synthesized in the 5' to 3' direction moving along the template DNA in the 3' to 5' direction.
- *See Figure 12.4.*
- Not all promoters are identical. Some bind RNA polymerase more effectively; this causes them to be transcribed more frequently, when other conditions allow.

- Prokaryotes have one type of RNA polymerase that transcribes mRNA, tRNA and rRNA.
- Eukaryotes have three different RNA polymerases: RNA polymerase I, II and III.
- RNA polymerase II makes all *mRNA* in eukaryotes.
- In eukaryotic cells, other proteins must bind to the DNA around the promoter to prepare a “docking site”.

B. RNA polymerase elongates the transcripts

- After binding, RNA polymerase unwinds about 20 base pairs at a time and reads the template in the 3' to 5' direction.
- The RNA transcript is antiparallel to the DNA template strand.
- Energy for synthesis comes from the phosphodiester linkages in the *ribonucleoside triphosphates*, the monomeric building blocks of RNA.
- One of the three phosphates remains in the RNA polymer.
- Transcription errors are high relative to DNA polymerase, a mistake occurs for every 10^4 to 10^5 bases incorporated.

C. Transcription terminates at particular base sequences

- Particular base sequences in the DNA specify termination.
- Mechanisms for termination vary.
 - One is when the part of the RNA molecule already synthesized forms a certain secondary structure and simply falls away.
 - Another involves special proteins that recognize a site on newly synthesized RNA molecules.

IV The Genetic Code

- DNA codes for RNA.
- RNA is read in three-base contiguous segments.
- Each three-base segment is called a codon.
- The number of different codons possible is 64, because any four bases can form a three-base codon.
- ($4^3 = 64$)
- All codons determine just 20 amino acids, and the start and stop signals found in all *mRNA* molecules.
- Three of the possible codons are stop codons (UAA, UAG and UGA).
- Stop codons direct the ribosomes to stop reading the mRNA; i.e., they end translation.

A. The genetic code is redundant but not ambiguous

- After subtracting start and stop codons, the remaining 60 codons code for 19 different amino acids.
- Many amino acids have more than one codon. This is called redundancy.
- The code is *not* ambiguous. Each codon is definitely assigned a certain amino acid.
- The tRNA molecules which have the correct amino acids attached determine the assignment.
- The genetic code is nearly global, applying to all species on our planet.

- Exceptions are within mitochondria and chloroplasts.
- *See Figure 12.5 for the genetic code.*

B. Biologists broke the genetic code by using artificial messengers

- Decoding breakthroughs started in 1961.
- Nirenberg prepared an artificial mRNA in which all bases were uracil.
- Incubated with required additional components, poly-phenylalanine was synthesized.
- Other codons were deciphered from this point.
- An additional technique finished the deciphering.
 - Simple synthetic mRNA's, three nucleotides long, could bind to ribosomes.
 - This complex then caused the tRNA-amino acid to bind.
 - Using this and radioactive labeling, the code was fully deciphered.
 - *See Figure 12.6.*

V Preparation for Translation: Linking RNA's, Amino Acids, and Ribosomes

- Translation occurs at ribosomes.
- In prokaryotes, ribosomes bind to mRNA as it is being synthesized.
- In eukaryotes, mRNA is made in the nucleus and translation occurs in the cytoplasm.
 - The tRNA's must read mRNA correctly.
 - The RNA's must carry the correct amino acids.

A. Transfer RNA's carry specific amino acids and bind to specific codons

- A tRNA molecule has 75 to 80 nucleotides.
- It has a three-dimensional shape maintained by complementary base pairing and hydrogen bonding.
- *See Figure 12.7.*
- At the 3' end of every tRNA molecule is a site to which its specific amino acid binds, covalently.
- Midpoint in the sequence is three bases called the anticodon.
- The anticodon is the contact point between the tRNA and the mRNA.
- The anticodon is complementary (and antiparallel) to the mRNA codon.
- The codon and anticodon complementary base pair.
- There are fewer anticodon codes than mRNA codons.
- This is possible because some codon-anticodon interactions tolerate a mismatch at the 3' third base of the mRNA. (*See Figure 12.5*)
- This is called wobble.
- The three-dimensional shape of the tRNA's allows them to combine with the binding sites of the ribosome.

B. Activating enzymes link the right tRNA's and amino acids

- The correct amino acids are attached to correct tRNA's by a family of activating enzymes called aminoacyl-tRNA synthetases.
- Each activating enzyme is specific for one amino acid and its tRNA.
- The enzyme has a three-part active site that binds a specific amino, ATP and a specific tRNA.

- The reactions have two-steps:
 - $\text{Enzyme} + \text{ATP} + \text{AA} \rightarrow \text{enzyme} - \text{AMP} - \text{AA} + \text{PPi}$.
 - $\text{Enzyme} - \text{AMP-AA} + \text{tRNA} \rightarrow \text{enzyme} + \text{AMP} + \text{tRNA} - \text{AA}$.
 - *See Figure 12.8*

C. The ribosome is the staging area for translation

- Each ribosome has two subunits, a larger and smaller one.
- The large one in eukaryotes has 3 different associated rRNA molecules and 45 different proteins.
- The smaller subunit has one rRNA and 33 different protein molecules.
- When not translating the two subunits are separate.
- Ribosomes of prokaryotes are somewhat smaller. Their ribosomal proteins and rRNA's are different.
- The different proteins and rRNA are held together by ionic and hydrogen bonds, and hydrophobic forces, not covalent bonds.
- The structure can self-assemble if disassembled in soap.
- Ribosomes are non-specific and combine with any mRNA and all tRNA's.
- The large subunit has four sites where tRNA molecules bind.
 - *See Figure 12.9.*
 - The T site is where the tRNA first lands. It is brought to the site by the T or transfer, factor.
 - The A site is where the tRNA anticodon binds to the mRNA codon.
 - The P site is where the tRNA adds its amino acid to the growing peptide chain.
 - The E (exit) site is where the uncharged (no amino acid attached) tRNA goes before leaving the ribosome.

VI Translation: RNA-Directed Polypeptide Synthesis

RNA-directed assembly of a protein:

A. Translation begins with an initiation complex

- An initiation complex forms, which includes the first tRNA and its amino acid, a small subunit of the ribosome and an mRNA molecule.
- This complex is bound to a region upstream (toward the 5' end) of where reading of the mRNA begins.
- Methionine is the first amino acid in all proteins. (However, some proteins are trimmed after synthesis and the methionine is thereby removed.)
- The large subunit then associates.
- The process is directed by initiation factors, which use GTP as an energy source.
- *See Figure 12.10.*

B. The polypeptide elongates from the N terminus

- Ribosomes move in the 5' to 3' direction on the mRNA. They synthesize the peptide in the N-term to C-term direction.
 - The large subunit catalyzes two reactions:
 - Breakage of the bond between the tRNA in the P site and its amino acid (on the polypeptide).

- Peptide bond formation between this (tRNA attached) amino acid and the tRNA in the A site.
- This is called peptidyl transferase activity.
- One of the rRNA's in the large subunits appears to participate in the catalysis of this reaction.
- *See Figure 12.11.*

C. Elongation continues and the peptide chain grows

- The A site tRNA then moves to the P site.
- The next charged tRNA enters the A site.
- The peptide chain is then transferred to the A site.
- These steps are assisted by elongation factors.

D. A release factor terminates translation

- When a stop codon-UAA, UAG or UGA enters the A site, a release factor and a water molecule enters the A site, instead of an amino acid.
- *See Figure 12.12.*

VI Regulation of Translation

A. Some antibiotics work by inhibiting translation

- Antibiotics are defense molecules.
- They are produced by some fungi and bacteria.
- They have been used to combat human bacterial infectious disease.
- Antibiotics destroy microbial invaders, but do not harm the human host.
- Some antibiotics work by inhibiting protein synthesis.
- Because of differences in eukaryotic ribosomes, the human ribosomes are unaffected.
- *See Table 12.2.*

B. Polysome formation increases the rate of protein synthesis

- Polysomes are mRNA molecules with more than one ribosome attached.
- These make protein more rapidly.
- *See Figure 12.13.*

VII Posttranslational Events

- Some proteins are modified after synthesis.
- New chemical groups might be added, folding might be assisted by other proteins, or proteins might get trimmed.

A. Chemical signals in proteins direct them to their cellular destinations

- *See Figure 12.14.*
- As the polypeptide chain forms, it spontaneously folds.
- The amino acid sequence also contains an “address label” indicating where in the cell the polypeptide belongs.
- All protein synthesis begins on free ribosomes in the cytoplasm.
- In eukaryotes, as the peptide chain is made, information on the nascent portion gives one of two sets of instructions:
 - Finish transcription and be released to the cytoplasm.

- Stall translation, go to the endoplasmic reticulum, and finish synthesis at the ER surface.
- Those destined to finish synthesis may contain information in their amino acid sequence that specifies where they belong.
 - Some belong in the nucleus.
 - Some belong in mitochondria.
 - Some belong in peroxisomes.
- Some of the transport to destination requires chaperonin proteins and docking proteins at the membrane that the protein must cross.
- Those destined for the endoplasmic reticulum generate an approximately 25 amino acid hydrophobic leader sequence that signals to a signal recognition particle, which is composed of protein and RNA.
- *See Figure 12.15.*
 - The association of the signal to the signal receptor particle stalls any additional translation.
 - This stall continues until the ribosome attaches to a specific receptor protein on the surface of the ER.
 - Translation continues with the protein moving through a pore in the mitochondrial membrane.
 - Some proteins have signals that direct the embedding of the protein into ER membrane.
 - This is when membrane proteins of the ER, Golgi, lysosomes and plasma membrane get positioned.
- Other signals direct the protein to the golgi, lysosomes, or to the outside the cell.

B. Many proteins are modified after translation

- *See Figure 12.16.*
- Proteolysis is the cleavage of the protein to make a shortened finished protein.
 - Insulin is an example of a protein that gets trimmed.
 - The signal to go to the ER is often cleaved after the protein gets there.
 - HIV needs a protease to cleave a protein. One treatment for HIV inhibits this enzyme.
- Glycosylation involves the addition of sugars to the protein.
 - Signals in the amino acid sequence of the protein direct the addition of the sugars in the ER.
 - Additional modifications occur in the golgi.
- Phosphorylation is the addition of phosphate groups to certain amino acids of certain proteins. These additions are often temporary and often affect the activity of the protein.

VIII Mutations: Heritable Changes in Genes

- Mutations are heritable changes in the DNA.
- Heritable means the change is passed to daughter cells.
 - Somatic mutations are passed on during mitosis, but the affected cells never become gametes.
 - Germ-line mutations are mutations that occur in cells that might give rise to gametes.

- Some mutations cause visible phenotypic change. Others cause metabolic changes that might not yet be detectable.
- Some mutations exert their effect under certain, referred to as, restrictive conditions.
 - These are called conditional mutants.
 - They are unaffected at the permissive conditions, but express the mutant phenotype at the restrictive condition.
 - Temperature-sensitive mutants are an example.
- All mutations are of two types:
 - Point mutations are mutations of single genes.
 - Chromosomal mutations are changes in the arrangements of chromosomal segments.

A. Point mutations are changes in single bases

- Point mutations result from the addition or subtraction of a nucleotide base or the substitution of one base for another.
- Point mutations can occur as a result of mistakes during DNA replication, or by environmental mutagens, such as chemicals and radiation.
- Because of redundancy in the genetic code, some point mutations result in no change in the amino acids in the protein.
- These are called silent mutations.
- Some mutations cause an amino acid substitution.
 - These are called missense mutations.
 - An example in humans is sickle-cell disease, a defect in the β -globin subunits of hemoglobin.
 - The red blood cells collapse when oxygen levels are low.
- Missense mutations might reduce the functioning of a protein or disable it completely.
- Nonsense mutations are base substitutions that cause a change from a codon that instructs the incorporation of an amino acid to a codon that terminates translation.
- A frame-shift mutation is when a single base is inserted or deleted in a gene.
- This causes the most disruption when the event occurs at or near the beginning of the template.
- This type of mutation shifts the code, changing many of the codons to different codons.

B. Chromosomal mutations are extensive changes in the genetic material

- *See Figure 12.18.*
- DNA molecules can break and reform.
- This can cause four different types of mutations: deletions, duplications, inversions and translocations.
- Deletions are a loss of a chromosomal segment.
- Duplications are a repeat of a segment.
- Breaking and rejoining can lead to inversions.
 - Inversions are when segments get reattached in the opposite orientation.
 - When homozygous for inversion, no problem is usually encountered.
 - When heterozygous for an inversion, crossovers can cause duplications, deletions and other problems.

- Translocations are when a portion of one chromosome attaches to another.
 - Translocations can be reciprocal (*see Figure 12.18d*) or nonreciprocal.
 - Translocations can make synapses in meiosis difficult and can lead to aneuploidy.
 - Sometimes two acrocentric chromosomes join to create a single metacentric.
 - This changes the number of chromosomes.
 - Translocation homozygotes become genetically isolated. This is important to mammalian speciation.

C. Mutations can be spontaneous or induced

- *See Figure 12.19.*
- Spontaneous mutations:
 - Nucleotides occasionally change their structure (called a tautomeric shift).
 - When a base temporarily changes to its unusual tautomer at the same time as when replication is occurring, it pairs with the alternate purine, if it is a purine or the alternate pyrimidine if it is a pyrimidine.
 - This causes DNA polymerase to make errors in replication.
 - Meiosis is imperfect. Nondisjunction can occur. Random chromosome breaks rejoin incorrectly.
- Induced mutations:
 - These are caused by some outside agent.
 - Some chemicals alter covalent bonds in nucleotides.
 - Nitrous acid deaminates cytosine, converting it to uracil.
 - DNA polymerase mistakes uracil for thymine and puts an A in during replication instead of the G that would have been incorporated otherwise.
 - Benzo a-pyrene, a product of incomplete combustion, which is found in all smoke, adds a large chemical group to guanine, making it unavailable for base pairing.
 - Any base might be inserted to fill the gap.
 - Radiation damages DNA.
 - Ionizing radiation (x-rays and gamma rays) produces highly reactive compounds and atoms called free radicals.
 - These can alter bases or break the sugar-phosphate backbone, causing chromosomal abnormalities.
 - Ultraviolet radiation is absorbed by pyrimidines in the DNA, and when two thymines or two cytosines are next to each other on the same strand of a double stranded DNA molecule, a covalent bond can form.
 - Their inter-strand covalent bonds make the DNA un-replicable.
 - The long term benefit of mutations is that they account for all the differences between and within organisms, excluding the effect of different environments.

D. Mutations are the raw material of evolution

- Mutations are rare events.

- Frequency of mutations is much lower than one mutation per 10^4 genes per DNA duplication. Sometimes they are as rare as one per 10^9 genes per duplication.
- Different organisms vary in mutation frequency.
- Mutations can be detrimental, neutral or occasionally beneficial.
- Humans have 1000X the DNA of a prokaryote.
- This is at least partially due to duplication of DNA sequences and then divergence of the sequences over time.
- Random accumulation of mutations in the extra copies of genes can lead to the production of new useful proteins.