From Gene to Protein
by Dr. Ty C.M. Hoffman

Slide 1
The Beadle and Tatum experiment involved causing mutations in bread mold. By observing that various mutants were unable to grow on minimal medium but were able to grow if the medium was supplemented with the correct substance, Beadle and Tatum formulated the "one gene - one enzyme" hypothesis to explain the function of genes. We now know that protein coding genes serve as the instructions for how to produce all of the cell's proteins, not just the enzymes.

Slide 2
Gene expression is the process in which the code stored in a gene (DNA nucleotides) is used to indirectly assemble a polypeptide. Gene expression occurs in two major steps:

- In transcription, the sequence of nucleotides (known as a gene) within a DNA molecule is read. A strand of RNA nucleotides is assembled wherein each RNA nucleotide is complementary to the corresponding DNA nucleotide in the gene. Thus, C in DNA specifies G in RNA; G specifies C; T specifies A; and A specifies U (because RNA uses uracil instead of thymine). The resulting strand is messenger RNA (mRNA), and it is called the mRNA transcript.
- In translation, the mRNA transcript that was produced during transcription is used as a set of instructions for how to assemble a polypeptide. The transcript joins with a ribosome, which reads the transcript, and the polypeptide is assembled one amino acid at a time.

A prokaryotic cell lacks membrane-bounded organelles, so it has just one compartment. In these cells, both transcription and translation occur in the cytoplasm. In eukaryotic cells, transcription occurs in the nucleoplasm (where the DNA is found), and translation occurs in the cytoplasm (where ribosomes are found). In addition, the mRNA produced in eukaryotic cells must be modified before it can function for translation.

Slide 3
Transcription features a one-to-one correspondence between DNA nucleotides in the gene and RNA nucleotides in the RNA transcript. However, the four nucleotide types used in RNA are insufficient to individually specify all twenty biological amino acids during translation. Instead, a sequence of three mRNA nucleotides (collectively called a codon) specifies a single amino acid. Since there are three nucleotide positions within a codon, and each position can be occupied by any of four types of nucleotide, there are sixty-four possible codons. This is more than enough to specify the twenty biological amino acids.

Slide 4
The genetic code is universal in that all organisms (with very few exceptions) use the same code for specifying amino acids with mRNA codons. The genetic code is redundant in that more than one codon can specify the same amino acid. This is because there are more types of codons than there are types of amino acids used by organisms. The genetic code is not ambiguous, however, because each codon always specifies the same amino acid. Four of the codons serve special functions. One operates as a start codon (signaling initiation of translation), and three codons operate as stop codons (signaling the termination of translation).

Slide 5
Transgenesis is the process by which a gene from some organism is transferred to some unrelated organism. In these examples, the tobacco plant and the pig acquire a new trait (the ability to glow) that was present in the organism from which the gene was taken. This demonstrates the universality of the genetic code, because a gene specifies the same polypeptide, regardless of what type of organism expresses that gene.

Slide 6
Transcription occurs in three steps:

- During initiation, RNA polymerase must bind to a sequence of DNA nucleotides, collectively called the promoter, that lie at the beginning of the gene.
- Elongation is the step that builds the mRNA transcript, one nucleotide at a time, as RNA polymerase reads the DNA nucleotide sequence. The sequence of RNA nucleotides is determined by complementarity rules.
- When RNA polymerase reaches the end of the gene, the final step (termination) removes the polymerase molecule from the DNA and removes the mRNA transcript from the polymerase molecule.

Slide 7
In eukaryotic cells, RNA polymerase is not generally able to bind by itself to the promoter. First, some number of transcription factors (which are proteins) must bind to the DNA near the promoter, allowing the RNA polymerase to bind. The polymerase and the transcription factors are collectively called the transcription initiation complex.

Slide 8
In eukaryotic cells, transcription directly produces pre-mRNA, so called because it must undergo post-transcriptional processing before it is able to be used for translation. Part of this modification is the addition of a 5' cap and a poly-A tail.

Slide 9
Another part of post-transcriptional processing that occurs in eukaryotic cells is the removal of introns. These introns are sequences of nucleotides within the pre-mRNA that will not be used as codons. They are therefore non-coding sequences. The coding sequences that remain after the introns have been removed are called exons, and they must be joined end-to-end to make the mature mRNA.

Slide 10
Removal of introns from eukaryotic pre-mRNA involves a combination of proteins and small nuclear RNA (snRNA). The snRNA and the proteins form a complex called a spliceosome. A spliceosome recognizes the introns in a pre-mRNA transcript, cuts them out, and connects the exons, producing the mature mRNA. The excised introns are digested into nucleotides and recycled to make more RNA.

Slide 11
Different exons within a transcript correspond to domains within the polypeptide that is created during translation. Domains are parts of the overall polypeptide. Each domain has a specific shape that gives the polypeptide a specific function. By changing which parts of the gene will correspond to exons in the transcript, a single eukaryotic gene can specify the production of multiple different mRNA molecules (and therefore different polypeptides), depending on the conditions within the cell.

Slide 12
Translation takes place at a ribosome, which is formed when two ribosomal subunits (a small subunit and a large subunit) come together, clamping the mRNA transcript between them. The ribosome then travels along the length of the transcript, one codon at a time, and each codon that is read specifies which kind of transfer RNA (tRNA) molecule is allowed to bring in the next amino acid to be added to the growing polypeptide. A triplet of nucleotides on each tRNA, called the anticodon, is able to recognize a specific and complementary codon in the mRNA transcript. There are therefore many different versions of tRNA, each having its own anticodon. Any given tRNA is able to carry only one type of amino acid, ensuring proper correspondence between mRNA codon and choice of amino acid, according to the unambiguous genetic code.

Slide 13
Transfer RNA (tRNA) is a single polynucleotide strand that folds back on itself to allow for hydrogen bonding between nucleotides of that single strand. There are three hairpin loops, the middle of which includes a sequence of three nucleotides that are collectively called the anticodon. The anticodon is positioned at one end of the three-dimensional structure of tRNA. At the other end are the 3' and 5' ends of the strand. The 3' end is able to attach to a specific amino acid (the amino acid that is specified in the genetic code by the mRNA codon that is complementary to the anticodon of the tRNA). Just as there are multiple possible codons in mRNA, there are multiple possible anticodons, so there are multiple kinds of tRNA. Each tRNA is able to carry only its specific kind of amino acid. This maintains the non-ambiguity of the genetic code.

Side 14
A tRNA that is not carrying an amino acid is said to be uncharged. Charging a tRNA involves putting the right kind of amino acid onto the 3' end of the tRNA. This requires an enzyme called aminoacyl tRNA synthetase. There are multiple versions of this enzyme, each version corresponding to a particular kind of tRNA molecule. At the expense of ATP, aminoacyl tRNA synthetase charges tRNA with the appropriate kind of amino acid.

Slide 15
A complete ribosome consists of two subunits that come together during translation. Each ribosomal subunit is a mixture or ribosomal RNA (rRNA) and protein. The ribosome contains three slots on its surface, each corresponding to one codon on the mRNA transcript. As translation proceeds, the ribosome travels down the length of the mRNA transcript, shifting by one codon at a time, so that a new codon is read by the ribosome each time the ribosome moves.

- The A slot (the aminoacyl tRNA binding site) is where the next codon is read by the ribosome. As a codon is positioned at the A slot, only an aminoacyl tRNA with an anticodon that is complementary to that codon is able to enter the A slot.
- The P slot (the peptidyl tRNA binding site) is occupied by a tRNA that is carrying a string of amino acids (i.e., it's carrying a the growing polypeptide). When the ribosome moves to the next codon, the growing polypeptide is transferred to the single amino acid that is attached to the aminoacyl tRNA that occupies the A slot. This makes the former aminoacyl tRNA a peptidyl tRNA, and it make the former peptidyl tRNA an uncharged tRNA.
- The E slot (the exit slot) receives the uncharged tRNA that previously occupied the P slot before the ribosome moved. The uncharged tRNA leaves the ribosome and will be recharged with its appropriate amino acid by the appropriate aminoacyl tRNA synthetase.

Slide 16
Translation, like transcription, occurs in three steps (initiation, elongation, and termination). Initiation involves the binding of an mRNA transcript to a small ribosomal subunit. This positions the start codon at the position of the P slot, allowing a methionine-carrying tRNA to bind. At the expense of a GTP molecule, a large ribosomal subunit can then join the transcript, the tRNA, and the small subunit to form what is called the translation initiation complex.

Slide 17
In the elongation phase of translation, the polypeptide is enlarged one amino acid at a time. The tRNA that occupies the P slot on the ribosome is carrying the string of amino acids that have been joined so far. A new tRNA (with the appropriate anticodon and therefore carrying the appropriate amino acid) recognizes the mRNA codon at the A slot, and enters the slot, bringing in the next amino acid. The string of amino acids on the tRNA in the P slot is then transferred to the amino acid on the tRNA in the A slot, leaving the P-slot tRNA uncharged (empty). As the ribosome shifts along the transcript by one codon, the uncharged tRNA is put into the E slot and exits the ribosome. At the same time, the tRNA carrying the growing strand is put into the P slot, and the A slot is now open, reading the next codon. Assembly of a polypeptide is expensive, requiring energy in the form of two GTP molecules for each amino acid that is added during elongation.

Slide 18
Termination of translation occurs when any of three stop codons appears at the A slot. In this case, no tRNA enters the A slot; instead, a special protein called a release factor enters, and this causes disintegration of the translation complex. The two ribosomal subunits separate, the mRNA transcript is released, and the completed polypeptide is freed from the tRNA that was carrying it.

Slide 19
A polyribosome is a complex consisting of one mRNA transcript and multiple ribosomes, each of which are translating that transcript simultaneously but at different locations along the length of the transcript. A polyribosome allows for the production of several identical copies of the polypeptide in a short amount of time, using the same transcript.

Slide 20
In all cases, translation begins in the cytosol, at a free ribosome. For some polypeptides, the first several amino acids to emerge from the ribosome serve as a signal that the polypeptide under construction should be completed in the rough endoplasmic reticulum rather than being completed in the cytosol. This signal peptide is able to bind to a signal recognition particle, and that binding causes the ribosomal complex to be associated with the membrane of the rough endoplasmic reticulum, with the emerging polypeptide threaded through a pore into the lumen of the endoplasmic reticulum. When translation is completed, the polypeptide will be released into the ER lumen, and the ribosomal subunits and the transcript will dissociate in the cytosol.

Slide 21
A mutation is an accidental change in the DNA nucleotide sequence of a gene. A change at even one nucleotide position can have drastic effects, as is illustrated by sickle cell disease. Just one of the DNA triplets (CTT) in the normal gene mutates into CAT. Transcription occurs as usual for either the normal person or the one with the mutation. But transcription of the normal CTT triplet specifies GAA as an mRNA codon, whereas the mutant DNA triplet (CAT) specifies GUA. During translation, the normal codon (GAA) specifies glutamic acid as the amino acid, but the mutant codon (GUA) specifies valine as the amino acid. This incorrect amino acid gives hemoglobin a different shape, making hemoglobin much less able to effectively carry oxygen within red blood cells of people with sickle cell disease.

Slide 22
A point mutation is a mutation caused by a change in a single nucleotide position (i.e., one DNA base pair). Point mutations are classified differently, depending on whether the mutation causes a change in the length of the gene.

- A substitution is a mutation that does not change the length of the gene, because one base pair is substituted for another. A substitution could be a silent mutation (having no effect, because the mutated codon still codes for the same amino acid), a missense mutation (causing placement of a different amino acid at the corresponding place in the polypeptide), or a nonsense mutation (causing a truncated polypeptide because the mutated codon serves as a stop codon).
- An insertion or a deletion is a mutation that changes the length of the gene. An insertion adds a base pair, and a deletion removes a base pair. Either way, a frameshift results, because the frame of reference that determines how codons are read is shifted. Although an insertion or a deletion is a point mutation (and therefore changes only one base pair), the resulting frameshift affects every codon from that point onward, so the resulting mutant polypeptide is substantially different from the wild type polypeptide.

Slide 23
Because a prokaryotic cell does not have a nucleus, transcription and translation not only occur in the same compartment, but they can overlap each other in time. As soon as transcription produces enough nucleotides (though the transcript is not complete), a ribosome can attach to the incomplete transcript and begin translation. Therefore, gene expression in a prokaryote is quicker compared to that in a eukaryote.

Slide 24
The steps involved in eukaryotic gene expression are summarized.