

Overview: The Flow of Genetic Information

- The information content of DNA is in the form of specific sequences of nucleotides
- The DNA inherited by an organism leads to specific traits by dictating the synthesis of proteins
- Proteins are the links between genotype and phenotype
- **Gene expression**, the process by which DNA directs protein synthesis, includes two stages: transcription and translation

Concept 17.1: Genes specify proteins via transcription and translation

- How was the fundamental relationship between genes and proteins discovered?

Evidence from the Study of Metabolic Defects

- In 1902, British physician Archibald Garrod first suggested that genes dictate phenotypes through enzymes that catalyze specific chemical reactions
- He thought symptoms of an inherited disease reflect an inability to synthesize a certain enzyme
- Linking genes to enzymes required understanding that cells synthesize and degrade molecules in a series of steps, a metabolic pathway

Nutritional Mutants in Neurospora: Scientific Inquiry

- George Beadle and Edward Tatum exposed bread mold to X-rays, creating mutants that were unable to survive on minimal media
- Using crosses, they and their coworkers identified three classes of arginine-deficient mutants, each lacking a different enzyme necessary for synthesizing arginine
- They developed a one gene–one enzyme hypothesis, which states that each gene dictates production of a specific enzyme

The Products of Gene Expression: A Developing Story

- Some proteins aren't enzymes, so researchers later revised the hypothesis: one gene–one protein
- Many proteins are composed of several polypeptides, each of which has its own gene
- Therefore, Beadle and Tatum's hypothesis is now restated as the one gene–one polypeptide hypothesis
- Note that it is common to refer to gene products as proteins rather than polypeptides

Basic Principles of Transcription and Translation

- RNA is the bridge between genes and the proteins for which they code
- **Transcription** is the synthesis of RNA using information in DNA
- Transcription produces **messenger RNA (mRNA)**
- **Translation** is the synthesis of a polypeptide, using information in the mRNA
- **Ribosomes** are the sites of translation
- In prokaryotes, translation of mRNA can begin before transcription has finished
- In a eukaryotic cell, the nuclear envelope separates transcription from translation
- Eukaryotic RNA transcripts are modified through RNA processing to yield the finished mRNA
- A **primary transcript** is the initial RNA transcript from any gene prior to processing
- The *central dogma* is the concept that cells are governed by a cellular chain of command: DNA → RNA → protein

The Genetic Code

- How are the instructions for assembling amino acids into proteins encoded into DNA?

- There are 20 amino acids, but there are only four nucleotide bases in DNA
- How many nucleotides correspond to an amino acid?

Codons: Triplets of Nucleotides

- The flow of information from gene to protein is based on a **triplet code**: a series of nonoverlapping, three-nucleotide words
- The words of a gene are transcribed into complementary nonoverlapping three-nucleotide words of mRNA
- These words are then translated into a chain of amino acids, forming a polypeptide
- During transcription, one of the two DNA strands, called the **template strand**, provides a template for ordering the sequence of complementary nucleotides in an RNA transcript
- The template strand is always the same strand for a given gene
- During translation, the mRNA base triplets, called **codons**, are read in the 5' to 3' direction
- Codons along an mRNA molecule are read by translation machinery in the 5' to 3' direction
- Each codon specifies the amino acid (one of 20) to be placed at the corresponding position along a polypeptide

Cracking the Code

- All 64 codons were deciphered by the mid-1960s
- Of the 64 triplets, 61 code for amino acids; 3 triplets are “stop” signals to end translation
- The genetic code is redundant (more than one codon may specify a particular amino acid) but not ambiguous; no codon specifies more than one amino acid
- Codons must be read in the correct **reading frame** (correct groupings) in order for the specified polypeptide to be produced

Evolution of the Genetic Code

- The genetic code is nearly universal, shared by the simplest bacteria to the most complex animals
- Genes can be transcribed and translated after being transplanted from one species to another

Concept 17.2: Transcription is the DNA-directed synthesis of RNA: a closer look

- Transcription is the first stage of gene expression

Molecular Components of Transcription

- RNA synthesis is catalyzed by **RNA polymerase**, which pries the DNA strands apart and hooks together the RNA nucleotides
- The RNA is complementary to the DNA template strand
- RNA synthesis follows the same base-pairing rules as DNA, except that uracil substitutes for thymine
- The DNA sequence where RNA polymerase attaches is called the **promoter**; in bacteria, the sequence signaling the end of transcription is called the **terminator**
- The stretch of DNA that is transcribed is called a **transcription unit**

Synthesis of an RNA Transcript

- The three stages of transcription
 - Initiation

- Elongation
- Termination

RNA Polymerase Binding and Initiation of Transcription

- Promoters signal the transcriptional **start point** and usually extend several dozen nucleotide pairs upstream of the start point
- **Transcription factors** mediate the binding of RNA polymerase and the initiation of transcription
- The completed assembly of transcription factors and RNA polymerase II bound to a promoter is called a **transcription initiation complex**
- A promoter called a **TATA box** is crucial in forming the initiation complex in eukaryotes

Elongation of the RNA Strand

- As RNA polymerase moves along the DNA, it untwists the double helix, 10 to 20 bases at a time
- Transcription progresses at a rate of 40 nucleotides per second in eukaryotes
- A gene can be transcribed simultaneously by several RNA polymerases
- Nucleotides are added to the 3' end of the growing RNA molecule

Termination of Transcription

- The mechanisms of termination are different in bacteria and eukaryotes
- In bacteria, the polymerase stops transcription at the end of the terminator and the mRNA can be translated without further modification
- In eukaryotes, RNA polymerase II transcribes the polyadenylation signal sequence; the RNA transcript is released 10–35 nucleotides past this polyadenylation sequence

Concept 17.3: Eukaryotic cells modify RNA after transcription

- Enzymes in the eukaryotic nucleus modify pre-mRNA (**RNA processing**) before the genetic messages are dispatched to the cytoplasm
- During RNA processing, both ends of the primary transcript are usually altered
- Also, usually some interior parts of the molecule are cut out, and the other parts spliced together

Alteration of mRNA Ends

- Each end of a pre-mRNA molecule is modified in a particular way
 - The 5' end receives a modified nucleotide **5' cap**
 - The 3' end gets a **poly-A tail**
- These modifications share several functions
 - They seem to facilitate the export of mRNA to the cytoplasm
 - They protect mRNA from hydrolytic enzymes
 - They help ribosomes attach to the 5' end

Split Genes and RNA Splicing

- Most eukaryotic genes and their RNA transcripts have long noncoding stretches of nucleotides that lie between coding regions
- These noncoding regions are called intervening sequences, or **introns**
- The other regions are called **exons** because they are eventually expressed, usually translated into amino acid sequences
- **RNA splicing** removes introns and joins exons, creating an mRNA molecule with a continuous coding sequence

- In some cases, RNA splicing is carried out by spliceosomes
- **Spliceosomes** consist of a variety of proteins and several small nuclear ribonucleoproteins (snRNPs) that recognize the splice sites

Ribozymes

- **Ribozymes** are catalytic RNA molecules that function as enzymes and can splice RNA
- The discovery of ribozymes rendered obsolete the belief that all biological catalysts were proteins
- Three properties of RNA enable it to function as an enzyme
 - It can form a three-dimensional structure because of its ability to base-pair with itself
 - Some bases in RNA contain functional groups that may participate in catalysis
 - RNA may hydrogen-bond with other nucleic acid molecules

The Functional and Evolutionary Importance of Introns

- Some introns contain sequences that may regulate gene expression
- Some genes can encode more than one kind of polypeptide, depending on which segments are treated as exons during splicing
- This is called **alternative RNA splicing**
- Consequently, the number of different proteins an organism can produce is much greater than its number of genes
- Proteins often have a modular architecture consisting of discrete regions called **domains**
- In many cases, different exons code for the different domains in a protein
- Exon shuffling may result in the evolution of new proteins

Concept 17.4: Translation is the RNA-directed synthesis of a polypeptide: a closer look

- Genetic information flows from mRNA to protein through the process of translation

Molecular Components of Translation

- A cell translates an mRNA message into protein with the help of **transfer RNA (tRNA)**
- tRNAs transfer amino acids to the growing polypeptide in a ribosome
- Translation is a complex process in terms of its biochemistry and mechanics

The Structure and Function of Transfer RNA

- Molecules of tRNA are not identical
 - Each carries a specific amino acid on one end
 - Each has an **anticodon** on the other end; the anticodon base-pairs with a complementary codon on mRNA
- A tRNA molecule consists of a single RNA strand that is only about 80 nucleotides long
- Flattened into one plane to reveal its base pairing, a tRNA molecule looks like a cloverleaf
- Because of hydrogen bonds, tRNA actually twists and folds into a three-dimensional molecule
- tRNA is roughly L-shaped
- Accurate translation requires two steps

- First: a correct match between a tRNA and an amino acid, done by the enzyme **aminoacyl-tRNA synthetase**
- Second: a correct match between the tRNA anticodon and an mRNA codon
- Flexible pairing at the third base of a codon is called **wobble** and allows some tRNAs to bind to more than one codon

Ribosomes

- Ribosomes facilitate specific coupling of tRNA anticodons with mRNA codons in protein synthesis
- The two ribosomal subunits (large and small) are made of proteins and **ribosomal RNA (rRNA)**
- Bacterial and eukaryotic ribosomes are somewhat similar but have significant differences: some antibiotic drugs specifically target bacterial ribosomes without harming eukaryotic ribosomes
- A ribosome has three binding sites for tRNA
 - The **P site** holds the tRNA that carries the growing polypeptide chain
 - The **A site** holds the tRNA that carries the next amino acid to be added to the chain
 - The **E site** is the exit site, where discharged tRNAs leave the ribosome

Building a Polypeptide

- The three stages of translation
 - Initiation
 - Elongation
 - Termination
- All three stages require protein “factors” that aid in the translation process

Ribosome Association and Initiation of Translation

- The initiation stage of translation brings together mRNA, a tRNA with the first amino acid, and the two ribosomal subunits
- First, a small ribosomal subunit binds with mRNA and a special initiator tRNA
- Then the small subunit moves along the mRNA until it reaches the start codon (AUG)
- Proteins called initiation factors bring in the large subunit that completes the translation initiation complex

Elongation of the Polypeptide Chain

- During the elongation stage, amino acids are added one by one to the preceding amino acid at the C-terminus of the growing chain
- Each addition involves proteins called elongation factors and occurs in three steps: codon recognition, peptide bond formation, and translocation
- Translation proceeds along the mRNA in a 5' to 3' direction

Termination of Translation

- Termination occurs when a stop codon in the mRNA reaches the A site of the ribosome
- The A site accepts a protein called a release factor
- The release factor causes the addition of a water molecule instead of an amino acid
- This reaction releases the polypeptide, and the translation assembly then comes apart

Polyribosomes

- A number of ribosomes can translate a single mRNA simultaneously, forming a **polyribosome** (or **polysome**)

- Polyribosomes enable a cell to make many copies of a polypeptide very quickly

Completing and Targeting the Functional Protein

- Often translation is not sufficient to make a functional protein
- Polypeptide chains are modified after translation or targeted to specific sites in the cell

Protein Folding and Post-Translational Modifications

- During and after synthesis, a polypeptide chain spontaneously coils and folds into its three-dimensional shape
- Proteins may also require post-translational modifications before doing their job
- Some polypeptides are activated by enzymes that cleave them
- Other polypeptides come together to form the subunits of a protein

Targeting Polypeptides to Specific Locations

- Two populations of ribosomes are evident in cells: free ribosomes (in the cytosol) and bound ribosomes (attached to the ER)
- Free ribosomes mostly synthesize proteins that function in the cytosol
- Bound ribosomes make proteins of the endomembrane system and proteins that are secreted from the cell
- Ribosomes are identical and can switch from free to bound
- Polypeptide synthesis always begins in the cytosol
- Synthesis finishes in the cytosol *unless* the polypeptide signals the ribosome to attach to the ER
- Polypeptides destined for the ER or for secretion are marked by a **signal peptide**
- A **signal-recognition particle (SRP)** binds to the signal peptide
- The SRP brings the signal peptide and its ribosome to the ER

Concept 17.5: Mutations of one or a few nucleotides can affect protein structure and function

- **Mutations** are changes in the genetic material of a cell or virus
- **Point mutations** are chemical changes in just one base pair of a gene
- The change of a single nucleotide in a DNA template strand can lead to the production of an abnormal protein

Types of Small-Scale Mutations

- Point mutations within a gene can be divided into two general categories
 - Nucleotide-pair substitutions
 - One or more nucleotide-pair insertions or deletions

Substitutions

- A **nucleotide-pair substitution** replaces one nucleotide and its partner with another pair of nucleotides
- **Silent mutations** have no effect on the amino acid produced by a codon because of redundancy in the genetic code
- **Missense mutations** still code for an amino acid, but not the correct amino acid
- **Nonsense mutations** change an amino acid codon into a stop codon, nearly always leading to a nonfunctional protein

Insertions and Deletions

- **Insertions** and **deletions** are additions or losses of nucleotide pairs in a gene
- These mutations have a disastrous effect on the resulting protein more often than substitutions do

- Insertion or deletion of nucleotides may alter the reading frame, producing a **frameshift mutation**

Mutagens

- Spontaneous mutations can occur during DNA replication, recombination, or repair
- **Mutagens** are physical or chemical agents that can cause mutations

Concept 17.6: While gene expression differs among the domains of life, the concept of a gene is universal

- Archaea are prokaryotes, but share many features of gene expression with eukaryotes

Comparing Gene Expression in Bacteria, Archaea, and Eukarya

- Bacteria and eukarya differ in their RNA polymerases, termination of transcription, and ribosomes; archaea tend to resemble eukarya in these respects
- Bacteria can simultaneously transcribe and translate the same gene
- In eukarya, transcription and translation are separated by the nuclear envelope
- In archaea, transcription and translation are likely coupled

What Is a Gene? *Revisiting the Question*

- The idea of the gene has evolved through the history of genetics
- We have considered a gene as
 - A discrete unit of inheritance
 - A region of specific nucleotide sequence in a chromosome
 - A DNA sequence that codes for a specific polypeptide chain
- In summary, a gene can be defined as a region of DNA that can be expressed to produce a final functional product, either a polypeptide or an RNA molecule