Lecture Outline: Transcription and Translation

I. **Gene Expression: Overview and Foundational Concepts**

A. Definition of Gene Expression

1. A gene "doing its job," being **turned on** or **activated**.
2. Main job of DNA besides replication is transcription.
3. Transcription followed by translation leads to a polypeptide with a specific cellular function.

B. Historical Context: "One Gene-One Enzyme" Hypothesis Evolution

1. Original Hypothesis and Bread Mold Experiment
   a. Famous experiment using bread mold (Neurospora) due to ease of growth.
   b. Mutations intentionally caused by bombarding with UV light (random process).
   c. **Wild type**: naturally most prevalent form; grows on **minimal medium** (bare minimum for life).
   d. Mutants: differed genetically due to mutations.
      (1) Class 1 mutants grew on minimal medium plus ornithine.
      (2) Class 2 mutants grew on minimal medium plus citrulline.
      (3) Class 3 mutants grew on minimal medium plus arginine.
   h. **Ornithine, citrulline, and arginine** are intermediates in a biochemical pathway.
   i. Each step in a biochemical pathway requires its own **enzyme**.
   j. Deduction: A gene individually codes for a specific enzyme.
   k. Original hypothesis: **One gene one enzyme hypothesis**.
2. Evolution of the Hypothesis
   a. **Incorrect portrayal of reality** for multiple reasons.

b. Not all genes code for enzymes; some code for other types of proteins.

c. Revised to **One gene one protein hypothesis** (still incorrect).

d. Some proteins are **multimeric proteins** (quaternary structure), requiring multiple polypeptides.

(1) Example: Hemoglobin requires two different genes to code for alpha and gamma polypeptides.

f. Revised to **One gene one polypeptide hypothesis** (closer, but still wrong).

g. Reasons for continued inaccuracy:

(1) Not all genes code for polypeptides (e.g., non-protein coding genes).

(2) In eukaryotes, a **single gene can code for multiple different polypeptides** due to alternative splicing.

C. Role of DNA and RNA in Gene Expression

1. DNA **indirectly** gives traits; **proteins directly** give traits.
2. DNA is directly involved in **transcription**.
3. Transcription produces an **RNA molecule** (e.g., messenger RNA, mRNA) that holds the code.
4. mRNA then holds the code for assembling the polypeptide (translation).
5. **Non-protein coding genes** produce other types of RNA (e.g., ribosomal RNA, transfer RNA) as their end product.

D. Comparison: Prokaryotic vs. Eukaryotic Gene Expression

1. Fundamental Difference: Presence or Absence of Nucleus
   a. **Prokaryotic cells**: No nucleus; DNA is in the cytoplasm.
   b. **Eukaryotic cells**: Have a nucleus; DNA is in the nucleus.
2. Location of Transcription and Translation
   a. **Eukaryotic cells**:
      (1) Transcription occurs in the **nucleus** (where DNA is).
      (2) Translation occurs in the **cytoplasm**.

(3) Processes occur in separate compartments.

e. **Prokaryotic cells**:

(1) Transcription and translation both occur in the **cytoplasm** (the only compartment).

(2) More streamlined process.

3. mRNA Processing

a. **Eukaryotic cells**: mRNA (**pre-mRNA**) requires **post-transcriptional modification** to become mature and exit the nucleus.

b. **Prokaryotic cells**: mRNA produced by transcription is immediately ready for translation.

E. Core Processes: Transcription and Translation Defined

1. **Transcription**: "Rewriting" or "copying" the code from DNA to RNA.

a. Language remains the same (nucleotides), just a different dialect (DNA vs. RNA).

b. DNA is directly involved.

2. **Translation**: "Change in language" from nucleotides (mRNA code) to amino acids (polypeptide).

a. Occurs after transcription.

b. DNA is not directly involved in translation after its code is transcribed.

II. **The Genetic Code**

A. Mathematical Basis for Codons

1. DNA and RNA both use 4 different nucleotides.

2. Transcription has a 1:1 correspondence between DNA and mRNA nucleotides.

3. Translation does not have a 1:1 correspondence between mRNA nucleotides and amino acids due to limited nucleotide types (4) and more amino acid types (20).

4. To specify 20 amino acids with 4 nucleotides:

a. One nucleotide: 4 possibilities (insufficient).

b. Two nucleotides: 4*4 = 16 possibilities (insufficient).

c. Three nucleotides: 4*4*4 = **64 possibilities** (sufficient).

5. A group of three consecutive nucleotides is called a **codon**.

B. Properties of Codons

1. Total 64 codons.
2. **Stop codons**: 3 codons (UAA, UAG, UGA) signal the end of translation and do not code for any amino acid.
3. **Start codon**: 1 codon (AUG) signals the beginning of translation and also codes for the amino acid **Methionine (Met)**.

a. All polypeptides initially start with Methionine.

4. **Redundant (Repetitive)**: More than one codon can specify the same amino acid (e.g., Leucine can be coded by 6 different codons).
5. **Not Ambiguous (Not Vague)**: Each codon specifies **only one** amino acid. This ensures the correct and consistent protein sequence.

C. Universality and Genetic Engineering

1. The genetic code is **universal**: almost all organisms use the same genetic code with very few exceptions.
2. This universality is strong evidence for a **single original cell lineage** (common ancestor).
3. Genetic engineering works because unrelated organisms (e.g., tobacco plant and firefly, pig and jellyfish) can understand and translate genes from one another, producing the same proteins, leading to observed traits like glowing.

III. **Transcription: The First Step**

A. Definition and Purpose

1. The process where the genetic information from a DNA gene is rewritten into an RNA molecule (transcript).
2. Directly involves DNA.

B. Stages of Transcription

1. Initiation (Starting)

a. The **gene (transcription unit)** is recognized.

b. A **promoter**, a specific DNA nucleotide sequence just before a gene, acts as the binding site.
c. **RNA polymerase**, a protein complex, binds to the promoter.
d. Binding of RNA polymerase initiates transcription, turning the gene "on".
e. RNA polymerase separates the two DNA strands.
f. Only **one DNA strand** (the **template strand**) is read. The other strand is temporarily moved aside.
g. RNA is synthesized by laying down RNA nucleotides complementary to the template strand.
h. RNA transcript grows in the **5' to 3' direction**.

2. Elongation (Making it Longer)
    a. RNA polymerase continues to move along the DNA template strand.
    b. RNA transcript grows longer, one nucleotide at a time.
3. Termination (Ending)
    a. RNA polymerase reaches the end of the gene.
    b. The transcription complex (DNA, RNA polymerase, RNA transcript) separates.
    c. The RNA transcript is released.
    d. The DNA strands naturally re-associate.
    e. The gene is temporarily "turned off" again.

C. Control of Transcription (Eukaryotes)

1. In eukaryotes, RNA polymerase usually cannot bind to the promoter by itself.
2. Other proteins, called **transcription factors**, must bind to the promoter first.
3. Transcription factors are themselves proteins, coded for by other genes.
4. The combination of the promoter, transcription factors, and RNA polymerase forms a **transcription initiation complex**.

5. This complex process allows for precise control of gene expression, turning genes on or off as needed, since every cell contains the entire genome but only expresses specific genes.

IV. **Post-Transcriptional Modification of Eukaryotic mRNA**

A. Overview

1. In eukaryotes, the direct product of transcription is **pre-mRNA** (primary messenger RNA).
2. Pre-mRNA is not yet in its final, mature form and requires modification before translation.
3. This process is called **post-transcriptional modification** and does not occur in prokaryotes.

B. Key Modifications

1. 5' Cap and Poly-A Tail
   a. A **5' cap** is added to the 5' end of the mRNA.
   b. A **Poly-A tail** (a long sequence of adenine nucleotides) is added to the 3' end.
   c. Functions:
      (1) Allow the mRNA to **exit the nucleus** through nuclear pores and reach the cytoplasm (where ribosomes are).
      (2) Help in the **binding of mRNA to ribosomes** for proper translation.
   f. Untranslated regions (UTRs) exist at both ends of the mRNA, outside the coding sequence, which are important for translation initiation but not themselves translated.
2. Splicing (Introns and Exons)
   a. Parts of the pre-mRNA are removed, and the remaining parts are joined together.
   b. **Exons**: "Expressed" regions that code for amino acids and will be part of the final, mature mRNA.
   c. **Introns**: "Intervening sequences" that are cut out or **spliced out** and are not translated.

d. The mature mRNA is shorter than the pre-mRNA after introns are removed and exons are "sewn together".

e. This process is performed by a molecular machine called a **spliceosome**.

f. Spliceosomes are complexes of protein and **small nuclear ribonucleoproteins (snRNPs)**, which contain **small nuclear RNA (snRNA)**.

g. The snRNA recognizes sequences in the introns for precise cutting.

h. Cutout introns are digested and their nucleotides recycled.

C. Alternative mRNA Splicing

1. Eukaryotes can change which sections are considered exons and introns under different circumstances.
2. This allows a **single gene to code for multiple different polypeptides**.
3. This mechanism provides more efficient use of the genome, as one gene can produce variations of a protein (e.g., enzyme family members optimized for different conditions).
4. It explains how humans can have roughly 20,000 genes but produce a greater number of different proteins.

V. **Translation: Assembling the Polypeptide**

A. Overview and Components

1. The process of converting the mRNA code into a sequence of amino acids (polypeptide).
2. Occurs in the cytoplasm in both prokaryotes and eukaryotes.
3. Does not involve DNA.
4. Components:
   a. **Messenger RNA (mRNA)**: The transcript carrying the genetic code (codons).
   b. **Ribosomes**: Molecular machines where translation occurs, composed of ribosomal RNA (rRNA) and protein.
   c. **Transfer RNA (tRNA)**: Molecules that transport specific amino

acids to the ribosome.

B. Transfer RNA (tRNA) Function and Structure

1. A single RNA strand that folds into a 3D bent-rod shape, often depicted as a cloverleaf in 2D.
2. Its structure is maintained by intramolecular hydrogen bonds due to self-complementarity.
3. **Anti-codon**: A specific three-nucleotide sequence located on one of the hairpin loops.
    a. It is **complementary** to a specific codon on the mRNA.
    b. Determines which amino acid the tRNA is allowed to carry.
4. **3' end**: The site where the specific amino acid attaches.
5. Each type of tRNA can only carry one specific type of amino acid, ensuring the correct protein sequence is assembled.
6. tRNA Categories
    a. **Uncharged tRNA**: Not carrying an amino acid (empty).
    b. **Charged tRNA**: Carrying at least one amino acid.
        (1) **Aminoacyl tRNA**: A charged tRNA carrying **exactly one** amino acid.
        (2) **Peptidyl tRNA**: A charged tRNA carrying **two or more** amino acids (a growing peptide chain).

C. Charging of tRNA

1. The process by which an uncharged tRNA becomes an aminoacyl tRNA.
2. Catalyzed by specific **aminoacyl-tRNA synthetase enzymes**.
3. This enzyme brings together the correct uncharged tRNA and its corresponding amino acid.
4. Requires energy, typically from **ATP hydrolysis**, to form the bond between the amino acid and the tRNA.
5. There are different kinds of aminoacyl-tRNA synthetases, each specific for a particular tRNA/amino acid pair, maintaining accuracy.
6. Once charged, the aminoacyl tRNA is released and becomes part of a

pool ready for translation.

7. Uncharged tRNAs (after delivering their amino acids) return to this enzyme pool to be recharged and reused.

D. Ribosome Structure and Binding Sites

1. Composed of a **small ribosomal subunit** and a **large ribosomal subunit**.
2. Subunits are modular and only come together when translation is occurring.
3. Ribosomes are made of **ribosomal RNA (rRNA)** and protein.
4. The rRNA within the ribosome can act as a biological catalyst, known as a **ribozyme**.
5. The large subunit has three conspicuous binding slots for tRNA molecules:
    a. **A slot (Aminoacyl-tRNA binding site)**: Where incoming aminoacyl tRNAs enter, carrying their single amino acid.
    b. **P slot (Peptidyl-tRNA binding site)**: Where the tRNA carrying the growing polypeptide chain is held.
    c. **E slot (Exit site)**: Where uncharged tRNAs exit the ribosome after releasing their amino acid.

E. Stages of Translation

1. Initiation (Starting)
    a. The **mature mRNA first binds to the small ribosomal subunit**.
    b. Untranslated regions (UTRs) on the mRNA help position it correctly.
    c. The **start codon (AUG)** on the mRNA is positioned in what will become the P slot.
    d. The first tRNA, carrying **Methionine** and having the complementary anti-codon (UAC), binds to the AUG in the P slot.
    e. The **large ribosomal subunit then joins** the complex.
    f. This entire assembly is called the **translation initiation complex**.
    g. This step requires energy, typically from **GTP hydrolysis**.
2. Elongation (Making it Longer) - A Cyclic Process

a. The A slot is initially empty.
b. The ribosome reads the next codon in the **A slot**.
c. The appropriate **aminoacyl tRNA** (carrying its specific amino acid) enters the A slot, binding to the codon via its anti-codon.
d. A **peptide bond forms** between the amino acid in the A slot and the growing polypeptide chain (held by the tRNA in the P slot). The polypeptide is transferred from the P slot tRNA to the A slot tRNA.
e. The tRNA in the A slot now becomes a **peptidyl tRNA**.
f. The ribosome then **shifts by one codon** (three nucleotides) along the mRNA.
g. The former A slot tRNA (now peptidyl tRNA) moves to the **P slot**.
h. The former P slot tRNA (now uncharged) moves to the **E slot** and is ejected from the ribosome.
i. The A slot becomes empty, ready for the next incoming aminoacyl tRNA.
j. This cycle repeats for each amino acid to be added.
k. Each cycle of elongation requires energy (typically **two GTPs per amino acid added**).

3. Termination (Ending)
   a. Elongation continues until one of the **three stop codons** (UAA, UAG, UGA) appears in the A slot.
   b. There are **no tRNAs corresponding to stop codons**.
   c. Instead, a special protein called a **release factor** binds in the A slot.
   d. The release factor signals the complex that the polypeptide is complete.
   e. The completed polypeptide is released from the tRNA in the P slot.
   f. The ribosomal subunits, release factor, and mRNA detach from each other.
   g. This step also requires energy (**GTP**).

F. Efficiency of Translation

1. **mRNA is reusable**: A single mRNA molecule can be translated multiple times to produce many copies of the same protein.
2. Cells actively destroy mRNA molecules (**mRNA degradation**) when the protein they code for is no longer needed, to conserve energy and prevent interference.
3. Destroying mRNA is not problematic because the cell can always transcribe more mRNA from the original gene (the "library").
4. A gene must be transcribed repeatedly throughout the cell's lifetime to meet protein demands.

VI. **Post-Translational Events and Protein Targeting**

A. Polyribosomes (Polysomes)

1. A single mRNA transcript can be translated simultaneously by multiple ribosomes.
2. Each ribosome is at a different stage of translation, forming a "polyribosome" or "polysome".
3. This is a highly efficient way to produce many identical copies of a protein quickly from one mRNA.
4. Occurs in both prokaryotic and eukaryotic cells.

B. Coupled Transcription-Translation (Prokaryotes Only)

1. In prokaryotes, transcription and translation can occur simultaneously.
2. Ribosomes can attach to and begin translating an mRNA molecule even before its transcription is complete.
3. This rapid, "coupled" process is possible because prokaryotes lack a nucleus, and both processes happen in the cytoplasm.
4. It significantly speeds up protein production, which is vital for unicellular organisms.

C. Protein Targeting and Post-Translational Modification

1. Free vs. Bound Ribosomes
    a. All translation begins on **free ribosomes** in the cytosol.
    b. **Free ribosomes**: float in the cytoplasm, not attached to membranes.

c. **Bound ribosomes**: attached to the membrane of the **rough endoplasmic reticulum (RER)**.

d. Proteins synthesized entirely on free ribosomes typically function in the cytosol and may not require further modification.

2. Signal Peptides and ER Targeting

a. Cells distinguish between proteins that need post-translational modification and those that don't.

b. For proteins requiring modification or destined for specific cellular locations (e.g., secretion, membrane insertion), the initial amino acid sequence emerging from the ribosome is a **signal peptide**.

c. A **Signal Recognition Particle (SRP)**, a protein, recognizes and binds to this signal peptide.

d. The entire translation complex (ribosome + mRNA + growing polypeptide + SRP) is then translocated to the RER membrane, often via the cytoskeleton.

e. The SRP then binds to an **SRP receptor protein**, an integral transmembrane protein in the ER membrane, which is part of a pore.

f. This docking positions the ribosome so that the growing polypeptide chain is threaded through the pore into the **ER lumen** (the space inside the ER).

g. Once inside the ER, proteins can undergo various **post-translational modifications**:

(1) Removal of the initial Methionine or the entire signal peptide.

(2) Folding into correct 3D shapes.

(3) Addition of functional groups (e.g., sugars to form glycoproteins).

(4) Assembly of multiple polypeptide chains into multimeric proteins (e.g., hemoglobin).

l. Proteins can remain in the ER lumen, be inserted into the ER membrane, or be packaged into vesicles for transport to other

organelles (like the Golgi apparatus) or for secretion out of the cell.

VII. **Mutations: Changes in Genetic Information**

A. Definition and Types

1. A **mutation** is a change in the DNA sequence.
2. **Point mutation**: The simplest type, affecting just a single nucleotide pair (one position on two DNA strands).
    a. Example: Sickle cell disease results from a single point mutation in the hemoglobin gene, changing one amino acid (Glutamic acid to Valine) and disrupting protein function.
3. **Large scale mutations**: Affect large chunks of a chromosome (thousands or millions of nucleotides), potentially moving them around, causing major effects.

B. Categories of Point Mutations

1. Base Pair Substitutions
    a. One nucleotide is replaced or "substituted" by a different nucleotide.
    b. The overall length of the DNA sequence remains unchanged.
    c. Types of effects:
        (1) **Silent mutation**: A substitution that results in a new codon that still codes for the **same amino acid** due to the redundancy of the genetic code. The protein sequence is unaffected.
        (2) **Missense mutation**: A substitution that changes one codon into a codon for a **different amino acid**. This alters the protein sequence at a single point, potentially affecting its function (could be neutral, harmful, or rarely beneficial for natural selection).
        (3) **Nonsense mutation**: A substitution that changes an amino acid codon into one of the **three stop codons**. This leads to premature termination of the polypeptide chain, often resulting in a non-functional protein and is generally more severe than a missense mutation.
2. Insertions or Deletions (Indels)
    a. **Insertion**: An extra nucleotide is added into the sequence,

increasing its length.

b. **Deletion**: A nucleotide is removed from the sequence without replacement, making it shorter.

c. These mutations cause a **frameshift**:

    (1) Because codons are read in groups of three (the "reading frame"), inserting or deleting even one nucleotide shifts the entire downstream sequence of codons.

    (2) This leads to changes in **all amino acids downstream** from the mutation.

f. Frameshift mutations typically have a more profound and severe effect on the resulting protein compared to substitutions, often leading to non-functional proteins or premature termination.