Gene Expression Unit Topics

- Discovery of DNA
- Nucleic Acid Structure
- DNA Replication
- Transcription & Translation
- Recombinant DNA

Discovery of DNA

The precipitate at the bottom of this flask is Deoxyribose Nucleic Acid. This chemical is the informational basis for all life. Its properties allow for the storage of instructions to build living things.

Every molecule, organelle, cell, organ, organ system, organism and population is built by this molecule.

It is the building block of the genes that control how bodies are shaped and how organisms react to environmental factors.

This molecule is evolution.
"Individuals are not stable things, they are fleeting. Chromosomes too are shuffled into oblivion, like hands of cards soon after they are dealt. But the cards themselves survive the shuffling. The cards are the genes. They merely change partners and march on.

They are the replicators and we are their survival machines. When we have served our purpose we are cast aside. But genes are denizens of geological time: genes are forever."

Richard Dawkins, Evolutionary Biologist and Oxford University professor.

As an example, blue eyes are a phenotype; a physical trait, controlled by a single gene.

A recent study showed that a mutation in one individual's OCA2 gene, which produces the pigment that gives color to eyes, created a gene for blue eyes. This occurred 8,000 years ago and the new gene was passed generation to generation.

Today approximately 560,000,000 people have blue eyes. Each individual carries 2 copies of the original mutation. The gene has long outlived the human that it originated in.

Nucleic acids were first isolated by the Swiss physician Friedrich Miescher who, in 1869, discovered a microscopic substance in the pus of discarded surgical bandages.

At the time it was an unknown cellular substance and was not considered important until many years later.

Frederick Griffith in 1928 conducted the following experiment using 2 different strains of the bacterium *Streptococcus pneumoniae*.

- S strain bacteria kills mice.
- R strain bacteria does not kill mice.
- Killed S strain does not kill mice.
- Killed S strain mixed with living R strain kills mice.

Dead mouse blood contains living S strain.
In Griffith's experiment why does the dead mouse contain living S strain when only dead S strain was injected? Theorize what may be happening.

Griffith's conclusion: Living R strain absorbs a chemical left from the dead S strain. This chemical transforms the living bacteria into the deadly S strain.

What we know now: Bacteria is capable of transformation. This is when DNA is taken in from the environment and incorporated into the bacteria's DNA. In this case the gene that produces the deadly toxin is absorbed.

1 Which strain of *S. pneumoniae* was virulent?
- A R strain
- B S strain
- C both
- D neither

2 What does Griffith's experiment illustrate?
- A Bacteria can transfer genes via sex pili
- B Phages increase the genetic variation of bacteria
- C DNA is the genetic material of cells
- D Bacteria can absorb genetic information from their environment

Closing in on the Genetic Material

Avery used a test tube assay. This is when a scientist compares differences in test tubes after treating each differently. The benefit is that you can discover more specific reactions. This approach will lead to more information than dead or living mice can provide.

Closing in on the Genetic Material

After Griffith's experiment most scientists believed that the chemical transforming bacteria was a protein, not a nucleic acid.

In the early 1940s experiments performed by Oswald T. Avery and his colleagues at the Rockefeller Institute for Medical Research challenged that assumption.

Closing in on the Genetic Material

First he heat-killed the S strain bacteria and mixed it with detergent. This caused the bacterial cells to break apart. Their membranes lysed and spilled out the cell's contents.

The upper portion of the test tube, the *lysate*, contains less dense materials like proteins, enzymes, and nucleic acids.
Closing in on the Genetic Material

The precipitate contained the large organelles and proteins of the cell. Avery isolated the lysate to use because it contained smaller molecules that were more likely to be the genetic material.

To be sure he took the lysate and mixed it with R strain to see if it would transform the bacteria to S strain and it worked.

It is easy to tell the difference from R and S because they look different when grown on a petri dish. (R for rough edge; S for smooth edge).

Next Avery put in an enzyme that digests proteins into the lysate and did the same experiment. What do you suspect is the result and why?

What could you do to confirm this result? In other words what would be another way to treat the lysate that would give usable data?
Closing in on the Genetic Material

Avery and his team devised a technique that used alcohol to isolate and purify nucleic acids from solution.

In a later experiment they mixed the purified nucleic acid from S strain with R strain bacteria. What is the expected result?

3. Avery's work retested Griffith's hypothesis using a test tube assay. What was the purpose of Avery's experiment?

- A. To test the validity of Griffith results
- B. To determine the macromolecule responsible for genetic information
- C. To determine the accuracy of a modern technique
- D. To illustrate that the R and S strains were two separate species of bacteria

The Definitive Proof

The Hershey–Chase experiments were a series of experiments conducted in 1952 by Alfred Hershey and Martha Chase that confirmed DNA was the genetic material.

By this time many new discoveries allowed these scientists to go beyond what others had been able to discover about nucleic acids.

Most importantly, intense research on viruses at the time expanded the knowledge of these tiny particles. Hershey and Chase concentrated on bacteriophages. These viruses that infected and killed bacteria were known to only be composed of 2 things: proteins and DNA

Secondly, a lot was being learned about radioactivity. Since they could not see the viruses, Hershey and Chase used a novel approach that took advantage of a new technique called radioactive labeling. This allowed them to track different parts of the virus by looking for radiation.

A geiger counter can find and measure radioactive particles

Hershey and Chase began by creating 2 kinds of radioactive virus using a labeling technique. Below is how they made virus A.
The Definitive Proof

Since proteins, not DNA, need sulfur to be constructed only the proteins in these new phages are radioactive.

The Definitive Proof

The procedure is repeated to make virus B with a change in the radioactive material.

Viruses grow via lytic cycle. When they make DNA they must use the radioactive phosphorus.

The Definitive Proof

Since DNA, not proteins, need phosphorus to be constructed only the DNA in these new phages are radioactive.

The Definitive Proof

Virus A is mixed with bacteria. If the bacteria were sampled and isolated, would it be radioactive? Apply what you know about the lysogenic cycle and genes to hypothesize.

The Definitive Proof

Virus B is mixed with bacteria. If the bacteria were sampled and isolated, would it be radioactive? Apply what you know about the lysogenic cycle and genes to hypothesize.

Summary of Hershey Chase experiments:
4 Hershey and Chase utilized phages in their research because phages...

- A have easily controlled reproductive cycles
- B are inexpensive to maintain
- C contain only nucleic acids and proteins
- D have a higher mutation rate than living organisms

5 After viral infection, radioactive phosphorus was found in the bacterial cells. This result occurred because...

- A Expression of the viral DNA incorporated radioactive phosphorus into the proteins
- B Only viral DNA was inserted into the host cell
- C The viral DNA was incorporated in the host DNA during the lysogenic cycle
- D DNA and proteins require phosphorus while only proteins require sulfur

6 Which of the following illustrates the sequence of events leading to the confirmation of nucleic acids as the genetic material of the cell?

- A Hershey-Chase experiments, Avery experiments, Miescher's discovery, Griffith experiments
- B Griffith experiments, Miescher's discovery, Avery experiments, Hershey-Chase experiments
- C Miescher's discovery, Griffith experiments, Avery experiments, Hershey-Chase experiments
- D Miescher's discovery, Avery experiments, Hershey-Chase experiments, Griffith experiments

---

The Double Helix

Working at King's College London in 1951, Rosalind Franklin and Maurice Wilkins produced x-ray diffraction images of DNA to try to discover its shape and understand its mechanisms.

The Double Helix

X-ray diffraction shoots subatomic particles into a substance. The collision of these particles with those of the molecules in the substance cause them to diffract, or ricochet, at specific angles. This gives insight to the structure of the substance.
The Double Helix

Wilkins shared this now famous “photo 51” (that was prepared by Franklin) with his colleagues James Watson and Francis Crick.

Watson and Crick deciphered from the photo that DNA was a double helix. They began to build models of the structure so they could speculate on how DNA can:

1) Self replicate
2) Code for all the traits of living things

RNA and DNA

RNA and DNA are the 2 nucleic acids necessary for living organisms.
Uracil is a nitrogenous base in RNA but not DNA.
Thymine is a nitrogenous base in DNA but not RNA.
RNA is single stranded and can fold into many shapes.
DNA is double stranded and can only be a double helix.

DNA

DNA is an informational molecule encoding the genetic instructions used in the development and functioning of all known living organisms.
This diagram highlights the major chemical features.
Notice that guanine and cytosine form 3 hydrogen bonds, while adenine and thymine for only 2 hydrogen bonds.

The two strands run in opposite directions to each other and are therefore \textbf{anti-parallel}, one backbone being 3’ (three prime) and the other 5’ (five prime). This refers to the direction the 3rd and 5th carbon on the sugar molecule is facing.
RNA makes the molecular machinery necessary for the function of DNA. It plays a major role in the replication of DNA and the reading of the information stored in DNA.

Three-dimensional representation of the small ribosomal subunit. RNA is in brown, protein in blue. The active site is in the middle (red). This molecule reads the genetic code.

7. Which of the following is not true regarding DNA and RNA?
   - A. Ribose has one more hydroxyl group than DNA
   - B. Both RNA and DNA backbones consist of a sugar-phosphate chain
   - C. Only DNA bases form hydrogen bonds
   - D. DNA forms only one shape. RNA forms many shapes

8. Which of the following is true regarding uracil?
   - A. It is only found in DNA
   - B. It forms 2 hydrogen bonds with adenine
   - C. It is more stable than thymine
   - D. It is only found outside the nucleus of eukaryotic cells

A chromosome is an organized structure of DNA and protein found in cells. It is a single piece of coiled DNA containing many genes.

Chromosomes vary widely between different organisms. The DNA molecule may be circular or linear, and can be composed of 100,000 to over 3,750,000,000 nucleotides in a long chain.

Typically, eukaryotic cells have large linear chromosomes and prokaryotic cells have smaller circular chromosomes.

This is a chromosomal map of a bacterium, H. orenii.

Like all bacteria, this circular DNA molecule contains all the genes that are needed to make the entire organism.

This particular bacterial genome is made of ~2,500,000 nucleotides. Each different color in the outer circle represents another gene.
Eukaryotic Chromosomes

In eukaryotes, nuclear chromosomes are packaged by proteins into a condensed structure called chromatin. This allows the very long DNA molecules to fit into the cell nucleus. Chromosomes are the essential unit for cellular division and must be replicated, divided, and passed successfully to their daughter cells to ensure the genetic diversity and survival of offspring.

Eukaryotic Chromosomes

Since eukaryotes are larger and more complex, eukaryotic chromosomes are much larger and require more complex methods for storage of their numerous genes. Special proteins called histones and scaffolds pack the DNA strand into tight coils.

Eukaryotic Chromosomal Structure

This diagram represents a eukaryotic chromosome after replication has occurred.

1. Chromatid – one of the two identical copies of the chromosome.
2. Centromere – the point where the two chromatids touch, and where the microtubules attach during cell division.
3. Short arm.
4. Long arm.
Chromosomes are often represented as genetic maps that show the loci of genes. This is a representation of human chromosome 7.

Each band represents a gene or a group of genes that code for a phenotype of the human. Humans have 23 pairs of chromosomes in each of their cells that contain multiple copies of ~40,000 genes.

Chromosomes can be diploid, 2 versions of each chromosome, or haploid, 1 version of each chromosome.

A karyotype is a photograph of the actual chromosomes of an individual human. A nucleus is isolated and the chromosomes are removed and arranged. They can be used to learn about possible chromosomal abnormalities.
Human Karyotype
What can you learn about this individual from their chromosomes?

Most likely a normal female.

Human Karyotype
What can you learn about this individual from their chromosomes?

Abnormal number of sex chromosomes. Klinefelter syndrome (XXY).

Human Karyotype
What can you learn about this individual from their chromosomes?


DNA Replication

The functions of a cell are determined by its DNA.

Cells have to reproduce many times. In complex organisms, trillions of copies are made from one original cell.

But when cells reproduce, they must replicate (or copy) their DNA.

The structure of DNA reveals how trillions of copies of the DNA in one of your cells can be made, and be almost exactly the same each time.

Watson & Crick

When Watson and Crick published the structure of DNA in a short article in 1953 they stated:

"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."

The fact that there are two DNA strands that are mirror images of one another suggested how copies could be made of each DNA sequence.
Each molecule of DNA is made of a template strand and a new strand. The template is used to make the new strand. The template strand is also known as the parent strand since it came from the original DNA molecule. The new strand is also known as the daughter strand.

The template strands of the DNA molecule separate and the new strands are made on the inside.

Nucleotides can only be added to the -OH end (3'), not the 5' so all new strands are made in the 5' - 3' direction.

DNA nucleotide monomers are made ahead of time and stored in the cell. DNA polymerase is the enzyme responsible for adding each new nucleotide to the growing strand.

The result of this process is 2 new DNA molecules each having an old template strand and new strand. This is called semi-conservative because it "conserves" some of the old DNA in each copy.

The 3' end of a DNA strand has a phosphate at the end.

- True
- False
12. Why does a DNA strand only "grow" in the 5' to 3' direction?

A. because DNA can only add nucleotides to the 3' end of the molecule
B. because DNA can only add nucleotides to the 5' end of the molecule
C. because mRNA can only read a DNA molecule from 5' to 3'
D. because mRNA can only read a DNA molecule from 3' to 5'

Replication Practice

3’ ATCGGTTAACCGTAAA 5’ template strand

5’ ______________________ 3’ new strand

What is the sequence of the new strand?

3’ GGTTACTAATCGAGCCCCT 5’ template strand

5’ ______________________ 3’ new strand

What is the sequence of the new strand?

13. If the parent DNA strand is 5' ATCGATACTAC 3', what will the daughter strand be?

A. 5' TAGCTATGATG 3'
B. 3' ATCGATACTAC 5'
C. 5' UAGCUAUGAUG 3'
D. 3' TAGCTATGATG 5'

The Molecular Process of Replication

A strand of DNA is replicated in segments. At intervals down the DNA molecule portions of the 2 strands separate creating replication bubbles. Either side of the replication bubble is know as a replication fork.

Replication Fork

DNA replication is a precise process that must minimize error. To do this cells use many enzymes in a complex process that uses template strands to create new DNA molecules.

Topoisomerase

Topoisomerase is an enzyme that controls winding of DNA during replication. A winding problem in DNA arises due to the intertwined nature of its double helical structure. During DNA replication, DNA becomes overwound ahead of a replication fork. This tension would eventually stop replication.
Helicase breaks the weak hydrogen bonds that hold nucleotides together opening up both strands to become templates for new strands. Helicase is the molecule that creates the replication fork.

Single stranded binding proteins ensure that the nucleotide pairs do not re-bind after helicase passes.

Since the strands are anti-parallel they are arranged in opposite directions. In order to replicate both strands in the same direction there are 2 different strategies, one for each template (leading and lagging).

The leading strand is simple since it runs 3' to 5'. DNA polymerase can follow behind helicase and simply copy the template as it is being exposed.

DNA polymerase can only read in the 3' to 5' direction. So on the lagging strand there has to be a way to make the new strand in reverse. It starts with an enzyme called primase that adds RNA nucleotides as a primer for DNA polymerase.

DNA polymerase can latch onto the RNA primers and begin to write fragment of the new strand. Since it is going away from the replication fork it only does a portion, then it jumps back in front of the portion it just did to start again.
The fragments formed by this process are called Okazaki fragments. When the RNA primers fall away from the strand gaps are left between the fragments that must be repaired.

DNA ligase finishes the job by filling in the gaps between the Okazaki fragments.

14 Hydrogen bonds hold base pairs together in DNA. Which enzyme severs these bonds during replication?

- A Ligase
- B Polymerase
- C Primase
- D Helicase

15 A mutation has caused a change in the shape of the topoisomerase enzyme. This would most likely affect...

- A the directionality of DNA synthesis
- B the formation of bonds between Okazaki fragments
- C the tension in the DNA double helix
- D the separation of the DNA strands

Polymerase Chain Reaction

PCR (Polymerase Chain Reaction) is a technique which uses the principles of DNA replication to amplify the amount of DNA available for testing and manipulation.

This reaction is carried out by a special machine that utilizes repeating cycles of heat, DNA polymerase, DNA primers and free nucleotides to build copies of the DNA fragment.

This technology enables small amounts of DNA to be turned into large amounts.
Polymerase Chain Reaction

1. DNA is heated to high temperature, the DNA strands denature, separating the double helix.
2. DNA is cooled, primers and polymerase in the mixture anneal to the DNA.
3. The temperature is increased slightly to increase the rate of replication reactions.

The cycle is repeated, doubling the amount of DNA each cycle.

17. A single DNA molecule is placed in a PCR machine. After 20 cycles, how many copies of DNA will be present?

18. Taq polymerase is typically used in polymerase chain reactions. This polymerase enzyme is found in thermophilic bacteria, *Thermus aquaticus*. What is the best explanation for the use of this enzyme?

A. Enzymes from thermophilic bacteria are stable at high temperatures.
B. Bacterial enzymes are easier to replicate than eukaryotic enzymes.
C. Ethical objections exist to the use of human macromolecules, such as DNA polymerase.
D. Only Taq polymerase is available commercially.

Transcription & Translation

Gene Expression

Gene expression is the molecular process of reading the order of nucleotides in a DNA molecule and making the coded product. This product is usually a protein but RNA is also coded for in genes. Gene expression occurs whenever a specific protein or RNA molecule is needed by the cell.

DNA to RNA to Protein

Expressing the information stored on a gene into a protein requires:

- Translating from the 4-letter language of DNA to RNA
- Then from the 4-letter language of RNA, to the 20-letter language of proteins (their amino acid sequence).
The mRNA "message" is read in 3-letter words called **codons**. Each codon codes for an amino acid or tells the process to stop.

There are 64 codons (4x4x4) but only 20 amino acids. So some codons code for the same amino acid.

### The Universal Genetic Code

- 61 of the codons code for an amino acid
- 3 of the remaining codons are "STOP" codons that do not code for an amino acid. They just signal that translation is over.
- 1 codon that codes for the amino acid "methionine" is also the "START" codon. Methionine is always the first amino acid in a protein.

This is called a "universal" code because **ALL LIFE uses the same genetic code**... from the smallest bacteria or virus to the largest animal or tree.

This tells us that **this code goes back billions of years**, in the first cell...or even before that.

If there were alternative codes that could work, they would have appeared in nature.

There are very minor alterations, but they are rare and insignificant in their effect.

The codon UAA specifies:

20 The codon GGG specifies:

21 Why is methionine the very first amino acid in all proteins?
The processes of replication, transcription and translation are so critical that they are called the **Central Dogma** of Biology.

A "Dogma" is a postulate; an idea; a philosophy.

It is "Central" because it is what life is based on.

---

**The Central Dogma**

The Central Dogma is a one way process. Changes in DNA affect mRNA and protein.

But changes in proteins or mRNA do not affect the DNA.

This will have important implications when we study genetics.

---

**The First Step - Transcription**

Transcription is the first step of gene expression, in which a particular segment of DNA is copied into RNA by the enzyme RNA polymerase.

---

**Gene Anatomy**

Transcription of genes is regulated by the cell. Genes are turned on and off in response to environmental signals.

---

**Control Region**

This control region is where transcription factors bind to the gene. When all the necessary factors are combined RNA polymerase can bind to the gene and initiate transcription.

---

**Transcription - Initiation**

Transcription initiated when RNA polymerase and cofactors bind to the promoter (a section of the control region). The RNA polymerase unwinds the DNA creating an initiation bubble. This is a space that grants RNA polymerase access to a single strand of the DNA molecule.
**Template vs. Non-Template Strands**

The RNA polymerase never attaches to the strand that actually contains the gene.

The strand with the genes is called the "non-template strand." This **IS NOT** the strand that is transcribed.

The other strand is the mirror image of the first, it carries the mirror image of the gene, not the gene itself. It is called the "template strand." This **IS** the strand where the RNA polymerase attaches.

---

**Transcription: DNA Strands**

This makes sense in that the RNA will be the mirror image of the DNA it is transcribed from. And the non-coding strand is the mirror image of the gene.

---

**Base Pairing**

Transcription is made possible by the fact that the different bases are attracted to one another in pairs based on the number of hydrogen bonds they can make.

RNA    DNA
A      bonds with T
U      bonds with A
G      bonds with C
C      bonds with G

Note: In DNA replication adenine paired with thymine, in DNA transcription uracil is now paired with adenine. Remember that RNA does not contain thymine as a nucleotide base.
RNA is made from the 5’ end to the 3’ end. Just like in DNA replication,

DNA ("template strand") 3’ TACGGCATT

RNA 5’ AUGCCGUAAU 3’

being made in 5’--------->3’ direction.

24 If the template strand of DNA is 5’ ATAGATACCATG 3’, which is the RNA strand produced from transcription

A 5’ UAUCUAUGGUAC 3’
B 5’ TATCTATGGTAC 3’
C 3’ UAUCUAUGGUAC 5’
D 3’ TATCTATGGTAC 5’

25 If the template strand of DNA is 5’ AAAGACACTATT 3’, which is the RNA strand produced from transcription

A 5’ UUUCUGUGAUAA 3’
B 5’ TTTCTGTGATAA 3’
C 3’ UUUCUGUGAUAA 5’
D 3’ TTTCTGTGATAA 5’

26 If the non-template strand of DNA is 3’ ACGATTACT 5’, which is the RNA strand produced through transcription

A 5’ TGCTAATGA 3’
B 3’ UGCUAAUGA 5’
C 3’ UGCUAAUGA 3’
D 3’ ACGAUUACU 5’

Transcription - Termination

RNA polymerase gets to a sequence on the DNA called a termination sequence. This sequence signals the RNA polymerase to STOP transcription.

RNA polymerase falls off the DNA. The new RNA strand separates from the DNA. The DNA recoils into a helix.

Controlling Gene Expression

Individual cells respond to environmental change by regulating their gene expression.

Prokaryotes and eukaryotes have evolved different mechanisms for regulating their gene expression.

Remember: one of the properties of life is "response to the environment"
Prokaryotic Gene Expression

Regulating Prokaryotic Gene Expression

The following are 2 examples of the regulation of gene expression in prokaryotes:

- Lac Operon
- Trp Operon

Operons: The Basic Concept

In prokaryotes, genes are often clustered into operons within the chromosome.

Operons consist of 3 parts...

- An **operator** - essentially an "on-off" switch
- A **promoter** - an area that attracts RNA polymerase
- The **genes** - which code for the protein needed by the cell

Repressors

An operon can be switched off by a protein called a **repressor**.

The repressor can be controlled through allosteric regulation with co-repressors and inducers.

- **Aco-repressor** is a small molecule that cooperates with a repressor to help switch an operon off.
- **An inducer** is small molecule that inhibits a repressor to help switch an operon on.

Inducible Operons

An **inducible operon** is one that is usually off; a molecule called an inducer inactivates the repressor and turns on transcription.

An example of an inducible operon is the **lac operon**, which contains genes coding for enzymes that break down lactose into glucose so the bacteria can use it for energy.

If no lactose is present then no enzyme needs to be made. The bacteria saves energy this way.

In this operon, the lactose molecule is the inducer.

Repressible Operons

A **repressible operon** is one that is usually on.

When a repressor binds to an operator, transcription is shut off.

The **trp operon** is a repressible operon. The trp operon codes for a number of genes responsible for the production of the amino acid tryptophan. If tryptophan is present in the environment, the trp operon is not used.

Tryptophan acts as the co-repressor in this operon.
27. The lac operon is an example of an _____.
   - A. inducible operon
   - B. repressible operon

28. In the presence of Trp _____.
   - A. the repressor is activated
   - B. the cell makes more Trp
   - C. the operon is turned on

29. In the presence of lactose _____.
   - A. the lac operon is turned off
   - B. the lac operon is turned on
   - C. the repressor becomes active

---

Gene Expression in Eukaryotes

Eukaryotes have much more complex chromosomes that require multiple levels of regulation including:
- "unpacking" of genes
- transcription factors
- RNA processing

---

Chromatin's Role in Gene Expression

When DNA is packed in chromatin it is not accessible to RNA polymerase so transcription cannot happen.

The main factor in the specialization of cells in multi-cellular organisms is what genes are "unpackaged" from the chromatin to be exposed to RNA polymerase.

---

Chromatin Modifying Enzymes

The genes that need to be expressed are unwound from histones by chromatin modifying enzymes in order to expose their nucleotide sequences.

Genes that are unnecessary to a particular cell will remain packed while the necessary ones are unpacked.
mRNA Processing

After transcription of eukaryotic DNA, the transcript is known as pre-mRNA. Enzymes in the nucleus modify pre-mRNA before the genetic messages are sent to the cytoplasm. This is known as mRNA processing.

During mRNA processing, both ends of the pre-mRNA are altered. Some interior sequences of pre-mRNA may be cut out, and other parts spliced together.

Alteration of mRNA Ends

The 5’ end of the pre-mRNA receives a molecule known as a nucleotide (or 5’) cap. This cap is a modified guanine molecule (the G in A, T, C, G).

The 3’ end of the pre-mRNA gets a poly-A tail. This tail is series of adenosine (A) nucleotides.

The modifications to the ends of the pre-mRNA have several functions:
- They facilitate the export of mRNA from the nucleus to the cytoplasm
- They protect mRNA from hydrolytic enzymes once it is in the cytoplasm
- They help ribosomes attach to the mRNA so they can be translated into a protein.

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 called exons (because they are eventually expressed), are usually translated into amino acid sequences.

RNA splicing removes introns and joins exons, creating an mRNA molecule with a continuous coding sequence.

mRNA Processing

This is an example of a pre-mRNA becoming a final transcript.
**Alternative RNA Splicing**

Some genes can code more than one kind of polypeptide, depending on which segments are treated as exons during RNA splicing.

**Alternative splicing** allows the number of different proteins an organism can produce to be much greater than its number of genes.

---

**DNA sequence**

AAATTTCGGGAAATTTCGGG

**Pre-mRNA**

(Cap)-UUUAAAAGGCCC-CUUAAAGGGCC-(Tail)

**Alternate splices**

(Cap)-UUU UUU AAA A-(Tail) OR (Cap)-GGC CGG GGC-(Tail)

**Resulting polypeptide (protein)**

Phe - Lys - Phe - Lys OR Gly - Pro - Gly

Alternate splicing can dramatically change the length and/or the sequence of the polypeptide chain that will be made.

---

30. The first step in eukaryotic gene expression is...

- [ ] A transcription
- [ ] B translation
- [ ] C RNA processing
- [ ] D unraveling the gene

31. Where does transcription occur in eukaryotic cells?

- [ ] A nucleus
- [ ] B nucleoid
- [ ] C cytoplasm
- [ ] D cell membrane

32. Once the DNA is unwound from the chromatin, which of the following is necessary to begin transcription?

- [ ] A RNA polymerase
- [ ] B ribosome
- [ ] C transcription factors
- [ ] D both A & C

33. A transcription unit that is 8,000 nucleotides long may use 1,200 nucleotides to make a protein consisting of 400 amino acids. This is best explained by the fact that

- [ ] A many introns are present in mRNA.
- [ ] B there is redundancy and ambiguity in the genetic code.
- [ ] C many nucleotides are needed to code for each amino acid.
- [ ] D many exons are present in mRNA.
A mutation in which of the following parts of a gene is likely to be most damaging to a cell?

- A intron
- B exon
- C both would be equally damaging.

Entrance into the Cytoplasm

After the finalized mRNA transcript is complete and correct, the pores in the nuclear envelope allow it to pass to the cytoplasm where it can be translated into proteins by ribosomes.

The nuclear pore is a protein structure that controls the traffic flow of the nucleus. Each nuclear pore is made up of hundreds of individual proteins that insure only mRNAs with proper caps and tails can make it to the cytoplasm.

Degradation of mRNA

Hydrolytic enzymes in the cytoplasm breakdown mRNA molecules. The length of time an mRNA survives in the cytoplasm relates to how much protein is made from it. Longer time in the cytoplasm means more translation by ribosomes.

The length of the poly-A tail is one of many factors that determines the time of survival in the cytoplasm. The longer the tail, the longer it's survival.

Summary of Gene Expression Regulation in Eukaryotes

- The gene must be unpacked from chromatin
- The right transcription factors must be present
- Cap and tail must be added to the mRNA
- Pre-mRNA must be edited (spliced)
- Nuclear pores allow passage to the cytoplasm
- mRNA comes into contact with a ribosome

Translation occurs

- Protein is used within the cell or exported to the environment

Translation

Translation is the process by which RNA strands are read to build proteins.

Translate means to convert something from one language to another, you can remember that the process of making protein from RNA is called translation because the "language" of nucleotides is being changed to the "language" of amino acids.

Three Types of RNA

Translation requires 3 types of RNA that are created using transcription.

1. mRNA or messenger RNA, carries the information for protein synthesis. This type of RNA is key to The Central Dogma.
2. rRNA or ribosomal RNA, is a catalyst for protein synthesis
3. tRNA or transfer RNA, helps in the assembly of amino acids during protein synthesis
The specific RNA that contains the protein's information from DNA is called Messenger RNA (mRNA); it carries the genetic message to ribosomes, where it is translated.

**Messenger RNA (mRNA)**

Ribosomal RNA (rRNA) and some additional proteins make up the ribosome. The ribosome includes two subunits: one small, and one large. During translation, the ribosome catalyzes the reaction that makes covalent bonds between amino acids, thus building the protein.

**Ribosomal RNA (rRNA)**

RNA, being single stranded, can fold in on itself. In tRNA, the RNA folds into a l-shape.

**Transfer RNA (tRNA)**

Transfer RNA (tRNA) carries amino acids to the ribosome so that the ribosome can covalently bond them together to form the protein.

35 What is the function of the ribosome?
- A to make an ionic bond between amino acids
- B to make a covalent/peptide bond between amino acids thus building the protein
- C to make hydrogen bonds
- D to make RNA

36 Why does tRNA fold into its specific shape?
- A The sequence and bonding of its amino acids
- B The sequence of and bonding of nucleotides
- C Its protein structure
- D A and B
- E A and C

37 What is the anticodon for GGU?
- A CCT
- B CCA
- C GGU
- D GGA
38. If the anticodon of a tRNA is AUG, what is attached to the amino acid attachment site of that tRNA molecule?

- A RNA
- B methionine
- C tyrosine
- D glycine

---

**Translation - An Overview**

All the pieces are ready to begin translation:
- a coded strand of mRNA
- a set of 20 amino acids
- ribosomes
- tRNA to match all the amino acids

---

**Proteins: Words**

Amino Acids :: Letters

The length and sequence of these amino acids allow all the proteins in the world to be created from only 20 amino acids.

This is very similar to how all the words can be created from only 26 letters in the alphabet.

---

**Translation - Initiation**

The small subunit of the ribosome attaches to the mRNA at the bottom of the start codon (at the 5' end).

Then the large subunit of the ribosome comes in over the top.

The result is that the mRNA is "sandwiched" between the mRNA at the start codon (and the second codon as well!)
**Translation - Initiation**

The tRNAs, hydrogen bonded to their specific amino acids, surround the ribosome.

As the leading edge of the mRNA, with the starting code AUG, is exposed in the A site, the tRNA with the code UAC enters the site and hydrogen bonds with it, carrying methionine into the ribosome.

The methionine is removed from the tRNA and stays in the ribosome to be bonded with the next amino acid. The tRNA leaves the ribosome so another tRNA can enter.

Each tRNA will carry the appropriate amino acid into the ribosome to be bonded in the proper sequence, since each tRNA anticoding site matches the coding site on the mRNA, which is located at the A site of the ribosome.

Because each tRNA has an anticoding sequence it complimentary base pairs with the codon on the mRNA.

---

**Translation - Elongation**

The 2nd tRNA with its amino acid is delivered into the A-site in the ribosome.

The ribosome catalyzes a covalent bond between the amino acids.

---

**Translation - Elongation**

The ribosome moves the mRNA using chemical energy.

The tRNA that was in the A-site moves to the P-site and the tRNA that was in the P-site separates from its amino acid.

Notice the protein emerging from the P-site!
Elongation continues by adding one amino acid after another.

Each amino acid is delivered to the A-site by its matching tRNA.

The ribosome makes a peptide bond between the 2 amino acids in the P and A sites.

The ribosome reaches a STOP codon. Remember that STOP codons do not code for amino acids. This signals the end of translation.

The protein is complete.

The 2 subunits (large and small) separate from each other.

UAA is 1 of the 3 possible STOP codons.

What is the first event of translation?

- A the tRNA comes in
- B the small subunit of the ribosome and the 1st tRNA brings in Methionine to the start codon
- C elongation happens
- D the large subunit of the ribosome comes in

What is the function of the ribosome in translation?

- A it makes a peptide/covalent bond using the energy from translocation
- B it makes hydrogen bonds between the codons
- C it makes covalent/peptide bonds between the codons
- D none of the above

What does termination in translation involve?

- A translocation of the ribosome
- B the ribosome gets to a stop codon and the small and large subunits of the ribosome separate
- C RNA polymerase falls off the DNA
- D a tRNA brings in an amino acid
Recombinant DNA

Biotechnology

Understanding the structure, replication, and expression of genes has allowed scientists to create biotechnologies which improve human lives.

One of the first procedures to use DNA to successfully was recombinant DNA technology.

In this procedure genes from one organism are spliced into the genome of another. Since all organisms use the same genetic code the cells that contained the recombined DNA will produce the protein encoded by the new gene.

Diabetes

Diabetes is a disease in which a person has high blood sugar either because the pancreas does not produce enough insulin, or because cells do not respond to the insulin that is produced.

Untreated, diabetes causes many problems. Most severe is damage to the kidneys resulting from the increased solute of the blood for extended periods.

Early Treatment of Diabetes

At first doctors treated by injecting bovine insulin harvested from cows blood. This had some effect but the protein is not exactly the same as the human insulin protein. The symptoms would eventually overcome the patient.
Recombinant DNA

In the late 1970s scientists started to look for a way to make human insulin in a laboratory. Their efforts produced the first product ever made using genes from multiple organisms.

They recombined fragments of DNA from humans with the bacterial chromosome of *E. coli*. The new bacteria was then able to produce human insulin.

Humulin

The result was a new product called Humulin. The first human hormone to ever be produced by another organism.

Recombinant DNA Technology

DNA pieces can be recombined to make unique, man made sequences. There are 7 main steps.

Recombinant DNA Technology

Step 1: Find the piece of DNA in the genome, the gene of interest.

Today this step is done by computers attached to robotic DNA sequencers that fragment, analyze and find a gene based on user input.

Recombinant DNA Technology

Step 2: “Cut” the gene of interest from the genome

This is made possible by restriction enzymes. In nature these enzymes are used by bacteria as weapons against invading viruses.

For example: EcoRI is a restriction enzyme that makes a staggered cut when it reads the sequence GAATTC. The staggered ends are called sticky ends because they leave a few unpaired nucleotides that will easily stick to another piece of DNA with the same sticky end.
Recombinant DNA Technology

**Step 3: Isolate the gene of interest**

If we look at the insulin gene again we can see that the sequence between the two EcoRI cut sites has a unique length.

- EcoRI cut site
- Insulin Gene (gene of interest)
- EcoRI cut site

DNA fragment

- 5,000 nucleotides (bp)
- 10,000 nucleotides (bp)
- 15,000 nucleotides (bp)
- to end of fragment

So in this digest there are DNA fragments that are 5k, 10k, and 15k nucleotides long. The gene of interest here is the 5k piece.

Recombinant DNA Technology

**Step 4: Make more of the gene of interest (amplification)**

Once the gene of interest is isolated in the gel, the band that contains the gene can be cut from the gel, but this is a very small sample. More copies of the can be made using polymerase chain reaction (PCR).

Recombinant DNA Technology

**Step 5: “Paste” the gene of interest into the host’s DNA**

Sticking to the insulin example, the technique utilized to get the insulin gene into the E. Coli bacteria involved using a plasmid, the small circular pieces of DNA that bacteria use to trade pieces of genetic information.

- A plasmid with an EcoRI cut site is “digested” using the same restriction enzyme that was used to cut out the insulin gene.

Recombinant DNA Technology

**Step 6: Put the recombined piece of DNA into the host organism**

Now that the gene of interest is in a plasmid, it can be mixed with bacterial cells and be taken up into the bacterial chromosome.

Remember, all living things use the universal genetic code. The bacterial cells will read the newly acquired gene, transcribe it into mRNA and its ribosomes will translate the mRNA into a protein.

The bacterial cells will reproduce and express the gene. Each time a recombinant bacterial cell divides by binary fission it will make a new copy of the gene.
Recombinant DNA Technology

Step 7: Collect the protein product

The protein can be extracted from bacterial cultures using various techniques. It can then be delivered to the patient.

Currently there is no cure for diabetes, but with advancements in insulin therapy patients can now avoid many of the life threatening complication.