Genetics
& Cell biology

◉ Sheet ○ Slides

Number

9

Done by:

Sarah Jamos

Corrected by:

Leena Al-Nsour

Doctor

Mamoun

Transcription

Some definitions:

Chromosome (chromo= color , some=body): the whole genetic material

= DNA + protein

DNA = Nitrogen base + sugar + phosphate

Gene: the entire nucleic acid sequence that is necessary for the synthesis of an RNA (not for protein synthesis specifically because tRNA, rRNA and others don't get translated to proteins, but they still have a synthesizing gene). It includes coding and non-coding regions that regulate the activity of the gene itself. = Specialized regions within the DNA that get transcribed.

Cistron: A genetic unit that encodes a polypeptide. Two types:

- *Monocistronic gene*: Cistron that encodes a *single* mRNA, and this mRNA makes a *single* polypeptide. This polypeptide can make a protein or can be cut down to smaller polypeptides and each can function as a hormone. Most ***eukaryotic*** genes are monocistronic

- Polycistronic gene: Cistron that encodes a single mRNA, and this single mRNA can make *multiple and different* polypeptides. Most prokaryotic genes are polycistronic. In bacteria, genes that encode enzymes, which are involved in related functions, are located next to each other. Example: Trp operon[1] for the synthesis of tryptophan in bacteria. Trp operon synthesizes different enzymes and proteins that all participate in tryptophan synthesis pathway.

Gene Amplification:

The mechanisms mentioned above are physiological, but they can turn into pathological by gene amplification. Gene amplification is an increase in copy number of a restricted region of a chromosome increasing the quantity of DNA in these regions. Normally, gene amplification happens to increase the amount of a certain protein, but accidently a single gene on a certain chromosome gets amplified a lot causing a puffy appearance of that chromosomes causing cancer. Cancer cells get advantage of this mechanism to

[1] *Operons are clustered polycistrinic genes that make one mRNA. This mRNA produces several polypeptides and enzymes each with a different function, but they all participate in the same pathway.*

escape resistance from methotrexate whereby the target gene, dihydrofolate reductase, is amplified. Another example: Breast cancer is divided into 4 types, one of these types is *HER2 enriched*; higher HER2 receptors on the cell surface (by gene amplification) cause higher chance of cancer cell proliferation.

Introns vs. exons

- Only in Eukaryotic cells.
- Exons: The region of the gene within the gene that encodes the polypeptides
- Introns: Non-coding regions that separate different exons from each other.

When transcription happens, we end up with **Pre-mRNA** (also called primary transcript, Primary mRNA). This Pre-mRNA contains introns and exons.
Pre-mRNA splicing: taking out the introns from the mRNA leaving the exons, producing a mature mRNA (mature transcript or transcript) so that exons can get translated.

The surprise is that introns are NOT JUNK regions since they can code for a functional mRNA molecules.

Importance of introns:

1) They can code for functional RNA molecules such as nucleolar RNA that function in ribosomal processing as well as microRNA that regulate the expression of proteins.

2) They can contain regulatory sequences that control gene expression.

3) The exon-intron arrangement is important for *genetic recombination* (*genetic material* exchange to reduce the chance of error) and *alternative splicing*:

A. *Genetic recombination*:

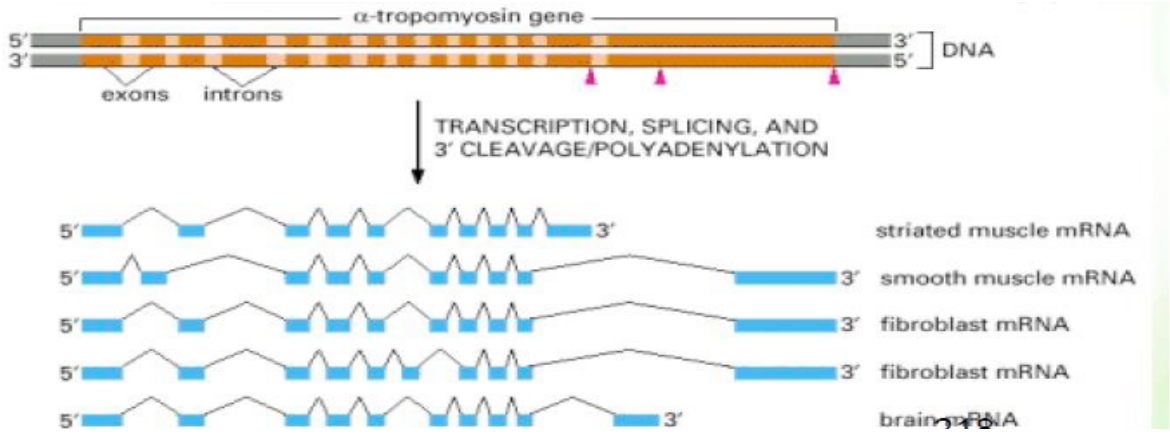
Usually introns are large regions compared to exons. When maternal and paternal chromosomes are lined up in the M phase, the genetic material exchange happens (genetic recombination). This recombination might happen within the introns.

Example: In one case, we take exon 1 and 2 from the maternal chromosome, and take exon 3 ,4 ..etc from the paternal chromosome and combine these exons together randomly. Another case, we take the exon

1,2,3,4 from the maternal chromosome and we take the introns from the paternal chromosome. This caused shuffling of the recombination leading to genetic diversity; every one has different and specific inherited exons from both the mother and the father (as a fingerprint), yet still look relatively alike.

Genetic recombination can be inaccurate (loss or addition of a single/multiple bases accidently)
A problem in recombining introns ---> not a big problem
A problem in recombining exons ---> FRAMESHIFT mutation
So its better that the genetic recombination happen in non coding regions (introns, transposone, VNTR..)

B. *Alternative splicing*:
Introns allows for alternative splicing. Alternative splicing is splicing the transcripts in different ways to produce different mRNAs and different proteins. Also known as *protein isoforms*
Example 1: **Tropomyosin** is a binding protein that contains a regulatory region and binding region that binds to actin. This tropomyosin differs from one cell type to another.
1. Tropomyosin in striated muscles: have exon1, exon3, exon5, exon6, exon7, and exon9.
2. Tropomyosin in smooth muscles: have exon1, exon2, exon5, exon6, and exon7.
To conclude: Different exons in different types of cells produce isoforms of tropomyosin.
All these isoforms have the same function (binding to actin) but *with different regulation.*



Example 2: **UDP-Glucuronosyl Transferase Gene**
UDP-Glucuronosyl transferase is an enzyme that transfers glucuronic acid to drug molecules to make them hydrophilic and increase their excretion.

UDP-Glucuronosyl transferase (UGT1A) has two domains:
-**Catalytic domain** (have the same exons in all cells)
-**Recognition domain** (have different exons between cells).

* All isoforms of UGT1A have exon2, exon3, and exon4 BUT differ in exon1; thus, exon 1 is responsible for the recognition domain.
* Having different enzyme isoform means identifying different drugs in different tissues (like cytochrome isoforms).

General mechanism of transcription

- Transcription is the process of making RNA from DNA.
- All we need is: DNA template[2] and RNA polymerase.

Q: which strand of DNA is used to produce the RNA? Depends on the presence of a region called the **promoter** (the starting point), but most importantly that RNA synthesis is going to occur in the direction of 5' to 3' in which the DNA template is read from 3' to 5' (same as DNA synthesis).

- The transcribed DNA strand is called: template, anti-sense or **negative strand**
- The other antiparallel DNA strand is called: coding strand[3], sense or **positive strand**
- The enzyme: RNA polymerase
- Substrates: ATP, UTP, CTP, GTP
- Energy source: by the hydrolysis of 2 phosphates from the substrate itself.
- The RNA chain produced by transcription is also known as the **transcript**.
- RNA polymerases catalyze the formation of the phosphodiester bonds between two nucleotides.

---

[2] *One of the two strands of the DNA double helix acts as a template for the synthesis of an RNA molecule.*

[3] *Called the coding strand because it has the same sequence as the synthesized RNA (Except the U an T base differences).*

- **Polysomes**: as the first RNA polymerase is synthesizing the RNA from the DNA, another RNA polymerase also starts synthesizing RNA and so on (Synthesis of multiple RNA from the same gene).This allows the simultaneous synthesis of many RNA chains from the same gene forming structures known as ***polysomes***.

DNA replication vs. transcription:

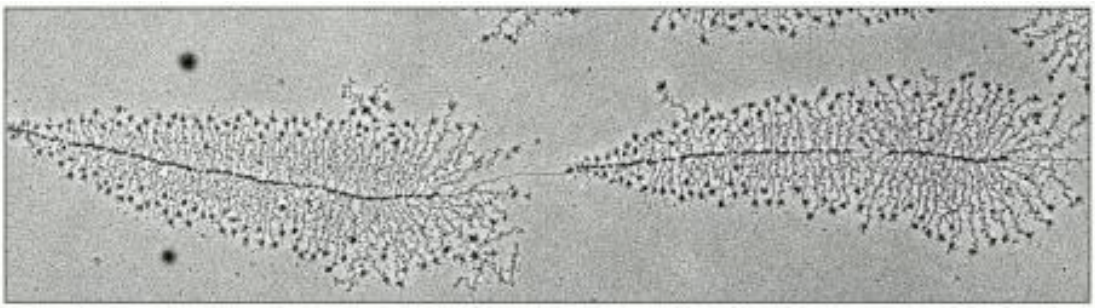
- The RNA strand does not remain hydrogen-bonded to the DNA template strand. (Only part of it is bonded to the DNA template)
- RNA polymerase read the A in DNA and inserts U in the growing chain of RNA rather than T.
- RNA molecules are much shorter than DNA molecules.
- Unlike DNA, RNA does not store genetic information in cells. Therefore ➔ it's ok if RNA polymerase made a mistake because we're not changing the DNA.

DNA polymerase vs. RNA polymerase:

- The most important and the only differences the doctor mentioned is: *DNA polymerase is more accurate than RNA polymerase. Meaning:* The consequences of an error in RNA transcription are much less significant than that in DNA replication. Although RNA polymerases are not as accurate as the DNA polymerases, they have a modest proofreading mechanism.

Other differences:

- RNA polymerase catalyzes the linkage of ribonucleotides, not deoxyribonucleotides.
- Unlike DNA polymerases, RNA polymerases can start an RNA chain without a primer.
- RNA polymerases make about one mistake for every 104 nucleotides.



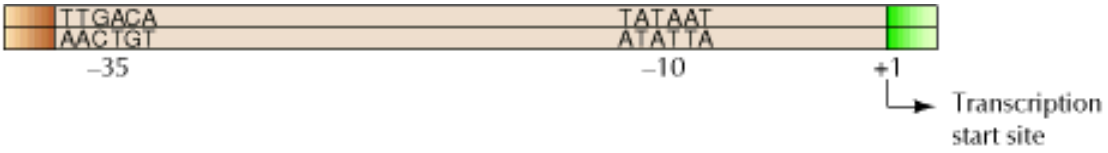
Note: In this gene the starting point of transcription is at the left side (since it has the shorter RNAs).

Transcription in Prokaryotes:

RNA polymerase:

- Core enzyme subunits: 2α - β - β' - σ
- σ (sigma) subunit is important for the binding to -10/-35 region of the promoter, but it is not required for the basic catalytic activity of the enzyme.
- In the absence of σ, RNA polymerase binds to DNA with low affinity and nonspecifically.
- The role of σ is to identify the correct sites for transcription initiation and direct the polymerase to promoters by binding specifically to both the -35 and -10 sequences. Thus, mutations in these regions will affect the sigma function; it will keep binding then sliding off the DNA.

Q: what if σ subunit itself got deleted/mutated? The RNA polymerase will stay functional and still can bind to the promoter but with a different efficiency (it will slide on the DNA)



Important regions in the gene:

- **-1,-2,-3 …etc.** (before the transcription initiation site)
- **(-10)/(-35) elements** (the promoter) they are two consensus sequences, which are the binding sites for the σ subunit ➔Located 10/35 base pairs **UPSTREAM** of the transcription site.
- They are important because Consensus[4] sequences have strong effects on promoter function. Thus, if a gene has a sequence that is different than of the consensus sequence (the *optimal sequence*) due to mutations in these regions, the expression of that gene is less efficient.
- RNA polymerase generally binds to promoters over approximately a 60-base-pair region, extending from -40 to +20.
- **+1 position** (after the promoter) also called the transcription start site/ transcription initiation site, is the first nucleotide read of the gene after the promoter ➔ **DOWNSTREAM.**
- **Open reading frame (ORF)**: The row that get transcribed and read by the RNA polymerase.
- Note: the promoter is not a part of the ORF

[4] A sequence of DNA having similar structure and function in different organisms.

Mechanism of Transcription

1) ***Initiation****:*

- RNA polymerase binds to the DNA. If it finds a -10/-35 regions that are free, then the σ subunit come and bind to it.

- The interaction between the DNA promoter and RNA polymerase is known as the closed promoter complex. In the closed promoter complex the DNA is still in the form of a double helix.

- Opened promoter complex: RNA polymerase has separated the two DNA strands in order to expose the template, and RNA synthesis has begun.

- It doesn't need a primer. All it need is an open promoter complex and the RNA polymerase will start putting nucleotides.

- After reading the first 10 or so nucleotides by the RNA polymerase, the σ subunit is released and binds to another RNA polymerase.

2) ***Elongation****:*

As the RNA polymerase is walking along the DNA:

- It opens up the DNA ahead of it "**unwinds DNA**".
- It Elongates the RNA.
- It "**rewinds the DNA**" in the back where it finished RNA synthesis.

The RNA is synthesized inside the bubble formed between the winded and unwinded regions of DNA.

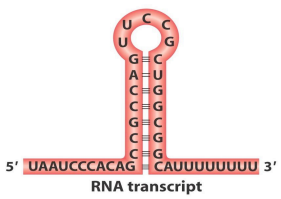
3) ***Termination*** (happens due to signals):

RNA synthesis continues until the polymerase encounters a termination signal where the RNA is released from the polymerase, and the enzyme dissociates from its DNA template.

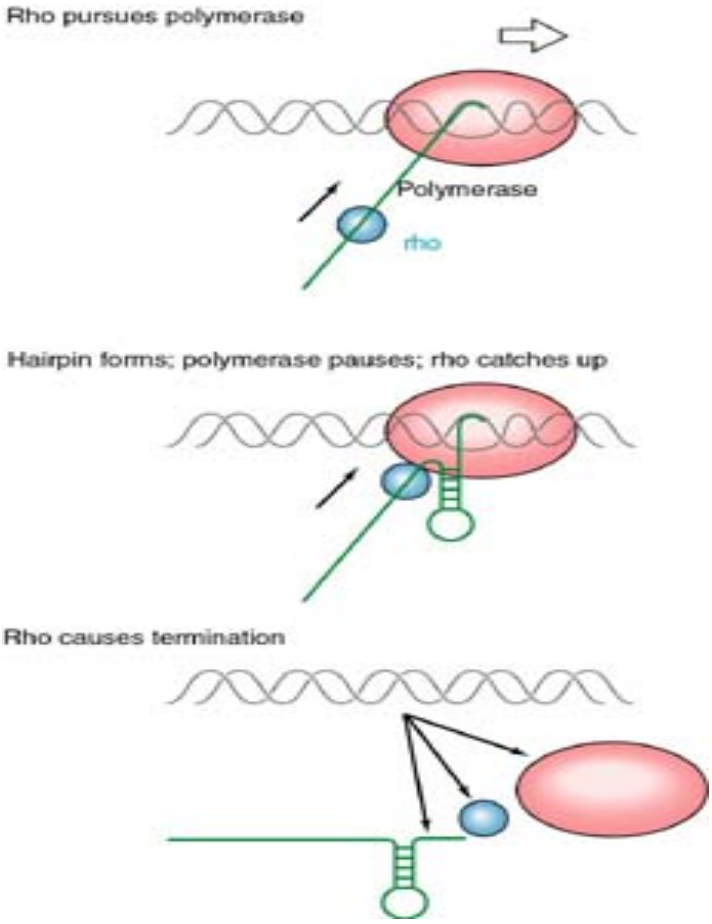
In Bacteria there are two types of signals:

1. Rho Independent signal: **GC's rich palindromic sequence** followed by an AUUUUU sequence. This sequence forms H-bonds between the base pairs of the same strand➔ forming a stable stem loop structure (hair pin like structure).

When the RNA polymerase reaches the stem loop structure, it gets kicked out, and it falls off.



2. **Rho protein dependent signal:**
Rho protein binds to the mRNA and starts walking on mRNA trying to follow the RNA polymerase in order to reach it and kick it out.
When the RNA polymerase reaches the stem loop structure the Rho protein will catch up with the RNA polymerase and kick it out; and termination occur.
Note: the stem loop structure is **less stable** here than the previous mechanism. (Meaning: this loop ALONE can slow down the RNA polymerase, but it is not sufficient for the transcription to terminate, it requires the Rho protein for the termination to occur).



Sorry for any mistakes!
Best wishes