University of Jordan
Faculty of Medicine

GENETICS & Molecular Biology

Number: 13

Done By: Amal Al-Akily.

Corrected By: Toqa Alhumaidi.

Subject: Protein domains & transcription factors, Epigenomics, regulation of gene expression & techniques of studying it.

Doctor: Ma'moun Ahram.

Price:

Date:

-This sheet was written according to section-3 record

-The slides of this lecture are from SLIDE 3 (143-169)

The last thing we talked about is that regulatory proteins have at least 2 **independent** domains :

1) DNA-binding domain
2) Activation domain

-that's a domain that is used to bind to a certain sequence in the DNA and another domain that would interact with the other proteins thereby facilitating the assembly of a transcription complex on the promoter; so at least 2 domains are independent

(slide143) DNA-Binding Domains

These domains have different structures:

-Zinc finger domains (i.e,Steroid receptors)

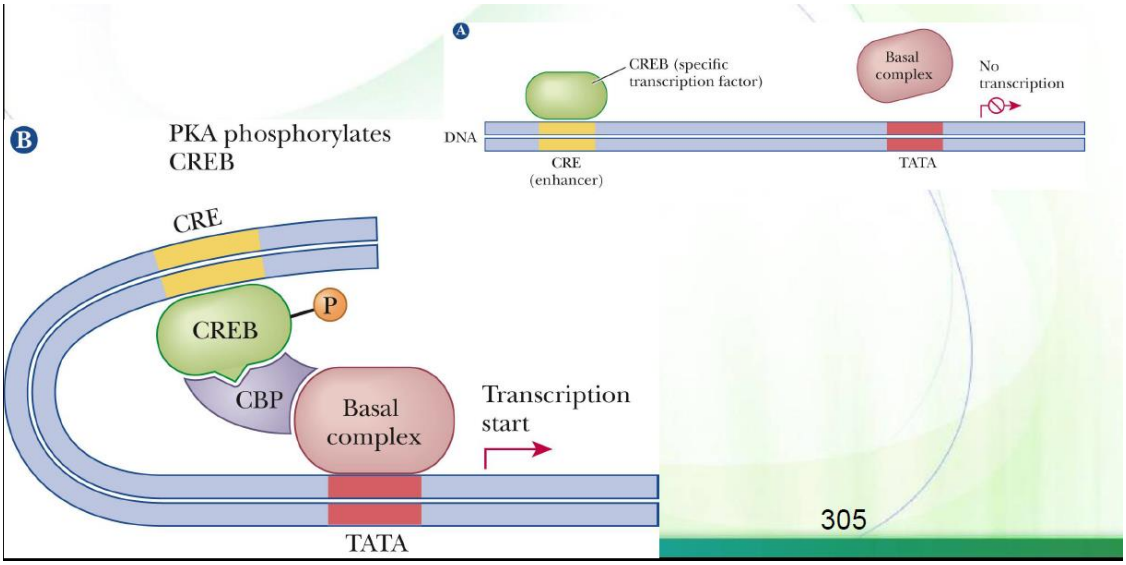
-Helix-turn-helix motif

-Leucine zipper (i.e,CREB)

-Helix-loop-helix

●as an example is a protein known as **CREB:**

CREB is a **C**yclicAMP-**R**esponse **E**lement **B**inding protein



▪so there is a certain sequence known as a response element and this protein (CREB) binds to it.

▪but it is not active, it doesn't do anything until it gets **phosphorylated.**

▪once it gets phosphorylated, the feature of DNA "the DNA looping" touches these proteins via a CREB (binding protein) and get activation of transcription.

▪CREB is phosphorylated by a protein kinase known as **protien kinase A**

▪protein kinase A phosphorylates CREB → CREB is now active → and you have activation of transcription.

-one of the things that protein kinase A does is that it phosphorylates a lot of proteins.

♦<u>BUT How does it get activated itself ?!</u>

By signal transduction (we will study that in endocrine system).

-an example: FIGHT OR FLIGHT response

what do you need when you want to run (flight) or when you want to fight ?

- you need **Energy**

  >> you need ATP (comes from metabolism →from the breakdown of glucose and lipids).

- Then :

→ the hormone **epinephrine** binds to its receptor.

→ the epinephrine receptor activates a series of proteins that would lead to activation of **protein kinase A.**

→ **protein kinase A** phosphorylates **CREB** .

→ **CREB** would then bind to cyclicAMP response element (**CRE**) of **different genes.**

>>so there are multiple genes that have this response element (CRE) and CREB would bind to **ALL** of these enhancer regions activating a transcription of these genes.

-One of these genes that are activated in this process are genes that would express the **enzymes** necessary for the breakdown of glucose and lipid metabolism.

-so it is a **single shot** [a single activation of one protein]

(activation of protein kinase A would activate so many enzymes ,not only gene expression, but also activation of the enzyme itself induces the metabolism or breaking down of glucose and lipids all at the same time)

The same thing happens with other genetic systems where you have activation of one protein that would induce expression of **multiple** genes and they all participate in a similar mechanism or similar phenomenon or for a common purpose.

so that is the story of CREB .

-An example of mutations that happen in transcription factors:

Huntington's Disease

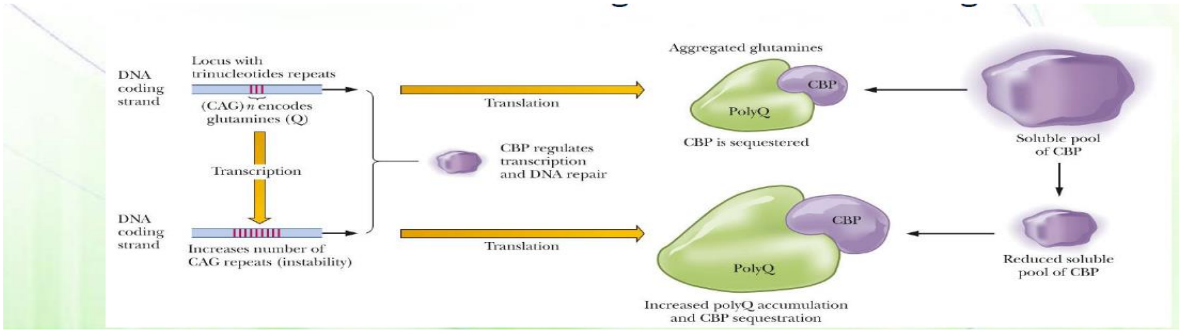
It is a degenerative disease (a progressive disease), meaning it gets worse with time as a person gets older, the person becomes more and more sick.

Usually people with this disease die in their 30s,early 40s due to paralysis and loss of control.

Huntington's disease is caused by a mutation in an transcription factor called **Huntingtin** (it is -tin rather than -ton).

Huntingtin : Is a transcription factor that has CAG repeats ,(certain number of these repeats).

-In Huntington's disease, what happens in this transcription factor is that the number of repeats **increases** with age.



-<u>what is the function (purpose) of the CAG repeat?</u>

The main function is binding of the **CBP (**CREB binding protein)

With **more** CAG repeats , you have binding with **more** CBP protein meaning that CREB would not have enough CBP to induce transcription of these genes because they are sequestered (hijacked) by Huntingtin.

- *Again,* what happens is the following:

Huntingtin protein (a transcriptional factor) has CAG repeats that bind to the CREB binding protein "CBP" → more CAG repeats→ interaction with more CBP (Huntingtin binds to a lot of CBP) → CREB doesn't have enough CBP to bind to (interact with) to induce expression → cell death and degeneration of cells.

Eukaryotic repressors

Eukaryotic repressors have different mechanisms of repressing transcription:

►they are composed of both a DNA binding domain and a protein binding domain:

so they can bind to the response element (to the DNA) and they can touch the basal complex (the transcriptional complex; polymerases & transcription factors),BUT it does not push the complex forward ,rather, it **holds** it in its place.

∴ Eukaryotic repressors inhibit the transcription by preventing the polymerase from moving forward.

►Repressors can have just a DNA binding domain so they bind to the DNA and they act as competitive inhibitors, they prevent the activators from binding to the DNA:
so the response element is occupied, so there is no binding of the activator and induction or activation stimulation of the basal complex

►Repressors can have just an activation domain meaning that they prevent the activator from binding to the basal complex (preventing binding to transcription factors).

SO the activator binds to the enhancer (the response element), the DNA loops, BUT it cannot touch the other transcriptional factors (it cannot stimulate them) because they are all have been interacting with the repressor →there is an inhibition or block of protein-protein interaction.

---

steroid receptors

-steroid receptors are examples of transcription factors that are also receptors.

■The way they function :

□steroids are small lipophilic hormone molecules and because of that they can diffuse through the plasma membrane easily.

□once they get into the cytoplasm, they bind to their receptor (i.e, Estrogen receptor ,Androgen receptor, Progesterone receptor, Glucocorticoid receptor ,…etc).

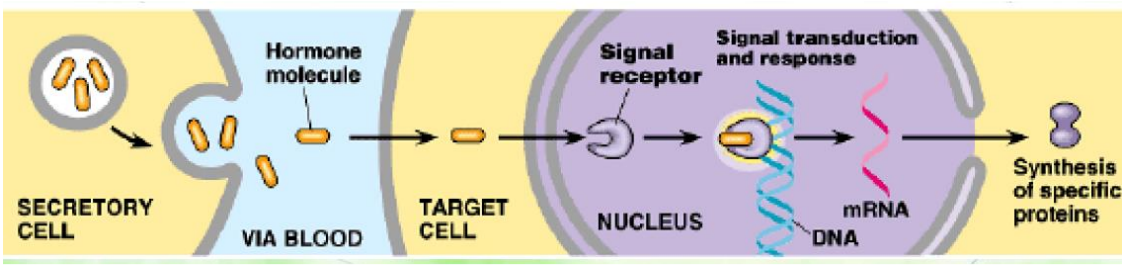
□then, this complex ,as a transcription factor, binds to a certain **response** element (certain binding sites on the DNA) activating or inhibiting gene expression.

■These receptors have at least three domains :

-a **DNA binding** domain.

-an **activation** domain for protein interacting domain.

-a **hormone** binding domain.



●so you have a protein that is large,

—that hormone comes in and binds to the receptor at the **hormone binding domain.**

—then the complex goes to the nucleus and binds to the DNA via the **DNA binding domain.**

—then it interacts with other transcriptional factors activating or inhibiting transcription using the **activation domain**.

▪▪▪note: remember that these domains are independent of each other.

so if we do genetic engineering as the following:

we took the Androgen binding domain then we placed it with the DNA binding domain of the Estrogen receptor and the activation domain of the Estrogen receptor we'll end up with:

-a protein that is engineered

-it has the hormone binding domain of the Androgen receptor.
-it has the DNA binding domain and the activation domain of the Estrogen receptor.

HOW WOULD THIS PROTEIN FUNCTION?
something like this will be in the exam *

-**Androgen** is what binds to the receptor but the protein will regulate genes that are normally regulated by **estrogen**.

▪▪▪note: Estrogen now can't recognize this engineered protein because it doesn't have its binding domain anymore

---

Transcriptional Regulatory Network

regulation of transcription is **not a single shot** mechanism, rather, it is a **network.**

(it is NOT that you have an activated transcriptional factor that binds to a certain genes to have a response)

-What happens is that you have activation of transcription of certain genes; these genes themselves can also be transcription factors! So what they do once they are produced, they would bind to regulatory regions of other genes inducing their expression. So you have induction of another set of genes, these can also be transcription factors that can regulate a third set of genes. So it is **a network**.

-Sometimes this **slow** response can take hours and days!.

-At the cellular level there are two types of responses in terms of speed:

1-a quick (fast) response.

2-a slow response.

■A quick response:

It means we are activating / inactivating proteins within seconds.
(i.e, phosphorylation & dephosphorylating)

■A slow response:

It involves gene expression because it takes time to induce :

-gene expression & production of certain proteins.

- processing of the RNA.

-processing of the proteins.

and then you have the function. Sometimes it is not the primary response (it is not the proteins that are produced from the first time), Sometimes it is the secondary response, Sometimes it can be the tertiary response and so on.

Epigenomics

All we talked about it is something that would be modified or induced by changes in the DNA sequence (change A for T and so on).

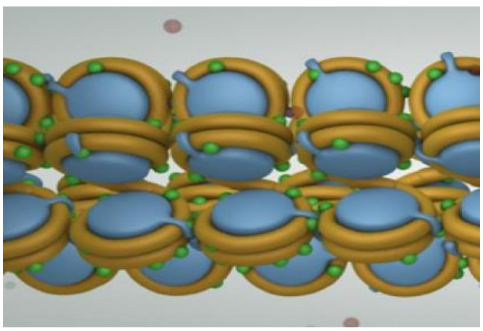
♦Now, we'll talk about another form of molecular regulation :

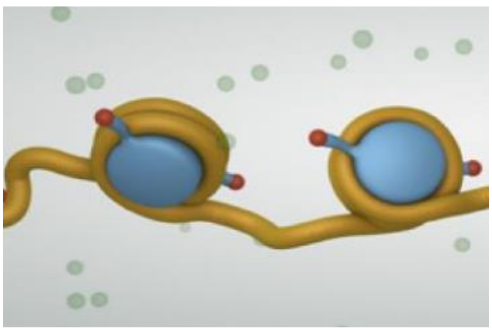
-There is a **higher** level of regulation which we call ***epigenomics*** or ***epigenetics*** *(epi means above / higher),which* means we can regulate gene expression without changing the sequence of the DNA among individuals.

♦ There are different ways of regulating gene expression.

*[Remember when we talked about **chromatin** in the first lectures; chromatin is Histones (proteins) and DNA. And the purpose of Histones is to rap and pack DNA].*

→So the DNA can be really packed or **loosely** packed.





● Transcription factor:

→when it comes in, it cannot find any DNA sequence that it can bind to because the DNA is packed (it is hidden / covered).
→ so in order to activate gene expression, the DNA piece that contains the gene must be **loose** - DNA fragment must be **accessible** for the transcription factor-.

-We can regulate gene expression by changing the structure of the DNA >> by packing it or making it unpacked (loose). This is one form of epigenetics.

So if you look at these identical twins.

They have the same sequence of DNA, but <u>Are they really identical at the molecular level ??</u>

-No, they are not.

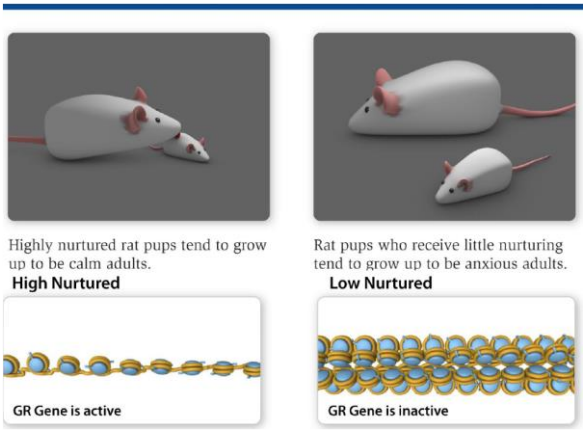
<u>Why?</u>

-let's say that this person on the right takes care of himself; he drinks a lot of milk, he takes vitamins, he plays sports, he exercises ,...etc, but the one on left smokes, drinks alcohol, drinks a lot of soda, no milk, no vitamins, no sport …etc .And that would change the **structure** of DNA →they are not identical at all in that regard.

This is an experiment:

▪▪▪*note:-The doctor said the story behind this experiment he told to sec. 1&2 are not true and he apologized for that.*

-we are not required to know the details of this experiment (just the concept).



▪ The story *(the details that you won't be asked about in the exam)* A scientist was doing research on mice and she noticed that mothers that nurtured their pups and took care of them -, their pups were **relaxed**
but pups that were not been taking care of - the mother kept a distance between her and the pups , they were agitated, nervous, anxious -they were **not relaxed** at all. She found out that there is no genetic difference rather it was **epigenetic** difference; more licking the pup had, the more relaxed the DNA was for the glucocorticoid receptor ,so the receptor was really active, it was actively transcribed making pups relaxed. On the other hand, pups that were not licked, their DNA for the glucocorticoid receptor was packed and the gene was not transcribed causing these pups to be agitated & be nervous.

How are chromosomal structures altered?

There are 4 mechanisms:

1)♦***Specialized proteins that bind to the DNA and they can change the structure of the chromatin*** **activating or inactivating certain genes.**

2)♦***Acetylation of Histones.***

3)♦***Histone remodeling factors that can change the structure of the chromosome***.

▪▪▪note: Histones are **highly** positively charged proteins because they have a lot of **Lysine** and **Arginine** residues.

● we are more concerned with the Lysine content:
Lysines have amino group ($-NH_3^+$) in their R group ,this group can be acetylated (the addition of acetyl group to the R group of amino acid).

How would that affect the interaction between Histones and DNA?

It **weakens** the interaction; because the interaction is based on **electrostatic** interaction between the positively charged Lysines and the negatively charged Phosphates of DNA .If you remove these positive charges, it means the interaction would become **weaker** and the DNA becomes **loose.**

- Acetylation is controlled by two enzymes :

  1) Histone Acetyl transferases:
  Enzymes that add the acetyl group and stimulate gene expression.

  2) Histone deacetylases
  It removes the acetyl group and strengthening the interaction between Histones and DNA so packing the DNA further.

[so here we have gene expression **activation** by Acetyl Transferases and **inactivation** by Deacetylases]

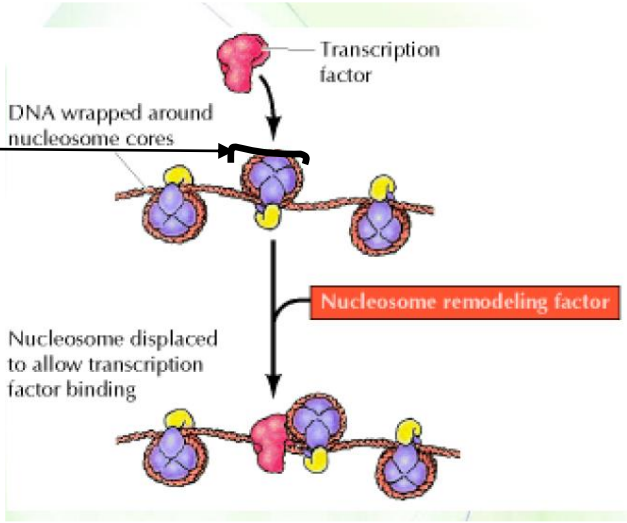
The gene expression is regulated by a system; you have the transcription factor **TFIID** *[remember:this is the factor that **bends** the DNA and sends a signal to other proteins : "here is a gene that needs to be transcribed "]* This transcription factor has a component to it and this component is an Acetyl Transferese ;so it bends the DNA & sends a signal to other proteins and **loosens** the DNA.

-some gene expression inhibitors have Deacetylases.

The nucleosome remodeling factors

They are not really removing Histones rather what they push them (they change the location of Histones / remodel the DNA).

Example: A transcription factor is looking for this DNA sequence but because it's rapped it can't bind to it, and here the *nucleosome remodeling factor* pushes the Histones aside a bit making this sequence **accessible** for the transcription factor.



The forth mechanism of epigenetics or epigenomics control which is

4) ♦ *DNA methylation*

▪Methylation of **Cytosine**. Some genes in their promoter region, they have what we call CPG islands (clusters of C's and G's, i.e, CGCGCGCG…).

The purpose of having a lot of C's is that they can be methylated, If they are methylated, it means the gene expression would be **reduced** & blocked.

so methylation of cytosine is associated with reduction of transcription

-so methylation inhibits transcription by certain proteins such as **MeCP2**

That is also associated with human **Deacetylase**.

-so this proteins have **dual** mechanisms in inhibiting gene expression ; *methylation of cytosine* & *Histone deacetylation*.

(It is like making sure that this gene is not expressed WHAT SO EVER).

---

significance of DNA methylation

In terms of significance of DNA methylation biologically or at the molecular level ,we have two phenomena:

1) X chromosome inactivation

■Females have two X chromosomes while males have one.

-Dose this mean that females would produce **double** the amount of proteins as males?

NO, it should be the same. HOW??

-During development, one of the X chromosomes becomes inactivated and it is a random process (either this or that but not the both). The females, in every one of her cells, there is only one x chromosome that is active.

So that now the number of x chromosomes would be equal between males and females.

How the x chromosome is inactivated?
By **methylating** it (the **whole** X chromosome is methylated producing what is known as the Barr body; the strong X chromosome).

The other phenomena is known as:

2) **Genomic imprinting**

■ We get two genes from both parents, and for certain genes, **only** the maternal gene must be expressed but not the paternal gene
for other genes, **only** the paternal gene must be expressed but not the maternal gene. Otherwise, a certain disease would be caused.
-the mechanism of genomic imprinting is not will understood.

How do different cells appear ? What makes a nerve cell a nerve cell when all our cells have the same genome (22000 genes)?

►It depends on **which** genes are expressed ,not all genes are expressed in every cell ,so let's say you have 6000 genes expressed in nerve cells vs. 7000 genes in muscle cells.
Some of the genes are common like for example Actin (Actin is needed in all cells),but there are certain genes that are needed only for nerve cells or **only** for muscle cells or only for skin cells and so on.

So it is the regulation of gene expression that determines the lineage, the nature, the development & the differentiation of the cell into a certain type.

This type of experimental field is known as **TRANSCRIPTOMICS**

-A transcript is mRNA & Transcriptomics is the study of genes that are expressed or the study of mRNA more specifically

- Genomics :the study of genes as whole.

▪▪▪note: it is hard to tell if the epigenetics can be transmitted from parents to children because the environment plays a really important role in regulating that.

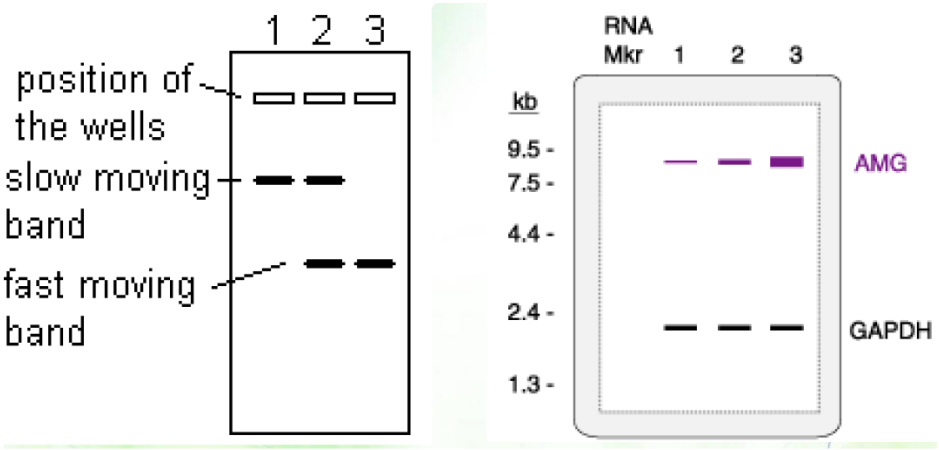
How can we study gene expression ?

♦There is a technique known as **Northern blotting.**

[remember Sothern blotting was developed by someone named Edwin Sothern and then scientists said "you know what, this is a lovely technique why don't we do the same thing for RNA].

- **Northern blotting** in general is separation of DNA in a gel according to size then transferring the RNA to a membrane (according to size) and then adding a probe (the probe would hybridize to a certain RNA molecule that is complementary to the probe).

♦There are 3 pieces of information that you can get from these gels here



1) Size:
size of RNA molecule (not the size of the gene).
-you are looking at the size of the RNA molecule (the part of the gene that is transcribed).

2) Which cells express the gene:
so we can tell if cells we have in different samples express the gene or not.

3) Density:
The amount of RNA , how much of the gene is expressed.
So you compare sample 3 to sample 1 and you say : it expresses **more** → the gene is more **active.**

-BUT can we say that in sample 3 we have more RNA because the gene is more active?

Nooo ,because the amount of RNA can also be affected by **stability** of RNA.

-So you can **increase** the amount of the RNA not by making more of it ,but more by reducing how much of it is degraded by freezing the stability of the RNA molecule.

▪▪▪*note: The doctor said: some of this will be in the exam; he will bring us a Northern blot and ask us questions about it and you have to think about it; (think about **ternative splicing**, **polyadenylation** , **alternative polyadenylation**).*

****Think about these things ,relate that to Northern blotting****

---

another technique **in situ hybridization**

*(-in situ means in place).

It is a similar technique:

▪we are looking at the expression of a certain genes by adding a probe **except** that we are looking at gene expression **in its place** (this technique is for this question :**where** is it expressed ?)

▪If we have a tissue section that has a lot of cells and we take all of the RNA from these cells, we can say there is more expression of this gene right here because I have more RNA.

▪But what are the cells that express the gene ?
We won't know because we got the RNA from different types of cells; epithelial, fibroblast, muscles, endothelial,... etc. (It is a tissue section). But if we add a probe we'll know.

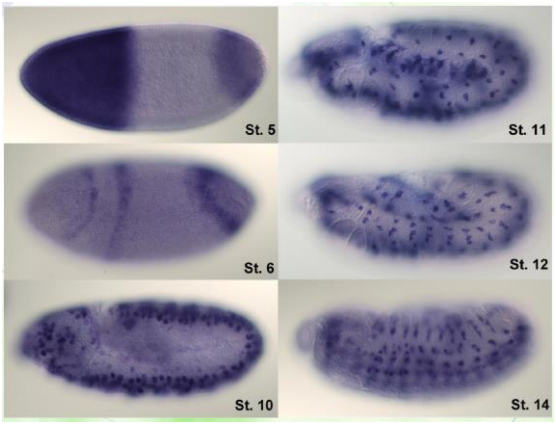
-so by doing in situ hybridization –in which we add a probe to a tissue section and we see **where** the signal is coming out from , and from which cells -.

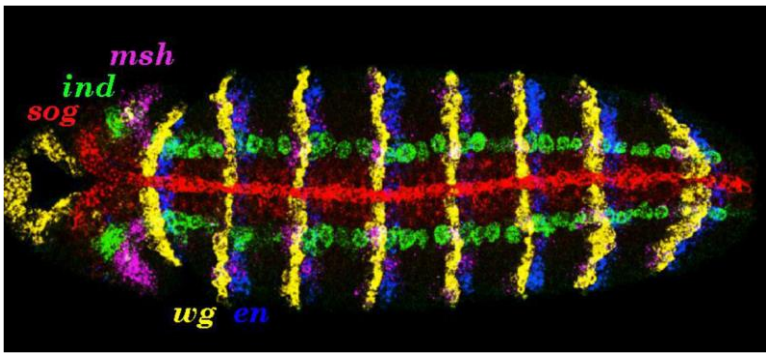
■<u>This is the embryo of Drosophila (fruit fly)</u> and we are looking at the expression of a certain genes in it:

- here we have sporadic expression everywhere.

-there is expression in the periphery>> so that tells me that this gene is important for the development of these cells in this stage (cells of the periphery).

so we are looking at where the gene is expressed in situ hybridization.





▪regarding the pic above ↑:

-the gene labeled in red is expressed as a whole line.

-the gene labeled in yellow is expressed in segments; so it is important for segmentation

▪so that tells us something about differentiation, symmetry,left and right, top and bottom ,head and legs,…etc. It tells us something about DEVELOPMENT.

-The end

-I'm sorry for any mistakes

لك شيء في هذا العالم فقم ")