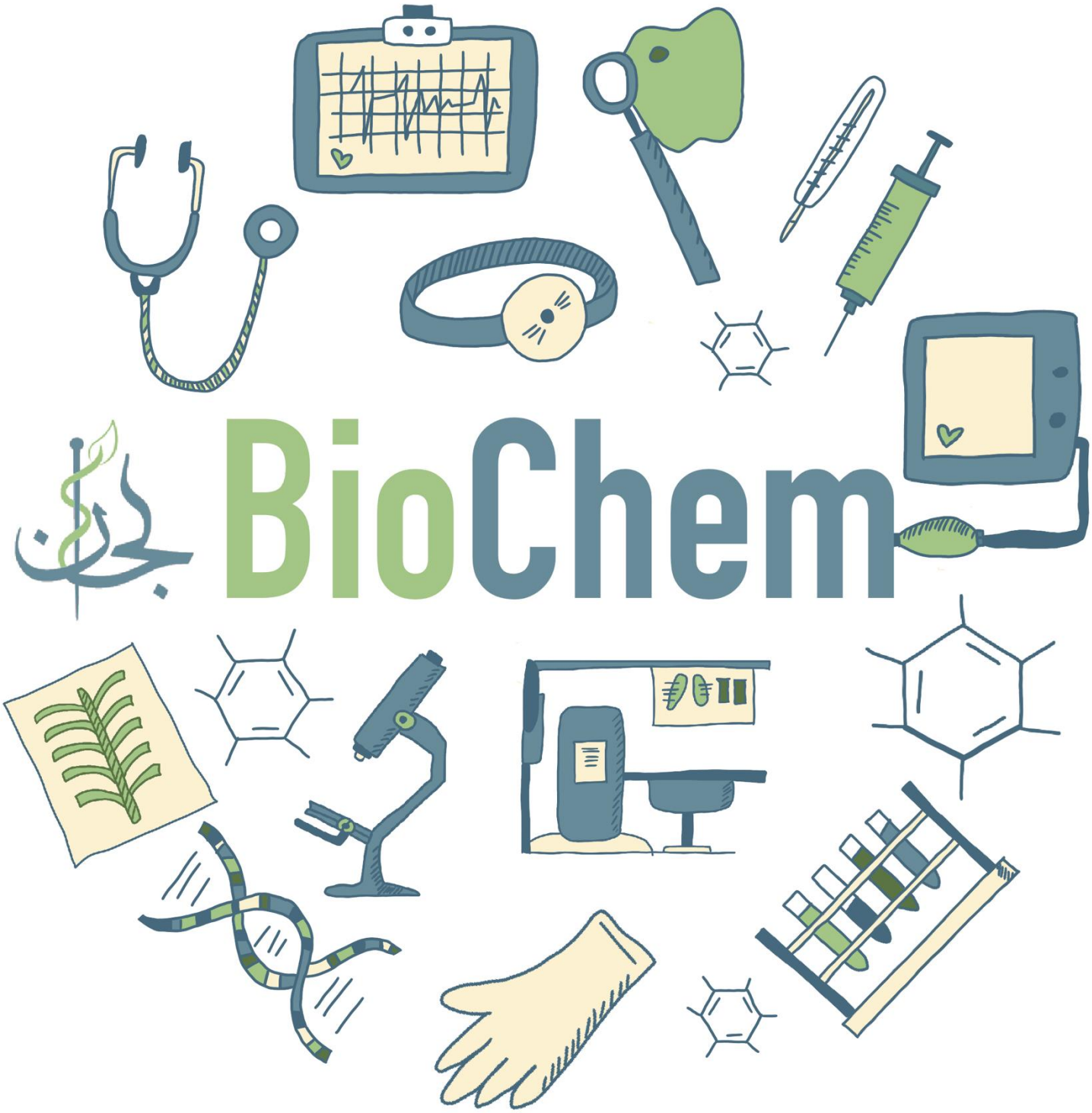


BioChem

Writer: Saif Aburumman & Laila Elayyan

Science: Albara' Al-azab

Grammar: Nada Odeh

Doctor: Prof. Mamoun Ahram

Recombinant DNA-based molecular techniques

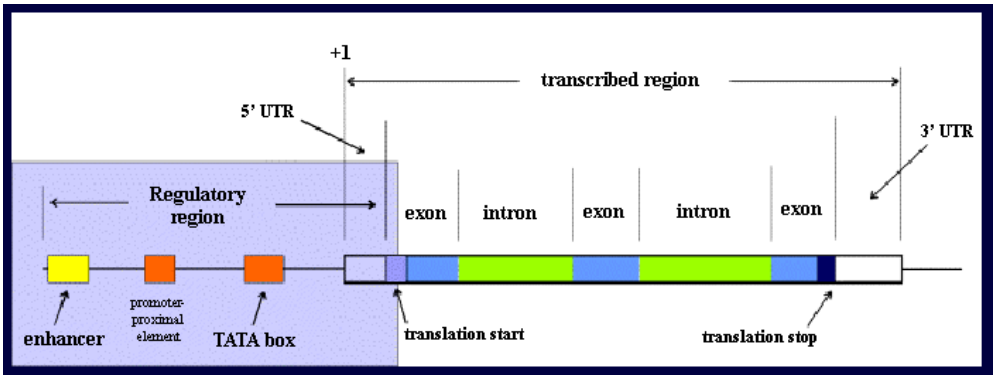
In this sheet we will cover three techniques, one involves analysis of regulatory sequences of transcription. We'll talk about techniques that allow us to analyze protein-protein interactions as well. All the techniques are based on the production of recombinant DNA.

Recombinant DNA: a DNA molecule that is made of two or more foreign independent DNA sequences. Using genetic engineering and restriction endonucleases we can integrate different pieces of DNA creating recombinant DNA.

Transcriptional regulatory sequences

There are three different types of these sequences:

1. **The core promoter sequence:** includes TATA box, which is the region where the RNA polymerase binds in order to start transcription at the transcription start site.
2. **Promoter proximal element (PPE):** a sequence that binds to a regulatory protein in order to coordinate and harmonize the expression of different genes having the same PPE, these genes usually participate in a similar function or pathway. PPE can either activate or inhibit transcription.
3. **Enhancer/silencer:** regions that can be located faraway from where the gene is, upstream, downstream, or even within the gene itself like in introns for example.



- The DNA loops (folds) which makes interactions between proteins that bind to the enhancer/silencer with proteins that bind to the core promoter or PPE possible.

Enhancers ⟶ activate transcription.

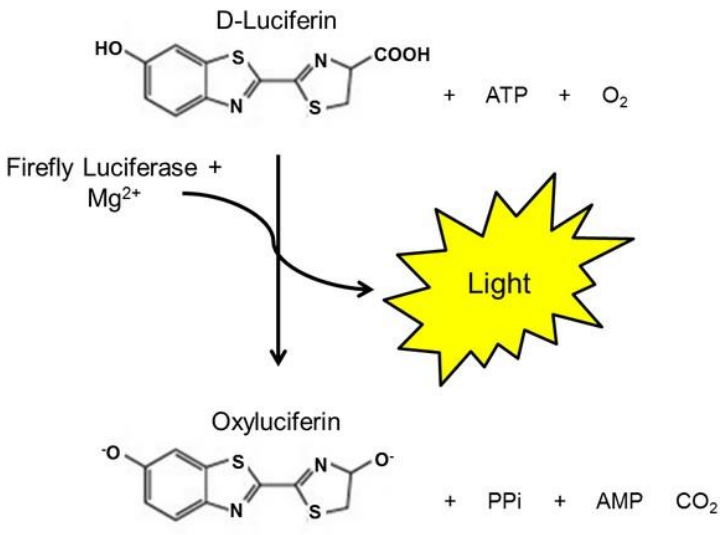
Silencers ⟶ inhibit (block) transcription.

Firefly luciferase

Let's say we have a gene and we want to know what regions are important for the transcription of this gene, we use recombinant DNA technology here and take advantage of an enzyme called firefly luciferase.

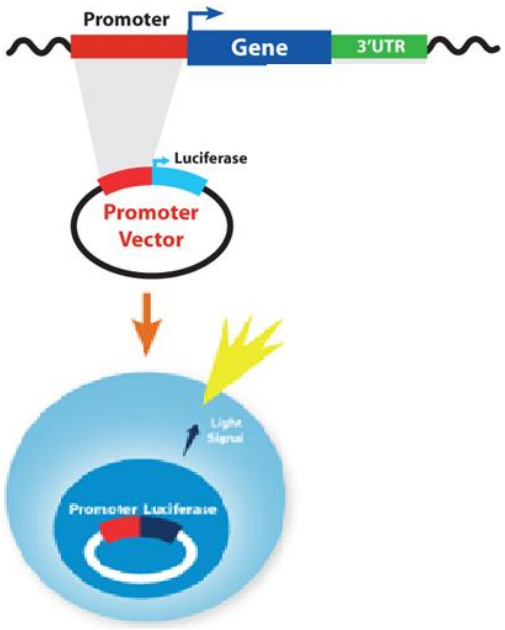
Firefly: an insect that has a tail that fluoresces (lights up).

Scientists found out that the tail lights up because there is an enzyme called firefly luciferase, this enzyme converts a molecule known as **D-Luciferin** into something else called **oxyluciferin** producing light.



Scientists have taken advantage of the luciferase enzyme, how?

Scientists take the luciferase gene and place it in a plasmid. However, this gene is not regulated by its own promoter, it's controlled by the promoter of the gene of interest. So, we are doing recombinant DNA technology and genetic engineering whereby we have the luciferase gene transcription controlled by the promoter that we want to study.



The luciferase gene is known as a "reporter gene" because it gives information about how active transcription is. (This activity is measured by the amount of light produced).

Mechanism	1. Adding the luciferase gene to the plasmid creating a recombinant DNA. 2. Inserting the plasmid into human cells. (human cells can accept plasmids)
Results	• Very active promoter → high number of luciferases produced → high amount of light. • Slightly active promoter → low number of luciferases produced → low amount of light. • Inactive promoter → no luciferases produced → no light.

EXAMPLE:

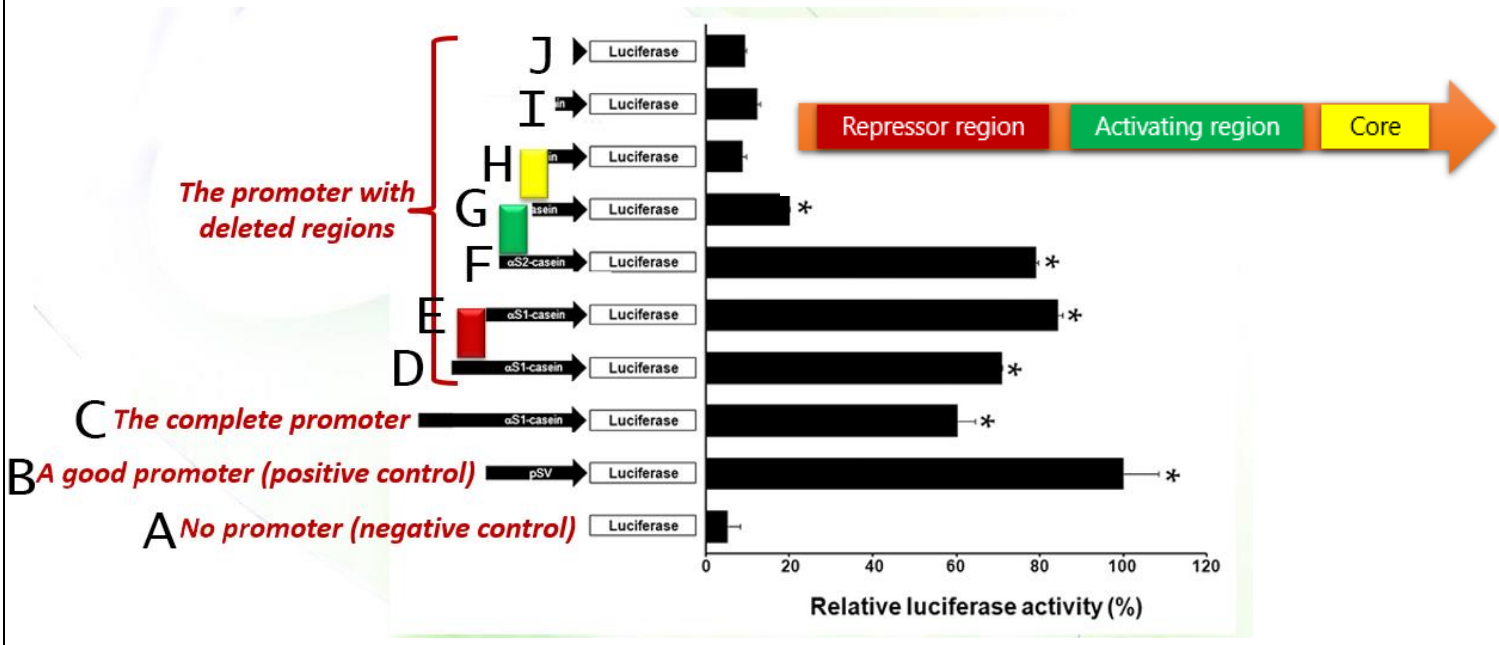
We have this promoter region containing:

1. **The core promoter:** important for the binding of RNA polymerase.
2. **The activating region (PPE):** an element that binds to an activating protein (an activator) and induces transcription.
3. **The repressor region:** an element that binds to an inhibitor which prevents transcription from resuming.

Two important terms:

1. Negative control: a sample that I know it should not work.
2. Positive control: a sample that I know it should work.
   - ❖ In any experiment I must have a negative control which should give me zero results and a positive control which should give me good results.

The figure below shows an experiment with its results, let's read the figure (starting from the bottom & going up).



A. Luciferase gene but without a promoter, meaning that there should be no production of the enzyme luciferase. However, we have a small signal, this signal is due to the leakage of transcription, some RNA polymerases can leak and bind to the region transcribing the luciferase protein.

10:00

B. We use a good promoter, a promoter that should work and give good results (positive control).
C. We use the region of interest (the whole promoter region of the gene that we want to analyze). Notice that we have some signal, this means that the promoter is functioning (there is transcription & as a result there is translation as well).

> ❖ Now we want to know the regions that are important for transcription, so we start chopping off the promoter (deleting sequences of the promoter from left to right).

D. We remove part of the repressor. Notice a little increase in the production of luciferase.
E. We remove the whole repressor. Notice a large increase in the production of luciferase, this increase tell us that there is a repressor region here.
F. We remove the region upstream the activating region (the nucleotides between the repressor and the activator). Notice a little reduction in transcription but not a whole much. The difference between E and F is little to none.
G. We remove the activating region. Notice the huge reduction (drop) in transcription measured by how much light is produced, this tells me that there is an activating region here.
H. We remove the core promoter region. Notice there is very small production of the protein thus tells us that there is a core promoter here.
I. We remove base pairs downstream to the core promoter (base pairs between core promoter and transcription start site), notice there is hardly any transcription.
J. Similar to I.

Note that the rate of transcription in H, I& J is almost identical to that in A.

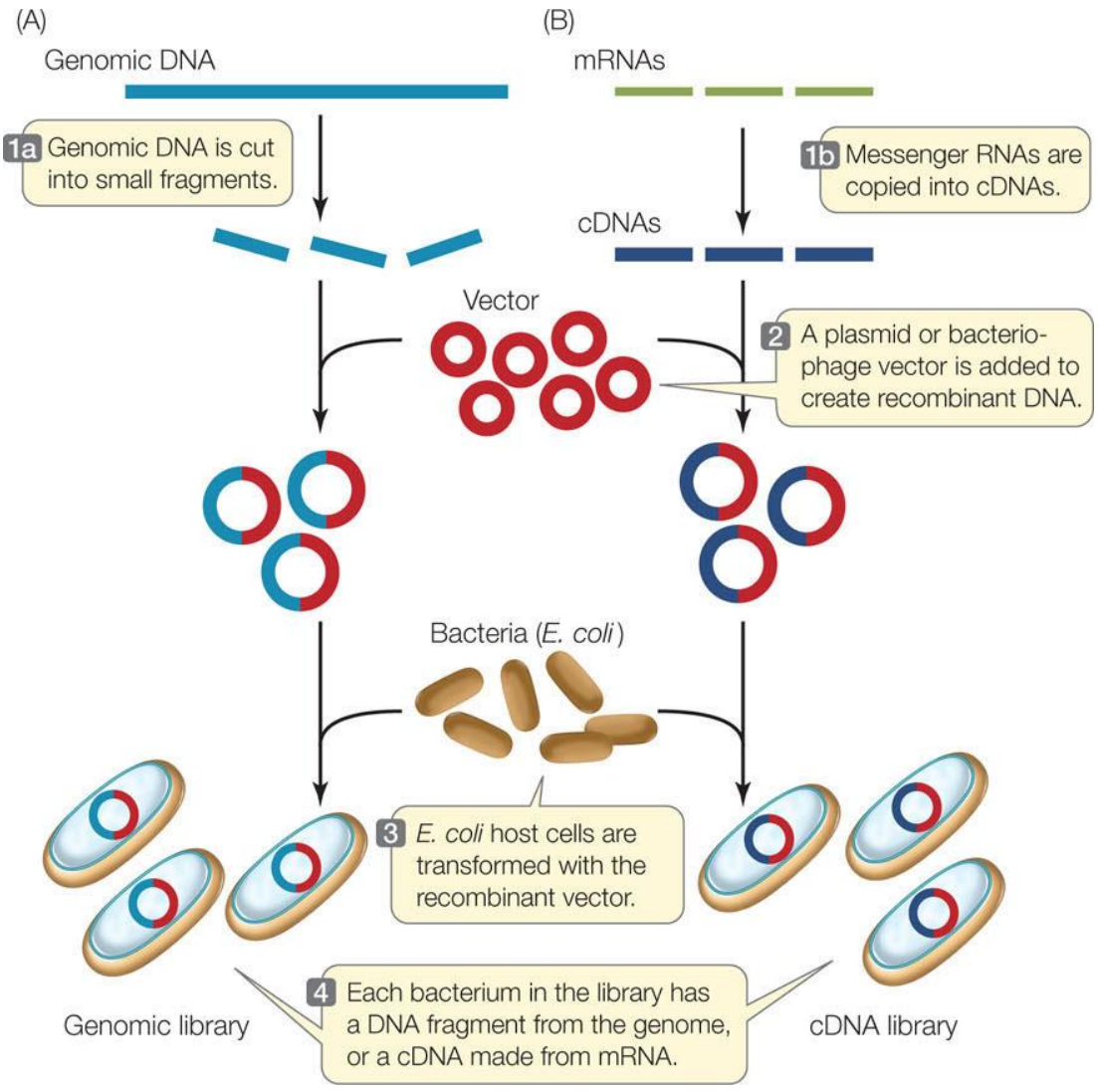
- This experiment is called reporter gene essay. (Essay means experiment or related to measurements). Enzyme essay for instance means a method by which I can measure enzyme activity.
- By this experiment we can tell whether a certain region is a repressor region, an activating region or a core promoter itself.
- In this experiment we combined recombinant DNA technology and use of enzymes in molecular biology.

The concept of DNA library

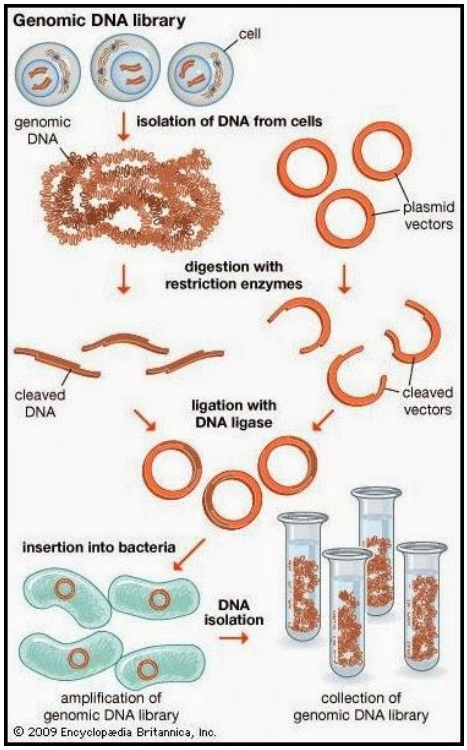
A library can be created for DNA fragments just like book libraries! A DNA library is similar to an ordinary library. It's a bunch of plasmids, each one of them contain a certain DNA insert. We keep these plasmids in a freezer either inside bacteria or as purified plasmids. So whenever you want to study a certain gene all you need to do is go to the computer and search for this particular gene, and it accurately informs you of the exact location (level, shelf, container etc..) and type of bacteria or plasmid containing that gene. In a DNA library we have different clones of bacteria or different plasmid molecules each one of them containing a specific piece of DNA insert.

How do we create a DNA library?

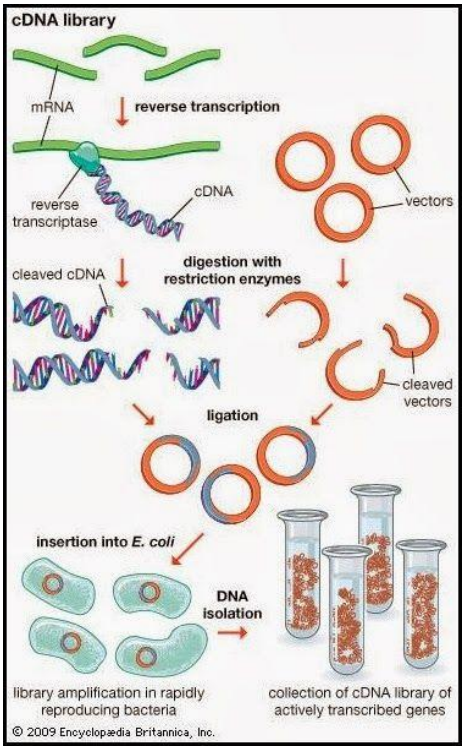
There are two types of libraries; a **genomic library** and a **cDNA library.**



A. **<u>Genomic library:</u>** which is basically when we take a whole genomic DNA (the whole DNA of a cell), fragment it into smaller pieces (fragments), and place each fragment into a plasmid so that each plasmid contains a specific piece of DNA. Then we take theses plasmids and insert them inside bacteria. Each bacterial cell will contain a plasmid and each plasmid will contain a specific insert. We grow bacterial cells individually on dishes, each bacterial cell would form a colony and we can then take these colonies one by one, grow them more, freeze them and identify which bacterial cell contains which plasmid and which plasmid contains which insert.



B. **<u>cDNA library</u>:** We take mRNA which is fully processed, meaning that it doesn't contain introns and consists of exons that are connected to each other, we convert mRNA into a cDNA representing a full expressed gene via reverse transcriptase. We put each cDNA into a plasmid so that each plasmid carries a gene. (the following steps are similar to those in genomic libraries). We insert these plasmids into bacterial cells, grow them on dishes, produce colonies etc…. Resulting in a literal library of DNA.



- cDNA library- c stands for complementary
- cDNA is made of mRNA not DNA.

Which method is preferred? cDNA libraries

Genomic libraries are much more complex than cDNA, because in genomic DNA we have EVERYTHING; introns, exons, enhancers, promoters, coding and non-coding genes, regions representing genes for RNA molecules, genes that code for proteins, telomere, centromeres and much more!

So, it's much easier and better to deal with a processed mRNA using the cDNA library methodology.

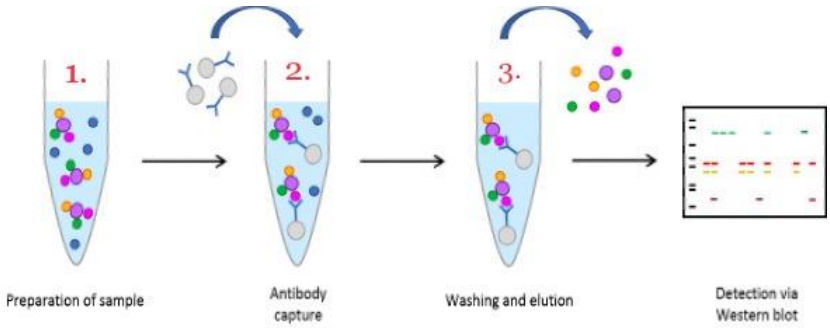
Protein-protein interaction

Protein-protein interaction is essential for cell function, but a protein doesn't behave by itself, it has to interact with other molecules. Proteins can interact with each other regulating each other's function.

(Co)-immunoprecipitation

Whenever you hear the word immuno you have to think about an immunoglobulin or an antibody. Antibodies are known to be very specific when binding to proteins. We can design antibodies/drugs against different proteins. The main idea behind this experiment is isolating proteins that interact with certain specific proteins in order to identify what these proteins are.

- An **antibody** is a protein that is produced by our immune cells in order to fight antigens or foreign bodies.
- Co means مصاحب.



Explanation of the diagram:

1. We have a sample of a bunch of cells and once we open up these cells, all of their proteins will be released. Since proteins interact with other proteins, you

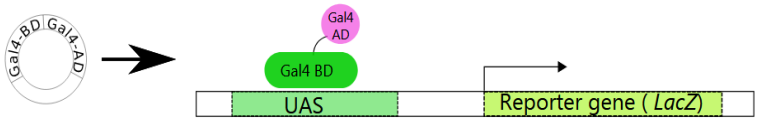
would have what is known as a protein complex (a complex of proteins bound to each other).

2. Antibody molecules that target a specific protein are conjugated to special beads. So, what we do is adding an immunoglobulin that is specific to a certain protein. (In this case you can see that it binds very specifically to the purple protein). Now we precipitate all the protein complexes that are bound to this immunoglobulin -only the protein of interest is precipitated as well as other proteins bound to it (co-precipitated)-. That way, we are specifically isolating or purifying the proteins that interact with our protein of interest.

3. We remove all the other proteins that are not bound to our complex. Then, we release the desired protein complex from the immunoglobulin so we can analyze the nature and identities of all these proteins using different techniques, an example is the western blot technique.

Yeast two-hybrid system

The yeast two-hybrid system is based on a genetic system that is activated by an element known as upstream activating sequence (UAS).



A. Regular transcription of the reporter gene

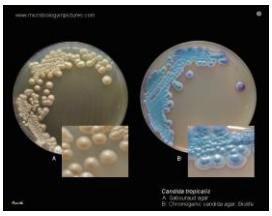
- UAS is controlled by a transcription factor called Gal4 that is made of two domains:
  1. A DNA-binding domain (BD).
  2. An activating domain (AD) that is responsible for the activation of transcription. It interacts with the RNA polymerase that sits on the promoter.

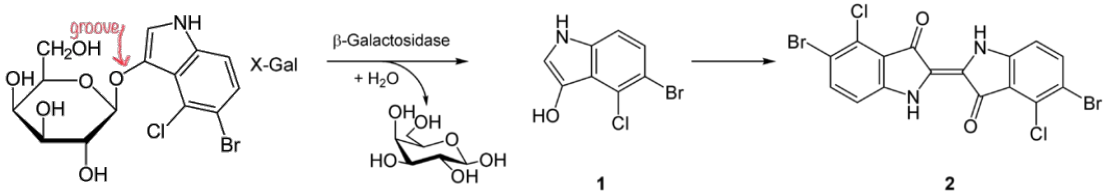
- Both domains must be close to each other in order to transcribe a reporter gene. The reporter gene used here is LacZ gene, a bacterial gene take from the Lac operon that produces β-Galactosidase.

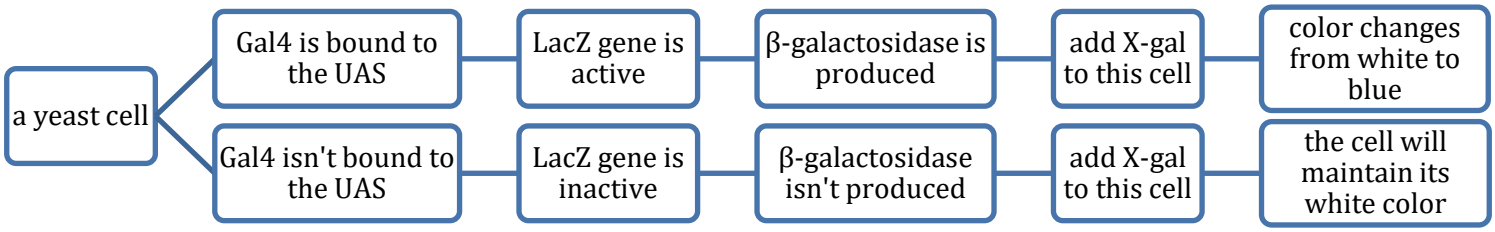
Domain: part of a protein that has its own three-dimensional structure. A domain maintains its structure and function even after it is cut from a protein.

LacZ gene

Yeasts cells are grown in the presence of a lactose analog called X-gal, this metabolite generates a blue product when cleaved.
When the LacZ gene is activated, β-galactosidase is produced, which cleaves (splices) X-gal into another molecule that's responsible for the blue color.







Cloning of hybrid proteins

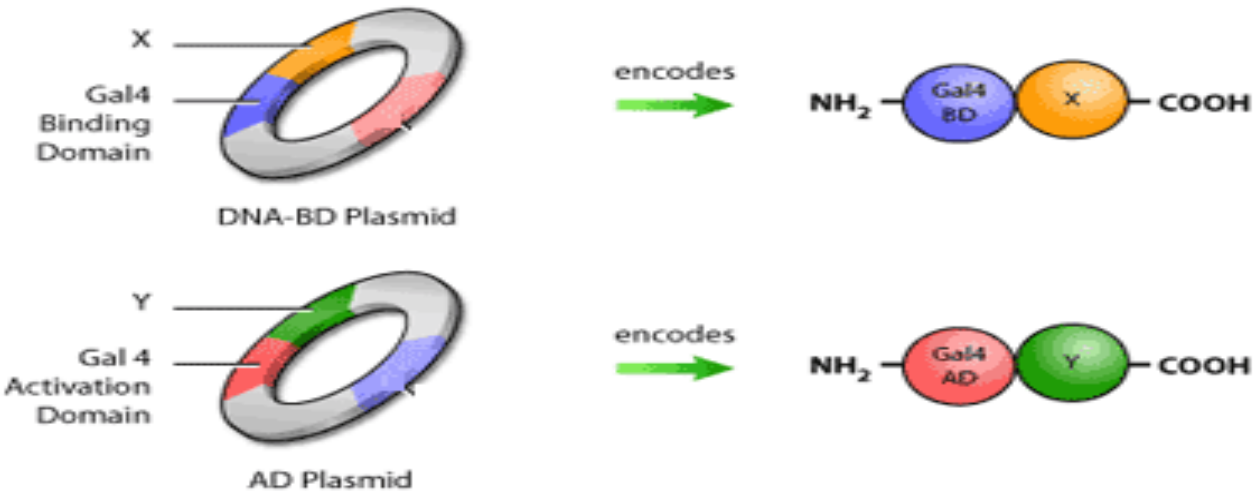


Diagram:

In order to discover unknown proteins (Y) that interact directly with a known protein (X), gene X is cloned so it is produced integrated with the DB domain and the unknown gene/s are separately cloned so they are produced integrated with AD.

The first plasmid:

Gene (X) is for the protein of interest, we want to find out which other proteins interact with this protein. Linked to it is the gene that produces the gal4 binding domain. Which encodes into a recombinant protein that consists of (X) and (gal BD).

When its inserted into a yeast cell no transcription of the lacZ gene will occur, why? Because of the absence of the activating domain (AD).

The second plasmid:

You can have any other different gene in this plasmid (Y1, Y2, Y3, Y4.....) in such a way that each plasmid will have one particular gene inserted into it. This gene is linked to the activating domain (AD).

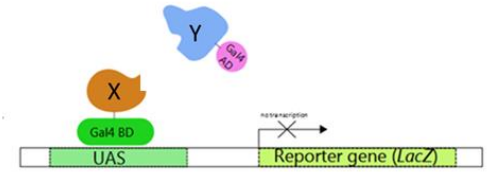
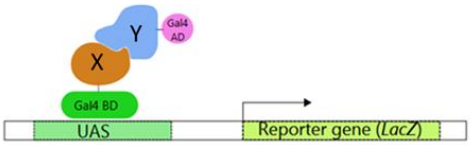
When this plasmid is inserted into yeast cells no transcription of the lacZ gene will occur, why? Because of the absence of the DNA binding domain (BD).

In order for the transcription to occur:

1. Both the activating and the binding domains have to be next to or close to each other.
2. Genes (X) and (Y) have to interact with each other.

Both recombinant plasmids (X and the unknown gene) are transferred into yeast cells. Each cell will produce two proteins:

1. Protein X linked to the DNA-binding domain.
2. A different Y protein linked to the activating domain.

no interaction between X and Y ▼ AD and BD not close ▼ no transcription ▼ white colonies (because there is no production of lacZ and no metabolism of the X-gal into the blue metabolites)	interaction between X and Y ▼ the domains are very close ▼ transcription will occur ▼ Blue colonies (Blue colonies are an indication that the unknown protein is able to interact with gene X)
	
	

After that we grow the colonies, isolate the plasmids, and find out what the protein is.

SHORT QUIZ

TEST BANK

1. In the yeast two-hybrid system, which of the following statements is accurate: A reporter gene

a) Is fused to the activation domain of a transcription factor
b) Is fused to the DNA binding domain of a transcription factor
c) Requires the presence of Histidine in the growth medium for its expression
d) Is expressed only if the tested protein interaction occurs

2. . In two hybrid screening system, the activator binds through _____ domain to a sequence upstream of the gene under its control, and _____ domain stimulates transcription.

a) DNA binding, activation
b) Activation, DNA binding
c) Activation, transcription
d) DNA binding, transcription

3. Luciferase genes are also used at times for detection. Choose the correct statement for them.

a) They are not obtained from fire flies
b) The detection requires provision of substrate which produces light
c) Enzymes such as beta-galactosidase requires substrate X-gluc to produce light
d) Lucifearse genes are preferred over fluorescent proteins

4. When the two domains are located in two different proteins, to preserve the same functionality, their close proximity and interaction have to be preserved as well.

a) True. B)false

5. Choose the correct answer about co-immunoprecipitation:
a) We use an antibody that can bind to all the proteins in the complex

a) We use an antibody that can bind to all the proteins in the complex
b) In the co-immunoprecipitation we know all the components of the complex
c) It is a way to detect protein-protein interactions
d) More than one of the above

ANSWERS

Q1	Q2	Q3	Q4	Q5
D	A	B	A	C