Ministry of Higher Education and Scientific Research

University ABOU BAKR BELKAID

Faculty of Natural Sciences and Earth and Universe Sciences

Departement of Biology

HANDOUT

Practical approaches in genetic engineering: Reminders and Exercises

By : **BERRAHOUI Samira**

For : **Biological sciences degree levels**

College year : 2023/2024

Summary

Genetic engineering is a revolutionary technology that allows the genetic makeup of an organism to be modified, either by removing, introducing or replacing one or more DNA fragments.

This handout, briefly describes the main tools and techniques commonly used to perform these manipulations.

Through theoretical reminders and practical exercises, the techniques of amplification, analysis and sequencing of genomes are covered: Molecular cloning, hybridization methods, PCR and sequencing.

These fundamental notions of genetic engineering, are addressed to undergraduate biology students, and more specifically to students of molecular biology and microbiology.

Key words: genetic engineering, recombinant DNA, cloning, amplification, hybridization, sequencing.

Summary

Introduction

Genetic engineering is a technological revolution and an artificialization of life which seems to give credit both to those who glorify it and to those who fear and condemn it.

The development of genetic engineering, that is to say all the techniques allowing the isolation, recombination and modification of genes, has made it possible to specify their structure and the regulation of their expression. However, the main goal of genetic engineering is the deliberate modification of an organism's DNA to improve or alter its characteristics.

Today we have a wide range of molecular tools and techniques for manipulating genes. Molecular cloning is often used in genetic engineering to copy and reproduce modified DNA molecules. This highly targeted method for endowing organisms with new genetic properties, involves the transfer of genes or other components of genetic material from bacteria, into the genetic material of plants or animals. Genetic engineering methods allow an approach highly targeted, only the gene corresponding to the desired new characteristic is transferred directly to the recipient organism

Techniques for manipulating, cloning and expressing genes were first developed in bacteria and then used routinely in numerous eukaryotic models. Today, transgenesis allows scientists to produce plants , animals and microorganisms that have desired characteristics for applications such as food and pharmaceutical production.

Genetic engineering is a very active field of research because the possible applications are multiple, but can be grouped into four main areas:

- Agronomic and environmental improvements: Many transgenesis works concern the introduction of genes for resistance to herbicides or insects, and to a lesser extent, to certain viruses and diseases. Combined with reasonable use of herbicides and pesticides, these transgenic plants will improve the efficiency of agriculture, while respecting the environment even better.

- Production of molecules of industrial interest: Biotechnologies open up numerous perspectives in the fields of industry, by producing new molecules (Molecular Farming) and by improving industrial processes and product quality.

- Production of molecules intended for human health: Genetically modified, bacteria, yeasts, tobacco, corn, or potato plants.....can produce therapeutic molecules or vaccines. The great

advantage of the production of these molecules is the absence of risks of contamination by viruses pathogenic for humans.

- Medical research and gene therapy: The medical applications of such scientific advances are already considerable in most areas of pathology: genetic, infectious, cancerous diseases. In 1991, the first gene transfer trials in humans, with the aim of correcting a constitutional enzyme deficiency, were carried out and provided a glimpse of the multiple applications of gene therapy.

This handout is aimed at students of the Bachelor of Biology, 3rd year of the Molecular Biology and Microbiology sectors. It can be used as support for tutorials in the genetic engineering module. Indeed, it reviews the main tools and techniques of genetic engineering and the manipulation of genes, covered in class.

Each chapter is made up of three parts; first a theoretical part intended to recall the fundamental notions and knowledge acquired in class. Then a series of some exercises aimed at highlighting the essential concepts to remember. The third section provides standard answers for each exercise.

Thus this handout aims to help the student to master the fundamental notions learned in class and to understand, through corrected exercises, the modes of action of the different tools and their interest in genetic engineering techniques. Furthermore, this handout attempts to explain through concrete applications the principles and different possible uses of genetic engineering techniques in the field of research and analysis of genomes.

This handout introduces students to recombinant DNA technology to better understand, in a master's degree, the manipulation of DNA in eukaryotic systems, such as transgenesis, reproductive cloning and therapeutic cloning, on the one hand. And on the other hand, it prepares them to master the applications of genetic engineering techniques in mutagenesis, interreference, and gene therapy among others.

However, students must have a certain number of prerequisites in molecular biology and biochemistry. They must know precisely the structure and physicochemical properties of DNA and RNA molecules, but also the mechanisms of replication and expression of the genome.

I. Restriction enzymes

1. Reminders

Genetic engineering developed when bacterial enzymes were discovered in 1965, capable of cutting DNA according to a precisely predetermined sequence. This discovery was confirmed in 1973 by Paul Berg and his collaborators.

These proteins, capable of precisely cutting and re-gluing DNA , give researchers the tools they lacked to map the genome . They also open the way to recombinant DNA technology which allows the insertion of a portion of DNA (one or more genes) into another DNA .

Restriction enzymes are endonucleases that recognize double-stranded, base-specific sequences in the DNA double helix (restriction sites),

- 4 to 15 base pairs

- usually palindromes, that is to say they are composed of identical nucleotide sequences on both strands but in antiparallel orientations.

They cleave the two strands of the duplex at specific locations (cutting site) by hydrolysis of the phosphodiester bond, to obtain a restriction fragment.

Restriction enzymes are proteins synthesized primarily by bacteria to protect against bacteriophage infections. At the same time, the bacteria has methylases capable of modifying (methylating) its own DNA so that it is not recognized by restriction enzymes.

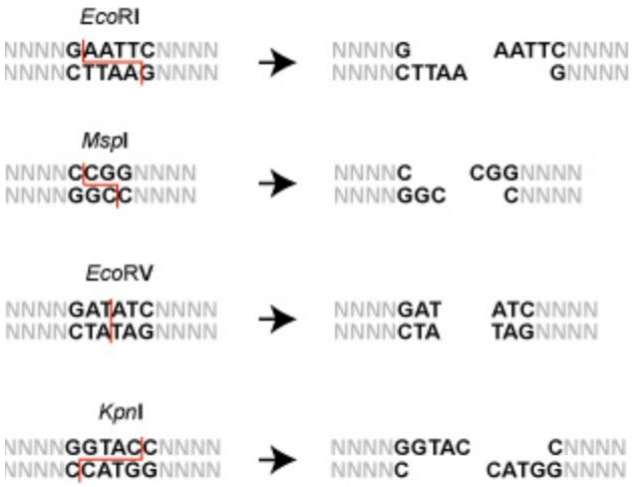
This mechanism of resistance to bacteriophages, called restriction, was studied by W. Arber at the University of Geneva in the 1960s. He won the Nobel Prize in Medicine in 1978 with D. Nathans and H. Smith for the discovery and applications. restriction enzymes.

The name of a restriction enzyme indicates its origin: *Sau* 3A comes from *Staphylococcus aureus* 3A, *Bam* HI comes from *Bacillus amyloliquefaciens* H, *Eco* RI comes from *Escherichia coli* , *Msp* I comes from *Moraxella* species, KpnI comes from *Klebsiella pneumoniae* .

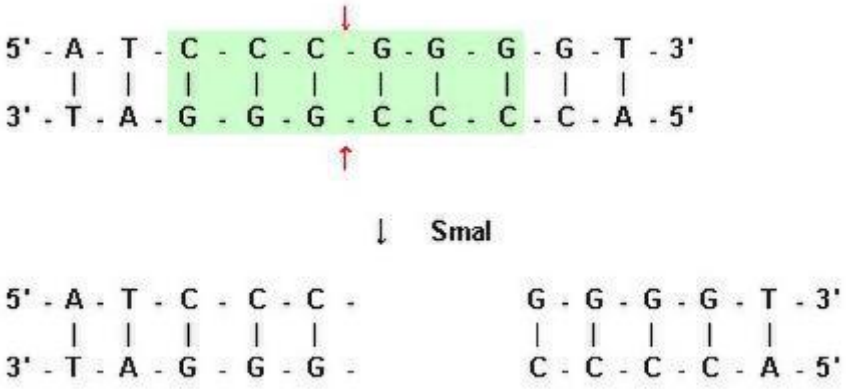
The first capital letter indicates the genus, the two lowercase letters the species, sometimes the strain is also indicated, the Roman numeral indicates the order of discovery of the enzyme in the same species.

Restriction enzymes can cut the restriction site either in the middle, generating ends of the same length (double strands) which are called blunt ends, or in an offset manner, and generate sticky or cohesive single-strand ends (those which are used in recombinant DNA technologies)

Sticky ends



Blunt ends



Restriction enzymes are essential in the cloning protocol, since they make it possible to generate complementary cohesive ends between the insert and the vector, thus allowing their recombination. To do this, the insert must be flanked by unique and preferably different restriction sites (in order to prevent the circularization of the DNA fragment by complementarity). These sequences containing a restriction site are called linkers. It is possible to generate complementary cohesive ends either by using the same enzymes to cut the insert and the vector, or by using compatible enzymes.

These molecular scissors are also essential for genome editing , which allows targeted genetic modifications to be carried out in any type of cell. Available since the 1980s, these tools have gained in efficiency and specificity over time. In 2012, the advent of the CRISPR-Cas9 system, characterized by its great simplicity and modest cost, revolutionized this approach: genome editing has now spread to all areas of science and medicine. It allows researchers to carry out the genetic modifications of their choice, in order to develop tailor-made cellular and animal models, to progress in the knowledge of the development of living organisms, of diseases, or even to test therapeutic molecules. The first clinical trials based on this approach have begun, aiming to treat monogenic diseases, certain cancers and even infectious diseases.

The notions of restriction sites, restriction fragments and cohesive and complementary ends are covered in the exercises below.

2. Applications

Exercise 1 : Briefly answer the following questions:

1- The name of the first enzyme synthesized by Streptomyces exfoliatus strain A

2- Enzymes which belong to class II among the following enzymes justifying your answer: DpnII /GATC, SexAI A/CCNGGT, Sau3AI /GATC, EcoRI G/AATTC, MabI A/CCNGGT, EcoRV GATC/ATC, BamHI G/GATCC

3- Compatible enzymes among these same enzymes

4- Isochisomeric enzymes among these same enzymes

5- Those which would be synthesized by Diplococcus meningitidis and Microbacteruim arborescens

Exercise 2 : The sequence of a double-stranded DNA, corresponding to a gene, is partially shown below:

5'ATACGGGATCCGAGCTCTCGATCGTCTGCAGAAATTCC 3'

Consider the restriction enzymes BamHI, Pst I, XhoI and Mbo I whose recognized sites are: BamH I: 5' G/GATCC 3' PstI: 5' CTGCA/G 3'; Xho I: 5'C/TCGAG 3'; Mbo I: 5' /GATC 3'.

1- Copy the DNA sequence and frame the restriction sites, indicating the position of the cuts
2- For each enzyme, write the sequences of the ends of the digested DNA molecules and specify the type of ends obtained

Exercise 3: Consider the following restriction enzymes:

BamHI G/GATCC; PstI CTGCA/G; MstII CCTNA/GG

Which generate protruding 3' cohesive ends?

Which generate protruding 5' cohesive ends?

Exercise 4: The following fragments are obtained after cleavage by the BamHI enzyme

5' G 3' 5' GATCC 3'

3' CCTAG 5' (1) 3' G 5' (2)

- The following fragments are obtained after cleavage by the MboI enzyme

5' To 3' 5' GATCAGC 3'

3' TCTAG 5' (3) 3' TCG 5' (4)

1- Reconstruct the restriction sites and cleavage sites of these two enzymes

2- Which of these ligatures are possible?

a- 1+4

b- 1+3

c- 2+3

d- 2+4

3. Corrected exercises

Exercise 1 :

1- Sex A 1 (**Streptomyces** genus, **ex** foliatus species, **A** strain **,** first enzyme discovered **)**

2- All these enzymes belong to class II because they recognize a palindromic sequence and cut at this sequence

3- Compatible enzymes recognize different sequences but generate complementary sticky ends:

DpnII, Sau3AI and BamHI: **GATC**

SexAI and MabI: **CCNGGT**

4- Isochizomeric enzymes recognize the same palindromic sequence:

DpnII and Sau3AI: /GATC

SexAI and MabI: A/CCNGGT

5- None is synthesized by Diplococcus meningitidis, MabI is synthesized by **M** icrobacteruim **ar** borescens

Exercise 2 :

1- Restriction sites

5'ATACG **G / GATCC** GAGCTCTC / **GATC** GT **CTGCA / G** AAATTCC 3'

3'TATGC **CCTAG/G** CTCGAGAG **CTAG** /CA **G/ACGTC** TTTAAGG 5'

2- Restriction fragments and cohesive ends

BamHI gives outgoing 5' ends

5'ATACGG 3' **5' GATC** CGAGCTCTCGATCGTCTGCAGAAATTCC 3'

3'TATGCC **CTAG 5'** 3' GCTCGAGAGCTAGCAGACGTCTTTAAGG 5'

PstI gives protruding 3' ends

5'ATACGGGATCCGAGCTCTCGATCGTC **TGCA 3'** 5' GAAATTCC 3'

3'TATGCCCTAGGCTCGAGAGCTAGCAG 5' **3' ACGT** CTTTAAGG 5'

No XhoI site so no interruption

MobI gives 5' outgoing ends

5'ATACGG 3' **5' GATC** CGAGCTCTC 3' **5' GATC** GTCTGCAGAAAATTCC 3'

3'TATGCC **CTAG** 5' 3' GCTCGAGAGCTAG 5' 3' CAGACGTCTTTAAGG 5'

Exercise 3:

BamHI generates outgoing 5' ends

5'G 3' **5' GATC** C 3'

3'C **CTAG 5'** 3' L 5'

PstI generates outgoing 3' ends

5'C **TGCA 3'** 5' L 3'

3'G 5' **3' ACGT** C 5'

MstII generates protruding 3' ends

5' CC **TNA 3'** 5' GG 3'

3'GG 5' **3' ANT** CC 5'

Exercise 4:

1- Restriction sites

- BamHI site (palindromic sequence)

  5' G/GATCC 3'

  3' CCTAG /G 5'

- MboI site (palindromic sequence)

  5' A/ **GATC** AGC 3'

  3' T **CTAG/** TCG 5'

2- Possible ligatures

e- 1+4

f- 2+3

Complementary ends

Ends 1 and 3, 2 and 4 are identical and therefore cannot match.

II. Cloning

1. Reminders

- **Cloning** is the natural or artificial identical multiplication of a living being, that is to say with exact conservation of the same genome for all descendants (clones).

Molecular cloning does not allow the cloning of an entire organism, but only a gene or a DNA fragment.

It is a molecular biology technique which consists of isolating a DNA fragment and multiplying it identically by inserting it into a DNA molecule called a vector (the vector-insert association is called recombinant DNA), allowing its amplification . This technique also makes it possible to produce the recombinant protein corresponding to the cloned gene.

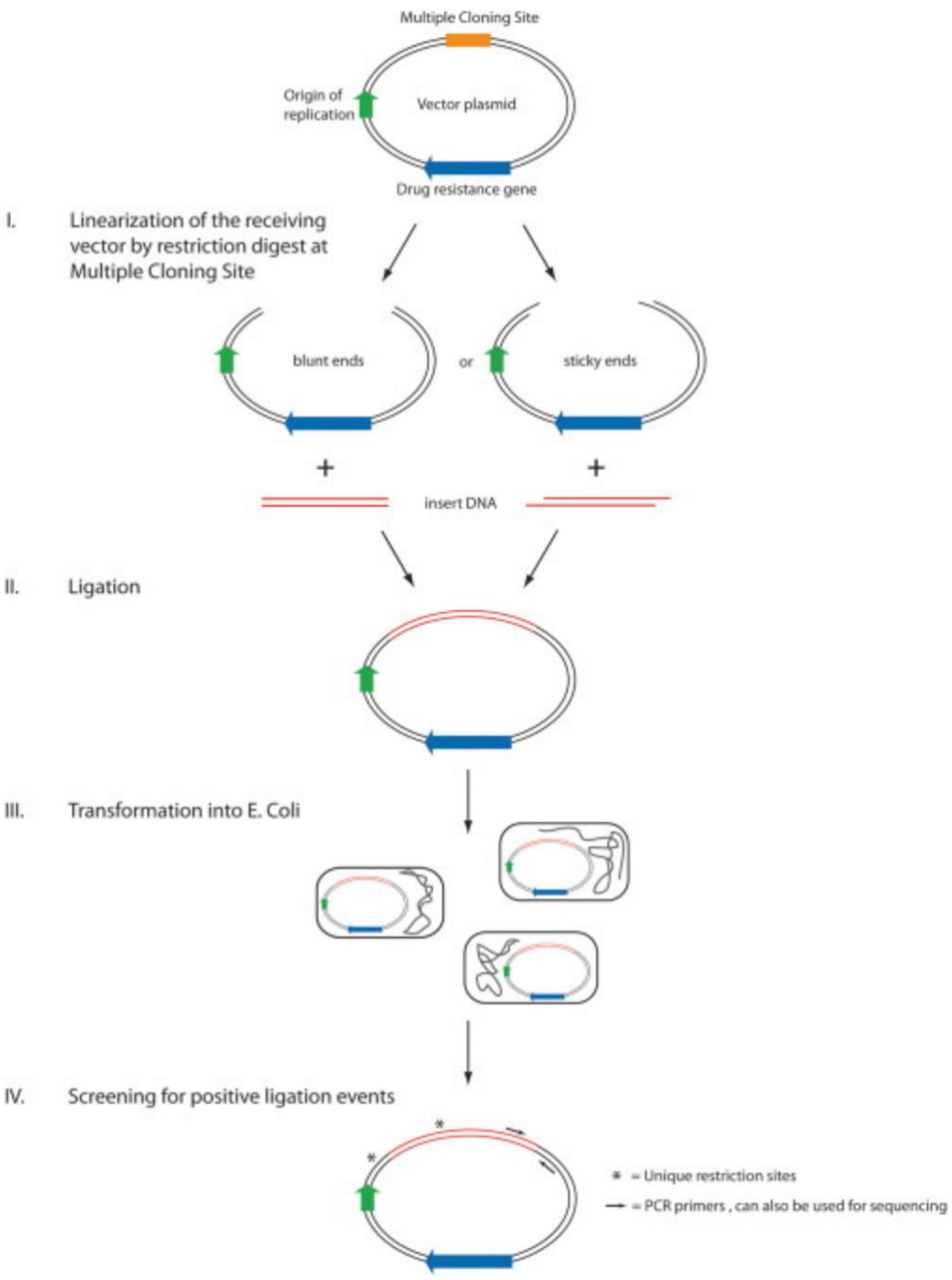
Carried out in the laboratory, this manipulation of the genome aims to artificially introduce foreign genetic material into another DNA sequence which will integrate it. Very common in research or development projects, the use of recombinant DNA allows the development of essential medical products and therapeutic proteins such as, for example, insulin intended for diabetics or a certain number of vaccines. Recombinant DNA also made it possible to identify the protein responsible for hemophilia.
In genetic engineering, this practice of DNA recombination makes it possible to identify the behavior and characteristics of specific genes. This results in research on genetically modified organisms (GMOs) and cloning for therapeutic purposes.

In 1977, the first human gene, coding for somatostatin, was cloned, allowing bacteria to produce human proteins. The era of genetic engineering and biotechnology had just begun.

Molecular cloning entails the preparation of the vector and insert DNAs, ligation of the insert into the vector, transformation of competent *E. coli*, and identification of positive clones.The four basic steps of molecular cloning are outlined in the figure below.

Traditionally, molecular cloning is defined as the isolation and amplification of a specific DNA fragment. Most of these fragments are created either by digesting an existing piece of DNA with restriction enzymes or by targeting it via PCR. Short inserts of ~ 100 bp can also be commercially synthesized as complementary single-stranded oligos, which are subsequently annealed to form a double-stranded fragment.

After successful isolation, the DNA of interest is ligated into a vector that can propagate in E. coli. The vectors used in the laboratory represent a modified version of small DNA molecules capable of self-replication in a host cell.



Molecular cloning strategy overview

- **Vectors** are molecules capable of self-replication thanks to the presence of an Ori sequence (origin of replication), which allows them to replicate hundreds of times independently of the genome of the host cell. These are therefore plasmids or viruses, having undergone modifications to be suitable for cloning.

Among the essential modifications:

- Presence of cloning sites, i.e. unique restriction sites to allow the opening of the vector and the insertion of the DNA of interest. Most often the vectors are equipped with polylinker or MCS (multi cloning sites) sequences. It is a DNA fragment whose sequence is a succession of unique restriction sites for numerous restriction enzymes, thus offering a range of choices for the use of different restriction enzymes.

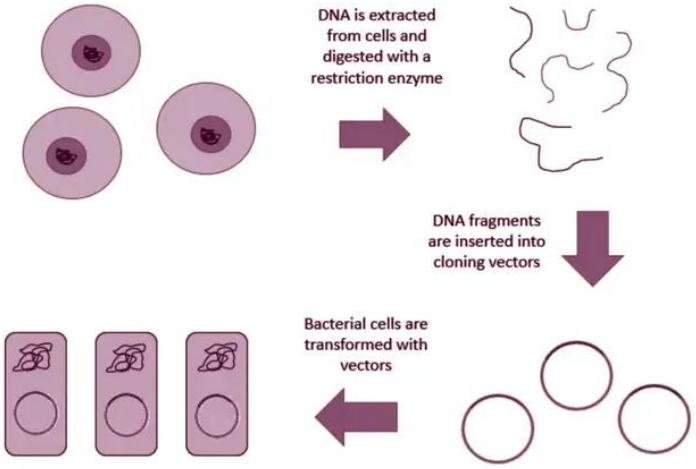
- Selection markers: These are genes conferring on the host cells transformed by the recombinant DNA, new characteristics allowing their selection: elimination of non-transformed cells (thanks to a resistance gene carried by the vector), and distinction between cells transformed by the recombinant DNA and those transformed by the non-recombinant vector (a second gene for resistance to an antibiotic, gene laz, etc.)

- The recombinant DNA technique also allows the construction of **DNA libraries** . A DNA library is a collection of cloned DNA molecules corresponding to a given species.

Two types of banks can be constructed: genomic DNA bank and complementary DNA bank.

The genomic DNA library is an assortment of clones containing fragments of the total genomic DNA of an organism. It therefore contains the entire genome of the organism, including coding and non-coding sequences. It is obtained by fragmentation of the genome extracted from the nucleus by restriction enzymes. It is the same for all the cells of an organism.

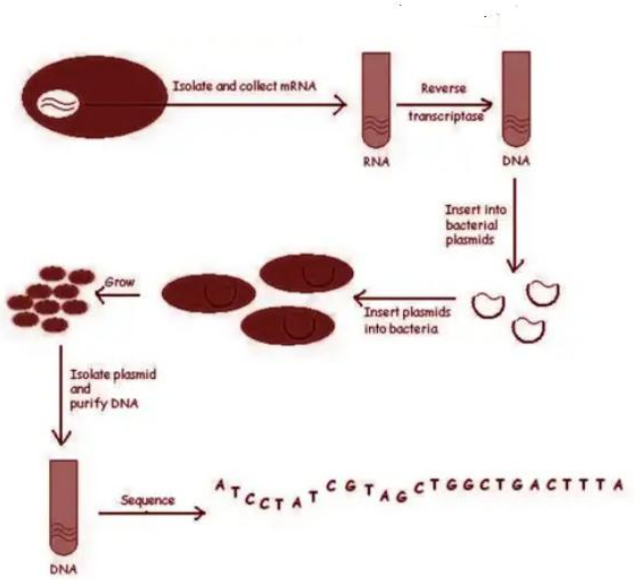
Genomic libraries are essential for understanding the function and structure of genes specific to genomics, genetic mapping and mutations of gene loci, as well as gene sequencing, discovery of new therapeutic genes, etc.



Construction of genomic DNA library

The complementary DNA library is an assortment of DNA created from the total mRNA of an organism, through the process of reverse transcription. This bank therefore contains only the genetic sequences, but devoid of regulatory and intronic regionsi. There are therefore several cDNA libraries for the same organism, which differ from one cell type to another, and from one stage of development to another, depending on the genes expressed (transcribed).

cDNA libraries are essential for studying gene functions, coding regions and gene expression, etc.

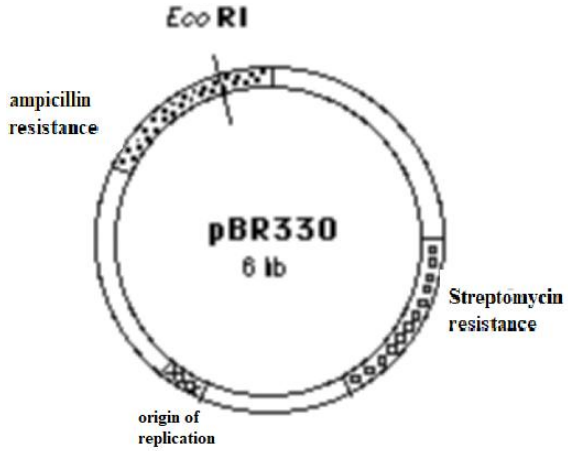


Construction of cDNA library

The classic cloning protocol as well as different strategies for selecting clones transformed by recombinant DNA are covered in the exercises below.

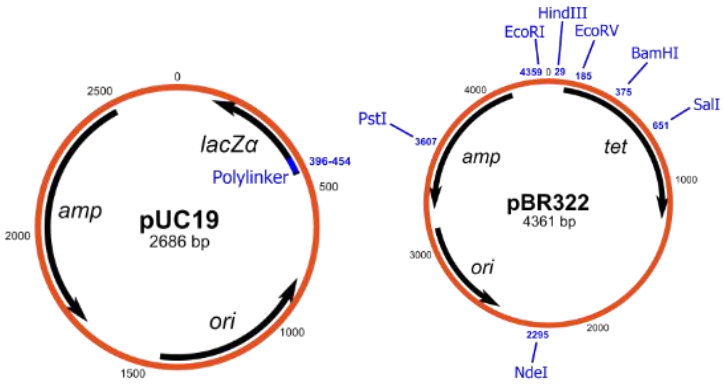
2. Applications

Exercise 1 : We want to study the mouse M gene. For this, we try to clone into the plasmid vector pBR330 (see diagram) an *Eco* RI- *Eco* RI fragment of mouse genomic DNA carrying the gene of interest.



- Propose a cloning protocol and indicate how you select recombinant clones.

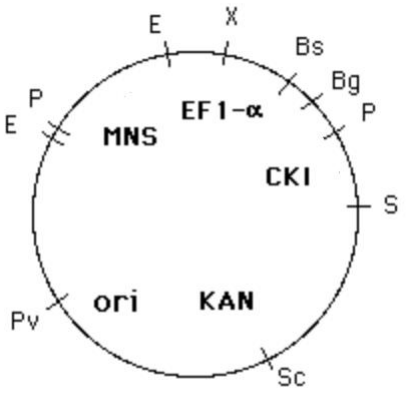
Exercise 2 : The maps of two molecular cloning plasmids are shown below:



1- Where are the cloning sites on the two plasmids?
2- How are the clones transformed by recombinant pBR322 selected?
3- How are the clones transformed by recombinant pUC19 selected?

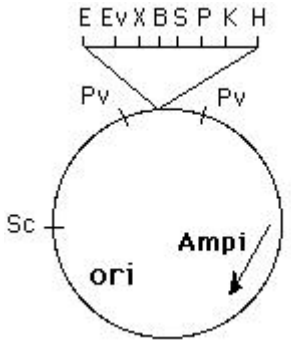
Exercise 3 : An *Eco* RI- *Sma* I fragment of tobacco genomic DNA was cloned on which three genes are found: MNS, EF1-alpha and CKI. We are in fact only interested in the EF1-alpha gene (between the E and Bg sites) which we wish to subclone into the pBLUESCRIPT vector.

Map of the pTOB vector carrying the genomic DNA fragment is as follows:



E: *Eco*RI, G*AATTC
X: *Xba*I, A*GATCT
Bg: *Bgl*II, T*GATCA
P: *Pst*I, CTGCA*G
Bs: *Bst*EII, G*GTNACC
Sc: *Sca*I, AGT*ACT
S: *Sma*I, CCC*GGG
Pv: *Pvu*I, CGAT*CG

pBLUESCRIPT map is as follows:



H: *Hind*III, A*AGCTT
B: *Bam*HI, G*GATCC
Ev: *Eco*RV, GAT*ATC
K: *Kpn*I, GGTAC*C

pBLUESCRIPT

KAN : kanamycin resistance
Ampi : ampicillin resistance
ori : origin of replication of the plasmid

1- Establish a strategy to achieve this subcloning?
2- How would you proceed to select the recombinant clones?

Exercise 4 : Based on the protocols for constructing DNA banks (see reminders), compare between a gDNA bank and a cDNA bank according to the following criteria: Content, size,

starting material, manufacturing protocol , vectors used for cloning and the possibility of expression in a prokaryotic system.

3. Corrected exercises

Exercise 1 :

- The insert is surrounded by two EcoRI sites, so it is digested with the EcoRI enzyme to generate the cohesive ends
- It is treated with alkaline phosphatase to avoid self-ligation (complementary cohesive ends)
- The plasmid is digested with the EcoRI enzyme to generate cohesive ends
- Ligation is carried out between the vector and the digested insert (which have complementary cohesive ends) using a ligase enzyme, the role of which is to establish phosphodiester bonds between the two DNAs, thus forming the recombinant DNA.
- We transform competent host cells (bacteria, most often *Escherichia coli*) and we cultivate them on medium supplemented with streptomycin in order to eliminate non-transformed cells
- We select the cells transformed by the recombinant DNA by transplanting them on ampicillin medium, those which do not grow contain the recombinant DNA (the amp gene has been deactivated by the insertion of the insert)

Exercise 2 :

4- The cloning sites correspond to the restriction sites, which are located on the polylinker for pUC19 and scattered on the tet and amp genes for pBR322

5- The clones transformed by the recombinant pBR322 are selected by subculture on medium containing the good antibiotic:

- if the cloning was carried out at a restriction site located on the amp gene, the transformed cells are first cultured on tet medium then subcultured on amp medium. Tet medium eliminates empty cells. Subcultured cells that do not grow on amp medium contain the recombinant DNA.
- if the cloning was carried out at a restriction site located on the tet gene, the transformed cells are first cultured on amp medium then transplanted onto tet medium. The amp medium eliminates empty cells. Subcultured cells that do not grow on tet medium contain the recombinant DNA.

6- The clones transformed by the recombinant pUC19 are selected directly on the amp+X-gal medium on which the transformed cells are cultured. The white cells contain the recombinant DNA. The amp eliminates the empty cells.

Exercise 3 :

1- The subcloning of the EF1-alpha gene is carried out according to the following protocol:

- Cut the pTOB plasmid using enzymes E and B to release the gene of interest.

- Cut the pBluescript vector with the enzymes E and Bg (enzyme compatible with B since the B site is not present on the pBluescript polylinker)

In both cases, the cohesive ends generated are not complementary, it is not useful to treat the DNA digested with alkaline phosphatase.

- Transform host cells and cultivate them in amp medium to eliminate non-transformed cells.

2- To select the clones containing the recombinant DNA, the pBluescript extracted from the cells transformed by the enzymes E and Bg is digested, then electrophoresis is carried out. The presence of two bands on the gel indicates the presence of the insert. A single band corresponds to the linearized vector.

Exercise 4 :

Criteria	gDNA	cDNA
Content	- collection of clones carrying the total genomic DNA of an organism. - includes the entire genomic DNA, including non-coding DNA (introns and regulator). - includes all the genes of the organism	- collection of clones carrying DNA complementary to the mRNA of an organism - contains only the coding sequences; does not contain introns or promoters - contains only expressed genes
Size	Big	small
Starting material	DNA extracted from the nucleus	mature mRNAs contained in the cytoplasm
Manufacturing	DNA fragmentation by restriction enzymes	mRNA retrotranscription
Vectors	The plasmid, the cosmid, the lambda phage, the BAC and the YAC in order to accommodate large fragments	The plasmid, the phagemids, the lambda phage, to accommodate small fragments
Expression in a bacteria	No The bacteria cannot carry out splicing (sequences containing introns)	Yes cDNA does not contain introns (does not require splicing)

III. RFLP and Restriction Cards

1. Reminders

In 1973 the first restriction map for simian virus DNA was constructed, by Danna et al. Since then, restriction maps (also sometimes called physical maps) have become fundamental data structures in molecular biology.

- **A restriction map** is a graphical representation of the location of restriction sites on a DNA molecule, after having subjected it to enzymatic digestion. It gives the number of sites for each enzyme used as well as the distance separating them (in pairs of nucleotides).

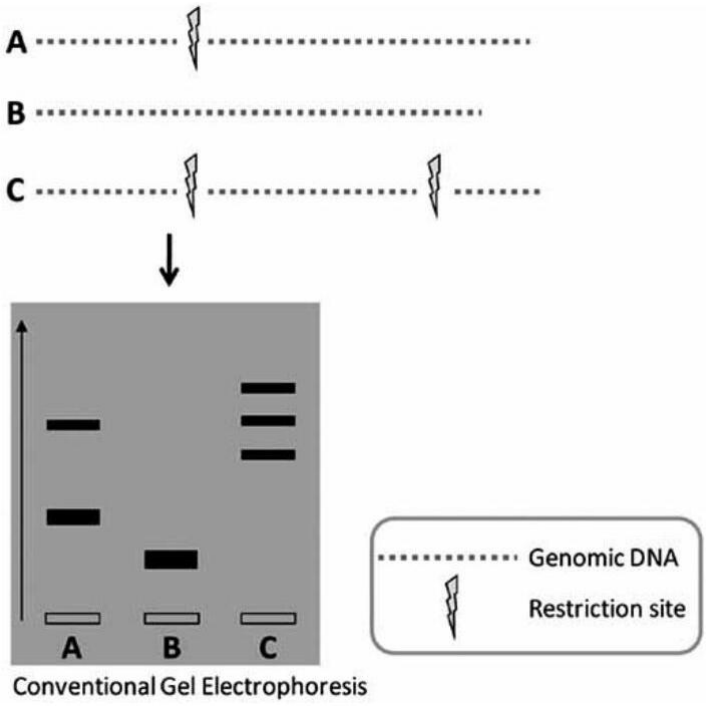
  The objective of restriction mapping is to reconstitute the DNA sequence by reordering the fragments resulting from digestions.

  The use of mathematical models to constitute the restriction map is based on the results obtained from partial or complete digestions of the sequence. By controlling the experimental conditions, either a total digestion or a partial digestion can be carried out. Under total digestion conditions, all cut sites are cut, whereas in partial digestion, a proportion of cut sites remain intact.

Double digestion is a much simpler experimental mapping technique than partial digestion. In this approach, a biologist maps the positions of the sites of two restriction enzymes through complete digestion of DNA, so that only fragments between consecutive sites are formed. To construct such a map, one can measure the fragment lengths (and not the order) of a complete digestion of DNA by each of the two enzymes separately, then by both enzymes applied simultaneously. The problem consists of determining the positions of the cuts from the lengths of the fragments.

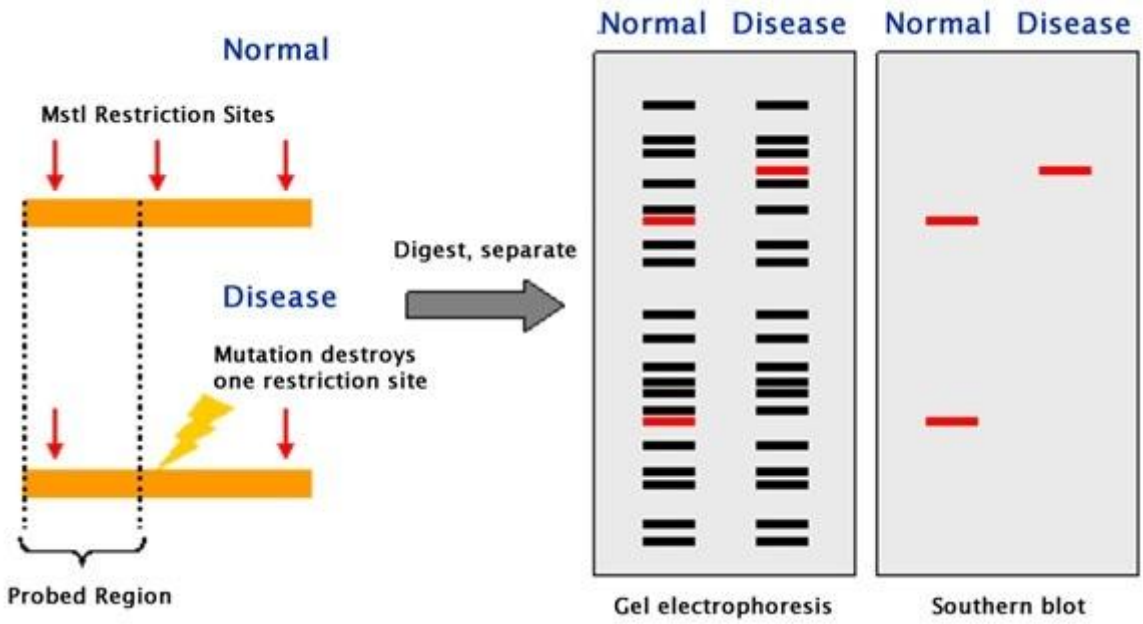
- Analysis of the restriction profile makes it possible to highlight variations in the presence and position of restriction sites within DNA. This technique called **RFLP (Restriction Fragment Length Polymorphism)** reveals point mutations causing the abolition or appearance of restriction sites.

RFLPs are very useful as markers, particularly for the mapping of chromosomes and the analysis of polymorphisms (paternity tests, genetic fingerprinting in criminology and leagle medicine, characterization of genetic diversity or reproduction models in animal populations).



Restriction Fragment Length Polymorphism

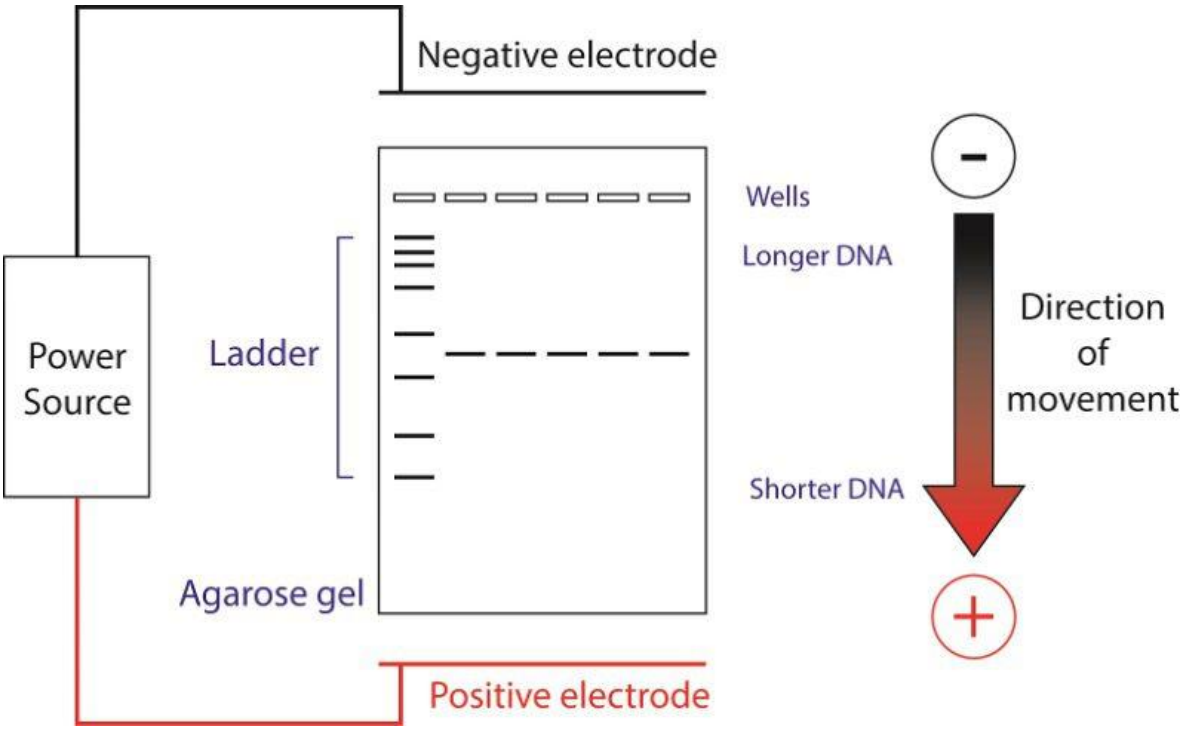
Genetic linkage of an RFLP to a human disease gene is also a useful tool for counseling prospective parents and diagnosing diseases.



Disease diagnosis by RFLP

- **Electrophoresis** , an essential technique in molecular biology, makes it possible to reveal the restriction profile after digestion. The restriction fragments migrate in the agarose gel which allows them to be separated, visualized and measured their size, thus allowing the reconstruction of the restriction map.

Electrophoresis is a separation technique. It is used most often for analytical purposes but also sometimes for purification. The principle consists of subjecting a mixture of DNA or RNA molecules to an electric field, which causes the migration of the (negatively) charged molecules at variable speeds depending on their size . The speed of migration is inversely proportional to the size of the nucleic acid molecules.



Electrophoresis technique

The RFLP technique and the interpretation of restriction profiles as well as their interest are discussed in the exercises below.

2. Applications

Exercise 1 : A 3Kb plasmid is digested by the XhoI enzyme which allows it to be linearized. Its digestion with SmaI gives two 1.5Kb fragments. Double digestion generates two fragments of the same size and one fragment of length equal to that of the fragments obtained after simple digestion with SmaI.

Based on these data, diagram the restriction map of this plasmid and the restriction gel obtained following a single then double digestion.

Exercise 2 : Consider double-stranded DNA, whose single-strand sequence is as follows:

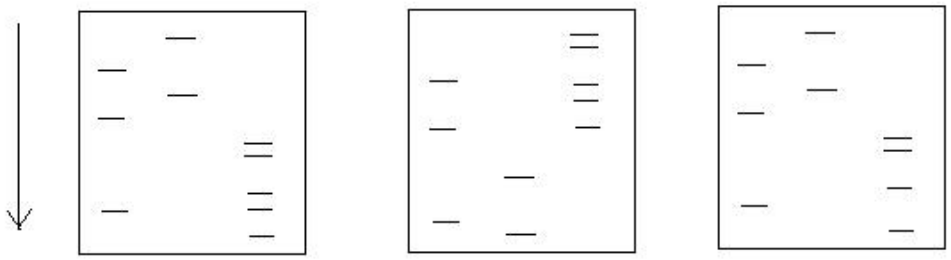
CAATTGAATTCATCGTAGTAGGCCGTACATGCGTGGAATTCGTAATGTGAATTCT

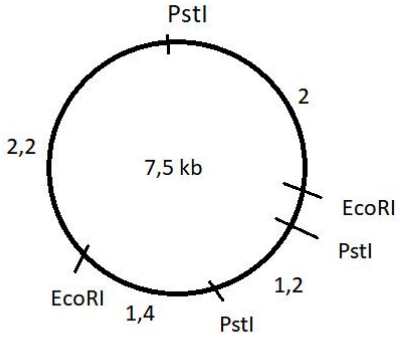
GAAATTTCCCTACATGGGCTACGGCTAGTAGCGCTGATGGGCCATCGCGCTGGG

CTGGAATTCGGTCTATT

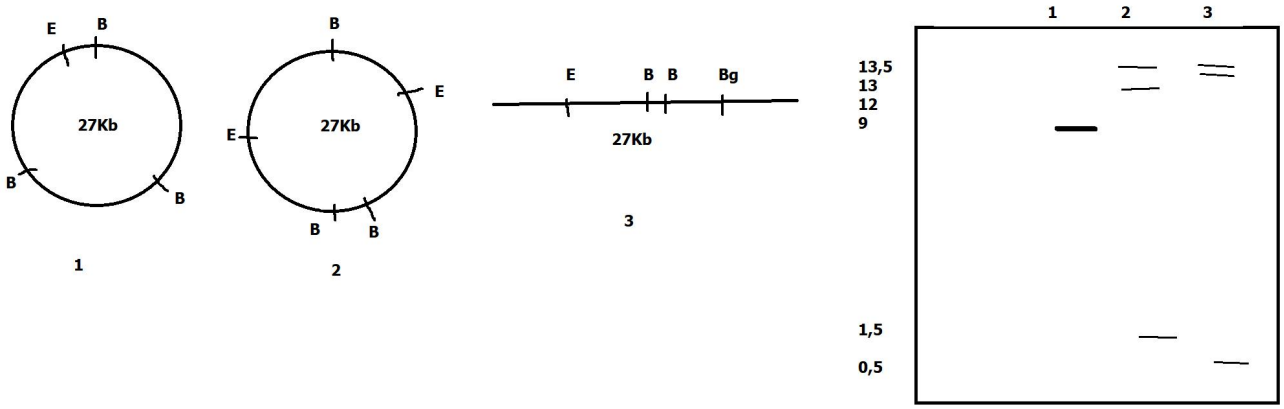
An enzymatic digestion is carried out using a restriction enzyme EcoRI (G/AATTC), represent the restriction profile of this fragment and justify your answer.

Exercise 3 : Which of these gels corresponds to the restriction profile corresponding to the restriction map shown below? Justify your answer then indicate for each track the enzyme used and specify the sizes of the bands (only for the correct gel)





Exercise 4 : Represented below are the restriction maps of 3 DNA molecules and their respective restriction profiles. The 3 fragments were digested by the same enzyme. Based on the figures, find which enzyme it is, justifying your answer, then complete the restriction maps by indicating the distances between the different restriction sites. Also indicate by arrow the direction of migration of the fragments on the gel, justifying your answer.



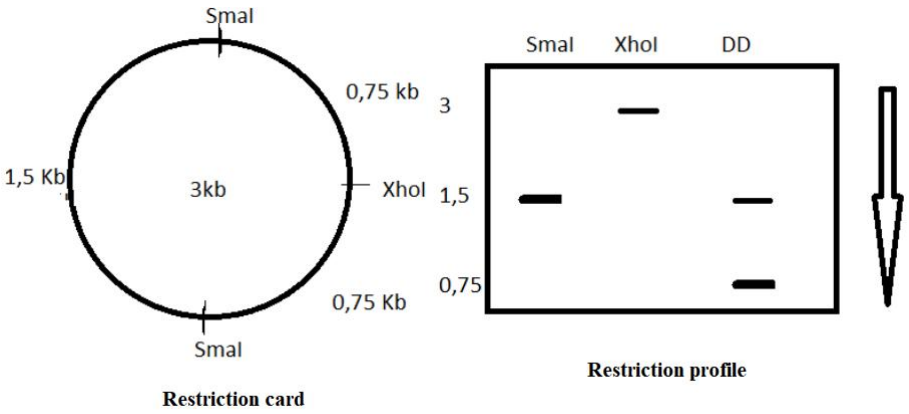
3. Corrected exercises

Exercise 1 : The plasmid is circular DNA. From the data we deduce the following:

- its digestion with the XhoI enzyme linearizes it so there is only one XhoI site on this plasmid

- digestion with SmaI gives two fragments of the same size, therefore two SmaI sites

- digestion with the two enzymes simultaneously, i.e. double digestion (DD) gives 3 fragments, two of which are of the same size and one of 1.5 kb. We deduce that the XhoI site is located at an equal distance between the two SmaI sites



Exercise 2 :

The restriction fragments obtained after digestion with EcoRI and their sizes in nucleotide pairs are:

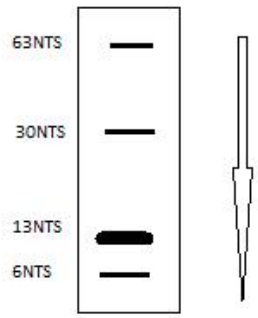
CAATT G **6nts**

AATTC ATCGTAGTAGGCCGTACATGCGTGG **30nts**

AATTCGTAATGTG **13nts**

AATTCTGAAATTTCCCTACATGGGCTACGGCTAGTAGCGCTGATGGGCCATCGCG CTGGGCTGG **63nts**

AATTCGGTCTATT **13nts**



The restriction profile of this fragment is therefore

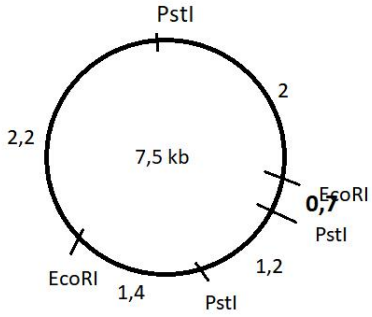
Exercise 3 :

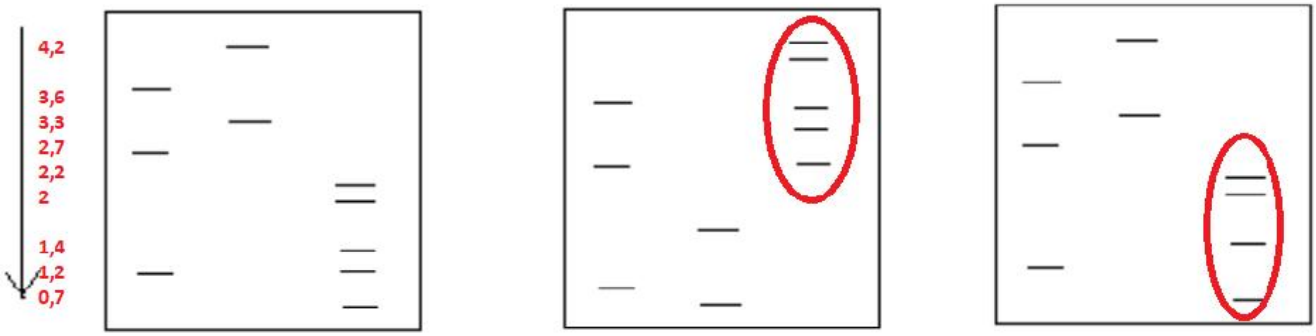
Digestion with EcoRI gives two fragments of respective lengths: 4.2 and 3.3 Kb

Digestion with PstI gives three fragments of respective lengths: 3.6 and 2.7 and 1.2 kb

Double digestion therefore gives five fragments of respective lengths: 2.2 - 2- 1.4- 1.2- 0.7 Kb

Profiles 2 and 3 do not correspond to these results, in terms of number of bands and fragment lengths, so it is the profile which is the correct one

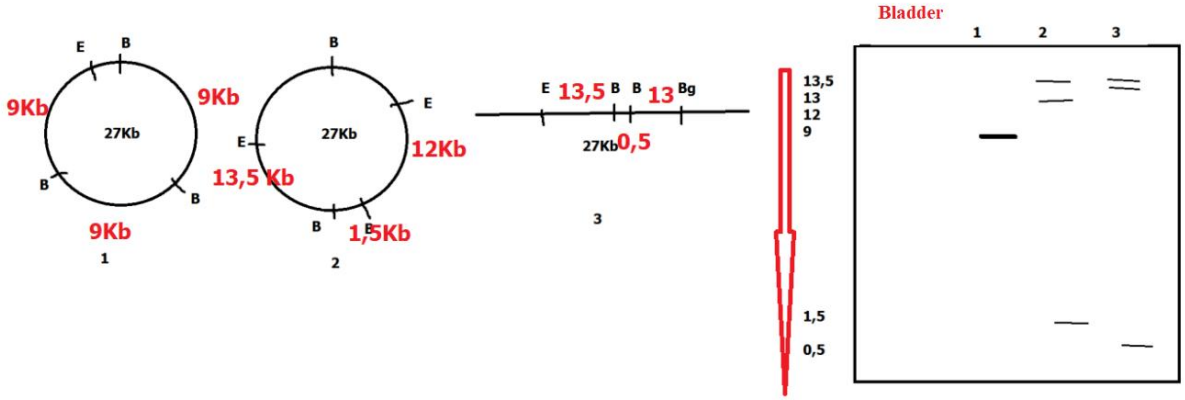




Exercise 4 :

The 3 fragments were digested with enzyme B. If it was enzyme E we would have obtained a 27KB band for DNA1, and for DNA 2 two bands instead of 3.

Since the migration speed is inversely proportional to the size of the fragment, the arrow should point downward in the figure



IV. Molecular hybridizations

1. Reminders

- **Molecular hybridization** is the property that a single-stranded nucleic acid molecule exhibits to associate spontaneously and in a specific and reversible manner with another single-stranded molecule which is complementary to it.

Molecular hybridization is enabled by the hydrogen bonds that can be established by the purine and pyrimidine bases which constitute the two strands of nucleic acids. The bond strength between the two strands of DNA (or RNA) depends on the number of hydrogen bonds that link the two strands together. The presence of GC pairs (3 H bonds) strengthens the bond between the two strands compared to the presence of A-T pairs (2 H bonds).

The principle of molecular hybridization involves detecting a molecule that we call a "target" (DNA, RNA) in a mixture of similar but not identical molecules (different in their sequences), using a probe.

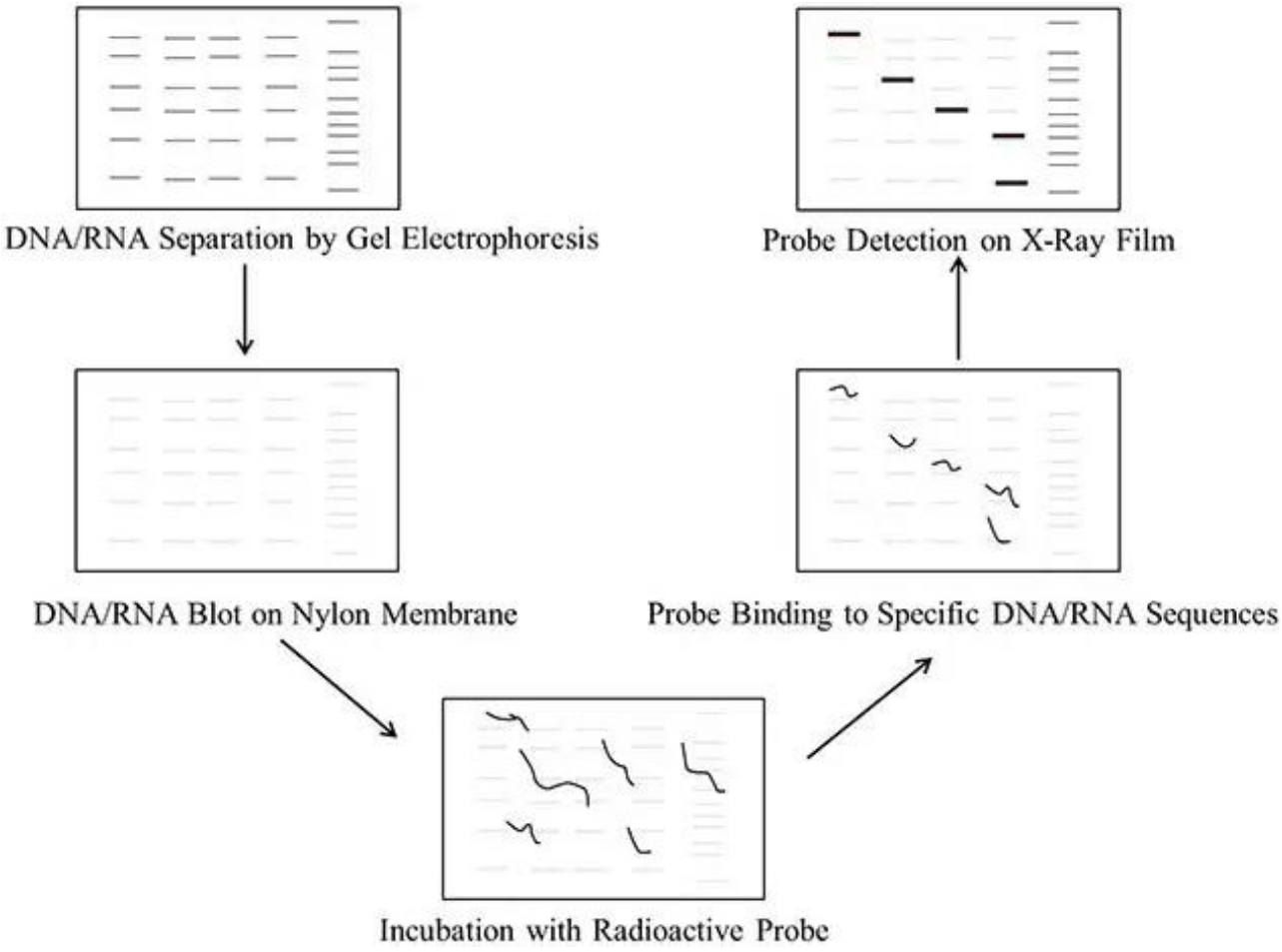
- **The probe** is a single-stranded oligonucleotide complementary and antiparallel to a region of the target nucleic acid. It should not be long to facilitate hybridization but long enough to be unique, i.e. only hybridize with a single target. The probe must be marked to allow identification of the hybridized nucleic acid. Labeling the probe with a radioelement or a fluorochrome makes it possible to visualize and quantify the hybridization.

Hybridization makes it possible, for example, to find out if a gene is present or absent in a cell, if the target molecule is DNA (Southern blot). Or study the expression of a gene under different conditions or detect the expression of a gene in a given tissue, if the target molecule is RNA (Northern blot)

Among the hybridization methods most used in genetic engineering, we can cite the Southern blot and its variant, the Northern blot.

- Initially described by EM Southern in 1975, the **Southern blot technique** consists of specifically detecting DNA fragments digested then transferred to a membrane, thanks to their hybridization with labeled probes. Northern blotting derives from this technique and analyzes RNA.

The Southern Blot and Northern Blot protocols are similar and begin with the electrophoretic separation of nucleic acid fragments on a gel, which are then transferred to a nitrocellulose membrane where they are immobilized. This allows the probes to bind to the immobilized target. Then, molecules of interest can be visualized with various methods.



Classic protocol of a hybridization method

The use of hybridization methods in medical analysis, the study of polymorphism or the study of gene expression, is discussed in the exercises below.

2. Applications

Exercise 1 : A genetic disease is caused by a recessive mutation in a gene of which here is the partial sequence:

…….GGTATTGGTATCCCGTAACCGATCTCG……..

A point mutation in this gene is responsible for the appearance of a BamHI site (G/GATCC)

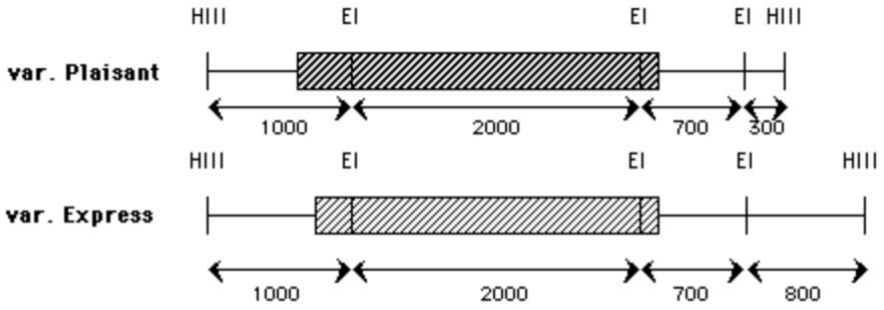
1- Give the sequence of the mutated allele and identify the mutation
2- Represent the restriction profile of sick individuals

Exercise 2 : The NPQ gene was isolated in two varieties of Barley Hordeum vulgare: var.Plaisant and var.Express, in a genomic DNA fragment limited by HindIII sites

Both fragments have been sequenced. They present 92% homology and 100% homology over a region of 2000bp.

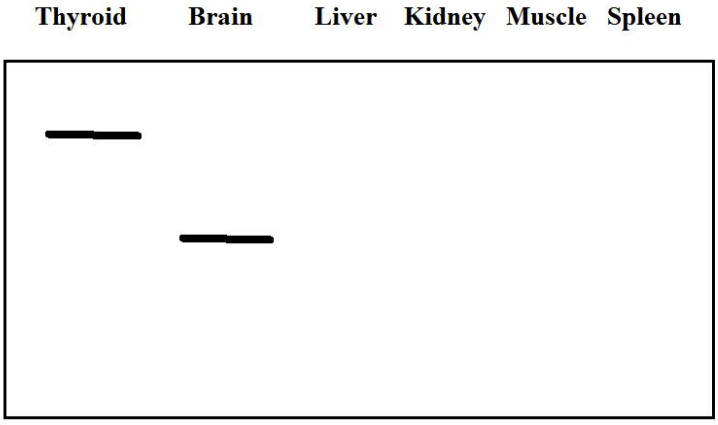
During their storage, the technician neglected to label the tubes containing the two subcloned samples.

- Propose an analysis method that can distinguish them based on the restriction maps of the two genes

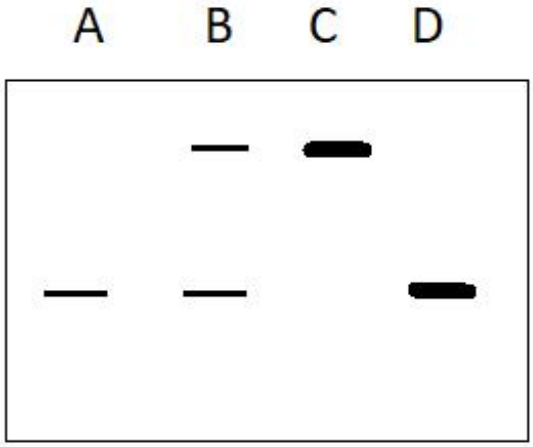


Exercise 3 : The RNAs extracted from the cytoplasm of different tissues are analyzed using a probe complementary to part of the 1st exon of the calcitonin gene.

Indicate the analysis method used and interpret the results of the experiment based on the gel obtained, and specify its purpose.



Exercise 4 : The following gel was obtained by Southern blot analysis in a family with familial hemophilia A. Knowing that individual A is the index individual, specify the sex, genotype and phenotype of the different members of this family



3. Corrected exercises

Exercise 1 :

3- The sequence of the mutated allele

…….GGTATT **GGATCC** CGTAACCGATCTCG…….. deletion of T

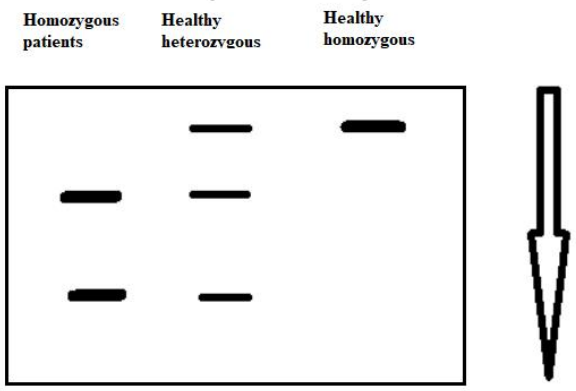
Or

……...GGTATTG **GGATCC** CGTAACCGATCTCG…….. substitution of T in G

4- The restriction profile of sick individuals

Recessive mutation so patients are homozygous for the mutated allele which contains the BamHI site

The molecular diagnosis is carried out by BamHI digestion then electrophoresis and hybridization with a specific probe, which can hybridize specifically with the two alleles.



Restriction profile

Exercise 2 :

We propose to carry out a Southern blot, that is to say digest the two samples with HIII and carry out a hybridization with a probe complementary to the homologous part of the two variants

The presence of a band at 4000 bp will indicate the var Plaisant, and a band at 4500 bp the var Express

Exercise 3 :

The samples analyzed are RNA so it is a Northern blot

Its goal is to identify tissues that express the calcitonin gene.

The results indicate that only the thyroid and the brain express the gene (transcribed).

The difference in size between the two mRNAs indicates that there is alternative splicing (allows us to have different RNAs including different proteins from the same gene).

Exercise 4:

Individual	Sex	Genotype	Ghenotype
A	male	XhY hemizygous	haemophiliac
B	female	XHXh heterozyous	Haemophiliac or healthy
C	female	XHXH homozygous	healthy
D	female	XhXh homozygous	haemophiliac

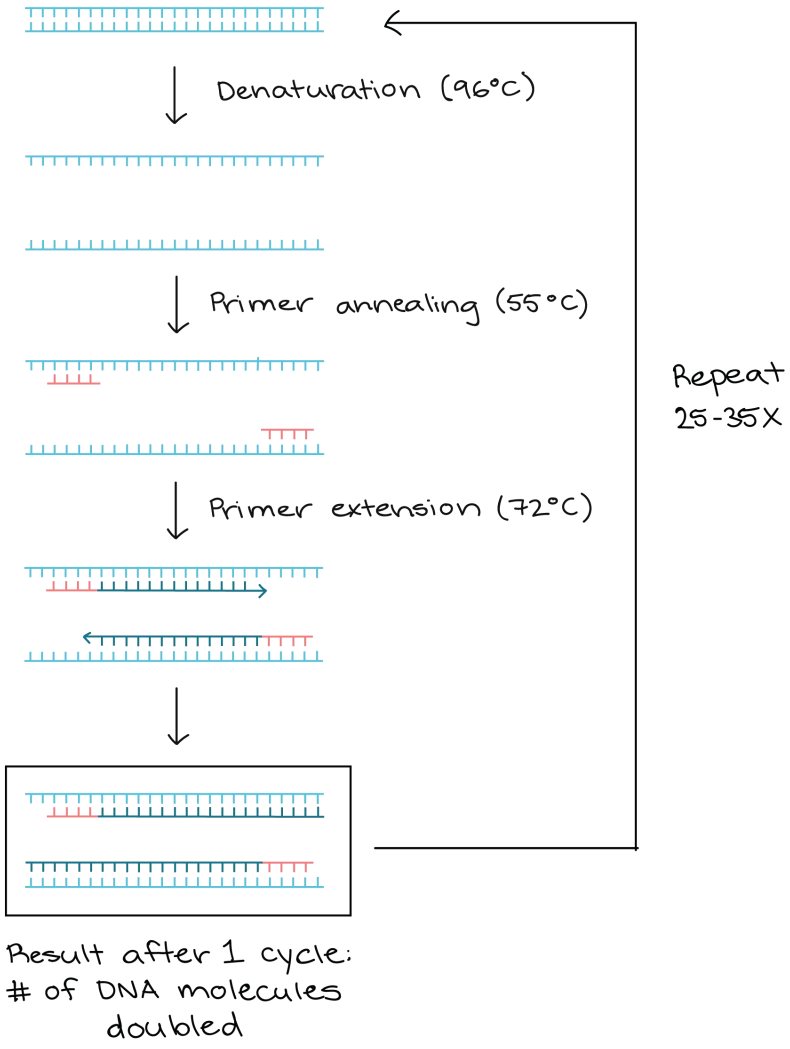
V. Polymerase Chain Reaction PCR

1. Reminders

PCR (Polymerization Chain Reaction) is a biochemical DNA synthesis reaction carried out *in vitro* and repeatedly, which makes it possible to amplify in large quantities a DNA fragment from a DNA template which contains this fragment.

This technique, invented by Kary Mullis in 1980, earned the scientist a Nobel Prize in Chemistry in 1993.

PCR is carried out in a device called a thermocycler, where several synthesis cycles take place in three stages: Denaturation of the double-stranded DNA by heat to open the double helix, hybridization with the primers, elongation of the primer by the DNA polymerase. Each step takes place at appropriate temperatures.



The steps of a PCR cycle

The reaction medium must therefore contain: the template DNA, the dNTPs in equal and sufficient quantity, a pair of primers and a DNA polymerase.

The DNA polymerase used must be heat-resistant; Taq polymerase is used, a polymerase produced by thermophilic bacteria *Themophilus acquaticus*

PCR mimics the replication mechanism, and therefore synthesis takes place in the 5'3' direction in a complementary and anti-parallel manner to the template strand. Each primer is complementary and antiparallel to the 3' end of the template strand. The primers therefore also serve to delimit the region of DNA to be amplified. The choice of primers is crucial to amplify the correct fragments. The primers should not be too long to facilitate hybridization, but must be long enough to be specific to the ends of the fragments that we want to amplify.

The product of a PCR reaction is a fragment of a determined size. The size of the fragment depends on the distance separating the primers used to carry out the amplification. To visualize the result of a PCR experiment, it is therefore sufficient to migrate the reaction mixture on an agarose gel and look for the presence of a fragment corresponding to the expected size.

PCR is particularly useful when genetic material is available in low quantity or in poor condition. It finds many applications in:

- cloning and sequencing genetic
- the diagnosis of genetic diseases but also viral infections (AIDS , Hepatitis C , SARS), bacterial (tuberculosis) or parasitic (toxoplasmosis), and cancers
- mutation detection
- detection of GMOs
- forensic identification of individuals
- determining parentage
- directed mutagenesis (using mutated primers)
- DNA marking (particularly for fundamental research)
- the study of fossils .

The choice of primer, their design and their crucial role in the PCR protocol are discussed in the exercises below.

2. Applications

Exercise 1 : Consider the gene of which part of the sequence is given below and which we wish to amplify:

5'ATACGGGATCCGAGCTCTCGATCGTCTGCAGC….TCCAGACTAAGCTATCCGAG TGC3'

1- What method do you suggest?
2- Indicate the sequences of the forward and reverse primers?
3- How could you optimize your primers?
4- What are the molecular tools necessary to carry out this technique?

Exercise 2 : We wish to amplify by PCR part of the sequence (given below) of the corn ß-6 tubulin gene.

1 ttttaagtta ctgtgtgctt gttgcaggat ctgtaactaa ttcctatgcg attctcttgt
61 ttgtagggcg aagatgaggg agatcctgca catccaggga gggcaatgtg gcaaccagat
121 tggcgccaag ttctgggagg tggtgtgcga tgaacatggc attgacccta cggggcggta
181 cactggcaat tccgaccttc agttggagcg tgttaatgtc tactacaatg aagcctcctg
241 cggacgcttt gttccccgcg ctgttctcat ggatcttgag cctgggacaa tggacagtgt
301 ccggaccgga ccctatgggc agatcttccg ccctgacaac tttgtgtttg ggcaatctgg
361 tgctggtaac aattgggcta agggccacta caccgagggt gctgagctca ttgactctgt
421 tctggatgtt gtgaggaagg aagctgagaa ctgtgactgc ttgcaaggat tccaagtatg
481 ccactccctt ggtggtggta ctggatctgg tatgggtacg ctgttgatct caaagatcag
541 ggaagagtac cctgaccgca tgatgcttac attctcagtt ttcccctcac cgaaagtatc
601 tgataccgtg gttgagccat acaatgccac tctttctgtc caccagttgg tcgagaatgc
661 tgatgagtgc atggttctcg ataacgaagc cctctatgac atctgcttca ggactcttaa
721 gctgaccacc cctagctttg gtgatctgaa ccatttgatc tctgcaacca tgagtggagt
781 cacctgctgc ctaaggttcc ctggtcagct gaactccgac ctcaggaagc tggcagtgaa
841 cctgatcccc ttccccgtc tccacttctt catggtcggc ttcgcgccgc tgacgtcccg
901 tggctcccag cagtaccggg ccctcacagt ccccgagctc acgcagcaga tgtgggatgc
961 caagaacatg atgtgtgccg ctgaccctcg ccatgggcgt tacctcaccg cctcggccat
1021 gttccgcggg aagatgagca ccaaggaggt tgacgagcag atgatcaacg tccagaacaa
1081 gaactcgtcc tacttcgtgg agtggatccc caacaacgtc aagtccagcg tgtgcgacat
1141 cccgcccagg ggcctgtcca tggcgtccac cttcatcggc aactcgacct ccatccagga
1201 gatgttccgg agggtgagcg agcagttcac tgccatgttc aggaggaagg ctttcttgca
1261 ctggtacacg ggcgagggca tggacgagat ggagttcacc gaggccgaga gcaacatgaa
1321 cgacctcgtg tcggagtacc agcagtacca ggacgcgact gccgacgagg aggagtacga
1381 ggacgaggag gaggtgcagg ccgatgacat gtgaggggag ggctgttatc gtgtgaagcc
1441 ttgtggtccc tagggcaagc ggacctcgat gagttcggtg ttcccttcg tgttgttgcc
1501 atctttctac tgctagcgta cccaccctcg tggcccattc cgtcgctgtt gacgtatgta
1561 tttttcttgt gctatggaac cttgcttttg gtacggtact atcctgctag tatgcttggc
1621 gtttgaggtt cctggcgtga atttaagcct tccgtatgca gtgattggag ttggagaccg
1681 gctgcttcgt ccaggcgaag caattgacag cgacgtgcta tacactcaag

1- Determine the sequence of two primers used to specifically amplify approximately 1 kilobase of the tubulin gene sequence extending from nt 450 to nt 1450

2- What are the primer sequences necessary to be able to clone this fragment between the BamHI and EcoRI sites?

Exercise 3 : Give the sequences (only the ends) of the template DNA amplified using these primers

Sense CCTATGGTACGGCTAAATCC Anti sense GGATCCGTATCCGCATACG

3. Corrected exercises

Exercise 1 :

1- Two amplification methods are possible:

- *In vivo* amplification is cloning
- *In vitro* amplification is PCR

2- Primer sequences (around twenty nts) for PCR

- Direction: same direction and same sequence as the 5' end of one of the two template strands. This primer hybridizes with the 3' end of the strand complementary to the one shown (PCR exploits the replication mechanism and therefore polymerization takes place in the 5' 3' direction, and in a complementary and antiparallel manner to the template strand):
  5' ATACGGGATCCGAGCTTCG 3'
- Antisense: antiparallel and complementary oligonucleotide at the 3' end of the template strand represented:
  5'GCACTCGGATAGCTTAGTCT 3'

3- The optimal primers hybridize correctly and in a specific and stable manner to the ends of the DNA fragment to be amplified. So we must :

- Tm (melting temperature) for the two comparable primers to allow optimal hybridization with the template strand under the same temperature conditions
- CG nts content between 40% and 60%
- Must not be complementary to each other

4- The molecular tools necessary for carrying out PCR are:

- DNA template
- A pair of primers
- Thermostable DNA polymerase (Taq polymerase)
- dNTPs (the 4 nucleotides in equal quantity)

Exercise 2 :

1- Primers of around twenty nucleotides

Direction 5'ACTGTGACTGCTTGCAAGGA 3' the same as the end 5'

Anti sense 5' GGGACCACAAGGCTTCACAC 3' complementary and anti parallel to the 3' end

PCR is replication-driven, so each strand is replicated in the 5'3' direction, and therefore the primer pair must be complementary and antiparallel to the 3' end of each strand

2- Primers necessary to be able to clone this fragment between the BamHI and EcoRI sites

Direction 5' **GGATCC** ACTGTGACTGCTTGCAAGGA 3'

Anti sense 5' **GAATTC** GGGACCACAAGGCTTCACAC 3'

We introduce the sequences of the restriction sites of the two enzymes into the primers at the 5' end of each primer, this does not hinder the hybridization and it allows an additional sequence to be added to the newly formed strand compared to the template DNA. , including the desired restriction site. The amplified sequence will therefore be flanked by the two sites.

Exercise 3 :

The primer has the same sequence and direction as the 5' end of the strand used for the design of the primers

The antisense primer is, however, complementary to the 3' end of this same strand.

The second strand being complementary and anti-parallel, the sequence sought is:

5' CCTATGGTACGGCTAAATCC ..... CGTATGCGGATAACGGATCC 3'

3' GGATACCATGCCGATTTAGG ..... GCATACGCCTATTGCCTAGG 5'

VI. Sequencing

1. Reminders

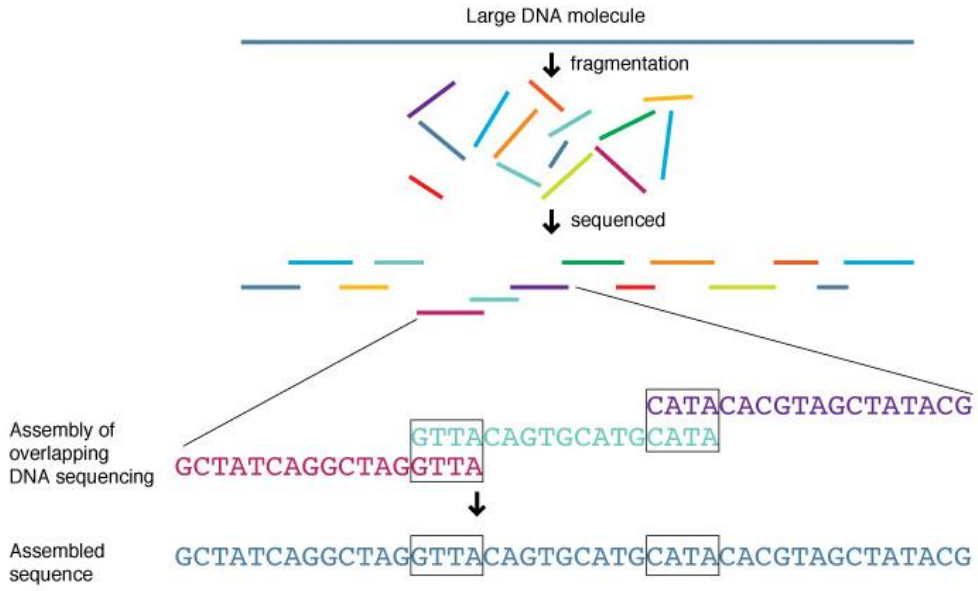
Since the end of the 1970s and the advent of molecular biology techniques, it has been possible to sequence a strand of DNA, that is to say to read the sequence, or sequence, of the nucleotides constituting this molecule. . This actually comes down to determining the sequence of bases, the only variable part of nucleotides.

However, current techniques only make it possible to read, at each sequencing operation, a thousand bases at most. However, the "sequenceable" part of the human genome includes 2.9 billion base pairs! It is therefore impossible to read an entire genome at once. It is, in any case, impossible to manipulate DNA molecules of several tens or even hundreds of millions of bases (the order of magnitude of those which constitute human chromosomes).

We speak of the "sequenceable part" of the human genome (2.9 Gb, for a total of 3.2 Gb). Indeed, it is not technically possible to determine the sequence of certain regions almost exclusively made up of repeated sequences, such as centromeres, telomeres or the short arms of certain chromosomes. There are two reasons for this: firstly, it is difficult to isolate DNA fragments of suitable size from these regions; on the other hand, it is not possible to reconstruct the complete sequence from pieces of practically identical sequences. As a result, only the sequence of the so-called euchromatic part of the genome can be effectively determined.

The basic principle, in any sequencing of a genome, therefore consists of randomly fragmenting this genome – or large pieces of DNA derived from the genome – to obtain pieces of DNA of a few thousand base pairs, easy To manipulate. The ends of many of these small fragments are then sequenced. The complete sequence of the genome – or large piece of genome – is then reconstructed from these unit sequences, or reads, based on the overlaps between the sequences (if the sequences are overlapping, it is because the DNA fragments from which they derive have part of their length in common; the break being random, the DNA molecules in the sample are not all broken in the same places).

Rodger Staden, invents the first DNA sequencing "software". In 1982, Sanger used it to assemble all 48,502 bp of the genome of the bacteriophage Lambda.



Overlapping sequences obtained by genome fragmentation

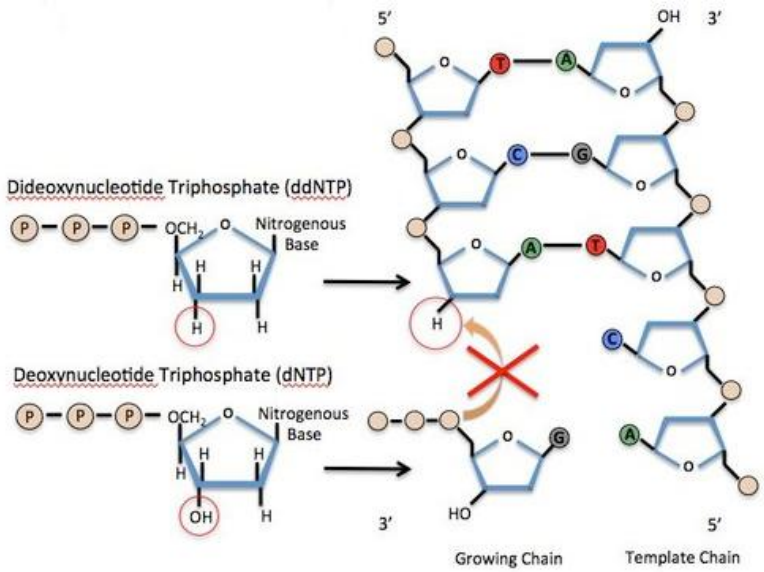
Reading DNA sequences makes it possible to study the biological information contained therein. Given the uniqueness and specificity of the DNA structure in each individual, the DNA sequence allows numerous applications, particularly in the field of medicine, for the diagnosis and treatment of numerous human diseases (examples : cancers, infectious diseases, hereditary diseases, etc.). It also provides information on the genome (structure, function, evolution) and allows the study of genome variations (biallelic polymorphisms, insertions, deletions, insertions/deletions (also called indels), gene rearrangement, copy number variation of genes, duplication (or more) Sequencing is currently used for the search for genetic variants associated with a pathology (for example, diabetes), genome methylation analysis (epigenetic and methylome studies), microbiological analysis. (species identification, taxonomy, epidemiological studies, genotyping for prognostic and/or therapeutic purposes)....in addition to applications in forensic medicine and anthropology.

The first two DNA sequencing techniques, that of Maxam-Gilbert and that of Sanger, were described in 1977. The Maxam-Gilbert technique is practically abandoned today, while that of Sanger remains to this day the main one. sequencing method used in laboratories.

Its principle is as follows: firstly, it is necessary to amplify the target DNA by PCR, then to denature it in order to obtain single-stranded DNA. Using a primer specific and complementary to the strand studied (sense or antisense), a DNA polymerase then carries out

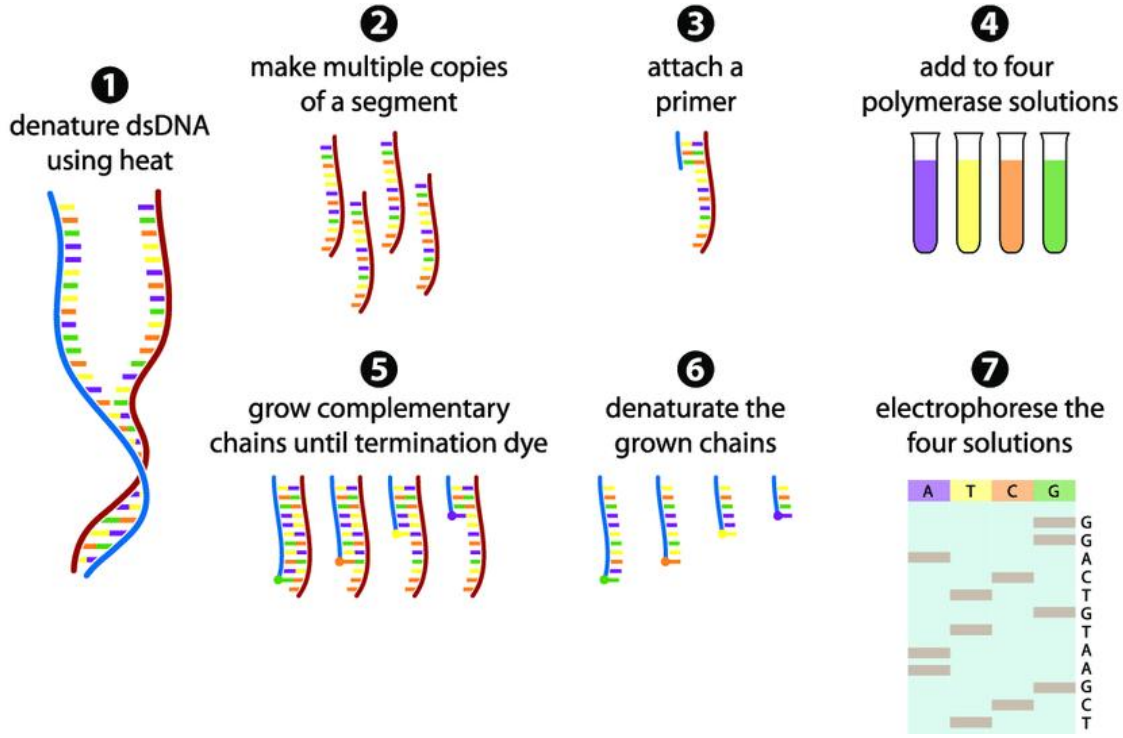
the synthesis of complementary DNA from this primer. The so-called universal primer is complementary to the known 3' end upstream of the fragment to be sequenced (on the vector).

From the 5′ end to the 3′ end, DNA polymerase adds complementary deoxyribonucleotide triphosphates (dNTPs) and randomly and inconsistently dideoxyribonucleotide triphosphates (ddNTPs), for example a ddGTP will sometimes be added instead of 'a dGTP. As the reaction takes place in a single tube, the ddNTPs (ddATP, ddGTP, ddCTP and ddTTP) are labeled using different fluorophores for each ddNTP ("four-color" fluorophores). When a ddNTP is incorporated in place of a dNTP, the DNA polymerase can no longer continue its polymerization. The extension reaction stops (in fact, the dideoxynucleotide does not have a 3′-hydroxyl group essential for the polymerization reaction of the enzyme).



The difference between dNTP and ddNTP

Statistically, during the reaction, for each "base" of the target DNA, at least once, a complementary ddNTP will be incorporated in place of a dNTP. Therefore, at the end of the reaction we will obtain fragments of different sizes. Analysis of the reaction is then carried out by electrophoresis.



Sanger method sequencing protocol

The Sanger sequencing method in 7 steps. (1) The dsDNA fragment is denatured into two ssDNA fragments. (2) A fragment of ssDNA is multiplied into millions of copies. (3) A primer that corresponds to one end of the fragment is attached. (4) The fragments are added to four polymerase solutions. Each solution contains the four types of bases but only one type of terminating nucleotide. (5) The chain grows until a termination nucleotide is randomly added. (6) The resulting dsDNA fragments are denatured to obtain a series of ssDNA of various lengths. (7) The fragments are separated by electrophoresis and the sequence is read.

The principle of sequencing by the Sanger method, the interpretation as well as the analysis of the results on the migration profile of truncated fragments, are discussed in the exercises below.

2. Applications

Exercise 1 : You have a strand of DNA to sequence (template), a primer, a truncated DNA polymerase (Klenow fragment), the four 2'-deoxyribonucleotides (dNTPs) and a set of different 2',3'-dideoxyribo-nucleotides (ddNTPs).

Strand to sequence 5' CATTCTAGCCATGCA 3'

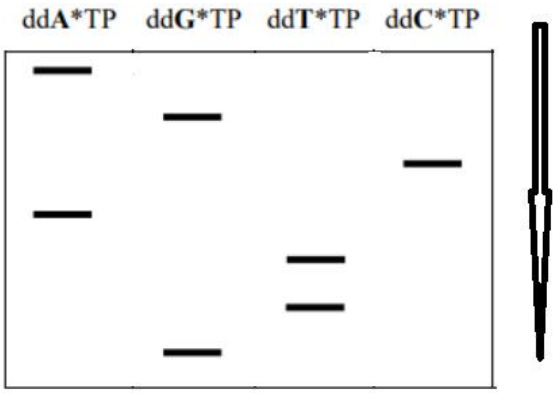
1- Indicate the composition of the different reaction media and, for each medium, the type and size of the newly synthesized fragments.
2- On a gel, represent the size of the fragments newly synthesized in each reaction medium, and indicate the direction of migration of the fragments
3- Write the sequences of the synthesized strand and the sequenced strand, indicating the reading direction of the established sequences.

Exercise 2 : Consider the following gel obtained during a sequencing experiment.

1- What do the tapes contain?

2- Explain the principle

3- Interpret the results



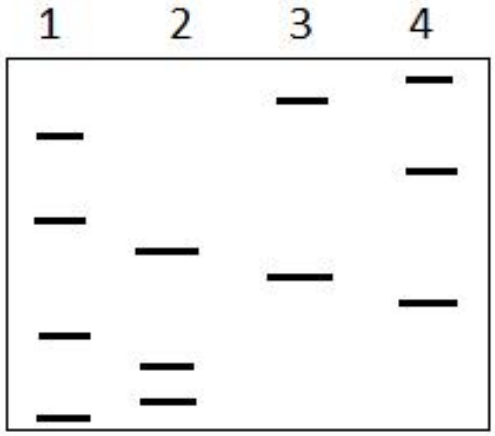
Exercise 3 : The following fragments were obtained during sequencing from 4 different reaction media

5' A3'
5' AT3'
5' ATT 3'
5' ATTA 3'
5' ATTAG 3'
5' ATTAGC 3'
5' ATTAGCT 3'
5' ATTAGCTA 3'
5' ATTAGCTAG 3'

5' ATTAGCTAGA 3'

5' ATTAGCTAGAC3'

5' ATTAGCTAGACG 3'



1- Based on the migration profile above and the truncated fragments, specify for each track the ddNTP contained in the reaction medium

3- Give the sequence of the sequenced fragment

3. Corrected exercises

Exercise 1 :

1- Each medium contains: template DNA + 4dNTPs + a universal primer + **1 ddNTP**

Medium 1: template DNA + 4dNTPs + a universal primer + **ddTTP**

Medium 2: template DNA + 4dNTPs + a universal primer + **ddATP**

Medium 3: template DNA + 4dNTPs + a universal primer + **ddCTP**

Medium 4: template DNA + 4dNTPs + a universal primer + **ddGTP**

Middle 1 contains the truncated fragments:

5' **T** 3'

5'TGCAT3'

5'TGCATGGC **T** 3'

5'TGCATGGCTAGAA **T** 3'

Middle 2:

5'TGC **A** 3'

5'TGCATGGCT **A** 3'

5'TGCATGGCTAG **A** 3'

5'TGCATGGCTAGA **A** 3'

Middle 3:

5'TG **C** 3'

5'TGCATGG **C** 3'

Middle 4:

5'T **G** 3'

5'TGCAT **G** 3'

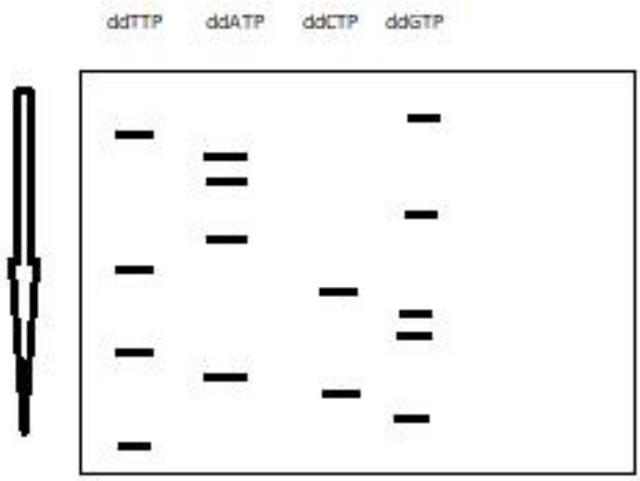
5'TGCATG **G** 3'

5'TGCATGGCTA **G** 3'

5'TGCATGGCTAGAAT **G** 3'

Each time a ddNTP is incorporated, the synthesis stops.

2- The gel obtained is:



3- The sequence of the synthesized strand is as follows: 5'TGCATGGCTAGAATG 3'

The sequence of the sequenced strand is that given in statement

The reading direction on the gel is indicated by the arrow

Exercise 2 :

1- The tapes contain the truncated fragments

2- Each truncated fragment ends with a known ddNTP, which makes it possible to reconstruct the sequence of the synthesized fragment, starting from the smallest fragment to the largest (see the migration gel).

3- DNA synthesis is done in the 5'3' direction, the smallest fragment being the one which migrates the fastest, so we can determine the sequence of the synthesized strand starting from the bottom to the top of the gel (from the most small fragment to the largest, each truncated fragment differs from the other by a new ddNTP which is incorporated into the chain in formation)

And therefore the synthesized fragment is 5' GTTACGA 3'

The sequenced fragment is cocomplementary and antiparal to it, so its sequence is as follows: 5'TCGTAAC 3'

Exercise 3 :

1- Based on the gel, we can classify the truncated fragments according to their size, and based on the last nucleotide, we can determine the ddNTP contained in the medium. Combining these data, we deduce that the ddNTPs contained in the reaction media are as follows:

(1) ddATP

(2) ddTTP

(3) ddCTP

(4) ddGTP

3- The sequence of the sequenced fragment 5'CGTCTAGCTAAT 3' (complementary and antiparallel to that which we can deduce from the gel, i.e. the sequence of the synthesized fragment)

References

Brown T. (2001) . Southern blotting. Curr Protoc Immunol; Chapter 10:Unit 10.6A.

Cariani E . , C. Bréchot (1988) . Detection of DNA sequences by Southern blot . Ric Clin Lab.; 18(2-3):161-70.

Chang M.M. (2022) . Plasmid-to-plasmid Southern blot analysis validates the presence of nucleotide binding site (nbs) sequences in cloned plasmids. Biochem Mol Biol Educ; 50(4):373-380.

Gebbie L . (2014) . Genomic Southern blot analysis. Methods Mol Biol; 1099:159-77.

Giani A.M. , GRGallo, L.Gianfranceschi, G.Formenti (2020) . Long walk to genomics: History and current approaches to genome sequencing and assembly. Computational and Structural Biotechnology Journal ; Vol 18 : 9-19

Haras D. , J. P. Amoros (1994) . Polymerase chain reaction, cold probes and clinical diagnosis. Health ;4 (1):43-52.

Jarcho J.(2001) . Restriction fragment length polymorphism analysis. Curr Protoc Hum Genet. Chapter 2:Unit 2.7

Krumlauf R. (1996) . Northern blot analysis. Methods Mol Biol; 58:113-28.

Lamoril J. , N. Ameziane , J.-C. Deybach , P. Bouiz egarene , M. Bogard (2008) . DNA sequencing techniques: a revolution in progress. First part ; 23(5): 260–279.

Lessard J.C (2013) .Molecular Cloning. Methods in Enzymology; 529: 85-98

Perl A. , J. Fehér (1990) . Polymerase chain reaction in clinical diagnosis. Orv Hetil ;131(13): 671-5.

Pingoud A. , J. Alves , R. Geiger (1993) . Enzyme restriction. Methods Mol Biol; 16:107-200.

Porchet N. , J.P. Aubert (2000) . Northern blot analysis of large mRNAs. Methods Mol Biol.; 125:305-12.

Southern E . (2006) . Southern blotting. Nat Protoc;1(2): 518-25.