Biotechnological Applications of Recombinant DNA

Dr. Selma Hamimed
Department of Molecular and Cellular Biology
Faculty of Natural Sciences and Life
University of Jijel
Email: selma.hamimed@univ-jijel.dz[1]

1.0 September 2024

1 file:///C:/Users/Selma/Downloads/Tle_site1-3/Cours%20%C3%A0%20reproduire/Cours%20version%20fran%C3%A7ais%20(1)/Cours%20%C3%A0%20reproduire_gen_1

Attribution - Pas d'Utilisation Commerciale : http://creativecommons.org/licenses/by-nc/4.0/fr/

Table des matières

I DNA sequencing and PCR as diagnostic tools

1. Objectives

- Understand the current state-of-the-art methods of DNA sequencing and PCR techniques.
- Analyze the applications of DNA sequencing and PCR in various diagnostic fields, such as medical diagnostics, forensic science, and environmental monitoring.
- Compare and contrast different DNA sequencing platforms (e.g., Sanger sequencing, next-generation sequencing) and PCR variants (e.g., real-time PCR, digital PCR), evaluating their strengths and limitations in diagnostic contexts.

2. What is PCR?

Polymerase Chain Reaction (PCR) is a revolutionary technique in modern molecular biology, developed by Nobel laureate Kary Mullis in 1984. This *in vitro* method allows for the precise amplification of a target DNA molecule, making it easier to handle and examine using routine molecular biological methods. PCR has significantly contributed to the advancement of biological sciences since its inception, with the first PCR machine introduced to the market in 1988. Its wide-ranging applications have led to the development of numerous PCR variants over the past few decades, playing a crucial role in projects such as the Human Genome Project.

The PCR process involves three main steps: denaturation, annealing, and polymerization. First, the target DNA strands are separated through denaturation. Then, oligonucleotide primers hybridize (anneal) to the single-stranded templates. Finally, a thermostable DNA polymerase, typically derived from thermophilic microorganisms, extends these primers by adding nucleotides (dNTPs) complementary to the template sequence. This cycle is repeated multiple times, with each cycle potentially doubling the amount of target DNA.

PCR's ability to make millions or billions of copies of a specific DNA region has made it an invaluable tool in various fields, including biotechnology, genomics, diagnostics, and systematics. Researchers can use PCR to amplify genes for functional studies, forensic scientists can analyze genetic markers from crime scenes, and the amplified DNA can be used for sequencing, gel electrophoresis, or cloning into plasmids for further experiments. The versatility and power of PCR have cemented its position as a cornerstone technique in modern molecular biology and biotechnology[18]*.

***Méthode* : PCR requirements**

1. Primer: Complementary to the **DNA** of interest (**High stability**).
2. Nucleotides (Deoxyribonucleotide Triphosphate).
3. DNA polymerase (*Thermus aquaticus*)
4. Heat for Denaturation
5. Cooling for Re-annealing

2.1. Steps of PCR

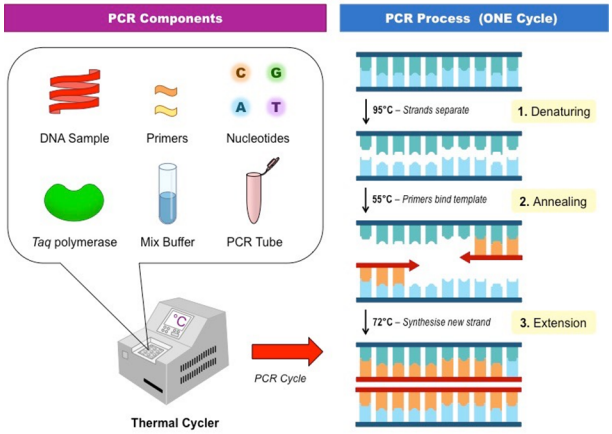
As shown in the figure below, the key ingredients of a PCR reaction are Taq polymerase, primers, template DNA, and nucleotides (DNA building blocks). The ingredients are assembled in a tube, along with cofactors needed by the enzyme, and are put through repeated cycles of heating and cooling that allow DNA to be synthesized. The basic steps are:

1. Denaturation (96 °C): Heat the reaction strongly to separate, or denature, the DNA strands. This provides single-stranded template for the next step.

2. Annealing (55-65°C): Cool the reaction so the primers can bind to their complementary sequences on the single-stranded template DNA.

3. Extension (72 °C): Raise the reaction temperatures so Taq polymerase extends the primers, synthesizing new strands of DNA.

This cycle repeats [25-35] times in a typical PCR reaction, which generally takes [2-4] hours, depending on the length of the DNA region being copied. If the reaction is efficient (works well), the target region can go from just one or a few copies to billions.



Steps of PCR

(cf. PCR - Polymerase Chain Reaction)

3. Quantitative PCR

Quantitative PCR (qPCR), also known as real-time PCR or quantitative real-time PCR, is an advanced variant of the traditional PCR technique. This method combines the amplification of a target DNA sequence with simultaneous quantification of the DNA concentration in the reaction. By linking the amplified product to fluorescence intensity through a fluorescent reporter molecule, qPCR allows researchers to observe the amount of PCR product in real-time as it doubles with each cycle, generating a characteristic amplification curve.

The power of qPCR lies in its ability to calculate the initial template concentration, making it a versatile analytical tool for various applications. These include evaluating DNA copy number, measuring viral load, detecting single nucleotide polymorphisms (SNPs), and performing allelic discrimination. When combined with reverse transcription (RT-qPCR), it becomes an invaluable method for quantifying mRNA expression, serving as the gold standard for confirming microarray gene expression data.

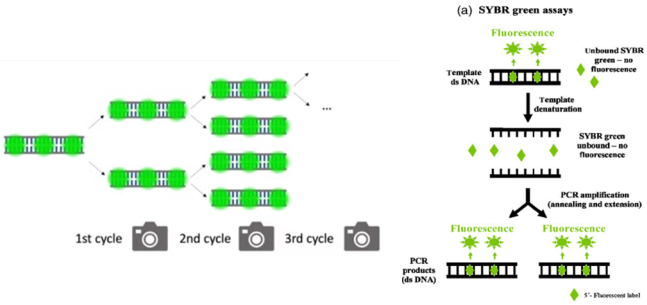
qPCR can be performed in two main ways: endpoint QPCR and real-time QPCR. In endpoint QPCR, fluorescence data are collected after the amplification reaction is complete, typically after 30-40 cycles, and this final fluorescence is used to back-calculate the initial template quantity. In contrast, real-time QPCR measures the fluorescent signal

while the amplification is still in progress, providing a more dynamic view of the reaction kinetics *19*[19]*. The qPCR is mainly performed using one of these two essays SYBR Green assay and TaqMan probe.

Quantitative PCR (cf. qpcrGuide-Stratagene.pdf)

3.1. SYBR Green assay

DNA binding dyes such as SYBR Green are cost effective and easy to use, especially for researchers who are new to using QPCR techniques. These same factors make SYBR Green a common choice for optimizing QPCR reactions. When free in solution, SYBR Green displays relatively low fluorescence, but when bound to double-stranded DNA its fluorescence increases by over 1000-fold. The more double-stranded DNA that is present, the more binding sites there are for the dye, so fluorescence increases proportionately to DNA concentration. This property of the dye provides the mechanism that allows it to be used to track the accumulation of PCR product. As the target is amplified, the increasing concentration of double stranded DNA in the solution can be directly measured by the increase in fluorescence signal (Figure below). Compared to probe-based methods, SYBR Green assays are relatively easy to design and optimize. All that is necessary is to design a set of primers, optimize the amplification efficiency and specificity, and then run the PCR reaction in the presence of the dye. One limitation of assays based on DNA-binding dye chemistry is the inherent non-specificity of this method. SYBR Green will increase in fluorescence when bound to any double-stranded DNA (dsDNA). Therefore, the reaction specificity is determined solely by the primers.

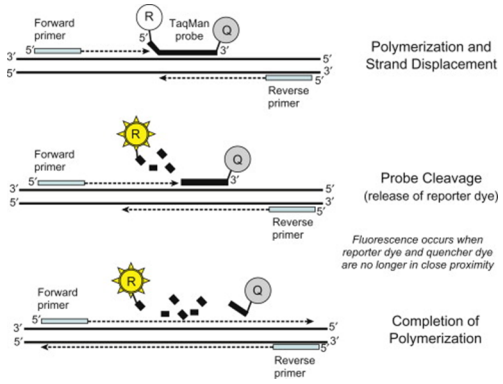


SYBR Green assay

3.2. TaqMan qPCR probes

TaqMan qPCR probes are the most widely used and published detection chemistry for QPCR applications. In addition to the PCR primers, this chemistry includes a third oligonucleotide in the reaction known as the probe. A fluorescent dye, typically FAM, is attached to the 5´ end of the probe and a quencher, historically TAMRA, is attached at the 3´ end. As long as the two molecules (reporter and quencher) are maintained in close proximity, the fluorescence from the reporter is quenched and no fluorescence is detected at the reporter dye's emission wavelength. The probe is designed to anneal to one strand of the target sequence just slightly downstream of one of the primers. As the polymerase extends that primer, it will encounter the 5´ end of the probe. Taq DNA polymerase has 5´–3´ nuclease activity, so when Taq DNA polymerase encounters the probe it displaces and degrades the 5´ end, releasing free reporter dye into solution. Following the separation of reporter dye and quencher, fluorescence can be detected from the reporter dye (Figure below).

During gene expression analysis multiple genes can compared and the number of cycles after which the signal intensity reaches the threshold is shown as Ct value. Where Low Ct values go along with high DNA concentrations and high Ct values mean that signals are detected late indicating low DNA concentrations.



TaqMan qPCR probes

4. Applications of PCR

Using PCR, a DNA sequence can be amplified millions or billions of times, producing enough DNA copies to be analyzed using other techniques. For instance, the DNA may be visualized by gel electrophoresis, sent for sequencing, or digested with restriction enzymes and cloned into a plasmid.

PCR is used in many research labs, and it also has practical applications in forensics, genetic testing, and diagnostics. For instance, PCR is used to amplify genes associated with genetic disorders from the DNA of patients (or from fetal DNA, in the case of prenatal testing). PCR can also be used to test for a bacterium or DNA virus in a patient's body: if the pathogen is present, it may be possible to amplify regions of its DNA from a blood or tissue sample.

4.1. Molecular Cloning via PCR

PCR has revolutionized molecular cloning, offering a more efficient alternative to traditional methods. This technique facilitates gene isolation without the need for conventional DNA library insertion into plasmid vectors and subsequent bacterial transformation, screening, and selection processes. The acellular nature of PCR negates the requirement for cellular amplification systems, streamlining the cloning procedure.

The efficacy of PCR-based molecular cloning is contingent upon two critical factors: the selection of an appropriate DNA template and the design of specific primers. This methodology necessitates prior knowledge of the target gene sequence or its flanking regions to enable the synthesis of suitable primer sets for full or partial gene amplification*20*[20]*.

Template DNA selection is crucial and can be approached in two ways:

1. Genomic DNA: This template encompasses the entire genome, including all genes with their exons, introns, and potentially regulatory regions. Amplification from genomic DNA results in the cloning of complete gene sequences.
2. Messenger RNA (mRNA): This approach focuses on coding sequences (transcripts) only. Due to RNA instability, mRNA is converted to complementary DNA (cDNA) via RT-PCR, a PCR variant utilizing reverse transcriptase. The resultant cDNA library serves as the template for subsequent PCR amplification of the gene of interest.

The mRNA-based approach presents several complexities: a) Transcript presence is cell type-, tissue-, or organ-specific due to differential gene expression. b) Gene expression is often regulated by physiological and environmental factors, potentially leading to absence of the target transcript in the cDNA library. c) Transcriptional regulation and alternative splicing may result in different protein isoforms from the same gene, depending on cell type and regulatory conditions.

Despite these challenges, cDNA cloning offers distinct advantages:

1. The nucleotide sequence directly corresponds to the amino acid sequence of the translated protein.
2. cDNA facilitates easier gene expression studies and functional evaluation of the corresponding protein(s) in cellular expression models.

To comprehensively characterize a gene, researchers often perform parallel PCR cloning on both genomic DNA and various cDNA libraries. This approach enables determination of the complete gene sequence, expression profile, and splicing regulation modalities, providing a holistic understanding of the gene's structure and function[20]*.

(cf. Acellular cloning)

4.2. Reverse transcriptase PCR (RT-PCR)

The isolation of messenger RNA (mRNA) and its subsequent conversion to complementary DNA (cDNA) is a pivotal technique in molecular biology, particularly for gene expression studies. This process involves several precise steps that ensure the accurate analysis of gene activity within cells[21]*.

- **RNA Extraction and mRNA Isolation:**

  The process begins with the extraction of total RNA from cells or tissues, followed by the isolation of mRNA. Affinity chromatography, which utilizes the mRNA's 3' poly-A tail, is commonly employed for this purpose. The poly-A tail binds selectively to oligo-dT sequences, allowing for the enrichment of mRNA from the total RNA population.

- **Reverse Transcription:**

  The isolated mRNA then acts as a template for the synthesis of complementary DNA (cDNA). This reaction is catalyzed by reverse transcriptase, an enzyme derived from retroviruses. The reverse transcriptase synthesizes a cDNA strand that corresponds to the mRNA sequence, converting the single-stranded mRNA into a double-stranded cDNA molecule.

- **mRNA Elimination:**

  Once the cDNA is synthesized, the original mRNA template is no longer needed and is subsequently removed. This is typically achieved through methods such as RNase treatment, alkaline hydrolysis, or thermal denaturation, ensuring that only the newly synthesized cDNA remains for further analysis.

- **cDNA Amplification:**

  The single-stranded cDNA is then amplified through polymerase chain reaction (PCR). During PCR, the cDNA is subjected to repeated cycles of heating and cooling in a thermal cycler, allowing for the exponential replication of the target sequences. This amplification step is critical for generating sufficient quantities of cDNA for downstream analysis.

In mammalian systems, including humans, a typical cell expresses between 10,000 and 15,000 genes. The abundance of mRNA transcripts varies greatly, with some genes producing hundreds to thousands of mRNA copies per cell, while others are expressed at very low levels. The resulting transcript profile provides a snapshot of gene expression within the cell and can reveal important insights into the biological state of the organism.

The accuracy of mRNA extraction and cDNA synthesis is paramount for meaningful gene expression analysis. Factors such as RNA integrity, the efficiency of reverse transcription, and potential PCR bias must be carefully controlled to ensure that the transcriptome is faithfully represented. Any errors in these steps can lead to skewed data, potentially obscuring the true gene expression landscape.

The isolation of mRNA and conversion to cDNA, coupled with gene expression profiling, offers a powerful tool for understanding cellular processes. It provides critical insights into both normal physiology and disease states, with applications spanning from basic research to clinical diagnostics. Accurate gene expression analysis allows researchers to probe the molecular underpinnings of health and disease, advancing both scientific knowledge and medical practice*22*[22]*.

(cf. Using Reverse Transcription Polymerase Chain Reaction (RT-PCR) in COVID-19 Testing)

4.3. PCR applied to identification

PCR is highly effective for identifying species, varieties, or individuals through genetic fingerprinting, an application based on our understanding of genome structure. This process involves amplifying specific nucleotide sequences unique to the species, variety, or individual. In eukaryotes, these sequences are abundant and provide a wide range of options for precise and selective identification. Unlike prokaryotes, eukaryotic genomes contain both coding and noncoding sequences. Coding sequences, which correspond to genes, are translated into proteins and are highly conserved among individuals of the same species, reflecting common traits and genetic features. Phenotypic differences between individuals are largely due to allelic variations, though these differences are minimal (about 1 base pair per 1000)*28*[28]*.

When comparing species, the phylogenetic distance between them influences the degree of homology in gene sequences coding for essential functions. As a result, coding sequences are not particularly useful for identification purposes. In contrast, noncoding sequences exhibit high polymorphism both between species and among individuals of the same species. This polymorphism provides a rich array of genetic markers for highly discriminative identification tests. Key markers include minisatellites, also known as variable number of tandem repeats (VNTRs), and microsatellites, or short tandem repeats (STRs). VNTRs consist of repeat sequences 10 to 40 base pairs long, while STRs are made up of shorter repeats, typically 1 to 5 base pairs. The repeated sequences themselves are identical between individuals, but the number of repetitions, and thus the length of VNTRs or STRs, varies widely, allowing for differentiation*27*[27]*.

VNTRs and STRs are abundant in eukaryotic genomes, and their polymorphism can be detected by PCR. Primers designed to hybridize to nonpolymorphic flanking regions of these repeats are used to amplify them, and the resulting PCR products are analyzed via electrophoresis or capillary sequencing. The ability to simultaneously amplify multiple STRs or VNTRs with different primer pairs enhances the discriminatory power of this approach, creating genetic profiles unique to each individual. The sensitivity of PCR also enables the amplification of micro- and minisatellites from minimal amounts of DNA, making DNA fingerprinting a crucial tool in judicial investigations, as well as in identifying individuals, species, or varieties in other contexts. The specificity of identification depends on the choice of markers, and PCR-based protocols are commonly employed for varietal identification*27*[27]*.

Two additional techniques for varietal identification include random amplification of polymorphic DNA (RAPD) and amplification of fragment length polymorphism (AFLP). RAPD uses pairs of short, random primers (about 10 base pairs) that hybridize randomly across the genome. The resulting PCR products produce an amplification profile specific to the variety from which the matrix DNA was derived. AFLP, a more efficient method, begins with the hydrolysis of genomic DNA by one or two restriction endonucleases. Adapters—defined sequences of about 15 nucleotides—are then ligated to the cohesive ends generated by the restriction enzymes. PCR amplification is subsequently performed using primers that hybridize to these adapters. AFLP produces results similar to RAPD but with greater clarity and reproducibility, making it the most reliable method for varietal identification currently available*27*[27]*.

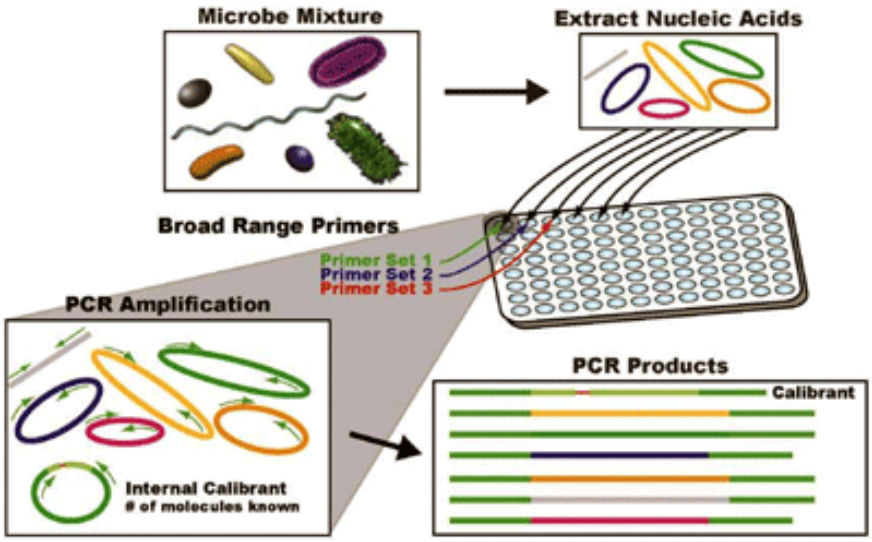


Image 1 Bacterial Identification

4.4. PCR applied to diagnosis

Polymerase Chain Reaction (PCR) has emerged as an indispensable diagnostic tool in modern medicine and molecular biology, offering unprecedented sensitivity and specificity in detecting genetic and infectious diseases. Its versatility and efficiency have revolutionized diagnostic approaches across various medical fields.

In the realm of genetic disease diagnostics, PCR enables the amplification of specific gene segments associated with hereditary disorders. This amplification facilitates the identification and characterization of deleterious mutations, providing crucial information about their position, size, and nature. The technique's high resolution allows for the detection of a wide spectrum of genetic alterations, including deletions, inversions, insertions, and point mutations. These can be analyzed either directly through electrophoretic separation of PCR products or in combination with other molecular techniques, offering a comprehensive genetic profile.

The application of PCR extends beyond genetic disorders to the domain of infectious disease diagnostics. It has proven particularly valuable in detecting viral, bacterial, and parasitic pathogens, including but not limited to Human Immunodeficiency Virus (HIV), Hepatitis C Virus (HCV), and Chlamydia species. The technique's exceptional sensitivity enables the detection of pathogens from minute biological samples, often surpassing the capabilities of traditional diagnostic methods. This is especially crucial in cases where the pathogen load is below the detection threshold of conventional techniques or in early stages of infection[23]*.

(cf. PCR applied to diagnosis)

***Exemple* : Key advantages of PCR in diagnostic applications**

- **High Sensitivity:** Ability to detect extremely low levels of target nucleic acids.
- **Specificity:** Precise identification of specific pathogens or genetic sequences.
- **Rapidity:** Quick turnaround time for results compared to many traditional methods.
- **Versatility:** Applicable to a wide range of genetic and infectious diseases.
- **Minimal Sample Requirement:** Effective with very small amounts of biological material.

The speed of diagnosis is unrivalled. Reduced from days/week to hours.

a) PCR for the detection of infectious diseases

Contamination with viruses or microorganisms, such as bacteria or parasites, results in the presence of their genetic material within the infected organism. PCR is an especially effective tool for detecting pathogens in biological samples due to its high sensitivity and specificity. The success of PCR-based diagnostics relies primarily on the careful selection of primers that selectively amplify a specific DNA sequence of the virus or microorganism. The matrix DNA must first be extracted from a tissue where the pathogen is present, after which a targeted sequence of the pathogen's genome can be amplified from a patient sample and analyzed by electrophoresis. The size of the amplified DNA fragment, which must match the expected size, ensures the reliability of the result and the diagnosis.

In the context of HIV testing, for instance, routine screening is typically based on the ELISA method, which detects HIV antibodies or viral antigens in the patient's serum via an immunoassay. Although this method is generally reliable and cost-effective, it has some limitations, such as frequent false positives due to cross-reactivities. Consequently, positive results are confirmed by another standard technique, Western blotting. However, there are cases, such as HIV-positive newborns whose mothers have AIDS, where PCR is particularly valuable. These newborns may carry maternal anti-HIV antibodies and therefore test seropositive, even if they are not infected with the virus themselves. In such situations, PCR is used to amplify a specific sequence of the HIV provirus from a lymphocyte extract, providing a more accurate diagnosis.

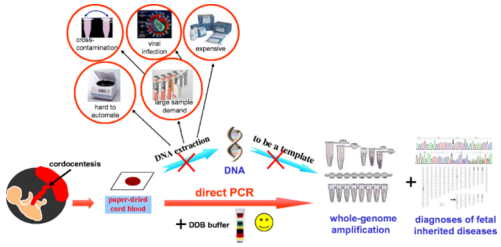
The same PCR-based principle is applied in detecting other pathogens, such as Toxoplasma in newborns whose mothers are infected. Additionally, PCR can be adapted for diagnosing HIV by using reverse transcription PCR (RT-PCR) to detect viral RNA in the patient's serum. Quantitative and semi-quantitative methods have also been developed to assess viral load, further enhancing the utility of PCR in monitoring infection severity[27]*.

4.5. PCR for the detection of genetic diseases

The goal of PCR in genetic diseases is often to detect a mutation in a gene sequence, and several scenarios can arise. The simplest cases involve insertions or deletions, where the mutation results in a change in the size of the gene or a part of it. If the mutation is known and well-documented, amplification of all or part of the gene through PCR is sufficient. For an insertion, the PCR product from a patient’s DNA will be longer compared to that of a healthy individual, while a deletion yields a shorter product. The analysis of these PCR products by electrophoresis, which evaluates their size, leads directly to a diagnosis. However, detecting inversions and point mutations is more challenging. In the case of an inversion, the difference in size between healthy and mutated DNA is zero, and for point mutations, it is nearly zero. Thus, size alone cannot be used as a criterion for diagnosis. To detect these mutations, additional techniques must be employed, such as Southern blotting, restriction fragment length polymorphism (RFLP), or mismatch detection. Southern blotting involves hybridizing an oligonucleotide probe, labeled with either a radioactive isotope or a fluorochrome, to the PCR product. The probe’s sequence is complementary and specific to the mutated region, making this method particularly suitable for detecting inversions[25]*.

RFLP can detect both inversions and point mutations by utilizing a restriction enzyme that hydrolyzes the PCR product at the site of the mutation. This approach is feasible only if the sequence contains a restriction site, whether in the mutated allele or the wild-type allele. The enzyme will hydrolyze either the PCR product derived from healthy DNA or that from diseased DNA, generating one or two DNA fragments that can be visualized by electrophoresis. Similarly, mismatch detection is well-suited for identifying inversions and point mutations. This technique involves mixing the PCR product from the patient’s DNA (sample DNA) with the PCR product from a healthy individual's DNA (reference DNA), followed by denaturation and rehybridization. If the sample DNA contains a mutation,

incomplete pairings will occur at the mutation site between the sample and reference DNAs. In the case of a point mutation, the mismatch involves a single base pair, whereas an inversion leads to mismatches across several base pairs. The mismatches are then degraded by S1 nuclease, an enzyme that specifically degrades single-stranded DNA. Alternatively, chemical cleavage using osmium tetroxide followed by piperidine can be employed, although this method is more appropriate for detecting point mutations*26*[26]*.



Diagnosis of genetic diseases

5. Exercice : Which of the following are true about quantitative PCR (qPCR)?

[solution n°1 p. 16]

quantitative PCR

- ☐ It uses fluorescent reporter molecules
- ☐ It can be performed as endpoint qPCR
- ☐ It always requires gel electrophoresis
- ☐ It can use SYBR Green or TaqMan probes

6. Exercice : What are the primary applications of PCR?

[solution n°2 p. 16]

Applications of PCR

- ☐ Amplifying DNA for sequencing
- ☐ Visualizing DNA by gel electrophoresis
- ☐ Creating proteins
- ☐ Diagnosing genetic diseases

7. What is DNA sequencing?

You may have heard of genomes being sequenced. For instance, the human genome was completed in 2003, after a many-year, international effort. But what does it mean to sequence a genome, or even a small fragment of DNA?

DNA sequencing is the process of determining the sequence of nucleotide bases (As, Ts, Cs, and Gs) in a piece of DNA. Today, with the right equipment and materials, sequencing a short piece of DNA is relatively straightforward.

Sequencing an entire genome (all of an organism's DNA) remains a complex task. It requires breaking the DNA of the genome into many smaller pieces, sequencing the pieces, and assembling the sequences into a single long "consensus." However, thanks to new methods that have been developed over the past two decades, genome sequencing is now much faster and less expensive than it was during the Human Genome Project.

***Fondamental* : Fundamental reasons for knowing the sequence of DNA molecule**

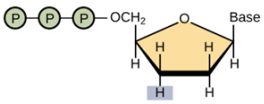
- To characterize the newly cloned DNA.
- To predicate its functions.
- To facilitate manipulation of the molecule.
- To confirm the identity of a clone or a mutation.
- To check the newly created mutation.

DNA Sequencing Methods: From Past to Present (cf. 8.pdf)

8. Modern Sanger's Chain Termination Sequencing

It is a PCR based method by modifying DNA replication reaction through growing chains by dideoxynucleotides. It is processed as follow*29*[29]*:

- A DNA polymerase enzyme.
- A primer, which is a short piece of single-stranded DNA that binds to the template DNA and acts as a "starter" for the polymerase.
- The four DNA nucleotides (dATP, dTTP, dCTP, dGTP).
- The template DNA to be sequenced.
- Each tube has one radio label ddNTP at a very low concentration each labeled with a different color of dye.



Dideoxynucleotide (ddNTP)

The DNA sample to be sequenced is combined in a tube with a primer, DNA polymerase, and the four standard DNA nucleotides (dATP, dTTP, dGTP, and dCTP). In addition, four dye-labeled, chain-terminating dideoxy nucleotides are added, but in much smaller quantities than the regular nucleotides. The mixture is first heated to denature the template DNA, separating the strands, and then cooled to allow the primer to bind to the single-stranded template. Once the primer is bound, the temperature is raised again, enabling DNA polymerase to synthesize new DNA starting from the primer. DNA polymerase adds nucleotides to the growing chain until it incorporates a dideoxy nucleotide, which prevents the addition of further nucleotides, terminating the strand.

This cycle is repeated multiple times, and by the end of the process, a dideoxy nucleotide will have been added at every possible position of the target DNA in at least one reaction. As a result, the tube contains DNA fragments of varying lengths, each terminating at a different nucleotide position. The ends of the fragments are labeled with dyes indicating the final nucleotide in each sequence.

Once the reaction is complete, the fragments are separated by capillary gel electrophoresis, in which they are passed through a thin tube filled with a gel matrix. Shorter fragments move through the gel faster than longer ones. As each fragment reaches the end of the tube, it is illuminated by a laser, allowing the dye label to be detected. The data recorded by the detector appear as peaks in fluorescence intensity, representing the DNA sequence. The DNA sequence is then read directly from the peaks in the resulting chromatogram.

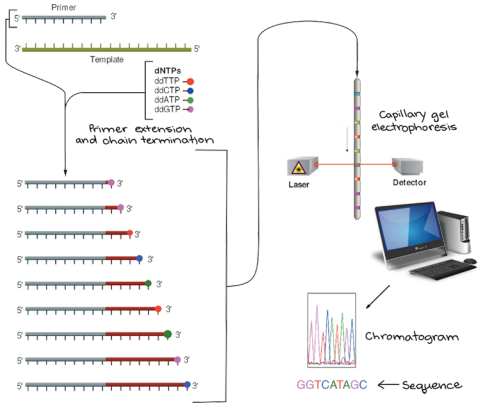


Image 2 Modern Sanger's Chain Termination Sequencing

9. Next-generation sequencing

The most recent set of DNA sequencing technologies are collectively referred to as next-generation sequencing.

There are a variety of next-generation sequencing techniques that use different technologies. However, most share a common set of features that distinguish them from Sanger sequencing:

- **Highly parallel:** many sequencing reactions take place at the same time.
- **Micro scale:** reactions are tiny and many can be done at once on a chip.
- **Fast:** because reactions are done in parallel, results are ready much faster.
- **Low-cost:** sequencing a genome is cheaper than with Sanger sequencing.
- **Shorter length:** reads typically range from 50-700 nucleotides in length.

Conceptually, next-generation sequencing is kind of like running a very large number of tiny Sanger sequencing reactions in parallel. Thanks to this is possible to sequence an entire human genome in one day.

***Méthode* : Steps of Next-generation sequencing**

- Generate a smaller DNA fragments, which will ligate with a certain oligonucleotide called adaptors.
- The double stranded fragments are denatured.
- All fragments have attached to different types of binding sites on each end (complementary sequences). All DNA fragments are now applied on ship called flow cell.
- DNA fragments will at this end hybridize with the flow cell as applied to the plate.
- The first step in flow cell is DNA amplification by PCR, the complementary sequence to our S-D DNA fragment is synthesized. This process seeing here is done simultaneously for all DNA fragments on the plate.
- DNA is denatured and the single-strand which not attached to the flow cell oligonucleotides get washed away.
- The bridge building where the second type oligonucleotide is hybridized to one of this sticky oligonucleotides on the plate.
- The bridge amplification is a process where DNA is amplified again where a polymerase synthesize the complementary sequence again.
- DNA is denatured again but both fragments stick to the plate.
- The DNA fragments will build a bridge again and will have bridge amplification.

- The whole process will repeat it for several rounds, so one DNA fragment we generate multiple copies by PCR. Reverse and forwards strands are presented from our copy.
- The reversed strands are cleaved from the plate and then the sequencing begin.
- The primer can bind to the oligonucleotide, which is the sequencing binding sites and specific nucleotides fluorescents
- The fluorescent excite by a laser and the fluorescent signal is obtained due to the color code.
- NGS generated billions of reads which are overlaid and compare to reference genome.

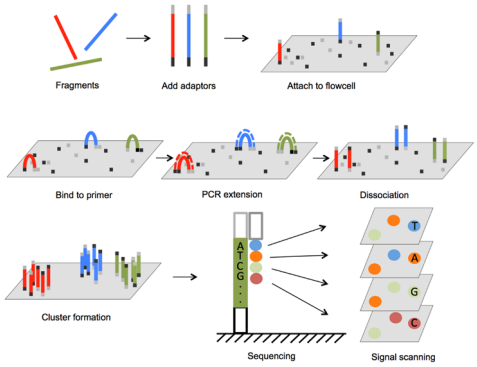


Image 3 Next-generation sequencing

10. Application of DNA sequencing

Next-generation sequencing (NGS) is widely used for disease biomarker discovery and validation. Whole exome sequencing, in particular, has been instrumental in diagnosing poorly characterized diseases by identifying novel mutations and biomarkers. Whole genome sequencing is well-suited for diagnosing rare diseases, especially when complex mutations, such as chromosomal translocations, are involved. NGS has become a key system for detecting diseases and generating personalized pharmacogenomic profiles.

RNA sequencing, exome sequencing, and whole genome sequencing are among the preferred NGS methods for discovering and validating biomarkers associated with various diseases and disorders. For instance, long non-coding RNAs (lncRNAs), microRNAs (miRNAs), and messenger RNAs (mRNAs) have been used as diagnostic biomarkers and even therapeutic targets in patients with diverse conditions. These include diseases such as autism, connective tissue disorders, cardiomyopathies (which affect the heart's ability to pump blood), and disorders of sex development, all of which can be diagnosed using NGS.

Application of Next Generation Sequencing (cf. lm-2021-41-1-25.pd)

(cf. DNA Sequencing Application in Disease Diagnosis & Treatment)

11. Exercice

[solution n°3 p. 16]

How can PCR be used to detect the presence of a viral infection in a patient's blood sample?

Infection detection

- ○ PCR amplifies the patient's entire genome to identify viral mutations.
- ○ PCR amplifies specific viral DNA or RNA sequences to detect the virus, even in small amounts.
- ○ PCR can only detect viruses if large amounts of viral particles are present.
- ○ PCR modifies the viral RNA into proteins for detection.

Which technique analyzes gene expression levels?

Application

How is DNA sequencing used to create personalized pharmacogenomic profiles?

Application

- ○ DNA sequencing identifies genetic variations that affect a patient's response to medications, allowing for personalized treatment.
- ○ DNA sequencing alters a patient's genetic code to improve drug efficacy.
- ○ DNA sequencing only helps in diagnosing diseases, not in drug-related treatments.
- ○ DNA sequencing is used to detect viral infections that influence drug metabolism.

Use of DNA Sequencing in Cancer Diagnostics

Application

- ☐ DNA sequencing is primarily used to detect the presence of cancer in tissue samples, not for guiding treatment.
- ☐ Understanding genetic mutations allows for the use of targeted therapies that are more effective for specific mutations.
- ☐ DNA sequencing can only identify genetic mutations related to hereditary cancers, not sporadic cases.
- ☐ DNA sequencing can identify mutations in cancer-related genes, aiding in the development of personalized treatment plans.

What is the process of copying DNA called?

Division

What technique is employed for rapid pathogen detection?

Application

Solutions des exercices

Solution n°1

[exercice p. 11]

quantitative PCR

- [x] It uses fluorescent reporter molecules
- [ ] It can be performed as endpoint qPCR
- [ ] It always requires gel electrophoresis
- [x] It can use SYBR Green or TaqMan probes

Solution n°2

[exercice p. 11]

Applications of PCR

- [x] Amplifying DNA for sequencing
- [x] Visualizing DNA by gel electrophoresis
- [ ] Creating proteins
- [ ] Diagnosing genetic diseases

Solution n°3

[exercice p. 15]

How can PCR be used to detect the presence of a viral infection in a patient's blood sample?

Infection detection

- ○ PCR amplifies the patient's entire genome to identify viral mutations.
- ◉ PCR amplifies specific viral DNA or RNA sequences to detect the virus, even in small amounts.
- ○ PCR can only detect viruses if large amounts of viral particles are present.
- ○ PCR modifies the viral RNA into proteins for detection.

Which technique analyzes gene expression levels?

Application

RT-PCR

How is DNA sequencing used to create personalized pharmacogenomic profiles?

Application

- ◉ DNA sequencing identifies genetic variations that affect a patient's response to medications, allowing for personalized treatment.

- ○ DNA sequencing alters a patient's genetic code to improve drug efficacy.
- ○ DNA sequencing only helps in diagnosing diseases, not in drug-related treatments.
- ○ DNA sequencing is used to detect viral infections that influence drug metabolism.

Use of DNA Sequencing in Cancer Diagnostics

Application

- ☐ DNA sequencing is primarily used to detect the presence of cancer in tissue samples, not for guiding treatment.
- ☑ Understanding genetic mutations allows for the use of targeted therapies that are more effective for specific mutations.
- ☐ DNA sequencing can only identify genetic mutations related to hereditary cancers, not sporadic cases.
- ☑ DNA sequencing can identify mutations in cancer-related genes, aiding in the development of personalized treatment plans.

What is the process of copying DNA called?

Division

DNA replication

What technique is employed for rapid pathogen detection?

Application

Real-time PCR

Glossaire

Polymerase chain reaction (PCR)

A method widely used to make millions to billions of copies of a specific DNA sample rapidly, allowing scientists to amplify a very small sample of DNA (or a part of it) sufficiently to enable detailed study.

Abréviations

DNA : Deoxyribonucleic acid ; is the molecule that carries genetic information for the development and functioning of an organism.

Bibliographie

Gupta, V., Sengupta, M., Prakash, J., & Tripathy, B. C. (2016). An Introduction to Biotechnology. Basic and Applied Aspects of Biotechnology, 1–21. doi:10.1007/978-981-10-0875-7_1

Zhao, X., Li, G., & Liang, S. (2013). Several Affinity Tags Commonly Used in Chromatographic Purification. Journal of Analytical Methods in Chemistry, 2013, 1–8.

Gilroy, C. A., Roberts, S., & Chilkoti, A. (2018). Fusion of fibroblast growth factor 21 to a thermally responsive biopolymer forms an injectable depot with sustained anti-diabetic action. Journal of Controlled Release, 277, 154–164. doi:10.1016/j.jconrel.2018.03.015

Goh, H. C., Sobota, R. M., Ghadessy, F. J., & Nirantar, S. (2017). Going native: Complete removal of protein purification affinity tags by simple modification of existing tags and proteases. Protein Expression and Purification, 129, 18–24.

Johnston, M. O. (2006). Mutations and New Variation: Overview. Encyclopedia of Life Sciences. doi:10.1038/npg.els.0004165

Durland J, Ahmadian-Moghadam H. Genetics, Mutagenesis. [Updated 2022 Sep 19]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560519/

Madhavan, A., Sindhu, R., Binod, P., Sukumaran, R. K., & Pandey, A. (2017). Strategies for design of improved biocatalysts for industrial applications. Bioresource Technology, 245, 1304–1313. doi:10.1016/j.biortech.2017.05

Ling MM, Robinson BH. Approaches to DNA mutagenesis: an overview. Anal Biochem. 1997 Dec 15;254(2):157-78.

Jabalameli HR, Zahednasab H, Karimi-Moghaddam A, Jabalameli MR. Zinc finger nuclease technology: advances and obstacles in modelling and treating genetic disorders. Gene. 2015 Mar 01;558(1):1-5.

Shahzad, S., Afzal, M., Sikandar, S., & Afzal, I. (2020). Polymerase Chain Reaction. Genetic Engineering - A Glimpse of Techniques and Applications. doi:10.5772/intechopen.81924

Dymond, J. S. (2013). Explanatory Chapter. Laboratory Methods in Enzymology: DNA, 279–289. doi:10.1016/b978-0-12-418687-3.00023-9

Morrow, J. F. (1979). [1] Recombinant DNA techniques. Recombinant DNA, 3–24. doi:10.1016/0076-6879(79)68003-5

Marcela AAV, Rafael LG, Lucas ACB, Paulo RE, Alessandra ATC, Sergio C. Principles and applications of polymerase chain reaction in medical diagnostic fields: A review. Brazilian Journal of Microbiology. 2009;40:1-11

Joyce C. Quantitative RT-PCR. A review of current methodologies. Methods in Molecular Biology. 2002;193:83-92

Rajeevan MS, Vernon SD, Taysavang N, Unger ER. Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. The Journal of Molecular Diagnostics. 2001;3(1):26-31

Stephen B, Mueller R. Realtime reverse transcription PCR (qRTPCR) and its potential use in clinical diagnosis. Clinical Science. 2005;109:365-379

Lin MH, Chen TC, Kuo TT, Tseng C, Tseng CP. Real-time PCR for quantitative detection of Toxoplasma gondii. Journal of Clinical Microbiology. 2000;38:4121-4125

Fortin NY, Mulchandani A, Chen W. Use of real time polymerase chain reaction and molecular beacons for the detection of Escherichia coli O157:H7. Analytical Biochemistry. 2001;289:281-288

Jeyaseelan K, Ma D, Armugam A. Real-time detection of gene promotor activity: Quantification of toxin gene transcription. Nucleic Acids Research. 15 June 2001;29(12):e58

Kadri, K. (2020). Polymerase Chain Reaction (PCR): Principle and Applications. Synthetic Biology - New Interdisciplinary Science. doi:10.5772/intechopen.86491

Shehata HR, Hassane B and Newmaster SG (2024) Real-time PCR methods for identification and stability monitoring of Bifidobacterium longum subsp. longum UABl-14 during shelf life. Front. Microbiol. 15:1360241. doi: 10.3389/fmicb.2024.1360241

Men, A. E., Wilson, P., Siemering, K., & Forrest, S. (2008). Sanger DNA Sequencing. NextGeneration Genome Sequencing: Towards Personalized Medicine, 1–11.https://doi.org/10.1002/9783527625130.ch1 (PDF) SANGER`S DIDEOXY CHAIN TERMINATION METHOD OF DNA SEQUENCING. Available from: https://www.researchgate.net/publication/378490879_SANGERS_DIDEOXY_CHAIN_TERMINATION_METHOD_OF_DNA_SEQUENCING [accessed Oct 16 2024].

Wilson, A. J., Morgan, E. R., Booth, M., Norman, R., Perkins, S. E., Hauffe, H. C., … Fenton, A. (2017). What is a vector? Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1719), 20160085.

Bertero, A., Brown, S., & Vallier, L. (2017). Methods of Cloning. Basic Science Methods for Clinical Researchers, 19–39. doi:10.1016/b978-0-12-803077-6.00002-3 10.1016/B978-0-12-8030

Chiew, K. L., Yong, K. S. C., & Tan, C. L. (2018). A survey of phishing attacks: Their types, vectors and technical approaches. Expert Systems with Applications, 106, 1–20. doi:10.1016/j.eswa.2018.03.050

Redman, M., King, A., Watson, C., & King, D. (2016). What is CRISPR/Cas9? Archives of Disease in Childhood - Education & Practice Edition, 101(4), 213–215. doi:10.1136/archdischild-2016-310459

Ghalayini M., Magnan M., Dion S., Zatout O., Bourguignon L., Tenaillon O., et al. (2019). Long-term evolution of the natural isolate of Escherichia coli 536 in the mouse gut colonized after maternal transmission reveals convergence in the constitutive expression of the lactose operon. Mol. Ecol. 28 4470–4485.

Kimple, M. E., Brill, A. L., & Pasker, R. L. (2013). Overview of Affinity Tags for Protein Purification. Current Protocols in Protein Science, 9.9.1–9.9.23.

Terpe, K. (2005). Protein Tags. In: Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine. Springer, Berlin, Heidelberg . https://doi.org/10.1007/3-540-29623-9_3650