University of Jordan
Faculty of Medicine

GENETICS & Molecular Biology

Number: 8

Done By: Toqa Al-humaidi

Corrected By: Asma' Kilani

Subject: DNA sequencing & PCR

Doctor: Mamoun Ahram

Price:

Date: 15-2-16

In this sheet we'll talk about techniques used commonly in labs :

- DNA sequencing
- Polymerase chain reaction (PCR)

DNA sequencing:

What is DNA sequencing?

Simply, it is to know the sequencing bases of DNA.

Why do we actually use it?

- We use this method to **determine the species** of an organism that we have a piece of its DNA → i.e, if you have unknown bacterial cell and you want to know how many origins of replication (consciences sequences) it has, you simply sequence DNA and compare it to known sequences looking for consciences sequences.
- To **identify genes** by looking at the DNA sequence.

  → How do we know that this DNA piece is a gene?

  We recognize the promoters or exons sequences in the gene for example.

- To **identify proteins** (structure & function) produced from a certain gene.

  Remember: Codons ⟶ amino acids ⟶ protein.

  →Example: if you compared the sequence of a certain gene with consciences sequences of membrane proteins (known) and you found a match, then the gene codes for membrane protein. (Not 100% accurate but it gives an idea of the structure of the protein).
- To **identify DNA mutations** (by comparing sequences to each other).
- To **identify genetic variations** among individuals and diseases. Also, this technique gives an idea of interrelationship between different populations on the level of DNA.
- As an indication of the **evolutionary conservation** among organisms generally.

Sequencing of viral & bacterial DNA was started in 1970s as viruses & bacteria have simpler genome, then they sequenced mitochondrial DNA. In 1999, they began human genome sequencing project! After completing this project three years ago, they found out that human only have ~22 thousand genes.

**the human genome is of 3 billion bases! **

(Table slide 5 not for memorization).

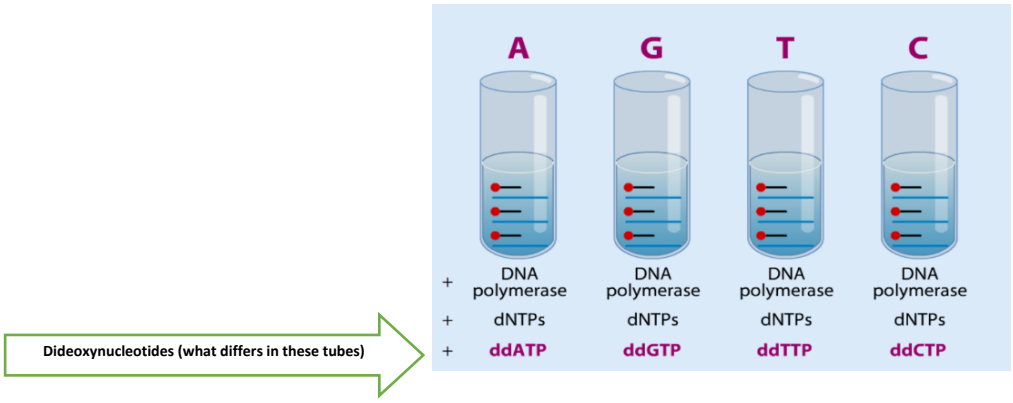
-DNA sequencing is based on using a specific modified nucleotide, (dideoxynucleotide → with missing 3 –OH groups not 2, the third is on C #3), look at slide 7, the left one is normal deoxynucleotide, but the one on the right is dideoxynucleotide, which means that we can't synthesize DNA using dideoxynucleotide (it will make the DNA synthesis stop). Instead, the DNA synthesis will be terminated.

What do we need in DNA sequencing?

- DNA template (that we know it's sequence)
- Primer (DNA or RNA primer).
- DNA polymerase.
- Substrates (deoxy A,G,C,T).
- Dideoxynucleotide (one substrate/each fragment).

How to sequence DNA?

We prepare four reaction tubes as what follows:



In these tubes, as there are nucleotides, polymerases & primers the DNA synthesis takes place, but the DNA polymerase can (by chance) add dideoxynucleotide instead of the deoxynucleotide leading to termination of the DNA synthesis.

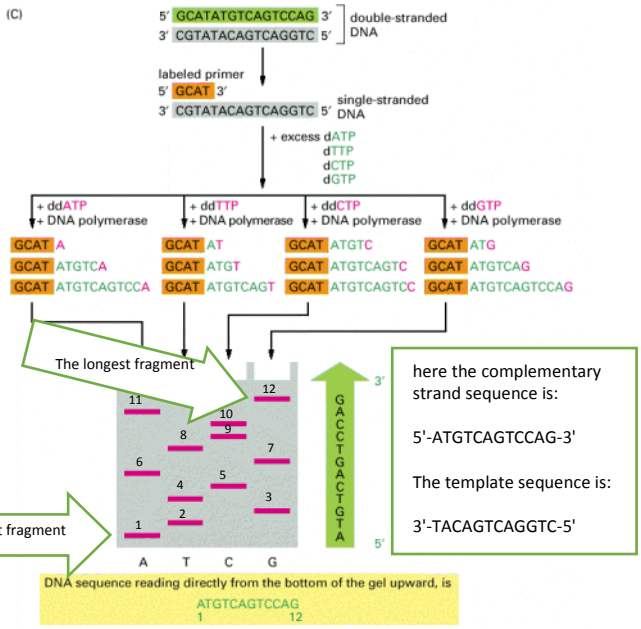
In the first tube for example, if the polymerase faces T in the template strand, it will either add dAdenosine OR ddAdenosine (it's a matter of chance). The synthesis of different DNA molecules is terminated at different points (points at which the polymerase added ddAdenosine).

Thus, fragments of different lengths will be produced. Collectively, all fragments of all tubes we get fragments that differ in one nucleotide in length.

The next step is electrophoresis, adding the content of each tube in one lane in the gel.

the gel has **high resolution to separate molecules of one nucleotide length difference**

The whole process:



Pay attention, the synthesized fragment is complementary to the DNA molecule with the unknown sequence, so if the fragment is terminated with ddAdenosine for example, the base in the template strand in that certain point is Thymine.

SO, what is the sequence of the DNA template?

We read the sequence of the complementary strand from the bottom to the top (shorter- longer / 5'-3') and by so we know the sequence of the template strand BUT (3'-5') as the two strands are antiparallel.

This type is radioactivity-based DNA sequencing (old method).

-There is another type of DNA sequencing: fluorescence-based DNA sequencing.

It's based on colors for example we label ddA → red, ddT → yellow, ddG → green, ddC → blue. So we combine them and allow the computer to read we will find a sequence of colors that we know the nucleotide base so we can read the DNA sequence.

- What if the gene we sequence is mutated?

  In this case,

  * We'll see two bands at the exact same line in the gel (extra one in from another reaction). For example:

  The normal gene complementary: 5'-ATTCCCGTACGTGGTC-3'

  The mutated gene complementary: 5'-ATTCCCGTACGTGGTT-3'

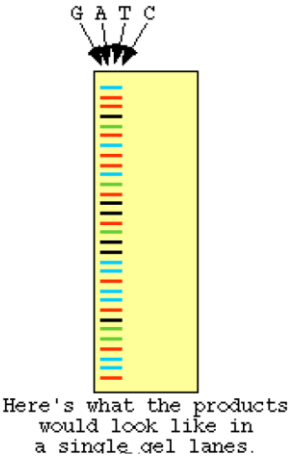
  These two strands will travel at the same time, in the same speed to the same place in the gel as C&T have almost equal sizes.

  *If a normal band disappeared then both alleles are mutated, so if I expect the complementary of C will be G and on the computer I had G and T overlap this means I have a mutation in one allele, but if I just found T which is abnormal with no G (normal) → there is a mutation in both alleles.

  So summary:

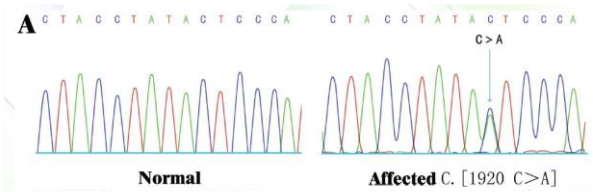
Fluorescence-based DNA sequencing

In this type, each dideoxynucleotide in the DNA fragments gives a different color (different wavelength), computer scans the gel and analyzes these colors and gives the template sequence. Here, the contents are put in the same lane.



Overlap:

We are diploid means we have two alleles with exactly the same sequence but sometimes we have genetic variation or mutation in one of them, the computer in the case will read two peaks.



When we have two peaks overlapping, then we have mutation in one allele.

BUT when we have **total change of the peak color** then both alleles are mutated.

Polymerase chain reaction (PCR):

Chain reaction → goes on and on and on again.

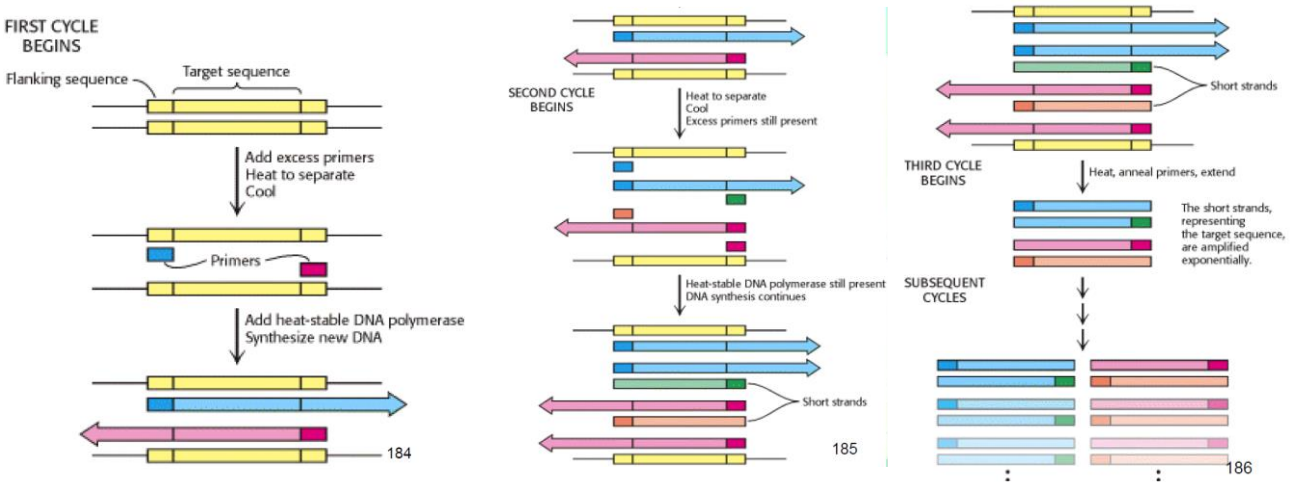
Used to amplify the DNA sample you have very fast (within hour)→ you start with one DNA fragment for example and ends with billions.

What do we need here? (No need for ddNucleotides here, unlike DNA sequencing).

- DNA template (that we want to amplify).
- DNA polymerase (bacterial, NOT human DNA polymerase).
- Substrates (normal nucleotides).
- Two primers (one for DNA sequencing).

The whole process is cycles and each cycle has **three steps**:

1- **Denaturation (by heating to 95°)** :
   We heat the DNA sample to 95° to separate the two strands (denature the molecule); 95° is constant in this step.
2- **Cooling down the sample to the annealing temperature**, about 60 ° this is the temperature at which the primers bind the DNA. We use variable annealing temperatures; because each primer binds specifically at certain temperature, and the reaction changes according to the melting temperature. The annealing temperature determines the specificity of the reaction; using very low annealing temperature allows primers to bind any DNA sequence (non-specific reaction), and using very high temperature prevents the annealing of primers, so they don't hybridize the DNA (NO reaction) → note that by raising the annealing temperature the produced fragment band on the gel disappears indicating that primers are not specific for the DNA.
3- **Increasing the temperature to 72°**, that is the temperature at which the activity of DNA polymerase is at its finest and the synthesis occurs, 72° is constant in this step.



Notice that the products of the first & second cycles are longer than what we want, only in the third cycle the target DNA fragment is produced. So, we actually start to amplify the sequence we want after the third cycle. Usually, 20-30 cycles are enough; we check the amplification by electrophoresis.

The purpose of PCR is amplifying the DNA.

**we said previously that we use bacterial DNA polymerase (of bacteria that live in hot springs / thermophilic bacteria i.e, Thermus aquaticus), why?? Because the enzymes of these bacteria are stable at very high temperature, so they won't denature at 95^{o} in the first step.

PCR has importance in paternity testing & Forensic medicine.