Genetics & molecular biology

◉ Sheet ○ Slide

Number:	5
Done by:	Dara Osama Suleiman
Corrected by:	Marah Bitar
Doctor:	Ma’moun Ahram

In this sheet we will discuss DNA sequencing and PCR which revolutionized the field of molecular biology, this was written according to the online video lecture 5 found on the website.

DNA sequencing

It is defined as **'**knowing the sequence of DNA; the process of determining the exact order of nucleotides in a genome.**'**

-Figuring out DNA sequence helps in **identifying genes** by consensus sequences (discussed earlier) which demarcate the beginning of a gene, exon, intron, etc.…

The sequence of nucleotide is meaningless until it is translated into something meaningful; amino acids.

-With the help of the table of genetic codon, the sequence of amino acids can be predicted and eventually the primary **structure of the protein** which in turn determines the function of the protein, eg; a kinase domain, receptor or a DNA binding protein.

-Knowing the DNA sequence also helps us to **identify mutations**. If someone is suspected of having sickle cell anemia, sequencing DNA of that gene helps us to identify if there's any variation in the DNA indicating a mutation.

-Moreover, **susceptibility to acquire a disease** can be also identified; **genetic variations** may indicate if a person is susceptible to hypertension in the next 10 years or even sudden heart failure, this is used to spread awareness and save lives in the future.

However; our environment and lifestyle also have an impact on our health not only our genes and they are factors which we can control. (A topic discussed later in epigenetics.)

-DNA sequencing also has **an evolutionary conservation** importance where you can compare DNA of different populations knowing their exact origins.

DNA sequencing started with **viruses** because they have a simple genome then simple and complex bacteria then yeast (unicellular eukaryotic system) then Caenorhabditis (a nematode where we learned differentiation by knowing the genes shared with our genome) then Drosophila (fruit fly) then humans.

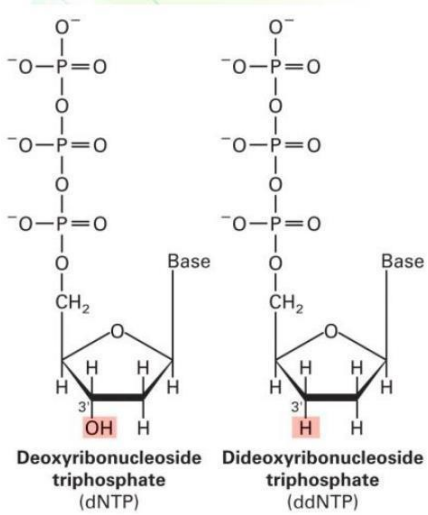
Human genome project (general information)

Took scientists 15 years and billions of dollars to finish it (1990-2006). The base sequence of human genome was identified.

-It was estimated that our DNA contained 100,000 gens. It turned out we only have about 21500 genes, which is similar to mice.

Method of DNA Sequencing (technique)

The most popular method is based on premature termination of DNA synthesis by **dideoxyribonucleotide;** a special substrate that doesn't form a phosphodiester bond with the upcoming nucleotide in the growing DNA so that synthesis is stopped. This is the purpose of its use in this method.



-We initially have millions of copies of the same DNA that needs to be sequenced.

-In order to carry out the DNA sequencing we need:

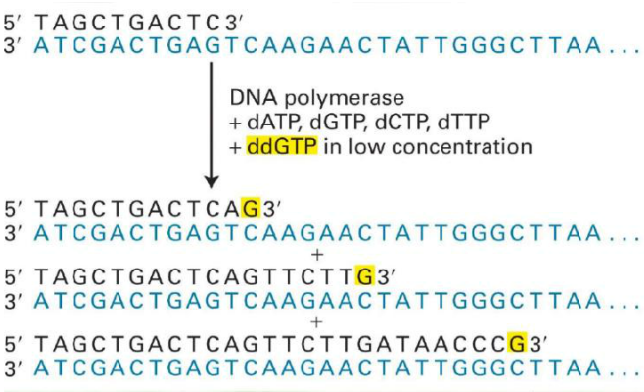
1- The Template targeted for sequencing.

2- DNA polymerase, which needs a primer and deoxynucleotides (dATP, dCTP, dGTP, dTTP).

3- Dideoxynucleotides (ddATP, ddCTP, ddGTP, ddTTP) **each labeled with a unique fluorescent tag.**

Steps of the reaction

This procedure relies heavily on probability.



1- Increase the temperature to cause denaturation of the DNA. Single strands are generated.

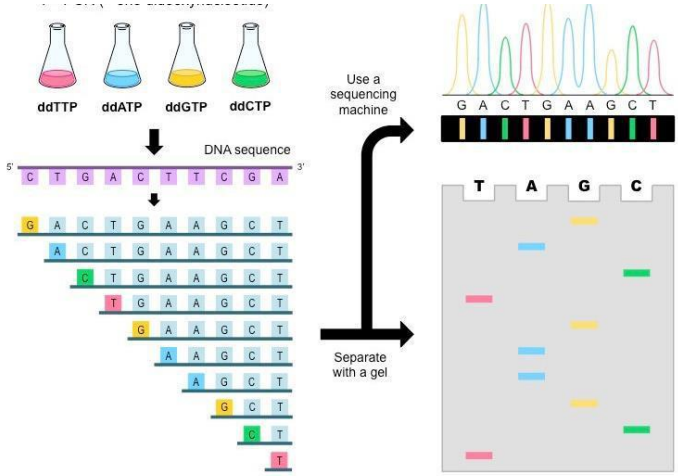
2-Decrease the temperature and the primer attaches to a specific place on the template that we want to sequence, and the reaction starts.

3- The DNA polymerase begins to synthesize by adding nucleotides, it's a matter of probability if it adds the dideoxy or deoxy. If dideoxy is added the reaction stops, if deoxy added then reaction continues.

We need to generate **n** DNA fragments of **varying lengths.** Each terminated with a labeled dideoxynucleotide, where n is the **number of nucleotide bases in the target DNA sequence**. (difference between each fragment is one nucleotide)

Each fragment will end with a **labeled nucleotide**.

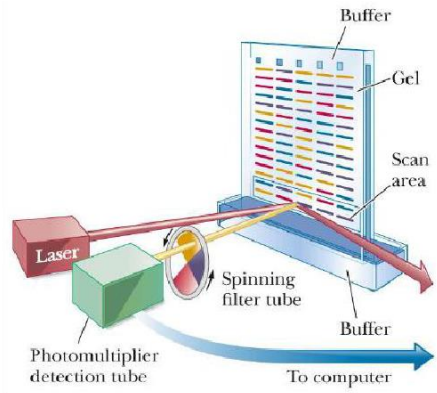
-We use in this technique labeled dideoxynucleotides; each base gives a certain color (fluorescence) so a fragment that ends with G will give a green peak, T red, etc.…

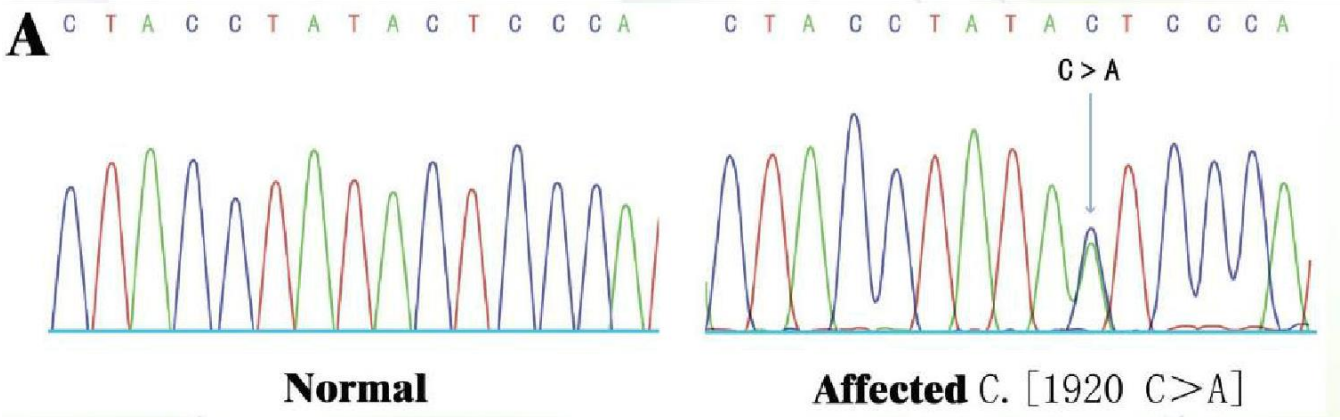


DNA fragments give color according to the last nucleotide in their sequence regardless of the length.

-All these fragments are separated on a gel which has a very high resolution, it can separate DNA fragments that differ with as little as 1 nucleotide. The shortest will be at the bottom(fastest). With the length increasing as we go up(slower).

-As the fragments move through the gel a laser beam directed towards the bands will reveal the fluorescence that is emitted from the fragment (the fluorescent output is displays in the form of a **chromatogram)** and according to the sequences length and colors emitted(we know each color represents a nucleotide) we will know the sequence of the DNA.





Note: the direction of reading is 5' to 3' and since the DNA is antiparallel we read one strand from left (5'end) to right (3'end), and for the other strand we can use 2 ways 1)read from right(5' end) to left (3' end) or 2) as the 1st strand from left to right but we should state that the left side is the 3' end.

EX: 2 ways to name the CCCTTTG strand

1) GTTTCCC (R to L)

2) 3'(CCCTTTG) 5' (L to R)

GGGAAAC
CCCTTTG

-In normal DNA sequence we have a peak for each position. However, if 2 peaks are present in at the same position, the person has two different nucleotides at that same site (in this case having two DNA segments that differ in only 1 nucleotide suggests that the parents maybe close relatives ex cousins).

At that peak, a different base will be added in each chromosome and the device will detect the 2 different colors at this position (**heterozygosity)**.

-If at one position one color was present but it is different than the normal sequence, this means that the person is homozygous but has a genetic variation or a mutation.

PCR: Polymerase Chain Reaction

-Polymerase chain reaction (PCR) allows the DNA from a selected region of a genome to be amplified a billionfold, effectively "purifying" this DNA away from the remainder of the genome.

-It was used when dinosaur bones were discovered. DNA was found in small amounts. It was isolated and amplified with PCR. Same with forensic science.

Components of the reaction

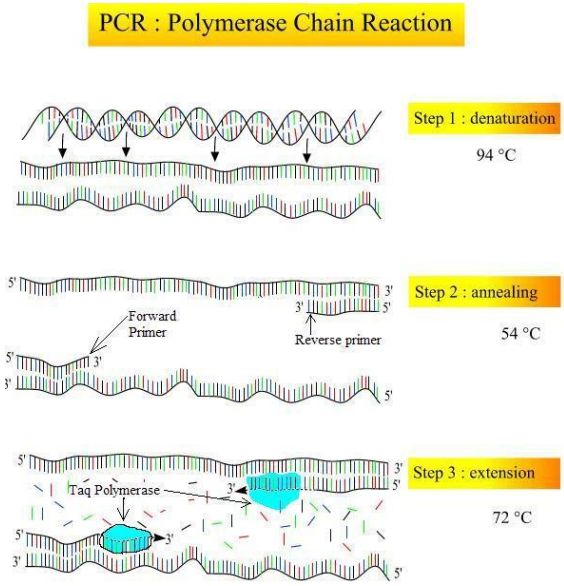
1- Template for amplification.

2- A heat stable DNA polymerase.

3- 2 primers (15-25 nucleotides-long) should surround the area needed to be amplified (left and right side).

4- All four deoxyribonucleoside triphosphates.

Steps for one cycle



1 –Denaturation done at very high temperature (95 degrees) to **separate** the two strands.

2 – **Annealing (binding) step**, decrease the temperature (50-60 degrees) so the primer can bind to the template. (ex the annealing temp is 55 degrees above 55 no binding happens and below that will have nonspecific binding).

3 – DNA synthesis (72 degrees, optimal temperature for Taq polymerase.)

-The cycle keeps on repeating with the same steps till we reach our point of interest. **(Usually 20-30 cycles.)**

-This is similar to cloning, talked about earlier, which takes time using bacteria while PCR takes 2 hours only.

Cycles

-20-30 cycles of reaction are required for DNA amplification.

(the products of each cycle serve as the DNA templates for the next-hence the term polymerase "chain reaction“).

-Every cycle doubles the amount of DNA.

-After 30 cycles, there will be over 250 million short products derived from each starting molecule.

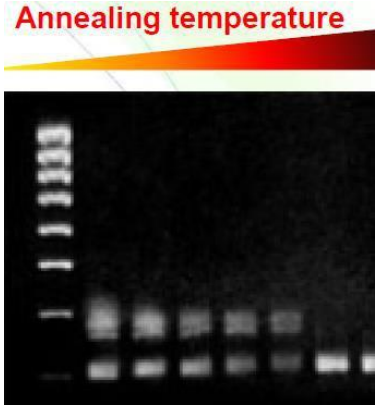
The DNA Polymerase

-A special DNA polymerase is used since the procedure takes place at very high temperatures. The polymerase was isolated from **thermophilic bacteria** “Thermus aquaticus” that live in hot water springs. It is thermostable up to 95°C.

Detection of Fragments

-The DNA fragment can be easily visualized as a discrete band of a specific size by agarose **gel electrophoresis.**

-The specificity of amplification depends on **the specificity of the primers** to not recognize and bind to sequences other than the intended target DNA sequences.



- **The Optimal temperature** for DNA synthesis (72 degrees) results in very specific amplification for the wanted sequence. If the temperature is too high, the primers can’t anneal to the DNA and **no hybridization occurs.**

-Too low temperatures cause **nonspecific amplification** due to imperfect hybridization of the primers.

– Imperfect hybridization and nonspecific amplification can occur on other parts of the DNA and result in the presence of **multiple bands**. This indicates the presence of a similar sequence on multiple regions. Gene families are discovered this way.

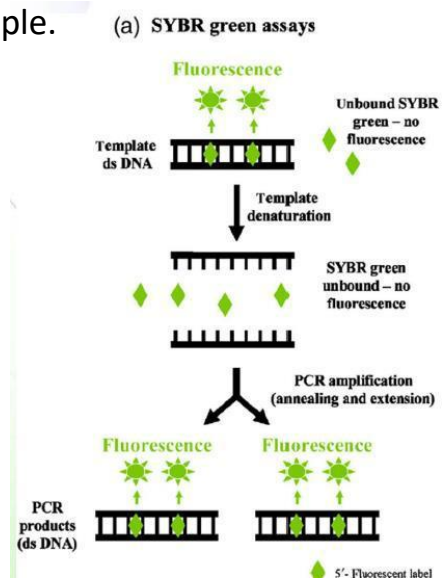
Uses of PCR

1-Paternal testing and criminal cases

An individual DNA profile is highly distinctive because many genetic loci are highly variable within a population.

2- Quantitative PCR (qPCR)

-Detects viral and bacterial load. Can be used to know how much viral DNA is present in the blood of a patient with flu for example.

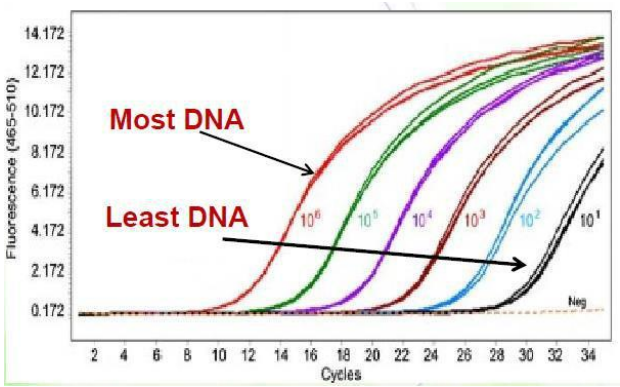


-A small molecule known as **SYBR green** fluoresces when it binds to **double stranded** DNA is used in this process.

-When there is more viral DNA in a blood sample, SYBR green binds to larger numbers of DNA molecules.

Ex let's say as a start 10 bacterial DNA molecules and 1000 viral DNA molecules are found in a sample of blood as the polymerase action starts the DNA particles start to double until it reaches up to a point where the fluoresce detector starts to detect the DNA particles.(in our experiment viral DNA will be detected first)

-the intensity of SYBR green detected by a special device reflects the amount of DNA in the sample (more DNA particles elects stronger intensity and DNA will be detected sooner).



Best of luck!