DNA Sequencing:

DNA Sequencing is the method that determines the order of the four nucleotides bases (adenine, thymine, cytosine, and guanine) that make up the DNA molecule and convey important genetic information.

DNA Sequencing applications:

DNA sequencing can used in numerous applied fields such as: *diagnostic* , *biotechnology*, *forensic biology* and *biological systematics*.

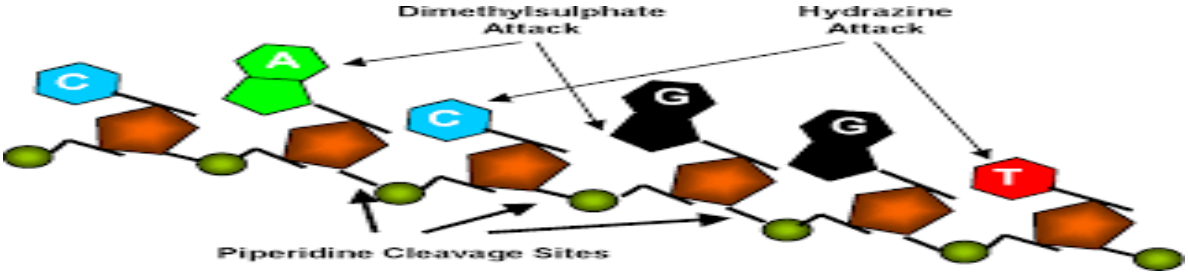
DNA Sequencing methods:

A sequencing can be done by different methods :
1. Maxam – Gilbert sequencing
2. Chain-termination methods
3. Dye-terminator sequencing
4. Automated DNA sequencing
5. Next generation sequencing

1- Maxam & Gilbert sequencing(chemical cleavage method,1977)

Allan Maxam and Walter Gilbert developed a method for sequencing single-stranded DNA by taking advantage of a two-step catalytic process involving **piperidine** and two chemicals that selectively attack **purines** and **pyrimidines**. Purines will react with **dimethyl sulfate** and pyrimidines will react with **hydrazine** in such a way as to break the glycoside bond between the ribose sugar and the base displacing the base (Step 1).

Piperidine will then catalyze phosphodiester bond cleavage where the base has been displaced (Step 2). Moreover, dimethyl sulfate and piperidine alone will selectively cleave guanine nucleotides but dimethyl sulfate and piperidine in **formic acid** will cleave both guanine and adenine nucleotides. Similarly, hydrazine and piperidine will cleave both thymine and cytosine nucleotides whereas hydrazine and piperidine in 1.5M **NaCl** will only cleave cytosine nucleotides).



The DNA sequencing then involved creating a single-stranded DNA substrate carrying a radioactive label on the 5' end.This labeled substrate would be subjected to four separate cleavage reactions, each of which would create a population of labeled cleavage products ending in known nucleotides.

The reactions would be loaded on high percentage polyacrylamide gels and the fragments resolved by electrophoresis. The gel would then be transferred to a light-proof X-ray film cassette, a piece of X-ray film placed over the gel, and the cassette placed in a freezer for several days. Wherever a labeled fragment stopped on the gel the radioactive tag would expose the film due to particle decay (**autoradiography**). Since electrophoresis, whether in an acrylamide or an agarose matrix, will resolve nucleic acid fragments in the inverse order of length, that is, smaller fragments will run faster in the gel matrix than larger fragments, the dark autoradiographic bands on the film will represent the 5'→3' DNA sequence when read from bottom to top . The process of **base calling** would involve interpreting the banding pattern relative to the four chemical reactions.

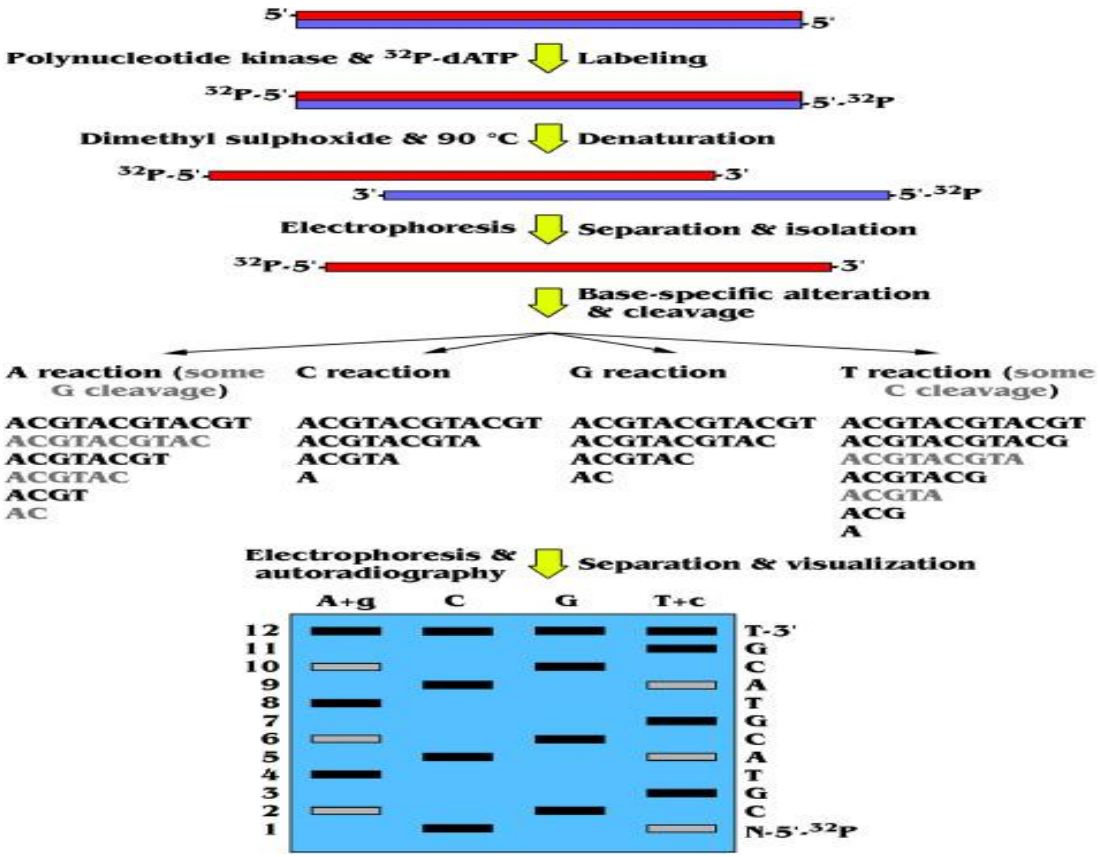


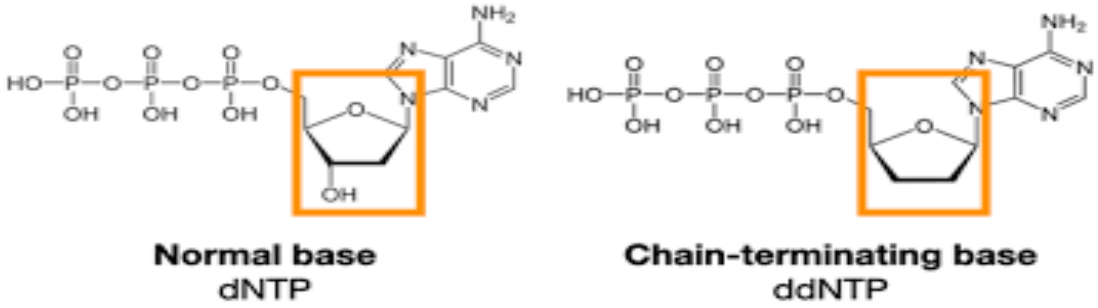
Figure: The Maxam-Gilbert manual sequencing scheme. The target DNA is radiolabeled and then split into the four chemical cleavage reactions. Each reaction is loaded onto a polyacrylamide gel and run. Finally, the gel is autoradiographed and base calling proceeds from bottom to top.

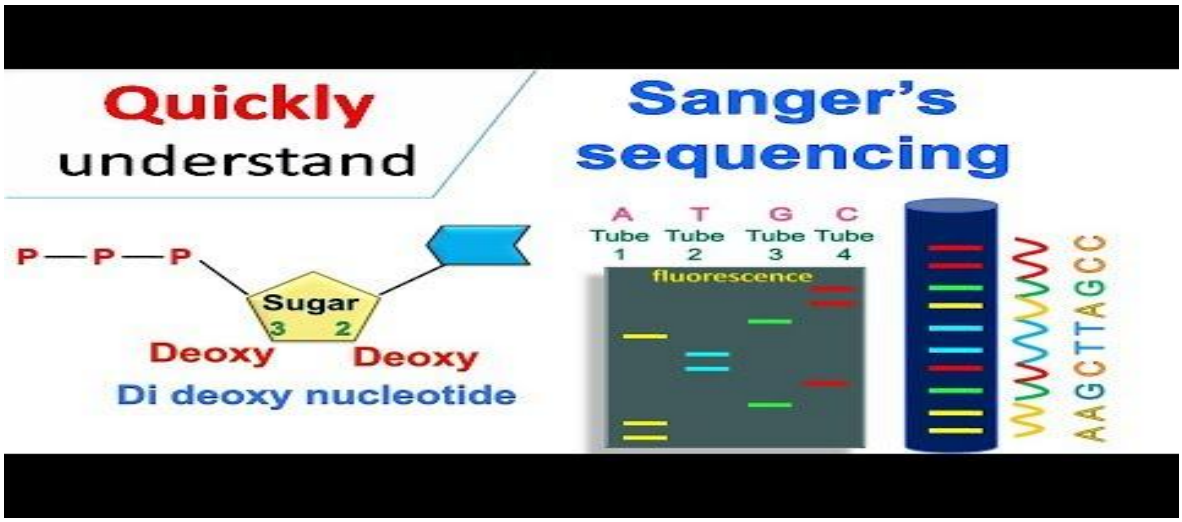
Challenges of the Maxam-Gilbert Method

The Maxam-Gilbert Method can only be used to typically resolve up to 400bp in a single run which is comparatively lower than other traditional sequencing methods such as Sanger sequencing. Additionally, the requirement of relatively large amounts of template DNA, coupled with the use of radioisotopes and hazardous chemicals, makes the Maxam-Gilbert method less favorable for routine use.

2-Sanger Method (Dideoxynucleotide chain termination)

In Sanger or dideoxy sequencing (also known as the chain termination method), dideoxynucleotides (ddNTPs) are used in addition to the normal deoxynucleotides (dNTPs): since dideoxynucleotides lack both 2′- and 3′-OH groups, 5′-to-3′ extension between 5′-phosphate and 3′-OH groups cannot take place, hence "terminating" the elongation of new DNA synthesis. Essentially, DNA synthesis takes place from a single-stranded template using a oligonucleotide primer (called the sequencing primer) and dNTPs. After a brief period of incubation this reaction is aliquotted into four different tubes, each of which contains a different labeled radioactively dideoxynucleotide terminator.





3-Dye termination sequencing

Most DNA sequencing is now automated. In the dye-terminator variant of Sanger sequencing ,the Sanger method chain termination reactions are still used, but pouring, running, & reading polyacrylamide gels has been replaced by automated methods. Instead of labeling the products of all 4 sequencing reactions the same (with a radioactive deoxynucleotide), each dideoxynucleotide is labeled with a different fluorescent marker. When excited with a laser, the 4 different kinds of products are detected and the fluorescence intensity translated into a data "**peak**".

Thus all four chain termination reactions can be performed in the same tube, and run on a single lane on a gel. A machine scans the lane with a laser. The wavelength of fluorescence from the label conjugated to the ddNTPs can be interpreted by the machine as an indication of which reaction (ddG, ddA, ddT, or ddC) a particular DNA band came from.

Note: The fluorescence output is stored in the form of a **chromatogram**.

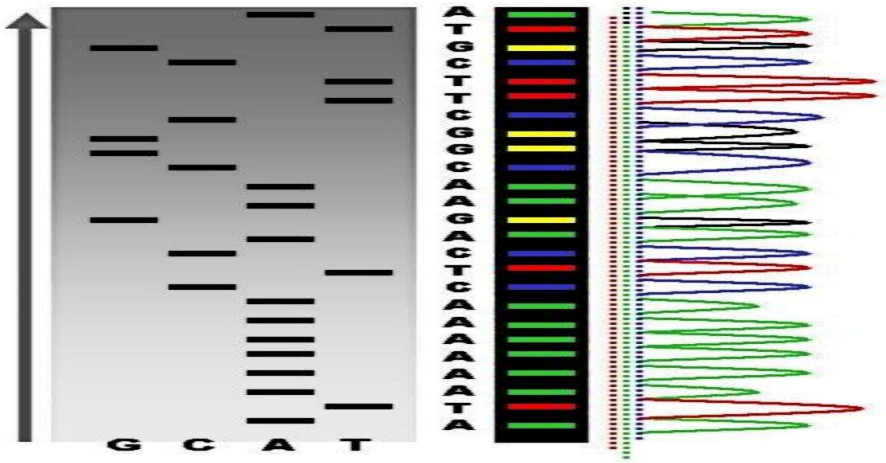


Figure. Sequence ladder by radioactive sequencing compared to fluorescent peaks.

4- Automated DNA sequencing

In 1986 Smith et al, published the first report of automation of DNA sequencing which established the dye-terminator variant of Sanger sequencing. This initial report showed that sequencing data could be collected directly to a computer.

Automated DNA sequencing instruments (**DNA sequencers**) can sequence upto 384 DNA samples in a single batch (run) in up to 24 runs a day.

Note: Sanger sequencing using dye-terminators became the dominant sequencing technique until the introduction of so-called *next-generation sequencing*/or *The 2nd generation of sequencing* technologies beginning in 2005(such as : Roche/454: pyrosequencing, Solexa (Illumina) &SOLID (ABI)). The obtainable sequence length be 1000 nucleotides.

5-Next Generation Sequencing(NGS)

Next-generation sequencing (**NGS**) is a technology for determining the sequence of DNA or RNA to study genetic variation associated with diseases or other biological phenomena. Introduced for commercial use in 2005, this method was initially called "massively-parallel sequencing", because it enabled the sequencing of many DNA strands at the same time, instead of one at a time as with traditional Sanger sequencing by capillary electrophoresis (CE). The key principles behind Sanger sequencing and **NGS** share some similarities. In **NGS**, the genetic material (DNA or RNA) is fragmented, to which oligonucleotides of known sequences are attached, through a step known as adapter ligation, enabling the fragments to interact with the chosen sequencing system. The bases of each fragment are then identified by their emitted signals. The main difference between Sanger sequencing and 2G NGS stems from sequencing volume, with NGS allowing the processing of millions of reactions in parallel, resulting in high-throughput, higher sensitivity, speed and reduced cost. A plethora of genome sequencing projects that took many years with Sanger sequencing methods could now be completed within hours using NGS.

Note: Using capillary electrophoresis-based Sanger sequencing, the Human Genome Project took over 10 years and cost nearly $3 billion. Next-generation sequencing, in contrast, enables scientists to analyze the entire human genome in a single sequencing experiment, or sequence thousands to tens of thousands of genomes in one year.

NGS applications:

1-Rapidly sequence whole genomes
2-Deeply sequence target regions
3-Utilize RNA sequencing (RNA-Seq) to discover novel RNA variants and splice sites, or quantify mRNAs for gene expression analysis
4-Analyze epigenetic factors such as genome-wide DNA methylation and DNA-protein interactions
5-Sequence cancer samples to study rare somatic
6-Identify novel pathogens

Next Generation Sequencing Techniques :

1-Second generation sequencing techniques

- 454 sequencing or pyrosequensing
- Illumina (Solexa) HiSeq and Mi Seq sequencing
- SOLID Sequencing
- Polony sequencing
- Massively parallel signature sequencing (MPSS)

2. Third generation sequencing techniques

- Single molecule real time sequencing Biosciences
- Heliscope Sequencing

3. Fourth generation sequencing techniques

- DNA nanoball sequencing
- Nano pore DNA sequencing