**DNA Sequencing: History, Steps, Methods, Applications and Limitations**

***“DNA sequencing is a process of determining or identifying the order of nucleotides present in a DNA sequence.”***

Nitrogenous bases, sugar and phosphate are three ingredients of the DNA in which Adenine, Thymine, Cytosine and Guanine are bases.  A functional piece of DNA is known as a gene which encodes proteinas.

For understanding the structure and function of a gene, It is very important to study the nucleotide sequence of it. DNA sequencing serves this purpose. In a sequential manner, the long chain of DNA is read by the sequencer machine.

Read more:

1. [Structure and function of DNA](https://geneticeducation.co.in/dna-story-the-structure-and-function-of-dna/)
2. [What Is a Gene?- Definition, Structure And Function](https://geneticeducation.co.in/what-is-a-gene-definition-structure-and-function/)

The present article is huge, in-depth, giant and contains all the information on the DNA sequencing, how to do sequencing and different methods of DNA sequencing. Please read the article till the end. The content of the article is,

**History of DNA sequencing:**

The story of DNA begins when *Watson* and *Crick* discovered the structure of DNA in the year 1953. In 1964, *Richard Holley* who performed the sequencing of the tRNA was the first attempt to sequence [the nucleic acid](https://geneticeducation.co.in/nucleic-acid-monomer-dimer-and-polymer/).

Using the technique of Holley and *Walter Fieser,*they sequenced the genome of bacteriophage MS2 (RNA sequencing). The sequenced molecules were RNA, Yet DNA sequencing was not performed.

In the year 1977, *Fredrick Sanger*postulated the first method for sequencing the DNA, named as a **chain termination method**.

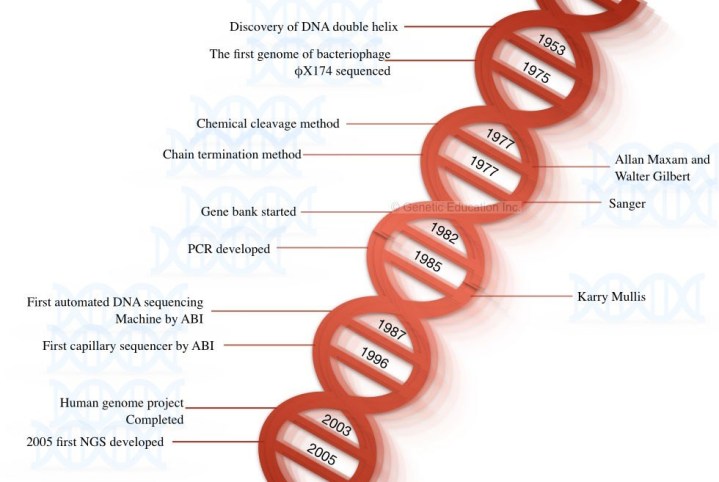
In the same year, the chemical method of DNA sequencing was explained by *Allan Maxam* and *Walter Gilbert*. The genome of bacteriophage X174 was sequenced in the same year using the chemical degradation method.

Because of the lack of automation, Both the methods (chemical degradation and chain termination) were tedious and time-consuming. The first semi-automated DNA method was developed by *Lorey* and *Smith* in the year 1986.

Later on, in the year 1987, the Applied Biosystem had developed a fully automated machine-controlled DNA sequencing method. After the development of the fully automated machines, the era of the 2000s become a golden period for the sequencing platforms.

Furthermore, in 1996, Applied Biosystem developed another innovative sequencing platform known as capillary DNA sequencing. After that, the human genome project was completed by using the combination of these methods in the year 2003.

Fast, accurate, reliable, and highly efficient next-generation sequencing platform was postulated in the year 2005 by Solexa/Illumina. Some of the milestone into the DNA sequencing is shown in the figure below,



**What is DNA sequencing?**

Studying the allelic variation is not enough in some cases, further, using the polymerase chain reaction like methods, new polymorphism can’t be identified. Overcome the present problem the DNA sequencing method was evolved.

**Definition:**

“A laboratory technique employed to known the correct DNA sequence by the sequential chemical reaction is known as DNA sequencing.”

laboratory processing and computational analysis both are the key processes in DNA sequencing. Once the chemical reaction is completed, the machine-generated data are sent to the computer lab.

In the process, the amplification done by each nucleotide is recorded in the form of signals which are collected by the machine and analyzed on the computer. This is the basic mechanism, the nucleotides are either radio or fluorescent-labeled.

Now let us understand the process stepwise.

**What are the steps in DNA sequencing?**

Steps mentioned below are the generalized representation of DNA sequencing, it may vary from platform to platform.

* Sample preparation (DNA extraction)
* PCR amplification of target sequence
* Amplicons purification
* Sequencing pre-prep
* DNA Sequencing
* Data analysis

**Sample prep:**

Here the DNA is our starting material, therefore we need to isolated DNA first. Animal, plant, bacterial, plasmid, or [environmental DNA](https://geneticeducation.co.in/what-is-environmental-dna/) can be used for it.

As the quality of DNA is the major concern in sequencing, we recommend using proteinase K or spin-column DNA extraction methods which has a high yield.

With the quantity, good quality of DNA also reduces the chance of reaction failure. Conclusively, a DNA sample with a 260/280 ratio of nearly ~1.80 (or 1.7 to 1.88) is considered as pure DNA. The quantity of DNA should be ~100ng for the present assay.

**PCR amplification of target sequence:**

Amplification is the pivotal step in sequencing. here only the gene of interest or the DNA sequence we wish to study is first amplified, the rest of the DNA is discarded. Nonetheless, in the [whole-genome sequencing](https://geneticeducation.co.in/3-of-the-best-genome-sequencing-methods/), the entire genome of an organism can be sequenced.

A set of flanking primers anneal at the outer regions of the DNA sequence of interest, therefore, unwanted DNA can’t amplify. Using the standard PCR conditions, the amplification process is done; 35 cycles, denaturation at 94°C, annealing at 55°C to 65°C and extension at 72°C.

After the amplification, the PCR product is run on 2% agarose gel along with the DNA ladder. Once the amplification achieved, DNA purification is performed.

Read more:

* [A Complete Guide of the Polymerase Chain Reaction](https://geneticeducation.co.in/a-complete-guide-of-the-polymerase-chain-reaction/)
* [Agarose gel electrophoresis](https://geneticeducation.co.in/agarose-gel-electrophoresis/)

**Amplicon purification:**

Why purification is needed?

This information is very interesting, no one will tell you about it.

If you check the purity of the PCR product at 260 and 280nm wavelength, it will always be 1.80, exactly! because the amplicon is the pure DNA fragment. If any contaminants are present in the DNA, amplification will not happen. Then why amplicon purification required? the answer is here,

**Unbound primers, primer-dimers, unused Taq DNA polymerase, unused DNA templates, and other unused**[**PCR reaction buffer components**](https://geneticeducation.co.in/what-are-the-different-components-used-in-the-pcr-reaction-buffer/)**can abort DNA sequencing. That is why amplicon purification is required.**

Here, the [alcohol purification](https://geneticeducation.co.in/different-methods-of-nucleic-acid-purification/) can’t work, we have to purify our PCR product with the spin column PCR amplicon purification kit to achieve maximum purification. After removing debris, the PCR amplicons send to the sequencing lab.

**Sequencing pre-preparation:**

Sample pre-preparation is a very crucial step during DNA sequencing. During this step, adaptor DNA sequences are ligated on both DNA ends. To the adaptor, the primer anneals for doing amplification. The amplification process in sequencing is similar to PCR, however, here the nucleotides are either radio or fluorescent-labeled.

With it, [high fidelity DNA polymerase](https://geneticeducation.co.in/choosing-the-right-dna-polymerase-for-your-pcr-experiment/) and other key ingredients are added in the reaction.

**DNA sequencing:**

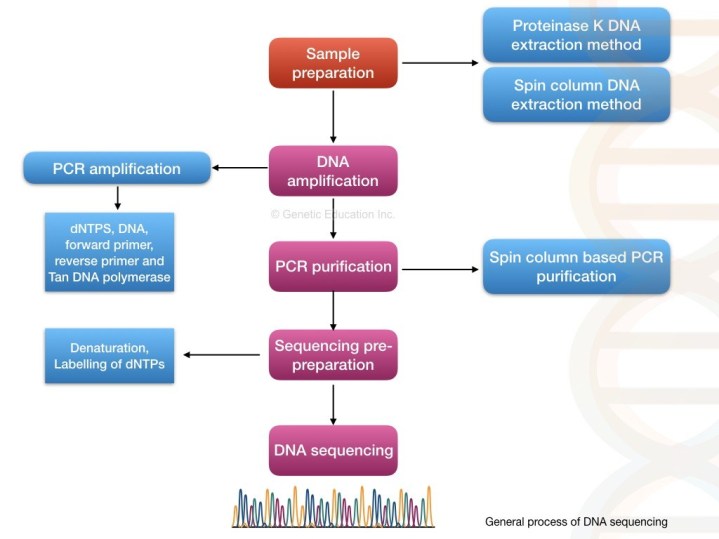
The prepared reaction tubes are placed into the sequencer machine. During, the reaction in the sequencer, denaturation, annealing, and extension occurs, simultaneously. Here, as we are using the labeled nucleotides, the signals produced during the reaction is recorded.

That signals of the addition of each complementary nucleotides are recorded by the machine and the data is sent to the computer.

**Data analysis:**

Once the whole DNA is sequenced, the result is saved into one unique file formate. The inbuilt software (provided by the manufacturer)  processed the data and compared it with the available data.

The sequence data is compared with other available data by the software to find out variations and other mutations present in a gene. The brief overview of the process is explained in the figure below,



*Steps in DNA sequencing.*

**Different methods of DNA sequencing:**

Various methods of DNA sequencing are explained here.

* Maxam and Gilbert method
* Chain termination method
* semiautomated method
* automated method
* Pyrosequencing
* The whole-genome shotgun sequencing method
* Clone by the clone sequencing method
* Next-generation sequencing method

Terminologies used into the article:

|  |  |
| --- | --- |
| **Term** | **Explanation** |
| Fragment library | A collection of the entire strand of the DNA (to be sequenced) fragments. |
| Gaps | Here the un-sequenced region of the DNA is called a gap. |
| Conting | A continuous sequence of the DNA assembled. |
| Read | The output data came from the sequencer machine to the computer for one particular sequence. |
| Coverage | The number of times the sequencing machine covered the DNA sequence. |

**Maxam-Gilbert sequencing:**

The *Maxam*and *Gilbert* method was developed in 1977. It is also referred to as a chemical cleavage method. By using this method, they had sequenced 24 nucleotides only. However, their method was published after two years of sangers method.

The brief principle of the present method is as stated,

***“The single-stranded DNA is cleaved at the specific location with the help of the chemicals at specific base location and the fragments of DNA is then run on polyacrylamide gel.”***

Obviously, DNA extraction is the very first step. After that, the DNA is denatured using the heat denaturation method and single-stranded DNA is generated.

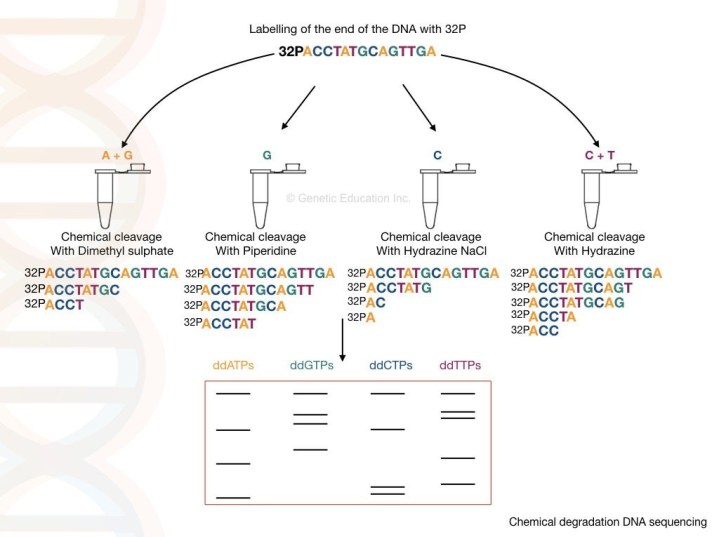
The phosphate (5’ P) end of the DNA is removed and labeled by the radiolabeled P32. The enzyme named phosphatase removes the phosphate from the DNA and simultaneously, the kinase adds the 32P to the 5’ end of it.

4 different chemicals are used to cleave DNA at four different positions; hydrazine and hydrazine NaCl are selectively attack pyrimidine nucleotides while dimethyl sulfate and piperidine attack purine nucleotides.

* Hydrazine: T + C
* Hydrazine NaCl: C
* Dimethyl sulfate: A + G
* Piperidine: G

An equal volume of 4 different ssDNA samples is taken into 4 different tubes each containing 4 different chemicals. The samples are incubated for sometimes and electrophoresed in polyacrylamide gel electrophoresis. The results of the chemicals cleavage of four different tubes are shown in the figure below.

Autobiography is used to visualize the separation of DNA fragments. Due to the radiolabelled 32P end of the DNA, the DNA bands visualized through autoradiography.



*Maxam- Gilbert chemical cleavage method of DNA sequencing*

**At our glance:**

The method is more accurate than Sanger sequencing. It’s best suitable for DNA footprinting and DNA structural studies. It is more advantageous over the Sanger method because the purified DNA is directly used for sequencing.

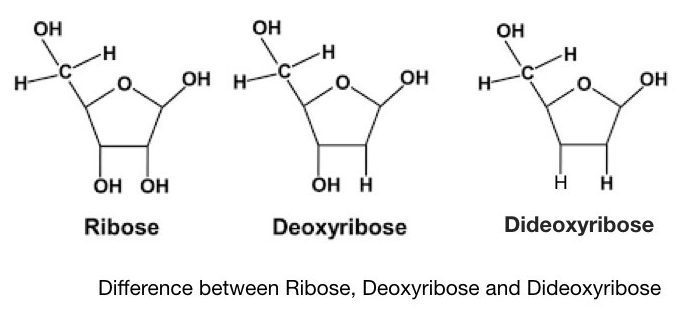
In the present automation time, the present method is used in DNA fingerprinting and [genetic engineering studies](https://geneticeducation.co.in/what-is-genetic-engineering-definition-types-process-and-application/).

Notwithstanding, plenty of disadvantages makes it harder to use regularly. The scalability of it is poor, only 400bp can be sequenced. Moreover, It is less popular because of the use of harmful radiolabeled chemicals.

**Sanger sequencing:**

Sanger and co-workers developed a chain termination method of DNA sequencing, after a few years of the Maxam and Gilberts method. The method is also known as the first-generation DNA sequencing method.

The **chain termination method** is also referred to as a **dideoxynucleotide sequencing** because of the use of the special types of ddNTPs. The ddNTPs are different from normal dNTPs. it possesses the hydrogen group instead of a hydroxyl group in the dNTPs.



*The image represents the difference between Ribose sugar, deoxyribose sugar, and dideoxyribose sugar.*

Phosphodiester bond can’t form between two adjacent nucleotides due to the absence of hydroxyl group in ddNTPs. The nucleotide chain can’t synthesis further and hence it is known as chain termination method.

The process of Sanger sequencing is broadly divided into 3 steps:

1. **DNA extraction**: using any of the DNA extraction protocols
2. **PCR amplification**: using the flanking primers, dNTPs,  ddNTPs, Taq DNA polymerase and PCR buffer.
3. **Identification of the amplified fragments**: using autoradiography, PAGE, or capillary gel electrophoresis.

Read more on DNA extraction methods: [Different types of DNA extraction methods](https://geneticeducation.co.in/different-types-of-dna-extraction-methods/)

The process of chain termination is started with the DNA extraction and purification. DNA extraction can be achieved using the proteinase K method or the phenol-chloroform DNA extraction method. Or we can use ready to use silica-column based kit method too.

The main aim of any DNA extraction method is to achieve the purity nearby ~1.8 and the quantity over 100ng.

|  |  |  |
| --- | --- | --- |
| Reaction | PCR reaction | modification |
| Reaction “A” | Taq DNA polymerase, dATPs, dGTPs, dCTPs, dGTPs and PCR buffer, primers | Labelled **ddATPs** |
| Reaction “G” | Taq DNA polymerase, dATPs, dGTPs, dCTPs, dGTPs and PCR buffer, primers | Labelled **ddGTPs** |
| Reaction “T” | Taq DNA polymerase, dATPs, dGTPs, dCTPs, dGTPs and PCR buffer, primers | Labelled **ddTTPs** |
| Reaction “C” | Taq DNA polymerase, dATPs, dGTPs, dCTPs, dGTPs and PCR buffer, primers | Labelled **ddCTPs** |

In the next step, the PCR amplification is performed by designing four different reactions

Each tube contains the same amount of the PCR reagents but in each tube, extra ddNTPs are added as shown into the table.

The flanking primers are such a primer that binds to the region near to our sequence of interest or the sequence what we are going to read.

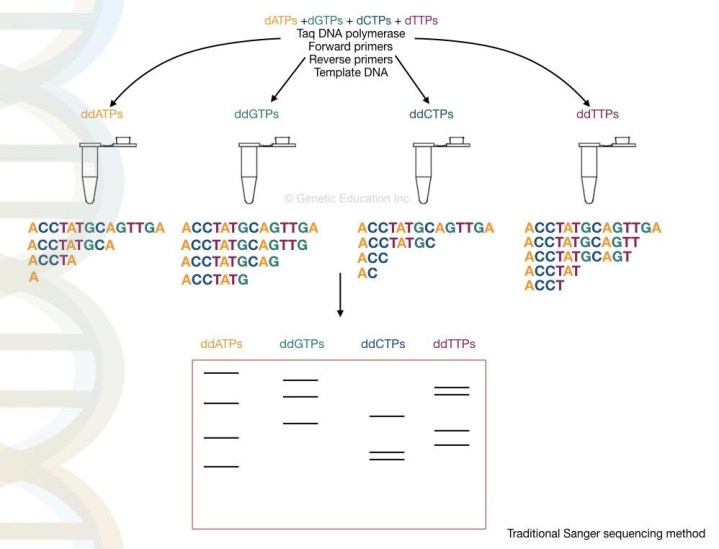
Now in the next step, the Taq DNA polymerase adds the dNTPs to the DNA strand.

Here, the high fidelity DNA polymerase is not needed. normal Taq polymerase expands the growing DNA strand by the addition of the dNTPs. Interestingly, once it adds the ddNTP instead of dNTP the chain expansion is stopped or terminated.

The termination process is complete in 4 different tubes for 4 different ddNTPs. For example, in the ddATP tube, it terminates the chain at all the positions where the dATPs are going to bind.

The amplified PCR products are loaded on to the polyacrylamide gel electrophoresis. The DNA fragments migrate into the gel based on the size of the fragments. The smaller fragments run faster towards the positive charge than the larger fragments.

Before the PAGE run, the amplified DNA fragments are further denatured by heat. Depending upon the types of labeling the gel is then analyzed under UV light or X-ray film. The banding pattern of the amplified product is shown in the figure below,



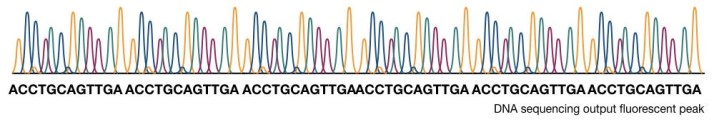
*Sangers chain termination method of DNA sequencing.*

**Advancement:**

The Sanger sequencing is the gold standard method for research as well as in the diagnosis, nowadays. Because of its easy setup and high reproducibility. its automation had done easily.

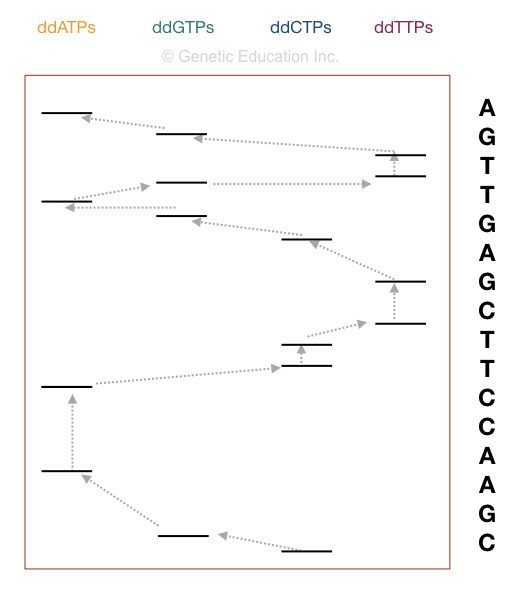
Traditionally, [the results are interpreted](https://geneticeducation.co.in/part-2-analysing-and-interpreting-agarose-gel-electrophoresis-results/) on PAGE manually but now, the scenario is changed. The process is fully or partially automated. A detector detects the fluorescence signals each time when the chain is terminated. Further, the signals are recorded and analyzed computationally.

The computational software generates various fluorescence peaks depending upon the amount of fluorescence emitted. The hypothetical illustration of the Sanger sequencing results are shown in the figure below,



**How to interpret the PAGE result for Sanger sequencing?**

This section is very important, I will teach you how you can interpret the results of the sequencing. Although it is a type of overview.



Now take a look at the figure above,

The principle of the electrophoresis stated that the smaller DNA fragments migrate faster than lager one. So the fragment on the positive side is the smallest one. Start reading the sequence from that.

The first point in sequencing the DNA is to match the size of DNA. if the bands obtained in gel and the nucleotide sequence of our DNA is similar, the reaction is completed properly. For instance, if your sequence length is 16, then 16 bands must be present into the gel.

Start reading from the bottom. The DNA sequence starts with “C” as the last fragment of DNA terminated with ddCTP. Arrange the sequencing shown in the figure, accordingly.

**Automated DNA sequencing:**

The manual Sanger method was tedious. However, recent advancement into the sequencing makes it easy and rapid to use. The semi-automated Sanger sequencing method is based on the principle of Sanger’s method with some minor variations.

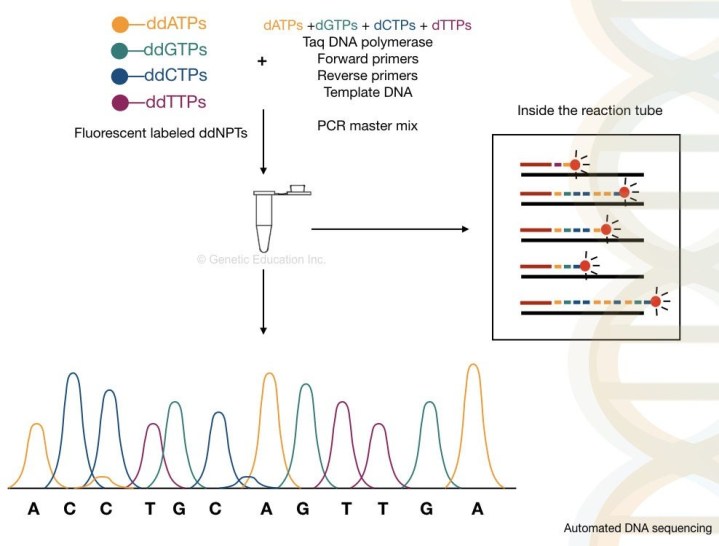
Instead of the 4 different reactions, the automated DNA sequencing carried out in the single tube. Which means on a gel the DNA runs in a single lane.

Here in the semi-automated DNA sequencing, the fluorescent-labeled set of primers are used, instead of ddNTPs. Thus four different primers give four different peaks.

The PAGE method isn’t capable of separating all the fragments in a single reaction. Therefore, alternatively, the capillary gel electrophoresis method is practiced. This method separates each and every single fragment precisely

By using the labeled primers or the dNTPs, the machine reads the sequence accurately, on a single lane capillary electrophoresis. We can sequence more than 300 samples in a single run on an automated DNA sequencing platform.

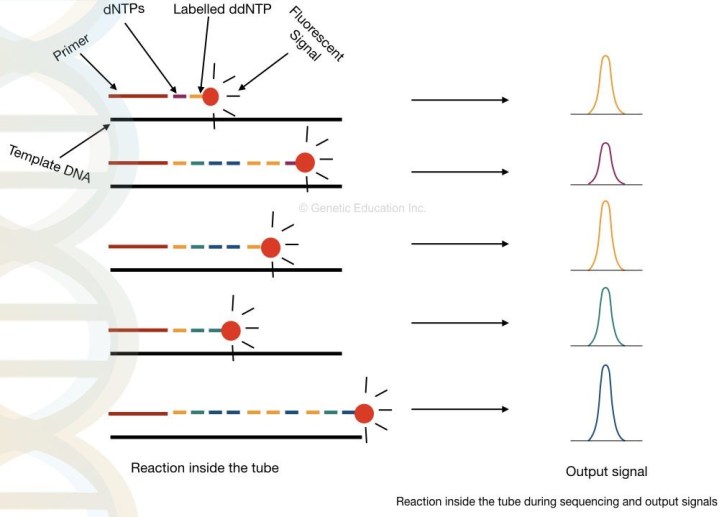
The capillary electrophoresis used to separate DNA molecules on the basis of the size, It is powerful enough to separate single basepair fragment. The chromatogram generated through the C.E sent the output as a fluorescent peak. Roughly, the process of it is explained below here,



*Automated DNA sequencing based on the sangers chain termination method.*

The advanced semi-automated Sanger sequencing method is more accurate, reliable and faster than the traditional method.

**What happened inside the tube?**



**At our Glance:**

The read capacity of the Sanger sequencing is higher as compared with the chemical degradation method. It can sequence 700 to 800bp sequence in a single run, therefore, it is more suitable for sequencing bacterial or other prokaryotic genomes.

It is more advanced and automated. Even the error rate is very low as compared with the conventional chain termination method. Still, it is time-consuming and a high-cost method.

***Fact:***

**It takes approximately, 50 years to sequence all the 3.2 billion bases of the human genome through the semi-automated Sanger sequencing methods.**

**Pyrosequencing:**

In 1993, *Bertil Pettersson, Mathias Uhlen* and *Pål Nyren* described the pyrosequencing method.

The method is based on the detection of the pyrophosphate released during the chain reaction of nucleotide addition. Here the order of the nucleotide is determined by the PPi released during the joining of two adjacent nucleotides (3’OH- 5’P).

In contrast with other methods, instead of a single polymerase, two additional enzymes are required in the pyrosequencing method. The three enzymes are:

* DNA polymerase (without exonuclease activity)
* Luciferase
* Sulfurylase

All three enzymes work in a sequential manner for the detection of the PPi. The real-time polymerase activity monitoring allows the detection of the released pyrophosphate in a cascade of the enzymatic reaction,

(DNA)n + dNTP ———————— (DNA)n+1 + PPi **(Polymerase)**

The addition of one dNTP removes one pyrophosphate from the DNA.

PPi +  APS —————————– ATP + SO4-2**(ATP sulfurylase)**

ATP + luciferin + O2 ——————— AMP + PPi + oxyluciferin + CO2 + photon **(luciferase)**

Here the reaction is completed into the three steps:

The enzyme polymerase adds the dNTPs to the single-stranded DNA. If the correct complementary base is added, the pyrophosphate released.

The enzyme sulfurylase converts the  PPi into the ATP (energy) with the help of the APS (adenosine 5´ phosphosulfate).

The ATP act as a substrate for the luciferase activity (more specifically “firefly luciferase”). With the help of the ATP substrate, the luciferase converts the luciferin into the oxyluciferin in the presence of oxygen and the photon of light is released.

Once the correct nucleotide is added, the amount of the light released by the enzymatic reaction is detected by the charged device coupled camera, photodiode, or photomultiplier tube. This is the basic fundamental of the pyrosequencing set up.

Based on the substrate used in the technique two types of pyrosequencing methods are available: solid-phase pyroseq and liquid phase pyroseq. we will discuss each type of pyrosequencing in some other article.

**At our Glance:**

The pyrosequencing method is more accurate than the Sanger sequencing which has the capacity to add up to 500 nucleotides. The major advantage of pyrosequencing is the speed of the reaction.

However, the method required more chemical steps than a chain termination method which makes it more complex. Furthermore, the read length is too short as compared with the automated sequencing.

**Whole-genome shotgun sequencing:**

Yet another modification of the Sanger’s chain termination method is the whole-genome shotgun sequencing. However, instead of a single gene or few base pairs, the present method is powerful enough to sequence the entire genome of an organism.

The principle of the shotgun is the same as Sanger’s method, one additional step of DNA fragmentation allows to read multiple fragments.

The entire genome of an organism is fragmented with the help of endonuclease enzymes or by the mechanical techniques. After that, the smaller fragments of DNA sequenced individually into the machine.

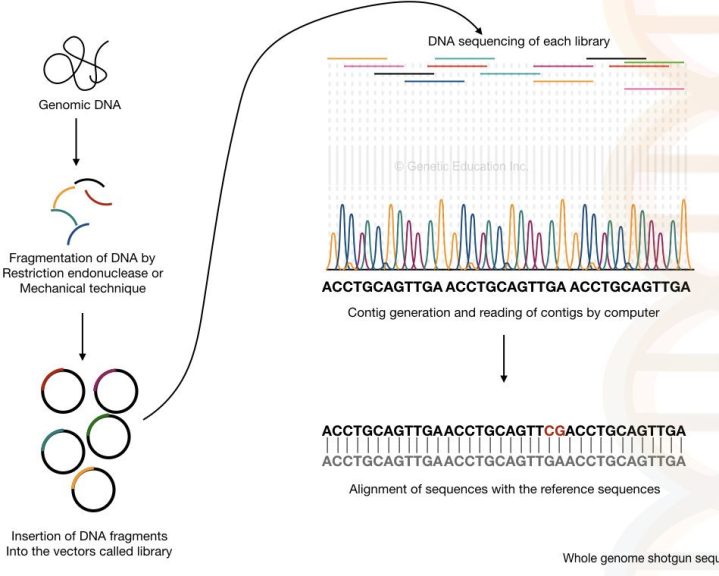
The computer-based software analyses each and every overlapping fragment and reassembled it to generate the complete sequence of the entire genome.

The method can be divided into four steps:

1. Fragmentation of a DNA: with the help of restriction endonucleases or physical method
2. Formation of libraries of the subfragments: the fragments are ligated in vectors and an entire library for various vectors are generated
3. Sequencing the subfragments: each library is sequenced individually.
4. Generation and reading the contigs: the overlapping fragments called contigs are read by the computer.

The fragments generated by the lysis or restriction digestion are around 2 to 20kb.

Importantly, the shotgun sequencing reads both the sequences (it sequence the double-stranded DNA) based on that contigs data, it identifies the gaps that remained unsequenced. The brief overview of present method is given below,



*Whole-genome shotgun sequencing method.*

***Do you know?***

***The shotgun sequencing concept was originally discovered by the Sanger F and his colleagues for sequencing the whole genome.***

**At our glance:**

The present method is faster and cheaper than the previous technique.

The technique becomes more aggressive if any reference sequence is available to align (it can find out the gaps and mutations as fast as possible).

Gene or chromosomal mapping and tedious enzymatic reaction steps are not required in the shotgun sequencing method.

It has the power to sequence the whole genome of an organism (more time required to sequence the complex genome such as humans and animals).

The major portion of the human genome is made up of non-functional repetitive DNA thus it is very difficult for shotgun sequencing to assembled the repetitive DNA sequences. And it is nearly impossible if any reference genome is not available.

The technique solely depends on computational analysis. A huge, powerful, supercomputer is required to work efficiently.

The gap can’t be further filled by the techniques.

The technique becomes feeble if reference genome sequence data is not available.

***“In the year 1981, The genome of cauliflower mosaic virus was sequenced by the shotgun sequencing method.”***

http://www.ftjcfx.com/image-9019789-11745480

1. [An Introduction To Genome-Wide Association Study (GWAS)](https://geneticeducation.co.in/an-introduction-to-genome-wide-association-study-gwas/).

**Clone by clone sequencing:**

For sequencing the whole genome, the clone by clone method is also used instead of whole genome shot-gun sequencing. notably, this present method had helped to successfully complete the human genome project.

The present method is more or less similar to the previous method with one additional step. first, instead of smaller fragments, large clumps of DNA fragments are constructed. through [gene mapping](https://geneticeducation.co.in/a-brief-introduction-to-gene-mapping/), the location of each fragment is noted.

Using the BACs- bacterial artificial chromosome, multiple copies of each fragment are generated for accurate sequencing. In the next step, every copied fragment are fragmented into smaller piece and inserted into the vectors.

The sequencing is performed as the shotgun and the overlapping fragments are assembled by the computer.

In the last step, the data of the gene or chromosome mapping generated previously employed to assemble the sequence. the sequences are arranged on each chromosome based on their location.

**At our Glance:**

Any gaps in sequencing can be identified easily because the fragment of the DNA is taken from the known locations.

Scientists can work on various chromosomes at once and hence a large number of DNA can be sequenced at the same time.

Techniques such as mapping, cloning, and restriction digestion make the present method more tedious and time-consuming than previous ones.

because of this reason, the present method is costlier.

Cloning the chromosomal part such as telomeres and centromeres are difficult to achieve which makes this technique even more toughest.

***“During the year, 1980 and 1990, the genomes of the nematode worm, C. elegans and the yeast, S. cerevisiae was sequenced using the clone by clone sequencing.”***

As the clone by clone sequencing method was used during the human genome project, it is a part of our history. And hence, its discovery was one of the valuable milestones in the history of genetics.

Interestingly, the shotgun sequencing method was evolved from the clone by clone sequencing method.

Read more:

* [DNA digital data storage](https://geneticeducation.co.in/dna-digital-data-storage/)
* [Agarose gel electrophoresis](https://geneticeducation.co.in/agarose-gel-electrophoresis/)

Every platform, we discussed here, have some drawbacks and limitations. Now, at last, the question strike in mind that, is any sequencing platform available which is fast, accurate, reliable, accurate and cost-effective? The answer is “yes”.

The next-gen sequencing, the next level of revolution in sequencing technology.

**Next-generation sequencing:**

The next-generation sequencing platform is different from the Sanger technique or chain termination method of DNA sequencing. Broadly, it amplifies millions of copies of a particular fragment in a massively parallel fashion and the “reads” are analyzed by the computational program.

The NGS process is a bit complex, However, it can be divided into 4 different steps:

1. Library preparation
2. Cluster generation
3. DNA sequencing
4. Data analysis

**1. Library preparation:**

The library preparation is the combination of two reactions viz, fragmentation and ligation. The fragmentation of cDNA or DNA fragments is done by restriction digestion. After that, the smaller DNA fragments are ligated with the known DNA sequence.

The known DNA sequences are called the **adapters** and the process is called adapter ligation. Once the adapters are ligated, the library of the smaller DNA fragments is generated.

The unbounded DNA fragments are washed by the washing buffer, the process of library preparation is called as **tagmentation.**

**2. Cluster generation:**

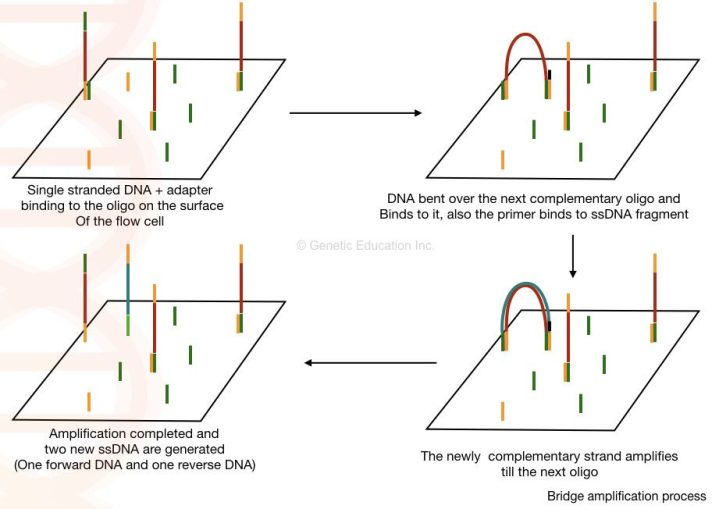
What is the role of adapters??

The short oligonucleotide sequences are immobilized on the solid surface which is complementary to our adapter sequences.

Once the library of our fragmented DNA is loaded into the cell, it is bound with the immobilized oligos on the solid surface and by the bridge amplification, the cluster of the DNA sequence is generated.

Here, in the bridge amplification, the DNA fragments bend over and bind to the next oligo which creates a bridge. A primer binds to this DNA sequence and amplified vertically.

Two new single-stranded DNA is generated by bridge amplification. See the image below,

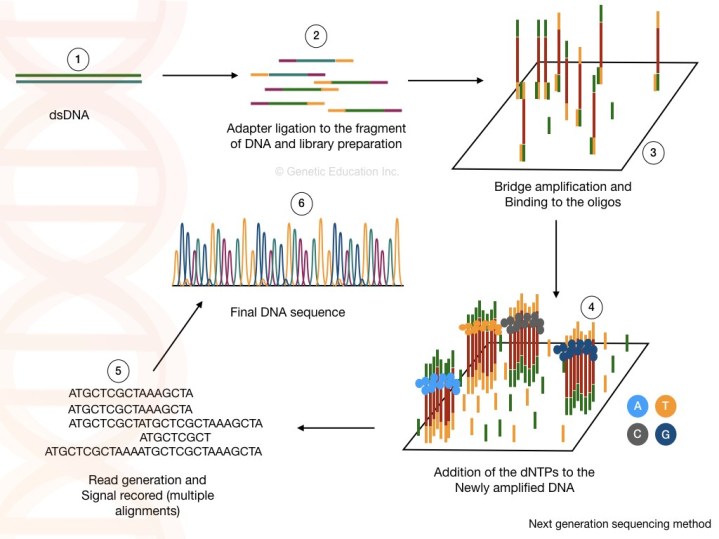


**3. Sequencing:**

As the polymerase adds the nucleotide into the bridge amplification, the amplification signals are recorded each time. This will generate multiple sequencing databases for the DNA sequence.

**4. Data analysis:**

The read generated by the sequencing can be aligned to the reference genome sequence and by doing this we can identify any addition, deletion or variation into the sequence.



**At our glance:**

The NGS is the most advanced, fast, accurate and 100% effective technique in modern-day science.

These are the common types of sequencing platforms used in the genomic labs. Besides this, several other sequencing methods are:

* Single-molecule real-time (RNAP) sequencing
* Single-molecule SMRT(TM) sequencing
* Helioscope(TM) single-molecule sequencing
* DNA nano ball sequencing
* SOLiD sequencing
* Illumina (Solexa) sequencing
* Polony sequencing
* massively parallel signature sequencing (MPSS)
* High throughput sequencing

**Applications of DNA sequencing**

In medical science, DNA sequencing can be used in the identification of genes responsible for hereditary disorders. New mutation can also be detected with the help of the DNA sequencing

In forensic science, it is used for parental verification, criminal investigation and identification of individuals through any of the available samples such as hair, nail, blood or tissue.

In the agriculture industries, identification of GMO species can be possible with the help of the DNA sequencing methods. Any of minor variations in the plant genome can be detected with the help of DNA sequencing.

It is used to construct maps such as whole chromosomal maps, restriction digestion maps, and genome maps. Read more on mapping: [A Brief Introduction to “Gene Mapping”](https://geneticeducation.co.in/a-brief-introduction-to-gene-mapping/).

Open reading frames, non-open reading frames and protein-coding DNA sequences can be identified by the present method.

DNA sequencing is used in exon/ intron, repeat sequence and tandem repeat identification and detection.

Furthermore, the present method is employed in gene manipulation and gene editing. New variations in nature can also be determined through sequencing.

Metagenomic studies are nowadays possible by sequencing methods such as pyrosequencing.

It is further used in the [Microbial identification](https://geneticeducation.co.in/microbial-genetics-a-rapid-advancement-in-microbiology/) and study of the new bacterial species. The sequencing technique advances the microbial identification by eliminating the traditional and time consuming culturing methods.

Nowadays microbial identification and characterization become more rapidly and accurately done using sequencing. By comparing the sequence of the target microbes with the available data, scientists can identify new mutations and new strains.

The sequencing techniques specifically, the NGS has a great application in the oncology and cancer studies. various cancer-causing genes are identified and characterized by the present method.

The advancement in evolutionary studies is only possible because of DNA sequencing. By comparing various genes and sequences, an evolutionary map can be generated. Also, new variations through evolution can be encountered.

sequencing helps in studying the asymptomatic high-risk population, prior to the occurrence of disease. And thus preventive steps can be taken earlier.

***Do You Know?***

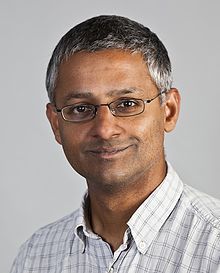


Image credit: https://en.wikipedia.org/wiki/Shankar\_Balasubramanian

***Sir Shankar Balasubramanian****is the only known Indian scientist who was involved in the development of a sequencing platform. He was the principal investigator in the development of Solexa/ Illumina Next-generation sequencing.*

**Limitations of DNA sequencing**

A sequencing platform is a computer algorithm based assistive technique that relies on computational data processing. For that, a huge, high-speed supercomputer is required.

Also, several sequences like tandem repeats, repetitive DNA, fragmented genes, and other duplicated regions can not be studied properly. The chance of errors in the pre-sample processing can cause big economical loss, these two are the major limitations of the DNA sequencing methods.

**Conclusion:**

Scopes of DNA sequencing in various fields are infinite. The biggest hope of it is its use in the diagnosis of multigenic genetic disorders. The present method can be a useful tool for prenatal and preimplantation genetic studies.

In the latest scientific trends, the development of gene therapies and the concept of personalized medicines can not be fulfilled without the NGS like robust techniques. However, more optimization and advancements are required to reduce the cots, time-duration and error rate in the [present genetic technology](https://geneticeducation.co.in/present-genetic-technologies-an-overview/).

**Source:**

1. Brown TA. Genomes. 2nd edition. Oxford: Wiley-Liss; 2002. Chapter 6, Sequencing Genomes.
2. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 2016;107(1):1-8.
3. Lehrach H. DNA sequencing methods in human genetics and disease research. *F1000Prime Rep*. 2013;5:34. Published 2013 Sep 2.