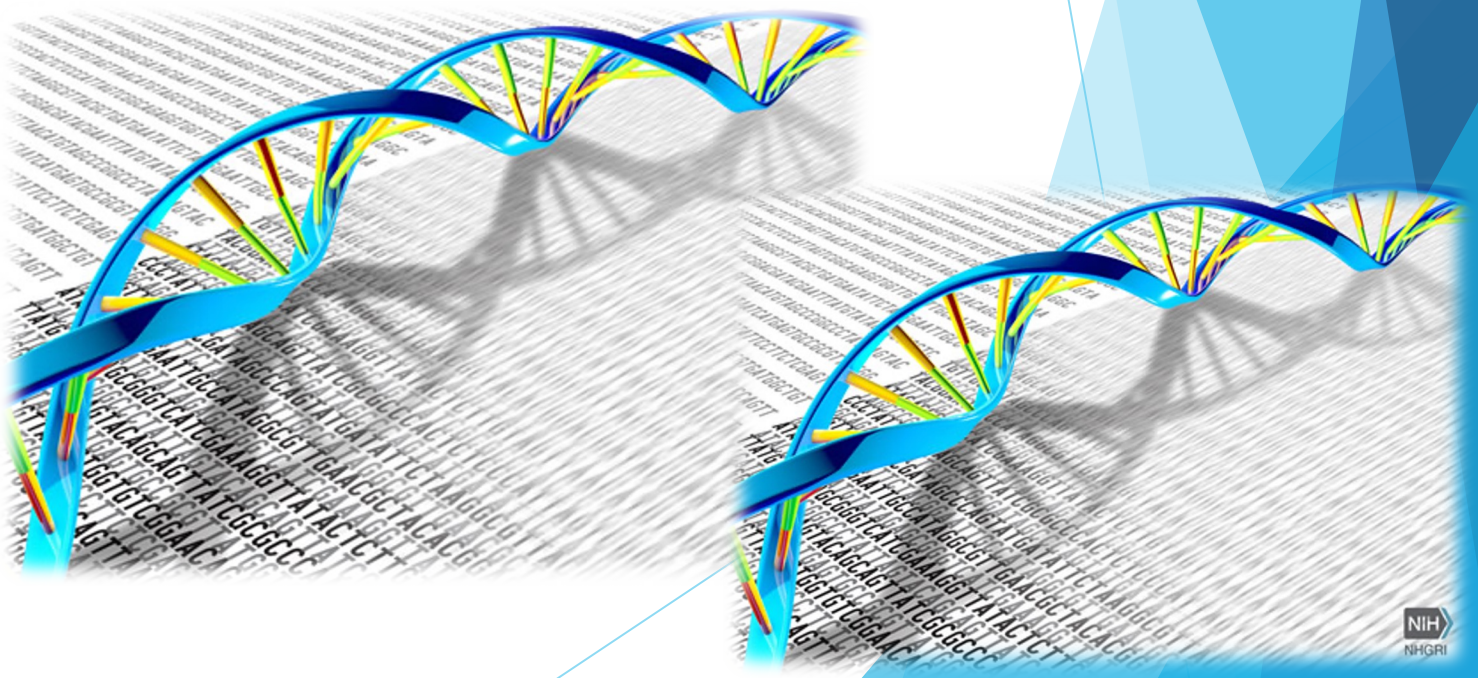
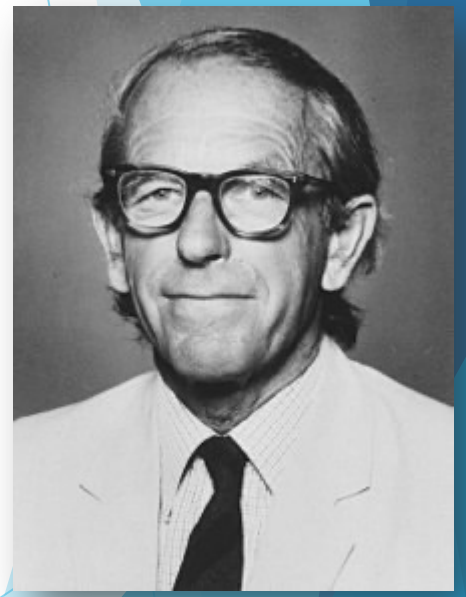
DNA Sequencing



History of DNA sequencing

- The sequencing of DNA molecules began in the 1970s with development of the **Maxam-Gilbert method**, and later the **Sanger method.**
- Originally developed by **Frederick Sanger** in 1975, most DNA sequencing that occurs in medical and research laboratories today is performed using sequencers employing variations of the Sanger method.



Introduction

- The term DNA sequencing refers to sequencing methods for determining the order of the nucleotide bases - adenine, guanine, cytosine, and thymine - in a molecule of DNA. Knowledge of DNA sequences has become indispensable for basic biological research, other research branches utilizing DNA sequencing, and in numerous applied fields such as:
- Diagnostic,
- Biotechnology,
- Forensic Biology And
- Biological Systematics.

Different Methods for DNA Sequencing

- **Basic Methods:**
  - Maxam-Gilbert sequencing
  - Chain-termination methods
- **Next-generation methods:**
  - Massively parallel signature sequencing (MPSS)
  - Polony sequencing
  - 454 pyrosequencing
  - Illumina (Solexa) sequencing
  - SOLiD sequencing
  - Ion Torrent semiconductor sequencing
  - DNA nanoball sequencing
  - Single molecule real time (SMRT) sequencing

DNA Sequencing

Two main methods

1. Sanger dideoxynucleotide chain termination method

Commonly used method

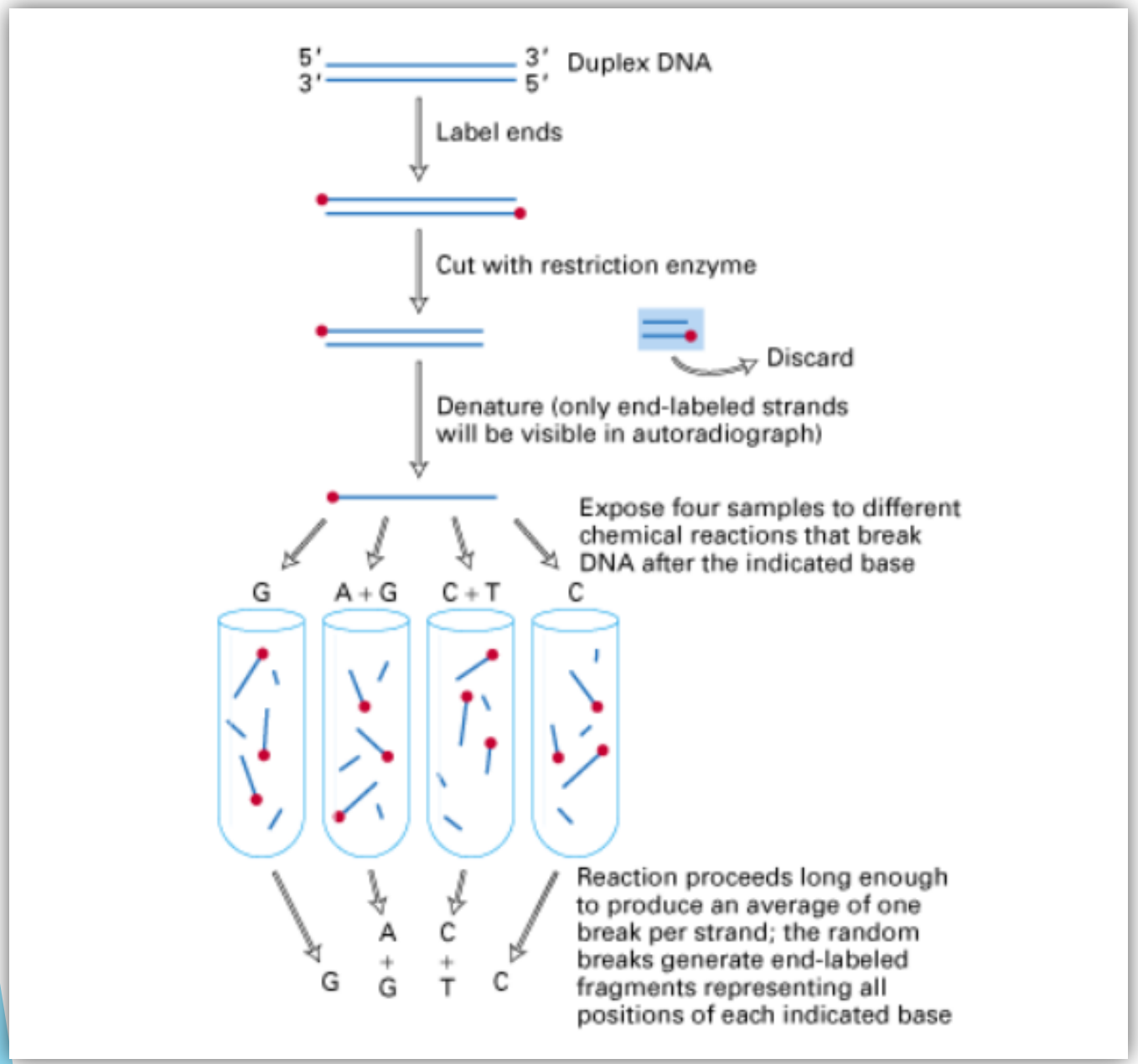
A. **Manual** method

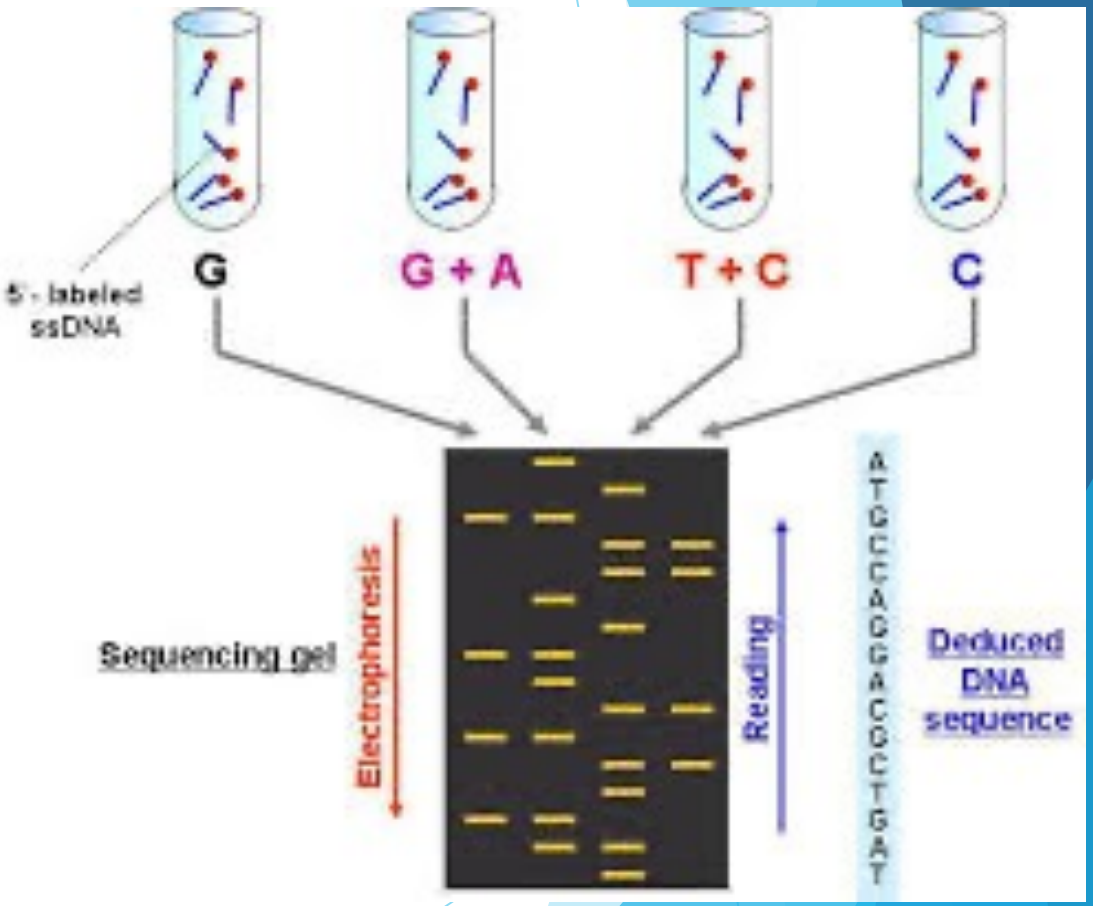
B. **Automated** method

2. Chemical cleavage method (maxam and gilbert method)

Not used nowadays

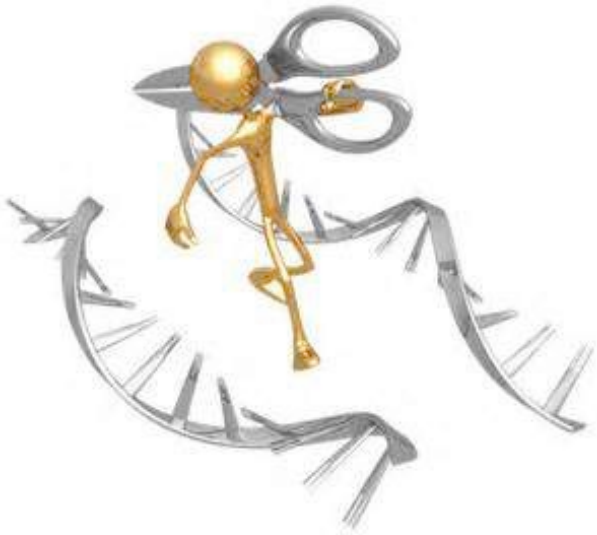
Maxam-Gilbert method





Restriction enzymes

- In the 1960, scientists discovered that bacteria have enzymes that cut, or “digest”, the DNA of foreign organisms and thereby protect the cells from invaders such as viruses.
- Scientists have now isolated several hundred of these enzymes, known as restriction enzymes, or restriction endonucleases.
- Each is able to recognize and cut at a specific DNA sequence, known as a recognition sequence.



Restriction enzymes cont..

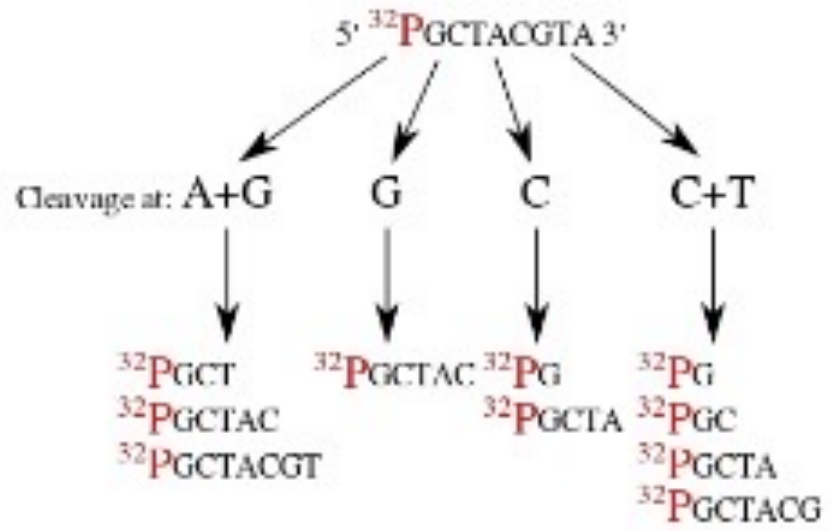
- The discovery of restriction enzymes made genetic engineering possible because researchers could use them to cut DNA into fragments that could be analyzed and used in a variety of procedures such as gene cloning, in genotyping…….etc.

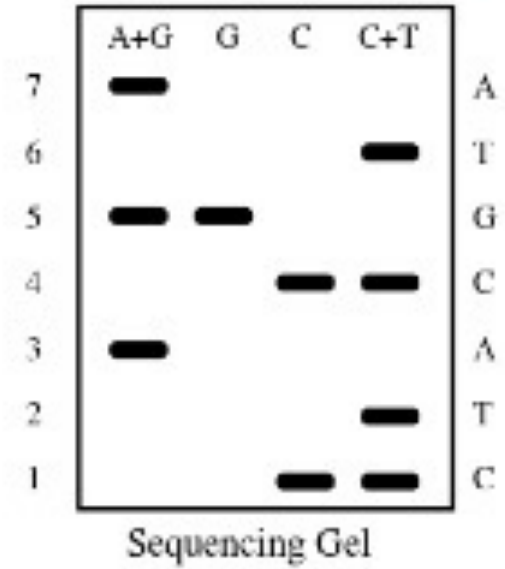
Definition

- A restriction enzyme (or restriction endonuclease)

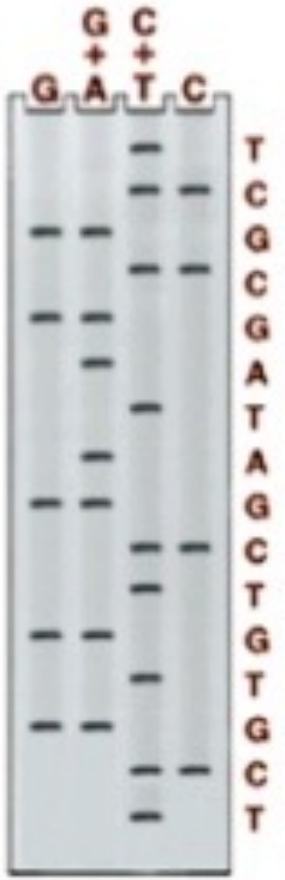
Is an enzyme that cuts double-stranded or single stranded DNA at specific recognition nucleotide sequences as restriction sites.

Maxam-Gilbert method





Sequencing Gel

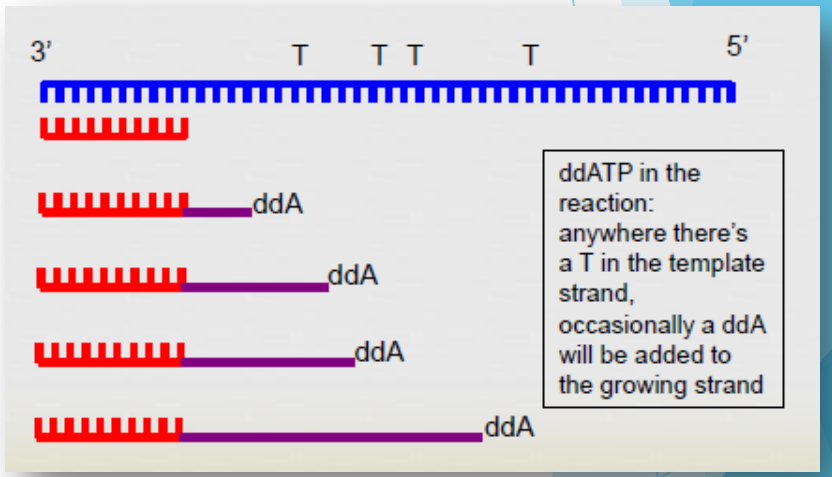


Manual Sanger method sequencing

- It is method to find out the nucleotides Sequence of unknown DNA strand.
- More recently, Sanger sequencing has been upgraded as "Next-Generation“ sequencing methods, especially for large scale genome analyses and for obtaining especially long DNA sequence reads (>500 nucleotides).

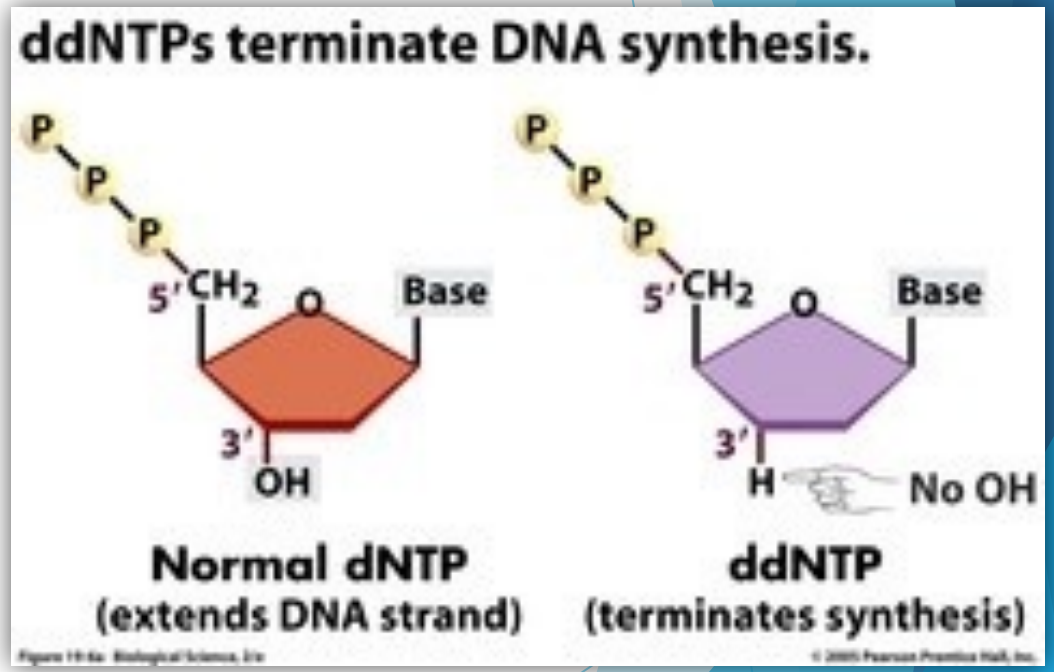
Basic Principle

- This method generally is an In-Vitro synthesis of DNA strand and by using terminators (di-deoxynucleotide) the growing strand terminates at specific site.
- Upon termination the strands are overlap to got original sequence of unknown DNA Strand.



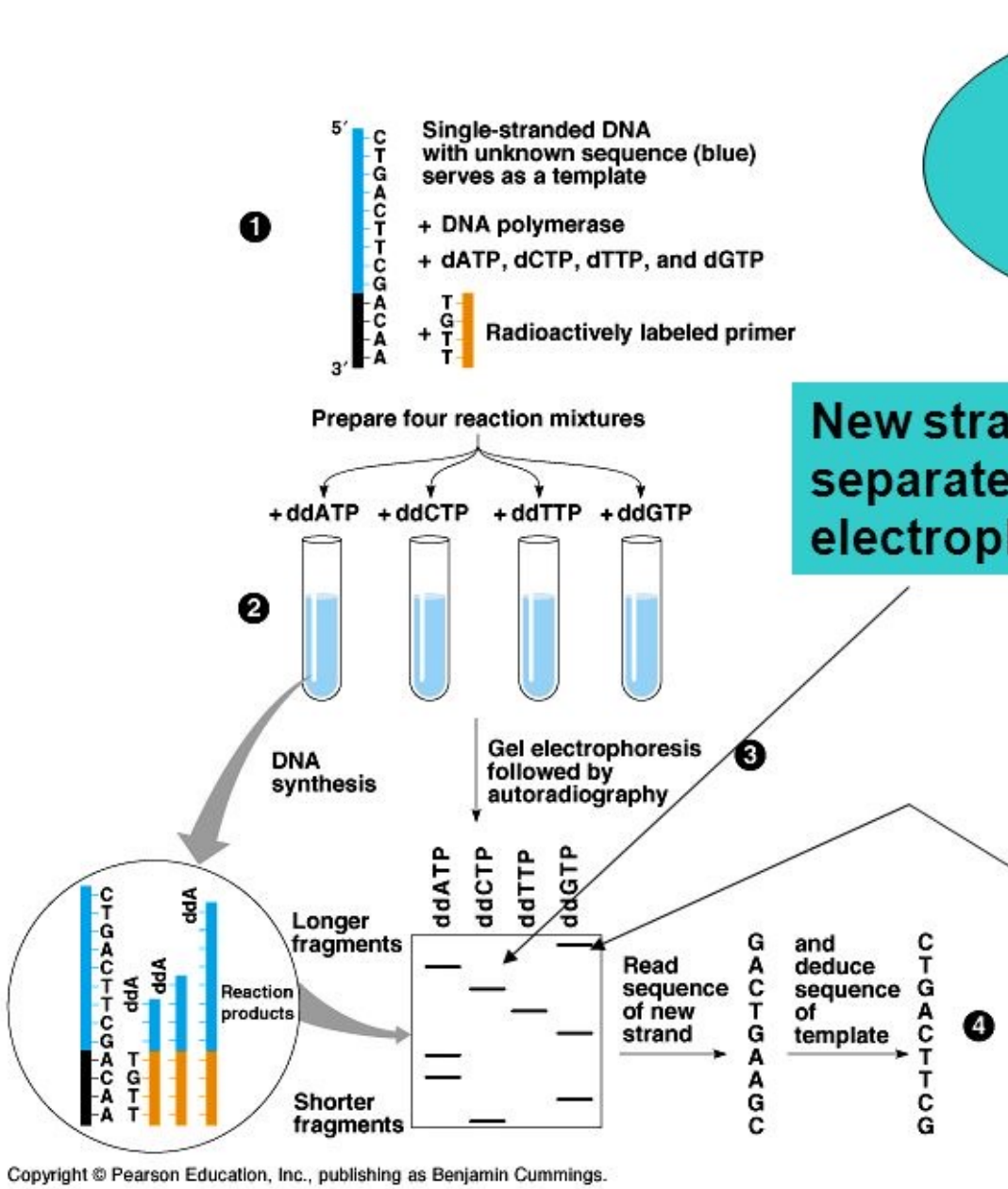
Requirements

- Single Stranded template
- Primers
- DNA polymerase
- Di-Deoxynucleotide
- ✓ The 3′-OH group necessary for formation of the phosphodiester bond is missing in ddNTPs)
- ✓ Every nucleotide have its specific ddNTP form i.e., ddATP, ddGTP…etc



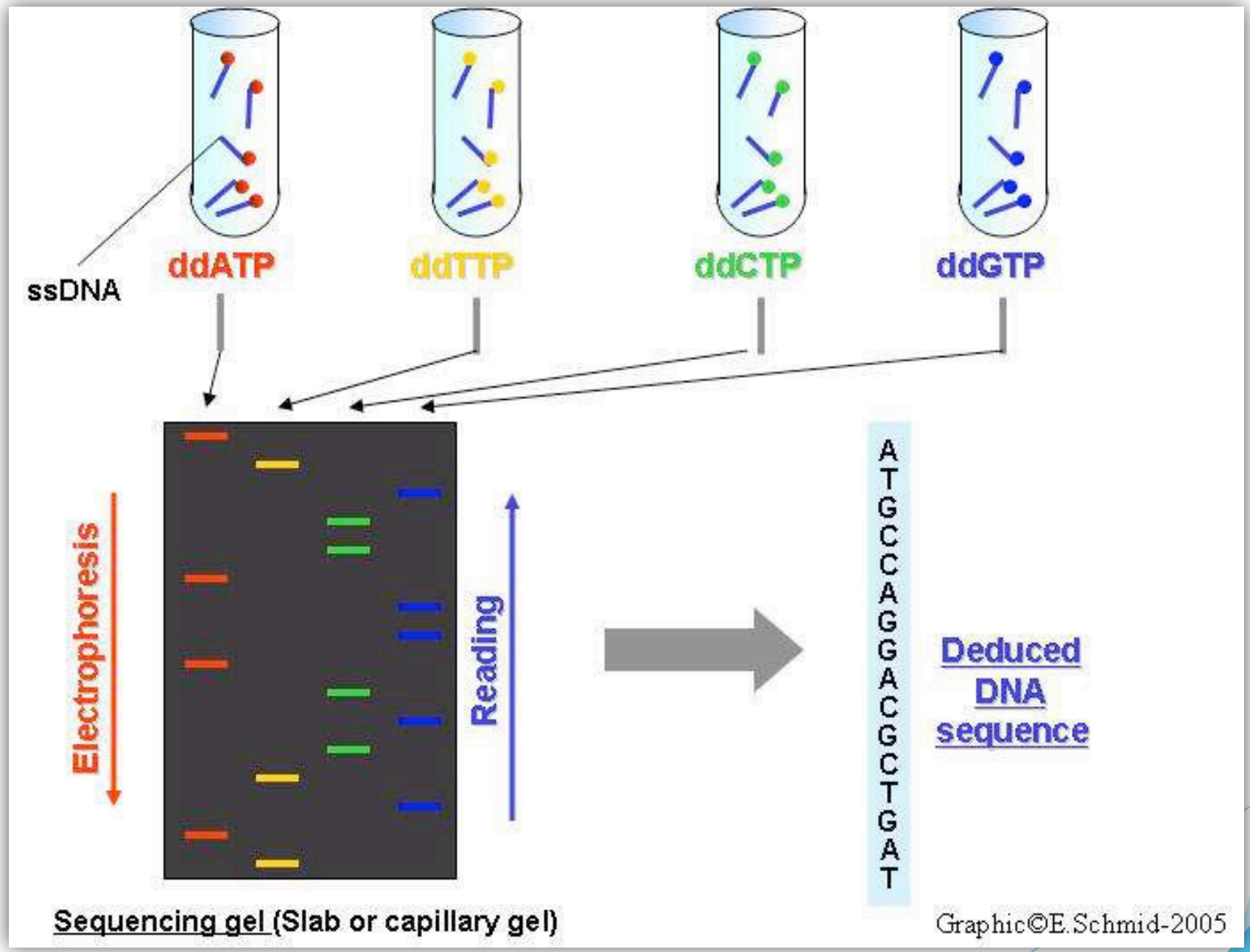
Procedure

- **Steps**:
- 1. Denaturation
- 2. Primer attachment and extension of bases
- 3. Termination
- 4. Gel electrophoresis

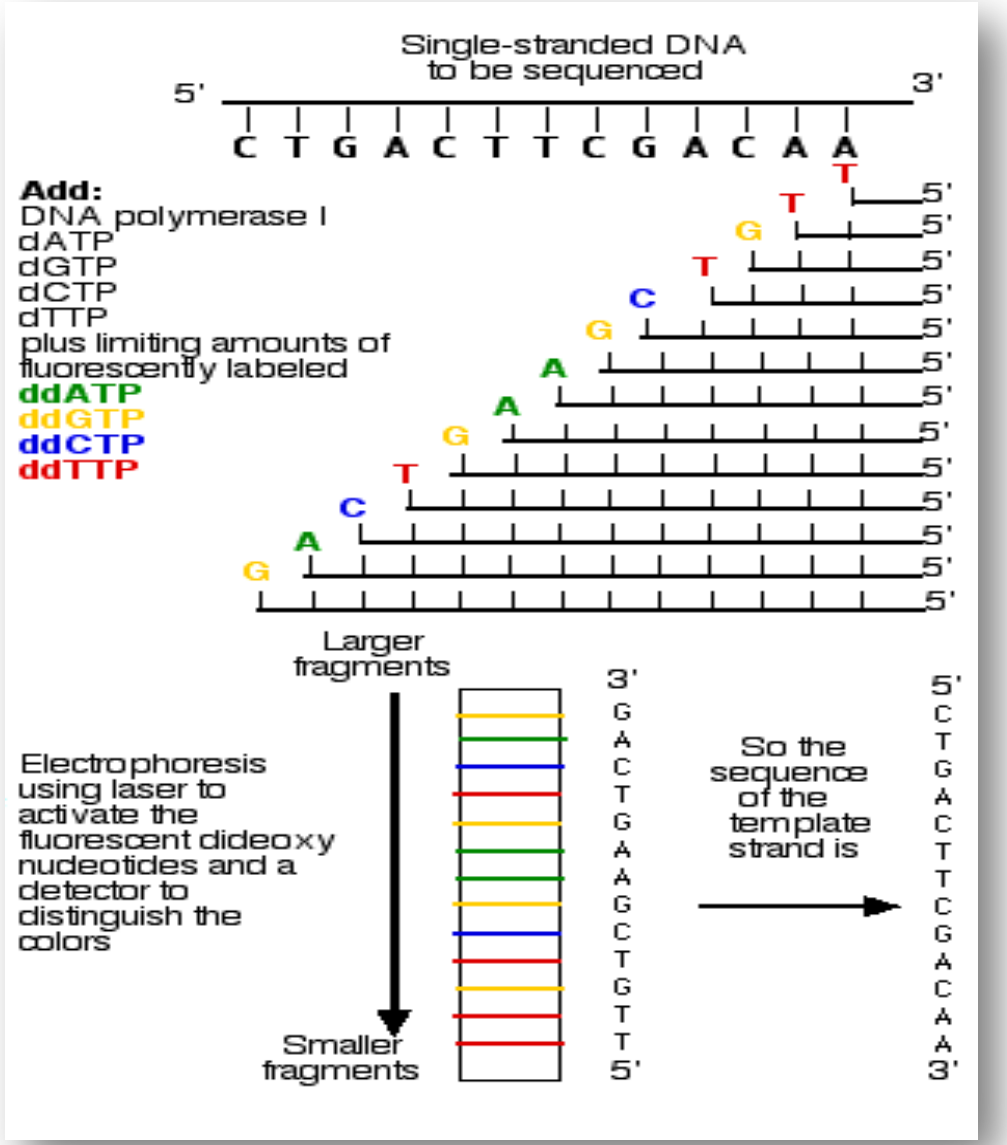


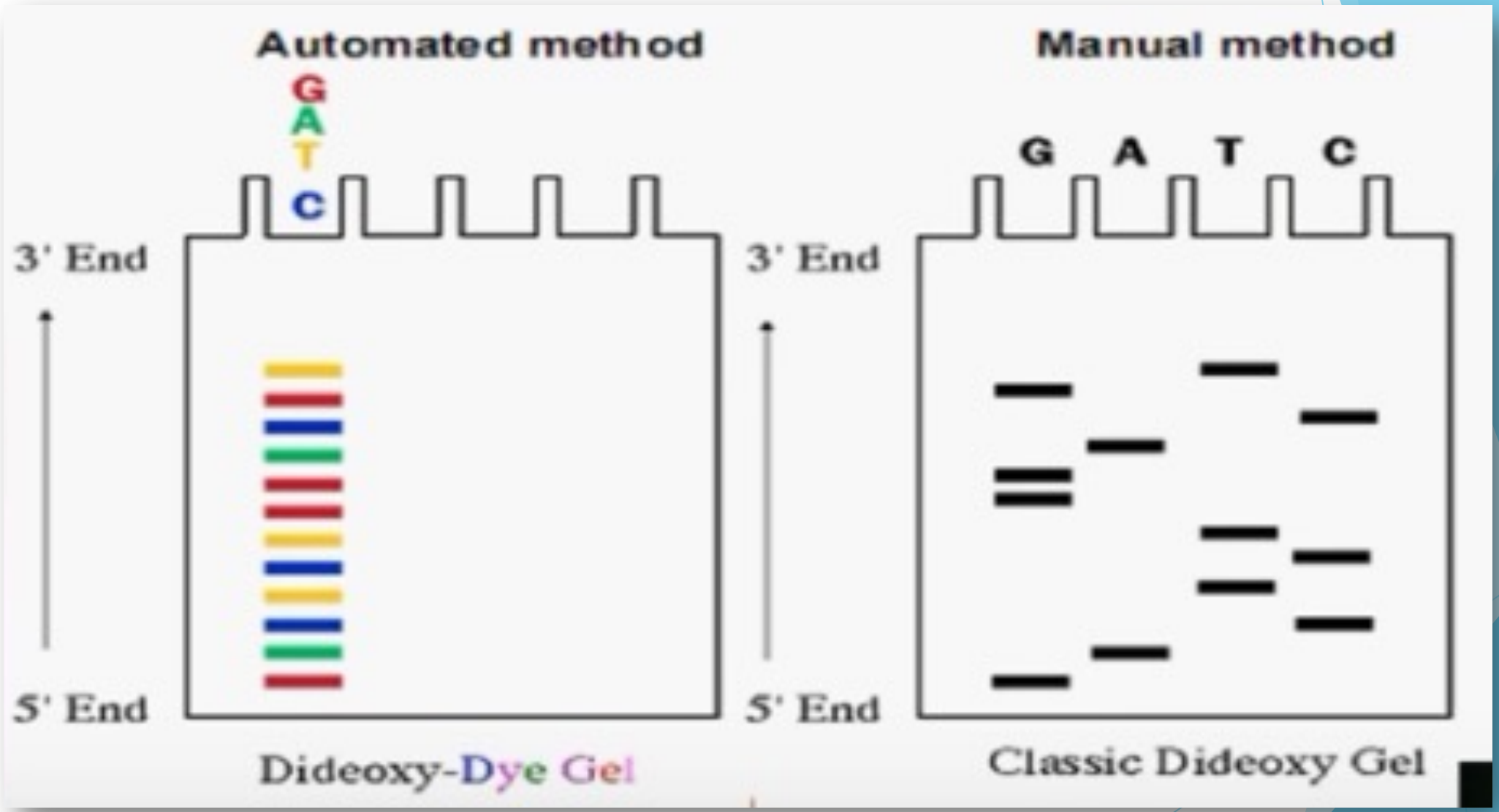
Procedure cont..

- The DNA template is treated with heat so that it becomes single stranded
- A short, single-stranded primer which is radioactively labelled is added to the end of the DNA template
- Add template DNA and primer in 4 Tubes.
- Now add ddNTPs In tubes in the way that single tube contain one type of ddNTP.
- Extension is start and band formed of various sizes.
- The fragments of DNA are separated by electrophoresis.
- Overlap these sequences to find out sequence of Target DNA.

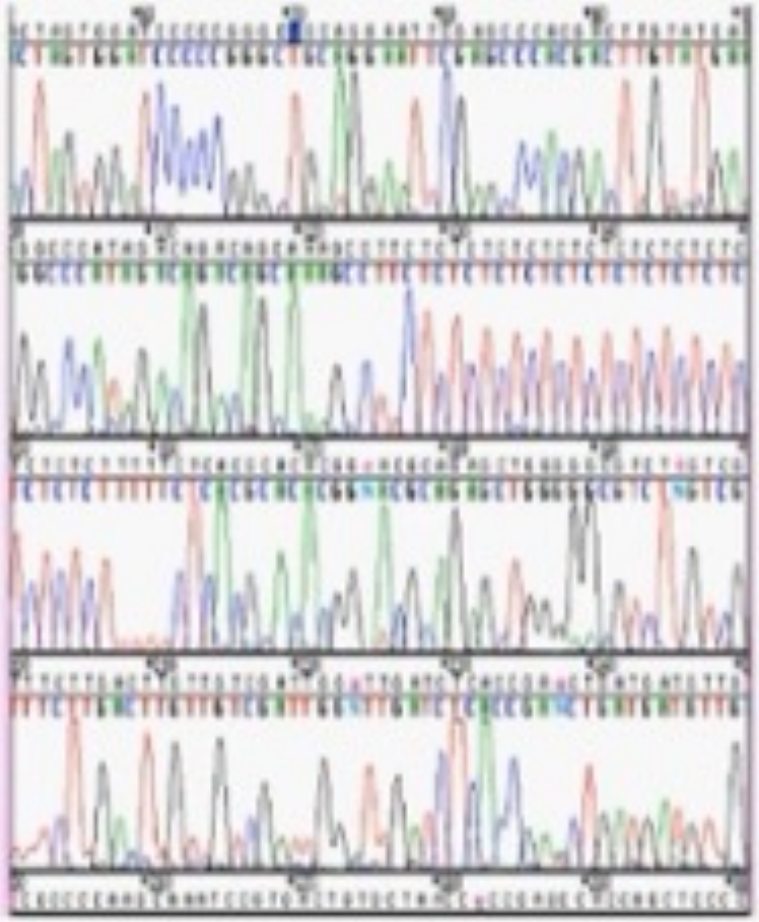

ssDNA
ddATP
ddTTP
ddCTP
ddGTP
Electrophoresis
Reading
ATGCCAGGACGCTGAT
Deduced DNA sequence
Sequencing gel (Slab or capillary gel)
Graphic©E.Schmid-2005

Automated Sanger method sequencing

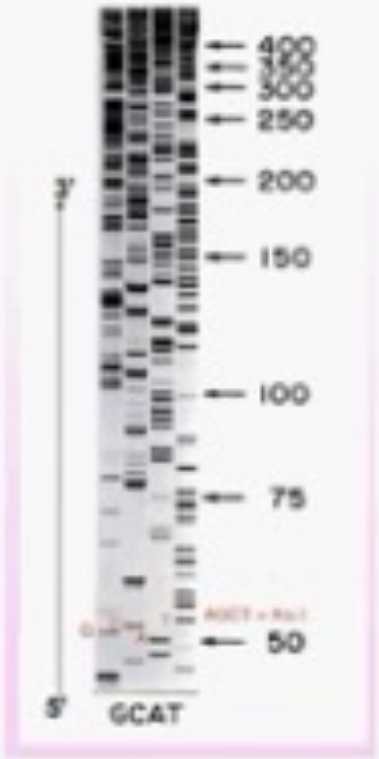



Automated method
G
A
T
C
3' End
5' End
Dideoxy-Dye Gel
Manual method
G A T C
3' End
5' End
Classic Dideoxy Gel

Automated DNA sequencing



Manual DNA sequencing

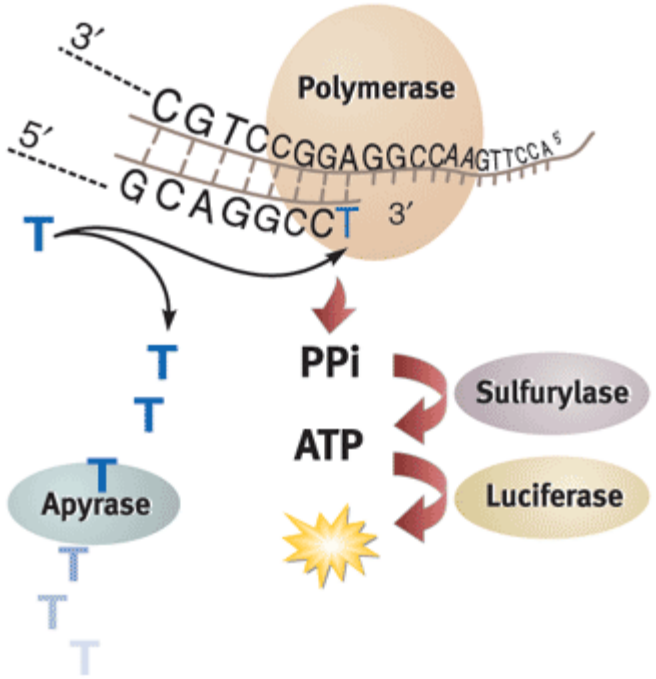


Automated Vs Manual

- **Automatic** sequencing ...
- is far less error-prone than manual sequencing,
- makes assembly WORLDS easier,
- is almost always cheaper than manual sequencing,
- produces more data than manual gels,
- can process more samples in a short period than you could possibly do manually.

WHAT IS PYROSEQUENCING

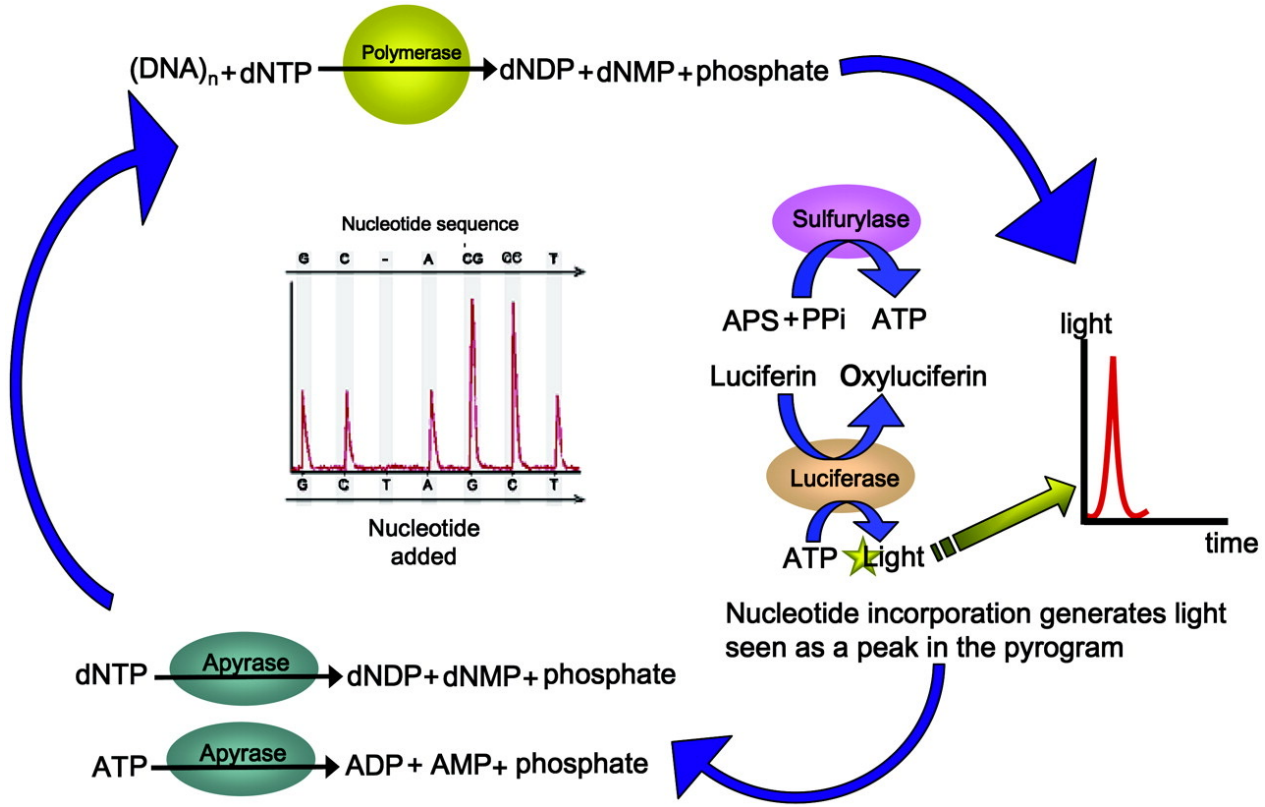
- Pyrosequencing is a method of sequencing that relies on light emitted in an enzymatic reaction set in motion by the release of pyrophosphate when a base is added during the sequencing process



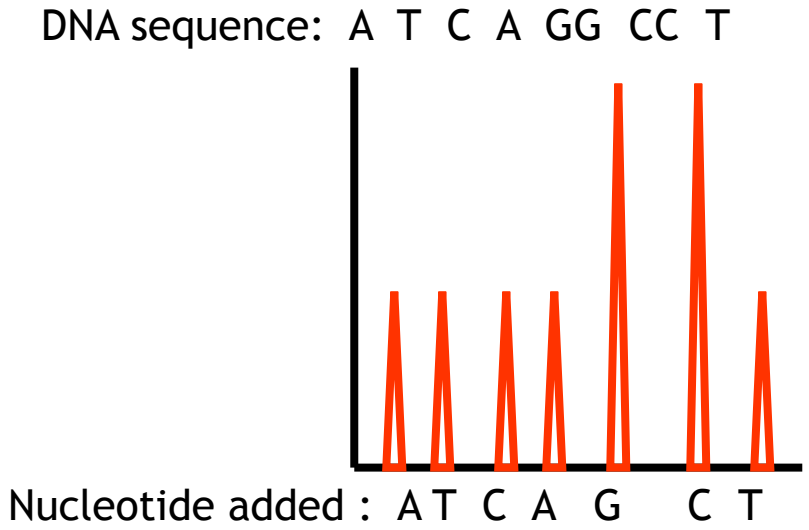
HISTORY

- The technique was developed by Mostafa Ronaghi and Pål Nyrén at the Royal Institute of Technology in Stockholm in 1996.
- A second solution-based method for Pyrosequencing was described in 1998 (In this alternative method, an additional enzyme apyrase is introduced).
- A third microfluidic variant of the Pyrosequencing method was described in 2005 by Jonathan Rothberg.
- This allowed for high-throughput DNA sequencing and an automated instrument was introduced to the market. This became the first next generation sequencing instrument starting a new era in genomics research

- Pyrosequencing is based on the generation of light signal through release of pyrophosphate (PPi) on nucleotide addition.
  - $DNA_n + dNTP \rightarrow DNA_{n+1} + PP_I$
- PPi is used to generate ATP from adenosine phosphosulfate (APS).
  - $APS + PP_I \rightarrow ATP$
- ATP and luciferase generate light by conversion of luciferin to oxyluciferin.



- Each nucleotide is added in turn.
- Only one of four will generate a light signal.
- The remaining nucleotides are removed enzymatically.
- The light signal is recorded on a pyrogram.



There are two different pyrosequencing strategies that are currently available:

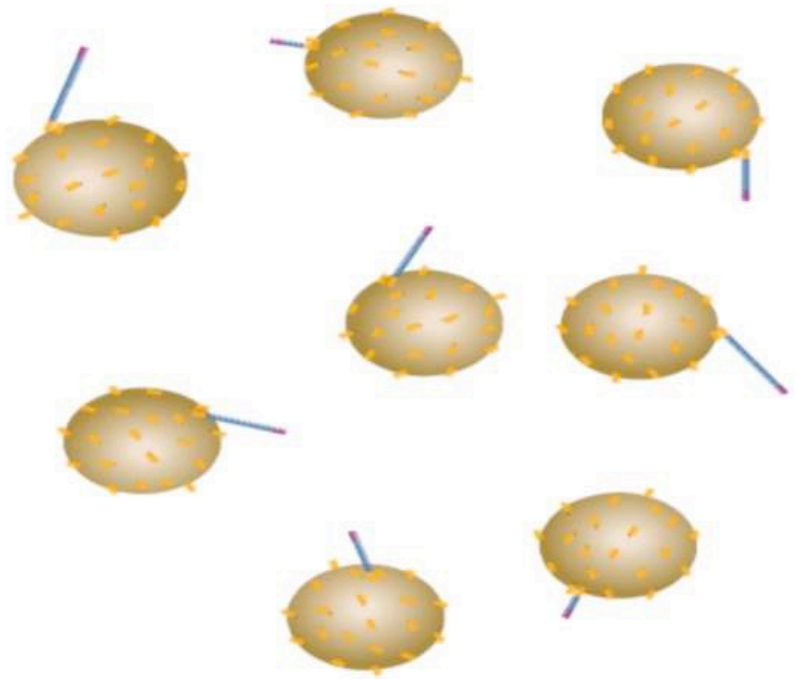
- **Solid-phase pyrosequencing**
  - The solid-phase method is based on a combination of the sequencing-by-synthesis technique and a solid-phase technique. The four nucleotides are dispensed sequentially in the reaction system and a washing step removes the unincorporated nucleotides after each addition.
- **Liquid-phase pyrosequencing**
  - Liquid phase pyrosequencing (apyrase, a nucleotide-degrading enzyme from potato, is introduced to make a four-enzyme system. Addition of this enzyme has eliminated the need for solid support and intermediate washing thereby enabling the pyrosequencing reaction to be performed in a single tube

Solid-phase pyrosequencing

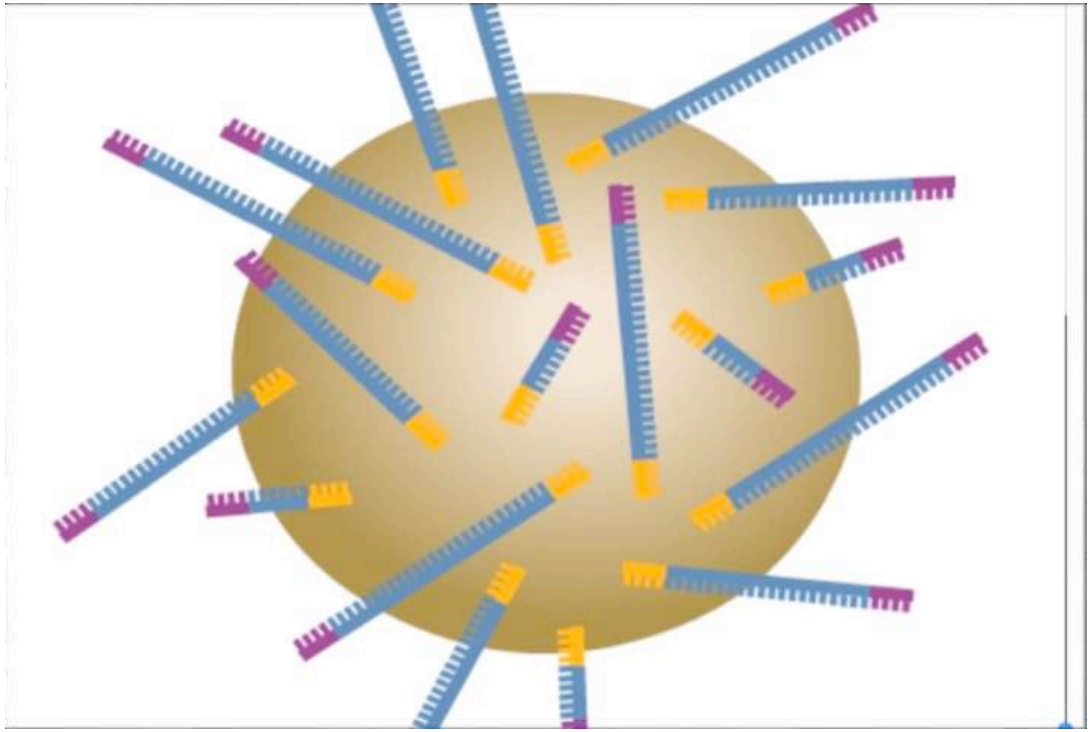
1. Enzyme cocktail. (DNA polymerase, ATP-Sulfurylase, Luciferase)
2. Single dNTPs.
3. APS (Adenosine 5’-phosphosulfate).
4. Luciferin
5. DNA Adaptors.

- Genomic DNA Which is then cut into 300-800 bp fragments as We Start with Double Stranded Genomic DNA (Our Sample).

- Then we add small DNA adaptors to our fragments… Then our double stranded DNA is denatured to a single strand.



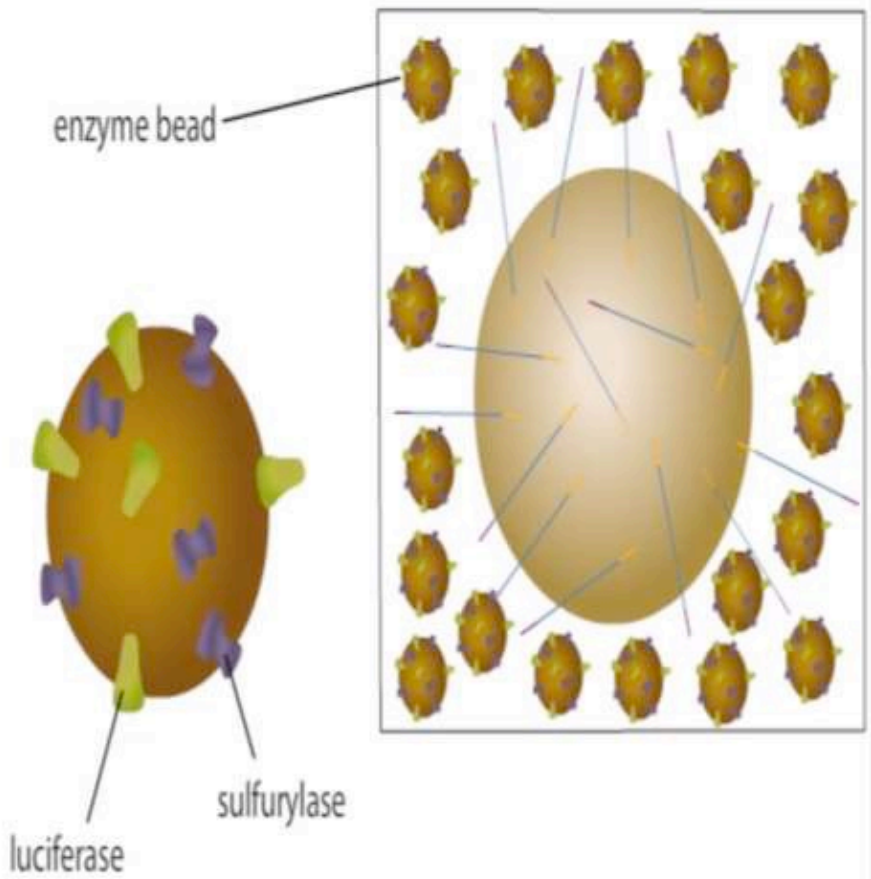
- Then we take tiny micro beads coated with DNA primers complementary to one of the adaptors on the DNA strand.
- Then single stranded DNA binds complementarily with the primer on bead.
- Under conditions that favor one DNA strand per bead.



- PCR is run for amplification, so that enough identical DNA strands are available to analyze.

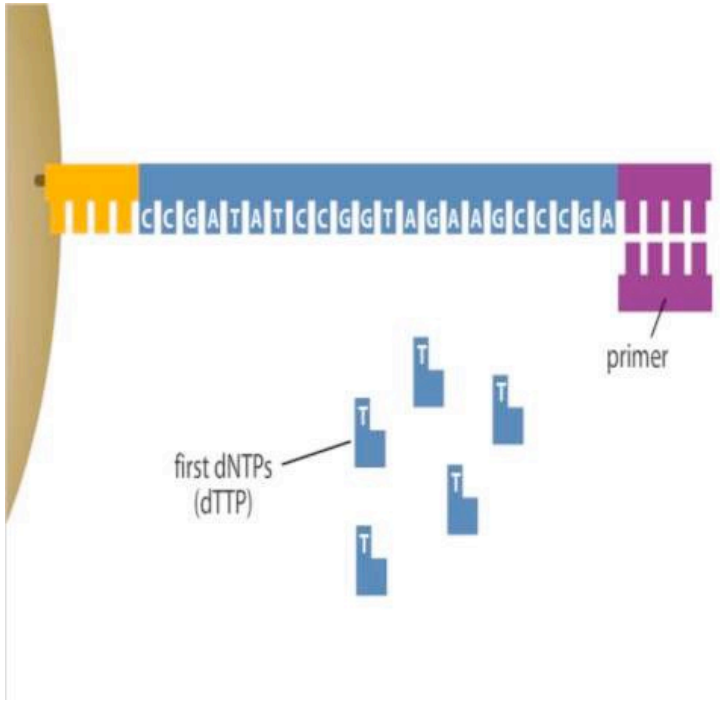
- Then beads are loaded into tiny wells with room for a single bead per well.
- Each well consists of a different amplified DNA fragment.

- Along with a single bead The well contains enzyme beads covered with 2 types of enzymes, Sulferylase & Luciferase.

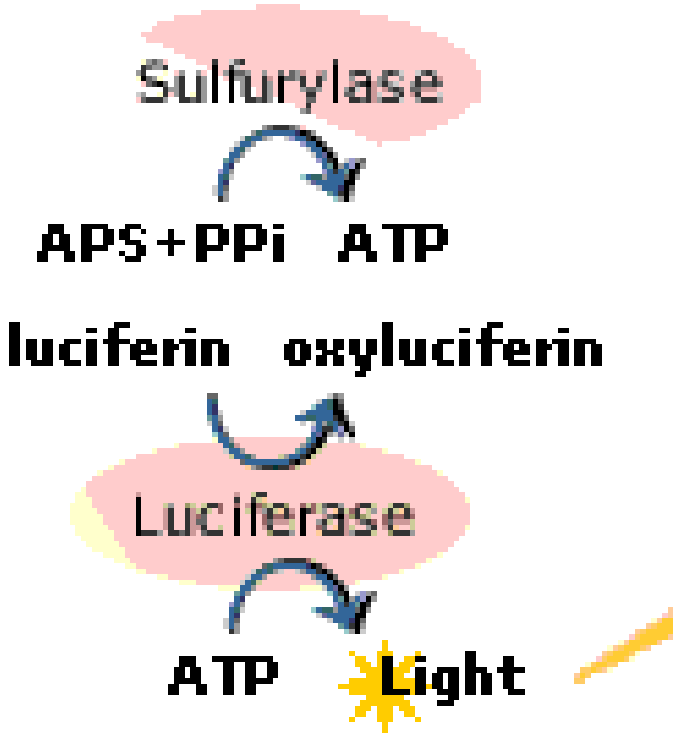


Pyrosequencing Technique:

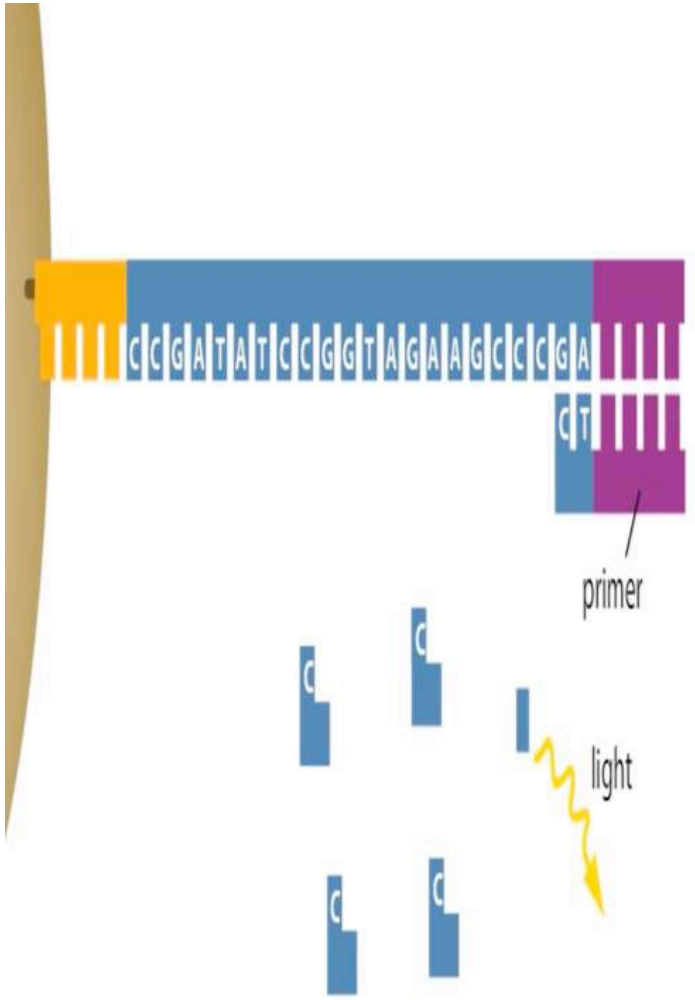
- First a primer is allowed to attach to the DNA.
- Then DNA Polymerase begins to add
- Nucleotides.
- a single type of dNTP is flowed across the wells.



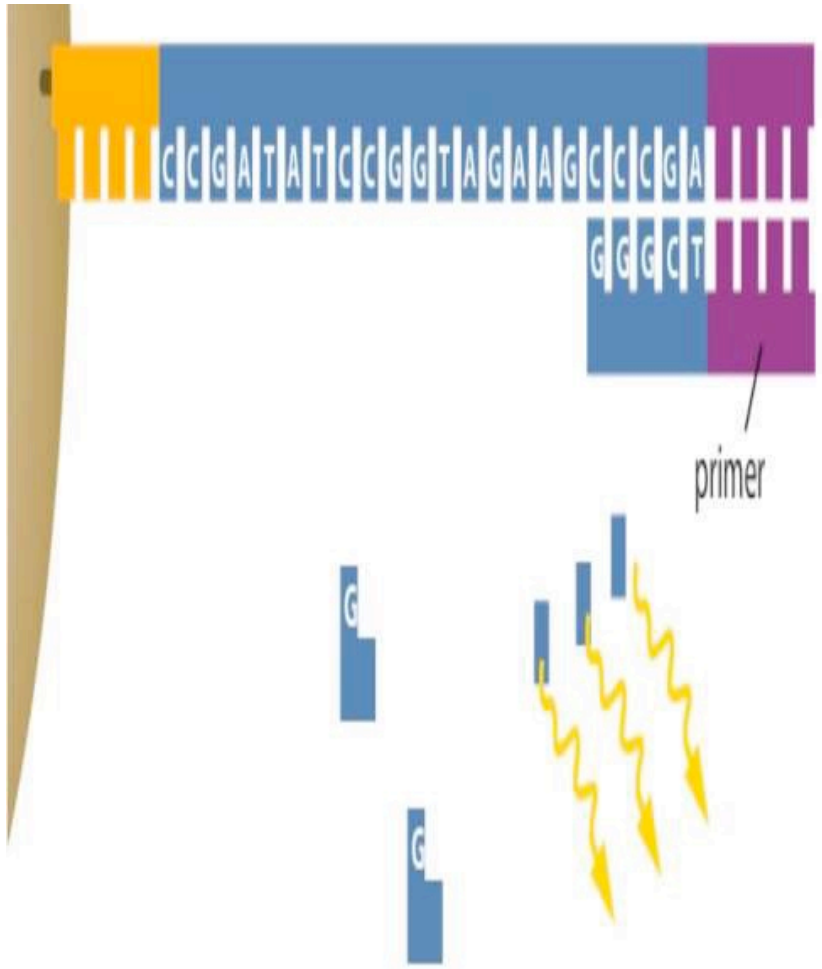
- ATP sulfurylase converts PPi to ATP in the presence of adenosine 5′ phosphosulfate. This ATP acts as a substrate for the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP.



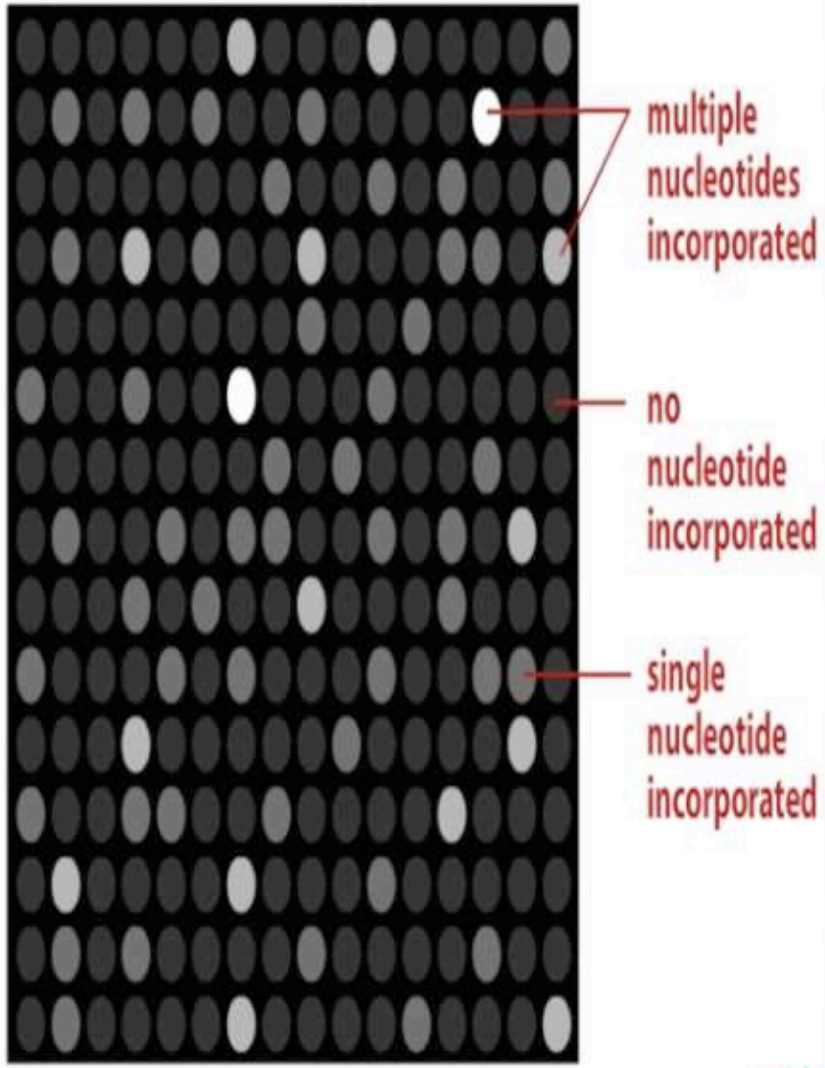
- The light produced in the luciferase-catalyzed reaction is detected by a camera and analyzed in a program, And the nucleotides are washed out of the wells.
- a new nucleotide set is added into the wells. But because Adenine doesn't pair with Guanine so no nucleotide will be added thus no light is emitted.
- The next set of nucleotides bears a cytosine which is incorporated, after which light is emitted too.



- The next step involves a set of dGTPs which are incorporated.
- 3 nucleotides are incorporated and each pyrophosphate ion results in emission of a photon.
- 3 times as much as light is emitted from this incorporation, indicating that 3 Cytosine bases appear consecutively in the template strand.



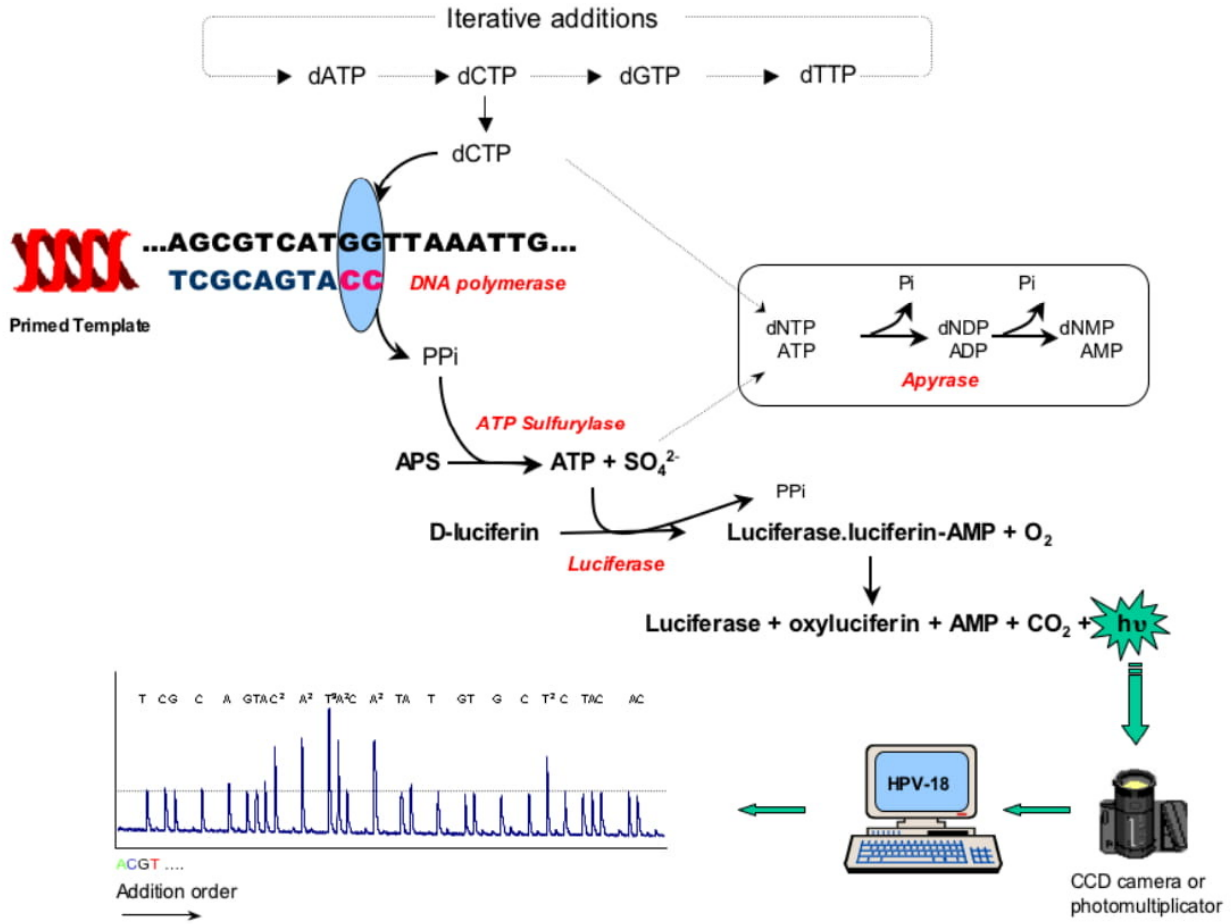
- The 4 dNTPs are flowed sequentially in the wells.
- This grid represents a moment of time in which a single set of dNTPs has been flowed across the wells.



- The data from a single well can be depicted on a chart, showing the level of light emitted from incorporation of each nucleotide.
- Then Software packages assemble these fragments of sequences into longer pieces.
- And in this way determine the overall sequence of the genome.

Liquid-phase pyrosequencing

- Pyrosequencing by the liquid-phase uses nucleotide-degrading enzyme, called apyrase
- The implementation of this enzyme in the Pyrosequencing system excluded the use of solid-phase separation, and that eliminated extra steps such as washes and repetitive enzyme additions.
- The liquid-phase Pyrosequencing method employs a cascade of four enzymes and the DNA sequencing is monitored in real-time.
- The sequencing reaction is initiated by annealing a sequencing primer to a single-stranded DNA template.



sequencing of a partially amplified DNA of human papillomavirus (HPV) by Pyrosequencing technology. The cascade of the fourenzyme catalyzed reactions is demonstrated where APS stands for adenosine 5'- phosphosulphate and hu represents light photons emitted by the bioluminescent reaction.

Alternative Sequencing Methods:

Bisulfite Sequencing

- Bisulfite sequencing is used to detect methylation in DNA.
- Bisulfite deaminates cytosine, making uracil.
- Methylated cytosine is not changed by bisulfite treatment.
- The bisulfite-treated template is then sequenced.

The sequence of treated and untreated templates is compared.

Methylated sequence: GTC ^{Me}GGC ^{Me}GATCTATC ^{Me}GTGCA …

Treated sequence: GTC ^{Me}GGC ^{Me}GATUTATC ^{Me}GTGUA …

DNA Sequence:

(Untreated) reference: …GTCGGCGATCTATCGTGCA…

Treated sequence: …GTCGGCGATUTATCGTGUA…

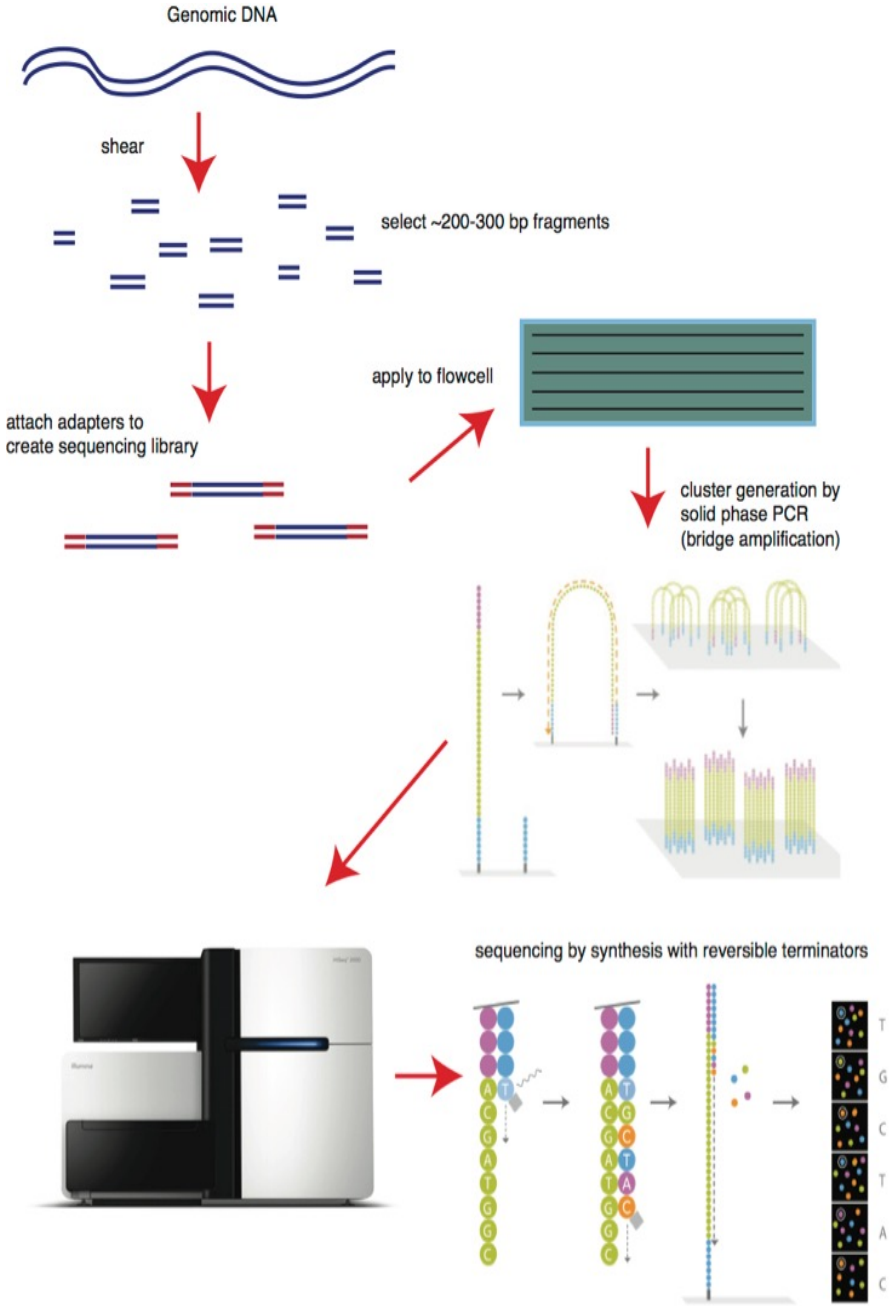
This sequence indicates that these Cs are methylated.

1-Illumina (solexa) sequencing

- Illumina sequencing technology works in three basic steps: amplify, sequence, and analyze.
- The process begins with purified DNA.
- The DNA gets chopped up into smaller pieces and given **adapters, indices**, and other kinds of molecular modifications that act as reference points during amplification, sequencing, and analysis.
- The modified DNA is loaded onto a specialized chip where amplification and sequencing will take place.

- Along the bottom of the chip are hundreds of thousands of oligonucleotides (short, synthetic pieces of DNA). They are anchored to the chip and able to grab DNA fragments that have complementary sequences.
- Once the fragments have attached, a phase called **cluster generation begins.** This step makes about a thousand copies of each fragment of DNA.
- Next, primers and modified nucleotides enter the chip. These nucleotides have reversible 3' blockers that force the polymerase to add on only one nucleotide at a time as well as fluorescent tags.
- After each round of synthesis, a camera takes a picture of the chip.

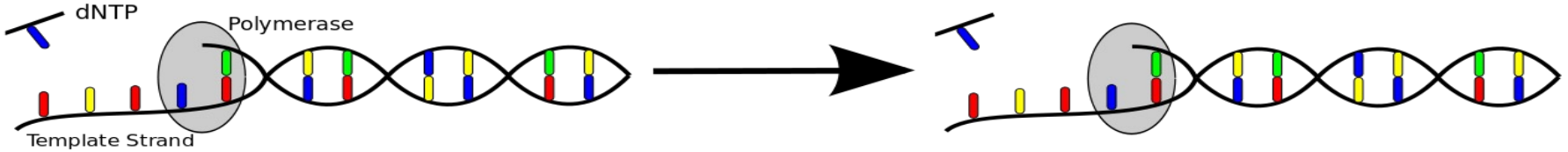
- A computer determines what base was added by the wavelength of the fluorescent tag and records it for every spot on the chip.
- After each round, non-incorporated molecules are washed away. A chemical deblocking step is then used in the removal of the 3’ terminal blocking group and the dye in a single step.
- The process continues until the full DNA molecule is sequenced.


Genomic DNA
shear
select ~200-300 bp fragments
apply to flowcell
attach adapters to
create sequencing library
cluster generation by
solid phase PCR
(bridge amplification)
sequencing by synthesis with reversible terminators
T
G
C
T
A
C

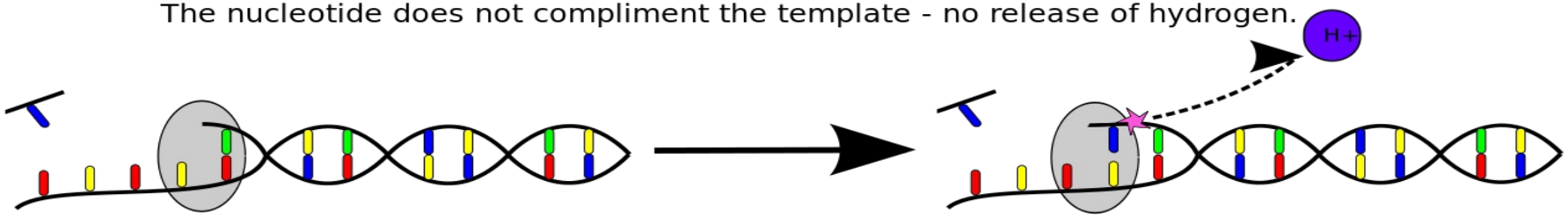
2- Ion semiconductor sequencing

- Is a method of DNA sequencing based on the detection of hydrogen ions that are released during the polymerization of DNA. This is a method of "sequencing by synthesis", during which a complementary strand is built based on the sequence of a template strand.
- A microwell containing a template DNA strand to be sequenced is flooded with a single species of deoxyribonucleotide triphosphate (dNTP).

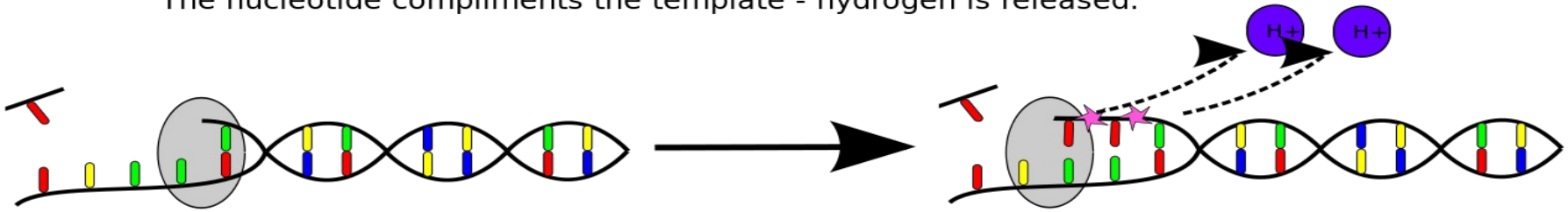
- If the introduced dNTP is complementary to the leading template nucleotide, it is incorporated into the growing complementary strand. This causes the release of a hydrogen ion that triggers an ISFET ion sensor, which indicates that a reaction has occurred.
- If homopolymer repeats are present in the template sequence, multiple dNTP molecules will be incorporated in a single cycle.
- This leads to a corresponding number of released hydrogens and a proportionally higher electronic signal.



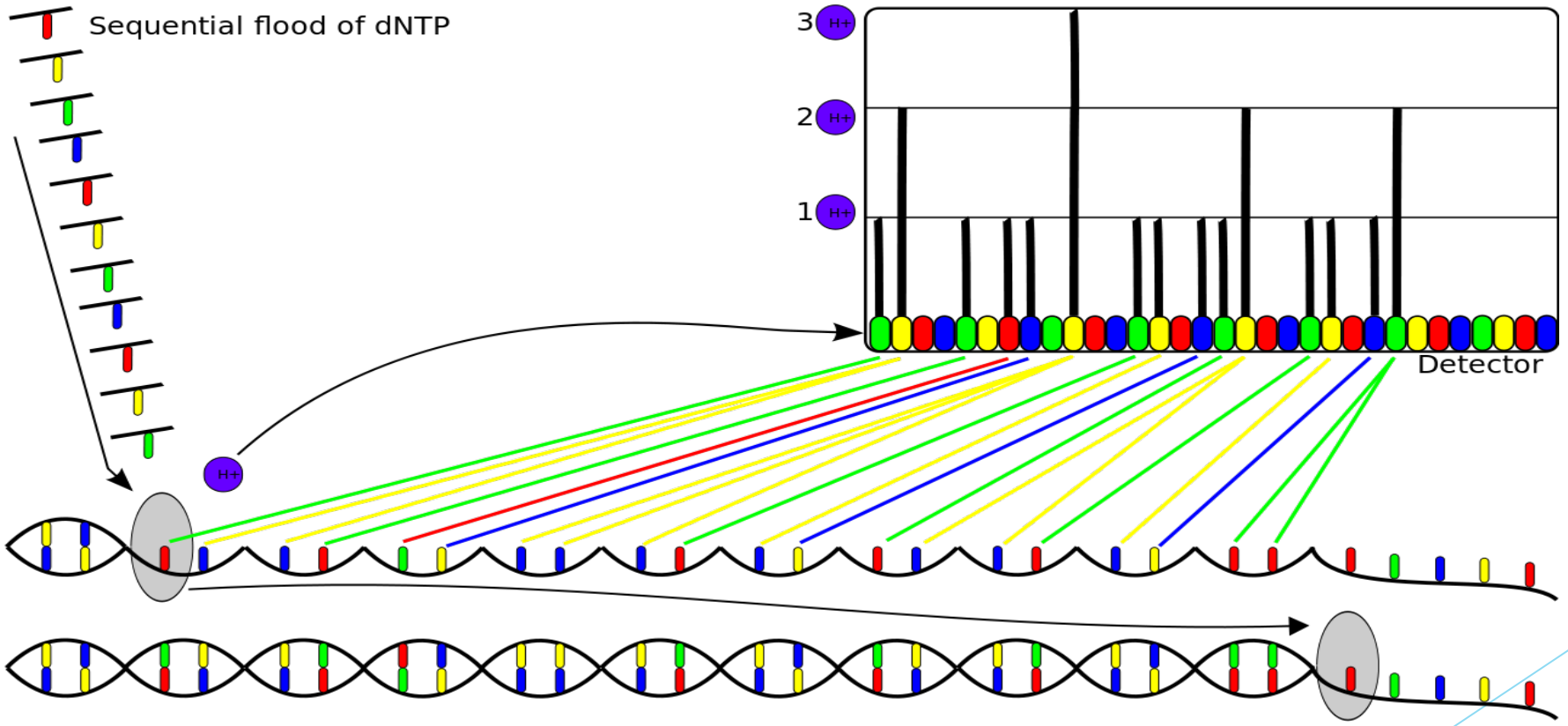
The nucleotide does not compliment the template - no release of hydrogen.



The nucleotide compliments the template - hydrogen is released.



The nucleotide compliments several bases in a row - multiple hydrogen ions are released.


Sequential flood of dNTP
3
2
1
H+
Detector