

## Chapter V:

# Nucleic acid sequence determination: **Sequencing**

# DNA sequencing

- Maxam and Gilbert chemical method
- Sanger enzymatic method
- Automated Sequencing
- Pyrosequencing
- Next generation of sequencing:
  - High-throughput sequencing

# DNA sequencing

- DNA sequencing involves determining the order of nucleotides in a given DNA fragment.
- Used to identify and diagnose genetic diseases and in microbiology.
- An important tool for species classification.

# Sanger sequencing

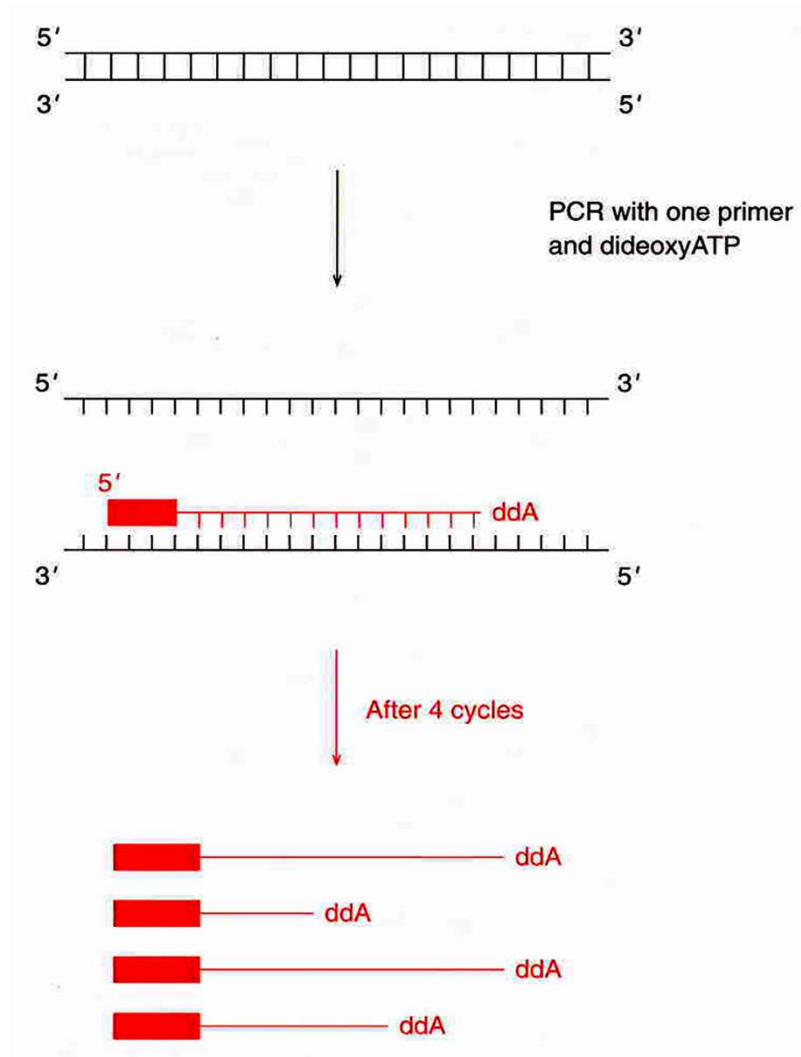
The most commonly used method (developed in 1977).

For Sanger sequencing you need:

- 1) DNA Matrix
- 2) A primer for DNA synthesis
- 3) DNA polymerase
- 4) deoxynucleotide triphosphates (dNTPs) and dideoxynucleotide triphosphates (ddNTPs) ( $^{32}\text{P}$ )



# Sequencing cycle:



95° C: Denatured DNA

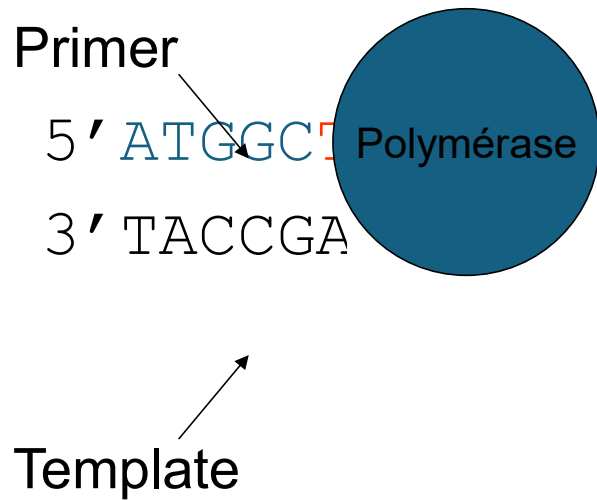
45° C: primer hybridization

60-72° C: primer elongation by thermostable DNA polymerase

Repeat 25 times (25 cycles)

# Sanger sequencing

- Polymerisation of DNA

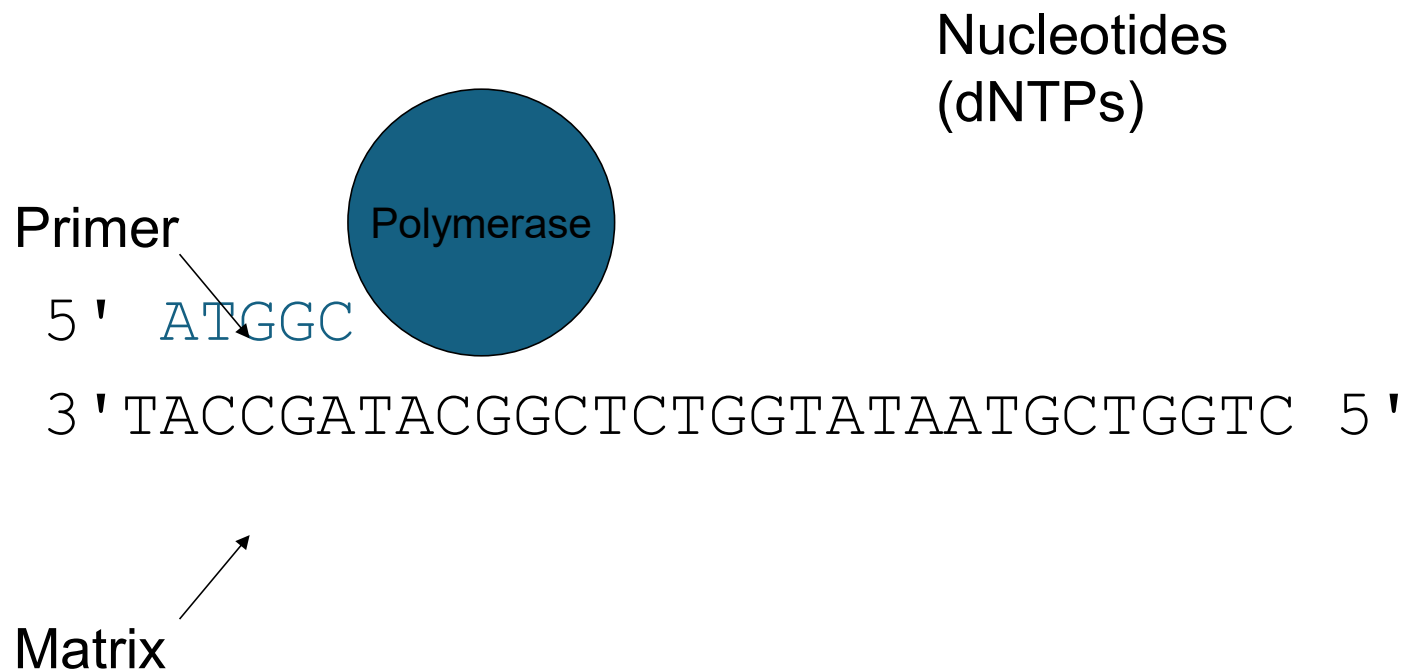


Nucleotides (dNTP)

A G T C C G A C G

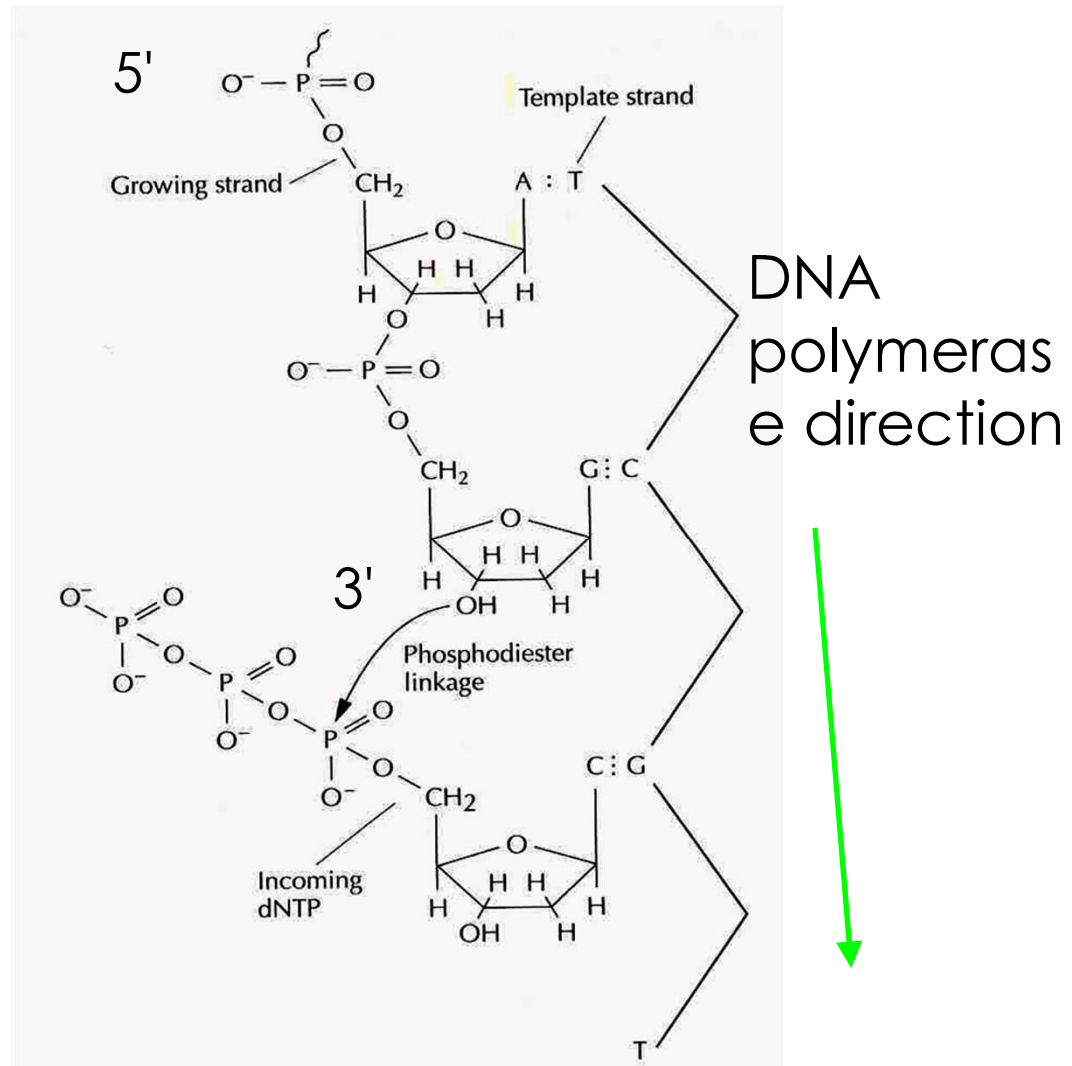
# Sanger sequencing

- DNA polymerization

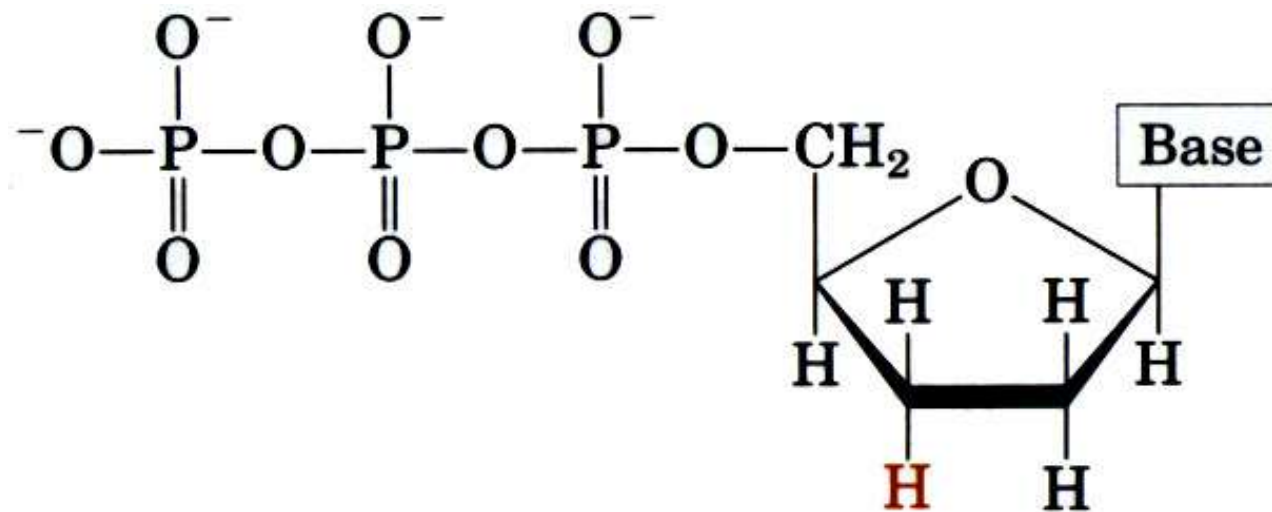


# Sanger sequencing

Primer elongation with DNA polymerase in the presence of the four dNTPs, with a limited amount of radioactive dideoxy NTPs (ddNTPs)



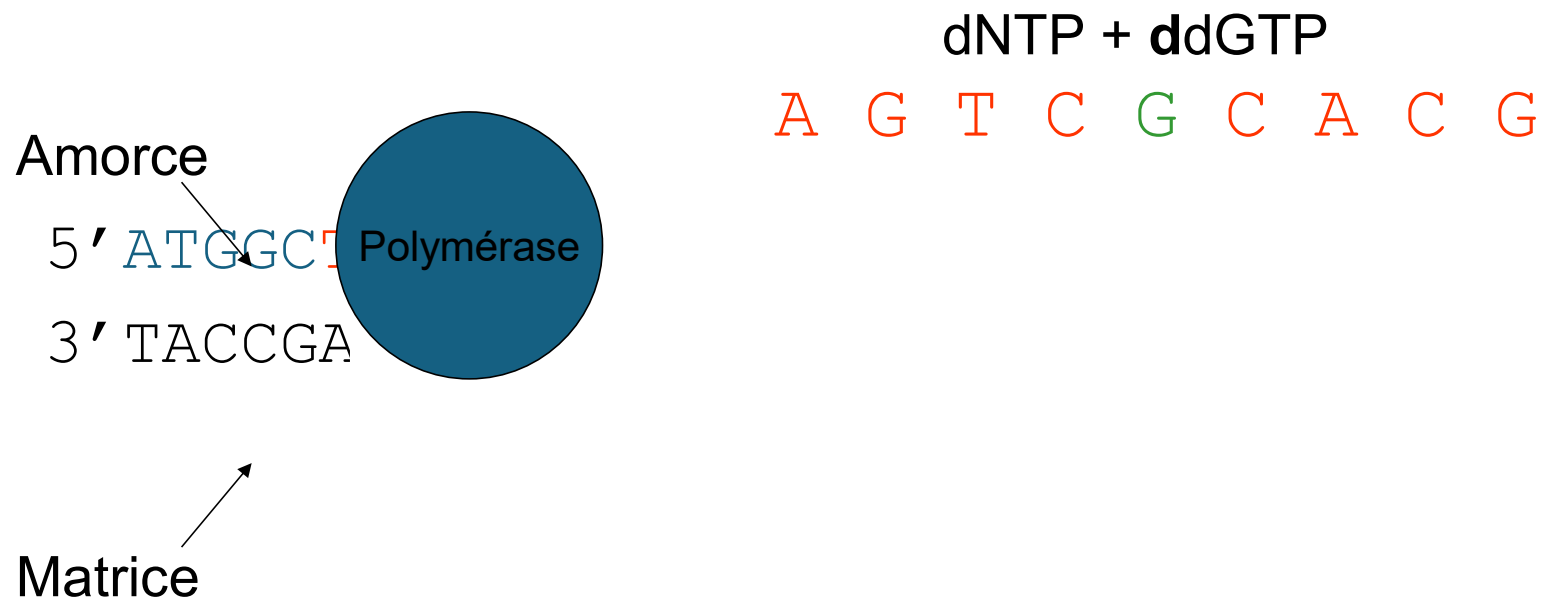
# Dideoxy Nucleotide



ddNTP Analog

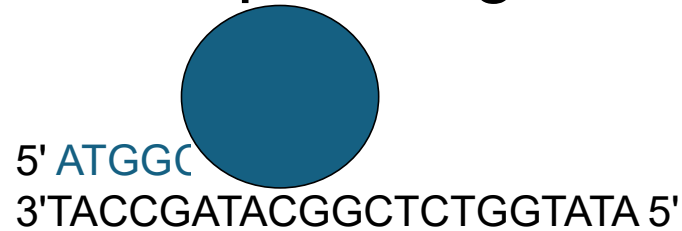
# Sanger sequencing

- Sequencing reaction (ddGTP)



# Sanger sequencing

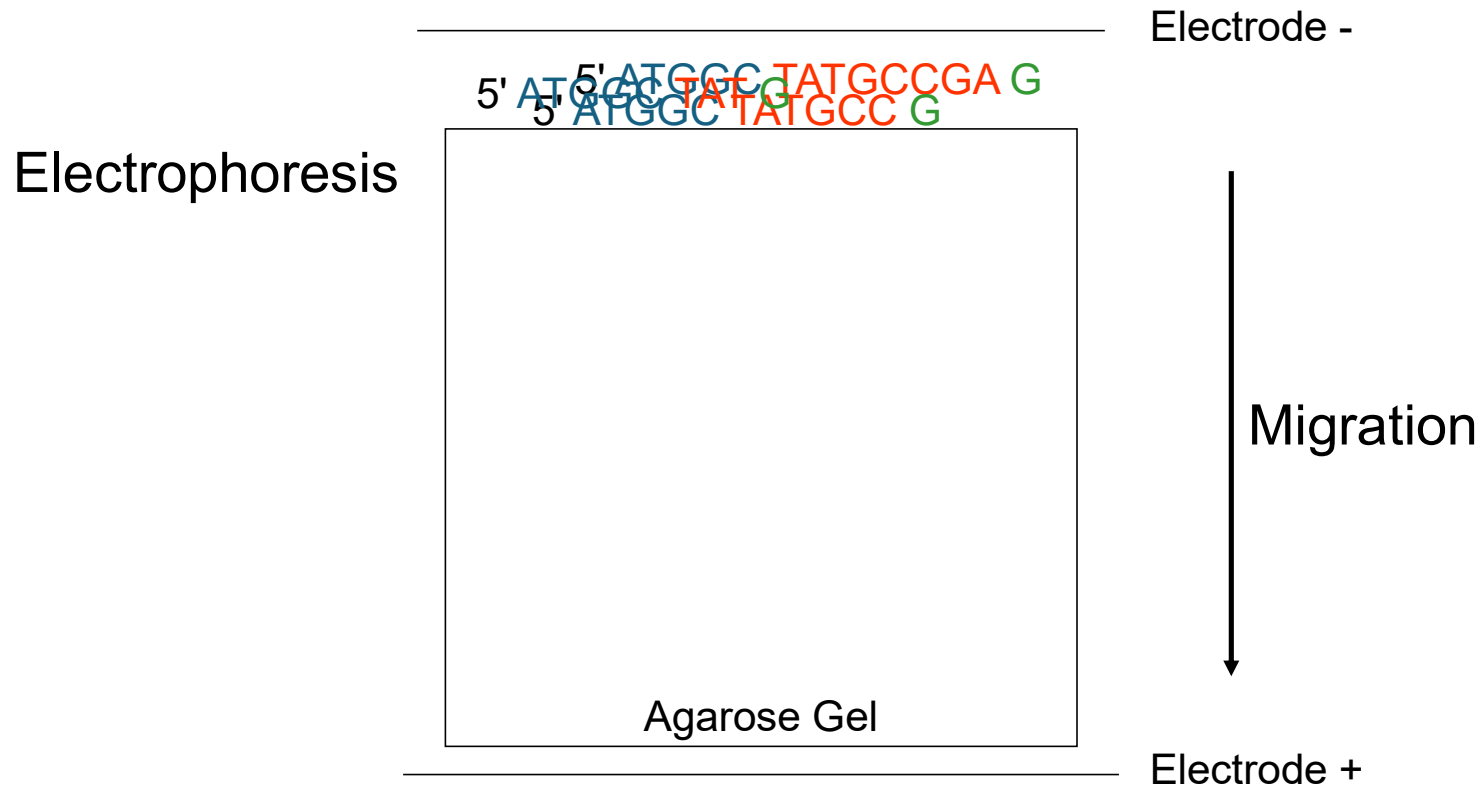
- Sequencing reaction (ddGTP)



Random incorporation of a ddNTP stops synthesis

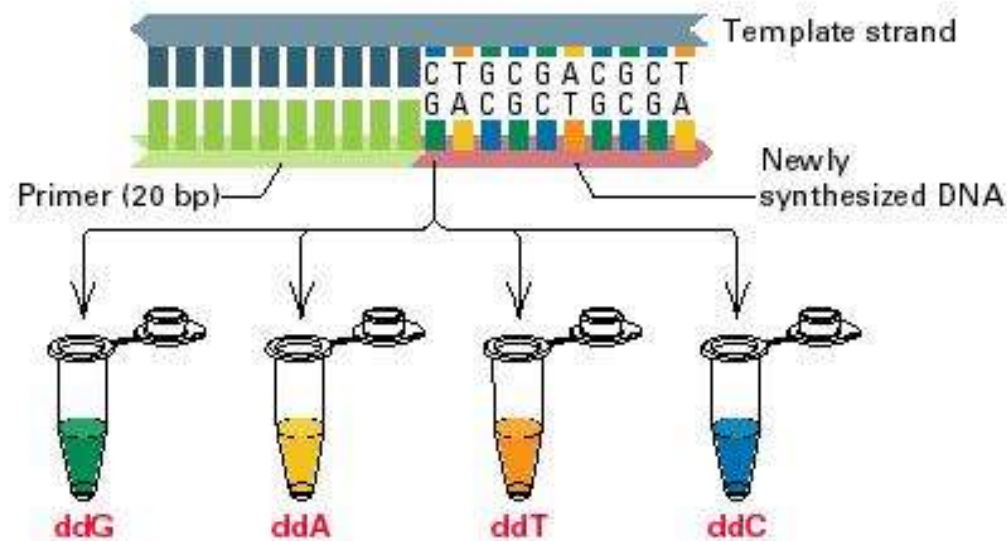
# Sanger sequencing

- Product separation



# Sanger's Method

1. Prepare a large quantity of DNA. Denature and add the primer, the 4 dNTPs, and DNA polymerase.
2. Make 4 aliquots and add one of the radioactive ddNTPs (S35 or P32) to each tube
3. Elongation and termination by incorporation of specific dideoxynucleotides.

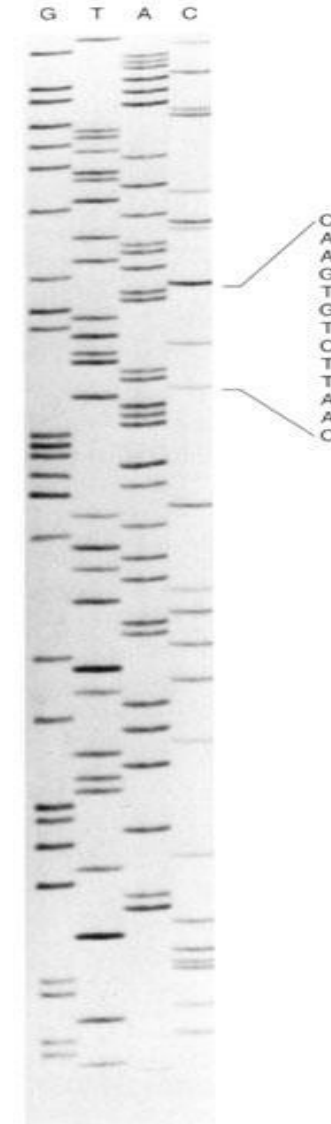




# Sanger's method

7. Transfer of the gel onto filter paper  
(Whatman 3MM).  
Development of the autoradiographic  
film.  
Computer capture.

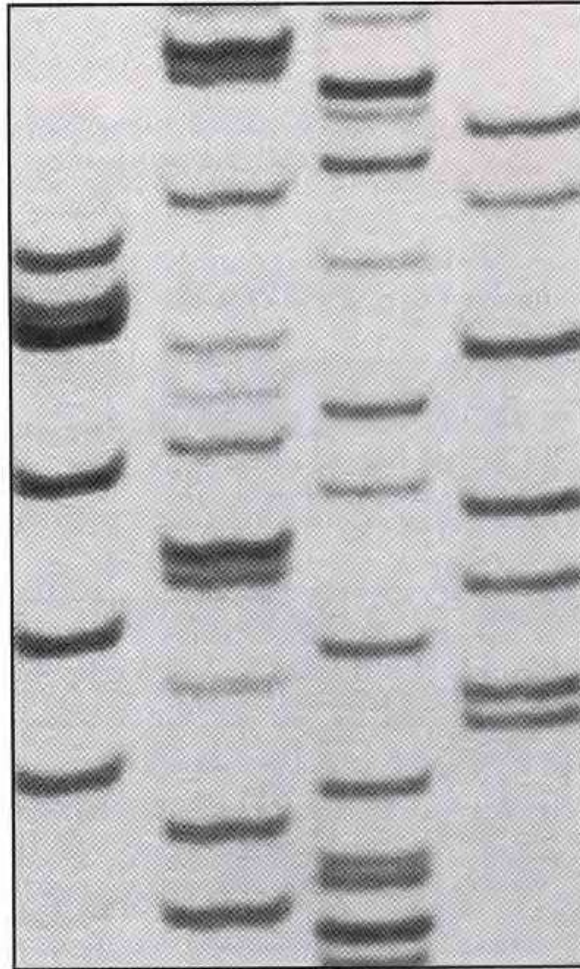
Typical read length: 400 nucleotides.  
Total experiment duration:  
approximately 3 days.



Sequencing gel autoradiograph

A	C	G	T
---	---	---	---

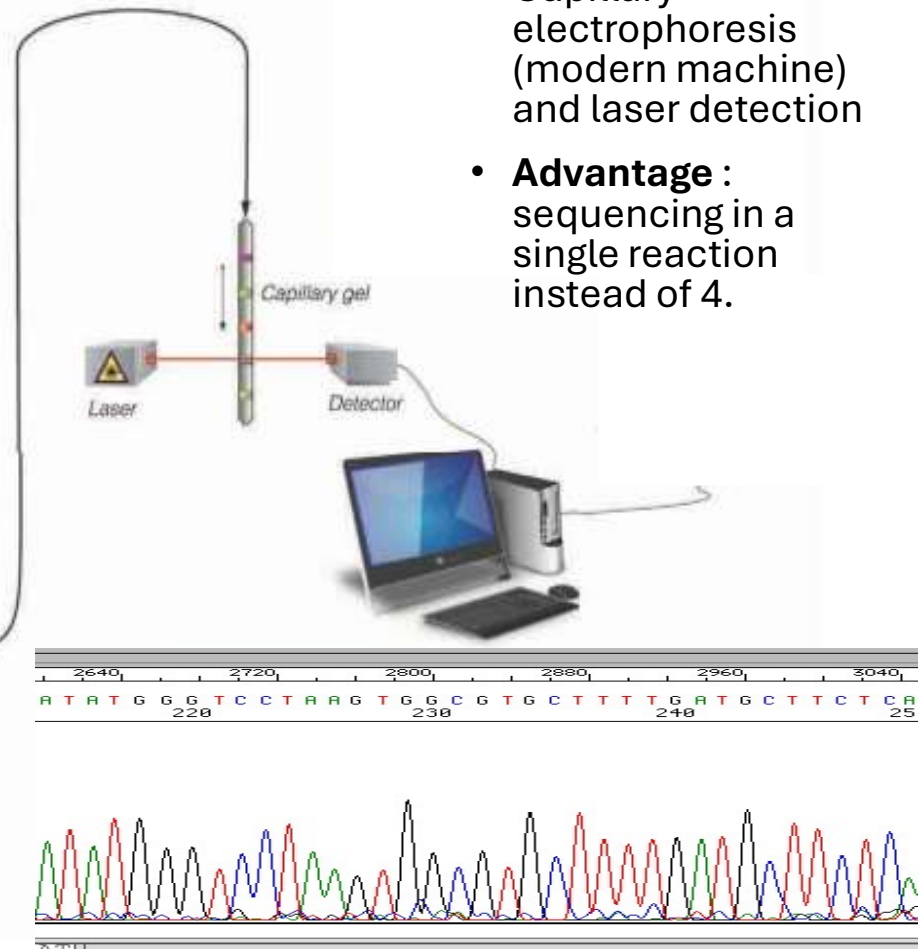
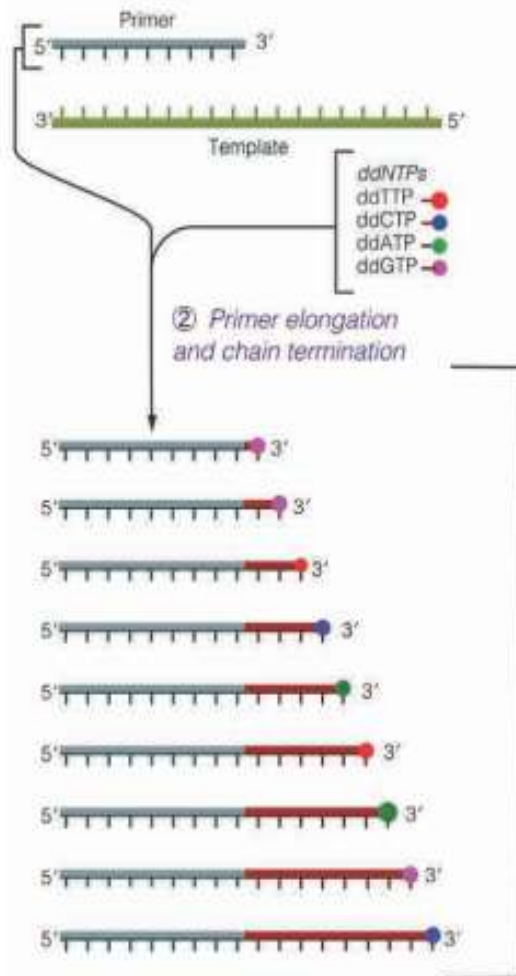
Electrophoresis



T  
C  
G  
C  
A  
G  
T  
C  
C  
T  
A  
G  
C  
T  
T  
A  
G  
C  
G  
G

# Automation of the Sanger method

- Each ddNTP is labeled with a different fluorophore.
- Capillary electrophoresis (modern machine) and laser detection
- **Advantage:** sequencing in a single reaction instead of 4.



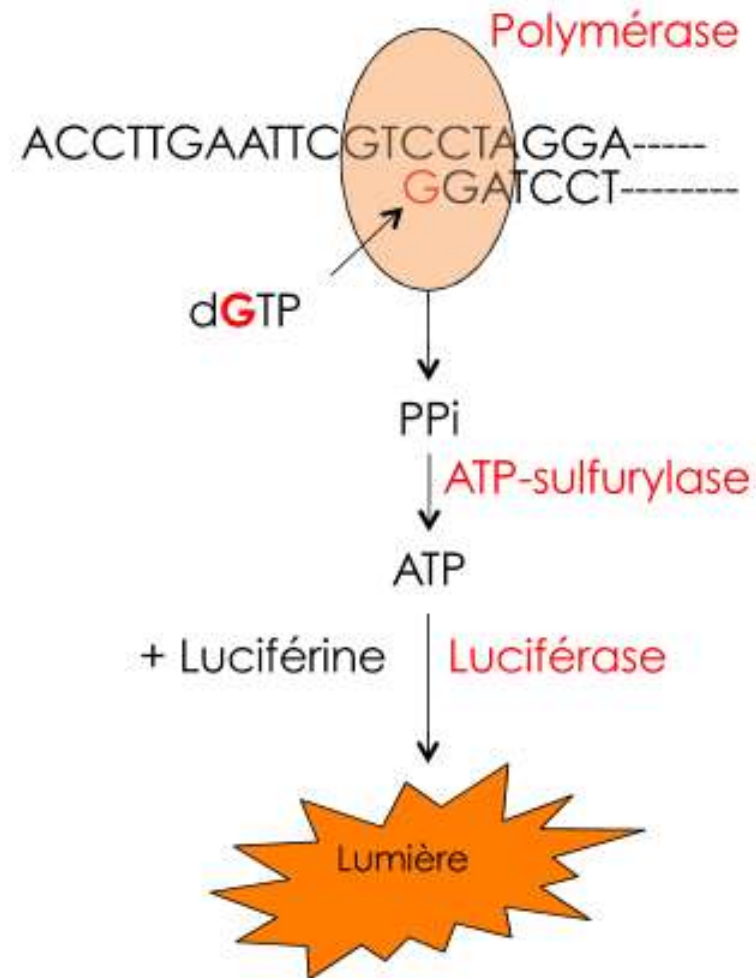
# Pyrosequencing

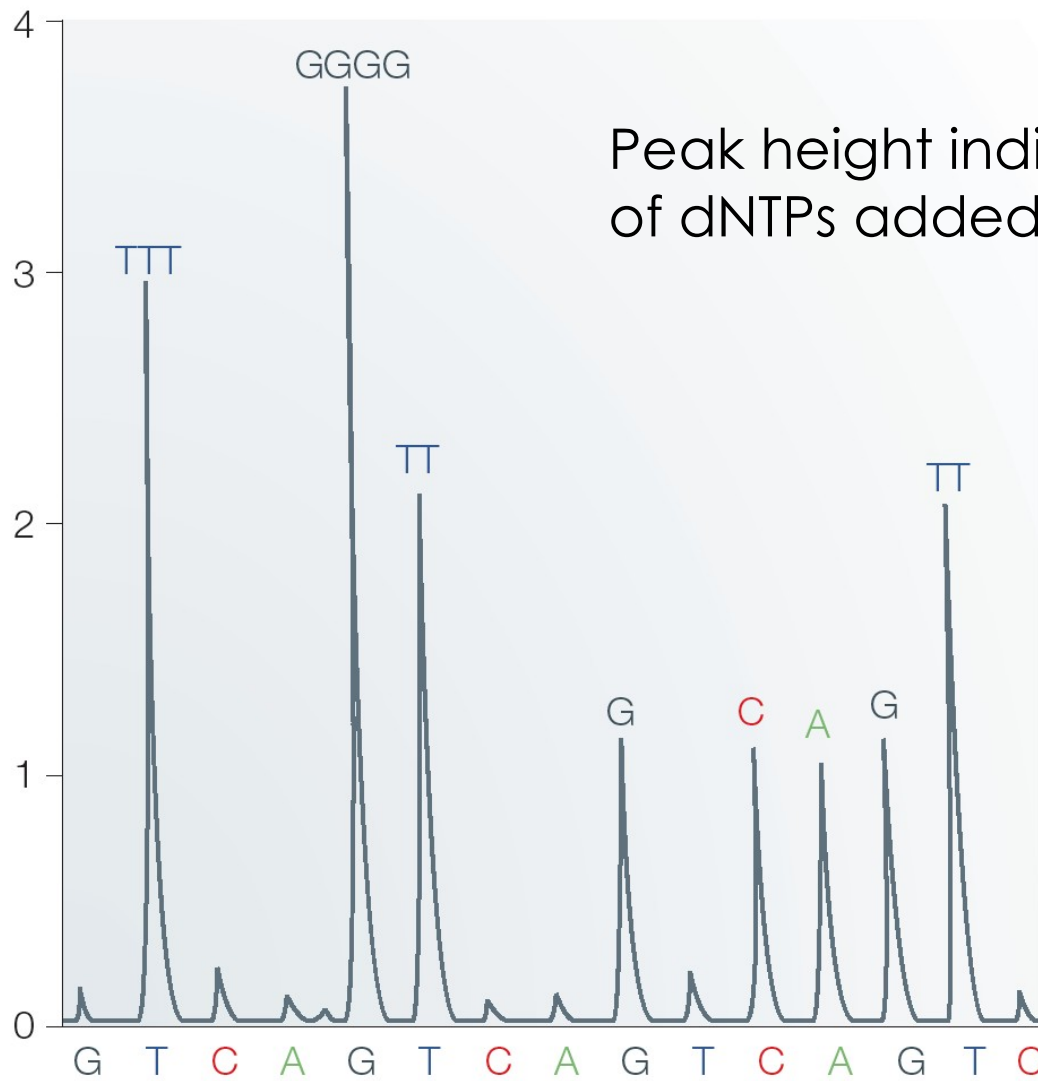
# Pyrosequencing

- Based on a principle of sequencing by synthesis, as opposed to sequencing by "termination" (Sanger method)
- Sequencing of a single-strand DNA by synthesis of the complementary strand, base by base, detecting at each step the activity of the polymerase by another chemiluminescent enzyme: luciferase.

# Pyrosequencing

- Nucleotides (dNTPs) added one after the other (≠Sanger sequencing)
- If it is the correct nucleotide: incorporation and release of a pyrophosphate (PPi)
- PPi to ATP by the action of ATP-sulfurylase
- ATP provides the energy needed for the conversion of luciferin by luciferase. This reaction generates visible light, the intensity of which is proportional to the amount of ATP.





sequence: TTT GGGG TT G C A G TT

Next-generation sequencing: High-throughput sequencing

# High-throughput sequencing technologies



Whole genome sequencing

# High-throughput sequencing: Principle

## Whole genome sequencing

Biological material: DNA or RNA

Common steps across different technologies:

- Enzymatic DNA fragmentation

- Preparing a DNA library  
by linking adapters

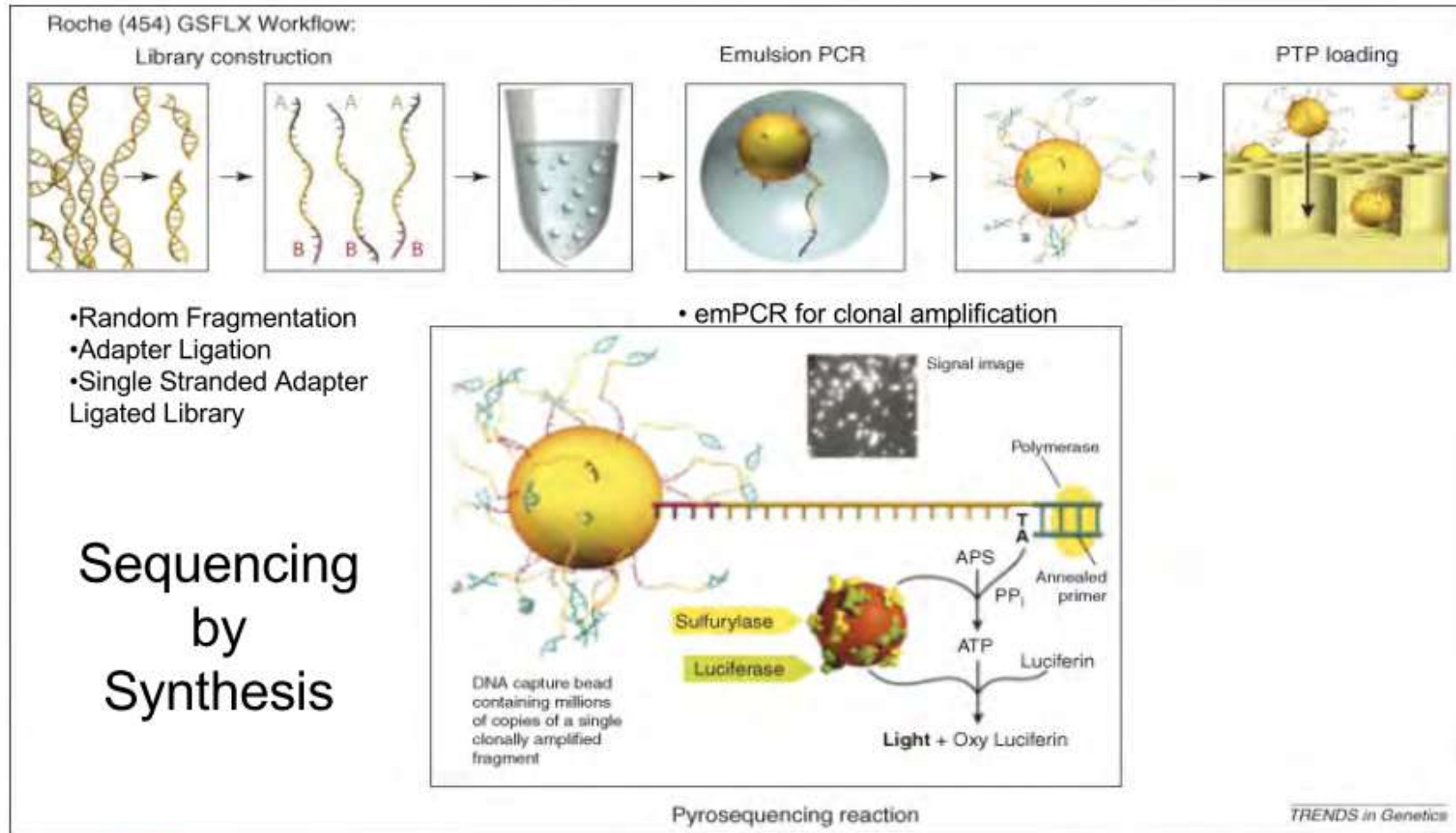
- Amplification

- Sequencing generating signals (luminescent)  
(or fluorescent)

- Detection of emitted signals and conversion  
in sequence

# High-throughput sequencing: Principle

## Exp: Roche 4.5.4 GS20 technology



# High-throughput sequencing: Principle

Exp: Roche 4.5.4 GS20 technology

The screenshot displays the ZOOM lite software interface for next-generation sequencing. The main window shows a 'Mapping Result' for a sample job. The interface includes a menu bar (File, Control, Tool, Help), a toolbar with icons for file operations and analysis, and a tree view on the left showing the job structure. The central panel displays a grid of sequencing reads aligned to a reference sequence. The reads are color-coded by base: C (blue), T (green), G (red), and A (yellow). A specific read is highlighted in green: `ctctgtctctgtgcttccgttaagttaggacttaggcttagaagctcgatctgatgcaggttcag`. Below the reads, a search bar shows the coordinates `chr19_44308119_to_44361966` and a zoom level of 11. The bottom panel provides 'read information' for the selected read, including the read name (2842), reference offset (16...), and mapping direction (Po...). A 'Copy the read sequence' button and a 'Find its mate pair' button are also visible. A progress bar at the bottom indicates the overall progress of the job.

Running monitor

type : job  
total tasks : 1  
overall progress: 100%

Job Properties

Subtask	Subprogress	Time
TASK.2009-1...	██████████	00:00:03

Mapping Result [sample job[LINQUE]]

```
CTCTGT      CTTCCTGTAAGTTAGGACTTAGGCTAGAAGCTNGG      CAGGTTCA
TTTT      GCTGCCIGTAAGTTAGGACTTAGGCTAGAAGCTCG      ATGCAGGTTCA
CTCT      TNCCTCCGTAAGTTAGGACTTAGGCTAGAAGCTC      GATGCAGGTTCA
NTI      TGTGCTTCCGTAAGTTAGGACTTAGGCTAGAATC      CTGATGCAGGTTCA
      TGTCTCTGTGCTTCCGTAAGTTAGGACTTAGGCT      GATCTGATGCAGGTTCA
      TCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGGT      GATCTGATGCAGGTTCA
      TCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGGT      NCGATCTGATGCAGGTTCA
      NNTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGGT      CTCGATCTGATGCAGGTTCA
      CICTGTCTCTGTGCTTCCGTAAGTTAGTACTTAGG      GCTGNATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGG      AGCTCGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAG      AANCTCNATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAAGTTAGGACTTA      ACAACCCTGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAAGTTAGGACTTA      TAGGCTAGAAGCTCGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAA      GACTTAGGCTAGAAGCTCGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTTCC      TGAGGACTTAGGCTAGAAGCTCGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTT      AAGTTAGGACTTAGGCTAGAAGCTCGATCTGATGCAGGTTCA
      CTCTGTCTCTGTGCTT      CTGTAAGTTAGGACTTAGGCTAGAAGCTCGATCTG      GGTTCAGGTTCA
      CTCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGGCTAG      GATCTGATGCAGGTTCA
      CNCTGTCTCTGTGCTTCCGTAAGTTAGGACTTAGAAGCTCGATCTGATGCAGGTTCA
      ctctgtctctgtgcttccgttaagttaggacttaggcttagaagctcgatctgatgcaggttcag
      ctctgtctctgtgcttccgttaagttaggacttaggcttagaagctcgatctgatgcaggttcag
```

10543 10900 16600

11 chr19\_44308119\_to\_44361966 16626 - 16689

read information

Read name: 2842  
Reference offset: 16... Mapping direction: Po...

Copy the read sequence

Find its mate pair

```
TTAGGACTTAGGCTAGAAGCTCGATCTGATGCAGG
T GAGGACTTAGGCTAGAAGCTCGATCTGATGCAGG
████████████████████████████████████████
```