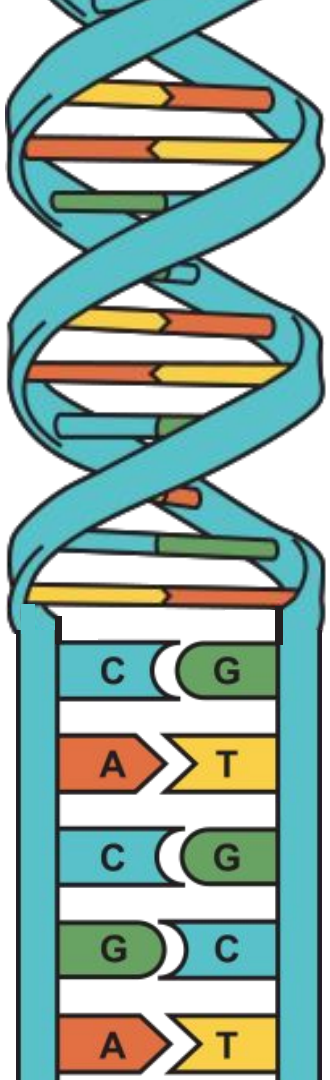


DNA SEQUENCING: FROM SANGER TO NGS

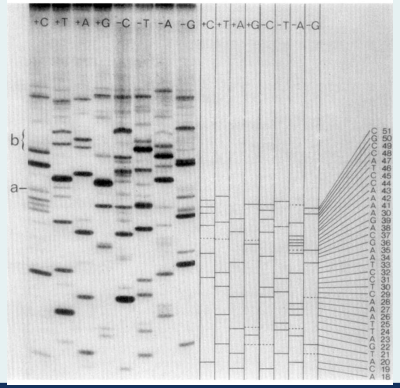
Presented by: Aliyah, Clare, and Jules

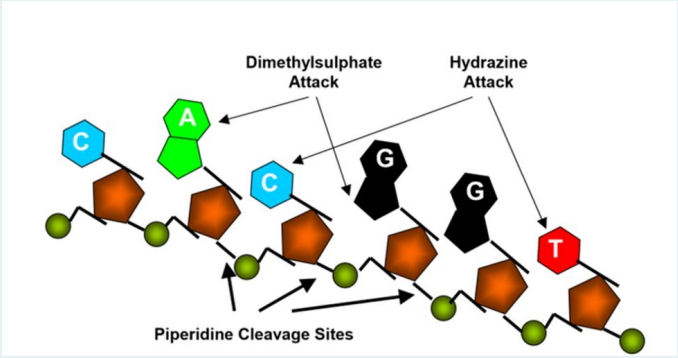
What is DNA sequencing?

- DNA sequencing determines exact nucleotide order: A, T, G, and C
- Organism’s complete DNA sequence is its “genome”
- Sequencing allows scientists to read and interpret genetic information
- Four-letter code forms blueprint for all biological processes
- Understanding sequences fundamental to modern molecular biology

Historical Background: DNA Sequencing Before 1977

- Pre-1977: Two dominant methods — Sanger & Coulson 'plus and minus' method (1975); Maxam & Gilbert chemical degradation (1977)
- Plus/minus method: rapid but neither method alone was completely accurate — both had to be used together
- Maxam-Gilbert: applicable to double-stranded DNA but required strand separation for each restriction fragment — more laborious
- Sanger's challenge: design a simpler, more accurate method using DNA polymerase to read sequence through chain termination





Scientific Significance: Sanger's 1977 paper introduced a method that would become the dominant sequencing technology for 30 years and earn him a second Nobel Prize in Chemistry (1980).

The Chain-Termination Principle

Sanger et al. (1977) — Logic of the Experiment

LOGIC OF THE EXPERIMENT

- Atkinson et al. (1969): ddTTP incorporated into growing chain in place of dTMP — chain cannot extend because the 3'-OH is absent
- Key insight: if ddNTP is randomly incorporated among normal dNTPs, synthesis terminates at every position where that base occurs
- Four parallel reactions — one per ddNTP (ddTTP, ddATP, ddGTP, ddCTP or arabinoside analogues) — produce nested fragment ladders
- Ratio of terminator:normal dNTP ~1:100 (dideoxy) or 1:5000 (arabinoside) ensures only partial termination — full ladder produced
- Fragments share same 5' end (primer) but differ at 3' end: each band position = one nucleotide position in the sequence

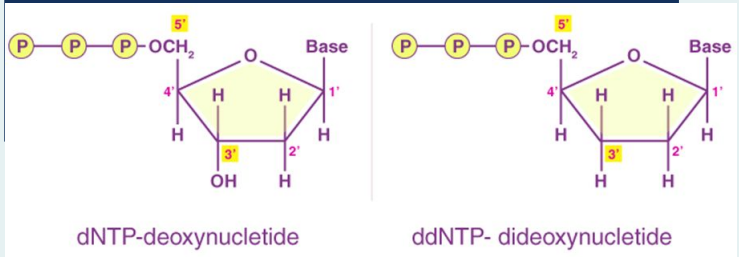
The Key Molecule

dNTP (normal):
Has 3'-OH → chain can extend

ddNTP (terminator):
No 3'-OH → chain stops

Two types used:
• 2',3'-Dideoxy derivatives (ddNTPs)
• Arabinoside analogues (araATP, araCTP)

Both act as specific chain-terminating inhibitors of *E. coli* DNA Polymerase I



Sanger et al. (1977), p. 5463. Based on Atkinson et al. (1969) Biochemistry 8: 4897.

Preparing the Terminators: Chemistry of ddNTP Synthesis

Tosylation as the Key Enabling Reaction

ddTTP & ddATP

ddTTP: already commercially available.
ddATP: prepared by McCarthy et al. method; converted to triphosphate by phosphorylation; purified on DEAE-Sephadex.

ddGTP — Tosylation is the Key Step

Starting material:
N-isobutyryl-5'-O-monomethoxytrityldeoxyguanosine
→ TOSYLATION of the 3'-OH group
→ Sodium methoxide converts to 2',3'-didehydro derivative
→ Reduction to dideoxy; then phosphorylation to triphosphate

ddCTP

Prepared from N-anisoyl-5'-O-monomethoxytrityldeoxycytidine using the same tosylation-based method.
Yields very low — solution used directly without final purification.
ddCTP concentration determined experimentally.

What is Tosylation?

Tosylation (sulfonylation) converts an alcohol's -OH group into a tosylate ester (-OTs) using p-toluenesulfonyl chloride + pyridine. This makes the -OH a far better leaving group, enabling subsequent nucleophilic substitution — in this case, conversion to the didehydro intermediate essential for ddNTP synthesis.

Sanger et al. (1977), p. 5463–5464. Robins et al. (1966) Biochemistry 5: 224. Tipson (1944); Kabalka et al. (1986).

Tosylation of Alcohols: From Organic Chemistry to Sequencing

Tipson (1944) · Kabalka, Varma & Varma (1986)

- Tosylation: reaction of an alcohol with p-toluenesulfonyl chloride (TsCl) in pyridine → converts -OH to -OTs (tosylate ester)
- Tosylate is an excellent leaving group: enables nucleophilic substitution reactions that are otherwise impossible on free alcohols
- Tipson (1944): established general method — alcohol dissolved in dry pyridine, TsCl added at 0°C; yields >75% for most substrates; chlorination avoided
- Kabalka et al. (1986): optimized conditions — 1:1.5:2 ratio of alcohol/TsCl/pyridine in chloroform; primary alcohols complete in 2.5 h; 98% yield
- Sanger's application: tosylation of 3'-OH of protected deoxyguanosine was the essential step enabling synthesis of ddGTP — the scarcest and most difficult terminator

Bridge to Automated Sequencing:

Reliable synthesis of all four ddNTPs — made possible by tosylation chemistry — was the prerequisite for packaging them into commercial kits. This chemical reproducibility, combined with Hood et al.'s (1987) fluorescent labelling, enabled fully automated Sanger sequencing instruments (Applied Biosystems 370A). Without stable, scalable ddNTP production, automation would not have been possible.

The Sequencing Procedure

Methods from Sanger et al. (1977)

1. Template Preparation

5 µg φX174 replicative form restriction fragments separated by acrylamide gel electrophoresis; mixed with viral or complementary strand DNA (0.6 µg) in HX10 buffer.

2. Annealing

Sealed in capillary tube; heated 100°C for 3 min (denaturation); incubated 67°C for 30 min (annealing). Template-primer complex formed.

3. Extension Reaction

Four separate 2 µl samples mixed with 2 µl 'mix' (1.5× H buffer + [α-^{32}P]dATP + 3 dNTPs + one ddNTP) + 0.2 units Klenow DNA Pol I. Room temperature, 15 min.

4. dATP Chase

1 µl of 0.5 mM unlabeled dATP added; incubated 15 more min. Critical: prevents false termination at A residues caused by low [α-^{32}P]dATP concentration.

5. Electrophoresis

Reaction products denatured and loaded on 12% denaturing polyacrylamide gel. 40 mA for 14 hr. Four lanes per primer — one per terminator.

Single-Site Ribo-Sub Method

Alternative splitting method (N.L. Brown, unpublished): incorporates single ribonucleotide (rCTP) in Mn^{2+} buffer; split by RNase A or alkali. Avoids problems when restriction enzyme has second site near primer.

Sanger et al. (1977), pp. 5463–5464. Klenow = DNA Pol I large fragment (Boehringer Mannheim).

Results — Autoradiographs of φX174 Sequencing Gels

Figures 1, 2 & 3 — Sanger et al. (1977)

Figure 1 (A12d & A14 primers)
Complementary strand. Inhibitors: ddGTP, ddATP, ddTTP, araCTP. 12% gel, 40 mA, 14 hr. Sequence readable from nt 88 (from primer 5' end) for ~80 nt with reasonable accuracy. Next 50 nt: uncertainty in runs. Validates method against previously published φX174 sequence.

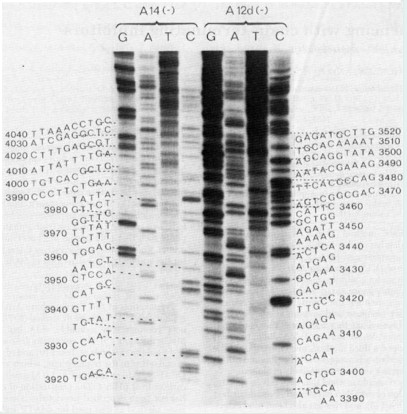


Figure 2 (R4 primer)
Complementary strand. ddCTP replaces araCTP (araC removed by 3'-exonuclease at 37°C restriction step). 120 nt readable from position 61 from splitting site. Identified error in provisional sequence: trinucleotide C-G-C insertion between positions 4380–4381. Confirms region containing viral strand replication origin.

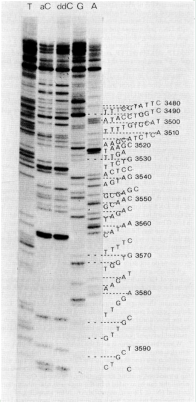
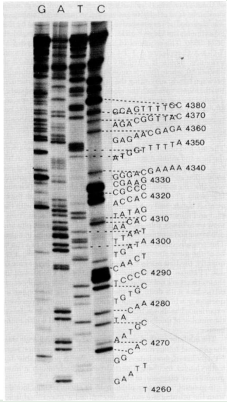


Figure 3 (A8 primer — Ribo-sub method)
Viral strand. Single-site ribo-substitution method. ~110 nt from residue 33. Confirmed most of provisional sequence but revealed clear revision at positions 3524–3530 (A-T-C-A-A-C, not A-T-T-C---A-C). Insertion of C between 3547–3548 confirmed at higher temperature. Secondary structure 'pile-up' at 3545–3550.



Sanger et al. (1977), Figs. 1–3, pp. 5464–5465. φX174 reference sequence: Sanger et al. (1977) Nature 265: 687–695.

Discussion: Advantages, Limitations & Critical Review

Advantages Over Plus/Minus Method

- Simpler: no preliminary extension or Sephadex purification step required
- Only commercially available Klenow fragment needed
- Cleaner results with fewer artefact bands
- Intermediate nucleotides in 'runs' appear as bands — eliminates run-length ambiguity
- 15–200 nt readable per primer; occasionally up to 300 nt

Limitations Identified by Sanger et al.

- ddGTP and ddCTP yields very low — neither commercially available
- Secondary structure 'pile-ups': DNA forms base-paired loops under gel conditions — bands cluster at same position
- Confirmation by complementary strand sequencing recommended
- araCTP and araATP unstable at 37°C (3'-exonuclease degrades them)

Critical Review

The chain-termination method was elegant and simple in concept but chemically demanding in practice — particularly the synthesis of ddGTP and ddCTP, where yields were so low that Sanger acknowledged the results 'can hardly be regarded as adequate chemical characterization.' The method's dependence on single-strand templates and restriction enzyme primers also limited applicability. Nevertheless, the logic — that random termination positions map directly to sequence — was immediately verified on the φX174 genome and proved more reliable than any prior rapid method. The paper also honestly identified artefacts (pile-ups, faint bands in runs), demonstrating rigorous self-critique.

Sanger et al. (1977), pp. 5466–5467 (Discussion).

Automation of the Sanger Sequencing method

- Radioactive bases replaced with fluorescent dyes that are read with machine
- Safer, automated, longer sequences possible
- Machines commercially available in late 1980s
- Gel must be checked over

Fluorescence detection in automated DNA sequence analysis

Lloyd M. Smith, Jane Z. Sanders, Robert J. Kaiser, Peter Hughes, Chris Dodd, Charles R. Connell*, Cheryl Heiner*, Stephen B. H. Kent & Leroy E. Hood

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA
* Applied Biosystems, Inc., Foster City, California 94404, USA

Research Articles

A System for Rapid DNA Sequencing with Fluorescent Chain-Terminating Dideoxynucleotides

JAMES M. PROBER,* GEORGE L. TRAINOR, RUDY J. DAM, FRANK W. HOBBS, CHARLES W. ROBERTSON, ROBERT J. ZAGURSKY, ANTHONY J. COCUZZA, MARK A. JENSEN, KIRK BAUMEISTER

The Fluorescent Dyes

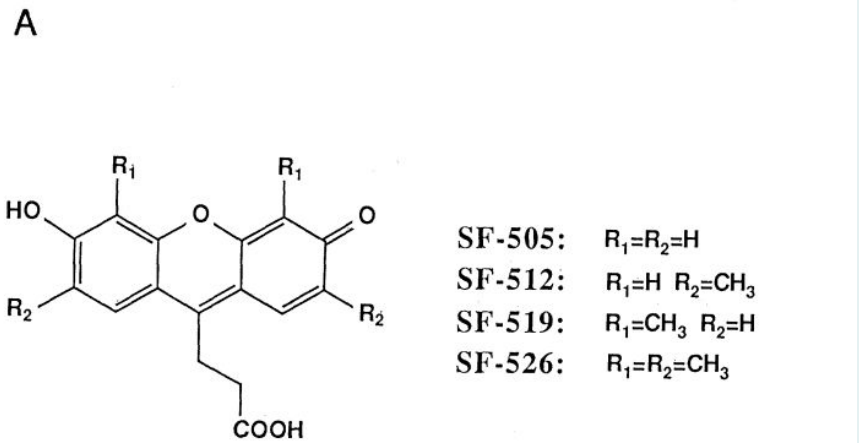


Fig. 1. Succinylfluorescein dyes. (**A**) Chemical structure of the four dyes used to label dideoxynucleotide triphosphates for use as chain-terminators in modified dideoxy DNA sequencing protocols. (**B**) Normalized absorption spectra of the dyes shown in (A). Absorption coefficient at the maximum for

(Prober et al. 1987)

Criteria

1. Absorption and emission maxima in visible region of spectrum
2. Emission maxima well resolved
3. Highly fluorescent
4. Do not impair reaction

Automated sequencing: required materials

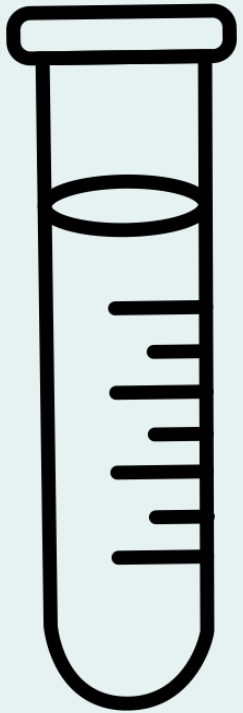
DNA Polymerase

Nucleotides

ddNTP

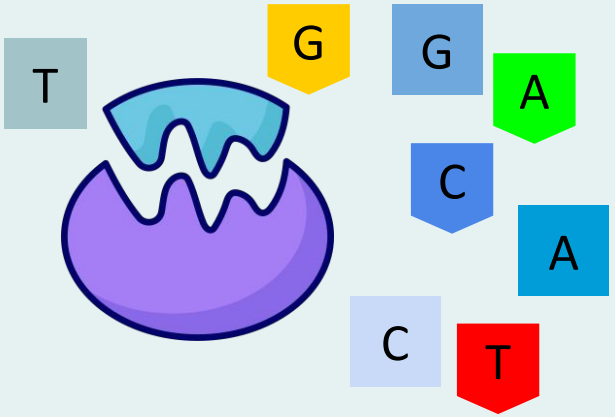
DNA Primer

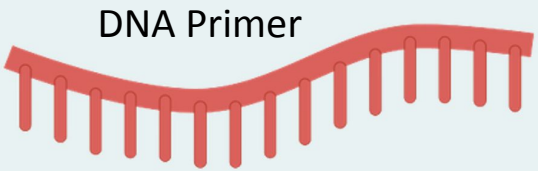
Template DNA

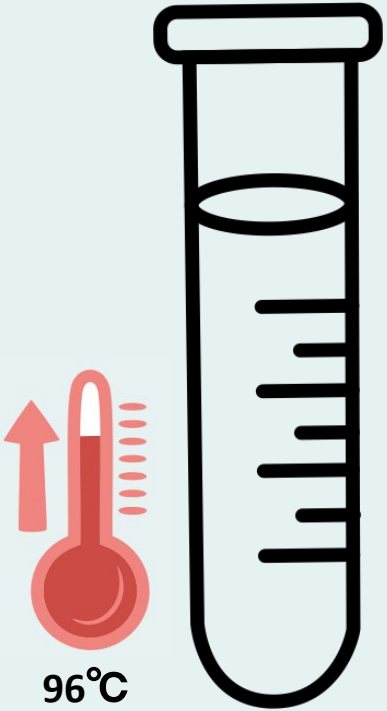


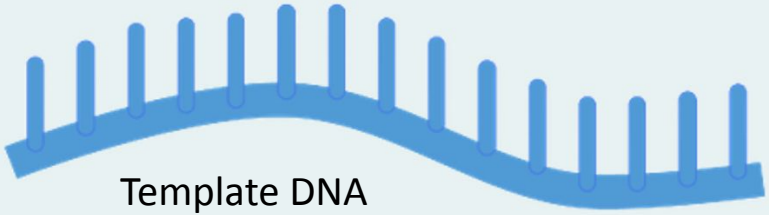
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



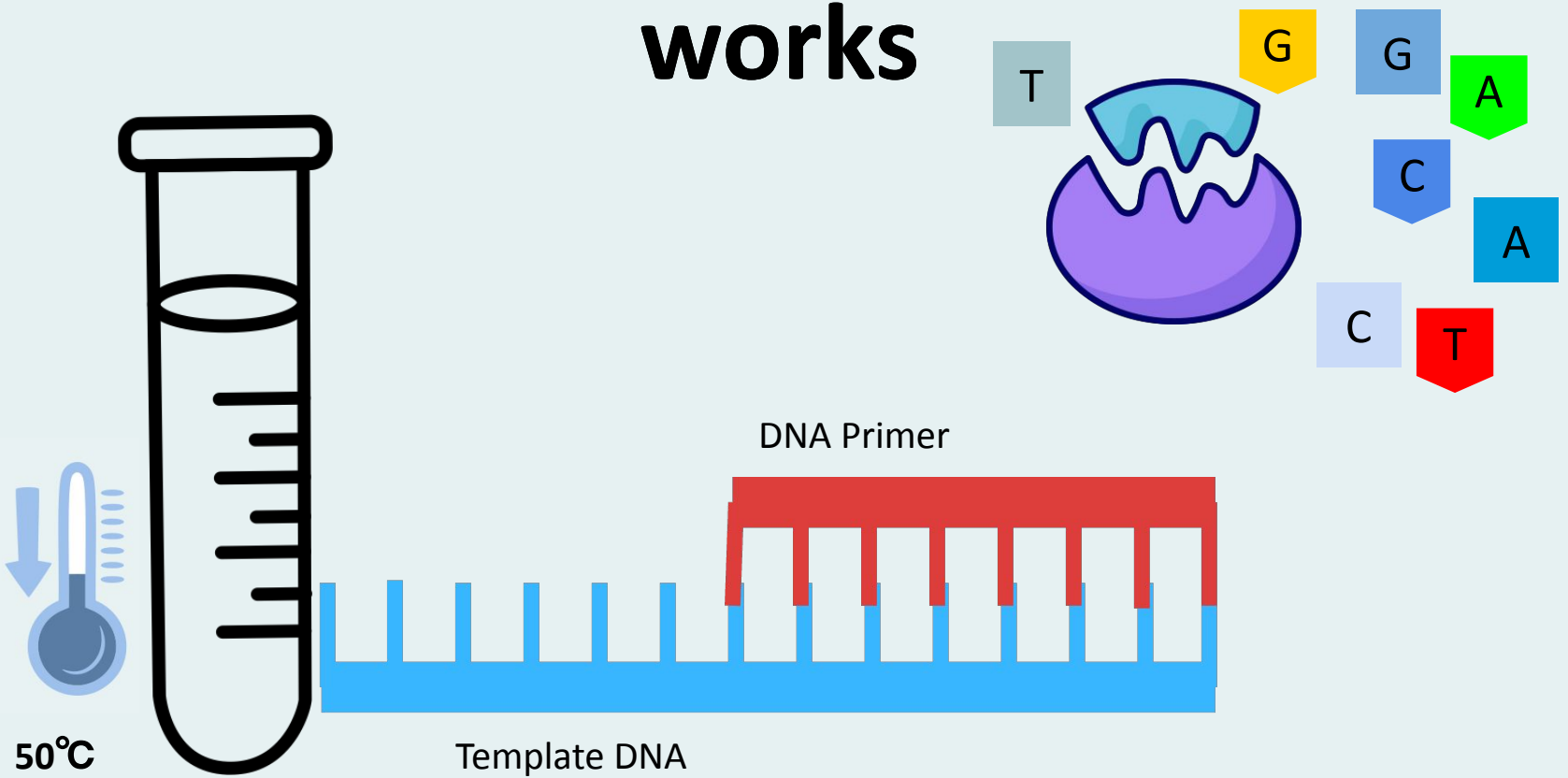






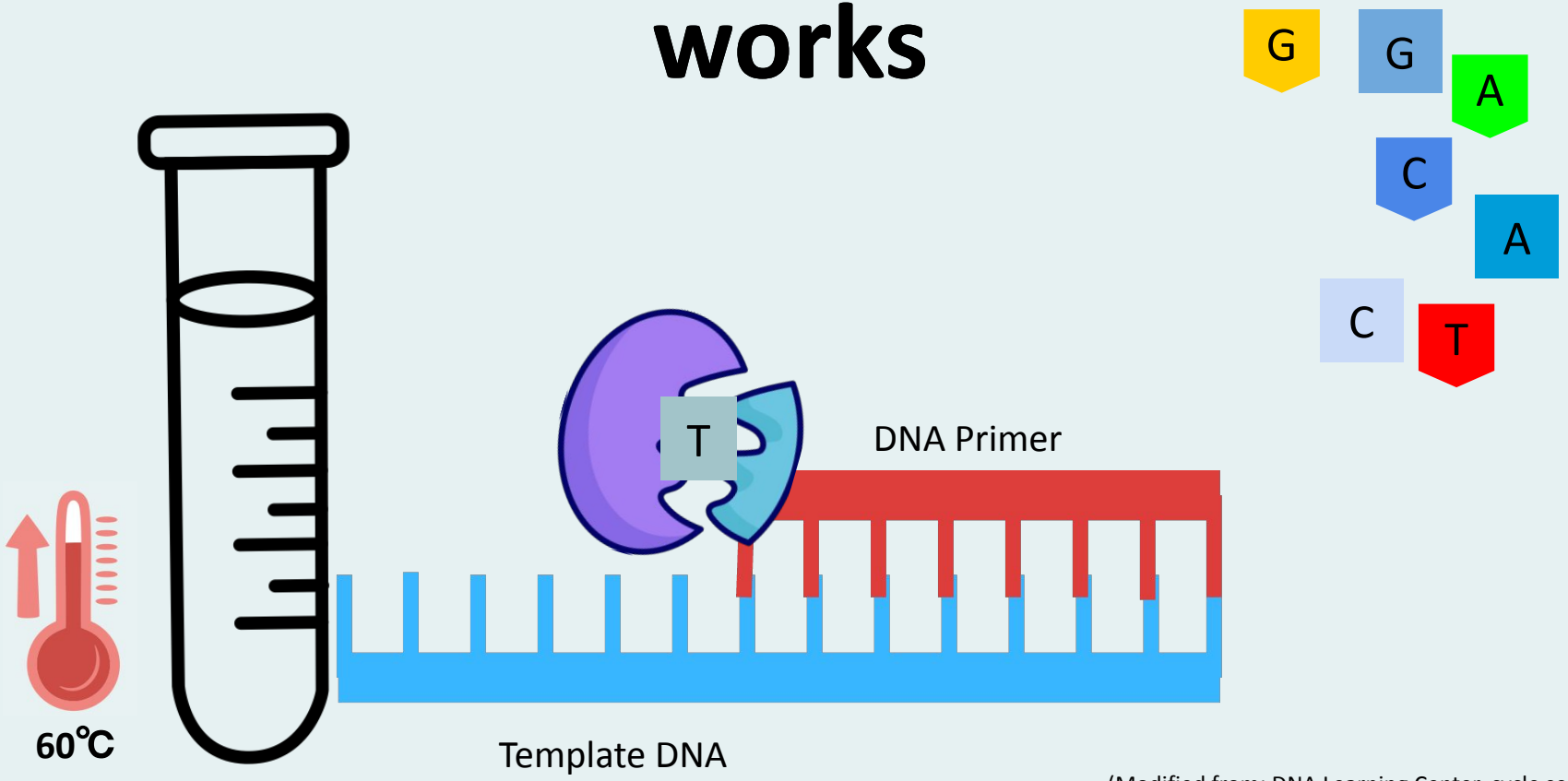
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



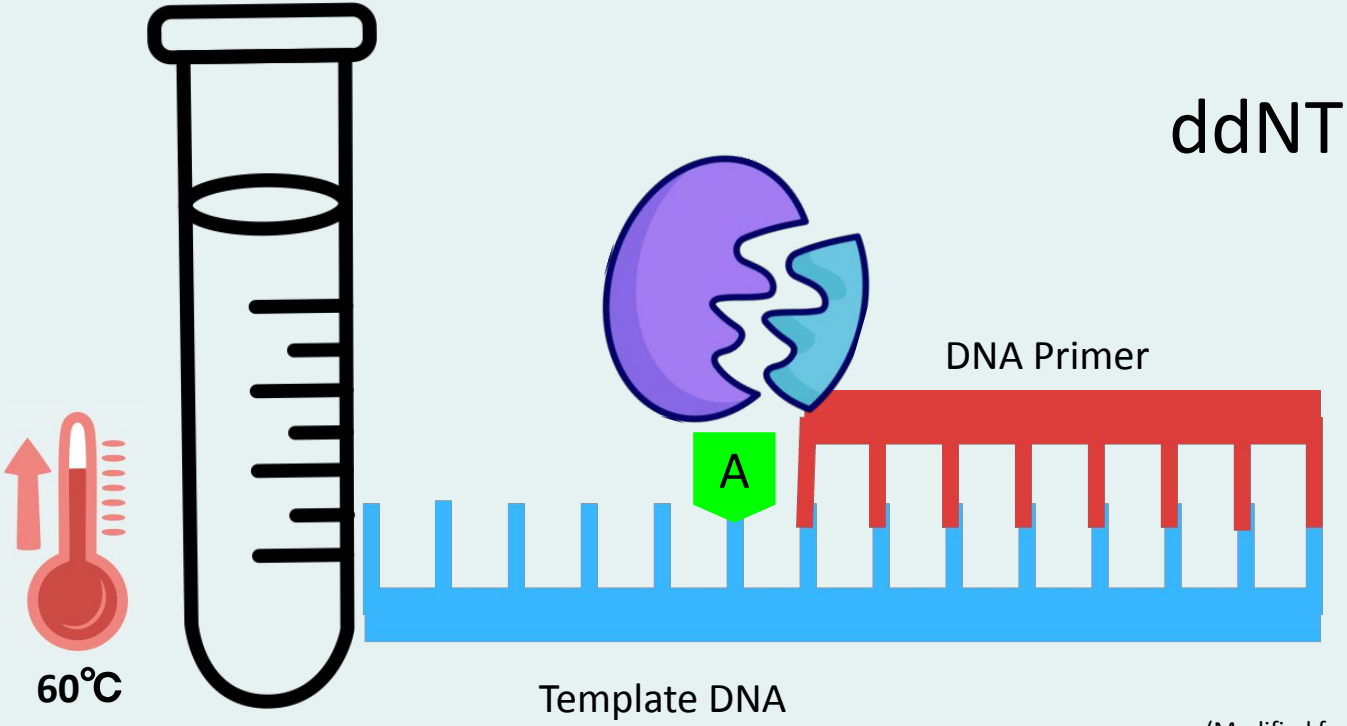
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



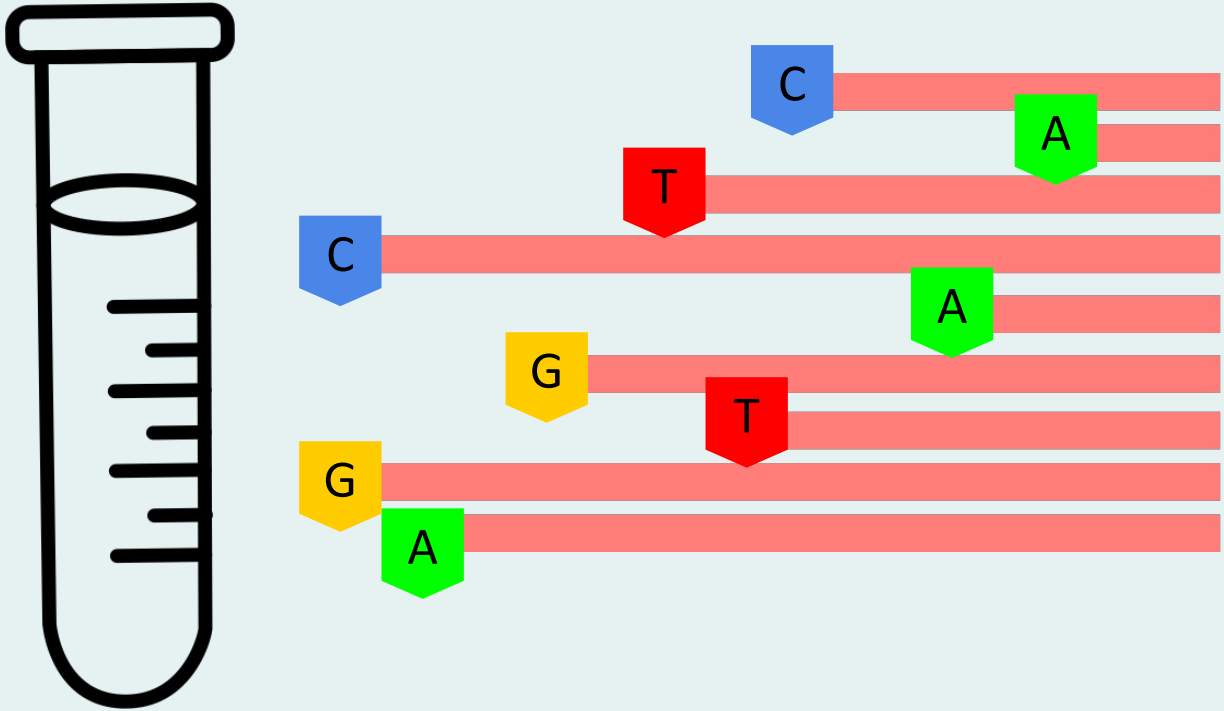
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



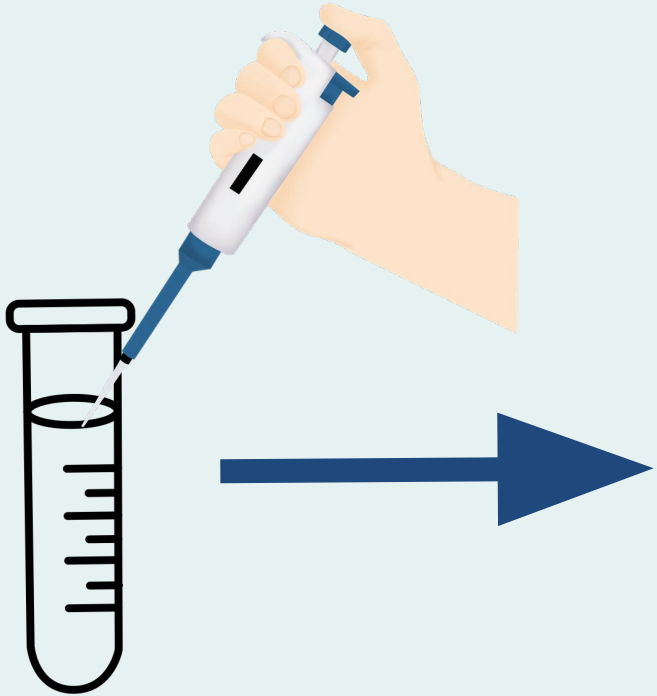
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



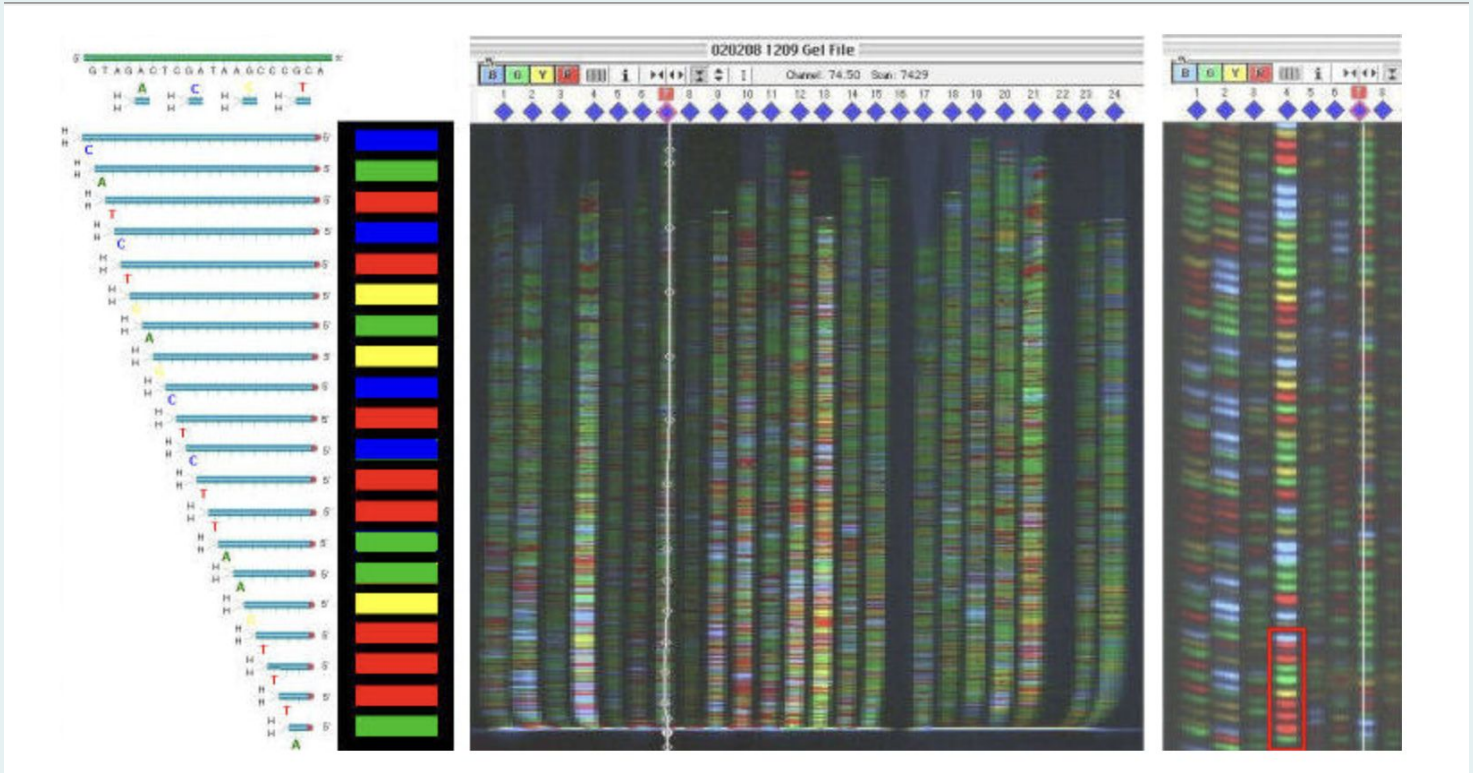
(Modified from: DNA Learning Center, cycle sequencing)

Automated sequencing: how it works



(Science museum group collection)

Gel images from Automated DNA sequencer



Gel image

Magnified view

(Lecture materials, 2026)

Gel images from Automated DNA sequencer

- A C G T
- Read bottom to top
- What is the DNA sequence in the red box?

Magnified view

(Lecture materials, 2026)

Gel images from Automated DNA sequencer

- A C G T
- Read bottom to top
- What is the DNA sequence in the red box?

ATTTGAATTC!

Magnified view

(Lecture materials, 2026)

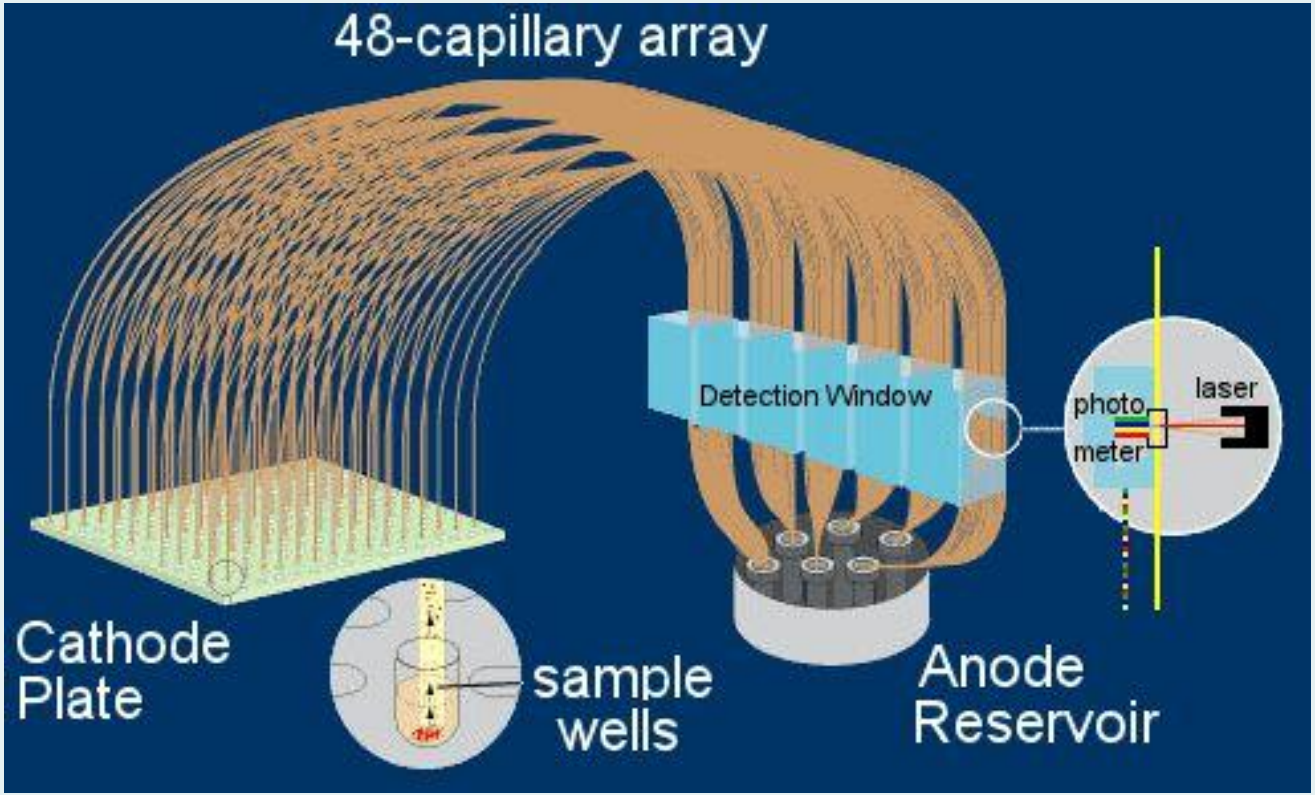
Chromatogram from Automated DNA sequencer

(Lecture materials, 2026)

Capillary use in automated sequencing

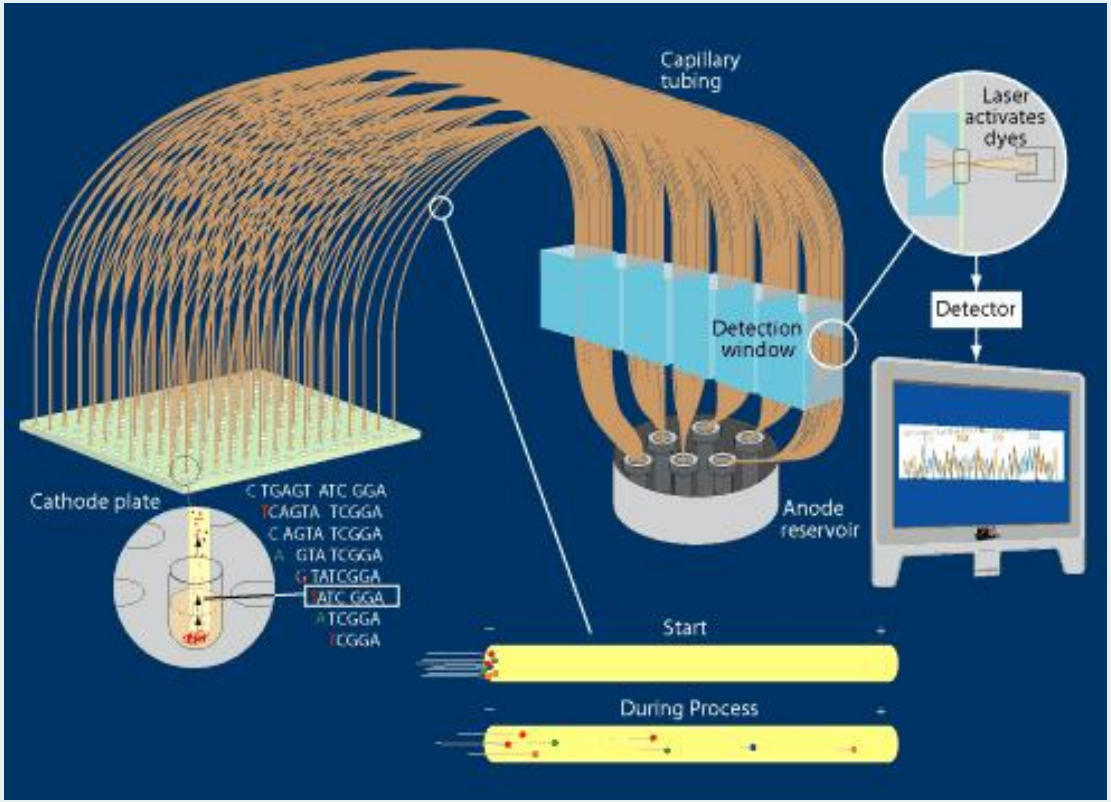
- 1990: Sanger sequencing evolves to use long, thin capillary fibre rather than gel to separate DNA fragments
- Machine makes coloured bases fluoresce, reads and records sequence
- Quicker and higher resolution than gel-based techniques
- Used for Human Genome Project!

Capillary electrophoresis



(Lecture materials, 2026)

Capillary electrophoresis



(Lecture materials, 2026)

Key differences

	Traditional Sanger	**Automated Sanger**
Method	ddNTP chain termination	ddNTP chain termination
DNA extraction	Manual	Kits, automated
Sequencing reaction	4	1
Labelling	Radiolabeling	Fluorescent

(Modified from: DNA Sequencing Methods: from Past to Present)

Automated:

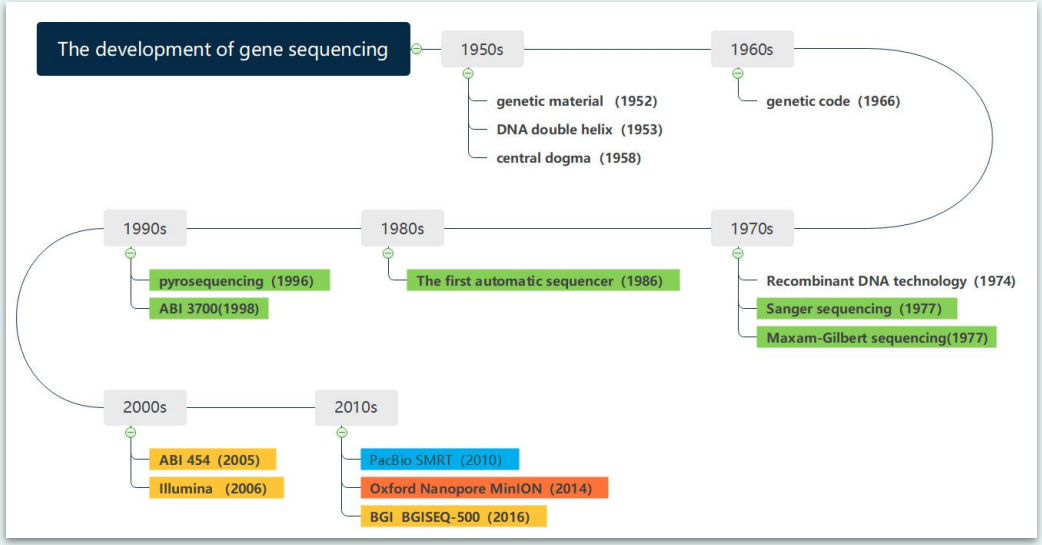
- Lower error
- Higher throughput, accuracy, efficiency, read length
- Safer

The Human Genome Project

- Mission to decipher human genetic material and provide research tools to analyze this information
- Highly collaborative
- Used automated Sanger sequencing with capillary electrophoresis
- 2003: 3.2×10^9 base pair genome sequence completed (92% of genome)
- 2022: Remaining gaps filled
- Basis for many findings regarding disease

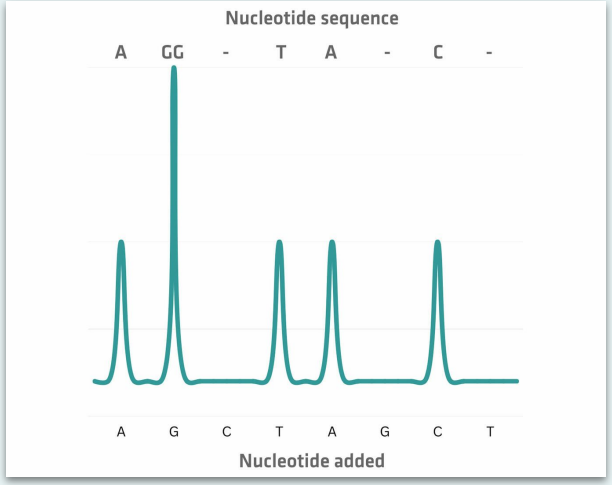
The Next Generation (TNG)

- TNG sequencing includes all sequencing methods post-automated Sanger (1st gen.)
- Later generations of sequencing are advantageous in high throughput, faster sequencing, and more cost effective.
- TNGs are as follows:
    - 2nd gen.:
        Pyrosequencing
        Illumina
    - 3rd gen.:
        PacBio
        Nanopore/ONT



2nd Generation: Pyrosequencing

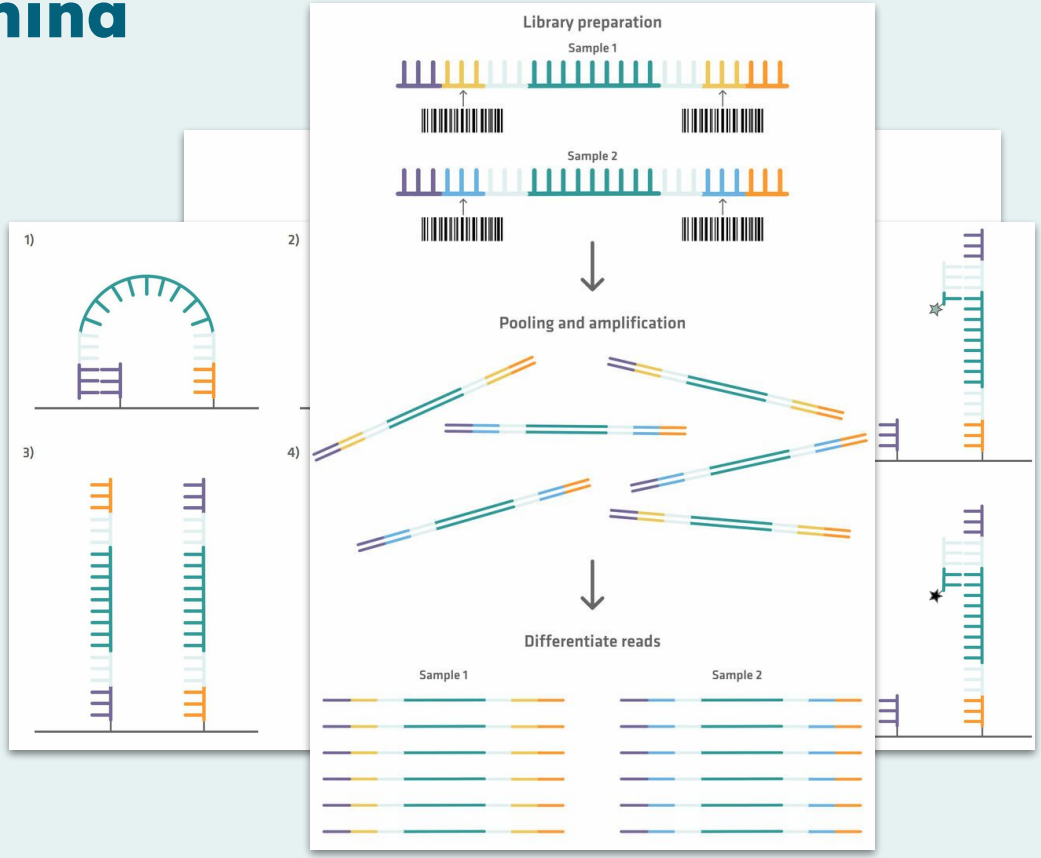
- Developed ~20 years after automated Sanger
- Became standard DNA sequencing method
- DNA sample amplified via PCR
- dNTPs are added one at a time
- If DNA polymerase incorporates dNTP, **pyrophosphate** is released
- Reaction to pyrophosphate produces light, peaks are recorded
- Apyrase added to degrade previous dNTP and next is added
- Nucleotide sequence determined from peaks



DNA sequencing methods: from Sanger to NGS, Mészáros, Éva.

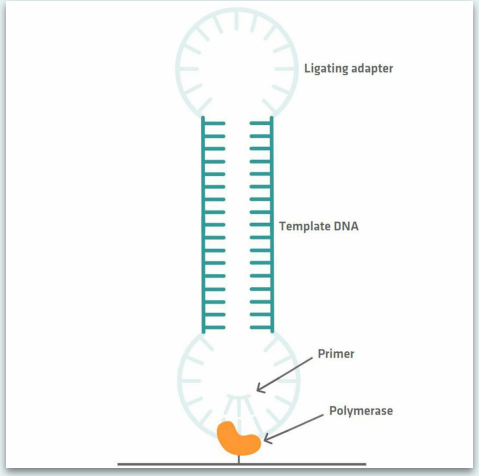
2nd Generation: Illumina

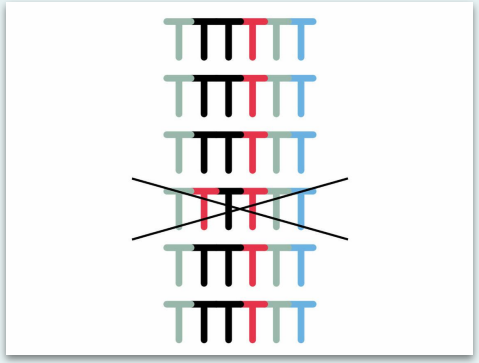
- Fixes problem from Pyro- where only one dNTP can be added at a time
- **Synthesis** method that can sequence millions of DNA fragments at once
- DNA samples fragmented, purified, and amplified
- Sequence determined simultaneously, then compared to “normal” reference genome



DNA sequencing methods: from Sanger to NGS, Mészáros, Éva.

3rd Generation: PacBio



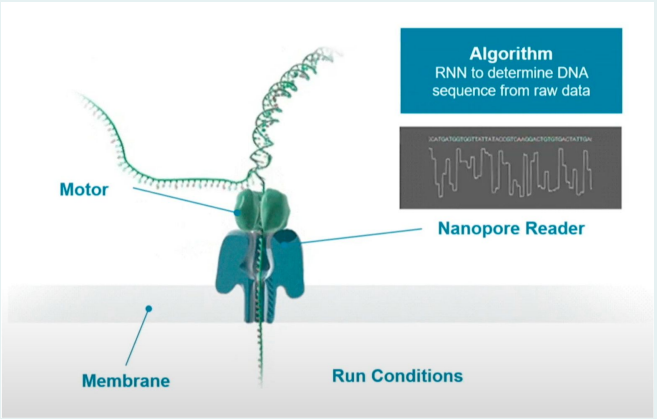


- PacBio fixes copying errors and time issues from Illumina
- Also called SMRT (**single-molecule real-time**) sequencing
- Library preparation does not require PCR
- SMRTbell library with circular DNA templates
- Polymerase enzymes immobilized in wells; Zero-Mode Waveguides (ZMWs)
- Millions of ZMWs form SMRT cell for PacBio sequencing
- Uses fluorescent dye specific for each nucleotide

DNA sequencing methods: from Sanger to NGS, Mészáros, Éva.

3rd Generation: Nanopore/ONT

- Oxford Nanopore Technology (ONT) also uses SMRT sequencing
- **Nanopores** and electrical currents instead of polymerase enzymes and labelled bases
- Adapter sequences added so motor protein can bind to DNA fragments
- Flow cell has nanopores in electro-resistant lipid membrane
- Electric current through pores threads negatively charged DNA
- DNA disrupts current differently for each nucleotide



Nanopore technology and its applications in Gene Sequencing, Lin, B, Hui, J, & Mao, H.

- Only nanopore tech is able of directly sequence RNA

From Sanger to TNG: Recap

Table 1. The comparison of different generation of gene sequencing.

	Reading Length (kb) N50	Estimated Cost per Gb (US $)	Throughput per Flow Cell (Gb)	Read Accuracy (%)
Sanger(1st)	<1 kb	13,000 [d]	/	>99.9
Illumina(2nd)	0.075–0.15 [a]	50–63	16–30	>99.9
PacBio(3rd)	10–20 [b]	43–86	15–30	>99
ONT(4th)	10–60 [c]	21–42	50–100	87–98

The estimated cost in a, b, c excludes the cost for labor, instrumentation, maintenance, and computer resources. [a] Illumina NovaSeq 550 works in single-end mode. [b] PacBio Sequel II works in HiFi mode. [c] ONT PromethION works in long mode. [d] The cost was calculated by the HCV genome from [21] and the cost includes the operating costs. Data in the table comes from [21–23].

Nanopore technology and its applications in Gene Sequencing, Lin, B, Hui, J, & Mao, H.

- 1st gen uses dideoxy terminators (ddNTPs)
- 2nd gen. uses sequencing-by-synthesis
- 3rd gen. (PacBio) uses SMRT
- 3rd/4th gen. (ONT) uses nanopore technology with electric currents

Thank you for listening!

Questions?

have a great day!

References:

Dovichi, N. J., & Zhang, J. (2000). How Capillary Electrophoresis Sequenced the Human Genome. *Angewandte Chemie International Edition*, *39*(24), 4463–4468.

Eren K, Taktakoğlu N, & Pirim I. (2022). DNA Sequencing Methods: From Past to Present. Eurasian J Med. 47-56.

Lin, B., Hui, J., & Mao, H. (2021, June 30). Nanopore technology and its applications in Gene Sequencing. MDPI. https://doi.org/10.3390/bios11070214

Mészáros, É. (2023, April 3). DNA sequencing methods: From Sanger to NGS. Integra. https://www.integra-biosciences.com/global/en/blog/article/dna-sequencing-methods-sanger-ngs

Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, & Baumeister K. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. Science. 1987 Oct 16;238(4825):336-41.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., & Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature (London)*, *321*(6071), 674–679.

References

PRIMARY PAPERS:

Hood, L. E., Hunkapiller, M. W., & Smith, L. M. (1987). Automated DNA sequencing and analysis of the human genome. Genomics, 1, 201–212.

Kabalka, G. W., Varma, M., Varma, R. S., Srivastava, P. C., & Knapp, F. F. (1986). Tosylation of alcohols. J. Org. Chem., 51, 2386–2388.

Lin, B., Hui, J., & Mao, H. (2021). Nanopore technology and its applications in gene sequencing. Biosensors, 11, 214.

Mészáros, É. (2023, April 3). DNA sequencing methods: From Sanger to NGS. Integra. https://www.integra-biosciences.com/global/en/blog/article/dna-sequencing-methods-sanger-ngs

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, 74(12), 5463–5467.

Tipson, R. S. (1944). On esters of p-toluenesulfonic acid. J. Org. Chem., 9, 235–241.

APPROVED REFERENCES:

-

-

-