DNA SEQUENCING

Techniques:

1. Maxam e Gilbert:first method

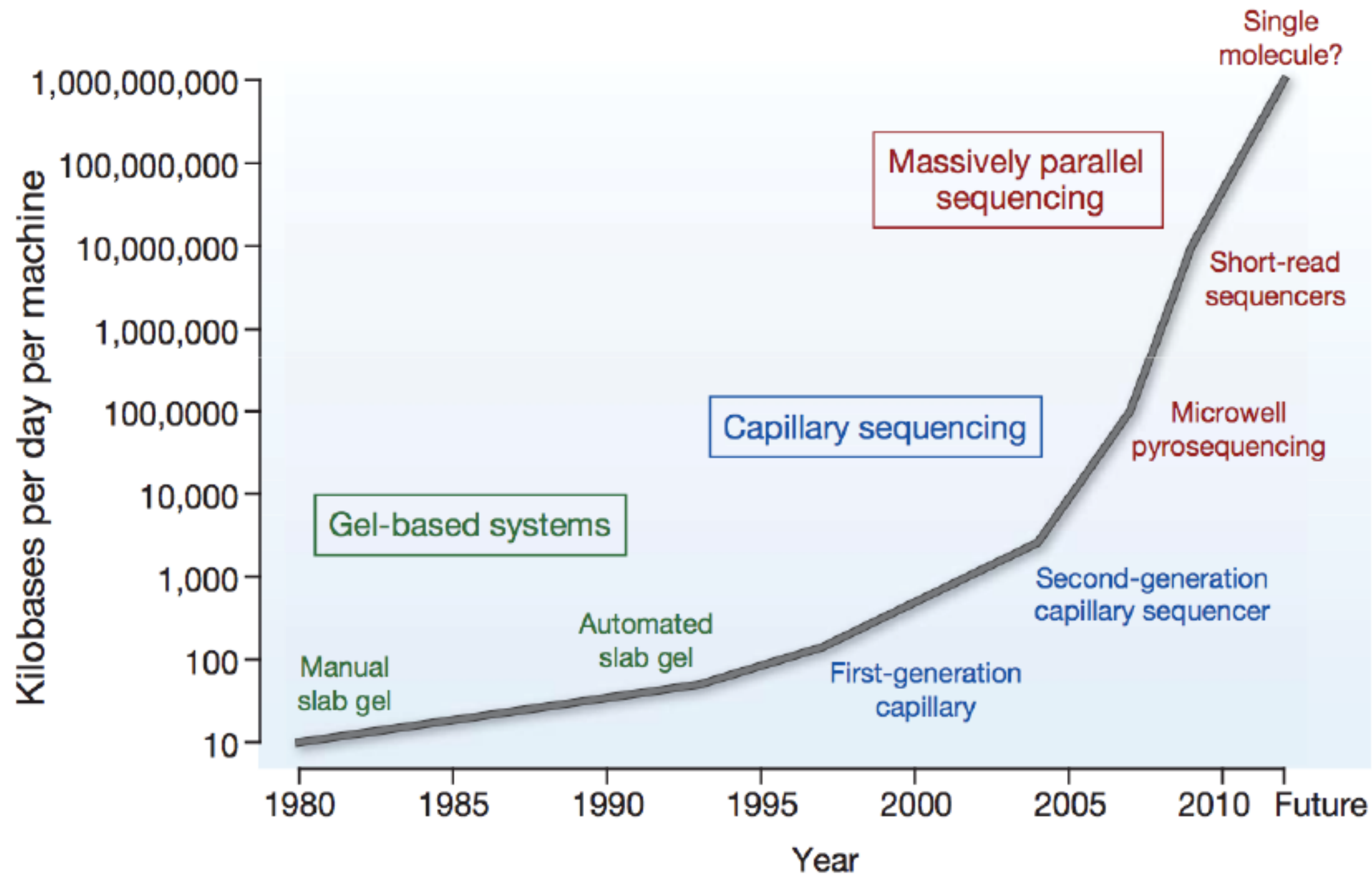
2. Sanger Sequencing:

basis for all seqeuncing tecniques

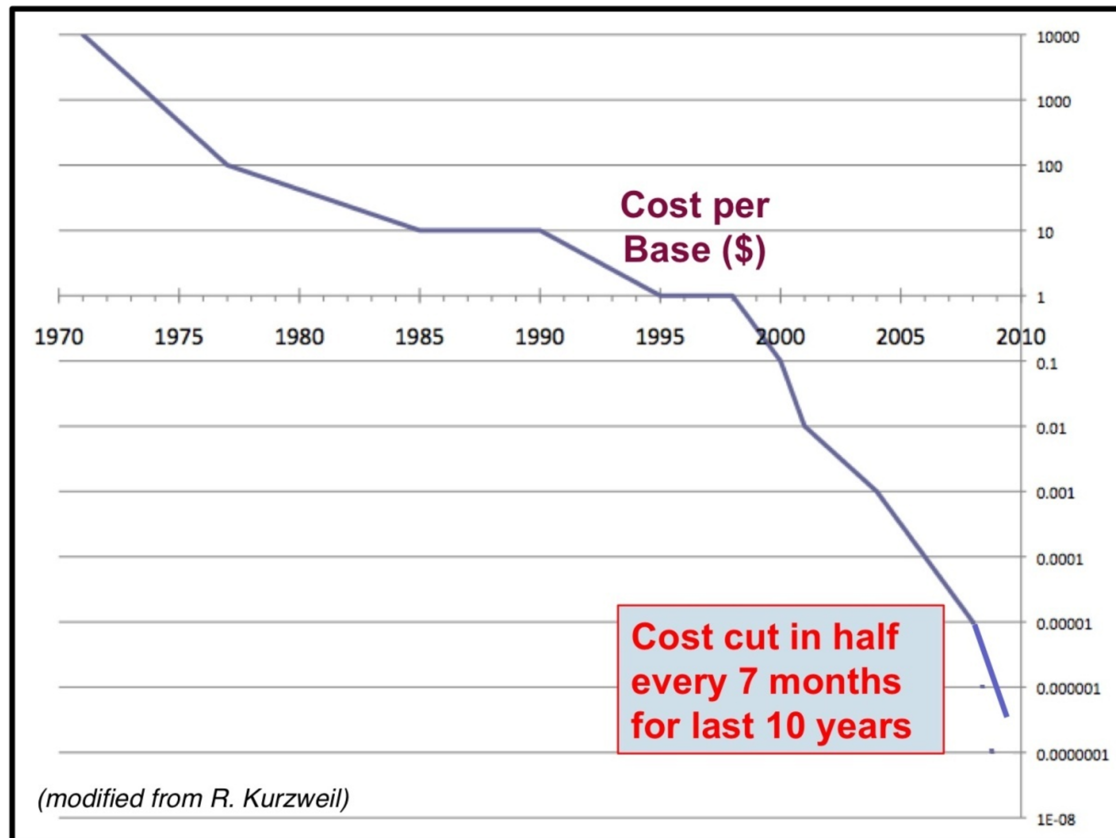
3. Massive Parallel Seqeuncing

DNA sequencing includes several methods and technologies that are used for determining the order of the nucleotide bases—adenine, guanine, cytosine, and thymine—in a molecule of DNA.

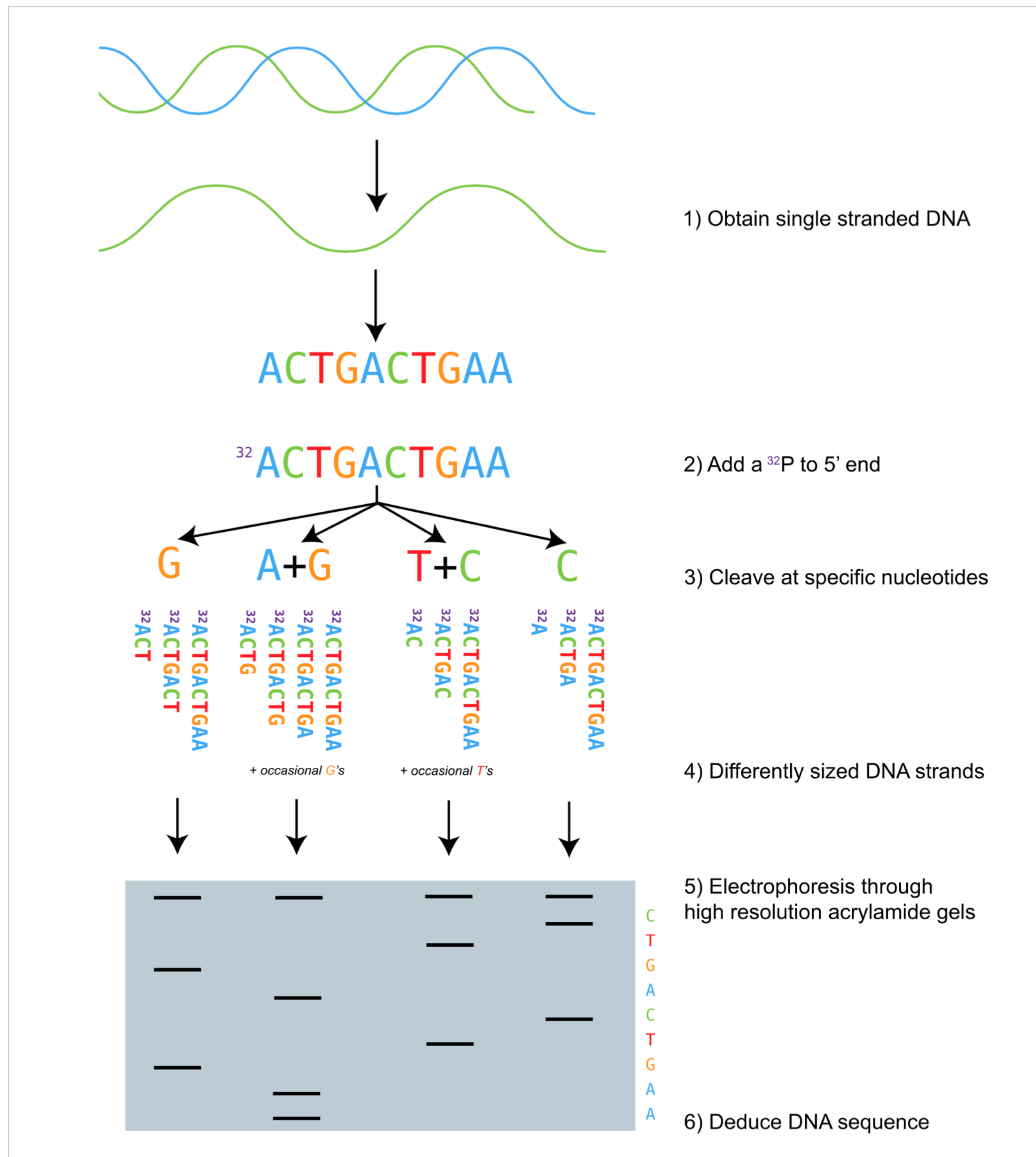
The History of DNA Sequencing Technology



Quickly reduced Cost



1. Maxam-Gilbert Method chemical sequencing



1. Denature a double-stranded DNA to single-stranded by increasing temperature.
2. adioactively label one 5' end of the DNA fragment to be sequenced by a kinase reaction using gamma-32P.
3. Cleave DNA strand at specific positions using chemical reactions. For example, we can use one of two chemicals followed by piperdine.

- Dimethyl sulphate selectively attacks purine (A and G)
- hydrazine selectively attacks pyrimidines (C and T).

The chemical treatments outlined by Maxam-Gilbert cleaved at G, A+G, C and C+T.
A+G means that it cleaves at A, but occasionally at G as well.
- Now in four reaction tubes, we will have several differently sized DNA strands.
- Fragments are electrophoresed in high-resolution acrylamide gels for size separation.
- These gels are placed under X-ray film, which then yields a series of dark bands which show the location of radiolabeled DNA molecules. The fragments are ordered by size and so we can deduce the sequence of the DNA molecule.

Pros
Maxam-Gilbert sequencing was at one point more popular than the Sanger method. Purified DNA could be used directly, while the Sanger method required that each read start be cloned for production of single-stranded DNA.

Cons
Cons included difficulties scaling up, and the handling of X-rays and radiolabeling, which were harmful to technicians.

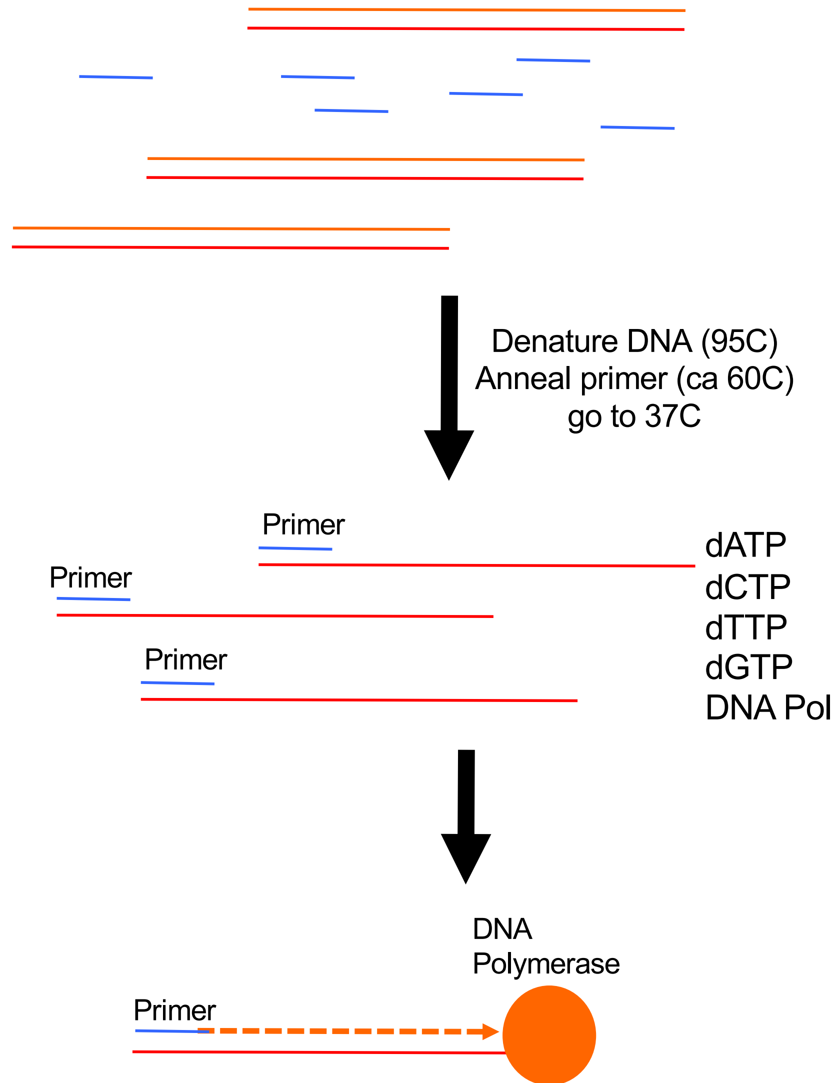
2. Sequenziamento di DNA mediante il metodo di Sanger

Sequenziamento con il metodo dei **dideossinucleotidi**

F. Sanger 13 agosto 1918 – 19 novembre 2013

Due premi Nobel. Uno per il sequenziamento dell'insulina ed uno per il sequenziamento del genoma del fago ϕ–X174

General concept in a sequencing reaction:
The synthesis of a new strand of DNA from a ss template DNA



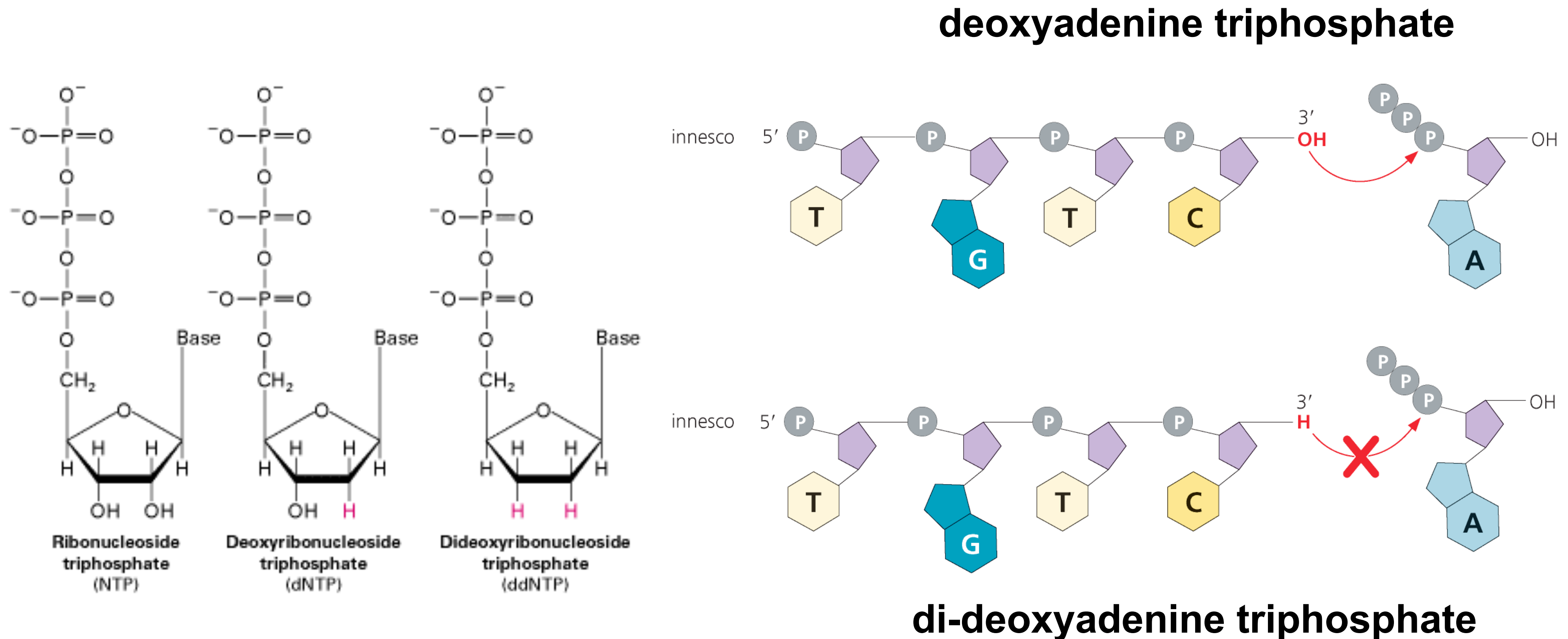
Mix primer oligonucleotides and
Many identical dsDNA molcules

add dATP, dTTP, dCTP, dGTP
and DNA polymerase

Polymerase elongates template DNA

PROBLEM: HOW CAN WE READ THE NEWLY SYNTHEZISED DNA SEQUENCE??

A great trick: using di-deoxyribonucleoside triphosphates to terminate the synthesis of DNA molecules

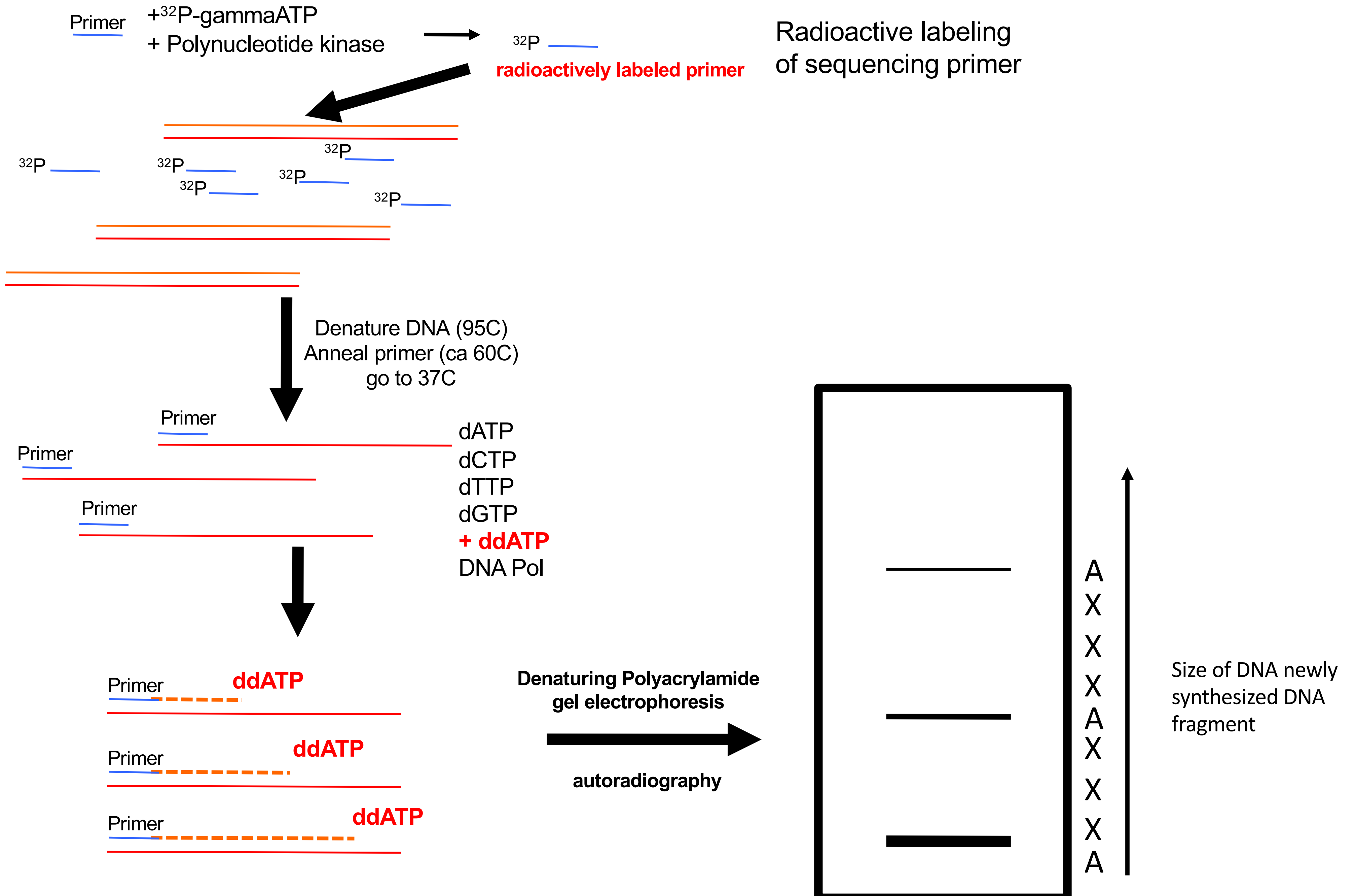


Concept: mixing a low amount of ddATPs into a high amount of dATP:
A pool of DNA molecules will be generated in which DNA molecules terminate at all possible A sites.

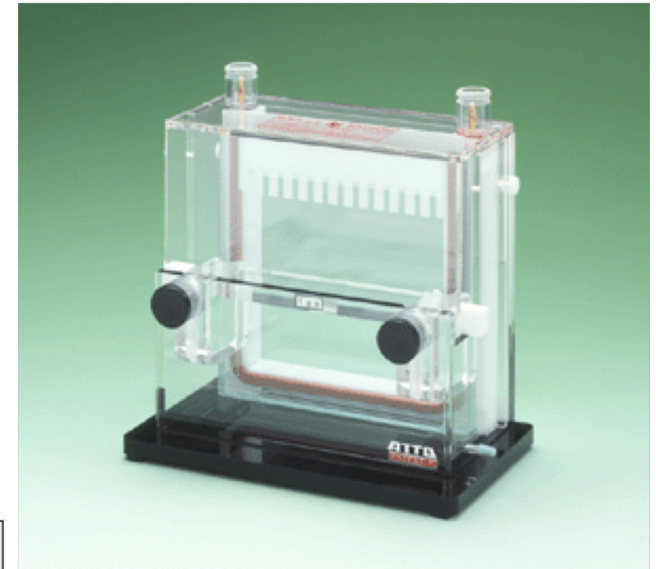
ddNTP enable me to terminate sequencing at a defined position in the newly synthesized DNA molecule

PROBLEM: How can detect sequencing products??

General concept in a sequencing reaction: The synthesis of a new strand of DNA from a ss template DNA



PolyAcrylamide Gel Electrophoresis



Sample wells

Cathode –

Buffer

Stacking gel

Running gel

Anode +

Buffer

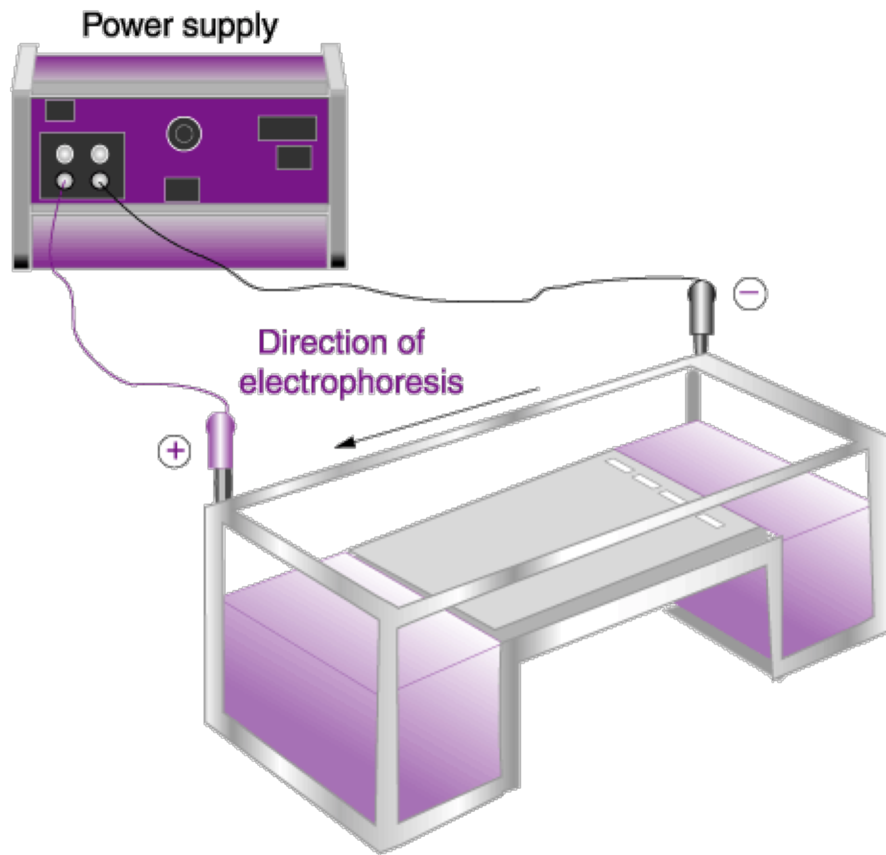
Plastic frame

PAGE

Le due basilari tecniche elettroforetiche per la separazione di frammenti di DNA (e di RNA).

- **Elettroforesi su gel di poliacrilammide (PAGE)**:
- il gel è ottenuto per polimerizzazione in soluzione acquosa tamponata di acrilammide con una piccola percentuale di bisacrilammide tra due vetri con intercapedine di 0,5–2 mm mantenuti verticalmente. Il gel è costituito da una rete tridimensionale covalente del polimero ed è sostanzialmente un gel irreversibile.

- **Elettroforesi su gel di agarosio**:
- il gel è ottenuto colando su un contenitore orizzontale una soluzione calda di agarosio (un polisaccaride algale), che per raffreddamento gelifica in seguito alla formazione di interazioni non covalenti tra le catene del polimero. Il gel è pertanto reversibile e può essere riportato a soluzione tramite successivo riscaldamento.



Agarose gel electrophoresis

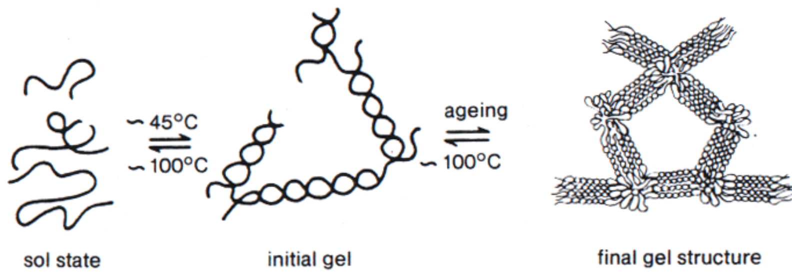
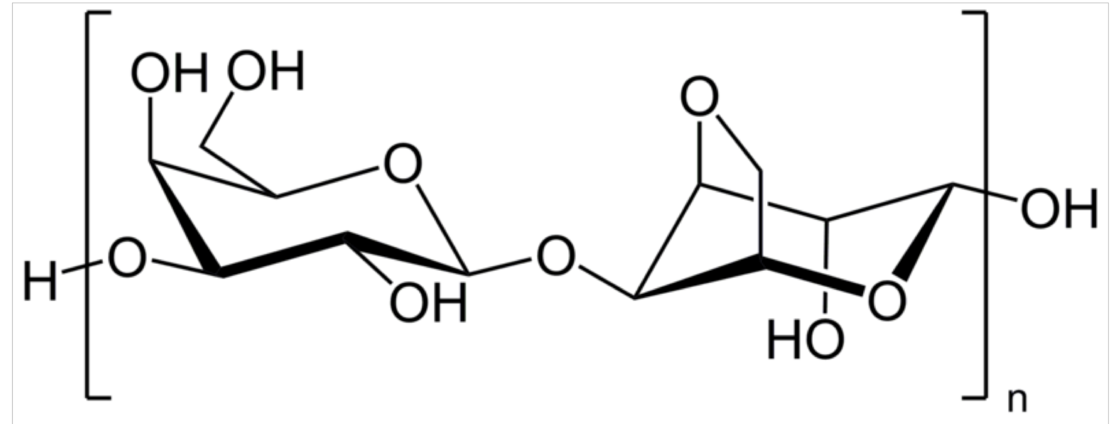
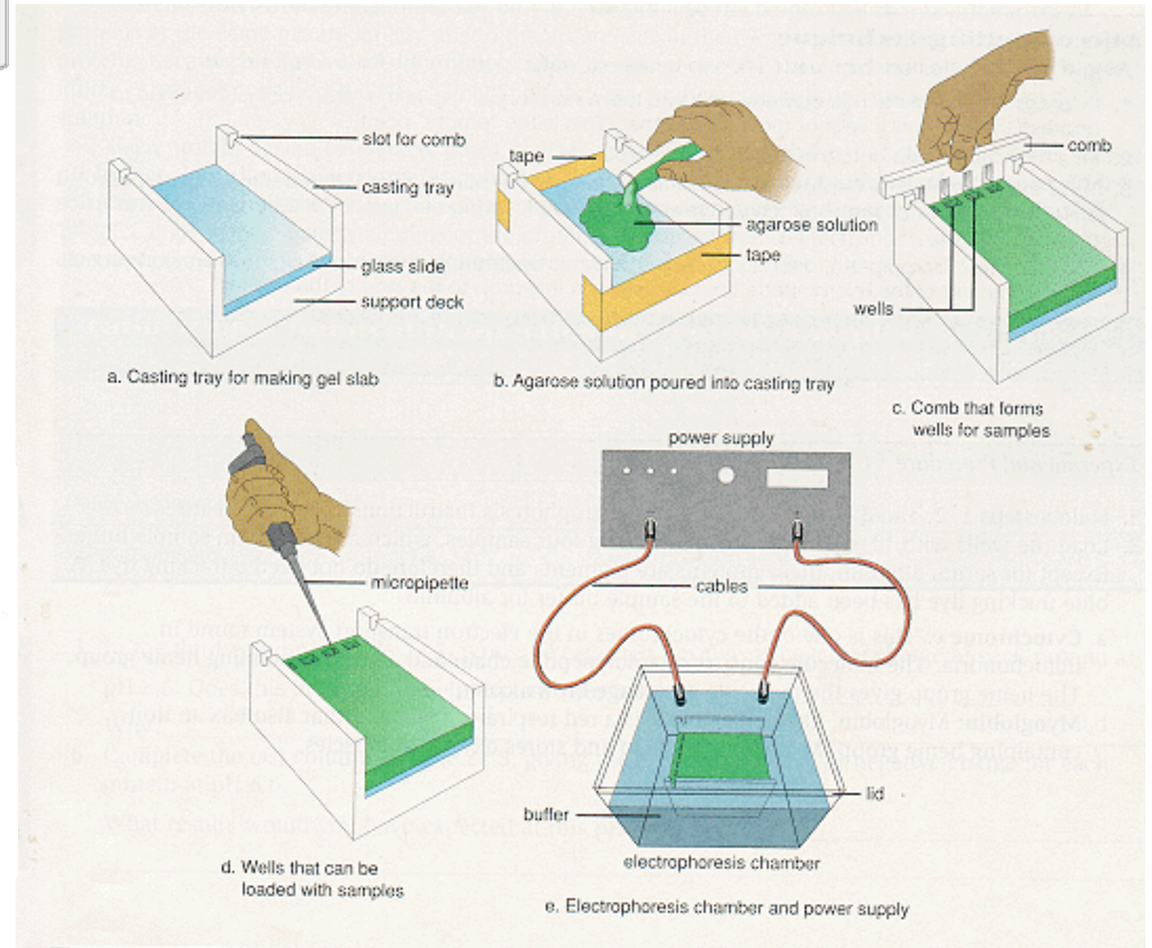
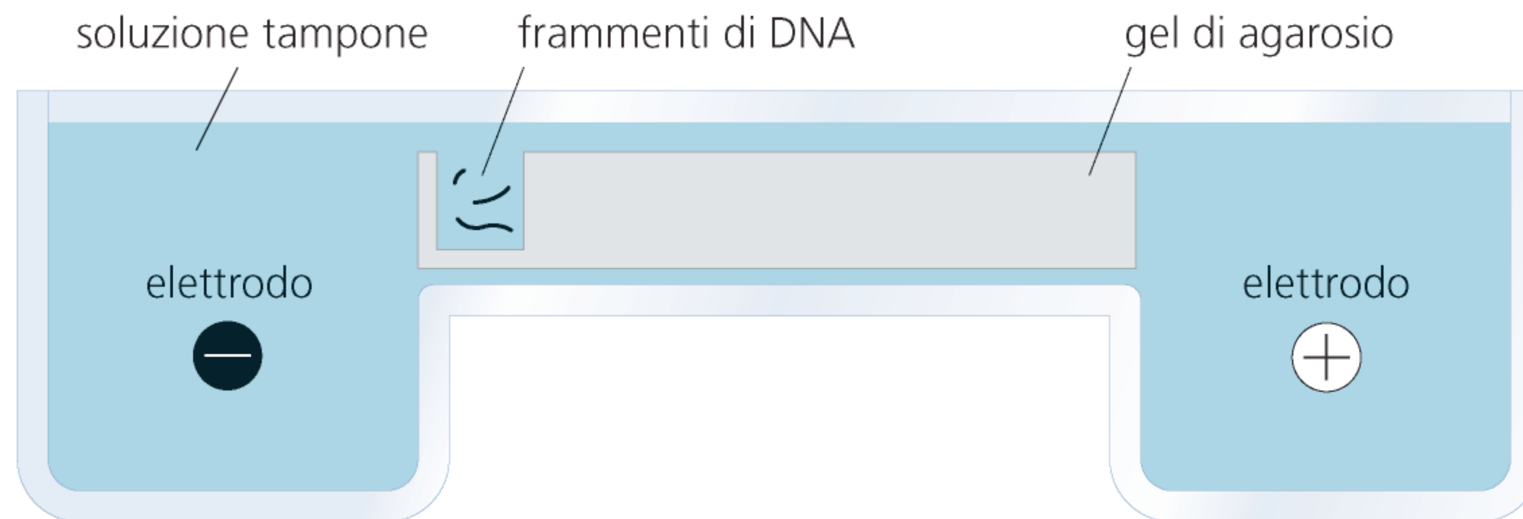


Fig. 25. Gel structure of agarose. (Låås, T. Doctoral thesis. Acta Universitatis Upsaliensis 1975. Reproduced by kind permission of the Author.)

Agarose repetitive unit

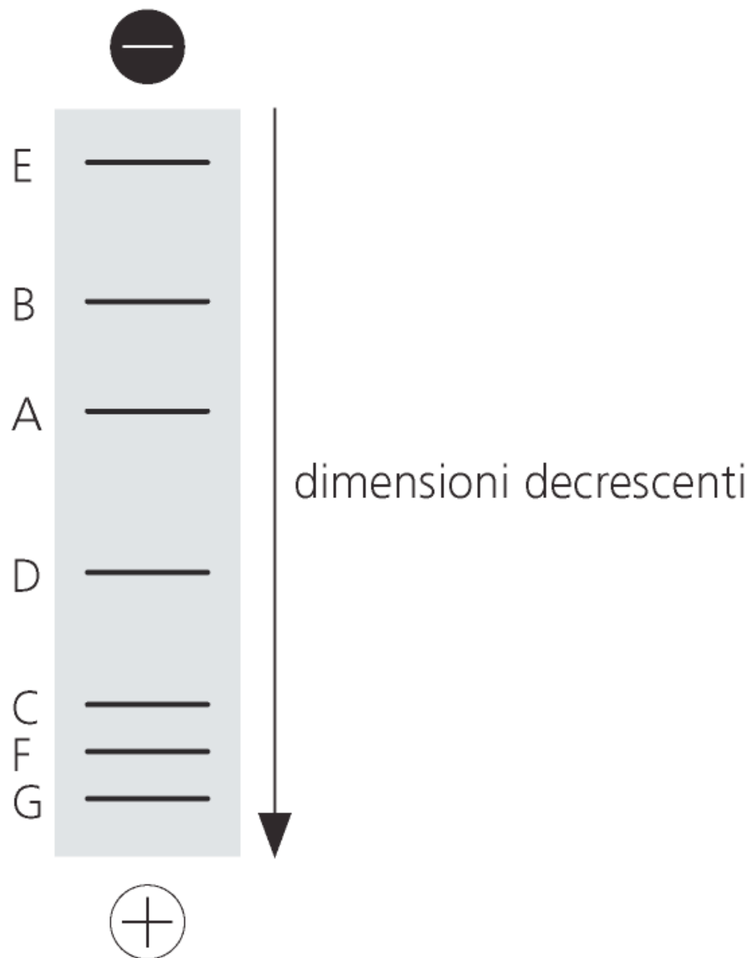


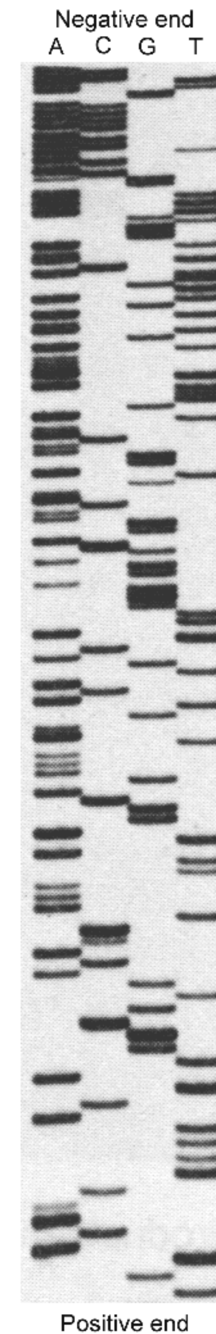



camera di elettroforesi
soluzione tampone
frammenti di DNA
gel di agarosio
elettrodo
elettrodo

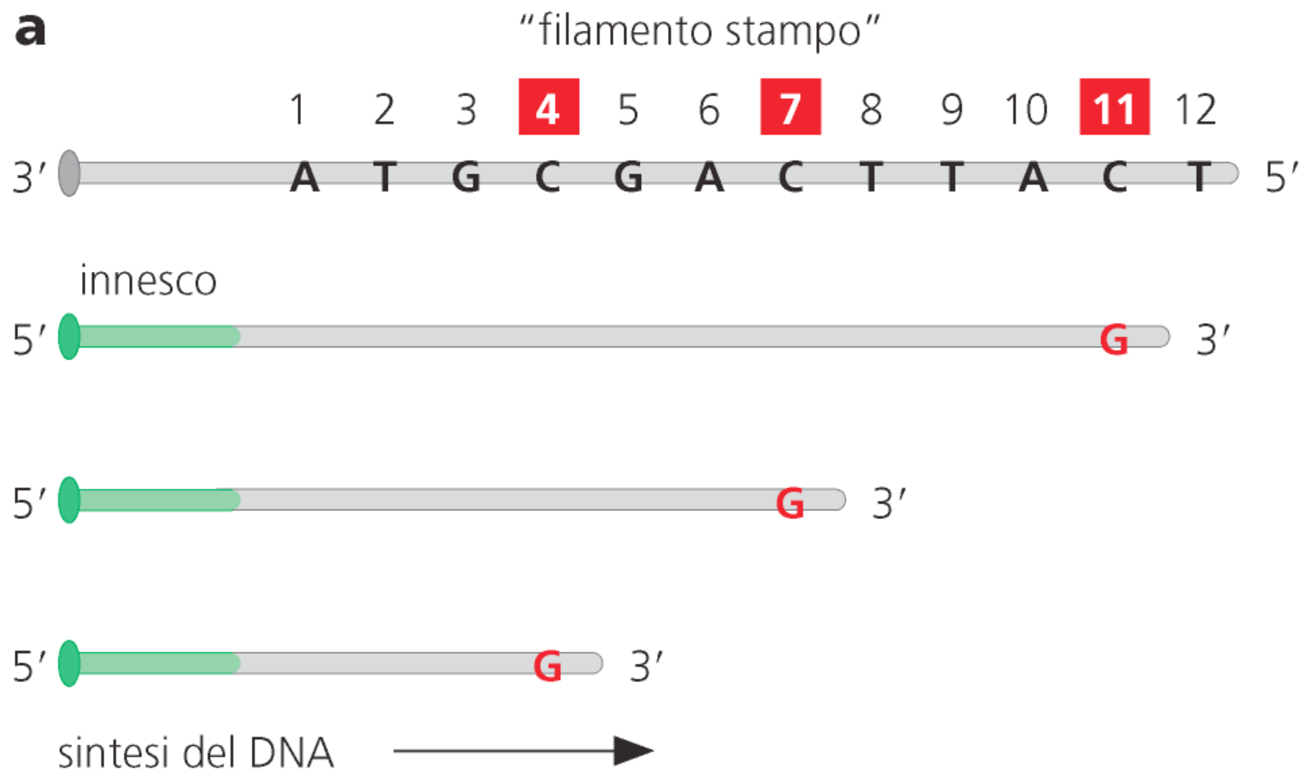
elettrodo
elettrodo
i frammenti di DNA piccoli
avanzano nel gel più
di quelli grandi

Primer is radioactively labelled!!
All fragments produced by DNA Polymerase can be visualized by autoradiography

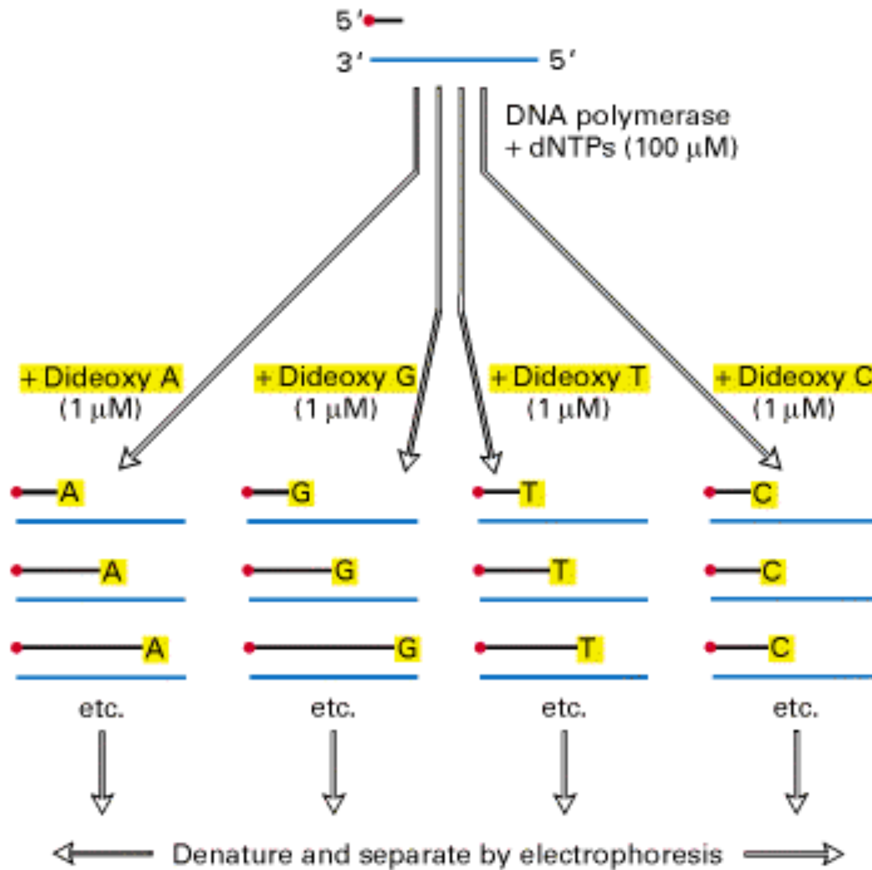


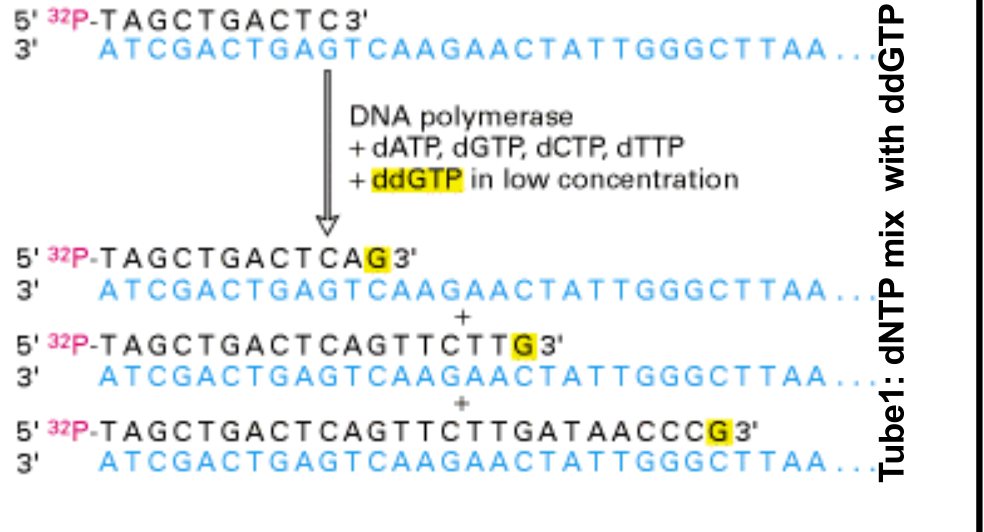


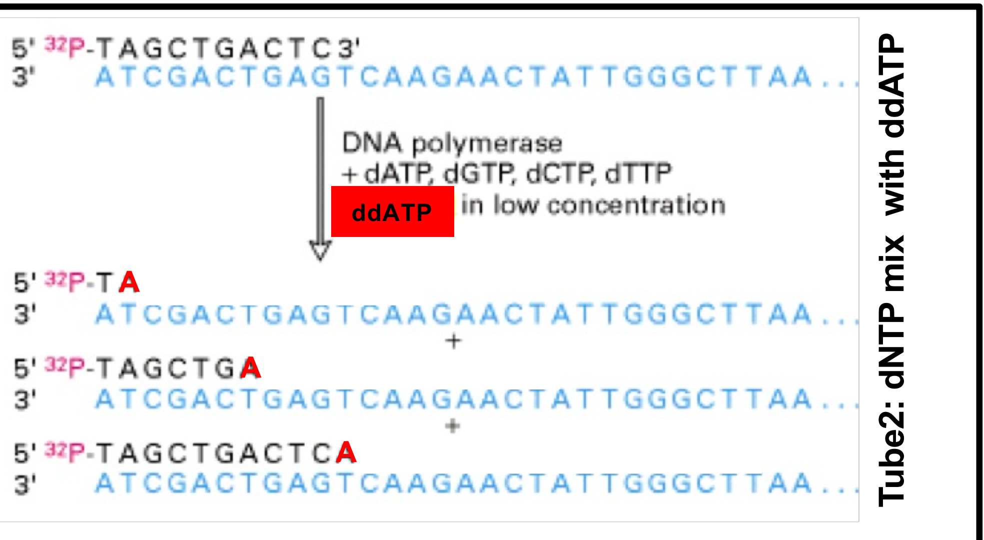
DNA syntesis is terminated by the use of ddNTPs



Classic Sanger sequencing of a DNA fragment requires 4 parallel Sequencing reactions

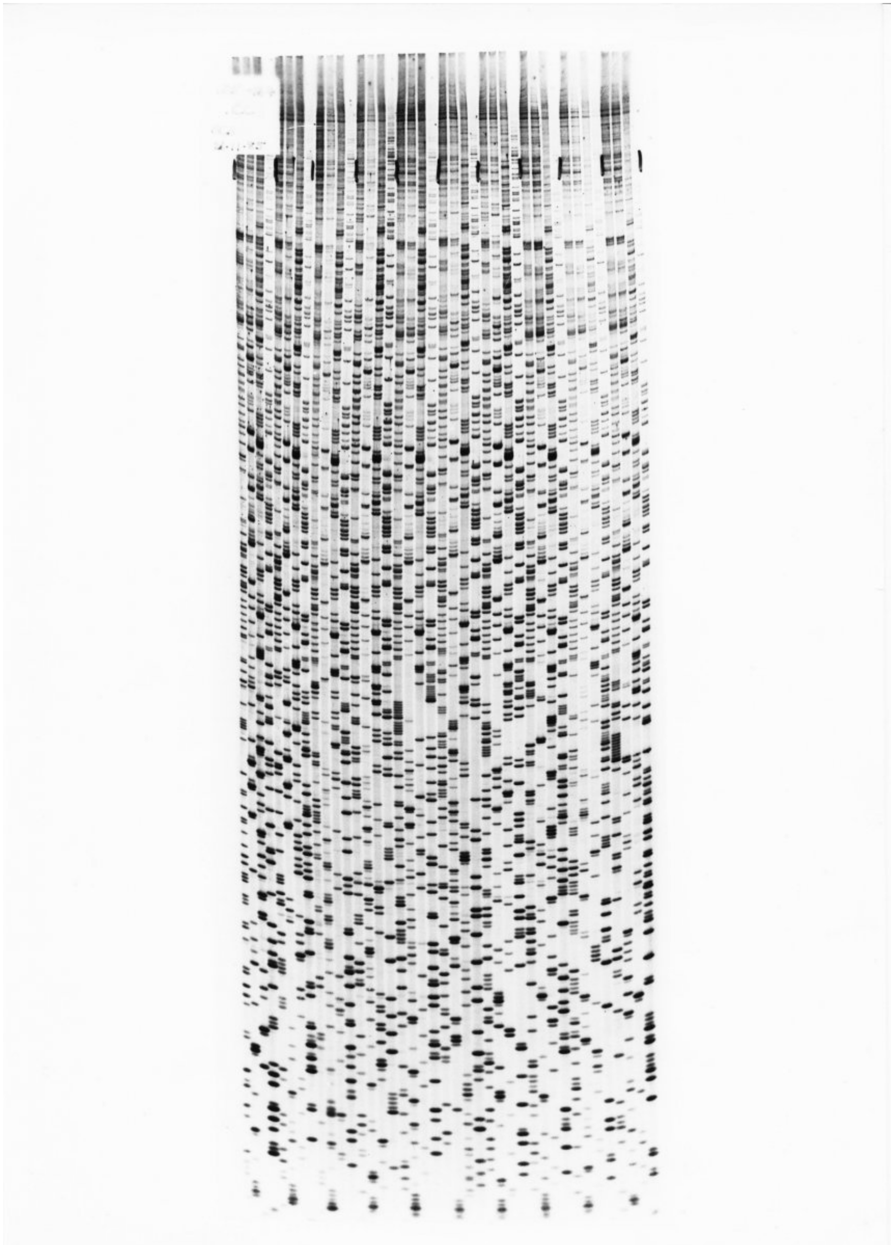


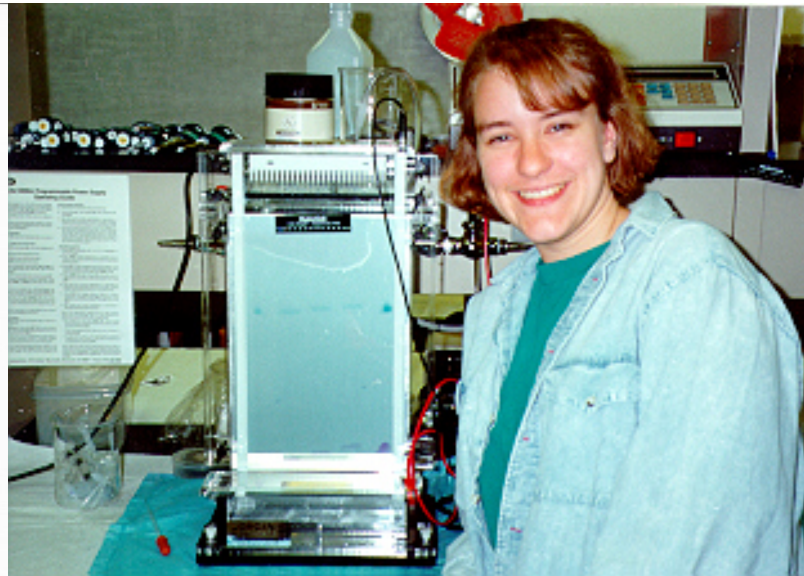


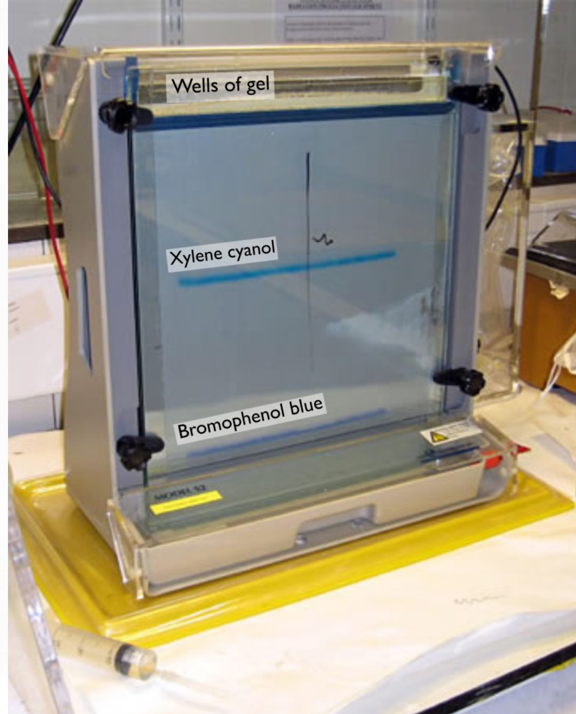


Tube3: dNTP mix with ddCTP

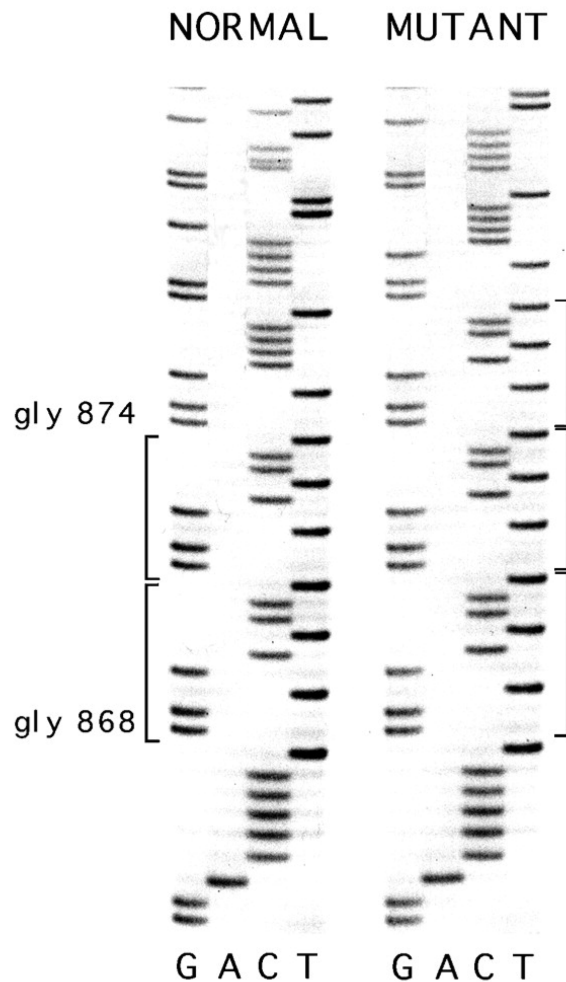
Tube4: dNTP mix with ddTTP







Standard until ca. 1995



Can you read the DNA sequence?

NORMAL: GGT GCT CCT GGT GCT CCT GGT GCC CCT GGC CCC GTT GGC CCT GCT

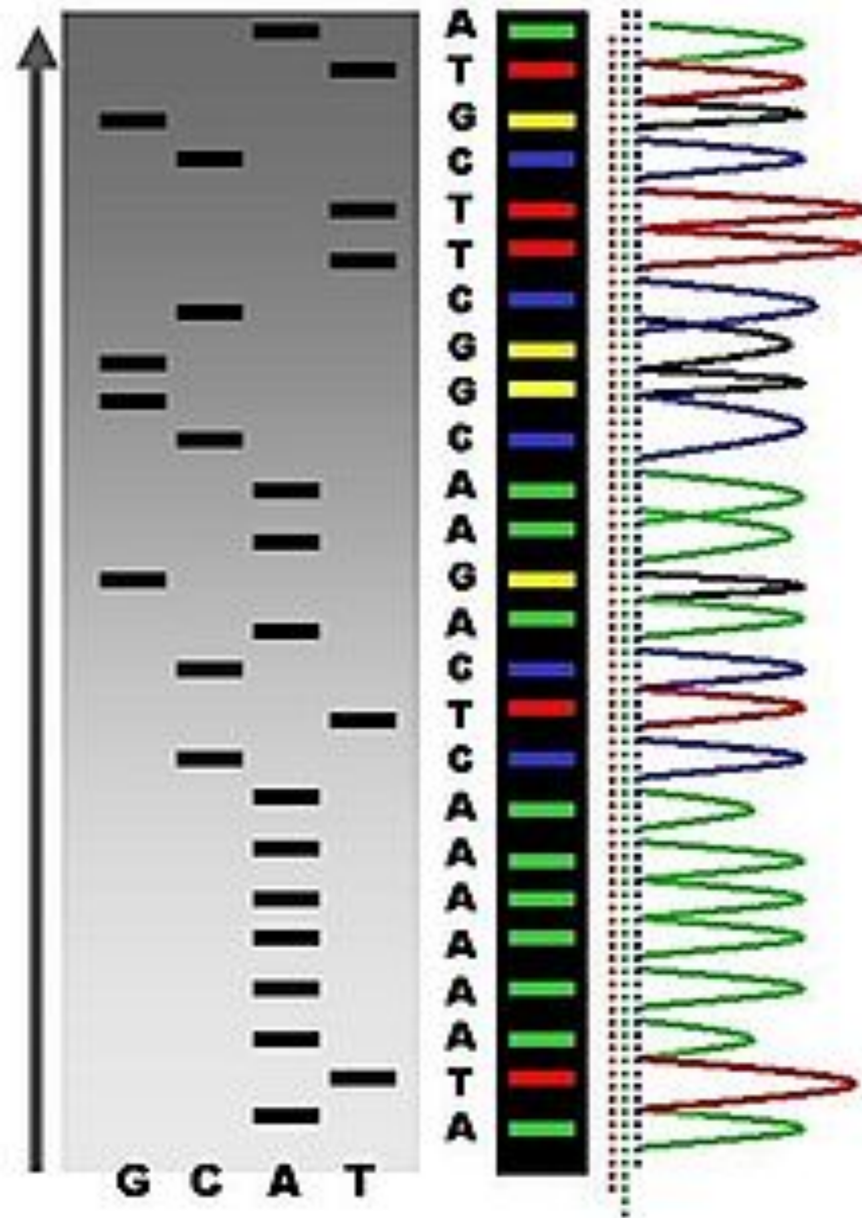
AMMINO ACID SEQEUNCE: G A P G A P G A P G P V G P A

MUTANT: GGT GCT CCT GGT GCT CCT GGT GCT CCT GGT GCC CCT GGC CCC GTT

AMMINO ACID SEQEUNCE: G A P G A P G A P G A P G P V

2.1. Automated sequencing based on Sanger technique

Dye-terminator Sanger sequencing



Classic radioactive

Dye-terminator Sanger sequencing

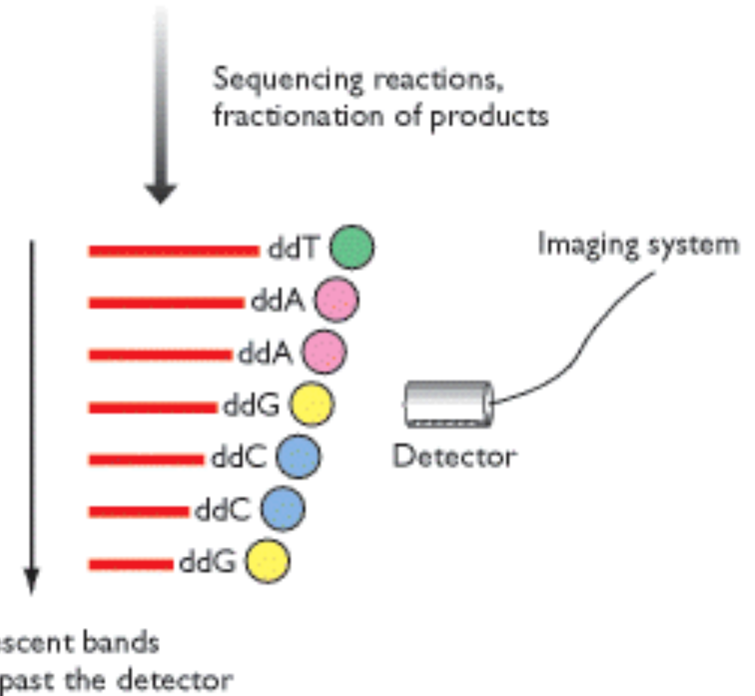
Dye-terminator Sanger sequencing

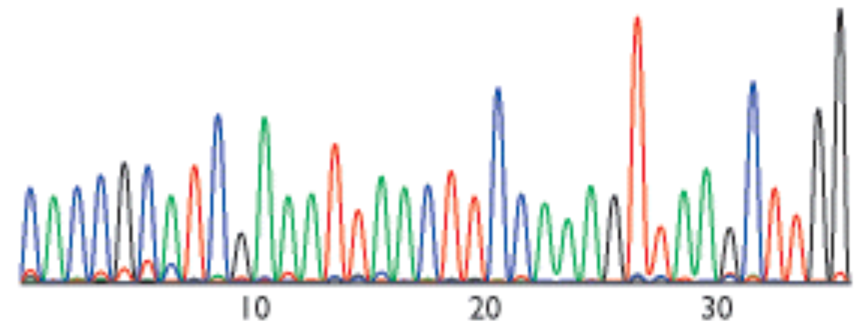
Tre diversi modi di marcare i frammenti di Sanger:

1) I frammenti di Sanger sono resi radioattivi per incorporazione di α-dNTP marcato

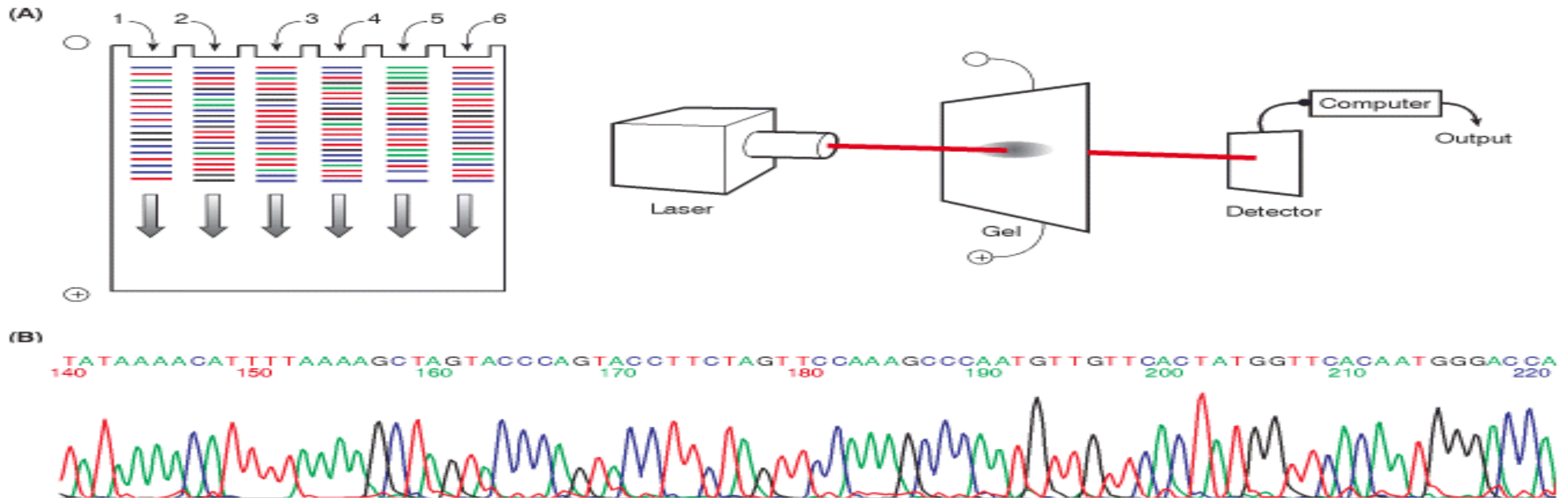
Questo metodo richiede quattro reazioni di polimerizzazione separate e quattro corsie elettroforetiche.

2) Ciascun ddNTP è reso fluorescente con un fluoroforo diverso. Questo metodo consente anche di effettuare tutte le reazioni in un'unica provetta.





Labeling each dideoxy requires only one lane

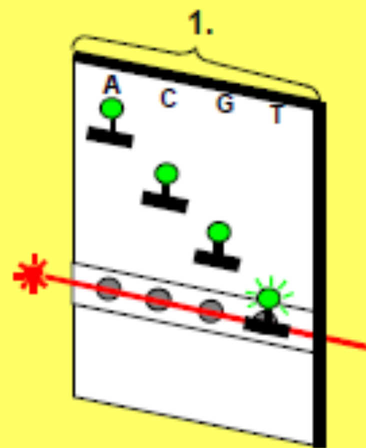


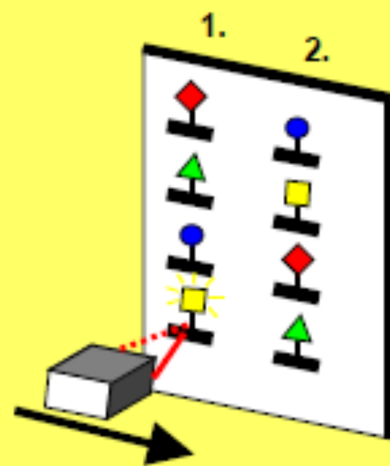
Model 3700 d00462_A05_Tas6up_033.ab1
Version 3.6
Basecaller-POP5opt.bcpTas6up
BC 1.1.b.2 Cap 33

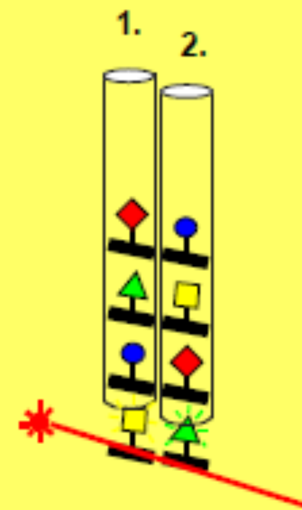
Signal G:172 A:243 T:195 C:173
DT3700POP5{BD}v3.mob
elu
Points 2767 to 13845 Pk 1 Loc: 2767

Tue, Sep 12, 2000 2:37 PM
Tue, Sep 12, 2000 1:21 AM
Spacing: 15.52{15.52}

Single-dye systems	4-dye systems	
slab-gel systems		capillary systems

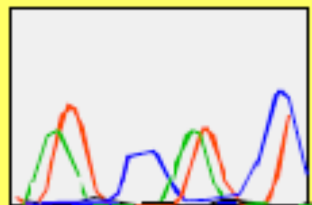


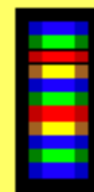




ELECTROPHORESIS

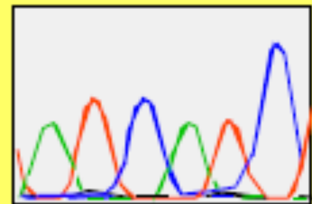
DATA COLLECTION

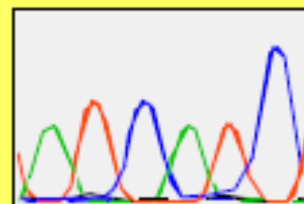


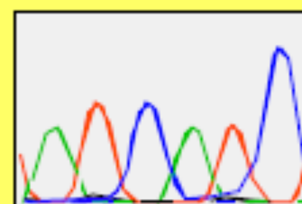


RAW DATA

PROSESSING





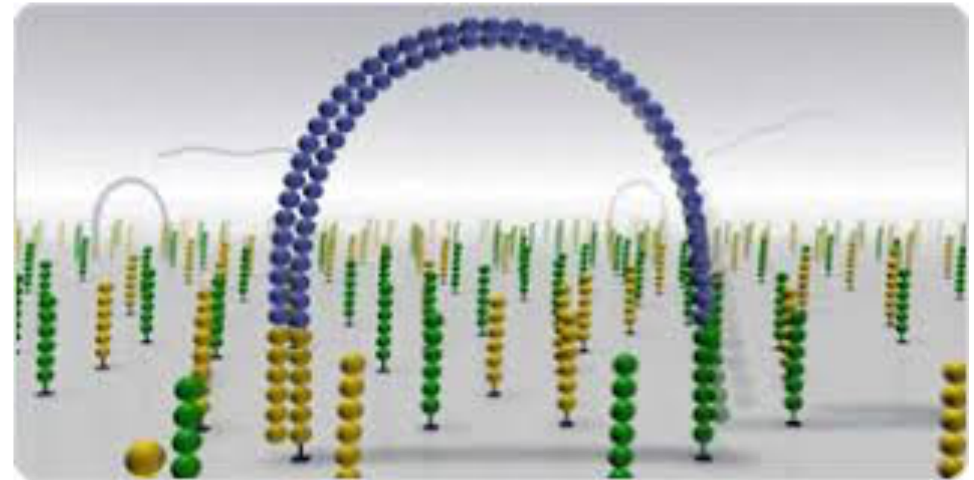


PROCESSED DATA

LOW CAPACITY → HIGH CAPACITY

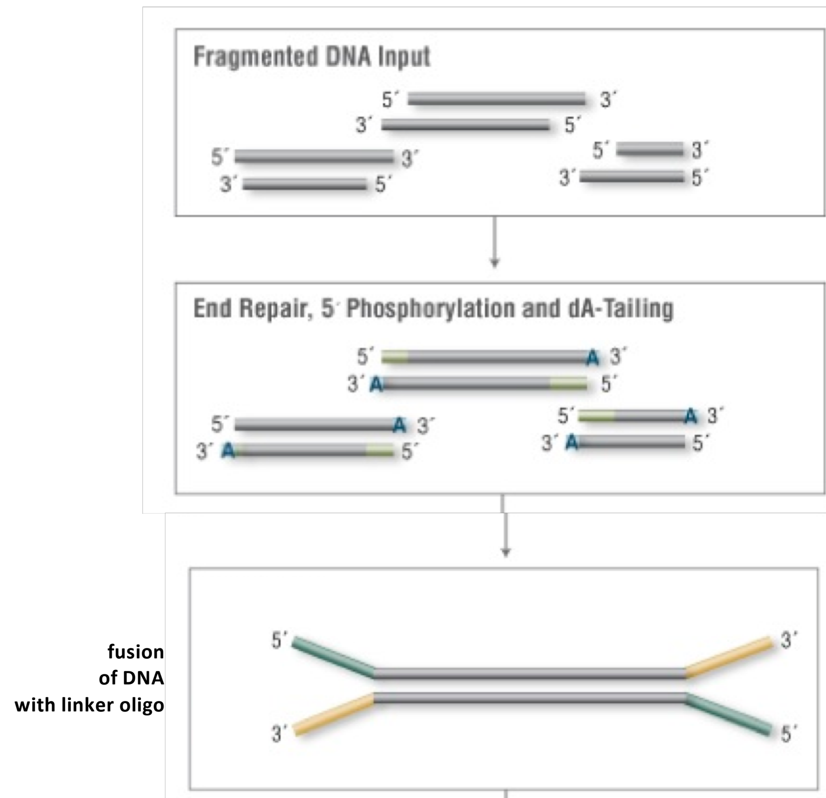
Sarén, A-M *et.al.* Kemia-Kemi 1996, 23, 724-727

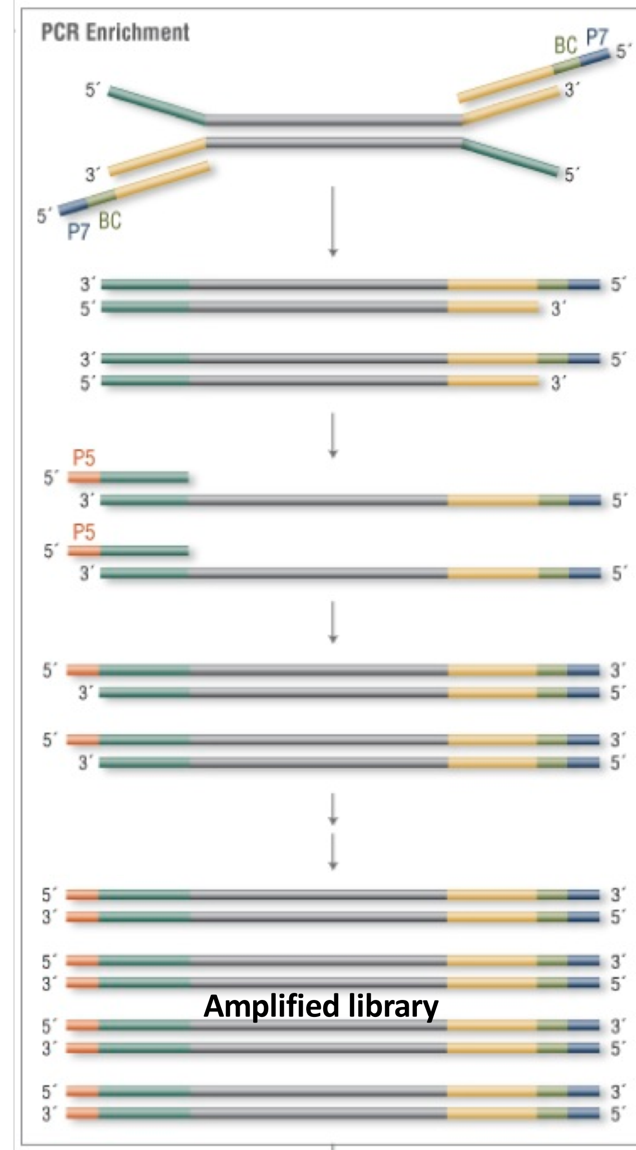
3. Massive parallel seqeuncing



Next generation sequencing of pools of DNAs

Use DNA to generate DNA libraries:
→ Genomic DNA (fragemented)
→ cDNA





Linkers serve as uniform primer binding sites. This allows the amplification of the entire DNA library using only 2 types of oligonucleotides

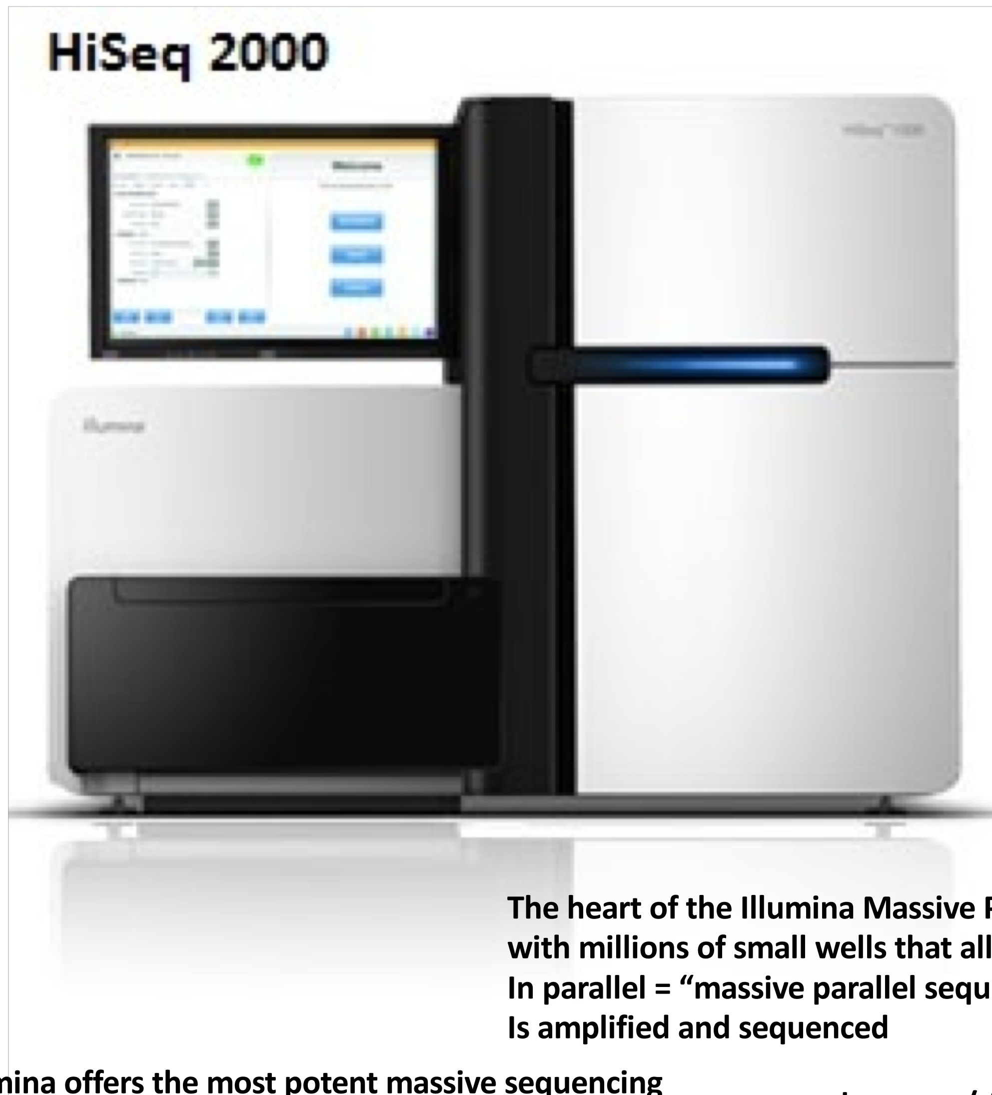
BC: barcode. Each biological sample has common P7 oligos (blue and yellow) and P5 oligos (red/green); however for each biological sample a defined BC sequence is chosen. This lings the sequencing result to the biological sample
→ Many samples can be sequenced at the same time

READY FOR MASSIVE PARALLEL SEQEUNCING

ChIP seq: Analysis of epigenetic information on the single nucleotide level
→ GENERASTION OF GENOME WIDE EPIGENTIC MAPS

Illumina Massively Parallel Sequencing

https://www.illumina.com/company/video-hub/pfZp5Vgsbw0.html



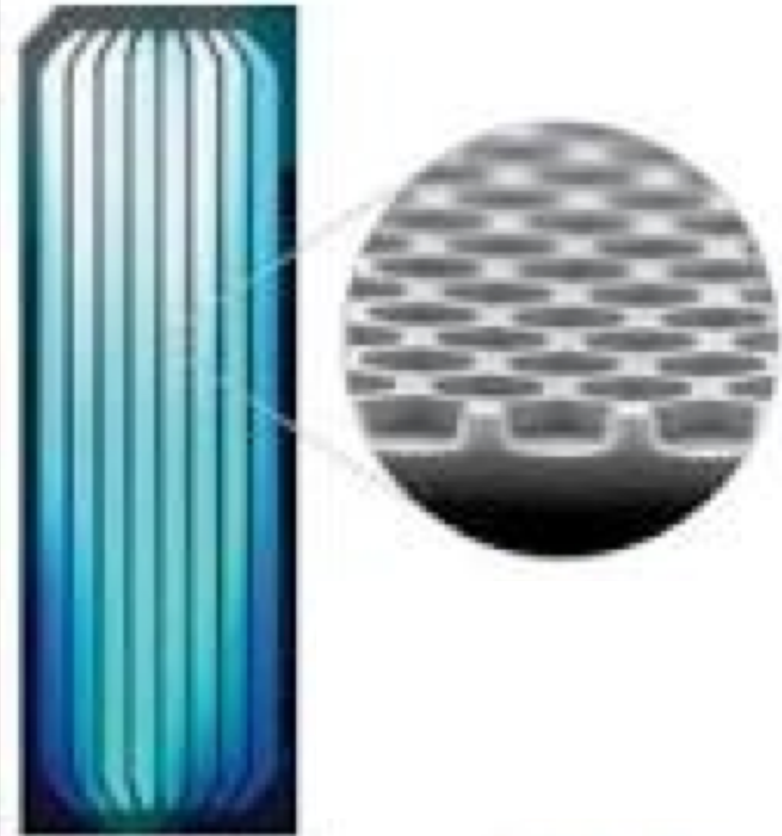
The heart of the Illumina Massive Parallel Sequencer is the "FLOW-CELL". A surface with millions of small wells that allow thousands of Sanger-sequencing reaction In parallel = "massive parallel sequencing". In each well a SINGLE MOLECULE of DNA Is amplified and sequenced

Illumina offers the most potent massive sequencing instruments – leader on the market

https://www.youtube.com/watch?v=pfZp5Vgsbw0

ChIP seq: Analysis of epigenetic information on the single nucleotide level
→ GENERASTION OF GENOME WIDE EPIGENTIC MAPS

CLUSTER AMPLIFICATION:



Flow cell contains surface with millions of wells

→Each well contains beads mounted with 2 species of oligonucleotides that hybridize with adaptor oligos of DNA library

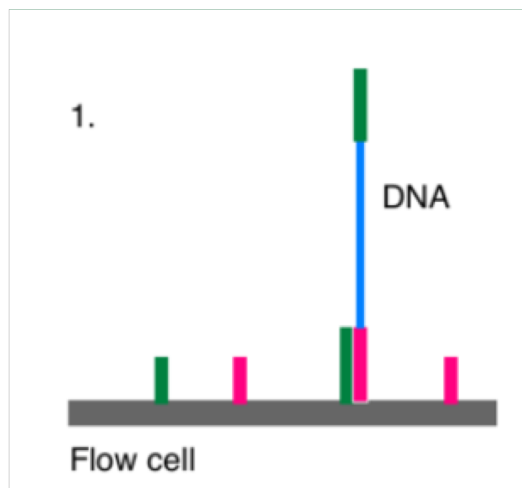
→DNA library will be loaded onto the flow cell in a determined concentration:

ONLY ONE MOLECULE OF DNA WILL BE PROCESSED FOR SEQUENCING IN A SINGLE WELL

CLUSTER AMPLIFICATION:

-making DNA library (~300bp fragments)

-ligation of adapters **A** and **B** to the fragments



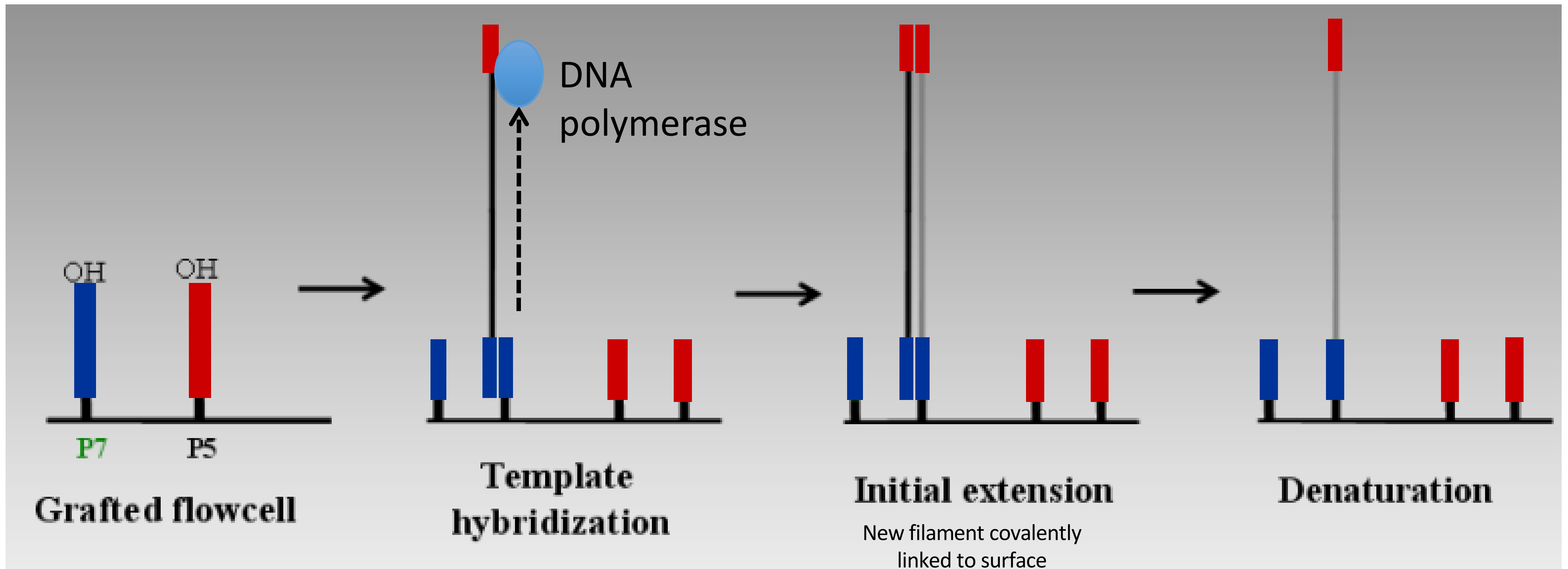
1 well, covered with millions of 2 types of oligos

1 well in a flow-cell with billions of wells

- ***complementary*** primers are ligated to the surface
- pairing with ChiP ed ssDNA at random position in the well of the flow cell

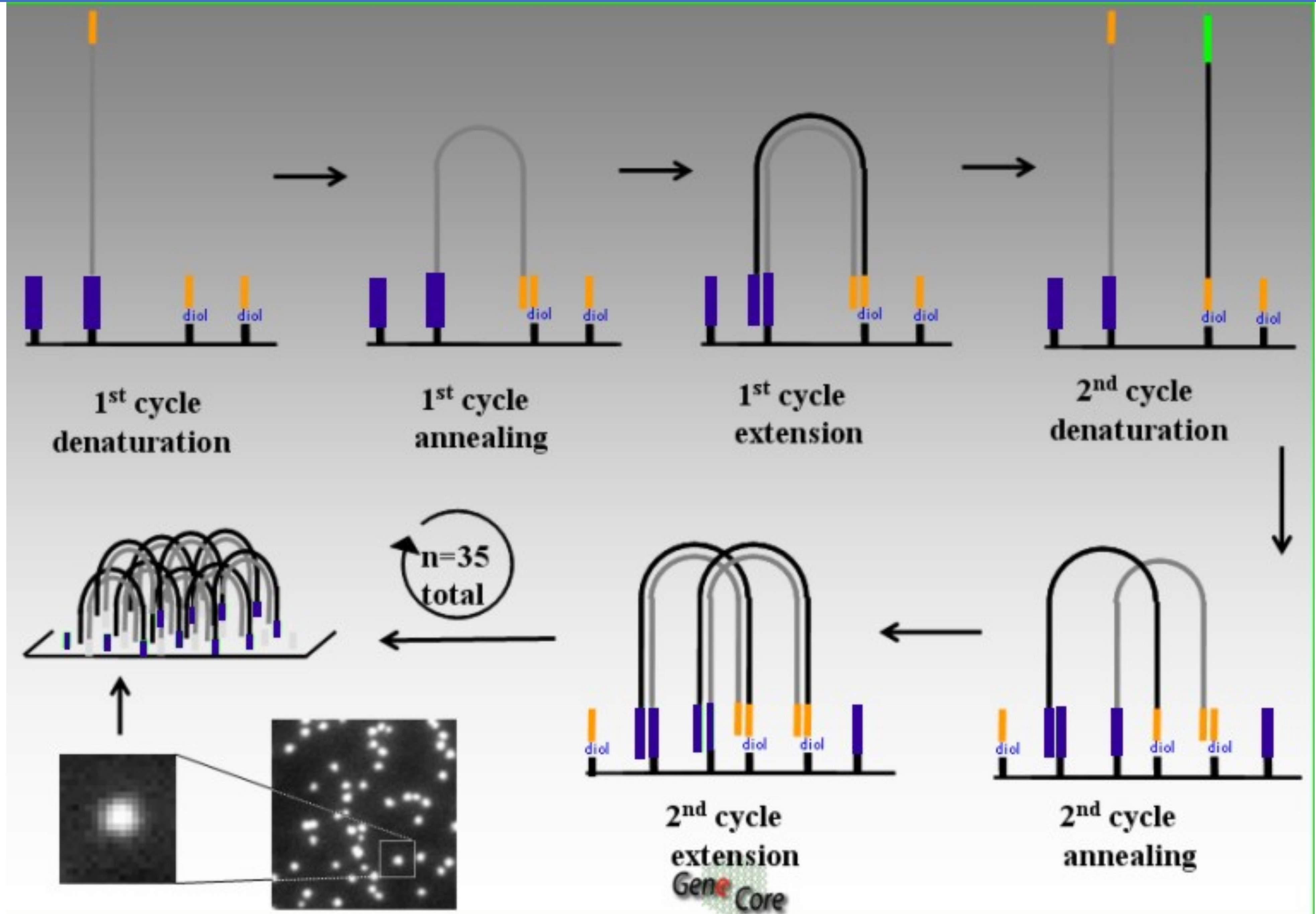
CLUSTER AMPLIFICATION:

Bridge amplification: takes place on surface of beads (each bead is mounted with 2 species of oligos; each oligo can hybridize to a DNA library fragment): initiation



On the surface: complementary oligos

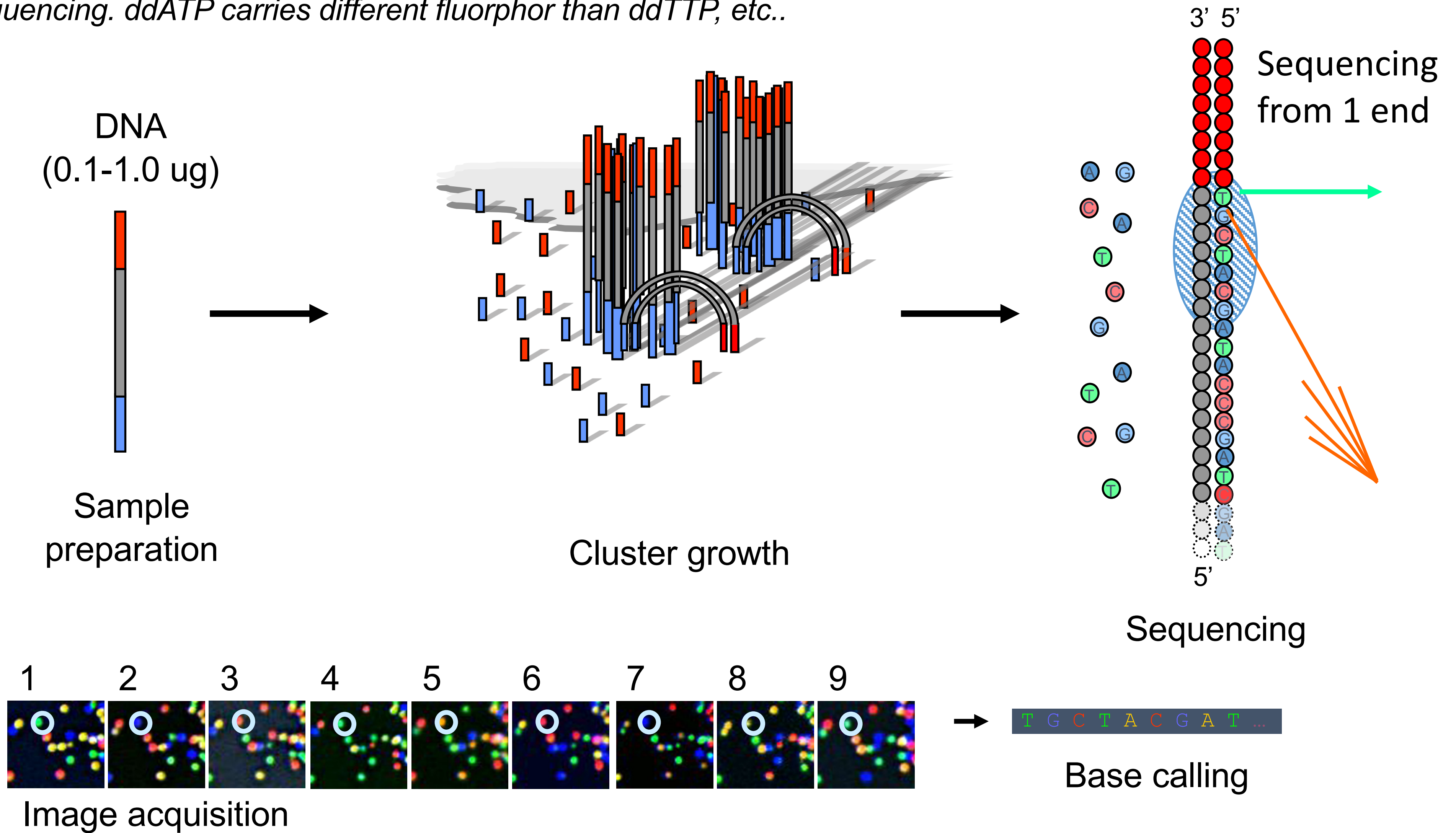
CLUSTER AMPLIFICATION:



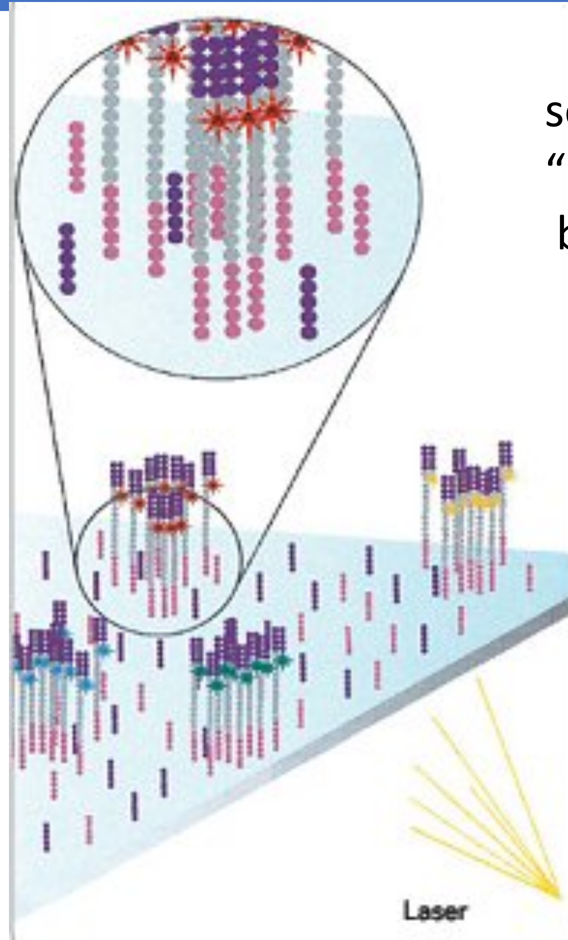
Illumina Sequencing Technology

Robust Reversible Terminator Chemistry Foundation

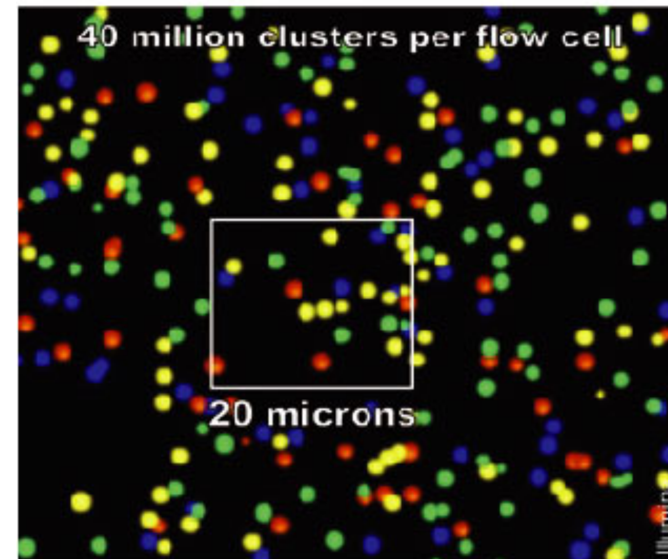
In each round of sequencing a fluorescently labelled ddNTP will be used for sequencing. ddATP carries different fluorphor than ddTTP, etc..



Illumina: massive parallel sequencing:



sequencing by synthesis:
“**reverible terminator**” nucleotides
blocked + fluorescently labeled



1. Synthesis = incorporation of fluorescent nucleotide: blocking synthesis
2. 4. Scanning of fluorescent signal
3. dye cleavage + elimination
4. wash step
1. Synthesis = incorporation of fluorescent nucleotide: blocking synthesis
READ LENGTH: ca: 150nt from each primer (2x150nt = 300nt)

illumina.com

Illumina: paired end sequencing increases information content

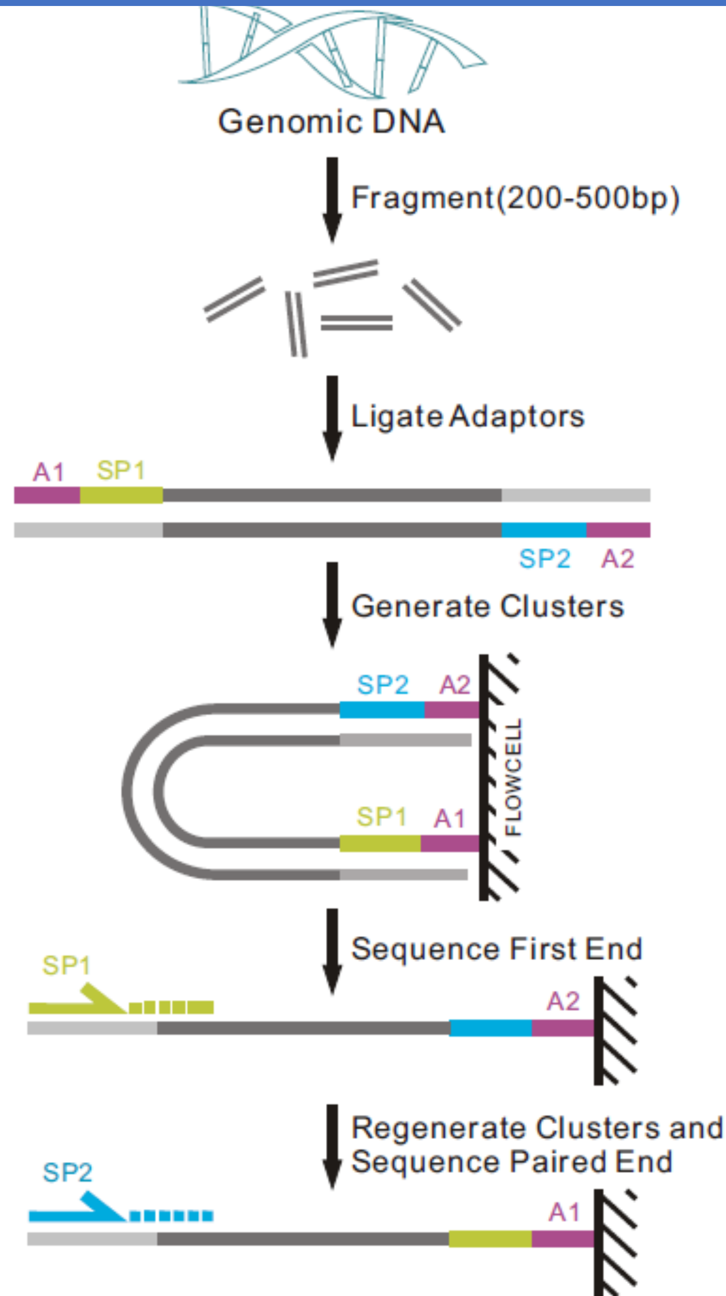


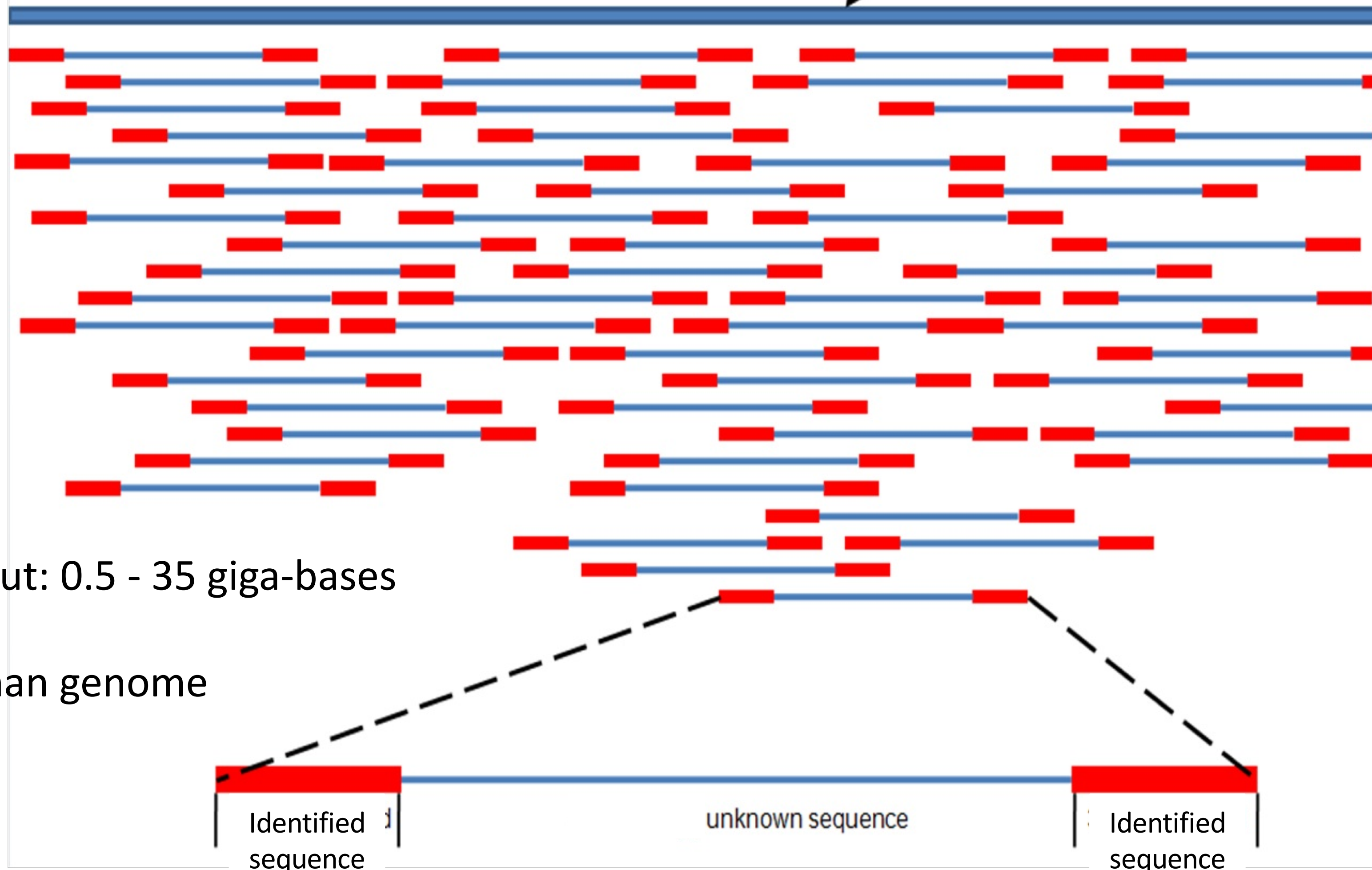
Figure 1-2-1 Pipeline of paired-end sequencing (www.illumina.com)

https://www.youtube.com/watch?v=9YxExTSwgPM

After 1° strand sequencing, A1 anneals to A1 in nanowell-->DNA synthesis --> template strand cleaved off → new strand sequenced

Data analysis: obtained sequence reads are aligned along genomic DNA sequence → high number of reads necessary to obtain full sequence coverage

Read length: 50 – max. 300 nt
Read does not necessarily cover entire library DNA fragment



Max. output: 0.5 - 35 giga-bases
$=3.5*10^{10}$
= 10x human genome

Sequence derived from one amplified cluster