

## **Workshop on Next Generation Sequence (NGS) Data Analysis and it's Applications in Livestock Research**

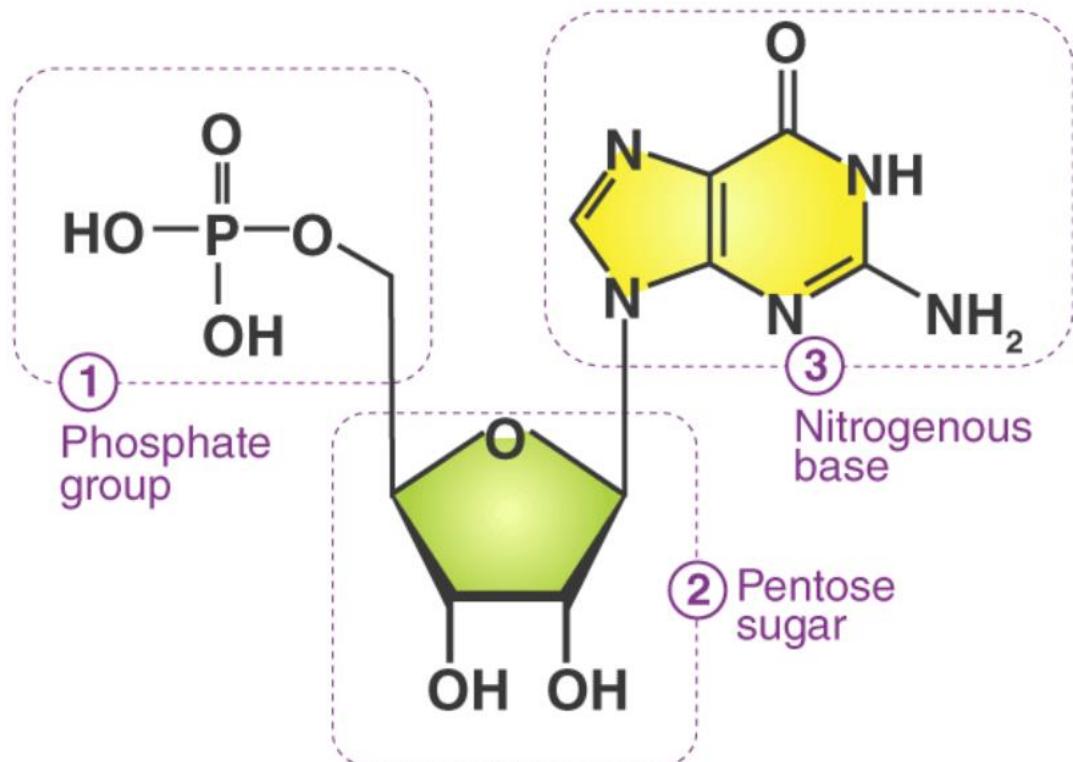
**Advances in sequencing technology and its application in livestock research**

**20<sup>th</sup> May, 2024**

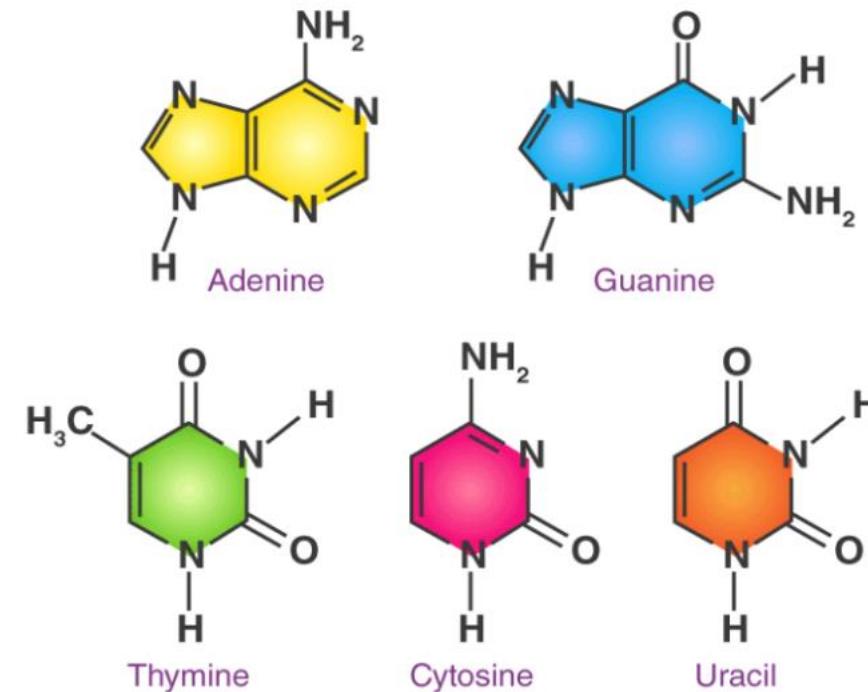
Dr. Yash Pal  
Scientist – C  
NIAB

# The Prologue

**Nucleotide:** The basic building block of nucleic acids (RNA and DNA)



## NITROGENOUS BASES



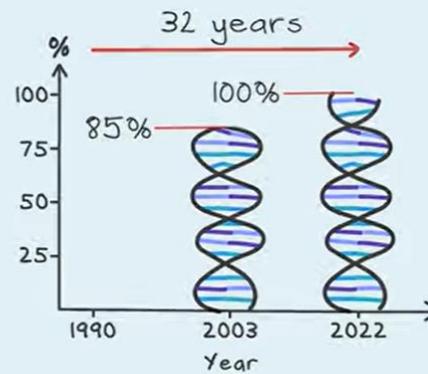
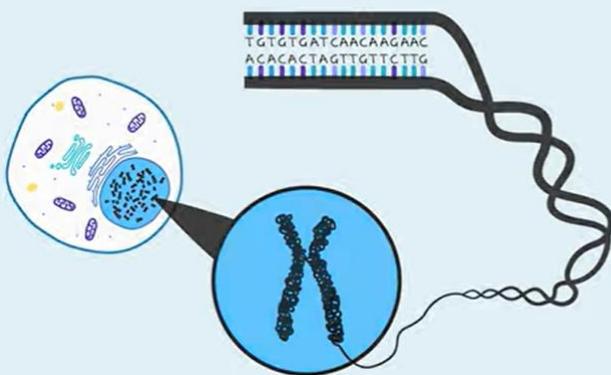
ATGC

# The Prologue

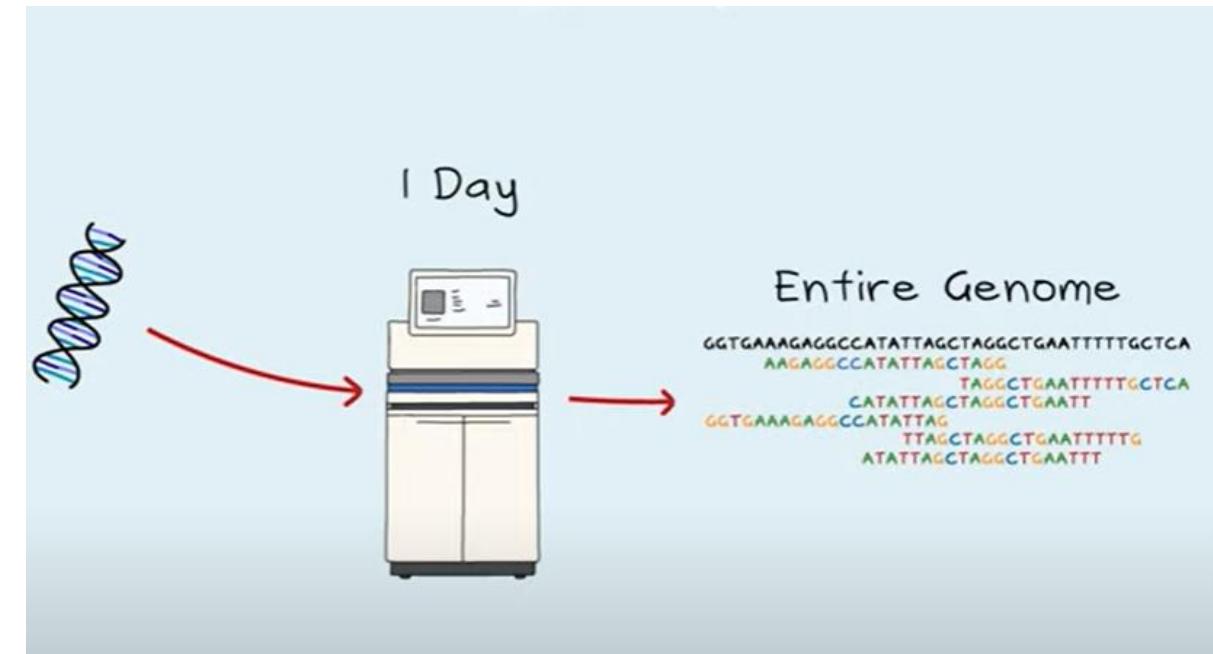
## Next Generation Sequencing (NGS)

### The Human Genome Project

3.2 Billion Bases

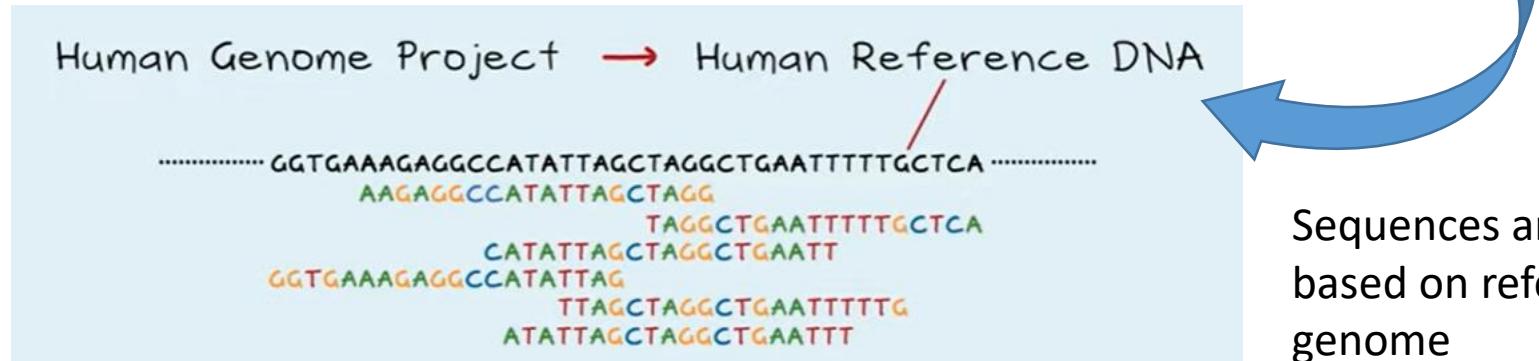
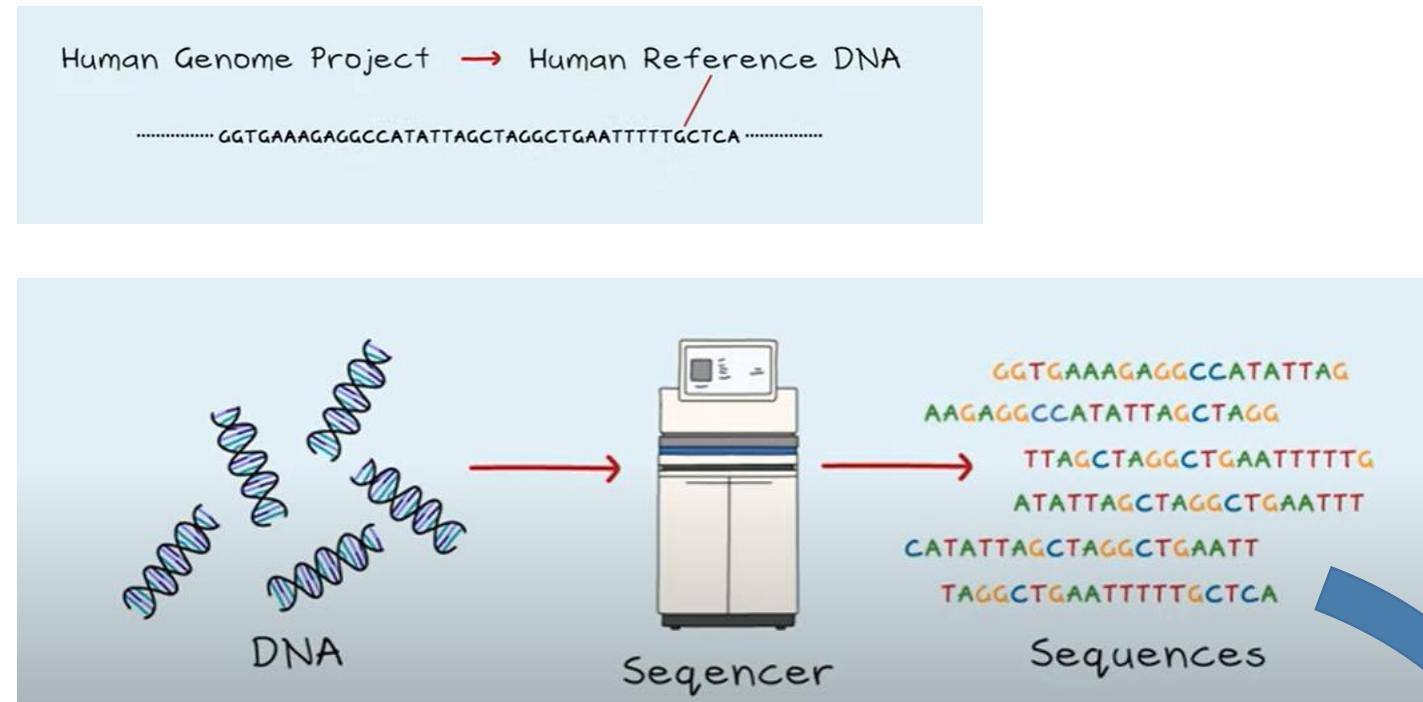


1 Day



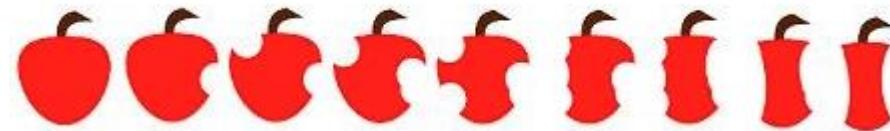
# The Prologue

The HGP gave us a  
Human reference DNA



# Introduction

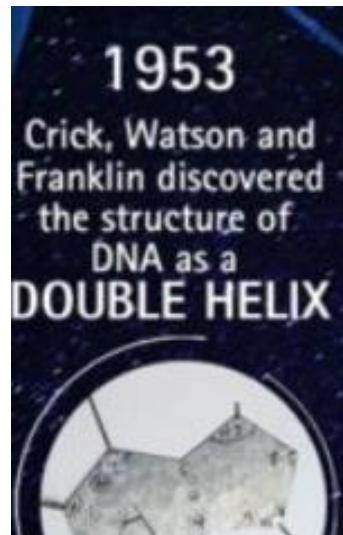
**Sequencing:** The process of determining the precise order of elements within a given sequence



1869



1953

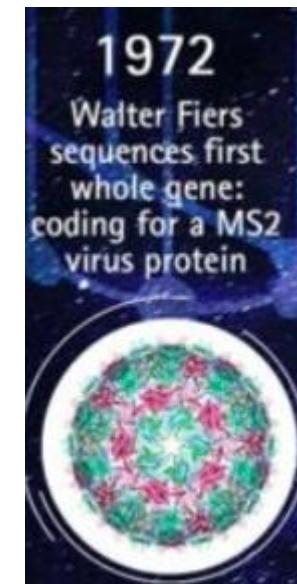


1965



Robert Holley sequenced the first tRNA, for which he was awarded the Nobel Prize in 1968 (Yeast as model org.)

1972

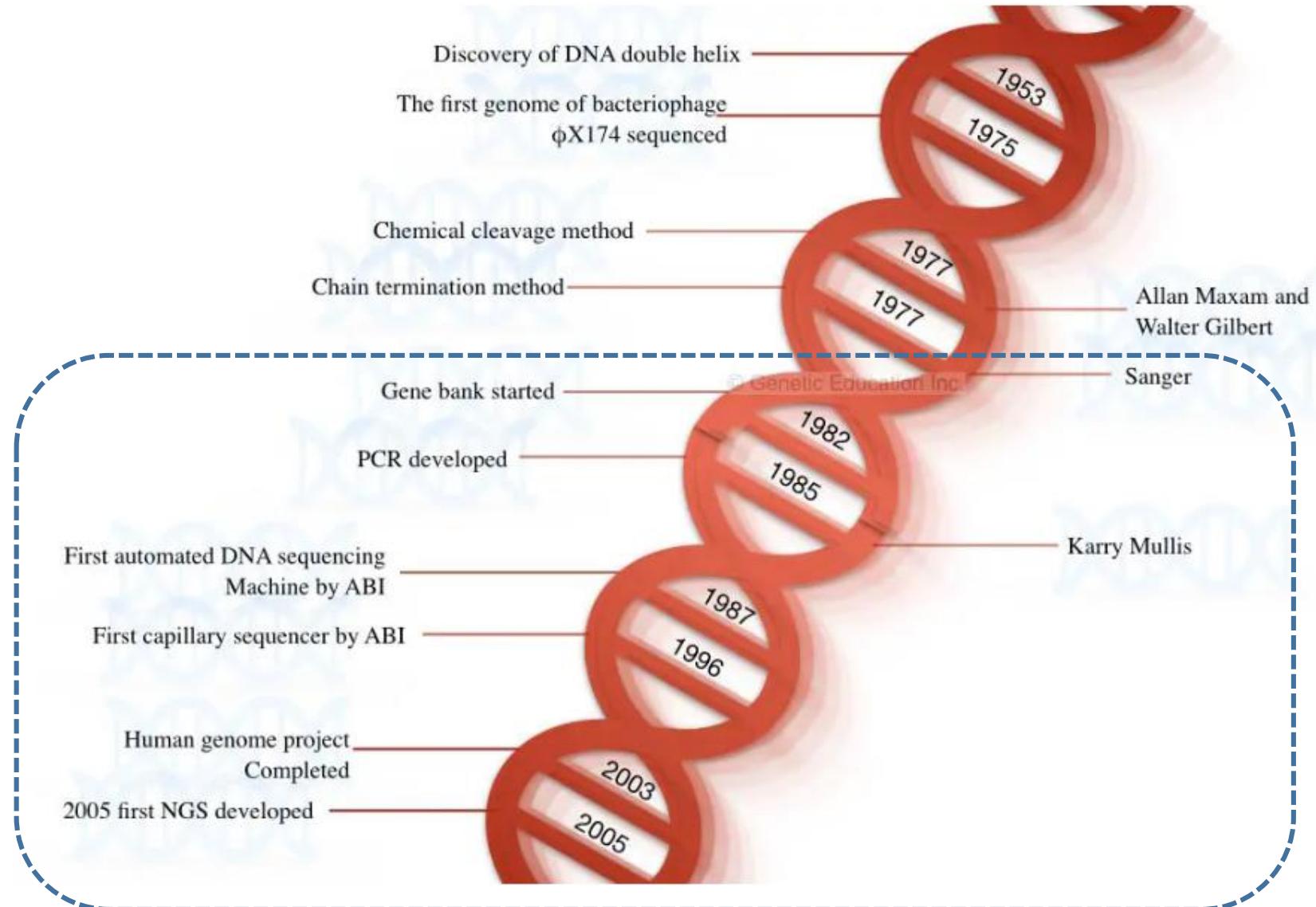


1977



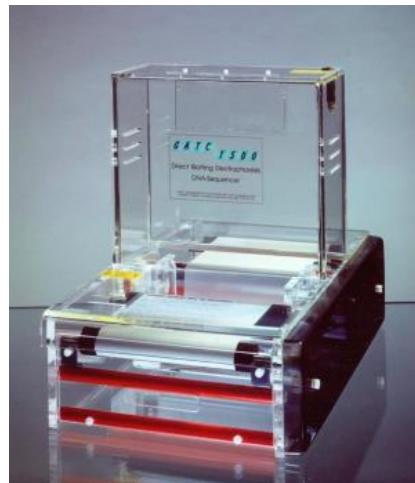
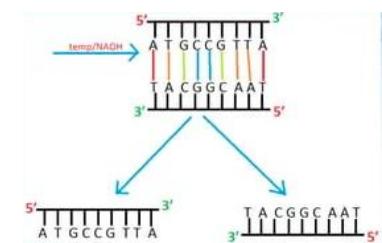
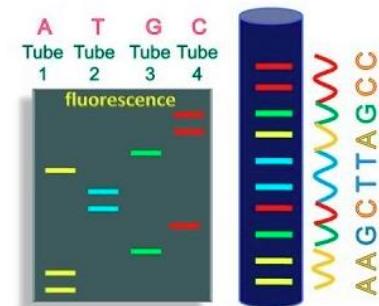
Fredrick Sanger Establishes the chain termination reaction method and sequenced first complete genome of bacteriophage PhiX174

# Introduction



## The breakthrough in DNA sequencing: The first generation

- In 1977, **Fredrick Sanger developed the chain termination method**, a DNA sequencing method utilizing radiolabeled partially digested fragments.
- **Maxam and Gilbert** also introduced a DNA sequencing method in 1977, based on chemical modification of DNA. Unlike Sanger sequencing, this method did not rely on DNA polymerase.
- Both Sanger sequencing and Maxam–Gilbert sequencing lacked automation and were time consuming and tiring

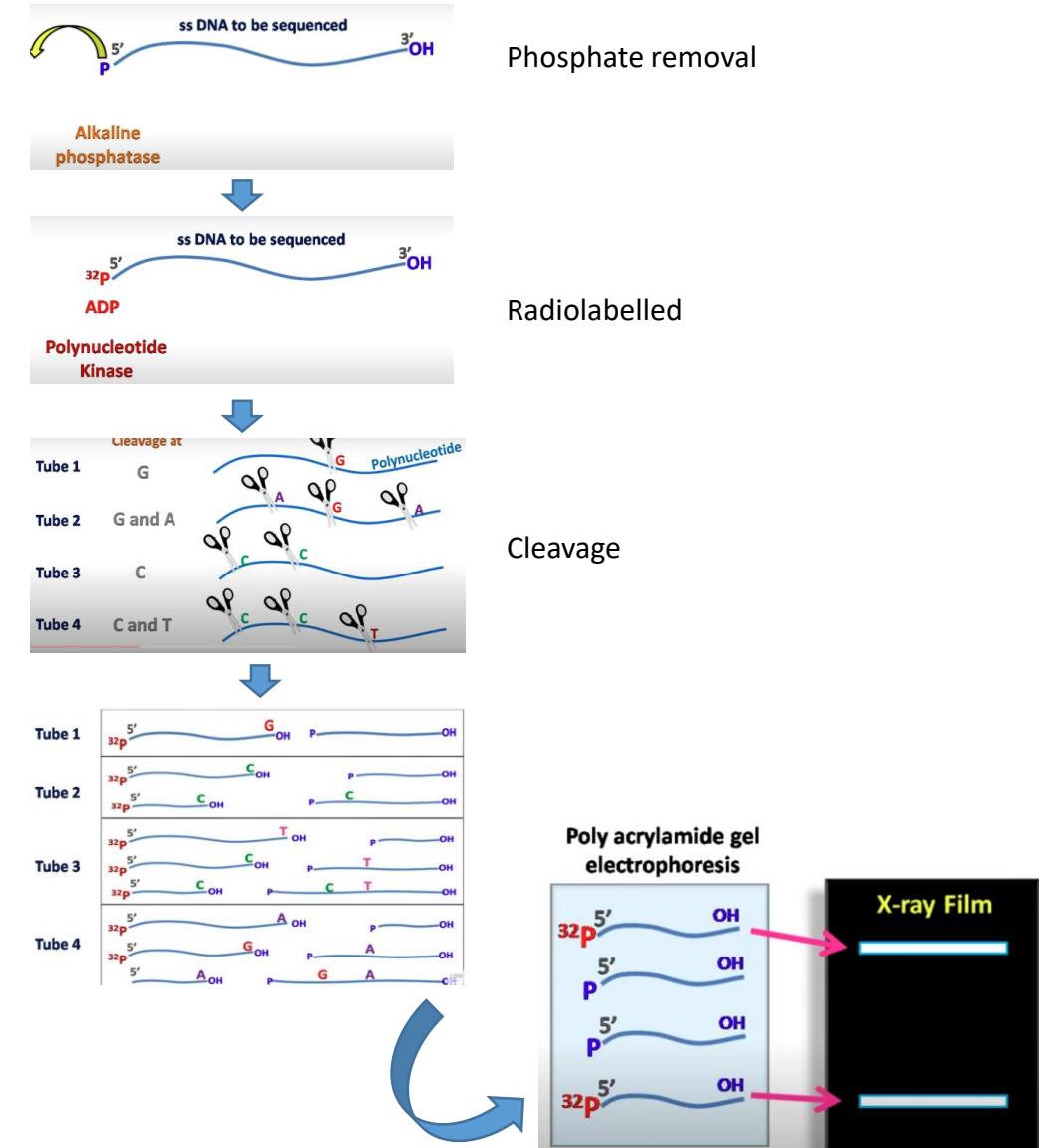
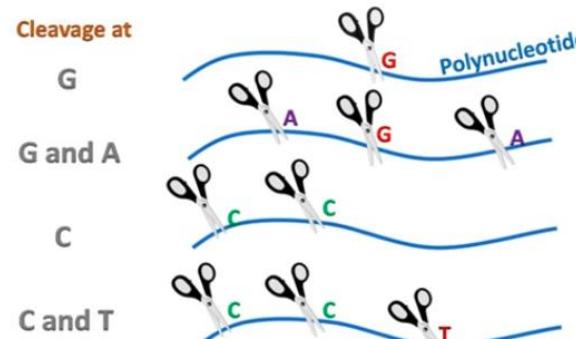


In 1984, Fritz Pohl established the first sequencing technology platform that did not rely on radioactive labelling: the GATC1500.

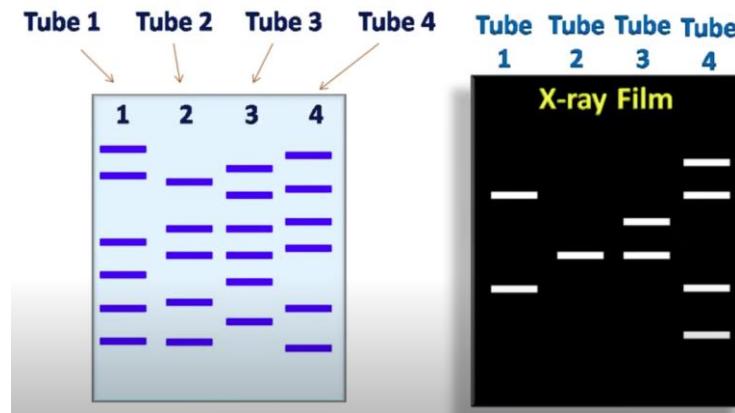
# Maxam-Gilbert method

- Chemical synthesis method

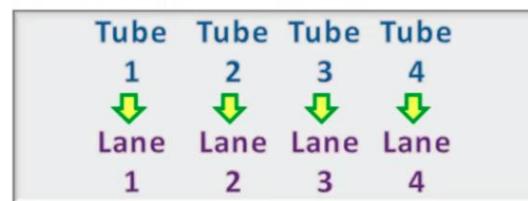
- Hydrazine** cleaves DNA at guanine (G) and adenine (A) residues.
- Dimethyl sulfate (DMS)** cleaves DNA at cytosine (C) and guanine (G) residues.
- Formic acid** cleaves DNA at purine residues (A and G).
- Piperidine** cleaves DNA at purine residues (A and G).



# Maxam-Gilbert method

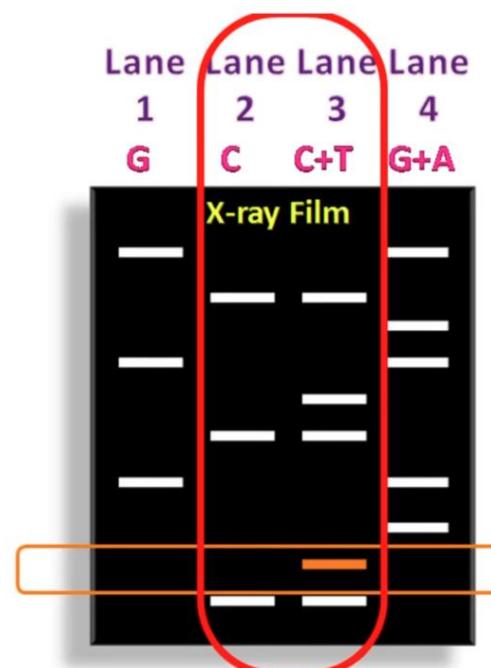


## Maxam Gilbert method

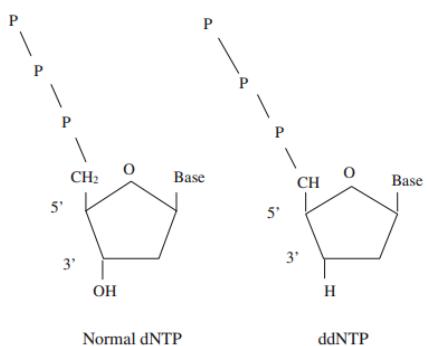
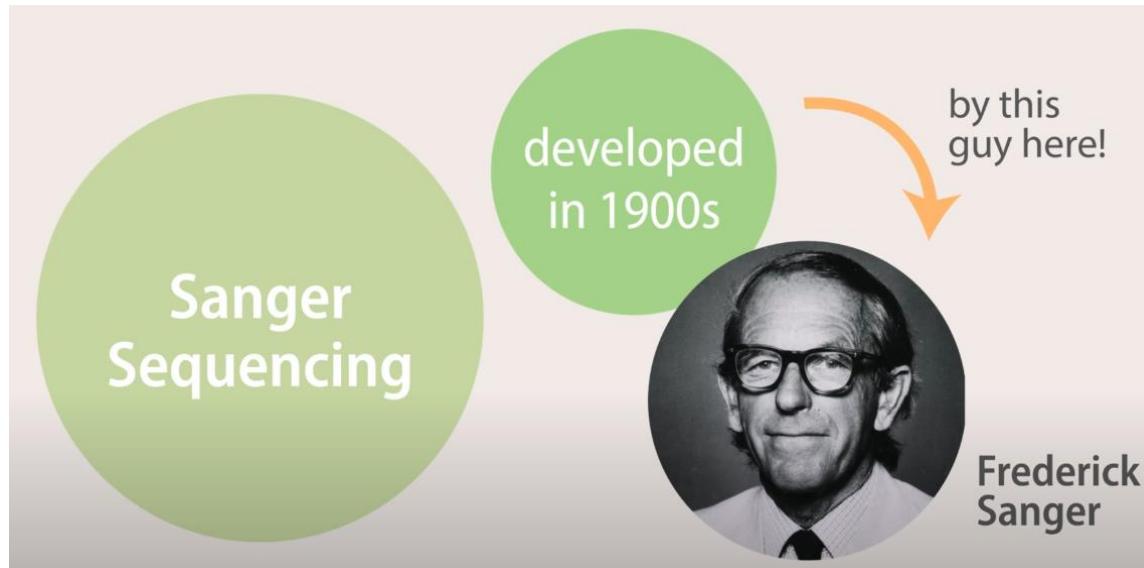


Reading the sequence

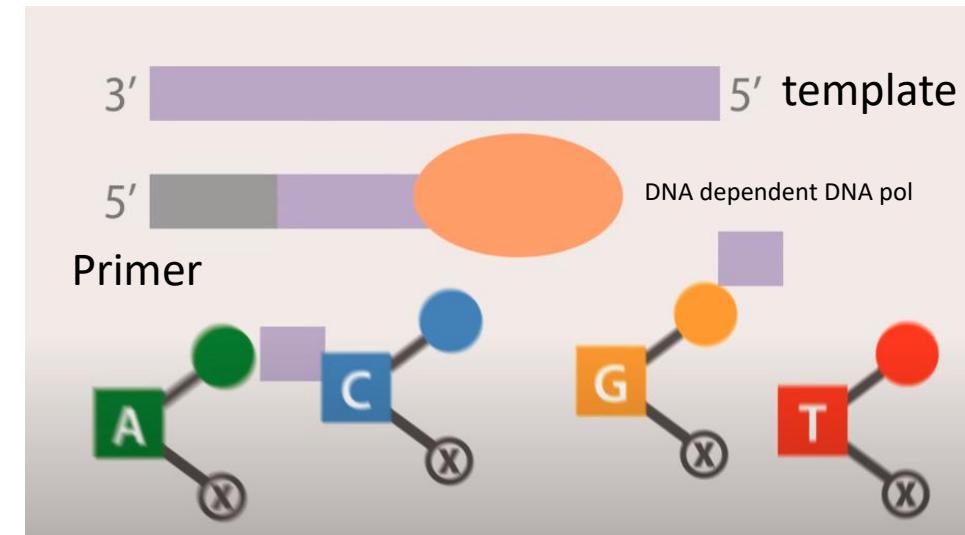
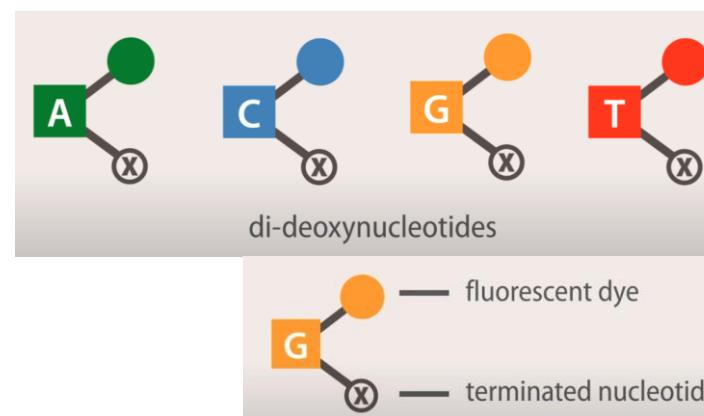
5'  
<sup>32</sup>P C  
1 2 3 4 5 6 7 8 9 10  
Nucleotide



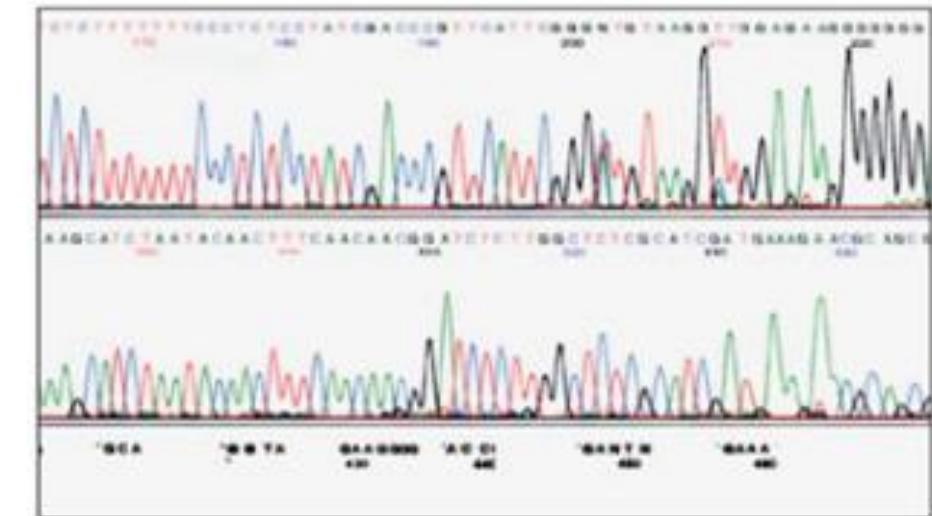
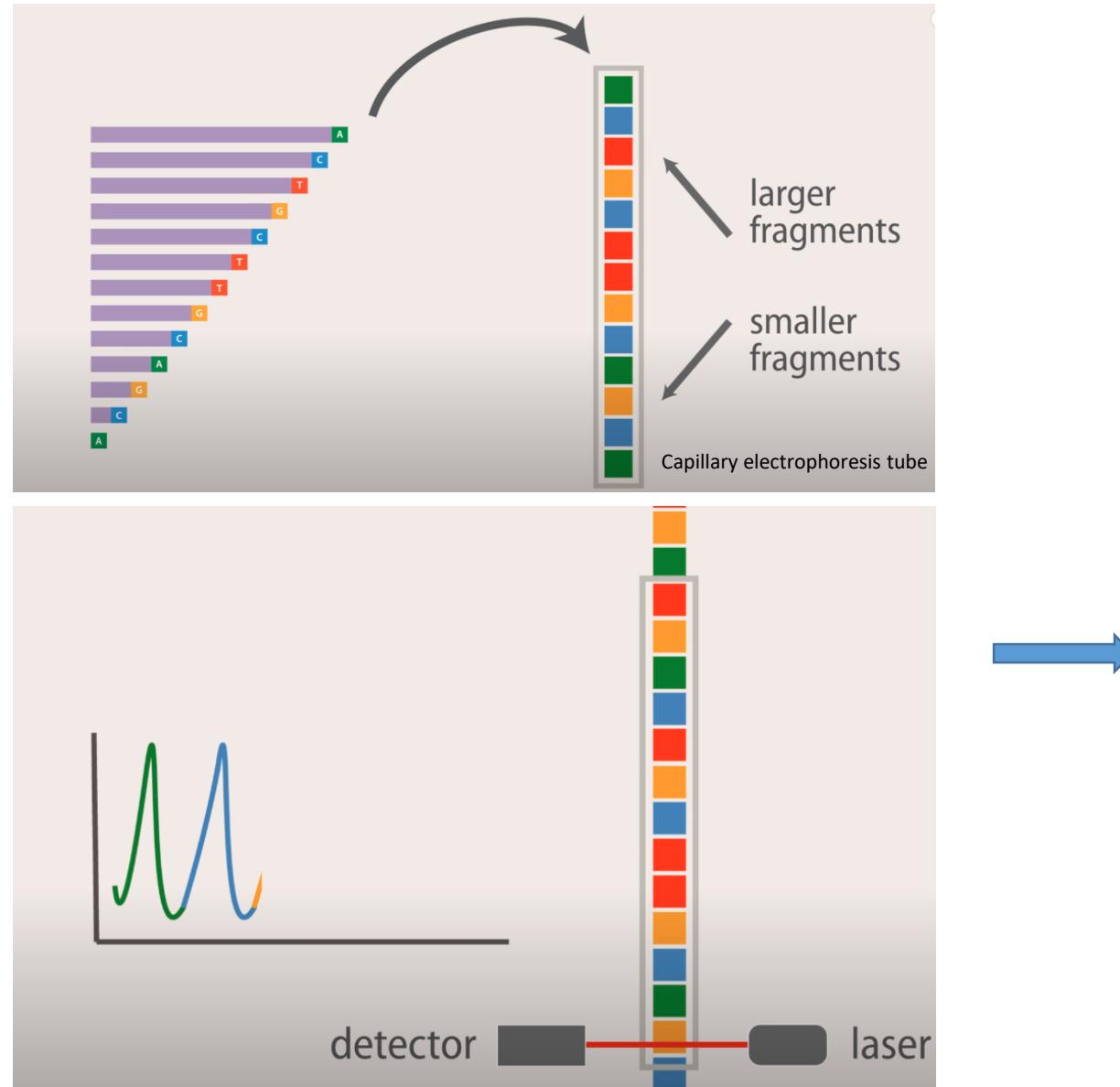
# Sanger Sequencing (Chain termination)



The hydroxyl group  
is missing from both  
the 2' and the 3'  
carbons



# Sanger Sequencing (Chain termination)



# Traditional methods of Sequencing and its limitations

- **Maxam-Gilbert Method**
  - Use of radioactive labels.
- **Sanger Method**
  - It utilize the fluorescent dye for labeling.
  - separation of extended fragments of DNA with the addition of di-deoxynucleotides (lack a 3'-OH group)  
Thus, chain termination.
- ❖ **Limitation**
  - ❖ Slow
  - ❖ High cost per run.

# Some useful terms

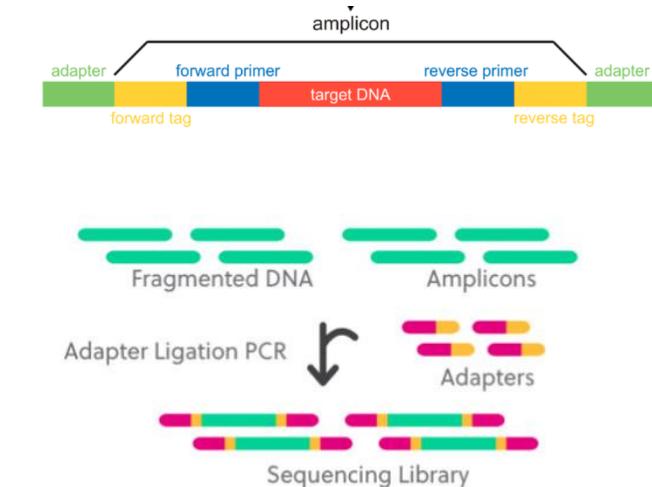
**Reads:** The output of an NGS sequencing reaction. A read is **a single uninterrupted series of nucleotides** representing the sequence of the template.



**Read Length:** The **length of each sequencing read**. This variable is always represented as an average read length since individual reads have varying lengths.

**Adapter:** Unique **sequences used to cap the ends of a fragmented DNA**. The adapter's functions are as follows:

- 1) allow hybridization to solid surface
- 2) provide priming location for both amplification and sequencing primers
- 3) provide barcoding for multiplexing different samples in the same run.



**Library:** A **collection of DNA fragments with adapters ligated to each end**. Library preparation is required before a sequencing run.

**Coverage:** Number of times each nucleotide is sequenced to ensure accuracy, typically requiring 30x coverage.

**Paired-End Sequencing:** Sequencing from **both ends** of a fragment while keeping track of the paired data.

**Mate-Paired reads:** A sample preparation step where large DNA fragments (~10kb) are circularized with an adapter sequence followed by degradation of the circular DNA

**Multiplexing (pooling):** Is a technique **used to simultaneously analyze multiple samples** in a single Next generation sequencing run. It involves the use of unique index (barcode) sequences that are added to each sample before sequencing.

## Start of the Next Generation Sequencing (2<sup>nd</sup> Generation)

- In 1996, **Mostafa Ronaghi**, **Mathias Uhlen** and **Pål Nyren** introduced a new DNA sequencing technique called **pyrosequencing**, and that is considered as the emergence of the second generation of DNA sequencing.
- It's **based on the measurement of luminescence generated as a result of pyrophosphate synthesis during sequencing** (sequencing-by-synthesis technology)

Jonathan Rothberg and colleagues implemented the pyrosequencing technology in an automated system (2005): The 454 system that was the first next generation sequencing platform to come to market.

NGS is also known as:

- Massively parallel sequencing
- Ultra deep sequencing



Roche 454 Sequencing System

Other notable platforms that are based on different technologies are

- **SOLiD system's "sequencing-by-ligation"** in 2007,
- **Ion Torrent** by Life Technologies in 2011 that uses "**sequencing-by-synthesis**" technology that detects hydrogen ions when new DNA is synthesised.



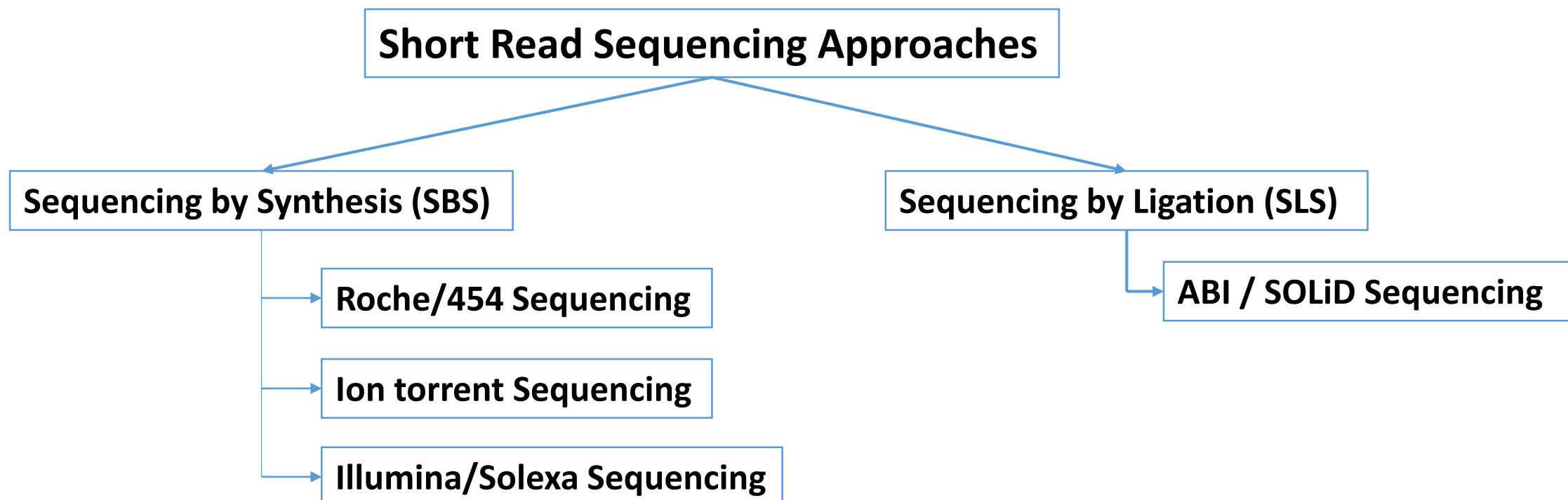
SOLiD system's



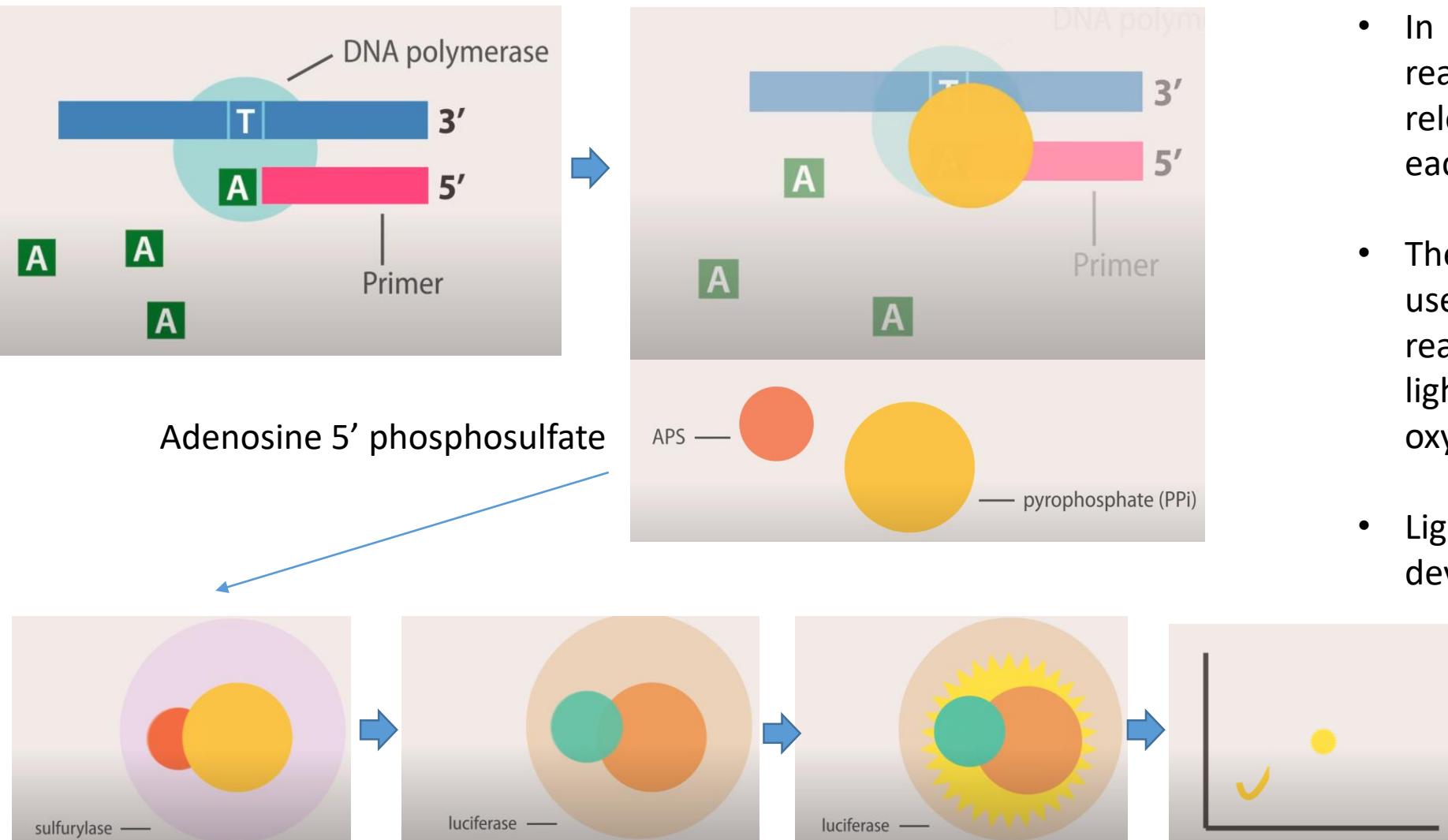
Ion Torrent

### Basic characteristics of second-generation sequencing technology:

- (1) Generation of many millions of short reads in parallel
- (2) Speed up sequencing process compared to the first generation
- (3) Low cost of sequencing
- (4) Sequencing output is directly detected without the need for electrophoresis

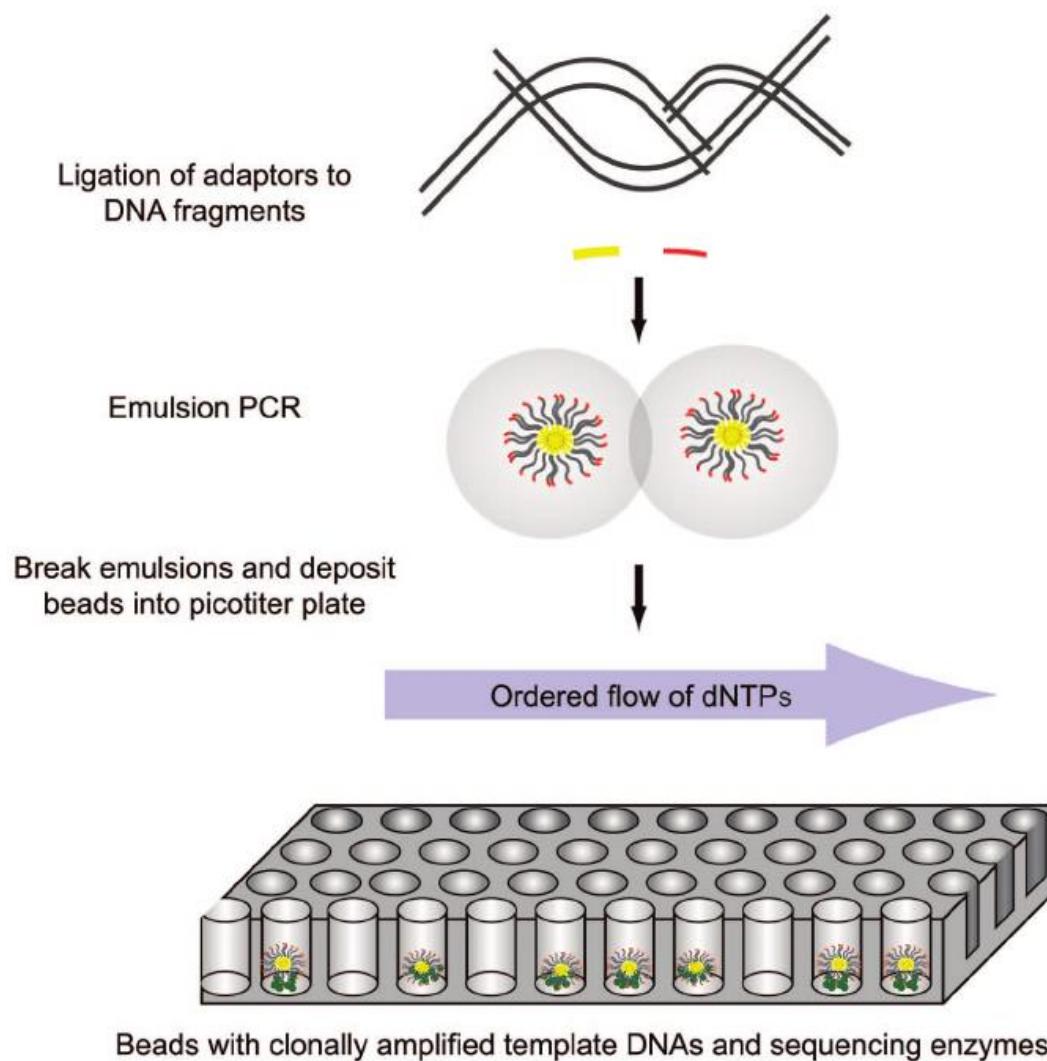


# Pyrosequencing

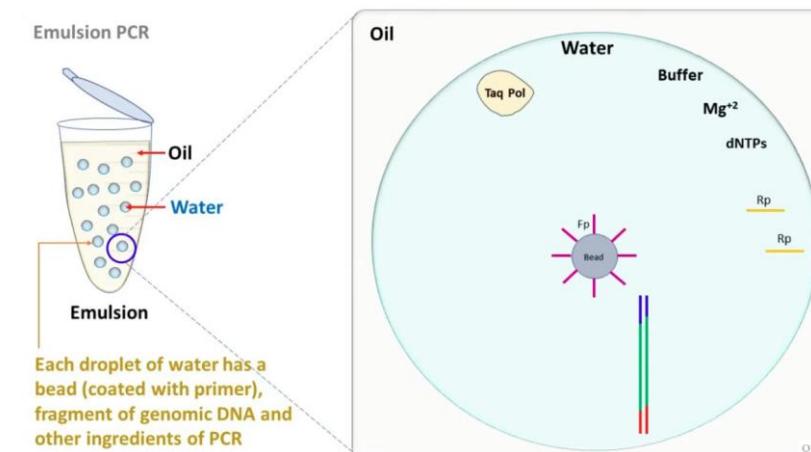


- In Pyrosequencing, sequencing reaction is monitored through the release of **pyrophosphate** during each nucleotide incorporation
- The released pyrophosphate is used in a series of chemical reactions resulting in generation of light (Luciferin is converted to oxyluciferin)
- Light detection by charge coupled device (camera)

# Roche 454 GS FLX Sequencing



- DNA fragments digested and ligated with adaptor and attached to the beads.
- Amplification using Emulsion PCR on surface of the beads.
- Beads transferred to picotiter plate and followed by pyrosequencing.



## Advantages of 454 Technology

**Longer Read Lengths:** Capable of sequencing reads between **700-1000 bp**, which aids in downstream bioinformatics by producing **longer contigs**, higher N50 length, and fewer gaps, especially beneficial for de novo sequencing projects.

**Paired-End Reads:** Facilitates the construction of better scaffolds, enhancing the overall quality and accuracy of the sequence assembly.

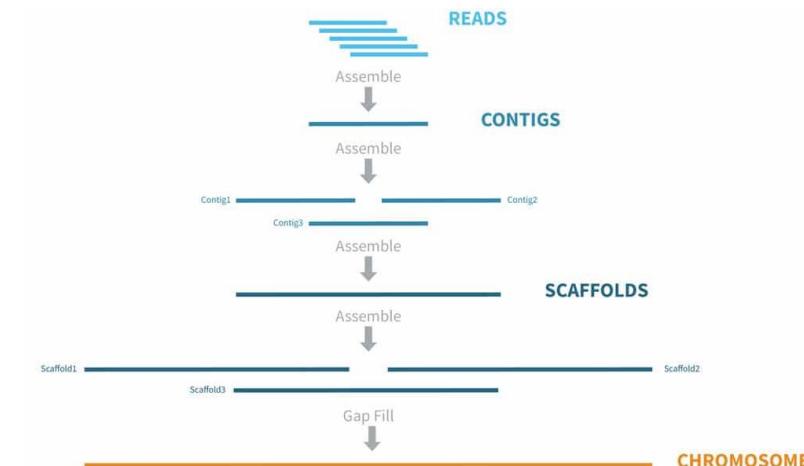
## Limitations of 454 Technology

**Homopolymer Sequencing Issues:** Difficulty in accurately sequencing homopolymer repeats longer than 6 bp due to simultaneous nucleotide incorporation.

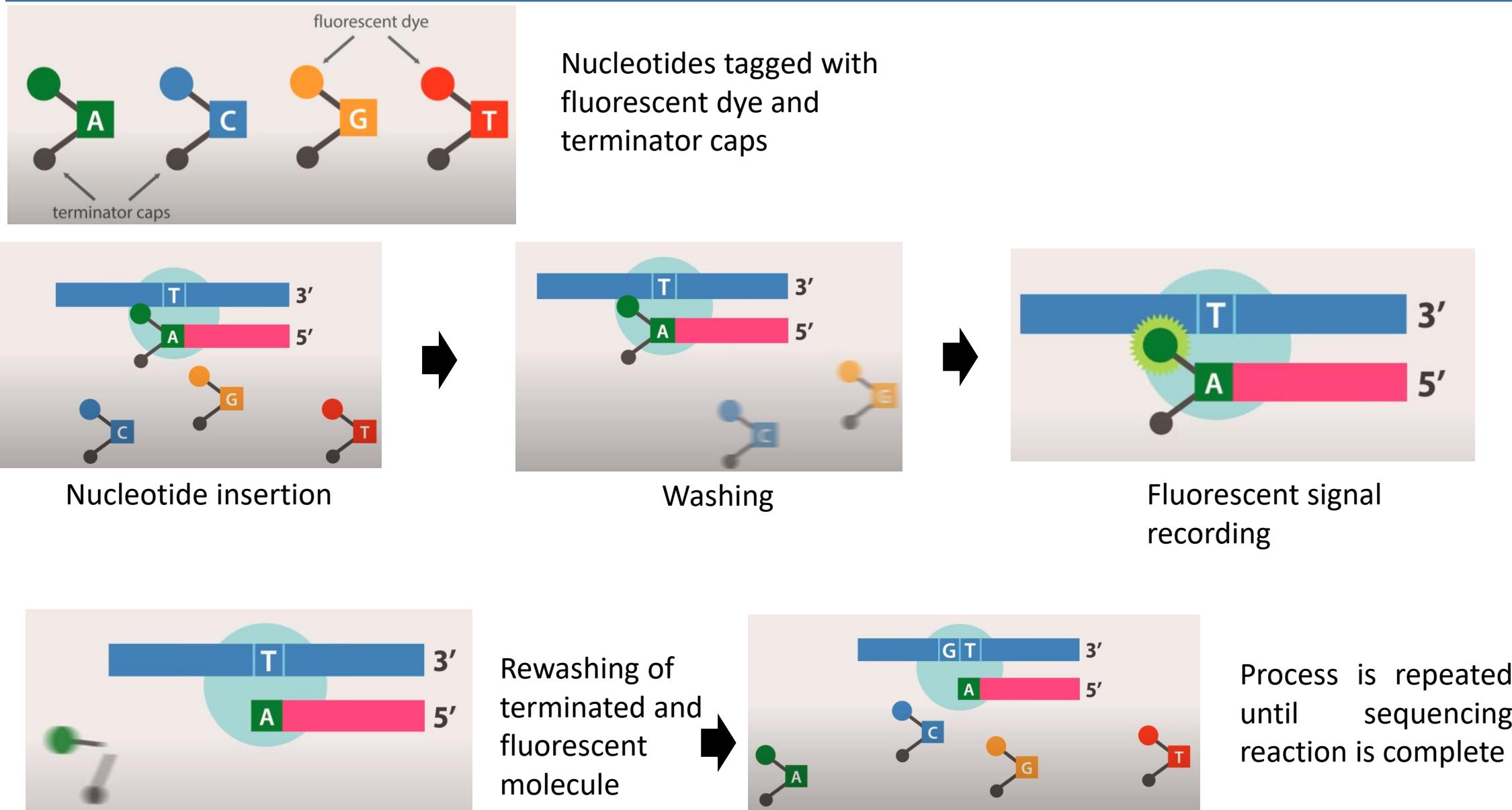
**Low Throughput and High Cost:** Generates relatively low bases per run (~700 Mb) compared to other NGS technologies, making it expensive and less suitable for high-depth re-sequencing projects.

## Pyrosequencing: Overview

- 1 Large read lengths generation
- 2 High reagent cost
- 3 High error rate over strings of 6+ homopolymers

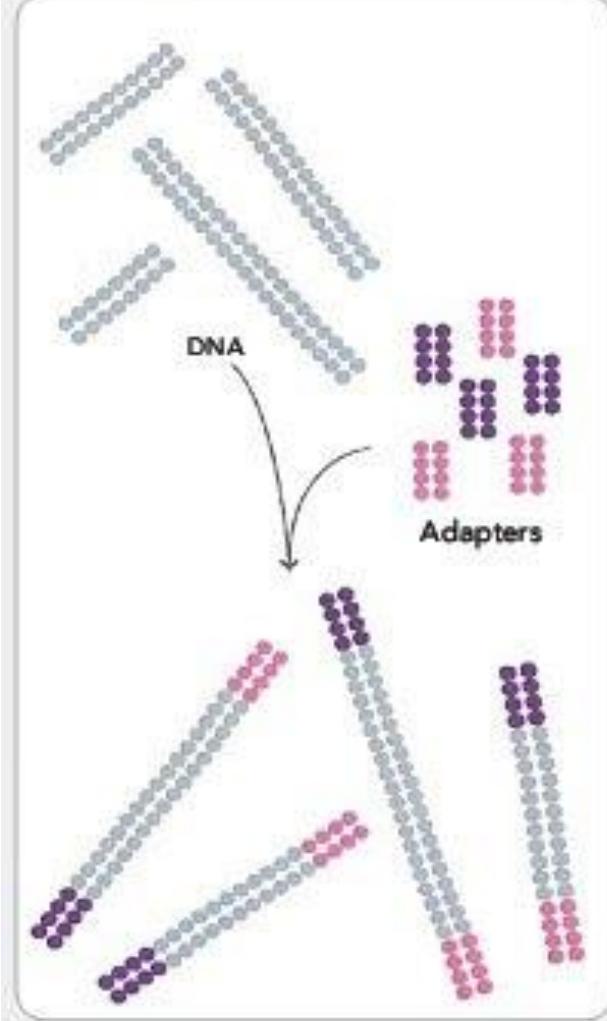


## Sequencing by synthesis (Basic overview)

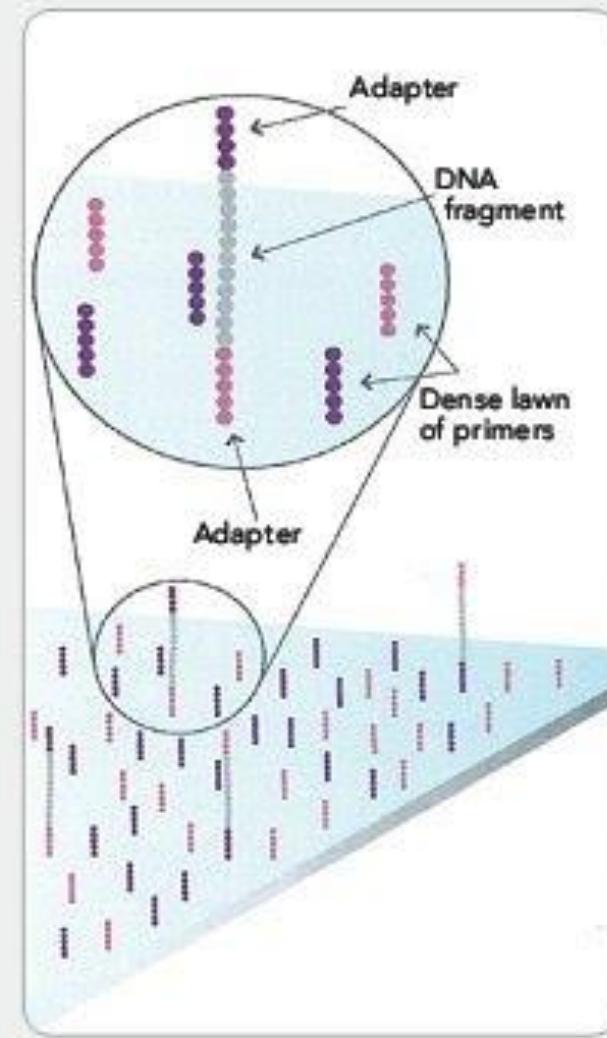


# Illumina Sequencing

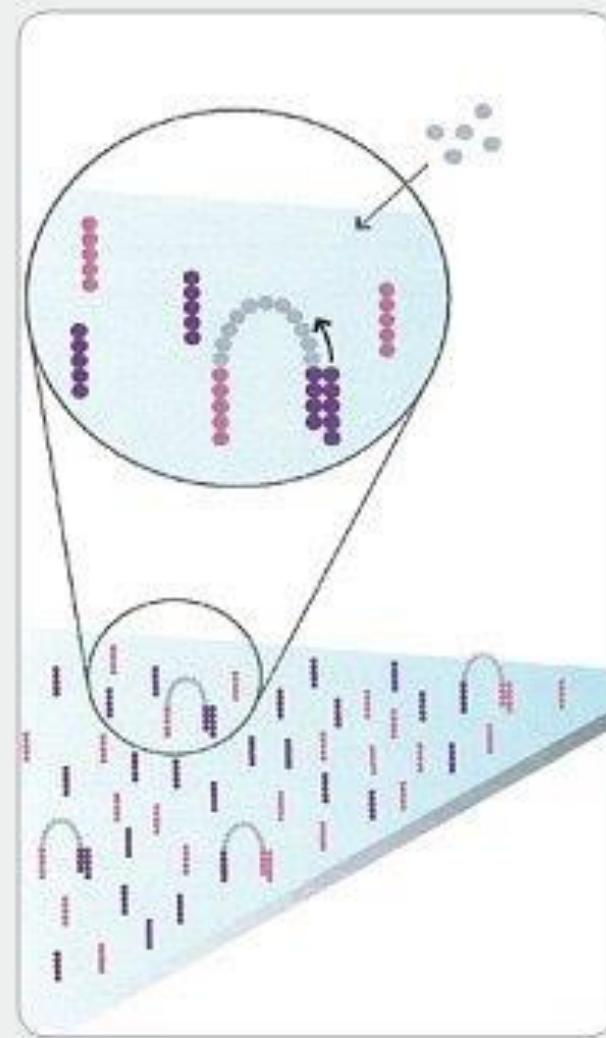
## 1. PREPARE GENOMIC DNA SAMPLE



## 2. ATTACH DNA TO SURFACE



## 3. BRIDGE AMPLIFICATION



Flow cell

- DNA is ligated to adapter
- Flow cell contains dense lawn of oligos complementary to adapter
- Unlabelled nucleotides and enzymes are added to initiate bridge amplification

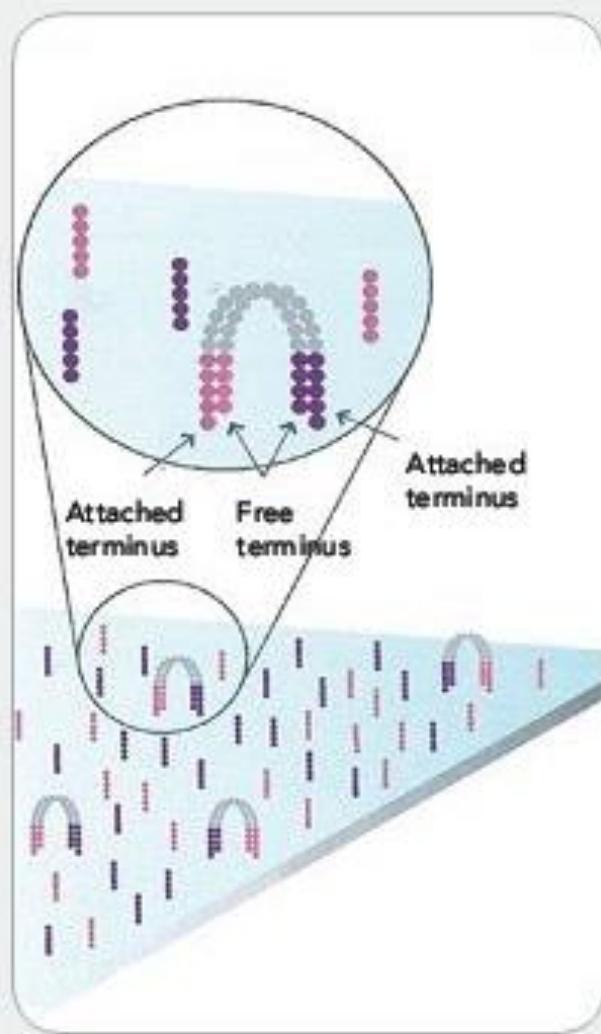
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

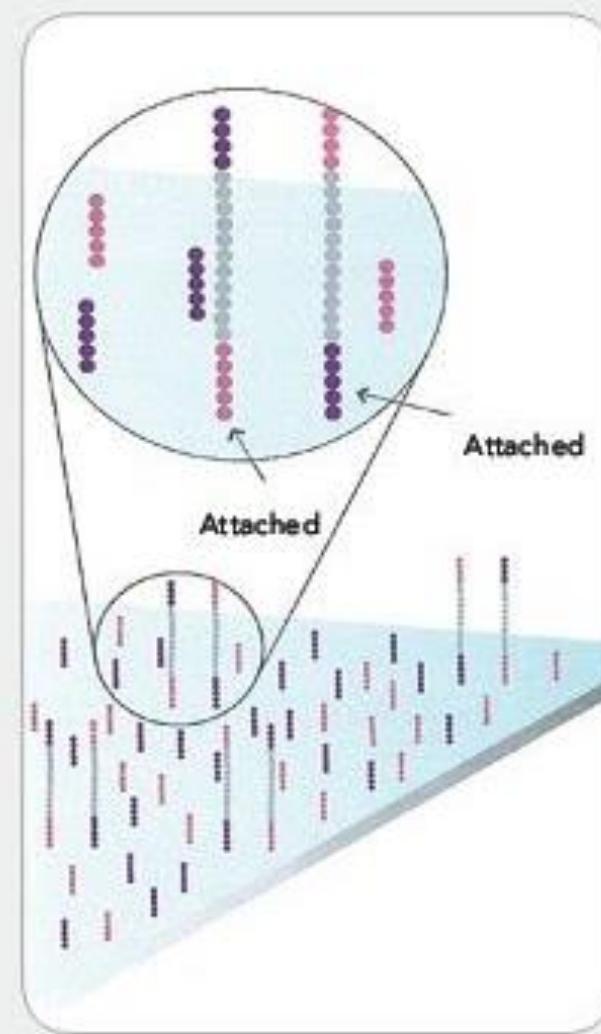
Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

# Illumina Sequencing

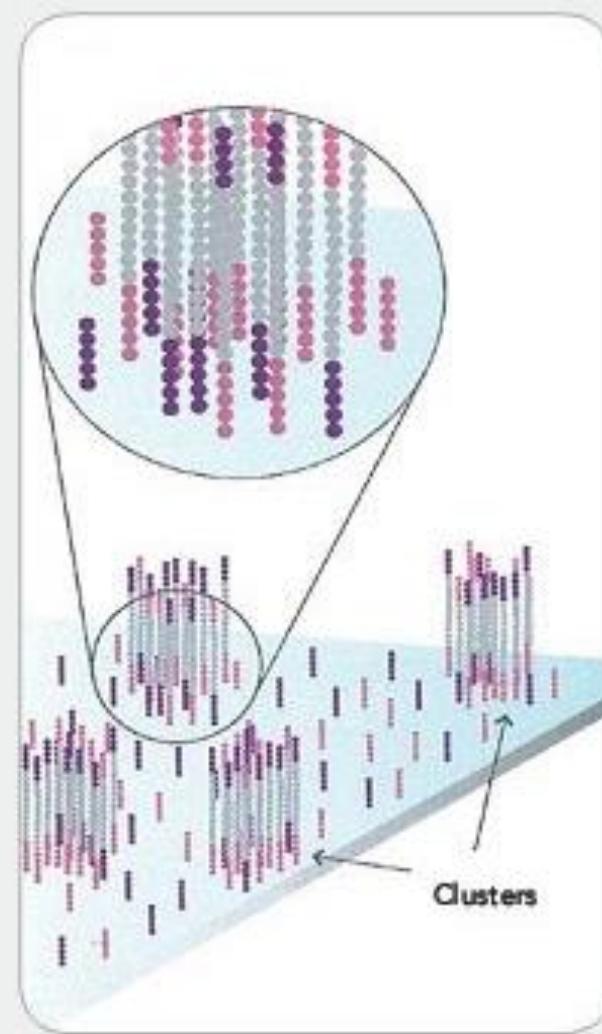
## 4. FRAGMENTS BECOME DOUBLE STRANDED



## 5. DENATURE THE DOUBLE-STRANDED MOLECULES



## 6. COMPLETE AMPLIFICATION



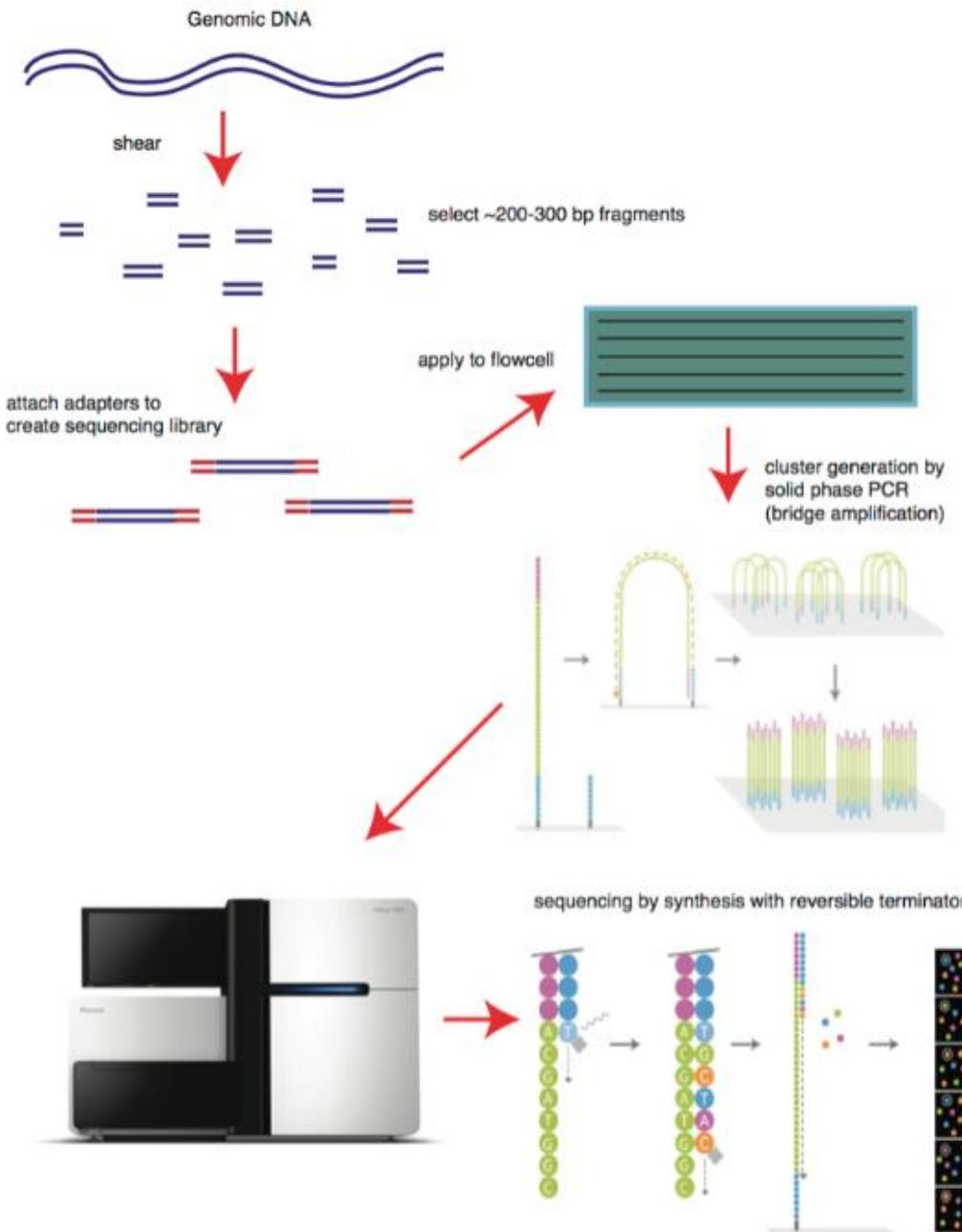
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

Denaturation leaves single-stranded templates anchored to the substrate.

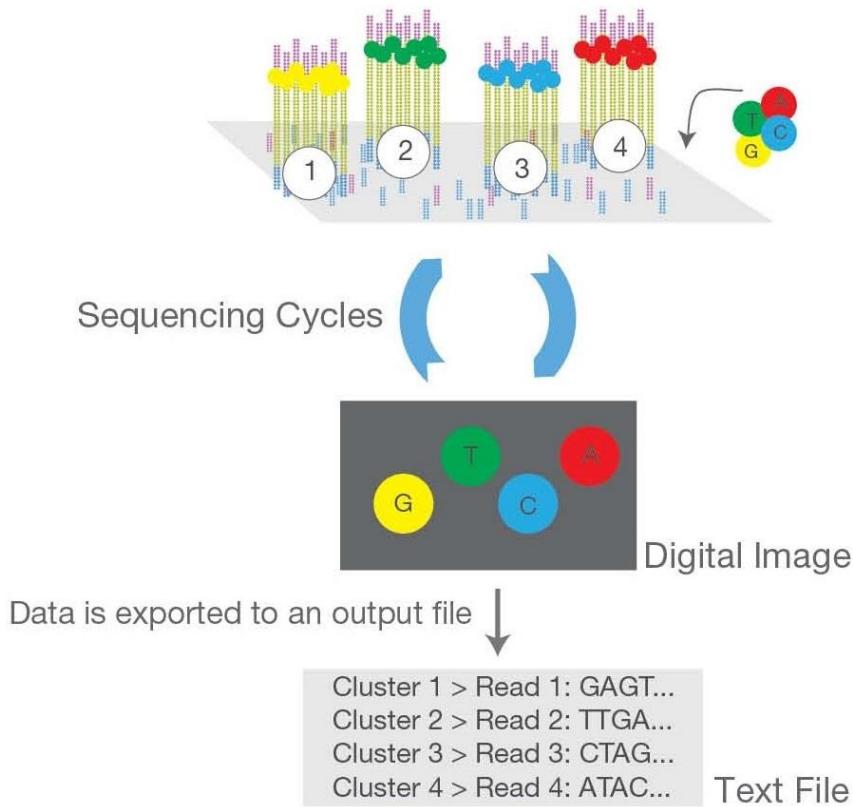
Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.

- ds bridges are created on solid phase substrate
- Single stranded template remains attached to substrate after denaturation

# Illumina Sequencing



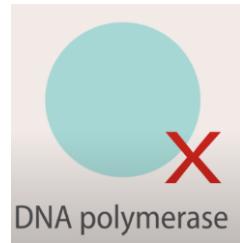
# Illumina Sequencing



Sequencing reagents, including fluorescently labeled nucleotides, are added and the first base is incorporated. The flow cell is imaged and the emission from each cluster is recorded. The emission wavelength and intensity are used to identify the base. This cycle is repeated “n” times to create a read length of “n” bases.

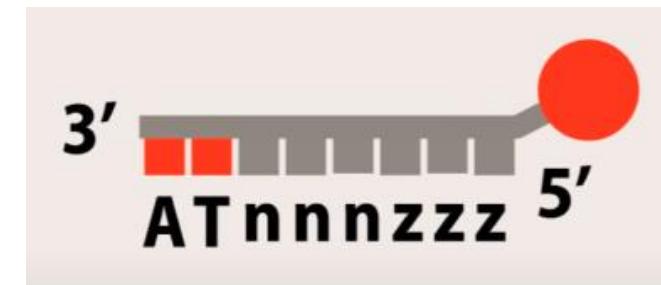
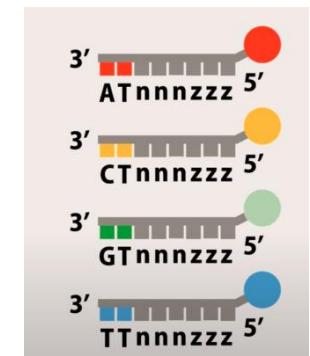
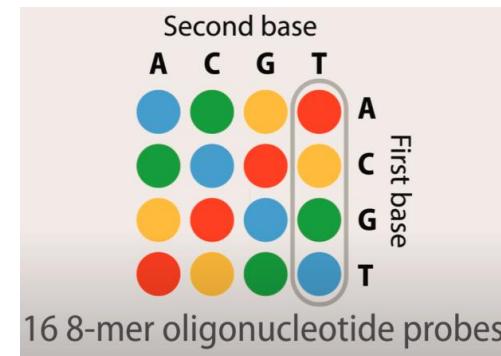
- DNA polymerase, connector primers and 4 dNTP with base-specific fluorescent markers are added.
- The 3'-OH of these dNTP are protected by chemical modification(Azidomethyl), so one base will be added at a time during sequencing.
- All unused free dNTP and DNA polymerase are washed after the synthesis reaction finished.
- Fluorescent signal is detected, and chemical reagent (TCEP) is added to quench the fluorescence signal and remove the dNTP 3'-OH protective group, so that the next round of sequencing reaction can be performed.

## Sequencing by Ligation (Ex: ABI SOLiD)



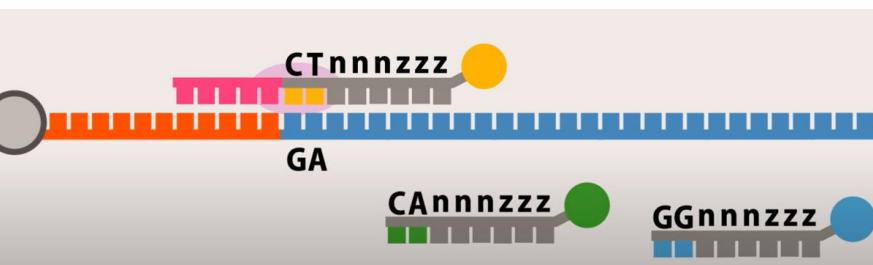
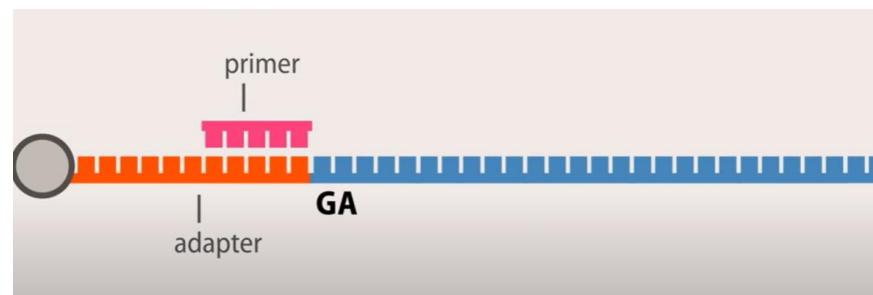
No DNA polymerase involvement

Involve 16 8-mer oligo probes



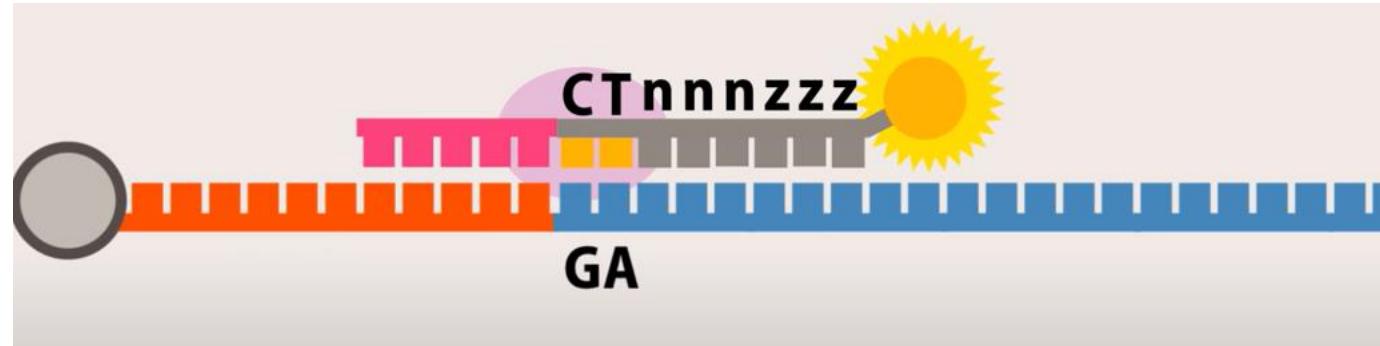
2 specific bases

6 degenerate bases

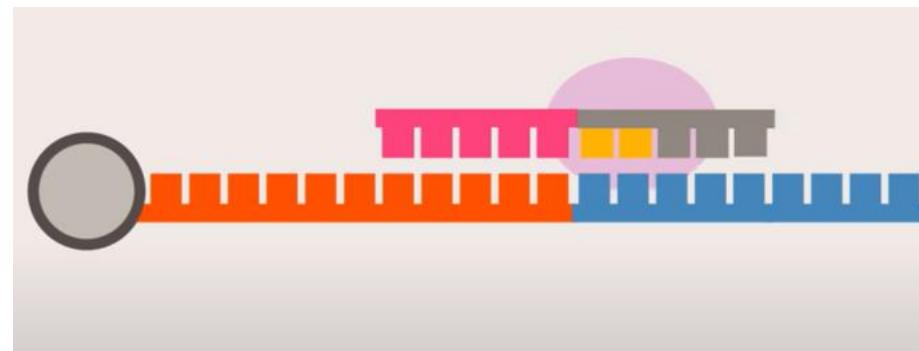


- The sequencing reaction commence by binding of the primer to the adapter sequence
- Followed by hybridization of appropriate probe
- Hybridization is guided by 2 specific bases
- And upon annealing is ligated to the primer sequence through a DNA ligase
- Unbound oligo's are washed away

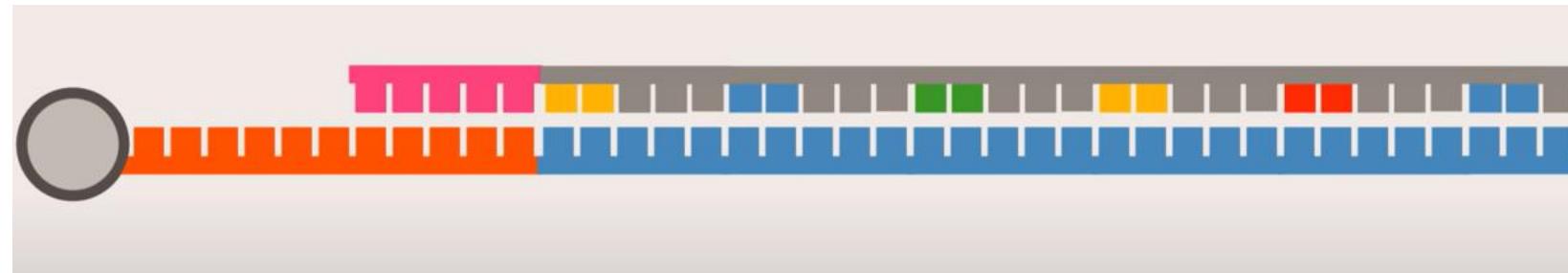
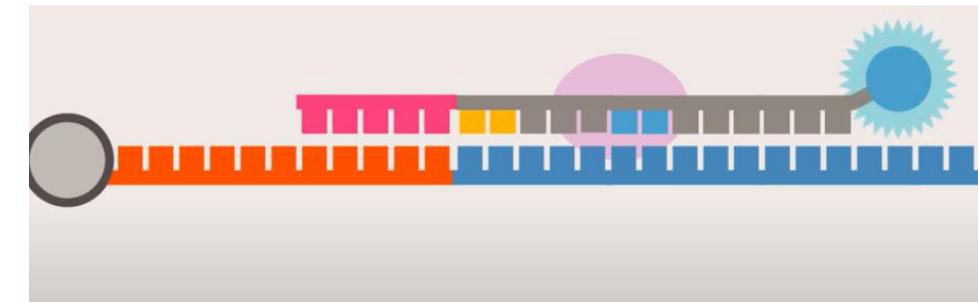
## Sequencing by Ligation



Signal is detected and recorded

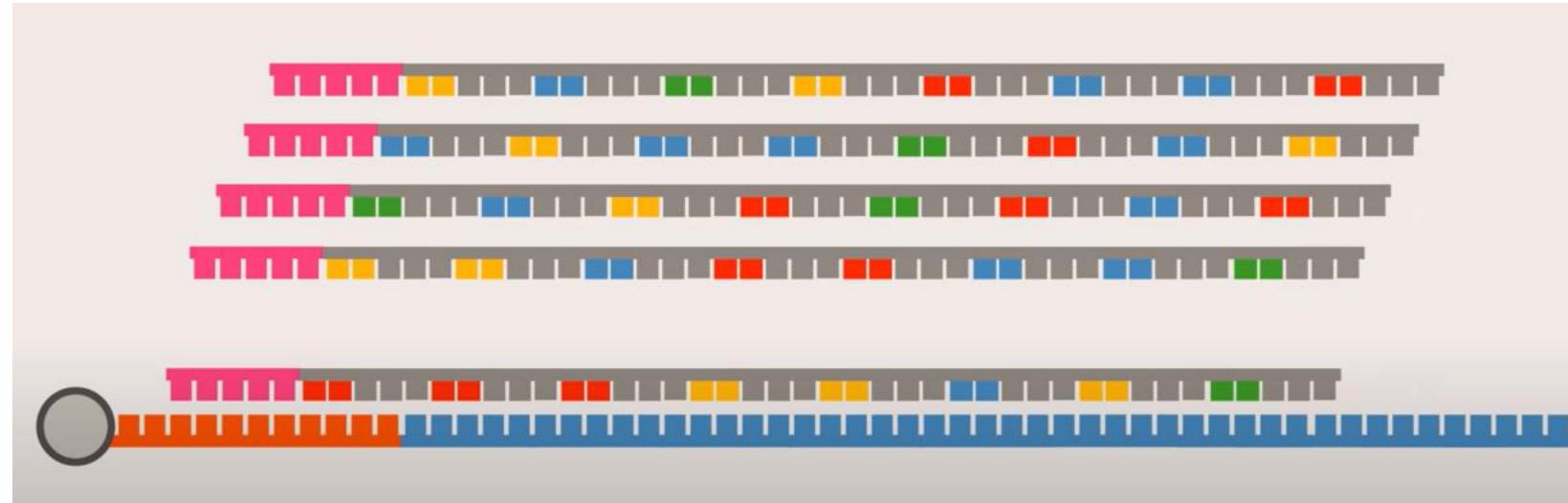


Fluorescent signal and last three bases are cleaved



After around 7 cycles of ligation DNA strand is denatured

## Sequencing by Ligation



- Primers, offset by one base from the previous primer is used to repeat these steps
- 5 sequencing primers are used in total

### Advantages

**High Accuracy:** Generates highly accurate sequencing data by using successive offset primers to sequence each nucleotide twice, reducing the likelihood of SNP miscalls.

### Limitations

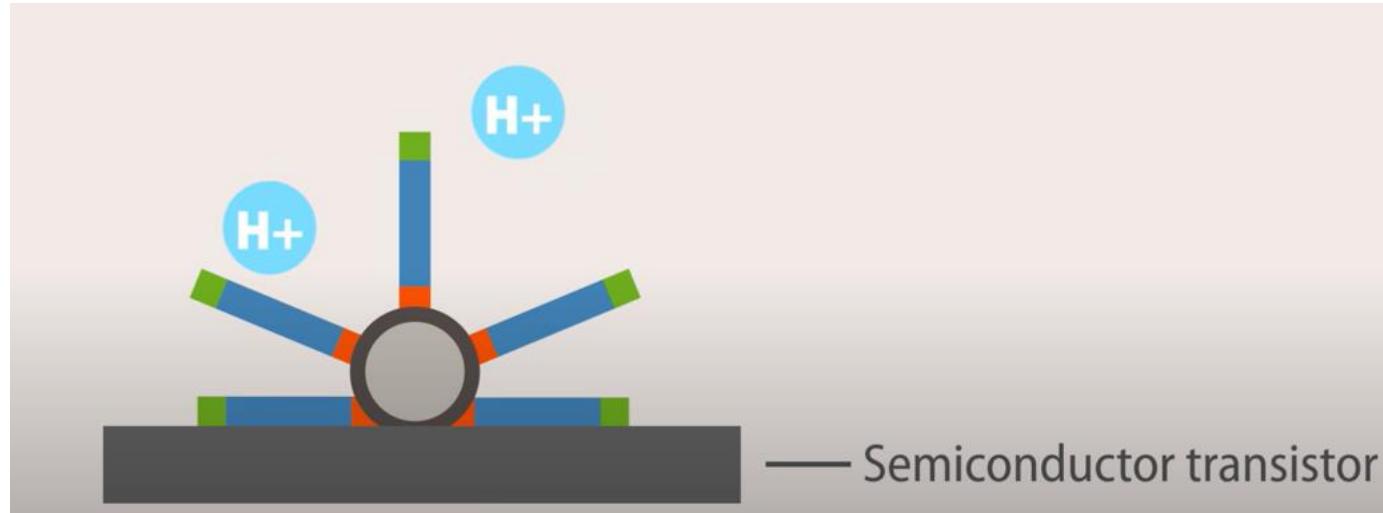
**Lower Data Output and Shorter Read Lengths:** Produces less data compared to Illumina and requires shorter read lengths, necessitating close genome sequencing for mapping.

**Longer Run Time:** Takes about 6-7 days to complete a whole run, particularly for larger genomes.

## Ion Semiconductor sequencing

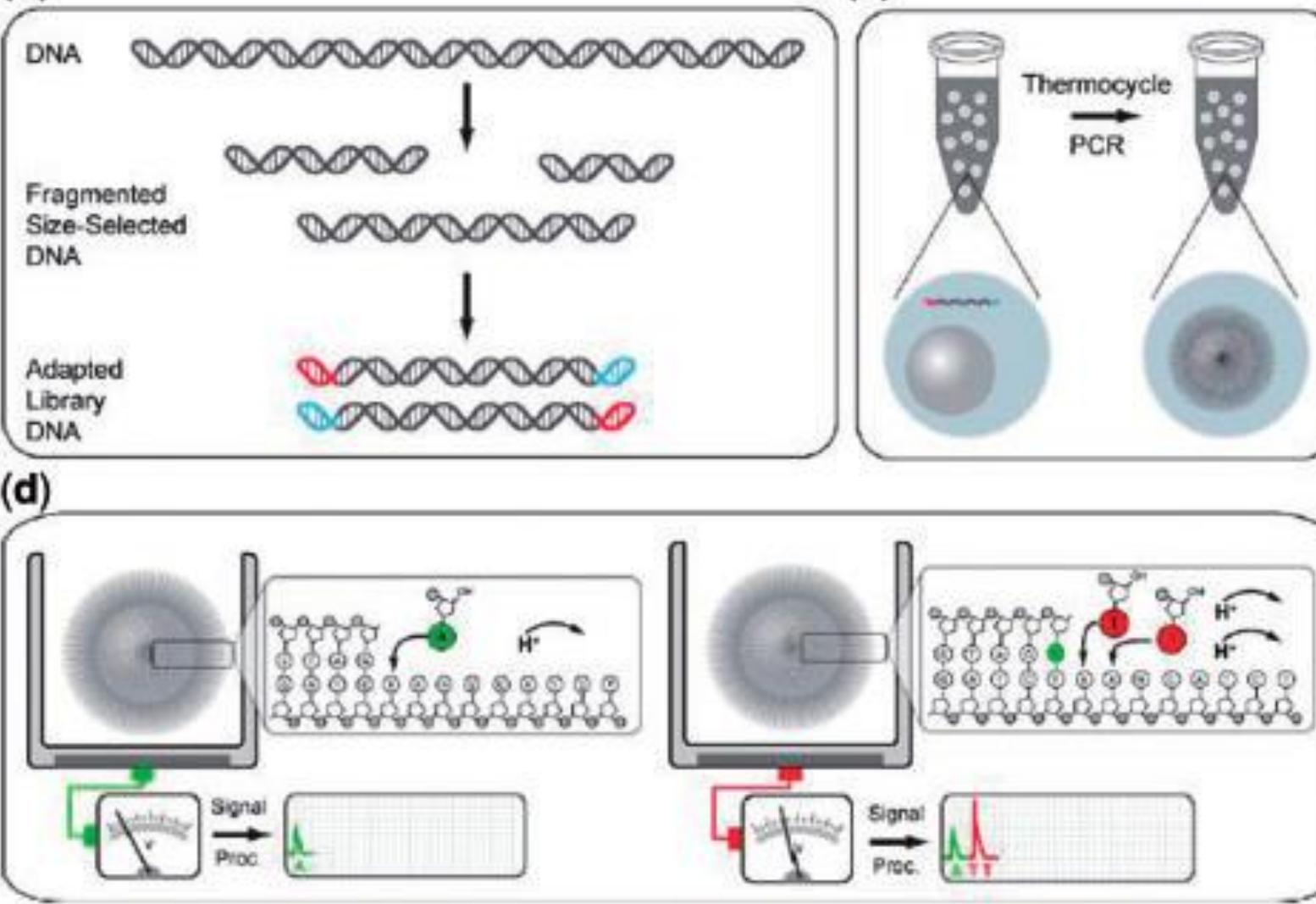


Utilizes the release of H<sup>+</sup> ions during the sequencing

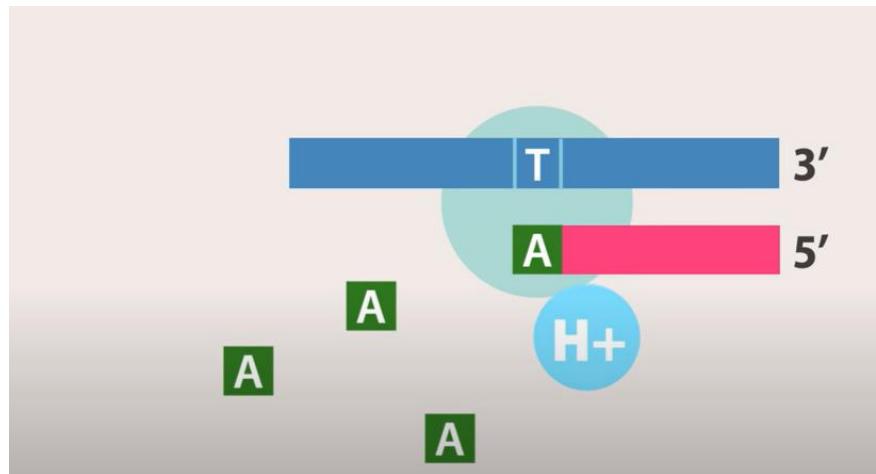


- Consist of semiconductor transistor –
- Capable of detecting changes in pH of the solution

## Ion torrent Sequencing



- Hydrogen ion released after incorporation of correct nucleotide
- Change of pH in the solution detected by sensor
- This change is converted into a voltage signal which is proportional to the number of nucleotides incorporated

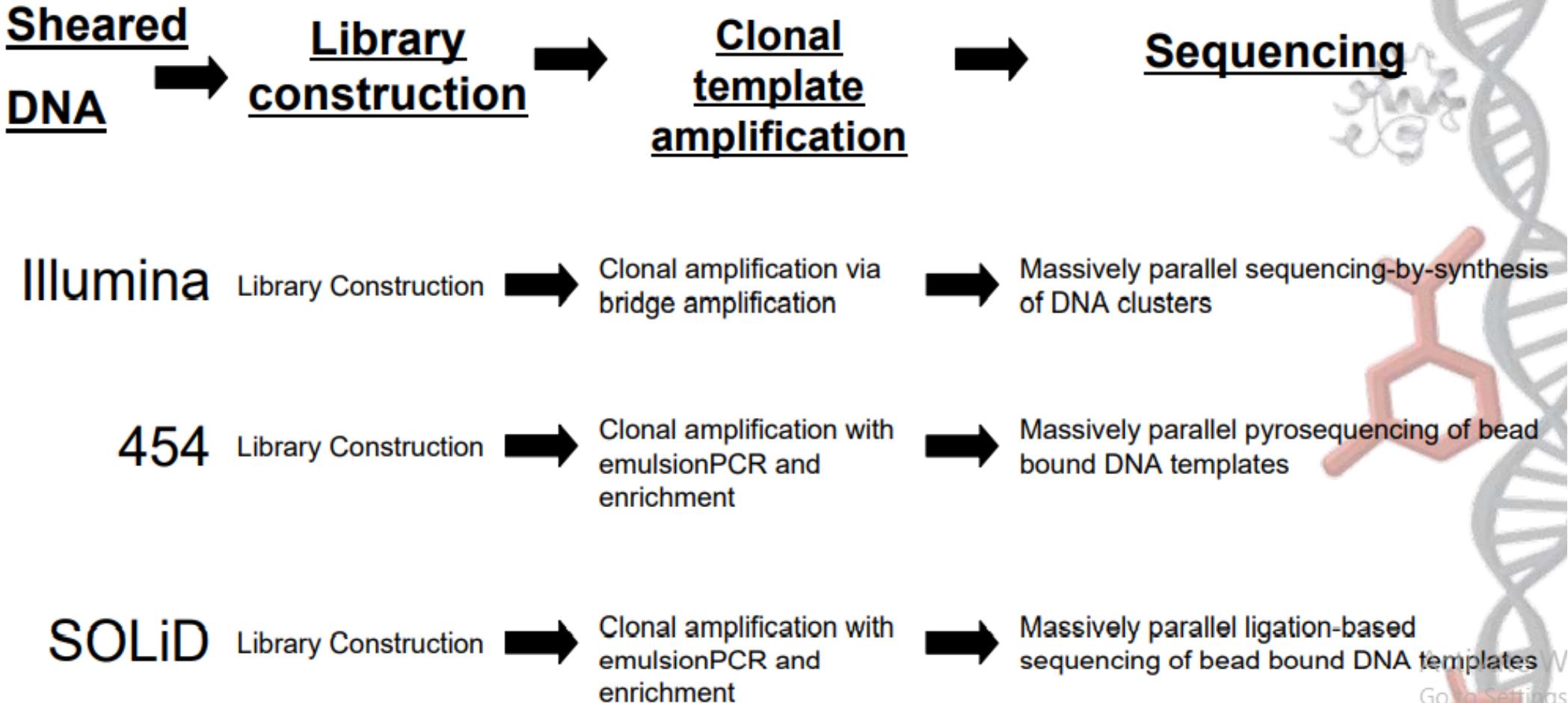


During nucleotide incorporation single hydrogen ion is released and is detected by the semiconductor

### Ion semiconductor sequencing: Overview

- 1 Similar to pyrosequencing, but measures the release of H<sup>+</sup> instead of pyrophosphate
- 2 More cost-effective and time-efficient

# Differentiating Next Gen technologies

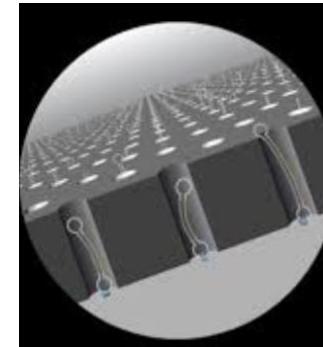


# Comparison of Next Gen Technologies

	GS FLX <u>Roche/454</u>	Genome Analyzer <u>illumina</u>	SOLiD
<b>Library Construction</b>	Fragment, Mate-Paired	Fragment, Mate-Paired, Paired-End	Fragment, Mate-Paired
<b>Sequencing Chemistry</b>	Sequencing by Synthesis	Sequencing by Synthesis	Sequencing by Ligation
<b>DNA Support</b>	25-35 µm bead	Flow cell surface	1 µ bead
<b>Amplification</b>	Emulsion PCR	Cluster amplification	Emulsion PCR
<b>Sequencing Reaction Surface</b>	High density well-plate	8-channel flow cell	Single slide imaged in panel

## The third generation of DNA sequencing

Pacific Biosciences, Inc. (**PacBio**) is the pioneer of third generation sequencing technology. In 2010, the introduction of its zero-mode waveguide (ZMW) raised the bar for DNA sequencing methods.



ZMW

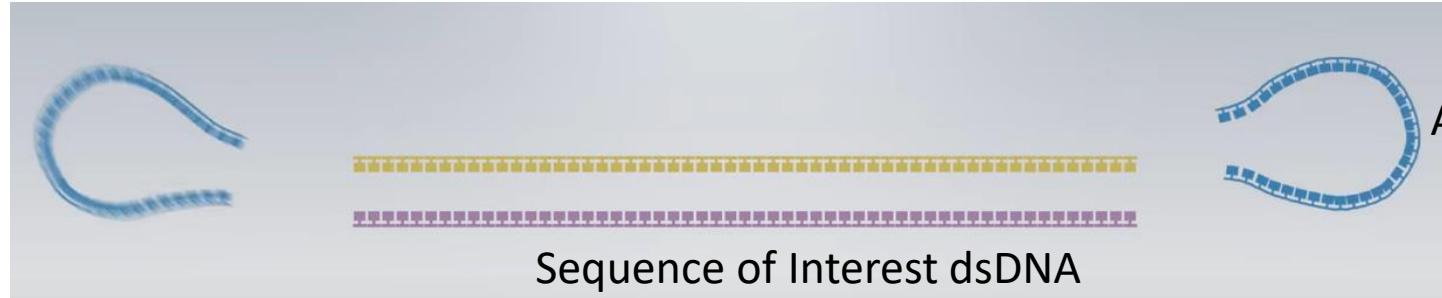
- ZMW (Zero mode wave-guides) utilises “nanoholes” that contain a single DNA polymerase
- **the incorporation of a single nucleotide can be observed directly**
- **Each nucleotide is labelled with a different fluorescent dye, and the signal that is emitted during incorporation is immediately recorded and read by very strong detectors attached below the ZMW**



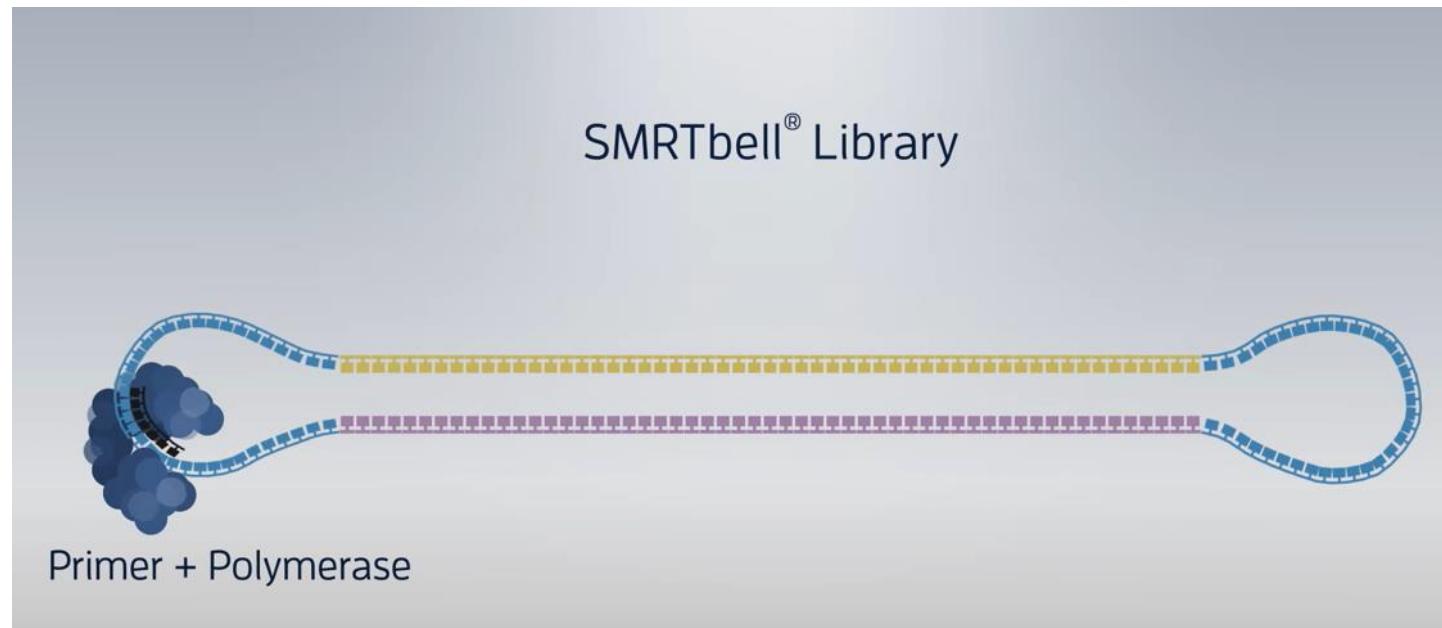
PacBio RSII sequencer

# PacBio Sequencing or SMRT sequencing

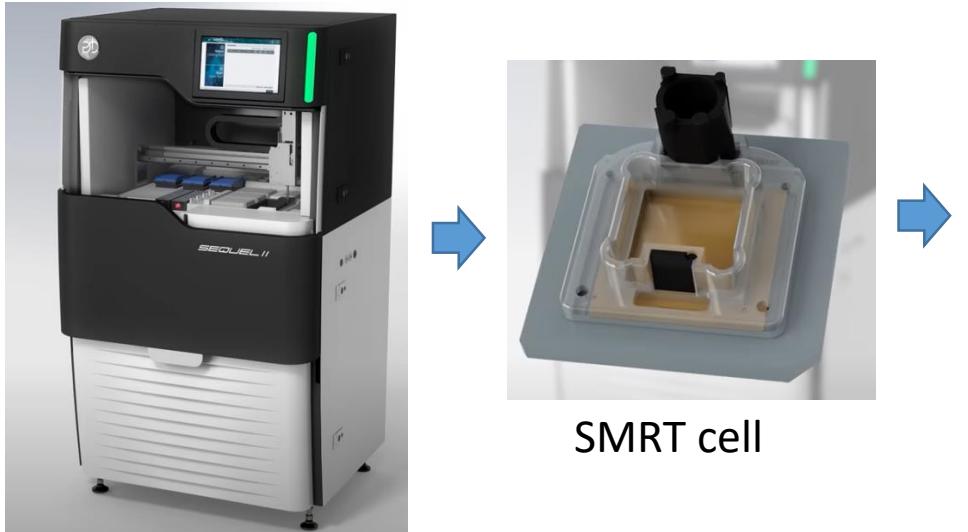
SMRT: Single molecule real time sequencing



- Adapter is ligated to ds DNA along with prime and polymerase

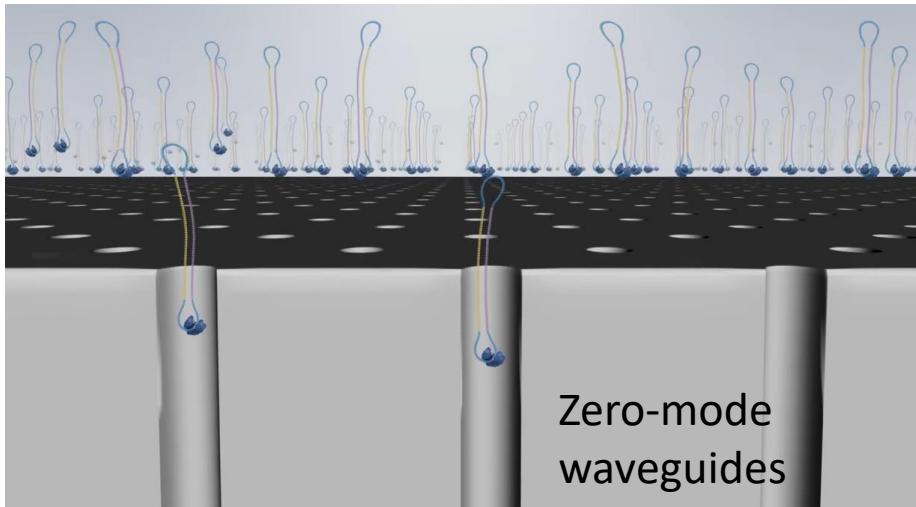


# PacBio Sequencing or SMRT sequencing

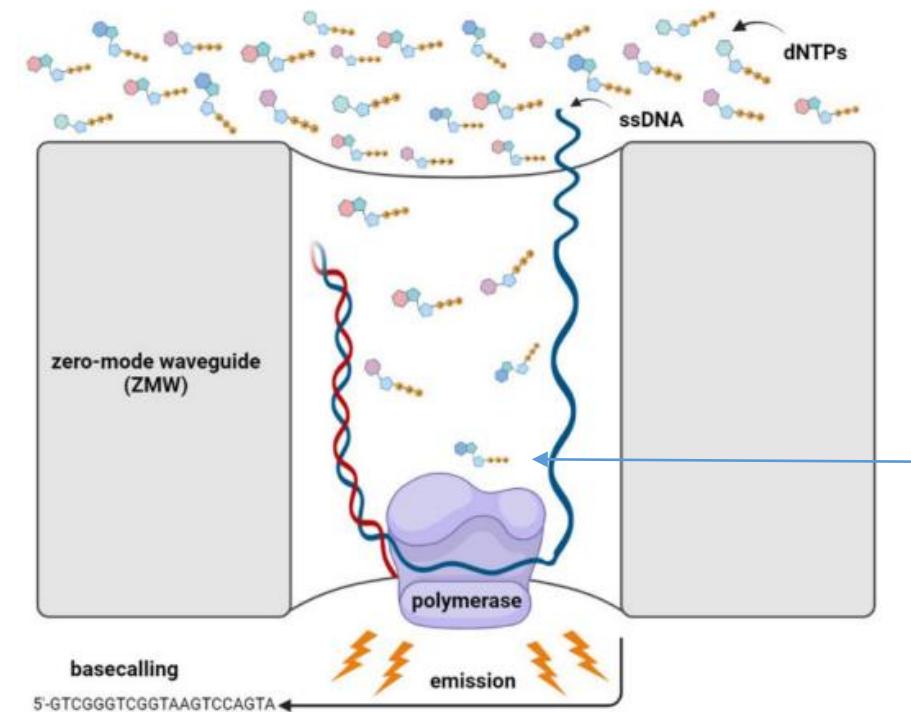


SMRT cell

- First nano-sensor based technology
- DNA Polymerase is immobilised in each well of a **SMRTcell** (specific silicon chip) while DNA is mobile.



Zero-mode  
waveguides

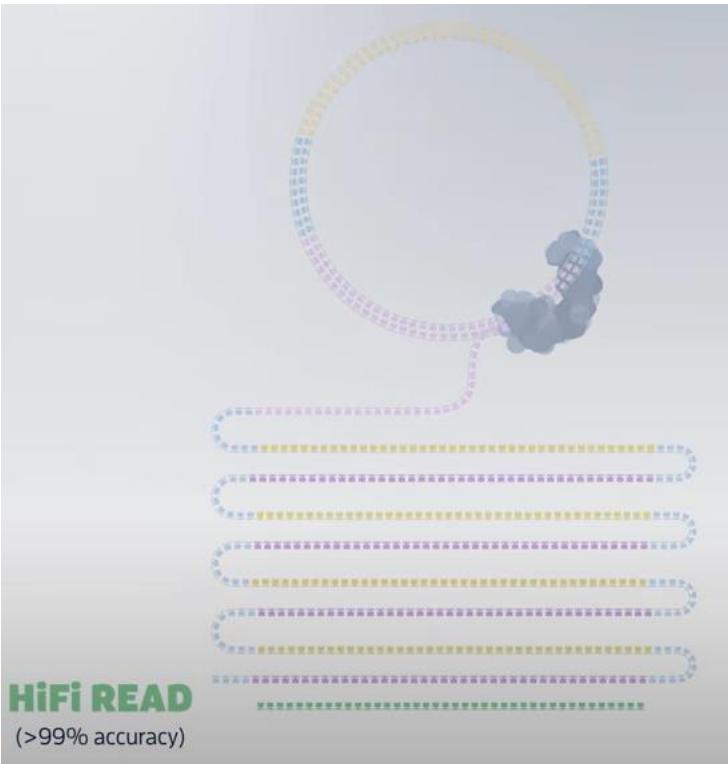


Each Nucleotide incorporation is measured in realtime

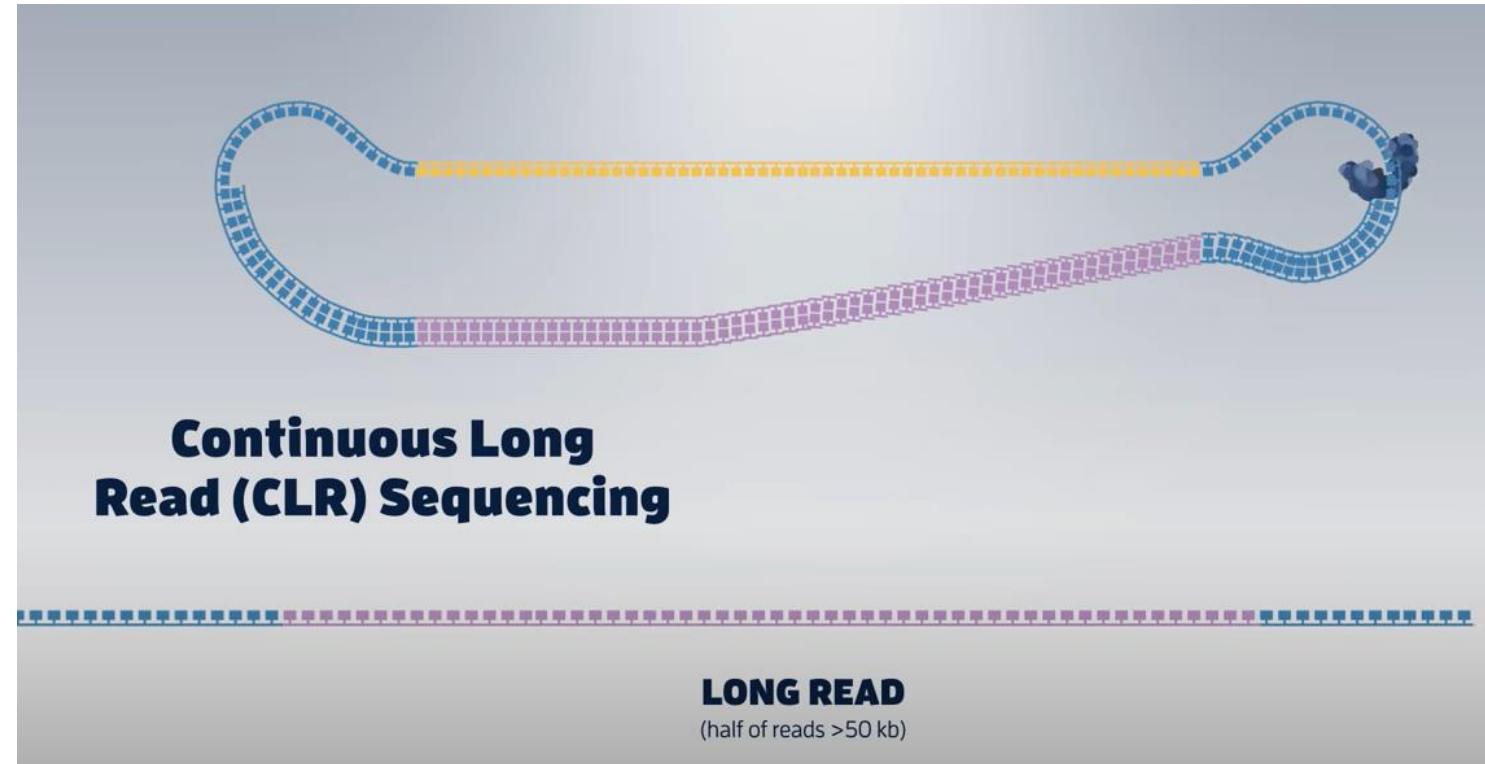
# PacBio Sequencing or SMRT sequencing

## Two Sequencing Modes

### Circular Consensus Sequencing (CCS)



### Continuous Long Read (CLR) Sequencing



## 3rd generation of DNA sequencing

### SINGLE-MOLECULE SEQUENCING SYSTEMS

GridION, MinION or Flongle are portable handheld system for RNA and DNA sequencing for reads of more than 2 Mb.



Oxford Nanopore Technologies MinION

The GridION was first introduced in 2012 and **uses the changes in electrical conductivity** that occur when DNA strands pass through biological nanopores in order to identify the nucleotide sequence. (Hayden, 2012; Lu *et al.*, 2016).



NovaSeq 6000



OhmX Analyzer™

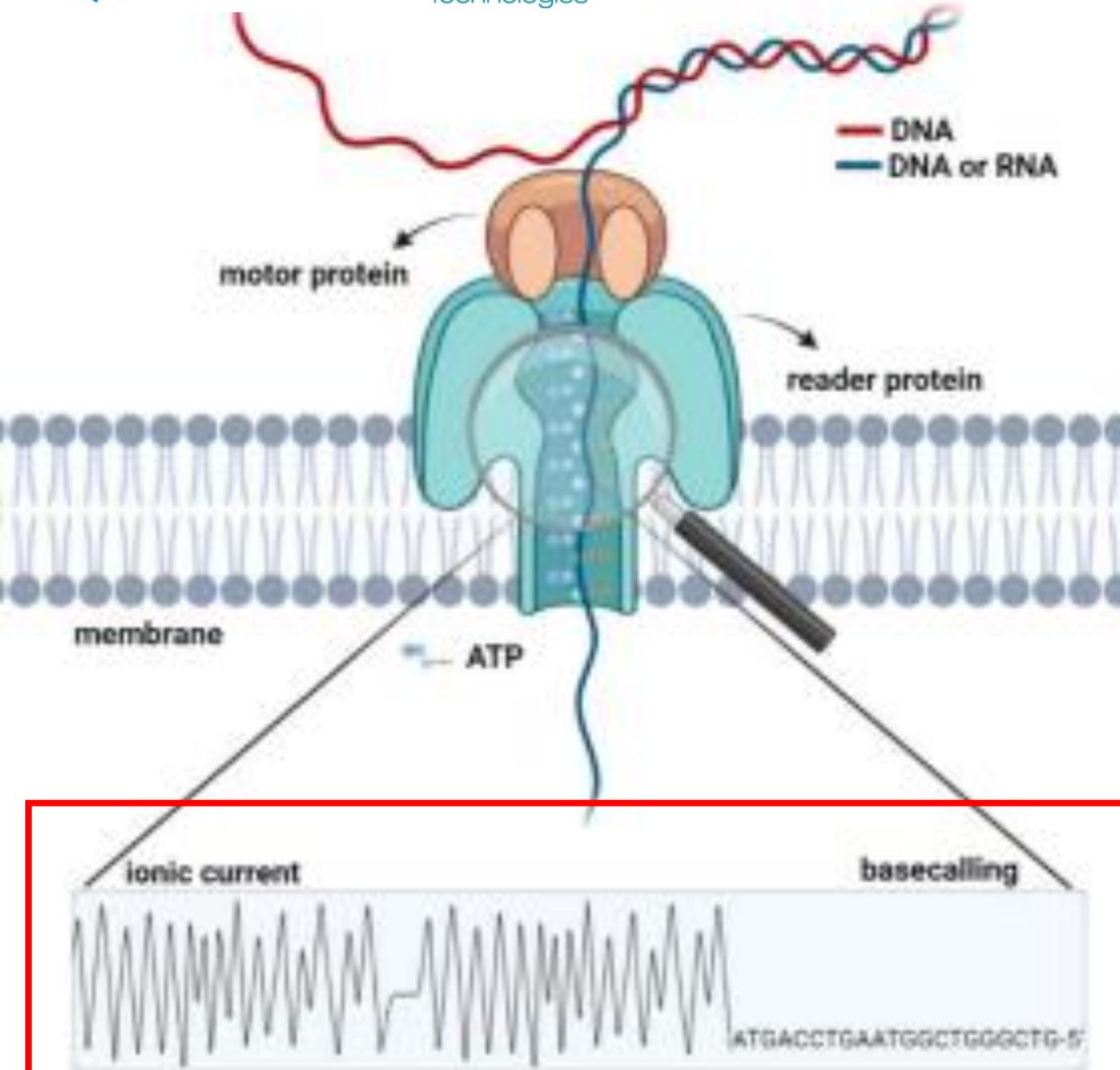
### Illumina's NovaSeq platforms

- NovaSeq 6000 (S4 flow cell) generates an output of up to 3000 Gb (Illumina, 2017).

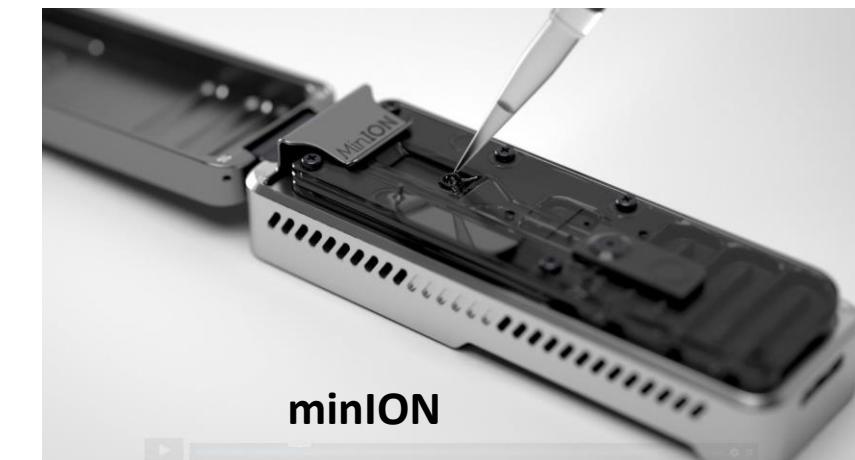
### **Nabsys' HD-Mapping**

Uses sequence-specific tags to label long DNA fragments that are then detected by nano-detectors and subsequently compiled to a map of the genome.

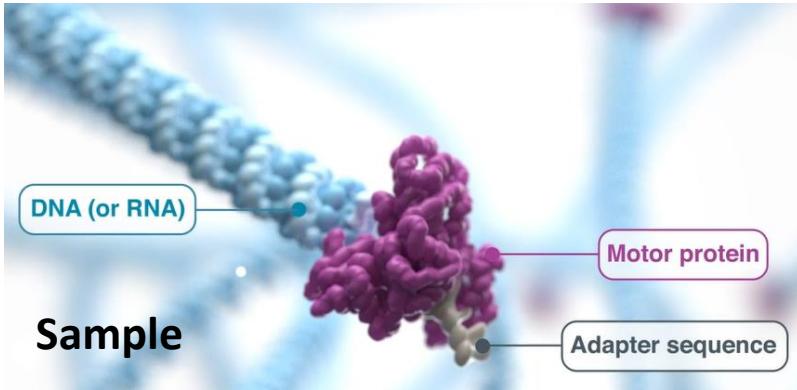
# Nanopore Sequencing

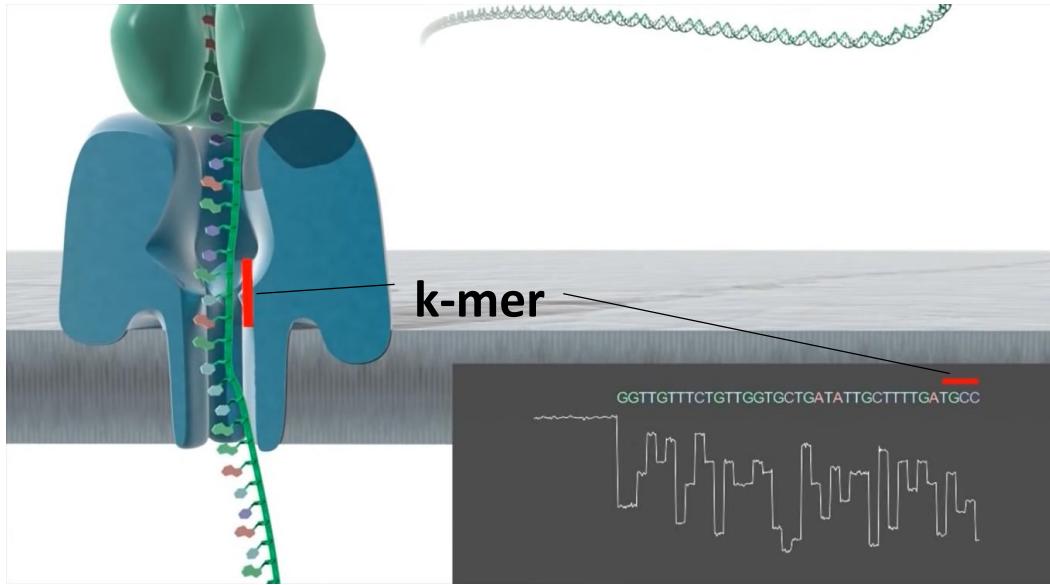


- It relies on detecting changes in the electric current, caused by the disorder of nanopore proteins when DNA or RNA strands pass through them.
- It provides sequencing in **real time**.
- It consists of **nanosensors**, a special structure of channels via which DNA penetrates.
- A motor protein, directs the single-stranded template and facilitates its passage through pore.

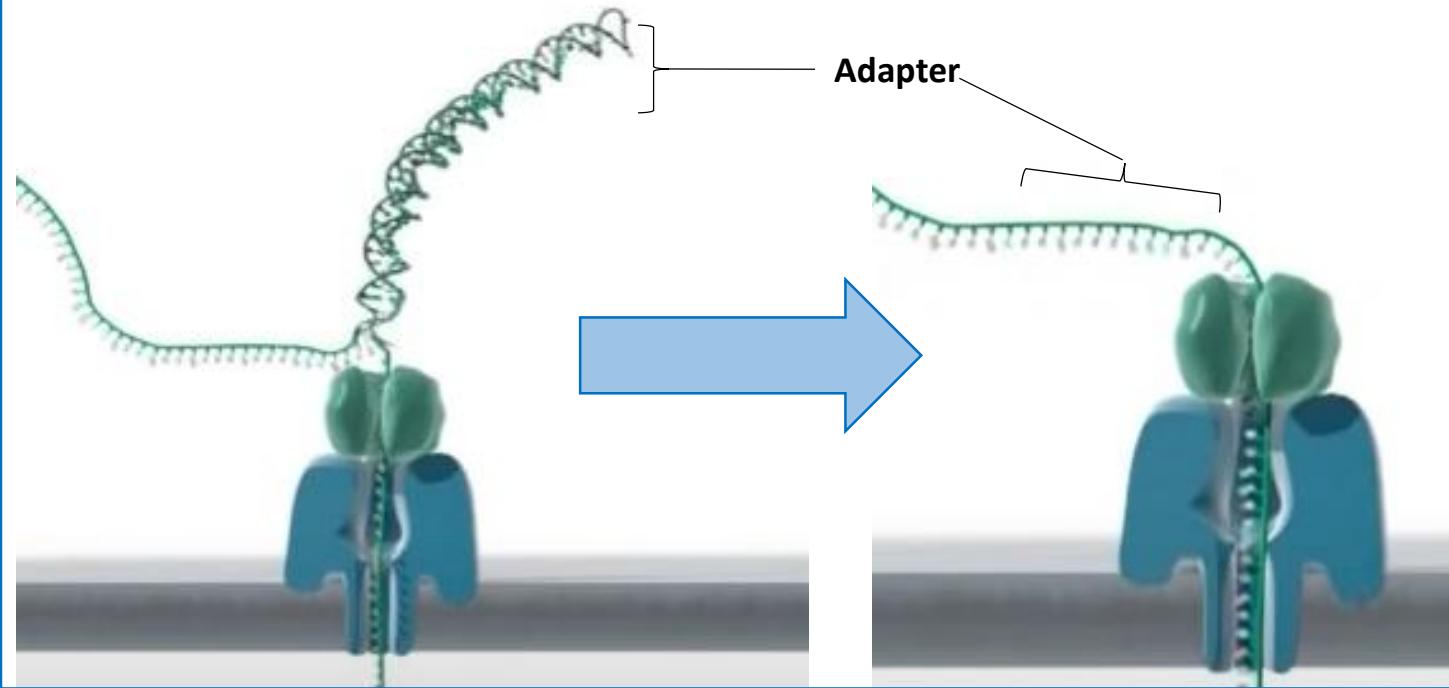


# Nanopore sequencing working





- ONT (oxford nanopore) systems do not identify individual bases and the observable current is determined by short nucleotide sequences of approx. 5 bases, also known as **k-mer**.
- 1000 different signals, one of each **k-mer** can be detected passing via nanopore.



- A primary adapter is used to promote the second strand through the pore after passage of first one and, thus, enabling the **sequencing of both strands**.

# Nanopore sequencing devices

## Portable sequencing



Flongle

minION

## Benchtop sequencing



minion Mk1C



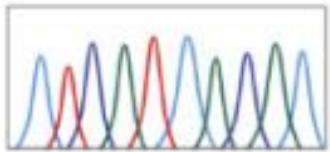
GridION



PromethION

## First generation

Sanger sequencing



1977



Maxam-Gilbert sequencing

## Next generation Second generation

Illumina GA



35bp  
 $30 \times 10^6$  reads

Ion torrent PGM



100bp  
 $27 \times 10^4$  reads

2005

2006 2007

2010

454 Pyrosequencing  
110bp  
 $20 \times 10^4$  reads

ABI's SoLiD system



35bp  
 $100 \times 10^6$  reads

## Third generation

Nanopore sequencing family



>4 Mb

2011



PacBio Sequel System

10-15 kb (average  
read length)

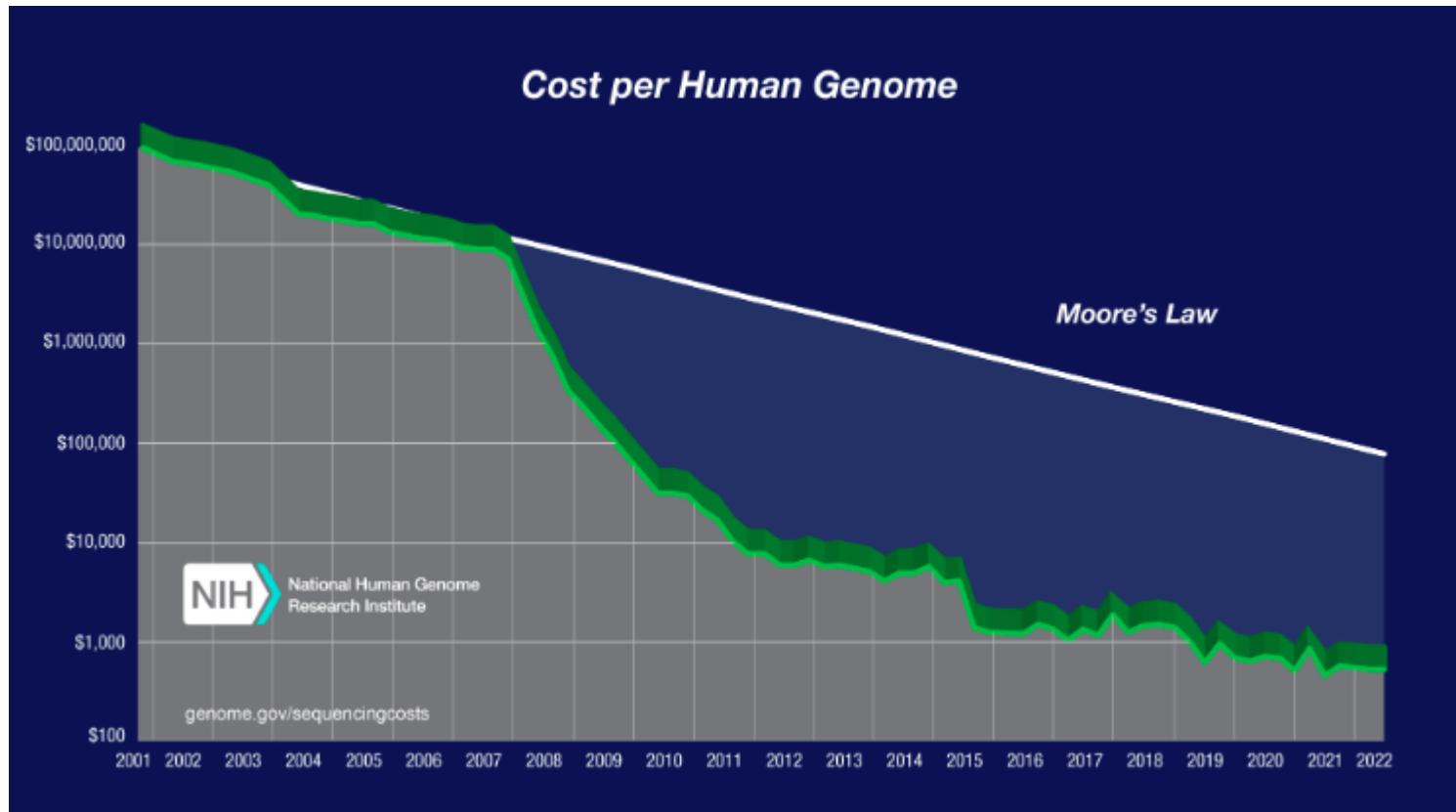
2014

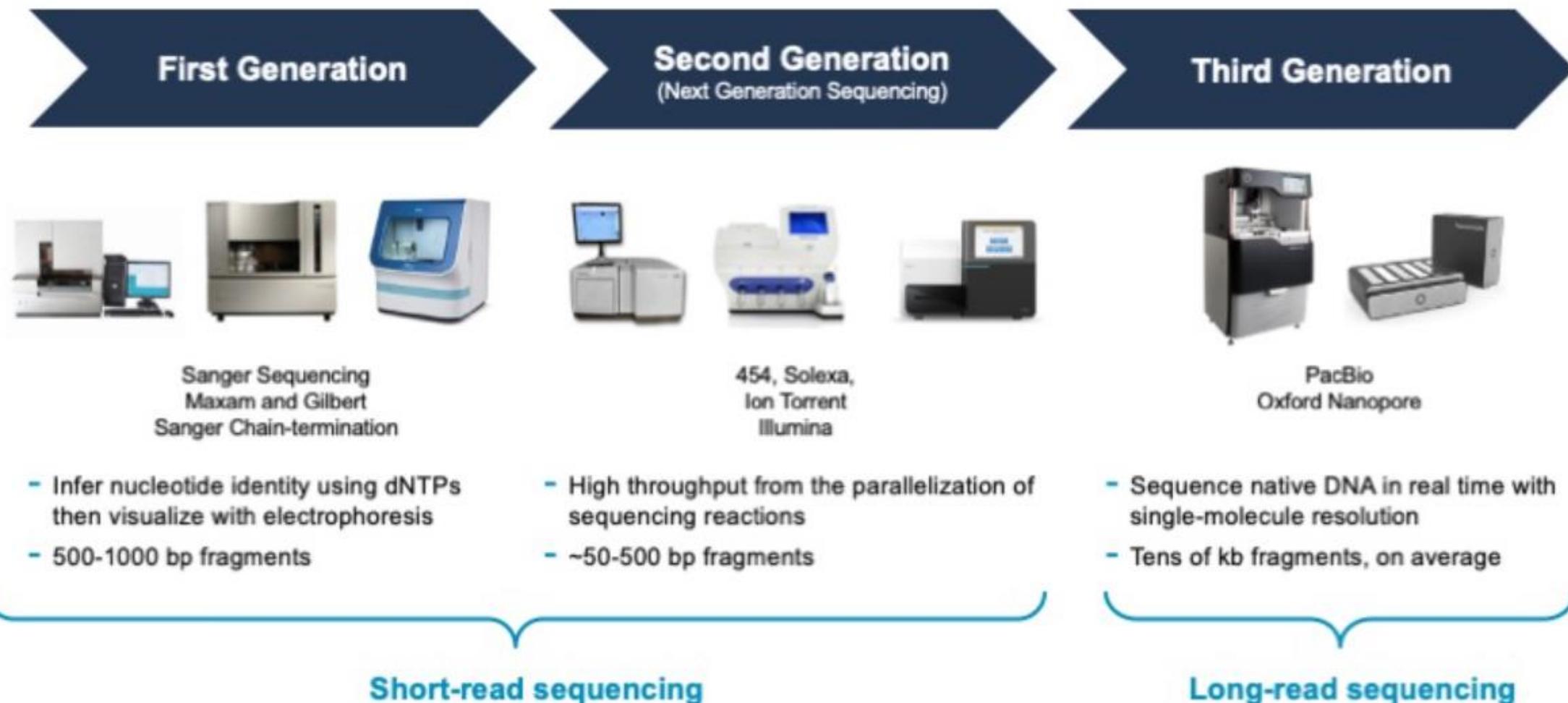
# Comparison

Platform	Sequencer	Average Read Length	Error Rate Per Read	Run Time	Maximum Throughput
PacBio	PacBio RS II	10–15 kb	10–15%	0.5–4 h	10 Gb
	Sequel System	10–15 kb	10–15%	≤20 h	10 Gb
	Sequel II System	10–15 kb	10–15%	≤30 h	500 Gb
	Sequel IIE System	10–15 kb	10–15%	≤30 h	500 Gb
ONT	MinION Mk1B	>4 Mb	~13%	1 min–72 h	50 Gb
	MinION Mk1C	>4 Mb	~13%	1 min–72 h	50 Gb
	GridION Mk1	>4 Mb	~13%	1 min–72 h	250 Gb
	PromethION 24	>4 Mb	~13%	1 min–72 h	7 Tb
	PromethION 48	>4 Mb	~13%	1 min–72 h	14 Tb

- **Third generation sequencing is currently more expensive** than the Second generation, has high sequencing error, and generates significantly less coverage. However, **the reads are long enough for whole genome sequencing** .
- Therefore, hybrid sequencing approach is favourable. Closed-genome scaffold generated by 3<sup>rd</sup> generation method and deep coverage by using 2<sup>nd</sup> generation techniques for accurate analysis.

## Cost Factor:





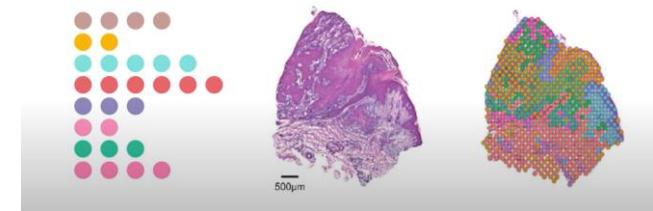
# Fourth Generation of Next-Generation Sequencing Technologies

This generation of the sequencing has the ability to *in situ* (perform sequencing directly in the cell) sequence the fixed tissue and cells by using second-generation methodology

## Spatial Transcriptomics Technology:

- Developed by Joakim Lundeberg and colleagues..
- Fresh-frozen tissue sections are imaged and permeabilized, allowing mRNA to diffuse onto the probes.
- Generates a sequencing library analyzed by NGS, mapping reads to spatial barcodes.

## Spatial Transcriptomics



## Advantages:

- High-throughput transcriptome-wide analysis with spatial context.
- Not yet at single cell resolution.

## Other Methods:

- Combinatorial labeling schemes (smFISH, MERFISH [multiplexed error-robust fluorescence *in situ* hybridization]) offer high resolution but limited throughput.

### **In Situ Sequencing (ISS):**

- Sequences RNA molecules directly in fixed cells/tissues.
- Uses padlock probes and rolling circle amplification (RCA) for targeted sequencing.
- ISS adapts to formalin-fixed, paraffin-embedded (FFPE) tissues.

### **Technical Challenges:**

- Autofluorescence and sample imaging limitations.
- Handling and interpreting large data sets from 2D/3D imaging.
- Physical limits of sequencing reactions within cells.

### **Future Directions:**

- Expansion microscopy to overcome physical limits.
- Improved data storage and compression methods.

# NGS Methods

## Genomics

WGS

Exome sequencing

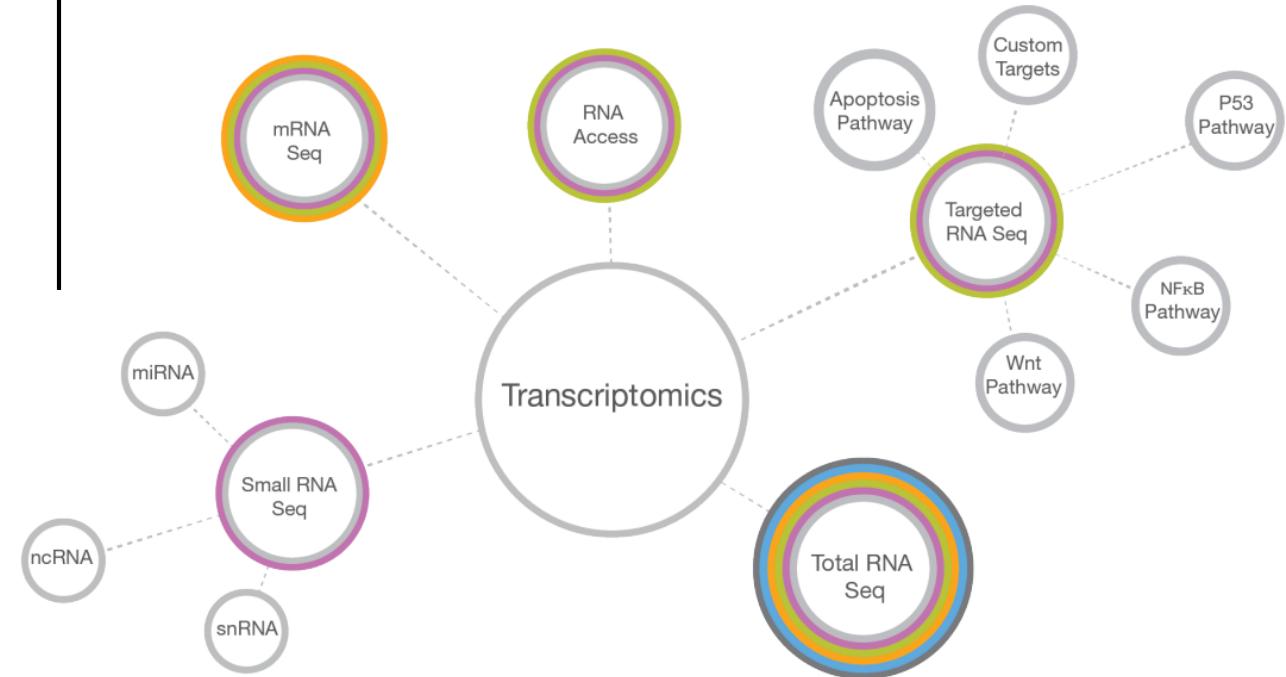
*De novo* Sequencing

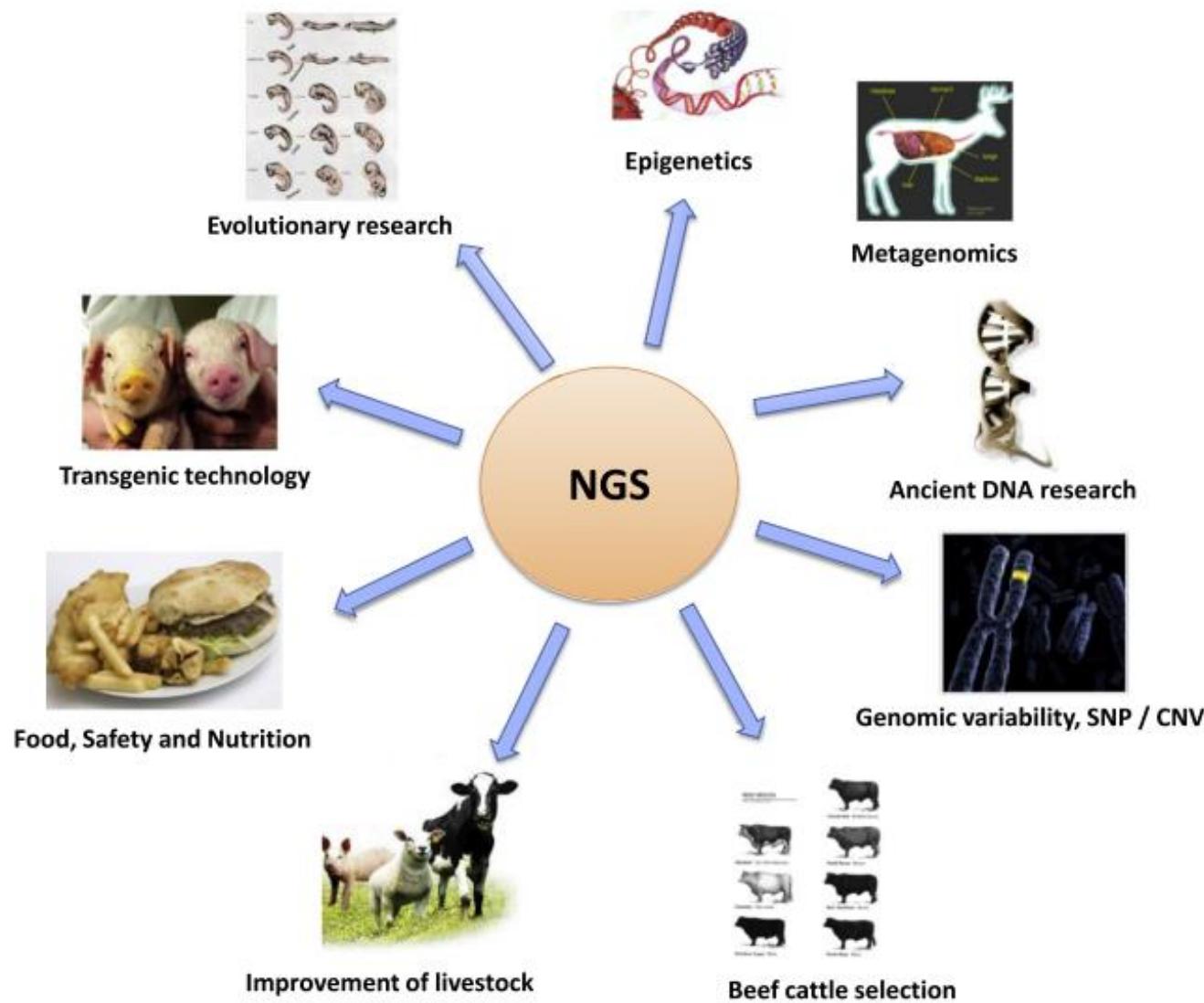
Targeted sequencing

## Epigenomics

Methylation Sequencing

## Transcriptomics





## Evolutionary Biology

**Study of Diversity:**  
Evolutionary processes shaping life on Earth.

**NGS Impact:** Next Generation Sequencing (NGS) enables **identification of genetic loci responsible for adaptation**, without requiring a closely related genetic model organism.

## Epigenetics

**Heritable Changes:**  
Epigenetics involves changes in gene expression without altering DNA sequences, such as DNA methylation and histone modification.

**Importance:** Crucial for understanding complex traits and diseases, influenced by environmental factors.

## Metagenome Sequencing

**Study of Mixed Genomes:**  
Analyzes genetic material from environmental samples, revealing microbial diversity.

**Gut Microbiomes:** NGS enhances understanding of gut flora interactions with host genotypes, impacting nutrient utilization and health.

## Ancient DNA

**Revolutionized Research:**  
NGS has vastly increased ancient DNA data, aiding in understanding species evolution despite challenges like high contamination.

## Genomic Variability: SNPs and CNVs

**Identifying Variations:** NGS helps identify Single Nucleotide Polymorphisms (SNPs) and Copy Number Variations (CNVs), which influence phenotypic traits and disease susceptibility.

## Beef Cattle Selection

**Advancements:** NGS enables efficient beef cattle selection, improving traits like meat quality through rapid, low-cost genome sequencing.

## Animal Breeding and Livestock Productivity

**Breeding Efficiency:** NGS improves selection accuracy for desirable traits, enhancing livestock productivity and product quality.

## Food Safety and Nutrition

**Food Biotechnology:** NGS aids in assessing genetic modifications in food, ensuring safety, and enhancing nutritional value.

## Transgenics

**Experimental Models:** Transgenic animals are used for disease research and therapeutic protein production, with NGS improving mutation detection and efficiency.

# Metagenomics and Its Application in Livestock Microbiome Research



## Rumen Microbiome

The rumen, a complex microbial ecosystem in ruminant animals like cattle and sheep, plays a crucial role in feed digestion and nutrient metabolism. Metagenomic studies have revealed the remarkable diversity and dynamic nature of the rumen microbiome, providing new avenues for improving feed efficiency, reducing methane emissions, and developing targeted interventions to optimize animal health and productivity.



## Gut Microbiome

The gut microbiome of livestock animals, including pigs and poultry, has been the focus of extensive research using metagenomic approaches. These studies have linked the composition and functionality of the gut microbiome to various aspects of animal health, such as nutrient utilization, immune function, and disease susceptibility, opening up opportunities for microbiome-based interventions to enhance livestock production and welfare.



## Disease Diagnostics

Metagenomic analysis of livestock samples, such as feces, soil, or water, can also facilitate the detection and characterization of pathogenic microorganisms, including bacteria, viruses, and parasites. This information can aid in the development of targeted disease prevention and control strategies, promoting the overall health and productivity of livestock populations.

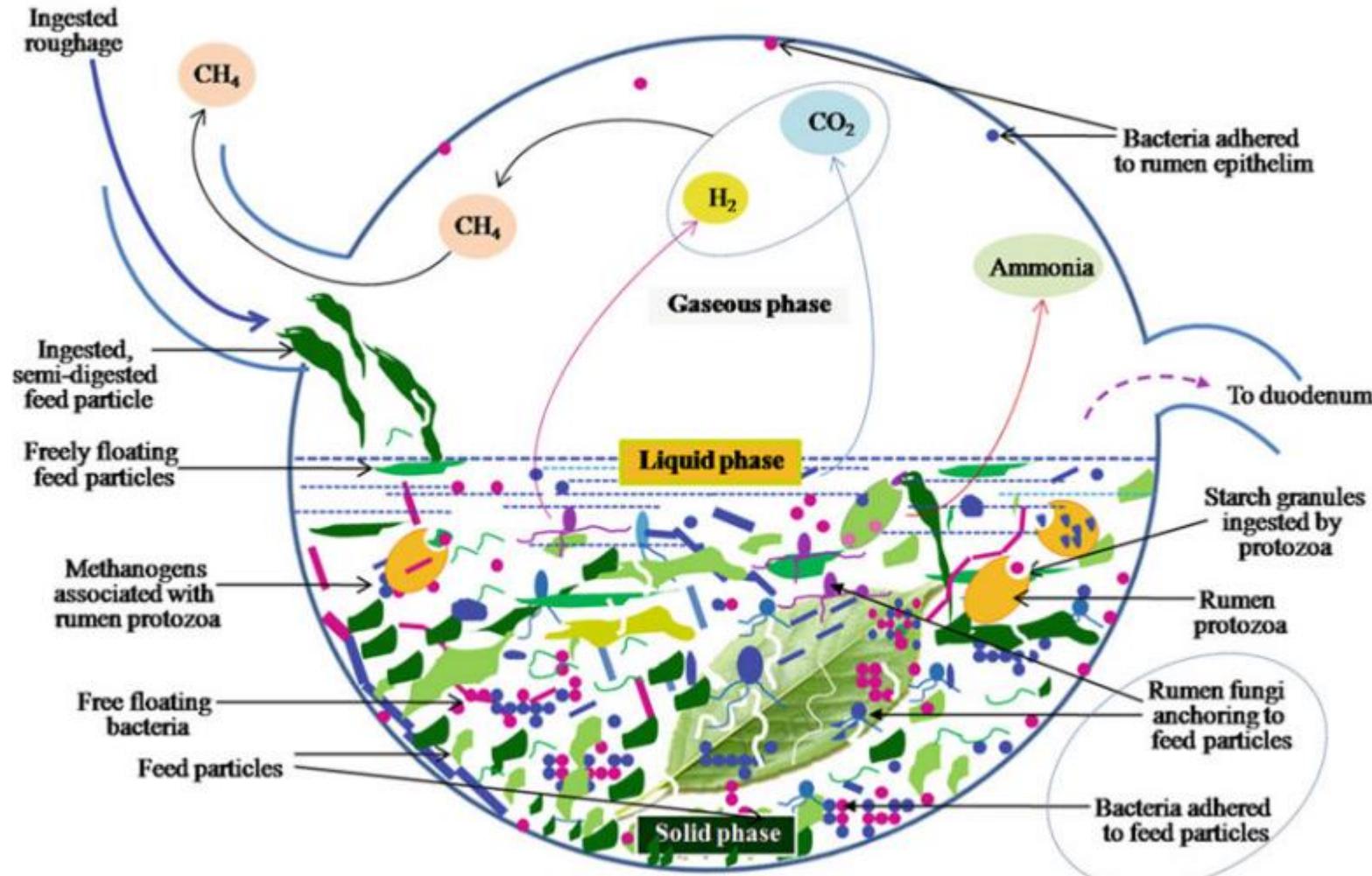
# How to choose the write sequencing method

	Sanger sequencing	NGS
Research application(s)	Microbial identification, plasmid sequencing to check that the correct DNA was inserted into a vector, NGS-detected variant validation, genome editing confirmation, cell line authentication, SNP genotyping, MLPA, microsatellite marker analysis, and more	Low-pass whole-genome sequencing (WGS); DNA sequencing (whole-exome); RNA sequencing (whole-transcriptome); targeted sequencing applications such as solid tumor profiling; discovery of novel variants/detection of translocations, copy number variants (CNVs), insertions and deletions (indels), and single-nucleotide variants (SNVs); and more
Read length	Up to 1,000 bp	Short reads
Throughput	Low	High
Number of samples/targets	1 or 2 genes or up to 96 targets	>96 samples and/or targets in a single run
Cost per sample	Low for small-scale projects	Low for large-scale projects
Ease of use	+	++
Turnaround time	<1 day	<1 day

Platform	Advantages	Drawbacks	Recommended applications
Sanger	<ul style="list-style-type: none"> <li>• Costs (low target number)</li> <li>• Established workflow</li> <li>• Simple data analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Sensitivity</li> <li>• Scalability</li> <li>• Sample input requirements</li> </ul>	TRS, validation of NGS data
Ion Torrent	<ul style="list-style-type: none"> <li>• Costs</li> <li>• Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Short length reads</li> <li>• Accuracy</li> </ul>	TRS, metagenomics
Illumina	<ul style="list-style-type: none"> <li>• Sensitivity</li> <li>• Amount of generated data with same DNA</li> <li>• High throughput</li> </ul>	<ul style="list-style-type: none"> <li>• Costs for low target numbers</li> <li>• Short length reads</li> </ul>	WGS, WES, TRS, RNAseq, epigenomics, metagenomics
BGI Group	<ul style="list-style-type: none"> <li>• Accuracy</li> <li>• No optical duplicates</li> </ul>	<ul style="list-style-type: none"> <li>• Short length reads</li> </ul>	WGS, WES, TRS
Pacific Biosciences	<ul style="list-style-type: none"> <li>• Long reads</li> <li>• High accuracy with CCS mode</li> <li>• Direct detection of epigenetic modifications</li> </ul>	<ul style="list-style-type: none"> <li>• Costs</li> <li>• Large amounts of starting material</li> <li>• Error rate with CLR mode</li> </ul>	WGS, TRS, RNAseq
Oxford Nanopore Technologies	<ul style="list-style-type: none"> <li>• Very long reads</li> <li>• Direct sequencing of RNA</li> <li>• Detection of RNA modifications</li> </ul>	<ul style="list-style-type: none"> <li>• Costs</li> <li>• Error rate</li> <li>• Large amounts of starting material</li> </ul>	WGS*, TRS, RNAseq, epigenomics, metagenomics

—Eucarya, Archaea, and Bacteria

$10^{10}$  bacteria/ml,  $10^6$  protozoa/ml, and  $10^3$  fungi/ml.



- Amplicon
- Whole genome metagenome
- Metatranscriptomics

Diagrammatic illustration of rumen, and the rumen components,

# Bioinformatic tools

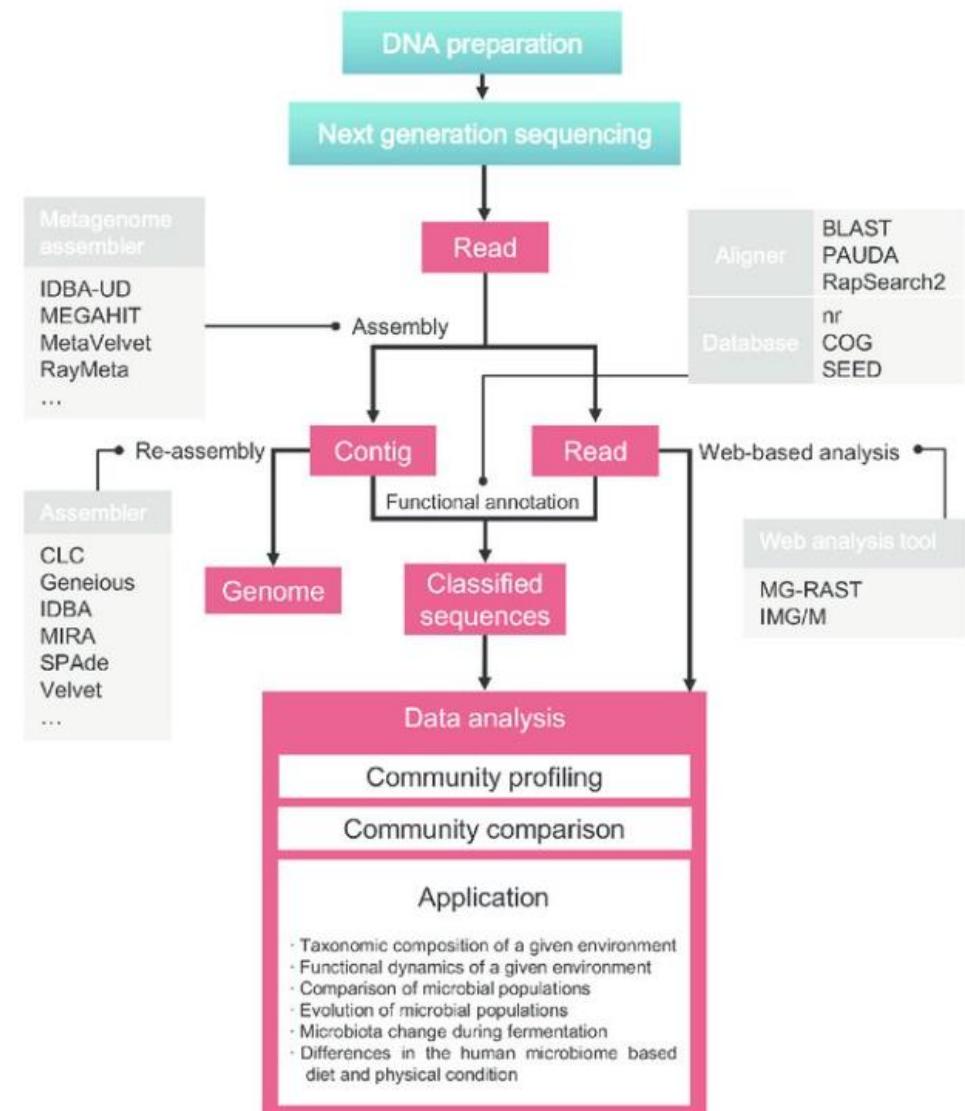


## METAGENOMICS

Analysis of metagenomics (HVR amplicon or whole 16S rRNA) data



Pipe E



WHAT I SAY

WHAT I THINK

SO ANY  
QUESTIONS?

THANK YOU  
END OF PRESENTATION



PLEASE DON'T ASK  
ANY QUESTIONS!  
PLEASE LET THIS BE OVER.

THANK YOU  
END OF PRESENTATION

