Next generation sequencing
Medical biotechnology

WHAT IS SEQUENCING?

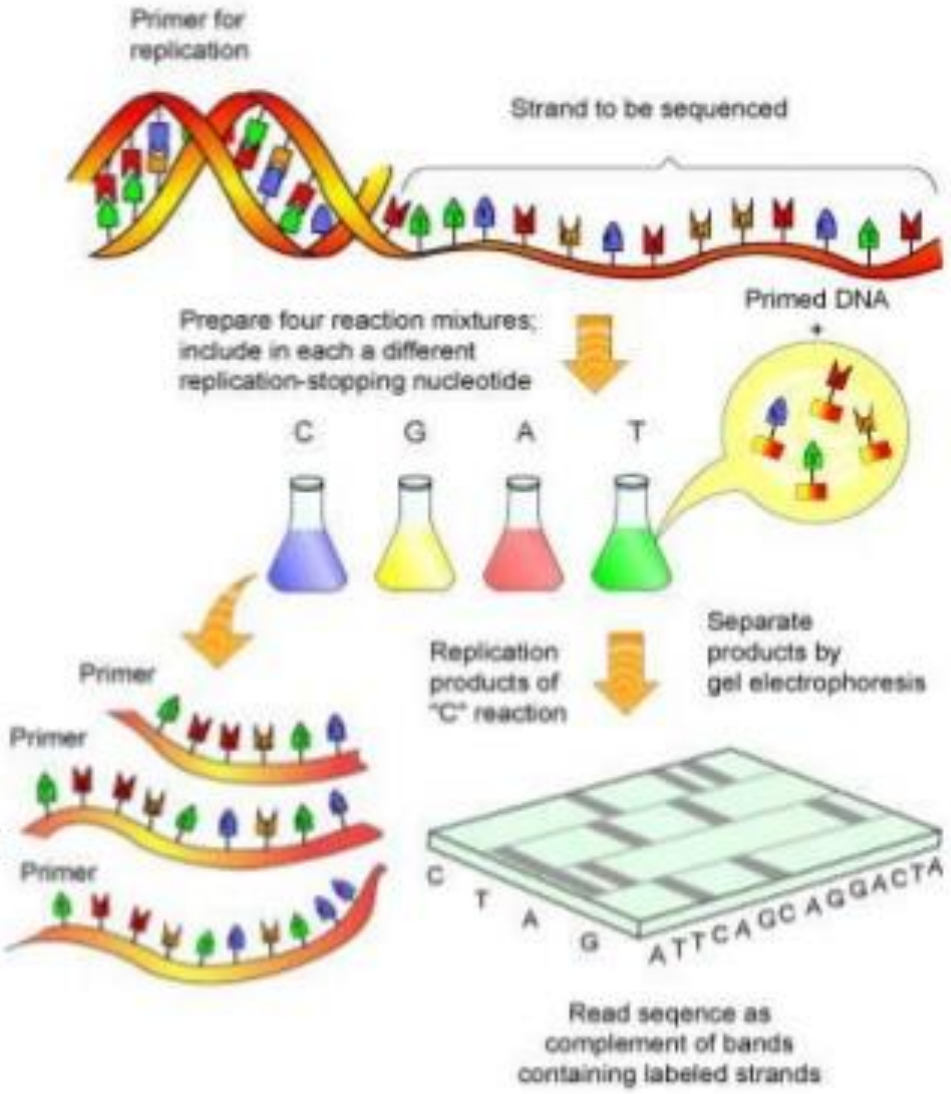
Genome sequencing is a method used to figure out the order of DNA nucleotides, or bases in a **genome**-the order of A, C, G, and T that make up an organism's DNA.

IMPORTANCE OF DNA SEQUENCING

- Better understanding of gene expression. Gene expression has significance in protein creation etc.
- It is capable to detect various diseases and genetic illnesses.
- Personalised medicine and disease discovery is possible.
- Forensics

Automated Sanger method

1. Bacterial cloning or PCR template purification
2. labelling of DNA fragments using the chain termination method with energy transfer
3. dye-labelled di-de oxynucleotides and a DNA polymerase
4. capillary electrophoresis
5. fluorescence detection that provides four-colour plots to reveal the DNA sequence.



NEXT GENERATION SEQUENCERS

Next-generation sequencing (NGS), also known as high-throughput sequencing, is the catch-all term used to describe a number of different modern sequencing technologies including:

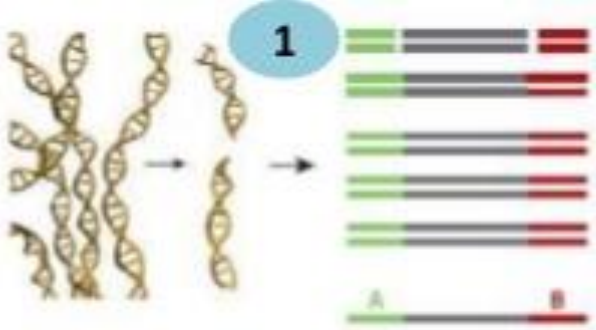
- Illumina (Solexa) sequencing
- Roche 454 sequencing
- SOLiD sequencing
- Ion torrent: Proton / PGM sequencing

NEXT GENERATION SEQUENCERS

- These recent technologies allow us to sequence DNA and RNA much more quickly and cheaply than the previously used Sanger sequencing, and as such have revolutionised the study of genomics and molecular biology.
- NGS has brought high speed not only to genome sequencing and personal medicine it has also changed the way we do genome research

OVERVIEW OF NEXT GENERATION SEQUENCING PROTOCOL

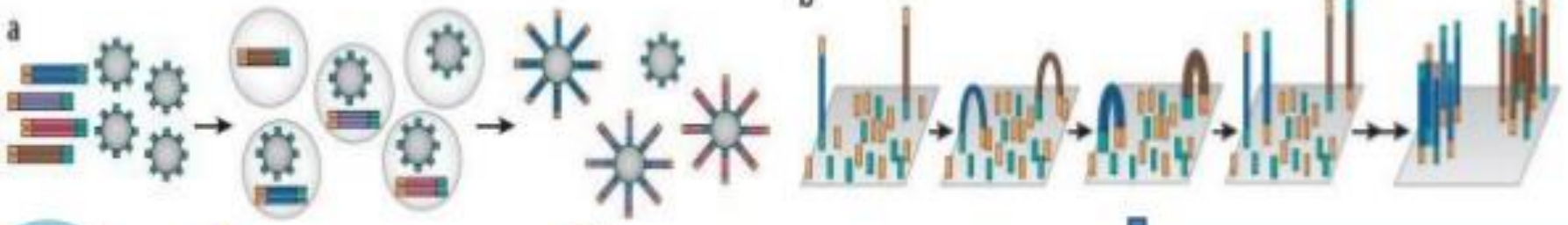
1. LIBRARY PREPARATION
2. CLONAL AMPLIFICATION
3. CYCLIC ARRAY SEQUENCING



2

Emulsion PCR

Bridge PCR

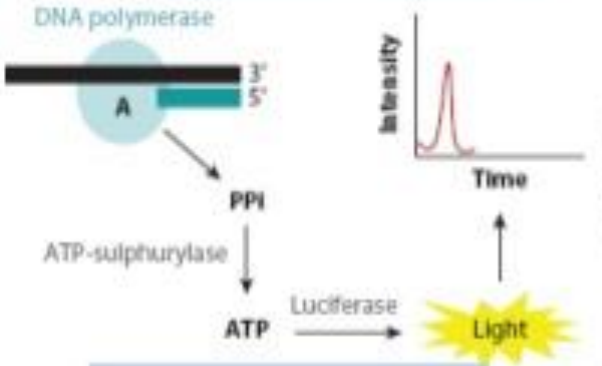


3

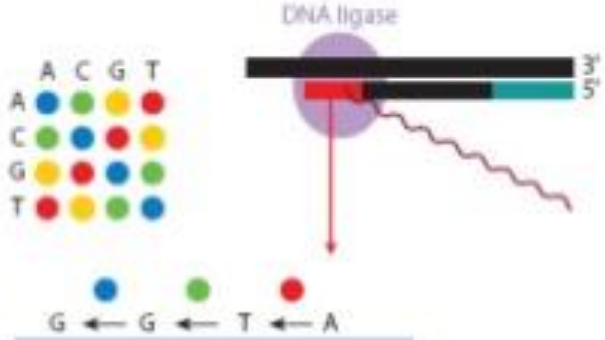
Pyrosequencing

Sequencing-by-ligation

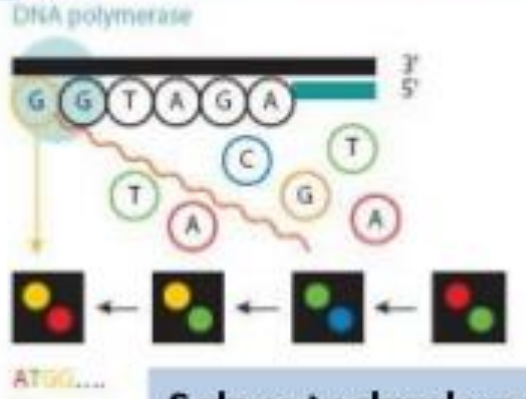
sequencing-by-synthesis



454 sequencing



SOLID platform



Solexa technology

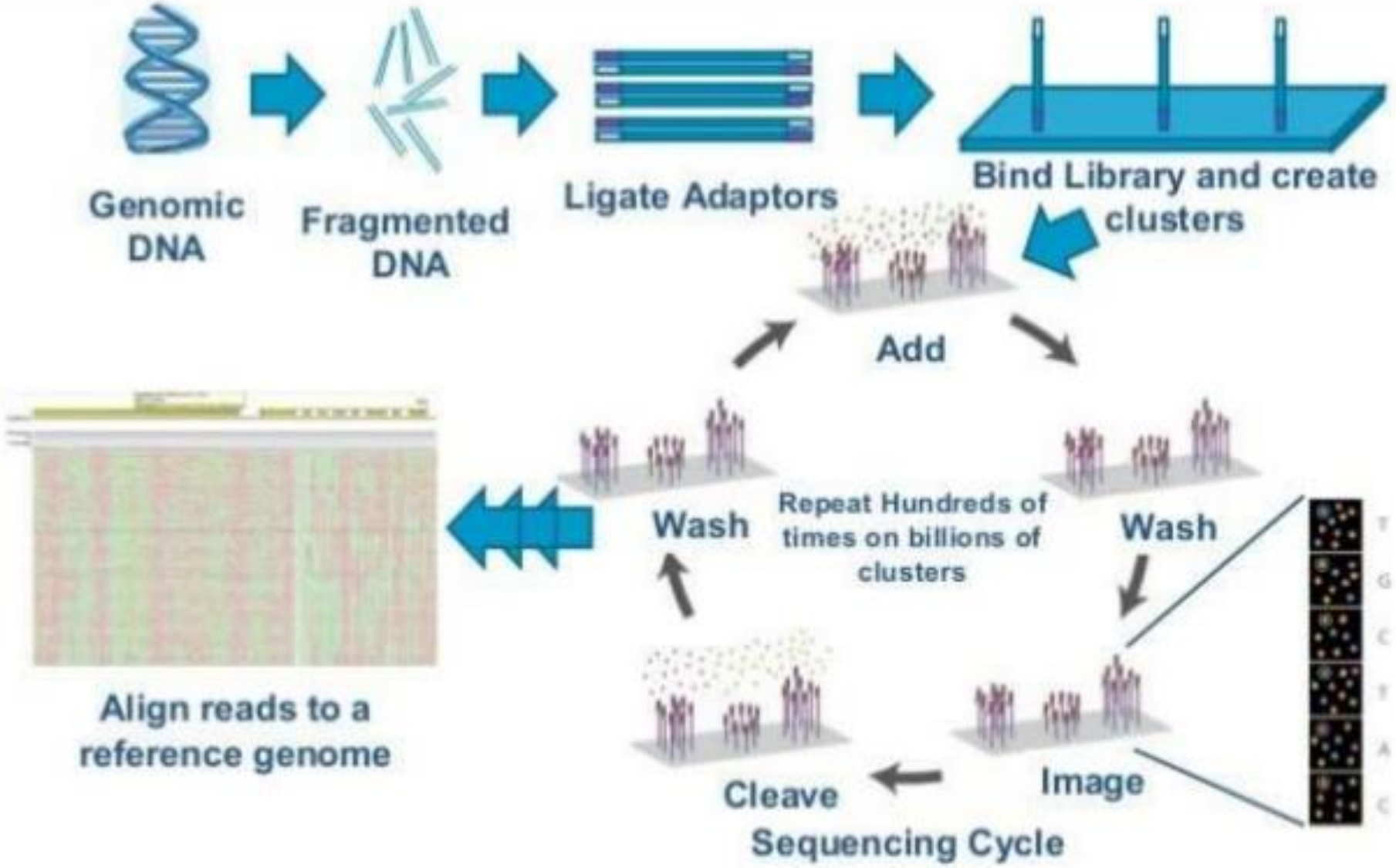
IILUMINA / SOLEXA SEQUENCING

- In NGS, vast numbers of short reads are sequenced in a single stroke.
- To do this, firstly the input sample must be cleaved into short sections. The length of these sections will depend on the particular sequencing machinery used.

IILUMINA / SOLEXA SEQUENCING

- In Illumina sequencing, 100-150bp reads are used. Somewhat longer fragments are ligated to generic adaptors and annealed to a slide using the adaptors. PCR is carried out to amplify each read, creating a spot with many copies of the same read. They are then separated into single strands to be sequenced.
- The slide is flooded with nucleotides and DNA polymerase. These nucleotides are fluorescently labelled, with the colour corresponding to the base. They also have a terminator, so that only one base is added at a time.

IILUMINA / SOLEXA SEQUENCING



454 SEQUENCING

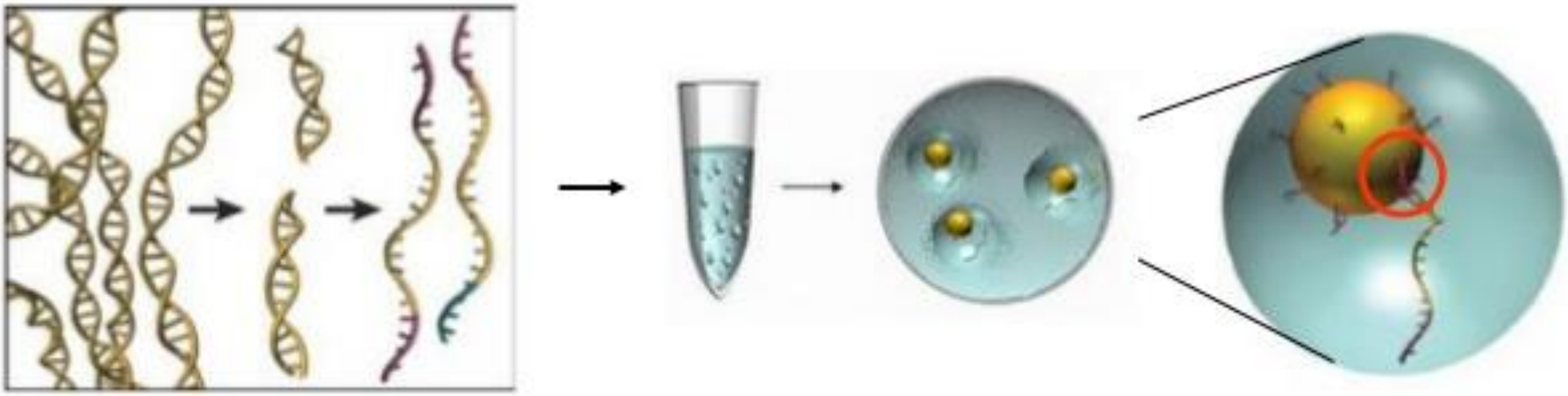
Generates more than 1,000,000 individual reads with improved Q20 read length of 400 bases per 10-hour instrument run

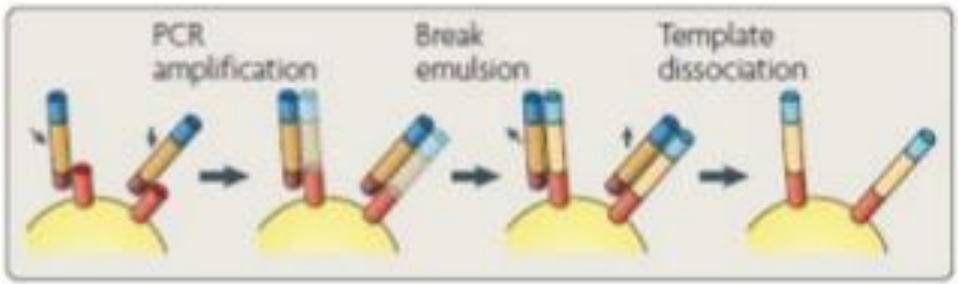
Methodology:

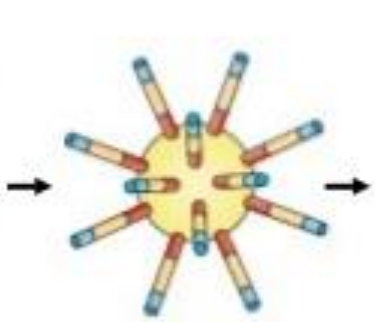
- Generation of a single- stranded template DNA library
- Emulsion-based clonal amplification of the library
- Data generation via sequencing-by-synthesis
- Data analysis using different bioinformatics tools

Roche/454 - Pyrosequencing

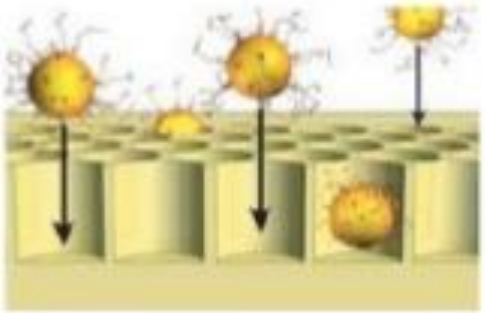
1. **Emulsion-based sample preparation (emPCR)**







Several thousand copies of the same template sequence on each bead



on average 1.6 million wells

Roche/454 - Pyrosequencing

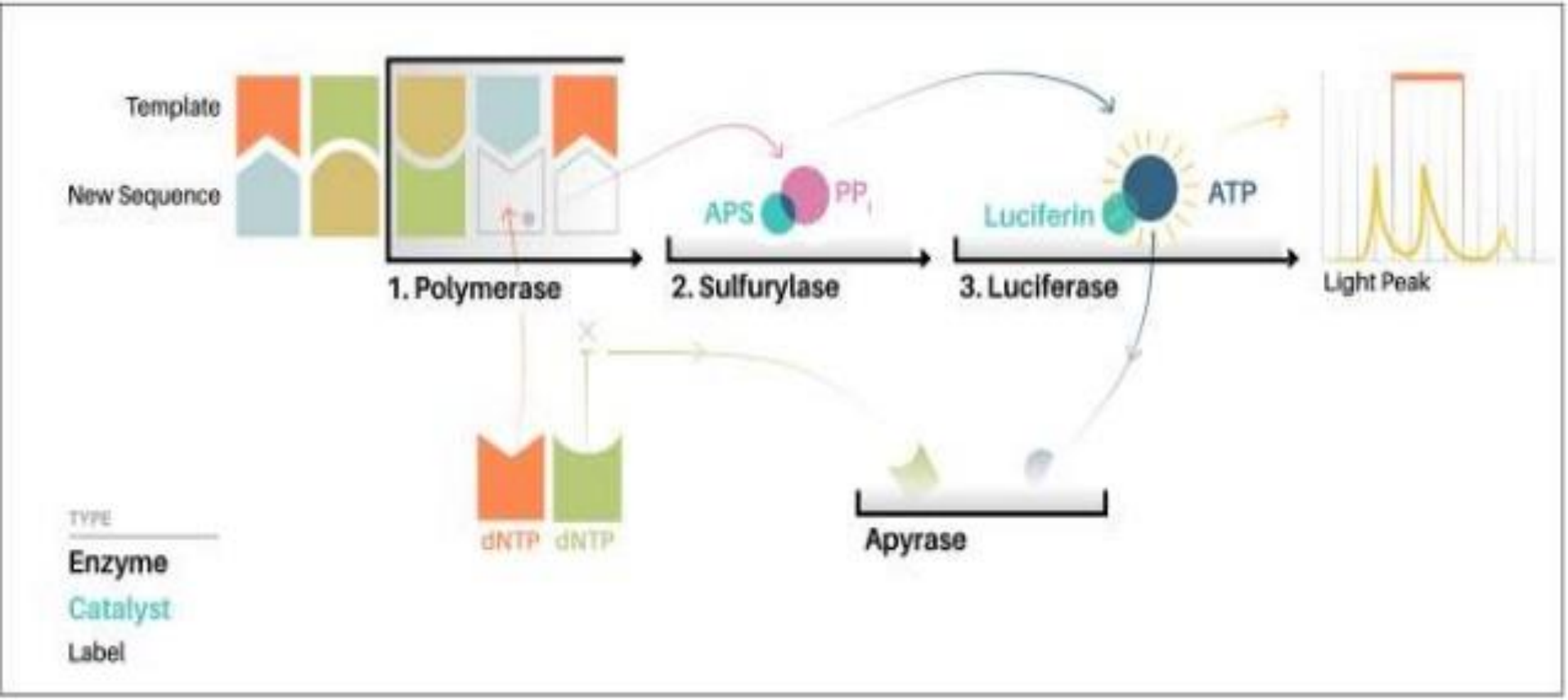
- **Applications**
  - Whole genome sequencing
  - Targeted resequencing
  - Sequencing-based Transcriptome Analysis
  - Metagenomics

- **Over 1300 publications...**

PYROSEQUENCING

- A method of DNA sequencing based on the "sequencing by synthesis" principle.
- It differs from Sanger sequencing, relying on the detection of pyrophosphate release (hence the name) on nucleotide incorporation, rather than chain termination with dideoxynucleotides.
- ssDNA template is hybridized to a sequencing primer
- Incubated with the enzymes DNA polymerase, ATP sulfurylase, luciferase and apyrase, and with the substrates adenosine 5´ phosphosulfate (APS) and luciferin.

PYROSEQUENCING



PYROSEQUENCING

- The addition of one of the four deoxynucleotide triphosphates (dNTPs)(in the case of dATP we add dATPαS which is not a substrate for a luciferase) initiates the second step.
- DNA polymerase incorporates the correct, complementary dNTPs onto the template.
- This incorporation releases pyrophosphate (PPi) stoichiometrically.

PYROSEQUENCING

- ATP sulfurylase quantitatively converts PPi to ATP in the presence of adenosine 5´ phosphosulfate.

- This ATP acts as fuel to the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP.

Advantages of Roche 454:

- Longer read lengths (700bp)
- Fast operation
- High accuracy(~99.9%)

Disadvantages of Roche 454:

- Error rate increases with the increase in the length of polybase
- High cost of sequencing
- Low throughput
- Low scalability

ABI SOLID Sequencing

AB SOLiD™ 3 System generates
over 20 gigabases and 400 M tags per run
(ca. 50 bp long)

Methodology:

- Library preparation
- Emulsion PCR/ Bead enrichment
- Bead deposition
- Sequencing by ligation
- Primer reset

Appication of SOLiD Sequencing

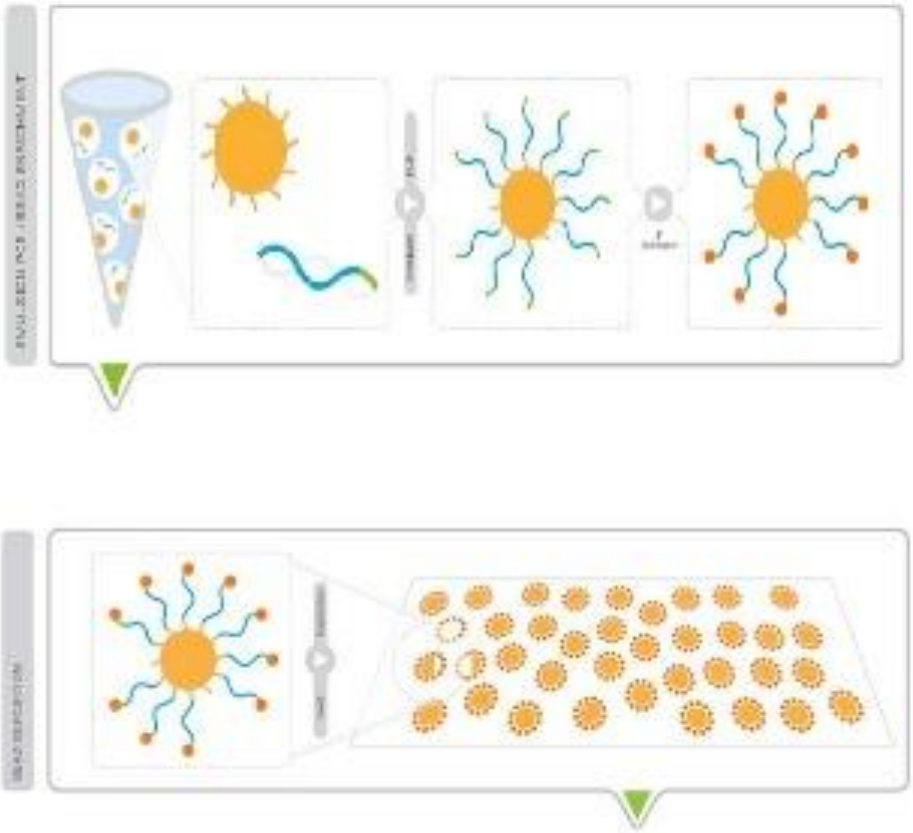
Application of SOLiD includes[1]:

1. Whole genome sequencing
2. Targeted sequencing
3. Transcriptome research (including gene expression profiling, small RNA analysis, and whole transcriptome analysis)
4. Epigenome (like ChIPSeq and methylation) analysis.

SOLiD

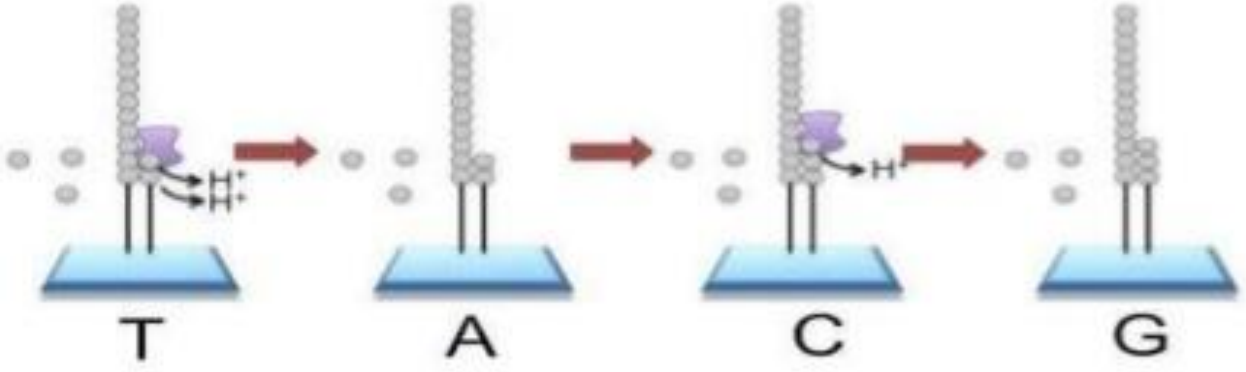
(support oligonucleoti

- Sequencing by Oligo/Ligation and Detection.
- Steps
  - **Library Preparation**
    - two types of libraries sequencing-fragment or mate-paired are prepared.
  - **Emulsion PCR/Bead Enrichment**
    - amplification of template fragments is done in same manner as 454.
  - **Bead Deposition**
    Deposit 3' modified beads onto a glass slide.

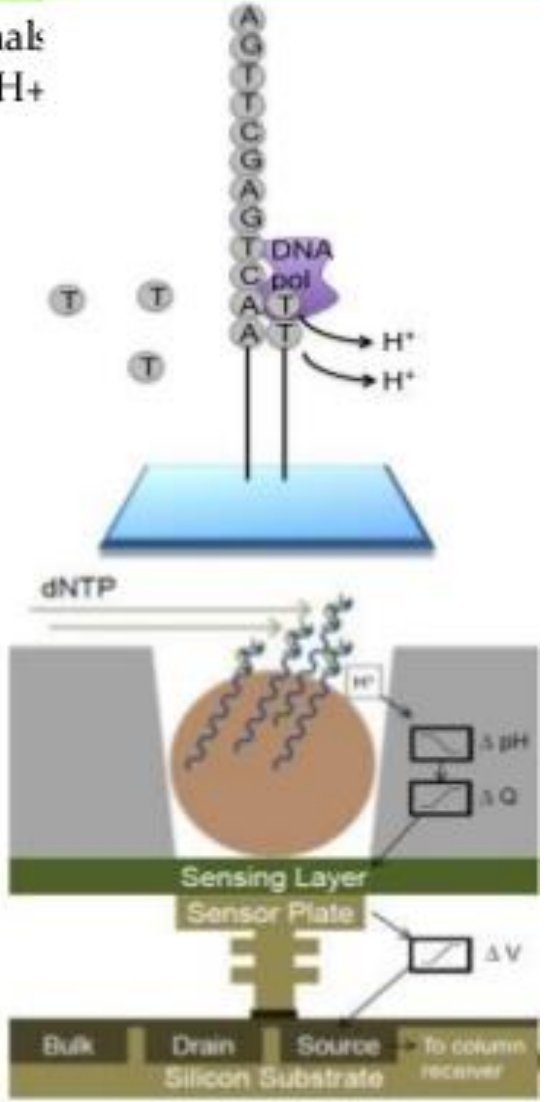


ION TORRENT SEQUENCING

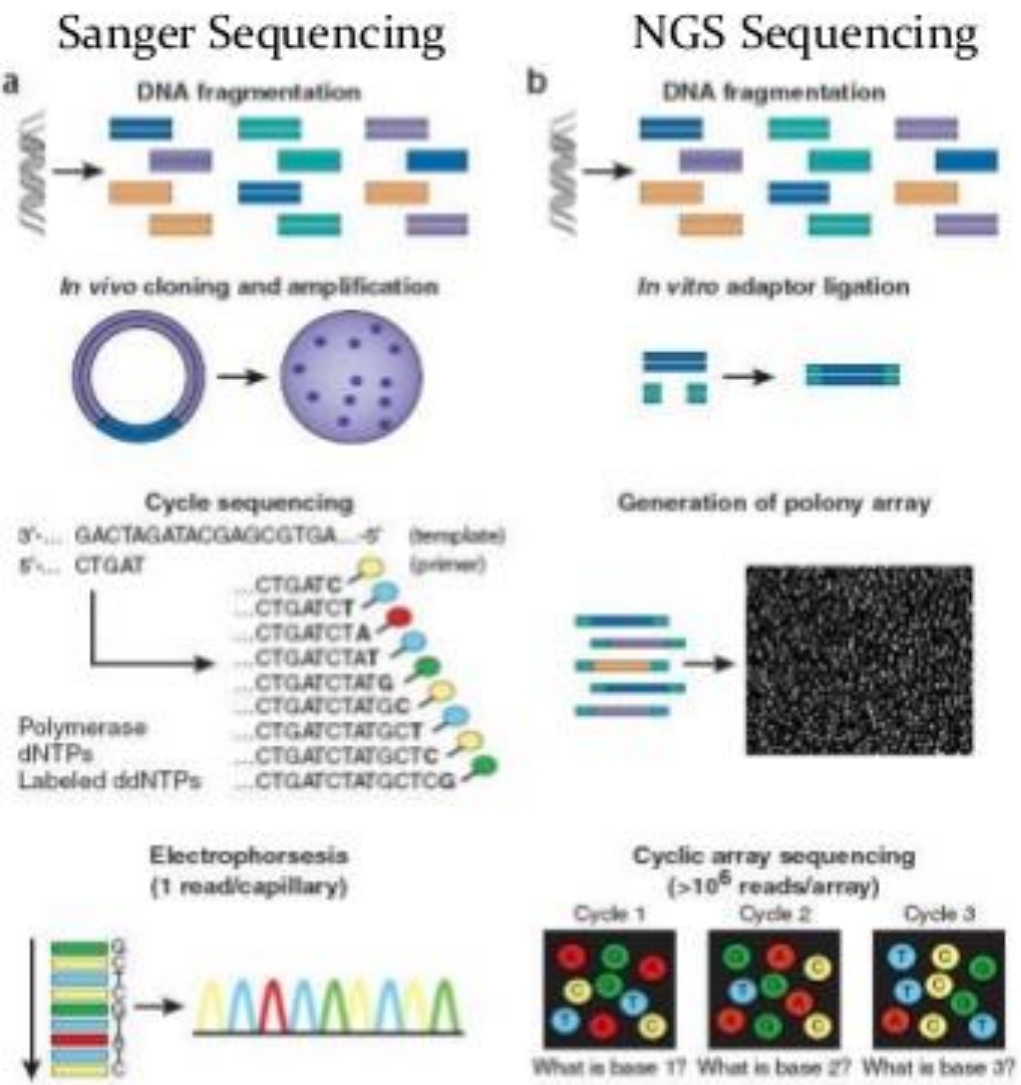
- Ion torrent and ion proton sequencing do not make use of optical signals exploit the fact that addition of a dNTP to a DNA polymer releases an H+
  - Run time: 3 h; no termination or deprotection steps
  - Clustal Amplification- Emulsion PCR
  - Read length: 100–300 bp
  - Throughput determined by chip size : 10Mb – 5 Gb
  - Cost: $1–$20/Mb



The pH change, if any, is used to determine how many bases (if any) were added with each cycle.



ADVANTAGES OF NGS



- No *invivo* cloning, Transformation, Colony picking
- High degree of Parallelism then Capillary Sequencing
- Low Reagent Cost
- Reduced Sample Size
- Less Time

Thank you