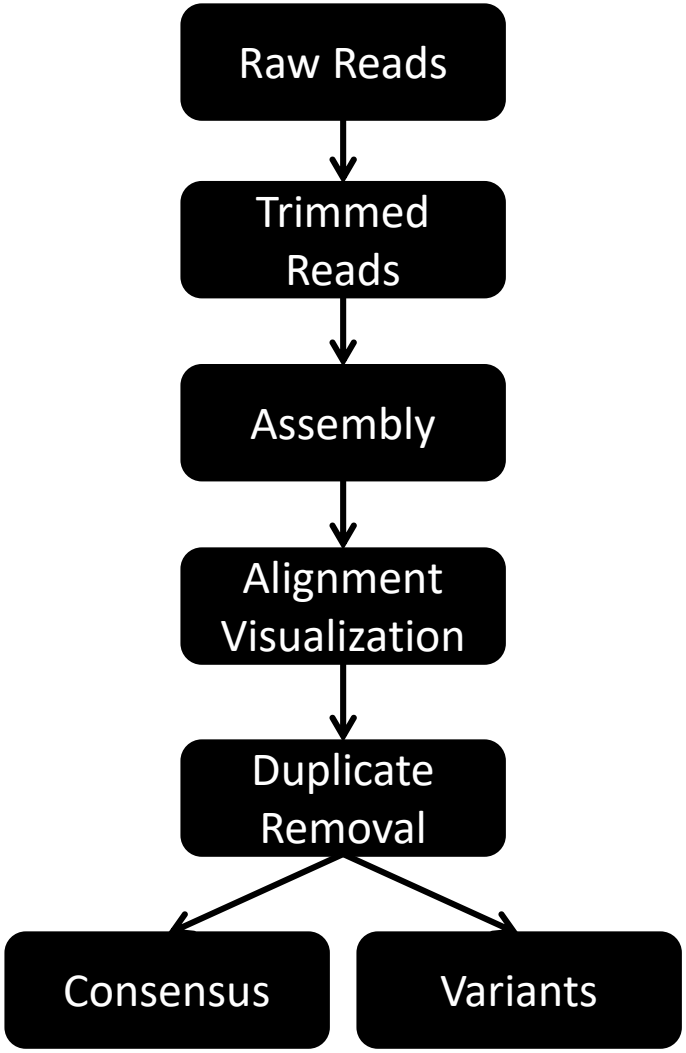
Introduction to Sequencing

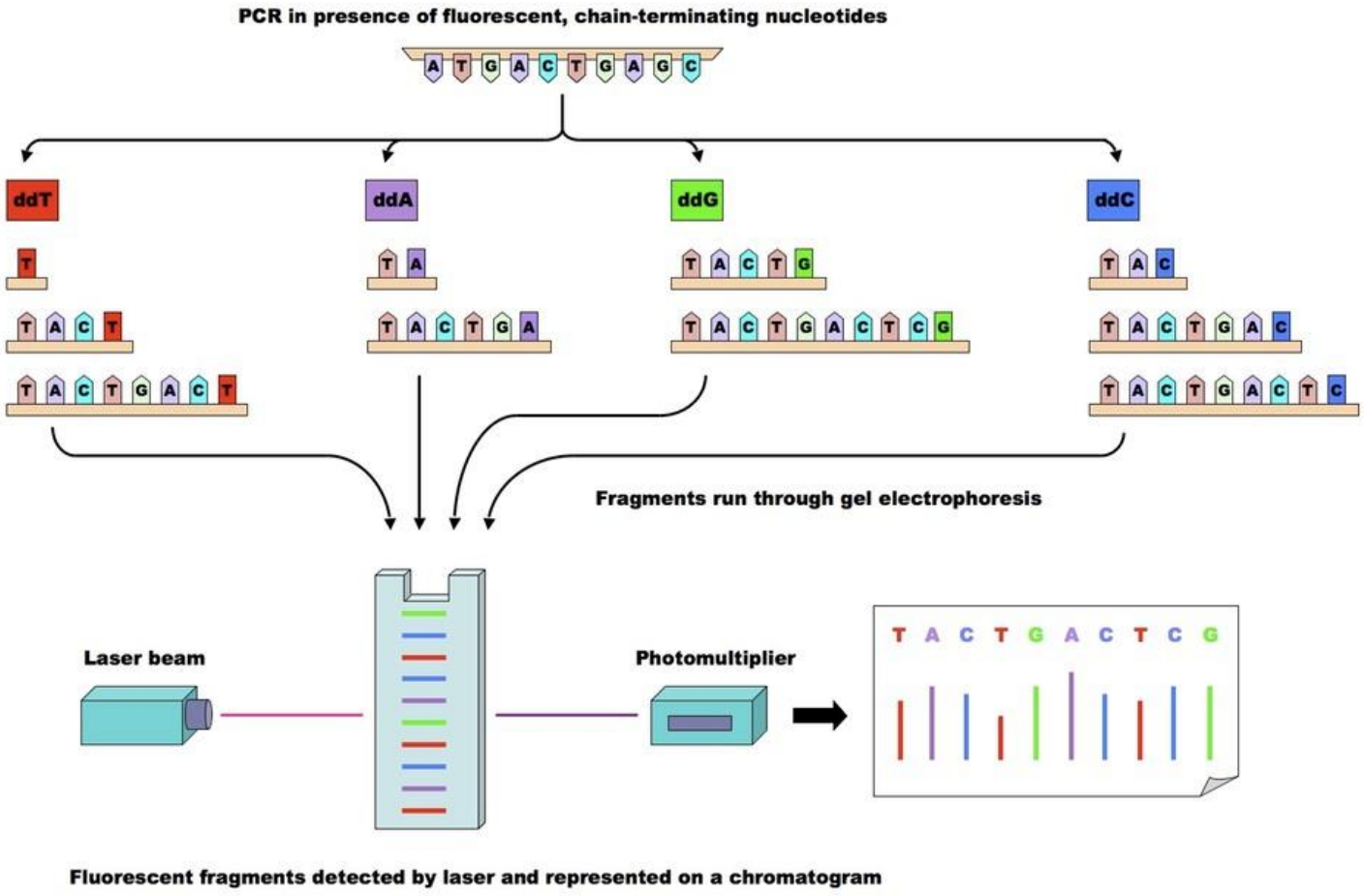
Outline of the Course



Brief History of Sequencing

- **1953** – Structure of DNA
- **1965** – Robert Holley sequenced Alanine tRNA
- **1975** – Sanger used plus minus system used to sequence ΦX174.
- **1977** – Maxam and Gilbert introduced the Chemical Clevage method of sequencing
- **1977** – Sanger introduced the Chain Termination or Dideoxy sequencing method
- This was improved on further to give the ABI sequencers that became the gold standard of First Generation technologies.

First Generation Sequencing

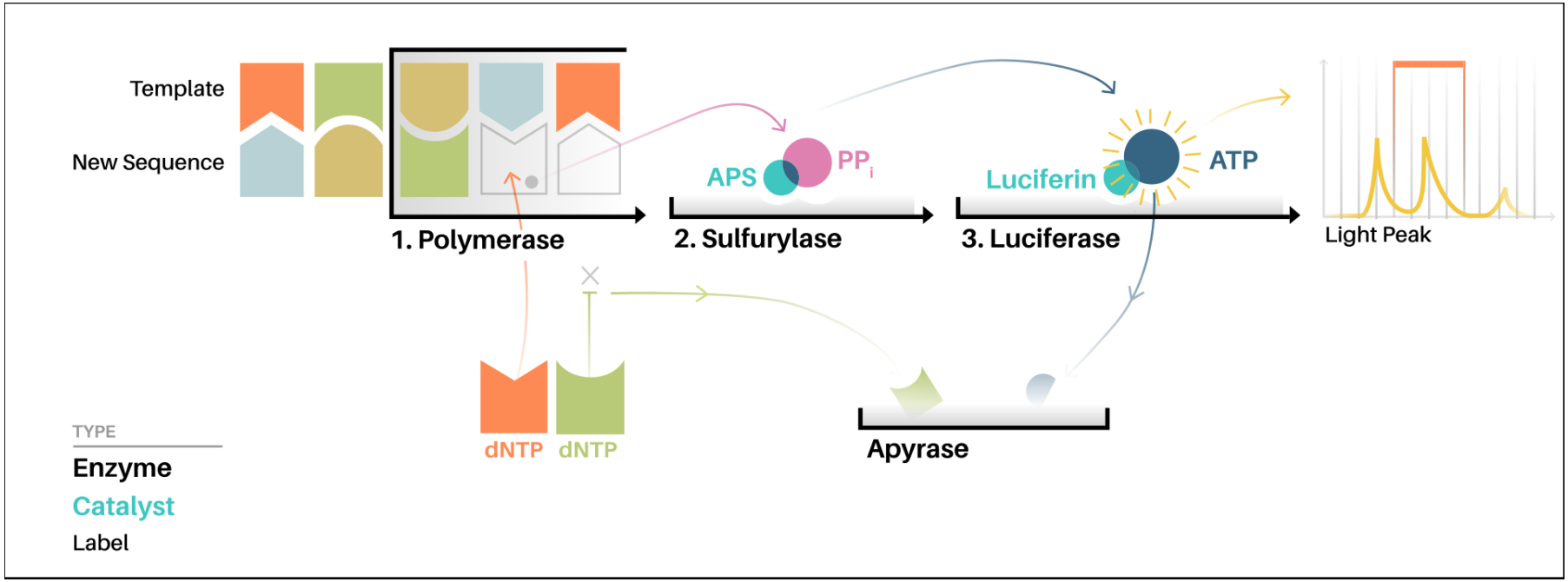


http://www.vce.bioninja.com.au/aos-3-heredity/molecular-biology-technique/sequencing.html

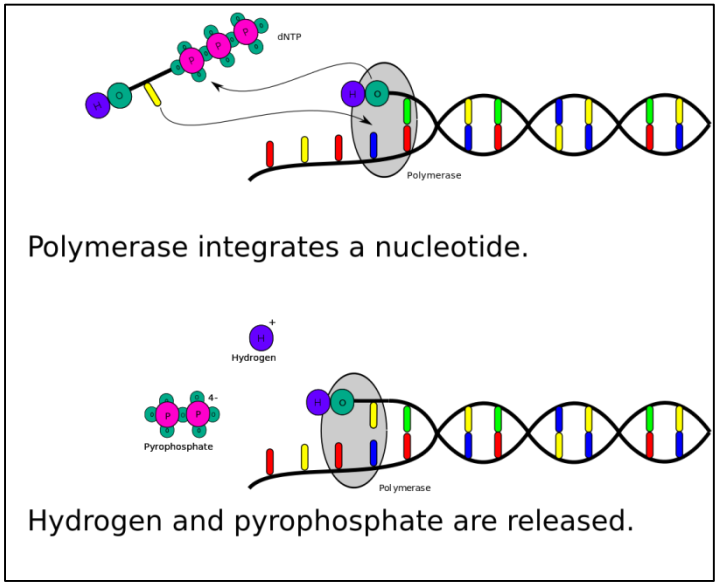
Second Generation Sequencing

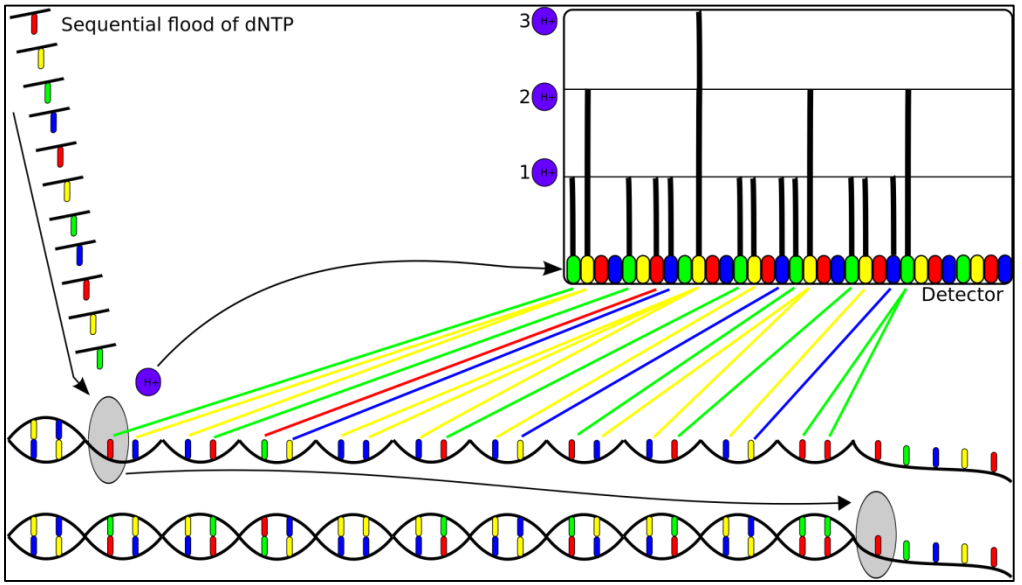
- **Sequencing by Synthesis (SBS)**
  - **Pyrosequencing (Roche 454)**
  - **Ion Semiconductor Sequencing (Ion Torrent)**
  - **Dye Based Sequencing (Illumina)**

Pyrosequencing (Roche 454)

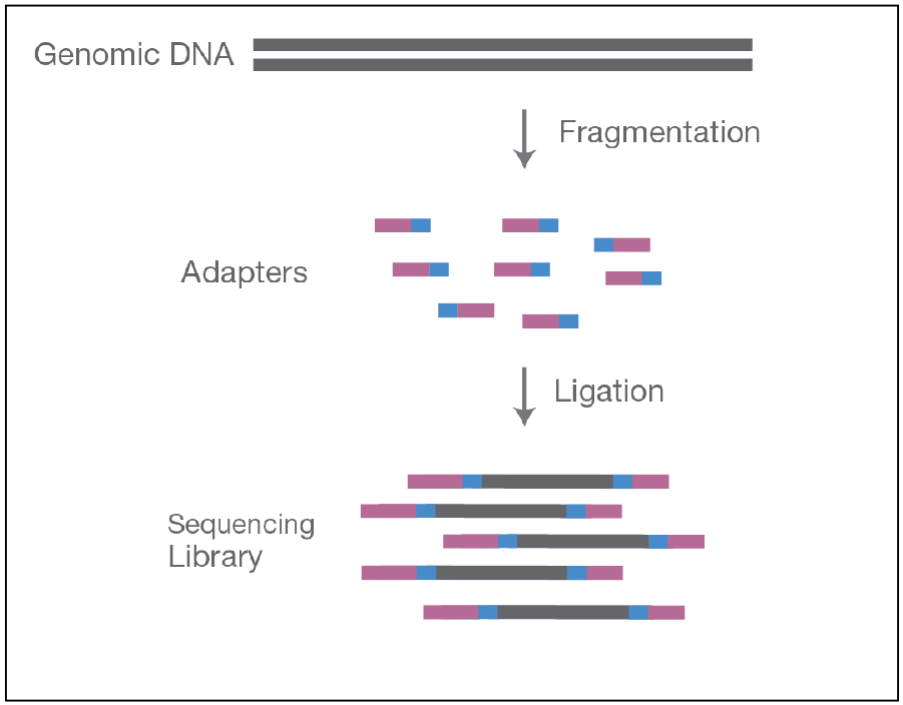


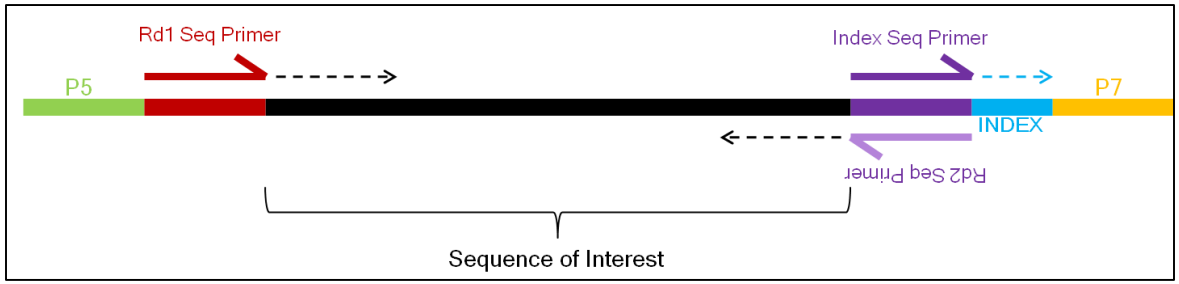
Ion Semiconductor Sequencing (Ion Torrent)



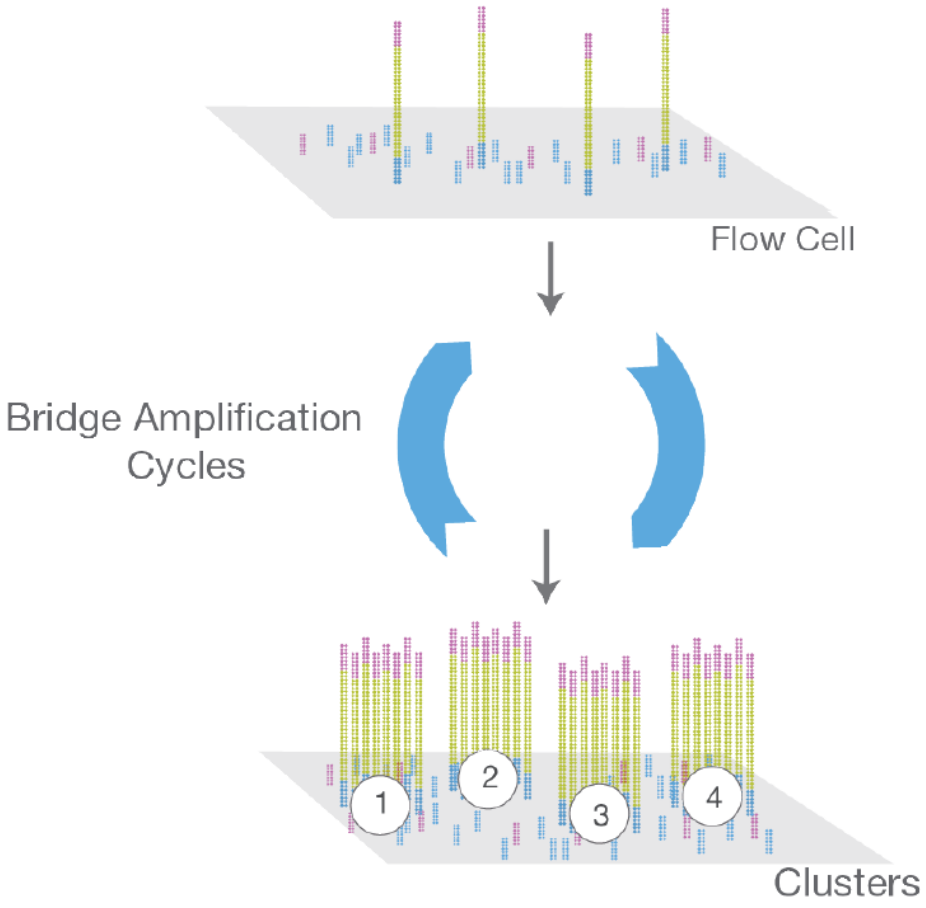


Dye Based Sequencing (Illumina) - Library

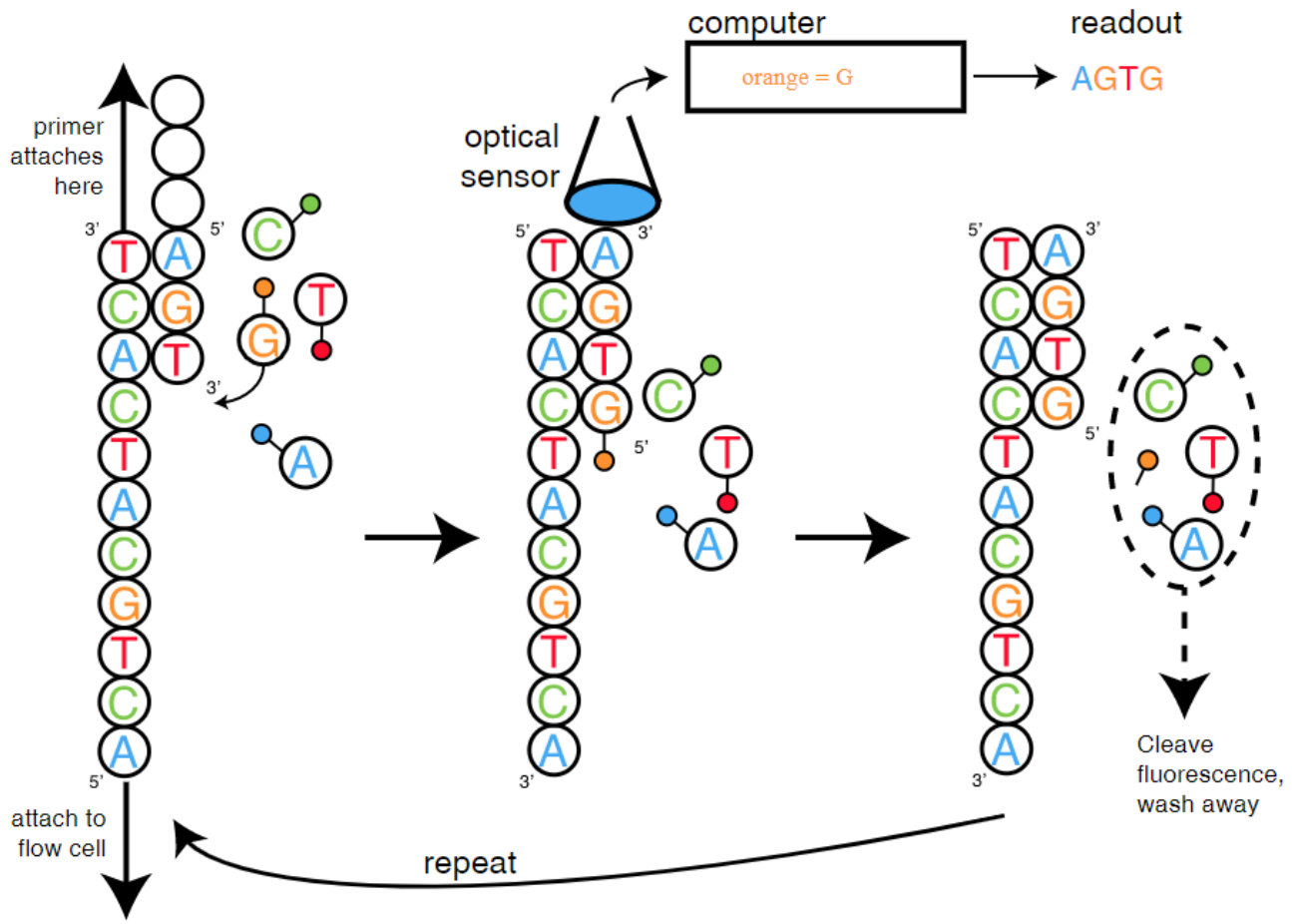




Dye Based Sequencing (Illumina) - Clusters

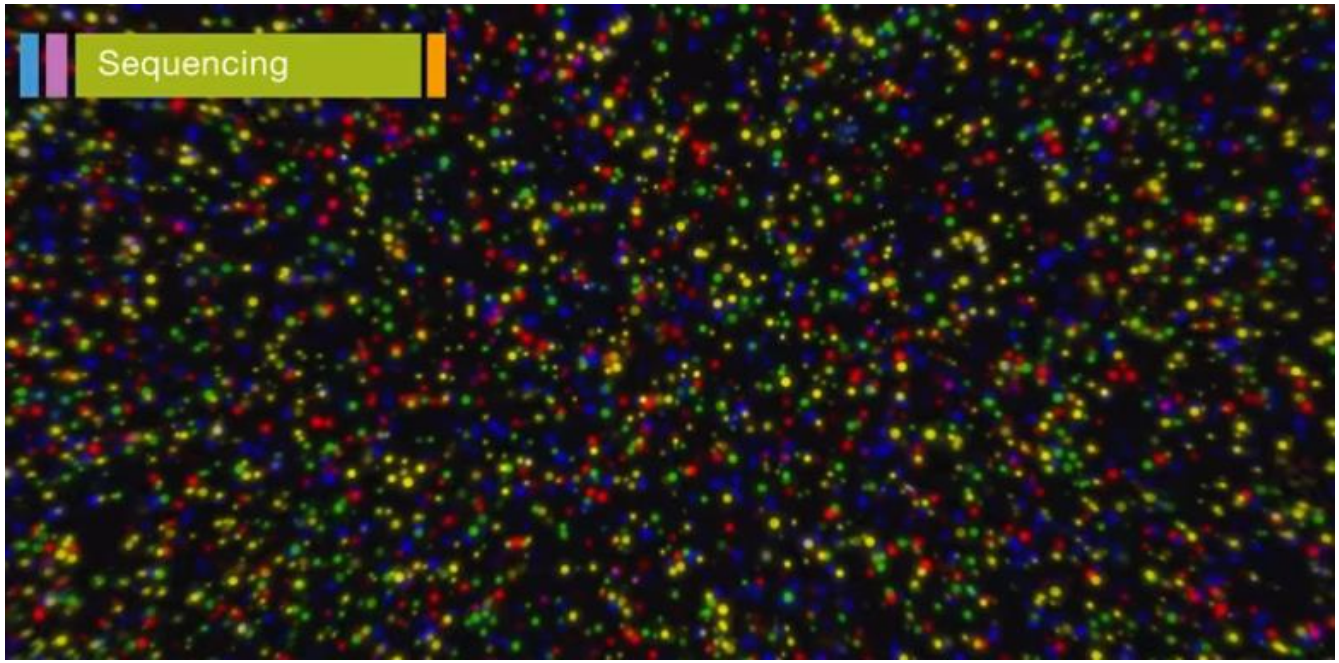


Dye Based Sequencing (Illumina) - Sequencing



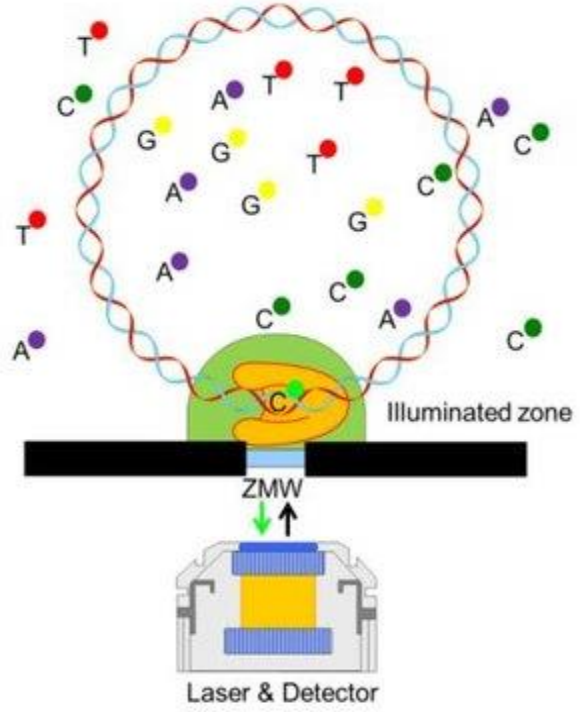
https://binf.snipcademy.com/lessons/ngs-techniques/illumina-solexa

Dye Based Sequencing (Illumina) - Sequencing

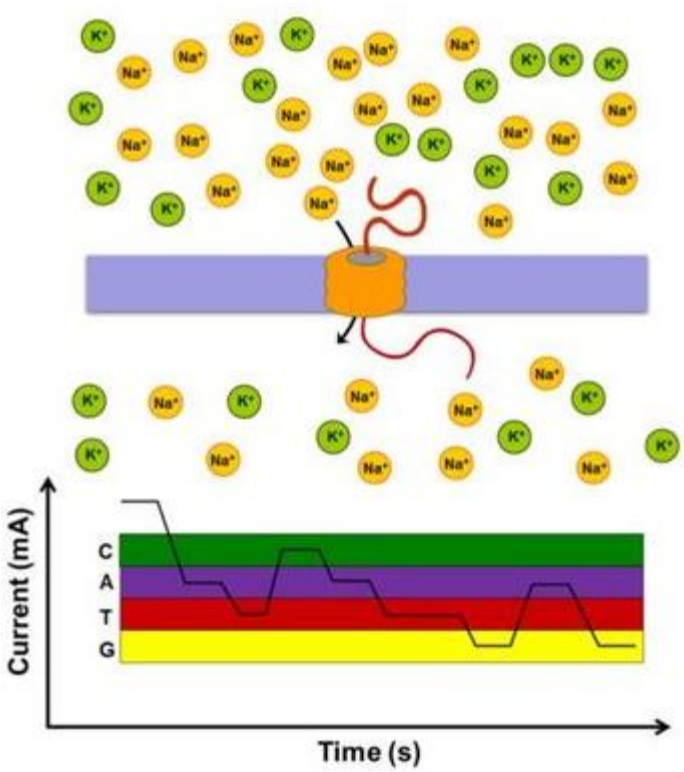


This data is stored in the Binary **BCL** files in real time

Third Generation Sequencing



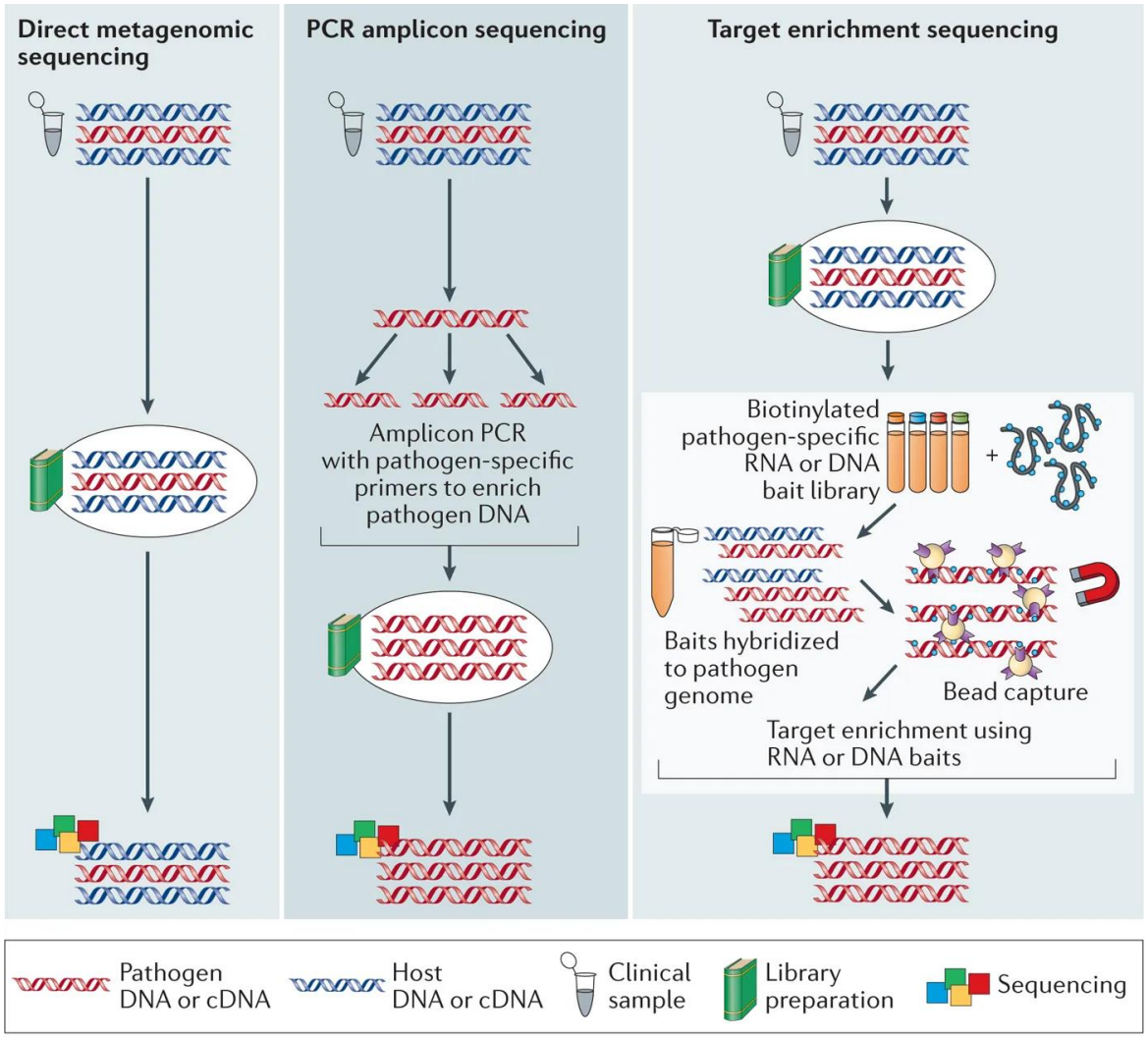
PacBio



Oxford Nanopore (MinIon)

https://www.researchgate.net/figure/Illustration-of-the-main-sequencing-techniques-A-Sanger-sequencing-B-Illuminas_fig5_331770390

Sequencing Strategies

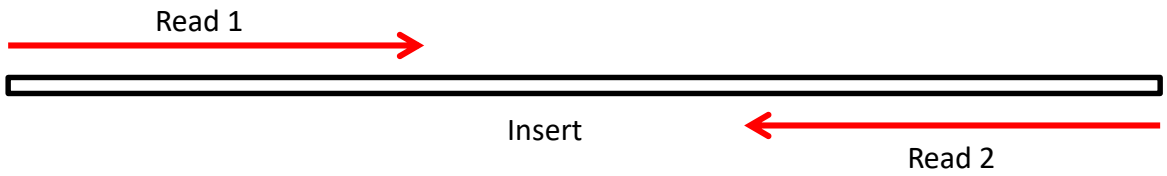


https://doi.org/10.1038/nrmicro.2016.182

Illumina Data

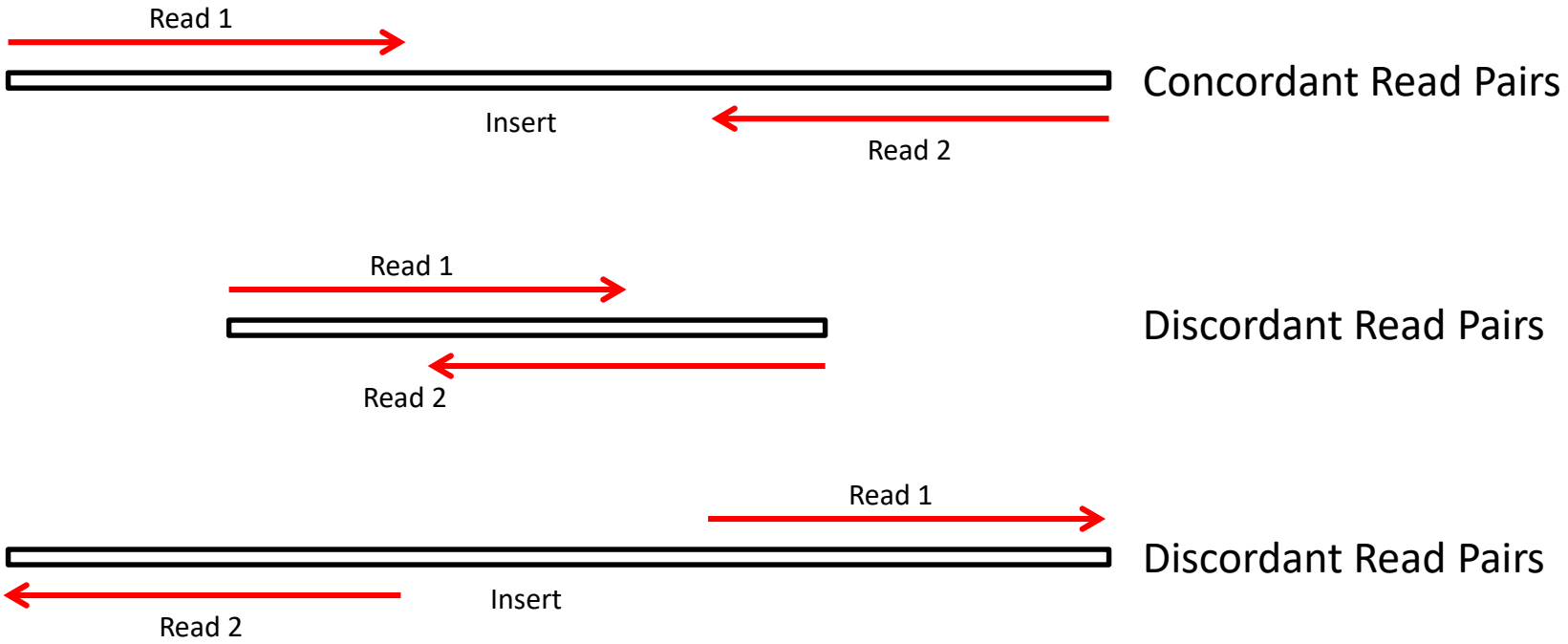
- BCL data (binary format) is converted to FASTQ (text).
- MiSeq and MiniSeq converts to FASTQ in the machine.
- Other Illumina platforms conversion can be done on Basespace or using Illumina bcl2fastq tool.
  (https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html)
- FASTQ files are used for all further downstream data analysis.

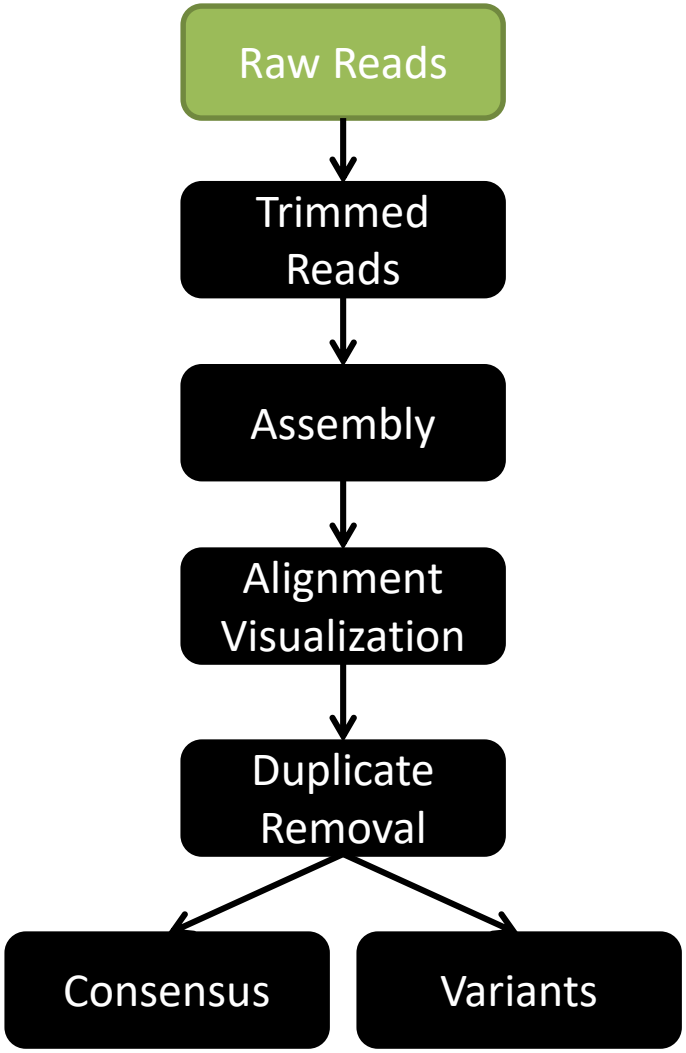
Illumina Data Paired End



- Two files produced R1 and R2
- Same number of reads in each file
- The ordering of reads is produced based on the scan of lanes/tiles

Illumina Data




Raw Reads
Trimmed Reads
Assembly
Alignment Visualization
Duplicate Removal
Consensus
Variants