

Určování sekvence DNA

Od manuálních po plně automatické metody sekvenování

1st generation:

Maxam-Gilbert (čtyři štěpící reakce - G, AG, C, TC)

Sangerova metoda (terminace syntézy druhého vlákna)



Next-generation:

454/Roche (pyrosekvence, PP se chemicky vaze na emisi svetla)

SOLiD/LifeTechnologies (ligace fluorescenčních oligonukleotidů)

Solexa/Illumina (polymerace fluorescenčních nukleotidů)

Complete Genomics (ligace k nanoshlukům DNA)



3rd generation:

Single-molecule sequencing:

Helicos (sekvenuje jednotlivé molekuly díky lepší optice)

SMRT/Pacific Biosciences (jamky pro jednu polymerázu)

IonTorrent (semiconductor sequencing)

Nanopore sequencing:

Oxford Nanopore

NABSys

NobleGen

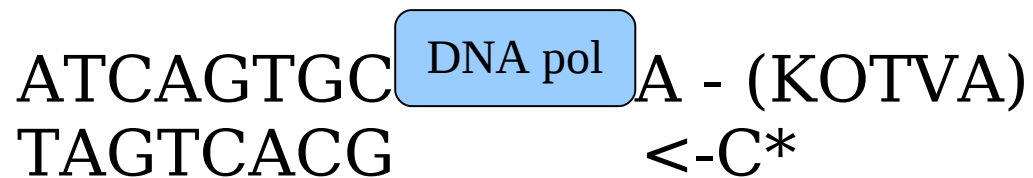


NOBLEGEN



SEKVENACE DNA - PŘEHLED METOD

polymerizace:

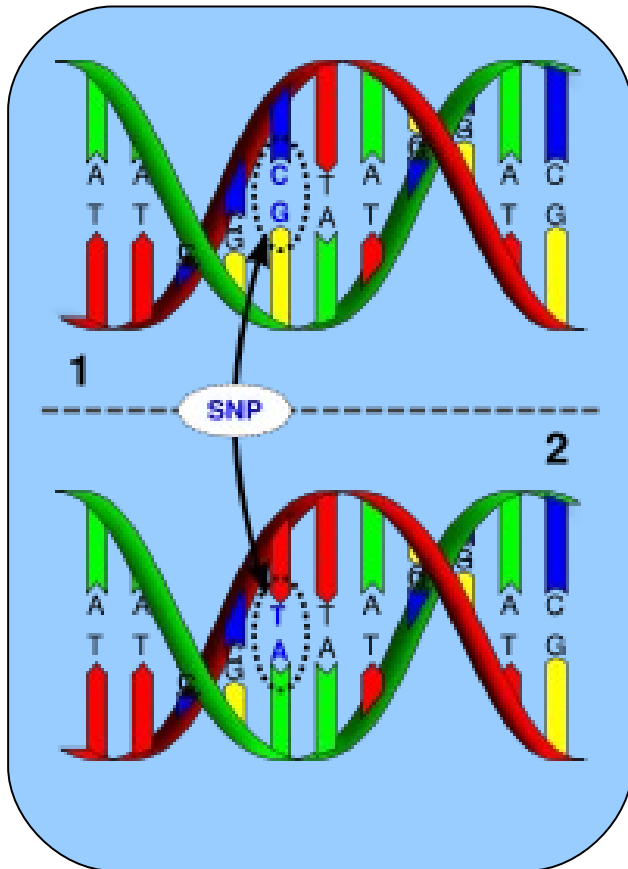


hybridizace:



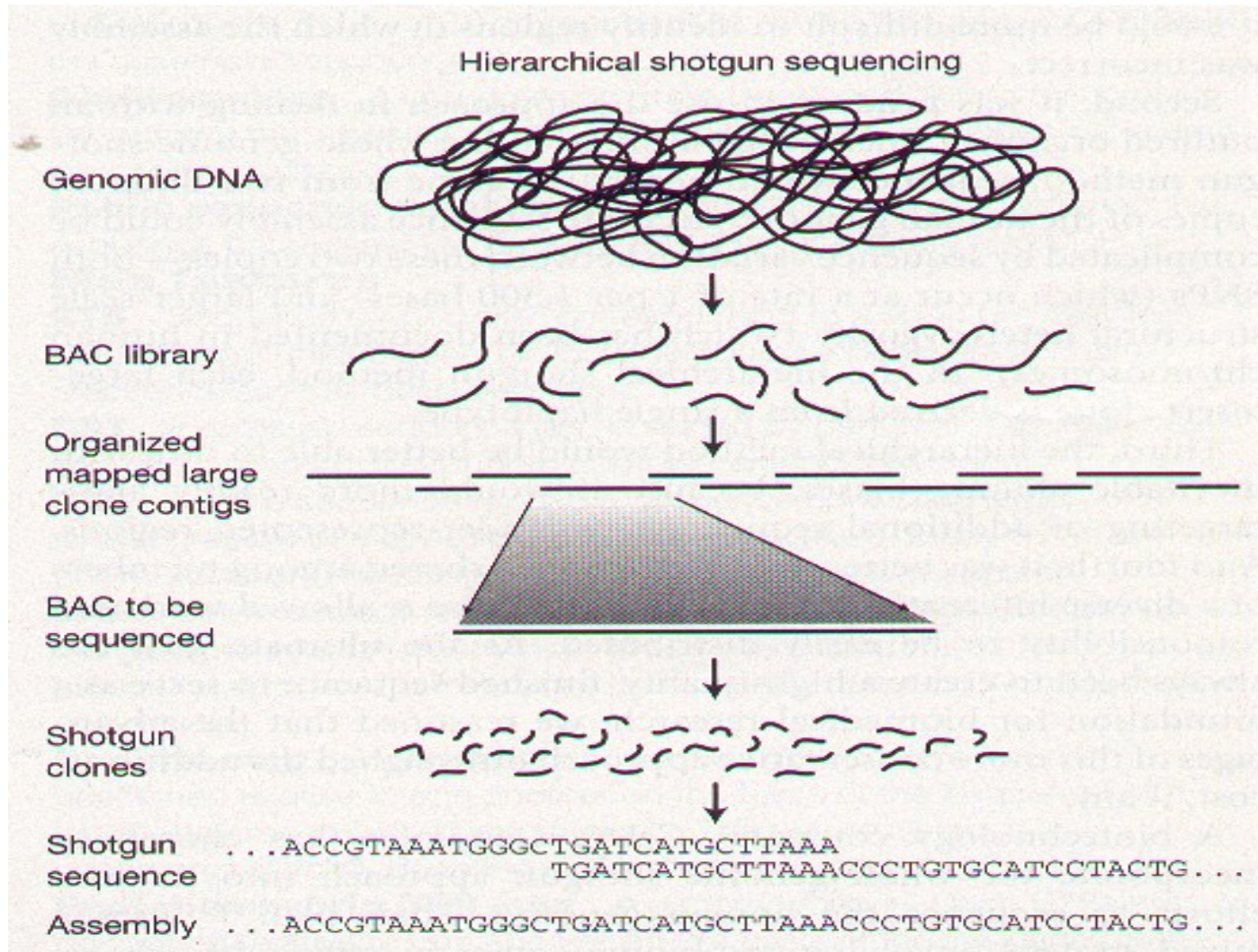
ligace:



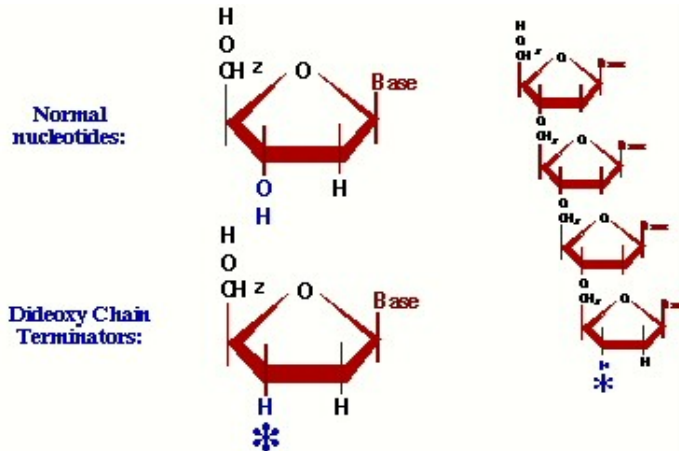
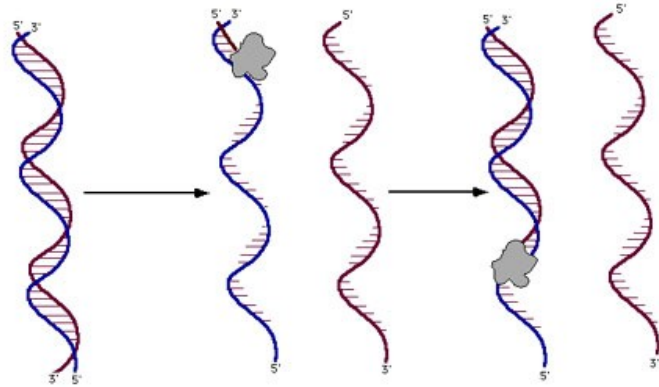


- de novo sekvenování
 - organizmy
 - populace
- resekvenování
- detekce polymorfizmu
 - SNP
 - přeskupení
- zjišťování metylace
- měření exprese (náhrada microarray)

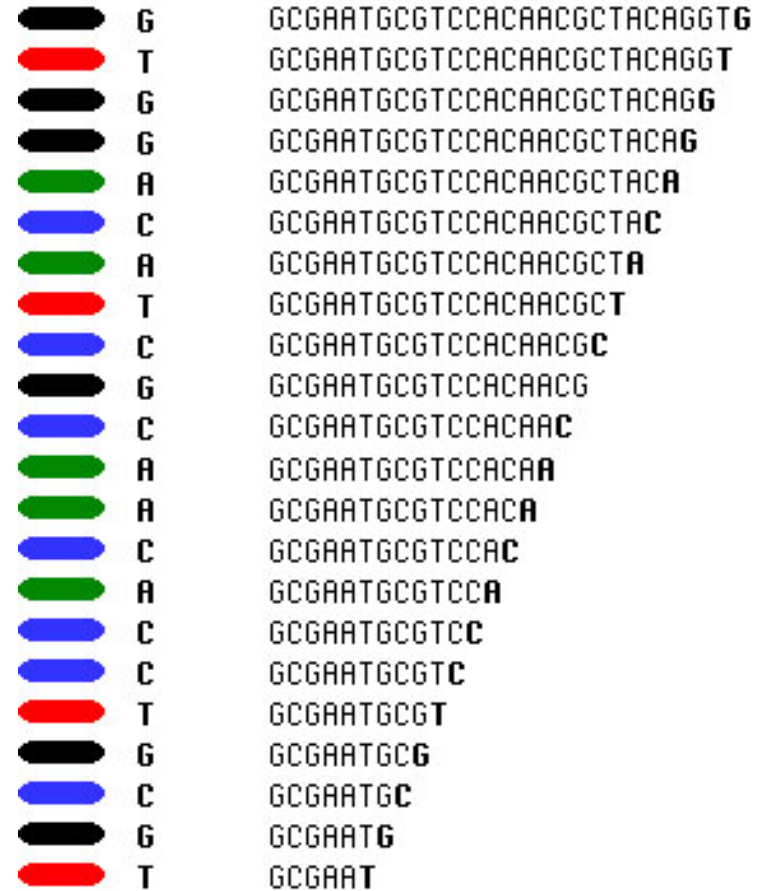
Tradiční způsoby sekvenace



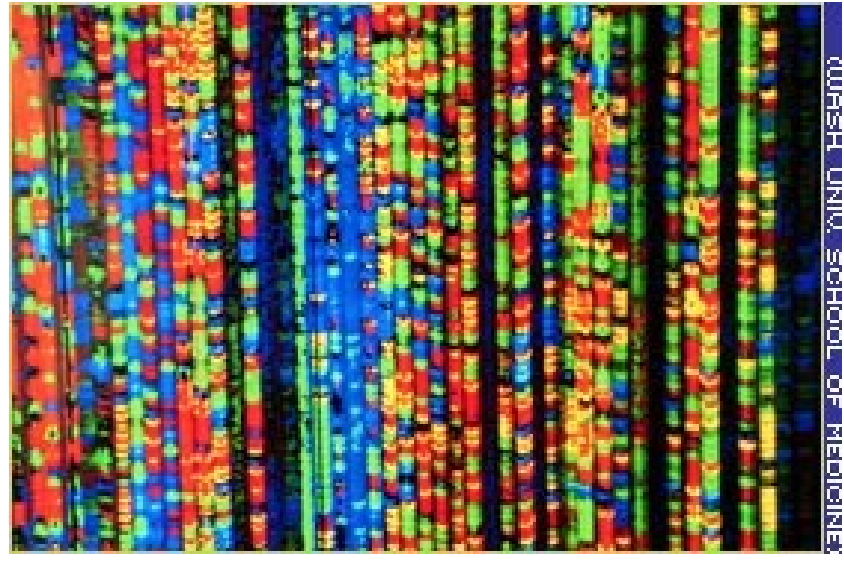
Sanger



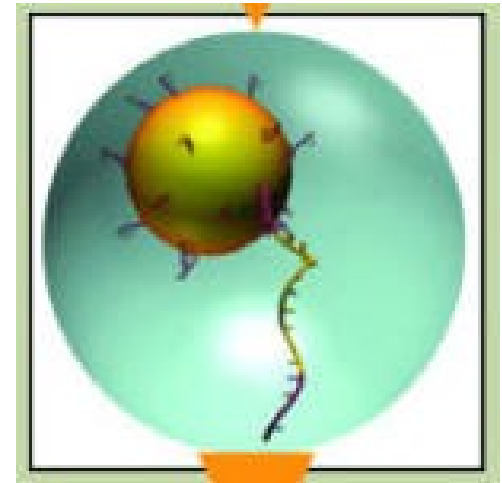
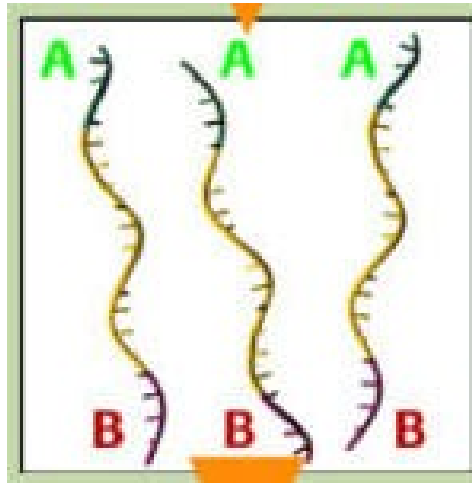
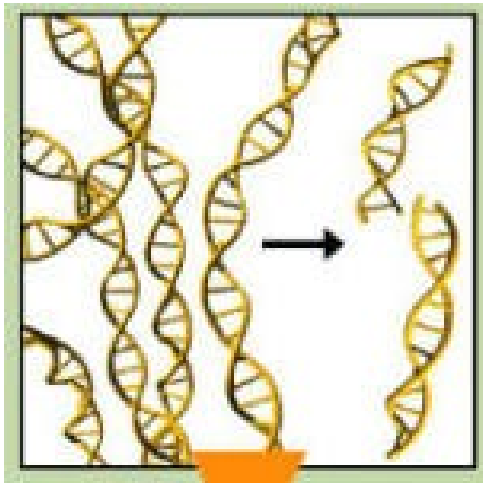
Gel:



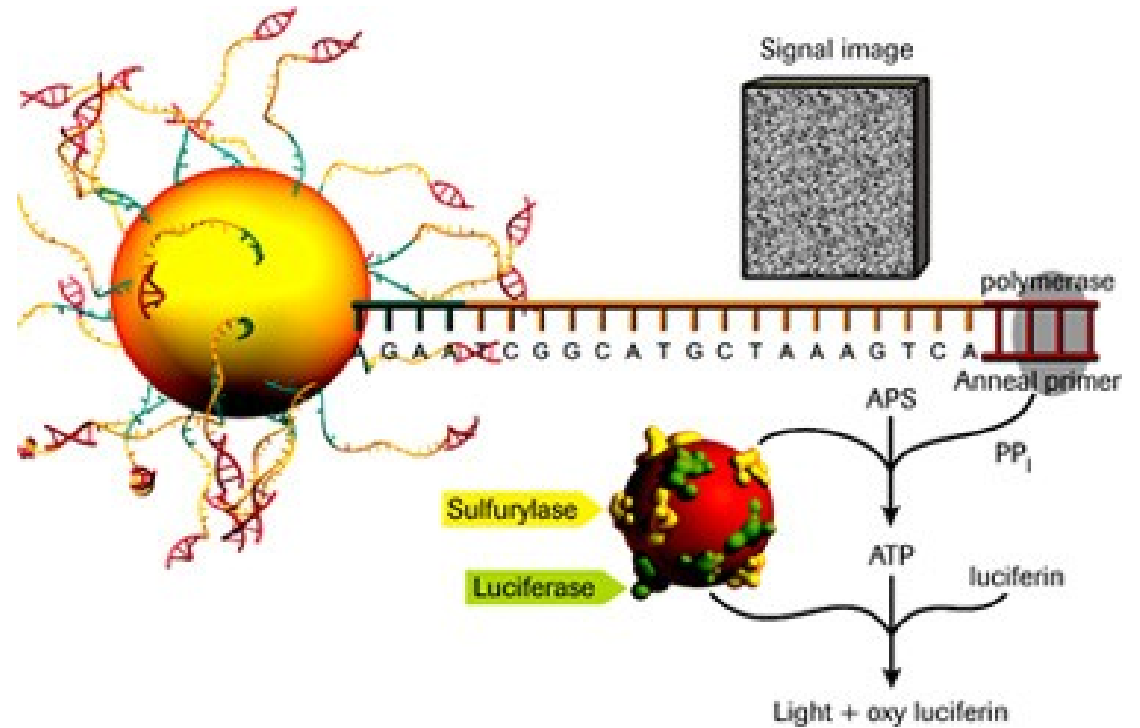
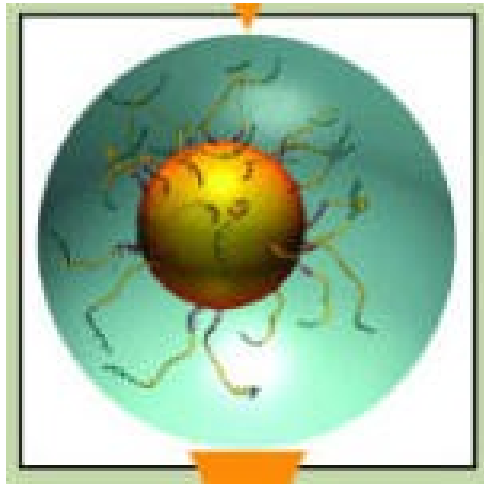
Sanger



454 sequencing (pyrosequencing)



454 sequencing (pyrosequencing)

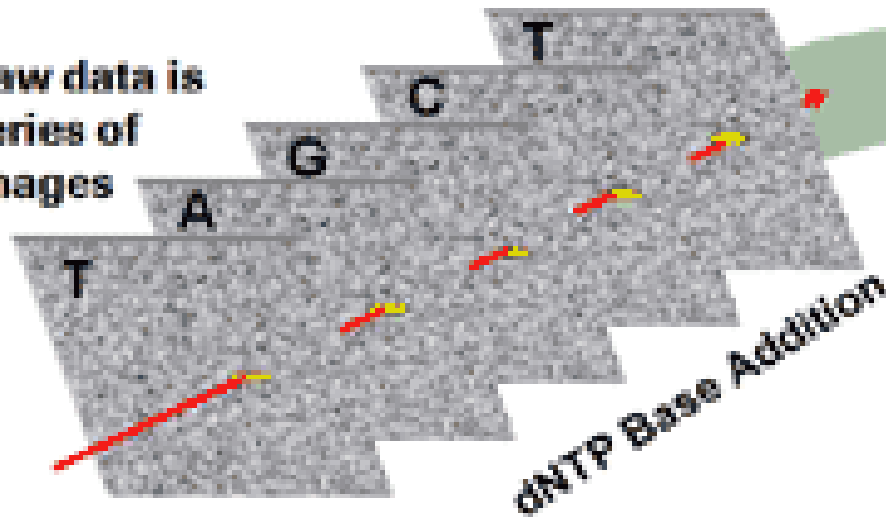


454 sequencing (pyrosequencing)

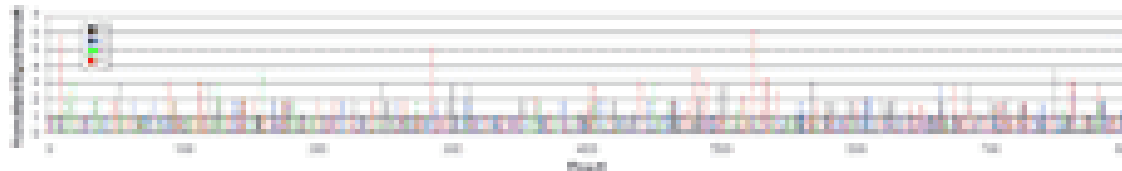
GS FLX Data

Image Processing Overview

1. Raw data is series of images

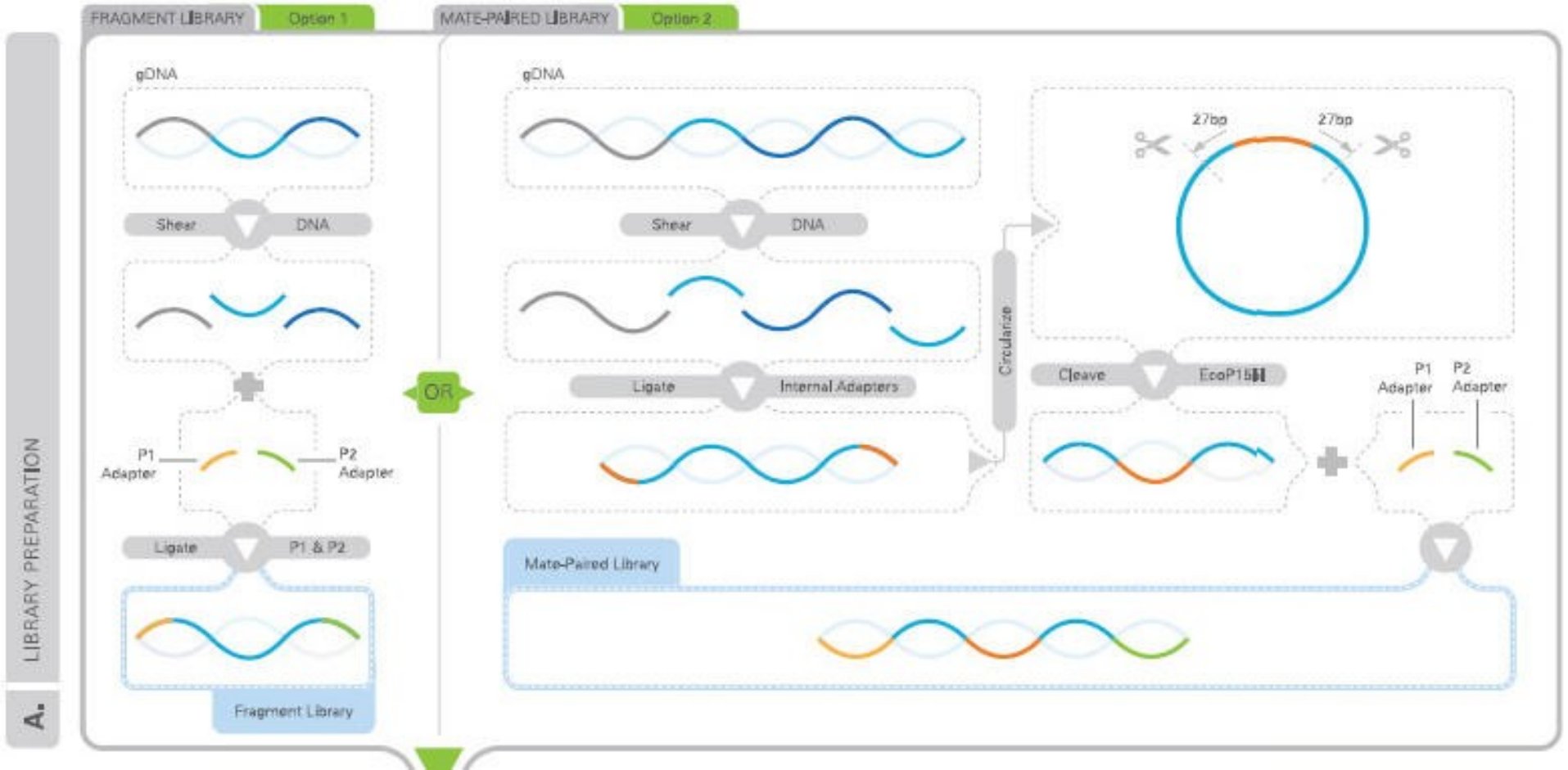


2. Each well's data extracted, quantified and normalized

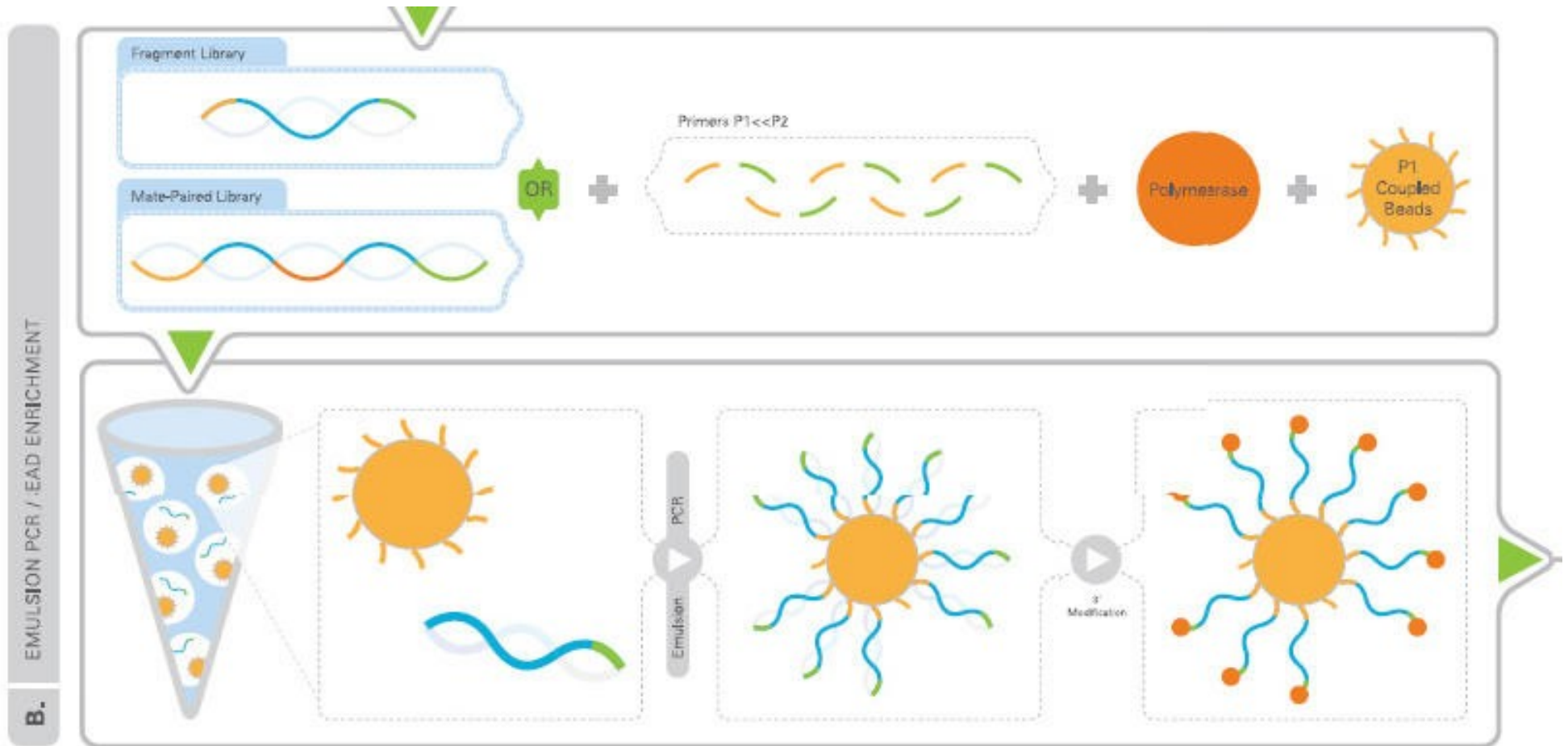


3. Read data converted into "flowgrams"

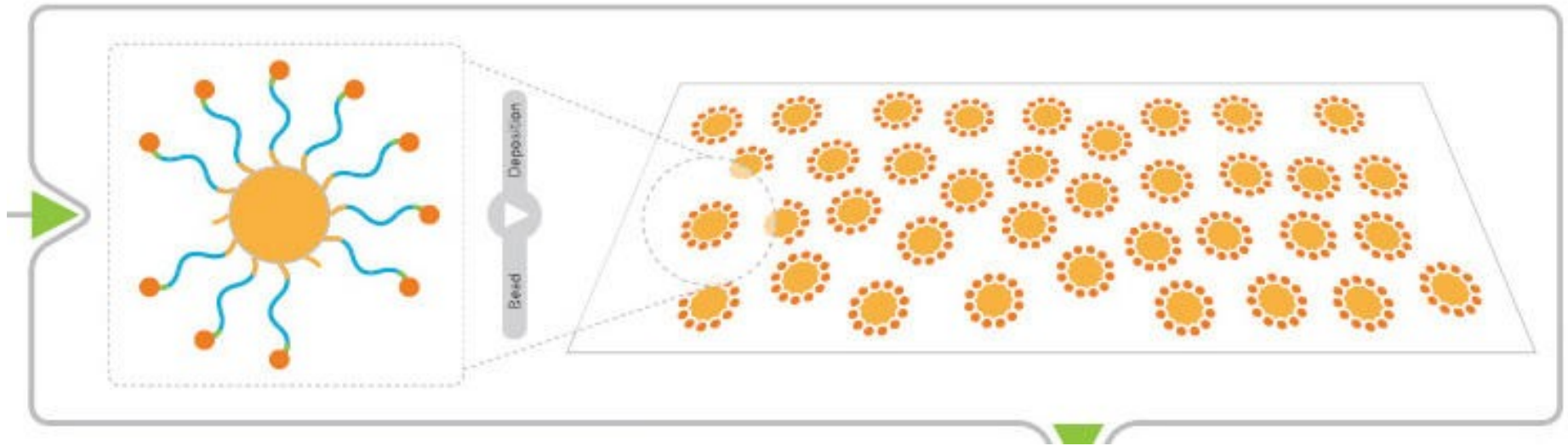
SOLiD (seq by ligation)



SOLiD (seq by ligation)



SOLiD (seq by ligation)



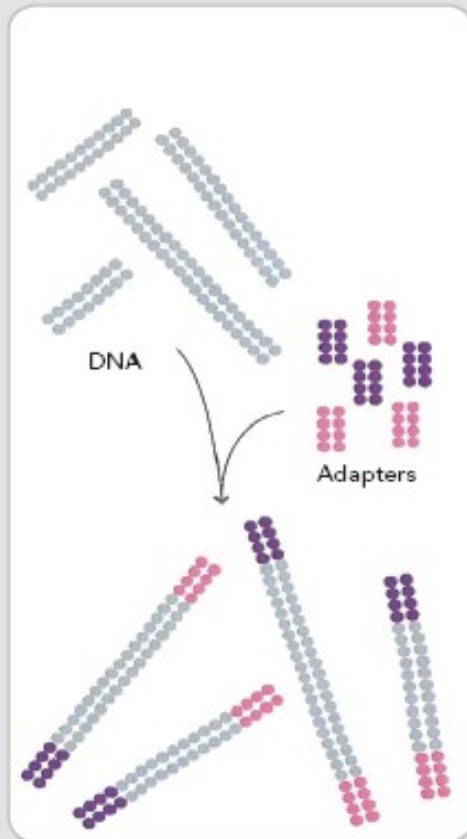
C.
BEAD DEPOSITION

Illumina



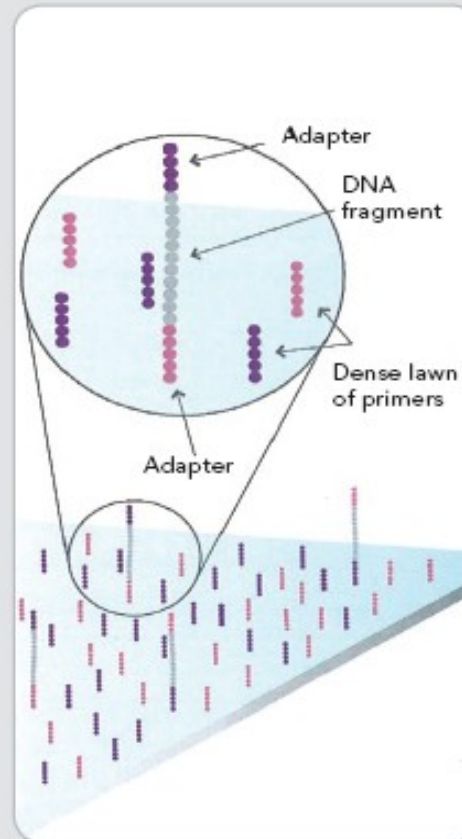
Illumina

1. PREPARE GENOMIC DNA SAMPLE



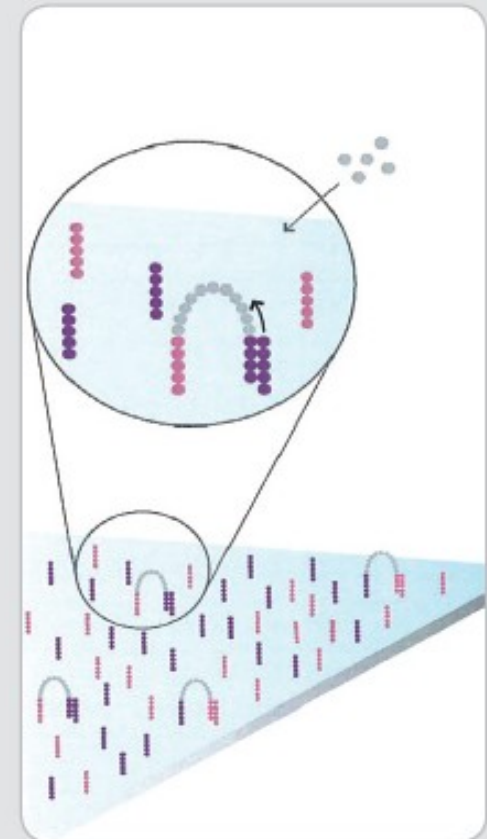
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE



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

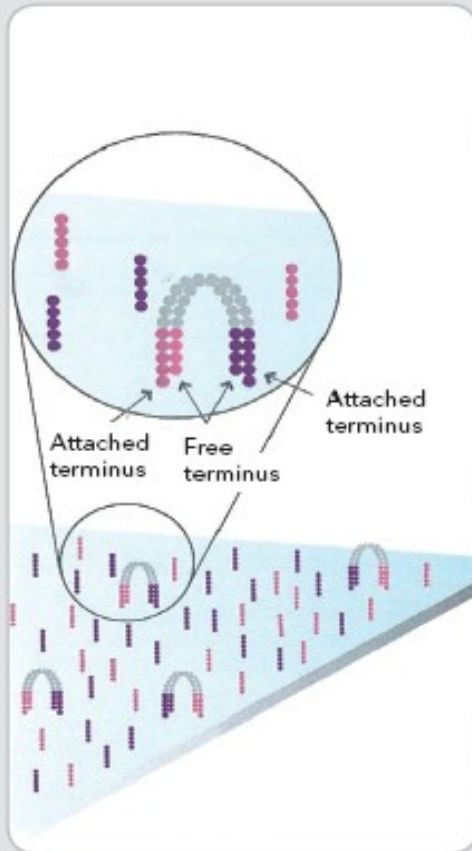
3. BRIDGE AMPLIFICATION



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

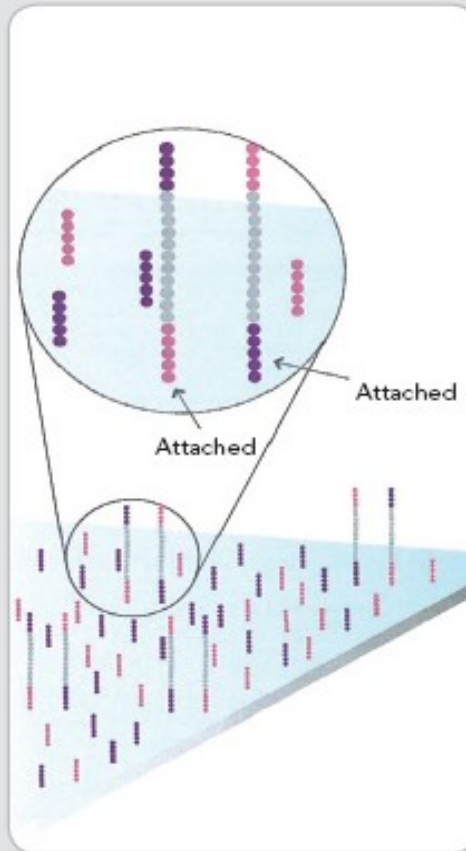
Illumina

4. FRAGMENTS BECOME DOUBLE-STRANDED



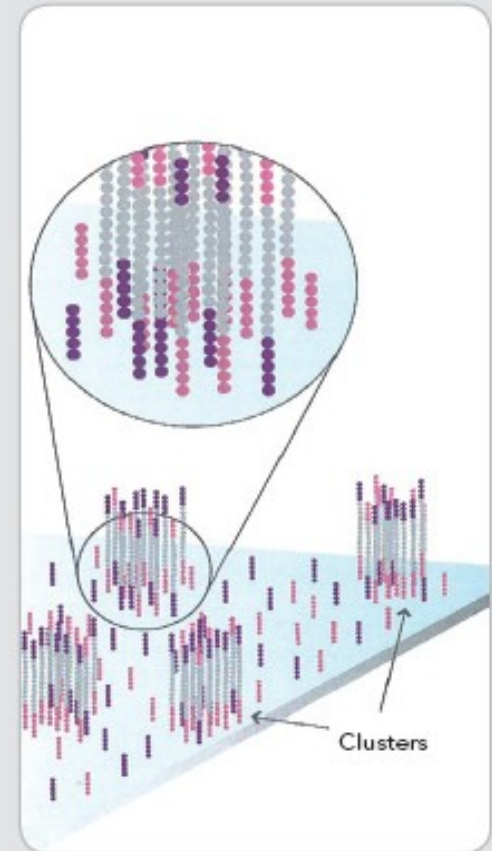
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES



Denaturation leaves single-stranded templates anchored to the substrate.

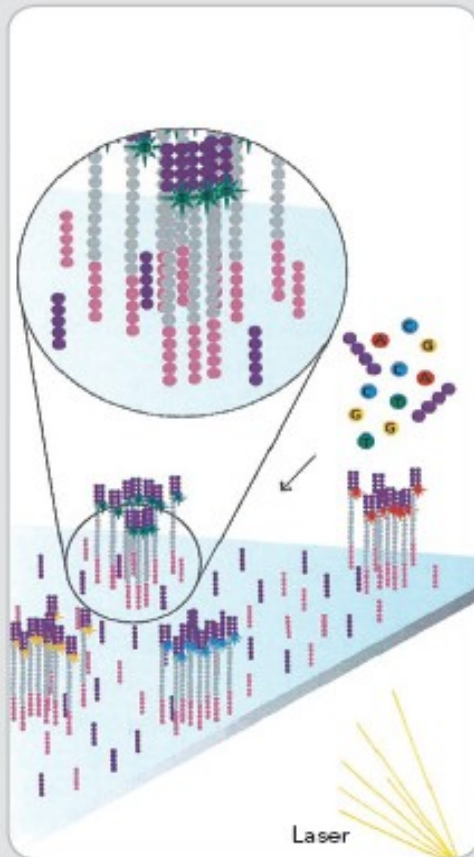
6. COMPLETE AMPLIFICATION



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

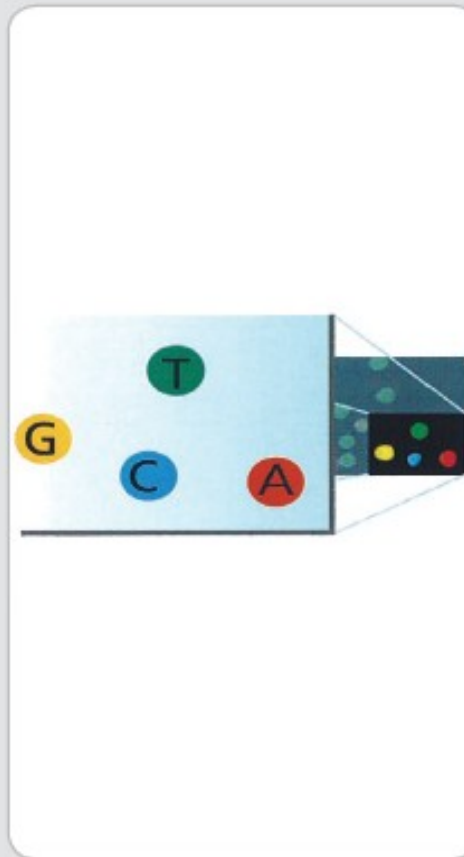
Illumina

7. DETERMINE FIRST BASE



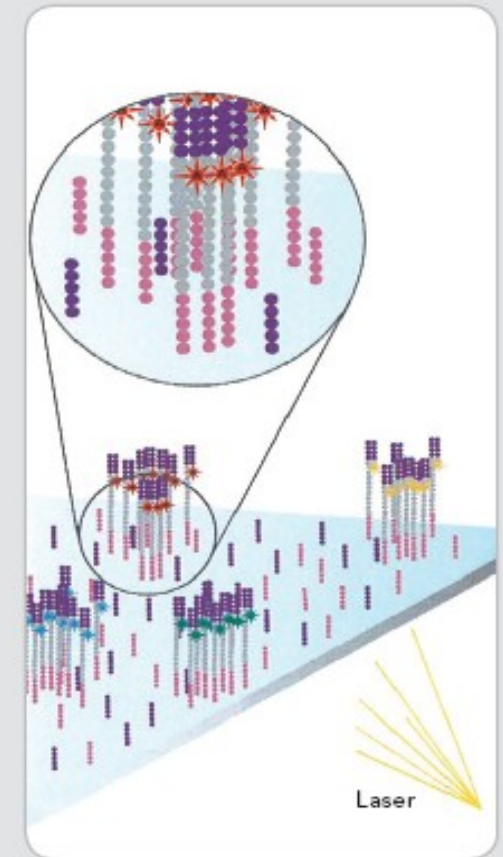
The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.

8. IMAGE FIRST BASE



After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.

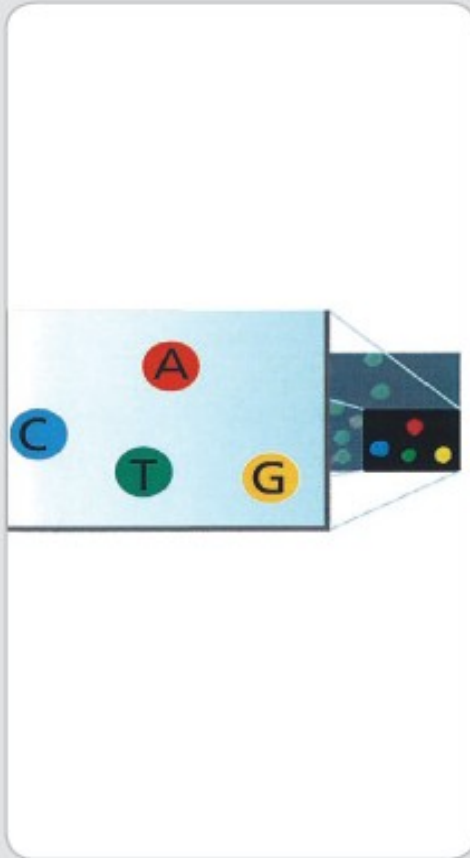
9. DETERMINE SECOND BASE



The next cycle repeats the incorporation of four labeled reversible terminators, primers, and DNA polymerase.

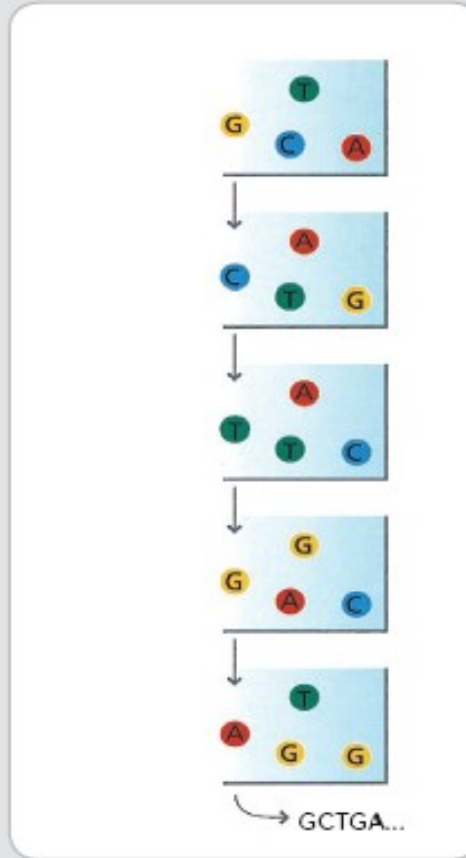
Illumina

10. IMAGE SECOND CHEMISTRY CYCLE



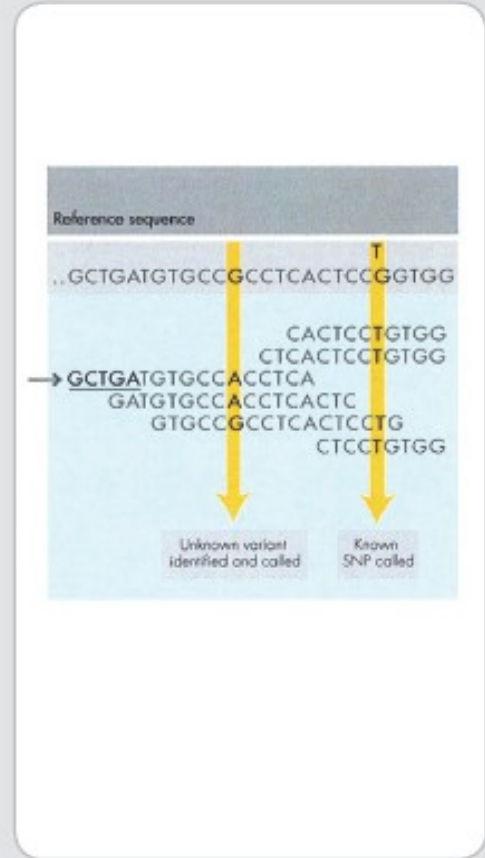
After laser excitation, the image is captured as before, and the identity of the second base is recorded.

11. SEQUENCING OVER MULTIPLE CHEMISTRY CYCLES



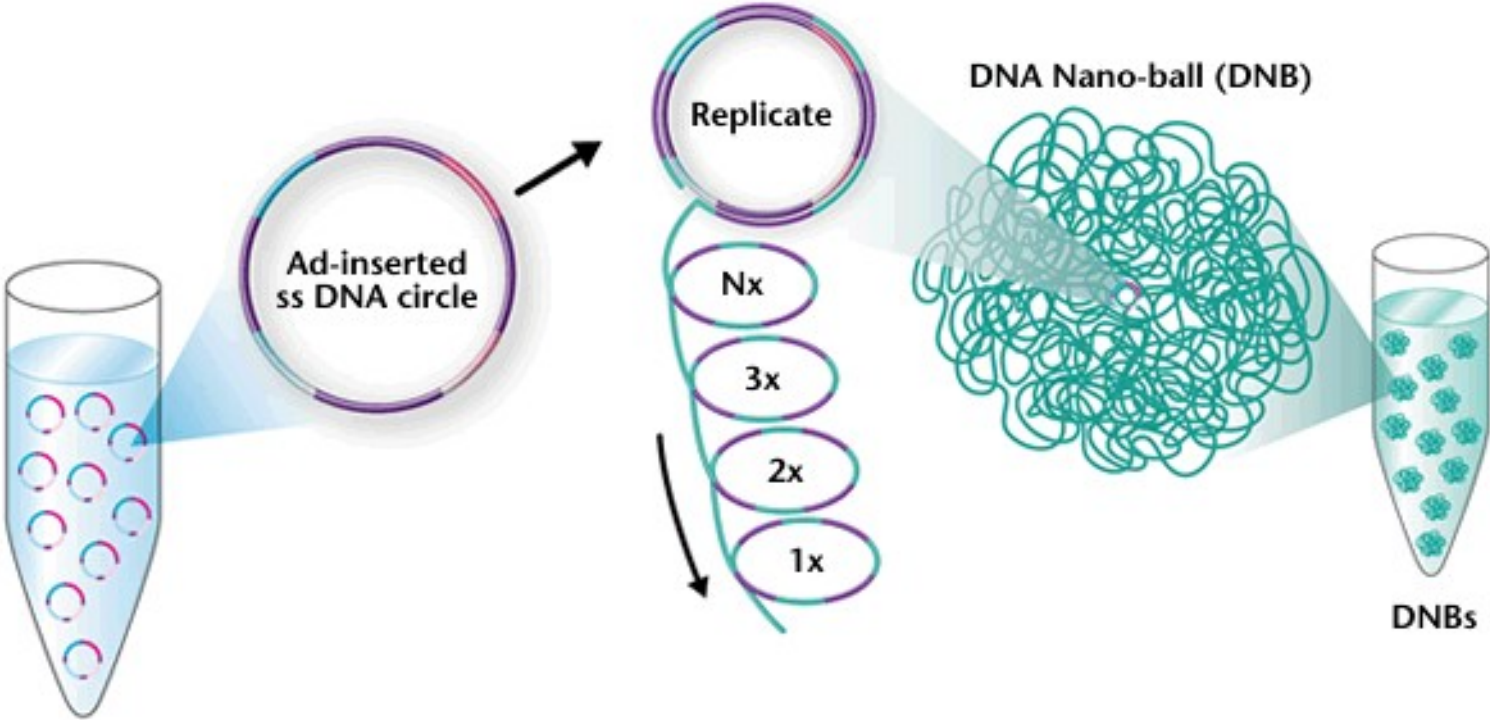
The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.

12. ALIGN DATA



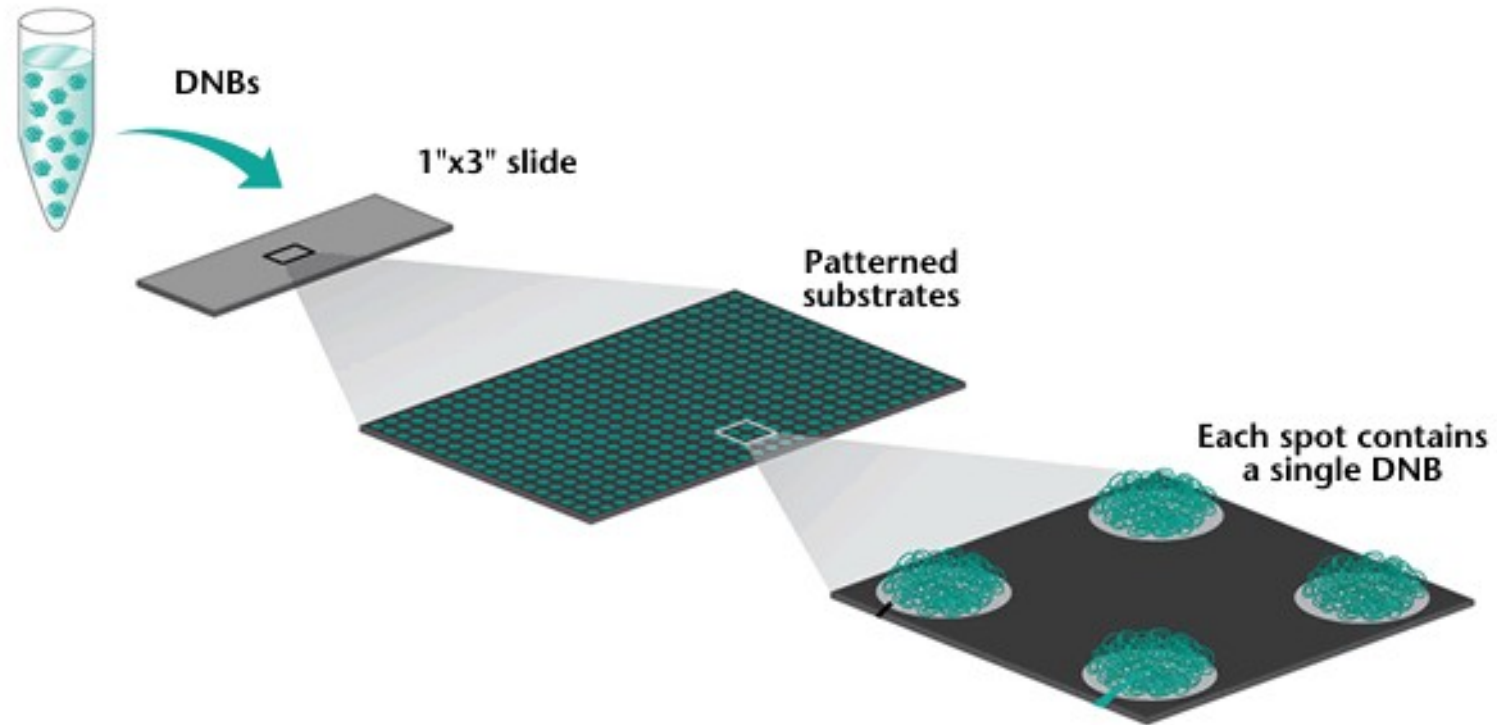
The data are aligned and compared to a reference, and sequencing differences are identified.

Complete Genomics - DNA nanoballs

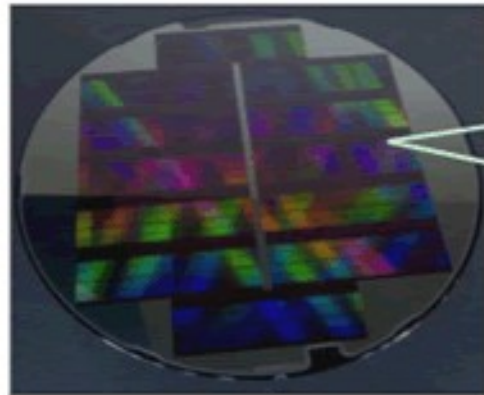


Complete Genomics

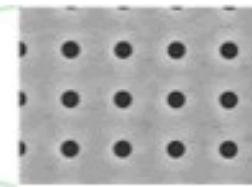
300 nm spots, 2.8 billion per slide



Complete Genomics



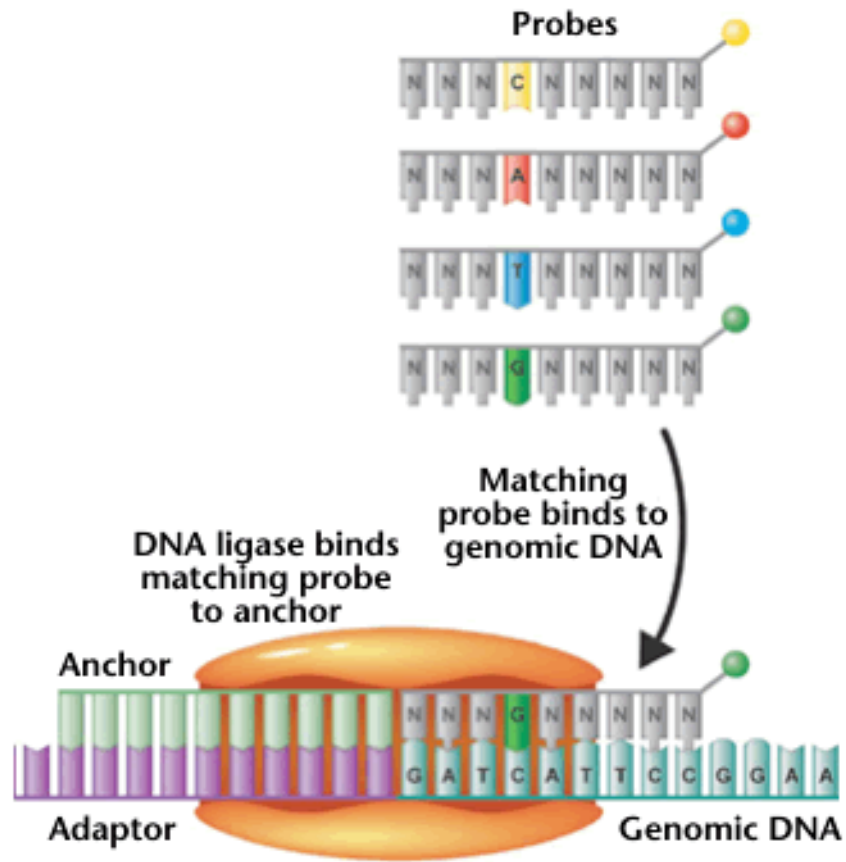
Silicon substrate consisting of 12 slides



Magnification of a section: Each spot will accommodate only a single DNB

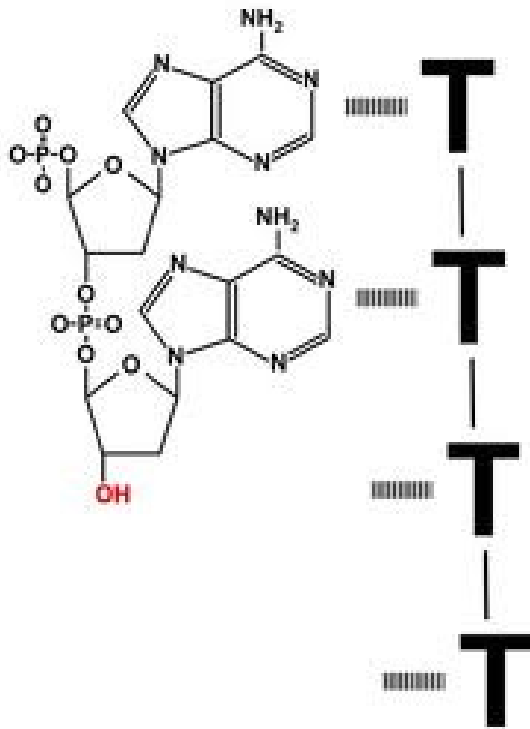
Complete Genomics

cPAL - Probe-Anchor Ligation sequencing, 70 bases per nanoball



Ion Torrent vyrábí přístroj pro polovodičové sekvenování. V srpnu byla firma převzata firmou LifeTechnologies.

Zachytává se elektrický signál z protonů (vodíkových Iontů) uvolněných po navázání dalšího nukleotidu)



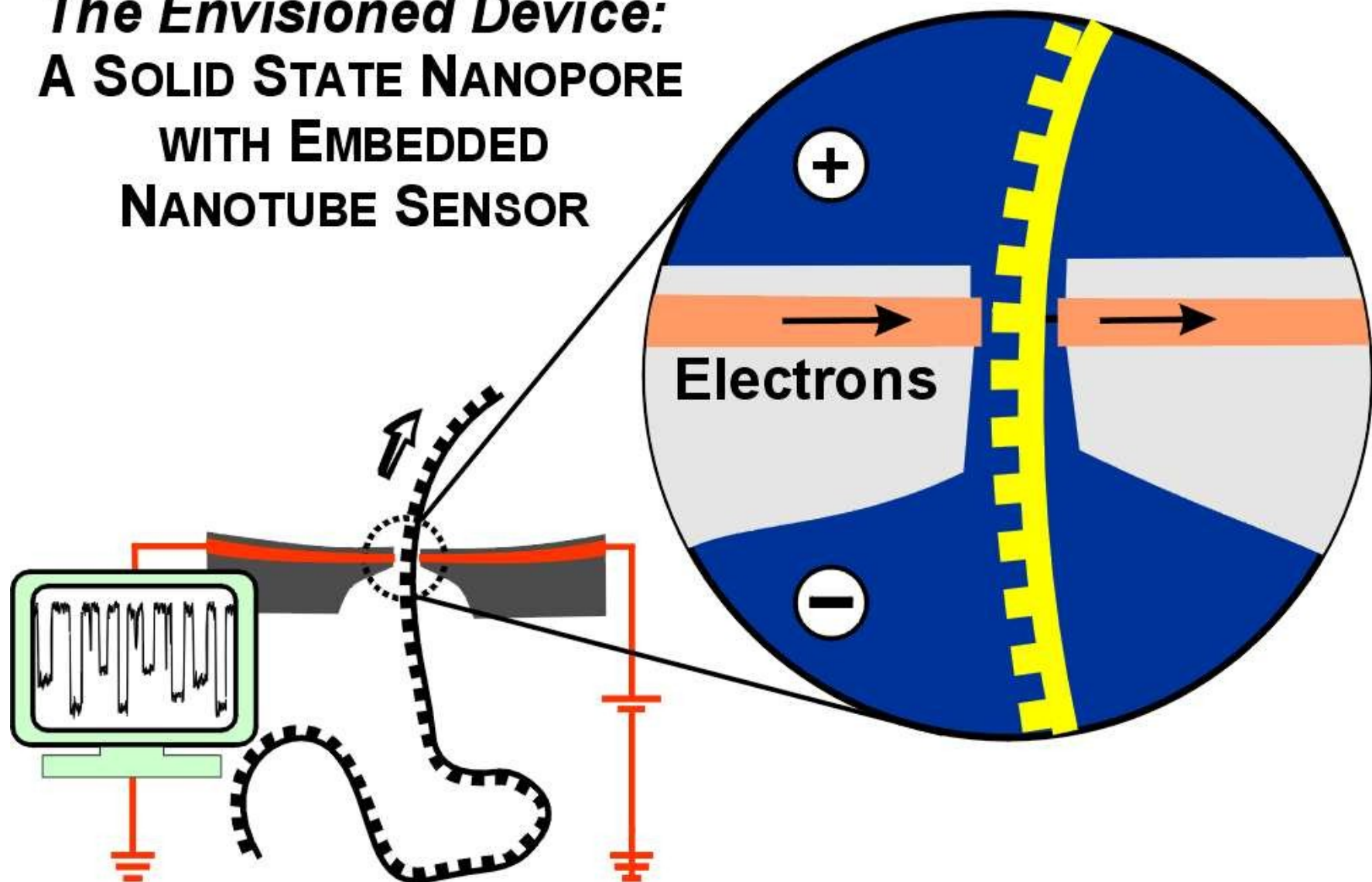
tSMS – single molecule sequencing



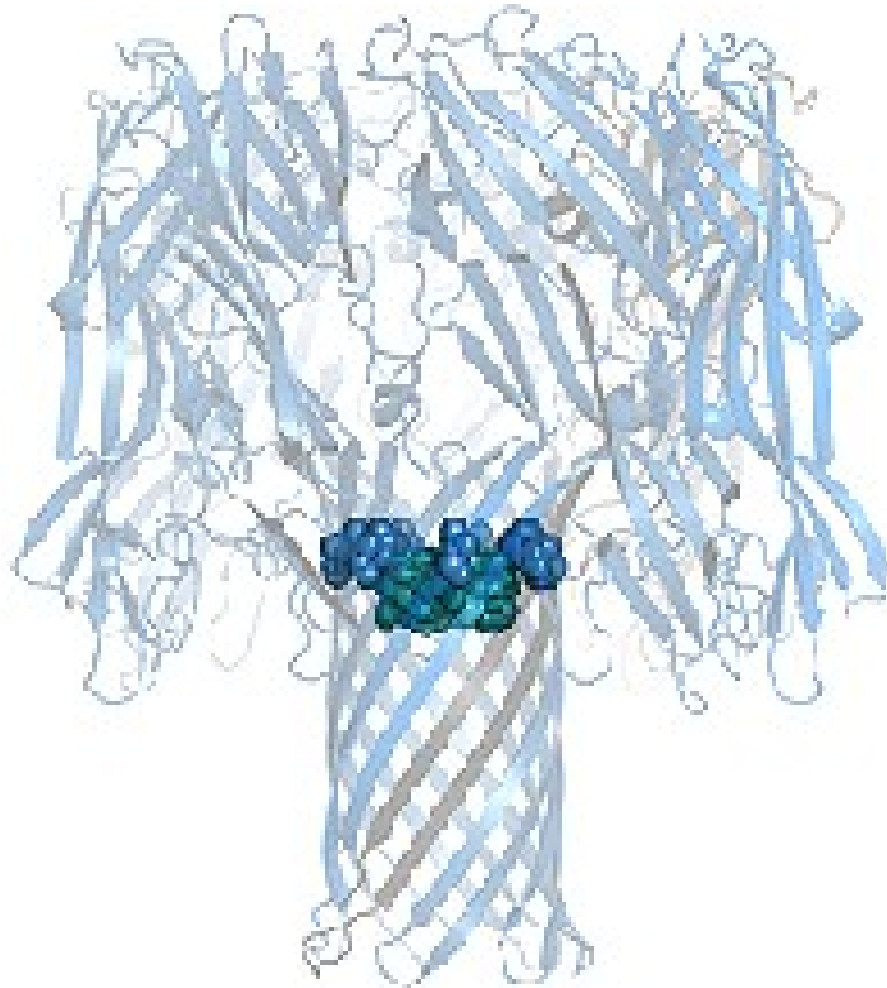
- 10^9 bp / day
- >35bp readlength

Sekvenování elektropórem

The Envisioned Device:
A SOLID STATE NANOPORE
WITH EMBEDDED
NANOTUBE SENSOR



Oxford Nanopore Technologies



Alpha haemolysin nanopore showing cyclodextrin adapter molecule (the DNA binding site).

Sanger Sequencing Centre



Důležité otázky

1. Jaký je počet nukleotidů v genomu:
 - bakterie E.coli
 - kvasinky
 - mušky Drosophila
 - člověka
 - pšenice
2. Jaká je průměrná délka zjištěných sekvencí u metod
 - první
 - druhé
 - třetígenerace?
3. Kolik trvá a stojí osekvenování jednoho lidského genomu?
4. Jaká část lidského genomu v procentech zůstává nesekvencována?

Důležité otázky

1. Jaký je počet nukleotidů v genomu:

bakterie E.coli **4.6 Mbp**

kvasinky **12 Mbp**

mušky Drosophila **123 MBp**

člověka **3200 Mbp**

pšenice **16000 Mbp**

2. Jaká je průměrná délka zjištěných sekvencí u metod

první

druhé

třetí

generace?

3. Kolik trvá a stojí osekvenování jednoho lidského genomu?

4. Jaká část lidského genomu v procentech zůstává nesekvencována?

Důležité otázky

1. Jaký je počet nukleotidů v genomu:

bakterie E.coli **4.6 Mbp**

kvasinky **12 Mbp**

mušky Drosophila **123 MBp**

člověka **3200 Mbp**

pšenice **16000 Mbp**

2. Jaká je průměrná délka zjištěných sekvencí u metod

první **800 bp**

druhé **100 bp**

třetí **50-1500 bp**

generace?

3. Kolik trvá a stojí osekvenování jednoho lidského genomu?

4. Jaká část lidského genomu v procentech zůstává nesekvencována?

Důležité otázky

1. Jaký je počet nukleotidů v genomu:

bakterie E.coli **4.6 Mbp**

kvasinky **12 Mbp**

mušky Drosophila **123 MBp**

člověka **3200 Mbp**

pšenice **16000 Mbp**

2. Jaká je průměrná délka zjištěných sekvencí u metod

první **800 bp**

druhé **100 bp**

třetí **50-1500 bp**

generace?

3. Kolik trvá a stojí osekvenování jednoho lidského genomu?

několik dnů - \$10000

4. Jaká část lidského genomu v procentech zůstává nesekvencována?

3-10%