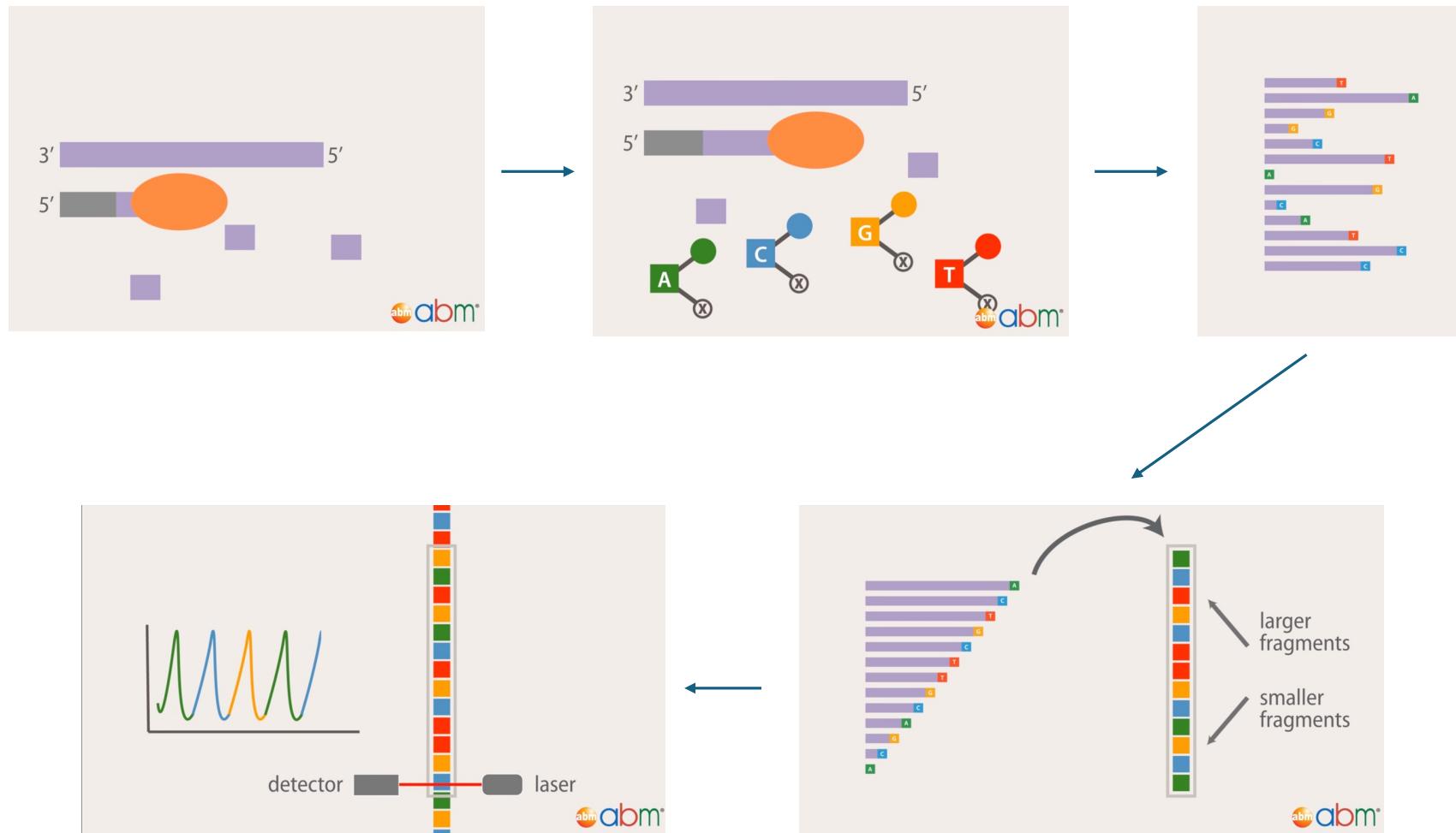
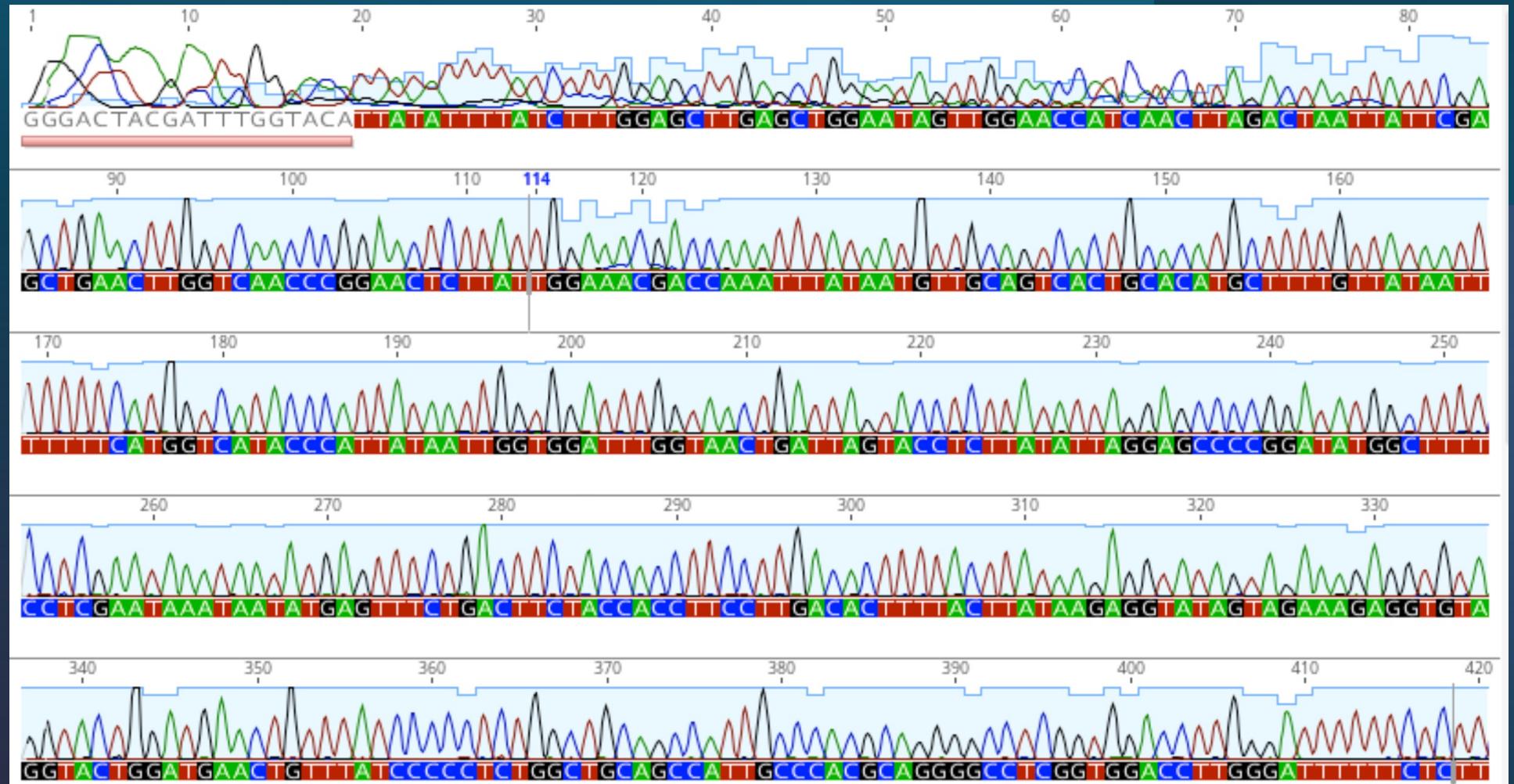
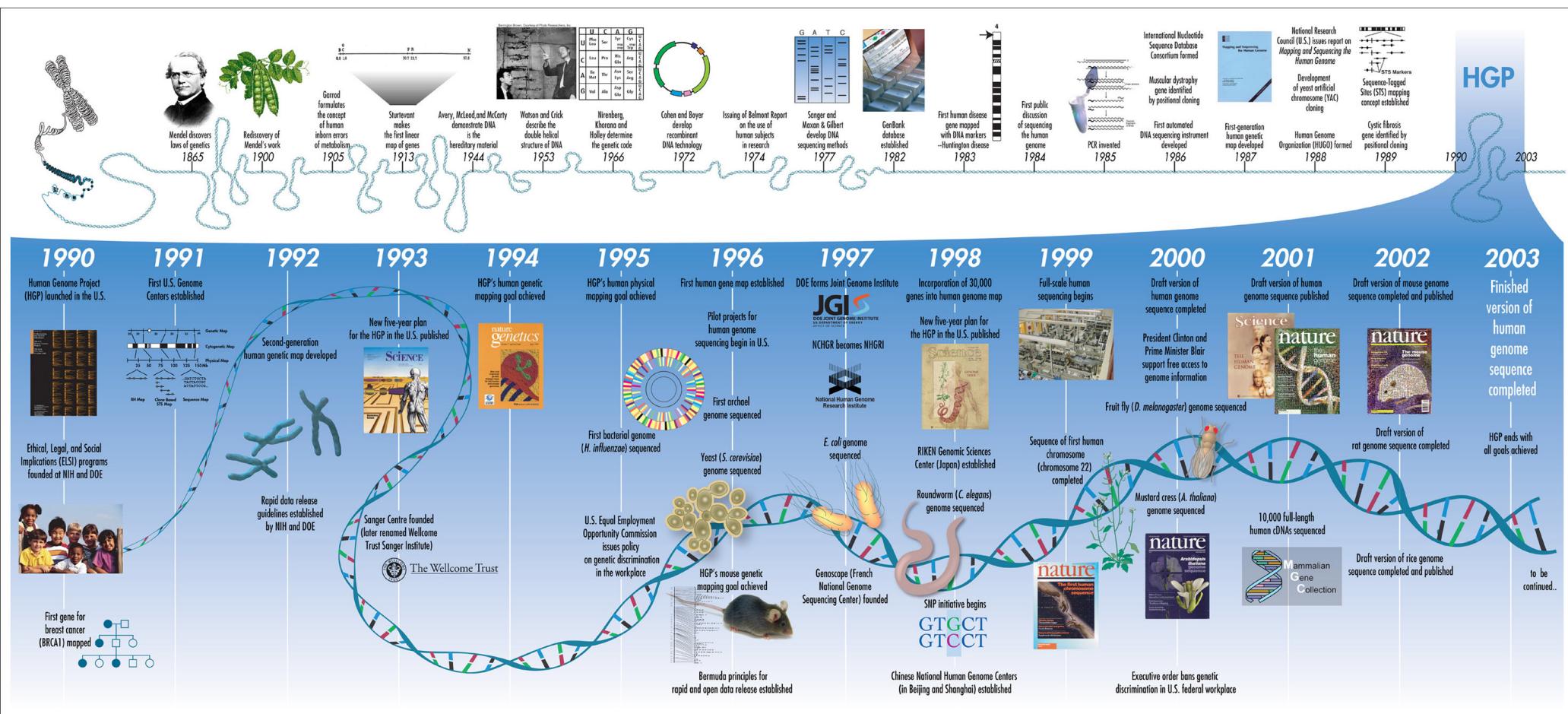


Secuenciación Sanger





Proyecto Genoma Humano



Secuenciación de Siguiente Generación Next-generation sequencing (NGS)



Minion



MinION MkI: portable, real time biological analyses

MinION

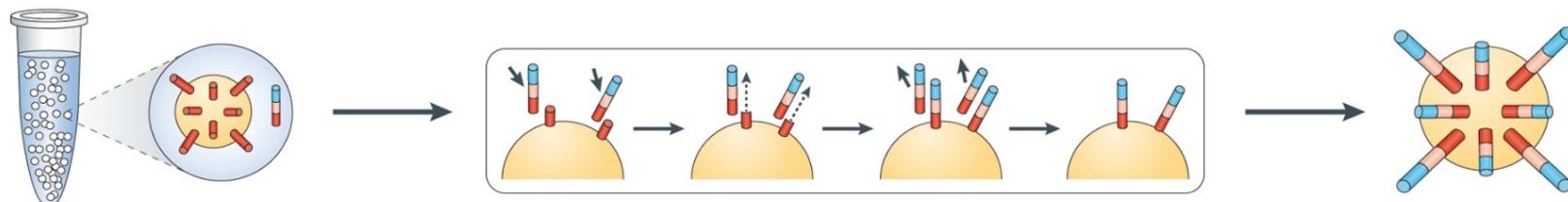
Secuenciación de Siguiente Generación

Next-generation sequencing (NGS)

- 1. Whole Genome Sequencing (WGS):** Visión completa de la composición genética de un organismo. Se utiliza para el descubrimiento de variantes, la genética de poblaciones y los estudios de asociación de enfermedades.
- 2. Targeted Sequencing (Exome Sequencing):** La secuenciación dirigida se centra en regiones específicas de interés. La secuenciación del exoma, por ejemplo, captura las regiones que codifican proteínas (exones) para identificar variantes relacionadas con enfermedades.
- 3. RNA Sequencing (RNA-Seq):** Cuantifica los niveles de expresión génica. Se utiliza para estudiar la regulación génica, el splicing alternativo y nuevas variantes de RNA.
- 4. Metagenomic Sequencing:** Analiza las comunidades microbianas a partir de muestras ambientales. Identifica especies, estudia microbiomas y explora la diversidad microbiana.
- 5. Epigenetic Sequencing:** Técnicas como la secuenciación por bisulfito revelan patrones de metilación del ADN, modificaciones de las histonas y accesibilidad de la cromatina.
- 6. Single-Cell Sequencing:** Permite perfilar la expresión génica en células individuales, proporcionando información sobre la heterogeneidad celular y los procesos de desarrollo

NGS: Lecturas cortas

a Emulsion PCR (454 (Roche), SOLiD (Thermo Fisher), GeneReader (Qiagen), Ion Torrent (Thermo Fisher))



Emulsion

Micelle droplets are loaded with primer, template, dNTPs and polymerase

On-bead amplification

Templates hybridize to bead-bound primers and are amplified; after amplification, the complement strand disassociates, leaving bead-bound ssDNA templates

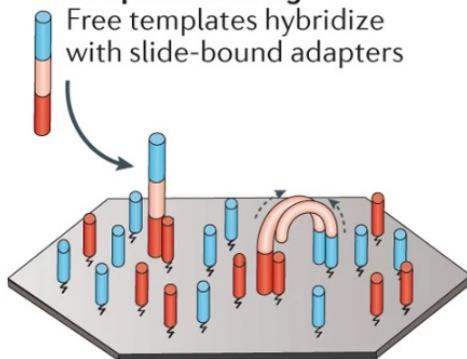
Final product

100–200 million beads with thousands of bound template

NGS: Lecturas cortas

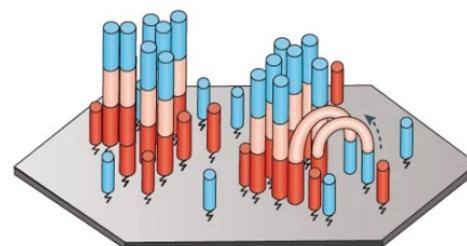
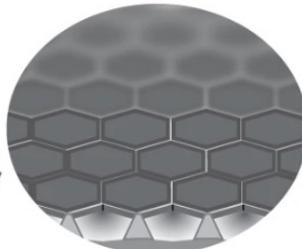
b Solid-phase bridge amplification (Illumina)

Template binding
Free templates hybridize with slide-bound adapters



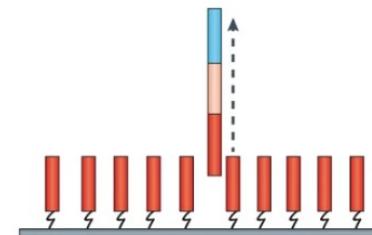
Bridge amplification
Distal ends of hybridized templates interact with nearby primers where amplification can take place

Patterned flow cell
Microwells on flow cell direct cluster generation, increasing cluster density

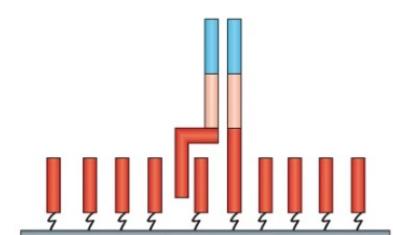


Cluster generation
After several rounds of amplification, 100–200 million clonal clusters are formed

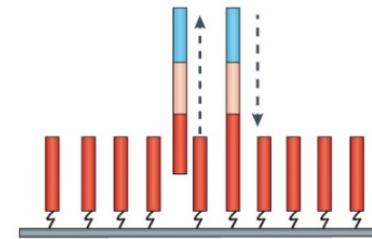
c Solid-phase template walking (SOLiD Wildfire (Thermo Fisher))



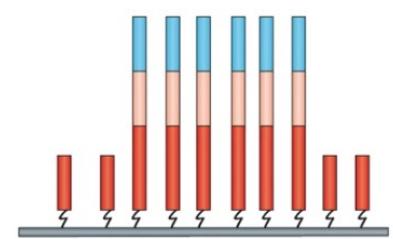
Template binding
Free DNA templates hybridize to bound primers and the second strand is amplified



Primer walking
dsDNA is partially denatured, allowing the free end to hybridize to a nearby primer



Template regeneration
Bound template is amplified to regenerate free DNA templates



Cluster generation
After several cycles of amplification, clusters on a patterned flow cell are generated

NGS: Lecturas cortas

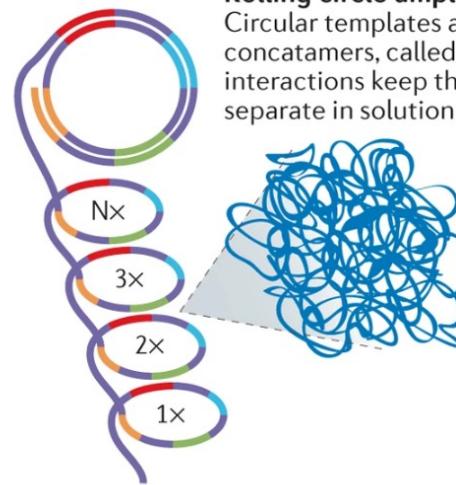
d In-solution DNA nanoball generation (Complete Genomics (BGI))



Cleavage
Circular DNA templates are cleaved downstream of the adapter sequence

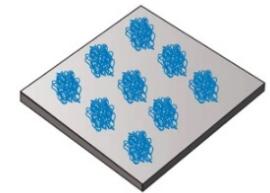


Iterative ligation
Three additional rounds of ligation, circularization and cleavage generate a circular template with four different adapters



Rolling circle amplification

Circular templates are amplified to generate long concatamers, called DNA nanoballs; intermolecular interactions keep the nanoballs cohesive and separate in solution



Hybridization
DNA nanoballs are immobilized on a patterned flow cell

Nature Reviews | Genetics

<https://doi.org/10.1038/nrg.2016.49>

NGS: Lecturas largas

A Real-time long-read sequencing

Aa Pacific Biosciences

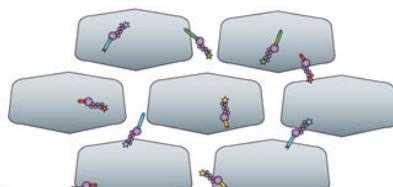
SMRTbell template

Two hairpin adapters allow continuous circular sequencing



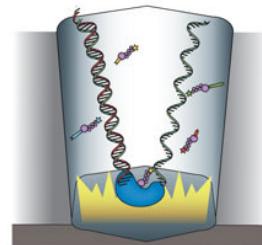
ZMW wells

Sites where sequencing takes place



Labelled nucleotides

All four dNTPs are labelled and available for incorporation



Modified polymerase

As a nucleotide is incorporated by the polymerase, a camera records the emitted light

PacBio output

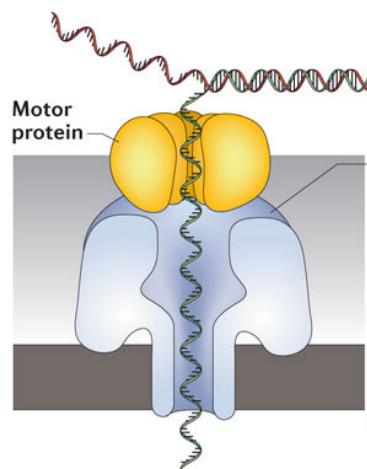
A camera records the changing colours from all ZMWs; each colour change corresponds to one base



Ab Oxford Nanopore Technologies

Leader-Hairpin template

The leader sequence interacts with the pore and a motor protein to direct DNA, a hairpin allows for bidirectional sequencing

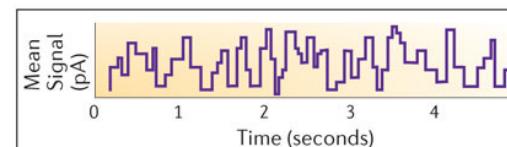


Alpha-hemolysin

A large biological pore capable of sensing DNA

Current

Passes through the pore and is modulated as DNA passes through



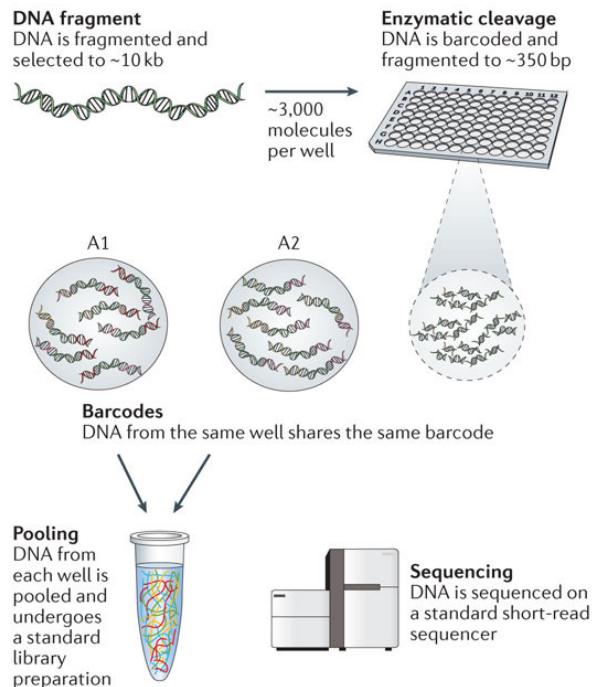
ONT output (squiggles)

Each current shift as DNA translocates through the pore corresponds to a particular k-mer

NGS: Lecturas largas

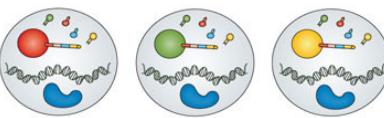
B Synthetic long-read sequencing

Ba Illumina



Bb 10X Genomics

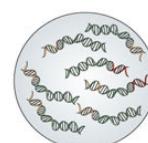
Emulsion PCR
Arbitrarily long DNA is mixed with beads loaded with barcoded primers, enzyme and dNTPs



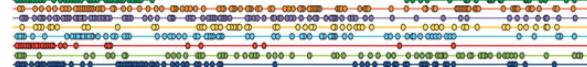
GEMs
Each micelle has 1 barcode out of 750,000



Amplification
Long fragments are amplified such that the product is a barcoded fragment ~350 bp



Pooling
The emulsion is broken and DNA is pooled, then it undergoes a standard library preparation



Linked reads

- All reads from the same GEM derive from the long fragment, thus they are linked
- Reads are dispersed across the long fragment and no GEM achieves full coverage of a fragment
- Stacking of linked reads from the same loci achieves continuous coverage

SECUENCIACIÓN DE NUEVA GENERACIÓN (NGS)

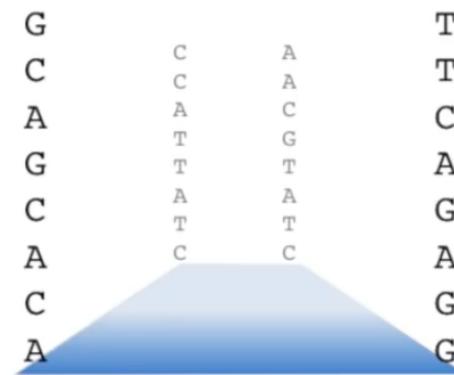


Examples of NGS systems



ILLUMINA

Actually, there are about 400,000,000 fragments laid out vertically in a grid.



The machine has fluorescent probes
that are color coded according to the
type of nucleotide they can bind to.

● = A

● = G

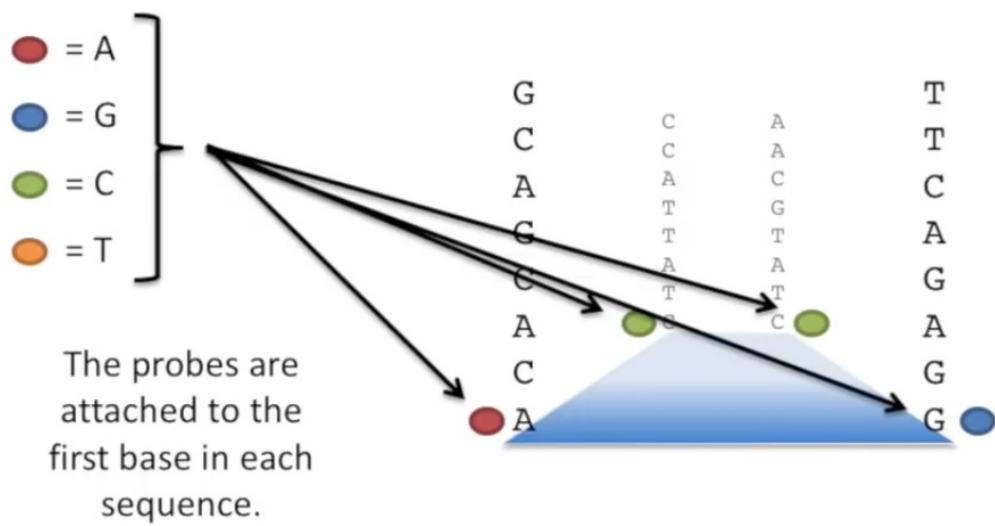
● = C

● = T

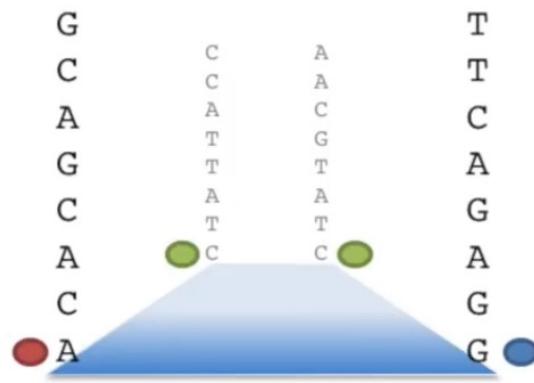
G		A	T
C	C	A	C
A	A	T	G
G	T	T	T
C	A	A	A
A	T	T	G
C	C	C	A
A			G



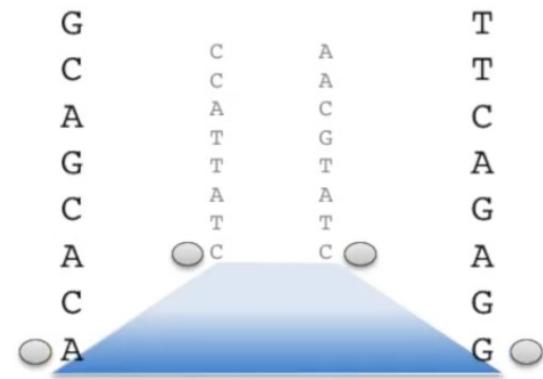
The machine has fluorescent probes that are color coded according to the type of nucleotide they can bind to.



Once the probes have attached, the machine takes a picture of the flow cell from above that looks like this...

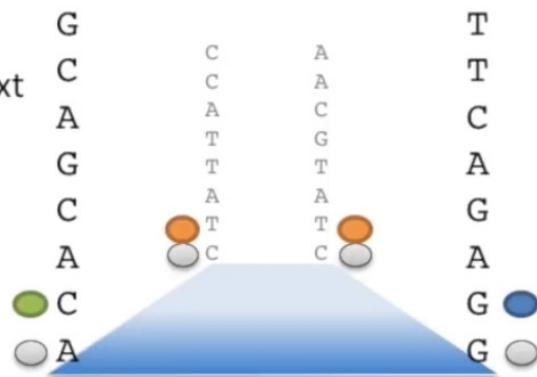


Then the machine washes the color off of the probes....

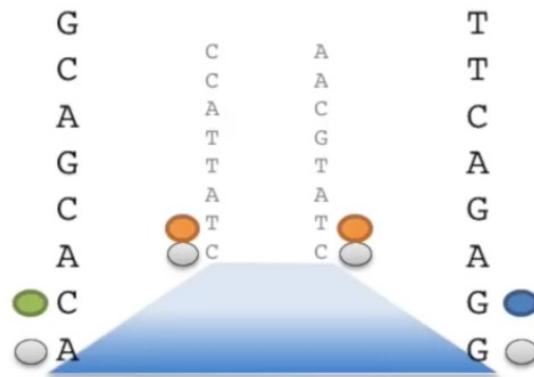
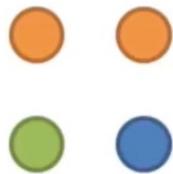


- = A
- = G
- = C
- = T

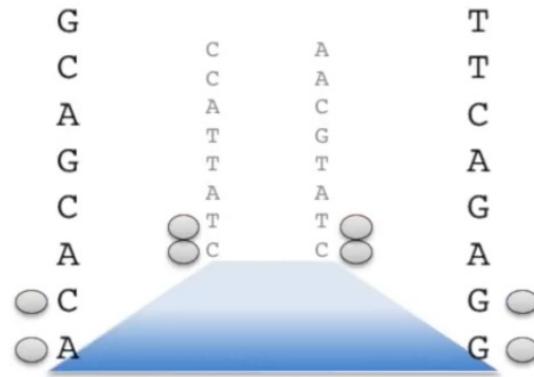
} Then probes are bound to the next base in each fragment.



The machine takes a picture
from above...



Then the machine washes the color off of the probes....



And the process repeats until the machine has determined each sequence of nucleotides.

● = A

● = G

● = C

● = T

● G

● C

● A

● G

● C

● A

● C

● A

C
C
A
T
G
T
A
T
A
C

A
C
G
T
T
A
A
T
C

T ●

T ●

C ●

A ●

G ●

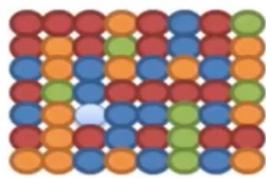
A ●

G ●

G ●

This is how it works with 4 DNA fragments.

This matrix still isn't 400,000,000 DNA fragments, but it illustrates one type of problem that can occur.



The raw data...

```
@NS500177:196:HFTTAFXX:1:11101:10916:1458 2:N:0:CGCAGCTG  
ACACGACGATGAGGTGACAGTCACGGAGGATAAGATCAATGCCCTCATTAAGCAGCCGGTGTAA  
+  
AAAAAEEEEEEEEE//AEEEEEEEEEEE/EE/<<EE/AAEEAEE//EEEAEAEAA<
```

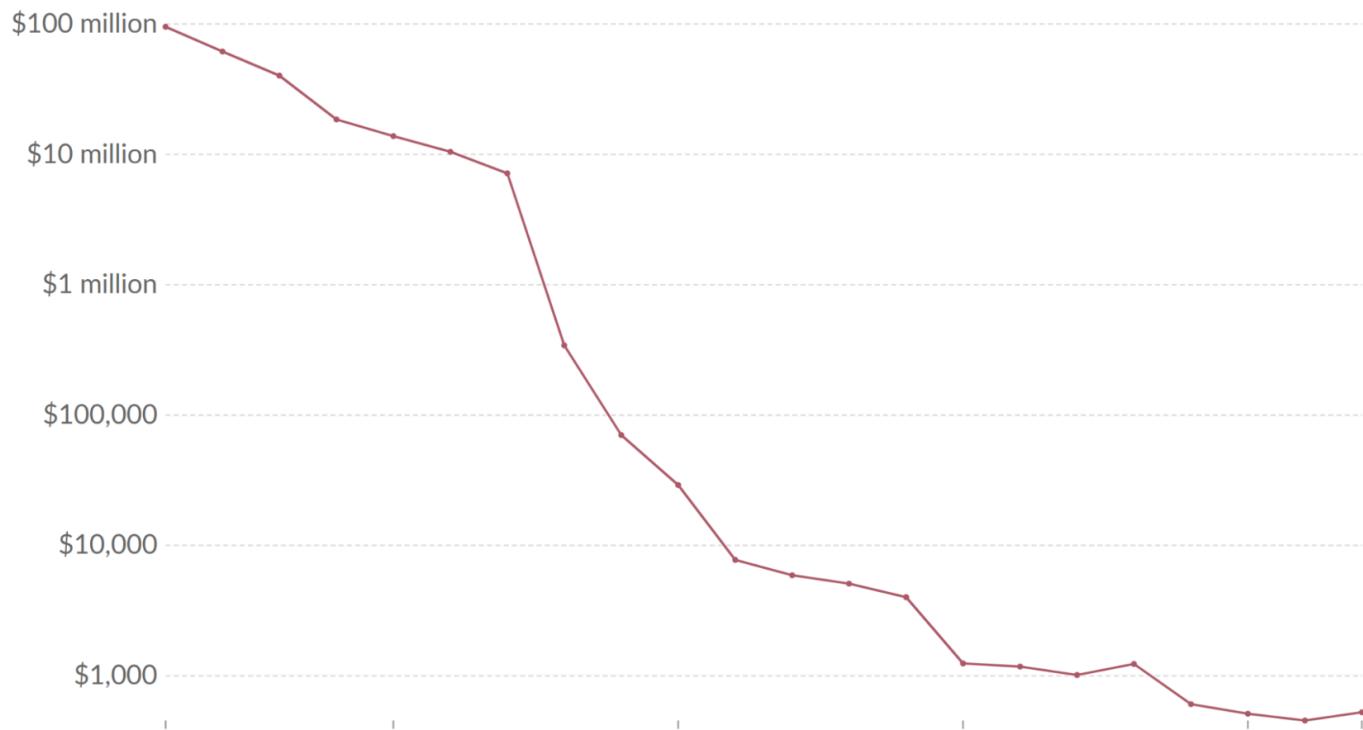
Each sequencing “read” consists of 4 lines of data.

A typical sequence run with 400,000,000 reads will generate a file containing 1.6 billion lines of data!!!

Cost of sequencing a full human genome

Our World
in Data

The cost of sequencing the full genetic information of a human, measured in US\$. This data is not adjusted for inflation.



Data source: National Human Genome Research Institute (2022)

OurWorldInData.org/technological-change | CC BY

Figure 1. The cost of sequencing a full human genome from 2001 to 2022.

El futuro de la NGS

- Acceso a equipos de NGS.
- Automatización (Sample-to-Answer)
- Diagnóstico



100,000 Genomes Project
Genomics England's very first initiative – sequencing 100,000 genomes from around 85,000 NHS patients affected by rare disease or cancer – is leading to groundbreaking insights and continued findings into the role genetics can play in healthcare.

EARTH BIOPROJECT
CREATING A NEW FOUNDATION FOR BIOLOGY
Sequencing Life for the Future of Life