



CCiT 

Centres Científics i Tecnològics
UNIVERSITAT DE BARCELONA

Aplicaciones y Tendencias en secuenciación de ADN

Agustín Arasanz Duque

**Unidad de Genómica
Centres Científics i Tecnològics
Universidad de Barcelona**

Aplicaciones y Tendencias en secuenciación de ADN

- **Introducción**
- **Técnicas de secuenciación y aplicaciones:**
 - 1ª Generación: Sanger
 - 2ª Generación: Secuenciación masiva en paralelo
 - 3ª Generación: Molécula Única
- **Tendencias**

Introducción: ADN

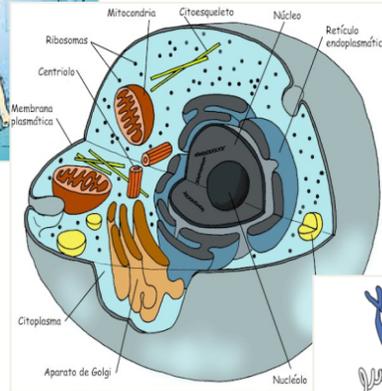
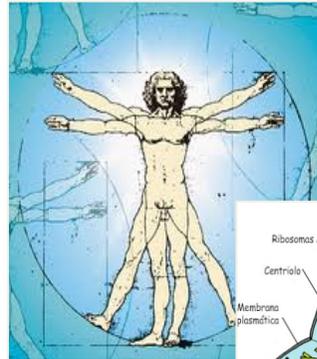
*ADN

Almacena información estructural i funcional de los organismos vivos.

Cambios en la secuencia = mutaciones. Puede dar lugar a enfermedades.

Secuencia de cada Organismo es única.

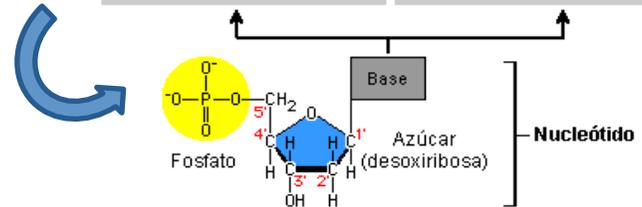
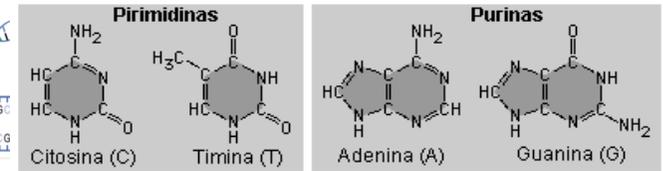
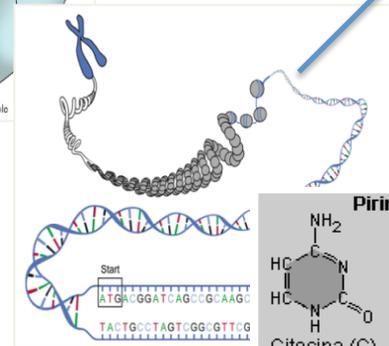
Se puede identificar individuos y compararlos



La longitud depende de la complejidad del organismo:

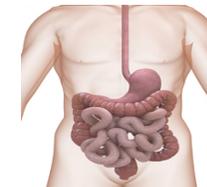
Homo sapiens:
 $3000 \times 10^6 = 3 \text{ Gb}$

Mycoplasma genitalium:
 $580 \times 10^3 = 0,58 \text{ Mb}$

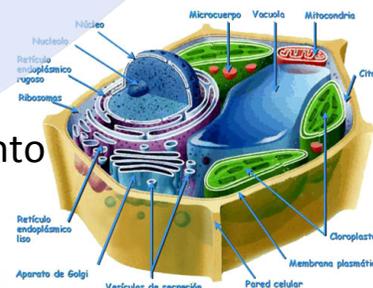


Introducción: Pasos previos al Análisis de ADN

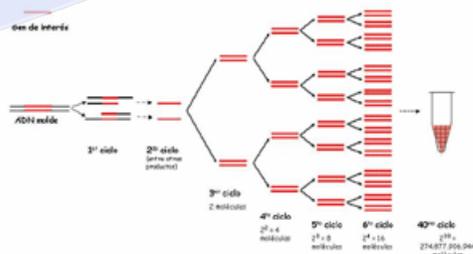
1. Obtención de la muestra



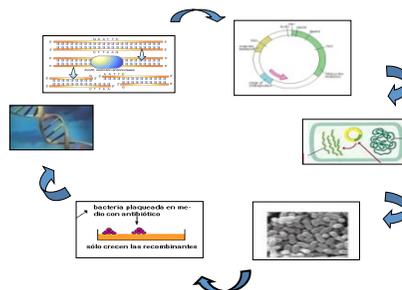
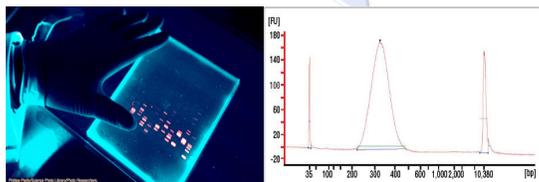
2. Aislamiento



3. Amplificación

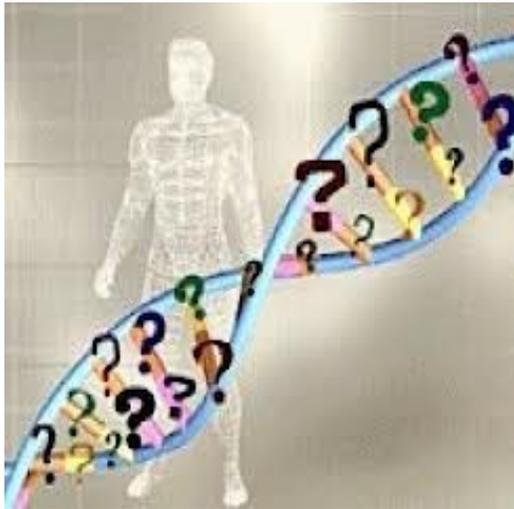


4. Control de Cantidad y Calidad

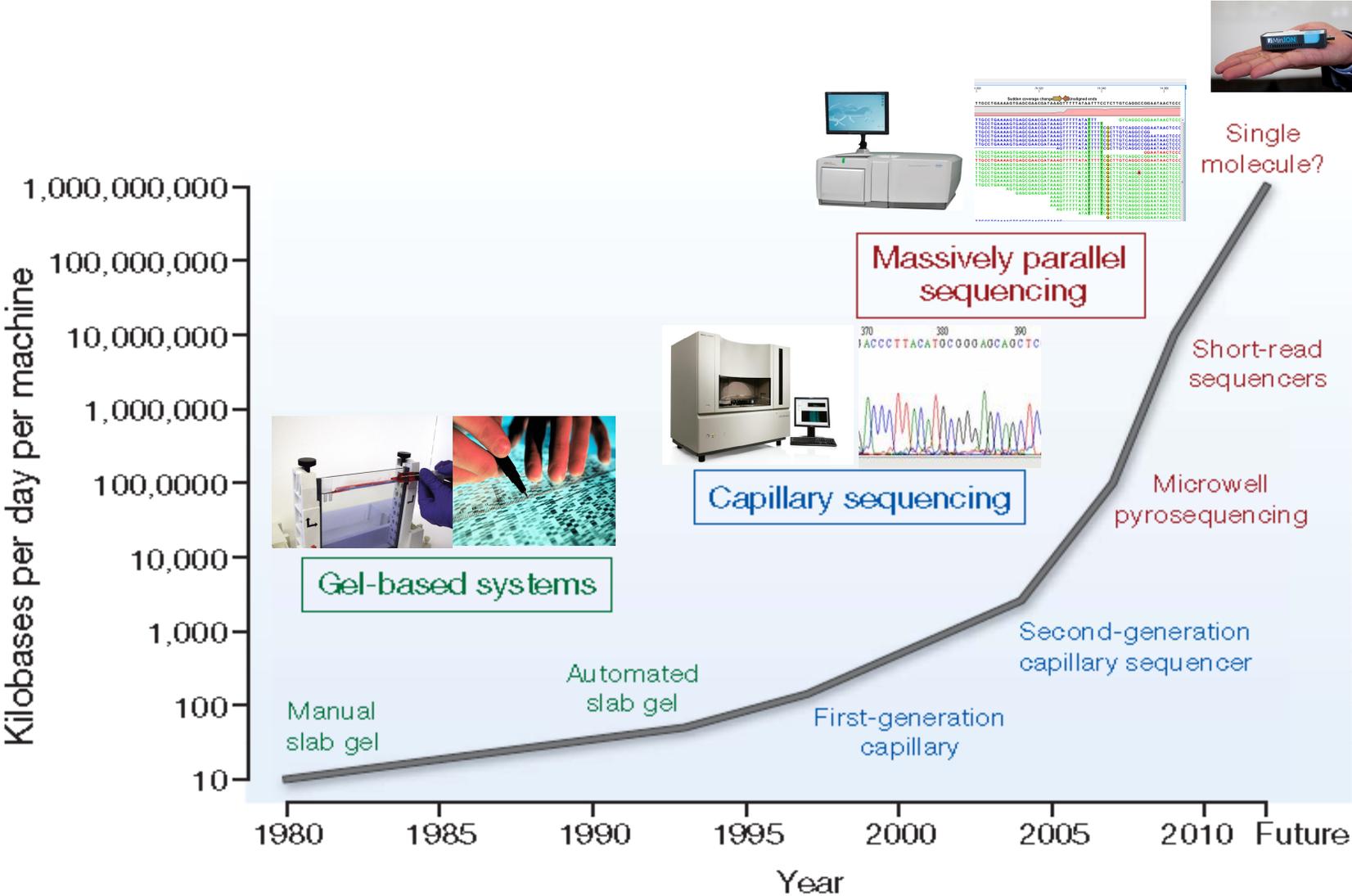


Introducción: Secuenciación ADN

Descifrar el orden de los nucleótidos de un fragmento de ADN



Introducción: Historia Secuenciación ADN



Secuenciación ADN: Tecnología

- Primera generación: Sanger / Método de terminación de cadena



- Terminadores (dd NTP)
- Electroforesis capilar
- 1975->

- Segunda generación: Secuenciación masiva en paralelo / NGS



- Alto rendimiento.
- Clonación in vitro
- Lecturas cortas
- 2000 ->

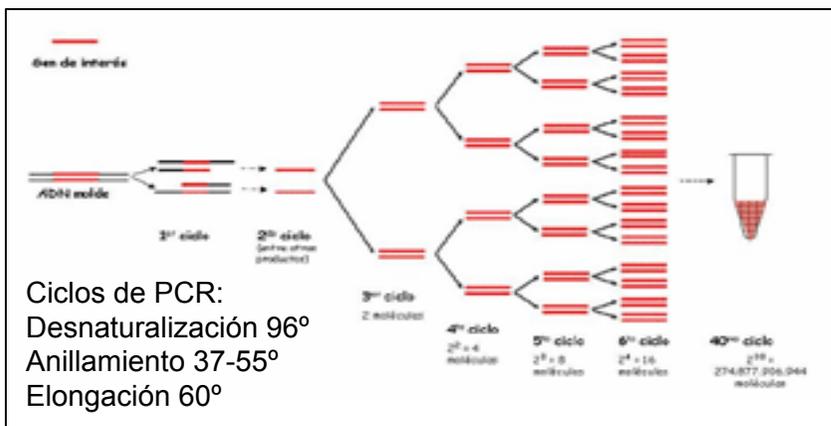
- Tercera generación: Secuenciación de molécula única



- Molécula simple
- No amplificación.
- 2005 ->

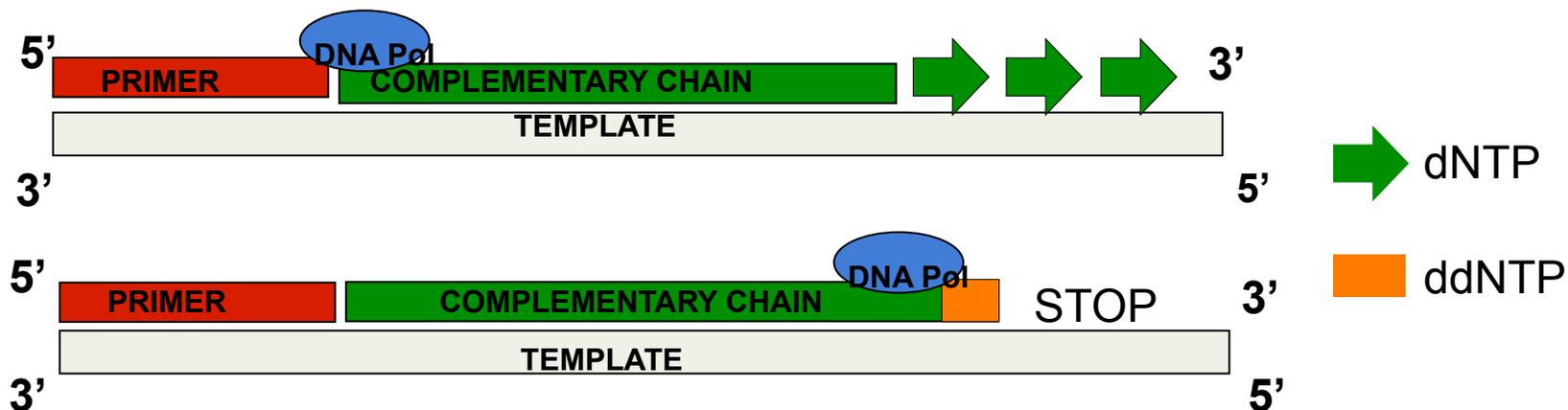
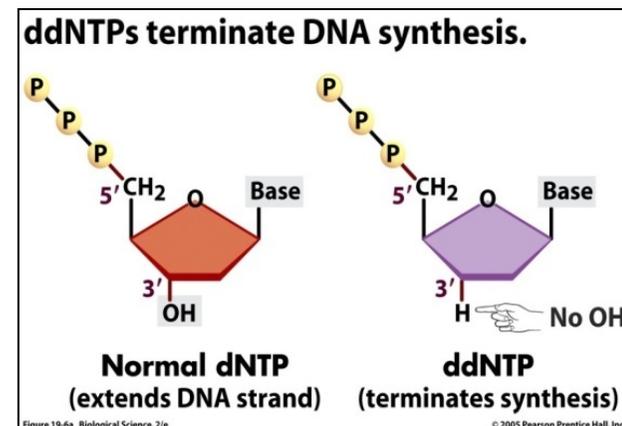
Primera generación

Sanger / Método de terminación de cadena



Reactivos (PCR)

- DNA
- dNTP
- ddNTP**
- TAQ polymerase
- Buffer
- Primer



Secuenciación Sanger / método de terminación de cadena

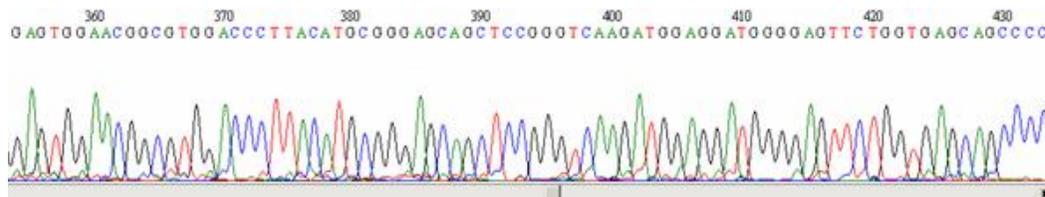
Electroforesis capilar automatizada



- Separación de los fragmentos DNA por electroforesis capilar
- Detección por fluorescencia.

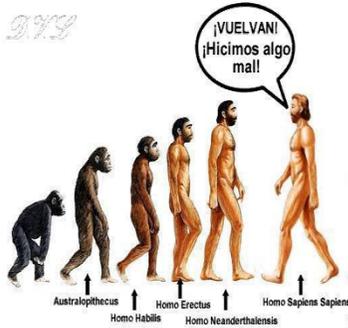
Rendimiento:

- Run: 384 muestras
- 12 runs por día.
- Longitud de fragmentos: 800 bp.
- Rendimiento por run: 300kb
- Logros: Proyecto Genoma Humano (3000×10^6 pb, 25000 genes). 10 años y 3B \$



Secuenciación Sanger / método de terminación de cadena y Electroforesis capilar

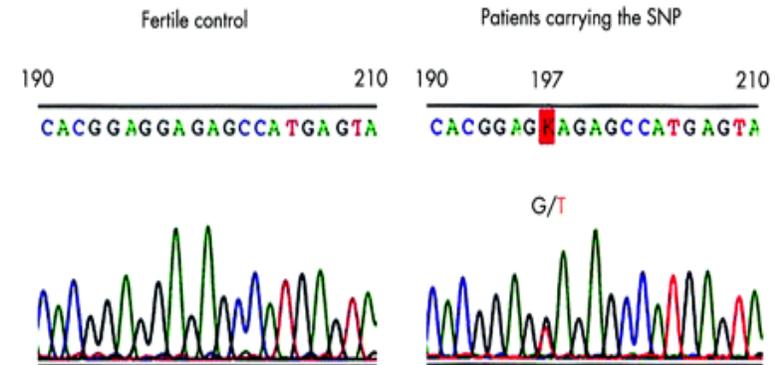
Aplicaciones



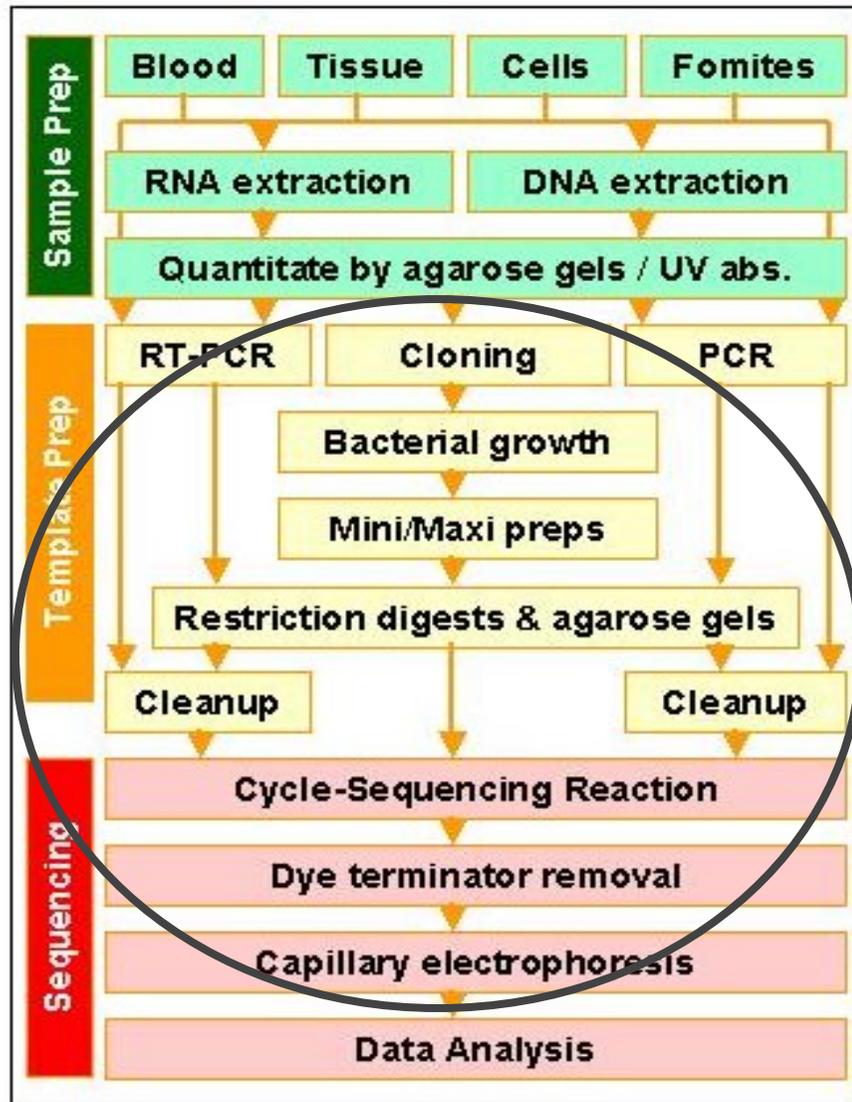
Ejemplo:

Genomic analysis of the *PRM1* gene (protamina1)

Causa genética de infertilidad en el hombre



Secuenciación Sanger: Proyectos a gran escala



The Human Genome Project (1990 -2003)



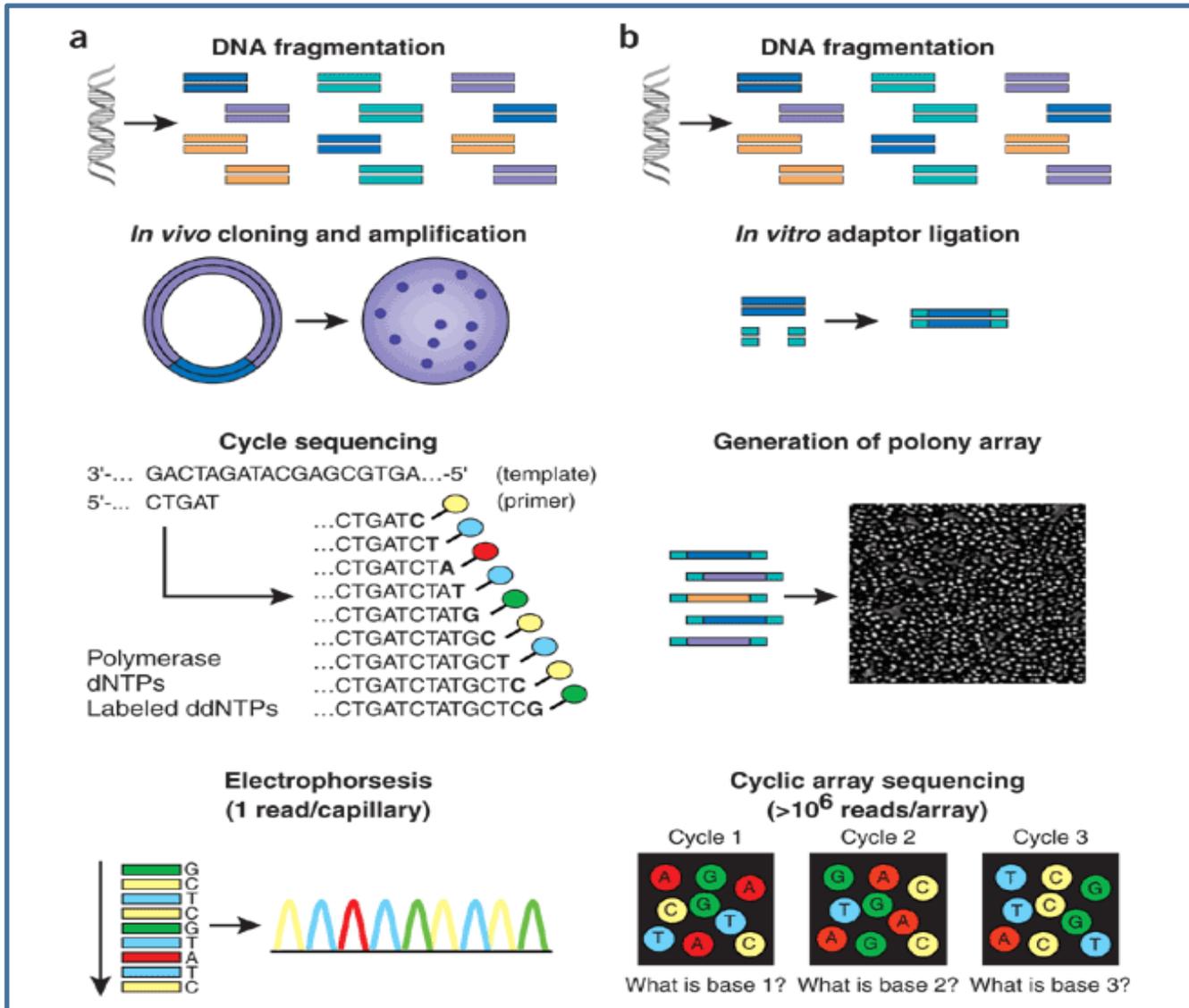
Human genome
3.000.000.000pb
25.000 genes

**EXPENSIVE AND
LONG PROCESS
70B USD
>10 YEARS**



2ª Generación

Sanger (a) vs Secuenciación Masiva en Paralelo (b)



- The ability to process millions of sequence reads in parallel rather than 384 at a time.

- NGS fragment libraries do not need vector based cloning and E. coli based amplification stages used in capillary sequencing.

- Shorter Read Lengths.

- Reduce time and cost.

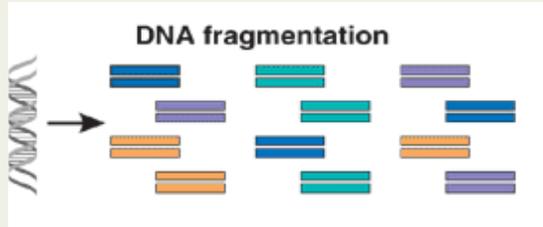
- High throughput:

Sanger: <300kb

NGS: > 500Mb

Secuenciación Masiva en Paralelo / NGS

1. Fragmentación y selección de tamaño del ADN

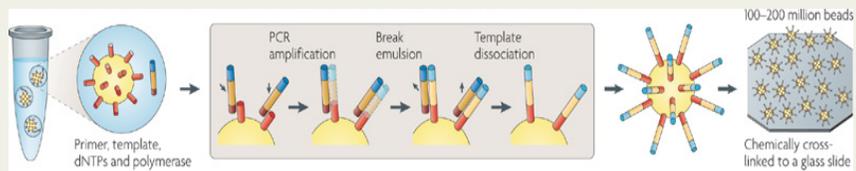


2. Ligación de adaptadores en los extremos.

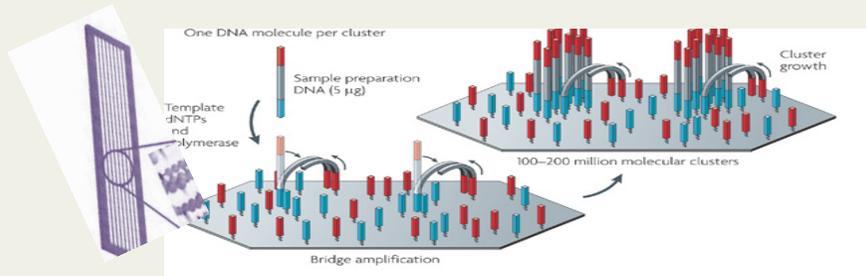


3. Clonación **masiva** en una superficie sólida (millones de fragmentos a la vez).

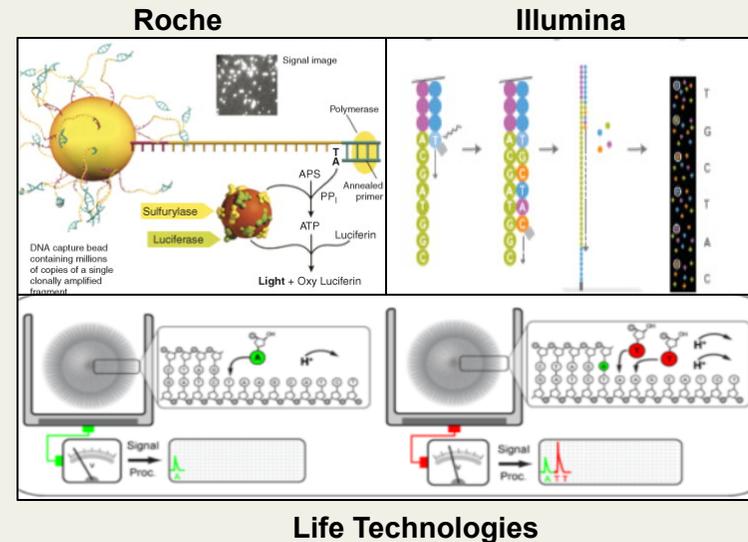
PCR en emulsión (Roche/Life Technologies)



PCR en Clústers (Illumina)



4. Secuenciación por síntesis de cada fragmento en **paralelo**



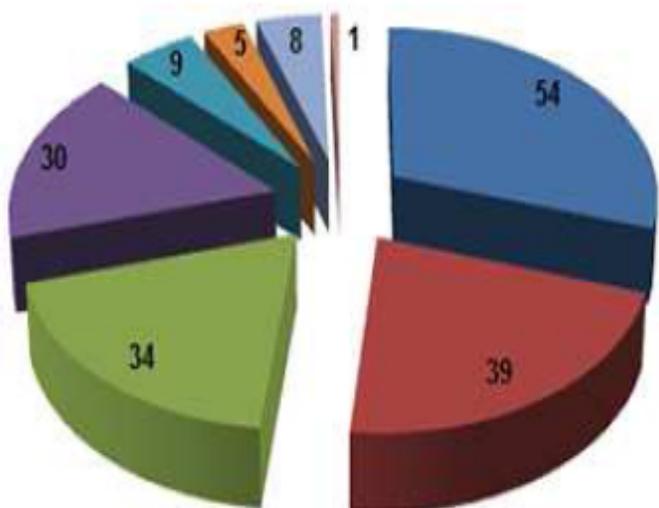
Secuenciación Masiva en Paralelo: Tecnología



> De 14 empresas de NGS

Secuenciación masiva en paralelo: Tecnología

What is your favourite high-throughput sequencing machine by platform?



- Illumina MiSeq
- Illumina HiSeq 2000/2500
- Roche 454
- Ion Torrent
- ABI SOLiD
- Pacific Biosciences
- Others
- Illumina Genome Analyzer

<http://nextgenseek.com>

Illumina



Roche



454 FLX+



GS Junior

Life Technologies

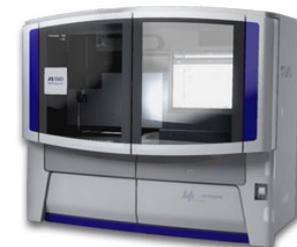


Ion Proton



Ion Torrent

ABI SOLiD



Secuenciación masiva en paralelo: Aplicaciones



¿Que tecnología escoger para mi proyecto



•Tipo de estudio ?

- Secuenciación de novo.
 - Resecuenciación.
 - Exoma.
 - Transcriptoma.
 - Expresión génica
 - Metagenómica.
 - Amplicones
 - Búsqueda de marcadores moleculares
 - Diagnóstico
 - Interacción DNA- Proteína (Chip-seq)
- Presupuesto?
- Tamaño del proyecto y/o número de muestras necesarias?

- Longitud de las secuencias
- Profundidad de cobertura (Coverage).
- Rendimiento

Secuenciación masiva en paralelo: Aplicaciones

Estudios donde la longitud de secuencia si que importa

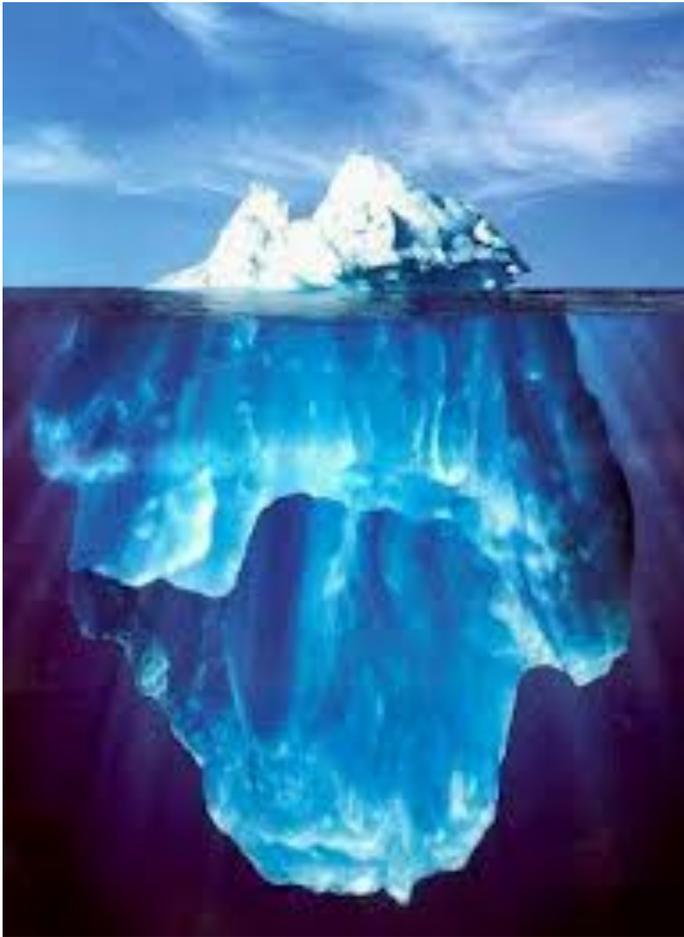


- Secuenciación *de novo* (mejor ensamblaje)
- Metagenómica (mejor clasificación taxonómica)
- Haplotipos y mutaciones estructurales (mejor detección)
- Detección de polimorfismos CNV's
- Amplicones (mejor diseño)
- Facilita el análisis bioinformático



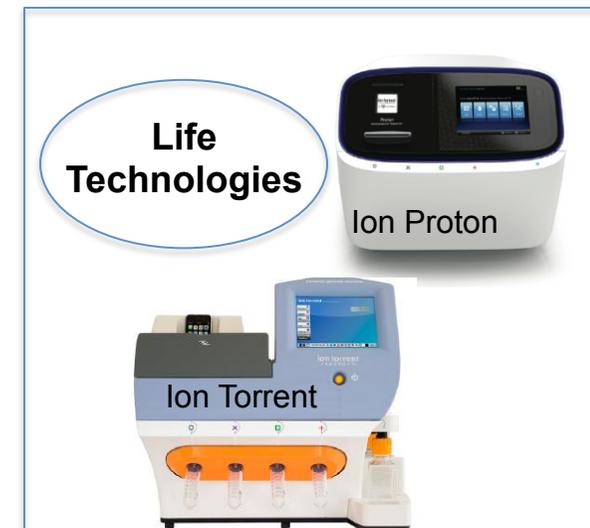
Secuenciación masiva en paralelo: Aplicaciones

Estudios donde la profundidad de cobertura (coverage) es esencial



- Resecuenciación (mejora la fiabilidad)
- Expresión génica (permite distinguir diferentes niveles de expresión)
- Exomas (Permite detectar variaciones poco frecuentes)
- Detección de polimorfismos SNP's

> X 5000!



Secuenciación masiva en paralelo: Aplicaciones

Rendimiento

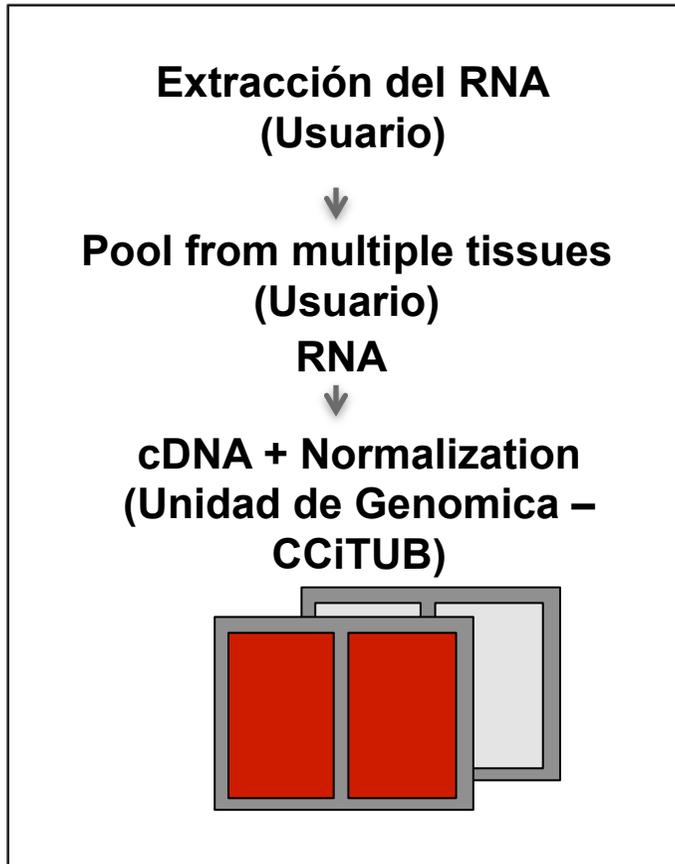
Tecnología	Plataforma	# Reads (Mill.)	Read longitud (bp)	Total output (Gb)	Coste por base	Técnica
Roche	GS FLX+	1	800	1	\$\$\$	emPCR, SBS, light detection
	GS Junior	0.100	800	0.05	\$\$\$	
Illumina	HiSeq 2000	6000	2x200	600	\$	Bridge PCR, SBS, fluorophore
	MiSeq	12	2x200	8	\$\$	
Life Technolgies	Ion PGM - 316	6	150	0.1	\$\$	emPCR, SBS, pH change
	Ion Proton	60	150	10	\$	

Secuenciación masiva en paralelo / NGS

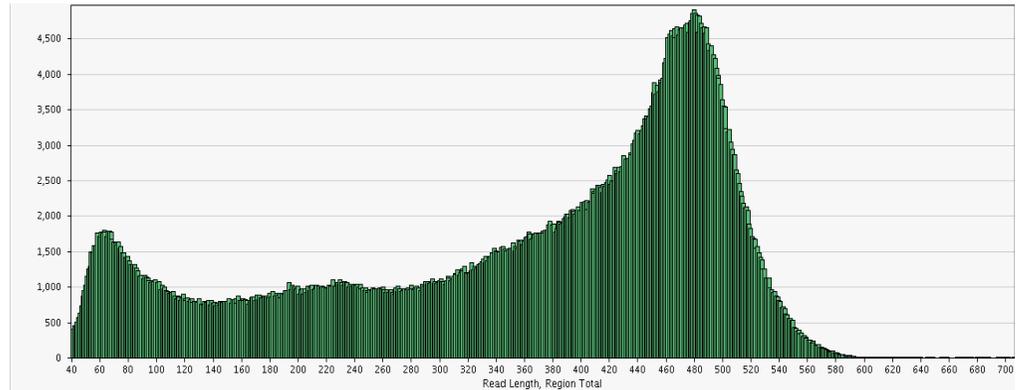
CCiT-UB- Unidad de Genómica
Tecnología 454FLX ROCHE



Ejemplo: Transcriptoma - Sparus aurata



Diseño Experimental



TCAG (Library)	Region		Total
	1	2	
Raw Wells	1,097,310	1,131,343	2,228,653
Key Pass Wells	1,041,150	1,079,018	2,120,168
Passed Filter Wells	453,313	417,261	870,574
Total Bases	161,924,266	147,408,872	309,333,138
Length Average	357.83	353.87	355.32
Length Std Deviation	141.08	145.76	
Longest Reads Length	1,957	1,080	1,957
Shortest Reads Length	40	40	40
Median Reads Length	406.0	405.0	405.0

Resultado

7500€

*Análisis
Bioinformático
incluido

Secuenciación masiva en paralelo / NGS

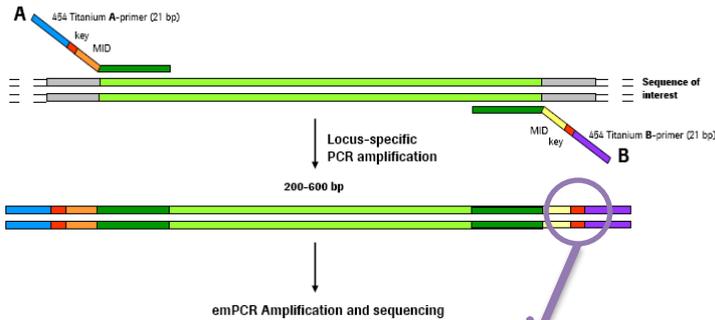
CCiT-UB- Unidad de Genómica
Tecnología GS JUNIOR



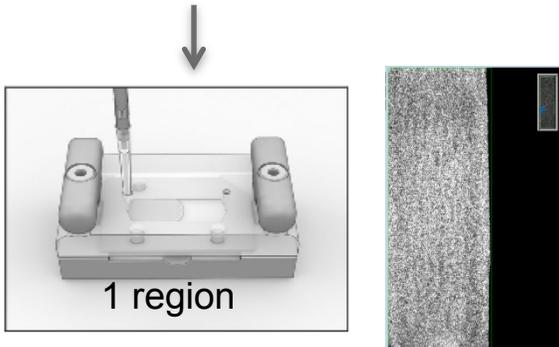
Ejemplo: Metagenómica (16S) _



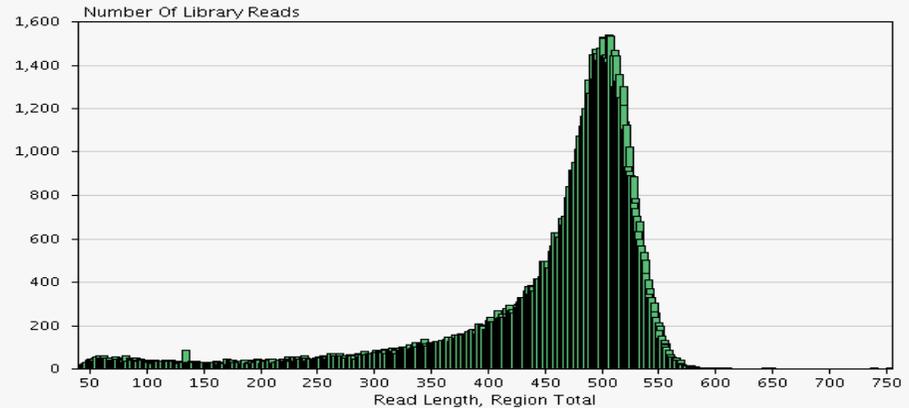
4 Soil DNA Samples + 4 primer-MID'S



4 Amplicon-MID



Diseño Experimental



GACT (Library)	Region	Total
Raw Wells	1	208,568
Key Pass Wells	1	182,595
Passed Filter Wells	1	132,550
Total Bases	1	59,932,241
Length Average	452.25	452.15
Length Std Deviation	96.82	
Longest Reads Length	755	755
Shortest Reads Length	40	40
Median Reads Length	486.0	486.0

Resultado

1000€
*Análisis
Bioinformático
incluido

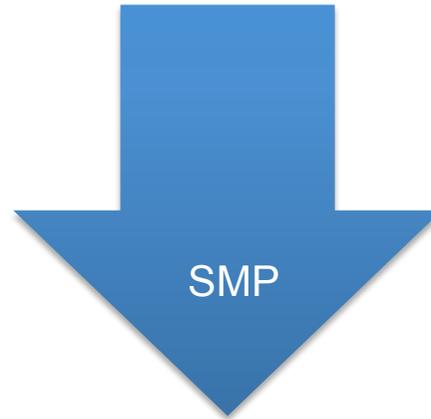


3era Generación

Secuenciación Molécula Única vs Secuenciación Masiva en Paralelo

Desde
2005...

Objetivo:
Secuenciación de
Genomas grandes
en poco tiempo y
menor coste



Ventajas:

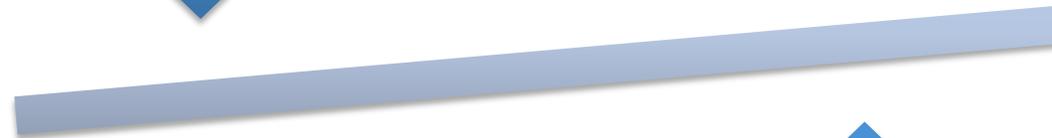
No requiere amplificación previa

Se secuencia a partir de una única molécula.

Requiere menor cantidad de muestra inicial

Fragmentos más largos (Pac Bios)

Resultado obtenemos a tiempo real.



Desventajas:

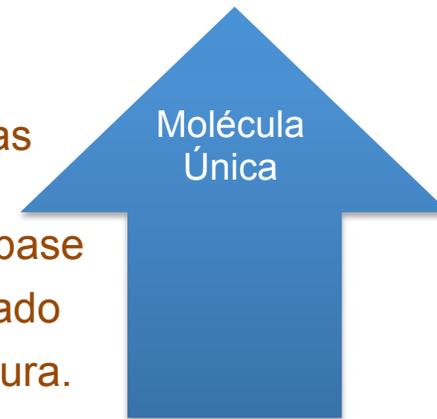
Elevado coste de las plataformas

Elevado coste por base

Ratio de error elevado

Tecnología no madura.

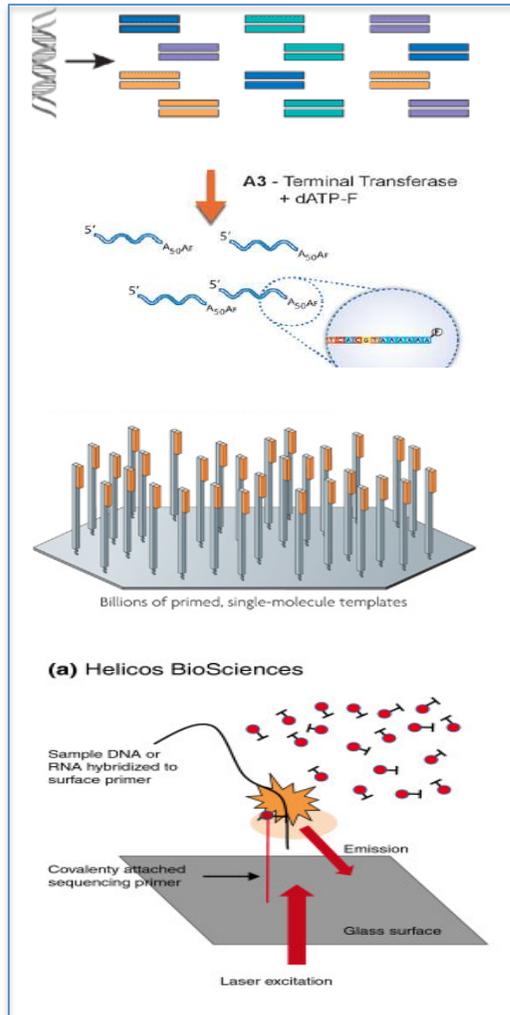
Poca oferta.



Secuenciación de Molécula Única: Tecnología

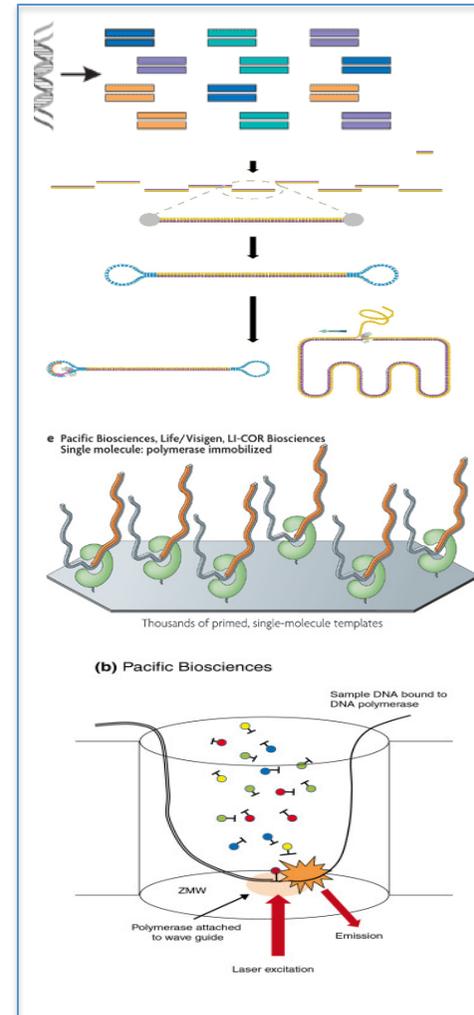
Helicos BioSciences

Single-molecule sequencing (tSMS™)



Pacific Biosciences

The technology, denoted Single-molecule Real Time Sequencing-by-synthesis (SMRT™)



Secuenciación de Molécula Única: Tecnología

Oxford Nanopore:

DNA can be sequenced by threading it through a microscopic pore in a membrane. Bases are identified by the way they affect ions flowing through the pore from one side of the membrane to the other.

1 One protein unzips the DNA helix into two strands.

2 A second protein creates a pore in the membrane and holds an "adapter" molecule.

3 A flow of ions through the pore creates a current. Each base blocks the flow to a different degree, altering the current.

4 The adapter molecule keeps bases in place long enough for them to be identified electronically.

MEMBRANE

DNA DOUBLE HELIX

Current (pA)

Time

Open pore

Blocked pore

400 sec

Residual pore current (pA)

Time (s)

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

30.4 30.6 30.8 31.0 31.2

Nanopores

Circuits ASIC et nanopores

Cartouches de séquençage

Système de séquençage Gridlon

Cluster de séquençage

Peut contenir jusqu'à 96 échantillons d'ADN à analyser en parallèle

Mini-système de séquençage sur clé USB

Logiciel de contrôle et d'exploitation

Aun en fase de pruebas!

NANOPORE

Características

- Mide corriente a través de nanoporos que es captada por un sensor.
- Cada Nt (A,T,G,C) produce una intensidad diferente.
- Secuencias largas.
- Tiempo real.
- Coste muy reducido
- Secuenciadores USB

Secuenciación Molécula Única: Aplicaciones



Estudios donde la largada de secuencia es importante

- Ensamblaje de Genomas completos
- Mutaciones estructurales
- Haplotipos
- Variaciones de número de copias. (CNV's)

Fragmentos
de 3 kb!



Pacific
Biosciences



Estudios donde la profundidad de cobertura sea lo esencial

- Resequenciación
- Exomas
- Expresión génica
- Polimorfismos puntuales. (SNP's)



Helicos
Biosciences

Tendencias

- Secuenciadores portátiles: *Nanopore technologies*
- Secuenciación a tiempo real sin requerir amplificación.
- Bajo coste (1 genoma humano < 1000 \$)
- Menor requerimiento de técnicos especialistas



Objetivo:

- Incluir la secuenciación en el diagnóstico clínico
- Medicina personalizada



¿PREGUNTAS ?

¡ Muchas Gracias por su asistencia !

CCIT-UB
Unidad de Genómica
(Campus Diagonal)

Amaya Amador
Agustín Arasanz
Berta Fusté
Ramón Seminago
Jaume Comas*

www.sequen.ub.edu

93 403 46 57



CCiT 

Centres Científics i Tecnològics
UNIVERSITAT DE BARCELONA