



Curso Teórico-Práctico

EPIDEMIOLOGÍA GENÓMICA



UNA HERRAMIENTA PARA
FORTALECER LA VIGILANCIA DE
AGENTES INFECCIOSOS



Día 2: *NGS, epidemiología aplicada, consenso y Calidad*
Bogotá / Septiembre 24 del 2024

Introducción a la secuenciación de nueva generación (NGS) y su importancia en la epidemiología Genómica

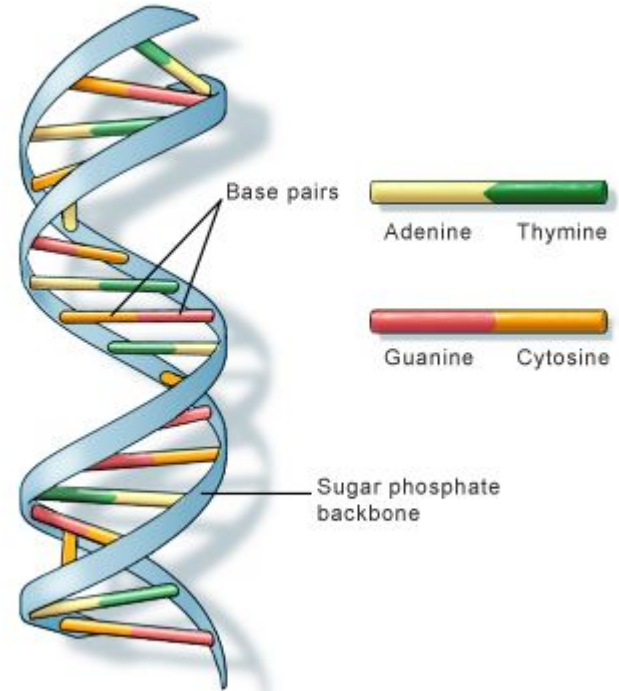
Daniel Martínez Vargas
Biólogo, M.Sc.

Grupo de Genómica de Microorganismos Emergentes, INS

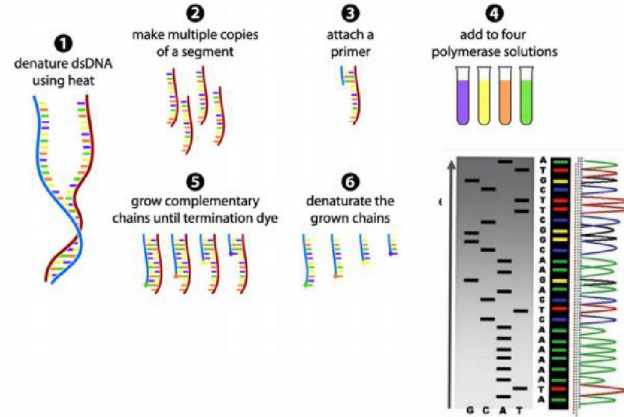
¿Qué es la secuenciación y por qué es útil?

¿Qué es la secuenciación y por qué es útil?

La secuenciación es el proceso de determinar el orden de las bases de nucleótidos en una o más moléculas de ADN o ARN.



Evolución de las tecnologías de secuenciación

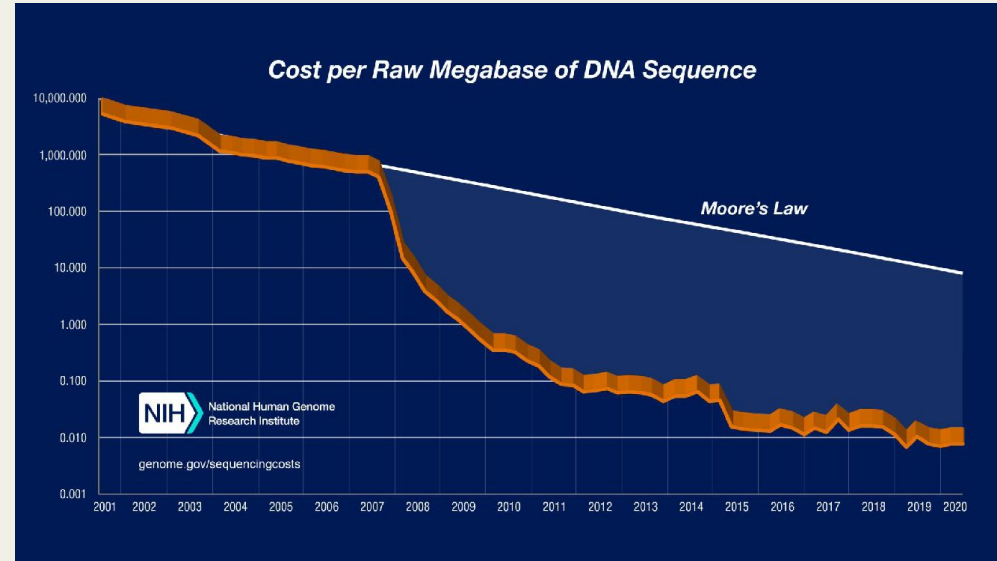


Por medio del PCR se realizan múltiples copias de una hebra usando un primer de secuenciación. La mezcla de dTTP y ddTP hacen que las cadenas de ADN se detengan en todas las posiciones

- Antes de la llegada de NGS, el método de secuenciación de Sanger se usaba ampliamente para la secuenciación de ADN. Sin embargo, el método de Sanger requería mucho tiempo, era costoso y tenía una capacidad limitada para secuenciar grandes volúmenes de ADN. Las tecnologías NGS surgieron como un gran avance a principios de la década de 2000, lo que permitió a los investigadores secuenciar millones de fragmentos de ADN en paralelo.

¿Qué es la secuenciación de próxima generación?

La secuenciación de próxima generación (NGS), también conocida como secuenciación de alto rendimiento, es un poderoso conjunto de tecnologías que nos permite secuenciar ADN y ARN con una velocidad y precisión sin precedentes. Ha revolucionado el campo de la genómica y ha contribuido en gran medida a nuestra comprensión de los procesos biológicos, la variación genética y los mecanismos de las enfermedades.



Pasos para la secuenciación

Diseño del estudio



Extracción de ADN/ARN



Preparación de la librería

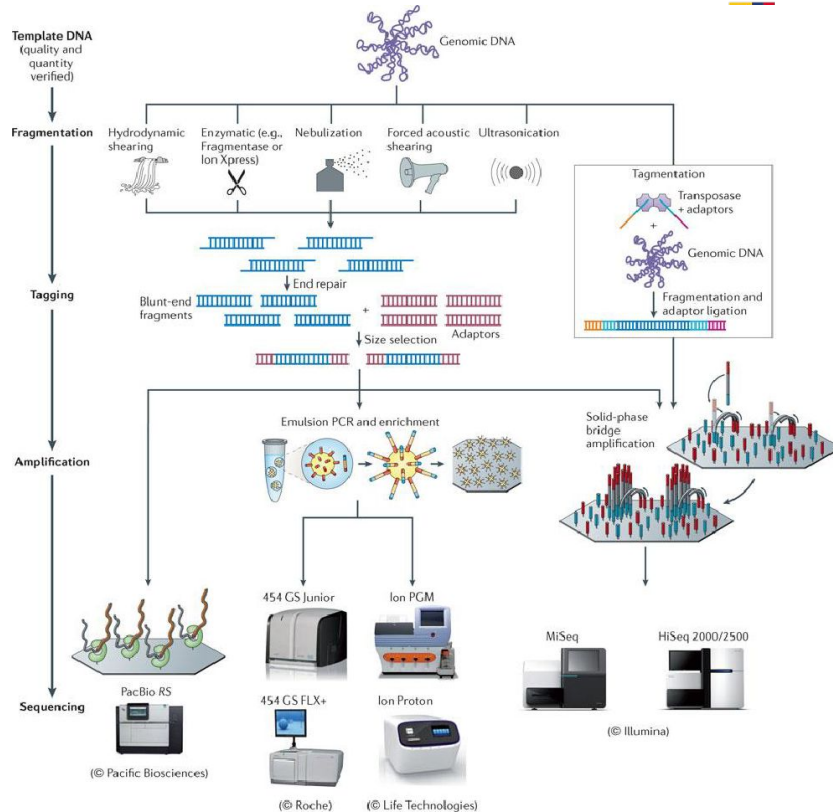


Secuenciación



Análisis de datos





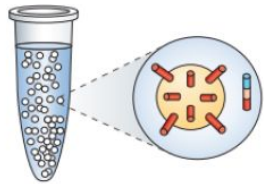
- Preparación de la librería: NGS comienza con la preparación de una librería de ADN o ARN, lo que implica fragmentar el material genético y unir adaptadores de ADN cortos a los extremos de los fragmentos.
- Plataformas de secuenciación: las plataformas NGS emplean varias tecnologías para secuenciar simultáneamente múltiples fragmentos de ADN. Algunas plataformas de uso común incluyen la secuenciación de Illumina, la secuenciación de Ion Torrent, la secuenciación de PacBio y la secuenciación de Oxford Nanopore.
- Generación de datos: las plataformas NGS generan grandes cantidades de datos de secuenciación. Estas lecturas representan pequeñas secciones de los fragmentos de ADN o ARN que se secuenciaron. La cantidad de lecturas y su calidad tienen un gran impacto en el análisis de datos posterior.

Secuenciación Masiva

Preparación de la librería

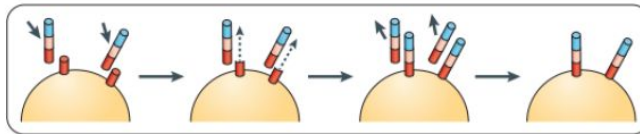
Estrategias de amplificación de fragmentos

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



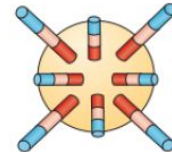
Emulsion

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

Secuenciación Masiva

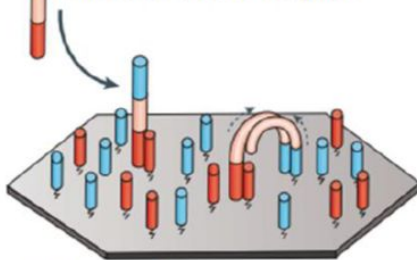
Preparación de la librería

Estrategias de amplificación de fragmentos

b Solid-phase bridge amplification (Illumina)

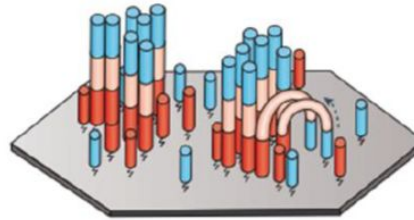
Template binding

Free templates hybridize with slide-bound adapters



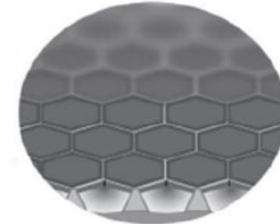
Bridge amplification

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



Cluster generation

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



Patterned flow cell

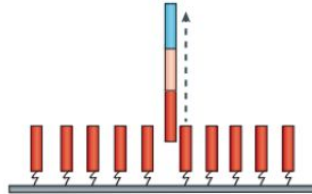
Microwells on flow cell direct cluster generation, increasing cluster density

Secuenciación Masiva

Preparación de la librería

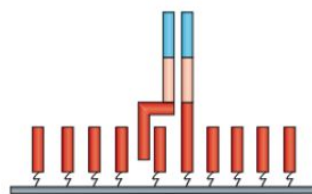
Estrategias de amplificación de fragmentos

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



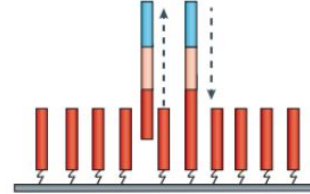
Template binding

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



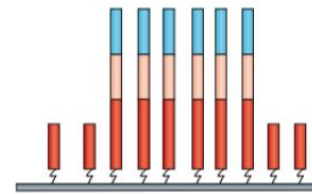
Primer walking

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

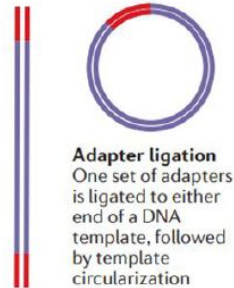
Secuenciación Masiva

Preparación de la librería

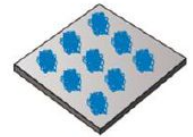
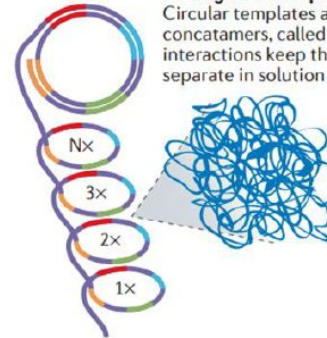
Estrategias de amplificación de fragmentos



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



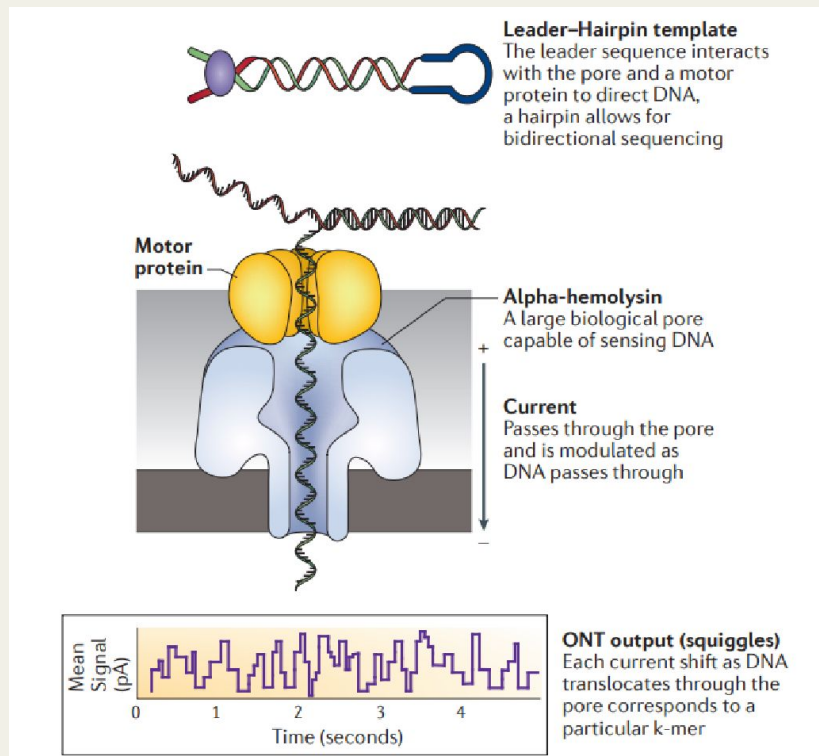
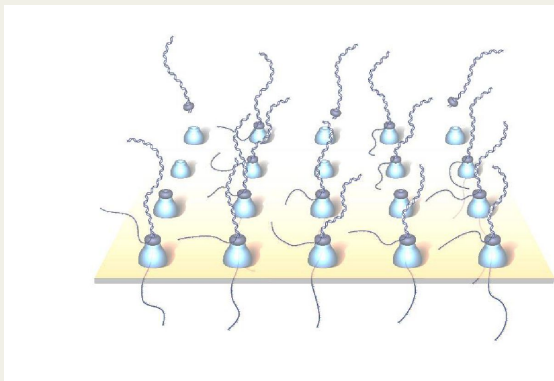
Cleavage
Circular DNA templates are cleaved downstream of the adapter sequence



Hybridization
DNA nanoballs are immobilized on a patterned flow cell

<https://youtu.be/xUVdJN0m38c>

Secuenciación sin amplificación: ONT



Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333-351.

Terminología y diferentes tipos de secuenciación microbiana

Secuenciación de Sanger

Amplificación por PCR de una región corta (~500 pb) seguida por secuenciación Sanger, esta es de bajo rendimiento. Útil para la investigación de una sola muestra y una sola región.

WGS (Secuenciación del genoma completo)

Secuenciación del genoma completo (ADN) de un solo organismo. El material inicial puede ser ADN de alto peso molecular o productos de PCR de amplicones.

Secuenciación de amplicones

Amplificación por PCR de una región(es) de un organismo(s), seguida de secuenciación NGS. Se utiliza normalmente para secuenciar regiones específicas u organismos específicos en una muestra mixta.

Secuenciación 16S

Un tipo específico de secuenciación de amplicones. Amplificación por PCR de la región 16S del ARN bacteriano seguida de secuenciación NGS. Típicamente para definir la composición bacteriana o fúngica de una muestra compleja de microbioma.

NGS metagenómica

Secuenciar todo el ácido nucleico (ya sea ADN o ARN) dentro de una muestra de forma imparcial (aleatoria, no dirigido). Normalmente se utiliza si el objetivo (los objetivos) es desconocido.

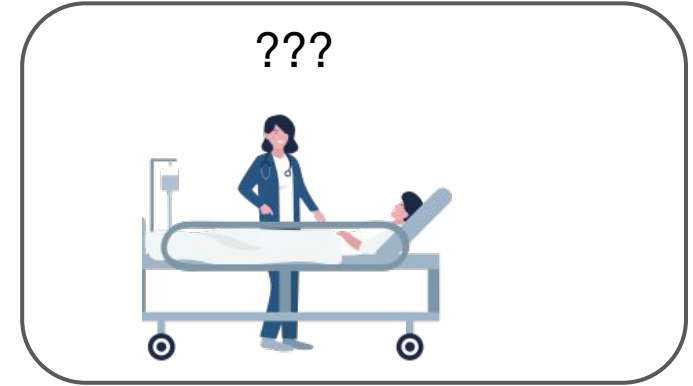
Comparación de plataformas de NGS

Manufacturer	Read length	Data output	Max. run time (hours)	Chemistry	Key applications**
Illumina (NovaSeq 6000)	300 PE	6 Tb (6000 Gb)	44	Sequencing by synthesis	SS-WGS and TGS, TGEP, 16sMGS, WES, SCP, LS-WGS, CA, MS, MGP, CFS, LBA
Thermo Fisher Scientific Ion Torrent (Ion GeneStudio S5 Prime)	600 SE	50 Gb	12	Sequencing by synthesis	WGS, WES, TGS
GenapSys (16 chips)	150 SE	2 Gb	24	Sequencing by synthesis	TS, SS-WGS, GEV, 16S rRNA sequencing, sRNA sequencing, TSCAS
QIAGEN (GeneReader)	100 SE	Not available	Not available	Sequencing by synthesis	Cancer research and identifying mutations
BGI/Complete Genomics	400 SE	6 Tb (6000 Gb)	40	DNA nanoball	Small and large WGS, WES and TGS
PacBio (HiFi Reads)	25 Kb	66.5 Gb	30	Real-time sequencing	DN sequencing, FT, identifying ASI, mutations, and EPM
Nanopore (PromethION)	4 Mb	14 Tb (14000 Gb)	72	Real-time sequencing	SV, GS, phasing, DNA and RNA base modifications, FT, and isoform detection

*Performance comparison is given as per manufacturer's description. **Applications by all sequencers of the respective manufacturer are listed. ***Full names are given in Abbreviations.

¿Para qué se utiliza la secuenciación en salud pública?

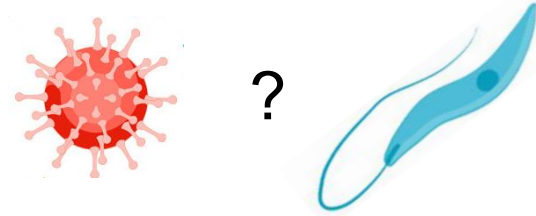
- Identificación de microbios
 - ¿Qué hay en la muestra?
- Información de secuencia de microbios
 - Subtipificación o tipificación de variantes
 - Factores de virulencia
 - Genes de resistencia a los antimicrobianos
- Vigilancia e investigaciones epidemiológicas
 - Cómo se relacionan las muestras
 - Seguimiento y monitoreo de brotes
 - Evaluación de la efectividad de las intervenciones
- Desarrollo de terapias y vacunas



¿Para qué se utiliza la secuenciación en salud pública?

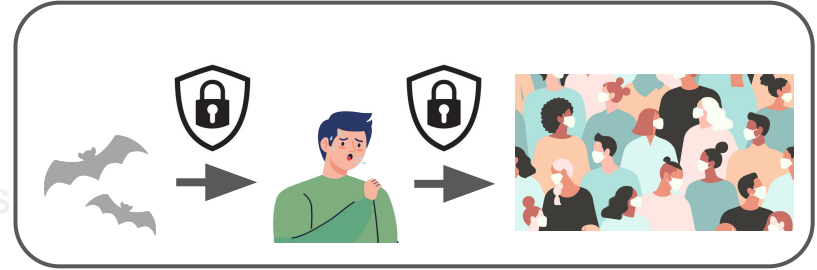
- Identificación de microbios
 - ¿Qué hay en la muestra?
- Información de secuencia de microbios
 - Subtipificación o tipificación de variantes
 - Factores de virulencia
 - Genes de resistencia a los antimicrobianos
- Vigilancia e investigaciones epidemiológicas
 - Cómo se relacionan las muestras
 - Seguimiento y monitoreo de brotes
 - Evaluación de la efectividad de las intervenciones
- Desarrollo de terapias y vacunas

Microbe characterization



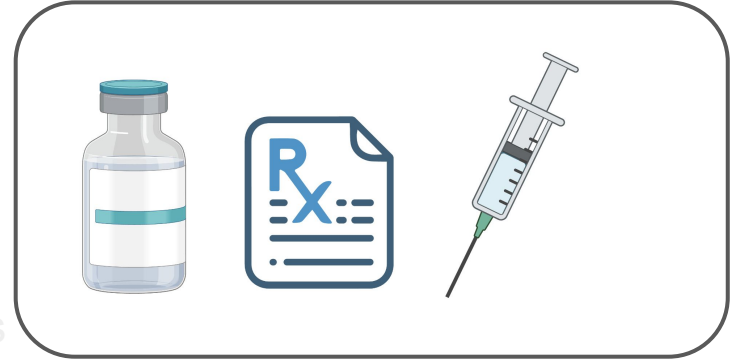
¿Para qué se utiliza la secuenciación en salud pública?

- Identificación de microbios
 - ¿Qué hay en la muestra?
- Información de secuencia de microbios
 - Subtipificación o tipificación de variantes
 - Factores de virulencia
 - Genes de resistencia a los antimicrobianos
- Vigilancia e investigaciones epidemiológicas
 - Cómo se relacionan las muestras
 - Seguimiento y monitoreo de brotes
 - Evaluación de la efectividad de las intervenciones
- Desarrollo de terapias y vacunas



¿Para qué se utiliza la secuenciación en salud pública?

- Identificación de microbios
 - ¿Qué hay en la muestra?
- Información de secuencia de microbios
 - Subtipificación o tipificación de variantes
 - Factores de virulencia
 - Genes de resistencia a los antimicrobianos
- Vigilancia e investigaciones epidemiológicas
 - Cómo se relacionan las muestras
 - Seguimiento y monitoreo de brotes
 - Evaluación de la efectividad de las intervenciones
- Desarrollo de terapias y vacunas



¿Para qué se utiliza la secuenciación en salud pública?

- Identificación de microbios
 - ¿Qué hay en la muestra?
- Información de secuencia de microbios
 - Subtipificación o tipificación de variantes
 - Factores de virulencia
 - Genes de resistencia a los antimicrobianos
- Vigilancia e investigaciones epidemiológicas
 - Cómo se relacionan las muestras
 - Seguimiento y monitoreo de brotes
 - Evaluación de la efectividad de las intervenciones
- Desarrollo de terapias y vacunas

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7008580/>

El tipo de secuenciación a utilizar depende de la pregunta que se intenta responder y de los recursos disponibles.

Secuenciación del Genoma Completo (WGS)

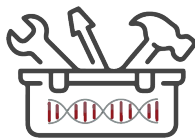
Material inicial	ADN de alto peso molecular de un aislado cultivado
Datos obtenidos	Secuenciación NGS del genoma del organismo cultivado
Fortalezas	<ul style="list-style-type: none">• No requiere el diseño de primers específicos• Buena cobertura del genoma de organismos incluso con genomas grandes
Usos	<ul style="list-style-type: none">• Obtener una cobertura del genoma completo de organismos incluso con genomas grande• Identificar o confirmar la identidad de un organismo cultivado• Identificar variantes genómicas dentro de un organismo
Limitaciones	<ul style="list-style-type: none">• Requiere un aislado cultivado<ul style="list-style-type: none">○ Requiere mucho tiempo○ No todos los organismos pueden cultivarse; muchos requieren condiciones especiales○ Más tiempo de trabajo debido a la necesidad de cultivo

Secuenciación de amplicones

Material inicial	<ol style="list-style-type: none">1. Muestra clínica directa u otro tipo de muestra; no requiere un aislado cultivado2. Primers complementarios a las regiones/ organismo de interés	
Datos obtenidos	Secuencia NGS de muchos productos PCR en paralelo; genoma del organismo de interés	
Fortalezas	<ul style="list-style-type: none">• Buena cobertura de todo el genoma de genomas relativamente pequeños• No requiere un cultivo aislado; puede ser una muestra mixta/directa• La mayoría de los datos resultantes son de la diana deseada	
Usos	<ul style="list-style-type: none">• Epidemiología genómica• Investigación de la transmisión	<ul style="list-style-type: none">• Identificación o seguimiento de variantes
Limitaciones	<ul style="list-style-type: none">• Requiere un buen diseño y validación de primers• Requiere conocimiento de la regiones objetivo• Difícil para genomas grandes	<ul style="list-style-type: none">• No permite mutaciones en las regiones en donde hay primers

Secuenciamiento del ARNr 16S/18S RNA

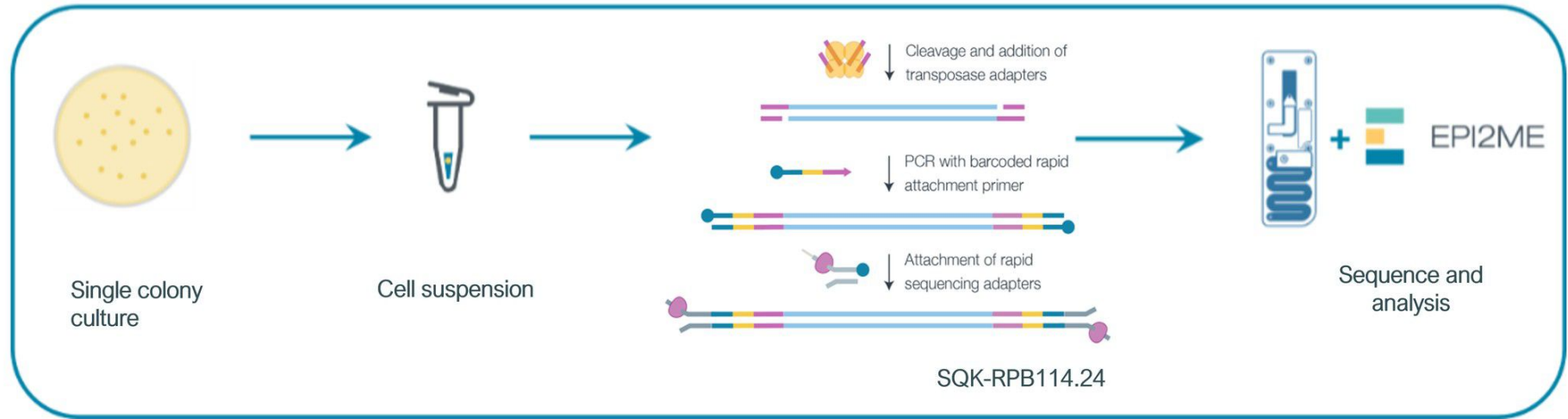
Material inicial	<ol style="list-style-type: none">1. Muestra clínica directa u otro tipo de muestra; no requiere un aislado cultivado2. Primers específicos 16S/18S	
Datos obtenidos	NGS de las secuencias de ARNr de la región 16S/18S de bacterias, hongos y microbios eucariotas	
Fortalezas	<ul style="list-style-type: none">• Buena cobertura del ARNr 16S/18S• No requiere un cultivo aislado; puede ser una muestra mixta/directa	
Usos	<ul style="list-style-type: none">• Caracterización e identificación de un microbioma o muestra ambiental• Estudios exploratorios para la identificación de microbios, no del genoma completo• Bueno para muestras con alto contenido y diversidad microbiana	
Limitaciones	<ul style="list-style-type: none">• No es una secuenciación del genoma completo• No identifica virus	<ul style="list-style-type: none">• La especificidad de la identificación de microbios está limitada por la singularidad de la región 16S/18S

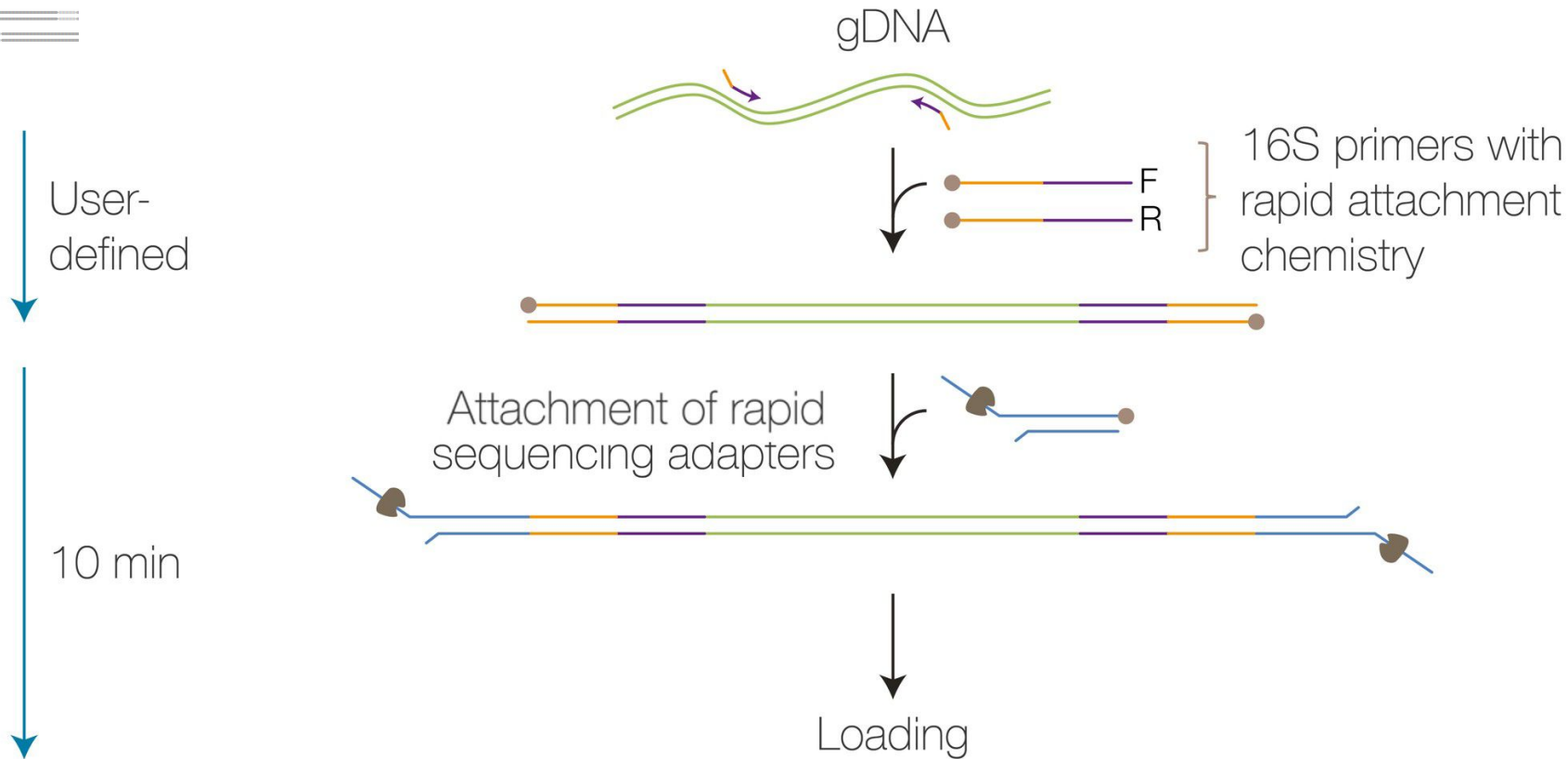
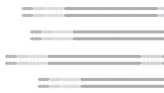


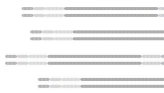
NGS para el estudio de enfermedades infecciosas

	WGS y transcriptoma	Amplicones	rARN 16S & 18S	mNGS
Material de inicio	RNA/DNA aislado de cultivo	RNA/DNA aislado de muestra	RNA/DNA aislado de muestra	RNA/DNA aislado de muestra
Primers	Primers aleatorios	Primers específicos	Primers 16S/18S	Primers aleatorios
Qué se secuencía	Organismo de interés	Organismo de interés	Solo rARN	ADN/ARN de todo lo que está presente en la muestra
Resultados	Secuencia de transcritos o genoma de organismos de interés	ADN/ARN de un organismo específico	rRNA de bacterias, hongos y eucariotas	ADN/ARN de todos los organismos
Casos	Investigación básica de organismo de interés	Genomas completos de muestras clínicas para epidemiología genética	Estudios exploratorios de muestras ambientales	Investigación de brotes, descubrimientos de patógenos

Preparación de librerías mediante ONT



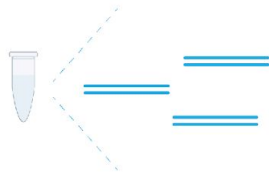




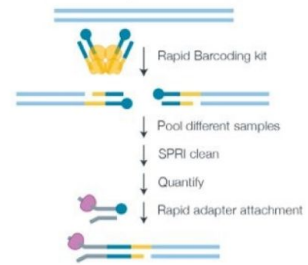
PCR



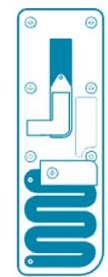
(Recommended)
Clean-up



Library Prep



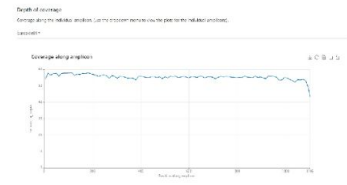
Sequencing

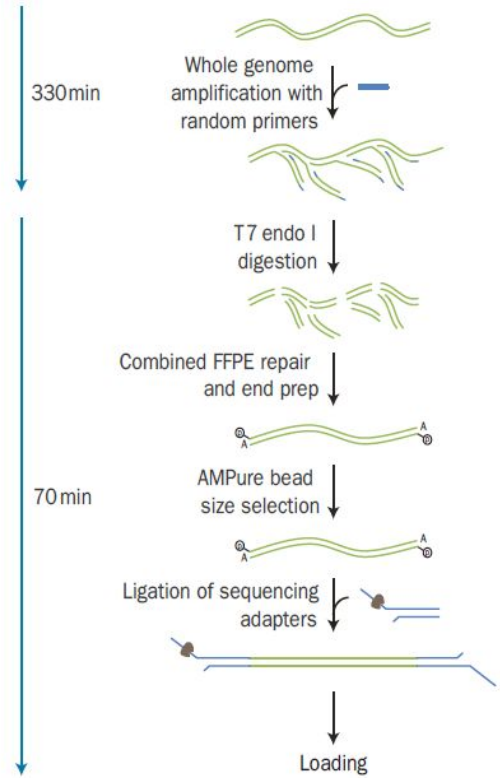


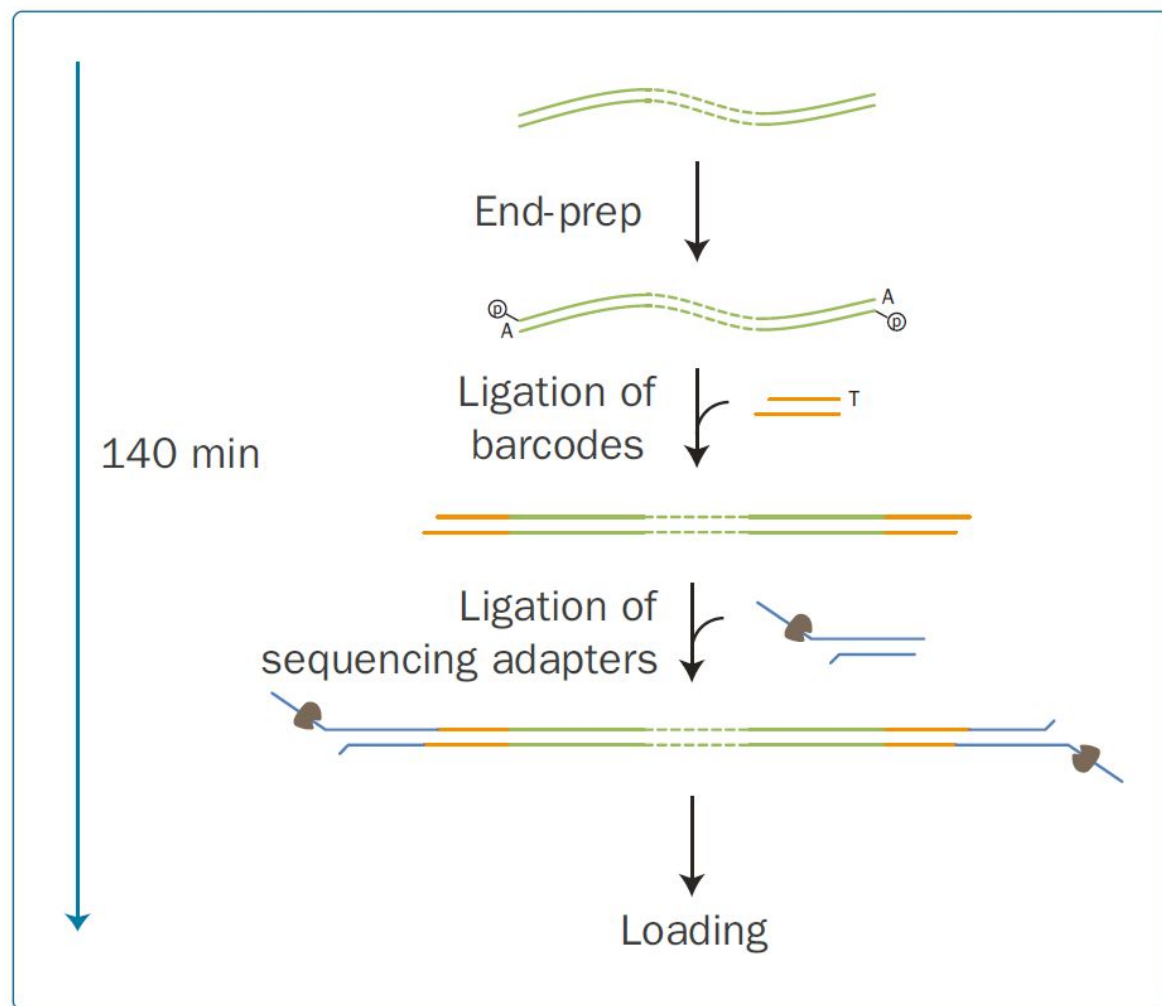
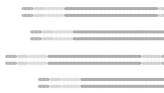
Downstream
analysis



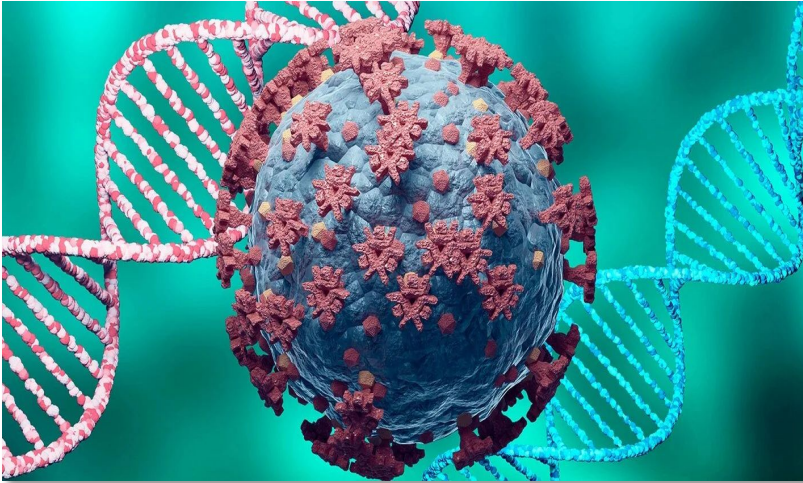
EPI2ME





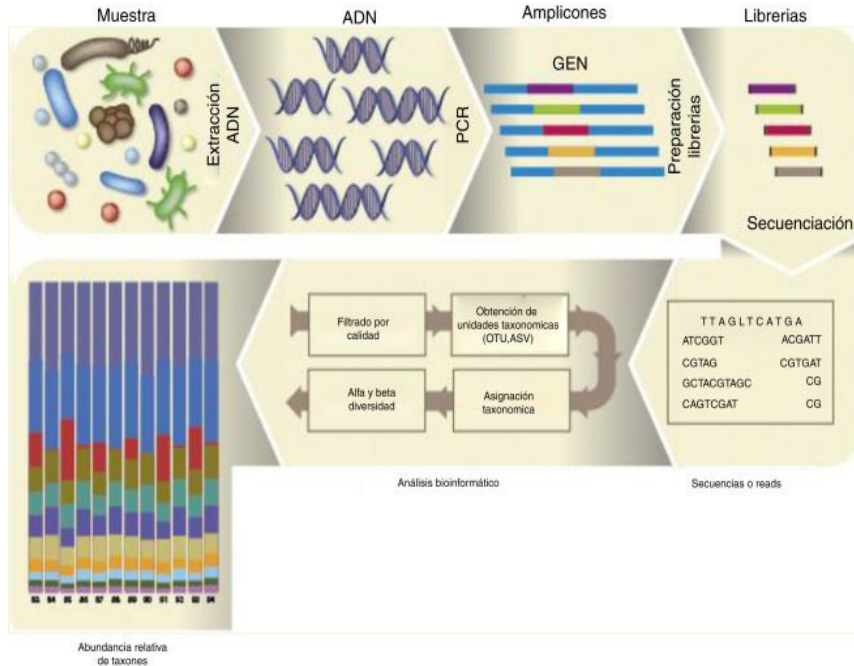


Epidemiología Genómica



- Disciplina que estudia la variabilidad genética de los patógenos y su relación con la epidemiología de las enfermedades.
- Analiza cómo la genética de los patógenos influye en su transmisión, virulencia, y respuesta al tratamiento.

<https://www.paho.org/es/noticias/21-7-2021-red-regional-vigilancia-genomica-rastrea-variantes-virus-sars-cov-2-toda-america>



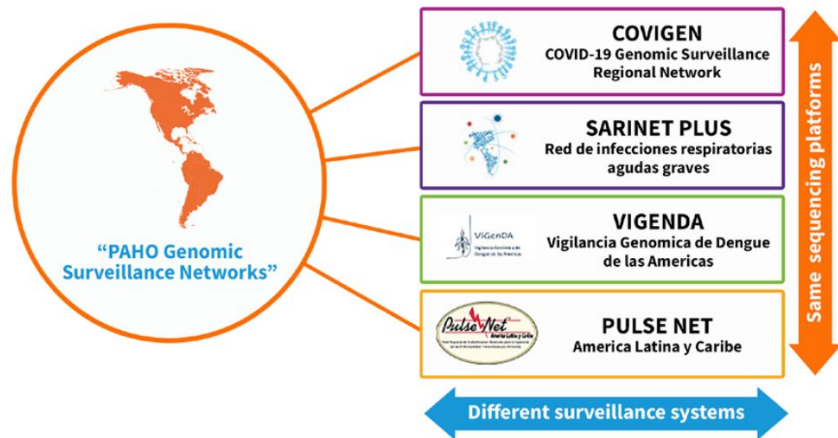
<https://doi.org/10.1016/j.ram.2019.06.003>

Ventajas de NGS en la Epidemiología Genómica

- **Resolución Genómica Completa:** Permite identificar mutaciones y variaciones genéticas con gran precisión.
- **Alta Sensibilidad:** Detecta patógenos en muestras con baja carga viral o bacteriana.
- **Rapidez y Escalabilidad:** Permite la secuenciación de múltiples muestras simultáneamente.

Impacto en la Salud Pública

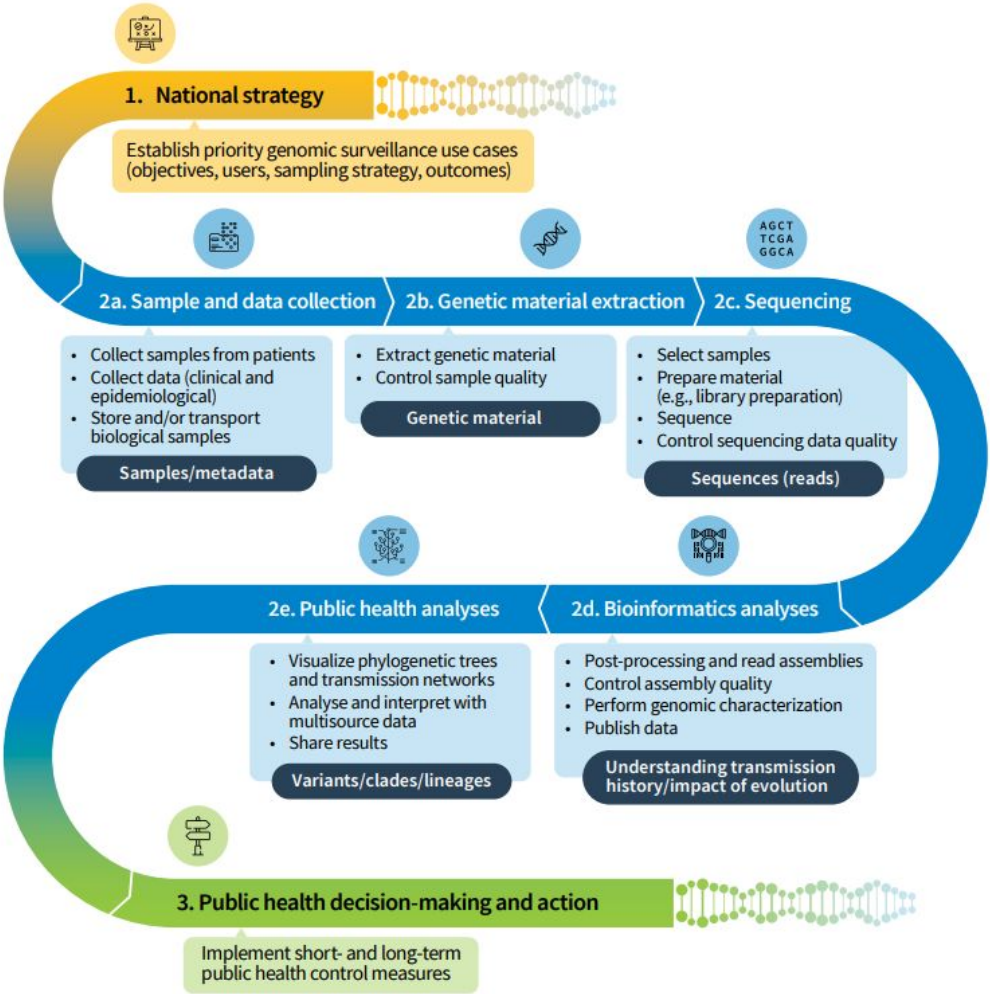
Mejora en la Toma de Decisiones



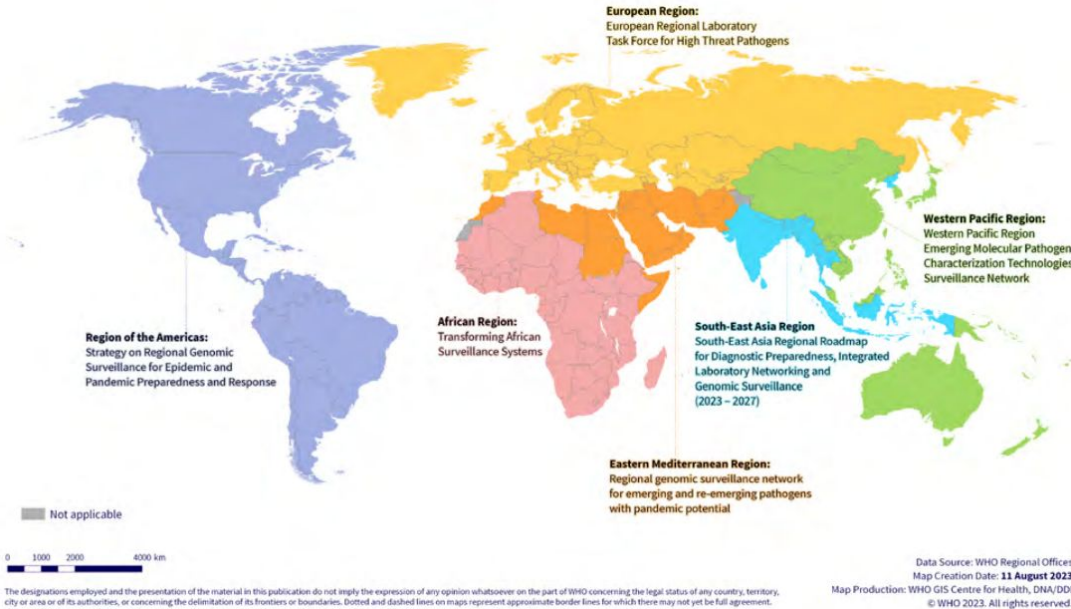
<https://iris.who.int/bitstream/handle/10665/374757/9789240084773-eng.pdf?sequence=1>

- **Datos en Tiempo Real:** Los datos genómicos permiten tomar decisiones informadas rápidamente durante un brote.
- **Personalización del Tratamiento:** Mejora la precisión en el diagnóstico y en la elección del tratamiento más efectivo.
- **Desarrollo de Vacunas:** Identificación de objetivos genéticos para el desarrollo de nuevas vacunas.

Descripción general de la cadena de valor de la vigilancia genómica con sus actividades y resultados correspondientes



Desafíos y Consideraciones



<https://iris.who.int/bitstream/handle/10665/374757/9789240084773-eng.pdf?sequence=1>

Retos de Implementación

- **Costos:** A pesar de la reducción en los costos, la NGS sigue siendo una tecnología cara.
- **Infraestructura:** Requiere equipamiento especializado y personal capacitado.
- **Interpretación de Datos:** La gran cantidad de datos generados requiere de bioinformáticos y analistas para su interpretación.

Futuro de la NGS en Epidemiología Genómica

Avances Tecnológicos

- Secuenciación más rápida y económica.
- Mayor integración de la NGS en la vigilancia epidemiológica rutinaria.
- Desarrollo de nuevas herramientas de análisis para gestionar la complejidad de los datos.



Goal

Genomic surveillance is strengthened and scaled for quality, timely and appropriate public health actions within local to global surveillance systems

Objectives



Objective 1

Improve access to tools for better geographic representation



Objective 2

Strengthen the workforce to deliver at speed, scale and quality



Objective 3

Enhance data utility for streamlined local to global public health decision-making and action



Objective 4

Maximize connectivity for timely value-add in the broader surveillance architecture



Objective 5

Maintain a readiness posture for emergencies

Strategic actions

- **Map and monitor capacity landscape**
- Deliver contextualized **technology and innovation solutions**
- **Stimulate innovation and research** to address local to global needs
- **Shape a sustainable and quality market to maximize access**

- Training packages in **genomics and bioinformatics**
- Promote **community of practice to disseminate & share good practices**
- **External quality assurance programs** for genomics and analytics
- **Knowledge exchange** to build or strengthen capacity

- **Meta data** standards
- **Data sharing** principles
- **Data sharing** agreements
- **Harmonized** protocols, norms, standards and reference materials

- **Facilitated data, specimen and information sharing**
- **Increase network linkages** at local, regional and global levels
- Targeted collaboration with **One Health partners**

- **Surge exercises** to test systems
- **Joint projects** for maintaining capabilities and capacities
- **After action reviews** following events to assess & address lessons learnt

Conclusiones

Importancia Clave de la NGS

- La NGS es una herramienta esencial en la epidemiología genómica.
- Su capacidad para proporcionar datos genómicos detallados permite una mejor comprensión y control de las enfermedades infecciosas.
- La inversión en NGS y su integración en la salud pública es crucial para la preparación ante futuras pandemias.

Gracias

Contacto:

dmartinezv@ins.gov.co

Algunas de las diapositivas usadas en esta presentación hacen parte del material disponible en Help Center de CZID

<https://chanzuckerberg.zendesk.com/hc/en-us>