

# Fundamentos y aplicaciones biomédicas de las principales tecnologías de secuenciación: una revisión de literatura

Anny Jineth Camargo Mancipe<sup>1</sup> , Karen Nattaly Valero González<sup>2</sup> , Alida Marcela Gómez Rodríguez<sup>1</sup> , Diego Fernando Camargo Mancipe<sup>1</sup> , Carlos Fernando Suárez Martínez<sup>3</sup> ,  
Laura Esperanza Cuy Chaparro<sup>1</sup> 

## RESUMEN

**Introducción:** el objetivo de la secuenciación es determinar la composición de los nucleótidos presentes en el ADN o el ARN. Desde la finalización del proyecto genoma humano, surgieron diversas tecnologías de secuenciación rápida como Roche 454, SOLiD, Illumina, Ion Torrent, PacBio y Oxford Nanopore, más precisas y costoeficientes, que permiten desarrollar proyectos a gran escala y estudiar genes y genomas, la composición de microbiomas, enfermedades metabólicas y enfermedades genéticas que afectan a la población.

**Objetivo:** describir los fundamentos de los métodos de secuenciación de ADN y sus aplicaciones en las ciencias biomédicas.

**Métodos:** revisión descriptiva de las principales estrategias de secuenciación de ADN de primera, segunda y tercera generación y su aplicación en el entorno biomédico, a partir de la búsqueda de artículos en bases de datos electrónicas especializadas en investigación científica. Se encontraron 118 documentos, de los cuales se excluyeron 6, por no cumplir con los criterios de inclusión, y se seleccionaron 112, por cumplir con todos los requisitos.

**Conclusiones:** el surgimiento de los métodos de secuenciación de siguiente generación arroja una gran cantidad de datos, incluidos genomas secuenciados completamente de varias especies, con un rendimiento extenso, tiempos reducidos y costoeficiencia, que lleva a la completa transformación de las ciencias de la vida y logra un progreso sin precedentes en el análisis de genomas, la evaluación de la ecología microbiana y el diagnóstico de enfermedades.

**Palabras clave:** secuenciación de nucleótidos de alto rendimiento; análisis de secuencia, ADN; genoma humano; técnicas genéticas.

<sup>1</sup> Universidad de Boyacá (Tunja, Colombia).

<sup>2</sup> Universidad Nacional de Colombia (sede Bogotá, Colombia).

<sup>3</sup> Universidad del Rosario (Tunja, Colombia)

**Autora de correspondencia:** Anny Jineth Camargo Mancipe. Correo electrónico: [anncamargo@uniboyaca.edu.co](mailto:anncamargo@uniboyaca.edu.co)

**Citar este artículo así:**

Camargo Mancipe AJ, Valero González KN, Gómez Rodríguez AM, Camargo Mancipe DF, Suárez Martínez CF, Cuy Chaparro LE. Fundamentos y aplicaciones biomédicas de las principales tecnologías de secuenciación: una revisión de literatura. Rev Investig Salud Univ Boyacá. 2020;7(2):138-172. doi: <https://doi.org/10.24267/23897325.498>

## Biomedical Foundations and Applications of Major Sequencing Technologies: A Literature Review

### ABSTRACT

**Introduction:** The purpose of sequencing is to determine the composition of the nucleotides present in DNA or RNA. Since the completion of the human genome project, several sequencing technologies such as Roche 454, SOLiD, Illumina, Ion Torrent, PacBio and Oxford Nanopore have emerged as tools for rapid sequencing, with greater precision and cost-efficiency, allowing the development of large-scale projects and the study of genes and genomes, along with the composition of microbiomes and the study of metabolic and genetic diseases that affect the population.

**Objective:** To describe the foundations of the methods of DNA sequencing and their applications in the biomedical sciences.

**Methods:** Descriptive review of the main strategies of first, second and third generation DNA sequencing and their application in the biomedical environment. This review was carried out by searching articles in electronic databases specialized in scientific research. A total of 118 papers were found, of which 6 were excluded as they did not meet the inclusion criteria and 112 were selected as meeting all the requirements.

**Conclusions:** The emergence of next-generation sequencing methods yielding a wealth of data, including fully sequenced genomes of various species, with extensive throughput, reduced time and cost-effectiveness that has led to the complete transformation of the life sciences, achieving unprecedented progress in genome analysis, assessment of microbial ecology and disease diagnosis.

**Keywords:** high-throughput nucleotide sequencing; sequence analysis, DNA; genome humans; genetic techniques.

## Fundamentos e aplicações biomédicas das principais tecnologias de sequenciamento: uma re-visão da literatura

### RESUMO

**Introdução:** o objetivo do sequenciamento é determinar a composição dos nucleotídeos presentes no DNA ou RNA. Desde a conclusão do projeto do genoma humano, surgiram diversas tecnologias de sequenciamento rápida como Roche 454, SOLiD, Illumina, Ion Torrent, PacBio e Oxford Nanopore, mais precisas e econômicas, que permitem o desenvolvimento de projetos de grande escala e estudo de genes e genomas, composição de microbiomas, doenças metabólicas e genéticas que afetam a população.

**Objetivo:** descrever os fundamentos dos métodos de sequenciamento de DNA e suas aplicações nas ciências biomédicas.

**Métodos:** revisão descritiva das principais estratégias de sequenciamento de DNA de primeira, segunda e terceira geração e sua aplicação no ambiente biomédico, a partir da busca de artigos em bases de dados eletrônicas especializadas em pesquisa científica. Foram encontrados 118 documentos, dos quais 6 foram excluídos por não atenderem aos critérios de inclusão e 112 foram selecionados por atenderem a todos os requerimentos.

**Conclusões:** o surgimento de métodos de sequenciamento de próxima geração rende uma riqueza de dados, incluindo genomas totalmente se-quenciados de várias espécies, com produção extensa, tempos reduzidos e eficiência de custo, levando à transformação completa das ciências da vida e alcançando um progresso sem precedentes no genoma análise, avaliação de ecologia microbiana e diagnóstico de doenças.

**Palavras-chave:** sequenciamento de nucleotídeos de alto rendimento; análise de sequência, DNA; genoma humano; técnicas genéticas.

## INTRODUCCIÓN

Los avances alcanzados en las últimas décadas con las tecnologías de secuenciación de nucleótidos como la secuenciación genómica, la secuenciación del exoma completo y la secuenciación de ARN o RNA-seq (del inglés *RNA sequencing*), que ofrece la secuenciación de todo el transcriptoma, brindan a los investigadores y clínicos una oportunidad sin precedentes para entender e investigar las variaciones genéticas implicadas en el desarrollo de la enfermedad (1). El conocimiento del genoma humano y las tecnologías de secuenciación de ADN han permitido evaluar alteraciones cromosómicas como los polimorfismos que afectan un solo nucleótido. Estas y otras mutaciones genéticas en los individuos pueden usarse para predecir diagnósticos y subclasificar enfermedades; ejemplo de ello es el caso de algunas leucemias hematopoyéticas, en las cuales mutaciones en los genes NPM1 y CEBPA se asocian con la leucemia mieloide aguda (2) o mutaciones en los genes TET2, DNMT3a, ASXL1, EZH2, SF3B1, SRSF2, U2AF1, ZRSR2, TP53 o STAG2 se asocian con la presentación del síndrome mielodisplásico (3).

Otro ejemplo es la identificación de defectos genéticos, como en la hemofilia A (4), la hemofilia B y la enfermedad de Von Willebrand, donde, gracias a las tecnologías de secuenciación de siguiente generación, se puede estudiar un grupo de entre 50 y 100 genes y agilizar en gran medida el diagnóstico (5).

En cuanto al enfoque farmacogenómico, con el estudio genómico se evalúa la respuesta de diversos medicamentos basados en genes para enfermedades crónicas, como la diabetes mellitus, la hipertensión arterial, las enfermedades cardíacas, los trastornos autoinmunes, las enfermedades mentales, los tumores, entre otras, que sin duda es un gran paso en la medicina moderna (6).

Con el advenimiento de la primera tecnología de secuenciación, denominada *método de Sanger*, se sentaron las bases para desarrollar tecnologías de siguiente generación, como Roche 454, SOLiD, Ion Torrent, Illumina, y de tercera generación, como PacBio y Oxford Nanopore, que han abierto nuevas posibilidades para los análisis de ADN (7). Según Verified Market Research, el mercado de la secuenciación de próxima generación se valoró en 1,26 mil millones de dólares en 2018 y se proyecta que para el 2026 alcance unos 5,47 mil millones de dólares. La aparición de las tecnologías de segunda y tercera generación ha permitido reducir marcadamente los tiempos y los costos de secuenciación. Hace casi dos décadas, los científicos obtuvieron la primera secuencia del genoma humano a un precio de más de ~2,7 mil millones de dólares, un proyecto que tomó al menos 13 años. Actualmente, secuenciar un genoma completo con las nuevas tecnologías de secuenciación tarda unos pocos días y cuesta menos de 1000 dólares (8-10).

Los avances en las tecnologías y la mejora en las estrategias de secuenciación siguen bajando los costos y abren paso a nuevas aplicaciones en la investigación biomédica. Con ello es posible predecir el riesgo de enfermedad, evaluar la respuesta o efectos adversos a un medicamento con oportunidades de pronóstico y orientar hacia tratamientos más precisos (11).

Teniendo en cuenta lo anterior, el objetivo de este artículo es revisar las principales características de las técnicas de secuenciación de ADN y ARN de primera, segunda y tercera generación; sus aplicaciones en estudios de genómica y transcrip-tómica, y algunos de los avances logrados en el diagnóstico y tratamiento de enfermedades.

## METODOLOGÍA

Se llevó a cabo una revisión descriptiva sobre los principales métodos de secuenciación y sus aplicaciones. La literatura científica se buscó en las siguientes bases de datos electrónicas: PubMed, Elsevier, LILACS, PMC, MDPI y Redalyc, y con pesquisas secundarias en las listas de referencias bibliográficas de los artículos identificados.

Las palabras clave se validaron mediante el descriptor *Medical Subject Headings (MeSH)*, en idioma inglés, y *Descriptores en Ciencias de la Salud (DeCS)*, en español: *high-throughput nucleotide sequencing; sequence analysis, DNA; genome, human; genetic techniques*.

Se seleccionaron todos los artículos en inglés y español, publicados entre 1977 y 2020, que describieran los principales métodos de secuenciación de primera, segunda y tercera generación; así como sus importantes aplicaciones en el campo biomédico. Se encontraron 118 artículos, de los cuales se seleccionaron 112, por título, materiales y métodos y resumen, y se excluyeron 6, por no cumplir con los criterios de inclusión, entre ellos la fecha de publicación reciente o un resumen conciso y enfocado en el tema de estudio. En total, se analizaron 112 referencias que presentaban los temas detallados.

## DESARROLLO

### Métodos de secuenciación

Los métodos de terminación de cadena desarrollados por Sanger et al. (12) fueron la base para la primera generación de secuencias de ADN, y ello abrió paso a la secuenciación del genoma humano, un gran avance para la ciencia, pero con un alto impacto económico (13,14).

Posterior al advenimiento del método de Sanger, en 2005 las tecnologías de siguiente generación —o también llamadas de segunda generación— se abrieron paso en el mercado. Así, Roche 454 instauró el primer secuenciador de alto rendimiento, que logró la secuenciación de millones de lecturas de ADN sin la necesidad de clonar

(15). Sin embargo, la alta tasa de error llevó a la mejora de la química de secuenciación y a la aparición de otras tecnologías de alta resolución de lectura corta 50-400 pb, como Illumina, y de lectura larga de tercera generación, como PacBio y Oxford Nanopore (16).

En esta revisión se tratan los principales métodos de secuenciación a través de la historia. En cuanto a los de primera generación, se describe el método de Sanger; de segunda generación, el método Roche 454 e Illumina, y dentro de los métodos de tercera generación, las plataformas PacBio y Oxford Nanopore. Además, se describen brevemente sus principales aplicaciones en el campo biomédico.

### **Secuenciación de primera generación: método de Sanger**

Durante 1976 se describieron dos métodos de secuenciación de ADN: el método de terminación de cadena, desarrollado por Sanger y Coulson, y el método de escisión química, desarrollado por Maxam y Gilbert (17,18).

El método de Sanger, durante muchos años considerado el método de referencia (*gold standard*), produce secuencias de ADN de alta calidad con longitudes entre 500 y 1000 pares de bases. En este se utiliza una técnica enzimática de terminación de cadena donde los nucleótidos modificados

llamados *didesoxi-nucleótidos* (ddNTP) carecen de un grupo 3'-OH que le impida a la ADN polimerasa formar un enlace fosfodiéster entre el grupo 3'-OH libre del último nucleótido y el 5'-fosfato del siguiente, a fin de detener la adición de más nucleótidos y obtener, de este modo, secuencias de diferentes longitudes aleatorias (19).

En la secuenciación manual de Sanger, cada uno de los ddNTP (ddATP, ddTTP, ddGTP y ddCTP) son agregados en una reacción en cadena de polimerasa (PCR) diferente. La síntesis termina cuando se incorpora un ddNTP a una cadena de nucleótidos, separada mediante gel de electroforesis en su versión más básica. El fragmento más corto termina en el primer nucleótido desde el extremo 5' y, así sucesivamente, hasta establecer la secuencia de 5' a 3'. En contraste, en la secuenciación automatizada de Sanger, cada ddNTP se marca con un fluoróforo diferente, lo que permite mezclarlos en la misma reacción. Posteriormente, se lleva a cabo una electroforesis capilar y un computador, a través de un láser, excita cada fluoróforo para identificar cada ddNTP (20).

A pesar de algunas de sus limitaciones, como su alto consumo de tiempo (días) y elevados costos (~500 dólares por millón de bases + reactivos), debido a la necesidad de procedimientos adicionales para el aislamiento de fragmentos generalmente, por técnicas como PCR o clonación, el método de Sanger sigue siendo una herramienta

molecular útil para las pruebas genéticas, estudios de genética evolutiva, genómica funcional, identificación de variantes genéticas en diversas enfermedades, entre otros (16,19).

### Secuenciación de segunda generación

Las tecnologías de secuenciación de siguiente generación facilitan la manera de secuenciar grandes cantidades de ADN. Se caracterizan por amplificar millones de lecturas en paralelo; generalmente, utilizan pozos individuales como soporte para las reacciones de secuenciación (19).

Estas tecnologías ofrecen mejor rendimiento, bajo costo y menor tiempo de corrida; sin embargo, presentan mayores tasas de error (tasa de error secuenciación por Roche 454: ~1%, SOLiD: ~5%, illumina: ~0,1%) con respecto a la secuenciación de Sanger de primera generación (tasa de error: ~0,001) (21) (tabla 1). Dentro de las plataformas de secuenciación de próxima generación más importantes se encuentran: Roche 454, Ion Torrent e Illumina, las cuales pueden usarse para obtener secuencias de genomas completos o restringirse a regiones genéticas de interés. La preparación de las librerías se realiza de manera análoga; pero las plataformas 454, SOLiD e Ion Torrent usan PCR en emulsión, en tanto que la plataforma Illumina utiliza PCR en fase sólida (16).

### Secuenciación Roche 454

La plataforma 454 fue la primera tecnología de secuenciación de próxima generación implementada en el mercado. Esta combina la síntesis de ADN *in vitro* mediante PCR en emulsión (22) y la técnica de secuenciación por síntesis, denominada *pirosecuenciación* (23). Fue descrita por primera vez por Pal Nyren y su equipo, en 1987 (24). En 2005, fue comercializada por la empresa 454 de Life Sciences (25)B>, y en 2007, la compañía Roche adquirió la empresa. Sin embargo, dado su bajo rendimiento, la plataforma dejó de ser competitiva y comercializada en 2013 (26).

Durante la preparación de la librería, el ADN se fragmenta enzimáticamente por sonicación (24); después se incorporan adaptadores de doble cadena a los extremos de los fragmentos de longitud de 400-700 pb, mediante la acción de una ADN ligasa. Cada uno de estos fragmentos tendrá un adaptador complementario flanqueado en la perla del andamiaje. El adaptador localizado en el extremo opuesto servirá de anclaje de la ADN polimerasa y de inicio de la replicación. Una vez obtenida la librería, se desnaturalizan los fragmentos para obtener secuencias monocatenarias. El adaptador incorporado hibrida con el adaptador acoplado a la perla en una solución diluida que permite la inmovilización de una sola molécula de ADN en cada perla. Posteriormente, se coloca el conjunto de perlas en una emulsión de agua en

**Tabla 1.** Principales características de las plataformas de secuenciación

Ref.	Generación	Plataforma de secuenciación	Método de secuenciación	Longitud de lectura (pb)	Tasa de error (%)	Tiempo de corrida	Costo por millón de bases (dólares)	Ventajas	Desventajas
Metzker et al., 2010	Primera generación	Sanger	Terminación de cadena didesoxi	600-1000	0,001	~2h	~500	<ul style="list-style-type: none"> <li><input type="checkbox"/> -Lecturas largas ~1000 pb</li> <li><input type="checkbox"/> Baja tasa de error</li> <li><input type="checkbox"/> No se requiere amplificación por PCR</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Costos altos, por necesidad de procedimientos adicionales, como clonación</li> <li><input type="checkbox"/> Bajo rendimiento</li> </ul>
		Roche 454	Química de pirosecuenciación	700	1	~24h	~8,57		
	SOLiD	Ligación secuencial de sondas de oligonucleótidos	2 × 60	5	~8 días	~0,11	<ul style="list-style-type: none"> <li><input type="checkbox"/> Alto rendimiento</li> <li><input type="checkbox"/> Profundidad en las lecturas</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Lecturas más cortas que los obtenidos por Sanger (100-500 pb)</li> </ul>	
Derocles et al., 2018	Segunda generación	Ion Torrent	Síntesis por polimerasa	200	1	~2-4 h	~0,10	<ul style="list-style-type: none"> <li><input type="checkbox"/> Bajo costo y baja tasa de error</li> <li><input type="checkbox"/> Rápido tiempo de respuesta a las ejecuciones</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Se requiere amplificación por PCR</li> <li><input type="checkbox"/> Ensamblajes altamente fragmentados</li> </ul>
		Illumina	MiSeq	Síntesis	2 × 300	0,1	4-5 h		
HiSeq	Síntesis		2 × 150	0,1	2-4 días	~0,03-0,15			
NovaSeq	Síntesis		2 × 150	0,1	16-36 h				
Kchouk et al., 2018	Tercera generación	PacBio	Secuenciación en tiempo real de una sola molécula (síntesis)	~10.000-15.000	13	20 min-4 h	~0,40-0,80	<ul style="list-style-type: none"> <li><input type="checkbox"/> Lecturas largas, &gt;14 kb</li> <li><input type="checkbox"/> En el caso de MinION, fácil portabilidad</li> <li><input type="checkbox"/> No requiere amplificación</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Alta tasa de error</li> </ul>

	Componente Morfológico					Componente Muscular					Componente Motor					Componente Cardiorrespiratorio				
	Saludable (%)	No saludable (%)	P valor	OR		Saludable (%)	No saludable (%)	P valor	OR		Saludable (%)	No saludable (%)	P valor	OR		Saludable (%)	No saludable (%)	P valor	OR	
				Lim inf	Lim sup				Lim inf	Lim sup				Lim inf	Lim sup				Lim inf	Lim sup
<b>Nivel educativo padre</b>																				
Ninguno-primaria-secundaria	75	69,9	0,62	0,27	2,19	60,9	75	*0,00	1,19	3,12	61,2	74,2	*0,01	1,11	2,97	66,9	26,1	0,15	0,87	2,24
Técnico tecnológico Universitario	25	30,1				39,1	25				38,8	25,8				33,1	26,1			
<b>Forma de desplazamiento</b>																				
Activo	55	60,2	0,64	0,32	1,07	59,1	60,3	0,84	0,60	1,50	57,3	61	0,51	0,53	1,37	63,5	55,9	0,15	0,88	2,12
Pasivo	45	39,8				40,9	39,7				42,7	39				36,5	44,1			

OR: Odds Ratio; Lim inf: Límite Inferior; Lim sup: Límite Superior; \* p valor <0,05: Asociación estadísticamente significativa; estadístico de prueba utilizado: Chi cuadrado

Fuente: autores

aceite (emPCR), que contiene la mezcla de reacción de amplificación para aumentar clonalmente cada fragmento (25).

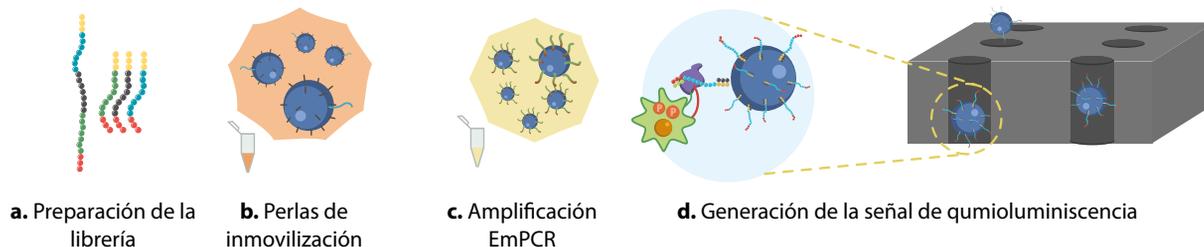
Una vez realizada la amplificación, se depositan las perlas en una placa de secuenciación que contiene aproximadamente 1,6 millones de pozos, de los cuales aproximadamente 450.000 presentan secuencias detectables, y 200.000, lecturas exitosas (27). La solución se diluye de tal manera que cada pozo de 75 pl contenga como máximo una perla para poder comenzar el proceso de secuenciación. Luego, la placa se coloca en una cámara de flujo a través del cual se difunden perpendicularmente los reactivos (25). Cuando el ddNTP correcto se incorpora, se detecta la señal de quimioluminiscencia generada por la liberación de una molécula de pirofosfato (21,22). En la reacción, la adenosina trifosfato, en presencia de APS (3'-fosfoadenosina-5'-fosfosulfato), convierte el pirofosfato liberado en adenosina trifosfato, que reacciona con la

enzima luciferasa. Así, la luciferina se transforma en oxilciferina y emite una señal de luz. La señal es capturada mediante una cámara de fibra óptica (25) y la intensidad de la luz será proporcional a la cantidad de pirofosfato escindido (27) (figuras 1a y 1b).

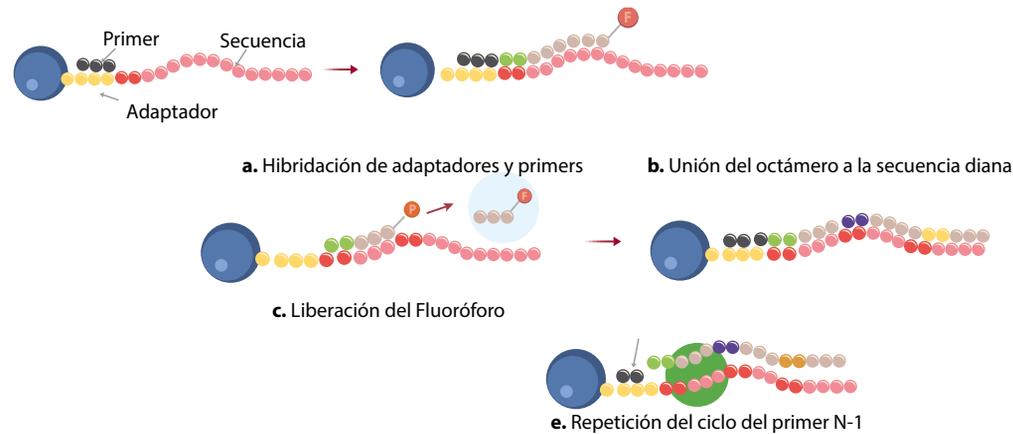
### Figura 1a. Secuenciación Roche 454

a) Preparación de la librería. Fragmentación del ADN para generar secuencias de longitud de 400-700 pb e incorporar adaptadores a los fragmentos obtenidos. b) Inmovilización-hibridación de adaptadores de los fragmentos de la librería con las secuencias complementarias de las perlas de andamiaje. c) Amplificación clonal de los fragmentos en las perlas suspendidas en una emulsión de agua en aceite (emPCR). d) Generación de la señal de quimioluminiscencia. La liberación de una molécula de pirofosfato desencadena la conversión de la luciferina a oxilciferina, generando una emisión de luz.

Figura 1. Métodos de secuenciación de segunda generación



**Figura 1.** Métodos de secuenciación de segunda generación



**Figura 1b. Secuenciación por ligación (SOLiD).** a) Hibridación de adaptadores y *primers*. Acoplamiento de los adaptadores a la perla de andamiaje e hibridación del cebador a la zona complementaria del adaptador. b) Unión del octámero a la secuencia diana. Hibridación de los octámeros obtenidos en la librería a las secuencias diana complementarias. c) Liberación del fluoróforo. Detección de la señal de fluorescencia determinada por las dos primeras bases del octámero acoplado. d) Proceso de elongación de la cadena blanco. e) Repetición del ciclo N-1. Continuación de la secuenciación a partir de cebadores de longitud N-1, N-2, etc.

La tecnología 454 de Roche fue capaz de secuenciar más de 25 millones de bases, con una precisión promedio del 96% para librerías genómicas en un

tiempo de corrida de 4 h (25). Fue usada comúnmente para secuenciación de genomas y para muestras de metagenomas (28). Se lograron ensamblajes *de novo*, es decir, sin genomas de referencia con coberturas del 96% y precisión del 99,06%. Adicionalmente, dentro de las ventajas de la pirosecuenciación no se requerían geles, marcadores fluorescentes o didesoxinucleótidos, lo que generaba una mayor rapidez, comparada con el método de Sanger. No obstante, presentó problemas con la determinación precisa de homopolímeros, debido a la dificultad para detectar la intensidad de luminiscencia conforme aumentaba la longitud de la secuencia. Por otra parte, su alta tasa de error (1%) y su elevado costo de secuenciación por millón de bases (~8,57 dólares) hicieron que la llegada de lecturas de 200-300 bp de MiSeq a un menor costo, junto con las lecturas de >8000 bp de PacBio, esencialmente, acabaran

con esta tecnología. Al final, la mayoría de los errores de esta plataforma fueron las inserciones o eliminaciones, que son mucho más difíciles de manejar, en comparación con los errores de sustitución que se presentan en la secuenciación de Illumina (24,27).

### ***Secuenciación por ligación y detección de nucleótidos (SOLiD)***

La plataforma SOLiD realiza secuenciación masiva de fragmentos en paralelo (29). Fue adquirida por Applied Biosystems Instrument, en 2006 (30), y posteriormente fue comercializada por Life Technologies Corporation (31).

Para la preparación de la librería se procede de manera similar a la descrita en la plataforma Roche 454 (véase figura 1a). El ADN es cortado en fragmentos de 80-130 pb (32) y los fragmentos obtenidos son adaptados en las posiciones 5' y 3' con secuencias complementarias para ser hibridadas a perlas paramagnéticas. Aproximadamente, 120M de perlas son utilizadas y colocadas en una emulsión para la amplificación de las secuencias, y al finalizar la secuenciación, el 60% de las perlas iniciales generarán secuencias de buena calidad (33,34).

La plataforma maneja una librería de sondas de 8 nucleótidos marcadas con fluoróforos. En total, se utilizan 4 fluoróforos diferentes para generar

la señal de fluorescencia que representará la combinación de dos nucleótidos. Dadas las 16 posibles combinaciones, se utiliza una técnica denominada espacio-color para la codificación de la señal y la determinación de la posición cada nucleótido (35). Un iniciador se acopla a la platinilla blanco para comenzar la secuenciación, el octámero se une y en cada paso de replicación la sonda se libera produciendo fluorescencia. Así, se deja el extremo 5' libre para continuar con la elongación. Este procedimiento se repite durante 10 ciclos continuos. Después de culminada la primera ronda, el producto obtenido de la extensión es degradado, para dar paso al mismo proceso con una sonda de longitud N-1. Ello permite conocer las posiciones faltantes de la anterior amplificación, y de esta manera el proceso se repite con sondas de longitud N-2, N-3 y así sucesivamente hasta determinar la totalidad de la cadena (35) (figura 1b). Se ha empleado en el análisis de expresión, splicings (corte y empalme de ARN), análisis de polimorfismos y de genomas completos (33). Genera 80-360 Gbp de datos por corrida, con longitudes de las lecturas de 75-85 bp (29). La precisión puede alcanzar un 99,99%; sin embargo, los tiempos de corrida son muy largos (~8 días). Desafortunadamente, emplea más de una semana para el procesamiento y genera lecturas de longitudes muy cortas  $\sim 2 \times 60$  pb, lo que la hace inadecuada para muchas aplicaciones (36).

## Secuenciación por Illumina

La plataforma Illumina, que actualmente ocupa una gran parte del mercado de la secuenciación de próxima generación y que implica la fragmentación aleatoria de ADN y la ligación de fragmentos de ADN mediante adaptadores presentes en una lámina, se describió entre 2006 y 2008 (37,38). Esta metodología se basa en la secuenciación por síntesis de la cadena complementaria, amplificación en puente y detección a partir de fluorescencia de nucleótidos (39,40).

Su baja tasa de error  $<1\%$ , comparada con la secuenciación de Pacific Biosciences (12%-15%) y Oxford Nanopore Technology (5%-20%), junto con una cobertura de secuencia mayor ("secuenciación masiva en paralelo"), es una gran ventaja de esta tecnología (19). Durante cada ciclo se adiciona una mezcla de cuatro nucleótidos (adenina, guanina, timina y citocina), individualmente marcados con un fluoróforo específico para ese tipo

de base y bloqueados en el 3' deoxinucleótido (dNTP). Posteriormente, cada base insertada se determina mediante microscopía de fluorescencia de reflexión interna total, que utiliza de dos a cuatro canales láser (41).

Los fragmentos de ADN están ligados a adaptadores que, después de la desnaturalización, se inmovilizan en una lámina que contiene adaptadores complementarios tanto directos como reversos. Posteriormente, cada fragmento monocatenario crea una estructura de puente que, gracias a una mezcla de reactivos de PCR compuesta por cebadores, cuatro nucleótidos terminadores reversibles cada uno marcado con un fluoróforo y el ADN polimerasa, inicia la amplificación por PCR y genera 100-200 millones de grupos clonales de amplificación (figura 2). Los recientes ajustes de química avanzada han permitido lecturas finales emparejadas de hasta  $2 \times 300$  pb de cada fragmento de ADN (42).

Figura 2. Secuenciación por síntesis - Illumina



**Figura 2. Secuenciación por síntesis (Illumina).**

a) Adición de adaptadores e identificadores a la plantilla de ADN, preparación de la librería. b) Hibridización de adaptadores: se adicionan las plantillas de ADN a la celda de flujo y ello permite hibridar el adaptador a los oligos complementarios en la celda. c) Amplificación en puente: una polimerasa crea la secuencia complementaria del fragmento hibridado, la hebra se dobla y el segundo adaptador se une a otro oligo complementario para amplificar en puente. d) Amplificación clonal: generación de grupos de amplificación o amplificación clonal.

La plataforma Illumina tiene una alta tasa de precisión (>99,5%), dada la química de terminador reversible que permite que cada ciclo se complete minimizando la incorporación incorrecta de cada base, lo que mejora la precisión, la cobertura de la secuencia consenso y el llamado de polimorfismos que afectan un solo nucleótido, para obtener datos precisos a muy alto rendimiento y bajo costo (~0,03-0,15). Esto hace que este tipo de secuenciación sea menos susceptible a los errores de homopolímeros observados en plataformas como Roche 454 o Ion Torrent (39).

Sin embargo, algunas de las desventajas de la secuenciación Illumina son los ensamblajes altamente fragmentados, sobre todo los ensamblajes *de novo*, debido a la dificultad de resolver

secuencias repetitivas (43), la disminución de la señal y una tasa de error bruto del 1%-1,5% (42) (véase tabla 1).

**Secuenciación por síntesis (Ion Torrent)**

La plataforma Ion Torrent es el primer método de secuenciación por síntesis con detección electroquímica. Traduce la información del cambio de pH generado durante la síntesis en información digital, en lugar de utilizar sistemas ópticos para detectar la señal de secuenciación (44,45). La tecnología fue lanzada al mercado en 2010 por Life Technologies, los inventores de Roche 454 (46), y a partir de ahí ha sido muy utilizada, debido al rápido procesamiento de las lecturas (47).

La secuenciación es dirigida hacia las celdas de un microchip que contiene sensores electrónicos que permiten detectar iones de hidronio para generar la señal de secuenciamiento (31). Existen dos secuenciadores que utilizan la tecnología Ion Torrent: el chip PGM (Ion Personal Genome Machine), que utiliza un promedio de 11,1 M sensores, y el Ion Protón, que contiene más de 165 M de sensores (48).

La preparación de la librería es similar a la de Roche 454 (véase figura 1). El ADN es fragmentado hasta obtener secuencias de 200-400 pb y se realiza el ligamiento de los adaptadores, lo que permite la unión a las perlas, las cuales se colocan

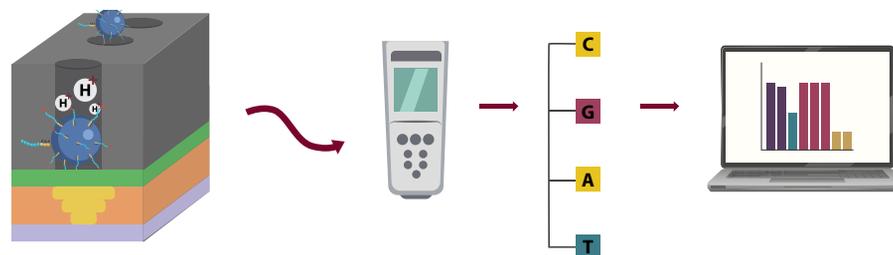
posteriormente en una emulsión para la reacción de amplificación del fragmento (31,49). En esta plataforma no se incorporan fluoróforos para visualizar la secuenciación; el ensamblaje es con nucleótidos sin modificar, utilizando como señal el cambio en el pH de la solución, que es detectado electrónicamente por un transistor, dada la liberación de un ion hidronio. La detección se genera con el clivaje del pirofosfato y la liberación de un átomo de hidrógeno en la posición 3' tras la incorporación de un nucleótido (50). El sistema normaliza las señales y reconoce el voltaje correspondiente a cada nucleótido (31,49) (figura 3).

**Figura 3. Secuenciación Ion Torrent.** a) Amplificación en chip semiconductor: la amplificación de los fragmentos ligados a las perlas de andamiaje libera iones de hidronio. b) Lectura del potencial de iones de hidrógeno: el chip semiconductor contiene sensores electrónicos que permiten detectar los iones hidronios liberados, debido a la incorporación de un nucleótido. c) Procesamiento y análisis datos: el sistema normaliza los

datos y detecta una señal específica para cada uno de los nucleótidos.

El sistema Ion Protón es una tecnología que genera secuencias de alta calidad, utiliza tiempos cortos de corrida, es fácilmente accesible y económico (50). Se ha utilizado en estudios metagenómicos, caracterización microbiana y estudios de cáncer (51). El dispositivo Ion Protón presenta una buena tasa de rendimiento, ya que genera aproximadamente 10 G de datos y 80 M de lecturas, con un promedio de longitud de lectura de 200 bp en 4 h (31,52). Por otro lado, el PGM produce de 70 Mbp a 2 Gbp (0,5-0,7 M lecturas de 200 a 400 bp en un tiempo de 2-5 h), dependiendo del chip (31,53). Se observa mejor calidad y costos en la plataforma Ion Protón que la PGM; sin embargo, en los datos de secuenciación resultantes se observan inserciones asociadas a repeticiones de homopolímeros de un solo nucleótido (49,54) y se han obtenido tasas de error de inserción y eliminación de secuencias del 1,5% (53).

**Figura 3.** Secuenciación Ion Torrent.



a. Amplificación en chip semi-conductor    b. Lectura de potencial de iones Hidrógeno    c. Procesamiento y análisis datos

## Secuenciación de tercera generación

En 2004 comenzó el desarrollo de las tecnologías de secuenciación de tercera generación, que generan secuencias de más de 1000 pb. Estas utilizan como plantilla una molécula sencilla de ADN que es secuenciada en tiempo real (SMRT, por sus siglas en inglés), sin requerir pasos de limpieza entre las lecturas (55,56). Así, las tecnologías de secuenciación de tercera generación evitan una preparación extensiva de la muestra y contribuyen a disminuir el costo y a incrementar la velocidad de secuenciación (57).

### ***Tecnología de Molécula Única en Tiempo Real (SMRT®)***

La primera secuenciación en tiempo real fue desarrollada por Jonas Korlach y Stephen Turner, en la Universidad de Cornell, en 2004, con el objetivo de crear un microscopio con mayor eficacia que cualquiera de las tecnologías de secuenciación disponibles hasta el momento. La secuenciación SMRT, de Pacific Bio Sciences, inmoviliza directamente la polimerasa, induciendo un aumento de la velocidad de procesamiento de la enzima (58). Esta tecnología se caracteriza por la detección de la señal mediante fluorescencia, el procesamiento de fragmentos largos, los ciclos cortos de lectura y los bajos costos, comparada con otras plataformas (59).

El chip de secuenciación SMRT contiene 150.000 micropocillos ZMW (Zero-Mode-Waverguide) que permiten la observación directa de la elongación de la cadena de ADN (60). Los ZMW están fabricados en una película de metal de dióxido de silicio, tienen un diámetro de 100 nm y un volumen de detección de fluorescencia de 20 zeptolitros ( $10^{-21}$  L) (59).

En cada pozo se lleva a cabo la secuenciación por síntesis de una sola molécula (59), no obstante las fallas en el proceso de anclaje de la polimerasa sobre la base del pozo o del anclaje de la molécula de ADN a la enzima, que generan que solo 35.000-70.000 de los pozos utilizados arrojen lecturas (60).

En cada ZMW se difunde una hebra de ADN, denominada en inglés *SMRTbell*. La plantilla es una hebra de doble cadena circular previamente flanqueada con adaptadores monocatenarios en forma de horquilla. La polimerasa inmovilizada en la base del pozo reconoce el sitio de unión a la horquilla y permite comenzar la replicación (61).

Dada la característica circular del SMRTbell, se pueden realizar múltiples pases de replicación de una molécula, teniendo en cuenta el rendimiento y la vida útil de la polimerasa (60,62). El tiempo de residencia de la polimerización en el sitio activo ZMW depende de la velocidad de catálisis (63). Los desoxinucleótidos trifosfato (dNTP) son modificados con la

inserción de distintos fluoróforos, de tal manera que no se interfiera o se detenga el proceso de replicación. La señal de fluorescencia se enlaza al fosfato terminal de los dNTP, mientras la ADN polimerasa induce el clivaje del enlace  $\alpha$ - $\beta$ -fosforilo para liberar el grupo pirofosfato unido al fluoróforo (64). La asignación de la señal o base llamada se obtiene por la fluorescencia detectada en el volumen delimitado por el ZMW, mediante un sistema óptico confocal que identifica moléculas individuales del fluoróforo (62,65,66) (figura 4).

**Figura 4. Secuenciación SMRT (PacBio).** a) Chip de secuenciación SMRT: contiene 150.000 micropocillos ZMW. Cada micropozo permite la reacción de secuenciación de una sola cadena. b) Anclaje de SMRTbell a la enzima polimerasa: hibridación de la hebra de ADN de doble cadena circular flanqueada con adaptador en forma de horquilla a la polimerasa inmovilizada en el pozo.

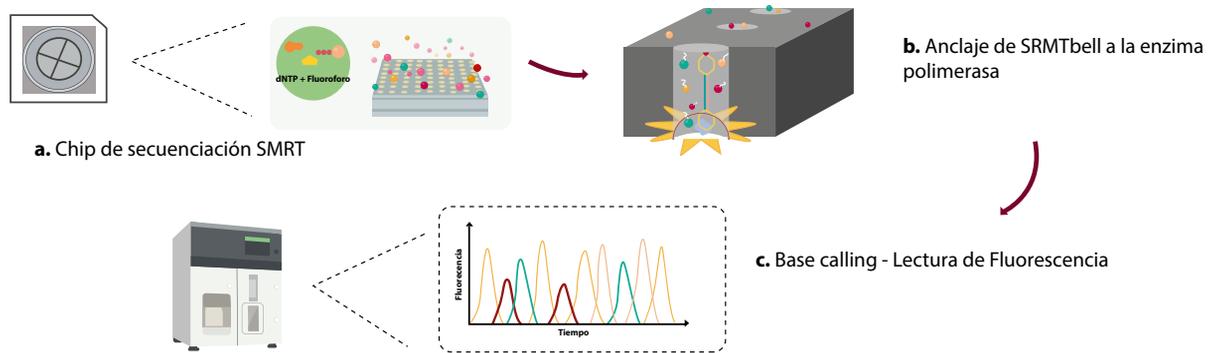
Una vez anclada, da inicio a la amplificación y se genera la señal de fluorescencia con la inserción de nucleótidos modificados con fluoróforos. c) *Base calling*: lectura de la señal obtenida mediante un sistema de detección confocal.

La secuenciación PacBio proporciona lecturas largas. La última tecnología implementada PacBio Sequel maneja lecturas promedio mayores a 10 kb con longitudes superiores a 60 kb y el rendimiento por celda es de 0,5-1 mil millones de bases; sin embargo, esta tecnología maneja una alta tasa de error (13%), comparada con tecnologías de secuenciación de segunda generación (véase tabla 1) (67).

### **Secuenciación Oxford Nanopore Technology**

El primer boceto de secuenciación por nanoporos (25 de junio de 1989), desarrollado por David Deamer, George Church y Daniel Branton,

**Figura 4.** Secuenciación SMRT (PacBio)



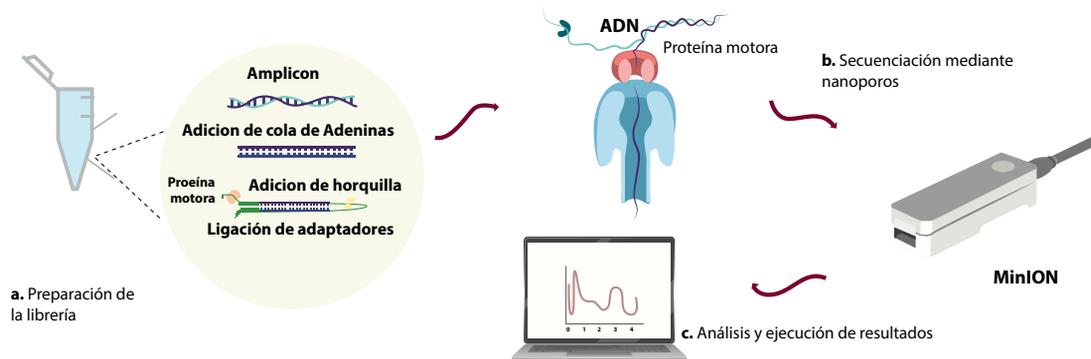
muestra el dibujo de 3 bases de una cadena de ADN que atraviesan un nanoporo bajo la aplicación de voltaje, donde cada base afecta la corriente iónica y permite determinar las secuencias de las bases (68). La secuenciación de nanoporos se distingue de otros enfoques, ya que en este una hebra larga de ADN pasa a través de un poro de proteína estabilizado en una membrana de polímero, sobre la cual se coloca un voltaje que detecta cambios de corriente iónica a medida que un nucleótido atraviesa el poro, sin necesidad de síntesis activa de ADN (69).

Similar a la secuenciación de segunda generación, se requiere la preparación de una biblioteca que consta de la reparación de ADN o producto de PCR para formar extremos romos, la adición de colas poli-A en los extremos 3' y la adición de adaptadores. Generalmente, las bibliotecas tienen dos adaptadores: el líder, en forma de Y, y otro en forma de horquilla, o adaptador HP. Las

proteínas motoras son las encargadas de guiar el ADN a través del nanoporo, donde la hebra monocatenaria inicia su paso por el adaptador Y en el extremo 5', seguida de la hebra complementaria, dirigida por la proteína motora del adaptador de horquilla.

Se aplica un voltaje a través de la membrana y, posteriormente, a medida que pasa el ADN, por el poro se genera un cambio de flujo de corriente iónica, que permite identificar cada base en tiempo real. Si solo es leída la información de una hebra de ADN, se denomina *lectura 1 dimensional*, lo que ofrece un mayor rendimiento; sin embargo, si se leen las dos hebras, se denomina *lectura 2 dimensional*, que representa una mayor calidad de secuencia (43) (figura 5). La celda del secuenciador MinION contiene 2048 nanoporos que se pueden controlar en grupos de 512 nanoporos mediante un circuito, con lo cual se secuencian 512 moléculas de ADN simultáneamente (70).

**Figura 5.** Secuenciación Oxford Nanopore-MinION



**Figura 5. Secuenciación Oxford Nanopore-MinION.** a) Preparación de la librería: adición de cola de adeninas a la hebra molde de ADN junto con los adaptadores de horquilla. b) Secuenciación mediante nanoporos: el ADN pasa a través del nanoporo y genera un cambio de flujo de corriente iónica. c) Análisis y ejecución de resultados.

El secuenciador MinION de Oxford Nanopore tiene un costo aproximado de 1000 dólares, incluido el dispositivo y los reactivos iniciales, un valor relativamente bajo con respecto a otros instrumentos de secuenciación masiva en paralelo. Por otra parte, su peso oscila entre 90 y 100 gr y mide aproximadamente  $3,3 \times 10,5 \times 2,3$  cm; esto lo hace ser un dispositivo portátil, veloz y asequible, que se puede conectar a cualquier computador (69). Tales características lo hacen atractivo para una gran comunidad científica (70,71).

Una de las principales desventajas de los secuenciadores de lectura larga es que el secuenciador de nanoporos tiene una tasa de error de entre el 5% y el 38%, que varía según la preparación de la biblioteca y el tipo de molécula (69,72). Sin embargo, Oxford Nanopore ha estado trabajando en mejorar la química de secuenciación y el *software* de llamado de base, que con ayuda de las nuevas herramientas bioinformáticas mejora la identidad de lectura secuenciada (73).

## Aplicaciones de las tecnologías de secuenciación

Un mayor conocimiento de la biología humana y del desarrollo de enfermedades mediante la caracterización de marcadores biológicos, incluidos genes o proteínas, o el estudio de sus modificaciones, ya sea por factores internos o externos, ha ayudado a la comprensión de patologías y su adherencia o resistencia a diferentes tratamientos (74,75).

Las tecnologías de secuenciación han traído innumerables avances biomédicos. En cuanto a la genómica, estas tecnologías incursionaron en el cribado neonatal, diagnóstico de enfermedades hereditarias, análisis de genes para la identificación de polimorfismos de nucleótido único, inserciones y deleciones y variaciones en el número de copias, discapacidades intelectuales y enfermedades crónicas (76-78). Así mismo, han aumentado la sensibilidad y eficacia de pruebas genéticas para enfermedades mendelianas, enfermedades raras, detección de mutaciones y defectos en el genoma, asociados con el desarrollo de cáncer, la principal causa de morbilidad y mortalidad en todo el mundo (79,80).

Estas herramientas, además, ayudan a los médicos a tomar mejores decisiones en cuanto al diagnóstico y tratamiento de distintas enfermedades. Es el caso de la invención en tratamientos

con anticuerpos monoclonales, como trastuzumab y pertuzumab, dirigidos contra el receptor del factor de crecimiento epidérmico humano 2, un receptor de tirosina quinasa transmembrana de 185 kDa codificado por el gen ERBB2 (cromosoma 17q12) en pacientes con cáncer de mama y cáncer gástrico (81).

Otros estudios realizados con herramientas de secuenciación de ADN han logrado avances médicos en el cáncer gástrico, donde se encontraron mutaciones somáticas involucradas en la aparición y progresión de estos tumores (80,82). Es el caso de mutaciones en el gen CDH1, responsables del desarrollo de cáncer gástrico difuso hereditario temprano, al igual que mutaciones SMAD4 o BMPR1A en la poliposis gastrointestinal y PTEN en el síndrome de Cowden (83-85). Otras aplicaciones en cáncer resaltan los avances generados con tecnologías de secuenciación, como en el diagnóstico y tratamiento de cáncer de pulmón (86), cáncer colorrectal (87,88), cáncer prostático avanzado (89), neoplasias de células B como mieloma múltiple (90), entre otras patologías.

En cuanto al transcriptoma que abarca el estudio del ARN, a través de la funcionalidad reguladora de genes (72,91), las tecnologías de secuenciación de ADN han tenido un gran auge en los últimos años (92). Avances en técnicas de secuenciación como PacBio SMS y Oxford Nanopore permiten estudiar completamente secuencias de

ADN complementario de longitud completa sin la necesidad de fragmentar el ADN, como ocurre en la secuenciación por Illumina, porque se elimina la necesidad de hacer ensamblaje, lo que identifica transcripciones completas (93).

Las aplicaciones de transcriptómica incluyen el estudio genético del carcinoma hepatocelular (94), la evaluación de la plasticidad tumoral e interacciones estromales en la metástasis cerebral y la confirmación de la regulación al alza de genes que miden la metástasis cerebral como SERPINI1 y L1CAM, que informan sobre características transcriptómicas de células tumorales (95), análisis de la expresión cuantitativa de genes y variantes genéticas relacionados con la diabetes tipo 2, glucemia alterada en ayunas (96,97), asociación del índice de masa corporal con los perfiles transcriptoma de adipocitos viscerales humanos en el cáncer colorrectal (98), transcriptómica del virus del papiloma humano como biomarcador de citología cervical de alto grado (99) o análisis de transcriptoma de verrugas inducidas por el virus del papiloma humano (100). Otros estudios han permitido la identificación de biomarcadores en etapa temprana y descubrir objetivos terapéuticos para el desarrollo de nuevas terapias.

Por último, es importante destacar que el uso de estas tecnologías para la secuenciación del genoma completo constituye una herramienta poderosa para comprender la dinámica de trans-

misión de los brotes. Es el caso de los estudios realizados para controlar la transmisión del SARS-CoV-2, los cuales permiten comprender de forma más precisa los patrones de transmisión y apoyan la toma de decisiones en el ámbito de salud pública (101).

Fundamentados en lo anterior, Oude Munnink et al. et al. (102) describieron la incursión del SARS-CoV-2 en los Países Bajos, por medio de un enfoque de secuenciación basado en amplicones. Ello reveló que el antepasado más reciente de los virus que circulaban en los Países Bajos se remonta a finales de enero y principios de febrero; además, se reportaron casos relacionados con viajes y un número creciente de casos locales que proporcionaron un apoyo adicional a la toma de decisiones y restricciones (101). Por otra parte, en un estudio más reciente que utilizó la secuenciación del genoma completo de los brotes en 16 granjas de visones, reportó que el 68% de los residentes de la granja evaluados, empleados o contactos, tenían evidencia de infección por SARS-CoV-2, lo que demostró la transmisión del SARS-CoV-2 de animal a humano dentro de las granjas de visones (103).

En Colombia, López-Álvarez et al. (104) reportaron la secuencia del genoma de un aislado del coronavirus 2 del SARS-CoV-2, obtenido de frotis nasofaríngeo de un paciente con síntomas de la enfermedad que estaba infectado en Cali

(Colombia), usando Oxford Nanopore-MinION. El análisis demostró 7 mutaciones, en comparación con el genoma de referencia y la agrupación del aislamiento Cali-01 y Bogota-78390 en el clado G; mientras que el aislado Antioquia-79256 permaneció en el clado S.

Estos y otros estudios demuestran la importancia de combinar el análisis epidemiológico con la secuenciación profunda del genoma viral mediante herramientas de secuenciación de segunda o tercera generación, para entender la propagación del SARS-CoV-2 y obtener conocimientos fundamentales sobre la dinámica mutacional y las propiedades de transmisión. Ello permitirá contribuir al esfuerzo internacional en curso para rastrear y comprender la pandemia del SARS-CoV-2 (105).

## DISCUSIÓN

Los métodos de secuenciación han revolucionado nuestra comprensión de la biología humana en los últimos 40 años. A medida que avanza la tecnología, estos serán más rápidos y precisos, con tasas de error mucho más bajas y con una reducción en la formación de artefactos, a un costo mucho menor (19). Las ventajas de la secuenciación de siguiente generación, como la construcción *in vitro* de la librería de secuenciación, amplificación clonal de moléculas de ADN, asignación de códigos de barras a cada muestra (también llamado proceso de multiplexación), estabilidad,

rendimiento y gran velocidad permitirá a los investigadores estudiar sistemas biológicos de una manera mucho más precisa (42).

En la actualidad, la aparición de técnicas que pasan moléculas de ADN a través de poros muy pequeños de membranas biológicas mediante la ayuda de proteínas motoras o mediante la implementación de tecnologías de sensor de estado sólido que utiliza sustratos de metal con poros nanométricos ofrecerán nuevas oportunidades de secuenciación de ADN en tiempo real, de forma portátil y rápida que ofrezcan diagnósticos oportunos para la prevención y tratamiento de enfermedades que aquejan a la población (19).

En el ámbito biomédico, estas tecnologías de secuenciación fortalecen los enfoques de la medicina moderna, en cuanto a la predicción y prevención de enfermedades. Las complejas preguntas de tipo genómico en las nuevas investigaciones exigen una mayor profundidad de información en la secuenciación rápida de genomas completos, secuenciación con alta profundidad de regiones objetivo del genoma, análisis de factores epigenéticos (como metilaciones de ADN en todo el genoma), estudios de la diversidad microbiana o en el medio ambiente, estudios de enfermedades genéticas raras no diagnosticadas, entre otros, que las tecnologías tradicionales no logran abarcar (11).

Como se mencionaba, estas tecnologías traen consigo avances y resultados prometedores; sin embargo, es importante resaltar que generan volúmenes de datos muy altos (aproximadamente Roche 454: 15 GB; Illumina: 1 TB, y SOLiD: 15 TB), lo que les plantea a los investigadores grandes desafíos, en cuanto al uso y procesamiento de datos, pues se requiere un sistema de almacenamiento y procesamiento mayor. Los análisis basados en puntajes Phred para rechazar lecturas de baja calidad, la alineación y el ensamblaje de lecturas cortas son sustancialmente más difíciles que para los datos obtenidos por Sanger, por lo cual una gran cantidad de *softwares* se han puesto a disposición de la comunidad científica, a pesar de que sus características varían con la aplicación. Actualmente, se están implementando diversas herramientas para análisis de datos de secuenciación de segunda y tercera generación que les saquen el máximo provecho a la cantidad de datos obtenidos, lo cual deriva en grandes retos para el campo de la bioinformática (106-108).

Así mismo, el futuro de la secuenciación genómica parece enfrentar desafíos adicionales, entre los que se incluyen la capacidad de integrar un conjunto grande de técnicas en una sola plataforma, que revolucione de una manera positiva los grandes conjuntos de datos de la biología moderna, masivamente paralelos y de alto rendimiento, a fin de ofrecerles una oportunidad a diversos equipos interdisciplinarios, como médicos,

científicos, bioinformáticos, matemáticos, biólogos, estadísticos, entre otros, para formular y generar enfoques en el tratamiento de diversas patologías, entre ellas el cáncer, con una mayor apropiación en la medicina clínica (109).

Las investigaciones en las cuales se generan grandes cantidades de datos genotípicos y fenotípicos que se pueden vincular con los participantes individuales de un estudio plantean una serie de cuestiones éticas (110) en tres áreas principales: privacidad, consentimiento informado y devolución de resultados (111).

Las instituciones o investigadores que utilizan tecnologías de secuenciación deberán asegurarse de que existan medidas de seguridad y almacenamiento adecuadas para la gran cantidad de datos; deberán estar atentos a la necesidad de obtener un consentimiento informado, basándose principalmente en la comunicación de beneficios y riesgos al participante; deberán estar atentos a proteger a los participantes del estudio de posibles usos indebidos o repercusiones perjudiciales de sus datos genómicos, y deberán tener en cuenta las preferencias del individuo con respecto a la recepción de resultados. Todo esto sugiere que aún existe un largo camino por recorrer, en cuanto a los aspectos éticos del uso de secuenciación genómica y manipulación de gran cantidad de datos por parte de los científicos o personal de la salud (112).

## CONCLUSIÓN

Las tecnologías de siguiente generación han traído consigo un gran avance en el estudio de los sistemas biológicos; sin embargo, grandes desafíos serán el diseño de algoritmos para el análisis adecuado de los datos, la incorporación de costos más asequibles y el logro de mejores tiempos de corrida, con una menor tasa de error.

La disminución de los costos, la mejora de las plataformas y de los softwares para el análisis de los datos de secuenciación de próxima generación, junto con la accesibilidad para su uso, no solo llevará al desarrollo de herramientas analíticas importantes para el procesamiento de datos obtenidos, sino también al abordaje de una amplia aplicabilidad, especialmente para datos clínicos y de diagnóstico de enfermedades en el campo biomédico, porque ayudarán a los científicos y clínicos a resolver desafíos biológicos con la apertura de conocimientos para el planteamiento de nuevas terapias.

Finalmente, con el uso de tecnologías de secuenciación también se podrán rastrear los orígenes y comprender la evolución de los agentes infecciosos, investigar la propagación y las cadenas de transmisión de los brotes, facilitar el desarrollo de pruebas de diagnóstico molecular (eficaces y rápidas) y contribuir a la búsqueda de tratamientos y vacunas oportunas.

## LIMITACIONES DEL TRABAJO

Algunos de los artículos seleccionados no presentaban acceso abierto para su lectura, por lo cual fueron de difícil consulta. La gran cantidad de datos e información relevante asociada al tema de investigación conlleva extender la presentación de esta revisión.

## AGRADECIMIENTOS

Los autores queremos agradecer a la Universidad de Boyacá, por su apoyo en esta trayectoria, y a la Dra. Marina Muñoz, PhD, por sus valiosos aportes en la construcción de este artículo.

## FINANCIACIÓN

Los autores declaran financiación de la Universidad de Boyacá de los medios necesarios para llevar a cabo este artículo de revisión.

## CONFLICTO DE INTERESES

Los autores declaran que no existen conflictos de intereses.

## REFERENCIAS

1. Mattick JS, Dziadek MA, Terrill BN, Kaplan W, Spigelman AD, Bowling FG, et al. The impact of genomics on the future of medicine and

health. *Med J Aust.* 2014;201(1):17-20. <https://doi.org/10.5694/mja13.10920>

2. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Thiele J, editores. WHO classification of tumours of haematopoietic and lymphoid tissues. 4.<sup>a</sup> ed. Ginebra: IARC Publications; 2017. p. 885-7. <https://doi.org/10.1002/9781118853771.ch51>

3. Greenberg PL, Stone RM, Al-Kali A, Barta SK, Bejar R, Bennett JM, et al. Myelodysplastic syndromes. *J Natl Compr Cancer Netw.* 2017;15(1):60-87. <https://doi.org/10.6004/jnccn.2017.0007>

4. Inaba H, Shinozawa K, Amano K, Fukutake K. The genetic analysis of hemophilia a: the application of next-generation sequencing in the analysis of causative variants deep inside the intron of the F8 gene. *Blood.* 2016;128(22):1393. <https://doi.org/10.1182/blood.V128.22.1393.1393>

5. Goodeve AC, Pavlova A, Oldenburg J. Genomics of bleeding disorders. *Haemophilia.* 2014;20:50-3. <https://doi.org/10.1111/hae.12424>

6. Collins FS, McKusick VA. Implications of the human genome project for medical science. *JAMA.* 2001;285:540-4. <https://doi.org/10.1001>

- jama.285.5.540
7. Pereira R, Oliveira J, Sousa M. Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. *J Clin Med*. 2020;9(1):132-50. <https://doi.org/10.3390/jcm9010132>
  8. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. <https://doi.org/10.1038/35057062>
  9. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001;291:13-51. <https://doi.org/10.1126/science.1058040>
  10. Wetterstrand KA. DNA sequencing costs: data [internet]. National Human Genome Research Institute. Disponible en: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>
  11. Laurentino S, Heckmann L, Di Persio S, Li X, zu Hörste GM, Wistuba J, et al. High-resolution analysis of germ cells from men with sex chromosomal aneuploidies reveals normal transcriptome but impaired imprinting. *Clin Epigenetics*. 2019;213(11):127. <https://doi.org/10.1186/s13148-019-0720-3>
  12. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977 Dec;74(12):54-67. <https://doi.org/10.1073/pnas.74.12.5463>
  13. Lander ES. Initial impact of the sequencing of the human genome. *Nature*. 2011;470:187-97. <https://doi.org/10.1038/nature09792>
  14. Tripp S, Grueber M. Economic impact of the human genome project [internet]. Battelle Memorial Institute, Technology Partnership Practice; 2011 may. Disponible en: <https://www.battelle.org/docs/default-source/misc/battelle-2011-misc-economic-impact-human-genome-project.pdf?sfvrsn=6>
  15. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Erratum: Corrigendum: genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2006;441:120-45. <https://doi.org/10.1038/nature04726>
  16. Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics*. 2016;107:1-8. <https://doi.org/10.1016/j.ygeno.2015.11.003>
  17. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and

- future. *Nature*. 2017;550:345-53. <https://doi.org/10.1038/nature24286>
18. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A*. 1977 Feb;74(2):560-4. <https://doi.org/10.1073/pnas.74.2.560>
  19. Slatko BE, Gardner AF, Ausubel FM. Overview of next-generation sequencing technologies. *Current Protocols in Molec. Biol*. 2018;122:59. <https://doi.org/10.1002/cpmb.59>
  20. França LT, Carrilho E, Kist TB. A review of DNA sequencing techniques. *Q Rev Biophys*. 2002;35:169-200. <https://doi.org/10.1017/S0033583502003797>
  21. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. *Genomics*. 2008;92:55-64. <http://doi.org/10.1016/j.ygeno.2008.07.001>
  22. Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*. 1996;242:84-9. <https://doi.org/10.1006/abio.1996.0432>
  23. Nyrén P, Lundin A. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Anal Biochem*. 1985;151:504-9. [https://doi.org/10.1016/0003-2697\(85\)90211-8](https://doi.org/10.1016/0003-2697(85)90211-8)
  24. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem*. 2009;55:641-58. <https://doi.org/10.1373/clinchem.2008.112789>
  25. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437:376-80. <https://doi.org/10.1038/nature03959>
  26. Nyrén P, Pettersson B, Uhlén M. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal Biochem*. 1993;208:171-5. <https://doi.org/10.1006/abio.1993.1024>
  27. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol*. 2007;8:143. <https://doi.org/10.1186/gb-2007-8-7-r143>
  28. Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, et al. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A*. 2006;103:15-20. <https://doi.org/10.1073/pnas.0605127103>

29. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* 2008;18(7):51-63. <https://doi.org/10.1101/gr.076463.108>
30. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol.* 2008;26(10):1135-45. <https://doi.org/10.1038/nbt1486>
31. Kulski JK. Next-generation sequencing—an overview of the history, tools, and “Omic” applications. En: Kulski J, editor. *Next generation sequencing: advances, applications and challenges.* IntechOpen; 2016. p. 3-60. <https://doi.org/10.5772/61964>
32. Shirota M, Kinoshita K. Discrepancies between human DNA, mRNA and protein reference sequences and their relation to single nucleotide variants in the human population. *Database.* 2016;baw 124. <https://doi.org/10.1093/database/baw124>
33. Cloonan N, Forrest A, Kolle G, Gardiner B, Faulkner GJ, Brown MK, et al. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods.* 2008;5:6-9. <https://doi.org/10.1038/nmeth.1223>
34. Li H, Homer N. A survey of sequence alignment algorithms for next-generation sequencing. *Brief Bioinform.* 2010;11:473-83. <https://doi.org/10.1093/bib/bbq015>
35. Garrido-Cárdenas JA, García-Maroto F, Álvarez-Bermejo JA, Manzano-Agugliaro F. DNA sequencing sensors: an overview. *Sensors.* 2017;17:588-590. <https://doi.org/10.3390/s17030588>
36. Barba M, Czosnek H, Hadidi A. Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses.* 2014;6:106-36. <https://doi.org/10.3390/v6010106>
37. Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res.* 2006;34(3):6-10. <https://doi.org/10.1093/nar/gnj023>
38. Turcatti G, Romieu A, Fedurco M, Tairi A-P. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Res.* 2008;36:25-34. <https://doi.org/10.1093/nar/gkn021>

39. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456:53-9. <https://doi.org/10.1038/nature07517>
40. Adessi C, Matton G, Ayala G, Turcatti G, Mermod J-J, Mayer P, et al. Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res*. 2000;28:12-8. <https://doi.org/10.1093/nar/28.20.e87>
41. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nature Rev Gen*. 2016;17:333-40. <https://doi.org/10.1038/nrg.2016.49>
42. Cao Y, Fanning S, Proos S, Jordan K, Srikumar S. A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Front Microbiol*. 2017;8:18-29. <https://doi.org/10.3389/fmicb.2017.01829>
43. Lu H, Giordano F, Ning Z. Oxford Nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics*. 2016;14(5):265-79. <https://doi.org/10.1016/j.gpb.2016.05.004>
44. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475:348-52. <https://doi.org/10.1038/nature10242>
45. Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW. Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Comput Biol*. 2013;9(4):e1003031. <https://doi.org/10.1371/journal.pcbi.1003031>
46. Rusk N. Torrents of sequence. *Nat Methods*. 2010;8:44-58. <https://doi.org/10.1038/nmeth.f.330>
47. Magierowski S. Ion torrent in a little detail [internet]. 2014 julio 6 [citado 2020 abr 8]. Disponible en: <http://www.cse.yorku.ca/~magiero/Pubs/iontorrent.pdf>
48. Kim S, Jung H, Han SH, Lee S, Kwon J, Kim MG, et al. Comparison of two high-throughput semiconductor chip sequencing platforms in noninvasive prenatal testing for Down syndrome in early pregnancy. *BMC Med Gen*. 2016;9:22. <https://doi.org/10.1186/s12920-016-0182-9>
49. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison

- of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 2012;13:341-9. <https://doi.org/10.1186/1471-2164-13-341>
50. Merriman B, D Team IT, Rothberg JM. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*. 2012;33(23):3397-417. <https://doi.org/10.1002/elps.201200424>
51. Hsiao Y-P, Lu C-T, Chang-Chien J, Chao W-R, Yang J-J. Advances and applications of Ion Torrent personal genome machine in cutaneous squamous cell carcinoma reveal novel gene mutations. *Materials*. 2016;9:464-78. <https://doi.org/10.3390/ma9060464>
52. Chen S, Li S, Xie W, Li X, Zhang C, Jiang H, et al. Performance comparison between rapid sequencing platforms for ultra-low coverage sequencing strategy. *PLoS One*. 2014;9:3-12. <https://doi.org/10.1371/journal.pone.0092192>
53. Wang Y, Wen Z, Shen J, Cheng W, Li J, Qin X, et al. Comparison of the performance of Ion Torrent chips in noninvasive prenatal trisomy detection. *J Hum Genet*. 2014;59:393-406. <https://doi.org/10.1038/jhg.2014.40>
54. Marine RL, Magaña LC, Castro CJ, Zhao K, Montmayeur AM, Schmidt A, et al. Comparison of Illumina MiSeq and the Ion Torrent PGM and S5 platforms for whole-genome sequencing of picornaviruses and caliciviruses. *J Virol Methods*. 2020;280:113865. <https://doi.org/10.1016/j.jviromet.2020.113865>
55. McCarthy A. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. *Chem Biol*. 2010;17:675-86. <https://doi.org/10.1016/j.chembiol.2010.07.004>
56. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, et al. Single-molecule DNA sequencing of a viral genome. *Science*. 2008;320(5872):106-9. <https://doi.org/10.1126/science.1150427>
57. Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Hum Mol Genet*. 2010;19:227-40. <https://doi.org/10.1093/hmg/ddq416>
58. Korlach J, Bjornson KP, Chaudhuri BP, Cicero RL, Flusberg BA, Gray JJ, et al. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol*. 2010;472:431-55. [https://doi.org/10.1016/S0076-6879\(10\)72001-2](https://doi.org/10.1016/S0076-6879(10)72001-2)

59. Pacific Bioesciences Technology Backgrounder. Pacific biosciences develops transformative DNA sequencing technology [internet]. 2008 nov 24 [citado 2020 may 5]. Disponible en: [https://www.ndsu.edu/pubweb/~mcclean/plsc411/Pacific%20Biosciences-technology\\_backgrounder.pdf](https://www.ndsu.edu/pubweb/~mcclean/plsc411/Pacific%20Biosciences-technology_backgrounder.pdf)
60. Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics*. 2015;13(5):278-89. <https://doi.org/10.1016/j.gpb.2015.08.002>
61. Travers KJ, Chin C-S, Rank DR, Eid JS, Turner SW. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Res*. 2010;38:159-65. <https://doi.org/10.1093/nar/gkq543>
62. Boldogkői Z, Moldován N, Balázs Z, Snyder M, Tombácz D. Long-read sequencing—a powerful tool in viral transcriptome research. *Trends Microbiol*. 2019;27:578-92. <https://doi.org/10.1016/j.tim.2019.01.010>
63. Metzker ML. Sequencing technologies—the next generation. *Nat Rev Genet*. 2010;11:31-46. <https://doi.org/10.1038/nrg2626>
64. Korlach J, Bibillo A, Wegener J, Peluso P, Pham TT, Park I, et al. Long, processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides. *Nucleosides Nucleotides Nucleic Acids*. 2008;27(9):1072-82. <https://doi.org/10.1080/15257770802260741>
65. Foquet M, Samiee KT, Kong X, Chauduri BP, Lundquist PM, Turner SW, et al. Improved fabrication of zero-mode waveguides for single-molecule detection. *J Appl Phys*. 2008;103:34-8. <https://doi.org/10.1063/1.2831366>
66. Lundquist PM, Zhong CF, Zhao P, Tomaney AB, Peluso PS, Dixon J, et al. Parallel confocal detection of single molecules in real time. *Opt Lett*. 2008;33(9):1026-8. <https://doi.org/10.1364/OL.33.001026>
67. Weirather JL, de Cesare M, Wang Y, Piazza P, Sebastiano V, Wang X-J, et al. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. 2017;6:100. <https://doi.org/10.12688/f1000research.10571.2>
68. Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing. *Nat Biot*. 2016;34:518. <https://doi.org/10.1038/nbt.3423>
69. Kono N, Arakawa K. Nanopore sequencing: review of potential applications in functional genomics. *Dev Growth Differ*. 2019;61(5):316-26. <https://doi.org/10.1111/dgd.12608>

70. Quick J, Ashton P, Calus S, Chatt C, Gossain S, Hawker J, et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. *Genom Biol.* 2015;16:114. <https://doi.org/10.1186/s13059-015-0677-2>
71. Judge K, Harris SR, Reuter S, Parkhill J, Peacock SJ. Early insights into the potential of the Oxford Nanopore MinION for the detection of antimicrobial resistance genes. *J Antimicrob Chemother.* 2015;70(10):2775-8. <https://doi.org/10.1093/jac/dkv206>
72. Raphael BJ, Hruban RH, Aguirre AJ, Moffitt RA, Yeh JJ, Stewart C, et al. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell.* 2017;32:185-203. <https://doi.org/10.1016/j.ccell.2017.07.007>
73. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genom Biol.* 2016;17:239. <https://doi.org/10.1186/s13059-016-1103-0>
74. Gibney E, Nolan C. Epigenetics and gene expression. *Heredity.* 2010;105:4-13. <https://doi.org/10.1038/hdy.2010.54>
75. Bohacek J, Mansuy IM. Epigenetic inheritance of disease and disease risk. *Neuropsychopharmacology.* 2013;38:220-36. <https://doi.org/10.1038/npp.2012.110>
76. Haraksingh RR, Snyder MP. Impacts of variation in the human genome on gene regulation. *J Mol Biol.* 2013;425(21):3970-7. <https://doi.org/10.1016/j.jmb.2013.07.015>
77. Mefford HC, Batshaw ML, Hoffman EP. Genomics, intellectual disability, and autism. *N Engl J Med.* 2012;366(8):733-43. <https://doi.org/10.1056/NEJMra1114194>
78. Botkin JR, Belmont JW, Berg JS, Berkman BE, Bombard Y, Holm IA, et al. Points to consider: ethical, legal, and psychosocial implications of genetic testing in children and adolescents. *Am J Hum Genet.* 2015 Jul 2;97(1):6-21. <https://doi.org/10.1016/j.ajhg.2015.05.022>
79. Pociot F, Akolkar B, Concannon P, Erlich HA, Julier C, Morahan G, et al. Genetics of type 1 diabetes: what's next? *Diabetes.* 2010;59:1561-71. <https://doi.org/10.2337/db10-0076>
80. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, et al. A brief history of human disease genetics. *Nature.* 2020;577:179-89. <https://doi.org/10.1038/s41586-019-1879-7>

81. Beigh MM. Next-generation sequencing: the translational medicine approach from “bench to bedside to population”. *Medicin*. 2016;3:14. <https://doi.org/10.3390/medicines3020014>
82. Ross DS, Zehir A, Cheng DT, Benayed R, Nafa K, Hechtman JF, et al. Next-generation assessment of human epidermal growth factor receptor 2 (ERBB2) amplification status: clinical validation in the context of a hybrid capture-based, comprehensive solid tumor genomic profiling assay. *J Mol Diagn*. 2017;19(2):244-54. <https://doi.org/10.1016/j.jmoldx.2016.09.010>
83. Pan X, Ji X, Zhang R, Zhou Z, Zhong Y, Peng W, et al. Landscape of somatic mutations in gastric cancer assessed using next-generation sequencing analysis. *Oncol Lett*. 2018;16:4863-70. <https://doi.org/10.3892/ol.2018.9314>
84. Colvin H, Yamamoto K, Wada N, Mori M. Hereditary gastric cancer syndromes. *Surg Oncol Clin*. 2015;24(4):765-77. <https://doi.org/10.1016/j.soc.2015.06.002>
85. Luo W, Fedda F, Lynch P, Tan D. CDH1 gene and hereditary diffuse gastric cancer syndrome: Molecular and histological alterations and implications for diagnosis and treatment. *Front Pharmacol*. 2018;9:1421. <https://doi.org/10.3389/fphar.2018.01421>
86. Nemtsova MV, Kalinkin AI, Kuznetsova EB, Bure IV, Alekseeva EA, Bykov II, et al. Clinical relevance of somatic mutations in main driver genes detected in gastric cancer patients by next-generation DNA sequencing. *Sci Rep*. 2020;10:1-11. <https://doi.org/10.1038/s41598-020-57544-3>
87. Zhang Y, Shen W-X, Zhou L-N, Tang M, Tan Y, Feng C-X, et al. The value of next-generation sequencing for treatment in non-small cell lung cancer patients: the observational, real-world evidence in China. *Bio Med Res. Int*. 2020;20:20. <https://doi.org/10.1155/2020/9387167>
88. Zhou W, Xu J, Zhao Y, Sun Y. SAG/RBX2 is a novel substrate of NEDD4-1 E3 ubiquitin ligase and mediates NEDD4-1 induced chemosensitization. *Oncotarget*. 2014;5:67-76. <https://doi.org/10.18632/oncotarget.2246>
89. Wang Y, Liu H, Hou Y, Zhou X, Liang L, Zhang Z, et al. Performance validation of an amplicon-based targeted next-generation sequencing assay and mutation profiling of 648 Chinese colorectal cancer patients. *Virchows Archiv*. 2018;472:959-68. <https://doi.org/10.1007/s00428-018-2359-4>
90. Suh J, Jeong CW, Choi S, Ku JH, Kim HH, Kim KS, et al. Targeted next-generation sequencing for locally advanced prostate cancer in the Korean

- population. *Investig Clin Urol.* 2020;61:127-35. <https://doi.org/10.4111/icu.2020.61.2.127>
91. Bolli N, Genuardi E, Ziccheddu B, Martello M, Oliva S, Terragna C. Next-generation sequencing for clinical management of multiple myeloma: ready for prime time? *Front Oncol.* 2020;10. <https://doi.org/10.3389/fonc.2020.00189>
92. Abeshouse A, Ahn J, Akbani R, Ally A, Amin S, Andry CD, et al. The molecular taxonomy of primary prostate cancer. *Cell.* 2015;163:1011-25. <https://doi.org/10.1016/j.cell.2015.10.025>
93. Lightbody G, Haberland V, Browne F, Taggart L, Zheng H, Parkes E, et al. Review of applications of high-throughput sequencing in personalized medicine: barriers and facilitators of future progress in research and clinical application. *Brief Bioinf.* 2019;20:1795-811. <https://doi.org/10.1093/bib/bby051>
94. Wang B, Olson A, Kumar V, Ware D. Reviving the transcriptome studies: an insight into the emergence of single-molecule transcriptome sequencing. *Front Genet.* 2019;10:384. <https://doi.org/10.3389/fgene.2019.00384>
95. Jin Y, Lee WY, Toh ST, Tennakoon C, Toh HC, Chow PK-H, et al. Comprehensive analysis of transcriptome profiles in hepatocellular carcinoma. *J Transl Med.* 2019;17:1-16. <https://doi.org/10.1186/s12967-019-2025-x>
96. Wingrove E, Liu ZZ, Patel KD, Arnal-Estapé A, Cai WL, Melnick M-A, et al. Transcriptomic hallmarks of tumor plasticity and stromal interactions in brain metastasis. *Cell Rep.* 2019;27:1277-92. <https://doi.org/10.1016/j.celrep.2019.03.085>
97. Jenkinson CP, Göring HH, Arya R, Blangero J, Duggirala R, DeFronzo RA. Transcriptomics in type 2 diabetes: bridging the gap between genotype and phenotype. *Genom Data.* 2016;8:25-36. <https://doi.org/10.1016/j.gdata.2015.12.001>
98. Cui Y, Chen W, Chi J, Wang L. Comparison of transcriptome between type 2 diabetes mellitus and impaired fasting glucose. *Med Sci Mon Int Med J Exp Clin Res.* 2016;22:4699. <https://doi.org/10.12659/msm.896772>
99. Del Cornò M, Baldassarre A, Calura E, Conti L, Martini P, Romualdi C, et al. Transcriptome profiles of human visceral adipocytes in obesity and colorectal cancer unravel the effects of body mass index and polyunsaturated fatty acids on genes and biological processes related to tumorigenesis. *Front Immunol.* 2019;10:265. <https://doi.org/10.3389/fimmu.2019.00265>

100. Pérot P, Biton A, Marchetta J, Pourcelot A-G, Nazac A, Marret H, et al. Broad-range papillomavirus transcriptome as a biomarker of papillomavirus-associated cervical high-grade cytology. *J Mol Diagn.* 2019;21:768-81. <https://doi.org/10.1016/j.jmoldx.2019.04.010>
101. Al-Eitan LN, Tarkhan AH, Alghamdi MA, Al-Qarqaz FA, Al-Kofahi HS. Transcriptome analysis of HPV-induced warts and healthy skin in humans. *BMC Med genomics.* 2020;13:1-10. <https://doi.org/10.1186/s12920-020-0700-7>
102. Munnink BBO, Nieuwenhuijse DF, Stein M, O'Toole A, Haverkate M, Mollers M, et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat Med.* 2020;26-11:1802. <https://doi.org/10.1038/s41591-020-0997-y>
103. Munnink BBO, Sikkema RS, Nieuwenhuijse DF, Molenaar RJ, Munger E, Molenkamp R, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science.* 2020;371(6525):172-7. <https://doi.org/10.1126/science.abe5901>
104. Munnink BBO, Sikkema RS, Nieuwenhuijse DF, Molenaar RJ, Munger E, Molenkamp R, et al. Jumping back and forth: anthropozoonotic and zoonotic transmission of SARS-CoV-2 on mink farms. *bioRxiv.* 2020. <https://doi.org/10.1101/2020.09.01.277152>
105. López-Álvarez D, Parra B, Cuéllar WJ. Genome sequence of SARS-CoV-2 isolate Cali-01, from Colombia, obtained using Oxford Nanopore MinION sequencing. *Microbiol Resour Announc.* 2020;9:26. <https://doi.org/10.1128/MRA.00573-20>
106. Popa A, Genger J-W, Nicholson MD, Penz T, Schmid D, Aberle SW, et al. Genomic epidemiology of superspreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. *Sci Transl Med.* 2020;12:573. <https://doi.org/10.1126/scitranslmed.abe2555>
107. Derocles SA, Bohan DA, Dumbrell AJ, Kitson JJ, Massol F, Pauvert C, et al. Biomonitoring for the 21st century: integrating next-generation sequencing into ecological network analysis. *Adv Ecol Res.* 2018;58:1-62. <https://doi.org/10.1016/bs.aecr.2017.12.001>
108. Kchouk M, Elloumi M. An error correction and de novo assembly approach for nanopore reads using short reads. *Curr Bioinform.* 2018;13:241-52. <https://doi.org/10.2174/1574893612666170530073736>

109. Castiblanco J. A primer on current and common sequencing technologies. En: Autoimmunity: from bench to bedside [internet]. Bogotá: Editorial Universidad del Rosario; 2013 [citado 2020 jul 22]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK459463/>
110. García-García G, Baux D, Faugère V, Moclyn M, Koenig M, Claustres M, et al. Assessment of the latest NGS enrichment capture methods in clinical context. *Sci Rep.* 2016;6:1-8. <https://doi.org/10.1038/srep20948>
111. Foster MW, Sharp RR. Ethical issues in medical-sequencing research: implications of genotype–phenotype studies for individuals and populations. *Hum Mol Genet.* 2006;15:45-9. <https://doi.org/10.1093/hmg/ddl049>
112. Martinez-Martin N, Magnus D. Privacy and ethical challenges in next-generation sequencing. *Exp Rev Precis Med Drug Develop.* 2019;4:95-10. <https://doi.org/10.1080/23808993.2019.1599685>



Esta obra está bajo una licencia de Creative Commons  
Reconocimiento-NoComercial 4.0 Internacional