

Artículo recibido 4 de julio de 2025.

Artículo aceptado 22 de septiembre de 2025.

Artículo publicado 5 de octubre de 2025.

Evolución y aplicaciones de las diferentes tecnologías de secuenciación de ácidos nucleicos

Marichich, Paula B.^{1,2}; Chalup, Laura M.^{1,2}; Seijo, Guillermo J.^{2,3}; Samoluk, Sebastián S.^{2,3}

¹Universidad Nacional del Chaco Austral (UNCAus) Comandante Fernández N° 755, CP 3700, Presidencia Roque Sáenz Peña, Chaco.

paulamarichich@gmail.com, laurachalup@gmail.com

²Instituto de Botánica del Nordeste (UNNE-CONICET) Sargento Juan Bautista Cabral 2131, W3402BKG, Corrientes;

³Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA-UNNE) Av. Libertad 5470, CP 3400, Corrientes

jgseijo@yahoo.com, ssamoluk@exa.unne.edu.ar

ORCID Paula B, Marichich 0009-0005-4645-8994

ORCID Laura M, Chalup 0000-0001-5262-5051

ORCID Guillermo J, Seijo 0000-0001-8699-2044

ORCID Sebastián S, Samoluk 0000-0002-0132-3210

Resumen

El modelo de la estructura de ADN propuesto por Watson y Crick en 1953 marcó un punto de inflexión en el desarrollo de la biología molecular y las ciencias genómicas. La secuenciación de ácidos nucleicos se consolidó como una herramienta esencial para explorar la estructura y función de los genomas. La publicación del genoma humano, basado principalmente en la tecnología de secuenciación de Sanger, constituyó un estímulo para la formación de consorcios internacionales destinados al desarrollo de diferentes proyectos genómicos. Sin embargo, los altos costos, el bajo rendimiento y los

extensos tiempos necesarios por la tecnología de Sanger la volvieron ineficiente para satisfacer dichas demandas. En las últimas dos décadas, múltiples desarrollos tecnológicos permitieron la aparición de diversas plataformas de secuenciación de alto rendimiento, capaces de generar lecturas nucleotídicas de millones de fragmentos de ADN en forma paralela, con alta precisión y menor costo. El análisis de los datos generados, gracias al desarrollo conjunto de una amplia variedad de herramientas bioinformáticas especializadas, ha generado importantes conocimientos genómicos aplicables a salud, agricultura, biotecnología, microbiología y sustentabilidad ambiental. En este artículo se presenta una revisión general sobre los principios, evolución y aplicaciones de estas tecnologías en genómica estructural y funcional.

Palabras clave: Plataformas de secuenciación, ácidos nucleicos, genómica estructural y funcional

Abstract

Evolution and applications of different nucleic acid sequencing technologies

The DNA structure model proposed by Watson and Crick in 1953 marked a turning point in the development of molecular biology and genomic sciences. Nucleic acid sequencing became established as a key tool for exploring genome structure and function. The publication of the human genome, primarily based on Sanger sequencing technology, served as a catalyst for the formation of international consortia aimed at advancing various genomic projects. However, the high costs, low throughput, and long processing times associated with Sanger technology rendered it inefficient for meeting these growing demands. Over the past two decades, numerous technological advancements have led to the emergence of high-throughput sequencing platforms, capable of producing nucleotide reads from millions of DNA fragments in parallel, with high accuracy and reduced cost. The analysis of the resulting data —enabled by the parallel development of a broad array of specialized bioinformatic tools— has generated significant genomic insights with applications in health, agriculture, biotechnology, microbiology, and environmental sustainability. This article presents a general overview of the principles, evolution, and applications of these technologies in structural and

functional

genomics.

Keywords: Sequencing platforms, nucleic acids, structural and functional genomic

INTRODUCCIÓN

Desde mediados del siglo XX, el estudio de los ácidos nucleicos —ADN y ARN— ha transformado profundamente nuestra comprensión de los mecanismos hereditarios, la evolución de las especies y los procesos celulares fundamentales. Estos avances han sido posibles gracias, en gran parte, al desarrollo progresivo de tecnologías capaces de “leer” la secuencia de nucleótidos que componen estos ácidos nucleicos. La secuenciación, entendida como la determinación del orden exacto de las bases nitrogenadas en una molécula de ADN o ARN, se ha convertido en una herramienta central en las ciencias biológicas, con aplicaciones que abarcan desde la medicina personalizada hasta la agricultura, la biotecnología, la ecología, y la investigación forense.

Los primeros métodos prácticos para secuenciar ADN fueron desarrollados en la década de 1970. En 1977, Maxam y Gilbert propusieron una técnica basada en la modificación química específica de las bases nitrogenadas, seguida de la escisión controlada de la molécula (Maxam y Gilbert, 1977). Ese mismo año, Frederick Sanger introdujo una alternativa revolucionaria: el método de terminación de cadena, que utiliza dideoxinucleótidos (ddNTPs) para interrumpir la síntesis del ADN en puntos específicos (Sanger et al., 1977). Este enfoque resultó ser más eficiente y más fácil de automatizar, lo que le permitió convertirse en la técnica estándar durante más de tres décadas. Gracias a esta metodología, se lograron importantes logros científicos, como la secuenciación del primer genoma completo —el del bacteriófago phiX174, de poco más de 5.300 nucleótidos— y, décadas más tarde, la culminación del Proyecto Genoma Humano (Waterston et al., 2001; Venter et al., 2001). Sin embargo, las limitaciones inherentes al método de Sanger —como su baja escalabilidad, la necesidad de clonar fragmentos y el costo elevado— motivaron el desarrollo de tecnologías más eficientes. A partir de la década de 2000, emergieron las denominadas tecnologías de secuenciación de nueva generación (NGS, por sus siglas en inglés), entre las que destacan plataformas como Illumina/Solexa y SOLiD. Estas técnicas introdujeron el

concepto de secuenciación masiva en paralelo, permitiendo leer millones de fragmentos de ADN simultáneamente y reduciendo drásticamente tanto los tiempos de procesamiento como los costos asociados (Mardis, 2008). Como consecuencia, se aceleró la expansión de nuevas áreas de investigación como la genómica estructural, con aplicaciones tanto en organismos modelo como en organismos no modelo (John et al., 2021).

En la última década, las llamadas tecnologías de tercera generación han llevado la secuenciación un paso más allá. Plataformas como PacBio SMRT (Single Molecule Real Time) y Oxford Nanopore Technologies permiten leer moléculas individuales de ADN sin necesidad de amplificación previa, generando lecturas ultra largas y abriendo nuevas posibilidades en el ensamblado de genomas complejos, la detección de variantes estructurales, regiones repetitivas complejas, isoformas de ARN y modificaciones epigenéticas (Rhoads y Au, 2015; Jain et al., 2016). Además, estas tecnologías han posibilitado aplicaciones emergentes como la secuenciación directa de ARN y el análisis portátil en tiempo real, lo cual es especialmente útil en entornos clínicos, en investigación de campo o en situaciones de respuesta rápida ante brotes infecciosos (Hilt y Ferrieri, 2022).

A la par del avance tecnológico, se ha producido una acelerada disminución en los costos de secuenciación. Entre 2008 y 2015, el costo de secuenciar un genoma humano completo descendió de aproximadamente 340.000 a apenas 4.200 dólares, una caída mucho más rápida que la prevista por la Ley de Moore, que históricamente se usaba como modelo de referencia para la evolución del poder computacional (Figura 1) (Muir et al., 2016). Esta tendencia ha democratizado el acceso a los estudios genómicos, permitiendo que cada vez más laboratorios y proyectos puedan incorporar análisis de secuenciación a gran escala.

Hoy en día, las plataformas de secuenciación disponibles varían considerablemente en cuanto a sus fundamentos técnicos, mecanismos de lectura, precisión, longitud de lectura, costos, requerimientos instrumentales y campos de aplicación (Metzker, 2010; Zhang et al., 2011; Heather y Chain, 2016; Hu et al., 2021). Comprender estos aspectos resulta esencial para que cualquier investigador, estudiante o profesional del área pueda seleccionar de manera crítica la tecnología más adecuada

según sus objetivos. En esta revisión se propone un abordaje integrado que recorre la evolución de las diferentes tecnologías de secuenciación —desde los métodos pioneros hasta las plataformas actualmente utilizadas—, destacando sus fundamentos, ventajas y limitaciones, así como sus aplicaciones actuales y perspectivas futuras.

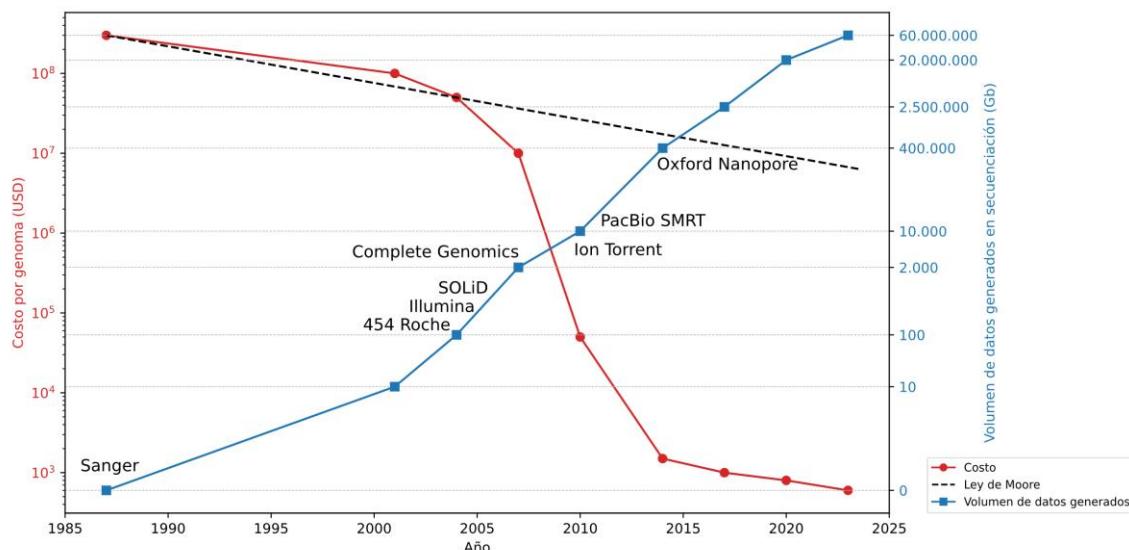


Figura 1: Patrones históricos en la reducción de costos por genoma y el aumento del volumen de datos generados por plataformas de secuenciación. Los datos que se muestran en la figura se obtuvieron del Instituto Nacional de Investigación del Genoma Humano.

LA SECUENCIACIÓN SANGER: EL PUNTO DE PARTIDA

La técnica de Sanger, también conocida como “método de terminación por dideoxinucleótidos”, fue desarrollada por Frederick Sanger en 1977, lo que le valió su segundo Premio Nobel de Química en 1980. Este método de secuenciación se basa en la incorporación competitiva de desoxirribonucleótidos (dNTPs) y dideoxinucleótidos (ddNTPs) durante la síntesis de una cadena complementaria de ADN catalizada por una ADN polimerasa. La clave del método reside en que los ddNTPs, al carecer de un grupo hidroxilo en la posición 3', impiden la elongación de la cadena, generando fragmentos de diferentes longitudes que terminan en un nucleótido específico. Inicialmente, el método de Sanger empleaba ddNTPs marcados radiactivamente, y la secuencia se deducía manualmente interpretando los patrones de fragmentos generados en cuatro

reacciones separadas y corridas en geles de acrilamida. En la actualidad, la versión automatizada se realiza en un solo tubo con ddNTPs marcados con fluoróforos diferentes, y posterior electroforesis capilar, permitiendo una lectura más rápida, precisa y digitalizada de la secuencia del fragmento de ADN (Zhou et al., 2010). Aunque esta técnica ofrece alta precisión (99.99%) y fue crucial para el desarrollo del Proyecto Genoma Humano, su escalabilidad es limitada, con una capacidad máxima de ~1 Mb por corrida y altos costos por base leída (Drmanac et al., 2010; Satam et al., 2023).

LA REVOLUCIÓN DE LA SEGUNDA GENERACIÓN: SECUENCIACIÓN MASIVA EN PARALELO

Las limitaciones de la secuenciación Sanger dieron lugar, a partir de mediados de los años 2000, a una nueva familia de tecnologías denominadas de “nueva generación” o NGS (“Next-Generation Sequencing”), cuyo rasgo distintivo es la capacidad de generar millones de lecturas simultáneas en un solo ensayo (Metzker, 2010; McCombie et al., 2019; Satam et al., 2023). Todas ellas se caracterizan por la ligación de adaptadores de fragmentos de ADN, una amplificación clonal de los fragmentos, y la secuenciación de los mismos mediante un mecanismo de síntesis dependiente de una enzima ADN polimerasa. Los adaptadores no solo permiten la unión del ADN a las superficies de secuenciación, sino que también incluyen secuencias índices (barcodes), que posibilitan mezclar múltiples muestras en una misma corrida y luego distinguirlas con precisión. Esto mejora notablemente el rendimiento y reduce costos (Gansauge y Meyer, 2013; Hu et al., 2021).

La primera tecnología de secuenciación de nueva generación (NGS) fue el sistema 454 GS20, lanzado en 2005 por la empresa 454 *Life Sciences*, fundada por Jonathan Rothberg y posteriormente adquirida por Roche en 2007. Esta plataforma se basa en la técnica de pirosecuenciación acoplada a PCR en emulsión, lo que hace posible detectar en tiempo real la incorporación de nucleótidos mediante la liberación de pirofosfato. Este subproducto genera una señal lumínica cuya intensidad es proporcional al número de bases incorporadas, posibilitando la inferencia directa de la secuencia. El ADN fragmentado se fija a micropelículas funcionalizadas con cebadores específicos y es amplificado dentro de microgotas de emulsión mediante PCR. Cada microesfera, que

contiene múltiples copias idénticas del fragmento original, se deposita en un pocillo individual de una placa PicoTiter. Esta matriz facilita la detección paralela de las señales lumínicas emitidas durante la síntesis del ADN. El sistema 454 representó un hito frente al método de Sanger, ya que permitió reducir los costos y acortar los tiempos de secuenciación necesarios para alcanzar una cobertura genómica adecuada (Zhao & Grant, 2011). En 2008, se lanza una versión mejorada, el GS FLX Titanium, que ofrecía ahora una longitud de lectura de 700 pb frente a los 150 pb del sistema original 454. Finalmente, en 2010, Rothberg lanzó Ion Torrent, que reemplazó la detección óptica por un chip que medía cambios de pH durante la síntesis de ADN. Este sistema eliminó la necesidad de cámaras y láseres, utilizando semiconductores que detectan variaciones de pH como consecuencia de la incorporación de nucleótidos. Aunque más compacto y económico, mantenía limitaciones similares de precisión, particularmente en regiones homopoliméricas (Satam et al., 2023).

En 2006, la empresa Illumina adquirió la empresa británica Solexa, que había desarrollado una innovadora tecnología de secuenciación por síntesis (SBS). Esta tecnología se basa en la amplificación de fragmentos de ADN mediante PCR en puente (*bridge amplification*) sobre una celda de flujo, donde se lleva a cabo la secuenciación mediante la incorporación de nucleótidos marcados con fluoróforos. Cada incorporación genera una señal fluorescente específica que es capturada ópticamente mediante escáneres de alta resolución, permitiendo la identificación base por base a lo largo de millones de clústeres en paralelo (Metzker, 2010; Loman et al., 2012; Hu et al., 2021). Los primeros equipos de secuenciación de Illumina, como *Genome Analyzer* (GA), producían lecturas muy cortas (~35 pb), aunque tenían la ventaja de permitir la secuenciación de ambos extremos de los fragmentos de ADN (*paired-end*), lo que mejoró significativamente el ensamblado y la detección de variantes. Actualmente, las plataformas más modernas como NovaSeq o NextSeq pueden generar más de 600 Gb por corrida, con lecturas entre 75 y 300 pb, dependiendo del equipo y el kit utilizado. Illumina comercializa múltiples plataformas adaptadas a distintos volúmenes y propósitos, como MiniSeq, MiSeq, NextSeq, e iSeq para laboratorios de menor escala y

NovaSeq para proyectos masivos¹. Gracias a su alta precisión (<1% de error), versatilidad y bajo costo por base (~0,07 USD por millón), la tecnología de Illumina se ha convertido en el estándar dominante en secuenciación masiva (Kulski, 2016).

Applied Biosystems en 2007, comercializó la tecnología SOLiD, una estrategia basada en la ligación de oligonucleótidos y detección mediante terminadores reversibles. En este sistema, el ADN se fragmenta, se ligan los adaptadores específicos y se une a microesferas, donde se amplifica por PCR en emulsión. Las perlas se disponen sobre una lámina de vidrio, para permitir la secuenciación por ligación con sondas marcadas con fluoróforos. Cada ciclo de ligación es seguido de un reseteo del molde, permitiendo que cada posición sea leída desde distintos puntos de partida, lo cual genera redundancia y mejora la precisión del llamado de bases. Esta tecnología fue pionera en ofrecer lecturas cortas con alta precisión. Las primeras versiones producían 35 pb y 3 Gb por ensayo (Moti et al., 2022). Su mayor fortaleza radica en la elevada precisión lograda gracias a la doble lectura de cada base, aunque presenta desventajas como la limitación en la longitud de lectura y los extensos tiempos de corrida. Aunque superada por Illumina en rendimiento y longitud de lectura, SOLiD mantenía costos competitivos y era útil para ciertos tipos de análisis donde la precisión era prioritaria (Heather & Chain, 2016; Satam et al., 2023).

Otra tecnología destacada basada en la ligación fue la de *Complete Genomics*, conocida como nanoesferas de ADN. En este enfoque, la amplificación clonal del ADN no se realiza mediante PCR convencional. En cambio, se utiliza amplificación por círculo rodante (RCA), generando largas moléculas de ADN con múltiples repeticiones de la secuencia molde. Estas cadenas se pliegan en estructuras compactas conocidas como nanoesferas, que se fijan a una superficie sólida para su secuenciación. La secuenciación se llevaba a cabo mediante la técnica cPAL (*combinatorial Probe-Anchor Ligation*), que utiliza sondas fluorescentes de 10 nucleótidos, con una base específica en una posición determinada, que se hibridaban y ligaban a oligonucleótidos complementarios a los adaptadores. La señal fluorescente registrada indicaba la base correspondiente, y tras cada lectura se iniciaba una nueva ronda en otra posición. En 2015, la empresa fue adquirida por BGI (*Beijing Genomics Institute*) quien integró esta

¹ <https://www.illumina.com/systems/sequencing-platforms.html>

tecnología en su sistema *Revolocity*, diseñado para la secuenciación de genomas y exomas con alta cobertura (~50x). Más recientemente, se han desarrollado plataformas derivadas como DNBSEQ y CoolMPS, que compiten con Illumina en términos de rendimiento y costos por base (Zhao & Grant, 2011; Heather & Chain, 2016; Kulski, 2016).

A pesar de sus ventajas en rendimiento y costo por base, las tecnologías de segunda generación presentan limitaciones importantes: las lecturas cortas dificultan la reconstrucción de genomas altamente repetitivos o fragmentados, y la necesidad de PCR puede introducir sesgos, especialmente en regiones ricas en AT o GC (Goodwin et al., 2016). Estas limitaciones abrieron el camino hacia la aparición de tecnologías de tercera generación, enfocadas en la lectura directa de moléculas individuales de ADN.

SECUENCIACIÓN DE TERCERA GENERACIÓN: TECNOLOGÍAS DE LECTURA LARGA EN TIEMPO REAL

La necesidad de obtener genomas cada vez más completos o de alta resolución, así como de resolver variantes estructurales complejas, condujo con el tiempo al desarrollo de métodos de secuenciación cada vez más sofisticados. Esto marcó un nuevo hito en la evolución de las tecnologías genómicas, con la aparición de métodos que permiten lecturas largas, secuenciación directa y mayor resolución estructural. Estas nuevas tecnologías superan limitaciones de la segunda generación, como la baja longitud de lectura, los errores en regiones repetitivas y la dependencia de amplificación previa. Estas plataformas constituyen lo que hoy se conoce como tecnologías de “tercera generación” o “*long-read sequencing*” (Hu et al., 2021). Las dos más relevantes, tanto por su impacto académico como por su crecimiento comercial, son *Pacific Biosciences* (PacBio) y *Oxford Nanopore Technologies* (ONT).

La primera tecnología en consolidarse fue desarrollada por PacBio en 2010 con *Single Molecule Real-Time sequencing* (SMRT). Esta tecnología permite observar en tiempo real la incorporación de nucleótidos en una única molécula de ADN durante la síntesis, sin amplificación. Esto es posible gracias al uso de un chip denominado *SMRT Cell*. Estas estructuras contienen miles de nanocavidades denominadas *zero-mode waveguides* (ZMWs), las cuales permiten focalizar la señal fluorescente de la actividad

de la ADN polimerasa en un volumen extremadamente reducido, permitiendo detectar eventos individuales de incorporación de nucleótidos con gran sensibilidad y resolución temporal (Eid et al., 2009; Metzker, 2010). El ADN molde se circulariza mediante adaptadores SMRTbell, lo que da lugar a la polimerasa recorrer repetidamente la misma secuencia y generar múltiples lecturas (subreads), que se combinan luego en una lectura de alta fidelidad o HiFi. Este enfoque permite corregir errores aleatorios asociados a lecturas individuales largas, mejorando significativamente la precisión. Las lecturas HiFi alcanzan longitudes de 10 a 25 kb con una precisión de ~99.8% comparable o superior a la de Illumina (Wenger et al., 2019; Hu et al., 2021). Gracias a esta combinación única de precisión y cobertura, las lecturas HiFi se han convertido en una herramienta clave para ensamblados *de novo*, análisis estructurales complejos y estudios en regiones genómicas altamente repetitivas o difíciles de mapear.

En 2015, *Oxford Nanopore Technologies* (ONT) propuso un cambio radical en las técnicas de secuenciación de ADN al reemplazar la detección por fluorescencia con la medición de variaciones en una corriente iónica. Su principio se basa en hacer pasar una molécula de ADN o ARN a través de un nanoporo proteico, insertado en una membrana artificial, mientras se mide la corriente iónica. Cada nucleótido que atraviesa el poro altera la señal eléctrica de manera característica, y estas perturbaciones son interpretadas por algoritmos de *basecalling* para inferir la secuencia (Kasianowicz et al., 1996; Huang et al. 2015). Entre las plataformas brindadas por ONT, se encuentra el MinION, un dispositivo portátil de 90 g (Figura 2). Su componente clave es una celda que contiene 2048 nanoporos direccionables de forma individual, ofreciendo la posibilidad de paralelizar la secuenciación de múltiples moléculas. Previo a la secuenciación, se ligan adaptadores al ADN para facilitar su reconocimiento por enzima helicasa que controla el avance del ADN a través del poro. Un adaptador tipo horquilla une ambas hebras de la molécula, permitiendo una secuenciación continua de la hebra sentido y antisentido. Durante el paso del ADN, los cambios en la corriente iónica generados por cada combinación de base se segmentan en eventos eléctricos discretos que son interpretados por algoritmos de *basecalling*. Su portabilidad ha permitido su uso en terreno, incluso en condiciones extremas, como estaciones polares o zonas de brotes epidémicos (Jain et al., 2016; Quick et al., 2016). También se encuentran

disponibles plataformas no portables (GridION y PromethION) que superan a MinION en capacidad y escalabilidad. Mientras que MinION permite correr una sola *flow cell* y depende de una computadora externa, GridION puede ejecutar hasta 5 *flow cells* en paralelo con basecalling incluido, y PromethION escala aún más, permitiendo secuenciar hasta 48 *flow cells* simultáneamente, ideal para estudios a gran escala como genómica o transcriptómica masiva (Kumar et al., 2019; Oxford Nanopore Technologies, 2018).



Figura 2: Vista comparativa del secuenciador MinION versión Mk1C (Oxford Nanopore Technologies) junto a un marcador indeleble que destaca la portabilidad de este dispositivo.

Las plataformas de secuenciación de ONT se destacan por generar lecturas ultra largas, de hasta 2 Mb, lo que las convierte en herramientas valiosas para ensamblados *de novo*, resolución de regiones repetitivas y detección de variantes estructurales complejas (Jain et al., 2016). Otra ventaja significativa es la posibilidad de secuenciar ARN directamente, sin necesidad de retrotranscripción a ADNc, lo que permite detectar modificaciones post-transcripcionales en tiempo real (Garalde et al., 2018). Además, Schreiber et al. (2013) demostraron que los sistemas de nanoporos pueden detectar variaciones covalentes en citocinas, lo que posiciona a esta tecnología como una herramienta eficaz para el análisis epigenético directo. Aunque las lecturas individuales presentan una menor precisión (90-98%), especialmente en regiones homopoliméricas, los avances en algoritmos de *basecalling* y en las quimiotecas han mejorado notablemente esta limitación (Bayega et al., 2021), y el uso de identificadores

moleculares únicos (UMIs) ha permitido aumentar aún más la precisión del análisis final (Karst et al., 2021).

APLICACIONES DE NGS

El desarrollo de plataformas de secuenciación de nueva generación (NGS), junto con métodos innovadores para la preparación de bibliotecas de distintas fracciones de ácidos nucleicos, han producido un gran cambio en el campo de las ciencias genómicas. La generación masiva y a costos relativamente bajos de datos genómicos, han impulsado el diseño de herramientas bioinformáticas cada vez más sofisticadas para su análisis. Como resultado, ha sido posible consolidar las bases de la genómica estructural y funcional, ampliando nuestro conocimiento sobre la organización, expresión y regulación de los genomas en una gran diversidad de áreas como la medicina, la agricultura y la ecología. Entre las aplicaciones genómicas más relevantes se destacan la secuenciación de genomas completos, exomas y transcriptomas (RNA-seq), el análisis de interacción proteína-ADN mediante ChIP-seq, la caracterización de patrones epigenéticos como la metilación del ADN, estudios metagenómicos y metatranscriptómicos, y el mapeo de la accesibilidad de la cromatina.

Secuenciación del genoma completo

La secuenciación del genoma completo (WGS, por sus siglas en inglés) es una herramienta fundamental para analizar el ADN de un organismo en su totalidad, incluyendo tanto regiones codificantes como no codificantes (Ng y Kirkness, 2010). El establecimiento de genomas de referencia, iniciado con el Proyecto Genoma Humano (1990- 2003), fue clave para impulsar esta metodología (Chial, 2008). Desde entonces, los avances en tecnologías de secuenciación de alto rendimiento han ido mejorando significativamente la calidad del ensamblaje, obteniéndose las versiones más completas y recientes del genoma humano, las versiones GRCh38 y T2T-CHM13 (Aganezov et al., 2022). El establecimiento del genoma humano como referencia marcó un hito que ha servido de punto de partida para el desarrollo de múltiples iniciativas genómicas a nivel global. Proyectos como el *International Wheat Genome Sequencing Consortium*, el *Rice Genome Project* y el *Peanut Genome Initiative* han centrado sus esfuerzos en la obtención de genomas completos de especies de interés agronómico, impulsando así el

avance en la mejora genética y la agricultura de precisión. Asimismo, la secuenciación genómica de individuos y la comparación contra un genoma de referencia previamente conocido es ampliamente utilizada para identificar variaciones genéticas en numerosos campos como la medicina y la agrobiotecnología.

Secuenciación de exomas

La secuenciación del exoma completo es una técnica que permite analizar específicamente las regiones codificantes del genoma, conocidas como exones, mediante su captura con sondas específicas. Al representar menos del 2% del genoma eucariota, la secuenciación del exoma resulta una alternativa más económica y focalizada que el WGS. Su aplicación ha sido fundamental para identificar mutaciones codificantes asociadas a enfermedades humanas (Bamshad et al., 2011). Sin embargo, también ha sido ampliamente utilizada en otras especies. Por ejemplo, en maíz (*Zea mays*), se ha estimado que los polimorfismos en regiones génicas explican gran parte de la variación natural en caracteres cuantitativos (Li et al., 2012). En trigo, experimentos de secuenciación de exones fueron realizados con el fin generar una base de datos de mutantes de trigo, que contiene mutaciones inducidas por diferentes agentes mutagénicos, con el propósito de proveer una mejor comprensión de la biología y fomentar el mejoramiento agronómico de este cultivo (Xiong et al., 2023). En el maní cultivado (*Arachis hypogaea*), la secuenciación de los exomas de más de 50 accesiones de la especie diploide silvestre *Arachis duranensis* ha sido utilizada en análisis genómicos comparativos para identificar la población de esta especie que más probablemente habría participado como parental materno en el proceso de aloploidización que dió origen al maní (Bertioli et al., 2019).

Secuenciación de ARN

El transcriptoma comprende el conjunto completo de transcritos presentes en un tipo celular o tejido determinado, en una etapa específica del desarrollo o bajo una condición fisiológica o ambiental particular. El análisis transcriptómico ofrece una visión integral de los mecanismos moleculares involucrados en diversos procesos, revelando información clave sobre la estructura, función y regulación de los genes (Wei et al., 2011). Actualmente, la caracterización del transcriptoma se realiza mediante

plataformas de secuenciación de alto rendimiento originalmente diseñadas para ADN. Por ello, es necesario convertir enzimáticamente el ARN en ADN complementario (ADNc) antes del proceso de secuenciación. Esta estrategia ha sido ampliamente utilizada para el ensamblaje y anotación de transcriptomas, la identificación de genes diferencialmente expresados, el estudio de variantes de *splicing* alternativo y el análisis de ARN no codificantes. La combinación de análisis de datos transcriptómicos y datos de secuenciación genómica han sido ampliamente utilizados para conocer las bases moleculares responsables de la aparición de ciertos rasgos fenotípicos en diferentes organismos. En este sentido, en el sector agropecuario, el análisis conjunto de datos genómicos y transcriptómicos ha permitido identificar genes asociados a la tolerancia al estrés, lo que ha facilitado el desarrollo de cultivos más resilientes frente al cambio climático (Roy et al., 2024), y ha profundizado nuestra comprensión de los mecanismos moleculares en plantas (Purugganan y Jackson, 2021) y al diagnóstico eficiente de enfermedades vegetales (Saeed et al., 2023). En medicina, han impulsado el desarrollo de la genómica personalizada, permitiendo diagnósticos más precisos de enfermedades raras y comunes, el diseño de inmunoterapias, y la implementación de tratamientos guiados por el perfil genómico de los pacientes (Popova y Carabetta, 2024). Por ejemplo, estudios combinados de genomas y transcriptomas hacen posible la identificación de mecanismos de resistencia o sensibilidad a los distintos fármacos, como el caso de mutaciones o sobreexpresión de genes como P53, HER2, BRCA1, BARD1 o genes relacionados con proteínas de shock térmico (Hsp70), que están asociadas a resistencia a ciertos fármacos en cánceres como el de ovario o mama (Khodadadian et al., 2020).

Inmunoprecipitación de la cromatina y secuenciación

Las modificaciones de histonas, como la acetilación, metilación, fosforilación, entre otras, son marcas epigenéticas críticas que regulan la estructura de la cromatina y, en consecuencia, la expresión génica. Para identificar regiones del ADN asociadas a estas modificaciones, se emplea la técnica de Chip-Seq, que combina inmunoprecipitación con anticuerpos específicos y secuenciación de los fragmentos recuperados mediante NGS (Park, 2009; Furey, 2012). Por ejemplo, la secuenciación de fragmentos de ADN capturados mediante la utilización de anticuerpos específicos de la

trimetilación y de la acetilación de lisina 9 de la histona H3 (H3K9me3) permiten detectar regiones asociadas con la heterocromatina y con potenciadores y promotores activos, respectivamente (Nakato y Sakata, 2021). Cabe destacar que el uso de Chip-seq ha sido fundamental para la anotación de los estados de la cromatina en una amplia variedad de líneas celulares y tejidos, gracias al trabajo de grandes consorcios como ENCODE, el Roadmap Epigenomics Consortium y el International Human Epigenome Consortium (Nakato y Sakata, 2021).

Análisis de metilación del ADN

La metilación del ADN es una de las modificaciones epigenéticas más estudiadas y se refiere a la adición de un grupo metilo en el carbono 5 de la base citosina (5-metilcitosina, 5mC), especialmente en dinucleótidos CpG. Las tecnologías de NGS han permitido analizar los patrones de metilación con resolución de nucleótido individual. Entre los métodos más utilizados para el análisis de metilación mediante NGS se destacan la secuenciación del genoma completo con bisulfito, que permite una cobertura global de los sitios CpG, y la secuenciación con bisulfito de representación reducida (RRBS), que utiliza enzimas de restricción para enriquecer regiones ricas en CpG y reducir costos de secuenciación (Satam et al., 2023). Más recientemente, la secuenciación de ONT se ha presentado como una alternativa innovadora. Esta tecnología detecta directamente modificaciones epigenéticas sin necesidad de tratamientos químicos, mediante el análisis de los cambios en la señal eléctrica que se producen cuando las bases modificadas, como la 5mC, atraviesan un nanoporo. Estos enfoques han revolucionado el estudio epigenético del ADN, proporcionando herramientas fundamentales para comprender cómo la metilación actúa como un regulador clave de procesos celulares y del desarrollo en eucariotas superiores (Schreiber et al., 2022).

Metagenómica y metatranscriptómica

La metagenómica se refiere al análisis de comunidades microbianas presentes en muestras ambientales mediante la caracterización cualitativa y cuantitativa de secuencias de ADN (Hugenholtz & Tyson, 2008). Este enfoque evita la necesidad de aislar y cultivar los microorganismos en laboratorio, permitiendo estudiar directamente

comunidades complejas presentes en agua, sedimentos, plantas, contenidos intestinales, muestras bentónicas y heces, provenientes de ambientes acuáticos y terrestres. El desarrollo de tecnologías de NGS ha sido clave para obtener directamente datos genómicos a partir de estas muestras (Sogin et al., 2006). Actualmente, se emplean dos estrategias principales en los análisis metagenómicos: la secuenciación dirigida (targeted sequencing), basada en la amplificación de genes marcadores como el 16S rRNA (bacterias y arqueas), el 18S rRNA (eucariotas) o el ITS (hongos); y la metagenómica de secuenciación masiva (shotgun metagenomics), que implica la secuenciación directa de todo el ADN presente en la muestra, sin amplificación previa. En paralelo, la metatranscriptómica ha emergido como una herramienta complementaria que permite explorar la expresión génica dentro de comunidades microbianas en contextos ambientales (Aplakidou et al., 2024).

Tanto la metagenómica como la metatranscriptómica tienen aplicaciones claves en múltiples disciplinas. La secuenciación masiva de ADN ha permitido monitorear biodiversidad marina y terrestre (Othman et al. 2021), evaluar la calidad ambiental (Cameron et al. 2021) y para la identificación de especímenes biológicos *in situ* (Pomerantz et al., 2022) abriendo nuevas perspectivas para análisis ecológicos y evolutivos. En salud humana, la metagenómica ha permitido caracterizar el microbioma intestinal y su vínculo con enfermedades como la obesidad o la enfermedad inflamatoria intestinal (Turnbaugh et al., 2006; Qin et al., 2010), mientras que la metatranscriptómica ha revelado patrones de expresión génica microbiana durante infecciones o tratamientos antimicrobianos (Franzosa et al., 2014). En el ámbito de la salud pública, fueron esenciales para la detección temprana y el monitoreo de variantes del SARS-CoV-2 (John et al., 2021), la vigilancia de la resistencia antimicrobiana (Struelens & Palm, 2024). En agricultura, se han empleado para identificar microorganismos promotores del crecimiento vegetal y estudiar su expresión bajo condiciones de estrés abiótico (Sessitsch et al., 2012; Naylor et al., 2017). En biotecnología industrial, han facilitado el descubrimiento de enzimas novedosas con aplicaciones en biodegradación de plásticos y producción de biocombustibles (Takasuka et al., 2014; Berlemont y Martiny, 2015). Además, en seguridad alimentaria, estas herramientas se emplean para la trazabilidad de productos, la detección de patógenos y la vigilancia de brotes (Masenya et al., 2024).

Mapeo de accesibilidad de la cromatina

La caracterización de los elementos reguladores en “cis”, como promotores y potenciadores, resultan fundamentales para comprender los mecanismos que regulan la transcripción génica. Uno de los métodos utilizados para localización de estas regiones en el genoma es la técnica DNase-seq, que combina la digestión del ADN genómico con nucleasa ADNasa I seguida de secuenciación NGS. Esta técnica fue una de las primeras herramientas genómicas desarrolladas para identificar estas regiones, y se aplicó ampliamente en proyectos colaborativos como *ENCODE* y *Roadmap Epigenomics* (Karabacak Calviello et al., 2019). Por otra parte, Buenrostro y colaboradores (2013) desarrollaron un método que utiliza la enzima transposasa Tn5 modificada para fragmentar y etiquetar regiones accesibles de la cromatina, las cuales son posteriormente secuenciadas mediante NGS. Entre múltiples aplicaciones de esta técnica, en *Tritipyrum*, por ejemplo, ATAC-seq permitió mapear regiones accesibles del genoma en respuesta a estrés salino, particularmente motivos de unión a factores de transcripción de las familias MYB, AP2/EREBP, bZIP, bHLH y WRKY (Tian et al., 2024).

PERSPECTIVAS FUTURAS

Las perspectivas futuras en el campo de la secuenciación apuntan a una integración aún más profunda entre tecnologías genómicas y aplicaciones clínicas, agrícolas y ecológicas. Se prevé que el desarrollo de métodos más rápidos, económicos y precisos —incluyendo plataformas híbridas y enfoques de lectura ultra-larga— permitirá una secuenciación verdaderamente ubicua y democratizada. En particular, aplicaciones emergentes como la secuenciación a nivel de célula única y la epigenómica de alta resolución abrirán nuevas posibilidades para caracterizar la heterogeneidad celular, los mecanismos regulatorios y la adaptación a diferentes entornos biológicos. Estas tendencias, junto con la consolidación de datos multi-ómicos, marcarán los próximos avances en la genómica de precisión

CONCLUSIONES

La evolución de las tecnologías de secuenciación ha transformado profundamente las ciencias genómicas. Desde la precisión y simplicidad del método de Sanger, pasando por la masividad y reducción de costos de la nueva generación, hasta la resolución estructural y versatilidad de las tecnologías de tercera generación. Estas plataformas han permitido consolidar la genómica estructural y funcional, generar genomas de referencia de alta calidad, impulsar la medicina personalizada, y abrir nuevas oportunidades en agricultura, biotecnología, microbiología y ciencias ambientales. Si bien el futuro señala la expansión hacia la secuenciación de célula única, la epigenómica y la integración multi-ómica, los avances actuales ya constituyen un cambio de paradigma en la forma de explorar y comprender la información genética.

REFERENCIAS BIBLIOGRÁFICAS

- Aganezov S, Yan SM, Soto DC, Kirsche M. (2022) A complete reference genome improves analysis of human genetic variation. *Science* 376-533.
- Amodio, N., Raimondi, L., Juli, G., Stamato, M. A., Caracciolo, D., Tagliaferri, P., Tassone, P. (2018). MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. *Journal of hematology & oncology*, 11, 1-19.
- Aplakidou E, Vergoulidis N, Chasapi M et al (2024) Visualizing metagenomic and metatranscriptomic data: a comprehensive review. *Comput Struct Biotechnol J* 23:2011–2033.
- Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews Genetics*, 12(11), 745-755.
- Bayega, A., Oikonomopoulos, S., Gregoriou, M. E., Tsoumani, K. T., Giakountis, A., Wang, Y. C., Ragoussis, J. (2021). Nanopore long-read RNA-seq and absolute

quantification delineate transcription dynamics in early embryo development of an insect pest. *Scientific reports*, 11(1), 7878.

Berlemont, R., Martiny, A. C. (2015). Genomic potential for polysaccharide deconstruction in bacteria. *Applied and environmental microbiology*, 81(4), 1513-1519.

Cameron, E. S., Schmidt, P. J., Tremblay, B. J. M., Emelko, M. B., & Müller, K. M. (2021). Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Scientific reports*, 11(1), 22302.

Chial, H. (2008). DNA sequencing technologies key to the Human Genome Project. *Nature Education*, 1(1).

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Schmutz, J. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nature genetics*, 51(5), 877-884.

Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., Reid, C. A. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 327(5961), 78-81.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., ... & Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 323(5910), 133-138.

Franzosa, E. A., Morgan, X. C., Segata, N., Waldron, L., Reyes, J., Earl, A. M., Huttenhower, C. (2014). Relating the metatranscriptome and metagenome of the human gut. *Proceedings of the National Academy of Sciences*, 111(22), E2329-E2338.

Furey, T. S. (2012). ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nature Reviews Genetics*, 13(12), 840-852.

Gansauge, M. T., Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nature protocols*, 8(4), 737-748.

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., Turner, D. J. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods*, 15(3), 201-206.

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.

Gong, Y., Li, Y., Liu, X. *et al.* (2023) A review of the pangenome: how it affects our understanding of genomic variation, selection and breeding in domestic animals? *J Animal Sci Biotechnol* 14, 73.

Heather, J. M., Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1), 1-8.

Hilt, E. E., Ferrieri, P. (2022). Next generation and other sequencing technologies in diagnostic microbiology and infectious diseases. *Genes*, 13(9), 1566.

Hu, T., Chitnis, N., Monos, D., Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human immunology*, 82(11), 801-811.

Huang, S., Romero-Ruiz, M., Castell, O. K., Bayley, H., Wallace, M. I. (2015). High-throughput optical sensing of nucleic acids in a nanopore array. *Nature nanotechnology*, 10(11), 986-991.

Hugenholtz, P., Tyson, G. W. (2008). Metagenomics. *Nature*, 455(7212), 481-483.

Jain, M., Olsen, H. E., Paten, B., Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology*, 17, 1-11.

John, T. M., Jacob, C. N., Kontoyiannis, D. P. (2021). When uncontrolled diabetes mellitus and severe COVID-19 converge: the perfect storm for mucormycosis. *Journal of fungi*, 7(4), 298.

Karabacak Calviello, A., Hirsekorn, A., Wurmus, R., Yusuf, D., Ohler, U. (2019). Reproducible inference of transcription factor footprints in ATAC-seq and DNase-seq datasets using protocol-specific bias modeling. *Genome biology*, 20, 1-13.

Karst, S. M., Ziels, R. M., Kirkegaard, R. H., Sørensen, E. A., McDonald, D., Zhu, Q., Albertsen, M. (2021). High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nature methods*, 18(2), 165-169.

Kasianowicz, J. J., Brandin, E., Branton, D., Deamer, D. W. (1996). Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences*, 93(24), 13770-13773.

Khodadadian, A., Darzi, S., Hagh-Dareh, S., Sadat Eshaghi, F., Babakhanzadeh, E., Mirabutalebi, S. H., Nazari, M. (2020). Genomics and Transcriptomics: The Powerful Technologies in Precision Medicine. *International Journal of General Medicine*, 13, 627–640.

Kingsmore, S. F., Smith, L. D., Kunard, C. M., Bainbridge, M., Batalov, S., Benson, W., Defay, T. (2022). A genome sequencing system for universal newborn screening, diagnosis, and precision medicine for severe genetic diseases. *The American Journal of Human Genetics*, 109(9), 1605-1619.

- Kirkness, E. F., Haas, B. J., Sun, W., Braig, H. R., Perotti, M. A., Clark, J. M., Pittendrigh, B. R. (2010). Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proceedings of the National Academy of Sciences*, 107(27), 12168-12173.
- Kumar, A., Sharma, V. K., Kumar, P. (2019). Nanopore sequencing: The fourth-generation sequencing. *J Entom Zool Stud*, 7, 1400-1403.
- Kulski, J. (Ed.). (2016). Next generation sequencing: advances, applications and challenges. 1-95.
- Li X, Zhu C, Yeh CT, Wu W, Takacs EM, Petsch KA, Tian F, Bai G, Buckler ES, Muehlbauer GJ, Timmermans MC, Scanlon MJ, Schnable PS, Yu J (2012b) Genic and nongenic contributions to natural variation of quantitative traits in maize. *Genome Res* 22:2436–2444
- Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology*, 30(5), 434-439.
- Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.*, 9(1), 387-402.
- Masenya, K., Manganyi, M. C., & Dikobe, T. B. (2024). Exploring cereal metagenomics: unravelling microbial communities for improved food security. *Microorganisms*, 12(3), 510.
- Matheson, P., McGaughan, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports*, 12(1), 13987.

- Maxam, A. M., Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences*, 74(2), 560-564.
- McCombie, W. R., McPherson, J. D., Mardis, E. R. (2019). Next-generation sequencing technologies. *Cold Spring Harbor perspectives in medicine*, 9(11), a036798.
- Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature reviews genetics*, 11(1), 31-46.
- Moti, T. B., Guyassa, C., Dima, C., Home, M. (2022). Illumina sequencing technology review. *image*, 10, 25-31.
- Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., Salichos, L., Gerstein, M. (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome biology*, 17, 1-9.
- Nakato, R., Sakata, T. (2021). Methods for ChIP-seq analysis: a practical workflow and advanced applications. *Methods*, 187, 44-53.
- National Human Genome Research Institute. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). <http://www.genome.gov/sequencingcosts>. Accessed 22 Feb 2016.
- Naylor, D., DeGraaf, S., Purdom, E., Coleman-Derr, D. (2017). Drought and host selection influence bacterial community dynamics in the grass root microbiome. *The ISME journal*, 11(12), 2691-2704.
- Ng PC, Kirkness EF (2010) Whole genome sequencing. *Methods Mol Biol* 628:215–226.
- Othman, N., Haris, H., Fatin, Z., Najmuddin, M. F., Sariyati, N. H., Md-Zain, B. M., Abdul-Latiff, M. A. B. (2021, April). A review on environmental DNA (eDNA)

- metabarcoding markers for wildlife monitoring research. In IOP conference series: Earth and environmental science (Vol. 736, No. 1, p. 012054). IOP Publishing.
- Oxford Nanopore Technologies. (2018). MinION Mk1B IT requirements. Recuperado de <https://nanoporetech.com/document/requirements>
- Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews genetics*, 10(10), 669-680.
- Pomerantz, A., Sahlin, K., Vasiljevic, N., Seah, A., Lim, M., Humble, E., Prost, S. (2022). Rapid in situ identification of biological specimens via DNA amplicon sequencing using miniaturized laboratory equipment. *Nature Protocols*, 17(6), 1415-1443.
- Popova, L., Carabetta, V. J. (2024). The use of next-generation sequencing in personalized medicine. In *High Throughput Gene Screening: Methods and Protocols* (pp. 287-315). New York, NY: Springer US.
- Purugganan, M. D., y Jackson, S. A. (2021). Advancing crop genomics from lab to field. *Nature genetics*, 53(5), 595-601.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., ... & Wang, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *nature*, 464(7285), 59-65.
- Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., Carroll, M. W. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*, 530(7589), 228-232.
- Rhoads, A., Au, K. F. (2015). PacBio sequencing and its applications. *Genomics, proteomics & bioinformatics*, 13(5), 278-289.

Roy, A., Dutta, S., Das, S., Choudhury, M. R. (2024). Next-Generation Sequencing in the Development of Climate-Resilient and Stress-Responsive Crops—A Review.

The Open Biotechnology Journal, 18(1).

Saeed, M., Jamil, Z., Shehzad, T., ul Hasan, S. Z., Bibi, R., Malik, S. N., Ahmed, R. (2023). Role of Next Generation Sequencing (NGS) in Plant Disease Management: A Review. *Journal of Applied Research in Plant Sciences*, 4(01), 512-517.

Sanger, F., Nicklen, S., Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences*, 74(12), 5463-5467.

Satam, H., Joshi, K., Mangrolia, U., Waghoo, S., Zaidi, G., Rawool, S., Malonia, S. K. (2023). Next-generation sequencing technology: current trends and advancements. *Biology*, 12(7), 997.

Schreiber, M., Gao, Y., Koch, N., Fuchs, J., Heckmann, S., Himmelbach, A., Dreissig, S. (2022). Recombination landscape divergence between populations is marked by larger low-recombining regions in domesticated rye. *Molecular Biology and Evolution*, 39(6), msac131.

Schreiber J, Wescoe ZL, Abu-Shumays R, Vivian JT, Baatar B, Karplus K, et al. (2013) Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc Natl Acad Sci U S A*. 110:18910–5.

Sessitsch, A. Hardoim, P; Döring, J., Weilharter, A., Krause, A., Woyke, T., Reinhold-Hurek, B. (2012). Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Molecular Plant-Microbe Interactions*, 25(1), 28-36.

- Sharma, P., Kumar, S., Pandey, A. (2021). Bioremediated techniques for remediation of metal pollutants using metagenomics approaches: a review. *Journal of Environmental Chemical Engineering*, 9(4), 105684.
- Sogin ML, Morrison HG, Huber JA *et al.* (2006) Microbial diversity in the deep sea and the underexplored ‘rare biosphere’. *Proceedings of the National Academy of Sciences, USA*, 103, 12115–12120.
- Struelens, M. J., Ludden, C., Werner, G., Sintchenko, V., Jokelainen, P. (2024). Real-time genomic surveillance for enhanced control of infectious diseases and antimicrobial resistance. *Frontiers in Science*, 2, 1298248.
- Takasuka, T. E., Walker, J. A., Bergeman, L. F., Meulen, K. A. V., Makino, S. I., Elsen, N. L., & Fox, B. G. (2014). Cell-free translation of biofuel enzymes. *Cell-Free Protein Synthesis: Methods and Protocols*, 71-95.
- Tian, L., Chen, C. J., Song, Y. N., Xu, K., Li, N. E., Zhang, X. H., Li, Y. (2024). Comprehensive genetic analysis reveals the mutational landscape of ABCA4-associated retinal dystrophy in a Chinese cohort. *Gene*, 891, 147832.
- Turnbaugh, P. J., Quince, C., Faith, J. J., McHardy, A. C., Yatsunenko, T., Niazi, F., Gordon, J. I. (2010). Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proceedings of the National Academy of Sciences*, 107(16), 7503-7508.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Kalush, F. (2001). The sequence of the human genome. *science*, 291(5507), 1304-1351.
- Waterston, R. H., Lander, E. S., Sulston, J. E. (2002). On the sequencing of the human genome. *Proceedings of the National Academy of Sciences*, 99(6), 3712-3716.

Wei, W., Qi, X., Wang, L., Zhang, Y., Hua, W., Li, D., Zhang, X. (2011).

Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC genomics*, 12, 1-13.

Wilkes, R. P. (2023). Next-generation diagnostics for pathogens. *Veterinary Clinics: Food Animal Practice*, 39(1), 165-173.

Williams, R. C., Hillary, L. S., Garcia-Delgado, A., Jameson, E., Kevill, J. L., Wade, M. J., Jones, D. L. (2025). Harnessing the Power of Next-Generation Sequencing in Wastewater-Based Epidemiology and Global Disease Surveillance. *Food and Environmental Virology*, 17(1), 1-7.

Xiong, H., Guo, H., Fu, M., Xie, Y., Zhao, L., Gu, J., Liu, L. (2023). A large-scale whole-exome sequencing mutant resource for functional genomics in wheat. *Plant Biotechnology Journal*, 21(10), 2047-2056.

Zanini, S. F., Bayer, P. E., Wells, R., Snowdon, R. J., Batley, J., Varshney, R. K., Golicz, A. A. (2022). Pangenomics in crop improvement—from coding structural variations to finding regulatory variants with pangenome graphs. *The Plant Genome*, 15(1), e20177.

Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics*. 2011 Mar 20;38(3):95-109.

Zhao, J., FA Grant, S. (2011). Advances in whole genome sequencing technology. *Current pharmaceutical biotechnology*, 12(2), 293-305.

Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. The next-generation sequencing technology and application. *Protein Cell*. 2010 Jun;1(6):520-36.