

Secuenciación de tercera generación en el diagnóstico hematológico

Javier Corral^{1,2*}, María E. de la Morena-Barrio^{1,2}, Carlos Bravo-Pérez^{1,2}, Javier Cuenca-Guardiola³, Juan J. Rojo-Carrillo¹, María Llamas-López¹, Pedro Garrido-Rodríguez^{1,2}, Esther Navarro¹, José Padilla¹, Raúl Teruel^{1,2}, Jesualdo Fernández-Breis³, Francisca Ferrer-Marín^{1,2}, José Rivera^{1,2}, Ma. Luisa Lozano^{1,2} y Belén de la Morena-Barrio^{1,2}

¹Servicio de Hematología, Hospital Universitario Morales Meseguer, Centro Regional de Hemodonación, Instituto Murciano de Investigación Sanitaria (IMIB) - Pascual Parrilla, Universidad de Murcia; ²Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER); ³Departamento de Informática y Sistemas, Universidad de Murcia, CEIR Campus Mare Nostrum, IMIB - Pascual Parrilla, Facultad de Informática, Campus de Espinardo. Murcia, España

Introducción

La secuenciación de tercera generación tiene fortalezas de gran proyección en el diagnóstico hematológico, como la obtención de lecturas muy largas (media 10Kb), la diferenciación de alelos y la identificación de marcas epigenéticas y epitranscriptómicas. En esta ponencia, además de indicar las principales características de esta tecnología, mostramos ejemplos concretos de su utilidad en hematología generados por nuestro propio grupo. En relación con la secuenciación de ADN, nuestros resultados avalan la validez de este método en la identificación y la caracterización con resolución nucleotídica de todo tipo de variantes estructurales; además, define grandes haplotipos (útiles desde la caracterización del efecto fundador de mutaciones hasta la discriminación del bialelismo), ha identificado la inserción de retrotransposones como nuevo mecanismo molecular implicado en diferentes enfermedades que no tenían defecto molecular asociado, y ha determinado firmas epigenéticas ligadas a variantes genéticas que explican el desarrollo de ciertas enfermedades. Las aplicaciones de esta tecnología al campo de la secuenciación de ARN pasan por conseguir la lectura completa de todos los transcritos, desde la cola de poli-A, lo que además de facilitar la identificación de nuevos transcritos permite un estudio de expresión diferencial mucho más robusto (basado en transcritos, no en genes), e identifica firmas

epitranscriptómicas cuyo impacto fisiopatológico todavía está poco estudiado.

La revolución de la secuenciación en biomedicina

La secuenciación de nueva generación (NGS, *next-generation sequencing*) ha revolucionado el campo de la biomedicina al permitir el análisis masivo, rápido y preciso de material genético. A diferencia de los métodos tradicionales de secuenciación, la NGS permite estudiar millones de fragmentos de ADN o ARN de manera simultánea, lo que ha facilitado avances significativos en la comprensión de enfermedades complejas, como el cáncer, enfermedades raras, trastornos genéticos y patologías infecciosas. Gracias a la NGS, y especialmente a su globalización por la accesibilidad de los equipamientos y la drástica reducción de los costes, es posible la secuenciación de todo el genoma y el transcriptoma. De esta forma, se han identificado variantes genéticas, se ha estudiado la expresión génica y se han detectado biomarcadores con una resolución sin precedentes, lo que ha impulsado el desarrollo de la medicina personalizada. En el siglo XXI, gracias a los avances en secuenciación se aspira a lograr un diagnóstico, un pronóstico y una decisión terapéutica cada vez con mayor precisión y más individualizados, que permitan una medicina preventiva eficaz y el uso de tratamientos que se adapten al perfil

*Correspondencia:

Javier Corral
E-mail: javier.corral@carm.es

Recibido: 06-07-2025

Aceptado: 06-08-2025

DOI: 10.24875/SANGRE.M25000076

Disponible en internet: 02-10-2025

Sangre. 2025;44(Supl 2):17-24

www.revistasangre.com

0036-4355 / © 2025 Sociedad Española de Hematología y Hemoterapia. Publicado por Permanyer. Éste es un artículo *open access* bajo la licencia CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

genético de cada paciente, mejorando su eficacia terapéutica y reduciendo los efectos adversos y los costes farmacéuticos. En resumen, la NGS representa una tecnología fundamental para el avance de la investigación biomédica y la transformación del cuidado de la salud, haciendo posible una medicina más precisa y personalizada.

Sin embargo, los sistemas de NGS de segunda generación basados en secuencias cortas tienen algunas limitaciones, como la dificultad de identificar variantes estructurales (VE), o la secuenciación de regiones repetitivas y con elevada homología. Tampoco consiguen identificar haplotipos ni discriminan la distribución de variantes en los dos alelos del gen, y no logran identificar variaciones epigenéticas.

A principios del siglo XXI, dos empresas, Pacific Biosciences y Oxford Nanopore Technologies, desarrollaron las técnicas de secuenciación de tercera generación basadas en la obtención de lecturas largas, con las que se ha conseguido ir rellenando los huecos todavía extensos que tenía el genoma humano, y lograr el que hasta ahora es el genoma de referencia más completo: T2T¹.

En esta revisión nos centraremos en la secuenciación con nanoporos, señalada por *Nature Genetics* en 2022 como método del año por las posibilidades y las fortalezas que tiene², mostrando las posibilidades de este método en hematología con ejemplos que nuestro grupo ha ido resolviendo en estos años.

Principios de la secuenciación con nanoporos

El principio de este método es muy simple: traducir los cambios de corriente que provoca el paso de una cadena de ácidos nucleicos cuando atraviesa un nanoporo proteico (Fig. 1).

La primera fortaleza del método es que no tiene ninguna restricción de longitud. Por ello, se pueden conseguir lecturas ultralargas de ADN. En nuestras manos, la longitud media idónea para conseguir la mayor eficacia de la célula de secuenciación (*flow cell*) es de 5-10 kb, pero hemos conseguido lecturas de más de 2,5 Mb. En el campo de la transcriptómica, la longitud de lectura media que conseguimos es de 750 bases, lo que supone que se obtienen lecturas completas de los transcritos, desde su cola de poli-A hasta el extremo 5'. Los beneficios de generar lecturas grandes son múltiples:

- Facilita el ensamblaje de las lecturas y asegura su correcto alineamiento, lo que sin duda es una extraordinaria ventaja al identificar y caracterizar VE.
- Permite identificar múltiples variantes genéticas en cada lectura, lo que posibilita la generación de haplotipos extensos que, entre otros beneficios, facilitan la diferenciación precisa de los alelos presentes en la muestra.

La segunda fortaleza de este sistema de secuenciación es que permite diferenciar cada nucleótido incluso con modificaciones epigenéticas. En otras palabras, diferencia una citosina de una metilcitosina o una hidroximetilcitosina, por lo que con un solo proceso se consiguen secuenciación y análisis epigenético, secuenciando tanto ADN como ARN.

La tercera fortaleza del sistema es la accesibilidad a los equipos. El coste del equipamiento más sencillo, el MinION, es testimonial, asumible por cualquier grupo de investigación u hospital. Con un ordenador relativamente potente y por menos de 1000 euros, que incluyen el *starting pack* con casi todos los reactivos necesarios y una *flow cell*, se podría realizar la secuenciación en cualquier lugar. También el equipo más avanzado, el PromethION, que permite la secuenciación completa del genoma humano, tiene un coste significativamente menor que los equipos similares de cadenas cortas.

Limitaciones de la secuenciación con nanoporos

Este sistema también tiene limitaciones, entre las que cabe destacar algunas importantes:

- Baja cobertura si se trabaja con ADN nativo para el estudio del genoma humano. Incluso en el equipamiento con mayor capacidad de secuenciación, la cobertura media empleando una sola *flow cell* es de 30x. No obstante, la secuenciación en tiempo real que permite esta tecnología ha facilitado el desarrollo de un atractivo sistema de enriquecimiento, el *adaptive sampling*³, que permite al equipo, definiendo las coordenadas de una o varias regiones de interés (hasta 3 Mb), tomar la decisión con las primeras 400 pb secuenciadas de continuar su lectura, si pertenece a la zona, o invertir la corriente y rechazarla, si no es de la zona. Con esta estrategia se consiguen enriquecimientos de lecturas largas de hasta 160x en la zona deseada, y cobertura de lecturas cortas en el resto del genoma (Fig. 2).

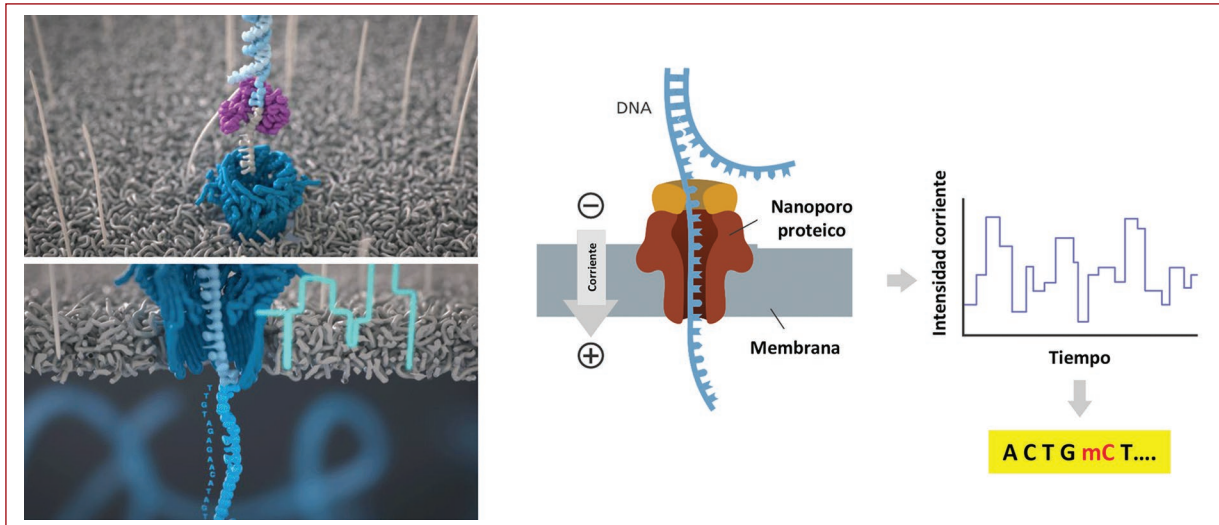


Figura 1. Proceso de secuenciación con nanoporos. El ADN o el ARN se acopla a un adaptador que le permite unirse a un nanoporo proteico embebido en una membrana lipídica y que actúa como proteína motora que va introduciendo la molécula de ácido nucleico por el poro. El paso de la cadena por el poro altera la corriente eléctrica que existe en la membrana, generando señales de intensidad, que son las que se interpretan para identificar cada nucleótido, diferenciando la existencia de modificaciones epigenéticas.

- Errores de secuenciación. Aunque los avances tanto en la química de las *flow cells* como en los procedimientos de identificación de bases han mejorado notablemente y superan el 99%, el grado de fiabilidad de este proceso sigue siendo inferior al de los sistemas de secuenciación de cadena corta, especialmente con pequeñas inserciones y deleciones⁴.
- El elevado coste del material fungible encarece el precio de cada base secuenciada, aunque se pueden incorporar códigos de barras y emplear varias muestras por *flow cell*, a costa de reducir la cobertura.
- Volumen de datos y necesidad computacional. La secuenciación requiere una alta capacidad computacional, especialmente si se emplea *adaptive sampling*, pues genera un gran volumen de datos (la secuenciación de un genoma completo supone más de 1 TB de datos), y su procesamiento necesita un equipamiento informático de altas prestaciones.
- Dificultad de interpretación de los resultados. Aunque se están desarrollando *pipelines* de análisis bioinformático que facilitan tanto el alineamiento como la identificación de VE, variantes puntuales, cambios epigenéticos y *phasing*, todavía son complejos el análisis y la interpretación de los resultados que genera este sistema de secuenciación.

Utilidad de la secuenciación con nanoporos en hematología

Identificación y caracterización con resolución nucleotídica de VE

Las VE son contribuyentes esenciales al desarrollo tanto de enfermedades raras como de neoplasias. De hecho, estas alteraciones genéticas, a pesar de ser numéricamente menores que las variantes puntuales o las pequeñas deleciones o inserciones, son las que mayor impacto tienen en cuanto a número de bases afectadas en todo el genoma y a potencial efecto funcional. Existen diferentes tecnologías que identifican VE, tanto de genes específicos (*Multiplex Ligation-dependent Probe Amplification* [MLPA], hibridación fluorescente *in situ* [FISH]) como del genoma completo (CGH, SNP array, *Optical Genome Mapping* [OGM]). Pero estos sistemas, incluso el más desarrollado (OGM), aunque definen la ganancia o la pérdida de material genético, no alcanzan resolución nucleotídica y no son capaces de definir con precisión la VE, especialmente si es compleja. Nuestro grupo ha contrastado en cinco patologías hematológicas, incluyendo enfermedades raras (deficiencia de antitrombina, deficiencia de factor XI, trombostenia de Glanzmann y síndrome de Hermansky-Pudlak) y leucemias agudas, la utilidad de la secuenciación con nanoporos para la identificación y la completa caracterización de VE implicadas en estas

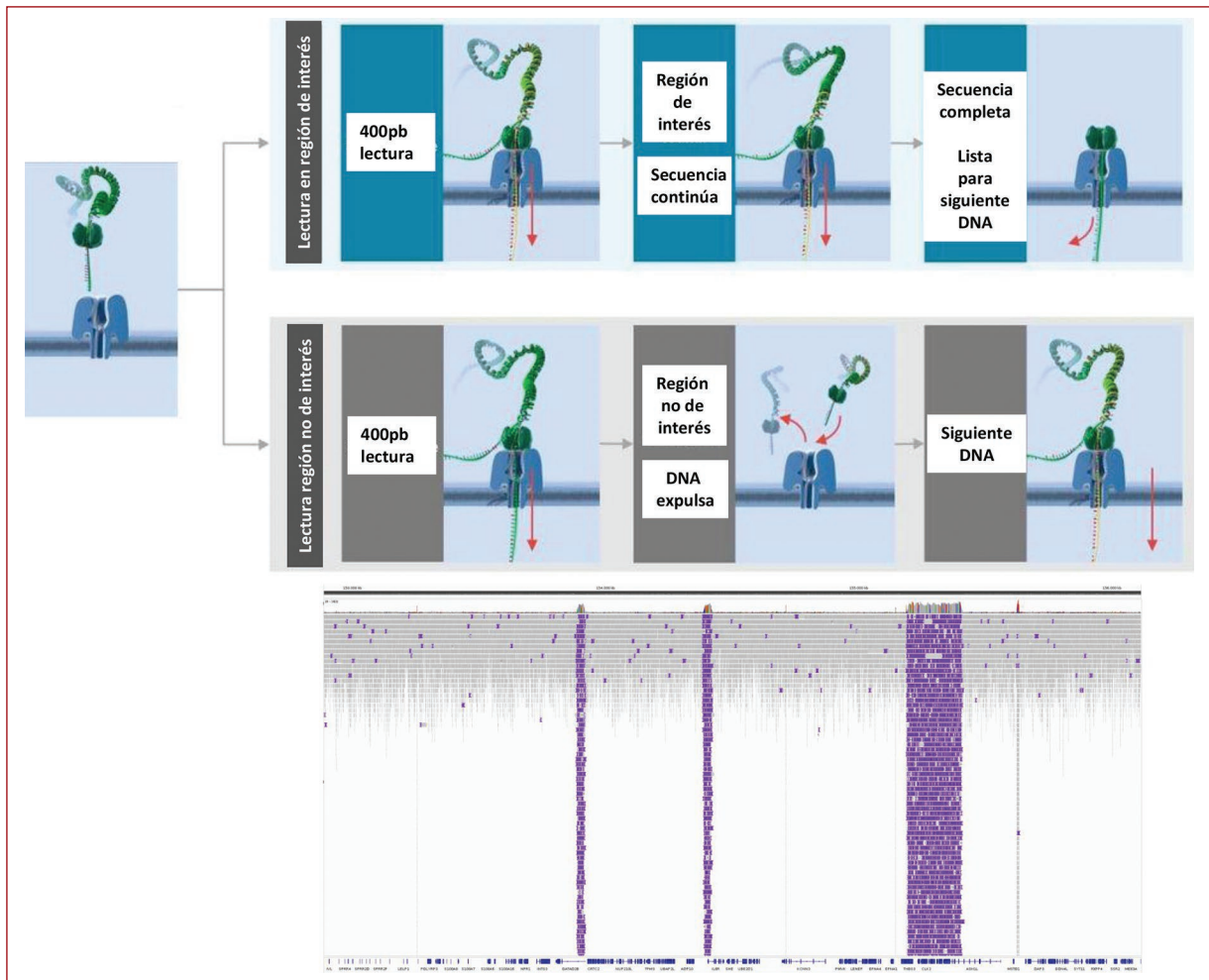


Figura 2. Enriquecimiento de secuenciación de regiones seleccionadas mediante *adaptive sampling*. Esta herramienta permite analizar las primeras 400 pb de cada secuencia para, si se localizan en la región de interés previamente definida, continuar la secuencia del resto de la cadena de ADN que está en el poro, pero si no se localizan en la región, se invierte el campo eléctrico y la cadena que se había introducido en el poro se expulsa, quedando este libre para aceptar una nueva cadena de ADN. De esta forma se consigue un enriquecimiento de lecturas largas en la región de interés, y lecturas cortas en el resto del genoma.

enfermedades, de cuya evidencia teníamos constancia por otros métodos (FISH, MLPA) o fueron identificadas *de novo*⁵⁻⁸. En todos los casos, se empleó *adaptive sampling* para enriquecer el gen o genes candidatos.

La secuenciación con nanoporos identificó las VE que presentaban los 16 casos analizados, independientemente de su tamaño (desde 190 pb hasta 8 Mb) y de su tipo (deleciones, duplicaciones, inversiones o VE complejas). Sin embargo, lo que es más interesante es que este estudio caracterizó con resolución nucleotídica los puntos de ruptura, definiendo con precisión la extensión y las características de cada VE, y definiendo los elementos que flanquean los puntos de ruptura. De esta forma, se pueden diseñar *primers* específicos para

amplificación de cada punto de ruptura, con lo que además de poder validar el resultado de la secuenciación se facilitan estudios complementarios, como genotipar la VE en estudios familiares. Esta información también permitió definir la existencia de microhomologías y elementos repetitivos flanqueantes de los puntos de ruptura para la mayoría de las VE evaluadas, lo que refuerza el papel de estos elementos en la aparición de VE y permite establecer los mecanismos moleculares implicados en su génesis. En un caso con inversión del cromosoma 16 identificada en un paciente con leucemia aguda, además de verificar que se trataba de una inversión atípica nos permitió predecir el posible ARN (y la proteína) de fusión entre *CBFB* y *MYH11*, y diseñar

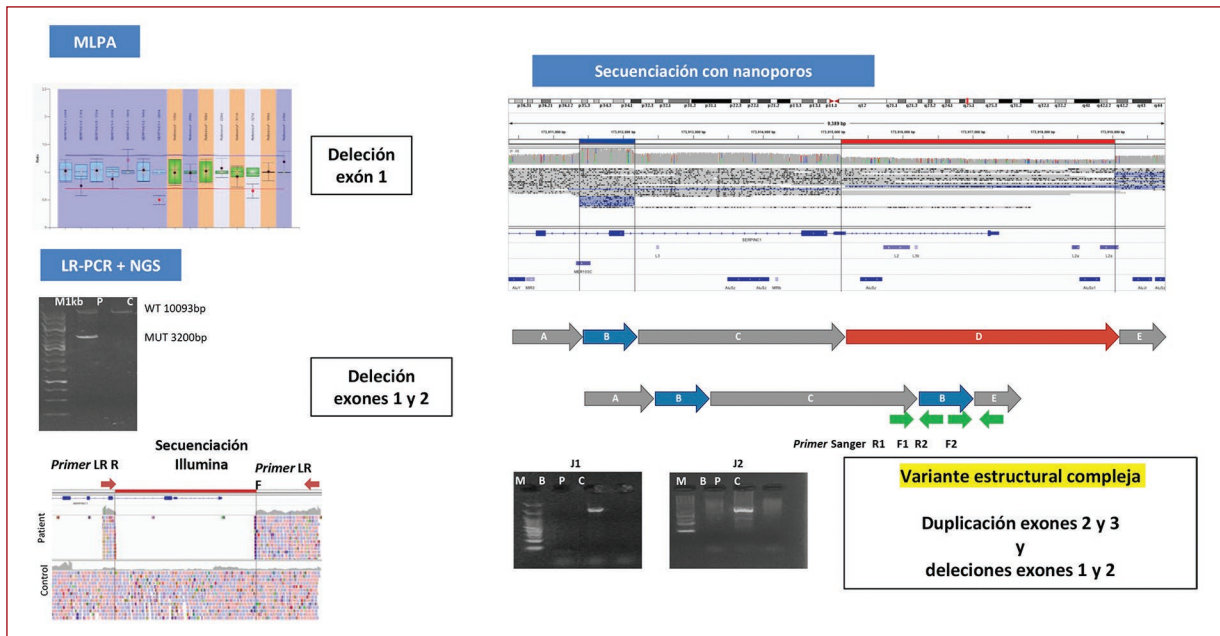


Figura 3. Fortaleza de la secuenciación con nanoporos para identificar y caracterizar variantes estructurales complejas que explican resultados contradictorios obtenidos con otros métodos. La secuenciación del gen *SERPINC1* de un paciente con deficiencia de antitrombina no identificó ninguna variante relevante. El estudio con MLPA identificó una delección heterocigota del exón 1. La amplificación de la zona cubierta por esta delección confirmó la delección, pero al secuenciar el producto de PCR se comprueba que, además del exón 1, el paciente presenta delección del exón 2, que no detectaba la MLPA. La secuenciación de cadenas largas con nanoporo genera lecturas que cubren toda la zona y permiten conocer la arquitectura genética final, y establecer que el paciente realmente tiene una variante estructural compleja resultado de dos procesos: duplicación de los exones 2 y 3, y delección de los exones 1 y 2. Como este procedimiento alcanza resolución nucleotídica, se pueden diseñar *primers* para amplificación genómica específica que permitieron validarla.

los *primers* adecuados para su validación y seguimiento de la enfermedad mínima residual⁷.

Finalmente, destacamos la identificación y la caracterización de VE complejas en tres patologías: deficiencia de antitrombina, trombastenia de Glanzmann y deficiencia de factor XI. Se trata de organizaciones genómicas complejas como resultado de diferentes tipos de VE que se originan en un mismo paciente (no se sabe si en uno solo o en diferentes procesos). Gracias a la resolución que alcanza la secuenciación con nanoporos se obtiene la organización genómica final resultante de los diferentes procesos, que podrían ser diagnosticados de manera incorrecta si se emplearan otros métodos. El ejemplo representativo es la delección del exón 1 que sugiere la MLPA en un caso con deficiencia de antitrombina, que con pruebas de reacción en cadena de la polimerasa (PCR) largas y secuenciación Illumina indicó una delección de los exones 1 y 2, pero que en realidad, como mostró la secuenciación con nanoporos, presentaba una duplicación de los exones 2 y 3 con posterior delección de un exón 2⁵ (Fig. 3).

En conclusión, nuestros estudios demuestran que la secuenciación con nanoporos es una herramienta robusta para la identificación y la caracterización de VE, destacan la heterogeneidad y la complejidad de estas grandes alteraciones genéticas, demostrando que las VE complejas pueden ser mucho más frecuentes que lo esperado, y sugieren la implicación de elementos repetitivos para la formación de las VE por mecanismos de reordenamiento genético comunes.

Identificación de grandes haplotipos

Este método ofrece la capacidad de definir haplotipos extensos y diferenciar las características genéticas de los alelos presentes en una muestra, información importante en diferentes contextos que a continuación trataremos de ilustrar con ejemplos de nuestro grupo.

EFFECTO FUNDADOR

La identificación de mutaciones recurrentes asociadas a una enfermedad podría explicarse por dos

mecanismos: la presencia de una zona caliente mutacional o la existencia de una mutación fundadora. Este fue el reto con que nos encontramos al identificar una variante puntual p.Thr147Ala en *SERPINC1* en 12 pacientes con deficiencia de antitrombina procedentes de diferentes países africanos. La ausencia de familiares disponibles para estudios de STR flanqueantes limitaba la capacidad de demostrar un efecto fundador. La forma de solventar este problema, y de manera rápida, elegante y económica, fue realizando la secuenciación con nanoporos. Para este proyecto empleamos dos PCR largas que solapaban con la variante y cubrían 14 Kb. Las PCR de los pacientes se marcaron con código de barras y empleando una sola *flow cell* se secuenciaron en 10 minutos, alcanzando una cobertura media de 2000x. Mediante una sencilla *pipeline* que permitía definir el *phasing* discriminamos los dos alelos y establecimos las variantes genéticas de cada uno de ellos. En todos los casos, el alelo mutado contenía otros 13 polimorfismos intragénicos, alguno con baja frecuencia. De este modo pudimos demostrar la existencia de la primera variante recurrente con efecto fundador causante de deficiencia de antitrombina en el continente africano⁹.

BIALELISMO

Determinar si dos o más variantes genéticas se encuentran en *cis* o *trans*, en el mismo alelo o en otro, tiene importancia en enfermedades recesivas. Las lecturas largas que genera este sistema permiten, en el mismo proceso, además de identificar las variantes, establecer si están en el mismo o en diferente alelo. Este fue el procedimiento que empleamos para diagnosticar trombostenia de Glanzmann o síndrome de Hermansly-Pudlak en casos con una variante puntual y una VE que afectaban a diferentes alelos de los genes *ITGB3* y *HPS3*, respectivamente⁶. También en enfermedades neoplásicas, definir el bialelismo de variantes puede tener valor pronóstico, como ocurre con las variantes de *TP53* en síndromes mielodisplásicos. Con una sola PCR larga se puede determinar si las dos variantes patogénicas identificadas en un paciente se localizan en el mismo o en diferentes alelos.

Combinación de datos de epigenética y genética

Una de las mayores fortalezas de la secuenciación con nanoporos es la capacidad de disponer de información epigenética y genética en el mismo proceso, e

incluso poder asociar las características epigenéticas con variantes genéticas. Un ejemplo del valor de esta información ha sido la identificación del mecanismo que explica por qué una mujer portadora de una variante patogénica en *PGK1*, gen del cromosoma X asociado con una anemia hemolítica muy grave (el hermano afecto, con grave neuropatía, murió joven como consecuencia de la anemia), desarrolló anemia hemolítica tras la adquisición de una mutación somática en *CALR* que provocó una trombocitemia esencial. El estudio del ADN nativo con nanoporos del gen *PGK1* en células CD15+ mostró una hipermetilación de la isla CpG de *PGK1* ligada al alelo silvestre que justifica que, en estas células, el único alelo que se transcribe es el mutado, como comprobamos al secuenciar con nanoporos el ADNc. Estos resultados sugieren que el evento clonal expandió progenitores mieloides con una firma epigenética, y explican cómo un desorden ligado al cromosoma X es dominante y específico de tejido en una mujer¹⁰.

Además de ser especialmente útil en enfermedades de *imprinting*, y para caracterizar la inactivación del cromosoma X, la información epigenética que se obtiene mediante secuenciación con nanoporos podría explicar mecanismos patogénicos en casos en los que los sistemas de secuenciación convencionales no han sido resolutivos, tanto en enfermedades monogénicas como tumorales.

Identificación de nuevos mecanismos patogénicos

Nuestro grupo ha demostrado la utilidad de este sistema de secuenciación para identificar un nuevo mecanismo patogénico en la deficiencia de antitrombina: la inserción de retrotransposones, elementos genéticos móviles muy abundantes en el genoma humano que presentan alta homología y tienen regiones repetitivas. El estudio molecular del gen candidato, *SERPINC1*, con secuenciación Sanger, NGS y MLPA, y el análisis de glicosilación en 350 casos, identificaron la base molecular en la mayoría de los casos. La secuenciación con nanoporos de 14 casos en los que no se encontraban alteraciones que justificaran la deficiencia de antitrombina identificó cuatro casos con dos retrotransposones (un SVA-E en tres casos y un SVA-F1 en el cuarto caso) insertados en distintos intrones del gen *SERPINC1*. Estos elementos son de difícil amplificación por PCR, y solo empleando *primers* flanqueantes e internos al propio retrotransposón, cuya secuencia se obtuvo de los datos generados mediante ensamblaje

de novo, se pudieron validar las inserciones y confirmar su segregación con la deficiencia de antitrombina en los estudios familiares¹¹. Además, pudimos comprobar que la inserción de los retrotransposones no afecta la metilación del gen afectado, por lo que especulamos que el mecanismo patogénico de estas inserciones radica en el cambio del patrón de procesamiento de intrones. Con este hallazgo, mostramos un nuevo mecanismo patogénico implicado en la deficiencia de antitrombina relativamente frecuente (hasta el 1% de los casos de nuestra cohorte) y de difícil diagnóstico molecular empleando los sistemas clásicos. Las colaboraciones con otros grupos del Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER) han mostrado que este mecanismo está implicado en otras enfermedades raras, resolviendo casos con diagnóstico incompleto hasta entonces¹².

El estudio genómico de elementos móviles en casos con deficiencia de antitrombina ha mostrado la deficiente información disponible sobre retrotransposones. El desarrollo de una *pipeline* para la identificación y la anotación de retrotransposones empleando secuenciación con nanoporos¹³ nos ha permitido identificar miles de inserciones no reportadas, mayoritariamente frecuentes, y polimorfismos responsables de la variabilidad genética, pero también inserciones que afectan a genes con potencial impacto patogénico, como la inserción de un elemento Alu en un exón del gen *BRCA2*¹⁴. Esta tecnología es idónea para establecer un mapa definitivo de elementos móviles en el genoma que facilite la interpretación de nuevos hallazgos, especialmente para definir su impacto patogénico.

Variaciones intrónicas profundas

Las variantes intrónicas profundas son otra fuente de variabilidad genética con potencial impacto patogénico que no se ha explorado con intensidad, y que la secuenciación con nanoporos puede relanzar. Estudiando casos con deficiencia de antitrombina sin base molecular, la secuenciación con nanoporos identificó variantes intrónicas patogénicas en dos casos. En uno identificamos dos inserciones en fase que afectan al intrón 1 y la región promotora del gen *SERPINC1* que posiblemente tengan efectos de regulación transcripcional. En otro caso identificamos cuatro variantes puntuales intrónicas profundas, también en fase, cuyo análisis individual no predice ningún efecto sobre el procesamiento del intrón, pero que al analizarse conjuntamente, como realmente se encuentran en el gen, causan la aparición de un sitio aceptor críptico.

Análisis de ARN

Queda finalmente aportar unas pinceladas sobre las posibilidades que ofrece la secuenciación del ARN nativo con nanoporos, la primera tecnología capaz de secuenciarlo directamente¹⁵. La lectura completa de todos los transcritos, desde la cola de poli-A, junto a la capacidad de identificar variantes epitranscriptómicas, puede revolucionar el campo de la secuenciación de ARN al permitir realizar análisis transcriptómicos mucho más robustos (sustentados en transcritos, no en genes); identificar nuevos transcritos con nuevas funciones fisiopatológicas, como los antisentido identificados en nuestro estudio¹⁶; identificar los transcritos generados por cada alelo; mostrar nueva diversidad genética o conocer el impacto de las variaciones epitranscriptómicas. Es un campo con un futuro prometedor que instamos a explorar en diferentes ámbitos¹⁷.

Conclusiones y futuras direcciones de la secuenciación con nanoporos

Las fortalezas de la secuenciación de cadenas largas con nanoporos hacen que esta tecnología tenga excelentes perspectivas de futuro, en especial si se consigue reducir los costes y aumentar la cobertura. Además, la secuenciación con nanoporos tiene tres retos a los que ya se está enfrentando y que pueden ofrecer en el futuro cercano excelentes resultados:

- La identificación de interacciones tridimensionales de rango amplio. Con el sistema PoreC se pretende obtener información de la organización de la cromatina y del plegamiento del ADN¹⁸.
- Definir con precisión secuencias repetitivas de diferentes tamaños, implicadas en estructuras complejas como los telómeros o en diferentes enfermedades.
- Secuenciación de proteínas¹⁹, aunque este campo parece el de más largo recorrido.

Agradecimientos

Este trabajo ha sido posible gracias al proyecto PMP21/00052 (ISCIII & Next Generation EU) y la Ayuda Cátedra Cajal (Universidad de Zaragoza).

Referencias

1. Nurk S, Koren S, Rhie A, Rautiainen M, Bizikadze AV, Mikheenko A, et al. The complete sequence of a human genome. *Science*. 2022; 376:44-53.
2. Method of the Year 2022: long-read sequencing. *Nat Methods*. 2023;20:1.
3. Payne A, Holmes N, Clarke T, Munro R, Debebe BJ, Loose M. Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nat Biotechnol*. 2021;39:442-50.

4. Ying YL, Hu ZL, Zhang S, Qing Y, Fragasso A, Maglia G, et al. Nanopore-based technologies beyond DNA sequencing. *Nat Nanotechnol.* 2022;17:1136-46.
5. de la Morena-Barrio B, Orlando C, Sanchis-Juan A, García JL, Padilla J, de la Morena-Barrio ME, et al. Molecular dissection of structural variations involved in antithrombin deficiency. *J Mol Diagn.* 2022;24:462-75.
6. Zamora-Cánovas A, de la Morena-Barrio B, Marín-Quilez A, Sierra-Aisa C, Male C, Fernández-Mosteirín N, et al. Targeted long-read sequencing identifies and characterizes structural variants in cases of inherited platelet disorders. *J Thromb Haemost.* 2024;22:851-9.
7. Bravo-Perez C, Cifuentes-Riquelme R, Padilla J, de la Morena-Barrio ME, Ortuño FJ, Garrido-Rodríguez P, et al. The whole is greater than the sum of its parts: Long-read sequencing for solving clinical problems in haematology. *J Cell Mol Med.* 2024;28:e17961.
8. de la Morena-Barrio B, Palomo A, Padilla J, Martín-Fernández L, Rojo-Carrillo JJ, Cifuentes R, et al. Impact of genetic structural variants in factor XI deficiency: identification, accurate characterization, and inferred mechanism by long-read sequencing. *J Thromb Haemost.* 2023;21:1779-88.
9. Orlando C, de la Morena-Barrio B, Pareyn I, Vanhoorelbeke K, Martínez-Martínez I, Vicente V, et al. Antithrombin p.Thr147Ala: the first founder mutation in people of African origin responsible for inherited antithrombin deficiency. *Thromb Haemost.* 2021;121:182-91.
10. Morales ML, Cano H, de la Morena-Barrio B, Vives-Corrons JL, Cuenca-Zamora EJ, Garrido-Rodríguez P, et al. CALR^{ins5}-mediated clonal hematopoiesis causes severe hemolytic anemia in a female PGK1^{Ser320Asn} carrier. *Blood Cancer J.* 2025;15:8.
11. de la Morena-Barrio B, Stephens J, de la Morena-Barrio ME, Stefanucci L, Padilla J, Miñano A, et al. Long-read sequencing identifies the first retrotransposon insertion and resolves structural variants causing antithrombin deficiency. *Thromb Haemost.* 2022;122:1369-78.
12. Fernández-Suárez E, González-Del Pozo M, Méndez-Vidal C, Martín-Sánchez M, Mena M, de la Morena-Barrio B, et al. Long-read sequencing improves the genetic diagnosis of retinitis pigmentosa by identifying an Alu retrotransposon insertion in the EYS gene. *Mob DNA.* 2024;15:9.
13. Cuenca-Guardiola J, de la Morena-Barrio B, Corral J, Fernández-Breis JT. Advanced analysis of retrotransposon variation in the human genome with nanopore sequencing using RetroInspector. *Sci Rep.* 2025;15:14489.
14. Cuenca-Guardiola J, Morena-Barrio B, Navarro-Manzano E, Stevens J, Ouwehand WH, Gleadall NS, et al. Detection and annotation of transposable element insertions and deletions on the human genome using nanopore sequencing. *iScience.* 2023;26:108214.
15. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol.* 2021;39:1348-65.
16. Rojo-Carrillo JJ, Garrido-Rodríguez P, Llamas-López M, Cifuentes-Riquelme R, Padilla J, Ramos-Molina B, et al. Landscape of antisense genes in the human genome and identification of new human hepatic antisense RNAs by long-read sequencing. *BMC Genomics.* 2024;25:1148.
17. Monzó C, Liu T, Conesa A. Transcriptomics in the era of long-read sequencing. *Nat Rev Genet.* 2025 Mar 28. doi: 10.1038/s41576-025-00828-z. Online ahead of print.
18. Deshpande AS, Ulahannan N, Pendleton M, Dai X, Ly L, Behr JM, et al. Identifying synergistic high-order 3D chromatin conformations from genome-scale nanopore concatemer sequencing. *Nat Biotechnol.* 2022;40:1488-99.
19. Lu C, Bonini A, Viel JH, Maglia G. Toward single-molecule protein sequencing using nanopores. *Nat Biotechnol.* 2025;43:312-22.