



# SECUENCIACIÓN COMPLETA DEL GENOMA, EN EL CAMINO DE UN DIAGNÓSTICO GENÓMICO INTEGRAL

Para la mayoría de las personas con enfermedades genéticas, poder conocer la causa molecular de su enfermedad supone, lo que ellos denominan, una Odisea Genómica. Un largo peregrinaje, a través de numerosas consultas médicas y análisis y diagnósticos erróneos, lo que implica, importantes retrasos para alcanzar un diagnóstico correcto, manejos clínicos inadecuados y ausencia de asesoramientos familiares y reproductivos específicos.



JAVIER GARCÍA PLANELLS,  
PhD. Director Científico de Health in Code.

*La implementación del Genoma Clínico, como opción de primera línea, nos abre las puertas a una nueva era con grandes expectativas y grandes retos. Una era en la que, una gestión eficiente de los datos generados y del conocimiento disponible, nos permita mitigar las consecuencias de esa larga odisea y obtener diagnósticos más precisos, precoces y eficientes.*

El Genoma Humano es el conjunto de instrucciones que tienen cada una de las células de un individuo para ejercer sus funciones vitales. Unas instrucciones complejas, que constan de más de 3.000 millones de caracteres (A, C, T y G), en los que, el más mínimo cambio o modificación, puede dar lugar a alguna de las 7.000 enfermedades genéticas que conocemos actualmente y cuyo número aumenta cada día. Estas enfermedades genéticas son muy numerosas y heterogéneas, afectando a entre un 3-6% de la población mundial. Muchas de ellas tienen una frecuencia poblacional muy baja, también conocidas como enfermedades raras, y están causadas por mutaciones concretas en determinados genes. Otras, son más frecuentes o comunes y están causadas por mutaciones en diversos genes, o, incluso, es necesaria la combinación de

diferentes circunstancias, genéticas o ambientales, para desencadenar la enfermedad. Por todos estos motivos, el diagnóstico genético es un proceso extremadamente complejo, pudiendo demorarse por encima de los 5 años, teniendo que acudir a múltiples visitas médicas con diversos especialistas, incluso obteniendo diagnósticos incorrectos en un importante porcentaje de los casos.

¡Una verdadera Odisea!

No obstante, si lo analizamos en perspectiva, el diagnóstico genético ha evolucionado de una manera muy rápida, especialmente en los últimos años. Sobre los años 70 del siglo pasado comenzamos a aplicar el cariotipo para detectar el Síndrome de Down. Estas primeras técnicas citogenéticas nos permitían identificar alteraciones numéricas y estructurales de los cromosomas hasta el nivel de resolución que nos permitían los microscopios ópticos. Desde entonces, estas técnicas han evolucionado permitiendo detectar desde grandes alteraciones numéricas, como las monosomías o trisomías, hasta pequeñas alteraciones estructurales (SV: *Structural Variation*) o del número de copias (CNVs: *Copy Number Variation*), bien mediante la optimización de las técnicas citogenéticas (FISH)



o el desarrollo de técnicas moleculares (CMA: *Chromosomal Microarray*), *MLPA*, *Optical Mapping*) dirigidas a la detección de este tipo frecuente de evento mutacional responsable de muchas enfermedades genéticas. También en torno a los años 80, el desarrollo de las técnicas de biología molecular (PCR, clonación, secuenciación,...) nos permitió analizar otras causas frecuentes de mutación, pequeñas alteraciones que afectaban a la secuencia o estructura de determinados genes. De esta manera podíamos diagnosticar aquellas enfermedades monogénicas, de las cuales disponíamos de suficiente conocimiento clínico y científico como para poder orientar clínicamente el análisis genético, y aplicar la técnica específica asociada a la mutación o mutaciones mayoritarias de un determinado gen.

Desde entonces, el diagnóstico genético ha ido mejorando de manera proporcional al conocimiento genético y a los grandes avances de las distintas técnicas de análisis genético orientadas al análisis preciso de los diferentes tipos de mutaciones genéticas conocidas. No obstante, sigue siendo un largo proceso que implica una evaluación clínica y científica profunda de cada caso y una aplicación secuencial de las técnicas de análisis genético en función de la sospecha clínica, el tipo de mutación o mutaciones más frecuentes y el acceso a las técnicas disponibles en el entorno sanitario cercano al paciente. Aun así, muchos pacientes con un diagnóstico clínico preciso quedan sin diagnóstico debido a limitaciones técnicas o de conocimiento científico.

En el 2003, el proyecto Genoma Humano supuso un hito importantísimo tanto para la Genómica, como para la Medicina. Por un lado, pudimos disponer de un primer mapa genómico de referencia que permitió incrementar significativamente el descubrimiento de nuevos genes y sus asociaciones con las enfermedades genéticas más frecuentes. Pero, además, conseguimos un impresionante desarrollo tecnológico gracias a la secuenciación masiva o *Next Generation*

*Sequencing* (NGS). Esta técnica supuso un cambio importante de paradigma, facilitando el análisis masivo de datos genómicos con una reducción progresiva de los costes de secuenciación.

La NGS nos proporciona grandes cantidades de datos procedentes de diferentes genes o regiones genómicas mediante un único ensayo. Esto nos ofrece una gran capacidad de análisis y descubrimiento de nuevas variantes, genes y asociaciones clínicas, incrementando exponencialmente el conocimiento científico y la capacidad potencial de diagnóstico genético derivada de ello.

Para poder usar esta potente tecnología en el contexto del diagnóstico genético debemos garantizar una alta sensibilidad, o capacidad de detectar aquello que queremos identificar; una alta especificidad, o la capacidad de no generar incertidumbre evitando detectar aquello que no queremos identificar o no sabemos interpretar; y su utilidad clínica. Por este motivo, las primeras aplicaciones clínicas de la NGS se orientaron hacia paneles de genes o regiones que conocemos que están asociadas a la indicación clínica del estudio, de manera que, acotamos los datos genómicos al conocimiento científico del momento para incrementar su sensibilidad, especificidad y utilidad clínica y, además, limitamos los costes de secuenciación y de almacenamiento de datos. Estas estrategias rápidamente se convirtieron en las técnicas de elección por su coste/eficacia ante los trastornos más heterogéneos y complejos que no habían podido resolverse mediante las técnicas convencionales orientadas a mutaciones, genes o enfermedades monogénicas más frecuentes. Con la reducción de los costes de secuenciación y almacenamiento y el incremento del conocimiento científico, el tamaño de los paneles de genes también creció hasta llegar a la secuenciación completa del exoma (WES: *Whole Exome Sequencing*).

Actualmente, el análisis del exoma, o WES, es la técnica más utilizada, salvo en aquellos casos en los que el diagnóstico



clínico es muy claro y concreto y disponemos de una técnica específica que proporciona un rendimiento diagnóstico mayor al menor coste. El exoma humano es el conjunto de “exones” o regiones codificantes de todos los genes conocidos (unos 19.000), lo que supone aproximadamente el 1,5% del genoma total. Desde el punto de vista clínico, reducimos significativamente la cantidad de datos analizados a aquellas regiones que sabemos que tienen una mayor probabilidad de encontrar una mutación. Por tanto, desde el punto de vista de coste/eficacia, podemos detectar el 80% de las variantes patogénicas que conocemos actualmente analizando, tan solo, el 1,5% de toda la información genómica. En estos casos en los que se utiliza el exoma como técnica de primera línea el rendimiento diagnóstico oscila entre el 25 y el 50% en función de la indicación clínica y la referencia bibliográfica.

A pesar de la gran potencia de análisis del WES, presenta algunas limitaciones que requieren el uso complementario de otras técnicas tradicionales. Por una parte, existen algunos tipos de mutaciones que, debido a limitaciones propias de la técnica, no pueden detectarse, como es el caso de la mayoría de mutaciones dinámicas y otras mutaciones para las que la precisión del WES es muy baja para emitir un informe diagnóstico con garantías o notablemente inferior a la de la técnica convencional, como es el caso de mutaciones en pseudogenes o genes parálogos, las alteraciones estructurales o incluso las CNVs, especialmente las de pequeño tamaño. Pero, sobre todo, el exoma no permite detectar mutaciones en aquellas regiones no incluidas en el diseño, como por ejemplo, todas las regiones no codificantes que suponen más del 98% del Genoma, incluyendo mutaciones conocidas en regiones intrónicas o regiones reguladoras que son cada vez más frecuentes conforme avanza el conocimiento científico.

Actualmente, la secuenciación completa del Genoma, o WGS, ya es una alternativa diagnóstica real y posiblemente esa prueba genética global que permita analizar cualquier tipo

de mutación genética en cualquier lugar del genoma.

Entre sus principales ventajas debemos destacar que es una única técnica que permite detectar la gran mayoría de las mutaciones conocidas con una alta precisión: variaciones en la secuencia de ADN (SNVs), alteraciones del número de copias (CNVs), variaciones estructurales (SV), mutaciones dinámicas, mutaciones en pseudogenes y genes parálogos, mutaciones en el genoma mitocondrial,.....; tanto en regiones codificantes como no codificantes y, especialmente, proporciona toda la información genética de un individuo, permitiendo su reanálisis conforme el conocimiento científico avanza. En esta información podemos incluir, no solo aquellas variantes genéticas causales de enfermedades genéticas de alta penetrancia y expresividad, sino también, variantes asociadas al metabolismo de determinados fármacos o compuestos que nos permitan predecir su eficacia o toxicidad, o variantes que nos permitan establecer riesgos poligénicos asociados a enfermedades más comunes como el cáncer o las enfermedades cardiovasculares antes de que aparezcan y podamos actuar frente a ellas.

Entre los principales retos y limitaciones del WGS nos encontramos con el exceso de datos generado. Por una parte, sus implicaciones directas de costes de secuenciación y almacenamiento, y que probablemente evolucionen de manera similar a lo ocurrido durante el tránsito de paneles a exoma, descrito anteriormente. Por otra parte, el exceso de información genética puede llevarnos a una incertidumbre diagnóstica por falta de conocimiento específico sobre determinadas variantes genéticas o regiones. Además, el exceso de información generalmente conlleva importantes repercusiones éticas y asistenciales sobre qué analizar y qué informar, en el contexto de los derechos de nuestros pacientes a saber y no saber sobre su información genética. Durante estos últimos años, hemos podido aprender a gestionar grandes volúmenes de datos, variantes y conocimiento, que tendremos que reaprender a dimensionar con el Genoma. Hemos establecido criterios para la clasificación de variantes en base a la evidencia disponible en cada momento y a sus potenciales implicaciones funcionales, gestionado y actualizado los datos y el conocimiento y, sobre todo, hemos podido informar a nuestros pacientes a través de un asesoramiento genético especializado que les permita consentir y tomar sus decisiones sobre qué información recibir en función de sus circunstancias.

El análisis del genoma, como test genómico integral, no solo nos va a proporcionar un incremento considerable del rendimiento diagnóstico a corto y medio plazo, sino que además, nos acerca a una Medicina Personalizada de Precisión en la cual, debemos ser capaces y responsables de gestionar eficientemente unos datos genómicos y una información que resulte útil a pacientes e individuos, en función de unas circunstancias en constante evolución y repleta de retos técnicos, clínicos, científicos y sociales.