

# DIAGNÓSTICO GENÉTICO DE ENFERMEDADES RARAS OCULARES

**Dr. D. José María Millán Salvador**

Investigador del Instituto de Investigación Sanitaria La Fe (Valencia)

Las distrofias de retina (DR) son un conjunto heterogéneo de enfermedades hereditarias y degenerativas de la retina que afectan a los fotorreceptores y causan pérdida de la visión. Este conjunto de enfermedades raras tiene distintas causas genéticas y en el momento presente los genes conocidos implicados en su etiología son más de 250 (RetNet, <https://sph.uth.edu/retnet/sum-dis.htm>) que explican aproximadamente la mitad de los casos y se estima que deben existir al menos otros tantos genes responsables de los casos restantes. Las terapias en desarrollo para esta patología son a menudo dependientes del gen y mutación causantes.

Por todo ello la caracterización molecular de las DR suponen un desafío por su complejidad pero una necesidad para poder abordar su prevención, a través del consejo genético, su diagnóstico precoz y en un futuro su tratamiento.

El desarrollo de la secuenciación de nueva generación (Next Generation Sequencing, NGS por sus siglas en inglés) ha supuesto un avance tremendo en el campo de la investigación y el diagnóstico genético molecular, particularmente en las enfermedades raras con elevada heterogeneidad genética y en las que no se conocen aun todos los genes causantes como ocurre con las DR

La NGS permite la “lectura” de un gran número de genes para un paciente determinado en un tiempo y a un coste económico y técnico impensable el siglo pasado.

Desde su aparición, inicialmente en el campo de la investigación, hacia el año 2006 y a partir de 2010 en el campo del diagnóstico, hasta la actualidad, la NGS ha ido mejorando su rendimiento a la vez que disminuía su coste. Como ejemplo, el primer genoma humano que se secuenció completamente en el año 2003 costó unos 1000 millones de

dólares y se tardó 13 años en secuenciarlo y analizarlo mientras que el objetivo de las empresas biotecnológicas para el futuro próximo es conseguir la secuenciación de un genoma humano por 1000 dólares en varios días de trabajo (\$1000 project).

## **Tipos de Secuenciación Masiva**

El genoma completo de un individuo está compuesto por regiones que codifican para proteínas (exones) y otras (intrones y ADN intergénico). La cantidad total de ADN presente en el genoma humano es de unos 3000 millones de nucleótidos y entre 20.000 y 25.000 genes. Aunque el conjunto de genes codificantes (exoma o conjunto de exones) del individuo solo supone el 1,5% del genoma, en él se localizan el 85% de las mutaciones patológicas causantes de enfermedades raras.

Es conveniente distinguir entre las técnicas de secuenciación masiva no enfocadas, esto es la secuenciación de todo el genoma (WGS) o el exoma (WES), de aquellas enfocadas a una enfermedad, mediante paneles con un alto nº de genes conocidos.

Es importante tener en consideración estos aspectos porque mientras la secuenciación dirigida se está extendiendo cada vez más, en el ámbito del diagnóstico y sus resultados son de más fácil manejo desde el punto de vista de su análisis, interpretación y aspectos éticos, en el caso de la WES/WGS la cantidad de datos a analizar, interpretar e informar a los pacientes es ingente y plantea problemas de coste, técnicos, de conocimiento y ético-legales, por lo que su uso está por ahora bastante restringido al campo de la investigación aunque es previsible su incorporación a la clínica en un futuro inmediato, en la medida que estos aspectos se vayan resolviendo.

## **Aspectos Técnicos**

En la secuenciación masiva debemos distinguir 4 fases o aspectos. (figura 1).

La primera de ellas es la captura y preparación de las librerías y la técnica de amplificación del DNA que se hace utilizando diseños específicos realizados de modo adecuado a la enfermedad en estudio y que es preciso actualizar constantemente, y que usa una parte analítica

en el laboratorio (wet lab) desarrollada por distintas compañías. Aunque difieren en varios aspectos, el esquema principal de trabajo es conceptualmente similar para todos ellos. El ADN se fragmenta (la fragmentación puede hacerse por diversos métodos), posteriormente se le añaden unas secuencias adaptadoras a los extremos de estos fragmentos y se hibridan a sondas representativas de todo el genoma o regiones del genoma que queremos secuenciar. Este proceso se conoce como captura o realización de la librería. A continuación, esos fragmentos de ADN “capturados” se amplifican clonalmente para ser utilizados como moldes a secuenciar (Enriquecimiento).



Figura 1: Las cuatro fases de la secuenciación por NGS se representan cada una en un color diferente.

En segundo lugar, estas secuencias de ADN capturadas y amplificadas

se deben secuenciar en un equipo adecuado (secuenciador).

Existen distintos métodos de secuenciación (pirosecuenciación, hibridación-ligación o el registro de cambios en el pH durante la incorporación de nucleótidos). En cualquier caso, de esta forma se consiguen una serie de secuencias de pequeño tamaño (entre 75 y 400 pares de bases) que una vez filtradas, ensambladas y comparadas con una secuencia consenso nos darán la lectura de la región genómica que queríamos analizar.

Dependiendo del estudio que se quiera realizar, la potencia y complejidad de manejo de estas plataformas hace inviable el estudio. Así, el análisis de un genoma o un exoma el estudio de unos pocos genes

La aparición de los secuenciadores y plataformas de arrays de 2ª generación permiten secuenciar simultáneamente más de cientos de millones de lecturas cortas (de entre 30 a 70 bases hasta varios cientos de bases, más recientemente) que se mapean frente a un genoma de referencia a una cobertura muy redundante, esto es varios cientos/miles de veces.

Esta estrategia está permitiendo no solo secuenciar el ADN de un genoma o exoma completo sino hacer otros estudios como secuenciación del ARN (transcriptoma), estudios epigenómicos etc (RNA-Seq, ChIP-Seq, bisulfite sequencing)

Además, recientemente han aparecido versiones de equipos de NGS con características más limitadas en cuanto a su rendimiento y capacidad de secuenciación pero con mayor facilidad de manejo y enfocados a un segmento de mercado distinto. Algunos ejemplos son los equipos 454 GS Junior de Roche, MiSeq Personal Sequencer de Illumina e Ion Torrent de Life Technologies. Estas plataformas se conocen como benchtop o de bancada y son las utilizadas comúnmente para el estudio de determinados paneles de genes enfocados al diagnóstico de una enfermedad o grupo de enfermedades.

Recientemente, han aparecido otras tecnologías que están basadas en la emisión de protones y que pretenden conseguir la secuenciación de una cadena de ADN en tiempo real lo que reduciría aún más el coste y el tiempo de procesado. Sin embargo, estas tecnologías conocidas como Next-Next Generation Sequencing todavía no superan a las

mencionadas anteriormente.

En tercer lugar, se debe realizar el procesamiento bioinformático de las secuencias. Este paso es muy importante y supone en estos momentos el cuello de botella de esta tecnología, particularmente en el caso de la WGS/WES.

Una parte muy importante en el esquema de trabajo de un experimento de NGS es el análisis computacional (figuras. 2)2A y B). Las ciencias informáticas han tomado una relevancia crítica en la NGS en el sentido de que sus capacidades son esenciales para manejar y analizar datos biológicos. La NGS produce una cantidad de datos sin precedentes que un ordenador común no puede manejar. Aunque para algunas plataformas existen herramientas de manejo de datos y análisis en un único programa, cualquier tarea no trivial a realizar con los datos requerirá al menos de una persona con conocimientos en bioinformática. En el futuro las compañías de software y los proveedores de equipos de NGS desarrollarán programas con los que no será imprescindible tener conocimientos en bioinformática para analizar datos de secuenciación masiva, aunque este hecho podría limitar al usuario a solo aprovechar las funciones predefinidas en ese hipotético software.

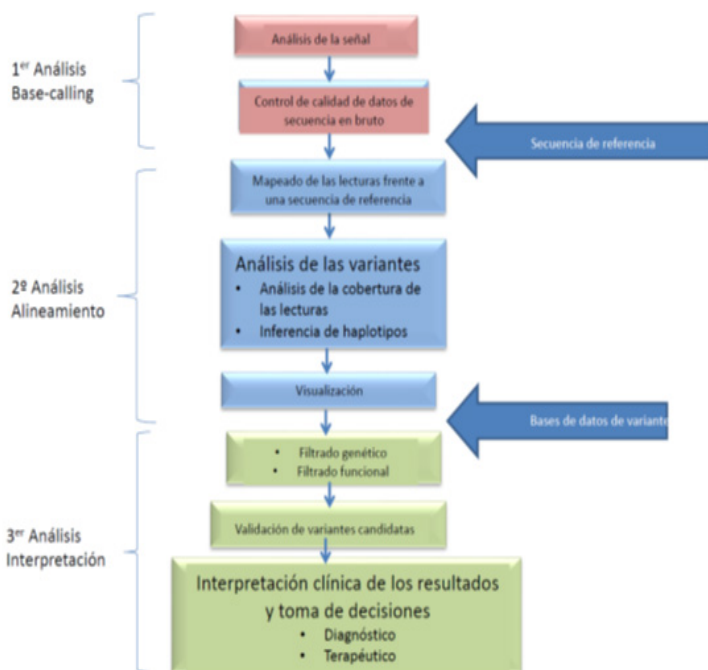


Figura 2:

Y por último se deben comprobar las mutaciones encontradas,

validando los resultados con otras técnicas, así como interpretar su papel patogénico mediante otras pruebas (estudio en la familia, búsqueda en las bases de datos, análisis funcionales in silico y en vivo), proporcionando un análisis de los datos experimentales que tenga sentido biológico y que pueda informarse adecuadamente al paciente y sus familiares.

Ello requiere de una gran experiencia en genómica y en la propia enfermedad por parte del equipo profesional que lleve a cabo el estudio y depende de la incorporación de toda la información relevante que exista.

Por tanto, nos parece evidente que, a día de hoy, para analizar los datos experimentales obtenidos e integrarlos con la información disponible en las numerosas bases de datos de información biológica, los conocimientos en ciencias computacionales son absolutamente necesarios.

## **Ventajas e inconvenientes de los distintos tipos de NGS**

Las técnicas de NGS pueden utilizarse tanto en el campo del diagnóstico (principalmente paneles) como de la investigación de las DR (paneles y WGS/WES).

En el campo del diagnóstico la NGS por paneles permitirá estudiar todos los genes conocidos relacionados con las distrofias retinianas (en este caso) en un tiempo y con una relación coste/efectividad muy elevados, independientemente del diagnóstico clínico lo que supone una ventaja en aquellos casos en los que el diagnóstico clínico no esté bien definido o sea solapante entre distintas entidades clínicas.

En el caso de la investigación, la WGS/WES permitirá encontrar nuevos genes relacionados con la enfermedad en aquellos pacientes (alrededor del 50% actualmente) en los que no se encuentra la mutación o mutaciones responsables en los genes conocidos.

Sin embargo, la NGS no es la solución definitiva para todos los casos de DR. Existen todavía problemas técnicos para la detección de cierto tipo de defectos genéticos como deleciones inversiones o traslocaciones. La captura de regiones genómicas ricas en GC o en elementos repetitivos y, sobre todo, la detección de mutaciones en las regiones intrónicas, que se sabe actualmente son una causa no infrecuente de DR.

Además la NGS actualmente debe ser validada por la tecnología tradicional (secuenciación por Sanger), supone un esfuerzo de interpretación de las variantes encontradas por parte de expertos experimentados en el campo clínico en cuestión y sigue sin permitir identificar mutaciones de efecto patológico en regiones no codificantes cuya patogenicidad solo puede ser dilucidada mediante estudios funcionales difíciles si los genes implicados únicamente se expresan en tejidos inaccesibles como la retina.

Por último, en el caso de la WES/WGS se pueden encontrar variantes con significado fenotípico que pueden afectar a la vida futura del paciente y sus familiares y que ocasionan situaciones éticas y legales que deben ser valoradas por expertos en el campo de la bioética.

La tabla 1 refleja las diferencias entre la NGS enfocada (paneles) y los estudios ómicos (WGS/WES) en cuanto a tiempo, costes y disponibilidad del estudio enfocado y % de casos que es esperable caracterizar

	<b>Estudio Enfocado (paneles)</b>	<b>Estudio ómico (WES/WGS)</b>
<b>Coste</b>	bajo	elevado
<b>Estudio</b>	individual	familiar
<b>Dependencia técnica</b>	Experiencia en el campo	Experiencia en el campo y bioinformática
<b>Tipo de Resultados</b>	No hay resultados inesperados*	Resultados ajenos al objetivo clínico del estudio
<b>Rutina diagnóstica</b>	Sí	No
<b>Identificación de nuevos genes</b>	No	Sí
<b>Identificación de nuevos mecanismos de enfermedad</b>	No	Sí

\*En ocasiones sí se pueden producir resultados inesperados o no deseados

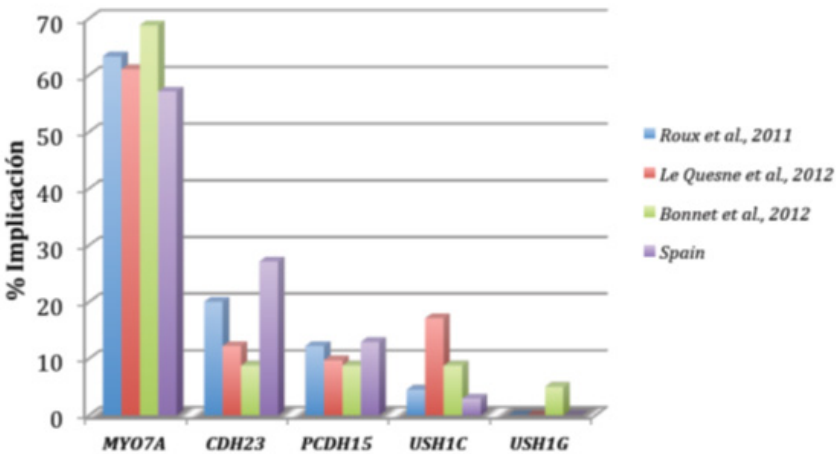
Tabla 1: Comparación entre NGS (paneles) y WES/WGS: Indicaciones, ventajas e inconvenientes

# Experiencia en el laboratorio del Hospital la Fe de Valencia

En nuestro laboratorio se trabaja tanto con distrofias de retina aisladas como con síndrome de Usher, la distrofia de retina sindrómica más frecuente y que asocia retinosis pigmentaria, hipoacusia neurosensorial y, en ocasiones, alteración del equilibrio.

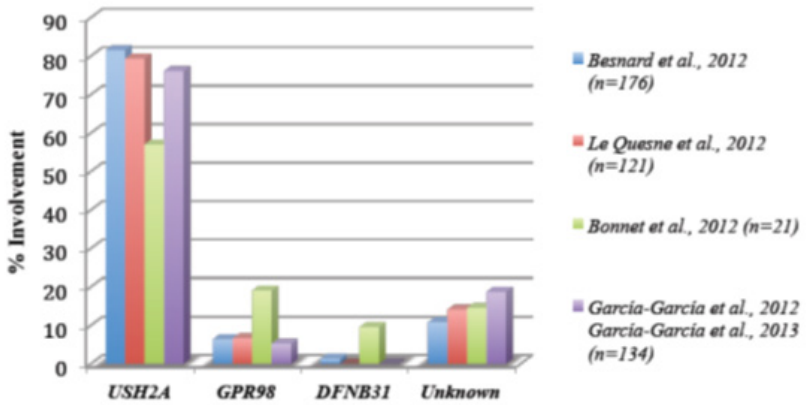
Mediante un panel de NGS que incluye todos los genes conocidos relacionados con el síndrome de Usher conseguimos una eficacia diagnóstica del casi el 100% de los casos de síndrome de Usher tipo 1 (la forma más grave de la enfermedad) y de más del 80% de los casos de síndrome de Usher tipo 2. En los casos en los que el diagnóstico clínico no está claro, la eficacia desciende hasta un 40% (figura 3). Además, nos ha permitido conocer qué genes son los más frecuentes en nuestra población.

## USH1 (97,5%)





## USH2 (82%)



## USH NO CLASIFICADO:

identificada la causa genética de la enfermedad en el 40% (12 de 30) de las familias

**USH2A: 8 (26,7%)**

**PCDH15: 2 (6,7%)**

**CDH23: 1 (3,3%)**

**MYO7A: 1 (3,3%)**

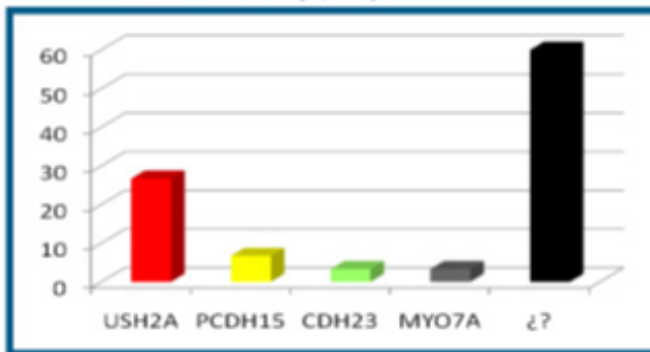
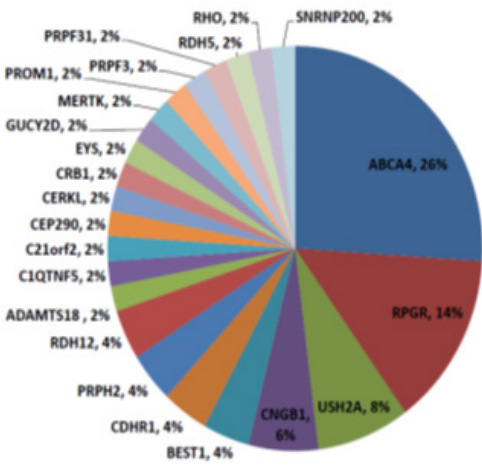
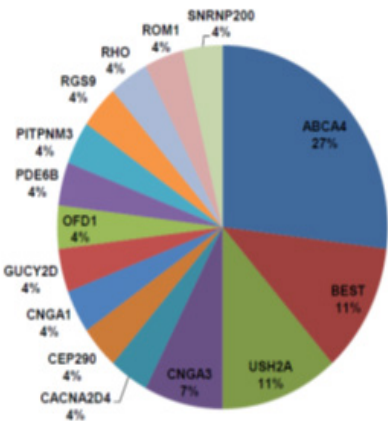


Figura 3

Respecto a las distrofias de retina no sindrómica, aplicando un panel de 117 genes, hemos conseguido el diagnóstico de el 50% de los casos mientras que en 26 de ellos no hemos podido completar el diagnóstico al ser genes implicados en distrofias autosómicas recesivas y no haber encontrado la segunda mutación. En otros 23 pacientes no se encontró ninguna mutación por lo que se están secuenciando mediante WES para intentar identificar un nuevo gen implicado en la enfermedad (figura 4).



- > AR: 68%, AD:18%, L-X: 14%
- > 21/75 mutaciones novel



- > ¿Mutaciones intrónicas profundas?
- > ¿CNV?



- > Secuenciación del gen / RNA-Seq / CGH-array

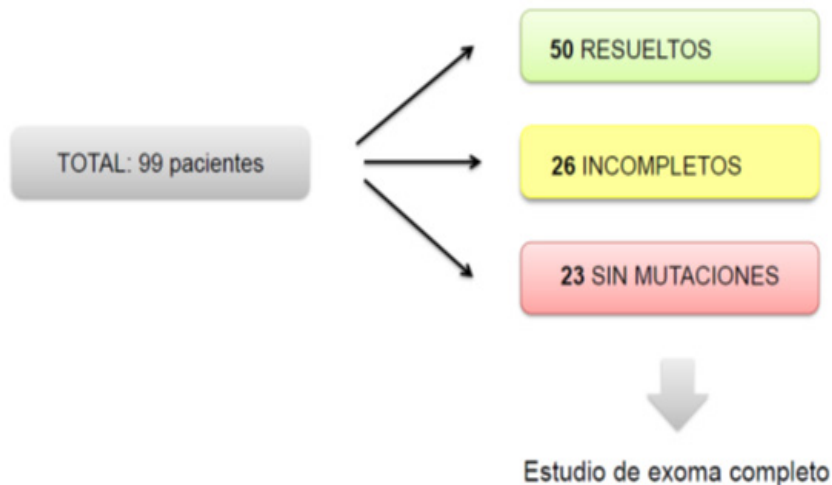


Figura 4:

## CONCLUSIONES

El diagnóstico molecular de las distrofias hereditarias de la retina se ha visto históricamente dificultado por su enorme heterogeneidad clínica y genética.

En la mayoría de los casos, ante un diagnóstico oftalmológico, el genetista se enfrenta a un elevado nº de genes candidatos que analizar. (Tabla 2).

Estudio	Genes en el panel	% de casos diagnosticados	Distrofia de Retina Estudiada
<i>Bowne et al (2011)</i>	46	24%	ADRP
<i>Shanks et al (2013)</i>	73	25% // 53%	DR en general// de inicio temprano
<i>Neveling et al (2012)</i>	111	36%	RP
<i>Xia Wang et al 2013)</i>	163	39%	LCA
<i>Audo et al; (2012)</i>	254	57%	DR en general

ADRP: Retinosis pigmentaria autosómica dominante; DR: Distrofias retinianas; LCA: Amaurosis congénita de Leber

Tabla 2: Ejemplos de la eficacia de la NGS en estudios realizados recientemente mediante NGS enfocada

Se recomienda en el momento presente utilizar la NGS (incluyendo aquí el exoma clínico) como técnica diagnóstica de 1er nivel. Ello permite caracterizar molecularmente aproximadamente la mitad de los casos.

En el resto de los casos o como un 2º/3er tercer paso, y dependiendo de la estructura familiar, tipo de patología y disponibilidad /precio de la técnica y de su análisis bioinformático se recomendaría realizar la técnica de WGS/WES.

La NGS es una herramienta eficaz en este tipo de enfermedades genéticamente heterogéneas ya que permite el estudio de un elevado número de genes al mismo tiempo reduciendo el tiempo del análisis y el coste técnico y económico.

NOTA: Esta ponencia está basada parcialmente en el artículo publicado en la revista VISIÓN (FARPE). Ayuso C, Millán JM. Secuenciación masiva en distrofias retinianas. Visión. 43: 17-22 (2013)