

Nuevos avances en la identificación de pacientes con riesgo de muerte súbita

Nuevas herramientas diagnósticas en la genética de la muerte súbita

Catarina Allegue^a, Oscar Campuzano^a, Sergio Castillo^b, Mònica Coll^a, Anna Iglesias^a y Ramon Brugada^{a,*}

^aCentro de Genética Cardiovascular, UdG-IdibGi, Instituto de Investigación Biomédica de Girona, Girona, España

^bGendiag SL, Barcelona, España

Palabras clave:

Muerte súbita
Nuevas tecnologías
Diagnóstico genético
Canalopatías
Miocardiopatías

RESUMEN

El estudio a escala genómica de las enfermedades cardiovasculares permite entenderlas mejor para optimizar y dirigir terapias personalizadas. En el desarrollo de la muerte súbita cardiaca destacan dos grandes grupos de enfermedades de gran heterogeneidad genética y fenotípica: las enfermedades estructurales o miocardiopatías y las arritmogénicas o canalopatías. El principal problema de ambos tipos de enfermedad es que a menudo causan la muerte súbita en individuos previamente asintomáticos en los que la muerte súbita cardiaca es la primera manifestación de cardiopatía en un alto número de casos. Por consiguiente, y dado que se trata de enfermedades hereditarias, hay un elevado riesgo para los familiares que, pese permanecer asintomáticos, podrían ser portadores de variantes genéticas de riesgo. Se han descrito alrededor de 100 genes implicados en las enfermedades asociadas a muerte súbita cardiaca. La capacidad de la secuenciación convencional de genes Sanger es limitada con respecto a las nuevas opciones tecnológicas en constante desarrollo, tales como los *arrays* de resecuenciación y especialmente la secuenciación masiva en paralelo, *next generation sequencing*. Gracias a la mejora de las diferentes químicas en las distintas opciones tecnológicas, los proveedores centran sus esfuerzos en el aumento de la capacidad de generación de datos de estos equipos, así como en la rebaja de los costes de los reactivos necesarios para estos análisis, con objeto de facilitar a la comunidad científica el acceso a estas tecnologías. Dadas la gran cantidad y la complejidad de los datos genéticos derivados de la ultrasecuenciación, que requieren un análisis minucioso de sus implicaciones médicas en todas las ramas de la medicina, es necesaria la creación de centros especializados en el estudio de enfermedades concretas, en el manejo de datos genéticos a gran escala y en prestar asesoramiento genético a las familias.

New Genetic Diagnostic Tests for Sudden Cardiac Death

ABSTRACT

The study of cardiovascular disease at the genetic level will provide the greater understanding needed for the design and optimization of personalized treatment. The development of sudden cardiac death primarily involves two large groups of disorders that have considerable genetic and phenotypic heterogeneity: the structural disorders or cardiomyopathies and the arrhythmogenic disorders or channelopathy. The principle challenge with both types of disorder is that, in a substantial number of cases, they cause sudden death in previously asymptomatic individuals in whom sudden cardiac death is the first manifestation of cardiac disease. Consequently and given the hereditary nature of the condition, families with asymptomatic family members will be at risk as they could be carriers of a high-risk genetic variant. Around 100 genes have been implicated in conditions associated with sudden cardiac death. The capacity of conventional gene sequencing techniques based on the Sanger method is limited compared with that of the new approaches constantly being developed, such as resequencing arrays and, especially, massively parallel sequencing or next-generation sequencing. Thanks to improvements in the chemicals used in the various sequencing techniques, suppliers have now concentrated their efforts on increasing the data-generating capacity of the technology as well as on reducing the cost of the reagents needed for assays, with the aim of making it easier for the scientific community to access these new techniques. Given the quantity and complexity of the genetic data produced by ultrasequencing, which necessitates meticulous analysis to identify the clinical implications for each branch of medicine, it will be necessary to create specialized centers that can carry out research into specific diseases, manage the large-scale genetic data produced, and provide genetic counseling to families.

Keywords:

Sudden death
New technology
Genetic diagnosis
Channelopathy
Cardiomyopathy

*Autor para correspondencia: Centre de Genètica Cardiovascular, Institut d'Investigació Biomèdica Girona-IdIBGi, Facultat de Medicina, Universitat de Girona, Girona, España.
Correo electrónico: rbrugada@idibgi.org (R. Brugada).

Abreviaturas

MCD: miocardiopatía dilatada
 MCH: miocardiopatía hipertrófica
 MSC: muerte súbita cardiaca
 NGS: *next generation sequencing*
 SQTC: síndrome de QT corto
 SQTL: síndrome de QT largo

MUERTE SÚBITA

La muerte súbita (MS) se define como un fallecimiento natural e inesperado que tiene lugar durante la primera hora tras el comienzo de los síntomas. En defunciones no presenciadas, se considera MS la que ocurre 24 h tras haberse visto por última vez a la víctima en buen estado de salud. La mayor parte de las MS son de origen cardiaco (MSC). La MSC es una de las principales causas de muerte en los países occidentales, en torno a 800.000 muertes anuales. Aunque su incidencia varía según los distintos estudios y franjas de edad, se estima que la incidencia de MSC se sitúa alrededor de 30-200/100.000 habitantes/año, de modo que representa un problema de salud mayor¹. En España se producen al año unas 9.000 MSC de pacientes de entre 25 y 74 años, aunque en pocas ocasiones se diagnostican como tales en el boletín estadístico de defunción². Entre los menores de 31 años, la incidencia de MSC se ha estimado en 1,2/100.000 habitantes en una serie del País Vasco³. Estas muertes están especialmente vinculadas a la práctica de deporte, a la que se considera uno de los mayores factores de riesgo junto con la historia familiar de MS. La dificultad del cálculo de la incidencia radica en la propia definición de la MSC. En muchos casos los informes médicos no están disponibles, no se realiza autopsia o la causa de muerte que se da en el certificado de defunción es poco específica^{4,5}.

En un 80% de los casos, la MSC es consecuencia de enfermedad coronaria (infarto). Pese a que en la población joven (< 35 años) la frecuencia de MS de causa isquémica (aterosclerosis) está aumentando progresivamente, se considera que en esta población la mayor parte de las muertes son de causa no isquémica. En este grupo de individuos jóvenes destacan dos grandes tipos de enfermedades asociadas a MSC: las estructurales (miocardiopatías), que se pueden identificar en la autopsia, y las canalopatías, que cursan con autopsias negativas en las que la causa de la muerte no se puede determinar. El principal problema de ambos tipos de enfermedad es que a menudo causan MS en individuos aparentemente sanos. Desgraciadamente, la MSC puede ser la primera manifestación de la enfermedad cardiaca. Tanto las miocardiopatías (miocardiopatía hipertrófica [MCH], miocardiopatía dilatada [MCD], miocardiopatía arritmogénica del ventrículo derecho [MAVD], entre otras) como las canalopatías (síndrome de Brugada [SBr], síndrome de QT largo [SQTL], síndrome de QT corto [SQTC], taquicardia ventricular polimórfica catecolaminérgica [TVPC], entre otras) pueden ser de etiología hereditaria y, por lo tanto, tener una base genética. En este sentido, merece la pena destacar que se ha descrito un porcentaje importante de casos de MSC en los que el análisis genético ha puesto de manifiesto la presencia de dobles o incluso triples mutaciones, localizadas en los mismos genes o en genes diferentes y asociados a fenotipos distintos⁶⁻⁸. A día de hoy se conocen alrededor de 100 genes implicados en las enfermedades asociadas a MSC. Pese al enorme avance en la genética de estas enfermedades, en un porcentaje elevado de pacientes no se logra un diagnóstico genético. Dado el gran número de genes implicados y que la lista de genes está en constante crecimiento, el análisis genético de un número elevado de pacientes resulta una tarea costosa. En esta dirección se publicó un documento internacional donde se recogen recomenda-

ciones sobre los genes que se debe analizar en casos concretos⁹. Aunque este documento recomienda un estudio genético en gran parte de los casos asociados a MSC, ciertas familias no serían analizadas. Por consiguiente, y dado que se trata de enfermedades hereditarias, hay elevado riesgo para los familiares que, pese permanecer asintomáticos, podrían ser portadores de variantes genéticas de riesgo. Lamentablemente, en muchas ocasiones la primera manifestación de la enfermedad es la MSC. En los casos en que la autopsia no es concluyente a la hora de esclarecer la causa de la muerte, investigar la MS inexplicada resulta esencial para la evaluación del riesgo y la prevención de otro episodio de MSC en familiares portadores.

EVOLUCIÓN DE LAS TECNOLOGÍAS DE SECUENCIACIÓN GÉNICA Y SU APLICACIÓN AL ESTUDIO DE LA MSC

Tecnología Sanger

La aparición de la tecnología Sanger de secuenciación del ADN supuso un extraordinario avance en el estudio de las enfermedades genéticas¹⁰. Hasta ese momento se utilizaban tecnologías basadas en el uso de enzimas de restricción (RFLP).

La secuenciación de Sanger («secuenciación con terminadores de cadena») evolucionó rápidamente para pasar a ser durante las décadas de los ochenta y los noventa la tecnología de referencia para el diagnóstico de alteraciones génicas en el genoma humano. En pocos años se avanzó desde la amplificación de secuencias de tamaño corto (100 pb) a la generación de secuencias lineales de hasta 800 pb. Asimismo, el avance en el desarrollo de nuevos equipamientos de secuenciación permitió reducir progresivamente tanto el tiempo de amplificación como los costes de amplificación por base secuenciada. Actualmente el coste de reactivos por cada mil bases secuenciadas (1 kb) se ha establecido alrededor de unos 0,4 dólares (<http://www.genome.gov/sequencingcosts>).

La secuenciación de Sanger, si bien se caracteriza por una elevada precisión (sensibilidad del 99,9% con un *Phred score* de 30) y una longitud de carrera del orden de 80 kb, tiene no obstante una capacidad de generación de datos reducida en comparación con otras opciones tecnológicas de estudio del ADN que se han desarrollado en los últimos años (*arrays* de resecuenciación y, especialmente, las tecnologías de secuenciación masiva en paralelo, *next generation sequencing* [NGS]). Merece la pena destacar que la generación del primer borrador del genoma humano se llevó a cabo mediante secuenciación convencional de Sanger (ABI 3700 con 96 capilares en Proyecto Genoma Humano). Dicho proyecto se completó en 13 años con un coste aproximado de unos 2.700 millones de dólares^{11,12}.

La secuenciación Sanger sigue siendo la tecnología de referencia para el análisis de regiones de extensión discreta del genoma (enfermedades monogénicas o causadas por un número pequeño de genes), así como para la validación de la presencia de alteraciones (variaciones puntuales e inserciones y deleciones) derivadas de otras tecnologías. Sin embargo, como se ha mencionado, su principal limitación es la relativamente baja capacidad de generación de datos, lo que da lugar a que su aplicación al estudio de una extensión grande del genoma sea un proceso lento y costoso.

Se ha publicado gran número de estudios basados en la aplicación de la secuenciación de Sanger como un mecanismo eficaz para la detección de numerosas mutaciones asociadas a MSC. La mayoría de los fenotipos principales como SBr, SQTL, MCH o MCD se han estudiado en menor o mayor medida por esta tecnología¹³⁻¹⁵. Estas publicaciones, entre otras muchas, han asentado el conocimiento de sus bases moleculares.

ARRAYS DE GENOTIPIFICACIÓN

El conocimiento previo obtenido mediante tecnología Sanger ha permitido el desarrollo de otras tecnologías de análisis de ADN como

los chips de genotipificación (Affymetrix, Illumina, Agilent) o los tests de detección múltiple de variaciones conocidas (Sequenom, Taqman, Open Array de LifeTech, Fluidigm) en estudios y programas de cribado rápido^{16,17}.

A principios del presente siglo, la compañía Affymetrix desarrolló la tecnología de resecuenciación en fase sólida, que se basa en la capacidad de la doble hélice de ADN para, en presencia de secuencias complementarias, unirse específicamente mediante hibridación; es capaz de analizar en un único test hasta 300 kb del genoma en un soporte sólido (*array*)^{18,19}.

La utilización de este tipo de herramienta supone un avance importante en el análisis del ADN, pues permite el estudio simultáneo en un solo test de varias decenas de genes (enfermedades poligénicas) o incluso cientos de regiones del ADN (paneles de cribado mutacional múltiple); de esta manera, no pocos investigadores en los últimos años han diseñado sus propios *arrays* para el estudio y el diagnóstico de la presencia de mutaciones en paneles de genes asociados a enfermedades diversas²⁰⁻²⁵.

En el ámbito del estudio de las causas genéticas de la MSC, algunos centros o investigadores desarrollaron sus propios diseños para utilizar esta tecnología en el estudio simultáneo de todos los genes asociados a un mismo fenotipo. En la mayoría de los casos, estos proyectos se centraron en el estudio de los genes asociados a MCD o MCH²⁶⁻²⁸. Estos estudios confirman la utilidad de esta tecnología como mecanismo rápido y económico de detección de mutaciones y recomiendan combinarla con la secuenciación tradicional para obtener un estudio completo y detallado de los genes candidatos. No obstante, coincidían en la necesidad de mejora y desarrollo de los algoritmos de análisis de estos datos con objeto de minimizar la presencia de indeterminaciones y errores en la secuencia generada.

Si bien esta tecnología tenía la ventaja de poder generar mucha información en un único test, en un tiempo relativamente corto y a un coste razonable, tenía una serie de limitaciones, como la ausencia de protocolos de procesamiento estandarizados, el coste asociado al diseño de la plantilla a partir de la cual se fabrican los *arrays* (que disminuye las posibilidades de rediseño como oportunidad de mejora de la prueba) y una reducida capacidad de detección de alteraciones del tipo inserciones y deleciones. Es muy posible que por todo ello su implementación en laboratorios de investigación o de diagnóstico genético no haya llegado a las cotas presumiblemente esperadas en el momento del lanzamiento de la tecnología.

NEXT GENERATION SEQUENCING

Actualmente, el abordaje del estudio del genoma ha cambiado de forma espectacular desde el lanzamiento, en 2005, del primer equipo de secuenciación masiva paralela o *next generation sequencing* (NGS) (Roche 454). Desde entonces se han desarrollado diferentes opciones tecnológicas para el análisis genético por ultrasecuenciación: 454 de Roche, HiSeq2000 de Illumina, SOLID 5500 de LifeTech, PGM de Ion Torrent, MiSeq de Illumina, Junior de Roche, Pac Bio, Helicos.

En pocos años se ha evolucionado hasta la comercialización de plataformas con una capacidad de generación de datos muy elevada, incluso superior a los 180 Gb por carrera (Illumina HiSeq2000, SOLID 5500, Helicos, entre otros). Así pues, presentan una capacidad de generación de datos/día cifrada en unas 22.000 veces más que un secuenciador convencional²⁹. De manera general, la base científica de estos equipos NGS es:

- Selección de las regiones de interés del genoma (*targeted resequencing* o *exome sequencing*) mediante soluciones de enriquecimiento (basadas en chips o en solución como SureSelect de Agilent y EZ Exome/SeqCap de Nimblegen, o en solución solamente como RainDance, TruSeq de Illumina o basadas en el uso de PCR como es el caso de Qiagen y Fluidigm) que permiten la separación específica de estas

regiones del resto del genoma, aumenta la especificidad de la prueba y disminuye los costes asociados a su estudio.

- Gran capacidad de generación de datos, basada en la producción de varios millones de secuencias de tamaño discreto (entre 35 y 400 pb en los principales equipos de NGS) en varios días de reacción (de 1 a 12 días), dando lugar a que cada una de las posiciones sometidas a estudio sean finalmente secuenciadas varias decenas de veces, lo que en cierta medida facilita los procesos analíticos que se llevan a cabo a partir de los datos brutos (alineamiento, *variant calling*, anotación, entre otros). La generación de secuencias de tamaño relativamente pequeño dificulta enormemente el proceso de alineamiento, hecho que no sucedía, por ejemplo, con la secuenciación convencional Sanger. Para solucionar estos problemas, es necesario generar un número elevado de lecturas en cada posición genómica, punto crucial en el análisis con fines diagnósticos. La ausencia de una cobertura adecuada genera problemas en la detección de variaciones, por lo que disminuye la fiabilidad de los resultados. Este problema es más relevante, si cabe, ante la presencia de mutaciones del tipo deleciones o inserciones.

Los proveedores tecnológicos, con la mejora de sus químicas basadas en diferentes tecnologías (*pyrosequencing* en 454 de Roche, terminadores reversibles en Illumina o secuenciación por ligación en el caso de SOLID de LifeTech), han centrado sus esfuerzos en aumentar la capacidad de generación de datos de estos equipos, así como en reducir significativamente los costes de los reactivos necesarios para estas pruebas, con objeto de facilitar a la comunidad científica el acceso a estas tecnologías. De esta forma, los equipos han multiplicado por 10 su capacidad de generación de datos en apenas 5 años y el coste por millón de bases secuenciadas (1 Mb) ha pasado en ese tiempo de unos 0,9 dólares al precio actual de unos 0,09 dólares, alcanzado un coste teórico de análisis del genoma codificante en unos 10.000 dólares (suponiendo un grado de cobertura, esto es, un número de lecturas por cada posición, de 30 veces o 30x)²⁹. Por lo tanto, se ha reducido hasta 100 veces el coste respecto al del primer genoma secuenciado con tecnología 454 de Roche, cifrado en 1 millón de dólares³⁰.

El objetivo de lograr la secuenciación del genoma completo por 1.000 dólares resulta más alcanzable gracias al desarrollo y la evolución de los protocolos actuales y las nuevas tecnologías, como la secuenciación de molécula única a tiempo real o la secuenciación mediante nanoporos.

Next generation sequencing en el diagnóstico

Podemos decir que se ha alcanzado un punto en el que no hay limitación aparente en la capacidad de generación de datos de estas nuevas herramientas de análisis, y todo estudio parece abordable. Sin embargo, el uso de estas tecnologías no se ha trasladado apenas al campo del diagnóstico, por un lado, posiblemente debido al coste del equipamiento necesario (entre 400.000 y 1.000.000 dólares según la tecnología), la complejidad experimental asociada al procesamiento de las muestras y la ausencia de protocolos optimizados y estandarizados (en parte debido a la continua evolución de las químicas de secuenciación) y, por otro, a las enormes dificultades asociadas al manejo y el análisis de los datos obtenidos (alineación, ensamblaje y anotación de variantes). La interpretación y el almacenamiento de la gran cantidad de datos que se generan en cada carrera suponen también un problema para los centros especializados³¹.

Otro punto que merece también una mención especial es el tiempo de generación de los datos, al menos en la mayoría de los equipos más distribuidos comercialmente (plataformas Illumina y SOLID). Actualmente es del orden de 10 a 12 días, por lo que, cuando la necesidad de un diagnóstico es cuestión de unas horas, estas aproximaciones no son válidas todavía.

Si bien antes mencionábamos que entre los objetivos de los proveedores de tecnologías de secuenciación masiva estaba dotar a sus

equipos de una química de mayor calidad, más rápida y menos costosa, no han desarrollado en la misma medida sus soluciones de análisis de datos, de forma que, en muchas ocasiones, los investigadores han de usar programas *libres* (*open source*) a partir de los cuales optimizar su propio proceso de análisis en aras de conseguir unos resultados robustos y reproducibles³². La ausencia de programas y algoritmos universales es otro de los hechos que dificulta el salto de esta tecnología al campo del diagnóstico; en algunas ocasiones incluso se ha puesto de manifiesto que la aplicación de diferentes soluciones informáticas de análisis a los mismos datos brutos ha dado lugar a unos resultados, entendidos como un conjunto de variaciones en el genoma, que no se solapan completamente³².

Se trata de una tecnología relativamente joven y que todavía tiene mucho margen de mejora en determinados apartados como alguno de los mencionados anteriormente: diseño de soluciones de enriquecimiento más eficaces y específicas, desarrollo de algoritmos bioinformáticos sencillos, potentes y eficaces para el análisis de los datos (sobre todo en la detección de inserciones y deleciones), estandarización de protocolos experimentales, fijación de parámetros de calidad que aseguren la generación de unos datos sólidos, reproducibles y fiables y que disminuya el tiempo de reacción de secuenciación.

Algunas de las principales compañías tecnológicas en este campo, como Roche e Illumina, están trabajando activamente para mejorar estos aspectos, y ya en los últimos meses de 2010 y durante 2011 han puesto en el mercado equipos mucho más pequeños, más económicos y con una duración de la carrera mucho menor (GS Junior de Roche, 35 Mb en 10 h; MiSeq de Illumina, 1,5 Gb en 27 h). Asimismo, con objeto de aumentar en la medida de lo posible la reproducibilidad de los resultados, se han reducido los protocolos de trabajo y se han eliminado notablemente de ellos las etapas en que se produce intervención humana. Además, se ha realizado un esfuerzo considerable a la hora de implementar algoritmos de análisis primarios de los datos, lo que reduce la necesidad de recursos en este apartado.

Con esta misma filosofía, pero con una química muy diferente, a mediados de 2011 se presentó el PGM de Ion Torrent (LifeTech), equipo de reducido tamaño y coste económico, capaz de generar alrededor de 1,5 Gb de secuencias en lecturas de unos 200 pb y en un tiempo récord de 2 h. La clave de esta velocidad en la generación de datos es el uso de una química muy sencilla, basada en la detección mediante un sensor de las variaciones de pH causadas por los protones que se liberan con la incorporación de cada nucleótido en el proceso de elongación de la cadena secuenciada. El equipo no necesita escáneres ni cámaras y está libre de cascadas enzimáticas, por lo que el coste del equipo, los reactivos y su mantenimiento es mucho menor que el de las demás tecnologías.

Cabe pensar que este tipo de equipos facilitará el salto al campo del diagnóstico basado en técnicas de NGS. Sin embargo, sigue siendo necesaria una optimización específica del procedimiento de análisis de datos para poder llevar a cabo esta transferencia de confianza del campo de la investigación hacia el diagnóstico clínico.

APLICACIÓN DE LA TECNOLOGÍA DE ALTO RENDIMIENTO AL DIAGNÓSTICO GENÉTICO DE LA MUERTE SÚBITA

En los últimos años se ha descrito un elevado número de genes asociados a diversos fenotipos con alto riesgo de MSC, como SBr (13 genes), SQTC/SQTL (14 genes), MAVD (9 genes), MCD (más de 25 genes) o MCH (más de 16 genes), que hacen de este tipo de enfermedades las candidatas idóneas para la utilización de tecnologías de ultrasecuenciación, ya que su estudio exhaustivo con técnicas convencionales de secuenciación tendría un coste muy elevado y sería muy lento.

Se han publicado varios trabajos que han puesto en marcha programas de diagnóstico genético basados en el uso de estas tecnologías^{33,34} y han demostrado su utilidad en el análisis completo

de todos los genes candidatos. Los resultados muestran que el uso de tecnologías de NGS en este campo se asocia a una notable reducción de los costes (en torno a 20 veces menor) y del tiempo de procesado en comparación con tecnologías de secuenciación convencionales sin perder precisión. Además, numerosas empresas privadas desarrollan su propia aproximación de servicios diagnósticos basados en la aplicación de estas tecnologías que suelen seguir un mismo patrón. Se lleva a cabo el análisis de un panel general de genes que engloba todos los fenotipos. Este panel permite llevar a cabo un análisis específico del fenotipo de interés a partir de los datos procedentes de los genes relacionados^{31,35-37}.

Distintos factores justifican la necesidad de evolucionar tecnológicamente en el estudio de la genética de la MS: el salto de la genética de la MS a la práctica clínica y a la prevención en familiares, la gran heterogeneidad alélica y de *locus* con un gran y creciente número de genes asociados a la MSC, así como las limitaciones asociadas, en cuanto a tiempo y coste económico, de la tecnología Sanger.

Así, durante los últimos años, nuestros esfuerzos se han dirigido a poner a disposición de la comunidad médica unas herramientas de diagnóstico genético de la MSC rápidas, precisas y coste-efectivas.

Los pilares en que se basa la realización de estos análisis son los siguientes:

- Establecimiento de unos procedimientos experimentales de calidad y altamente reproducibles.
- Generación de datos precisos y validados por tecnología de referencia en todos los casos.
- Tiempo de entrega de resultados en consonancia con las dificultades del análisis y costes analíticos ajustados.
- Emisión de informe analítico de calidad.
- Generación de un informe de recomendaciones específico que facilite al clínico la interpretación de los resultados y le aporte información de utilidad para el manejo específico de sus pacientes.

La evolución normal de todos los centros que se dedican al diagnóstico con fines clínicos ha sido, primero, el análisis de un número discreto de genes por secuenciación convencional Sanger^{38,39}, para pasar después hacia *microarrays* y/o NGS. Nuestro grupo participó, en colaboración con el departamento de I+D de Gendiag.exe, también en este proceso con el desarrollo de diferentes tecnologías (chip de resecuenciación y NGS):

- Un chip de resecuenciación para el análisis de hasta 24 genes implicados en la presencia de enfermedades arritmogénicas del corazón (SBr, SQTC/SQTL, MAVD y TVPC, entre otros). El chip de resecuenciación se desarrolló con tecnología Affymetrix y permitía el análisis de las regiones codificantes de 24 genes asociados a MSC de causa arritmogénica, en total 420 exones y sus nexos intrón-exón, con una extensión final de unas 82 kb y con un grado de solapamiento de las sondas de al menos 1×. El grado de cobertura alcanzado fue del 99,96% de las bases, con un valor medio de precisión del 99,6%, concordancia de los datos generados del 99,9% y reproducibilidad del 99,9%. El gran avance que por entonces habían experimentado ya las tecnologías de NGS (referido fundamentalmente a su reducción de costes y su incremento en la capacidad analítica) nos condujo finalmente a centrar todos los esfuerzos en transferir el conocimiento adquirido a un servicio basado en tecnologías de secuenciación masiva paralela.

- El desarrollo de un servicio basado en ultrasecuenciación (NGS) para el estudio completo de más de un centenar de genes que han sido asociados con enfermedades estructurales y arritmogénicas del corazón; este servicio permite el análisis de las causas moleculares de la mayoría de los fenotipos descritos en este campo, como SBr, SQTC/SQTL, MCD, MCH, SWPW, MAVD, TVPC, síndrome de la muerte súbita infantil (SIDS), defectos de la conducción, enfermedad de Marfan y síndrome de Fabry, entre otros.

La mayoría de los datos generados hasta el momento, especialmente los relativos al *call rate*, la precisión frente a la tecnología de referencia, la reproducibilidad y la sensibilidad diagnóstica, están por encima de lo esperado y lo publicado por otros grupos tanto académicos como empresariales³⁴⁻³⁷. La robustez de los resultados es un dato clave en la tecnología de diagnóstico clínico.

En este sentido, el análisis de las primeras 20 muestras incluidas en la validación del panel inicial de 55 genes (1.210 exones + nexos intrón-exón) ha revelado los siguientes datos:

- *Call rate* a un grado de cobertura de 20× cifrado en el 99,82%.
- Sensibilidad y especificidad diagnósticas (capacidad para detectar variantes génicas) frente a secuencia de referencia del 100% en ambos casos.
- Una concordancia entre réplicas del mismo lote y de lotes distintos que se ha determinado en el 99,8%.

VALORACIÓN DE LOS ANÁLISIS GENÉTICOS

En estos últimos años, la tecnología ha avanzado de forma espectacular, más rápidamente que la capacidad que los investigadores tenían para desarrollar herramientas de la máxima calidad diagnóstica. Ha quedado claramente establecido que hay dos grandes niveles de análisis en las enfermedades cardíacas potencialmente genéticas: *a)* el uso de análisis de resecuenciación Sanger en las enfermedades causadas por un pequeño número de genes, y *b)* el uso de ultrasecuenciación (también de exoma o genoma completo) para el análisis de pacientes con enfermedades causadas por varios genes o para el análisis de enfermedades aún indeterminadas (MS de origen desconocido).

Se debe evaluar la validez de los resultados para el posterior proceso diagnóstico. A grandes rasgos, la valoración de los análisis genéticos masivos debe tener en cuenta tres grandes aspectos:

1. Valores de cobertura. Hace referencia al número de ocasiones que se ha podido genotipificar y analizar una posición concreta del genoma, un nucleótido. Cuando se usa la tecnología de ultrasecuenciación, cada base se analiza un número variable de veces pero, al igual que en la secuenciación Sanger convencional, es posible que no se pueda analizar regiones genómicas concretas o el análisis no pueda llevarse a cabo con el mismo nivel de cobertura. La cobertura mínima aceptada para el uso de estos datos como herramienta diagnóstica depende en parte de la tecnología y de la aplicación concreta (resecuenciación dirigida, exoma completo, etc.), por lo que hasta el momento no se han establecido los valores umbral aceptados para diagnóstico y para investigación. Una baja cobertura se puede aceptar en estudios de investigación, donde el error en la lectura del ADN puede ser menos crítica, pero no en aplicaciones diagnósticas.

2. El análisis bioinformático posterior. Una vez superada la primera fase de análisis, tiene lugar una fase de valoración posterior de los resultados en la que los datos de cobertura de las variantes detectadas orientan sobre el umbral a partir del cual la detección de un nucleótido se considerará válida o no. Este es un proceso de validación crítico, que requiere estudios previos y la creación de algoritmos a partir del análisis de un número elevado de muestras para detectar con precisión las variantes reales en cada caso particular.

3. Actualmente son pocos los laboratorios que tienen la capacidad tecnológica para cumplir con los criterios técnicos tan estrictos para ofrecer este tipo de análisis de las enfermedades cardiovasculares y que estén preparados para el análisis, la interpretación y el almacenamiento de los datos resultantes. Es imprescindible reforzar el concepto de que la genética es altamente compleja y dinámica, tanto por el continuo avance de las tecnologías —que complica la rentabilización de la inversión necesaria para la realización de este tipo de análisis— como por la necesidad de una dilatada experiencia que facilite la interpretación de los resultados; el estudio de un número elevado de

muestras aumenta la eficiencia de este tipo de servicios. Hay que recordar, además, que a menudo se debe tomar decisiones clínicas dirigidas al diagnóstico familiar.

El campo ha evolucionado de forma espectacular. El clínico hoy tiene la posibilidad de solicitar un estudio genético exhaustivo de forma comercial. Hay que tener en cuenta que el contenido de un informe genético de este alcance (por el elevado número de genes analizados) puede ser complejo para los profesionales que no estén familiarizados, y las repercusiones psicológicas, médicas y preventivas del informe pueden ser dramáticas en algunos casos para el paciente y sus familias. Por lo tanto, resulta recomendable que estas decisiones se tomen en centros especializados con expertos en la enfermedad y la valoración de los resultados y con el necesario apoyo de consejo genético.

CONCLUSIONES

Los avances tecnológicos aplicados a la investigación biomédica permiten ahondar en el conocimiento del genoma humano. Entender las enfermedades cardiovasculares a escala genómica puede permitirnos caracterizar específicamente a cada paciente para optimizar y dirigir terapias personalizadas. En todas las ramas de la medicina, incluida la cardiología, la genética tiene el potencial de ser una herramienta que complemente los datos clínicos proporcionando información acerca de la predisposición a enfermedades, lo que permitirá mejorar nuestras decisiones sobre las medidas de prevención o terapéuticas para cada paciente. Pero la tecnología genética avanza rápidamente y la información obtenida en cada muestra por las nuevas tecnologías requiere un análisis minucioso de sus implicaciones médicas que complica la capacidad de interpretación del médico no especializado. Es necesario que esta interpretación se realice en centros especializados tanto en la investigación de estas enfermedades como en el manejo de datos genéticos a gran escala.

CONFLICTO DE INTERESES

Ramón Brugada es consultor científico de la empresa Gendiag.

BIBLIOGRAFÍA

1. Oliva A, Brugada R, D'Aloja E, Boschi I, Partemi S, Brugada J, et al. State of the art in forensic investigation of sudden cardiac death. *Am J Forensic Med Pathol.* 2011;32:1-16.
2. Marrugat J, Elosua R, Gil M. Epidemiology of sudden cardiac death in Spain. *Rev Esp Cardiol.* 1999;52:717-25.
3. Morentin B, Suarez-Mier MP, Aguilera B, Bodegas A. Myocardial disease mortality in children and young adults. A population-based observational study. *Rev Esp Cardiol.* 2006;59:238-46.
4. Adabag AS, Luepker RV, Roger VL, Gersh BJ. Sudden cardiac death: epidemiology and risk factors. *Nat Rev Cardiol.* 2010;7:216-25.
5. Adabag AS, Peterson G, Apple FS, Titus J, King R, Luepker RV. Etiology of sudden death in the community: results of anatomical, metabolic, and genetic evaluation. *Am Heart J.* 2010;159:33-9.
6. Maron BJ, Maron MS, Semsarian C. Double or compound sarcomere mutations in hypertrophic cardiomyopathy: a potential link to sudden death in the absence of conventional risk factors. *Heart Rhythm.* 2012;9:57-63.
7. Bauce B, Nava A, Boffagna G, Basso C, Lorenzon A, Smaniotto G, et al. Multiple mutations in desmosomal proteins encoding genes in arrhythmogenic right ventricular cardiomyopathy/dysplasia. *Heart Rhythm.* 2010;7:22-9.
8. Millat G, Bouvagnet P, Chevalier P, Dauphin C, Jouk PS, Da Costa A, et al. Prevalence and spectrum of mutations in a cohort of 192 unrelated patients with hypertrophic cardiomyopathy. *Eur J Med Genet.* 2010;53:261-7.
9. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm.* 2011;8:1308-39.
10. Sanger F, Nicklen S, Coulson AR. ADN sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74:5463-7.
11. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409:860-921.
12. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science.* 2001;291:1304-51.

13. Benson DW, MacRae CA, Vesely MR, Walsh EP, Seidman JG, Seidman CE, et al. Missense mutation in the pore region of HERG causes familial long QT syndrome. *Circulation*. 1996;93:1791-5.
14. Tanaka T, Nagai R, Tomoike H, Takata S, Yano K, Yabuta K, et al. Four novel KVSQTCL1 and four novel HERG mutations in familial long-QT syndrome. *Circulation*. 1997;95:565-7.
15. Varnava A, Baboonian C, Davison F, De Cruz L, Elliott PM, Davies MJ, et al. A new mutation of the cardiac troponin T gene causing familial hypertrophic cardiomyopathy without left ventricular hypertrophy. *Heart*. 1999;82:621-4.
16. Hernessniemi JA, Karhunen PJ, Oksala N, Kähönen M, Levula M, Rontu R, et al. Interleukin 18 gene promoter polymorphism: a link between hypertension and pre-hospital sudden cardiac death: the Helsinki Sudden Death Study. *Eur Heart J*. 2009;30:2939-46.
17. Allegue C, Gil R, Sanchez-Diz P, Torres M, Quintela I, Carracedo A, et al. A new approach to long QT syndrome mutation detection by Sequenom MassARRAY system. *Electrophoresis*. 2010;31:1648-55.
18. Donger C, Georges JL, Nicaud V, Morrison C, Evans A, Kee F, et al. New polymorphisms in the interleukin-10 gene – relationships to myocardial infarction. *Eur J Clin Invest*. 2001;31:9-14.
19. Warrington JA, Shah NA, Chen X, Janis M, Liu C, Kondapalli S, et al. New developments in high-throughput resequencing and variation detection using high density microarrays. *Hum Mutat*. 2002;19:402-9.
20. Hacia JG. Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet*. 1999;21(1 Suppl):42-7.
21. Hacia JG, Collins FS. Mutational analysis using oligonucleotide microarrays. *J Med Genet*. 1999;36:730-6.
22. Hacia JG, Fan JB, Ryder O, Jin L, Edgemon K, Ghandour G, et al. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat Genet*. 1999;22:164-7.
23. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res*. 2009;37(Web Server issue):W600-605.
24. Hartmann TB, Mattern E, Wiedemann N, Van Doorn R, Willemze R, Niikura T, et al. Identification of selectively expressed genes and antigens in CTCL. *Exp Dermatol*. 2008;17:324-34.
25. Kothiyal P, Cox S, Ebert J, Aronow BJ, Greinwald JH, Rehm HL. An overview of custom array sequencing. *Curr Protoc Hum Genet*. 2009;Chapter 7:Unit 7 17.
26. Waldmuller S, Muller M, Rackebrandt K, Binner P, Poths S, Bonin M, et al. Array-based resequencing assay for mutations causing hypertrophic cardiomyopathy. *Clin Chem*. 2008;54:682-7.
27. Fokstuen S, Lyle R, Munoz A, Gehrig C, Lerch R, Perrot A, et al. A ADN resequencing array for pathogenic mutation detection in hypertrophic cardiomyopathy. *Hum Mutat*. 2008;29:879-85.
28. Zimmerman RS, Cox S, Lakdawala NK, Cirino A, Mancini-DiNardo D, Clark E, et al. A novel custom resequencing array for dilated cardiomyopathy. *Genet Med*. 2010;12:268-78.
29. Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics*. 2011;38:95-109.
30. Metzker ML. Sequencing technologies – the next generation. *Nat Rev Genet*. 2010;11:31-46.
31. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem*. 2009;55:641-58.
32. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet*. 2011;12:443-51.
33. Brion M, Quintela I, Sobrino B, Torres M, Allegue C, Carracedo A. New technologies in the genetic approach to sudden cardiac death in the young. *Forensic Sci Int*. 2010;203:15-24.
34. Meder B, Haas J, Keller A, Heid C, Just S, Borries A, et al. Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circ Cardiovasc Genet*. 2011;4:110-22.
35. Voelkerding KV, Dames S, Durtschi JD. Next generation sequencing for clinical diagnostics—principles and application to targeted resequencing for hypertrophic cardiomyopathy: a paper from the 2009 William Beaumont Hospital Symposium on Molecular Pathology. *J Mol Diagn*. 2010;12:539-51.
36. Dames S, Durtschi J, Geiersbach K, Stephens J, Voelkerding KV. Comparison of the Illumina Genome Analyzer and Roche 454 GS FLX for resequencing of hypertrophic cardiomyopathy-associated genes. *J Biomol Tech*. 2010;21:73-80.
37. Gowrisankar S, Lerner-Ellis JP, Cox S, White ET, Manion M, LeVan K, et al. Evaluation of second-generation sequencing of 19 dilated cardiomyopathy genes for clinical applications. *J Mol Diagn*. 2010;12:818-27.
38. Campuzano O, Beltrán-Álvarez P, Iglesias A, Scornik F, Pérez G, Brugada R. Genetics and cardiac channelopathies. *Genet Med*. 2010;12:260-7.
39. Brugada R. Sudden death: managing the family, the role of genetics. *Heart*. 2011;97:676-81.