

Nuevas tecnologías aplicadas en la detección de alteraciones genéticas

Idoia Martínez De LaPiscina Martín, Gustavo Pérez de Nanclares Leal

Investigación en Genética y Control de Diabetes y Enfermedades Endocrinas. Instituto de Investigación Sanitaria Biocruces, Hospital Universitario Cruces, CIBERDEM, CIBERER, UPV-EHU. Barakaldo.

INTRODUCCIÓN

El avance en el conocimiento del genoma humano y el desarrollo tecnológico ocurrido en los últimos 20 años han logrado determinar las bases genéticas de muchas enfermedades, y con ello abrir una puerta a encontrar los medios para su prevención, diagnóstico y tratamiento.

El rápido progreso de las técnicas de secuenciación está revolucionando los estudios genéticos de las enfermedades congénitas. La llegada de la secuenciación masiva (NGS, *Next Generation Sequencing*) ha transformado el estudio de la genética humana, reduciendo asimismo los costes en comparación a la secuenciación tradicional (secuenciación Sanger) ⁽¹⁾. Se ha modificado la visión del estudio genético clásico, ampliando considerablemente el número de genes candidatos a analizar para un mismo fenotipo.

El objetivo de este trabajo es hacer una revisión de la distinta metodología, sus aplicaciones y posibilidades, tanto para el diagnóstico clínico como para la identificación de nuevas causas genéticas. Probablemente, la NGS y los análisis bioinformáticos que ésta requiere se conviertan en el estándar para el cribado de alteraciones que puedan causar enfermedades genéticas. Diversos estudios han demostrado su eficacia y rentabilidad ⁽²⁾, y a medida que su costo disminuya, su expansión será posible a otras poblaciones de pacientes que se beneficien del diagnóstico y del consejo genético.

DE LA SECUENCIACIÓN TRADICIONAL A LA SECUENCIACIÓN MASIVA

Durante muchos años, la genética clínica partió del estudio de las enfermedades monogénicas, donde

se establece el paradigma “un gen-una enfermedad”, es decir, existe una relación causa-efecto entre la alteración de un determinado gen y una enfermedad concreta. Posteriormente, se han identificado un número mayor de genes asociados a enfermedades heterogéneas genética y clínicamente, donde un mismo gen está ligado a distintas enfermedades o donde un mismo fenotipo está asociado a varios genes.

Es por ello que la estrategia del estudio genético ha cambiado. Antes de la era de la genómica actual, se ha venido utilizando la secuenciación Sanger para el análisis de uno o unos pocos genes. El algoritmo diagnóstico consistía en iniciar el estudio del gen más frecuentemente mutado asociado a la enfermedad y, en caso de detectar la variante patogénica, confirmar el diagnóstico. Si no era así, se continuaba con la secuenciación de diferentes genes candidatos hasta detectar o no la mutación causante de la enfermedad ⁽³⁾. Este procedimiento implicaba una serie de limitaciones en relación coste-eficiencia cuando se aplicaba a una patología genéticamente heterogénea en la que no existe un gen candidato, sino varios ⁽⁴⁾.

En la última década, mientras que la secuenciación Sanger se ha mantenido como el “gold standard” para la confirmación de variantes, la NGS ha revolucionado la genética, permitiendo la secuenciación simultánea de miles o millones de fragmentos de ADN en un único proceso ⁽⁵⁾. Se han desarrollado distintas técnicas de captura, es decir de amplificación, de ADN ^(6, 7) que permiten abarcar desde pocas kilobases hasta la captura del exoma completo, y que, combinados con la NGS, hacen posible realizar un screening mutacional en cientos de genes a la vez, proporcionando una gran cantidad de información sobre la genética de un individuo y acortando el tiempo de diagnóstico molecular de las enfermedades genéticas.

La tecnología NGS permite la secuenciación de un panel de genes diana (TS, *Targeted Sequencing*), el genoma entero (WGS, *Whole- Genome Sequencing*) o únicamente la región codificante del mismo (WES, *Whole- Exome sequencing*) en cuestión de días, dependiendo de la tecnología y protocolo utilizado. Esto último también determinará la sensibilidad y precisión en la detección de variantes. Por todo ello, es importante conocer las distintas herramientas y análisis de las que se dispone para aplicarlo al diagnóstico clínico (Tabla1) ^(8,9):

1. Panel de genes (TS)

Esta técnica permite capturar zonas concretas del genoma (en este caso, genes de interés) que son secuenciados simultáneamente de una manera económica.

Ventajas

Además del beneficio del menor coste de la técnica, los paneles de secuenciación se plantean como el primer análisis genético por su rapidez y la alta profundidad de lectura del objetivo ⁽⁹⁾. Por otra parte, la interpretación de las alteraciones que se descubren es más sencilla y se minimizan los hallazgos accidentales ⁽¹⁰⁾.

La aplicación de paneles de secuenciación masiva al diagnóstico genético requiere del estudio de los ge-

nes asociados a la patología, del diseño bioinformático del panel y de su validación, que permita evaluar los parámetros de calidad del panel, como la reproducibilidad, cobertura media, sensibilidad, especificidad, detección de delecciones e indels (“inserción o delección”), previa a su aplicación en el diagnóstico de enfermedades genéticas ^(11, 12).

Desventajas

Una de las principales limitaciones de los paneles de genes es la dificultad para detectar determinadas alteraciones, especialmente las variaciones en el número de copias (CNVs), en regiones con número insuficiente de lecturas o con menor cobertura. Esta cobertura menor al 100% ocurre sobre todo en los primeros exones de genes con regiones ricas en GCs ⁽¹³⁾. Por otra parte, de acuerdo con las recomendaciones del Colegio Americano de Genética y Genómica Médica (ACMG, *American College of Medical Genetics and Genomics*), solo aquellos genes que han sido científicamente asociados con una enfermedad deben ser incluidos en el panel de genes ⁽¹²⁾. Por ello, el resultado para un paciente concreto puede no ser concluyente, y requerir de otro panel de genes, otro método adicional como los arrays de hibridación genómica compuesta (acGH) para la detección de CNVs o el análisis genético mediante WES o WGS.

Dada la rapidez con la que se identifican nuevos genes asociados a una enfermedad, es necesaria la ac-

Tabla 1. Indicaciones para el análisis de un solo gen, panel de genes o WES/WGS y ejemplos.

Estudios de un solo gen	
Enfermedad producida por mutaciones específicas en un solo gen	Neoplasia Endocrina Múltiple tipo 2 (MEN2)
Enfermedad con heterogeneidad mínima	Sínd. de insensibilidad a los andrógenos
Limitaciones de NGS, p.e., detección de repetición de trinucleótidos	Síndrome de Kennedy
Panel de genes	
Enfermedad con heterogeneidad genética	46, XY DSD Hiperparatiroidismo primario Diabetes tipo MODY
Trastornos solapantes que se encuentran en el diagnóstico diferencial de la patología genética	
Trastornos que comparten manifestaciones clínicas, pero tienen diferencias en la presentación global de la enfermedad	
Enfermedades producidas por alteración de una vía común	
WES/WGS	
Enfermedades con máxima heterogeneidad genética	Dismorfologías sindrómicas
2 o más fenotipos relacionados en un paciente	
Ausencia de características clínicas claves en el momento de la prueba	
Fenotipo indistinto y la causa subyacente no se identifica	

Modificado de Xue Y et al (2015) ⁽⁸⁾.

tualización de los paneles, especialmente en aquellas enfermedades en las que las bases moleculares no están del todo definidas. Incluir nuevos genes en el "pool" que va a amplificarse mediante NGS evitaría tener que secuenciar todos los genes conocidos mediante WES o WGS⁽⁹⁾.

2. Secuenciación del exoma completo (WES)

A pesar de que el exoma supone solo alrededor del 2% del genoma humano, el 85% de las variantes descritas causantes de enfermedad se encuentran en él⁽¹⁴⁾. El objetivo del WES es la secuenciación de los exones codificantes de todos los genes conocidos. Existen también variantes de WES comercialmente denominadas "exoma clínico" (p.e., Agilent SureSelect Focused Exome, Illumina TruSight One, Roche NimbleGen SeqCap EZ MedExome), que incluyen unos 5.000 genes asociados a enfermedad. Es decir, únicamente cubren el 20% de todo el exoma y requieren, al igual que los paneles de genes, análisis adicionales si no se encuentra la alteración causante de la clínica⁽¹⁵⁾.

Ventajas

Sin embargo, WES no solo incluye los exones de aquellos genes asociados al fenotipo en estudio, sino que también permite la identificación de nuevos genes en enfermedades con base molecular desconocida. El análisis de datos obtenido por WES puede limitarse a los genes de interés, y posteriormente ampliarse si es necesario, lo que reduce el hallazgo de alteraciones accidentales o secundarias. No es de extrañar que se esté implementando el uso de WES como el primer análisis genético a realizar, ya que hace posible el análisis de tríos (secuenciación del caso índice y sus progenitores) de una forma rentable, facilitando la interpretación de los datos y aumentando el rendimiento diagnóstico considerablemente (20-30%)^(16, 17).

Desventajas

La principal desventaja de secuenciar el exoma completo es que puede no cubrir partes del exoma que han sido pobres en enriquecidas, aunque esta cobertura podría mejorarse combinando kits de enriquecimiento⁽¹³⁾. Además, al igual que ocurre en los paneles de genes, las herramientas que han sido desarrolladas para la detección de CNVs mediante WES siguen sin ser tan precisas como las utilizadas para WGS⁽¹⁸⁾.

3. Secuenciación del genoma completo (WGS)

Ventajas

La secuenciación del genoma completo tiene la ventaja de una cobertura continua y la identificación de variantes a lo largo de todo el genoma, permitiendo

detectar variantes no exónicas, mejorar la detección de las CNVs y un mayor rendimiento sobre WES^(13, 18, 19).

La interpretación de resultados, como en WES, podría dirigirse inicialmente a las regiones de interés, ya que, a día de hoy, la interpretación de variantes en regiones intrónicas, intergénicas o reguladoras es difícil a nivel de ADN⁽²⁰⁾. Tales variantes tendrán gran importancia en un futuro cercano, incrementando el valor de los datos obtenidos mediante WGS, como ocurre hoy en día en la terapia farmacológica con algunos polimorfismos en regiones promotoras que afectan a la diana del medicamento⁽²¹⁾.

Desventajas

La mayor limitación en el uso de WGS viene determinada por la longitud de lectura de la secuencia, que al ser de alrededor de 200pb, puede llevar a que la alineación ocurra en una región del genoma que no corresponda.

¿QUÉ METODOLOGÍA ELEGIR?

Cuando se comparan todas las opciones, parece claro que la secuenciación del genoma completo (WGS) es teóricamente el mejor enfoque, ya que produce un mayor conjunto de datos sobre el genoma de un individuo (Tabla 2). Ofrece importantes ventajas: cobertura más uniforme que WES y permite además la detección de variantes intrónicas y de CNVs, incluyendo pérdidas de un solo exón⁽²²⁾. Sin embargo, obtener una cobertura suficiente incrementa los costes, y el tiempo de análisis limita su implementación en el diagnóstico genético de rutina, ya que se detectan miles de variantes. Por el contrario, WES es considerablemente más barato que WGS y es una eficiente estrategia⁽²³⁾.

El análisis de paneles genéticos representa una alternativa menos costosa, pero que solo tendrá éxito si el gen que causa la enfermedad está incluido en el panel. La ventaja es, además del corto tiempo de respuesta, que permite una mayor cobertura a un coste menor que los enfoques de todo el exoma o genoma y que la focalización restringida reduce la posibilidad de hallazgos inespecíficos o accidentales⁽⁹⁾. Esto puede solucionarse en WES y WGS dirigiendo el análisis a unos genes seleccionados y conocidos y, si no se encuentra nada patogénico, podrían añadirse genes novedosos y relevantes haciendo una pequeña modificación en el análisis bioinformático, ampliando así la búsqueda.

En WGS, la cobertura de lectura de todo el genoma puede permitir la detección de variaciones en el número de copias, repeticiones de tripletes y pequeñas delecciones, que pueden contribuir sustancialmente a la carga de una enfermedad en las que estas alte-

Tabla 2. Comparación de las 3 técnicas utilizadas mediante secuenciación masiva.

	NGS			
	Sanger	TS	WES	WGS
Longitud de lectura (pb)	1000pb	~300	~150	~150
Profundidad de lectura	NA	200-100x	100x	30-60x
Tasa de error en lectura (%)	0.001	0.1	0.1	0.1
Amplificación previa a secuenciación	Si	Si	Si	No
Ventajas	Precisión	Fácil interpretación de resultados, coste-eficacia, rapidez de respuesta	Mayor información, coste-eficacia	Uniformidad, cobertura independiente de regiones GC
Limitaciones	Bajo rendimiento	Cobertura incompleta por regiones GC, pocas sondas, regiones con mapabilidad<1	Cobertura incompleta por regiones GC, pocas sondas, regiones con mapabilidad<1	Cobertura incompleta en regiones con mapabilidad<1

TS, Targeted Sequencing; WGS, Whole- Genome Sequencing; WES, Whole- Exome sequencing; NA: No aplica. Modificado de Caspar SM et al (2017)¹⁰.

raciones genéticas son importantes. En los últimos años, los avances en la NGS han sido importantes en términos de rapidez, longitud de lectura y rendimiento. Al mismo tiempo los costes han disminuido, permitiendo la aplicación de la secuenciación masiva al diagnóstico clínico. Hace 10 años la secuenciación de un millón de pares de bases costaba alrededor de 1000\$, mientras que ahora no alcanza 0,10 \$. En la actualidad, varias casas comerciales permiten la secuenciación del genoma completo por menos de 1000\$, un logro que permitirá su aplicación clínica a gran escala, fomentando la comprensión de las enfermedades genéticas y, en última instancia, contribuir a la medicina personalizada de manera importante²⁴.

A la hora de calcular los costes hay que tener en cuenta, además del proceso de secuenciación, el equipamiento y su mantenimiento, el personal técnico para la preparación de muestras, interpretación de datos y realización de informes y, finalmente, el almacenamiento de los datos generados. La secuenciación de los paneles de genes y el exoma completo son las alternativas menos costosas en comparación con el genoma completo (Tabla 3), aunque esto no implica que deban ser los elegidos para la práctica clínica, ya que el rendimiento diagnóstico también debe ser tenido en cuenta. Este será considerablemente más alto en WGS que en WES y panel de genes, dependiendo de la población de pacientes. Además, si comparamos WGS y WES las diferencias en el rendimiento serán mínimas, teniendo en cuenta que en el estudio del exoma solo se secuencian las regiones que codifican

proteínas, en el que se cree que ocurren el 85% de las mutaciones²⁵.

Otro de los parámetros a tener en cuenta a la hora de elegir la estrategia a seguir son la sensibilidad (tasa de falsos negativos) y la especificidad (tasa de falsos positivos). En principio, la tasa de falsos positivos no es un problema en la secuenciación masiva, porque los hallazgos son confirmados, de momento, mediante secuenciación Sanger. Es, por tanto, el porcentaje de falsos negativos el principal parámetro a tener en cuenta. Cuando una enfermedad es causada principalmente por pocos genes, pero altamente penetrantes, como sucede, por ejemplo, en el cáncer de mama hereditario, encontrar un falso negativo sería inaceptable, mientras que en enfermedades más heterogéneas se acepta que haya falsos negativos, porque se está realizando un diagnóstico genético que de otra forma, como por secuenciación Sanger, sería impensable²⁶.

Por tanto, para cada población de pacientes, la decisión sobre qué estrategia de secuenciación utilizar está basada en los costes, profundidad de lectura y calidad de la secuencia.

INTERPRETACIÓN DE LOS RESULTADOS

La secuenciación masiva da como resultado alrededor de 10.000 y 1.000.000 de variantes en WES y WGS, respectivamente. Un filtrado adecuado reduce estos números mediante la distinción de las alteraciones en patogénicas o benignas²⁷.

Tabla 3. Coste (euros por muestra) de las distintas aplicaciones de NGS.

	TS	WES	WGS
Extracción de ADN	31,53	31,53	31,53
Fungible para preparación de muestra	242,62	296,68	27,61
Fungibles para secuenciación	4,56	262,24	1057,81
Personal	8,97	70,08	70,08
Procesamiento de datos	0,50	10,0	100,0
Almacenamiento de datos	0,05	0,75	30,0
Interpretación de datos	31,23	62,65	93,97
Costes totales por muestra	332,90	791,75	1669,02

Modificado de Van Nimwegen KJ *et al* (2016)⁽²⁴⁾.

La asociación entre una variante y la enfermedad se puede encontrar en la literatura (ncbi.nlm.nih.gov/pubmed) y en bases de datos como HGMD (<http://www.hgmd.cf.ac.uk/ac/index.php>) y ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>). Además, los SNV (*single nucleotide variant*) e indels se priorizan en función de la frecuencia en la población (ExAc, exac. broadinstitute.org), conservación filogenética (SiPhy, portals.broadinstitute.org/genome_bio/siphy/index.html) y el potencial efecto en la función o estructura de la proteína (PROVEAN, provean.jcvi.org/; Polyphen-2, genetics.bwh.harvard.edu/pph2/; MutPhen, mutpred. mutdb.org/). Para la priorización de CNVs, sin embargo, existen menos herramientas y se utiliza sobre todo el tamaño y frecuencia en la población (DGV, dgv. tcag.ca/). Por otro lado, aunque se sabe que las variantes en regiones de corte y empalme (*splice sites*) son clínicamente relevantes, su priorización es más complicada y sería necesario el desarrollo de mejores herramientas *'in silico'*⁽²⁸⁾.

La interpretación más completa de las alteraciones encontradas se llevará a cabo no solo consultando las bases de datos, sino realizando los análisis funcionales apropiados que permitan evaluar el efecto de las variantes sobre, por ejemplo, la expresión génica. Los análisis de segregación o de tríos proporcionarán evidencia adicional de la asociación de las variantes con la enfermedad, y si son *'de novo'* o heredadas. Sin embargo, a pesar de la interpretación, las variantes de significado incierto (VUS) permanecerán como una de las entidades más frecuentes⁽²⁷⁾.

CONCLUSIONES

El análisis genético mediante NGS aumenta el conocimiento acerca de las enfermedades de herencia mendeliana y son, por tanto, herramientas poderosas para el diagnóstico genético y el estudio etiológico de múltiples patologías. Desde el punto de vista de la investigación, el abordaje de tríos permite la identificación de nuevos genes asociados a enfermedades.

En la actualidad, debido al impulso de los avances tecnológicos, la capacidad de secuenciar es mayor que la capacidad para interpretar las variantes que se detectan. De hecho, los efectos de las VUS y de las variantes en regiones intrónicas plantean un desafío para su interpretación. Además, los métodos utilizados en la alineación o el filtrado de variantes influyen considerablemente en su detección. Por tanto, se recomienda utilizar más de un protocolo de análisis independiente y reevaluar los casos no resueltos para minimizar los falsos negativos.

Por otro lado, confiar exclusivamente en las herramientas de interpretación *'in silico'* es peligroso, no son el reemplazo del análisis de los expertos en la clínica y genética de una enfermedad o de los estudios funcionales⁽²⁹⁾. La secuenciación masiva no ha alcanzado su máximo potencial en la aplicación clínica debido a cuestiones técnicas y regulatorias, como el almacenamiento de los datos genómicos del paciente.

Finalmente, la secuenciación masiva hace posible el análisis genético de todos los genes descritos en pocas semanas, disminuyendo el tiempo de estudio que previamente era de meses, mejorando la efectividad y eficiencia en el diagnóstico, al identificar un alto número de variantes de ADN que pueden actuar como mutaciones patogénicas responsables de la enfermedad o como modificadores de la expresividad fenotípica⁽³⁰⁾. La integración de estos hallazgos genéticos en el manejo clínico de los pacientes permitirá en un futuro establecer posibles tratamientos personalizados, sin olvidar el asesoramiento genético familiar sobre las bases moleculares concretas.

GLOSARIO

- Análisis *'in silico'*: Análisis computacional que se realiza de las variantes encontradas tras la secuenciación. Se utilizan distintos software de predicción que indican la posible patogenicidad de una alteración.

- Captura: Método de selección del ADN que se quiere amplificar para su análisis mediante NGS, p.e. Ion AmpliSeq Cancer Hotspot Panel, para la amplificación de regiones de genes frecuentemente mutados en cáncer.
- CNV o *Copy Number Variation*: Duplicación o eliminación de un número considerable de pares de bases. Se considera un tipo de variación estructural y se detecta mediante p.e. aCGH.
- Cobertura: Número de veces que un determinado nucleótido es leído durante el experimento de NGS.
- Enriquecimiento: Proceso mediante el cual se recuperan las regiones capturadas para su posterior amplificación.
- Librería: Conjunto de amplicones o secuencias de interés de un paciente.
- Mapabilidad: Localización y representación de los genes y alteraciones encontrados, respecto al genoma.
- Secuenciación Sanger: Secuenciación tradicional, gen por gen, de manera progresiva. Se secuencian genes candidato asociados a un determinado fenotipo.
- Secuenciación mediante panel de genes o *Targeted sequencing* (TS): Análisis múltiple en paralelo de varios genes relacionados con fenotipos parecidos o solapantes.
- Secuenciación del exoma o *Whole- Exome Sequencing* (WES): Secuenciación del exoma conocido. Consiste en secuenciar los exones de los genes conocidos, lo que supone alrededor del 2% de todo el genoma.
- Secuenciación del genoma o *Whole- Genome Sequencing* (WGS): Secuenciación de todo el genoma de un individuo, incluyendo cualquier tipo de secuencia: exones, intrones, regiones reguladoras, promotores, etc.

3. Schully SD, Lam TK, Dotson WD, Chang CQ, Aronson N, Birkeland ML, et al. Evidence synthesis and guideline development in genomic medicine: current status and future prospects. *Genet Med.* 2015;17(1):63-7.

4. Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat.* 2013;34(7):1035-42.

5. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016;17(6):333-51.

6. Hernan I, Borràs E, de Sousa Dias M, Gamundi MJ, Mañé B, Llort G, et al. Detection of genomic variations in BRCA1 and BRCA2 genes by long-range PCR and next-generation sequencing. *J Mol Diagn.* 2012;14(3):286-93.

7. Wei X, Dai Y, Yu P, Qu N, Lan Z, Hong X, et al. Targeted next-generation sequencing as a comprehensive test for patients with and female carriers of DMD/BMD: a multi-population diagnostic study. *Eur J Hum Genet.* 2014;22(1):110-8.

8. Xue Y, Ankala A, Wilcox WR, Hegde MR. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genet Med.* 2015;17(6):444-51.

9. Caspar SM, Dubacher N, Kopps AM, Meienberg J, Henggeler C, Matyas G. Clinical sequencing: From raw data to diagnosis with lifetime value. *Clin Genet.* 2018;93(3):508-19.

10. Hehir-Kwa JY, Claustres M, Hastings RJ, van Ravenswaaij-Arts C, Christenhusz G, Genuardi M, et al. Towards a European consensus for reporting incidental findings during clinical NGS testing. *Eur J Hum Genet.* 2015;23(12):1601-6.

11. Yohe S, Hauge A, Bunjer K, Kemmer T, Bower M, Schomaker M, et al. Clinical validation of targeted next-generation sequencing for inherited disorders. *Arch Pathol Lab Med.* 2015;139(2):204-10.

12. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med.* 2013;15(9):733-47.

13. Meienberg J, Bruggmann R, Oexle K, Matyas G. Clinical sequencing: is WGS the better WES? *Hum Genet.* 2016;135(3):359-62.

14. Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, et al. Distinguishing protein-coding and noncoding ge

Referencias Bibliográficas

1. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell.* 2013;155(1):27-38.
2. Hu X, Li N, Xu Y, Li G, Yu T, Yao RE, et al. Proband-only medical exome sequencing as a cost-effective first-tier genetic diagnostic test for patients without prior molecular tests and clinical diagnosis in a developing country: the China experience. *Genet Med.* 2017.

nes in the human genome. *Proc Natl Acad Sci U S A.* 2007;104(49):19428-33.

15. Eldomery MK, Coban-Akdemir Z, Harel T, Rosenfeld JA, Gambin T, Stray-Pedersen A, et al. Lessons learned from additional research analyses of unsolved clinical exome cases. *Genome Med.* 2017;9(1):26.

16. Alazami AM, Patel N, Shamseldin HE, Anazi S, Al-Dosari MS, Alzahrani F, et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* 2015;10(2):148-61.

17. Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu YF, McSweeney KM, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med.* 2015;17(10):774-81.

18. Tattini L, D'Aurizio R, Magi A. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. *Front Bioeng Biotechnol.* 2015;3:92.

19. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A.* 2015;112(17):5473-8.

20. Kremer LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun.* 2017;8:15824.

21. Owen RP, Gong L, Sagreiya H, Klein TE, Altman RB. VKORC1 pharmacogenomics summary. *Pharmacogenet Genomics.* 2010;20(10):642-4.

22. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature.* 2014;511(7509):344-7.

23. Rauch A, Wieczorek D, Graf E, Wieland T, Ende S, Schwarzmayr T, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet.* 2012;380(9854):1674-82.

24. van Nimwegen KJ, van Soest RA, Veltman JA, Nelen MR, van der Wilt GJ, Vissers LE, et al. Is the \$1000 Genome as Near as We Think? A Cost Analysis of Next-Generation Sequencing. *Clin Chem.* 2016;62(11):1458-64.

25. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A.* 2009;106(45):19096-101.

26. Sun Y, Ruivenkamp CA, Hoffer MJ, Vrijenhoek T, Kriek M, van Asperen CJ, et al. Next-generation diagnostics: gene panel, exome, or whole genome? *Hum Mutat.* 2015;36(6):648-55.

27. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-24.

28. Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum Mutat.* 2012;33(8):1228-38.

29. Miosge LA, Field MA, Sontani Y, Cho V, Johnson S, Palkova A, et al. Comparison of predicted and actual consequences of missense mutations. *Proc Natl Acad Sci U S A.* 2015;112(37):E5189-98.

30. 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(2):110-22.