

<i>Nereis. Revista Iberoamericana Interdisciplinar de Métodos, Modelización y Simulación</i>	9	39-48	Universidad Católica de Valencia San Vicente Mártir	Valencia (España)	ISSN 1888-8550
--	---	-------	---	-------------------	----------------

Diagnóstico de cáncer de páncreas mediante secuenciación de nueva generación (NGS)

Pancreatic cancer diagnosis by new generation sequencing (NGS)

Fecha de recepción y aceptación: 13 de diciembre de 2016 y 3 de febrero de 2017

Julián Ochando-Noguera¹ y Mónica Díez-Díaz^{1*}

¹ Departamento de Ciencias Aplicadas y Tecnológicas. Facultad de Veterinaria y Ciencias Experimentales. Universidad Católica de Valencia San Vicente Mártir.

* Correspondencia: Calle Guillem de Castro, 94. 46001 Valencia. España. *E-mail*: monica.diez@ucv.es



RESUMEN

El cáncer de páncreas es una enfermedad devastadora, altamente letal, que presenta múltiples desafíos para los estudios genómicos. El índice de supervivencia medio de cinco años es del 60 %. Las pruebas y los procedimientos para detectar, diagnosticar y estadificar el cáncer de páncreas se realizan al mismo tiempo, por lo que presentan innumerables complicaciones para los pacientes. Los avances en la secuenciación de nueva generación (NGS) se han adaptado a la oncología clínica, identificando nuevas mutaciones y cánceres poco frecuentes, permitiendo así la medicina personalizada. Un análisis genómico completo de la enfermedad abre nuevas vías terapéuticas convirtiéndose en la esperanza de pacientes con mal pronóstico.

PALABRAS CLAVE: *cáncer, páncreas, secuenciación de nueva generación, NGS.*

ABSTRACT

Pancreatic cancer is a devastating disease, highly lethal that presents multiple challenges for genomic studies. The average survival rate at 5 years is 60 %. Tests and procedures to detect, diagnose and stage pancreatic cancer are performed at the same time so they have many complications for patients. The advances in the technology next-generation sequencing (NGS) have adapted to clinical oncology, identifying new mutations and rare cancers, allowing personalized medicine. A complete genomic analysis of the disease opens new therapeutic pathways becoming the hope of patients with poor prognosis.

KEYWORDS: *cancer, pancreas, next-generation sequencing, NGS.*

INTRODUCCIÓN

El cáncer de páncreas es una patología en la que se producen células malignas en los tejidos de este órgano. Presenta uno de los índices más altos de mortalidad dentro de los distintos tipos de cáncer, con una tasa de supervivencia de cinco años. En el páncreas se constituyen tumores de distintos tipos dependiendo del tipo celular, exocrino o endocrino [1]. Es imprescindible diferenciar entre el cáncer de páncreas exocrino y endocrino. Cada uno muestra factores de riesgo diferentes así como causas, signos y síntomas distintos, se diagnostican con técnicas diferentes, se asignan tratamientos distintos y sus pronósticos se diferencian entre sí. El tipo de cáncer de páncreas más común es el tumor exocrino. Por otro lado los tumores de páncreas endocrinos son poco frecuentes, representando menos del 4 % de todos los adenocarcinomas pancreáticos.

En el cáncer de páncreas los genes implicados son KRAS, CDKN2A, TP53 y SMAD4. La localización citogenética del gen KRAS se encuentra en el brazo corto del cromosoma 12. KRAS proporciona las instrucciones para generar una



proteína denominada K-Ras que está implicada principalmente en la regulación de la división celular [2]. Como parte de la ruta de señalización RAS/MAPK, la proteína transmite señales extracelulares al núcleo de la célula. Estas señales instruyen a la célula para crecer y dividirse o para madurar y asumir funciones especializadas, es decir, diferenciarse. La proteína K-Ras es una GTPasa, lo que significa que convierte una molécula de GTP en otra molécula denominada GDP. KRas actúa como un interruptor, se activa y reprime por moléculas de GTP y GDP, respectivamente. Para transmitir señales, la proteína K-Ras debe estar activada unida a una molécula de GTP. K-Ras se desactiva cuando el GTP se convierte en GDP, es decir, cuando la proteína está unida al GDP no transmite las señales al núcleo de la célula. KRAS pertenece a una clase de genes conocidos como oncogenes. Cuando mutan, los oncogenes tienen el potencial de causar que las células normales se vuelvan cancerosas. El gen KRAS pertenece a la familia de oncogenes Ras, que también incluye otros dos genes: HRAS y NRAS. Las proteínas producidas a partir de estos tres genes son GTPasas. Estas proteínas juegan un papel muy importante en la división celular, en la diferenciación celular y en la apoptosis.

El gen CDKN2A genera diferentes variantes de transcritos que difieren en sus primeros exones [3]. Tres variantes de splicing alternativo codifican proteínas distintas, dos de las cuales codifican isoformas estructuralmente relacionadas para funcionar como inhibidores de la quinasa CDK4. El transcrito restante incluye un sustituto en el primer exón situado a 20 Kb upstream del resto del gen; este transcrito contiene un marco de lectura abierto alternativo (ARF), que especifica una proteína que no guarda relación estructuralmente con los productos de las otras variantes. Este producto ARF funciona como un estabilizador de la proteína supresora de tumores p53, ya que puede interactuar y secuestrar la proteína ubiquitina ligasa E3, una proteína responsable de la degradación de p53. A pesar de las diferencias estructurales y funcionales, las isoformas del inhibidor CDK y el producto ARF codificados por este gen, comparten una funcionalidad común en el control de la fase G1 del ciclo celular. Este gen es mutado o suprimido frecuentemente en una amplia variedad de tumores y se sabe que es un gen supresor de tumores importante.

El gen TP53 codifica una proteína supresora de tumores que contiene la activación de la transcripción, la unión al DNA y los dominios de oligomerización. La proteína codificada responde a diversos estreses celulares para regular la expresión de los genes diana, induciendo de esta forma la detención del ciclo celular, la apoptosis, la senescencia, la reparación del DNA o cambios en el metabolismo [4]. Las mutaciones en este gen están asociadas con una gran variedad de cánceres humanos, incluyendo cánceres hereditarios como el síndrome de Li-Fraumeni. El splicing alternativo de este gen y el uso de promotores alternativos resultan en múltiples variantes de la transcripción e isoformas. Las isoformas adicionales han demostrado ser el resultado del uso de los codones de iniciación de la traducción alternativa.

El gen SMAD4 codifica un miembro de la familia de proteínas de transducción de señales Smad. Las proteínas Smad son fosforiladas y activadas por las quinasas receptoras transmembrana serina/reonina en respuesta a la vía de señalización TGF- β . El producto de este gen forma complejos homoméricos y complejos heterómeros con otras proteínas Smad activadas, que se acumulan en el núcleo y regulan la transcripción de genes diana [5]. Esta proteína se une al DNA y reconoce una secuencia palindrómica de 8 pb (GTCTAGAC) denominada elemento de unión Smad (SBE). Las proteínas Smad están sometidas a una regulación compleja por modificaciones postraduccionales. Las mutaciones o deleciones en este gen han demostrado dar lugar a cáncer de páncreas.

DIAGNÓSTICO DE CÁNCER DE PÁNCREAS EN LA ACTUALIDAD

Para diagnosticar el cáncer de páncreas se realizan estudios por imágenes que permiten visualizar el páncreas y toda el área circundante. La estadificación es el proceso por el cual se determina si las células tumorales se extienden dentro y alrededor del páncreas [6]. Los procedimientos para detectar, diagnosticar y estadificar el cáncer de páncreas normalmente se realizan simultáneamente. Para programar el tratamiento se necesita conocer la fase en la que se encuentra el cáncer y poder decidir así si se puede eliminar por cirugía. Las técnicas que permiten diagnosticar el cáncer de páncreas son la colangiopancreatografía retrógrada endoscópica (CPRE) y la biopsia. Sin embargo, aunque estas dos técnicas son las principales en la confirmación del diagnóstico de cáncer de páncreas, presentan el inconveniente de ser invasivas [7], por lo que pueden presentar complicaciones a los pacientes.

La CPRE es un procedimiento empleado para tomar radiografías de los conductos que llevan la bilis del hígado a la vesícula biliar y de la vesícula biliar al intestino delgado. En ocasiones, algunos adenocarcinomas pancreáticos reducen la entrada de



estos conductos y obstaculizan el movimiento de la bilis, lo que genera ictericia. Se introduce un endoscopio desde la boca, pasando por el esófago y el estómago hasta la primera zona del intestino delgado. A continuación, se incorpora un catéter a través del endoscopio, hasta los conductos pancreáticos. Se suministra un contraste a través del catéter hacia los conductos tomándose una radiografía. No obstante, en la mayoría de tipos de cáncer, solo una biopsia permite formular un diagnóstico concluyente [8]. La biopsia se basa en la extirpación de células o tejidos para que un patólogo las pueda analizar al microscopio y confirmar si hay evidencias de cáncer. Hay distintos modos de efectuar una biopsia para cáncer de páncreas. Se puede colocar una aguja fina o una aguja ancha en el páncreas durante una ecografía o una radiografía para separar células.

En los últimos años, los avances en las tecnologías de secuenciación NGS han permitido ir más allá en la comprensión del cáncer de páncreas y las lesiones precursoras que dan lugar a estos tipos de cáncer [9]. Es importante resaltar que en la actualidad las técnicas NGS suponen una oportunidad significativa para la detección temprana y el tratamiento entre la primera alteración genética en una célula del páncreas y el desarrollo del cáncer en estado avanzado.

MÉTODOS DE SECUENCIACIÓN DE NUEVA GENERACIÓN (NGS)

Las plataformas de secuenciación NGS permiten una amplia variedad de procedimientos, lo que posibilita a los investigadores estudiar el genoma, el transcriptoma o el epigenoma de cualquier organismo. Los métodos de secuenciación difieren principalmente en cómo se obtienen las muestras de DNA o RNA (por ejemplo, el organismo, el tipo tisular o las condiciones experimentales) y por las opciones de análisis de datos utilizados. Después de que se hayan preparado las librerías de secuenciación, la etapa de secuenciación sigue siendo fundamental con independencia del método. Existen una serie de kits de preparación de librerías estándar que ofrecen protocolos para la secuenciación del genoma completo, mRNA-Seq, secuenciación dirigida (como la secuenciación del exoma o la secuenciación 16S), regiones seleccionadas a medida y regiones de unión a proteínas. Aunque el número de procedimientos NGS está en crecimiento constante, a continuación se presenta una breve descripción de los métodos más comunes.

GENÓMICA

Secuenciación completa del genoma

Los estudios de asociación del genoma completo basados en microarrays (GWAS) han sido el método más común para la identificación de enfermedades asociadas a lo largo de todo el genoma. Aunque los microarrays GWAS pueden cuestionar más de cuatro millones de marcadores por muestra, el método más completo de examinar los 3,2 mil millones de bases del genoma humano es la secuenciación completa del genoma (WGS). La caída rápida del precio de la secuenciación y la capacidad de WGS para producir grandes volúmenes de datos rápidamente hacen que sea una herramienta eficaz para la investigación genómica [10]. Aunque WGS se asocia comúnmente con la secuenciación de genomas humanos, la escalabilidad y la naturaleza flexible de la tecnología hace que sea igualmente útil para la secuenciación de cualquier especie, como genomas de plantas o enfermedades relacionadas con genomas microbianos.

Secuenciación de exoma

Posiblemente el método de secuenciación dirigida más utilizado sea la secuenciación del exoma. El exoma representa menos del 2 % del genoma humano, pero contiene la mayoría de las variantes que causan enfermedades conocidas, por lo que la secuenciación completa del exoma es una alternativa rentable a la secuenciación del genoma completo. Con la secuenciación del exoma, la fracción codificante de proteínas del genoma se captura y se secuencia de manera selectiva. Se pueden identificar eficientemente las variantes mediante una amplia gama de aplicaciones, incluyendo la genética poblacional, las enfermedades genéticas y los estudios sobre cáncer.



Secuenciación de novo

La secuenciación de novo hace referencia a la secuenciación de un genoma nuevo donde no hay secuencia de referencia disponible para el alineamiento. Las lecturas de las secuencias se ensamblan como contigs y la calidad de la cobertura de los datos de novo de la secuencia depende del tamaño y la continuidad de los contigs (es decir, el número de gaps en los datos). Otro elemento significativo en la generación de secuencias de novo de alta calidad es la diversidad de los tamaños de inserto de la librería. La combinación de secuencias de insertos cortos paired-end y de insertos largos mate pair es el método más potente para una cobertura máxima en todo el genoma. Las combinaciones de los tamaños de insertos permiten la detección de los tipos de variantes estructurales y son esenciales para identificar con precisión los reordenamientos más complejos. Las lecturas de insertos cortos se secuencian a profundidades mayores y pueden llenar los gaps no cubiertos por los insertos largos, que a menudo se secuencian a profundidades de lectura más bajas. Por lo tanto, los resultados del uso de un método combinado producen ensamblajes de mayor calidad. Paralelamente a las mejoras tecnológicas NGS, han surgido muchos avances algorítmicos en los ensambladores de secuencias para datos de lecturas cortas. Los investigadores pueden realizar ensamblajes de novo de alta calidad utilizando lecturas NGS y herramientas de ensamblajes de lecturas cortas disponibles públicamente. En muchas ocasiones, los recursos informáticos existentes en el laboratorio son suficientes para realizar ensamblajes de novo.

Secuenciación dirigida

Con la secuenciación dirigida, se aíslan y se secuencian un subconjunto de genes o regiones del genoma. La secuenciación dirigida permite a los científicos ahorrar tiempo, gastos, análisis de datos en áreas específicas de interés y la secuenciación con una cobertura mucho más alta. Por ejemplo, un estudio típico WGS alcanza niveles de cobertura de 30X-50X por genoma, mientras que un proyecto de resecuenciación dirigida puede cubrir fácilmente la región diana a 500X-1000X o superior. Esto permite a los investigadores una mayor cobertura para identificar variantes raras o variantes que serían demasiado raras y caras de identificar con WGS o por secuenciación basada en electroforesis por capilaridad. Los paneles de secuenciación dirigida se pueden comprar, preseleccionar el contenido o pueden ser diseñados a medida. Hay disponibles una amplia variedad de kits de preparación de librerías para secuenciación dirigida, incluyendo kits con conjuntos de sondas centradas en áreas específicas de interés como cáncer, cardiomiopatía, autismo o conjuntos de sondas personalizadas. Con diseños personalizados, los investigadores pueden dirigirse a las regiones del genoma relevantes para sus intereses específicos de investigación.

Transcriptómica

Los métodos de preparación de la librería para la secuenciación del RNA (RNA-Seq) habitualmente empiezan con la preparación de la muestra total de RNA seguido de una etapa de eliminación del ribosoma. La muestra total de RNA se convierte en cDNA antes de la preparación de la librería NGS estándar. RNA-Seq se centra en mRNA, RNA pequeño, RNA no codificante o miRNAs, y se realiza mediante la inclusión de etapas adicionales de aislamiento o enriquecimiento antes de la síntesis de cDNA.

Secuenciación de mRNA y RNA total

La secuenciación del transcriptoma es un avance importante en el estudio de la expresión génica, ya que permite una imagen de todo el transcriptoma en vez de un subconjunto predeterminado de genes. La secuenciación completa del transcriptoma ofrece una visión completa de un perfil transcripcional celular en un momento biológico dado y mejora en gran medida el poder de los métodos de detección del RNA. Como ocurre con cualquier método de secuenciación,



un rango dinámico prácticamente ilimitado permite la identificación y cuantificación de transcritos comunes y raros. Las competencias adicionales incluyen alineamientos de las lecturas de secuenciación mediante uniones de splicing, así como la detección de isoformas, nuevos transcritos y fusiones génicas. Los kits de preparación de la librería están disponibles tanto para RNA-Seq total como para métodos de RNA-Seq.

Secuenciación de RNA dirigida

La secuenciación de RNA dirigida es un método para medir los transcritos de interés de la expresión diferencial, la expresión específica de alelo, así como la detección de fusiones génicas, isoformas, cSNPs y uniones de splicing. Es un método eficaz para la investigación sobre las vías específicas de interés o para la validación de microarrays de expresión génica o resultados de secuenciación completa del transcriptoma.

Secuenciación de RNAs pequeños y RNAs no codificantes

Los RNAs pequeños, RNAs no codificantes o miRNAs son nucleótidos de 18-22 pb que juegan un papel en la regulación de la expresión génica de genes a menudo como represores o silenciadores. El estudio de los miRNAs ha crecido ya que se ha hecho evidente su papel en la regulación de la transcripción y la traducción [11, 12].

Epigenómica

Si la genómica implica el estudio de las alteraciones hereditarias o adquiridas en la secuencia del DNA, la epigenética es el estudio de los cambios hereditarios en la actividad de los genes causados por mecanismos distintos a cambios en la secuencia del DNA. Los mecanismos de actividad epigenética incluyen la metilación del DNA, la regulación mediada por RNA pequeño, las interacciones DNA-proteína, la modificación de las histonas y mucho más.

Secuenciación de la metilación

Un enfoque crítico en la epigenética es el estudio de los estados de metilación de la citosina (5-mC) mediante áreas específicas de la regulación, como la heterocromatina o los promotores. La metilación de la citosina puede modificar significativamente la expresión génica temporal y espacial; y remodelar la cromatina. Aunque hay muchos métodos para el estudio de la metilación génica, la secuenciación de la metilación aprovecha las ventajas de la tecnología NGS y el análisis del genoma completo mientras que la evaluación de los estados de metilación es a nivel de un solo nucleótido. Se emplean ampliamente dos métodos de secuenciación de la metilación: la secuenciación con bisulfito de genoma completo (WGBS) y la representación reducida de la secuenciación con bisulfito (RRBS). Con WGBS-Seq, la química del bisulfito de sodio convierte las citosinas no metiladas en uracilos, que luego se convierten en timinas en las lecturas de la secuencia o en la salida de los datos. En RRBS-Seq, el DNA es digerido con MspI; una enzima de restricción afectada por la metilación. Los fragmentos con un rango de tamaño de 100-150 pb son aislados para enriquecer las islas CpG y las regiones de DNA que contienen el promotor. Las librerías de secuenciación se construyen utilizando los protocolos NGS estándar.

Secuenciación ChIP

Las interacciones proteína-DNA o proteína-RNA tienen un impacto significativo en muchos procesos biológicos y etapas de enfermedades. Estas interacciones pueden ser investigadas con NGS mediante la combinación de los ensayos de inmunoprecipitación de la cromatina (ChIP) y métodos NGS. Los protocolos ChIP-Seq comienzan con la etapa de inmunoprecipitación de la cromatina (los protocolos ChIP-Seq varían ampliamente, ya que deben ser específicos para la especie, el tipo tisular y las condiciones experimentales).



Perfiles ribosomales

Los perfiles ribosomales son un método basado en deep sequencing de fragmentos de mRNA de ribosomas protegidos. La purificación y la secuenciación de estos fragmentos proporcionan una imagen de todos los ribosomas activos en una célula en un tiempo determinado. Esta información puede establecer qué proteínas están siendo traducidas activamente en una célula y puede ser útil para la investigación del control de la traducción, la medición de la expresión génica, la determinación de la tasa de síntesis proteica o la predicción de la abundancia de proteínas. Los perfiles ribosomales permiten la monitorización sistemática de los procesos de la traducción celular y la predicción de la abundancia de proteínas. La determinación de las regiones de un transcrito que están siendo traducidas puede ayudar a definir el proteoma de organismos complejos. Con NGS, los perfiles ribosomales permiten un análisis detallado y preciso de la producción de proteínas in vivo.

Secuenciación de una sola molécula en tiempo real (SMRT)

En un intento de superar los desafíos inherentes en el campo de la genómica, se ha desarrollado una nueva tecnología que ha impulsado los límites de la secuenciación. El resultado ha sido la secuenciación SMRT, que aprovecha el proceso natural de la replicación del DNA y permite la observación en tiempo real de la síntesis del DNA. Con esta tecnología exclusiva, se ha dotado y se han entregado a los científicos los resultados necesarios para impulsar el descubrimiento genético. La secuenciación SMRT se basa en dos innovaciones principales: la cámara zeromode waveguide (ZMWs) y los nucleótidos fluorescentes [13]. La ZMWs permite que la luz ilumine solo el fondo del pocillo en el que se inmoviliza el complejo DNA polimerasa/molde. Los nucleótidos fluorescentes permiten la observación del complejo inmovilizado mientras la DNA polimerasa produce una hebra totalmente natural de DNA (figura 1).

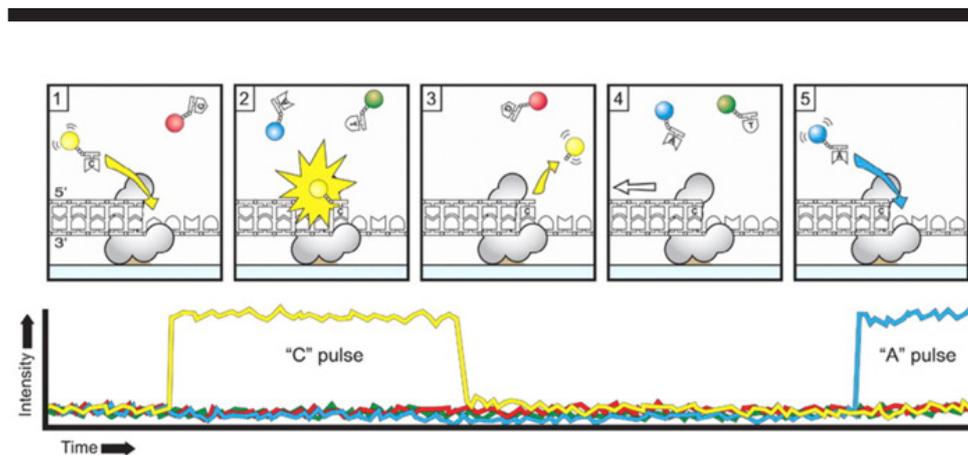


Figura 1. Cada uno de los cuatro nucleótidos se marca con un fluoróforo diferente (indicado en rojo, amarillo, verde y azul, respectivamente, para G, C, T y A), de modo que tienen espectros de emisión diferentes. Como un nucleótido se retiene en el volumen de detección por la polimerasa, un pulso de luz provoca que se identifique la base. (1) Un nucleótido marcado con fluorescencia se asocia con el molde en el sitio activo de la polimerasa. (2) La producción de la fluorescencia del color correspondiente a la base incorporada (amarillo para la base de C como se muestra aquí) se eleva. (3) El producto colorante-conector-pirofosfato se escinde del nucleótido y difunde fuera de la ZMW, y termina el impulso de la fluorescencia. (4) La polimerasa avanza a la siguiente posición. (5) La siguiente asociación de nucleótidos con el molde en el sitio activo de la polimerasa inicia el siguiente impulso de fluorescencia, que aquí corresponde a la base A [14].



La secuenciación SMRT es ideal para una gran variedad de aplicaciones en investigación y ofrece muchas prestaciones, como: longitud media de lectura más larga, máxima precisión de consenso, cobertura uniforme, caracterización epigenética simultánea y resolución de una sola molécula.

LONGITUD MEDIA DE LECTURA MÁS LARGA

Como el campo de la genómica evoluciona, existe una conciencia creciente en la comunidad científica de la importancia de los datos de lectura larga. Las lecturas de secuencias largas mejoran el mapeo de resecuenciación y simplifican el ensamblaje de novo. Los sistemas PacBio permiten secuenciar directamente el DNA y logran lecturas largas de secuenciación con una cobertura uniforme. La tecnología de secuenciación SMRT produce las longitudes medias de lectura más largas disponibles en la industria (promedio > 10.000 pb; lecturas > 60.000 pb). Estas lecturas más largas permiten: ensamblar genomas de novo de alta calidad, elaborar una lista de isoformas completa, alinear las secuencias sin ambigüedades, observar los alelos en fase, abarcar elementos repetitivos, regiones complejas y resolver variantes estructurales.

MÁXIMA PRECISIÓN DE CONSENSO

La secuenciación SMRT ofrece resultados excepcionales de secuenciación, con una precisión de consenso superior al 99,999 % (Q50). Con esta extraordinaria precisión, el sesgo de secuenciación baja y el mapeo de las lecturas de secuenciación aumentan. SMRT proporciona la información necesaria para calcular y detectar las variantes con confianza.

COBERTURA UNIFORME

La mayoría de los sistemas de secuenciación presentan sesgo, lo que provoca dificultades en la secuenciación de regiones de DNA ricas en AT o en GC, secuencias altamente repetitivas, homopolímeros largos y secuencias palindrómicas. Estas secuencias conducen a menudo a resultados incompletos, a veces falta hasta un 15 % del genoma. La secuenciación SMRT no requiere una etapa de amplificación, lo que conduce a una cobertura uniforme a través de todo el genoma. Esto permite secuenciar palíndromos y regiones de baja diversidad del genoma. Estas lecturas largas también permiten la división de las regiones complejas.

CARACTERIZACIÓN EPIGENÉTICA SIMULTÁNEA

Dado el papel central de la epigenética en una amplia variedad de procesos naturales y de enfermedades, la capacidad para detectar estos y otros cambios es clave para una comprensión más completa de la complejidad biológica. Con los softwares y con la secuenciación SMRT, los investigadores pueden detectar automáticamente las modificaciones del DNA y medir la velocidad de incorporación de bases de DNA durante la secuenciación. La caracterización epigenética aporta profundidad a la investigación y al análisis genético, ofreciendo la posibilidad de encontrar los motivos de reconocimiento de la metiltransferasa a lo largo de un genoma bacteriano; identificar las ubicaciones de metilación de la adenina y de la citosina; detectar secuencias integradas y modificaciones químicas; resolver modificaciones específicas de cadena, y llevar a cabo la detección sin hipótesis.

RESOLUCIÓN DE UNA SOLA MOLÉCULA

La distinción de moléculas de DNA estrechamente relacionadas en la misma muestra requiere la resolución de una sola molécula con alta precisión y con longitudes de lectura largas. Para estas aplicaciones se ha desarrollado un modelo



que genera la secuencia consenso intramolecular denominado secuenciación consenso circular (CCS). La secuenciación SMRT es la única tecnología que genera este nivel de precisión consenso de una sola molécula con longitudes de lectura de varias kilobases y se ha utilizado para diversas aplicaciones de poblaciones complejas. Con este método de secuenciación avanzado se tiene la capacidad de identificar variantes somáticas con precisión y resolver comunidades complejas.

APLICACIONES NGS AL CÁNCER DE PÁNCREAS

Las tecnologías NGS han permitido la detección eficiente y precisa de nuevas mutaciones somáticas. NGS se ha empleado con éxito para identificar nuevas mutaciones en una gran variedad de cánceres, incluyendo el cáncer de páncreas. Mediante la secuenciación de genoma completo (o de exoma completo), se han encontrado numerosas aberraciones genéticas y dianas terapéuticas potenciales. Estos hallazgos han proporcionado información clave en los mecanismos de tumorigénesis [15]. Aproximadamente del 5 al 10% de los cánceres son hereditarios. En la actualidad, el método más utilizado en análisis genéticos es la secuenciación Sanger, considerándose el estándar para la detección de mutaciones. Sin embargo, debido a que los genes relacionados con cánceres hereditarios son de gran tamaño y no hay una localización para una mutación en particular, este método tradicional en análisis genéticos de cáncer hereditario ha demostrado requerir mucho tiempo, alto coste y bajo rendimiento. El desarrollo NGS proporciona muchas oportunidades en análisis genéticos. Esto es debido a que permite realizar análisis de múltiples genes a la vez, mejorando en gran medida la tasa de detección de variaciones. Muchos pacientes con cáncer hereditario han dado negativo en variaciones genéticas, pero con NGS se han encontrado mutaciones causales.

Además de la identificación de mutaciones somáticas genéticas, otro uso NGS es la medicina personalizada. Hasta la fecha, muchos estudios han aplicado NGS al tratamiento personalizado del cáncer [16]. Con la mejora de la tecnología NGS, es posible disponer de un perfil molecular global para un paciente en un periodo de tiempo y con un coste clínicamente aceptable. La aplicación NGS en oncología clínica está progresando, pero se necesitan más ensayos clínicos de investigación y de validación para poner en práctica esta estrategia en el tratamiento personalizado del cáncer. Las mutaciones inusuales en DNA circulante han sido utilizadas para detectar las mutaciones somáticas en el diagnóstico y el tratamiento del cáncer. Hay muchos métodos disponibles para detectar estas mutaciones, como el ácido nucleico peptídico, la extensión del primer y perlas basadas en PCR en emulsión. Sin embargo, es difícil identificar mutaciones poco frecuentes en los genes supresores de tumores como TP53, el cual está altamente mutado a lo largo de todo el gen. NGS puede ser un método rentable para detectar y medir la frecuencia alélica y otras mutaciones de genes tumorales en plasma. Forshew et al. [17] desarrollaron taggedamplicon deep sequencing (TAmSeq), método que utiliza NGS y primers diseñados para amplificar aproximadamente 6.000 bases que cubren las regiones seleccionadas de los genes relacionados con cáncer, incluyendo EGFR, KRAS y TP53. Mediante el uso de muestras de plasma, demostraron que el método es capaz de identificar mutaciones en frecuencias alélicas del 2 al 65 %, evidenciando así que es viable secuenciar grandes regiones de DNA circulante por NGS.

CONCLUSIÓN

Como se ha descrito en el diagnóstico de cáncer de páncreas, la CPRE y la biopsia son procedimientos invasivos. Entre los riesgos más destacados de la CPRE [18] se incluyen la infección, la pancreatitis, las reacciones alérgicas a los sedantes, la hemorragia o sangrado excesivo y la perforación del tubo digestivo o de los conductos. Algunas complicaciones pueden ser lo suficientemente importantes como para requerir un tratamiento urgente e incluso una operación. En cuanto a los inconvenientes de la biopsia [8] encontramos la hemorragia, sangrado en el lugar de la biopsia e infección. Debido a la gravedad de todas estas complicaciones la finalidad de este trabajo es proponer los métodos de secuenciación masiva para el diagnóstico del cáncer de páncreas.

Con la técnica SMRT se utiliza el desarrollo natural de la replicación del DNA y se posibilita la observación en tiempo real de la síntesis del DNA. Esta tecnología proporciona una longitud media de lectura más larga, mejorando



el mapeo de resecuenciación, y simplifica el ensamblaje de novo. Una precisión máxima de consenso que permite reducir el sesgo de secuenciación y aumenta el mapeo de las lecturas. Por otra parte, ofrece una cobertura uniforme para secuenciar palíndromos y regiones de baja complejidad del genoma. Por último, facilita una caracterización epigenética simultánea, dado el papel de la epigenética en gran diversidad de procesos naturales y de enfermedades como el cáncer; detectar estos cambios es esencial para una percepción completa de la variedad biológica.

La aparición y la aplicación de métodos de secuenciación masiva han ampliado nuestro conocimiento del genoma del cáncer en sus diversas manifestaciones y han permitido el análisis integrado de datos a gran escala para proporcionar una imagen completa de la enfermedad. En la práctica, una explicación completa de las diferencias somáticas en el genoma y en el transcriptoma del tumor para cada paciente con cáncer de páncreas, junto con la adecuada interpretación en un contexto médico, podría ayudar a informar al oncólogo del paciente sobre nuevas vías terapéuticas [19]. En definitiva, estos avances crean esperanza a los pacientes que actualmente tienen un mal pronóstico y reducen la carga de enfermedad en un entorno en el que la cirugía es el único mecanismo en la actualidad para prolongar la supervivencia.

AGRADECIMIENTOS

Agradecemos a Antonio Pineda la ayuda proporcionada para la preparación del presente trabajo.

BIBLIOGRAFÍA

- [1] ¿Qué es el cáncer de páncreas? [Internet]. Cancer.org. 2016. Disponible en: <http://www.cancer.org/espanol/cancer/cancerdepancreas/guiadetallada/cancer-de-pancreas-what-is-what-is-pancreatic-cancer>
- [2] Reference G. KRAS [Internet]. Genetics Home Reference. 2016. Disponible en: <https://ghr.nlm.nih.gov/gene/KRAS#resources>
- [3] Genecards.org. 2016. Disponible en: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=CDKN2A>
- [4] Genecards.org. 2016. Disponible en: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=TP53>
- [5] Genecards.org. 2016. Disponible en: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=SMAD4>
- [6] Pruebas para detectar el cáncer de páncreas [Internet]. Cancer.org. 2016. Disponible en: <http://www.cancer.org/espanol/cancer/cancerdepancreas/guiadetallada/cancer-de-pancreas-early-diagnosis>
- [7] Protocolo de actuación en el cáncer de páncreas. Hospital Universitario Reina Sofía.
- [8] Versión en inglés revisada por Subodh K. Lal a. Biopsia del tracto biliar: MedlinePlus enciclopedia médica [Internet]. Nlm.nih.gov. 2016. Disponible en: <https://www.nlm.nih.gov/medlineplus/spanish/ency/article/003894.htm>
- [9] Pancreatic adenocarcinoma pathology: changing “landscape”. *Journal of Gastrointestinal Oncology*. 2015; 6(4): 358-374.
- [10] Grad YH, Lipsitch M, Feldgarden M, *et al*. Genomic epidemiology of the Escherichia coli O104:H4 outbreaks in Europe, 2011. *PNAS*. 2012; 109: 3065-3070.
- [11] Wang Y, Kim S, Kim IM. Regulation of metastasis by microRNAs in ovarian cancer. *Front Oncol*. 2014; 10: 143.
- [12] Dior Up, Kogan L, Chill HH, Eizenberg N, Simon A. Emerging roles of microRNA in the embryo-endometrium cross talk. *Semin Reprod Med*. 2014; 32: 402-409.
- [13] Revolutionize Genomics with SMRT Sequencing [Internet]. 1st ed. 2016. Disponible en: <http://www.pacb.com/smrt-science/smrt-sequencing/>
- [14] Rhoads Au K. PacBio Sequencing and Its Applications. *Genomics, Proteomics & Bioinformatics*. 2015; 13(5): 278-289.
- [15] Tran B, Dancy J, Kamel-Reid S, McPherson J, Bedard P, Brown A *et al*. Cancer Genomics: Technology, Discovery, and Translation. *Journal of Clinical Oncology*. 2012; 30(6): 647-660.
- [16] Guan Y, Li G, Wang R, Yi Y, Yang L, Jiang D *et al*. Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. *Chin J Cancer*. 2012; 31(10): 463-470.
- [17] Forsheo T, Murtaza M, Parkinson C, Gale D, Tsui D, Kaper F *et al*. Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. *Science Translational Medicine*. 2012; 4(136):136ra68-136ra68.



-
- [18] Colangiopancreatografía retrógrada endoscópica [Internet]. Niddk.nih.gov. 2016. Disponible en: <http://www.niddk.nih.gov/health-information/informacion-de-la-salud/Pruebas-de-diagnostico/colangiopancreatografia-retrograda-endoscopica-CPRE/Pages/diagnostic-test.aspx#9>
- [19] Mardis E. Applying next-generation sequencing to pancreatic cancer treatment. *Nature Reviews Gastroenterology & Hepatology*. 2012; 9(8): 477-486.

