



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



⑪ Número de publicación: **2 998 543**

⑫ Int. Cl.:  
**C12Q 1/68**  
(2008.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑬ Fecha de presentación y número de la solicitud internacional: **14.04.2017 PCT/EP2017/059075**

⑭ Fecha y número de publicación internacional: **19.10.2017 WO17178655**

⑮ Fecha de presentación y número de la solicitud europea: **14.04.2017 E 17718883 (6)**

⑯ Fecha y número de publicación de la concesión europea: **09.10.2024 EP 3443115**

⑭ Título: **Método y kit para la generación de bibliotecas de ADN para secuenciación masivamente paralela**

⑬ Prioridad:  
**15.04.2016 IT UA20162640**

⑮ Fecha de publicación y mención en BOPI de la traducción de la patente:  
**20.02.2025**

⑬ Titular/es:  
**MENARINI SILICON BIOSYSTEMS S.P.A.  
(100.00%)  
Via Giuseppe di Vittorio, 21 B/3  
40013 Castel Maggiore, IT**

⑭ Inventor/es:  
**MANARESI, NICOLÒ;  
BUSON, GENNY y  
TONONI, PAOLA**

⑭ Agente/Representante:  
**CARVAJAL Y URQUIJO, Isabel**

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

**ES 2 998 543 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Método y kit para la generación de bibliotecas de ADN para secuenciación masivamente paralela

5 Campo técnico de la invención

La presente invención se refiere a un método para secuenciación de genoma completo de paso bajo y perfilado de número de copias de genoma completo.

10 Antecedentes

Con células individuales es útil llevar a cabo una Amplificación del Genoma Completo (WGA) para obtener más ADN con el fin de simplificar y/o hacer posible llevar a cabo diferentes tipos de análisis genéticos, incluyendo secuenciación, detección de SNP, etc.

15 La WGA con una LM-PCR basada en un sitio de restricción determinista (como se describe, por ejemplo, en WO/2000/017390) se conoce de la técnica (en la presente denominada simplemente DRS-WGA). Se ha demostrado que DRS-WGA es una mejor solución para la amplificación de células individuales (Ref: Lee YS, et al: *Comparison of whole genome amplification methods for further quantitative analysis with microarray-based comparative genomic hybridization*. Taiwan J Obstet Gynecol. 2008, 47(1):32-41.) y también más resistente a la degradación del ADN debido a la fijación (ref. Stoecklein N.H. et al: *SCOMP is Superior to Degenerated Oligonucleotide Primed-PCR for Global Amplification of Minute Amounts of DNA from Microdissected Archival Samples*. American Journal of Pathology 2002, Vol. 161, núm. 1).

20 25 Se ha utilizado un kit comercial de DRS-WGA basado en LM-PCR (*kit Ampli1<sup>MR</sup> WGA*, Silicon Biosystems) en Hodgkinson C.L. et al., *Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer*, Nature Medicine 20, 897-903 (2014). En este trabajo, se realizó un análisis de número de copias mediante secuenciación de genoma completo de paso bajo en material de WGA unicelular. Sin embargo, para el flujo de trabajo estándar utilizado en este documento, la creación de bibliotecas Illumina requirió varios pasos, que incluyeron i) digestión de adaptadores de WGA, ii) fragmentación de ADN y pasos de flujo de trabajo Illumina estándar tal como iii) EndRepair, iv) Adición de cola de adenina v) ligación de adaptadores con código de barras, más los pasos habituales de vi) agrupamiento de muestras de bibliotecas NGS con código de barras y vii) secuenciación. Como se muestra en el artículo mencionado anteriormente (figura 5b), WBC presentó pocas llamadas de número de copia presumiblemente falsas positivas, aunque los CTC en general mostraron muchas más aberraciones.

30 35 40 *Ampli1<sup>MR</sup> WGA* es compatible con hibridación genómica comparativa en microarreglos (aCGH); de hecho, varios grupos (Moehlendick B, et al. (2013) *A Robust Method to Analyze Copy Number Alterations of Less than 100 kb in Single Cells Using Oligonucleotide Array CGH*. PLoS ONE 8(6): e67031; Czyz ZT, et al (2014) *Reliable Single Cell Array CGH for Clinical Samples*. PLoS ONE 9(1): e85907) demostraron que es adecuado para el análisis del número de copias de alta resolución. Sin embargo, la técnica aCGH es costosa y requiere mucha mano de obra, por lo que pueden ser deseables diferentes métodos, tal como la secuenciación de genoma completo de paso bajo (LPWGS) para la detección de alteraciones somáticas del número de copias (CNA).

45 Baslan et al (*Optimizing sparse sequencing of single cells for highly multiplex copy number profiling*, Genome Research, 25:1-11, 9 de abril de 2015), lograron el perfil de número de copias del genoma completo a partir de la amplificación del genoma completo por DOP-PCR, utilizando varios pasos enzimáticos, incluida la digestión de adaptador de WGA, la ligación de adaptadores Illumina, la amplificación por PCR.

50 Yan et al. Proc Natl Acad Sci EE.UU. 29 de diciembre de 2015; 112 (52): 15964-9, enseñan el uso de MALBAC WGA (Yikon Genomics Inc), para el diagnóstico genético preimplantacional simultáneo de anomalías cromosómicas y enfermedad monogénica.

55 60 US8206913B1 (Kamberov et al, Rubicon Genomics) enseña un enfoque donde se adopta una PCR de cebado de oligonucleótido degenerado especial (DOP-PCR). Esta referencia también contiene una descripción general de los diferentes métodos de WGA y el estado de la técnica. US8206913B1 está en la base del kit comercial PicoPlex.

Hou et al., *Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing*, GigaScience (2015) 4:37, reportan la comparación de rendimiento de varios métodos de WGA, que incluye MALBAC y amplificación por desplazamiento múltiple (MDA). LPWGS y WGS se utilizan en el documento. La preparación de la biblioteca se obtiene con flujos de trabajo.

65 Se ha demostrado que DRS-WGA es mejor que DOP-PCR para el análisis de perfiles de número de copias a partir de cantidades diminutas de material de FFPE microdisseccionado (Stoecklein et al., *SCOMP is superior to degenerated oligonucleotide primed-polymerase chain react ion for global amplification of minute amounts of DNA*

5 *from microdissected archival tissue samples*, Am J Pathol. 2002 Jul;161(1):43-51; Arneson et al., *Comparison of whole genome amplification methods for analysis of DNA extracted from microdissected early breast lesions in formalin-fixed paraffin-embedded tissue*, ISRN Oncol. 2012;2012:710692. doi: 10.5402/2012/710692. Epub 14 de marzo de 2012), cuando se usa CGH (hibridación genómica comparativa) en microarreglos, CGH en metafase, así como para otro ensayo de análisis genético tal como pérdida de heterocigosidad.

10 W02014068519 (Fontana et al.) enseña un método para detectar mutaciones de productos de DRS-WGA en loci donde la mutación introduce, elimina o altera un sitio de restricción.

15 10 W02015083121A1 (Klein et al.) enseña un método para evaluar la integridad de genoma de una célula y/o la calidad de un producto de DRS-WGA mediante una PCR múltiple, como se detalla e informa adicionalmente en Polzer et al. EMBO Mol Med. 30 de octubre de 2014; 6 (11): 1371-86.

20 15 Aunque la DRS-WGA proporciona los mejores resultados en términos de amplificación uniforme y equilibrada, los protocolos actuales basados en aCGH o CGH en metafase son laboriosos y/o costosos. Se ha propuesto la secuenciación de todo el genoma de paso bajo como un método de alto rendimiento para analizar varias muestras con mayor procesividad y menor costo que aCGH. Sin embargo, los métodos conocidos para la generación de una biblioteca de secuenciación masivamente paralela para productos WGA (tal como DRS-WGA) aún requieren protocolos que incluyen varios pasos y reacciones enzimáticas.

25 20 Más allá de la aplicación al análisis de CTC citado anteriormente, también para otras aplicaciones de análisis de células individuales, tal como el diagnóstico prenatal en blastocitos, así como para células fetales circulantes recolectadas de sangre materna, sería deseable tener un método más simplificado, que combine la reproducibilidad y la calidad de DRS-WGA con la capacidad de analizar variantes de número de copias (CNV) de genoma completo. Además, sería deseable determinar un perfil de número de copias de genoma completo también a partir de una cantidad diminuta de células, FFPE o biopsias de tejido.

30 25 WO 2014/071361 divulga un método para preparar una biblioteca para la secuenciación que comprende agregar oligos adaptadores de tallo bucle al ADN genómico fragmentado. Luego, los bucles se escinden, lo que da como resultado fragmentos de genoma flanqueados por adaptadores bicatenarios. Luego, los fragmentos se amplifican con cebadores que comprenden un código de barras y se utilizan para la secuenciación de ADN en una plataforma de secuenciación Ion Torrent.

35 30 Este método tiene una serie de inconvenientes, los más importantes de los cuales son:

- 35 - el método implica una serie de pasos posteriores que implican varias reacciones y varias enzimas;  
- el método no es aplicable como tal en el ADN derivado de una muestra unicelular.

40 40 Breve descripción de la invención

45 45 Un objeto de la presente invención es proporcionar un método para secuenciación de genoma completo de paso bajo que sea más simple y requiera menos pasos con respecto a los métodos de la técnica anterior. En particular, un objeto de la presente invención es proporcionar un método que incluya menos reacciones enzimáticas que las informadas generalmente en la literatura.

50 50 Otro objeto de la invención es proporcionar un método para generar un perfil de número de copias de todo el genoma usando el método para secuenciación de genoma completo de paso bajo de acuerdo con la invención.

55 55 Preferentemente, la biblioteca utilizada en el método de secuenciación de genoma completo de paso bajo debe ser compatible con una plataforma de secuenciación seleccionada, por ejemplo, plataforma Ion Torrent o plataforma Illumina.

60 60 La presente invención se refiere a un método para generar una biblioteca de secuenciación masivamente paralela para la secuenciación de genoma completo a partir de productos de amplificación de genoma completo como se define en las reivindicaciones adjuntas. La invención se refiere además a un método para generar un perfil de número de copias de todo el genoma usando el método para secuenciación de genoma completo de paso bajo de la invención.

65 65 También se proporcionan secuencias de cebadores y protocolos operativos.

70 70 Preferentemente, la reacción de generación de biblioteca comprende la introducción de un código de barras de secuenciación para multiplexar varias muestras en la misma corrida de NGS. Preferiblemente, la WGA es una DRS-WGA y la biblioteca se genera con una reacción de PCR de un solo tubo y un solo paso.

75 75 Breve descripción de los dibujos

La figura 1 muestra un producto de partida para que se utilice en una primera realización de la invención, que consiste en una biblioteca de ADN generada por DRS-WGA, de la cual se ilustra un fragmento individual de una manera puramente esquemática;

- 5 La figura 2 muestra un producto de partida para que se utilice en una segunda realización de la invención, que consiste en una biblioteca de ADN generada por MALBAC, de la cual se ilustra un fragmento individual de una manera puramente esquemática;
- 10 La figura 3 muestra de manera esquemática una realización del paso de reamplificación del método para generar una biblioteca de secuenciación masivamente paralela divulgada en la presente aplicada al fragmento de una biblioteca de ADN generada por DRS-WGA como se muestra en la figura 1 y dirigida a proporcionar una biblioteca de ADN compatible con una plataforma de secuenciación del tipo de la plataforma de secuenciación Ion Torrent o Illumina;
- 15 La figura 4 muestra de manera esquemática el flujo de trabajo del protocolo que incluye un paso de reacción de reamplificación aplicada al fragmento de un DRS-WGA como se muestra en la figura 1, y posteriormente una selección de biblioteca de fragmentos. Este método proporciona directamente una biblioteca de ADN compatible con la plataforma de secuenciación ILLUMINA;
- 20 La figura 5 muestra de manera esquemática la biblioteca de ADN monocatenario final obtenida de acuerdo con un método para generar una biblioteca de secuenciación masivamente paralela divulgada en la presente aplicada a un fragmento de DRS-WGA siguiendo los pasos mostrados en la figura 4; además, la figura 5 ilustra la biblioteca de ADNmc secuenciada final y los cebadores de secuenciación personalizados diseñados; a partir de unos cientos de células tumorales clasificadas digitalmente de FFPE con el sistema DEPArray (Bolognesi et al.) se genera una biblioteca DRS-WGA;
- 25 La figura 6 muestra los resultados de secuenciación de una secuenciación de genoma completo de paso bajo realizada a partir de unos cientos de células tumorales clasificadas digitalmente de FFPE con sistema DEPArray en una biblioteca de ADN preparada de acuerdo con el método divulgado y secuenciada por plataforma PGM;
- 30 La figura 7 muestra los resultados de secuenciación de la secuenciación de todo el genoma de paso bajo realizada por el protocolo PGM en bibliotecas de ADN preparadas en dos células tumorales diferentes;
- 35 La figura 8 muestra los resultados de secuenciación de una secuenciación de genoma completo de paso bajo realizada mediante un protocolo 1 ILLUMINA en bibliotecas de ADN preparadas de acuerdo con el método divulgado y compara los resultados obtenidos de una célula de WBC normal y una célula anormal (tumoral); y
- 40 La figura 9 muestra los resultados de secuenciación de una secuenciación de genoma completo de paso bajo realizada por un protocolo 2 ILLUMINA de acuerdo con un aspecto de la invención en bibliotecas de ADN preparadas de acuerdo con el método divulgado.

#### Descripción detallada

##### 45 Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en la presente tienen el mismo significado como se entiende comúnmente por un experto en la técnica a la que pertenece esta invención.

50 A menos que se mencione lo contrario, las técnicas descritas en la presente para su uso con la invención son metodologías estándar bien conocidas por personas expertas en la técnica.

55 Por el término "sitio de digestión (DS)" o "sitio de restricción (RS)" se entiende la secuencia de nucleótidos (típicamente 4-8 pares de bases (pb) de longitud) a lo largo de una molécula de ADN reconocida por la enzima de restricción en cuanto a dónde corta a lo largo de la cadena de polinucleótidos.

Por el término "sitio de escisión" se entiende el sitio en una cadena de polinucleótidos en cuanto a donde la enzima de restricción escinde nucleótidos mediante la hidrólisis del enlace fosfodiéster entre ellos.

60 Por el término "amplicón" se entiende una región de ADN producida por una amplificación por PCR.

Por el término "amplicón de DRS-WGA" o, en resumen, "amplicón de WGA" se entiende un fragmento de ADN amplificado durante DRS-WGA, que comprende una secuencia de ADN entre dos RS flanqueadas por los adaptadores ligados.

- Por el término "ADN original" se entiende el ADN genómico (ADNg) antes de la amplificación con la DRS-WGA.
- 5 Por el término "adaptador" o "adaptador de WGA" o "cebador de PCR de WGA" o "adaptador de secuencia universal de biblioteca de WGA" se entiende el oligonucleótido adicional ligado a cada fragmento generado por la acción de la enzima de restricción, en el caso de DRS-WGA, o la secuencia de polinucleótidos conocida presente en la sección 5' de cada molécula de la biblioteca de ADN de WGA como resultado de la extensión y el proceso de PCR, en el caso de MALBAC.
- 10 Por el término "alteración de número de copias (CNA)" se entiende un cambio somático en los números de copias de una región genómica, definida en general con respecto al mismo genoma individual.
- 15 Por el término "variación de número de copias (CNV)" se entiende una variante de la línea germinal en números de copias de una región genómica, definida en general con respecto a un genoma de referencia. A lo largo de la descripción, CNA y CNV se pueden usar indistintamente, ya que la mayoría del razonamiento se puede aplicar a ambas situaciones. Se debe pretender que cada uno de esos términos se refiera a ambas situaciones, a menos que se especifique lo contrario.
- 20 Por el término "secuenciación masiva-paralela de próxima generación (NGS)" se entiende un método de secuenciación de ADN que comprende la creación de una biblioteca de moléculas de ADN separadas espacialmente y/o en el tiempo, secuenciadas clonalmente (con o sin amplificación clonal previa). Ejemplos incluyen la plataforma Illumina (Illumina Inc), la plataforma Ion Torrent (ThermoFisher Scientific Inc), la plataforma Pacific Biosciences, MinION (Oxford Nanopore Technologies Ltd).
- 25 Por el término "secuencia diana" se entiende una región de interés en el ADN original.
- 30 Por el término "biblioteca de ADN de WGA primaria (pWGAlib)" se pretende una biblioteca de ADN obtenida a partir de una reacción de WGA.
- 35 Por el término "ciclos de amplificación basados en recocido y bucle múltiples (MALBAC)" se entiende un método de amplificación de genoma completo cuasilineal (Zong et al., Genome-wide detection of single-nucleotide and copy-number variations of a single human cell, *Science*. 21 de diciembre de 2012; 338(6114):1622-6. doi: 10.1126/science.1229164). Los cebadores de MALBAC tienen una secuencia aleatoria 3' de 8 nucleótidos, para hibridarse con el molde, y una secuencia común 5' de 27 nucleótidos (GTG AGT Gat GGT TGA GGT AGT GTG GAG). Después de la primera extensión, los semiamplícones se utilizan como moldes para otra extensión que produce un amplicón completo que tiene extremos 5' y 3' complementarios. Después de algunos ciclos de amplificación cuasi lineal, el amplicón completo se puede amplificar exponencialmente con ciclos de PCR posteriores.
- 40 Por el término "purificación de la biblioteca de ADN" se entiende un proceso por el cual el material de la biblioteca de ADN se separa de los componentes de reacción no deseados tales como enzimas, dNTP, sales y/u otras moléculas que no forman parte de la biblioteca de ADN deseada. Ejemplos de procesos de purificación de bibliotecas de ADN son la purificación con tecnología basada en perlas magnéticas tal como Agencourt AMPure XP o perlas de inmovilización reversible en fase sólida (SPRI) de Beckman Coulter o con purificación de columna de espín tal como columnas de espín Amicon de Merck Millipore.
- 45 Por el término "selección de tamaño de biblioteca de ADN" se entiende un proceso mediante el cual se altera la distribución de pares de bases de diferentes fragmentos que componen la biblioteca de ADN. En general, una porción de la biblioteca de ADN incluida en un determinado intervalo se conserva sustancialmente, mientras que los componentes de la biblioteca de ADN fuera de ese intervalo se descartan sustancialmente. Ejemplos de procesos de selección de tamaño de biblioteca de ADN son la escisión de geles electroforéticos (por ejemplo, ThermoFisher Scientific E-gel) o la doble purificación con un sistema de purificación basado en perlas magnéticas (por ejemplo, perlas Beckman Coulter SPRI).
- 55 Por el término "selección de biblioteca de ADN" se entiende un proceso por el cual se lleva a cabo la purificación de la biblioteca de ADN o la selección de tamaño de la biblioteca de ADN o ambas.
- 60 Por el término "reamplificación de NGS" se entiende una reacción de PCR en la que toda o una parte sustancial de la biblioteca de ADN de WGA primaria se amplifica adicionalmente. El término NGS se puede omitir por simplicidad en todo el texto, y se hará referencia simplemente a "reamplificación".
- 65 Por el término "adaptador de secuenciación (SA)" se entiende una o más moléculas que son instrumentales para secuenciar el inserto de ADN, cada molécula puede comprender ninguno, uno o más de los siguientes: una secuencia de polinucleótidos, un grupo funcional. En particular, se pretende una secuencia de polinucleótidos que se requiere que esté presente en una biblioteca de secuenciación masivamente paralela para que el secuenciador genere correctamente una secuencia de salida, pero que no lleve información, (como ejemplos no limitantes: una secuencia de polinucleótidos para hibridar un ADNmC a una celda de flujo, en caso de secuenciación Illumina, o a

una esfera de iones, en caso de secuenciación Ion Torrent, o una secuencia de polinucleótidos requerida para iniciar una reacción de secuenciación por síntesis).

5 Por el término "código de barras de secuenciación" se entiende una secuencia de polinucleótidos que, cuando se secuencia dentro de una lectura de secuenciador, permite que esa lectura se asigne a una muestra específica asociada con ese código de barras.

10 Por el término "funcional para una plataforma de secuenciación seleccionada" se entiende una secuencia de polinucleótidos que tiene que ser empleada por la plataforma de secuenciación durante el proceso de secuenciación (por ejemplo, un código de barras o un adaptador de secuenciación).

15 Por el término "secuenciación de genoma completo de paso bajo" se entiende una secuenciación de genoma completo a una profundidad de secuenciación media inferior a 1.

20 Por el término "profundidad de secuenciación media" se entiende aquí, en una base por muestra, el total de la cantidad de bases secuenciadas, mapeadas al genoma de referencia dividido por el tamaño total de genoma de referencia. El número total de bases secuenciadas y mapeadas se puede aproximar al número de lecturas mapeadas por la longitud de lectura promedio.

25 Por el término "ADN bicatenario (ADNb)" se entiende, de acuerdo con las reglas de emparejamiento de bases (A con T y C con G), dos cadenas complementarias de polinucleótidos separadas unidas por hidrógeno mediante la unión de las bases nitrogenadas de las dos. ADN monocatenario (ADNm): Las dos cadenas de ADN pueden formar dos moléculas de ADN monocatenarias, es decir, una molécula de ADN compuesta por dos moléculas de ADNm acopladas con emparejamiento de bases de Watson-Crick.

30 Por el término "ADN monocatenario (ADNm)" se entiende una cadena de polinucleótidos, por ejemplo, derivada de un ADN bicatenario o que se puede emparejar con un ADN monocatenario complementario, es decir, una molécula de ADN polinucleotídico que consiste en solo una cadena única contrariamente a las dos cadenas típicas de nucleótidos en forma helicoidal.

35 Por "igualar" se entiende el acto de ajustar la concentración de una o más muestras para hacerlas iguales.

40 Por "normalizar" se entiende el acto de ajustar la concentración de una o más muestras para que correspondan a una proporción deseada entre ellas (igualando es el caso especial donde la proporción es 1). En la descripción, por simplicidad, los términos normalización y ecualización se usarán indistintamente, ya que obviamente son conceptualmente idénticos.

45 Por "perlas paramagnéticas" se entiende perlas magnéticas conjugadas con estreptavidina (por ejemplo, Dynabeads® MyOne<sup>MR</sup> Streptavidin Cl, ThermoFisher Scientific). La expresión "condiciones diseñadas" cuando se refiere a la incubación de las perlas paramagnéticas se refiere a las condiciones requeridas para el paso de activación, que consiste en lavar las perlas magnéticas conjugadas con estreptavidina dos veces con el siguiente amortiguador: Tris-HCl 10 mM (pH 7,5), EDTA 1 mM, NaCl 2 M.

#### Flujos de trabajo

50 La siguiente tabla resume algunos posibles flujos de trabajo:

Tabla 1

Paso	wf1	wf2	wf3	wf4	wf5	wf6	wf7
Purificar/seleccionar tamaño	0	0	0	0	0	0	X
Reamp de NGS	SA	BC+SA	BC+SA	SA	BC+SA	BC	BC+SA
Purificar			X				X
Cuantificar			X				X
Agrupar		X	X				X
Selección de tamaño	X	X	X				
Purificar	0	0	0	0	0	0	
Secuencia	X	X	X	X	X	X	X

50 Leyenda: 0= paso opcional, SA=introducción de adaptadores de secuenciación, BC=introducción de códigos de barras, X= paso necesario, wf=flujo de trabajo

## Material de entrada de proceso

5 Toda la presente descripción se refiere a una biblioteca de ADN de WGA primaria. Los mismos flujos de trabajo se  
 pueden aplicar a la biblioteca de ADN de WGA primaria que se sometió adicionalmente a procesos adicionales,  
 tales como, por ejemplo, síntesis de ADNbc o reamplificación de biblioteca con cebadores de WGA estándar (por  
 ejemplo, como sea posible con el kit *Ampli1<sup>MR</sup> ReAmp/ds*, Menarini Silicon Biosystems spa, Italia). Por simplicidad,  
 nos referimos aquí solo a las bibliotecas primarias de ADN de WGA, sin tener en cuenta esos procesos adicionales.  
 10 Se debe pretender que todo ese tipo de muestras de entrada se puedan usar como entrada de muestra adecuada,  
 también para lo que se informa en la reivindicación.

## Purificación inicial

15 Cuando están presentes cantidades no despreciables de cebadores de WGA primarios en el producto de salida de  
 WGA primario, puede ser ventajoso tener una selección de biblioteca de ADN inicial que incluya una purificación  
 de biblioteca de ADN. De hecho, dado que los cebadores incluyen, en el extremo 3', una secuencia correspondiente  
 20 a la secuencia común encontrada en los cebadores WGA primarios, la presencia de cantidades no despreciables  
 de cebadores WGA primarios residuales puede competir con los cebadores de reamplificación utilizados para  
 obtener la biblioteca de secuenciación masivamente paralela, disminuyendo el rendimiento de moléculas de  
 biblioteca de ADN que tienen, como se deseé, los cebadores de reamplificación (o sus complementarios inversos)  
 en ambos extremos.

## Cuantificación para ecualización del número de lecturas en todas las muestras

25 Cuando las variaciones en la cantidad de biblioteca de ADN reamplificada son relativamente grandes entre las  
 muestras a agrupar y secuenciar juntas, puede ser ventajoso cuantificar la cantidad de biblioteca de ADN de cada  
 muestra, con el fin de dividir en alícuotas esas bibliotecas e igualar el número de lecturas secuenciadas para cada  
 muestra.

30 Los apareamientos erróneos entre la longitud de lectura del secuenciador y el pico de tamaño de WGA puede dar  
 como resultado una ecualización imprecisa.

35 Varios secuenciadores masivamente paralelos (que incluyen plataformas Ion Torrent e Illumina) emplean la  
 secuenciación de fragmentos de ADN que tienen un pico de distribución de tamaño comprendido entre 50 y 800  
 pb, tal como, por ejemplo, aquellos que tienen un pico de distribución a 150 pb, 200 pb, 400 pb, 650 pb de acuerdo  
 con las diferentes composiciones químicas utilizadas. Como la distribución de tamaño de pWGAlib puede tener un  
 40 pico de fragmentos más grandes, tal como aproximadamente 1 kpb, y cantidades mucho más pequeñas de ADN  
 a 150 pb, 200 pb, 400 pb, la cuantificación de las cantidades de la biblioteca de ADN reamplificado en el rango  
 deseado puede ser imprecisa si se lleva a cabo en la biblioteca de ADN reamplificado a granel sin selección de  
 45 tamaño previa del intervalo de fragmentos deseado. Como resultado, la cuantificación de ADN en masa y la  
 ecualización de varias muestras en el agrupamiento pueden dar como resultado variaciones relativamente grandes  
 del número real de lecturas por muestra, ya que el número de fragmentos dentro del rango de tamaño del  
 secuenciador varía estocásticamente debido a la imprecisión en la distribución de fragmentos de ADN en la  
 50 biblioteca (por lo tanto, incluso las cantidades totales perfectamente ecualizadas de la biblioteca de ADN dan como  
 resultado variaciones significativas del número de fragmentos secuenciados).

## Aumentar la cantidad de biblioteca de ADN dentro de longitud de lectura de secuenciador para mejorar ecualización

50 [por selección de tamaño antes de reamplificación]

55 Cuando la distribución de tamaño de producto de WGA primario se debe alterar para aumentar la proporción de  
 cantidad de biblioteca de ADN dentro del intervalo de longitud de lectura del secuenciador con respecto a la  
 biblioteca de ADN total, puede ser ventajoso tener una selección de biblioteca de ADN inicial que comprenda una  
 selección de tamaño de biblioteca de ADN.

60 [por reamplificación preferencial]

65 De manera alternativa, o además, puede ser ventajoso llevar a cabo la reacción de reamplificación en condiciones  
 que favorezcan la amplificación preferencial de fragmentos de biblioteca de ADN en el intervalo deseado.

## Reamplificación preferencial por elección de polimerasa o acortamiento del ciclo de extensión

70 Las condiciones de reacción que favorecen fragmentos más cortos pueden comprender la reacción de PCR de  
 reamplificación con una polimerasa que amplifica preferentemente fragmentos más cortos, o ciclos de PCR  
 iniciales mediante los cuales una fase de extensión más corta evita que los fragmentos largos se repliquen en toda

su longitud, generando fragmentos de biblioteca incompletos. Los fragmentos de biblioteca incompletos carecen de la porción de extremo 3' complementaria inversa a la sección 3' de cebadores de reamplificación y, por lo tanto, excluyen el fragmento de pasos de replicación adicionales con los cebadores de reamplificación, interrumpiendo la amplificación exponencial del fragmento incompleto, consintiendo la generación de solo un número lineal (con 5 ciclos) de fragmentos de amplificación incompletos originados por los fragmentos de biblioteca de ADN de WGA primarios más largos.

Ejemplo de flujos de trabajo de acuerdo con la tabla 1

- 10 Wf1) se puede aplicar a LPWGS de una biblioteca de WGA en IonTorrent PGM (por ejemplo, en un chip 314, procesando una muestra individual que no requiere códigos de barras de muestra). La reamplificación con dos cebadores permite la introducción de los dos adaptadores de secuenciación, sin códigos de barras.
- 15 Wf2) se pueden aplicar a LPWGS de múltiples muestras de WGA primarias en Ion Torrent PGM o Illumina MiSeq, cuando las muestras de entrada originales para la WGA primaria se derivan de tipos homogéneos de material no amplificado, por ejemplo, células individuales, que se sometieron al mismo tratamiento (por ejemplo, fresco o fijo), no apoptótico. Por lo tanto, no es necesaria cuantificación, ya que el rendimiento de WGA primario es 20 aproximadamente el mismo en todos. Las bibliotecas adaptadas al secuenciador con código de barras se agrupan, luego se selecciona el tamaño para aislar los fragmentos con el tamaño apropiado dentro de la longitud de lectura del secuenciador, se purifican y se secuencian. Si la selección de tamaño se lleva a cabo mediante gel, se lleva a cabo una purificación posterior. Si la selección de tamaño se lleva a cabo, por ejemplo, con purificación de perlas SPRI de doble cara, la salida resultante ya está purificada y no son necesarios más pasos de purificación.
- 25 Wf3) se puede aplicar a LPWGS de múltiples muestras WGA primarias en Ion Torrent PGM o Illumina MiSeq donde las muestras de entrada originales para el WGA primario se derivan de tipos no homogéneos de material no amplificado. Por ejemplo, parte de células individuales, parte de agrupamientos de células, que se sometieron a diferentes tratamientos (por ejemplo, algunos frescos, algunos fijos), con diferente calidad de ADN original (algunos no apoptóticos, algunos apoptóticos, con índices de integridad del genoma heterogéneos; véase Polzer et al. EMBO Mol Med. 30 de octubre de 2014; 6 (11): 1371-86). Por lo tanto, la cuantificación es necesaria ya que el 30 rendimiento de WGA primario puede diferir significativamente entre las muestras. Con respecto a Wf2, se lleva a cabo una cuantificación. Antes de la cuantificación, es ventajoso purificar con el fin de hacer que el paso de cuantificación sea más confiable ya que, por ejemplo, los cebadores residuales y los dNTP o dímeros de cebadores se eliminan y no afectan la cuantificación.
- 35 Las bibliotecas adaptadas al secuenciador con código de barras se agrupan, luego se selecciona el tamaño para aislar los fragmentos con el tamaño apropiado dentro de la longitud de lectura del secuenciador, se purifican y se secuencian. Si la selección de tamaño se lleva a cabo mediante gel, se lleva a cabo una purificación posterior. Si la selección de tamaño se lleva a cabo, por ejemplo, con purificación de perlas SPRI de doble cara, la salida resultante ya está purificada y no son necesarios más pasos de purificación.
- 40 Wf4) se puede aplicar a la preparación de una biblioteca de secuenciación masivamente paralela para la secuenciación de nanoporos de Oxford. Dado que el nanoporo puede acomodar longitudes de lectura más largas, la selección de tamaño puede ser innecesaria, y la secuenciación se puede llevar a cabo en sustancialmente todas las longitudes de fragmento en la biblioteca.
- 45 Wf5) se puede aplicar a la preparación de múltiples bibliotecas de secuenciación masivamente paralela para la secuenciación con nanoporos de Oxford. Con respecto a wf4, los cebadores de reamplificación incluyen además un código de barras de muestra para multiplexar más muestras en la misma corrida. Dado que el nanoporo puede acomodar longitudes de lectura más largas, puede ser innecesaria la selección del tamaño.
- 50 Wf6) se puede aplicar a la preparación de múltiples bibliotecas de secuenciación masivamente paralela para un secuenciador Oxford con nanoporos que no requiere la adición de adaptadores de propósito especial. Con respecto a wf5, los cebadores de reamplificación no incluyen un adaptador de secuenciación, sino solo un código de barras de muestra para multiplexar más muestras en la misma corrida. Dado que el nanoporo puede acomodar longitudes de lectura más largas, puede ser innecesaria la selección del tamaño.
- 55 Wf7) se puede aplicar a la preparación de múltiples bibliotecas masivamente paralelas para la secuenciación de bibliotecas de ADN de DRS-WGA obtenidas de muestras no homogéneas después de tratamientos heterogéneos y que tienen una calidad de ADN diferente en un sistema de longitud de lectura más corto, tal como IonProton utilizando la composición química de secuenciación de 200 pb. Dado que la cantidad de biblioteca de ADN de WGA primaria de alrededor de 200 pb es muy pequeña en comparación con el ADN total en la biblioteca de ADN de WGA primaria, puede ser ventajoso llevar a cabo una selección de tamaño eliminando todos o sustancialmente todos los fragmentos de pWGAlib fuera de la longitud de lectura de secuenciación, enriqueciendo los fragmentos de pWGAlib de alrededor de 200 pb.

La reamplificación se lleva a cabo entonces con cebadores de reamplificación que incluyen códigos de barras y adaptadores de secuenciación compatibles con el sistema IonProton. Por lo tanto, el producto de reamplificación se purifica y cuantifica para cada muestra, y se agrupan diferentes alícuotas de diferentes muestras para igualar el número de lecturas para cada código de barras de muestra, y luego se secuencian para llevar a cabo LPWGS.

5 Para aquellos expertos en la técnica es evidente que son posibles diferentes combinaciones de los pasos incluidos en los flujos de trabajo como se mencionó anteriormente sin desviarse del alcance de la invención, que depende de la reamplificación de la biblioteca de ADN de WGA primaria con cebadores especiales como se divulga en la presente.

10 Preparación de bibliotecas de secuenciación masivamente paralela a partir de un producto de WGA

En la presente divulgación, se proporciona un método que comprende los pasos de

15 a. proporcionar una biblioteca de ADN de WGA primaria (pWGAlib) que incluye fragmentos que comprenden una sección de secuencia de WGA 5' (5SS) conocida, una sección de secuencia de WGA (MSS) media y una sección de secuencia de WGA 3' (3SS) conocida complementaria inversa a la sección de secuencia de WGA 5' conocida, la sección de secuencia de WGA 5' (5SS) conocida que comprende un adaptador de secuencia universal de biblioteca de WGA y la sección de secuencia de WGA media (MSS) que comprende al menos una sección de inserción (IS) correspondiente a una secuencia de ADN del ADN no amplificado original antes de WGA, la secuencia de WGA media comprende opcionalmente, además, una sección intermedia 5' flanqueante (F5) y/o una sección intermedia 3' flanqueante (F3);

20 b. reamplificar la biblioteca de ADN de WGA primaria usando al menos un primer cebador (1PR) y al menos un segundo cebador (2PR); en donde

25 el por lo menos un primer cebador (1PR) comprende una primera sección 5' de cebador(1PR5S) y una primera sección 3' de cebador (1PR3S), la primera sección 5' de cebador (1PR5S) que comprende por lo menos un primer adaptador de secuenciación (1PR5SA) y por lo menos un primer código de barras de secuenciación (1PR5BC) en la posición 3' del por lo menos un primer adaptador de secuenciación (1PR5SA) y en la posición 5' de la primera sección de cebador 3' (1PR3S), y la primera sección de cebador 3'(1PR3S) que se hibrida con la sección de secuencia 5' conocida (5SS) o la sección de secuencia 3' (3SS) conocida; el por lo menos un segundo cebador (2PR) comprende una segunda sección 5' de cebador (2PR5S) y una segunda sección 3' de cebador (2PR3S), la segunda sección 5' de cebador (2PR5S) que comprende por lo menos un segundo adaptador de secuenciación (2PR5SA) diferente del por lo menos un primer adaptador de secuenciación (1PR5SA), y la segunda sección 3' de cebador (2PR3S) que se hibrida con la sección de secuencia 5' (5SS) conocida o la sección de secuencia 3' (3SS) conocida.

30 40 La sección de secuencia 5' (5SS) conocida consiste preferentemente en un adaptador de secuencia universal de biblioteca de WGA. Como un ejemplo, DRS-WGA (tal como el kit Menarini Silicon Biosystems *Ampli1<sup>MR</sup>* WGA), así como MALBAC (Yikon Genomics), producen pWGAlib con una sección de secuencia 3' conocida complementaria inversa de la sección de secuencia 5' conocida como se solicita para la entrada del método divulgado en la presente.

45 45 Por lo tanto, el adaptador de secuencia universal de biblioteca de WGA es preferentemente un adaptador de secuencia universal de biblioteca de DRS-WGA (por ejemplo, SEQ ID NO: 282) o un adaptador de secuencia universal de biblioteca MALBAC (por ejemplo, SEQ ID NO: 283), más preferentemente un adaptador de secuencia universal de biblioteca DRS-WGA.

50 50 Preferentemente, el segundo cebador (2PR) comprende además al menos un segundo código de barras de secuenciación (2PR5BC), en la posición 3' del por lo menos un segundo adaptador de secuenciación (2PR5SA) y en la posición 5' de la segunda sección 3' de cebador (2PR3S).

55 55 Debido a la presencia de los códigos de barras de secuenciación, se lleva a cabo un método para secuenciación de genoma completo de paso bajo de acuerdo con una realización de la invención, que comprende los pasos de:

60 60 c. proporcionar una pluralidad de bibliotecas de secuenciación masivamente paralela con códigos de barras y agrupar muestras obtenidas utilizando diferentes códigos de barras de secuenciación (BC);

65 65 d. secuenciar la biblioteca agrupada.

El paso de agrupar muestras utilizando diferentes códigos de barras de secuenciación (BC) comprende además los pasos de:

70 70 e. cuantificar el ADN en cada una de las bibliotecas de secuenciación masivamente paralela con código de barras;

f. normalizar la cantidad de bibliotecas de secuenciación masivamente paralela con código de barras.

5 El paso de agrupar muestras utilizando diferentes códigos de barras de secuenciación (BC) comprende además el paso de seleccionar fragmentos de ADN comprendidos dentro de al menos un intervalo seleccionado de pares de bases. Este intervalo seleccionado de pares de bases se centra en diferentes valores en vista de la selección 10 cadena abajo de la plataforma de secuenciación. Por ejemplo, para la plataforma de secuenciación Illumina, el intervalo de pares de bases se centra en 650 pb y preferentemente en 400 pb. Para otras plataformas de secuenciación, por ejemplo, Ion Torrent, el intervalo de pares de bases se centra en 400 pb y preferiblemente en 200 pb y más preferiblemente en 150 pb o en 100 pb o en 50 pb.

15 De acuerdo con una realización adicional de la invención, el método para secuenciación de genoma completo de paso bajo como se mencionó anteriormente comprende además el paso de seleccionar fragmentos de ADN que comprenden tanto el primer adaptador de secuenciación como los segundos adaptadores de secuenciación.

15 Preferentemente, el paso de seleccionar fragmentos de ADN que comprenden el primer adaptador de secuenciación y los segundos adaptadores de secuenciación se lleva a cabo poniendo en contacto la biblioteca de secuenciación masivamente paralela con al menos una fase sólida que consiste en/que comprende, por ejemplo, 20 perlas paramagnéticas funcionalizadas. En una realización del método de la invención, las perlas paramagnéticas se funcionalizan con un revestimiento de estreptavidina.

25 En un método para la secuenciación de genoma completo de paso bajo de acuerdo con la invención, uno del por lo menos un primer cebador (1PR) y el por lo menos un segundo cebador (2PR) se biotilan en el extremo 5', y se obtienen fragmentos seleccionados eluyendo de los fragmentos de ADNmc no biotinilados de las perlas.

30 Como se puede ver en la figura 4, en el caso anterior, la biblioteca de ADNbc de WGA reamplificada comprende: 1) fragmentos de ADNbc no biotinilados, fragmentos de ADNbc biotinilados en una cadena y fragmentos de ADNbc biotinilados en ambas cadenas. El método de la invención comprende los pasos adicionales de:

35 g. incubar la biblioteca de ADNbc de WGA reamplificada con las perlas paramagnéticas funcionalizadas en condiciones diseñadas que causan la unión covalente entre la biotina y la estreptavidina asignada en las perlas paramagnéticas funcionalizadas;

35 h. lavar los fragmentos de ADNbc no biotinilados no unidos;

35 i. eluir de las perlas paramagnéticas funcionalizadas los fragmentos de ADNmc retenidos.

40 También se divulga un kit de preparación de biblioteca de secuenciación masivamente paralela que comprende al menos:

40 - un primer cebador (1PR) que comprende una primera sección 5' de cebador (1PR5S) y una primera sección 3' de cebador (1PR3S), la primera sección 5' de cebador (1PR5S) que comprende al menos un primer adaptador de secuenciación (1PR5SA) y al menos un primer código de barras de secuenciación (1PR5BC) en la posición 3' del al menos un primer adaptador de secuenciación (1PR5SA) y en la posición 5' de la primera sección 3' de cebador (1PR3S), y la primera sección 3' de cebador (1PR3S) que se hibrida con una sección de secuencia 5' (5SS) conocida que comprende un adaptador de secuencia universal de biblioteca de WGA o una sección de secuencia 3' (3SS) conocida complementaria inversa a la sección de secuencia 5' conocida de fragmentos de una biblioteca de ADN de WGA primaria (pWGAlib), los fragmentos comprenden además una sección de secuencia media (MSS) 3' de la sección de secuencia 5' (5SS) conocida y 5' de la sección de secuencia 3' (3SS) conocida;

50 - un segundo cebador (2PR) que comprende una segunda sección 5' de cebador (2PR5S) y una segunda sección 3' (2PR3S), la segunda sección 5' de cebador (2PR5S) que comprende al menos un segundo adaptador de secuenciación (2PR5SA) diferente del por lo menos un primer adaptador de secuenciación (1PR5SA), la segunda sección 3' que se hibrida con la sección de secuencia 5'(5SS) conocida o la sección de secuencia 3' (3SS) conocida de los fragmentos.

55 En particular, el kit de preparación de biblioteca de secuenciación masivamente paralela comprende:

60 a) el cebador SEQ ID NO:97 (Tabla 2) y uno o más cebadores seleccionados del grupo que consiste en SEQ ID NO:1 a SEQ ID NO: 96 (Tabla 2); o

65 b) el cebador de SEQ ID NO: 194 (Tabla 2) y uno o más cebadores seleccionados del grupo que consiste en SEQ ID NO: 98 a SEQ ID NO:193 (Tabla 2); o

65 c) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:195 a SEQ ID NO:202 (Tabla 4), y al

menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 203 a SEQ ID NO:214 (Tabla 4); o

5 d) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:215 a SEQ ID NO:222 (Tabla 6), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 223 a SEQ ID NO:234 (Tabla 6); o

10 e) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:235 a SEQ ID NO:242 (Tabla 7), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 243 a SEQ ID NO:254 (Tabla 7); o

15 f) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:259 a SEQ ID NO:266 (Tabla 8), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 267 a SEQ ID NO:278 (Tabla 8).

El kit de preparación de biblioteca de secuenciación masivamente paralelo puede comprender:

15 - al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:235 a SEQ ID NO:242 (Tabla 7), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:243 a SEQ ID NO:254 (Tabla 7); un cebador de secuenciación personalizado de SEQ ID NO: 255; y un cebador de SEQ ID NO:256 o SEQ ID NO:258; o

20 - al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 259 a SEQ ID NO:266 (Tabla 8), al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:267 a SEQ ID NO:278 (Tabla 8); y cebadores de SEQ ID NO:279 y SEQ ID NO:280;

diseñado para llevar a cabo un proceso de secuenciación de lectura individual óptimo.

25 El kit anterior puede comprender además un cebador seleccionado de SEQ ID NO:257 (Tabla 7) y SEQ ID NO:281 (Tabla 8) diseñado para llevar a cabo un proceso de secuenciación de extremos emparejados óptimo en una plataforma de secuenciación seleccionada.

30 Preferentemente, el kit de preparación de biblioteca de secuenciación masivamente paralelo comprende además una ADN polimerasa termoestable.

35 La presente invención finalmente se refiere también a un método para perfilar el número de copias de genoma completo, que comprende los pasos de

35 a. secuenciar una biblioteca de ADN desarrollada usando el kit de preparación de biblioteca de secuenciación como se describió anteriormente,

40 b. analizar la densidad de lectura de secuenciación en diferentes regiones del genoma,

45 c. determinar un valor de número de copias para las regiones del genoma comparando el número de lecturas en esa región con respecto al número de lecturas esperadas en la misma para un genoma de referencia.

50 Secuenciación de genoma completo de paso bajo de CTC individuales

45 El perfilado de CNA por LPWGS es más tolerante a un índice de integridad del genoma más bajo, donde aCGH puede no dar resultados lo suficientemente limpios. De hecho, las sondas aCGH se diseñan para posiciones fijas en el genoma. Si esas posiciones no se amplifican estocásticamente debido a la reticulación del ADN, la sonda correspondiente no generará la cantidad adecuada de señal después de la hibridación, lo que dará como resultado un pixel ruidoso en la relación de señal entre el ADN de prueba y el ADN de referencia.

55 Por el contrario, usando LPWGS, los fragmentos se basan solo en la selección de tamaño. Si ciertos fragmentos no se amplifican estocásticamente debido a, por ejemplo, la reticulación del ADN o las roturas inducidas por la apoptosis, todavía puede haber fragmentos adicionales del mismo tamaño susceptibles de amplificación en regiones genómicas cercanas que caen en el mismo intervalo de paso bajo. Por consiguiente, la relación señal-ruido es más resistente al índice de integridad del genoma de la biblioteca, como se muestra claramente, por ejemplo, en las figuras 6 a 9.

Preparación de bibliotecas de secuenciación masivamente paralela a partir de DRS-WGA

60 La selección de tamaño implica un submuestreo del genoma dentro de las regiones compuestas por fragmentos DRS-WGA de sustancialmente la misma longitud (neta de inserción de adaptadores) que el tamaño del par de bases de la biblioteca de secuenciación.

65 Sin embargo, se ha encontrado sorprendentemente que este submuestreo no afecta la calidad del perfil de número de copia, incluso cuando se usan algoritmos estándar para la llamada a variantes de número de copia.

Ventajosamente, se selecciona el DRS-WGA (como kit *Ampli<sup>MR</sup> WGA*), que tiene un sitio de restricción determinista de TTAA. De esta manera, los fragmentos más cortos son más densos en regiones de bajo contenido de GC del genoma, y la densidad de fragmentos se correlaciona negativamente con un mayor contenido de GC.

5 Secuenciación de genoma completo de paso bajo de cantidades mínimas de células FFPE clasificadas digitalmente

10 A partir de unos cientos de células tumorales clasificadas digitalmente de FFPE con sistema DEPArray (Bolognesi et al.), generamos una biblioteca DRS-WGA. La biblioteca se utilizó para generar una biblioteca de secuenciación masivamente paralela para Ion/PGM de acuerdo con la invención, como se muestra en la figura 6. La biblioteca masivamente paralela se secuenció a <0,05 de profundidad media.

**Ejemplo 1:**

15 Protocolo para LPWGS en Ion Torrent PGM después de DRS-WGA

1) Amplificación de genoma completo de sitio de restricción determinista (DRS-WGA)

20 El ADN unicelular se amplificó utilizando el kit *Ampli<sup>1</sup><sup>MR</sup> WGA* (Silicon Biosystems) de acuerdo con las instrucciones del fabricante.

25 El kit *Ampli<sup>1</sup><sup>MR</sup> WGA* está diseñado para proporcionar amplificación de genoma completo a partir de ADN obtenido de una célula individual. Después de la lisis celular, el ADN se digiere con una enzima de restricción, preferentemente MseI, y una secuencia adaptadora universal se liga a fragmentos de ADN. La amplificación está mediada por un único cebador de PCR específico para todos los fragmentos generados, con un tamaño de intervalo de 200 a 1000 pb de longitud, que se distribuyen en todo el genoma.

30 2) Reamplificación de los productos WGA

35 Se diluyen cinco  $\mu$ L de ADN amplificado con WGA mediante la adición de 5  $\mu$ L de agua libre de nucleasa y se purifican utilizando el sistema AgenCourt AMPure XP (Beckman Coulter) para eliminar los oligonucleótidos no unidos y el exceso de nucleótidos, sales y enzimas.

40 La purificación de ADN basada en perlas se realizó de acuerdo con el siguiente protocolo: Se agregaron 18  $\mu$ L de perlas (volumen de muestra 1,8X) a cada muestra. Las perlas y los productos de reacción se mezclaron mediante agitación excéntrica breve y luego se centrifugaron para recolectar las gotitas. Las reacciones mixtas luego se incubaron fuera del imán durante 15 min a TA, después de lo cual luego se transfirieron a un imán lateral DynaMag-96 (Life Technologies) y se dejaron reposar durante 5 min. El sobrenadante se descartó y las perlas se lavaron con 150  $\mu$ L de EtOH al 80% recién hecho. Después de una segunda ronda de lavado con EtOH, se dejaron secar las perlas en el imán durante 5-10 min. Las perlas secas entonces se resuspendieron fuera del imán en 15  $\mu$ L de amortiguador LowTE y se incubaron durante 10 min, seguido de 5 min de incubación en el imán. Se transfirieron doce microlitros del eluado a otro tubo y posteriormente se cuantificaron mediante ensayo HS de ADNbc en el fluorómetro Qubit<sup>®</sup> 2.0 para preparar alícuotas de 10  $\mu$ L que contenían 25 ng de ADN purificado con WGA.

45 La reamplificación con código de barras se realizó en un volumen de 50  $\mu$ L utilizando el kit de PCR *Ampli<sup>1</sup><sup>MR</sup>* (Menarini Silicon Biosystems). Cada reacción de PCR se compuso de la siguiente manera: 5  $\mu$ L de amortiguador de reacción de PCR *Ampli<sup>1</sup><sup>MR</sup>* (10X), 1  $\mu$ L de 25  $\mu$ M de un cebador de SEQ ID NO:1 a SEQ ID NO:96

50 [1] (5'-CCATCTCATCCCTGCGTCTCCGACTCAG[BC]-)AGTGGGATTCTGCTGTCAGT-3')  
donde [BC] = Secuencia de código de barras, 1  $\mu$ L de 25  $\mu$ M del cebador SEQ ID NO:97

55 [2] (5'-CCTCTCTATGGGCAGTCGGTGATAGTGGGATTCTGCTGTCAGT-3')

1,75  $\mu$ L de *Ampli<sup>1</sup><sup>MR</sup>* PCR dNTP (10 mM), 1,25  $\mu$ L de BSA, 0,5 de *Ampli<sup>1</sup><sup>MR</sup>* polimerasa PCR Taq, 37,5  $\mu$ L de agua *Ampli<sup>1</sup><sup>MR</sup>* y 25 ng del ADN purificado por WGA.

60 El termociclador Applied Biosystems<sup>®</sup> 2720 se estableció como sigue: 95°C durante 4 min, 1 ciclo de 95°C durante 30 seg, 60°C durante 30 seg, 72°C durante 2 min, 10 ciclos de 95°C durante 30 seg, 60°C durante 30 seg, 72°C durante 2 min (extendido por 20 seg/ciclo) y extensión final a 72°C durante 7 min.

65 La figura 3 muestra esquemáticamente el proceso de reamplificación.

Los productos de WGA reamplificados con código de barras se purificaron con perlas AMPure XP 1,8x (90  $\mu$ L) y se eluyeron en 35  $\mu$ L de amortiguador de TE bajo de acuerdo con los pasos descritos anteriormente.

65 3) Selección de tamaño

- 5 Los productos de WGA reamplificados con código de barras, correspondientes a una biblioteca de fragmentos con adaptadores Ion Torrent proporcionados, se calificaron mediante el kit Agilent DNA 7500 en el 2100 Bioanalyzer® (Agilent) y se cuantificaron utilizando el kit de ensayo Qubit® dsDNA HS con el fin de obtener un agrupamiento final.
- 10 El agrupamiento equimolar se creó combinando la misma cantidad de 7 bibliotecas individuales con diferentes adaptadores A-LIB-BC-X, produciendo el grupo final con la concentración de 34 ng/µL en un volumen final de 42 µL. La concentración del agrupamiento se confirmó mediante el método Qubit®.
- 15 El sistema E-Gel® SizeSelect<sup>MR</sup> en combinación con el gel de agarosa prefabricado Size Select 2% (Invitrogen) se ha utilizado para la selección del tamaño de los fragmentos de interés, de acuerdo con las instrucciones del fabricante.
- 20 15 Se cargaron veinte µL del grupo final en dos carriles de un gel E y usando un estándar de tamaño (ADN 50 pb, Ladder, Invitrogen), se seleccionó un intervalo de sección entre 300 y 400 pb del gel.
- 25 Despues de seleccionar el tamaño, la limpieza se realizó con 1. 8x (90 µL) perlas AMPure XP. La biblioteca final se eluyó en 30 µL de amortiguador de TE bajo de acuerdo con los pasos descritos anteriormente y se evaluó utilizando un chip de alta sensibilidad 2100 Bioanalyzer (Agilent Technologies).
- 30 4) Secuenciación de Ion Torrent PGM
- 35 La preparación del molde se realizó de acuerdo con la guía de usuario Ion PGM<sup>MR</sup> Hi-Q OT2 kit-400 pb.
- 40 25 En resumen, los fragmentos de biblioteca se amplificaron clonalmente en partículas de esferas iónicas (ISP) a través de PCR en emulsión y luego se enriquecieron para ISP positivos para el molde. Las reacciones de PCR en emulsión de PGM se realizaron con el kit Ion PGM<sup>MR</sup> Hi-Q OT2 (Life Technologies) y las emulsiones y amplificaciones se generaron utilizando el Instrumento Ion OneTouch (Life Technologies). Despues de la recuperación, el enriquecimiento se realizó uniendo selectivamente las ISP que contenían fragmentos de biblioteca amplificados a perlas magnéticas revestidas con estreptavidina, eliminando los ISP vacíos a través de pasos de lavado y desnaturizando las cadenas de biblioteca para permitir la recolección de los ISP positivos para la plantilla.
- 45 35 Los pasos de enriquecimiento descritos se realizaron utilizando el sistema Life Technologies ES (Life Technologies).
- 50 40 El chip Ion 318v2<sup>MR</sup> se cargó siguiendo las instrucciones del protocolo "Simplified Ion PGM<sup>MR</sup> Chip loading with the Ion PGM<sup>MR</sup> weighted chip bucket" (MAN0007517).
- 45 45 Todas las muestras se procesaron por Ion Personal Genome Machine (PGM) (Life Technologies) utilizando el kit de secuenciación Ion PGM<sup>MR</sup> Hi-Q<sup>MR</sup> (Life Technologies) y configurando el formato de corrida de flujo 520.
- Finalmente, los fragmentos secuenciados se asignaron a muestras específicas con base en su código de barras único.

Tabla 2

50 Cebadores de reamplificación de NGS para la plataforma Ion Torrent (PGM/Proton)

a) SEQ ID NO list\_first\_primer\_[PGM/DRS-WGA]

# ES 2 998 543 T3

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:1	A -BC-LIB_1	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCTAAGGTAACAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:2	A -BC-LIB_2	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGGAGAACAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:3	A -BC-LIB_3	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGAAGAGGATT CAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:4	A -BC-LIB_4	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTACCAAGATCAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:5	A -BC-LIB_5	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGAAGAACAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:6	A -BC-LIB_6	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGCAAGTT CAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:7	A -BC-LIB_7	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTT CGTATT CAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:8	A -BC-LIB_8	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTT CGATAAACAGTGGGATT CCTGCTGTCAGT-3'
SEQ ID NO:9	A -BC-LIB_9	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTGAGCGAACAGTGGGATT CCTGCTGTCAGT-3'

# ES 2 998 543 T3

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:10	A -SC-U8_10	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGACCGAACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:11	A -SC-U8_11	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGCGAACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:12	A -SC-U8_12	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTAGGGTTCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:13	A -SC-U8_13	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTAACGGACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:14	A -SC-U8_14	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGAGTTCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:15	A -SC-U8_15	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTAGAGGTCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:16	A -SC-U8_16	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTGGATGACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:17	A -SC-U8_17	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTATTCTCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:18	A -SC-U8_18	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGAGGCAATTGCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:19	A -SC-U8_19	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTTAGTCGGACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:20	A -SC-U8_20	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGATECATCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:21	A -SC-U8_21	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTGGCAATTACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:22	A -SC-U8_22	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGAGACGCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:23	A -SC-U8_23	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTGCCACGAACAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:24	A -SC-U8_24	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGAACCTCACTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:25	A -SC-U8_25	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTGAGGATACTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:26	A -SC-U8_26	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTTACAACCTCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:27	A -SC-U8_27	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGAACCATGGCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:28	A -SC-U8_28	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGATECGGAATCAAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:29	A -SC-U8_29	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTGACCAACTCACTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:30	A -SC-U8_30	5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCGAGGTATCAAGTGGGATTCTGCTGTCAGT-3'

## ES 2 998 543 T3

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:31	A -SC-LB_31	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCCAAAGCTGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:32	A -SC-LB_32	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCCAAACACACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:33	A -SC-LB_33	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCCATTGAACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:34	A -SC-LB_34	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCCATCGTCAAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:35	A -SC-LB_35	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTAASCCATTGTCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:36	A -SC-LB_36	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGAAGGAATGTCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:37	A -SC-LB_37	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGCTTGAAGAATGTCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:38	A -SC-LB_38	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTGGAGGACGGACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:39	A -SC-LB_39	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTAACAAATCGGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:40	A -SC-LB_40	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGCTGAAATAATCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:41	A -SC-LB_41	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:42	A -SC-LB_42	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGAGCACGAATCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:43	A -SC-LB_43	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGCTGACACCCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:44	A -SC-LB_44	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTGGAGGCCAGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:45	A -SC-LB_45	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTGGAGCTTCCCTCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:46	A -SC-LB_46	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCAGTCCGAACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:47	A -SC-LB_47	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTAAGGCAACACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:48	A -SC-LB_48	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCTAAGAGACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:49	A -SC-LB_49	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTCATAACACAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:50	A -SC-LB_50	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGCGAACAAATGGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:51	A -SC-LB_51	5'-CCATCTCATCCCTGCCTGCTGCTCCGACTCAGTGGGATTCCCTGCTGTCACT-3'

## ES 2 998 543 T3

SEQ ID NO	Nombre de cebador	Siguencia de cebador
SEQ ID NO:52	A-BC-US_52	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCCGCACTGGAACAGTGGGATTCTGCTGTCACT-3'
SEQ ID NO:53	A-BC-US_53	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCTCGCAATCCTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:54	A-BC-US_54	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCCGAGAATCGCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:55	A-BC-US_55	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCCACCTCCTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:56	A-BC-US_56	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCAGCAATTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:57	A-BC-US_57	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCTGGCAACCGCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:58	A-BC-US_58	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCTAGAACACAGTGGGATTCTGCTGTCACT-3'
SEQ ID NO:59	A-BC-US_59	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCTTGAATGTTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:60	A-BC-US_60	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCTCACTCTTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:61	A-BC-US_61	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCTCGGATCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:62	A-BC-US_62	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCTGCTTCACAGTGGGATTCTGCTGTCACT-3'
SEQ ID NO:63	A-BC-US_63	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCCTTAGAGTTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:64	A-BC-US_64	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCTGAGCTGGGATTCTGCTGTCACT-3'
SEQ ID NO:65	A-BC-US_65	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCTGGCACATCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:66	A-BC-US_66	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCCGCAATCATCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:67	A-BC-US_67	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCTACCACTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:68	A-BC-US_68	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTCAAGAACAGTTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:69	A-BC-US_69	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTTCAATTGGCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:70	A-BC-US_70	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCCTACTGGTCACTGGGATTCTGCTGTCACT-3'
SEQ ID NO:71	A-BC-US_71	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGTGAGGCTCCGACAGTGGGATTCTGCTGTCACT-3'
SEQ ID NO:72	A-BC-US_72	5'-CCATCTCATCCCTGCGTGTCTCGACTCAGCGAAGGCCACACAGTGGGATTCTGCTGTCACT-3'

## ES 2 998 543 T3

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:73	A -8C-U8_73	5'-CCATCTCATCCTGCCTGCTCCGACTCAGTCCTGTCAGTGGGATTCCTGCTGTCACT-3'
SEQ ID NO:74	A -8C-U8_74	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGATCGGTTCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:75	A -8C-U8_75	5'-CCATCTCATCCTGCCTGCTCCGACTCAGTCAGGAATACAGTCGGATTCTGCTGTCACT-3'
SEQ ID NO:76	A -8C-U8_76	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGAAGAACCTCACTGGGATTCCTGCTGTCACT-3'
SEQ ID NO:77	A -8C-U8_77	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGAAGCGATTCACTGGGATTCCTGCTGTCACT-3'
SEQ ID NO:78	A -8C-U8_78	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGCAATTCTCACTGGGATTCCTGCTGTCACT-3'
SEQ ID NO:79	A -8C-U8_79	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCCTGGTTGTCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:80	A -8C-U8_80	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGAAGGCAGGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:81	A -8C-U8_81	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGCTGCAATTGGCACTGGGATTCCTGCTGTCACT-3'
SEQ ID NO:82	A -8C-U8_82	5'-CCATCTCATCCTGCCTGCTCCGACTCAGTTGCGATCTCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:83	A -8C-U8_83	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTAGGACATTCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:84	A -8C-U8_84	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTTCCATAACAGTCGGATTCCCTGCTGTCACT-3'
SEQ ID NO:85	A -8C-U8_85	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCCAGCCTCAACAGTCGGATTCCCTGCTGTCACT-3'
SEQ ID NO:86	A -8C-U8_86	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTTGGTTATTCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:87	A -8C-U8_87	5'-CCATCTCATCCTGCCTGCTCCGACTCAGTTGGCTGGACATTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:88	A -8C-U8_88	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGAACACTTCAGTCGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:89	A -8C-U8_89	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTTGAATCTCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:90	A -8C-U8_90	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTAACCGSCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:91	A -8C-U8_91	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCGGAAGGATGCACTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:92	A -8C-U8_92	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTAGGAACCGCAGTGGGATTCCCTGCTGTCACT-3'
SEQ ID NO:93	A -8C-U8_93	5'-CCATCTCATCCTGCCTGCTCCGACTCAGCTTGTCCAATCACTGGGATTCCCTGCTGTCACT-3'

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:94	A-BC-LIB_94	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:95	A-BC-LIB_95	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:96	A-BC-LIB_96	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'

b) SEQ ID NO list\_second\_primer\_[PGM/DRS-WGA]

SEQ ID-NO	nombre de cebador	Secuencia de cebador
SEQ ID NO:97	P1-LIB	5'-CCTCTCTATGGCAGTCGGTATGGGATTCCCTGCGTCAGT-3'

5

c) SEQ ID NO list\_first\_primer\_[PGM/MALBAC]

SEQ ID NO	Nombre de cebador	Secuencia de cebador
SEQ ID NO:98	A-BC-MALBAC_1	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:99	A-BC-MALBAC_2	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:100	A-BC-MALBAC_3	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:101	A-BC-MALBAC_4	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:102	A-BC-MALBAC_5	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:103	A-BC-MALBAC_6	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:104	A-BC-MALBAC_7	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:105	A-BC-MALBAC_8	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:106	A-BC-MALBAC_9	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:107	A-BC-MALBAC_10	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:108	A-BC-MALBAC_11	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:109	A-BC-MALBAC_12	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:110	A-BC-MALBAC_13	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:111	A-BC-MALBAC_14	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:112	A-BC-MALBAC_15	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:113	A-BC-MALBAC_16	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:114	A-BC-MALBAC_17	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:115	A-BC-MALBAC_18	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:116	A-BC-MALBAC_19	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:117	A-BC-MALBAC_20	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'
SEQ ID NO:118	A-BC-MALBAC_21	5'-CCATCTCATCCCTGCCTGCTCCGACTCAAGTCGACAAAGCAGTGGGATTCTGCTGTCAGT-3'

SEQ ID NO:119	A-SC-MALBAC_22	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTGGAGACCTCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:120	A-SC-MALBAC_23	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGCACGAACTTGAGTGAATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:121	A-SC-MALBAC_24	5'-CCATCTCATCCTCGCTGCTCGACTCAGAACCTCATCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:122	A-SC-MALBAC_25	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGAGATAACCTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:123	A-SC-MALBAC_26	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTACACCTCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:124	A-SC-MALBAC_27	5'-CCATCTCATCCTCGCTGCTCGACTCAGAACCTCGGCTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:125	A-SC-MALBAC_28	5'-CCATCTCATCCTCGCTGCTCGACTCAGATCGGAAATCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:126	A-SC-MALBAC_29	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGCACCTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:127	A-SC-MALBAC_30	5'-CCATCTCATCCTCGCTGCTCGACTCAGCGAGGTTATCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:128	A-SC-MALBAC_31	5'-CCATCTCATCCTCGCTGCTCGACTCAGTCAACGCTCGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:129	A-SC-MALBAC_32	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTACACACGTCATGATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:130	A-SC-MALBAC_33	5'-CCATCTCATCCTCGCTGCTCGACTCAGTCTCATGAACTGTCATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:131	A-SC-MALBAC_34	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGCATGTTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:132	A-SC-MALBAC_35	5'-CCATCTCATCCTCGCTGCTCGACTCAGTAACCTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:133	A-SC-MALBAC_36	5'-CCATCTCATCCTCGCTGCTCGACTCAGAACGAAATGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:134	A-SC-MALBAC_37	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGAGGAAATGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:135	A-SC-MALBAC_38	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGAGGACGGACGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:136	A-SC-MALBAC_39	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGAACAAATCGGCTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:137	A-SC-MALBAC_40	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGACATAATCGTGAATGATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:138	A-SC-MALBAC_41	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTCCACTTCCGCTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:139	A-SC-MALBAC_42	5'-CCATCTCATCCTCGCTGCTCGACTCAGAACGAAATGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:140	A-SC-MALBAC_43	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTTACACCGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:141	A-SC-MALBAC_44	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTGGAGGCCAGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:142	A-SC-MALBAC_45	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGAGCTTCTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:143	A-SC-MALBAC_46	5'-CCATCTCATCCTCGCTGCTCGACTCAGTCGAACGTGAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:144	A-SC-MALBAC_47	5'-CCATCTCATCCTCGCTGCTCGACTCAGTAACGAACTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:145	A-SC-MALBAC_48	5'-CCATCTCATCCTCGCTGCTCGACTCAGTTCTAACACCTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:146	A-SC-MALBAC_49	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTCTAACATAACGTAATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:147	A-SC-MALBAC_50	5'-CCATCTCATCCTCGCTGCTCGACTCAGCGAACAAATGGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:148	A-SC-MALBAC_51	5'-CCATCTCATCCTCGCTGCTCGACTCAGTGGCTTATGCTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:149	A-SC-MALBAC_52	5'-CCATCTCATCCTCGCTGCTCGACTCAGCCCATGGAACGCTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:150	A-SC-MALBAC_53	5'-CCATCTCATCCTCGCTGCTCGACTCAGCGAACAAATGGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:151	A-SC-MALBAC_54	5'-CCATCTCATCCTCGCTGCTCGACTCAGCGAACAAATGGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:152	A-SC-MALBAC_55	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTAACCTCTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:153	A-SC-MALBAC_56	5'-CCATCTCATCCTCGCTGCTCGACTCAGAACATTAAATGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:154	A-SC-MALBAC_57	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGGCAACCGCGTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:155	A-SC-MALBAC_58	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGATGAAACACCTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:156	A-SC-MALBAC_59	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGATGTTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:157	A-SC-MALBAC_60	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTAGCTCTGATGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:158	A-SC-MALBAC_61	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGATGTTGAGTGATGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:159	A-SC-MALBAC_62	5'-CCATCTCATCCTCGCTGCTCGACTCAGCTGATGTTGAGTGATGTTGAGGTAGTGTGGAG-3'

SEQ ID NO:160	A-SC-MALBAC_63	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTTAGAGTTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:161	A-SC-MALBAC_64	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTTAGAGTTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:162	A-SC-MALBAC_65	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTGGACATCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:163	A-SC-MALBAC_66	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCGCAATCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:164	A-SC-MALBAC_67	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTACCCCTCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:165	A-SC-MALBAC_68	5'-CCATCTCATCCTGGTGTCTCGGACTCAGTCAGGAACTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:166	A-SC-MALBAC_69	5'-CCATCTCATCCTGGTGTCTCGGACTCAGTCAGGAACTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:167	A-SC-MALBAC_70	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTACTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:168	A-SC-MALBAC_71	5'-CCATCTCATCCTGGTGTCTCGGACTCAGTCAGGAACTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:169	A-SC-MALBAC_72	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCGAAAGCCACACGTAACTCGATGGTTGAGGTACTGTGGAG-3'
SEQ ID NO:170	A-SC-MALBAC_73	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCTCGCCCTGTCTGGAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:171	A-SC-MALBAC_74	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGATGGTGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:172	A-SC-MALBAC_75	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGATGGTGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:173	A-SC-MALBAC_76	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAAGAACCTCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:174	A-SC-MALBAC_77	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAATCTCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:175	A-SC-MALBAC_78	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAATCTCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:176	A-SC-MALBAC_79	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:177	A-SC-MALBAC_80	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:178	A-SC-MALBAC_81	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:179	A-SC-MALBAC_82	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:180	A-SC-MALBAC_83	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:181	A-SC-MALBAC_84	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCTTCATAACCTCGATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:182	A-SC-MALBAC_85	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCCAGCGCTCAACCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:183	A-SC-MALBAC_86	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTATTCCTGGAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:184	A-SC-MALBAC_87	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGGACCTCGATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:185	A-SC-MALBAC_88	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAACACTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:186	A-SC-MALBAC_89	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:187	A-SC-MALBAC_90	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTAACCGGGCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:188	A-SC-MALBAC_91	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAAGGATCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:189	A-SC-MALBAC_92	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGAAACCGCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:190	A-SC-MALBAC_93	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:191	A-SC-MALBAC_94	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGCTGGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:192	A-SC-MALBAC_95	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCGGACACAGTCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'
SEQ ID NO:193	A-SC-MALBAC_96	5'-CCATCTCATCCTGGTGTCTCGGACTCAGCTTAAGCGCGTGAAGTCATGGTTGAGGTAGTGTGGAG-3'

d) SEQ ID NO list\_second\_primer\_[PGM/MALBAC]

SEQ ID-NO	nombre de cebador	Secuencia de cebador
SEQ ID NO:194	P1-MALBAC	5'-CCTCTCTATGGGCAGTCGGTGAATGTGAGTGATGGTTGAGGTAGTGTGGAG-3'

**Ejemplo 2:**

Protocolo para LPWGS en Ion Torrent Proton después de DRS-WGA

## 1. Amplificación de genoma completo de sitio de restricción determinista (DRS-WGA)

El ADN unicelular se amplificó utilizando el kit *Ampli1<sup>MR</sup>* WGA (Menarini Silicon Biosystems) de acuerdo con las instrucciones del fabricante, como se detalla en el ejemplo anterior.

## 5 2. Síntesis de ADN bicatenario

Se convirtieron cinco  $\mu$ L de ADN amplificado con WGA en ADN bicatenario (ADNb) utilizando el kit *Ampli1<sup>MR</sup>* ReAmp/ds, de acuerdo con el protocolo de fabricación. Este proceso asegura la conversión de moléculas de ADN 10 monocatenario (ADNm) en moléculas de ADN bicatenario.

## 15 3. Purificación de productos de ADNb

Se diluyeron seis  $\mu$ L de productos de síntesis de ADNb agregando 44  $\mu$ L de agua libre de nucleasa y se purificaron 20 mediante perlas Agencourt AMPure XP (Beckman Coulter) para eliminar los oligonucleótidos no unidos y el exceso 25 de nucleótidos, sales y enzimas. La purificación de ADN basada en perlas se realizó de acuerdo con el siguiente protocolo: 75  $\mu$ L (relación: 1,5X de volumen de muestra) de perlas Agencourt AMPure XP a cada muestra de 50  $\mu$ L y se mezclaron mediante agitación excéntrica. Las reacciones mixtas se incubaron luego fuera del imán durante 15 minutos a temperatura ambiente (RT), después de lo cual se colocaron en una placa magnética hasta que la solución se aclaró y se formó un sedimento ( $\approx$  5 minutos). Luego, el sobrenadante se retiró y se descartó sin alterar el sedimento (se pueden dejar aproximadamente 5  $\mu$ L), las perlas se lavaron dos veces con 150  $\mu$ L de EtOH al 70% recién hecho dejando el tubo en la placa magnética. Después de retirar cualquier solución de etanol residual de la parte inferior del tubo, el sedimento de perlas se secó brevemente al aire. Se agregaron 22  $\mu$ L de Tris Ultrapuro 10 mM, pH 8,0 y amortiguador EDTA 0,1 mM (TE bajo) y la reacción mixta se incubó a temperatura ambiente durante minutos fuera de la placa magnética, seguido de 5 minutos de incubación en la placa magnética. Se transfirieron 20  $\mu$ L del eluato a un nuevo tubo.

De lo contrario, se utilizó un paso alternativo 3 (descrito a continuación) para producir una distribución uniforme de fragmentos alrededor de un tamaño promedio.

## 30 30 Paso alternativo 3) Doble purificación de productos de ADNb

SPRIselect es una composición química basada en SPRI que acelera y simplifica la selección del tamaño del ácido nucleico para la preparación de la biblioteca de fragmentos para la secuenciación de próxima generación. Este 35 paso se podría realizar de forma alternativa al paso 3. Se diluyeron seis  $\mu$ L de productos de síntesis de ADNb agregando 44  $\mu$ L de agua libre de nucleasa y se purificaron mediante perlas SPRIselect (Beckman Coulter) para eliminar los oligonucleótidos no unidos y el exceso de nucleótidos, sales y enzimas y para producir una distribución uniforme de fragmentos alrededor de un tamaño promedio. La purificación de ADN basada en SPRI se realizó de acuerdo con el siguiente protocolo: 37,5  $\mu$ L (relación: 0,75X de volumen de muestra) de perlas SPRIselect a cada 40 muestra de 50  $\mu$ L y se mezclaron mediante agitación excéntrica. Las reacciones mixtas se incubaron luego fuera del imán durante 15 minutos a RT, después de lo cual se colocaron en una placa magnética hasta que la solución se aclaró y se formó un sedimento ( $\approx$  5 minutos). Luego, el sobrenadante se recuperó y se transfirió a un nuevo tubo. La segunda ronda de purificación se realizó agregando 37,5  $\mu$ L de perlas SPRIselect al sobrenadante y se mezclaron mediante agitación excéntrica. Las reacciones mixtas se incubaron luego fuera del imán durante 15 45 minutos a RT, después de lo cual se colocaron en una placa magnética hasta que la solución se aclaró y se formó un sedimento ( $\approx$  5 minutos). Luego, el sobrenadante se retiró y se descartó sin alterar el sedimento (se pueden dejar aproximadamente 5  $\mu$ L), las perlas se lavaron dos veces con 150  $\mu$ L de EtOH al 80% recién hecho dejando el tubo en la placa magnética. Después de retirar cualquier solución de etanol residual de la parte inferior del tubo, el sedimento de perlas se secó brevemente al aire. Se agregaron 22  $\mu$ L de amortiguador TE bajo y la reacción 50 mixta se incubó a temperatura ambiente durante 2 minutos fuera del imán, seguido de 5 minutos de incubación en la placa magnética. Se transfirieron 20  $\mu$ L del eluato a un nuevo tubo.

## 4. Reamplificación con código de barras

55 La reamplificación con código de barras se realizó en un volumen de 50  $\mu$ L utilizando el kit de PCR *Ampli1<sup>MR</sup>* (Menarini Silicon Biosystems). Cada reacción de PCR se compuso de la siguiente manera: 5  $\mu$ L de amortiguador de reacción de PCR *Ampli1<sup>MR</sup>* (10X), 1  $\mu$ L de 25  $\mu$ M de un cebador de SEQ ID NO:1 a SEQ ID NO:96.

60 [1] (5'-CCATCTCATCCCTGCGTCTCCGACTCAG[BC]AGTGGGATTCCCTGCTGTCAGT-3')  
donde [BC] = Secuencia de código de barras, 1  $\mu$ L de 25  $\mu$ M del cebador de SEQ ID NO:97

65 [2] (5'-CCTCTCTATGGGCAGTCGGTGATAGTGGGATTCCCTGCTGTCAGT-3')  
1,75  $\mu$ L de *Ampli1<sup>MR</sup>* PCR dNTP (10 mM), 1,25  $\mu$ L de BSA, 0,5 de *Ampli1<sup>MR</sup>* polimerasa PCR Taq (inicio RÁPIDO), 37,5  $\mu$ L de agua *Ampli1<sup>MR</sup>* y 2  $\mu$ L del ADN purificado ds. Estos son los mismos cebadores utilizados para Ion Torrent PGM, informados en la Tabla correspondiente de cebadores de reamplificación de NGS para la biblioteca Ion

Torrent (DRS WGA para PGM/Proton) mostrada anteriormente.

El termociclador Applied Biosystems® 2720 se estableció como sigue: 95°C durante 4 min, 11 ciclos de 95°C durante 30 segundos, 60°C durante 30 segundos, 72°C durante 15 segundos, luego una extensión final a 72°C durante 30 segundos.

5 5. Purificación de productos de ADNbc reamplificados con código de barras

10 Los productos de ADNcd reamplificados con código de barras se purificaron con una relación de perlas AMPure XP 1,5X (75 µL), de acuerdo con el paso 5 descrito anteriormente, y se eluyeron en 35 µL de amortiguador de TE bajo. El eluado se transfirió a un nuevo tubo y posteriormente se cuantificó mediante ensayo HS de ADNcd en el fluorómetro Qubit® 2.0 para obtener un agrupamiento de muestras equimolares finales. El agrupamiento equimolar se creó combinando la misma cantidad de cada biblioteca con diferentes adaptadores A-LIB-BC-X, produciendo el agrupamiento final con la concentración de 34 ng/µL en un volumen final de 42 µL.

15 6. Selección de tamaño

20 El sistema E-Gel® SizeSelect<sup>MR</sup> en combinación con el gel de agarosa prefabricado Size Select 2% (Invitrogen) se utilizó para la selección del tamaño de los fragmentos de interés, de acuerdo con las instrucciones del fabricante.

25 Se cargaron veinte µL del agrupamiento final en dos carriles de un gel E y usando un estándar de tamaño (ADN 50 pb, Ladder, Invitrogen), se seleccionó un intervalo de sección entre 300 y 400 pb del gel. Después de seleccionar el tamaño, la limpieza se realizó con perlas AMPure XP 8X (90 µL) de acuerdo con el paso 3 descrito anteriormente. La biblioteca final se eluyó en 30 µL de amortiguador de TE bajo.

25 7. Secuenciación de Ion Torrent Proton

30 El agrupamiento equimolar, después de la etapa de purificación, se calificó mediante el kit de alta sensibilidad de ADN Agilent en el 2100 Bioanalyzer® (Agilent) y se cuantificó utilizando el kit de ensayo Qubit® dsDNA HS. Finalmente, el conjunto equimolar se diluyó a una concentración final de 100 µM.

35 La preparación del molde se realizó de acuerdo con la guía de usuario Ion ML<sup>MR</sup> Hi-Q<sup>MR</sup> Chef. El sistema Ion Chef<sup>MR</sup> proporciona preparación automatizada de moldes de alto rendimiento y carga de chips para su uso con el secuenciador Ion Proton<sup>MR</sup>. El secuenciador Ion Proton<sup>MR</sup> realiza la secuenciación automatizada de alto rendimiento de las bibliotecas cargadas en el chip Ion ML<sup>MR</sup> utilizando el kit de secuenciación Ion Proton<sup>MR</sup> Hi-Q<sup>MR</sup> (Life Technologies). Finalmente, los fragmentos secuenciados se asignaron a muestras específicas con base en su código de barras único.

40 **Ejemplo 3:**

40 Protocolos para secuenciación de genoma completo de paso bajo en Illumina MiSeq

Protocolo 1

45 • Amplificación de genoma completo de sitio de restricción determinista (DRS-WGA):  
El ADN unicelular se amplificó utilizando el kit *Ampli1<sup>MR</sup>* WGA (Silicon Biosystems) de acuerdo con las instrucciones del fabricante. Se diluyeron cinco µL de ADN amplificado con WGA mediante la adición de 5 µL de agua libre de nucleasa y se purificaron utilizando el sistema Agencourt AMPure XP (relación 1,8x). El ADN se eluyó en 12,5 µL y se cuantificó mediante ensayo HS de ADNcd en el fluorómetro Qubit® 2.0.

50 • Reamplificación con código de barras

55 La reamplificación con código de barras se realizó como se muestra esquemáticamente en la figura 4, en un volumen de 50 µL utilizando el kit de PCR *Ampli1<sup>MR</sup>* (Menarini Silicon Biosystems). Cada reacción de PCR se compuso de la siguiente manera: 5 µL de amortiguador de reacción de PCR *Ampli1<sup>MR</sup>* (10X), 1 µL de un cebador de SEQ ID NO:195 a SEQ ID NO:202 (25 µM), 1 µL de un cebador de SEQ ID NO:203 a SEQ ID NO:214 cebador(25 µM), 1,75 µL de dNTP de PCR *Ampli1<sup>MR</sup>* (10 mM), 1,25 µL de BSA, 0,5 de polimerasa PCR Taq, 25 ng del ADN purificado por WGA y agua *Ampli1<sup>MR</sup>* para alcanzar un volumen final de 50 µL.

60 El termociclador Applied Biosystems® 2720 se estableció como sigue: 95°C durante 4 minutos, 1 ciclo de 95°C durante 30 segundos, 60°C durante 30 segundos, 72°C durante 2 minutos, 10 ciclos de 95°C durante 30 segundos, 60°C durante 30 segundos, 72°C durante 2 minutos (extendido por 20 segundos/ciclo) y extensión final a 72°C durante 7 minutos.

65 Los productos de WGA reamplificados con código de barras (que contienen secuencias adaptadoras de

secuenciación Illumina tomadas de la lista SEP IDS ILL PR1) se calificaron entonces mediante el kit Agilent DNA 7500 en el 2100 Bioanalyzer® y se cuantificaron mediante el fluorómetro Qubit® 2.0.

• Selección de tamaño

5 Luego, las bibliotecas se combinaron a una concentración equimolar y el grupo resultante, con una concentración de 28,6 ng/μL y un volumen final de 100 μL, se seleccionó por tamaño mediante doble purificación con perlas SPRI. En resumen, las perlas SPRI se diluyeron 1:2 con agua de grado PCR. Se agregaron 160 μL de perlas SPRI diluidas a los 100 μL de la combinación. Después de la incubación, se transfirieron 25 μL de sobrenadante a un nuevo vial y se agregaron 30 μL de perlas SPRI diluidas. El ADN se eluyó en 20 μL de TE bajo. El tamaño de fragmento se verificó mediante el chip de alta sensibilidad 2100 Bioanalyzer (Agilent Technologies) y la cuantificación de la biblioteca se realizó mediante qPCR utilizando el kit de cuantificación de la biblioteca Kapa.

10 • Secuenciación de MiSeq

15 7 nM del agrupamiento seleccionado por tamaño se desnaturizó 5 minutos con NaOH (concentración final de NaOH igual a 0.1N). La muestra desnaturizada luego se diluyó con HT1 para obtener una biblioteca desnaturizada 20 μM en NaOH 1 mM. Se cargaron 570 μL de biblioteca desnaturizada 20 pM y 30 μL de control PhiX desnaturizado 20 μM en un MiSeq (Illumina).

20 20 Se generaron lecturas de extremo individual de 150 bases usando la composición química v3 del Illumina MiSeq.

SEQ ID NO: list\_first\_primer\_[ILLUMINA/DRS-WGA] protocol1

25 La siguiente tabla ilustra la estructura de los cebadores compatibles con DRS-WGA para la plataforma Illumina (secuencias en la dirección 5' 3', 5' y 3' omitidas):

Tabla 3

	PS/Indice/Secuencia	BS	Secuencias read1	ad. de coda	secuencias	BS
LP_DL_D501	AATGATAACGGGACCCACCGAGATCACAC	TATAGCT	TACACTTCCCTACACGACGCTCTTCCGATCT			AGTGGGAATCTGCTGTCAGT
LP_DL_D502	AATGATAACGGGACCCACCGAGATCTACAC	ATAGAGG	TACACTTCCCTACACGACGCTCTTCCGATCT	T		AGTGGGAATCTGCTGTCAGT
LP_DL_D503	AATGATAACGGGACCCACCGAGATCTACAC	CCTATCT	TACACTTCCCTACACGACGCTCTTCCGATCT	CT		AGTGGGAATCTGCTGTCAGT
LP_DL_D504	AATGATAACGGGACCCACCGAGATCTACAC	GGCTCTG	TACACTTCCCTACACGACGCTCTTCCGATCT	GCC		AGTGGGAATCTGCTGTCAGT
LP_DL_D505	AATGATAACGGGACCCACCGAGATCTACAC	AGGGAA	TACACTTCCCTACACGACGCTCTTCCGATCT	GTC		AGTGGGAATCTGCTGTCAGT
LP_DL_D506	AATGATAACGGGACCCACCGAGATCTACAC	TATCTT	TACACTTCCCTACACGACGCTCTTCCGATCT	TAC		AGTGGGAATCTGCTGTCAGT
LP_DL_D507	AATGATAACGGGACCCACCGAGATCTACAC	CAGGAGT	TACACTTCCCTACACGACGCTCTTCCGATCT			AGTGGGAATCTGCTGTCAGT
LP_DL_D508	AATGATAACGGGACCCACCGAGATCTACAC	GTACTGAC	TACACTTCCCTACACGACGCTCTTCCGATCT	C		AGTGGGAATCTGCTGTCAGT
	P7r	P7r	Secuencias read2			BS
LP_DL_D701	CAACCGAAAGACGCCATACTGAGAT	CGAATAT	GTGACGCGAATTCAGACGGGCTCTTCCGATC	T		AGTGGGAATCTGCTGTCAGT
LP_DL_D702	CAACCGAAAGACGCCATACTGAGAT	TCTCCGA	GTGACGCGAATTCAGACGGGCTCTTCCGATC	T		AGTGGGAATCTGCTGTCAGT
LP_DL_D703	CAACCGAAAGACGCCATACTGAGAT	AAYGAGCG	GTGACGCGAATTCAGACGGGCTCTTCCGATC	CT		AGTGGGAATCTGCTGTCAGT
LP_DL_D704	CAACCGAAAGACGCCATACTGAGAT	GGAAATTC	GTGACGCGAATTCAGACGGGCTCTTCCGATC	GCC		AGTGGGAATCTGCTGTCAGT
LP_DL_D705	CAACCGAAAGACGCCATACTGAGAT	TTCTGAAT	GTGACGCGAATTCAGACGGGCTCTTCCGATC	GTC		AGTGGGAATCTGCTGTCAGT
LP_DL_D706	CAACCGAAAGACGCCATACTGAGAT	ACGAATT	GTGACGCGAATTCAGACGGGCTCTTCCGATC	TAC		AGTGGGAATCTGCTGTCAGT
LP_DL_D707	CAACCGAAAGACGCCATACTGAGAT	AGCTTAC	GTGACGCGAATTCAGACGGGCTCTTCCGATC	T		AGTGGGAATCTGCTGTCAGT
LP_DL_D708	CAACCGAAAGACGCCATACTGAGAT	GGCTTAT	GTGACGCGAATTCAGACGGGCTCTTCCGATC	C		AGTGGGAATCTGCTGTCAGT
LP_DL_D709	CAACCGAAAGACGCCATACTGAGAT	CATAGCCG	GTGACGCGAATTCAGACGGGCTCTTCCGATC	CT		AGTGGGAATCTGCTGTCAGT
LP_DL_D710	CAACCGAAAGACGCCATACTGAGAT	TTGGCGGA	GTGACGCGAATTCAGACGGGCTCTTCCGATC	GCC		AGTGGGAATCTGCTGTCAGT
LP_DL_D711	CAACCGAAAGACGCCATACTGAGAT	GGCGGAGA	GTGACGCGAATTCAGACGGGCTCTTCCGATC	TAC		AGTGGGAATCTGCTGTCAGT
LP_DL_D712	CAACCGAAAGACGCCATACTGAGAT	CTATCGCT	GTGACGCGAATTCAGACGGGCTCTTCCGATC	GTC		AGTGGGAATCTGCTGTCAGT

30 La siguiente tabla informa las secuencias de cebadores finales:

Tabla 4

35 SEQ ID NO: list\_first\_primer\_[Illumina\_prot1\_DRs\_WGA]

SEQ ID NO	Número de cebador	Secuencia completa de cebador
SEQ ID NO:195	LP_DL_D503	5'-AATGATAACGGGACCCACCGAGATCTACATTAAGCTTAACCTTTCCTACACGACGCTCTCCGATCTGCGATGTCAGT-3'

SEQ ID NO: list\_first\_primer\_[ILLUMINA/MALBAC] protocol1

- 5 La siguiente tabla ilustra la estructura de los cebadores compatibles con MALBAC-WGA para la plataforma Illumina  
(secuencias en la dirección 5' → 3', 5' y 3' omitidas):

Tabla 5

La siguiente tabla informa las secuencias de cebadores finales:

Tabla 6

- 5 SEQ ID NO list\_first\_primer\_[Illumina\_prot1/MALBAC]

SEQ ID NO:	Nombre de cebador	Secuencia completa de cebador
SEQ ID NO:215	SP_M1_D501	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:216	SP_M1_D502	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:217	SP_M1_D503	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:218	SP_M1_D504	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:219	SP_M1_D505	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:220	SP_M1_D506	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:221	SP_M1_D507	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'
SEQ ID NO:222	SP_M1_D508	5'-AATGATACGGGAAACACCGAGATCTAACTATAGGCTACACCTTCCCTACAGGACGGCTTCGACATCTGAGTGTGTTTGGCTAGGTAAGGAG-3'

SEQ ID NO list\_second\_primer\_[Illumina\_prot1/MALBAC]

SEQ ID NO:	Nombre de cebador	Secuencia completa de cebador
SEQ ID NO:223	SP_M1_D509	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:224	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:225	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:226	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:227	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:228	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:229	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:230	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:231	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:232	SP_M1_D510	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:233	SP_M1_D511	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'
SEQ ID NO:234	SP_M1_D512	5'-CAGCAGAGAGGCGATACGGGATGCTTAACTGAGTCTGGCTACGGGCTTCGAGCTTCGAGTGTGTTGAGGTTGAGTGTGAGGAG-3'

Limitaciones del protocolo 1

- 10 Las bibliotecas resultantes del protocolo Illumina 1 son pWGA lib de cadena doble con todas las parejas combinadas de adaptadores P5/P7 posibles
- 15 Dado que dentro de la celda de flujo la hibridación se produjo también por fragmentos con adaptadores de secuenciación homogéneos (P5/P5rc, P7rc/P7), la densidad de agrupamientos y/o calidad de los agrupamientos podría resultar ligeramente menor en comparación con el caso del protocolo Illumina 2.
- Protocolo 2
- 20 Se proporciona un segundo protocolo a modo de ejemplo. Este protocolo puede ser ventajoso para aumentar la calidad de los grupos en las celdas de flujo de Illumina, al seleccionar de pWGAlib solo fragmentos que abarcan ambos adaptadores de secuenciación (P5/P7), descartando fragmentos con adaptadores de secuenciación homogéneos (P5/P5rc, P7rc/P7).

Descripción de flujo de trabajo del protocolo 2 (Illumina/DRS\_WGA) como se ilustra esquemáticamente en la figura 4

5 Todos los productos de ADN amplificados con WGA están compuestos por moléculas de diferente longitud y con una etiqueta específica: la secuencia LIB en el extremo 5' la secuencia LIB complementaria en el extremo 3' de cada molécula de ADNmc individual (indicada en azul en la figura).

10 De acuerdo con este protocolo, ambas secuencias LIB de complemento inverso son las dianas para los cebadores Re-Amp de NGS (reamplificación).

Se han diseñado dos tipos de cebadores LPb\_DI\_D50X (intervalo entre cebador de SEQ ID NO:235 a SEQ ID NO:242) y cebador biotinilado LPb\_DI\_D70X (intervalo entre cebador de SEQ ID NO:243 a SEQ ID NO:254), respectivamente en verde-amarillo-azul y en rojo-rosa-azul en la figura.

15 Como se esperaba, ambos tipos de cebadores se pueden unir a la secuencia de LIB y a la secuencia de LIB complementaria, y de hecho surgen tres tipos de amplicones del proceso de Re-Amp de NGS, como se indica en la figura.

20 Este protocolo es proporcionado por LPb\_DI\_D70X (indicado en la figura como adaptador P7rc) que obtiene una etiqueta de biotina en el extremo 5'. Esta etiqueta específica se utiliza para seleccionar, mediante perlas de estreptavidina, el único fragmento sin etiqueta de biotina:

25 5'-P5-i5-LIB-inserto-LIBcomplementario-i7-P7-3'  
como se ilustra en la figura.

Para obtener ADNmc de la formación deseada (omitiendo por simplicidad las secciones de cebadores de lectura, el ADNmc deseado es: 5'-P5-i5-nnnnn-i7-P7-3'), los cebadores deben ser como:

30 (1PR) 5'-P5-i5-LIB-3' y  
(2PR) Biotina-5'-P7rc-i7rc-LIB-3' (La biotina se omitirá en lo que sigue por simplicidad de descripción, pero es evidente que estará presente en todos y solo los extremos 5' de los fragmentos que comienzan con P7rc).

A través de la reamplificación se obtiene:

35 inicio: (los fragmentos de ADNmc de WGA se forman todos como: 5'- LIB-nnn-LIBrc-3')

ciclo de extensión n=1):

1.5 5'-P5-i5-LIB-nnn-LIBrc-3',

1.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-3'  
2^n fragmentos [0% secuenciable]  
ciclo n=2):

2.5.5 5'-P5-i5-LIB-nnn-LIBrc-i5rc-P5rc-3'  
2.5.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i5rc-P5rc-3'

2.7.5 5'-P5-i5-LIB-nnn-LIBrc-i7-P7-3'

2.7.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i7-P7-3'

2^n=4 fragmentos[25% fragmentos secuenciables]  
-ciclo=3):

2.5.5.5 5'-P5-i5-LIB-nnn-LIBrc-i5rc-P5rc-3' = 2.5.5

2.5.5.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i5rc-P5rc-3' = 2.5.7

2.5.7.5 5'-P5-i5-LIB-nnn-LIBrc-i7-P7-3' = 2.7.5

2.5.7.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i7-P7-3' = 2.7.7

2.7.5.5 5'-P5-i5-LIB-nnn-LIBrc-i5rc-P5rc-3' = 2.5.5

2.7.5.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i5rc-P5rc-3' = 2.5.7

2.7.7.5 5'-P5-i5-LIB-nnn-LIBrc-i7-P7-3' = 2.7.5

2.7.7.7 5'-P7rc-i7rc-LIB-nnn-LIBrc-i7-P7-3' = 2.7.7

2^n=8 fragmentos [25% fragmentos secuenciables] 0 fragmentos secuenciables = 2^n/4 =

2^n/2^2=2^(N-2)

Ciclo m)... 2^(m-2) secuenciable

Al final, se forman los siguientes cuatro tipos de fragmentos después de la amplificación exponencial.

40 2.5.5 5'-P5-i5-LIB-nnn-LIBrc-i5rc-P5rc-3' (se→ lavará en la primera eliminación de líquido, en tanto que se

mantienen todos los fragmentos biotinilados en las perlas paramagnéticas o, si no se lava, se acoplará solo a un sitio de unión en la celda de flujo, pero no genera un agrupamiento de secuenciación ya que no se produce amplificación de puente).

- 5 2.5.7 Biotina-5'-P7rc-i7rc-LIB-nnn-LIBrc-i5rc-P5rc-3' (→se eliminará mediante perlas revestidas con estreptavidina)
- 2.7.5 5'-P5-i5-LIB-nnn-LIBrc-i7-P7-3' (→secuenciable)
- 10 2.7.7 Biotina-5'-P7rc-i7rc-LIB-nnn-LIBrc-i7-P7-3' (→se eliminará mediante perlas revestidas con estreptavidina).

**Ejemplo 4:**

- 15 1. Amplificación de genoma completo de sitio de restricción determinista (DRS-WGA)

El ADN unicelular se amplificó utilizando el kit *Ampli1<sup>MR</sup>* WGA (Silicon Biosystems) de acuerdo con las instrucciones del fabricante.

- 20 2 Reamplificación de los productos de WGA

Se diluyen cinco µL de ADN amplificado con WGA mediante la adición de 5 µL de agua libre de nucleasa y se purifican utilizando el sistema Agencourt AMPure XP (Beckman Coulter) para eliminar oligonucleótidos no unidos y el exceso de nucleótidos, sales y enzimas.

25 La purificación de ADN basada en perlas se realizó de acuerdo con el siguiente protocolo: Se agregaron 18 µL de perlas (volumen de muestra 1,8X) a cada muestra. Las perlas y los productos de reacción se mezclaron mediante agitación excéntrica breve y luego se centrifugaron para recolectar las gotitas. Las reacciones mixtas luego se incubaron fuera del imán durante 15 minutos a temperatura ambiente, después de lo cual luego se transfirieron a un imán lateral DynaMag-96 (Life Technologies) y se dejaron reposar durante 5 min. El sobrenadante se descartó y las perlas se lavaron con 150 µL de EtOH al 80% recién hecho. Despues de una segunda ronda de lavado con EtOH, se dejaron secar las perlas en el imán durante 5-10 min. Las perlas secas entonces se resuspendieron fuera del imán en 15 µL de amortiguador de TE bajo y se incubaron durante 10 min, seguido de 5 min de incubación en el imán. Se transfirieron doce microlitros del eluato a otro tubo y posteriormente se cuantificaron mediante ensayo 30 HS de ADNbc en el fluorómetro Qubit® 2.0 para preparar alícuotas de 10 µL que contenían 25 ng de ADN purificado 35 y el exceso de nucleótidos, sales y enzimas.

La reamplificación con código de barras se realizó en un volumen de 50 µL utilizando el kit de PCR *Ampli1<sup>MR</sup>* (Silicon Biosystems). Cada reacción de PCR se compuso de la siguiente manera:

40 5 µL de amortiguador de reacción de PCR *Ampli1<sup>MR</sup>* (10X), 1 µL de 25 µM de un cebador de SEQ ID NO:235 a SEQ ID NO:242

45 [3] 5'AATGATA CGGCGACCACCGAGATCTACAC[5]GCTCTCCGTAGTGGGATT CCTGCTGTCAGTTAA3') 1 µL de 25 µM de un cebador de SEQ ID NO:243 a SEQ ID NO:254

[4] (5'/Biosg/CAAGCAGAAGACGGCATACGAGAT[7]GCTCACCGAAGTGGGATT CCTGCTGTCAGTTAA3') 1,75 µL de *Ampli1<sup>MR</sup>* PCR dNTP (10 mM), 1,25 µL de BSA, 0,5 de *Ampli1<sup>MR</sup>* polimerasa PCR Taq y 25 ng del ADN purificado por WGA y 37,5 µL de agua *Ampli1<sup>MR</sup>*.

50 El termociclador Applied Biosystems® 2720 se estableció como sigue: 95°C durante 4 min, 1 ciclo de 95°C durante 30 seg, 60°C durante 30 seg, 72°C durante 2 min, 10 ciclos de 95°C durante 30 seg, 60°C durante 30 seg, 72°C durante 2 min (extendido por 20 seg/ciclo) y extensión final a 72°C durante 7 min.

- 55 3) Selección de tamaño

60 Los productos de WGA reamplificados con código de barras, correspondientes a una biblioteca de fragmentos con adaptadores Illumina proporcionados, se calificaron mediante el kit Agilent DNA 7500 en el 2100 Bioanalyzer® (Agilent) y se cuantificaron utilizando el kit de ensayo Qubit® dsDNA HS con el fin de obtener un agrupamiento final.

65 El grupo equimolar se creó combinando la misma cantidad de bibliotecas individuales con diferentes adaptadores de índice dual LPb\_DI, produciendo el agrupamiento final con la concentración de 35 ng/µL en un volumen final de 50 µL. La concentración del agrupamiento se confirmó mediante el método Qubit®.

- 5 Se ha seleccionado un intervalo de sección de fragmentos entre 200 pb y 1 Kb mediante doble purificación utilizando el sistema de perlas SPRI (Beckman Coulter) con una relación R:0,47X and L:0,85X, respectivamente. Con el fin de eliminar el fragmento de ADN grande, agregamos 82  $\mu$ L de SPRI diluido (42  $\mu$ L de perla SPRI + 42  $\mu$ L de agua de grado PCR) y 34,2  $\mu$ L de perla SPRI sin diluir al sobrenadante para eliminar los fragmentos de ADN pequeños.
- 10 La biblioteca final se eluyó en 50  $\mu$ l de amortiguador de TE bajo y se evaluó utilizando un chip de alta sensibilidad 2100 Bioanalyzer (Agilent Technologies).
- 15 4) Selección heterogénea de la biblioteca monocatenaria de adaptador P5/P7
- 20 Se ha realizado una selección de fragmentos utilizando el sistema Dynabeads® MyOne<sup>MR</sup> Streptavidin Cl, con el fin de disociar solo el ADN no biotinilado que contiene el adaptador P5/P7 y esto se podría obtener utilizando calor o NaOH, respectivamente. Más adelante se describen dos métodos.
- 25 Se lavaron veinte  $\mu$ L de Dynabeads® MyOne<sup>MR</sup> Streptavidin Cl en un tubo de 1,5 ml dos veces con la solución B&W IX (Tris-HCl 10 mM (pH 7,5); EDTA 1 mM; NaCl 2M).
- 30 Se agregaron cincuenta  $\mu$ L de biblioteca de agrupamiento fraccionada a la perla Dynabeads® MyOne<sup>MR</sup> Streptavidin Cl y se incubó durante 15 min, pipeteando cada 5 min para mezclar a fondo. Lavar dos veces las Dynabeads® revestidas de ADN en 50  $\mu$ L de 1 x SSC (NaCl 0,15 M, citrato de sodio 0,015 M) y resuspender las perlas con 50  $\mu$ L de 1 x SSC.
- 35 Despues de la incubación a 95 °C durante 5 minutos, el tubo se colocó en la placa magnética durante 1 min y los 50  $\mu$ L de sobrenadante se transfirieron a un nuevo tubo y se incubaron en hielo durante 5 min.
- 40 En este punto, el sobrenadante contiene una biblioteca de cadenas de ADN no biotiniladas con adaptador P5/P7. Para asegurar que el lavado fuera más estricto, se repitió el procedimiento de selección de estreptavidina por segunda vez.
- 45 En su lugar de usar calor, las Dynabeads® revestidas con ADN lavadas se podrían hacer resuspendiendo con 20  $\mu$ L de NaOH 0,15 M recién preparado.
- 50 Despues de la incubación a temperatura ambiente durante 10 min, el tubo se colocó en un soporte magnético durante 1-2 minutos y se transfirió el sobrenadante a un nuevo tubo.
- 55 El sobrenadante contiene su cadena de ADN no biotinilada.
- 60 La biblioteca monocatenaria se neutralizó mediante la adición de 2,2  $\mu$ L 10 x TE, pH 7,5 y 1,3  $\mu$ L de ácido acético 1,25 M.
- 65 La concentración final de la biblioteca cuantificada por el kit de ensayo de ADN monocatenario Qubit® fue de 5 ng/ $\mu$ L correspondiente a 25  $\mu$ M.
- 70 5) Secuenciación de MiSeq
- 75 4 nM del agrupamiento final se desnaturizaron 5 minutos con NaOH (concentración final de NaOH igual a 0,1N). La muestra desnaturizada luego se diluyó con HT1 para obtener una biblioteca desnaturizada 20 pM en NaOH 1 mM. Se cargaron 570  $\mu$ L de biblioteca desnaturizada 20 pM y 30  $\mu$ L de control PhiX desnaturizado 20 pM en un MiSeq (Illumina).
- 80 Se generaron lecturas de extremo individual de 150 bases usando la composición química v3 del Illumina MiSeq intercambiando el cebador estándar de lectura 1 y el índice de cebador estándar 1 con respectivamente 600  $\mu$ L de SEQ ID NO: 255 cebador (cebador de lectura personalizado 1) y 600  $\mu$ L de SEQ ID NO:256 o SEQ ID NO:258 cebador (índice de cebador personalizado 1a (i7) y 1b (i7))
- 85 SEQ ID NO list\_first\_primer\_[ILLUMINA\_DRS\_WGA] protocol2
- 90 La siguiente tabla informa las secuencias de cebadores finales del protocolo Illumina 2:

Tabla 7

SEQ ID	Nombre	Secuencia de cebador
SEQ ID NO list_first_primer_[Illumina_DRS_WGA_prot2]		
SEQ ID NO:235	1Pb_D1_D501	5'-AATGATAACGGCACCACCGAGATCTACACATAGCTCTGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:236	1Pb_D1_D502	5'-AATGATAACGGCACCACCGAGATCTACACATAGCGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:237	1Pb_D1_D503	5'-AATGATAACGGCACCACCGAGATCTACACCTATCTGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:238	1Pb_D1_D504	5'-AATGATAACGGCACCACCGAGATCTACACCGCTCTGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:239	1Pb_D1_D505	5'-AATGATAACGGCACCACCGAGATCTACACCGCTCTGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:240	1Pb_D1_D506	5'-AATGATAACGGCACCACCGAGATCTACACTAATCTTAGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:241	1Pb_D1_D507	5'-AATGATAACGGCACCACCGAGATCTACACAGGACGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:242	1Pb_D1_D508	5'-AATGATAACGGCACCACCGAGATCTACACGACTGACGCTCTCGTAGTGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO list_second_primer_[Illumina_DRS_WGA_prot2]		
SEQ ID NO:243	1Pb_D1_D701	/38long/CAAGCAGAAGACGGCATACGAGATCTACGTAATCTGCAACGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:244	1Pb_D1_D702	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:245	1Pb_D1_D703	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:246	1Pb_D1_D704	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:247	1Pb_D1_D705	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:248	1Pb_D1_D706	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:249	1Pb_D1_D707	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:250	1Pb_D1_D708	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:251	1Pb_D1_D709	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:252	1Pb_D1_D710	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:253	1Pb_D1_D711	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:254	1Pb_D1_D712	/38long/CAAGCAGAAGACGGCATACGAGATCTACGAGATCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO list_S8 Custom Sequencing Primer_[Illumina_DRS_WGA_prot2]		
SEQ ID NO:255	leíto de Lectura 1 personalizado	5'-GCTCTCGTAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:256	leíto 1a (D) de cebador personalizado	5'-TTAAGTGACAGCAGGAATCCCACTACGGAGAGC-3'
SEQ ID NO:257	leíto 2 de cebador personalizado (opuesto)	5'-GCTACCGAAAGTGGGATTCTGCTSTEAGTTAA-3'
SEQ ID NO:258	leíto 1b (D) de cebador personalizado	5'-TTAAGTGACAGCAGGAATCCCACTTCGGTGAGC-3'

SEQ ID NO list\_first\_primer\_[ILLUMINA/MALBAC] protocol2

La siguiente tabla informa las secuencias de cebadores finales compatibles con Illumina en caso de que el material de partida provenga de una biblioteca WGA-MALBAC:

Tabla 8

SEQ ID NO	Nombre	Secuencia de cebador
SEQ ID NO list_first_primer [Illumina/MALBAC_prst2]		
SEQ ID NO:258	1P_MII_D501	5'-AATGATAACGGCGACCAACCGAGATCTACACTATAGCTGTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:260	1P_MII_D502	5'-AATGATAACGGCGACCAACCGAGATCTACACATAGAGQCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:261	1P_MII_D503	5'-AATGATAACGGCGACCAACCGAGATCTACACCTATCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:262	1P_MII_D504	5'-AATGATAACGGCGACCAACCGAGATCTACACCTCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:263	1P_MII_D505	5'-AATGATAACGGCGACCAACCGAGATCTACACAGGGGAAGGTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:264	1P_MII_D506	5'-AATGATAACGGCGACCAACCGAGATCTACACTAATCTTAGTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:265	1P_MII_D507	5'-AATGATAACGGCGACCAACCGAGATCTACACAGGGACGCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:266	1P_MII_D508	5'-AATGATAACGGCGACCAACCGAGATCTACACGTACTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO list_second_primer [Illumina/MALBAC_prst2]		
SEQ ID NO:267	1P_MII_D701	/SBiosg/CAAGCAGAAGACGGCATACTGAGATCGAATATGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:268	1P_MII_D702	/SBiosg/CAAGCAGAAGACGGCATACTGAGATTCCTGGAGTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:269	1P_MII_D703	/SBiosg/CAAGCAGAAGACGGCATACTGAGATAATGAGCGGTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:270	1P_MII_D704	/SBiosg/CAAGCAGAAGACGGCATACTGAGATGGAAATCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:271	1P_MII_D705	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAAATCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:272	1P_MII_D706	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAAATCTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:273	1P_MII_D707	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:274	1P_MII_D708	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:275	1P_MII_D709	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:276	1P_MII_D710	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:277	1P_MII_D711	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:278	1P_MII_D712	/SBiosg/CAAGCAGAAGACGGCATACTGAGATACTGAACTGGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO list_BB Custom Sequencing Primer [Illumina/MALBAC_prst2]		
SEQ ID NO:279	Cebador de lectura 1M personalizado	5'-GTGAGTGATGGTTGAGGTTAGTGTGGAG-3'
SEQ ID NO:280	índice 1M de ceba- dor personalizado	5'-CTCCACACTACCTCAACCATCACTCAC-3'
SEQ ID NO:281	Lectura 2M de cebador persona- lizado (índice 2M)	5'-GCTCACCGAAGTGGGATTCTGTGTCACTTAA-3'

- 5 Ambas LIB complementarias inversas son las dianas para los cebadores Re-Amp de NGS (reamplificación) como se muestra en la figura 4. Además, se ha diseñado un cebador de secuenciación read1 personalizado (SEQ ID NO:255), con el fin de aumentar la complejidad de la biblioteca, porque las lecturas no comenzarán con el mismo nucleótido que podría afectar el rendimiento de secuenciación o evitar el uso de un pico de alta concentración para asegurar un conjunto más diverso de agrupamientos para cálculos de matriz, ajuste de fase y ajuste de prefase.
- 10 El cebador de secuenciación read1 personalizado (SEQ ID NO:255) contiene la secuencia LIB y es complementario a la secuencia complementaria inversa LIB, como se ilustra en la figura 4.

Además, los productos de Re-Amp (reamplificación) de NGS no tienen la secuencia canónica utilizada por los sistemas Illumina para leer el índice 1, por esta razón es necesario utilizar el índice de cebador de secuenciación personalizado 1 (i7) (SEQ ID NO:256 o SEQ ID NO:258) para permitir la lectura correcta del índice i7. Cabe destacar que el índice de cebador de secuenciación personalizado 1 contiene la secuencia LIB complementaria inversa.

Todos los ejemplos descritos anteriormente que incluyen los procedimientos PGM/Proton y el flujo de trabajo del protocolo de Illumina 1/2, se podrían realizar utilizando el cebador compatible con MALBAC enumerado en las tablas anteriores (SEQ ID NO:98 a SEQ ID NO:194 y SEQ ID NO:215 a SEQ ID NO:234 y SEQ ID NO:259 a SEQ ID NO:281).

#### Análisis de datos

Las lecturas secuenciadas se alinearon con el genoma de referencia humano hg19 utilizando el algoritmo BWA MEM (Li H. y Durbin R., 2010). Los duplicados de PCR, las alineaciones de QC secundarias/complementarias/que no pasan y las lecturas multimapador se filtraron utilizando Picard MarkDuplicates (<http://broadinstitute.github.io/picard/>) y samtools (Li H. et al, 2009). Los análisis de cobertura se realizaron usando BEDTools (Quinlan A. et al, 2010).

Se utilizó el algoritmo Control-FREEC (Boeva V. et al., 2011) para obtener llamadas de número de copias sin una muestra de control. Los recuentos de lecturas se corrigieron mediante el contenido de GC y la capacidad de asignación (opción uniqMatch) y el tamaño de la ventana se determinaron mediante software utilizando coefficientOfVariation = 0.06. La ploidía se estableció en 2 y no se utilizó el ajuste de contaminación.

Las gráficas para los perfiles de CNV se obtuvieron utilizando una secuencia de comandos de pitón personalizado como se muestra en las figuras de 6 a 9.

Aunque la presente invención se ha descrito con referencia al método para WGA *Ampli1* solamente, la técnica descrita, como parece obvio para un experto en la técnica, se aplica claramente *mutatis mutandis* también a cualquier otro tipo de WGA (por ejemplo, MALBAC) que comprende una biblioteca con regiones 5' y 3' autocomplementarias.

## REIVINDICACIONES

1. Un método para secuenciación de genoma completo de paso bajo que comprende los pasos de:

- 5 - proporcionar una pluralidad de bibliotecas de secuenciación masivamente paralela con código de barras obtenidas mediante un método para generar una biblioteca de secuenciación masivamente paralela que comprende los pasos de:  
 10 - proporcionar una biblioteca de ADN de WGA primaria (pWGAlib) que incluye fragmentos que consisten en una sección de secuencia 5'(5SS) conocida, una sección de secuencia media (MSS) y una sección de secuencia 3' (3SS) conocida complementaria inversa a la sección de secuencia 5' conocida, la sección de secuencia 5' (5SS) conocida que consiste en un adaptador de secuencia universal de biblioteca DRS-WGA o un adaptador de secuencia universal de biblioteca MALBAC, y la sección de secuencia media (MSS) que consiste en al menos una sección de inserción (IS), correspondiente a una secuencia de ADN del ADN no amplificado original antes de WGA, y opcionalmente una sección intermedia 5' flanqueante (F5) y/o una sección intermedia 3' flanqueante (F3);  
 15 - reamplificar la biblioteca de ADN de WGA primaria usando al menos un primer cebador (1PR) y al menos un segundo cebador (2PR); donde

20 el por lo menos un primer cebador (1PR) comprende una primera sección 5' de cebador (1PR5S) y una primera sección 3' de cebador (1PR3S), la primera sección 5' de cebador (1PR5S) que comprende por lo menos un primer adaptador de secuenciación (1PR5SA) y por lo menos un primer código de barras de secuenciación (1PR5BC) en la posición 3' del por lo menos un primer adaptador de secuenciación (1PR5SA) y en la posición 5' de la primera sección de cebador 3' (1PR3S), y la primera sección de cebador 3'(1PR3S) que se hibrida con la sección de secuencia 5' conocida (5SS) o la sección de secuencia 3' (3SS) conocida; el por lo menos un segundo cebador (2PR) comprende una segunda sección 5' de cebador (2PR5S) y una segunda sección 3' de cebador (2PR3S), la segunda sección 5' de cebador (2PR5S) que comprende por lo menos un segundo adaptador de secuenciación (2PR5SA) diferente del por lo menos un primer adaptador de secuenciación (1PR5SA), y la segunda sección 3' de cebador (2PR3S) que se hibrida con la sección de secuencia 5' (5SS) conocida o la sección de secuencia 3' (3SS) conocida

- 30 - agrupar muestras obtenidas utilizando diferentes códigos de barras de secuenciación (BC); y  
 - secuenciar la biblioteca agrupada.

35 2. El método de acuerdo con la reivindicación 1, donde el segundo cebador (2PR) comprende además al menos un segundo código de barras de secuenciación (2PR5BC), en la posición 3' del por lo menos un segundo adaptador de secuenciación (2PR5SA) y en la posición 5' de la segunda sección 3' de cebador (2PR3S).

40 3. El método de acuerdo con la reivindicación 1 o 2, donde el adaptador de secuencia universal de la biblioteca DRS-WGA es la SEQ ID NO:282 y el adaptador de secuencia universal de la biblioteca MALBAC es la SEQ ID NO:283 (MALBAC).

45 4. El método para secuenciación de genoma completo de paso bajo de acuerdo con cualquiera de las reivindicaciones 1 a 3, donde el paso de agrupar muestras utilizando diferentes códigos de barras de secuenciación (BC) comprende además los pasos de:  
 - cuantificar el ADN en cada una de las bibliotecas de secuenciación masivamente paralela con código de barras;

50 - normalizar la cantidad de bibliotecas de secuenciación masivamente paralela con código de barras.

55 5. El método para secuenciación de genoma completo de paso bajo de acuerdo con la reivindicación 4, donde el paso de agrupar muestras utilizando diferentes códigos de barras de secuenciación (BC) comprende además el paso de seleccionar fragmentos de ADN que tienen al menos un intervalo seleccionado de pares de bases, donde el intervalo de pares de bases se centra en 650 pb, o 400 pb, o 200 pb, o 150 pb, o 100 pb, o 50 pb.

60 6. El método para secuenciación de genoma completo de paso bajo de acuerdo con la reivindicación 5, que comprende además el paso de seleccionar fragmentos de ADN que comprenden el primer adaptador de secuenciación y los segundos adaptadores de secuenciación.

65 7. El método para secuenciación de genoma completo de paso bajo de acuerdo con la reivindicación 6, donde el paso de seleccionar fragmentos de ADN que comprenden el primer adaptador de secuenciación y los segundos adaptadores de secuenciación se lleva a cabo poniendo en contacto la biblioteca de secuenciación masivamente paralela con perlas paramagnéticas funcionalizadas con un revestimiento de estreptavidina, y donde uno del por lo menos un primer cebador (1PR) y el por lo menos un segundo cebador (2PR) se biotinilan en el extremo 5', y se obtienen fragmentos seleccionados eluyendo de los fragmentos de ADNmc no biotinilados de las perlas.

8. El método para secuenciación de genoma completo de paso bajo de acuerdo con la reivindicación 7, donde el por lo menos un segundo cebador está biotinilado en el extremo 5'.

9. El método para secuenciación de genoma completo de paso bajo de acuerdo con la reivindicación 7 o 8, que comprende además los pasos adicionales de:

- 5 - incubar la biblioteca de ADNbc de WGA reamplificada con las perlas paramagnéticas funcionalizadas en condiciones diseñadas, causando así la unión covalente entre la biotina y la estreptavidina asignada en las perlas paramagnéticas funcionalizadas;
- 10 - lavar los fragmentos de ADNbc no biotinilados no unidos;
- 10 - eluir de las perlas paramagnéticas funcionalizadas los fragmentos de ADNmc retenidos.

10 10. Un método para perfilar el número de copias de genoma completo, que comprende los pasos del método para secuenciación de genoma completo de paso bajo de acuerdo con cualquiera de las reivindicaciones anteriores,

15 15. - analizar la densidad de lectura de secuenciación en diferentes regiones del genoma,  
- determinar un valor de número de copias para las regiones del genoma comparando el número de lecturas en esa región con respecto al número de lecturas esperadas en la misma para un genoma de referencia.

20 11. Un método para perfilar el número de copias de genoma completo de acuerdo con la reivindicación 10, donde la biblioteca de secuenciación masivamente paralela se prepara mediante el uso de:

25 a) el cebador de SEQ ID NO: 97 (Tabla 2) y uno o más cebadores seleccionados del grupo que consiste en SEQ ID NO:1 a SEQ ID NO: 96 (Tabla 2); o

25 b) el cebador de SEQ ID NO:194 (Tabla 2) y uno o más cebadores seleccionados del grupo que consiste en SEQ ID NO:98 a SEQ ID NO:193 (Tabla 2); o

30 c) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 195 a SEQ ID NO:202 (Tabla 4), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:203 a SEQ ID NO:214 (Tabla 4); o

30 d) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:215 a SEQ ID NO:222 (Tabla 6), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:223 a SEQ ID NO:234 (Tabla 6); o

35 e) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:235 a SEQ ID NO:242 (Tabla 7), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:243 a SEQ ID NO:254 (Tabla 7), o

35 f) al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:259 a SEQ ID NO:266 (Tabla 8), y al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:267 a SEQ ID NO:278 (Tabla 8);

35 35. - analizar la densidad de lectura de secuenciación en diferentes regiones del genoma,  
- determinar un valor de número de copias para las regiones del genoma comparando el número de lecturas en esa región con respecto al número de lecturas esperadas en la misma para un genoma de referencia.

40 12. Un método para perfil de número de copias de genoma completo de acuerdo con la reivindicación 10, donde la biblioteca de secuenciación masivamente paralela se prepara mediante el uso de - al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:235 a SEQ ID NO:242 (Tabla 7); al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:243 a SEQ ID NO:254 (Tabla 7); un cebador de secuenciación personalizado de SEQ ID NO:255; y un cebador de SEQ ID NO:256 o SEQ ID NO:258; o

45 45. - al menos un cebador seleccionado del grupo que consiste en SEQ ID NO:259 a SEQ ID NO:266 (Tabla 8); al menos un cebador seleccionado del grupo que consiste en SEQ ID NO: 267 a SEQ ID NO: 278 (Tabla 8); y cebadores de SEQ ID NO:279 y SEQ ID NO:280;  
diseñado para llevar a cabo un proceso de secuenciación de lectura individual óptimo.

50 13. Un método para perfilar el número de copias de genoma completo de acuerdo con la reivindicación 12, donde la biblioteca de secuenciación masivamente paralela se prepara mediante el uso adicional de un cebador seleccionado de SEQ ID NO:257 (Tabla 7) y SEQ ID NO:281 (Tabla 8) diseñado para llevar a cabo un proceso de secuenciación de extremos emparejados óptimo en una plataforma de secuenciación seleccionada.

## DRS-WGA

## DIBUJOS

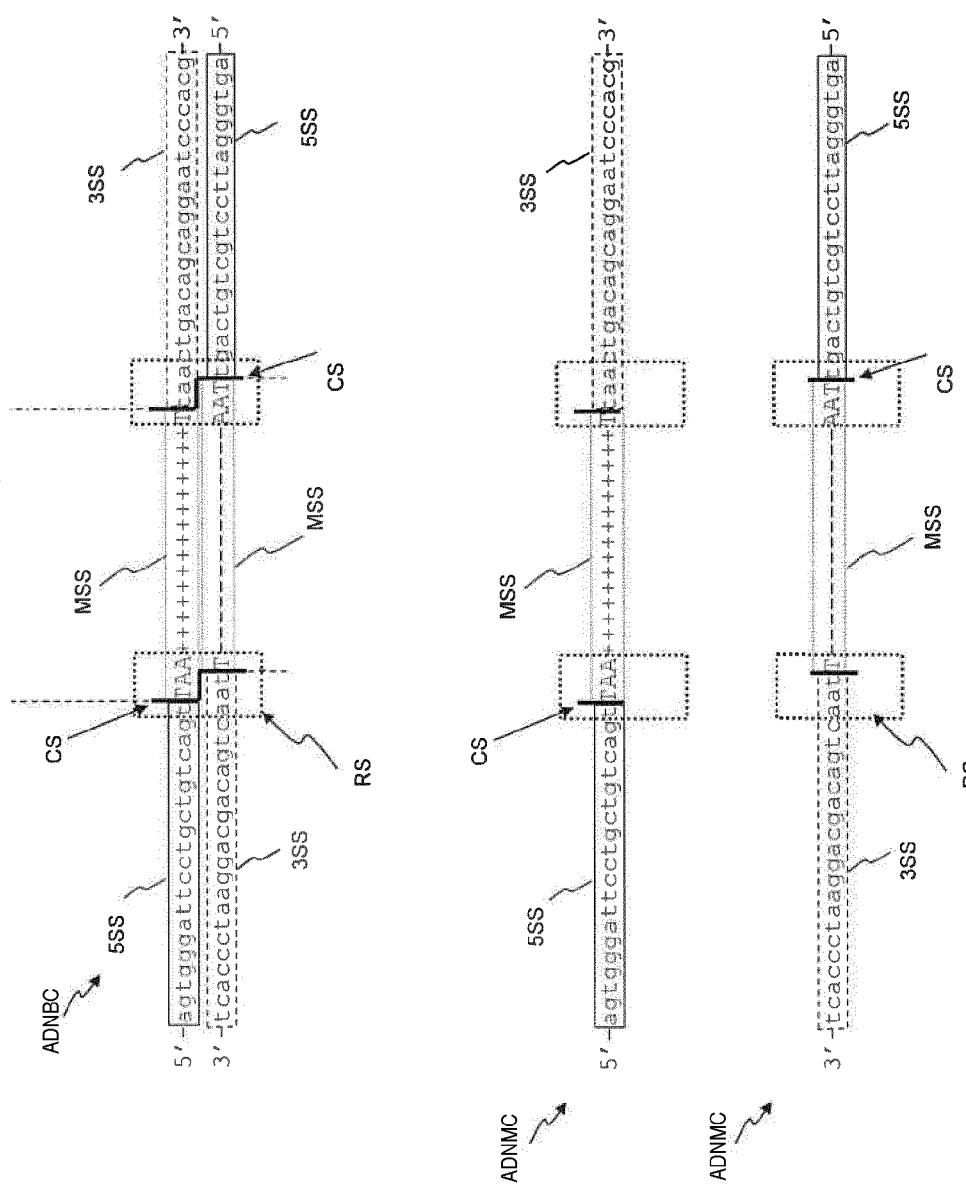


Fig.1

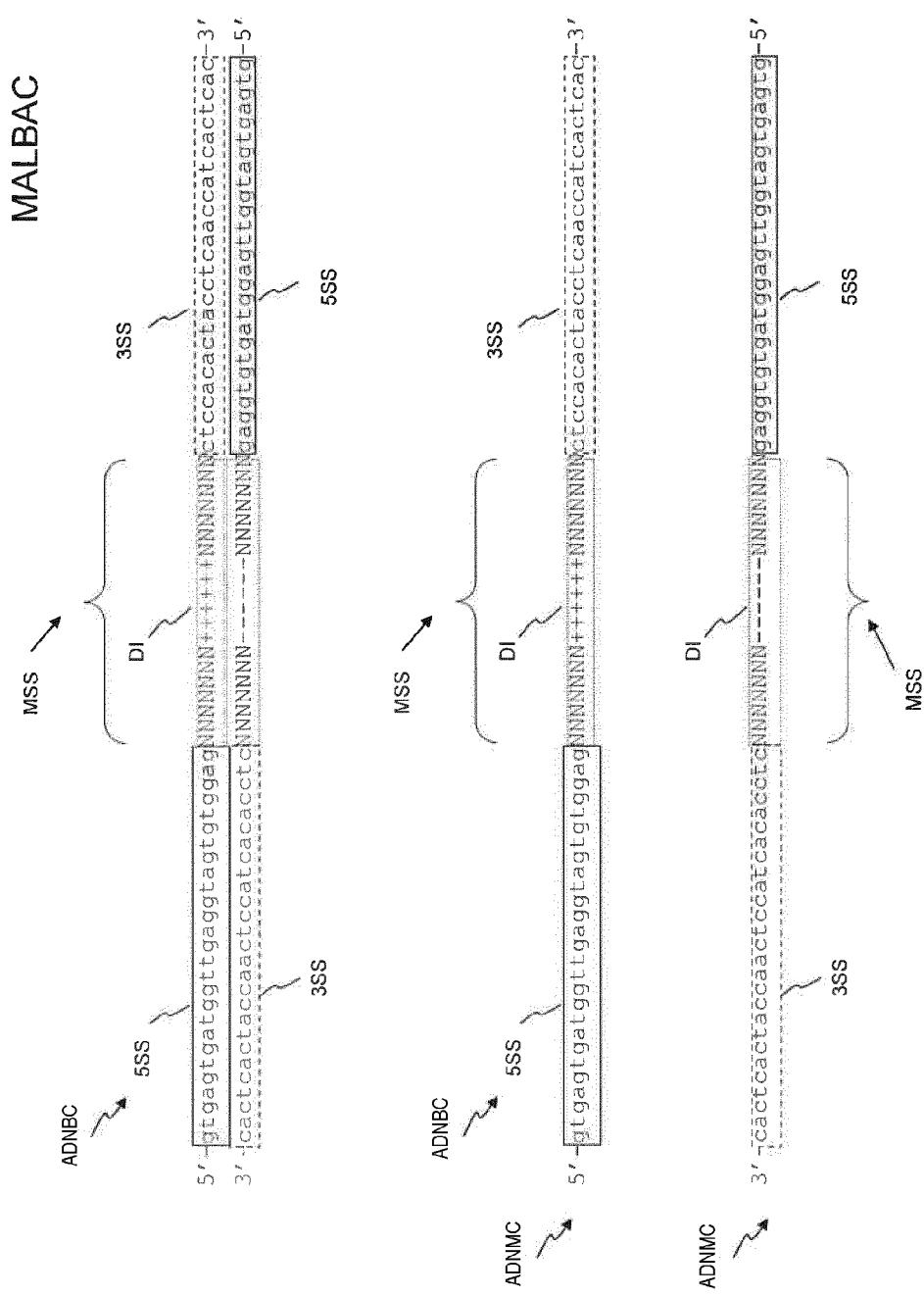


Fig. 2

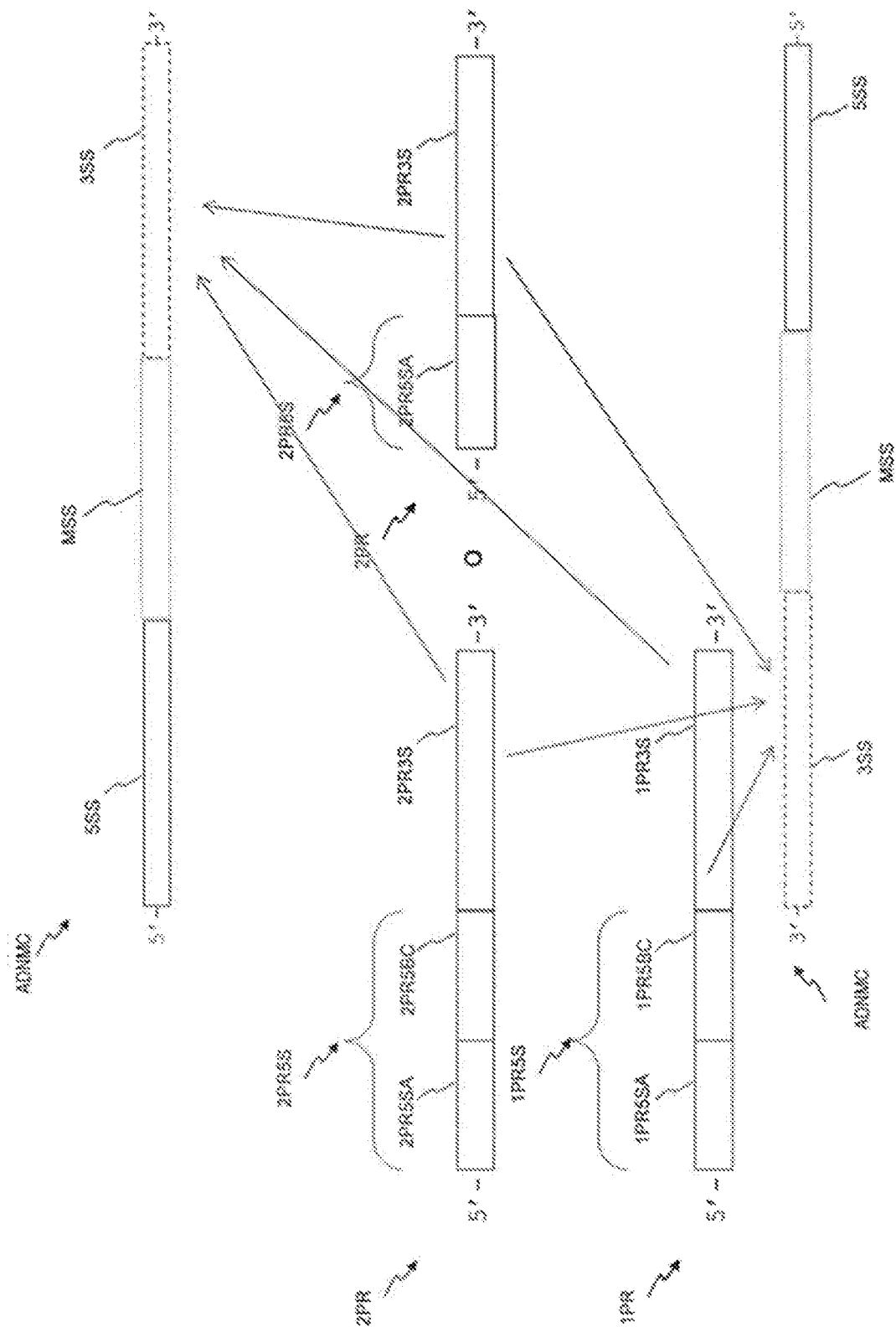


Fig. 3

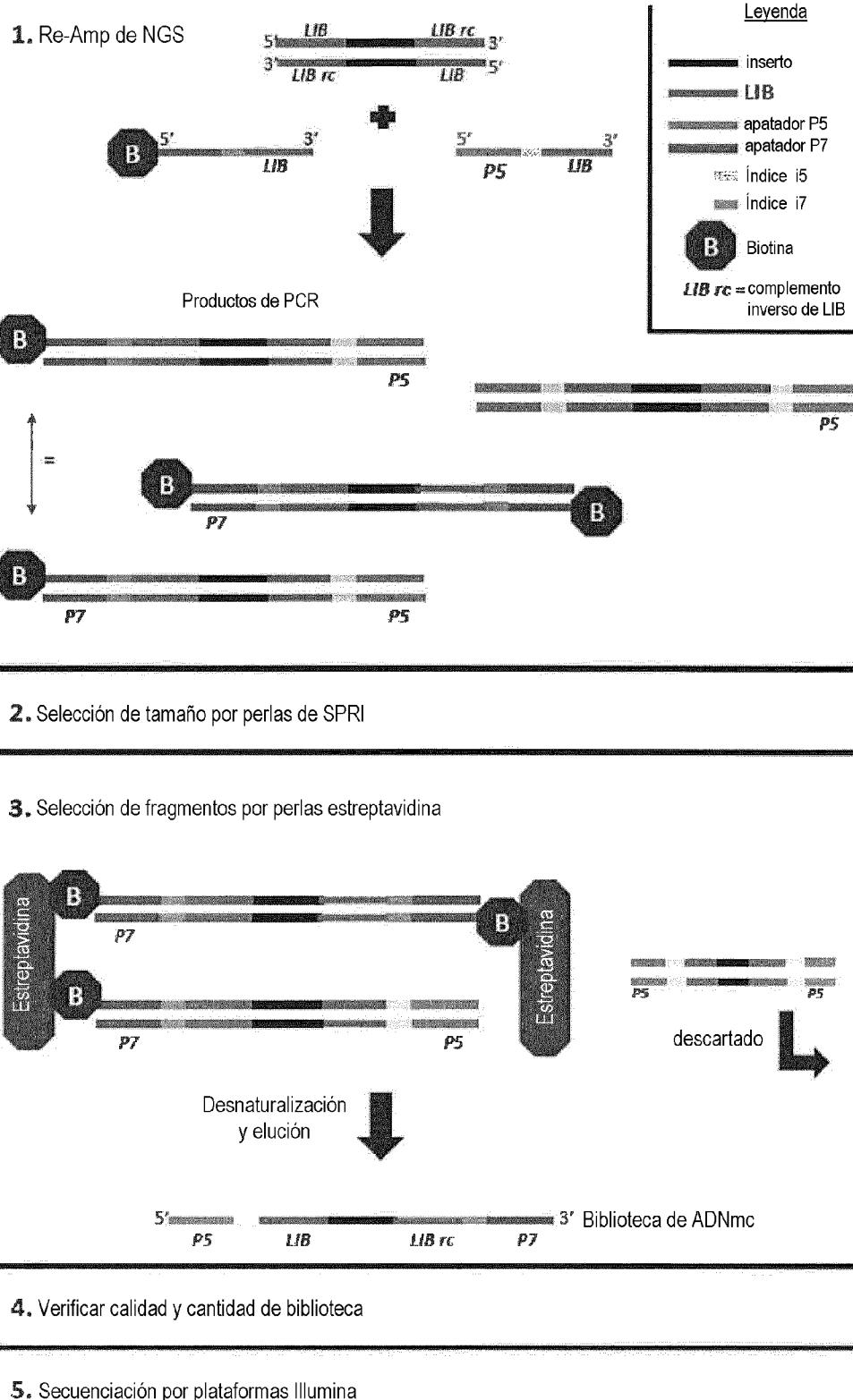


Fig. 4

## Illumina protocolo 2

Biblioteca de ADNmC final obtenido por lLumina protocolo 2  
(cebador personalizado biotinilado- índice dual)

5  
Fig.

Secuenciación de genoma completo de paso bajo de pequeñas cantidades de células de FFPE clasificadas digitalmente  
Flujo de trabajo: WGA de protocolo de FFPE/PGM

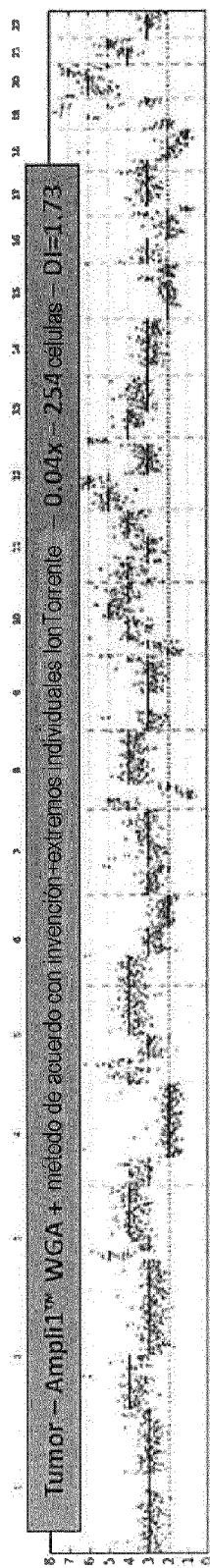


Fig. 6

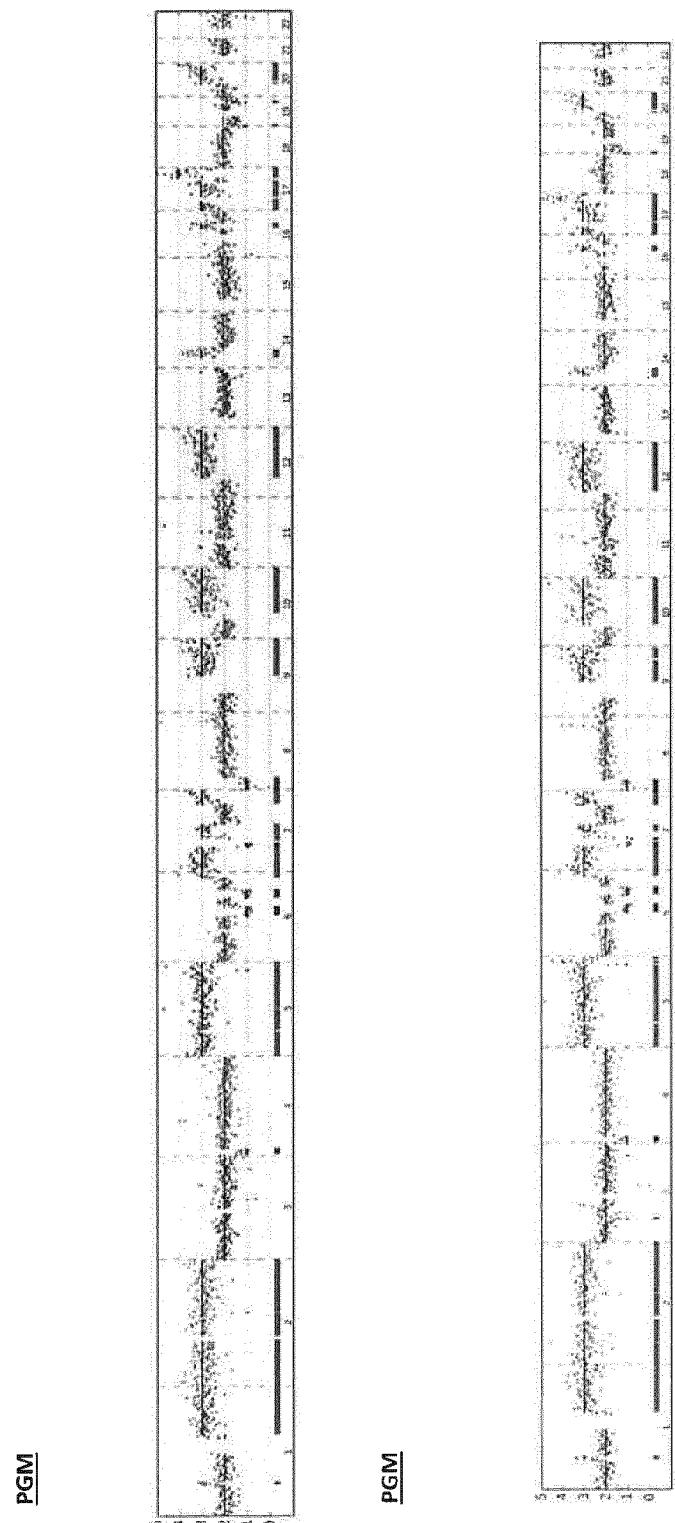


Fig. 7

ILLUMINA Protocol 1  
SW480



ILLUMINA Protocol 1  
WBC

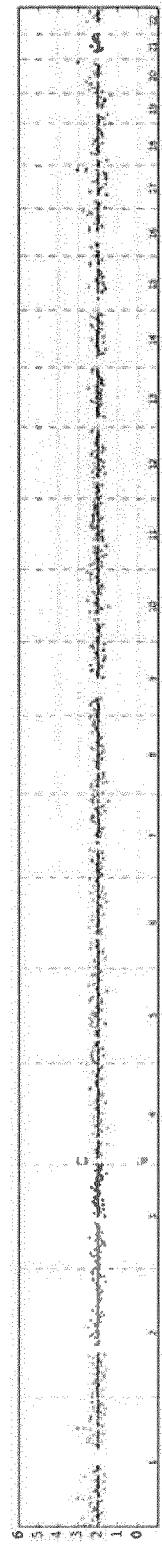


Fig. 8

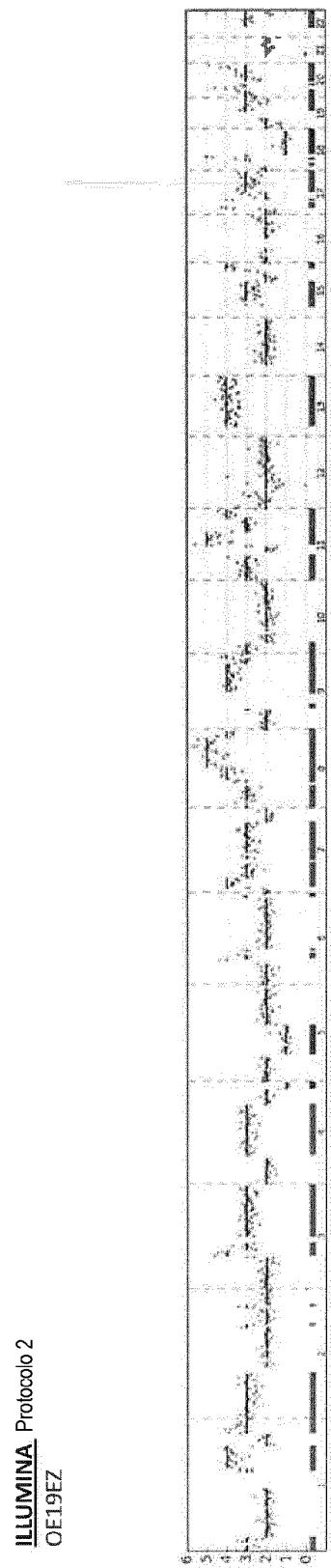


Fig. 9