

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 687 251**

51 Int. Cl.:

**C12N 15/11** (2006.01)

**C12Q 1/68** (2008.01)

**C12Q 1/70** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.10.2014 PCT/US2014/061127**

87 Fecha y número de publicación internacional: **23.04.2015 WO15058079**

96 Fecha de presentación y número de la solicitud europea: **17.10.2014 E 14853838 (2)**

97 Fecha y número de publicación de la concesión europea: **01.08.2018 EP 3058067**

54 Título: **Biomarcador de cáncer circulante y su uso**

30 Prioridad:

**18.10.2013 US 201361892796 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**24.10.2018**

73 Titular/es:

**NATIONAL TAIWAN UNIVERSITY (50.0%)  
No. 1, Sec.4, Roosevelt Road  
Taipei 10617, TW y  
CHUNG WO, ANDREW MAN (50.0%)**

72 Inventor/es:

**CHEN, PEI-JER;  
YEH, SHIOU-HWEI;  
LI, CHIAO-LING y  
CHEN, DING-SHINN**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 687 251 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

Biomarcador de cáncer circulante y su uso

**Referencia cruzada a solicitud relacionada**

5 La presente solicitud reivindica el beneficio de la solicitud provisional de los EE. UU. n.º 61/892796, presentada el 18 de octubre de 2013, cuyo contenido se adopta en la presente memoria como referencia.

**Campo técnico**

En general, los aspectos de la presente descripción se refieren al campo del uso de ácidos nucleicos circulantes en un sujeto como biomarcador para identificar y monitorizar el desarrollo de una enfermedad en el sujeto.

**Antecedentes**

10 La principal causa de tumores/cánceres se ha atribuido a alteraciones genéticas causadas por factores hereditarios o ambientales. Estas alteraciones genéticas, una vez que no han sido reparadas o se vuelven irreparables, se acumularán y eventualmente causarán que las células normales se conviertan en neoplásicas. A medida que un tumor/cáncer desarrolla su propio espectro de alteraciones genéticas exclusivo, monitorizar estas alteraciones puede proporcionar información sobre el tumor/cáncer.

15 Las células normales y las tumorales/cancerosas experimentan ciclos de renovación donde los cromosomas de células muertas se fragmentan y liberan en los fluidos corporales, tal como la circulación sanguínea. La secuenciación de estos fragmentos indica que este ADN libre circulante de la sangre o suero de pacientes con cáncer lleva las alteraciones genéticas del tumor/cáncer original. Este hallazgo señala el posible uso del ADN libre circulante.

20 El diseño convencional de uso de secuencias genómicas hospedantes que contienen alteraciones genéticas específicas como sondas para capturar secuencias de ácido nucleico específicas del cáncer/tumor a partir del ADN libre circulante total funciona para el cáncer avanzado, donde el tumor es suficientemente grande y se libera una cantidad significativa de secuencias de ácido nucleico específicas del tumor (más de 5 % del ADN circulante total) en la circulación. Dada su cantidad limitada (0,01 % ~ 1 % de la sangre total), es difícil detectar el ADN libre circulante incluso en un cáncer avanzado. En consecuencia, para un cáncer en estadio inicial o intermedio, la proporción de ADN circulante del cáncer/tumor es demasiado baja para detectarse de manera fiable. Además, las mutaciones específicas del cáncer/tumor son normalmente mutaciones en una sola base, inserciones o supresiones pequeñas que son muy difíciles de separar de las secuencias de ácido nucleico sin las mutaciones que se liberan de las células somáticas no tumorales. En otras palabras, no todo el ADN circulante lleva la información genética alterada; la mayor parte del ADN circulante está inalterado y es del genoma hospedante.

30 Estrategia convencional para la detección de ácido nucleico del cáncer/tumor en una muestra de ADN genómico. Murakami et al. presentaron una estrategia para detectar el ADN del HBV (siglas en inglés para "virus de la hepatitis B") en el hígado, suero y en células mononucleares de sangre periférica (Murakami, Y., Minami, M., Daimon, Y., & Okanoue, T. (2004). Hepatitis B virus DNA in liver, serum, and peripheral blood mononuclear cells after the clearance of serum hepatitis B virus surface antigen. *Journal of medical virology*, 72(2), 203-214.). En esta estrategia, se usó Alu-PCR para detectar el ADN de HBV circular cerrado covalentemente, ADN nuclear de HBV, ADN S de HBV y ADN X de HBV en el ADN genómico extraído de células mononucleares de sangre periférica y tejido hepático. El producto de la AluPCR se secuenció para determinar la presencia de las dianas mencionadas anteriormente. En esta estrategia, las regiones de ADN de HBV conocidas del genoma del paciente se amplificaron mediante PCR y se secuenciaron.

40 Para la detección ex-vivo de la integración del ADN de HBV, Lin et al. presentaron una estrategia para detectar los sitios de integración del ADN de HBV en líneas celulares de cáncer hepático (Lin, S., Jain, S., Block, T., Song, F., & Su, Y. H. (2013). Detection of clonally expanded HBV DNA integration sites as a marker for early detection of HBV related HCC.). Lin et al. describieron un ensayo de enriquecimiento de la diana y clonaron construcciones de ADN para identificar sitios de integración del ADN de HBV conocidos en la línea celular Hep3B. El ADN de HBV que contenía la posición de nucleótidos 1571 a 1960 se capturó mediante señuelos de ARN de HBV biotinilados y se clonó, y cada uno de los clones se secuenció. En esta estrategia, los sitios de integración del ADN de HBV en la línea celular Hep3B se detectaron a partir de clonación y secuenciación.

**Breve descripción de los dibujos**

50 Muchos aspectos de la presente descripción se pueden entender mejor en referencia a las siguientes figuras. Los componentes en las figuras no están dibujados necesariamente a escala, en cambio, se pone énfasis en ilustrar de manera clara los principios de la presente descripción.

La Figura 1 muestra esquemáticamente la progresión general de células infectadas por el virus.

La Figura 2 muestra esquemáticamente un método ejemplar para obtener ADN libre circulante diana.

La Figura 3 muestra la especificidad de la unión virus-hospedante.

La Figura 4 muestra los cambios en la cantidad de uniones virus-hospedante antes y después de la extirpación tumoral.

## 5 Descripción detallada

### 1. Introducción

La infección crónica por el virus de la hepatitis B (HBV) causa ciertos tumores/cánceres humanos, tales como el carcinoma hepatocelular (HCC, por sus siglas en inglés). Estos cánceres acumulan alteraciones genéticas en sus genomas. Entre dichas alteraciones, una exclusiva es la integración del genoma vírico en el genoma hospedante, que se produce normalmente en el estadio inicial de las infecciones. Superpuestas a estas mutaciones se encuentran otras mutaciones somáticas que continúan produciéndose y finalmente transforman las células en un tumor/cáncer.

Según se indicó, cuando las células de HCC se renuevan, el contenido genético fragmentado se liberará en los fluidos corporales. El ADN circulante que es el ADN que flota libremente en el sistema circulatorio, tal como la sangre, normalmente comprende fragmentos de ADN. Estos fragmentos incluyen aquellos del genoma del hospedante, del genoma vírico y/o de los sitios de integración del virus, tal como la unión virus-hospedante.

Las células infectadas, tales como los hepatocitos en un paciente infectado por HBV, proliferan si se vuelven cancerosas y lo mismo sucede con la cantidad de integrantes víricos que llevan las células infectadas. Por lo tanto, la cantidad de integrantes víricos es proporcional al tamaño del tumor/cáncer en general. Además, a medida que el ADN vírico se integra en el genoma hospedante en diferentes sitios, cada tumor/cáncer lleva un espectro exclusivo de sitios de integración vírica. Esta observación indica que los sitios de integración vírica y/o la unión virus-hospedante son específicos del cáncer/tumor y pueden usarse como biomarcadores para el diagnóstico del desarrollo del cáncer/tumor.

La Figura 1 muestra el proceso del desarrollo canceroso de células. Con referencia a la Figura 1, los hepatocitos en un sujeto tienen generalmente el mismo genoma hospedante. Con referencia a la Figura 1, los hepatocitos comprenden una pluralidad de hepatocitos A2, B2, C2, D2 y E2. Tras la infección por HBV 11, el HBV 11 puede integrar su genoma vírico 13 en el genoma hospedante de los hepatocitos infectados. Partes del genoma de HBV 13 se integran en el genoma hospedante, generando hepatocitos infectados con diferentes sitios de integración vírica y diferentes secuencias génicas víricas integradas. Como se muestra en la Figura 1, la secuencia vírica A1 se integra en la célula A2, la secuencia vírica B1 se integra en la célula B2, la secuencia vírica C1 se integra en la célula C2, la secuencia vírica D1 se integra en la célula D2 y la secuencia vírica E1 se integra en la célula E2. La integración de las secuencias de ADN de HBV crea uniones virus-hospedante en el genoma de la célula hospedante. Las células infectadas A2, B2, C2, D2 y E2 crecen, se desarrollan y acumulan alteraciones genéticas adicionales con el tiempo. Tanto las secuencias del hospedante como las víricas, alteradas o no, podrían conducir a la proliferación, estado estable o muerte celular. Con referencia a la Figura 1, la célula A2 lleva las alteraciones que inducen la transformación neoplásica y conducen a la proliferación o expansión clónica. Cabe señalar que una unión virus-hospedante puede conducir a la transformación neoplásica o puede ser una integración insignificante que no conduce a la proliferación.

Con referencia a la Figura 1, la célula infectada A2 prolifera, se expande en cantidad de células y se transforma en una célula neoplásica, que posteriormente forma una aglomeración celular cancerosa o tumoral. Las células B2, C2, D2, E2 no experimentan progresión neoplásica y permanecen como una población muy pequeña o mueren. Todas las células infectadas A2 llevan la misma información hereditaria, incluido el genoma hospedante, al menos genoma vírico parcial y las uniones virus-hospedante. Si las células infectadas A2 proliferan, la cantidad de uniones virus-hospedante específicas de la célula A2 aumentarán proporcionalmente en general. Las mismas uniones virus-hospedante están presentes en el mismo linaje celular canceroso, independientemente de si desencadenan el desarrollo canceroso o no. Según se representa esquemáticamente en la Figura 1, el clon canceroso experimenta una rápida proliferación y renovación, y algunas de las células infectadas A2 se rompen y mueren. Las hebras de ADN 12 de estas células infectadas rotas y muertas A2 se liberan en el sistema circulatorio o los fluidos corporales tales como la sangre. Estas hebras de ADN 12 se fragmentan, flotan libremente a través del sistema circulatorio y se vuelven parte del ADN libre circulante (ADNlc) en el torrente sanguíneo. Según se usa en la presente memoria, con referencia a la presente solicitud, se deberá entender claramente que los términos "ADN libre circulante", "ADN libre en circulación" y "ADNlc" hacen referencia al ADN que se obtiene del torrente sanguíneo o sistema circulatorio de un sujeto o paciente, en donde el ADN que se obtiene del torrente sanguíneo o sistema circulatorio del sujeto o paciente está sustancialmente libre de otros componentes celulares o está esencialmente libre de otros componentes celulares. Parte del ADNlc posteriormente se digiere o limpia a través de células funcionales tales como los macrófagos mientras que algunas permanecen en el torrente sanguíneo especialmente cuando el ADNlc está en grandes cantidades. Al examinar y/o detectar el ADNlc en el sistema circulatorio o los fluidos corporales, se puede obtener información sobre el desarrollo del cáncer/tumor.

## 2. Métodos

Los métodos para llevar a cabo la presente invención se describen a continuación. Cabe señalar que los métodos, el material y el proceso descritos más adelante son realizaciones ejemplares y no limitan el alcance de la invención de ningún modo.

5 Con referencia a la Figura 2, se muestra una vista esquemática del ADNlc diana aislado. El ADN circulante de células tumorales/cancerosas muertas se libera en la sangre en fragmentos. Dicho ADNlc se recoge y liga con adaptadores 21 y forma el ADNlc A, B, C y D. El ADNlc se amplifica mediante el uso de cualquier estrategia adecuada, por ejemplo, mediante el uso de un cebador complementario a la secuencia del adaptador 21 en una cantidad adecuada. Cabe señalar que los métodos de amplificación preferidos amplifican todo el ADNlc en una proporción similar o igual de manera que el ADNlc amplificado proporcione información legítima sobre la cantidad de ADNlc existente en la sangre. En la Figura 2, las secuencias derivadas del genoma vírico se designan con un área sombreada, mientras que las secuencias derivadas del genoma hospedante se designan con negro. El ADNlc amplificado se puede categorizar en ADNlc que tiene solo secuencias del genoma hospedante (ADNlc D), ADNlc que tiene solo secuencias del genoma vírico (no se muestra) y ADNlc que tiene secuencias del genoma vírico y hospedante y, por lo tanto, comprende uniones virus-hospedante 22 (ADNlc A, B y C). Según una estrategia preferida, todo el ADNlc se incuba con sondas polinucleotídicas 23 (derivadas de la secuencia genómica vírica) para posibilitar que se produzca la hibridación. Cabe señalar que las sondas 23 que se muestran en la Figura 2 pueden tener diferentes secuencias, aunque todas se dibujaron de manera similar. Con referencia a la Figura 2, solo el ADNlc que tiene la secuencia del genoma vírico sola y el ADNlc que tiene al menos una unión virus-hospedante (ADNlc A, B y C) formarían complejos de sonda-ADNlc 24. Estos complejos se aíslan del ADNlc que no se hibrida con la sonda. A continuación, el ADNlc diana, el ADNlc que tiene sola la secuencia del genoma vírico y el ADNlc que tiene al menos una unión virus-hospedante se obtienen a partir de los complejos y se separan. Se obtienen las secuencias del ADNlc diana. Se identifica el origen tisular del ADNlc diana en función del tropismo tisular y la especificidad de la infección vírica.

### 25 2.1 Sujetos

Se emplean sujetos humanos en las pruebas para ilustrar la presente invención. El sujeto 1 tiene un tumor 12x10x9 (cm) diagnosticado mediante tomografía computada. Según el informe histológico, cuando el sujeto 1 se emplea en esta prueba, el sujeto 1 se define como un paciente con HCC grado III. El sujeto 2 tiene un tumor 18x13,5x9 (cm) diagnosticado mediante tomografía computada. Según el informe histológico, el sujeto 2 se define como un paciente con HCC grado III. El sujeto 3 tiene un tumor 8x7,5x7 (cm) diagnosticado mediante tomografía computada. Según el informe histológico, el sujeto 3 se define como un paciente con HCC grado III. El sujeto 4 tiene un tumor de 2x2x2 (cm) y está en grado II. El sujeto 5 tiene un tumor con un tamaño menor que 2x2x2 (cm) y el estadio del desarrollo del cáncer no se ha determinado y/o no está disponible en el momento de la inscripción a la prueba.

### 2.2 Obtener el ADNlc en los sujetos

35 Se obtuvieron múltiples muestras de sangre de cada sujeto. En cada instancia, se extrajo la sangre, se recogió en un contenedor clínicamente adecuado y, en caso de necesidad, se almacenó en una condición adecuada para el análisis posterior. Cada muestra de sangre se procesa para obtener el suero, tal como mediante centrifugación. El ADNlc se extrae mediante un kit comercial, por ejemplo, el kit MagNA Pure LC Total Nucleic acid Isolation (Roche). Se obtienen los tejidos tumorales y se extraen los ADN genómicos de las células tumorales.

### 40 2.3 Suministro de sondas

Los polinucleótidos que tienen una secuencia genómica de HBV se usan como sondas aquí. Las sondas se pueden sintetizar u obtener a partir de la fragmentación del genoma vírico. Se describe la síntesis de las sondas. Se obtiene información sobre las secuencias del genoma de HBV enteras en el National Center for Biotechnology Information. Los polinucleótidos se sintetizan según la secuencia del genoma de HBV y abarcan la secuencia del genoma de HBV entera. Los polinucleótidos se sintetizan usando un kit comercial, por ejemplo, el kit Ion TargetSeq Custom Enrichment (Life Technologies). Todos los polinucleótidos tienen aproximadamente 50 a 200 o 50 a 120 residuos de longitud. Después de la síntesis de las sondas, cada sonda se etiqueta, por ejemplo, se biotinila, al menos un extremo del polinucleótido. La biotinilación de las sondas se puede llevar a cabo usando un kit comercial, por ejemplo, el kit Ion TargetSeq Custom Enrichment (Life Technologies). A continuación, las sondas se acoplan o enlazan a una perla, por ejemplo, a través de biotina.

### 2.4 Adaptadores de ligadura

Con el fin de amplificar proporcionalmente los ADNlc obtenidos del sujeto, ciertos ADN con secuencias conocidas se acoplan o ligan a al menos un extremo o ambos extremos del ADNlc. La ligadura de los adaptadores a al menos un extremo o ambos extremos del ADNlc pueden llevarse a cabo mediante el uso de TruSeq DNA Sample Preparation (Illumina), IonTorrent (Life Technologies) u otros reactivos equivalentes.

### 2.5 Amplificación de ADNlc diana

Después de que los ADNlc se ligan a los adaptadores, cada ADNlc en la muestra del sujeto se amplifica, por ejemplo, mediante el uso de TruSeq DNA Sample Preparation (Illumina) o IonTorrent (Life Technologies).

### 2.6 Captura y aislamiento de ADNlc diana

5 Las muestras de ADNlc de los sujetos se mezclan con perlas recubiertas con sondas biotiniladas y se incuban para posibilitar la hibridación entre el ADNlc y las sondas. El ADNlc que tiene al menos secuencias víricas parciales se aparea con las secuencias complementarias en las sondas y forma un complejo de perla-sonda-ADNlc. El ADNlc que no se une a las sondas flota libremente y no forma ningún complejo. Los complejos de perla-sonda-ADNlc se separan del ADNlc que no se une mediante, por ejemplo, centrifugación. Se obtienen los complejos y se retira el ADNlc diana de los complejos y se recoge. La captura del ADN circulante hibridado con las sondas se puede llevar a cabo mediante el uso del kit TargetSeq Hybridization & Wash Buffer (Life Technologies), o mediante cualquier otro reactivo equivalente.

### 2.7 Secuenciación e identificación del ADNlc diana

15 Los cebadores que tienen secuencias complementarias a las secuencias del adaptador se usan para secuenciar los ADNlc diana. El ADNlc diana se secuencía usando la plataforma IonTorrent, HiSeq 2500 (Illumina) o alguna otra plataforma de secuenciación.

## 3. Resultados

### 3.1 Secuencias de ADNlc diana

#### Sujeto 1

20 La Tabla 1 muestra las diez secuencias diana principales identificadas en las muestras de ADN obtenidas del tejido tumoral del Sujeto 1. Como se muestra, una secuencia de unión se inserta en el cromosoma hospedante (n.º de cromosoma hospedante) en una posición de integración específica (posición de integración) con un número de lecturas acumuladas (lecturas acumuladas). El número de lecturas acumuladas se obtiene mediante el resultado de la secuenciación. Se cuentan las secuencias que tienen la misma unión para proporcionar el número de uniones presentes en la muestra. Cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

Tabla 1

Datos de unión del tumor del sujeto 1				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
1	17	22247083	<u>GGTCTTACATAAGAGGACTCAGAAAATACTTTGTGATGAT</u>	290
2	17	22251295	<u>AACTCCTTTTGAGAGCGCAGTGTTCCGGTGCAGGTCCCCAG</u>	234
3	1	121360041	<u>ATCATCACAAAGTATTTTCTGAGTCCTCTTATGTAAGACC</u>	192
4	12	118876274	<u>TGAGGTGAGAGGATCTCTTGAGCACAGATGATGGGATAGG</u>	115
5	X	58568585	<u>AAACGTCCACTTGCAGATTTTATGTAATTGGAAGTTGGGG</u>	102
6	8	56895765	<u>AGCAGGAAAATATATGCCCCACCTTCCCTTTCTCTGACCC</u>	106
7	1	121475300	<u>AGGAAGACTGCCTACTCCCACAGGCCTGAAAGCGCTCCAA</u>	85
8	X	58563641	<u>AGCATTCTGGGCCAGGGTTCACTCAGGCTCAGGGCACATTG</u>	67

Datos de unión del tumor del sujeto 1				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
9	16	21525068	<u>GCATTTGGTGGTCTATAAGCACACCCGCCACACCAATCT</u>	38
10	18	77932557	<u>CAAGACCAGCCTGAGGATGACTGTCTCTTAGAGGTGGAGA</u>	26

5 La Tabla 2 muestra las secuencias diana identificadas en las muestras de ADNlc obtenidas de la sangre del sujeto 1. Las muestras de ADNlc del sujeto 1 se obtienen 13 días antes de la ablación del tumor. Como se muestra, cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

Tabla 2

Datos de unión del suero del sujeto 1				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
11	17	22251295	<u>CACTCCTTTTGAGAGCGCAGTGTTTCAGGTGCAGGGTCCCC</u>	94
12	1	121360041	<u>ATCATCACAAAGTATTTTCTGAGTCCTCTTATGTAAGACC</u>	82
13	1	13727	<u>AACAGAAAGATTCGTCCCCAAATCCAATCTGTCTTCCATC</u>	68
14	8	56895765	<u>AGCAGGAAAATATATGCCCCACCTTCCCTTTCTCTGCCCT</u>	62
15	17	22247083	<u>GGTCTTACATAAGAGGACTCAGAAAATACTTTGTGATGAT</u>	42
16	16	21525068	<u>GCATTTGGTGGTCTATAAGCACACCCGCCACACCAATCT</u>	31
17	8	56895953	<u>ATCATCCTGGGCTTTCTGCACTTCCCATAGGTAATCAAAG</u>	16
18	X	58563641	<u>AGCATTGGGGCCAGGGTTCACTCAGGCTCAGGGCACATTG</u>	9

10 Como se ilustra en las Tablas 1 y 2, al menos el n.º 3 (en la muestra tumoral) y el n.º 12 (en la muestra de suero), el n.º 1 y el n.º 15 y el n.º 2 y el n.º 11 tiene cada par las mismas secuencias de unión virus-hospedante. Los patrones similares (incluida la cantidad relativa de lecturas) de las secuencias de unión virus-hospedante identificadas en el ADN tumoral y el ADNlc indican que el ADNlc quimérico en el suero deriva del ADN tumoral. Al enriquecer selectivamente los ADNlc que llevan al menos una porción del genoma vírico en el suero, se identifican las uniones virus-hospedante para proporcionar información específica del tumor sobre el sujeto.

Sujeto 2

15 La Tabla 3 muestra las secuencias diana identificadas en las muestras de ADN obtenidas del tejido tumoral del sujeto 2. Como se muestra, cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

ES 2 687 251 T3

Tabla 3

Datos de unión del tumor del sujeto 2				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión Lecturas acumuladas	
1	3	111653312	ATGAAGCTATTTATAATAAAACAAACTTTATTAATCTAGTTTAAATG CCTTACTCTCTTTTTGCCTTCTGACTTCTTCCTTCTATTGAGAT CTCCT	4183
2	2	80278757	TTTCATTGTTGCTGTTTTCAAATTGATTTGGGATCCAGCCTGTTA TTCTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTGGGGTA CTTACC	3772
3	3	111653206	TCTCCCTTTAGACTTCAAACACTTCAAATATGACTTCACTACAAAG CTTATAGAATGCCAGCCTTCCACAGAGTATGTAATAATGCCTAG TTTTGAA	1269
4	2	80278655	CCAGCACATTTGTCTATAAATTTACATTCTTGGATATTAGCAAAT GCAAACAGACCAATTTATGCCTACAGCCTCCTAGTACAAAGACCTT TAACCTA	752
5	1	189879551	TCCAGTGTGGTGGGTTGAGCAGTATTATTGCATGGCCCAGTGGT GGTGGTTGATGTTCTGGAAGTAGAGGACAAACGGGCAACATACC TTGGTAGTCC	485
6	1	189879474	TGCAAGTGGTTGCAGTCTTTTTGCTTTGCCACCACCCTGGGCCA TGCAAACCTGCAGATTCTCTGCTCAAGGAACCTCTATGTTTCCCT CTTGTTGCT	174
7	20	60227034	CAGGAGGAGGTGATGGACCCACTGGGTGGTGAAGAACAGTTTCT CTTCCAAAATTACTTCCCACCCAGGTGGCCAGATTCATCAACTCAC CCCAACACAG	169
8	22	26941239	ATCTGTAAATGGGATCATCACACTTCTTTTATTGGGGTTTAAA TGAATACCCAAAGACAAAAGAAAATTGGTAATAGAGGTAAAAAGGG ACTCAAG	100
9	20	60227112	TGGCCGAGGCCATCTTCTAAATAAATGTGTGGAAGAGAACTGTT CTTCAGTATTGGTGTCTTTTGGAGTGTGGATTGCACTCCTCCCG CTTACAGAC	93

ES 2 687 251 T3

Datos de unión del tumor del sujeto 2				
N.º	N.º de cromosoma hospedante	de Posición de integración	de Secuencia de unión Lecturas acumuladas	
10	5	1295309	AGGACGGGTGCCCCGGTCCCCAGTCCCTCCGCCACGTGGAAG CGCGGTCCAGACCAATTTATGCCTACAGCCTCCTAGTACAAAGAC <u>CTTTAACCTAAT</u>	37

La Tabla 4 muestra las secuencias diana identificadas en las muestras de ADNlc obtenidas de la sangre del sujeto 2. Las muestras de suero del sujeto 2 se obtienen durante la ablación del tumor. Como se muestra, cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

5

Tabla 4

Datos de unión del suero del sujeto 2				
N.º	N.º de cromosoma hospedante	de Posición de integración	de Secuencia de unión Lecturas acumuladas	
11	3	111653312	ATGAAGCTATTTATAATAAAACAAACTTTATTAATCTAGTTTAAATGC <u>CTTACTCTCTTTTTGCCTTCTGACTTCTTTCCTTCTATTGAGATCTC</u> CT	3277
12	20	60227034	CAGGAGGAGGTGATGGACCCACTGGGTGGTGAAGAACAGTTTCTCT <u>TCCAAAATTACTTCCCACCCAGGTGGCCAGATTCATCAACTCACCCC</u> AACACAG	642
13	1	189879551	TCCAGTGTGGTGGGTTGAGCAGTATTATTGCATGGCCCAGTGGTGG <u>TGGTTGATGTTCTGGAAGTAGAGGACAAACGGGCAACATACCTTGG</u> TAGTCC	373
14	2	50012582	GTCCGTTGGTGGTGAAGTGGGCAAGATAATTGCATGGCCCAGTGGT <u>GGTGGTTGATGTTCTGGAAGTAGAGGACAAACGGGCAACATACCTT</u> GGTAGTC	372
15	15	48344568	AGATTGGTCTATAATTTCTTTTACTATCTTCAGTATTTGGTATCTTTG <u>GGAGTGTGGATTTCGCACTCCTCCCGCTTACAGACCACCAATGCC</u> CTATC	237
16	2	80278757	TTTCATTGTTGCTGTTTTCAAATTGATTTGGGATCCAGCCTGTTATT <u>CTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTGGGGTACTTT</u> ACC	230



Datos de unión del suero del sujeto 2				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
17	20	60227112	TGGCCGAGGCCATCTTCTAAATAAATGTGTGGAAGAGAACTGTTCT TCAGTATTTGGTGTCTTTTGGAGTGTGGATTGCGACTCCTCCCGCTTA <u>CAGAC</u>	209
18	1	189879474	TGCAAGTGGTTGCAGTTCTTTTGGCTTTGCCACCACCACTGGGCCATG CAAACCTGCACGATTCTGCTCAAGGAACCTCTATGTTTCCCTCTTG <u>TTGCT</u>	205
19	2	50012660	GTAAGCCATTGTGGCTTTCCTGACCAGCCCACCACCACTGGGCCAT GCAAACCTGCACGATTCTGCTCAAGGAACCTCTATGTTTCCCTCTT <u>GTTGCT</u>	205
20	2	80278655	CCAGCACATTTGTCTATAAATTTACATTCTTGGATATTAGCAAAATTGC AAACAGACCAATTTATGCCTACAGCCTCCTAGTACAAAGACCTTTAAC <u>CTA</u>	64

5 Como se ilustra en las Tablas 3 y 4, al menos el n.º 1 (en la muestra tumoral) y el n.º 11 (en la muestra de suero), el n.º 7 y el n.º 12, y el n.º 5 y el n.º 3 tienen ambas las mismas secuencias de unión virus-hospedante. Los patrones similares de las secuencias de la unión virus-hospedante identificadas en el ADN tumoral y el ADNlc muestran que el ADNlc quimérico en el suero deriva del ADN tumoral. Al enriquecer selectivamente los ADNlc diana en el suero, se identifican las uniones virus-hospedante para proporcionar información específica del tumor sobre el sujeto.

Sujeto 3

10 La Tabla 5 muestra las secuencias diana identificadas en las muestras de ADN obtenidas del tejido tumoral del sujeto 3. Como se muestra, cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

Tabla 5

Datos de unión del tumor del sujeto 3				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
1	5	1295930	GGAAATGGAGCCAGGCGCTCCTGCTGGCCGCGCACCGGGCGCCT CACACCAGAACATCGCATCAGGACTCCTAGGACCCCTGCTCGTGT <u>TACAGGCGGGG</u>	3024
2	8	111636420	TCAAGCAGAAAAACCATGAAGATTTAAAACTTGTAATATTTGAAT GTGGGCTCCACCCCAACAGTCCCCCGTGGGGAGGGGTGAACCCT <u>GGCCCGAAT</u>	635

ES 2 687 251 T3

Datos de unión del tumor del sujeto 3				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
3	14	52591737	<u>CTAAGGGACACTACAGGAAACCAGCCCCGAAGTGATTTCTTTTGAA ATTCCAAATCTTTCTGTCCCAATCCCCTGGGATTCTTCCCGATC ATCAGTTG</u>	354
4	9	138857330	<u>CCTCGAAGCCTGTGCCAACCTAGCCCATTCTCAGGCTCAGGGCC TCCTCACATCTGTGCCAGCAGCTCCTCCTCCTGCCTCCACCAATC GGCAGTCAGG</u>	190
5	1	68549419	<u>CATTGTTACTGTGATATGCTATAATTATTCTCACCTTATGTGTCCAA GGAATACTAACATTGAGATTCCCAGATTGAGATCTTCTGCGACGC GGCGATT</u>	188
6	9	31455679	<u>ATGGAGAATACAGCACATTATTAGGAGTAAGTTTCCTTAAACACATT TTGATTTTTTGTACAATATGTTCTGTGGCAATGTGCCCAACTCC CAATTAC</u>	172
7	17	71434403	<u>TTTGCCACCTTCTGCCACTTTGTAGATGCAAGATCTTGGGCAAGT TCCCGTGGGCGTTCACGGTGGTTTCCATGCGACGTGCAGAGGTG AAGCGAAGTG</u>	138
8	12	126230889	<u>CAGTGGAACAAAGCCACTGGGAAGTTCAAAGTGAAGAGAGAGCCCA CCACAAGTCTAGACTCTGTGGTATTGTGAGGATTTTTGTCAACAAG AAAAACCCC</u>	135
9	X	35911295	<u>AGTATATCATCAGTTATTTTTCAAGGTTTTCTAAGTAAACAGTTTCT CAACCTTTACCCCGTTGCTCGGCAACGGCCTGGTCTGTGCCAAGT GTTTGCTG</u>	124
10	10	75397400	<u>TCAGGGAGGGGATGTTGACTGCATTTTGGAGGTTTCAGGGCCTACT AACAAGTGTGCCAGCAGCTCCTCCTCCTGCCTCCACCAATCGGCA GTCAGGAAGG</u>	58

La Tabla 6 muestra las secuencias diana identificadas en las muestras de ADNlc obtenidas de la sangre del sujeto 3. Las muestras de suero del sujeto 3 se obtienen durante la ablación del tumor. Como se muestra, cada secuencia contiene al menos secuencia del genoma vírico parcial (subrayada) y secuencia del genoma hospedante parcial y forma una unión virus-hospedante.

5

ES 2 687 251 T3

Tabla 6

Datos de unión del suero del sujeto 3				
N.º	N.º cromosoma hospedante	de Posición de integración	Secuencia de unión	Lecturas acumuladas
11	5	1295930	GGAAATGGAGCCAGGCGCTCCTGCTGGCCGCGCACCGGGCGCC TCACACCAGAACATCGCATCAGGACTCCTAGGACCCCTGCTCGTG <u>TTACAGGCGGGG</u>	153
12	8	111636420	TCAAGCAGAAAAACCATGAAGATTTAAAACTTGTAATATTTGAAT GTGGGCTCCACCCCAACAGTCCCCCGTGGGGAGGGGTGAACCCT <u>GGCCCGAAT</u>	52
13	21	47565536	CCCGGGACCGACCCAGGAAGAGCCAGGGGCCCGGGTGATCCC TGCGGGGGTCTGGCTTTCAGTTATATGGATGATGTGGTATTGGGG <u>GCCAAGTCTGTA</u>	27
14	21	28573066	AATGAAAATCTCATTGATTTTTCACTTATAGGTTTTACCTTAGAGCT CCTCCTCTGCCTAATCATCTCATGTTTCATGTCCTACTGTTCAAGCC <u>TCCAAGC</u>	25
15	7	87842849	AGAATTGATACCTAAGCTGAGCAGAAATGAGGCCGACCATGAAGT GAGTGCCTAATCATCTCATGTTTCATGTCCTACTGTTCAAGCCTCCA <u>AGCTGTGCC</u>	24
16	7	148503201	CGTAGGAAAGACAAGGTGGCATTGATGGAAAGCAGTAGTTTTTGA GCCCTTCGCAGACGAAGGTCTCAATCGCCGCGTCGCAGAAGATC <u>TCAATCTCGGG</u>	19
17	1	162277132	TTAAAAAGGAGTTTTGTTTGTAGTCTATTCACTCATTTC AAGGAAC ATAGAAGAAGA ACTCCCTCGCCTCGCAGACGAAGGTCTCAATCGC <u>CGCGTCGC</u>	16
18	12	125048731	CAGTTCCTGGCTCCAAGCTCCCTCAAAGATGCCAGCTGGCCT TTCCCAAAGGCCTTGTAAGTTGGCGAGAAAGTAAAGCCTGTTTT <u>GCTTGATAC</u>	15
19	7	30412226	ACATGCCCTTCACTTCAGCCTGATGCTCCTGGCATAAGCTCAGCA ATTTTGGAGTGCGAATCCACACTCCAAAAGACACCAAATATTCAAG <u>AACAGTTTC</u>	13

Datos de unión del suero del sujeto 3				
N.º	N.º de cromosoma hospedante	Posición de integración	Secuencia de unión	Lecturas acumuladas
20	13	84505952	<u>AATTTCCCCTGAATAGCTGCAGTACTCACAGACACACTGGATGCTACTCACCTCTGCCTAATCATCTCATGTTTCATGTCCTACTGTTCAAGCCTCCAAGC</u>	13

Como se ilustra en las Tablas 5 y 6, los patrones similares de las secuencias de la unión virus-hospedante identificadas en el ADN tumoral y el ADNlc muestran que el ADNlc en el suero deriva del ADN tumoral. Al enriquecer selectivamente los ADNlc diana en el suero, se identifican las uniones virus-hospedante para proporcionar información específica del tumor sobre el sujeto.

### 3.2 Uniones virus-hospedante específicas del tumor

En la Figura 3, se procesa el ADN genómico del Sujeto 1, el Sujeto 4 y el Sujeto 5 y se analiza mediante reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) y PCR cuantitativa. Se obtiene el ADN genómico (ADNg) de los tejidos tumorales y los tejidos no tumorales. Se identifica una secuencia de ADN quimérico en el ADNg tumoral y se selecciona en cada sujeto para servir como marcador para realizar las pruebas. Específicamente, la secuencia de ADN quimérico del Sujeto 1 que se usa en este análisis es GGTCTTACATAAGAGGACTCAGAAAATACTTTGTGATGAT (secuencia del genoma vírico subrayada), Sujeto 4 ACTTCAAAGACTGTGTGTTTCTAATTATTTGGGGGACAT, y Sujeto 5 GTAGGCATAAATTGGTCTGTACCTCACTTCCCTGCTTTCC. Se determina la presencia de las tres uniones virus-hospedante específicas en el ADNg tumoral (T) y el ADNg no tumoral (N). Se usan porfobilinógeno desaminasa (PBGD) y miR-122 como testigo interno. También se incluye un testigo que no es plantilla (NTC). Como se ilustra en la Figura 3, la unión virus-hospedante específica del Sujeto 1 solo está presente en el ADNg tumoral (T), pero no en el ADNg no tumoral (N). Los mismos patrones se observan en el Sujeto 4 y el Sujeto 5, lo cual indica que las uniones virus-hospedante identificadas son específicas del tumor y se pueden usar como biomarcadores específicos del tumor.

### 3.3 Desarrollo tumoral y cantidad de uniones virus-hospedante

La Figura 4 muestra la relación entre el tamaño del tumor y la cantidad de secuencias de unión virus-hospedante específicas. Las secuencias de unión usadas en la Figura 4 para cada sujeto son las mismas que en la Figura 3. Se obtienen muestras de sangre en serie de cada sujeto al menos en el preoperatorio y en el posoperatorio. Con referencia a la Figura 4, ADNg se refiere al ADN genómico, NTC se refiere al testigo que no es plantilla, NT se refiere al ADNg de tejido no tumoral, T se refiere al ADNg de tejido tumoral, NA sérico se refiere al ADN obtenido del suero, Pre-OP se refiere al ADN sérico obtenido en el preoperatorio, Post-OP se refiere al ADN sérico obtenido en el posoperatorio, Sujeto 1\* se refiere a ADN sérico obtenido del Sujeto 1 y que se usa en el experimento con el Sujeto 5, Sujeto 4\* se refiere al ADN sérico obtenido del Sujeto 4 y que se usa en el experimento con el Sujeto 1, Sujeto 5\* se refiere al ADN sérico obtenido del Sujeto 5 y que se usa en el experimento con el Sujeto 4 y Normal se refiere al ADN sérico obtenido de un sujeto que no es paciente. Las muestras de suero del Sujeto 1 se obtienen en dos puntos de tiempo, 13 días antes de la ablación del tumor (operación) y 19 días después de la operación. Las muestras de suero del Sujeto 4 se obtienen 33 días antes de la operación y 30 días después de la operación. Las muestras de suero del Sujeto 5 se obtienen 24 días antes de la operación y 26 días después de la operación. Las muestras de suero del sujeto no paciente (Normal) también se incluyen como testigo en la Figura 4. Como se muestra en la Figura 4, en el panel izquierdo, la unión virus-hospedante específica de cada sujeto solo está presente en el ADNg tumoral y Pre-OP. Además, la unión virus-hospedante específica del Sujeto 1 solo está presente en las muestras de ADN del Sujeto 1, pero no en las muestras de ADN del Sujeto 4, lo cual sugiere que la unión virus-hospedante identificada es específica del sujeto. En referencia ahora al panel derecho de la Figura 4, la unión virus-hospedante específica en el suero del Sujeto 1 se detecta en una cantidad relativamente grande en el suero Pre-OP, mientras que la cantidad disminuye marcadamente en el suero Post-OP. La cantidad de la unión virus-hospedante específica en el suero Pre-OP y el suero Post-OP se determina mediante qPCR y se presenta en el panel derecho de la Figura 4. En el Sujeto 1, la cantidad de unión virus-hospedante específica en el suero Post-OP disminuye aproximadamente 32 veces en comparación con la del suero Pre-OP. Los mismos patrones se observan en el Sujeto 4 y el Sujeto 5, lo que muestra que las uniones virus-hospedante o la cantidad de uniones son específicas del tumor, específicas del sujeto, detectables en suero, reflejan el tumor y corresponden a cambios en el tamaño del tumor, tal como la disminución en el tamaño después de la operación.

Cabe señalar que, mediante el uso de la estrategia descrita en la presente invención, no se pueden detectar los

genes mutados p53 o beta-catenina en los ADNlc a pesar de que se identificaron las mutaciones en el tejido tumoral (no se muestran los datos). El resultado muestra que, mediante el uso del método de la presente invención, se enriquecen selectivamente y obtienen las uniones virus-hospedante específicas del tumor (inserción de secuencia del genoma vírico en el genoma hospedante) y mutaciones somáticas no convencionales para proporcionar información sobre el cáncer/tumor.

5

Las realizaciones que se mostraron y describieron anteriormente son solo ejemplos. Aunque se establecieron numerosas características y ventajas de la presente tecnología en la descripción precedente, junto con los detalles de la estructura y función de la presente descripción, la descripción es solamente ilustrativa y se pueden hacer cambios en los detalles, incluso en la forma, el tamaño y la disposición de las partes dentro de los principios de la presente descripción hasta, e incluso, la extensión total establecida por el significado general amplio de los términos usados en las reivindicaciones.

10

**REIVINDICACIONES**

1. Un método para identificar una cantidad de uniones virus-hospedante en los ADNlc de un sujeto infectado por el virus de la hepatitis B que comprende:
- obtener los ADNlc en un suero obtenido del sujeto;
- 5 ligar cada ADNlc con al menos un adaptador;
- amplificar todos los ADNlc ligados con el al menos un adaptador en una proporción similar mediante el uso de una pluralidad de cebadores, en donde cada uno de los cebadores es complementario a una secuencia del adaptador correspondiente;
- hibridar las sondas polinucleotídicas con los ADNlc ligados con los adaptadores;
- 10 capturar y aislar los ADNlc diana hibridados con las sondas polinucleotídicas;
- secuenciar los ADNlc diana, en donde los ADNlc diana tienen una o más secuencias de unión virus-hospedante, en donde cada secuencia de unión virus-hospedante comprende una secuencia del genoma vírico y una secuencia del genoma hospedante; y
- 15 acumular un número de lecturas relacionado con las uniones virus-hospedante en los ADNlc diana para identificar la cantidad de uniones virus-hospedante, en donde la cantidad de uniones virus-hospedante se define como el número de lecturas acumuladas, el número de lecturas se define como las veces en que la secuencia de la unión virus-hospedante aparece en los ADNlc diana, en donde las secuencias de la unión virus-hospedante detectadas comprenden las mismas secuencias, secuencias diferentes o una combinación de estas, y cada secuencia de la unión virus-hospedante se cuenta para proporcionar el número de lecturas; y en donde la cantidad de las uniones virus-hospedante se configura para indicar el estado tumoral relacionado con el virus de la hepatitis B.
- 20 2. El método de la reivindicación 1, que comprende, además, comparar la cantidad detectada de las uniones virus-hospedante del sujeto en diferentes puntos de tiempo, para identificar cambios en el tamaño del tumor.
3. El método de la reivindicación 1, en donde los ADNlc diana se enriquecen mediante las sondas polinucleotídicas complementarias a al menos una secuencia derivada del genoma vírico de la hepatitis B.
- 25 4. El método de la reivindicación 1, en donde las sondas polinucleotídicas abarcan todas las secuencias del genoma vírico de la hepatitis B.
5. El método de la reivindicación 1, en donde los ADNlc que se ligan con el adaptador correspondiente comprenden un extremo de cada ADNlc que se liga con el adaptador correspondiente, ambos extremos de cada ADNlc que se ligan con el adaptador correspondiente y una combinación de estos.
- 30 6. Un método para monitorizar un estado tumoral en un sujeto causado por la infección por el virus de la hepatitis B, que comprende:
- obtener los ADNlc en un suero obtenido del sujeto;
- ligar cada ADNlc con al menos un adaptador;
- 35 amplificar todos los ADNlc ligados con el al menos un adaptador en una proporción similar mediante el uso de una pluralidad de cebadores, en donde cada uno de los cebadores es complementario a una secuencia del adaptador correspondiente;
- hibridar las sondas polinucleotídicas con los ADNlc ligados con los adaptadores;
- capturar y aislar los ADNlc diana hibridados con las sondas polinucleotídicas;
- 40 secuenciar los ADNlc diana, en donde los ADNlc diana tienen una o más secuencias de unión virus-hospedante, en donde cada secuencia de unión virus-hospedante comprende una secuencia del genoma vírico y una secuencia del genoma hospedante;
- 45 acumular un número de lecturas relacionado con las uniones virus-hospedante en los ADNlc diana para identificar la cantidad de uniones virus-hospedante, en donde la cantidad de uniones virus-hospedante se define como el número de lecturas acumuladas, el número de lecturas se define como las veces en que la secuencia de la unión virus-hospedante aparece en los ADNlc diana y en donde las secuencias de la unión virus-hospedante detectadas comprenden las mismas secuencias, secuencias diferentes o una combinación de estas, y cada secuencia de la unión virus-hospedante se cuenta para proporcionar el número de lecturas; y
- monitorizar el estado tumoral en función de la cantidad de uniones virus-hospedante en los ADNlc diana.

7. El método de la reivindicación 6, en donde las sondas polinucleotídicas abarcan todas las secuencias del genoma vírico de la hepatitis B.
8. El método de la reivindicación 6, en donde los ADNlc que se ligan con el adaptador correspondiente comprenden un extremo de cada ADNlc que se liga con el adaptador correspondiente, ambos extremos de cada ADNlc que se ligan con el adaptador correspondiente y una combinación de estos.
9. El método de la reivindicación 6, en donde el tumor es un carcinoma hepatocelular inducido por el virus de la hepatitis B.
10. El método de la reivindicación 6, en donde los ADNlc se obtienen en diferentes puntos de tiempo, y en donde los diferentes puntos de tiempo se seleccionan a partir de una condición cancerosa, una condición pretratamiento, una condición postratamiento, una condición de recidiva del sujeto y cualquier combinación de estas.
11. El método de la reivindicación 10, que comprende, además, monitorizar el desarrollo tumoral del sujeto al comparar la cantidad detectada de las uniones virus-hospedante del sujeto en diferentes puntos de tiempo.
12. El método de la reivindicación 10, en donde los aumentos en la cantidad de la al menos una unión virus-hospedante en los ADNlc del sujeto indican que el desarrollo tumoral pasó de una condición postratamiento a una condición de recidiva y las disminuciones en la cantidad de la al menos una unión virus-hospedante en los ADNlc del sujeto indican que el desarrollo tumoral pasó de una condición pretratamiento a una condición postratamiento del sujeto.
13. El método de la reivindicación 10, en donde los aumentos en la cantidad de la al menos una unión virus-hospedante en los ADNlc del sujeto en la condición cancerosa indican el crecimiento de un tumor y las disminuciones en la cantidad de la al menos una unión virus-hospedante en los ADNlc circulantes del sujeto en la condición cancerosa indican un encogimiento del tumor.
14. Un método para diagnosticar una enfermedad en un sujeto infectado por el virus de la hepatitis B que comprende:
- detectar una cantidad de uniones virus-hospedante en ADNlc del sujeto, que comprende:
- obtener los ADNlc en un suero obtenido del sujeto;
- ligar cada ADNlc con al menos un adaptador;
- amplificar todos los ADNlc ligados con el al menos un adaptador en una proporción similar mediante el uso de una pluralidad de cebadores, en donde cada uno de los cebadores es complementario a una secuencia del adaptador correspondiente;
- hibridar las sondas polinucleotídicas con los ADNlc ligados con los adaptadores;
- capturar y aislar los ADNlc diana hibridados con las sondas polinucleotídicas;
- secuenciar los ADNlc diana, en donde los ADNlc diana tienen una o más secuencias de unión virus-hospedante, en donde cada secuencia de unión virus-hospedante comprende una secuencia del genoma vírico y una secuencia del genoma hospedante;
- acumular un número de lecturas relacionado con las uniones virus-hospedante en los ADNlc diana para identificar la cantidad de uniones virus-hospedante, en donde la cantidad de uniones virus-hospedante se define como el número de lecturas acumuladas, el número de lecturas se define como las veces en que la secuencia de la unión virus-hospedante aparece en los ADNlc diana y en donde las secuencias de la unión virus-hospedante detectadas comprenden las mismas secuencias, secuencias diferentes o una combinación de estas, y cada secuencia de la unión virus-hospedante se cuenta para proporcionar el número de lecturas; y
- diagnosticar la enfermedad en el sujeto infectado por el virus de la hepatitis B en función de la cantidad de uniones virus-hospedante en los ADNlc diana.
15. El método de la reivindicación 14, en donde la enfermedad es un cáncer causado por la infección crónica por el virus de la hepatitis B.

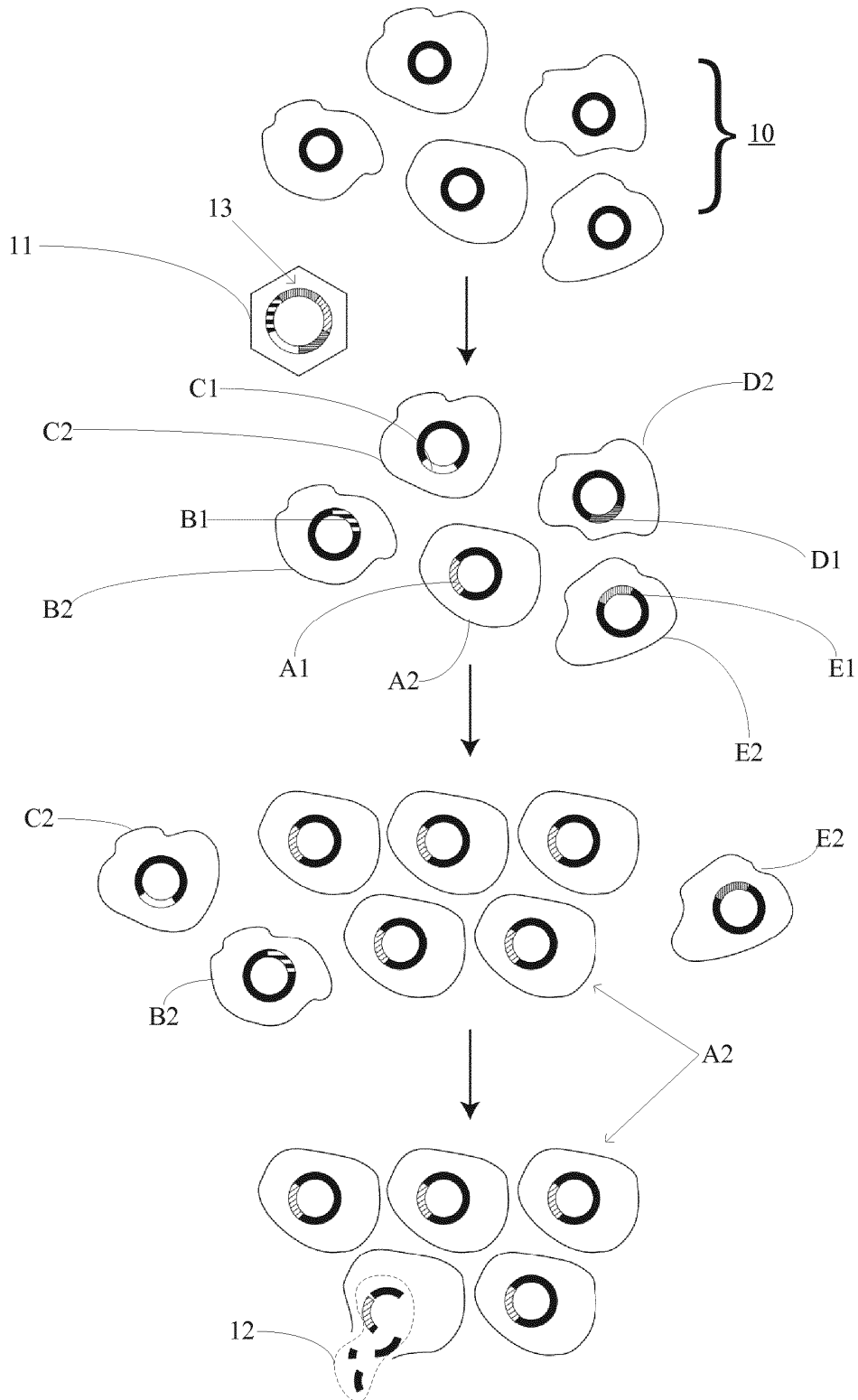


FIG.1



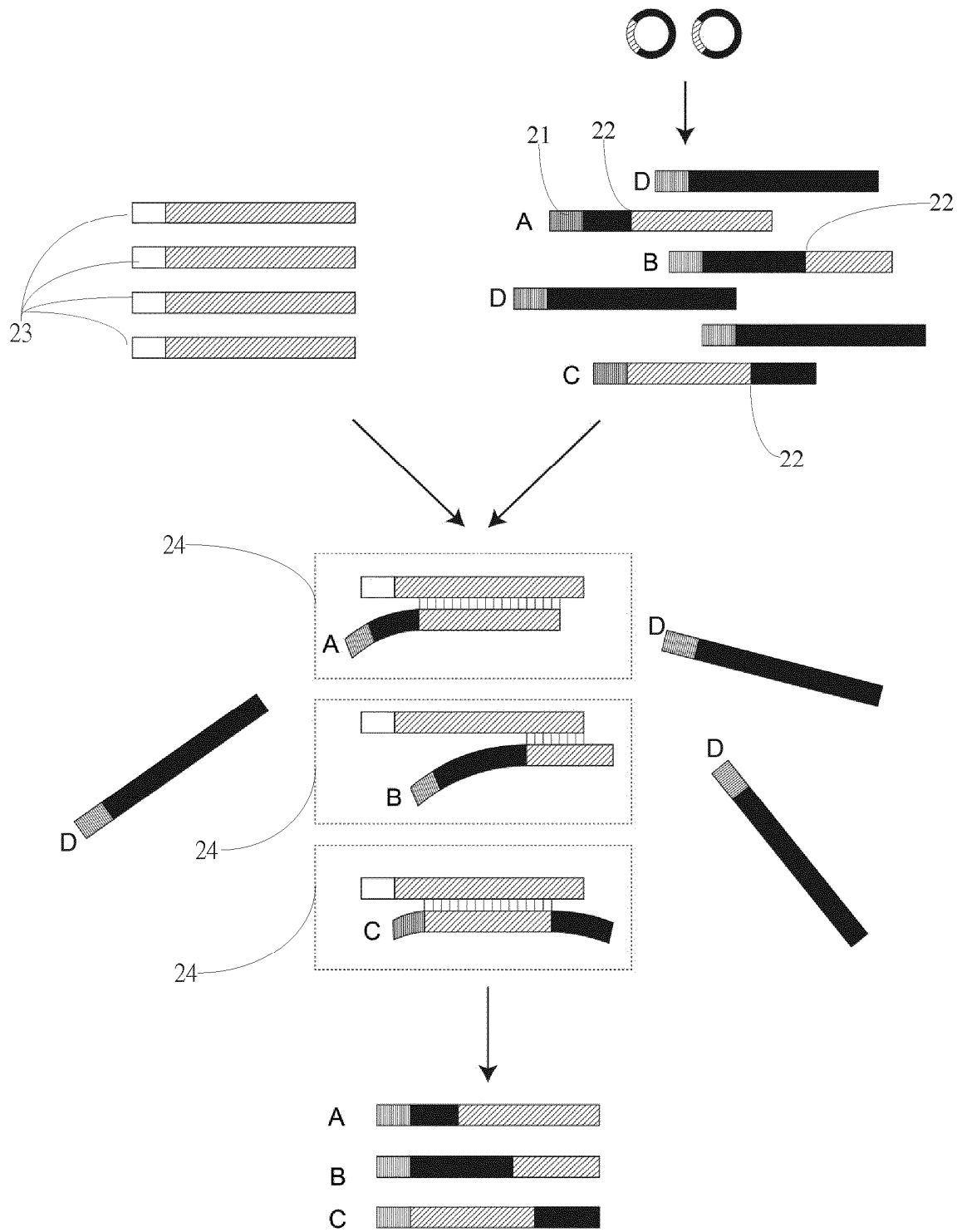


FIG.2

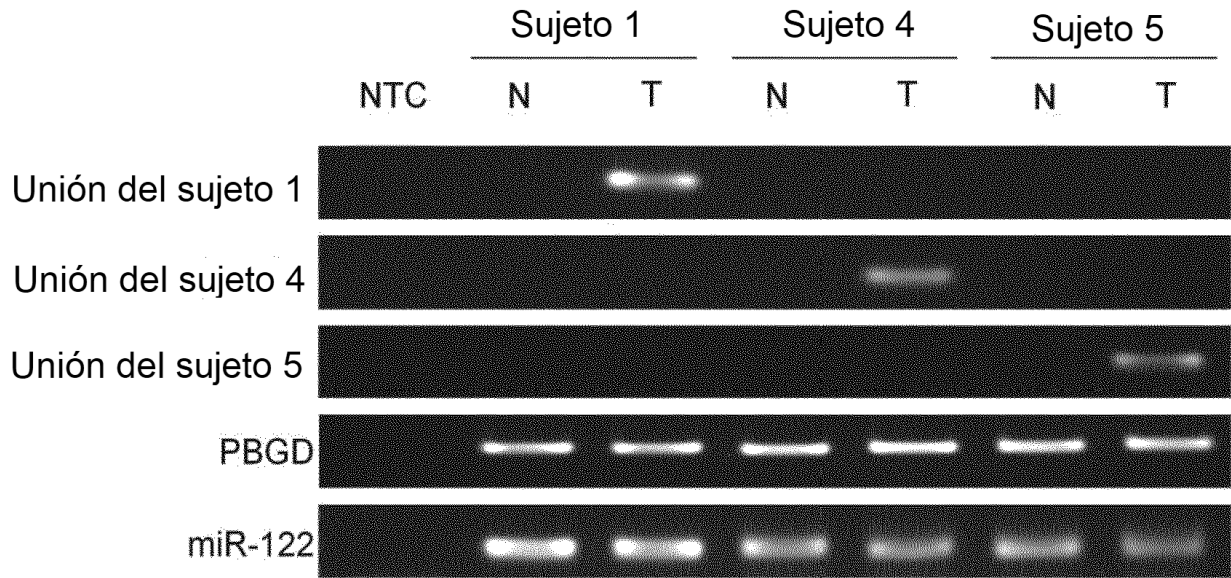


FIG.3

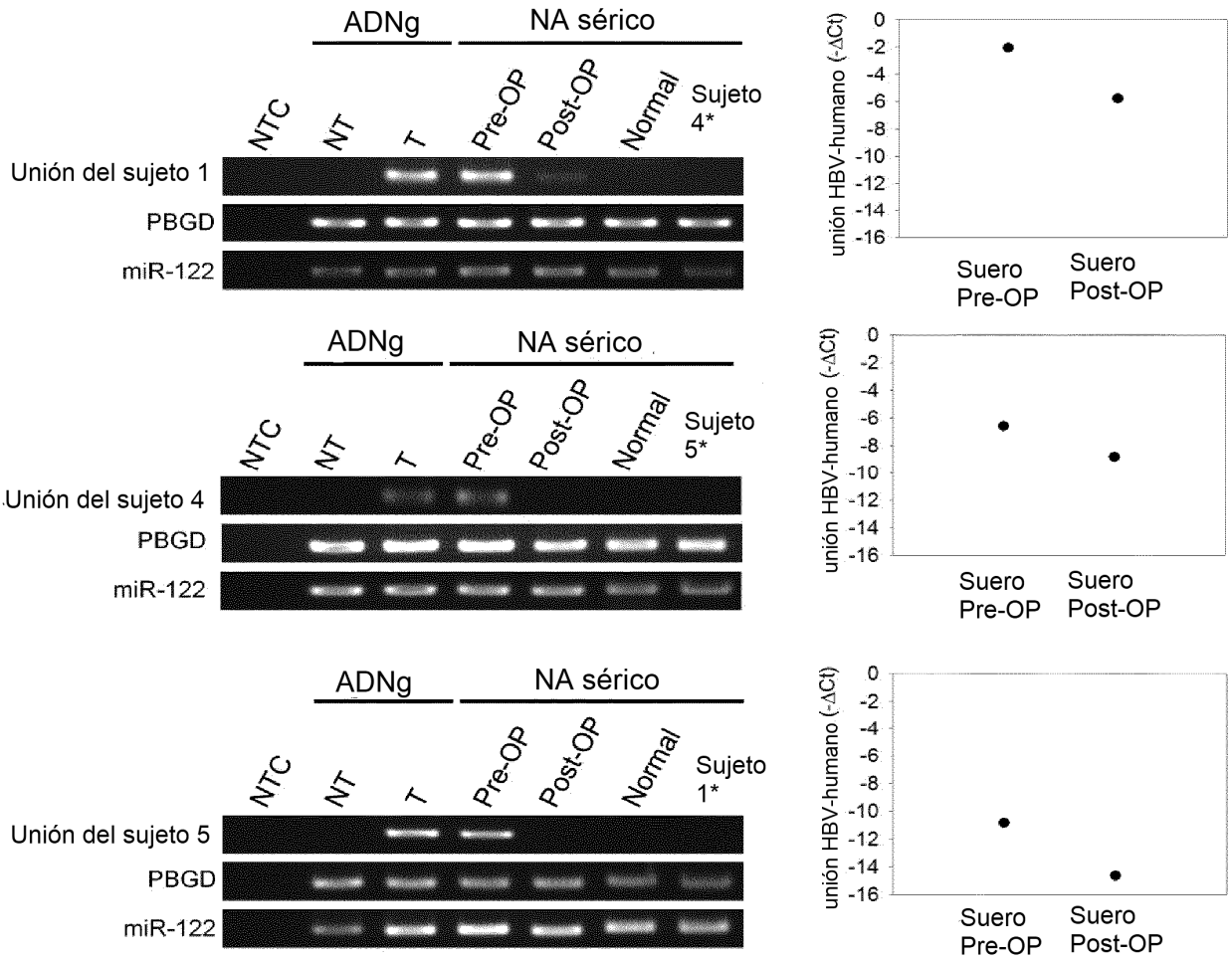


FIG.4