

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 971 286**

51 Int. Cl.:

**C12Q 1/6827** (2008.01)

**C12Q 1/6886** (2008.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.11.2017 PCT/GB2017/053615**

87 Fecha y número de publicación internacional: **07.06.2018 WO18100381**

96 Fecha de presentación y número de la solicitud europea: **30.11.2017 E 17817030 (4)**

97 Fecha y número de publicación de la concesión europea: **24.01.2024 EP 3548633**

54 Título: **Aplicación de interacciones cromosómicas epigenéticas en diagnóstico de cáncer**

30 Prioridad:

**01.12.2016 US 201662428893 P**

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

**04.06.2024**

73 Titular/es:

**OXFORD BIODYNAMICS PLC (100.0%)  
3140 Rowan Place, John Smith Drive, Oxford  
Business Park South  
Oxford, Oxfordshire OX4 2WB, GB**

72 Inventor/es:

**GRAND, FRANCIS HECTOR;  
RAMADASS, AROUL SELVAM;  
AKOULITCHEV, ALEXANDRE y  
HUNTER, EWAN**

74 Agente/Representante:

**FERNÁNDEZ POU, Felipe**

ES 2 971 286 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

Aplicación de interacciones cromosómicas epigenéticas en diagnóstico de cáncer

### 5 Campo de la invención

La invención se refiere a la detección de interacciones cromosómicas.

Antecedentes de la invención

10

La población mundial ha crecido hasta esperar servicios rápidos y eficaces en todos los aspectos de la tecnología, y particularmente en el campo de la medicina. Las mejoras en terapia personalizada a menudo se basan en el uso de un biomarcador tal como proteína, ARNm, marcadores de anticuerpo y SNP.

### 15 Compendio de la invención

Los inventores han investigado el papel de las interacciones cromosómicas en el pronóstico de cáncer. Su trabajo muestra que las interacciones cromosómicas específicas están asociadas a resultados particulares, tales como la capacidad de respuesta a la terapia. Como parte de este trabajo, los inventores también han desarrollado un nuevo ensayo epigenético que es particularmente adecuado para la detección de ácidos nucleicos ligados que se producen durante el transcurso de la detección de una interacción cromosómica.

20

Por consiguiente, la invención proporciona un método para determinar las interacciones cromosómicas epigenéticas que son relevantes para un ensayo epigenético con acompañante de pronóstico para cáncer que distingue entre subgrupos, que comprende poner en contacto un primer conjunto de ácidos nucleicos procedentes de los subgrupos con un segundo conjunto de ácidos nucleicos que representan una población índice de interacciones cromosómicas, y permitir que las secuencias complementarias se hibriden, en donde los ácidos nucleicos del primer y segundo conjuntos de ácidos nucleicos representan un producto ligado que comprende secuencias procedentes de ambas regiones cromosómicas que se han unido en la interacción cromosómica epigenética, y en donde el patrón de hibridación entre el primer y segundo conjunto de ácidos nucleicos permite la determinación de qué interacciones cromosómicas epigenéticas son específicas para subgrupos en la población, en donde los subgrupos difieren en al menos una característica relacionada con el pronóstico. Normalmente, los subgrupos difieren con respecto a la sensibilidad de respuesta a terapia con inhibidor de tirosina quinasa, y/o el cáncer está asociado a la expresión anormal de un gen de quinasa y/o un gen receptor de tirosina quinasa, y/o el cáncer está asociado a un cambio genético o epigenético en uno o más de los siguientes genes: FIP1L1, PDGFRA, Flt3, ABL1, FGFR1 y cKIT.

25

30

35

La invención también proporciona un método para determinar el pronóstico para un cáncer que comprende: (a) tipificar una interacción cromosómica que se ha identificado por medio del método anterior, y/o (b) detectar la presencia o ausencia de al menos una interacción cromosómica epigenética que es relevante para (i) capacidad de respuesta a la terapia con inhibidor de tirosina quinasa, y/o (ii) expresión anormal de un gen de quinasa y/o un gen receptor de tirosina quinasa, y/o (iii) uno o más de los siguientes genes: FIP1L1, PDGFRA, Flt3, ABL1, FGFR1 y cKIT.

40

La invención proporciona además un método para detectar cuantitativamente una secuencia ligada que es relevante para una interacción cromosómica usando una sonda que se puede detectar tras activación durante una reacción de PCR, en donde dicha secuencia ligada comprende secuencias de dos regiones cromosómicas que se unen en una interacción cromosómica epigenética, en donde dicho método comprende poner en contacto la secuencia ligada con la sonda durante una reacción de PCR, y detectar el grado de activación de la sonda, y en donde dicha sonda se une al sitio de enlace. El método permite la detección de interacciones particulares de manera compatible con MIQE usando una sonda de hidrólisis fluorescente de doble marcado.

45

50

El trabajo del solicitante ofrece por primera vez a la industria un ensayo de conformación cromosómica (3C) compatible con MIQE para determinar el pronóstico en una situación oncológica.

55

Breve Descripción de los Dibujos

La Figura 1 es un diagrama esquemático del proceso 3C. 3C significa captura de conformación de cromatina, o captura de conformación de cromosoma.

60

La Figura 2 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 1. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 9 en el eje Y de la parte izquierda) y la capacidad de plegado en u.a. se muestra en gris (Escala de 4 a 6,5 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 500 se muestra en el eje X.

65

La Figura 3 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 2. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 12 en el eje Y de la parte izquierda) y la capacidad de plegado en u.a. se muestra en gris (Escala de 3 a 6,5 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 700 se muestra en el eje X.

La Figura 4 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 3. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 9 en el eje Y de la parte izquierda), y la capacidad de plegado en u.a. se muestra en gris (Escala de 4 a 6 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 450 se muestra en el eje X.

La Figura 5 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 4. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 10 en el eje Y de la parte izquierda) y la capacidad de plegado en u.a. se muestra en gris (Escala de 3 a 7 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 500 se muestra en el eje X.

La Figura 6 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 5. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 9 en el eje Y de la parte izquierda) y la capacidad de plegado en u.a. se muestra en gris (Escala de 3 a 6,5 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 700 se muestra en el eje X.

La Figura 7 es un gráfico de curvatura y capacidad de plegado frente a la posición de secuencia para la secuencia ligada 6. La curvatura prevista en grados por cada 10,5 pb de giro helicoidal se muestra en negro (Escala de 0 a 12 en el eje Y de la parte izquierda) y la capacidad de plegado en u.a. se muestra en gris (Escala de 3 a 6,5 en el eje Y de la parte derecha). El número de secuencia en pares de bases de 0 a 600 se muestra en el eje X.

Las Figuras 8 y 9 muestran datos relativos a realizaciones terapéuticas.

#### Descripción detallada de la invención

La invención se refiere a identificar procesos aberrantes particulares, en particular con un ensayo cuantitativo robusto. La invención tiene diversos aspectos diferentes, incluyendo entre otros:

- un método para identificar interacciones cromosómicas epigenéticas relevantes para diferentes subgrupos
- un método para identificar el subgrupo de un individuo
- un método para detectar cuantitativamente una secuencia ligada que es relevante para una interacción cromosómica
- una sonda de hidrólisis que se puede detectar tras activación durante una reacción de PCR en la que la sonda se une al sitio de enlace de un ácido nucleico ligado
- un método para seleccionar cebadores y una sonda para su uso en una reacción de PCR
- un método para tratamiento de cáncer
- un ácido nucleico para su uso en el tratamiento de cáncer
- un método para identificar un agente terapéutico.

#### Cáncer relevante para la invención

La invención se refiere a interacciones epigenéticas relevantes para cáncer, por ejemplo cáncer provocado por diferentes mecanismos genéticos, tales como mutaciones en enzimas (incluyendo aquellas responsables de desmetilación de ADN), o causados por corte y empalme génico alternativo o mutaciones de punto de ruptura. El cáncer puede estar provocado por regulación por aumento de un receptor. El cáncer puede estar asociado a la expresión anormal de un gen de quinasa y/o un gen receptor de tirosina quinasa, por ejemplo, un gen de quinasa citoplasmática y/o un gen de serina/treonina quinasa. El cáncer puede estar asociado a un cambio genético o epigenético en uno o más de los siguientes genes: FIP1L1, PDGFRA, Flt3, ABL1, FGFR1 y cKIT. El cáncer puede estar asociado a un cambio genético o epigenético en proteínas que sean sensibles a los

miméticos de ATP, opcionalmente genes de tirosina quinasa, tales como ABL1, tirosina quinasa de Bruton, cKIT, FGFR1, PDGFRA y PDGFRB. Los genes de tirosina quinasa (ABL1, Brutons, cKIT, FGFR1 y PDGFRA) se clasifican frecuentemente como "oncogenes", funcionalmente cuando se expresan. El cambio genético o epigenético en proteínas que son sensibles a los miméticos de ATP se puede explicar en base a un cambio en la expresión de un gen que provoca la regulación por aumento de una proteína que es sensible a los miméticos de ATP. Por ejemplo, se cree que FIP1L1 regula por aumento directamente la expresión de ARNm de PDGFRA.

El cáncer puede ser leucemia, neoplasia mieloproliferativa, tumor, cáncer de pulmón o cáncer de ovario. El cáncer puede estar seleccionado entre síndrome hipereosinofílico idiopático (IHES), leucemias eosinofílicas crónicas y agudas, leucemia mieloide crónica (CML), leucemia mieloide aguda (AML), leucemia linfocítica aguda (ALL), carcinoma de pulmón no microcítico (NSCLC) PDGFRA-positivo, glioblastoma (incluyendo glioblastoma multiforme y astrocitoma), cáncer de próstata, cáncer de ovario avanzado, tumor estromal gastrointestinal (GIST) y carcinoma de células renales.

#### Interacciones Cromosómicas que se Tipifican

La invención se refiere a la tipificación de interacciones cromosómicas, por ejemplo, las asociadas a cáncer y/o capacidad de respuesta a terapia con inhibidor de tirosina quinasa. Las interacciones cromosómicas afectan o se asocian a un aspecto de la enfermedad, tal como en cualquiera de las afecciones patológicas mencionadas en la presente memoria. Las interacciones cromosómicas pueden afectar o estar asociadas a la susceptibilidad a la enfermedad, capacidad de respuesta a la terapia o probabilidad de recaída. Las interacciones cromosómicas pueden afectar o estar asociadas al pronóstico para un paciente y/o subgrupo de pacientes. En la presente memoria se divulgan interacciones cromosómicas específicas, genes y regiones, y en una realización se pueden tipificar esas interacciones cromosómicas, o interacciones cromosómicas en esos genes y regiones. Las regiones preferidas se divulgan a modo de intervalos en la Tabla 1a. Las secuencias ligadas específicas se divulgan en la presente memoria. La tipificación de interacciones cromosómicas puede requerir la detección de dichas secuencias ligadas u homólogos y/o fragmentos de las mismas.

Las interacciones cromosómicas, genes y regiones preferidos (ubicaciones genómicas) se muestran en las tablas, incluyendo las tablas 5, 6, 7, 8, 9, 10, 11, 12, 13 y 14. Otras interacciones cromosómicas preferidas son las que se pueden definir o detectar usando cualquier ácido nucleico mencionado en la presente memoria, por ejemplo, sondas o pares de cebadores.

La interacción cromosómica que se tipifica puede ser o no una que tenga lugar entre un gen (incluyendo la secuencia codificante) y su región reguladora, tal como un promotor.

La interacción cromosómica puede estar asociada a la regulación por disminución de un gen, por ejemplo, un cambio en la actividad (tal como un aumento o disminución) en un potenciador. La interacción cromosómica puede estar asociada a una mutación, tal como una mutación de punto de ruptura, que conduce a un producto genómico de fusión (cromosómico) (que puede conducir o no a la expresión de una proteína de fusión que tiene actividad modificada).

La mutación de punto de ruptura puede ser:

- cualquier mutación mencionada en la presente memoria, o
- dentro de, abarcando o incluyendo cualquier gen mencionado en la presente memoria, o
- una mutación de punto de ruptura que está asociada a o provoca cáncer.

La mutación de punto de ruptura puede ser al menos 1.000 nucleótidos, 10.000 nucleótidos, 50.000 nucleótidos, 100.000 nucleótidos, 500.000 nucleótidos, tal como al menos 800.000 nucleótidos de longitud, que representan la longitud del fragmento eliminado.

#### Pronóstico

El pronóstico como se usa en la presente memoria se refiere a la predicción del trascurso probable de una afección médica, tal como uno o más resultados. Un factor pronóstico es una característica clínica o biológica que se puede medir objetivamente y que proporciona información sobre el resultado probable de la enfermedad de cáncer, por ejemplo, en un individuo tratado o no tratado. El factor predictivo puede proporcionar información sobre el beneficio probable del tratamiento (por ejemplo, ya sea en términos de contracción del tumor o supervivencia). Dichos factores predictivos se pueden usar para identificar subpoblaciones de pacientes que tienen más probabilidades de beneficiarse de una terapia concreta. El marcador puede tener propiedades relacionadas con el pronóstico, así como ser predictivo del probable beneficio del tratamiento.

## Quinasas

La invención puede hacer referencia a un cambio en un gen de quinasa o gen receptor de quinasa, por ejemplo, un cambio en la expresión y/o un cambio en la forma epigenética y/o un cambio genético (secuencia genómica), y esto puede estar asociado a cáncer y/o provocar cáncer. Normalmente, por tanto, el cáncer puede estar asociado a la expresión anormal de un gen de quinasa y/o un receptor de tirosina quinasa. En una realización, el cáncer está asociado a la expresión anormal de tirosina quinasa citoplásmica (por ejemplo, ABL1) y/o quinasas citoplásmicas (por ejemplo, una serina/treonina quinasa). La expresión "expresión anormal" puede incluir regulación por aumento o regulación por disminución o puede incluir la expresión de una proteína diferente.

"Quinasa" en el presente contexto se puede definir funcionalmente como que experimenta un cambio conformacional durante la transferencia de ATP a otras proteínas, en el núcleo, citoplasma o en la membrana celular, o cualquier otra ubicación celular que incluya ubicaciones extracelulares. Ejemplos de familia de quinasas que sean relevantes incluyen tirosina quinasas receptoras (por ejemplo, PDGFRA) que están en la superficie celular y quinasas no receptoras, tales como tirosina quinasas citoplásmicas (por ejemplo ABL1) y serina/treonina quinasas.

## Interacciones Epigenéticas

Tal como se usa en la presente memoria, el término interacciones "epigenéticas" hace referencia normalmente a interacciones entre regiones distales en un cromosoma, siendo dichas interacciones en una realización dinámicas y alterando, formando o rompiendo, dependiendo del estado de la región del cromosoma.

En métodos particulares de la invención, las interacciones cromosómicas se detectan generando primero un ácido nucleico ligado que comprende una secuencia de ambas regiones de los cromosomas que son parte de las interacciones. En dichos métodos, las regiones se pueden reticular por cualquier medio adecuado. En una realización preferida, las interacciones se reticulan usando formaldehído, pero también se pueden reticular mediante cualquier aldehído, o éster de N-hidroxisuccinimida de ácido D-biotinoil-e-aminocaproico o éster de N-hidroxisuccinimida de ácido digoxigenin-3-O-metilcarbonyl-e-aminocaproico. Se puede usar paraformaldehído y puede entrecruzar cadenas de ADN que están separadas por 4 Angstroms.

La interacción cromosómica puede reflejar el estado de la región del cromosoma, por ejemplo, si se está transcribiendo o reprimiendo en respuesta al cambio de las condiciones fisiológicas. Se ha encontrado que las interacciones cromosómicas que son específicas de los subgrupos como se definen en la presente memoria son estables, proporcionando de este modo un medio fiable para medir las diferencias entre los dos subgrupos.

Además, las interacciones cromosómicas específicas de una característica (tal como una afección patológica) se producen normalmente al principio de un proceso biológico, por ejemplo en comparación con otros marcadores epigenéticos tales como metilación o cambios en la unión de proteínas de histona. De este modo, el método de la invención es capaz de detectar etapas tempranas de un proceso biológico. Esto permite una intervención temprana (por ejemplo, tratamiento) que, como consecuencia, puede resultar más eficaz. Además, existe escasa variación en las interacciones cromosómicas relevantes entre individuos dentro del mismo subgrupo.

## Localización y Causas de las Interacciones Epigenéticas

Las interacciones cromosómicas epigenéticas se pueden solapar e incluir las regiones de cromosoma que se ha demostrado que codifican genes relevantes o no descritos, pero igualmente pueden estar en regiones intergénicas. Las interacciones cromosómicas que se detectan en la invención podrían estar provocadas por cambios en la secuencia de ADN subyacente, por factores ambientales, metilación de ADN, transcripciones de ARN antisentido no codificantes, carcinógenos no-mutagénicos, modificaciones de histona, remodelación de cromatina e interacciones de ADN local específicas. La misma interacción cromosómica puede ser provocada por diferentes factores en diferentes sujetos. Por ejemplo, la interacción cromosómica particular puede estar provocada por cambios en la metilación de ADN, por ejemplo, debido a la mutación en IDH en una región diferente del cromosoma desde donde se observa la interacción. La misma interacción cromosómica puede estar provocada por un cambio en la secuencia subyacente, por ejemplo, eliminación en la misma región del cromosoma que donde se observa la interacción.

Los cambios que conducen a interacciones cromosómicas pueden estar provocados por cambios en la secuencia de ácido nucleico subyacente, que por sí mismos no afectan directamente al producto génico o modo de expresión génica. Dichos cambios pueden ser, por ejemplo, SNP dentro y/o fuera de los genes, fusiones génicas y/o eliminaciones de ADN intergénico, microARN y ARN no codificante. Por ejemplo, se sabe que aproximadamente un 20 % de los SNPs están en regiones no codificantes, y por tanto, el método que se describe también es informativo en una situación no codificante. Las regiones de cromosoma que se unen para formar la interacción pueden ser menores que 5 kb, 3 kb, 1 kb, 500 pares de bases o 200 pares de bases

separadas en el mismo cromosoma. Las regiones de cromosoma que se unen para formar la interacción pueden estar separadas por más de 5 kb, 10 kb, 50 kb, 100 kb, 200 kb, 500 kb o más de 800 kb en el mismo cromosoma, y pueden estar separadas por ejemplo de 800 kb a 1.200 kb, tal como de 900 kb a 1.000 kb. Cualquiera de dichos valores e intervalos divulgados en la presente memoria puede hacer referencia a un cromosoma normal (sin mutación de punto de ruptura) o a un cromosoma con una mutación de punto de ruptura.

La interacción cromosómica que se detecta está preferentemente dentro de cualquiera de los genes mencionados en la presente memoria. Sin embargo, también puede estar aguas arriba o aguas abajo de los genes, por ejemplo, hasta 50.000, hasta 30.000, hasta 20.000, hasta 10.000 o hasta 5.000 bases aguas arriba o aguas abajo del gen o de la secuencia codificante. Se considera la técnica anterior importante que es el documento de patente WO2017/106290, que muestra un método para detectar cáncer en un sujeto que comprende detectar en una muestra biológica obtenida a partir de la topología de cromatina alterada del sujeto dentro de una región de cromatina, que comprende un dinucleótido CpG dentro de un motivo de unión a CTCF, y en donde la región de cromatina comprende un oncogén y en donde la detección de la topología de cromatina alterada es indicativa de cáncer en dicho sujeto. El método comprende detectar la topología de cromatina alterada por medio de análisis de 5-metil citosina en el dinucleótido CpG y en donde el oncogén es PDGFRA, FIP1L1 o FGFR2. El documento WO2017/106290 comprende este método que se va a aplicar para cáncer, sin embargo, no menciona ni comprende las ubicaciones cromosómicas específicas de las Tablas 1-14 ni ningún número de identidad de secuencia especificado.

#### Tipos de Situación Clínica

El objetivo de la presente invención es permitir la detección de interacciones cromosómicas relevantes para una característica que define subgrupos en la población. Por ejemplo, esta tecnología permite la estratificación basada en biomarcadores para fenotipos específicos (por ejemplo, relacionados con afecciones médicas), es decir, reconocimiento de una firma de conformación cromosómica particular y/o cambio en esa firma particular. La característica que define el subgrupo es relevante para el pronóstico de cáncer. Los métodos de la invención se pueden usar en el contexto de características específicas relacionadas con la enfermedad, tales como sensibilidad a los tratamientos, identificación de la terapia/fármaco más eficaz, supervisión del trascurso de la enfermedad, identificación de predisposición a la enfermedad, identificación de la presencia de enfermedad residual y/o la probabilidad de recaída. Por tanto, los métodos se pueden usar o no para el diagnóstico de la presencia de una afección específica. En una realización, la invención se puede usar para monitorizar la respuesta al tratamiento, por ejemplo, de manera continua. Los métodos de la invención se pueden usar para clasificar locus, en donde los mecanismos de enfermedad son desconocidos, poco claros o complejos. La detección de interacciones cromosómicas proporciona una forma eficaz de seguir los cambios en los diferentes niveles de regulación, algunos de los cuales son complejos. Por ejemplo, en algunos casos se pueden activar alrededor de 37.000 ARN no codificantes por medio de un solo impulso.

#### Características Relevantes para el Pronóstico

La característica puede ser la capacidad de respuesta a terapia con inhibidor de tirosina quinasa. La característica puede ser la expresión anormal de un gen de tirosina quinasa y/o un gen de tirosina quinasa, por ejemplo, la característica puede ser la regulación por aumento de la expresión de tirosina quinasa. Las quinasas que se sabe que son relevantes para el pronóstico de cáncer incluyen tirosina quinasas receptoras (por ejemplo, PDGFRA) que están en la superficie celular, quinasas no-receptoras, tirosina quinasas citoplásmicas (por ejemplo ABL1) y serina/treonina quinasas. La característica puede ser la expresión anormal de un gen de quinasa que tiene capacidad de respuesta o codifica un producto génico que tiene capacidad de respuesta a un inhibidor conocido, por ejemplo, un inhibidor de tirosina quinasa. La característica puede ser un cambio genético o epigenético en un gen que tiene capacidad de respuesta a o codifica un producto génico que tiene capacidad de respuesta a inhibidores de tirosina quinasa. La característica puede ser un cambio genético o epigenético en uno o más de los siguientes genes: FIP1L1, PDGFRA, Flt3, ABL1, FGFR1 y cKIT.

La característica puede ser la capacidad de respuesta a fármacos que son miméticos de ATP. Los genes de tirosina quinasa (por ejemplo ABL1, Brutons, cKIT, FGFR1, PDGFRA y PDGFRB) codifican proteínas que son sensibles a los miméticos de ATP.

La característica puede ser la respuesta a un tratamiento específico y/o profilaxis (en particular a un tratamiento farmacéutico específico y/o profilaxis), por ejemplo un inhibidor de tirosina quinasa, opcionalmente donde el fármaco se sabe que tiene como objetivo las tirosina quinasas receptoras, por ejemplo, PDGFRA, preferentemente uno o más de un grupo que comprende Dasatinib, crenolanib, dovitinib e Imatinib. El inhibidor de tirosina quinasa puede ser Cediranib, Crenolanib, Dasatinib (Sprycel); Dovitinib, Imatinib (Gleevec); IMC-3G3 (Olaratumab, LY3012207) Lartruvo (PDGFRA); Masitinib (AB1010), Nilotinib (Tasigna); Pazopanib; PKC412 (midostaurina, CGP41251); Quizartinib; Regorfenib ; Sorafenib (Nexavar); Sunitinib (Sutent); Vatalanib; S116836. La característica puede predisponer a una condición específica, opcionalmente hipereosinofilia.

## Subgrupos y Tratamiento Personalizado

5 Como se usa en la presente memoria, "subgrupo" se refiere preferentemente a un subgrupo de población (un subgrupo de una población), más preferentemente un subgrupo en la población de un animal particular tal como un eucariota particular, o un mamífero (por ejemplo, un ser humano, un ser no humano, un primate no humano, o un roedor, por ejemplo, un ratón o una rata). Lo más preferentemente, "subgrupo" hace referencia a un subgrupo de la población humana.

10 La invención incluye detectar y tratar subgrupos particulares de una población. Dentro de dichos subgrupos, las características comentadas en la presente memoria (tales como capacidad de respuesta a tratamiento) están presentes o ausentes. Las diferencias de interacción epigenética en un cromosoma son, hablando en términos generales, diferencias estructurales que existen a nivel genómico. Los inventores han descubierto que estas difieren entre subconjuntos (por ejemplo, dos o al menos dos subconjuntos) de una población dada.

15 La identificación de estas diferencias permite a los médicos clasificar a sus pacientes como parte de un subconjunto de la población como se describe en el método. Por tanto, la invención proporciona a los médicos un método para personalizar la medicina para el paciente, basándose en sus interacciones cromosómicas epigenéticas, y proporcionar un régimen de tratamiento alternativo más eficaz, en particular para subgrupos de pacientes cuyo pronóstico se ve afectado como resultado de ser identificados como parte de dicho subgrupo.

20 Normalmente, el subgrupo es al menos un 1 %, 5 %, 10 %, 30 %, 50 % o 80 % de la población. La población puede ser la población general. La población puede ser el grupo de pacientes con cáncer, donde el cáncer es opcionalmente uno cualquiera de los siguientes cáncer: iHES, leucemias eosinofílicas crónicas y agudas, leucemia mieloide crónica (CML), leucemia mieloide aguda (AML), leucemia linfocítica aguda (ALL), carcinoma de pulmón no microcítico (NSCLC) PDGFRA-positivo, glioblastoma, incluyendo glioblastoma multiforme y astrocitoma, cáncer de próstata, cáncer de ovario avanzado, tumor estromal gastrointestinal (GIST) y carcinoma de células renales. Uno o más subgrupos pueden tener cualquiera de los tipos de cáncer enumerados anteriormente, y el método para determinar el pronóstico de cáncer se puede llevar a cabo en un sujeto con cualquiera de los tipos de cáncer anteriores.

30 Generación de Ácidos Nucleicos Ligados

Determinadas realizaciones de la invención utilizan ácidos nucleicos ligados, en particular ADN ligado. Estos comprenden secuencias de ambas regiones que se unen en una interacción cromosómica y por tanto proporcionan información acerca de la interacción. El método EpiSwitch™ descrito en la presente memoria utiliza la generación de dichos ácidos nucleicos ligados para detectar interacciones cromosómicas.

De este modo, el método de la invención puede comprender una etapa de generación de ácidos nucleicos ligados (por ejemplo, ADN) mediante:

- 40 (i) reticular *in vitro* de interacciones cromosómicas epigenéticas presentes en el locus cromosómico;
- (ii) aislar opcionalmente el ADN reticulado de dicho locus cromosómico;
- 45 (iii) someter dicho ADN reticulado a corte, por ejemplo mediante digestión de restricción con una enzima que lo corta al menos una vez (en particular una enzima que corta al menos una vez dentro de dicho locus cromosómico);
- (iv) unir dichos extremos de ADN escindido reticulado (en particular para formar bucles de ADN), e
- 50 (v) identificar la presencia de dicho ADN ligado y/o dichos bucles de ADN, en particular usando técnicas tales como PCR (reacción en cadena de polimerasa), para identificar la presencia de una interacción cromosómica específica.

55 Se puede usar PCR para detectar o identificar el ácido nucleico ligado, por ejemplo, el tamaño del producto de PCR producido puede ser indicativo de la interacción cromosómica específica que está presente, y por tanto se puede usar para identificar el estado del locus. El experto en la materia conoce numerosas enzimas de restricción que se pueden usar para cortar el ADN dentro del locus cromosómico de interés. Resulta evidente que la enzima particular usada depende del locus estudiado y de la secuencia del ADN localizado en el mismo.

60 Un ejemplo no limitante de una enzima de restricción que se puede usar para cortar el ADN como se describe en la presente invención es TaqI.

Modalidades tales como la tecnología EpiSwitch™.

65 La tecnología EpiSwitch™ hace referencia al uso de datos de marcadores EpiSwitch™ de micromatriz en la detección de firmas de conformación cromosómica epigenéticas específicas para fenotipos. Las realizaciones

tales como EpiSwitch™ que utilizan ácidos nucleicos ligados de la manera descrita en la presente memoria tienen varias ventajas. Tienen un bajo nivel de ruido estocástico, por ejemplo debido a que las secuencias de ácido nucleico del primer conjunto de ácidos nucleicos de la presente invención o bien hibridan o no hibridan con el segundo conjunto de ácidos nucleicos. Esto proporciona un resultado binario que permite una manera relativamente simple de medir un mecanismo complejo a nivel epigenético. La tecnología EpiSwitch™ también tiene un tiempo de procesamiento rápido y bajo coste. En una realización, el tiempo de procesamiento es de 3 a 6 horas.

#### Muestras y Tratamiento de Muestra

La muestra contiene ADN del individuo. Normalmente contiene células. En una realización, se obtiene una muestra por medios mínimamente invasivos, y puede ser, por ejemplo, sangre. Se puede extraer ADN y cortar con una enzima de restricción convencional. Esto puede determinar previamente qué conformaciones cromosómicas se retienen y se detectan con las plataformas EpiSwitch™. En una realización en la que la muestra es una muestra de sangre obtenida previamente del paciente, el método descrito resulta ventajoso porque el procedimiento es mínimamente invasivo. Debido a la sincronización de las interacciones cromosómicas entre tejidos y sangre, incluyendo la transferencia horizontal, se puede usar una muestra de sangre para detectar las interacciones cromosómicas en tejidos, tales como tejidos relevantes para la enfermedad. Para determinadas afecciones, tales como cáncer, el ruido genético debido a mutaciones puede afectar a la "señal" de interacción cromosómica en los tejidos relevantes y, por tanto, el uso de sangre resulta ventajoso.

#### Propiedades de los Ácidos Nucleicos Relevantes para la Invención

Los ácidos nucleicos que se preparan o detectan en las realizaciones de la invención pueden ser los mismos que, o tener cualquiera de las propiedades de, el primer y segundo ácidos nucleicos mencionados en la presente memoria. Los ácidos nucleicos de la invención comprenden normalmente dos partes, comprendiendo cada una de ellas una secuencia de una de las dos regiones del cromosoma que se unen en la interacción cromosómica. Normalmente, cada parte tiene una longitud de al menos 8, 10, 15, 20, 30 o 40 nucleótidos, por ejemplo de 10 a 40 nucleótidos. Los ácidos nucleicos preferidos comprenden una secuencia de cualquiera de los genes mencionados en cualquiera de las tablas. Los ácidos nucleicos preferidos se unen a cualquier sonda o secuencia de cebadores específicos mencionados en la presente memoria. Preferentemente, los ácidos nucleicos son ADN. Se comprende que cuando se define una secuencia específica, la invención puede usar la secuencia complementaria como se requiere en la realización particular.

#### El Segundo Conjunto de Ácidos Nucleicos - las Secuencias de "Índice"

El segundo conjunto de secuencias de ácido nucleico tiene la función de ser un conjunto de secuencias de índice, y es esencialmente un conjunto de secuencias de ácido nucleico que son adecuadas para identificar una secuencia específica de subgrupo. Pueden representar las interacciones cromosómicas "de fondo" y se podrían seleccionar o no de modo alguno. Son en general un subconjunto de todas las interacciones cromosómicas posibles.

El segundo conjunto de ácidos nucleicos se puede derivar por medio de cualquier método adecuado. Se pueden derivar por computación o se pueden basar en la interacción cromosómica en individuos. Normalmente representan un grupo poblacional más grande que el primer conjunto de ácidos nucleicos. En una realización particular, el segundo conjunto de ácidos nucleicos representa todas las posibles interacciones cromosómicas epigenéticas en un conjunto específico de genes. En otra realización particular, el segundo conjunto de ácidos nucleicos representa una gran proporción de todas las posibles interacciones cromosómicas epigenéticas presentes en una población descrita en la presente memoria. En una realización particular, el segundo conjunto de ácidos nucleicos representa al menos un 50 % o al menos un 80 % de interacciones cromosómicas epigenéticas en al menos 20, 50, 100 o 500 genes, por ejemplo en 20 a 100 o 50 a 500 genes.

El segundo conjunto de ácidos nucleicos representa normalmente al menos 100 posibles interacciones cromosómicas epigenéticas que opcionalmente modifican, regulan o median en modo alguno un estado/fenotipo de enfermedad en la población. El segundo conjunto de ácidos nucleicos puede representar interacciones cromosómicas que afectan un estado de enfermedad en una especie, por ejemplo, que comprende secuencias de ácidos nucleicos que codifican citoquinas, quinasas o reguladores asociados a cualquier estado de enfermedad, predisposición a enfermedad o fenotipo de enfermedad. El segundo conjunto de ácidos nucleicos comprende normalmente secuencias que representan interacciones epigenéticas relevantes y no relevantes para la característica que define el subgrupo.

En una realización particular, el segundo conjunto de ácidos nucleicos deriva al menos parcialmente de secuencias de origen natural en una población, y se obtienen normalmente por medio de métodos de sílice. Dichos ácidos nucleicos pueden comprender además mutaciones individuales o múltiples en comparación con una parte correspondiente de ácidos nucleicos presentes en los ácidos nucleicos de origen natural. Las



- mutaciones incluyen eliminaciones, sustituciones y/o adiciones de uno o más pares de bases de nucleótidos. En una realización particular, el segundo conjunto de ácidos nucleicos puede comprender una secuencia que representa un homólogo y/u ortólogo con al menos un 70 % de identidad de secuencia con la parte correspondiente de ácidos nucleicos presentes en las especies de origen natural. En otra realización particular,
- 5 se proporciona al menos un 80 % de identidad de secuencia o al menos un 90 % de identidad de secuencia con la parte correspondiente de ácidos nucleicos presente en la especie de origen natural.

#### Propiedades del Segundo Conjunto de Ácidos Nucleicos

- 10 En una realización particular, hay al menos 100 secuencias de ácidos nucleicos diferentes en el segundo conjunto de ácidos nucleicos, preferentemente al menos 1000, 2000 o 5000 secuencias de ácidos nucleicos diferentes, con hasta 100.000, 1.000.000 o 10.000.000 secuencias de ácidos nucleicos diferentes. Un número típico sería de 100 a 1.000.000, tal como de 1.000 a 100.000 secuencias de ácidos nucleicos diferentes. Todas, o al menos un 90 % o al menos un 50 %, de estas corresponderían a diferentes interacciones cromosómicas.
- 15 En una realización particular, el segundo conjunto de ácidos nucleicos representa interacciones cromosómicas en al menos 20 locus o genes diferentes, preferentemente al menos 40 locus o genes diferentes, y más preferentemente al menos 100, al menos 500, al menos 1000 o al menos 5000 locus o genes diferentes, tales como 100 a 10.000 locus o genes diferentes. Las longitudes del segundo conjunto de ácidos nucleicos son adecuadas para que se hibriden específicamente según el emparejamiento de bases de Watson Crick con el
- 20 primer conjunto de ácidos nucleicos para permitir la identificación de interacciones cromosómicas específicas de subgrupos. Normalmente, el segundo conjunto de ácidos nucleicos comprende dos partes correspondientes en secuencia a las dos regiones cromosómicas que se unen en la interacción cromosómica. El segundo conjunto de ácidos nucleicos comprende normalmente secuencias de ácidos nucleicos que tienen al menos
- 25 10, preferentemente 20, y preferentemente todavía 30 bases (nucleótidos) de longitud. En otra realización, las secuencias de ácido nucleico pueden ser como máximo 500, preferentemente como máximo 100, y preferentemente todavía como máximo 50 pares de bases de longitud. En una realización preferida, el segundo conjunto de ácidos nucleicos comprende secuencias de ácidos nucleicos de entre 17 y 25 pares de bases. En una realización, al menos un 100, un 80 % o un 50 % del segundo conjunto de secuencias de ácido nucleico
- 30 tienen longitudes como se ha descrito con anterioridad. Preferentemente, los diferentes ácidos nucleicos no tienen ninguna secuencia solapante, por ejemplo, al menos un 100 %, 90 %, 80 % o 50 % de los ácidos nucleicos no tienen la misma secuencia sobre al menos 5 nucleótidos contiguos.
- 35 Dado que el segundo conjunto de ácidos nucleicos actúa como un "índice", entonces se puede usar el mismo conjunto de segundos ácidos nucleicos con diferentes conjuntos de primeros ácidos nucleicos que representan subgrupos para diferentes características, es decir, el segundo conjunto de ácidos nucleicos puede representar una colección "universal" de ácidos nucleicos que se pueden usar para identificar interacciones cromosómicas relevantes para diferentes características.

#### 40 El Primer Conjunto de Ácidos Nucleicos

- El primer conjunto de ácidos nucleicos procede normalmente de individuos que se sabe que están en dos o más subgrupos distintos definidos por la presencia o ausencia de una característica relevante para un ensayo epigenético con acompañante de pronóstico, tal como cualquiera de las características mencionadas en la
- 45 presente memoria. Los primeros ácidos nucleicos pueden tener cualquiera de las características y propiedades del segundo conjunto de ácidos nucleicos mencionados en la presente memoria. El primer conjunto de ácidos nucleicos procede normalmente de una muestra de los individuos que se han sometido a tratamiento y procesado como se describe en la presente memoria, en particular las etapas de reticulación y escisión de EpiSwitch™. Normalmente, el primer conjunto de ácidos nucleicos representa la totalidad o al menos un 80 %
- 50 o un 50 % de las interacciones cromosómicas presentes en las muestras tomadas de los individuos.

- Normalmente, el primer conjunto de ácidos nucleicos representa una población más pequeña de interacciones cromosómicas a través de los locus o genes representados por el segundo conjunto de ácidos nucleicos en comparación con las interacciones cromosómicas representadas por el segundo conjunto de ácidos nucleicos, es decir, el segundo conjunto de ácidos nucleicos representa un fondo o conjunto de índices de interacciones
- 55 en un conjunto definido de locus o genes.

- En una realización, el primer conjunto de ácidos nucleicos está marcado, y normalmente los ácidos nucleicos procedentes de diferentes subgrupos se marcan con diferentes marcadores. Esto les permite ser distinguidos
- 60 en un método de detección.

#### Ácidos Nucleicos

- Los ácidos nucleicos descritos en la presente memoria pueden estar en forma de una biblioteca de ácidos
- 65 nucleicos que comprende al menos 200, al menos 500, al menos 1000, al menos 5000 o al menos 10000 ácidos nucleicos diferentes de la invención, tales como ácidos nucleicos "segundos". La invención proporciona

una biblioteca particular de ácidos nucleicos que normalmente comprende al menos 200 ácidos nucleicos diferentes. La biblioteca puede estar en la forma de ácidos nucleicos unidos a una matriz.

#### Hibridación

- 5 La invención requiere un medio para permitir que las secuencias de ácido nucleico total o parcialmente complementarias del primer conjunto de ácidos nucleicos y el segundo conjunto de ácidos nucleicos experimenten hibridación. En una realización, todo el primer conjunto de ácidos nucleicos se pone en contacto con todo el segundo conjunto de ácidos nucleicos en un único individual, es decir, en una única etapa de  
10 hibridación. No obstante, se puede usar cualquier ensayo adecuado.

#### Ácidos Nucleicos Marcados y Patrón de Hibridación

- 15 Los ácidos nucleicos mencionados en la presente memoria se pueden marcar, preferentemente usando una marcador independiente tal como un fluoróforo (molécula fluorescente) o marcador radioactivo que contribuye a la detección de hibridación exitosa. Determinados marcajes se pueden detectar con luz UV. El patrón de hibridación, por ejemplo en una matriz descrita en la presente memoria, representa diferencias en las interacciones cromosómicas epigenéticas entre los dos subgrupos, y de este modo proporciona un método para comparar interacciones cromosómicas epigenéticas y la determinación de qué interacciones  
20 cromosómicas epigenéticas son específicas para un subgrupo en la población de la presente invención.

- La expresión "patrón de hibridación" abarca ampliamente la presencia y ausencia de hibridación entre el primer y el segundo conjunto de ácidos nucleicos, es decir, qué ácidos nucleicos específicos del primer conjunto experimentan hibridación con qué ácidos nucleicos específicos del segundo conjunto, y por tanto no se limita  
25 a ningún ensayo o técnica particular, o la necesidad de tener una superficie o matriz en la que se pueda detectar un "patrón".

#### Condiciones Específicas y Selección de Subgrupos

- 30 Al centrarse en las interacciones cromosómicas, y los cambios en las interacciones cromosómicas, los inventores han identificado grupos de pacientes que podrían responder a intervenciones de diagnóstico terapéutico potencialmente salvadoras de vida, por ejemplo, usando inhibidores de tirosina quinasa (TKI), que pasaría por alto usando diagnósticos convencionales.
- 35 En una realización, el método de la invención detecta la capacidad de respuesta a la terapia con inhibidores de tirosina quinasa, por ejemplo, inhibidores que son miméticos de ATP. Preferentemente, se detecta la capacidad de respuesta al tratamiento con inhibidor de tirosina quinasa de cáncer, por ejemplo, en pacientes con eosinófilos elevados. Preferentemente, el cáncer es iHES, leucemias eosinofílicas crónicas y agudas, leucemia mieloide aguda (AML), carcinoma de pulmón no microcítico positivo para PDGFRA, glioblastoma,  
40 incluyendo glioblastoma multiforme y astrocitoma, cáncer de próstata, cáncer de ovario avanzado o tumor del estroma gastrointestinal (GIST). En dichas realizaciones, se prefiere la detección de interacciones cromosómicas en PDGFRA y/o FIP1L1, tal como cualquier interacción específica descrita en la presente memoria (por ejemplo, mediante referencias a cebadores y/o sondas específicos).
- 45 Preferentemente, se detecta la presencia o ausencia de cualquiera de las interacciones cromosómicas dentro de cualquiera de los genes relevantes mencionados en la presente memoria, por ejemplo, en al menos 1, 2, 3, 4, 5 o 6 de los genes. Preferentemente, en el método se determina la presencia o ausencia de interacciones cromosómicas representadas por cualquier secuencia de cebador y/o sonda específica descrita en la presente memoria. Estos números de genes o interacciones cromosómicas se pueden usar en cualquiera de las  
50 diferentes realizaciones mencionadas en la presente memoria.

- La invención proporciona un método que comprende detectar la presencia o ausencia de interacciones cromosómicas, normalmente 1 a 5, o 5 a 20, o 5 a 500 de dichas interacciones, con el fin de determinar la presencia o ausencia de una característica en un individuo. Preferentemente, las interacciones cromosómicas  
55 son las de cualquiera de los genes mencionados en la presente memoria.

#### Interacciones 3C y Tipos de Cáncer

- 60 Las diferentes realizaciones de la invención pueden hacer referencia a cualquier cáncer o gen mencionado en la presente memoria, por ejemplo, los enumerados a continuación.

- Los inventores han descubierto interacciones 3C en ABL1, tirosina quinasa de Bruton, cKIT, FGFR1, F1PL1, FLT3 y PDGFRA que están asociadas a cáncer, incluyendo cáncer de mama, cáncer de ovario, leucemias, glioma, linfomas y carcinomas.  
65

Normalmente, las interacciones relevantes para, o identificadas usando, los métodos descritos en la presente

memoria se asocian a cualquiera de las siguientes afecciones: leucemia mieloide aguda (AML), leucemia linfocítica crónica (CLL), leucemia mieloide crónica (CML), crisis blástica de CML, tumor del estroma gastrointestinal (GIST), síndromes mielodisplásicos (MDS), trastornos mieloproliferativos (neoplasias mieloproliferativas, MPN), agammaglobulinemia ligada al cromosoma X (XLA), cáncer de células B, NHL, linfoma, mieloma, neoplasias de células B maduras recurrentes, enfermedad crónica de injerto contra huésped, leucemia de células pilosas recidivante, tumores sólidos recidivantes o refractarios, linfoma de células del manto refractario, carcinoma esofagogástrico amplificado, artritis reumatoide, mastocitosis, leucemia de mastocitos, tumores de próstata, melanoma avanzado, melanoma maligno mutado, tumores estromales gastrointestinales no susceptibles de resección o metastásicos de 2ª línea (GIST), mesotelioma que expresa

5 cáncer, cáncer de ovario avanzado refractario a platino, cáncer de peritoneo primario o de tubo falopiano o epitelial ovárico, tumor ovárico de bajo potencial cancerígeno, carcinoma seroso papilar uterino, cáncer de pulmón microcítico en estadio avanzado, carcinoma nasofaríngeo localmente avanzado, linfoma no Hodgkin recidivante/refractario, cáncer de células renales metastásico, tumor neuroendocrino gastroenteropancreático metastásico, carcinoma ovárico de células claras recurrente, carcinoma esofágico de células escamosas,

10 cáncer de útero tratado previamente, carcinoma nasofaríngeo localmente avanzado, leucemia linfocítica crónica refractaria/linfoma linfocítico pequeño (CLL/SLL), tumores gastrointestinales estromáticos (GIST) con fallo a Sumitinib R. Las afecciones preferidas son glioma, leucemia mieloide aguda y cáncer de mama, iHES y leucemias crónicas.

15

En una realización, las interacciones relevantes para, o identificadas usando, los métodos descritos en la presente memoria se asocian a cualquiera de las siguientes afecciones: iHES, cáncer de pulmón no microcítico (NSCLC), cáncer colorrectal, melanoma metastásico, cáncer de células renales, linfoma de Hodgkin refractario, cáncer de ovario tratado previamente, carcinoma nasofaríngeo localmente avanzado, leucemia linfocítica crónica refractaria/linfoma linfocítico pequeño (CLL/SLL), tumores gastrointestinales estromáticos (GIST) con fallo a Sumitinib R. Las afecciones preferidas son glioma, leucemia mieloide aguda y cáncer de mama, iHES y leucemias crónicas.

20

25

Las interacciones en ABL1, gen de tirosina quinasa citoplasmática, por ejemplo localizado en 9q34, se asocian a las siguientes enfermedades: Leucemia Mieloide Aguda (AML), Leucemia Linfocítica Crónica (CLL), Cáncer de Mama, (GIST), Síndromes Mielodisplásicos (MDS) o Trastornos Mieloproliferativos (Neoplasias Mieloproliferativas, MPN).

30

Las interacciones en tirosina quinasa de Bruton, gen de tirosina quinasa citoplasmática, por ejemplo localizado en Xq22.1, se asocian a las siguientes enfermedades: agammaglobulinemia ligada al cromosoma X (XLA), cáncer de células B, NHL, linfoma (todas las clasificaciones), mieloma, leucemia linfocítica crónica, neoplasias de células B maduras recurrentes, enfermedad de injerto contra huésped crónica, leucemia de células pilosas recidivante, tumores sólidos recidivantes o refractarios, linfoma de células del manto refractario, carcinoma esofagogástrico amplificado y artritis reumatoide.

35

Las interacciones en cKIT, gen de receptor de tirosina quinasa, por ejemplo localizado en 4q12, se asocian a las siguientes enfermedades: mastocitosis y leucemia de mastocitos y tumores de próstata, melanoma avanzado, melanoma maligno mutado, tumores del estroma gastrointestinal (GIST) no susceptibles de resección o metastásicos de 2ª línea, mesotelioma que expresa cáncer. Cáncer de ovario avanzado, refractario a platino, cáncer de peritoneo primario o de tubo falopiano, epitelial de ovario recidivante, o tumor ovárico de bajo potencial maligno, cáncer de pulmón microcítico de estadio avanzado, carcinoma nasofaríngeo localmente

40

45

avanzado, linfoma no Hodgkin recidivante/refractario, cáncer de células renales metastásico, tumor neuroendocrino gastroenteropancreático metastásico, carcinoma ovárico de células claras, carcinoma esofágico de células escamosas, cáncer uterino persistente y tumores de mama.

Las interacciones en FGFR1, gen de tirosina quinasa del receptor, por ejemplo localizado en 8p11, se asocian a las siguientes enfermedades: cáncer de tiroides, adenoma oral, leucemia crónica, leucemia aguda, tumores de próstata y cáncer de mama, FGFR1-3 translocado, cáncer de cabeza y cuello recurrente mutado o amplificado, glioma maligno recurrente que expresa la fusión del gen FGFR-TACC, cáncer de pulmón no microcítico avanzado.

50

Las interacciones en FIP1L1, factor de especificidad de poliadenilación frecuentemente fusionado a PDGFRA (por ejemplo 4q12) se asocian a la siguiente enfermedad: iHES, leucemias crónicas.

55

Las interacciones en el gen de tirosina quinasa de receptor FLT3, por ejemplo localizado en 13q11-12.2, se asocian a las siguientes enfermedades: leucemia crónica, leucemia mieloide aguda y cáncer de mama, iHES, leucemias crónicas, cáncer de pulmón no microcítico (NSCLC), cáncer colorrectal, melanoma metastásico o cáncer de células renales, glioma, linfoma de Hodgkin refractario, cáncer de ovario tratado previamente, carcinoma nasofaríngeo localmente avanzado, leucemia linfocítica crónica refractaria/linfoma linfocítico pequeño (CLL/SLL), tumores del estroma gastrointestinal (GIST) con fallo a sumitinib.

60

Las interacciones en el gen de tirosina quinasa de receptor PDGFRA, por ejemplo localizado en 4q12, se asocian a las siguientes enfermedades: glioma, leucemias crónicas, leucemia mieloide aguda, cáncer de

65

mama, iHES.

El Individuo que se Somete a Ensayo

- 5 El individuo objeto de ensayo puede o no tener cualesquiera síntomas de cualquier afección o característica de enfermedad mencionada en la presente memoria. El individuo puede estar en riesgo de cualquier afección o característica. El individuo puede estar en riesgo debido a la característica. El individuo puede tener un peor pronóstico debido a la característica y puede requerir terapia específica para la característica. El individuo puede haberse recuperado o estar en el proceso de recuperación de la afección o característica. El individuo es preferentemente un mamífero, tal como un primate, un ser humano, un mamífero no humano o un roedor. El individuo puede ser hombre o mujer. El individuo puede tener 30 años de edad o más. El individuo puede tener 29 años de edad o menos.

- 10 Cualquiera de los genes, locus, regiones y ácidos nucleicos comentados en la presente memoria puede ser de, o proceder de, cualquiera de los tipos de individuos mencionados en la presente memoria.

Regiones Génicas Preferidas, Locus, Genes e Interacciones Cromosómicas

- 20 Para todos los aspectos de la invención se mencionan en la presente memoria regiones génicas preferidas, locus, genes e interacciones cromosómicas. Normalmente, en los métodos de la invención, las interacciones cromosómicas se detectan a partir de al menos 1, 2, 3, 4, 5 o 6 de los genes relevantes enumerados en la presente memoria. Preferentemente, se detecta la presencia o ausencia de al menos 1, 2, 3, 4, 5 o 6 de las interacciones cromosómicas específicas relevantes representadas por las secuencias de cebador y/o de sonda mencionadas en la presente memoria.

- 25 La región puede estar aguas arriba o aguas abajo de cualquiera de los genes mencionados en la presente memoria, por ejemplo, 50 kb aguas arriba o 20 kb aguas abajo, por ejemplo de la secuencia de codificación.

- 30 En una realización, las interacciones cromosómicas que se detectan están presentes en cualesquiera de las posiciones o en cualesquiera de las regiones mostradas en las tablas. En caso de que los productos ligados se detecten en el método, entonces se puede detectar la secuencia mostrada en cualesquiera de las secuencias de cebador o sonda divulgadas en la presente memoria. Por tanto, normalmente se podría detectar la secuencia de ambas regiones del producto ligado (es decir, de ambos sitios de la interacción cromosómica). En realizaciones preferidas se utilizan cebadores y/o sondas en el método que comprenden o consisten en la misma secuencia o en una secuencia complementaria a un cebador y/o sonda divulgados en la presente memoria. En algunas realizaciones se usan cebadores y/o sondas que comprenden una secuencia que es homóloga a cualesquiera de las secuencias de cebador y/o sonda mostradas en la presente memoria.

Detección de un Producto Ligado por PCR Cuantitativa (qPCR)

- 40 En una realización, se detecta cuantitativamente una secuencia ligada que es relevante para una interacción cromosómica, por ejemplo, usando detección de fluorescencia. La presente realización se puede llevar a cabo en cualquier secuencia ligada divulgadas en la presente memoria o que sea relevante para los métodos de tipificación/detección divulgados en la presente memoria, pero no se limita a dichas secuencias o métodos. De este modo, la realización relacionada con qPCR se puede llevar a cabo para detectar cualquier secuencia ligada que comprenda secuencias procedentes de regiones de cromosomas que se han unido en cualquier interacción cromosómica, que puede tener cualquiera de las características mencionadas en la presente memoria.

- 50 La invención proporciona además un método para detectar cuantitativamente una secuencia ligada que es relevante para una interacción cromosómica usando una sonda que se pueda detectar tras activación durante una reacción de PCR, en donde dicha secuencia ligada comprende secuencias de dos regiones cromosómicas que se unen en una interacción cromosómica epigenética, en donde dicho método comprende poner en contacto la secuencia ligada con la sonda durante una reacción de PCR, y detectar el grado de activación de la sonda, y en donde dicha sonda se une al sitio de enlace.

- 55 Las secuencias de cebador y sonda pueden ser cualquiera de las secuencias específicas divulgadas u homólogos y/o fragmentos de las secuencias específicas. Las secuencias de cebador son normalmente capaces de amplificar el producto ligado preparado en el proceso de detección de una interacción cromosómica, tal como un método 3C o EpiSwitch™, por ejemplo, cualquier método o producto ligado descrito en la presente memoria. Los cebadores son normalmente unidireccionales. Opcionalmente, las secuencias de cebador están diseñadas para ser únicas para el producto ligado de la interacción 3C en cuestión. La sonda normalmente es capaz de unirse específicamente a una secuencia en el producto ligado que incluye el sitio de unión. La sonda es normalmente una sonda de "hidrólisis" a partir de la cual se pueden escindir/liberar uno o más restos para activar un marcador durante la reacción de PCR.

La secuencia ligada se puede preparar mediante reticulación de una muestra de ácido nucleico, seguido de la ruptura de ácido nucleico en fragmentos más pequeños, por ejemplo, mediante digestión enzimática o mediante un proceso físico de cizalladura. A continuación se lleva a cabo una etapa de unión. La secuencia ligada normalmente procede de una muestra de un individuo o una estirpe celular que ha experimentado tratamiento y procesado como se describe en la presente memoria, en particular las etapas de reticulación y escisión de EpiSwitch™. La secuencia ligada puede ser un producto ligado como se define en o preparado por medio de cualesquiera de los métodos anteriores. El sitio de unión de la secuencia ligada puede comprender la secuencia de reconocimiento enzimático de restricción de la enzima de restricción usada para cortar el ácido nucleico reticulado. Una enzima de restricción preferida es TaqI o cualquier forma modificada de TaqI (que por ejemplo reconoce el mismo sitio de restricción). La molécula de ADN ligada puede comprender una secuencia de ácido nucleico de longitud de 10 a 100 bases nucleotídicas, o de 10 a 400 bases nucleotídicas, o de 10 a 500 bases nucleotídicas, o de 200 a 600 bases nucleotídicas, o de 200 a 800 bases nucleotídicas o de 200 a 1000 bases nucleotídicas.

Preferentemente, solo se usa un conjunto de cebadores (es decir, dos cebadores de PCR susceptible de amplificación conjunta) en la detección, y no se lleva a cabo una PCR anidada con más de un conjunto de cebadores. Preferentemente, los cebadores son unidireccionales. Incluso más preferentemente, la secuencia de cebador se elige de modo que sea única para la interacción 3C en cuestión. En una realización, el método de detección se lleva a cabo en presencia de otros ácidos nucleicos (tales como otros productos ligados) que no están amplificados y/o a los que la sonda no se une. Normalmente, otros ácidos nucleicos (tales como productos ligados) pueden representar al menos un 10 %, 30 % o al menos un 50 % de los ácidos nucleicos o productos ligados que están presentes.

#### Situaciones Detectadas por medio de qPCR

El método de qPCR descrito en la presente memoria se puede usar para detectar interacciones cromosómicas en cualquier situación en la que sean relevantes, por ejemplo, cualquier situación descrita. Normalmente se utiliza para detectar una interacción cromosómica asociada a un subgrupo, por ejemplo, como se define en la presente memoria. Se puede usar para detectar un cambio en una interacción cromosómica y/o una interacción cromosómica asociada a una mutación, cambio en la expresión de un gen o desregulación de un gen.

El método qPCR descrito en la presente memoria se puede usar para detectar interacciones cromosómicas en las que las secuencias de dos regiones cromosómicas que se unen en una interacción cromosómica epigenética se encuentran al menos 1000 pares de bases separados o al menos 10.000 pares de bases separados o al menos 100.000 pares de bases separados en el cromosoma. El método de qPCR descrito en la presente memoria se puede usar para detectar interacciones cromosómicas de largo alcance en las que las secuencias de dos regiones cromosómicas que se unen en una interacción cromosómica epigenética se encuentran a más de 100.000 pares de bases y la interacción, por tanto, genera un bucle de ADN mayor de 100 kilobases, opcionalmente mayor de 500 kilobases, preferentemente mayor de 900 kilobases. El método de qPCR descrito en la presente memoria se puede usar para detectar interacciones cromosómicas de largo alcance asociadas a un cambio genético o epigenético en genes que responden a la terapia con inhibidor de tirosina quinasa, tal como PDGFRA.

#### Diseño/Estructura de una sonda qPCR

La sonda para su uso en la reacción de qPCR se puede diseñar para unirse específicamente a la molécula de ADN ligada en las proximidades del sitio de enlace, en un sitio de la molécula de ADN ligada que comprende el sitio de enlace, que tiene opcionalmente una secuencia que es complementaria a la secuencia de cada una de dichas regiones cromosómicas. La sonda abarca la unión de ambos fragmentos, y por ejemplo tiene una secuencia que une al menos 5, 10 o 20 bases de nucleótidos 5' y 3' al sitio de enlace. La sonda puede comprender una secuencia de ácido nucleico de longitud 10 a 40 bases de nucleótidos, o preferentemente comprende secuencias de ácido nucleico de longitud 10 a 35 bases de nucleótidos, o más preferentemente comprende secuencias de ácido nucleico de longitud 20 a 30 bases de nucleótidos. La sonda puede tener una secuencia que tiene al menos un 70 % de homología/identidad con la secuencia de la molécula de ADN ligada u otros niveles de identidad como se define en la presente memoria.

#### Marcaje de una sonda qPCR

La sonda se marca generalmente con un marcador detectable que tiene un estado inactivo y activo, de modo que solo se detecta cuando se activa. El grado de activación estará relacionado con el grado de la plantilla (producto de unión) presente en la reacción de PCR. La detección se puede llevar a cabo durante toda o parte de la PCR, por ejemplo, durante al menos un 50 % u 80 % de los ciclos de la PCR.

La sonda puede comprender un fluoróforo unido covalentemente a un extremo del oligonucleótido, y un agente de inactivación unido al otro extremo del nucleótido, de modo que la fluorescencia del fluoróforo quede inactivada por el agente de inactivación. En una realización, el fluoróforo está unido al extremo 5' del

oligonucleótido, y el agente de inactivación está unido covalentemente al extremo 3' del oligonucleótido. Los fluoróforos que se pueden usar en los métodos de la invención incluyen FAM, TET, JOE, Yakima Yellow, HEX, Cyanine3, ATTO 550, TAMRA, ROX, Texas Red, Cyanine 3.5, LC610, LC 640, ATTO 647N, Cyanine 5, Cyanine 5.5 y ATTO 680. Los agentes de inactivación que se pueden usar con el fluoróforo apropiado incluyen TAM, BHQ1, DAB, Eclip, BHQ2 y BBQ650, opcionalmente en donde dicho fluoróforo está seleccionado entre HEX, Texas Red y FAM. Las combinaciones preferidas de fluoróforo y agente de inactivación incluyen FAM con BHQ1 y Texas Red con BHQ2. Otras combinaciones preferidas se muestran en la Tabla 3.

#### Uso de Sonda en un Ensayo de qPCR

Las sondas de hidrólisis de la invención son normalmente de gradiente de temperatura optimizado con controles negativos de concentración emparejada. Preferentemente, se optimizan las reacciones de PCR de etapa individual. Más preferentemente se calcula una curva convencional. Una ventaja de usar una sonda específica que se une a través de la unión de la secuencia ligada es que la especificidad para la secuencia ligada se puede lograr sin usar un enfoque de PCR anidada. Los métodos descritos en la presente memoria permiten una cuantificación precisa y precisa de objetivos de bajo número de copias. La secuencia ligada diana se puede purificar, por ejemplo, purificación en gel, antes de la optimización del gradiente de temperatura. La secuencia ligada objetivo se puede secuenciar. Preferentemente, las reacciones de PCR se llevan a cabo usando aproximadamente 10 ng, o de 5 a 15 ng, o de 10 a 20 ng, o de 10 a 50 ng, o de 10 a 200 ng de ADN de plantilla. Los cebadores directos e inversos se diseñan de manera que un cebador se une a la secuencia de una de las regiones cromosómicas representadas en la secuencia de ADN ligado, y el otro cebador se une a otra región cromosómica representada en la secuencia de ADN ligado, por ejemplo, siendo complementario a la secuencia.

#### Elección de Diana de ADN ligado

La invención proporciona un método para seleccionar cebadores y una sonda para su uso en un método de PCR como se define en la presente memoria que comprende seleccionar cebadores en base a su capacidad para unirse y amplificar la secuencia ligada y seleccionar las propiedades basadas en la secuencia de sonda de la secuencia objetivo a la cual se unirá, en particular la curvatura de la secuencia objetivo.

Las sondas se diseñan/eligen normalmente para unirse a secuencias ligadas que son fragmentos de restricción yuxtapuestos que abarcan el sitio de restricción. En una realización de la invención, se calcula la curvatura predicha de posibles secuencias ligadas relevantes para una interacción cromosómica particular, por ejemplo usando un algoritmo específico al que se hace referencia en la presente memoria. La curvatura se puede expresar como grados por vuelta helicoidal, por ejemplo 10,5° por vuelta helicoidal. Las secuencias ligadas están seleccionadas para dirigirse a donde la secuencia ligada tenga una puntuación máxima de propensión a la curvatura de al menos 5° por vuelta helicoidal, normalmente al menos 10°, 15° o 20° por vuelta helicoidal, por ejemplo de 5° a 20° por vuelta helicoidal. Preferentemente, la puntuación de propensión a la curvatura por vuelta helicoidal se calcula para al menos 20, 50, 100, 200 o 400 bases, tal como para 20 a 400 bases aguas arriba y/o aguas abajo del sitio de unión. Por tanto, en una realización, la secuencia objetivo en el producto ligado tiene cualesquiera de estos niveles de curvatura. Las secuencias objetivo también se pueden escoger basándose en la energía libre de estructura termodinámica más baja.

#### Realizaciones Preferidas para Preparación de Muestra y Detección de Interacción Cromosómica

Se describen en la presente memoria métodos para preparar muestras y detectar conformaciones cromosómicas. Se pueden usar versiones optimizadas (no convencionales) de estos métodos, por ejemplo como se describe en la presente sección.

Normalmente, la muestra contiene al menos  $2 \times 10^5$  células. La muestra puede contener hasta  $5 \times 10^5$  células. En una realización, la muestra contiene  $2 \times 10^5$  a  $5,5 \times 10^5$  células.

La reticulación de interacciones cromosómicas epigenéticas presentes en el locus cromosómico se describe en la presente memoria. Esto se puede llevar a cabo antes de que tenga lugar la lisis celular. La lisis celular se puede llevar a cabo durante 3 a 7 minutos, tal como 4 a 6 o aproximadamente 5 minutos. En algunas realizaciones, la lisis celular se lleva a cabo durante al menos 5 minutos y durante menos de 10 minutos.

Se describe en la presente la digestión de ADN con una enzima de restricción. Normalmente, la restricción de ADN se lleva a cabo a una temperatura de aproximadamente 55 °C a aproximadamente 70 °C, tal como de aproximadamente 65 °C, durante un período de aproximadamente 10 a 30 minutos, tal como aproximadamente 20 minutos.

Preferentemente se usa una enzima de restricción de corte frecuente que da como resultado fragmentos de ADN ligado con un tamaño promedio de fragmento de hasta 4000 pares de bases. Opcionalmente, la enzima de restricción tiene como resultado fragmentos de ADN ligado que tienen un tamaño promedio de fragmento

de aproximadamente 200 a 300 pares de bases, tal como aproximadamente 256 pares de bases.

En una realización, el tamaño normal de fragmento es de 200 pares de bases a 4.000 pares de bases, tales como de 400 a 2.000 o de 500 a 1.000 pares de bases.

5

En una realización del método EpiSwitch, no se lleva a cabo una etapa de precipitación de ADN entre la etapa de digestión de restricción de ADN y la etapa de unión de ADN.

10

La unión de ADN se describe en la presente memoria. Normalmente, la unión de ADN se lleva a cabo durante 5 a 30 minutos, tal como aproximadamente 10 minutos.

15

La proteína de la muestra puede ser digerida enzimáticamente, por ejemplo usando una proteinasa, opcionalmente Proteinasa K. La proteína se puede digerir enzimáticamente durante un periodo de aproximadamente 30 minutos a 1 hora, por ejemplo, durante aproximadamente 45 minutos. En una realización después de la digestión de proteína, por ejemplo, digestión de proteinasa K, no hay inversión de reticulación o etapa de extracción de ADN de fenol.

20

En una realización, la detección por PCR es capaz de detectar una copia individual del ácido nucleico ligado, preferentemente con una lectura binaria para la presencia/ausencia del ácido nucleico ligado.

#### Terapia e Identificación de Agentes Terapéuticos

25

Los inventores han descubierto cómo las interacciones cromosómicas se relacionan con la terapia de cáncer, incluyendo cómo los compuestos utilizados en el tratamiento de cáncer también pueden actuar directamente sobre conformaciones cromosómicas 3C. Por ejemplo, se ha encontrado que imatinib mimético de ATP, conocido por actuar directamente sobre tirosina quinasas activadas, actúa directamente sobre conformaciones cromosómicas 3C en genes sensibles a imatinib tales como cKIT.

30

La identificación o monitorización de un cambio en la conformación cromosómica asociado a la diana de un agente terapéutico conocido se puede usar por tanto para identificar compuestos que tienen como resultado el mismo cambio en la conformación y tienen el mismo efecto terapéutico. La identificación o monitorización de un cambio en la conformación cromosómica asociado a la diana de un agente terapéutico conocido se puede usar por tanto para identificar sujetos, o células, que sean susceptibles al tratamiento por el agente terapéutico u otros agentes terapéuticos que también puedan causar el cambio en la conformación cromosómica.

35

Además, la invención proporciona ácidos nucleicos terapéuticos que se pueden usar para prevenir o tratar cáncer alterando una interacción cromosómica relevante, por ejemplo como se menciona en la presente memoria. Dicho ácido nucleico puede ser ADN o ARN, y puede funcionar por "interferencia" tal como ARNi.

40

#### Homólogos y Fragmentos

45

La invención incluye el uso de homólogos y/o fragmentos de ácidos nucleicos mencionados en la presente memoria, incluyendo las secuencias específicas divulgadas en la presente memoria. Dichos homólogos y/o fragmentos pueden tener las mismas propiedades de unión que el ácido nucleico original, por ejemplo siendo capaces de unirse específicamente al producto ligado relevante.

50

Los homólogos de polinucleótido/ácido nucleico (por ejemplo, ADN) tienen normalmente al menos un 70 % de homología, preferentemente al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 %, al menos un 98 % o al menos un 99 % de homología, por ejemplo sobre una región de al menos 10, 15, 20, 30, 100 o más nucleótidos contiguos, o a través de la parte de ácido nucleico que es de la región del cromosoma implicado en la interacción cromosómica. La homología se puede calcular en base a la identidad de nucleótidos (a veces denominada "homología dura").

55

Por ejemplo, el Paquete UWGCG proporciona el programa BESTFIT que se puede utilizar para calcular la homología y/o % de identidad de secuencia (por ejemplo, utilizado en sus ajustes por defecto) (Devereux et al (1984) Nucleic Acids Research 12, p387-395). Los algoritmos PILEUP y BLAST se pueden usar para calcular la homología y/o % de identidad de secuencia y/o secuencias de alineación (tales como identificación de secuencias equivalentes o correspondientes (normalmente en sus ajustes por defecto), por ejemplo como se describe en Altschul S. F. (1993) J Mol Evol 36:290-300; Altschul, S, F et al (1990) J Mol Biol 215:403-10.

60

El software para la realización del análisis con BLAST se encuentra disponible públicamente en National Center for Biotechnology Information. Este algoritmo implica identificar primero los pares de secuencias con alta puntuación (HSP, por sus siglas en inglés) mediante la identificación de palabras cortas de longitud W en la secuencia problema que corresponde o satisface algún valor de umbral T valorado de forma positiva cuando se alinea con una palabra de la misma longitud en la secuencia de base de datos. La T hace referencia al umbral de puntuación de la palabra más próxima (Altschul y col., supra). Estos resultados iniciales de palabra

65

más próxima actúan como punto de partida para iniciar las búsquedas con el fin de encontrar HSP más largos que las contienen. Los resultados de la palabra se extienden en ambas direcciones a lo largo de cada secuencia hasta el grado en que se pueda incrementar el valor de alineación acumulativo. La extensión de los resultados de la palabra en cada dirección se detiene cuando: el valor de alineación acumulativo desciende en la cantidad X del valor máximo logrado; el valor acumulativo disminuye a cero o por debajo de cero, debido a la acumulación de una o más alineaciones de residuos de valor negativo; o se alcanza el final de alguna secuencia. Los parámetros W5, T y X del algoritmo BLAST determinan la sensibilidad y la velocidad del alineación. Para secuencias de aminoácidos, el programa BLAST utiliza como predeterminada una longitud de palabra (W) de 11, alineaciones (B) de matriz de puntuación BLOSUM62 (Henikoff y Henikoff (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919) de 50, expectativa (E) de 10, M=5, N=4 y una comparación de ambas hebras.

El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre las dos secuencias; véase, por ejemplo, Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873-5787. Una medida de similitud proporcionada por el algoritmo BLAST es la menor probabilidad de suma (P(N)), que proporciona una indicación de la probabilidad de que una coincidencia entre dos secuencias de polinucleótidos ocurra por casualidad. Por ejemplo, se considera que una secuencia es similar a otra secuencia si la menor probabilidad de suma al comparar la primera secuencia con la segunda secuencia es menor que aproximadamente 1, preferentemente menor que aproximadamente 0,1, más preferentemente menor que aproximadamente 0,01, y lo más preferentemente menor que aproximadamente 0,001.

La secuencia homóloga difiere normalmente en 1, 2, 3, 4 o más bases, tales como menos de 10, 15 o 20 bases (que pueden ser sustituciones, eliminaciones o inserciones de nucleótidos). Estos cambios se pueden medir a través de cualquiera de las regiones mencionadas anteriormente en relación con el cálculo de la homología y/o % de identidad de secuencia.

Los fragmentos normalmente tienen longitudes que son al menos un 90 %, 80 % o 50 % de la secuencia original. Pueden tener un truncamiento de al menos 5, 8 o 10 nucleótidos desde el extremo 5' y/o 3'.

#### Matrices

El segundo conjunto de ácidos nucleicos puede estar unido a una matriz, y en una realización hay al menos 15.000, 45.000, 100.000 o 250.000 segundos ácidos nucleicos diferentes unidos a la matriz, que preferentemente representan al menos 300, 900, 2000 o 5000 locus. En una realización, una, o más, o todas las diferentes poblaciones de segundos ácidos nucleicos están unidas a más de una región distinta de la matriz, en efecto repetida en la matriz permitiendo la detección de errores. La matriz puede estar basada en una plataforma de micromatrices Agilent SurePrint G3 Custom CGH. La detección de unión de los primeros ácidos nucleicos a la matriz se puede llevar a cabo mediante un sistema de color dual.

#### Agentes Terapéuticos:

En la presente memoria se mencionan agentes terapéuticos. La invención proporciona dichos agentes para su uso en la prevención o tratamiento de la afección relevante. Esto puede comprender administrar a un individuo que lo necesite una cantidad terapéuticamente eficaz del agente. Los métodos de la invención se pueden usar para seleccionar un individuo para el tratamiento. Los métodos de la invención, y en particular el método para llevar a cabo un ensayo epigenético con acompañante de pronóstico, pueden incluir una etapa de tratamiento en la que la persona identificada por medio del método se puede administrar a continuación con un agente que previene o trata la afección relevante, por ejemplo, un inhibidor de tirosina quinasa, preferentemente uno del grupo de inhibidores conocidos de tirosina quinasa que incluye Cediranib, Crenolanib, Dasatinib (Sprycel); Dovitinib, Imatinib (Gleevec); IMC-3G3 (Olaratumab, LY3012207) Lartruvo (PDGFRA); Masitinib (AB1010), Nilotinib (Tasigna); Pazopanib; PKC412 (midostaurina, CGP41251); Quizartinib; Regorafenib; Sorafenib (Nexavar); Sunitinib (Sutent); Vatalanib; S116836, más preferentemente uno o más de un grupo que comprende Dasatinib, crenolanib, dovitinib e Imatinib.

La formulación de agente dependerá de la naturaleza del agente. El agente se proporciona en forma de una composición farmacéutica que contiene el agente y un vehículo o diluyente farmacéuticamente aceptable. Los vehículos y diluyentes adecuados incluyen disoluciones salinas isotónicas, por ejemplo disolución salina tamponada con fosfato. Las composiciones de dosificación oral típicas incluyen comprimidos, cápsulas, disoluciones líquidas y suspensiones líquidas. El agente se puede formular para administración parenteral, intravenosa, intramuscular, subcutánea, transdérmica u oral.

La dosis de agente se puede determinar según diversos parámetros, especialmente según la sustancia utilizada; la edad, peso y condición del individuo a tratar; la vía de administración; y el régimen requerido. El médico será capaz de determinar la vía de administración y la dosificación requeridas para cualquier agente particular. No obstante la dosis adecuada puede ser de 0,1 a 100 mg/kg de peso corporal, tal como 1 a 40 mg/kg de peso corporal, por ejemplo, para una toma de 1 a 3 veces al día.



#### Formas de la Sustancia Mencionada en la Presente Memoria

5 Cualesquiera de las sustancias, tales como ácidos nucleicos o agentes terapéuticos, mencionadas en la presente memoria pueden estar en forma purificada o aislada. Los compuestos pueden estar en una forma que sea diferente de la encontrada en la naturaleza, por ejemplo, pueden estar presentes en combinación con otra sustancia con la que no aparecen en la naturaleza. Los ácidos nucleicos (incluyendo partes de secuencias definidas en la presente memoria) pueden tener secuencias que sean diferentes a las encontradas en la naturaleza, por ejemplo, que tengan al menos 1, 2, 3, 4 o más cambios de nucleótidos en la secuencia como se describe en la sección sobre homología. Los ácidos nucleicos pueden tener una secuencia heteróloga en el extremo 5' o 3'. Los ácidos nucleicos pueden ser químicamente diferentes de los que se encuentran en la naturaleza, por ejemplo, se pueden modificar de modo alguno, pero preferentemente todavía son capaces de apareamiento de bases de Watson-Crick. Cuando sea apropiado, los ácidos nucleicos se proporcionan en forma bicatenaria o monocatenaria. La invención proporciona todas las secuencias específicas de ácido nucleico mencionadas en la presente memoria en forma monocatenaria o bicatenaria, y por tanto incluye la hebra complementaria a cualquier secuencia divulgada.

20 La invención también proporciona un estuche para llevar a cabo cualquier proceso de la invención, incluyendo la detección de una interacción cromosómica asociada a un subgrupo particular. Dicho estuche puede incluir un agente de unión específico capaz de detectar la interacción cromosómica relevante, tal como agentes capaces de detectar un ácido nucleico ligado generado por medio de los procesos de la invención. Los agentes preferidos presentes en el estuche incluyen sondas capaces de hibridar con el ácido nucleico ligado o pares de cebadores, por ejemplo sondas hidrolizables como se describe en la presente memoria, capaces de amplificar el ácido nucleico ligado en una reacción de PCR o una reacción de qPCR.

25 La invención también proporciona un dispositivo que es capaz de detectar las interacciones cromosómicas relevantes. El dispositivo preferentemente comprende cualesquiera agentes de unión específicos, sonda o par de cebadores capaces de detectar la interacción cromosómica, tal como cualquier agente, sonda o par de cebadores tal como se describe en la presente memoria.

30 Los ácidos nucleicos mencionados en la presente memoria pueden estar en forma de ADN.

#### Realizaciones Específicas

35 La tecnología de plataforma EpiSwitch™ detecta firmas reguladoras epigenéticas de cambios regulatorios entre condiciones normales y anormales en lugares. La plataforma EpiSwitch™ identifica y supervisa el nivel epigenético fundamental de la regulación génica asociada a las estructuras reguladoras de alto orden de los cromosomas humanos también conocidas como firmas de conformación cromosómica. Las firmas cromosómicas son una etapa primaria distinta en una cascada de desregulación génica. Son biomarcadores de alto orden con un conjunto único de ventajas frente a plataformas de biomarcadores que utilizan biomarcadores de expresión génica y epigenética tardía, tales como metilación de ADN y perfil de ARN.

#### Ensayo de matriz EpiSwitch™

45 Las plataformas de detección de matriz personalizadas EpiSwitch™ vienen en 4 densidades de conformaciones cromosómicas únicas de 15K, 45K, 100K y 250K, cada fragmento químico se repite en las matrices 4 veces, haciendo que las densidades efectivas sean 60K, 180K, 400K y 1 millón, respectivamente.

#### Matrices EpiSwitch™ Diseñadas a Medida

50 La matriz EpiSwitch™ 15K puede detectar la totalidad del genoma incluyendo alrededor de 300 locus interrogados con la tecnología de descubrimiento EpiSwitch™ Biomarker. La matriz EpiSwitch™ está construida sobre la plataforma de micromatrices Agilent SurePrint G3 Custom CGH; esta tecnología ofrece 4 densidades, 60K, 180K, 400K y 1 millón de sondas. La densidad por matriz se reduce a 15K, 45K, 100K y 250K a medida que cada sonda EpiSwitch™ se presenta por cuadruplicado, permitiendo así la evaluación estadística de la reproducibilidad. El número promedio de marcadores EpiSwitch™ potenciales interrogados por locus genéticos es 50; como tal, los números de locus que se pueden investigar son 300, 900, 2000 y 5000.

#### Canalización de Matriz Personalizada EpiSwitch™

60 La matriz EpiSwitch™ es un sistema de color dual con un conjunto de muestras, después de la generación de la biblioteca EpiSwitch™, marcado en Cy5 y la otra de muestra (controles) a comparar/analizar marcada en Cy3. Las matrices se exploran utilizando el escáner Agilent SureScan y las características resultantes se extraen utilizando el software Agilent Feature Extraction. Los datos se procesan después usando los guiones de procesamiento de matriz EpiSwitch™ en R. Las matrices se procesan usando paquetes de color dual convencionales en Bioconductor en R: Lima\*. La normalización de las matrices se hace usando la función normalizada en las matrices en Lima\* y esto se hace a los controles positivos Agilent en chip y controles

positivos EpiSwitch™. Los datos se filtran en base a las llamadas Agilent Flag, se eliminan las sondas de control Agilent y se calculan los promedios de las sondas técnicas replicadas, para analizarlas utilizando Lima\*. Las sondas se someten a modelización en base a su diferencia entre los 2 escenarios que se comparan y a continuación se corrigen utilizando una Tasa de Descubrimiento Falsa. Las sondas con coeficiente de variación (CV)  $\leq 30\%$  que son  $\leq -1,1$  o  $\geq 1,1$  y pasan el valor p de  $p \leq 0,1$  FDR se usan para la detección adicional. Para reducir el conjunto de sondas, se lleva a cabo un Análisis de Factor Múltiple adicional utilizando el paquete FactorMineR en R.

\*Nota: LIMMA es Modelos Lineales y Procesos de Bases Empíricas para Evaluar la Expresión Diferencial en Experimentos de Micromatrices. Limma es un paquete R para el análisis de datos de expresión génica que surgen de la micromatriz o ARN-Sec.

El conjunto de sondas se selecciona inicialmente en base a los parámetros de valor p ajustado, FC y CV  $< 30\%$  (punto de corte arbitrario) para la selección final. Los análisis adicionales y el listado final se dibujan basándose únicamente en los dos primeros parámetros (ajuste del valor p; FC).

#### Canalización Estadística

Las matrices de detección EpiSwitch™ se procesan utilizando el Paquete Analítico EpiSwitch™ en R con el fin de seleccionar marcadores EpiSwitch™ de alto valor para su traducción a la plataforma PCR EpiSwitch™.

#### Etapas

Las sondas se seleccionan basándose en su valor p corregido (Tasa de Descubrimiento Falsa, FDR), que es el producto de un modelo de regresión lineal modificado. Las sondas por debajo del valor  $p \leq 0,1$  se seleccionan y luego se reducen adicionalmente por su relación epigenética (ER), las sondas ER tienen que ser  $\leq -1,1$  o  $\geq 1,1$  con el fin de poder ser seleccionadas para análisis adicional. El último filtro es un coeficiente de variación (CV), las sondas tienen que estar por debajo de  $\leq 0,3$ .

#### Etapas

Los 40 marcadores superiores de las listas estadísticas se seleccionan basándose en su ER para la selección como marcadores para la traducción de PCR. Los 20 marcadores superiores con la carga de ER negativa más elevada y los 20 marcadores superiores con la carga de ER positiva más elevada constituyen el listado.

#### Etapas

Los marcadores resultantes de la Etapa 1, las sondas estadísticamente significativas forman las bases del análisis de enriquecimiento usando enriquecimiento hipergeométrico (HE). El presente análisis permite la reducción de marcadores del listado de sondas significativas, y junto con los marcadores de la Etapa 2 forma el listado de sondas traducidas a la plataforma de PCR EpiSwitch™.

Las sondas estadísticas se procesan por HE para determinar qué ubicaciones genéticas tienen un enriquecimiento de sondas estadísticamente significativas, indicando qué ubicaciones genéticas son centros de diferencia epigenética.

Los locus enriquecidos más significativos basados en un valor p corregido se seleccionan para la generación del listado de sondas. Se seleccionan localizaciones genéticas por debajo del valor p de 0,3 o 0,2. Las sondas estadísticas que se someten a mapeado a estas ubicaciones genéticas, con los marcadores de la etapa 2, forman los marcadores de alto valor para la traducción de PCR EpiSwitch™.

#### Diseño y procesado de matrices

1. Los locus genéticos se procesan utilizando el software SII (actualmente v3.2) para:

a. Extraer la secuencia del genoma en estos locus genéticos específicos (secuencia génica con 50 kb aguas arriba y 20 kb aguas abajo)

b. Definir la probabilidad de que una secuencia dentro de esta región esté implicada en las CC

c. Cortar la secuencia utilizando un RE específico

d. Determinar qué fragmentos de restricción son susceptibles de interactuar en una determinada orientación

e. Clasificar la probabilidad de que diferentes CCs interactúen juntos.

2. Determinar el tamaño de matriz y por tanto el número de posiciones de sonda disponibles (x)

3. Extraer las interacciones x/4.

5 4. Para cada interacción, definir la secuencia de 30 pb al sitio de restricción desde la parte 1 y 30 pb al sitio de restricción de la parte 2. Comprobar que esas regiones no son repeticiones, si es así, excluir y anotar la siguiente interacción en el listado. Unir ambos 30 pb para definir la sonda.

10 5. Crear un listado de sondas x/4 más sondas de control definidas y replicar 4 veces para crear el listado que se va a crear en la matriz

6. Cargar el listado de sondas en el sitio web de diseño Agilent Sure para la matriz de CGH personalizada.

15 7. Utilizar un grupo de sondas para diseñar una matriz de CGH personalizada Agilent.

#### Procesado de Matriz

1. Procesar muestras usando EpiSwitch™ SOP para la producción de plantillas.

20 2. Limpiar con precipitación de etanol mediante un laboratorio de procesamiento de matrices.

3. Procesar las muestras según el estuche de marcaje de ADN completo Agilent SureTag - CGH basada en Agilent Oligonucleotide Array para análisis de ADN genómico de marcaje enzimático para sangre, células o tejidos

25 4. Explorar usando el escáner Agilent C usando el software de extracción de características Agilent.

#### Ejemplos

30 Ejemplo 1. Análisis de interacciones cromosómicas en cáncer

#### Introducción

35 Los métodos de diagnóstico oncológico actuales basados en ADN genómico son costosos, semicuantitativos y fallan a la hora de capturar todas las variantes de punto de corte. Los métodos de diagnóstico basados en RT-PCR anidada requieren células vivas, se deben confirmar usando otros métodos y tienen una etapa de procesamiento de ARN compleja. El método desarrollado en el presente trabajo captura todas las variantes de punto de ruptura, es cuantitativo, no tiene problemas de estabilidad de ARN, utiliza volúmenes de muestra pequeños y tiene un tiempo de procesamiento rápido. Por tanto, el método reduce la posibilidad de diagnóstico erróneo y de proporcionar el tratamiento incorrecto.

45 El presente trabajo también identifica una interacción 3C que tiene capacidad de respuesta terapéutica a un inhibidor de tirosina quinasa. Esto forma la base del ensayo pronóstico que se puede usar para detectar pacientes que se puedan tratar con inhibidores de tirosina quinasa. Los ensayos actuales no son precisos o sensibles, por ejemplo, debido a la variación del punto de ruptura. Un diagnóstico erróneo se traduce en que los pacientes no reciban terapia con inhibidores de tirosina quinasa, y en su lugar reciban esteroides a los que no muestran respuesta alguna. Por ejemplo, el pronóstico para los pacientes no tratados con iHES es malo.

50 Se investigaron interacciones epigenéticas relacionadas con cáncer y su utilidad en la estratificación de pacientes para el tratamiento con inhibidores de tirosina-quinasa. Utilizando una plataforma de captura de conformación cromosómica de alta resolución o análisis 3C conocida como EpiSwitch™ y PCR cuantitativa, se sometieron a mapeado, evaluación y cuantificación la yuxtaposición conformacional entre FIP1L1 y PDGFRA en estirpes celulares de glioma con y sin mutaciones IDH. La desregulación de PDGFRA por eliminación intersticial en 4q12 y fusión a FIP1L1 asociada a leucemia eosinofílica crónica llevó al grupo de la invención a investigar también si las mismas interacciones de conformación cromosómica estaban presentes en EOL-1 y otras estirpes celulares leucémicas.

60 Usando una captura de conformación cromosómica de alta resolución o una plataforma de análisis 3C conocida como EpiSwitch™, se identificó una selección de interacciones epigenéticas 3C relacionadas con la mutación. Para optimizar los cebadores los inventores usan un gradiente de temperatura en un ensayo de placa de 96 pocillos con el Sistema Qiagen QIAgility de PCR automatizado de alta precisión y someten a ensayo 72 cebadores en 36 PCR de gradiente de temperatura separadas. Se seleccionaron un total de 20 interacciones 3C individuales predichas por OBD. Se usaron múltiples conjuntos de cebadores para cada interacción predicha cuando fue posible. Se optimizaron PCRs de etapa individual usando 10 ng de plantilla de diana-3C de dos estirpes celulares de AML EOL-1 y HL-60. Usando detección SYBR y controles negativos de concentración coincidentes, se llevó a cabo la amplificación por PCR (de las dianas). Los controles negativos

incluyeron EOL-1 sin formaldehído procesado como muestra 3C (con digestión y enlace) y ADN genómico normal. Todos los productos de PCR se analizaron en LabChip. Los productos de PCR del tamaño esperado se purificaron usando gel (Invitrogen) y se secuenciaron en las orientaciones de cebador directo e inverso. Los cebadores se usaron a 3,2 pmol/μl por reacción con una concentración de plantilla de 1 ng/μl por 100 pb. Se

5 avanzaron productos que mostraban un 100 % de homología con el fragmento 3C predicho para el diseño de qPCR.

Para seleccionar las mutaciones IDH1 SNP rs11554137 e IDH1 R132H, se usaron 10 ng de plantilla 3C con cebadores genómicos IDH1 convencionales procedentes de publicaciones académicas; estas secuencias de

10 cebadores se diseñaron externamente y no se incluyen en la presente solicitud.

La qPCR de etapa individual se llevó a cabo usando qPCRs basados en sonda de hidrólisis con controles negativos emparejados por concentración. La detección de la interacción B7 de 462 pb se optimizó a continuación usando una sonda de hidrólisis diseñada frente a la región de unión del fragmento de ADN 3C

15 objetivo. Las sondas usadas en los ensayos de qPCR son específicas para los fragmentos de ADN 3C. Se usó una plantilla EOL-1 (10 ng) para optimizar las sondas en una qPCR basada en gradiente de temperatura (58-68 °C). Se seleccionaron cebadores para el desarrollo del ensayo de qPCR de sonda de hidrólisis. El producto B7 se utilizó para diseñar una qPCR específica de fragmento 3C. Se usó una sonda de hidrólisis de doble marcaje, marcada con 5'-FAM/BHQ1-3' o 5' Texas Red/BHQ2-3', para detectar la interacción secuenciada. La

20 sonda se diseñó para abarcar la región de unión del fragmento 3C, asegurando de este modo que la detección del producto 3C fuera específica para la interacción FIPL1L1 y PDGFRA. Se produjo una curva convencional de qPCR (copia  $10^{6-1}$ ) a partir del producto secuenciado.

Se usó una sonda de hidrólisis marcada con 5'-FAM-BHQ1-3' de doble marcaje para detectar la interacción secuenciada. Todas las muestras se sometieron a detección a 20 ng y se registró el número de copias de MMP1. Se usó el producto de PCR secuenciado B7 para generar curvas convencionales para detección de qPCR. Se usaron plantillas de qPCR, ajustadas a 20 ng de ADN de la biblioteca 3C, con controles negativos de concentración coincidente. Los controles incluyeron bibliotecas 3C procedentes de biopsias adiposas y

25 sangre normal. Los controles negativos adicionales incluyeron material de estirpe celular sin fijación de formaldehído, material de biblioteca digerido y ligado y ADN genómico normal. El presente estudio se llevó a cabo según las directrices de Sociedad Americana de Hematología. Las estirpes celulares fueron proporcionadas por ATCC. El material del paciente con consentimiento ético fue proporcionado por Tissue Solutions (Oxford BioDynamics TSD-OB2273).

De las seis secuencias de orientaciones de secuencia FIP1L1-PDGFRA 3C identificadas, dos reaparecieron (Tabla 1). Específicamente, la secuencia genómica marcada GBA461GB\_GBA496A6 se identificó cuatro veces. La secuencia genómica FH147946\_FH149AA7 se identificó dos veces. Usando PCR de etapa individual, se amplificaron estas seis interacciones en la estirpe celular de AML EOL-1. La interacción PDGFRA

35 3C B7 de FIP1L1 se encontró de manera reproducible en la estirpe celular de AML (leucemia eosinofílica mieloide aguda) EOL-1 (n = 2, ensayos de sonda de hidrólisis) usando PCR de etapa individual. La pérdida del aislamiento topológico entre FIP1L1 y PDGFRA en EOL-1 puede estar mediada en parte por eliminación.

La interacción de 462 pares de bases B7 se amplificó en dos estirpes celulares sensibles a imatinib: GDM-1 (n = 1) y HL-60 (n = 2). Se desconoce el mecanismo de pérdida de aislamiento topológico entre los genes FIP1L1 y PDGFRA, no obstante, ambas estirpes celulares responden a imatinib. Las estirpes celulares de glioblastoma DBTRG-05MG (n = 2 ensayos de qPCR) y U-373 (n = 1 ensayo de qPCR) también son positivas para la interacción B7. El número máximo de copias encontrado en DBTRG-05MG fue de 29,89 copias (20 ng de

45 plantilla, Tabla 2). Una muestra de glioma sin mutación IDH1 identificada o SNP rs11554137 presentó 1,19 (por 20 ng) de fragmentos FIPL1-PDGFRA 3C. Dos muestras de glioma con mutación IDH1 R132H identificada fueron positivas para fragmentos FIPL1-PDGFRA 3C.

Las curvas convencionales de qPCR (copia  $10^{6-1}$ ) se produjeron a partir de los productos de PCR secuenciados. El análisis de datos de qPCR se llevó a cabo usando el software gestor CFX (BioRad). La normalización se llevó a cabo con una interacción 3C estable en MMP1. Las curvas convencionales mostraron

55 amplificación selectiva de la interacción B7 en muestras de AML y glioblastoma usando qPCR de etapa individual.

El uso de conformación cromosómica detectó con éxito secuencias de ADN relacionadas con la pérdida de función aislante entre FIPL1 y PDGFRA en estirpes celulares de glioblastoma y AML. La pérdida de aislamiento entre los genes FIP1L1 y PDGFRA se traduce en una alteración 3C que se puede detectar en AML y glioblastoma usando qPCR. Estas interacciones resultarían indetectables usando la metodología de secuenciación convencional.

60

Ejemplo 2. El desarrollo de ensayos de PCR cuantitativa

65

Para cumplir con los requisitos de datos y técnicos para qPCR abordados en las directrices de MIQE, los

inventores desarrollan técnicas novedosas para la amplificación y detección de 3C utilizando qPCR. La PCR cuantitativa en tiempo real (qPCR) permite la determinación del número de copias de ADN con elevada sensibilidad en muestras objetivo. Se encontró que esto era particularmente adecuado para las demandas de análisis de productos de ácido nucleico relacionados con la detección de interacciones cromosómicas.

5

Las directrices de MIQE proporcionan especificaciones para la información mínima que se debe notificar para un experimento de qPCR con el fin de garantizar su relevancia, exactitud, interpretación correcta y repetibilidad. MIQE (Información Mínima para la Publicación de Experimentos Cuantitativos de PCR en Tiempo Real) se somete a modelización sobre directrices similares elaboradas para el análisis de micromatrices de ADN, experimentos proteómicos, especificación de secuencias genómicas, y aquellos bajo discusión para el trabajo de interferencia de ARN y metabolómica, siendo todos ellos iniciativas coordinadas bajo el paraguas de MIBBI (Información Mínima para Investigaciones Biológicas y Biomédicas)

10

Convencionalmente, en el análisis 3C en los casos en los que se usa un enfoque de PCR basado en fluorescencia, se diseñan a continuación múltiples conjuntos de cebadores en una región "constante" definida de ADN complejo del fragmento 3C. No obstante, dichos enfoques no se adhieren a los principios MIQE para qPCR, y son técnicamente una derivación de la PCR anidada. Dichos enfoques requieren que se conozca al menos la mitad de la interacción.

15

En las técnicas de detección convencionales, ADN polimerasa escinde únicamente las sondas que están hibridadas con la diana de ADN conocida y si el cebador de PCR inversa se une al mismo fragmento de biblioteca 3C. El aumento de fluorescencia tiene lugar únicamente si la diana es complementaria a la sonda y se amplifica durante PCR. Debido a estos requisitos, no se detecta amplificación inespecífica.

20

A partir de la cantidad de ADN genómico requerida para detectar productos de unión, los inventores estiman que algunos episodios de unión específicos de locus pueden ocurrir en solo 1/2.000 a 1/20.000 células de mamífero. Por tanto, la cuantificación fiable de los productos de unión con enfoques convencionales solo es posible si se añaden grandes cantidades de ADN (es decir, muchos equivalentes del genoma) a cada PCR en técnicas convencionales. No obstante, los inventores emplean rutinariamente 20-100 ng de plantilla 3C ( $\sim 3,2 \times 10^4$ - $1,6 \times 10^5$  equivalentes de genoma) por PCR con la nueva técnica que han desarrollado. La tecnología de los inventores requiere sondas hidrolizables 3C para unirse sobre la unión de ligadura. También este análisis de micromatriz para que sea compatible con el análisis de PCR.

25

30

¿Por qué la tecnología existente no funciona utilizando técnicas convencionales? Cuestiones técnicas relativas a la especificidad de detección y cuantificación del número de copia 3C.

35

El enfoque convencional requiere dos conjuntos de cebadores y dos rondas de PCR. Técnicamente, este es un derivado de PCR anidada con detección basada en fluorescencia. Otros enfoques imitan esto, pero con un conjunto de cebadores y entre cuarenta y cincuenta ciclos de PCR. Aunque se pueden apreciar diferencias entre las muestras, el enfoque no se puede describir como cuantitativo y contiene una serie de cuestiones técnicas graves:

40

(a) Se acepta PCR anidada como no cuantitativa, las diferencias de eficiencia entre los conjuntos de cebadores y las PCR que alcanzan la fase de meseta en algunas reacciones imposibilitan una cuantificación exacta y precisa de bajo número de copias.

45

(b) Los productos no específicos se pueden amplificar con cebadores anidados.

(c) Los ensayos a menudo no están secuenciados.

50

(d) Con el enfoque de PCR en dos rondas no se puede obtener una cuantificación absoluta "verdadera" con respecto al número de copias a partir de los datos.

(e) No se construyen curvas convencionales (que requieren una ronda de PCR), lo que se hace imposible una cuantificación exacta y precisa de número de copias bajo.

55

(f) Los estuches existentes no cumplen las directrices establecidas para ensayos de qPCR clínicos seguros.

(g) Las PCR TaqMan amplifican frecuentemente múltiples productos de PCR, lo que plantea serios problemas técnicos con respecto a la especificidad de producto.

60

(h) Con respecto a la letra (g), el uso de sondas de hidrólisis para detección de 3C, las sondas están situadas en la "región constante".

(i) La región constante no es un área específica de secuencia para los fragmentos 3C o el sitio TaqI; cualquier reivindicación de especificidad de que el producto de PCR detectado por la sonda de hidrólisis es un fragmento

65

3C resulta, por tanto, dudosa.

Los enfoques convencionales utilizan los siguientes controles:

- 5 Control 1. Preparar una plantilla de control que contenga todos los productos de unión en cantidades iguales. Utilizar esta plantilla para normalizar las diferencias en la eficacia del cebador.

- Control 2. Determinar las interacciones entre sitios separados aumentando las distancias genómicas para estimar la frecuencia de interacciones aleatorias. Un buen punto de partida sería someter a ensayo  
10 aproximadamente una docena de interacciones entre sitios separados por distancias genómicas que varían de 0 a 100 kb.

- Control 3. Cuando se comparan dos tipos de células diferentes o dos condiciones diferentes, se determinan interacciones en una región genómica que se supone que tiene una conformación similar en los dos casos.  
15 Utilizar este conjunto de datos para la normalización de manera que los datos obtenidos a partir de la región de interés se puedan comparar de forma cuantitativa.

- Se detecta una interacción específica cuando después de la normalización para la eficiencia del cebador (control 1), se observa un pico local en la frecuencia de interacción (control 2). El análisis de una región genómica no relacionada (control 3) se puede usar a continuación para comparar diferentes células o estados celulares con el fin de establecer si la interacción específica está correlacionada con actividades biológicas tales como transcripción del locus.  
20

- No obstante, la técnica que han desarrollado los inventores no requiere que dichos controles estén en su lugar.  
25 La amplificación y detección específicas se logra debido a la elección de dianas de unión de los cebadores y la sonda, combinados con el uso de un sistema de marcaje de agente de inactivación de fluorescencia. Por ejemplo, la Tabla 3 enumera algunas sondas fluorescentes y agentes de inactivación adecuados.

- Ejemplo 3. Características de secuencias de diana  
30

- La tecnología qPCR EpiSwitch™ se utiliza exclusivamente para detección de las sondas de hidrólisis objetivo que se extienden a través de un sitio de restricción TaqI en fragmentos de ADN yuxtapuestos unidos (Tabla 4). Normalmente, los sitios elegidos para la detección con sondas de hidrólisis se eligen para la eficiencia de la hibridación frente a la sonda y su propensión a la formación de estructuras secundarias y otras propiedades biofísicas que podrían interferir con el proceso de hibridación de la sonda, incluyendo propiedades de plegado y flexión y niveles mínimos de energía libre. Esto explica por qué, hasta ahora, cualquier uso de sondas de hidrólisis en el contexto de la detección 3C apuntaba a las sondas dentro de una parte del fragmento de unión y no se superponía a ambas. Esto socava en gran medida la especificidad de la detección solo para los productos unidos frente a todos los fragmentos unidos y no unidos dentro de la secuencia objetivo.  
35

- La curvatura de ADN y los niveles mínimos de energía libre en las yuxtaposiciones unidas de los sitios Check Point Charlie (CC) alrededor del sitio de restricción TaqI indican la conformación de estructura secundaria que diferencia las secuencias dirigidas convencionales para la hibridación de sonda de hidrólisis. Las sondas se diseñan en los fragmentos TaqI yuxtapuestos que abarcan el sitio de restricción. La curvatura de ADN se calcula mediante el algoritmo BEND<sup>1</sup> sumando los ángulos de enrollamiento, inclinación y torsión de la geometría de los dinucleótidos. La curvatura se expresa como grados por vuelta helicoidal, por ejemplo, 10,5° por vuelta helicoidal o 1° por par de bases. La energía libre mínima,  $\Delta G$ , en el sitio de restricción TaqI se calcula mediante el algoritmo de Zuker<sup>2</sup> con la estimación de la minimización de energía usando los parámetros de energía vecinos más cercanos<sup>3</sup>.  
40

- En la tecnología EpiSwitch™, los inventores exploran hasta 400 bases de secuencias aguas arriba y aguas abajo del sitio TaqI y calculan la curvatura de ADN. Como consecuencia de la unión conjunta de dos CC de estructura secundaria elevada, las secuencias dirigidas a los sitios de unión muestran una diferencia cuantitativa en la propensión a la curvatura, en comparación con las secuencias objetivo convencionales para hibridación de sonda de hidrólisis. La tecnología busca una puntuación máxima de propensión a curvatura de más de >5° por vuelta helicoidal. Los valores comparados con el valle de secuencias vecinas por debajo de 5° por vuelta helicoidal podrían mostrar que ADN puede adoptar una conformación curvada.  
45

- También, los inventores exploran hasta 20 bases de secuencias aguas arriba y aguas abajo del sitio de restricción TaqI y calculan la energía libre mínima,  $\Delta G$ , de ADN lineal a 60°C en las condiciones iónicas de Na<sup>+</sup> 0 mM y Mg<sup>++</sup> 4,5 mM con corrección de oligo como se describe en el algoritmo de Zuker. Se selecciona la secuencia con estructura termodinámica con los niveles de energía libre por debajo de < 1,15 kcal/mol a 60 °C para el diseño de la sonda.  
50

- Dichas secuencias que residen en una región reguladora de gen tal como promotores, intrones, agentes de terminación, indican una conformación de estructura secundaria, identificada por la plataforma EpiSwitch™  
55

mediante el cálculo de los valores de curvatura y energía libre. Se diseñaron sondas que abarcaban el sitio de restricción TaqI en las secuencias seleccionadas con propensión a la curvatura y los niveles mínimos de energía libre descritos con anterioridad.

- 5 Aunque el uso convencional de sondas de hidrólisis no utiliza para la hibridación dirigida dichas secuencias biofísicamente restringidas en las uniones ligadas de los sitios de CC, en el contexto del procesamiento y detección de EpiSwitch los inventores pueden usar sondas de hidrólisis específicamente para el solapamiento de sitio de enlace de TaqI con todos los controles de qPCR que demuestran una detección eficiente y proporcionada de los productos ligados por números de copias frente a controles no específicos. Este es el primer uso práctico de las sondas de hidrólisis dirigidas a la superposición de 3C, realizado debido y en el contexto del ensayo EpiSwitch, diseñado sobre la base del ensayo de gradiente de temperatura para la detección de sonda eficaz.

Los ejemplos siguientes ilustran la metodología:

#### 15 Secuencia 1

```
TTCCACGTGGCCTACACAGCATGTCAGGCCTGGGGGCAGAATCTTGCCATACTGTGCAGCCCAAATTTGAATGCC
AAAGGCTTTCGTTTGTCTCTGGGGGGCCACAGTCTAGGTCTAGTTCTGTGCAGGAGTTGTAATATTTGCTCTTCTCT
CCCTCCTCCAGCTCGCAGACCTCTGAAGAGAGTGCCATTGAGACGGGTTCCAGCAGTTCCACCTTCATCAAGAGAG
AGGACGAGACCATTGAAGACATCGAAAAGATAAAACAGGTGTTAGTGAGGATATGGGGAAATAAAACCTCATA
CACTTCTGGTGGGATTGTAAAATGGTGCAGCTGCTTTGAAGAACAGCCTGGCCATTTTCTCAAATGGTGAAACATG
GAGTTACCATGTGAGTCCTCCACCTTAGAGGAATGAAAACATATTCAGACAAAACTGTACGTGAATATTCATAG
CAGCATTATTCTTAATAGAAAGTGGAAGAAAGAAACCTCGCAGCTG
```

20

El perfil de propensión a la curvatura de la secuencia se representa a continuación. En este caso, el perfil muestra una magnitud de propensión a curvatura superior a 5° entre 50 y 100 bases y seguida por un valle de puntuaciones bajas. Esto indica una conformación de curvatura de ADN. El plegamiento de ADN y la estimación de energía libre de la secuencia que rodea la región TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de 1,13 kcal/mol a 60 °C, 4,5 mM de  $Mg^{++}$  condición iónica. Esta región corresponde a la región intrónica de FIP1L1 y extremo 3' de genes de PDGFRA que indican una región altamente regulada. Se seleccionó este fragmento TaqI yuxtapuesto y se diseñó una sonda que abarcaba el sitio de restricción. El sitio de restricción de TaqI de la secuencia se muestra en blanco sobre un fondo negro.

25

#### 30 Secuencia 2

```
TGGGAGTGGGTGGAGTGAGAACCTGGGAGAAGGCCAGCCCTTATATCCAGGCAGACAGCTCCAAGTGCCACCA
TGGATCAGCCAGTCTTGAGGGGTGATGCTATTCAGCTACAGATGGCTTGATCCTGAGTCATTTCTTCTTTCCAT
GCAGTGTGTCCACCGTGATCTGGCTGCTCGCAACGTCCTCTGGCACAAGGAAAAATTGTGAAGATCTGTGACTTT
GGCCTGGCCAGAGACATCATGCATGATTGAAAAGATAAAACAGGTGTTAGTGAGGATATGGGGAAATAAAACC
CTCATACACTTCTGGTGGGATTGTAAAATGGTGCAGCTGCTTTGAAGAACAGCCTGGCCATTTTCTCAAATGGTGA
AACATGGAGTTACCATGTGAGTCCTCCACCTTAGAGGAATGAAAACATATTCAGACAAAACTGTACGTGAATAT
TCATAGCAGCATTATTCTTAATAGAAAGTGGAAGAAAGAAACCTCGCAGCTGCATCAACTGATGAATGGATAGAT
TAAATGTGTTATATCCATACAGCGGAATATTATTTGGCAAGGACAATAAAATGAAGTACTGGTAGATGTTACAACA
CGGATGAACCTTACAAATGTGGAAGCTAAAGATGTCAGTCCGT
```

35

Un barrido de 647 pb de secuencias alrededor de TaqI de dos fragmentos yuxtapuestos muestra una magnitud de propensión a curvatura de más de 5° por vuelta helicoidal en la posición de aproximadamente 210 pb que rodea un valle de bajas puntuaciones. El plegamiento de ADN y la estimación de energía libre de la secuencia de 44 pb que rodea la región TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de 0,78 kcal/mol a 60 °C, 4,5 mM de  $Mg^{++}$  condición iónica.

Secuencia 3

GCAGCTGCGAGGTTTTCTTTTCCACTTTCTATTAAGAATAATGCTGCTATGAATATTCACGTACAAGTTTTGTCTG  
AATATGTTTTTCATTCCTCTAAGGTGGAGGACTCACATGGTAACCTCATGTTTCACCATTTGAGAAAATGGCCAGGC  
TGTCTTCAAAGCAGCTGCACCATTTTACAATCCCACCAGAAGTGATGAGGGTTTTATTTCCCATATCCTCACTA  
  
ACACCTGTTTTATCTTTTCGAGGTCCTTGCTGAACCTGGACCTATAAATGACGTCAATGATAGTGATCCCTACTGCA  
GAAATCTACAAGTGGCTATAAAGAACTCTGTAGGTAAGAAATTCTGTAAGATCAGAAAGTACAATGAATTCACCTC  
ATAATAAATTACTTGGTGGACACCAAATGGGTGCTAAATTGATTGGGTAGAAGGAATTGTATGCCCAAGCCACAT  
GGC

- 5 Un barrido de 462 pb de secuencias que se extienden aguas arriba y aguas abajo del sitio de restricción TaqI mostró aproximadamente entre 400 y 450 pb, la magnitud de propensión de más de 5° por vuelta helicoidal con valores de secuencias circundantes por debajo de 5° por vuelta helicoidal. El plegamiento de ADN y la estimación de energía libre de la secuencia de 44 pb que rodea la región TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de -0,93 kcal/mol a 60 °C, 4,5 mM de Mg<sup>++</sup> condición iónica.

10

Secuencia 4

AGTACTTCCTCTCCCTCCCATATTGTTAAAAATAGTTTACATTGCTTCCAGGCTGGGCTGGTGGAGTTGGCACGA  
GATGTCAGAGGAACCTGAGTCATGCTCAGGCCCAAGCCCTGTTGGCAGGCAGACCACTGCTTCTGGCCTTCCGT  
GACTATCTGAAAAAATCGTGAATGGCTAGAGCTACTCTTCACCTGCTGAACATTTTCAAAAAGAATTGAGAACTT  
CTGGATTAAATTGCCTTCTTCTCGAAAAGATAAAACAGGTGTTAGTGAGGATATGGGGAAATAAAACCCCTCATAC  
ACTTCTGGTGGGATTGTAAAAATGGTGCAGCTGCTTTGAAGAACAGCCTGGCCATTTTCTCAAATGGTGAAACATG  
GAGTTACCATGTGAGTCCTCCACCTTAGAGGAATGAAAACATATTCAGACAAAAACTGTACGTGAATATTCATAG  
CAGCATTATTCTAATAGAAAGTGAAAAAGAAAACCTCGCAGC

- 15 Un barrido de 499 pb de secuencias que se extienden aguas arriba y aguas abajo del sitio de restricción TaqI mostró dos picos de propensión a la curvatura a aproximadamente 160 y 240 pb. El plegamiento de ADN y la estimación de energía libre de la secuencia de 44 pb que rodea la región TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de 1,14 kcal/mol a 60 °C, 4,5 mM de Mg<sup>++</sup> condición iónica.

20 Secuencia 5

CCTCTCCACACAAACCTGCTACTGAGTACCTTCGCTAACTTAACCATTCAATCACCCCTGGAAGACCACCTACTAGC  
AGAAGGATTCTTAACAAATGTAAAGAAAGTAAGGACTTTACACTAACAAATACAAAATAACTCTCTCTTTGACAAT  
TCAAAAAACAAAAGATGTTGAACCTTTGACATTTACAGAATTAAATGTCAAATGTGACACAATACCATCACATCTGG  
CTAACTAACTCTTATGCTTTTTTTAGTAAGGAACAACCTTTGAGCCTCAATATCTTAATTCTTAAAATGATAAAGAAC  
ACTTAACTCAATTTGTTGAGATCAAATAAGGTAATGTAAAAGTGGGATTTTTATTTTACTTATTTATTCGAAAACC  
CTGGGACCCCTCCAGATGGGACTAACTGGGGAAAGTGGACAAGTTACAAACAAAGAACTCAAAGGAAAGTCAT  
TGGCACTGATCTCTAAGATGCTATCACATGTGATTGGTGGTTGATTTTATTAACAAATTATAAGCAAAGTACTACAA  
AGGTGGCTTTAAAAAGAAAATAAAGCAATTCACAGAACTACTTTTTCATGTAGCTTGTATGTGTGCTCCATGT

- 25 Un barrido de 609 pb de secuencias que se extienden aguas arriba y aguas abajo del sitio de restricción TaqI mostró dos picos de propensión a la curvatura a aproximadamente 200 y 500 pb con un valle de puntuaciones bajas. El plegamiento de ADN y la estimación de energía libre de la secuencia de 44 pb que rodea la región TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de 0,51 kcal/mol a 60 °C, 4,5 mM de Mg<sup>++</sup> condición iónica.



Secuencia 6

AAATGATGAGGCACGGGTGAATAAGATAGTTGGAAGTGACACATACAGATTCAAATAATCCCTAAAGGTGTAT  
GGTTTTCTTCCCTACAAATCCTACTTTCAGTTCTTTGTACATGGACCCAGAGTGGAAATTGCTGGATCATATGATA  
ATTCTATTTTAACTTCTTGATGGACCTCTGTTTTGTTTTGTTTTTACAGAGGCTGCAACATTTTATTCCTACTA  
ATAATGCACAGGGGTTTCGGATTCTCCACATCCTTGCCAACACTTGATATTCCTGGGTTTTTGATAATGGCCATC  
CTAATAGGTGTGGGGACATGAGGTTTTCAATATGCTTGTGGGACATCGAAAACCCTGGGACCCTTCAGATGGGA  
CTAACTGGGGAAAGTGGACAAGTTACAAACAAAGAACTCAAAGGAAAGTCATTGGCACTGATCTCTAAGATGCT  
ATCACATGTGATTGGTGGTTGATTTTATTAACAAATTATAAGCAAAGTACTACAAAGGTGGCTTTAAAAGAAAAT  
AAAGCAATTCACAGAACTACTTTTTCATGTAGCTTGATGTGTGCTCCATG

5 Un barrido de 586 pb de secuencias que se extienden aguas arriba y aguas abajo del sitio de restricción TaqI  
mostró dos picos de propensión a la curvatura a aproximadamente 300 y 500 pb con un valle de puntuaciones  
bajas. El plegamiento de ADN y la estimación de energía libre de la secuencia de 44 pb que rodea la región  
TaqI produjeron la energía libre de estructura termodinámica más baja,  $\Delta G$ , de 0,51 kcal/mol a 60 °C, 4,5 mM  
de  $Mg^{++}$  condición iónica.

Referencias:

- 10 1. Goodsell, D. S. y Dickerson, R. E. (1994) "Bending and curvature calculations in B-DNA" Nuc. Acids Res.  
22, 5497-5503
- 15 2. Zuker (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nuc. Acids Res. 31,  
3406-3415
3. SantaLucia, Jr (1998) A unified view of polymer, dumbbell and oligonucleotide DNA nearest-neighbor  
thermodynamics. Proc. Natl. Acad. Sci. USA 95, 1460-1465

Tabla 1a.

Pocillo	Tamaño pb	Posición FIP1L1 ENSEMBL BLAT	Posición PDGFRA ENSEMBL BLAT	Distancia (pb) entre fragmentos	Cebador 1	Cebador 2
A1	501	4: <u>53379986-53380174</u> *	4: 54294960-54295285 RF	914786	TTCCACGTTGGCTACCAACAG	CAGCTGGGAGGTTTCTTTT
A1	647	4: <u>53379986-53380310</u>	4: 54285685-54285944 RF	905367	TGGGAGTGGGTGGAGTGAGA	ACGGACTGACATCTTTAGCTTCC
B1 y B7	462	4: <u>53379986-53380176</u>	4: 54278879-54279091 RR	898703	GCAGCTGGGAGGTTTCTTT	GCCATGTGGCTTGGGCATAC
C1	499	4: <u>53379986-53380194</u>	4: 54281532-54281784 RF	901338	AGTACTTCTCTCTCCCTCCCA	GCAGCTGGGAGGTTTCTTTT
D7 y E7	609	4: 53361094-53361416	4: <u>54281778-54281990</u> FR	920362	CCCTCCACACAAACCTGCTA	ACATGGAGCACACATACAAGCTAC
D7	586	4: 53369589-53369898	4: <u>54281778-54282010</u> FR	911880	AAATGATGAGGCGACGGGTGAA	CATGGAGCACACATACAAGCTAC

\* Secuencia de cebador

Tabla 1b. Seis interacciones recurrentes de FIP1L1 y PDGFRA y tamaño de par de bases (PB)

N.º	Interacción	Pocillo N.º	Tamaño en pares de bases
1	FIP1L1 PDGFRA_4_GBA461GB_GBA466A6_FH15A743_FH161379_RF	A1	501
2	FIP1L1 and PDGFRA_4_GBA461GB_GBA466A6_FH149644_FH15A108_RF A1	A1	647
3	FIP1L1 and PDGFRA_4_GBA461GB_GBA466A6_FH145047_FH147946_RR	B1 y B7	B1: 462
4	FIP1L1 and PDGFRA_4_GBA461GB_GBA466A6_FH145047_FH147946_RF	C1	499
5	FIP1L1 y PDGFRA_4_GBAA681A_GBAA7581_FH147946_FH149AA7_FR	D7 y E7	D7: 609
6	FIP1L1 y PDGFRA_4_GBA31931_GBA38063_FH147946_FH149AA7_FR	D7	586

Nº.	Secuencia
1	TTCAAGTGGCTACACACAGCATGTGAGGGCTTGGGGGACAGAAATCTGGCCATACCTGTGGACGGCCAAATTTGAATGGCAAGGGCTTCTGTTTGTCTCTGGGGGGCCACAGCTTA GGCTAGTCTGTGTGCAGGAGTGTGAATATTTGGCTCTTCTCTCCCTCTCAGCTGGCAGACCTCTGAAGAGAGTGGCCATTGGAGACGGGTTCAGACAGTTCACCTTCATCATAG AGAGAGGACGAGACCATCTGAAGACATCTGAAGAATGAAGACAGGGTTAGTGGAGGATATGGGGAAATGAAGCCCTCATACACTTCTGGTGGGATTTGAATAATGGTGGCAGCT GCTTGAAGAACACCTGGCCATTTCTCAAAATGGTGAACATGGAGTTACCATATGTGAGTCTCCACCTTAGAGGGAATGAAGAACATATTCAGACAAAAAACTTGTGACGTGAAT ATTCATAGCAGCATTTCTTAATAGAAAGTGG
2	TGGGAGTGGGTGGAGTGAAGAACCTGGGAGAGAGGECAGCCCTTTATATCCAGGGACACAGCTCAAGTGGCCACCATGGATCAGCCAGTCTTGGCAGGGGTGATGCTATTCCAG CTACAGATGGGTGTGATCTGTGATCATTTCTCTTCTTCATGTCAGTGTGTGTCACCTGGATCTGGCCAAAGCTCTCTGGCAACAGGAAATTTGTGAAGATCTGTGGA CTTGGCTGGCCGACAGACATATGCAATGTTCGAAGAATGAAGACAGGTGTAGTGAAGATATGCGGAAATGAAGACCTCATACACTCTGGTGGGATTTGTAAATGGTGTG CAGCTGCTTGAAGAACAGCTGGCCATTTCTCAATGGTGAAGACATGGAGTTACCATGTGAGTCTCCACCTTAGAGGATGAAGACATATTCAGACAAAAAACTTGTGACG TGAATATTCATAGCAGCATTTCTTAATAGAAAGTGGGAAAGAAACCTGGCAGCTGCATCACTGATGATGAATGAATTAATGTGTATATCCATACAGGGGGAATAT TATTTGGCAAGGACAAATAAATGAAGTACTGGTAGATTTACACACGGGATGAATCTTACAAATGTGGAAAGCTAAGATGTCAGTCCGT
3	GGATGCGAGGCTTTCTTTCTTCTTATTAAGAAATAATGCTGTCTATGAAATTTTGTCTGAAATATGTTTCTTCTTCTTAAGGGTGAAGSAGACTACAT GAGTACTCCATGTTTCCACCATTTGAGAAATAAGGCCAGGCTGTCTTTCACAGACAGCTCACCATTGTTTACATATCCACAGAGTGTGATGAGGGTTTATTTCCCAATATCTCAC TAACACCTCTTTATCTTTTCGAGGCTCTTGCTGAACCTGGACCTATAAATGACGTCAATGATAGTGAATCTCTACTGCAAGATATCAAGTGGCTATAAAGAACTCTGTAGG TAAGAAATCTGTGAAGATCAGAAAGTACATGAAATTCACCTTCATATAATAATTACTGGTGGACACAAATGGGTGGCTAAATGATGAGTGGGTAGAGGAATGTGTATGTGCCAAGC CACATGGC
4	AGTACTTCTTCCCTCCATATTTGTTAAAAATAGTTTACATTTGCTTCCAGGGCTGGGCTGGTGGAGTGGGACAGAGATGTGAGAGGAACCTGAGTCAATGCTCAGGGCCCAA GGCTCTGTGGCAGGACGACCTGCTTCTGGGCTCTCGTGACTATCTGAAGAAATAATGCTGAATGGCTAGAGCTACTCTTCTTCACTGCTGAACATTTTCAAAAAGAAATTCAGAA CTTCTGGATTAATATGGCTTCTCTCGAAAGATGAAGACAGGGTGTAGTGAAGGATATGGGGAATGAAGCCCTCATACACTCTGGTGGGATTTGTAATAATGGTGGCAGCTGC TTTGAAGAACGCTGGCCATTTCTCAAAATGGTGAACATGGAGTTACCATGTGAGTCTCCACCTTAGAGGGAATGAAGACATATTCAGACAAAAAACTTGTGACGTGAATAT TCATAGCAGCATTTCTTAATAGAAAGTGGGAAAGAAACCTTCGCAGC
5	CTCTCCACACAAACCTGCTAGTGAATCTGGTATCACTTCACTTCAACCATTCACCTGGGAAAGACCATCTACTAGCAGAGAGGATTTTACAAATGTAAAGAAAGTAAAGGAC TTTACATAACAAATACAAAGATCACTCTCTTGCATCAATTAACAAAGATGTGGAACCTTTCACATTTACAGAAATTAATGTCAAATGTGACACACAATATCATCATCT GGCTAACTAACTTATGCTTTTGTAGTGAAGAACACTTTTGAAGCTCAATATCTAATTTTAAATGATGAAGACACTTACTCAATTTGTTGAGATCAATAATAGGTAA TGTAAAGTGGGATTTTATTTTACTTATTTATTCGAAACCTGGGACCTTCCAGATGGGACTAACTGGGGAAGCTGGCAAGTTCACAAAGAAAGAACTCAAAAGGAAA GTATGTGGCACTGATCTTAAGATGCTATACATGATGTGGTGGTGTGATTTATTAACAAATTAAGCAAAAGTACTACAAGGTGGCTTTAAAAAGAAAATAAGCAATTC ACAGAACTACTTTTCTATGCTGTATGTGTGCTCCATGT
6	AAATGATGGGCACGGGTGATAGATAGTGTGGAGGTGACATACAGATTTCAATTAATCCCTAAAGGGTGTATGGTTTTCTTCTCTCAAAATCCTTACTTTCAGTCTTTT GTACATGGACCAAGATGGAATGTGGAATCATATGATAAATCTATTTTAATCTTCTGATGGACCTCTGTTTTGTTTTTATACAGAGGCTGCAACATTTTATATCTCTA CTAATAATGGACAGGGGTTCGGATTTCTCCACATCTTGGCAACATGATATTTCTGGGTTTTTGTATTAATGGCCATCTAATAGGGTGTGGGACATGAGGTTTTCAATAT GCTTGTGGGACATGAAAACCTGGGACCTTCCAGATGGGACTACTGAGGAAAGTGGGAAAGTGGGACAGTTCACAAAGAAACTCAAGGAAAGTCAATGGCAGTCTGATCTTA AGATGCTATCAATGTGATGGTGGTGTGATTTATTAACAAATTAAGCAAAAGTACTACAAGGTGGCTTTAAAAAGAAAATAAGCAATTCATCTTTTCTATG TAGCTGTATGTGTGCTCCATG

Tabla 2

Indicación	Estípe celular o muestra clínica	IDH1 SNP estado rs11554137	Estado de IDH1 R132H	FIP1L1-PDGFR $\alpha$ 3C Copias de ADM en 20 ng (+/-)	MMMP1 3C copias en 20 ng
Leucemia Mielode Aguda (AML)	AML-193	Tipo salvaje	Tipo salvaje	+ (0,04)	0,37
Oligoastrocitoma anaplásico	BT-412	Tipo salvaje	R132H**	No en el ensayo	
Glioblastoma multiforme	DBTRG-05MG	Tipo salvaje	Tipo salvaje	+ (0,05)	0,35
AML (Eosinofílica)	EO4-1	Tipo salvaje	Tipo salvaje	+ (0,03)	0
Leucemia Mielode Aguda AML	GDM-1	Tipo salvaje	Tipo salvaje	+ (1,99)	7,3
AML	HL-60	rs11554137*	Tipo salvaje	+ (0,18)	1,24
Astrocitoma de Glioblastoma	KG-4	rs11554137	Tipo salvaje	---	4,2
Astrocitoma de Glioblastoma	U-373	Tipo salvaje	Tipo salvaje	+ (2,01)	1,24
	U87	Tipo salvaje	Tipo salvaje	No en el ensayo	-
Oligodendroglioma	47436	Tipo salvaje	R132H	---	1,89
Oligodendroglioma	47679	Tipo salvaje	Tipo salvaje	+ (1,18)	2
Oligodendroglioma	81600	Tipo salvaje	R132H	---	0,96
Oligodendroglioma	174694	Tipo salvaje	R132H	+ (0,04)	0,81
Oligodendroglioma	184263	Tipo salvaje	Tipo salvaje	+ (0,06)	5,19
Oligoastrocitoma	30736	Tipo salvaje	Tipo salvaje	---	1,42
Biopsia Adípica	91013A	Tipo salvaje	Tipo salvaje	---	6,05
Sangre Normal	17153	No analizado	No analizado	---	4,53
Sangre Normal	17335	Tipo salvaje	Tipo salvaje	---	2,18
Sangre Normal	17353	Tipo salvaje	Tipo salvaje	---	3,09

\*Previamente positivo, Dodemont y col., 2011 \*\*Previamente positivo, Flavahan y col., 2016

Tabla 3.

Fluoróforo	Abs [nm]	Em [nm]	Agente de Inactivación 3'
FAM	485	520	TAM, BHO1, DAB, Edap
TET	521	536	TAM, BHO1
JOE	520	548	TAM, BHO1, BHO2
Amarillo Yakima	530	549	BHO1, Edap
HEX	535	556	TAM, BHO1, BHO2, Edap, BBQ650
Cyanina3	552	570	BHO1, BHO2, BBQ650
ATTO 550	554	576	TAM, BHO2
TAMRA	544	576	BHO2
ROX	575	602	TAM, BHO2, BBQ650
Rodo Texas	583	603	BHO2, BBQ650
Cianina3,5	588	604	BHO2
LC 610	590	610	BHO2
LC 640	625	640	BHO2, BBQ650
ATTO 647N	644	669	BHO2, BBQ650
Cianina5	649	670	BHO2, BBQ650
Cianina5,5	675	694	BHO2, BBQ650
ATTO 680	680	700	BBQ650

Tabla 4. Detalles de las Sondas

Nombre de sonda	45047FPRC	OBOFH010_03_FAM	
Secuencia de Sonda	5'-Texas, RED-TCAGGACGGACCTCGAAAGATATAACA-BHQ2-3'	5'-FAM-ATTATTTCGAAACCCCTGGGAGCCO-BHQ1-3	
Cebador directo	5'-GCAGCTCGAGGGTTTCTTT-3'	5'-CCTCTCCACACACAACCTTGCTA-3'	
Cebador inverso	5'-GCCATGTGGCTTGGGCATAG-3'	5'-ACATGGAGCACACATACAAGCTAG-3'	
Identificador de Interacción	B7 462 pb, PCR 8.	D6/E7, 609 pb.	
Localización Genómica	4:54278879-54279091	4:53379986-53380176	4:54281778-54281990
Orientación	Directo	Inverso	Directo
Genes Solapantes	RP11-231C18.3, PDGFRA	FIP1L1, RP11-231C18.3	RP11-231C18.3, PDGFRA
Inicio-Fin de la consulta	388-400	1-191	320-532
Longitud	213	191	213
Puntuación	413	368	412
E-val	1,6E-115	5,6E-102	3,9E-115
% ID	100,00	99,48	100,00

## Ejemplo 4. Trabajo Adicional

La detección de la activación ectópica de tirosina quinasa se pierde actualmente por la metodología de diagnóstico convencional (por ejemplo, secuenciación) en múltiples indicaciones.

Se están desarrollando inhibidores de tirosina quinasa, por ejemplo dasatinib, imatinib e lbrutinib/Imbruvica (Janssen-Cilag Ltd) para numerosos tumores (crónicos) y trastornos constitucionales. Los ejemplos de indicaciones no malignas potenciales para los inhibidores de la Tirosina Quinasa de Bruton (BTK) incluyen artritis.

Los médicos requieren evidencia de una lesión molecular en una tirosina quinasa objetivo antes de la prescripción de un inhibidor de tirosina quinasa. Las intervenciones diagnósticas terapéuticas que pueden salvar la vida usando inhibidores de tirosina quinasa (TKI) se pierden usando diagnósticos convencionales.

Aspectos de la invención relacionados con el tratamiento de fármacos:

(a) Imatinib se dirige terapéuticamente a conformaciones 3C:

Los inventores demuestran que imatinib mimético de ATP conocido por actuar sobre tirosina quinasas activadas directamente también actúa directamente sobre conformaciones de cromosomas 3C en genes sensibles a imatinib tales como cKIT.

(b) GSK se dirige terapéuticamente a conformaciones 3C:

Se proporcionan datos para el inhibidor de IDH1 GSK864 que demuestran que este fármaco inhibe directamente las conformaciones cromosómicas en múltiples estirpes celulares. La estratificación de pacientes oncológicos para un tratamiento eficaz se puede llevar a cabo identificando mutaciones activadoras en genes de tirosina quinasa y cambios adquiridos en otras rutas. Los procedimientos diagnósticos que contribuyen a estratificar a los pacientes de esta manera son de gran beneficio potencial. Los cambios conformacionales en las regiones potenciadoras de FIP1L1 y PDGFRA regulan por aumento la expresión de PDGFRA en pacientes con glioma con mutaciones IDH1 R132H. La sobreexpresión ectópica de PDGFRA se puede tratar con imatinib.

Los siguientes datos se refieren a marcadores 3C y su uso en un ensayo de qPCR funcional. ABL1 y cKIT muestran la diferencia binaria entre la muestra no maligna y la estirpe celular.

RC = Complemento inverso. La secuencia de cebador inversa verdadera es adyacente a la secuencia RC.

Genes para los que se identificaron interacciones 3C e indicaciones en las que están implicados:

1. ABL1, gen de tirosina quinasa citoplasmática localizado en 9q34: Indicaciones de enfermedad: Leucemia Mieloide Aguda (AML), Leucemia Linfocítica Crónica (CLL), Cáncer de Mama. Leucemia Mieloide Crónica (LMC) y crisis blástica por LMC o Tumor del Estroma Gastrointestinal (GIST). Síndromes Mielodisplásicos (SMD) o Trastornos Mieloproliferativos (Neoplasias Mieloproliferativas, MPN).

2. Tirosina Quinasa de Bruton, gen de tirosina quinasa citoplasmática localizado en Xq22.1: Indicaciones de Enfermedad:

Agammaglobulinemia ligada al cromosoma X (XLA), neoplasia maligna de células B. NHL, Linfoma (todas las clasificaciones), Mieloma. Leucemia Linfocítica Crónica Pacientes con Neoplasias de Células B Maduras Recurrentes. Enfermedad Crónica De Injerto Contra Huésped. Leucemia De Células Pilosas Recidivante. Tumores Sólidos Recidivantes o Refractarios. Linfoma Refractario de Células de Manto. Carcinoma Esofagogástrico Amplificado y Artritis Reumatoide.

3. cKIT, gen receptor de tirosina quinasa localizado en indicaciones de enfermedad 4q12:

Mastocitosis y Leucemia de Mastocitos y Tumores de Próstata. Melanoma Avanzado, Melanoma Maligno Mutado, 2ª Línea. Tumores del Estroma Gastrointestinal (GIST) No Aptos para Resección o Metastásicos, que expresan Mesotelioma Maligno. Cáncer de Ovario Refractario a Platino Avanzado. Cáncer de Ovario Epitelial, de Trompa de Falopio o de Peritoneo Primario Recidivante, o Tumor Ovárico de Bajo Potencial Maligno. Pacientes con Carcinoma Seroso Papilar Uterino, Cáncer de Pulmón Microcítico en Etapa Extensa, Carcinoma Nasofaríngeo Localmente Avanzado, Linfoma -No Hodgkin en Recaída/Refractario, Cáncer De Células Renales Metastásico, Tumor Neuroendocrino Gastroenteropancreático Metastásico, Carcinoma Ovárico de Células Claras Recurrente, Carcinoma Esofágico de Células Escamosas, Cáncer Uterino Persistente y Tumores de Mama.

4. FGFR1, gen receptor de tirosina quinasa situado en 8p11: Indicaciones de enfermedad: cáncer de Tiroides,

adenoma Oral, Leucemia Crónica, Leucemia Aguda, Tumores de Próstata y Cáncer de Mama. Cáncer de Cabeza y Cuello Recurrente, Mutado o Amplificado, de FGFR1-3 Translocado, Glioma Maligno Recurrente que expresa la Fusión de Gen FGFR-TACC. Pacientes con Cáncer de Pulmón No Microcítico Avanzado.

- |    |   |
|----|---|
| 5  | 5. FIP1L1, factor de especificidad de poliadenilación fusionado frecuentemente a PDGFRA (4q12). El gen que no es de tirosina quinasa: iHES, Leucemia Crónica.   |
| 10 | 6. Gen de tirosina quinasa de receptor FLT3 localizado en 13q11-12.2: Indicaciones de enfermedad: Leucemia Crónica, Leucemia Mieloide Aguda y Cáncer de Mama, iHES, Leucemia Crónica. Carcinoma de Pulmón No Microcítico (NSCLC). Cáncer Colorrectal, Melanoma Metastásico o Cáncer de Células Renales, Glioma, Linfoma de Hodgkin Refractario. Cáncer De Ovario Tratado Previamente. Carcinoma Nasofaríngeo Localmente Avanzado. Leucemia Linfocítica Crónica Refractaria/Linfoma Linfocítico Pequeño (CLL/SLL). Tumores de Estroma Gastrointestinal (GIST) en los que ha fallado sumitinib. |
| 15 | 7. PDGFRA es un gen del receptor tirosina quinasa localizado en 4q12: Indicaciones de la enfermedad: Glioma. Leucemia Crónica. Leucemia Mieloide Aguda y Cáncer de Mama. iHES, Leucemia Crónica.  |

Tabla 5. Interacción ABL1 3C N°. 1.

Tabla 5: Interacción ABLT3C N.º 1:

Nombre de la interacción: ORF1\_9\_130700600\_130703071\_130719461\_130720884\_FR

Secuencias seleccionadas para sondas. Insertar una región de 25 pb a cada lado de la unión Taq I.

5'-CTTATAGCCTGTCTCTCTTGCTGATCGAGGTTGCAACGAGCTGAGATTGC-3'

Secuencia de cebador directo: 5'-TTGGAGTTCCAGGGTGATACTGTC-3'

Secuencia de cebador inverso: 5'-ACCAACGCTATTCTTGGCTTCCA-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
9:130719465-130719839	ABL1	Directo	96	470	375	715	2,00E-206	99,20
9:130702984-130703074	EXOSC2	Directo	5	95	91	172	6,00E-43	96,70
X:136026579-136026620	SLC9A6	Inverso	93	137	45	65	1,10E-10	93,33

Un ensayo basado en sangre que discrimina material normal y maligno (CML, CML-BC, CLL, Cáncer de Mama) usando muestras de sangre periférica. El control positivo es la estirpe celular K562 de crisis blástica de CML.

Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

>ORF1 9 130700606 130703071 130719463 130720684 FR

GCAGAGCTCCACCTCCTGGGTTCAAGTGATTCTCTCGCTCAGCCTCTCTGAGTAGCTGGGATTACAGGTGTGCCCAACCAT  
GCTCTGCTAATTATTCTATTFTTTAGTAGAGATGGGGTTTACCATTGTTGGTCAGGGCTGGTCTTTGAATTCCTGACCTTGTG  
ATCCACCGCGCTCGGGCTCCCAAAGTGCTGGGATTACAGACATGGAGACACTGCACTCTGCCCATTTCTTTTAAATCTCC

TTGGAATTAGCTGTTTGGTGGAT TGGTTGCGGTTTCTTGAGAGTCATAAATGATTTATTTGTGAATTCCT  
CTGGGCGCGTTTGGCTGTTGTTGTTGATTTCCCTTCCCTCTCTGTGTCTCCCTATAGCCCTGTCTCTCTGCTGATC  
GAGGTTGCAACGAGCTGAGATTGCACGCCACTGCACCTCCAGCCCTGGCGACGGAGTGAGACTCCATCTCANNNNNNNN  
NNNNNNNNNNNNNNNNNNNNNNNNNNNTGATGACAGCCCAACCTGTCACTTTTTTCAGATCCCTTTTATGAAAGAAATTGCT  
TAASTTGTGTCTGAAGACAA ATTCCTTTTTCTAAGCGGAATACCATGTTATTTTGTGTT  
ACAGATTGTTTGGCATCTTTCATAGCCCTGATCTTTCTAGAGTTGGTTTAATATCCATGATGTTAGATTGAAAAATCTGAA  
TTGAGAAATGACGGTGTGGAGCAGACATGGATC GTTGTGTTGATGGGTGAATCTG

Los datos de chip de laboratorio muestran la amplificación de una interacción 3C en el material normal (paciente 375) pero no en la estirpe celular K562 de crisis blástica de leucemia mieloide crónica (temperatura de recocido de cebador de 68 °C). Este patrón también se encontró a temperaturas de recocido de 67,5 °C y 66.4 °C. Tamaño esperado de producto: 515 pb.

Tabla 6. Interacción ABL 3C N°. 2.

Nombre de la interacción:	ORF1_9	130700600	130703071	130719461	130720884	RF
---------------------------	--------	-----------	-----------	-----------	-----------	----



Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.

5'-GGGCCAAGTGTGACTCTCAGGTTTTCGACCTGCCTCAGCCTCCCAAAGT-3'

Secuencia de cebador directo: 5'-ATGCAGGAGGTAAAAATGGAGGGT-3'

Secuencia de cebador inverso: 5'-CAAAGAGACAGCTCCGTCAGAGA-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
9:130720517-130720886	ABL1	Inverso	256	625	370	724	4,40E-209	100,00
9:130700586-130700859	EXOSC2	Inverso	1	273	275	497	7,90E-141	97,82

Un ensayo basado en sangre que discrimina material normal y maligno (CML, CML-BC, CLL, Cáncer de Mama) usando muestras de sangre periférica.

Ciclo qPCR (WI103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C

>ORF1 9 130700600 130703071 130719461 130720884 RF

[illegible]

AAGATATTCAFGACAACCAAGGGGAGAGGTAAGGTCACAGTGATGTGCTCTGGGTCAAAGATTGTTGAGCGCTGG  
 AFGGAGGTGTGGSGTCAAGGGSAGASSCTCAGAGGACAGGACAGTGTGTGSCCAAGGCTGCT  
 GTTGTTCCTTCATCACCTGGCCAGGCTGAGGTCCAGGCAGTGT  
 CTGAAATATGSAAGTAASTCGGGCTCTTGAIGTTCTGTGTGCTGACTGAGACTACAAGSCTATTTTGAATCCOCAT

5

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (paciente 375) pero no en el material maligno (derivado de la estirpe celular K562) a 67,5 °C. Tamaño esperado de producto: 676 pb

10    Tabla 7. Interacciones de Tirosina Quinasa 3C de Bruton N°. 1:

Nombre de la interacción: ORF2_X_101355470_101357976_101407183_101408846_FR										
Secuencias seleccionadas para sondas.										
5'-TTGAGGACAAGGACCTCGAGATACTGCCGAGAAATCC-3'										
Secuencia de cebador directo. 5'-ATGTTGTCTAACCTGTATGTTCT-3'										
Secuencia de cebador inverso. 5'-CTTCCTCTGCCGGGTGCTAA-3'										
Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.										
Localización Genómica	Gen(es) Solapante(s)	Orientación	Nombre de consulta	Inicio de consulta	Fin de consulta	Ori consulta	Longitud	Puntuación	E-val	% ID
X:101407180-101407241(3 pb)	RPL36A - HNRNP H2, GLA	Inverso	Consulta_1	5	66	Directo	62	123	3,00 E-25	100,00
X: (2 pb)	BTK	Inverso	Consulta_1	61	102	Directo	42	83,8	3,00 E-13	100,00

101357937-101357978										
5:154143558-154143584	MFAP3	Directo	Consulta_1	65	91	Directo	27	46,1	0,057	96,30
12:41941194-41941215		Inverso	Consulta_1	69	90	Directo	22	44,1	0,22	100,00
<p>Un ensayo basado en sangre que discrimina material normal y maligno (leucemia de células B) usando muestras de sangre periférica. Estirpe celular de control positivo RAMOS (línea de linfoma de Burkitt).</p> <p>Ciclo qPCR (WI103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C</p>										

Tumores malignos de células B = cualquier tumor maligno que se piense que implica Tirosina Quinasa de Bruton.

>ORF2\_X\_101355470\_101357976\_101407183\_101408846\_FR  
 ATGAGTCTGAAACAGAGAGAGAGGTCATGCTGTTGGTGTGGTGTAGGAGGTGGGATGCC  
 TCACACATCCTCACTTAAAGCCTCACACTTCCGGTGTGTATCTTTCTAGTACATTTTGAA

5

TCCCAGAAGACCTCCATGAACCCACATATTTCCAGAGGTTTCTCCTCTCACAAAATACTT  
 TTCAGGGACAAAAAGAGAAAATACTAATTGAAATACACTGCCTAACTGAATATATTTCCC  
 ACCTCCCAGCCCAGGCTTAGCTCACATCCTACCTTCTTGATGAAGCTGTTCCCCCATATC  
 CATGCCCTCACTAATGTCTCAGTTTTTAGAAATTTACAGTGCTTATAAACTTTCTTCTG  
 ATGTTGTCTAACTGTATGTTCTTTGAGGACAAAGGACCTGCTGAACTGCCGAGAAATCC  
 ACCACCCCAAAGGAGAAAAAATGGCAGCAATCTTAGTTCCCAATTCCTCCTACCCACCC  
 TAGATGTTCTAACGGTTAGTACCCCGCAGAGGAAAGCAGAAGTTCACAAGAAGGGTCTG

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en la estirpe celular solo a esta temperatura de recocido. Tamaño esperado de producto: 155 pb.

10

Tabla 8 Interacción de tirosina quinasa cKIT N°. 1:

Nombre de la interacción: ORF3_4_54711719_54715167_54736749_54740853_FR										
Secuencia seleccionada para las sondas (objeto de ensayo):										
TCACCCAGAATAAGGCTTCTCGATTCTAAGTTCTACAAGA										
Secuencia de cebador directo (s). cKIT_003 (5'-CAGTGCGGTGGGAGTACTGT-3')										
Secuencia de cebador inverso (s). cKIT_004 (5'-TTCAGCAGGTGCGTGTTTCTAG-3')										
Secuencia de cebador directo (s). cKIT_001 (5'-AGTGCGGTGGGAGTACTGTG-3') *										
Secuencia de cebador inverso (s). cKIT_002 (5'-ACATTTCTAGCAGGTGCGTG-3') *										
Ambos conjuntos de cebadores (001/002 y 003/004) amplifican la interacción. La posición de la secuencia de PCR BLAT para la secuencia de cebador es para cKIT003 y cKIT004 (también se han confirmado datos de PCR y secuenciación de una sola etapa para los cebadores adicionales).										
Localización Genómica	Gen(es) Solapante(s)	Orientación	Nombre de consulta	Inicio de consulta	Fin de consulta	Orientación Consulta	Longitud	Puntuación	E-val	% ID
4:54714861-54715168	ESTUCHE	Directo		1	307	Directo	308	585	2,80E-167	99,68
4:54736751-54736823	ESTUCHE	Directo		308	381	Directo	74	134	1,80E-31	98,65

Un ensayo basado en sangre que discrimina material normal y maligno (EOL1) usando muestras de sangre periférica. La estirpe celular de control positivo EOL1 y (leucemia eosinofílica) y BT142 (Glioma).

Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

\*Tener en cuenta que estas combinaciones de cebadores generan productos de etapa individual y no se han utilizado para PCR anidada.

[illegible]

5

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en material normal pero no en material maligno (derivado de la estirpe celular EOL1). El tamaño de producto esperado es de 429 pb.

Tabla 9. Interacción FGFR1 3C N°. 1:

Nombre de la interacción: ORF4_8_38427499_38430449_38480545_38486034_FF								
Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.								
5'-AAACCCCAGCAGCCCCTGCCCAAGTCGA GGGAGCAGCTCCCCACCCAGCCC-3'								
Secuencia de cebador directo: 5'-AGAAAGAGGCAAAGTTAGGAAGCA-3'								
Secuencia de cebador inverso: 5'-AAAACAGCAACCTATTTCAGAGAGC-3'								
Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.								
Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
8:38430233-38430452	FGFR1	Directo	9	228	220	422	2,90E-118	97,27
8:38485833-38486036		Inverso	224	427	204	403	2,00E-112	100,00
Un ensayo basado en sangre que discrimina material normal y maligno (cáncer de Tiroides, Adenoma Oral, Leucemia Crónica, Leucemia Aguda, Tumores de Próstata y Cáncer de Mama) usando muestras de sangre periférica. Estirpe celular de control positivo CCL30 (carcinoma de células escamosas).								
Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C								

10

ES 2 971 286 T3

>ORF4 8 38427499 38430449 38480545 38486034 FF(5)

[illegible]

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (paciente 375) pero no en el material maligno (estirpe celular CCL30) a 68 °C. Tamaño esperado de producto: 460 pb.

5

Tabla 10. Interacción FIP1L1 3C N°. 1:

Nombre de la interacción:

>FIP1L1PDGFRA\_4\_54231931\_54236063\_55143356\_55145047\_FR

Secuencia seleccionada para las sondas (objeto de ensayo).

5'-TATGCTTGTGGGACATCGACAAAAGCAATTATGC-3'

Secuencia de cebador directo. 5'-TGGAAGTGACACATACAGATTTCAA-3'

Secuencia de cebador inverso. 5'-CTTTCTGCACCGTAGCCAGC-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
4:53369634-53369898		Directo	10	274	265	517	3,00 E-143	99,62
4:54277188-54277335	AC058822.1, PDGFRA	Directo	271	418	148	293	7,00 E-76	100,00
10:61991285-61991356	ARID5B	Directo	123	194	72	87,7	8,00 E-14	90,28
10:10353728-103537327	NEURL1	Directo	35	82	48	87,7	8,00 E-14	97,92

Un ensayo basado en sangre que discrimina material normal y maligno (leucemia mieloide aguda) usando muestras de sangre periférica. Estirpe celular de control positivo EOL1 (estirpe de leucemia aguda eosinofílica).

Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

```
>FIPIL1PDGFRA_4_54231931_54236063_55143356_55145047_FR
CAGCTAAAGACTACTTTTCCAAGCCTCTTGCAGCTAGCTATGGACAAATGATGAGGCACG
GGTGAATAAGATAGTTGGAAGTGAACATACAGATTTCAATAATTCCCTAAAGGTGTAT
GGTTTTCTTTCCCTACAAATCCTACTTTTCAGTTCTTTTGTACATGGACCCAGAGTGGAAAT
TGCTGGATCATATGATAATTCTATTTTAACTTCTTGATGGACCTCTGTTTTTGTTTTGT
TTTTTACAGAGGCTGCAACATTTTATATTCCTACTAATAATGCACAGSGGTTCGGATTTTC
TCCACATCCTTGCCAACACTTGATATTTCTGGGTTTTTTGATAATGGCCATCCTAATAG
GTGTGGGGACATGAGGTTTTTCAATATGCTTGTGGGACA[REDACTED]CAAAAGCAATTATGCTAA
TTTCTTCCCTGTGGGCTCAATTCTTTTTTGACACGATGACTTGGAGGAGTCATTATGA
TTACTCCAAACAGGAAAGACACTCGCCAGCTGTCCGCCCCGAGAGATTTTCTGCTGCTG
CAGAAATCTGAGGAGGCGTCTGGAGTTTTTGGGTGTTAATGATTCTGCTGCCACAGGT
CGGGTCTTGGGGTCTGGAGCTTTCTGGAAGTGGTCAAGGAACAGCCTATGGATTAAAGC
CGGTCCCAACCTGTGATGAAAGTTGCAGTGAAGATGCTAAACGTAAGTGCTCCTTCCTG
GGGATTTTTTGAGCACGGGGATTTTTTGAGCATGGGGATATTAAGGGAATTTCTCAAAAT
CATGCAGCTAGTAAATAAGA
```

Tabla 11. Interacción FLT3 3C N°. 1:

Nombre de la interacción prevista: ORF5_13_28025286_28026644_28116701_28118054_RF								
Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.								
5'-CCTCGGCGACAGAGTGAGACTCTGTCGATCTCATTCTTTGGTTTCTGAAC-3'								
Secuencia de cebador directo: 5'-GGTATGTGCTGTAGTCCCAACT-3'								
Secuencia de cebador inverso: 5'-GAAGGATGGTGAAACGCTTACGG-3'								
Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.								
Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
13:28025285-28025376	FLT3	Directo	81	172	92	174	1,50E-43	96,74
13:28117987-28118056		Directo	15	84	70	139	5,30E-33	100,00
Un ensayo basado en sangre que discrimina material normal y maligno (Leucemia Mieloide Agua) usando muestras de sangre periférica.								
Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.								

>ORF5\_13\_28025286\_28026644\_28116701\_28118054\_RF

```
TAATTGCATCACTGCCTCCAGCC[REDACTED]TAAAGAAACAAAGCAGGTTCTTTCTCCT
CTCAGAAGCCCTCTCTCTTACTAGAGAGAGAGCTGTTCTCCTCTCTCTTTCTTCTATTAAACCTCCGCTCCT
AAAAACACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCCTAAAAAACTAGCCTGGGCAACATGG
TAAACCCCTGTCTCTACAAAAAACACAAATAAACTAGCCAGGCGCTGGT[REDACTED]TAT
TCAGGGAAGCTGAGGTGGGAGGTTTGCTTGAGACCGGGAAGTCCAGGCTGCAGTGAGCCGAGATGGCGCCACT
GCCTCTAGCCTCGGCGACAGAGTGAGACTCTGTGATCTCATTCTTGGTTTCTGAACTCAGCCCACTTTCC
CTCACATAATACAGATCTTCCCGAGAGTC[REDACTED]TTTTTCCAGTTTGCCTTTGT
AAAGCTCATTTCTTTTACTGTGAAATTAATAAATAGCAATAATACAAATCTCTTTTAAGGTATATAAAAATT
CTTCCC[REDACTED]TAGGAATTAAGTAGATCAGGATTCATTTTATTTCTAATATACGAGTTTAACCCA
ATTTACAATTTACATATCTTTAGGAAATGGCTTCATGCTTTCAGACCAAAATATTTCCAAAAGGAATAATTTT
AGTGAATTTAATATTCAG[REDACTED]TAAGTACTGACCACTGGGATACTGGTCACT
```

5

Datos de Lab Chip - fila B. Esto muestra la amplificación de la interacción 3C en el material normal (Paciente 376) pero no en el material maligno (derivado del paciente 277 con AML) a 67,5 °C. Tamaño esperado de

Tabla 12. Interacción EL T2 3C N° 2:

Nombre de la interacción prevista: ORF5\_13\_28025286\_28026644\_28116701\_28118054\_FF

Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.

5'-ACTTTACTGTGTCCCCATCACGTGTCGA CAGAGTCTCACTCTGTGCGCCGA-3'

Secuencia de cebador directo: 5'-CAAGCATCTCTCAGGATCCCCAT-3'

Secuencia de cebador inverso: 5'-ACTATTGAGGGAAGCTGAGGTGG-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
13:2802655 4-28026646	FLT3	Inverso	66	158	93	180	1,60E-45	98,92
13:2811799 0-28118056		Directo	3	69	67	133	3,30E-31	100,00
11:3560407 3-35604129		Directo	7	63	57	110	1,70E-24	98,25

Un ensayo basado en sangre que discrimina material normal y maligno (Leucemia Mieloide Agua) usando muestras de sangre periférica.

Ciclo qPCR (W103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

NN  
NNNAGAATGCTCCAAGGCTC  
AAC[REDACTED]GGAAATCCATTTCAGAGAAAGGTGCTACAAAGCCATGATGCTCAAGTAA  
TTAGCAAGAAAGTTTCAGAGGATGAGGAGGAGATCCAGCAAACCTCTACAGCCCTCATTTATGCCCGCTAAGACA  
ATAAATCCCTGCAGGACCA[REDACTED]GATGCTGCTGCTGCTGATGGATGGGAGGATAAACAACAGAGCAGAGACAAG  
GTACTGCTACTCTTTACTGTGTUCCCATCACGTGTGGACAGAGTGCTCACTCTGTGCGCGAGGCTAGAGTGCAGT  
GGCGCCATCTCGGCTCACTGCAGCCTGGACTTCCC GGCTCTCAAGCAAACCTC[REDACTED]  
[REDACTED]TGGGACTACAGGCACATACCACCAGGCCTGGCTAGTTTTATTGCTGTTTTTTGTAGAGACAGGGTTTTTACCA  
TGTGCCCCAGGCTAGTTTTTTTAGGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTGTTTTTAGG  
AGCGGAGGTTTTAATAGAAGAAAGAGAGAGAGAGAACAGCTCTCTCTCTAGTAAGAGAGAGCGGGCTTCTGAGAGG  
AAAGAACCTGCTTTGTTTCCTTA[REDACTED]GGCTGGAGTGCAGTGATGCAATTAA

5

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (Paciente

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (Paciente 376) pero no en el material maligno (derivado de paciente 277 con AML) a 68,0 °C. Tamaño esperado de producto: 204 pb.

Nombre de la interacción prevista: ORF5\_13\_28025286\_28026644\_28116701\_28118054\_RR

Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.
5'-GAGTTCAGAAACCAAAGAATGAGATCGA GTGATTGCTAGGTGATGGACCT-3'
Secuencia de cebador directo: 5'-GAAGGATGGTGAAACGCTTACGG-3'
Secuencia de cebador inverso: 5'-TGCACCCCAACAGGAGTTTCTA-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
13:28116700-28116784		Directo	50	134	85	164	1,40 E-40	98,82
13:28025280-28025338	FLT3	Inverso	1	57	59	93	2,40 E-19	96,61

Un ensayo basado en sangre que discrimina material normal y maligno (Leucemia Mieloide Agua) usando muestras de sangre periférica.

Ciclo qPCR (WI103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

[illegible]

- 5 Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (Paciente 376) pero no en el material maligno (derivado de paciente 277 con AML) a 68,0 °C. Tamaño esperado de producto: 173 pb.

Tabla 14. Interacción FLT3 3C N°. 4:

Nombre de la interacción: ORF5\_13\_28009061\_28014204\_28025286\_28026644\_FR

Secuencia seleccionada para las sondas (objeto de ensayo). Insertar una región de 25 pb a cada lado de la unión Taq I.

5'-GCTGGAGGATTGCTTGAGCTTGGGAGGCTCGA TCTCATTCTTTGGTTTCTGAAC-3'

Secuencia de cebador directo: 5'-GCTTCTGGAGGCCAGGGATTAT-3'

Secuencia de cebador inverso: 5'-AGCCCCTTTGATTCGTTCTGTTC-3'

Posición de la secuencia de PCR BLAT. Tener en cuenta que las dos primeras filas son las verdaderas homologías.

Localización Genómica	Gen(es) Solapante(s)	Orientación	Inicio de consulta	Fin de consulta	Longitud	Puntuación	E-val	% ID
13:28025286-28025660	FLT3	Directo	325	699	375	723	6,80E-209	100,00
13:28013880-28014203	FLT3	Directo	1	324	324	627	6,70E-180	99,07
X:47532217-47532398		Inverso	132	313	182	267	1,10E-71	84,07

Un ensayo basado en sangre que discrimina material normal y maligno (Leucemia Mieloide Agua) usando muestras de sangre periférica.

Ciclo qPCR (WI103). Producto secuenciado de ensayo 3C utilizado para la curva convencional. La sonda de hidrólisis se dirige a la unión del fragmento 3C.

[illegible]

Los datos de chip de laboratorio muestran la amplificación de la interacción 3C en el material normal (paciente 376) pero no en el material maligno (derivado de paciente 277 con AML) a 68 °C. Tamaño esperado del producto: 742 pb.

### Datos adicionales

La confirmación de los datos de gradiente de temperatura se proporciona a continuación. Las interacciones seleccionadas para cada gen en orden alfabético se presentan a continuación, se incluye la copia promedio para cada interacción en 20 ng (por reacción = un pocillo por placa de 96 pocillos).

1. Ensayo de qPCR basado en ABL1 ABL1-515 SYBR de material derivado de la estirpe celular K562 y sangre normal (10 ng/reacción).

K562 muestra un número promedio de copias reducido (12,36333, DT +/- 15,67294) en 20 ng, en comparación con las copias 33,24667, DT +/- 35,04153 por 20 ng en sangre normal, ( $p = 0,143245$  Ensayo T).

## 2. Tirosina Quinasa de Bruton

ORF2 X 101355470 101357976 101407183 101408846 FR

25 Las estirpes celulares (BT142, EOL1 y HL60) muestran un número promedio de copias reducido (20,99338 DT +/- 43,413166) en 20 ng, en comparación con las copias, 43,83277 DT +/- 72,808926 por 20 ng en sangre normal, (0,41426 Ensayo T).

3. Experimento cKIT: diferencia altamente significativa.

ORF3\_4\_54711719\_54715167\_54736749\_54740853\_FR

Las estirpes celulares malignas (BT142, EOL1 y HL60) muestran un número promedio de copias reducido (0,205135DT  $\pm$  0,453526) en 20 ng, en comparación con las copias 2,639932 (DT  $\pm$  1,742175) por 20 ng en sangre normal, ( $p = 0,001294$  Ensayo T). Esta es una diferencia altamente significativa entre las estirpes celulares y las muestras no malignas.

4. FGFR1,

ORF4 8 38427499 38430449 38480545 38486034 F

40 La estirpe celular maligna CCL30 muestra un número promedio de copias incrementado de 2,05 (DT +/- 0,843267) en 20 ng, en comparación con las copias 1,242222, (DT +/- 1,715319) por 20 ng en sangre normal, (p = 0,314743 Ensayo T).

## 5. Experimento FIP1L1

FIP1L1 PDGFRA 4 54231931 54236063 55143356 55145047 FR

Las estirpes celulares malignas (BT142, EOL1 y HL60) muestran un número promedio de copias reducido de 13,07325 (DT +/- 19,46042) en 20 ng, en comparación con las copias 32,61117 (DT +/- 38,58252) por 20 ng en sangre normal. (p= 0,147896 Ensayo T).

## 6. Experimento FLT3:

ORF5 13 28009061 28014204 28025286 28026644 F: Diferencia significativa



El material de paciente maligno muestra un número promedio de copias reducido 0,427619 (DT +/- 0,667599) en 20 ng, comparado con las copias 6,587778, DT +/- 7,672794 por 20 ng en sangre normal, ( $p = 0,043238$  Ensayo T).

5

#### Sensibilidad a fármacos e Imatinib

Imatinib (Glivec) muestra actividad frente a múltiples tirosina quinasas incluyendo ABL1, cKIT y PDGFRA. El éxito de los ensayos clínicos en leucemia mieloide crónica proporcionó el impulso para la identificación de una estrategia farmacogenética.

10

A continuación se describen ejemplos de datos novedosos con respecto a la acción de imatinib en el siguiente gen seleccionado sensible a imatinib. Se añadieron 20 ng de plantilla a cada pocillo para las muestras tratadas con imatinib.

15

Interacción de cKIT denominada anteriormente (ORF3\_4\_54711719\_54715167\_54736749\_54740853\_FR) e imatinib:

20

cKIT y datos a modo de ejemplo. Las células BT142 se han tratado con  $\mu\text{M}$  de imatinib. Se introdujeron 20 ng de plantilla en cada pocillo. Téngase en cuenta que la interacción cKIT 3C está completamente ausente en las estirpes celulares tratadas con imatinib (véase más adelante).

25

El producto en el pocillo C1 se ha confirmado como la interacción cKIT mencionada anteriormente (ORF3\_4\_54711719\_54715167\_54736749\_54740853\_FR). Estos datos ilustran la acción directa de imatinib sobre una conformación cromosómica cKIT 3C.

La interacción MMP1 en la estirpe celular BT142 se muestra a continuación. Esto confirma la calidad de las muestras. Este es un ensayo basado en sonda de hidrólisis.

Estirpe celular	Cq	Número de copia MMP1
BT142 ctrl	37,29354288	12,16144613
BT142 ctrl	35,68147305	36,19699893
BT142 ctrl	36,33204621	23,30823893
BT142 AMLtinib (IM) 1 $\mu\text{M}$	35,6402365	37,2211193
BT142 1 $\mu\text{M}$	36,14604001	26,43413752
BT142 1 $\mu\text{M}$	35,88943195	31,44606744

30

La acción de GSK864 sobre las estirpes celulares tratadas

Cuantificación de MMP-1 3C con controles sin fijación y muestras tratadas con formaldehído. Las estirpes celulares a modo de ejemplo usadas en la presente solicitud son:

35

(a) DBTRG-05 MG (glioblastoma).

(b) HL-60 (leucemia mieloide aguda).

40

Estas estirpes celulares se han tratado con el inhibidor de IDH1 GSK864 durante 24 h y 48 h. Se observa una diferencia binaria en la estirpe celular DBTRG, y diferencias cuantitativas en HL60 que muestran una acción directa del compuesto en las confirmaciones del cromosoma MMP1 (véase la Figura 8).

Estirpe celular	Cq	Número de copia en 20 ng
DBTRG 24 horas Ctrl	39,91101	2,898388
DBTRG 24 horas Ctrl	38,34145	8,990504
DBTRG 24 horas Ctrl	NaN	NaN
DBTRG 24GSK 1 $\mu\text{M}$	NaN	NaN
DBTRG 24GSK 1 $\mu\text{M}$	NaN	NaN
DBTRG 24GSK 1 $\mu\text{M}$	NaN	NaN
NaN = 0 copias		

45

Las estirpes celulares HL60 se han tratado con el inhibidor de IDH1 GSK864 durante 24 h y 48 h. Se observa

una diferencia cuantitativa binaria en el número de copias de MMP1 3C en la estirpe celular HL60 que muestra una acción directa del compuesto en las confirmaciones del cromosoma MMP1 (véase la Figura 9).

Estirpe celular	Cq	Número de copia en 20 ng
HL-60 Ctrl 48 horas	35,94292	50,70651
HL-60 Ctrl 48 horas	35,20558	86,30136
HL-60 Ctrl 48 horas	36,69911	29,39049
HL-60 GSK864 48 horas	40,32569	2,149149
HL-60 GSK864 48 horas	37,49586	16,54413
HL-60 GSK864 48 horas	37,10877	21,87217
20 ng = un pocillo de reacción.		

## 5 Compendio

Los presentes ensayos de etapa individual que usan detección SYBR green muestran discriminación entre sangre normal y material maligno (véanse las estirpes celulares descritas en las tablas anteriores). El número de copias de qPCR 3C se puede reducir significativamente en las estirpes celulares malignas en comparación con el material normal (por ejemplo, cKIT y FLT3) o puede existir ganancia, por ejemplo, de la estirpe celular de FGFR1 CCL30.

Se cree que el fármaco imatinib (Glivec: Novartis PLC) actúa contra la proteína BCR-ABL. Los inventores demuestran una acción novedosa directa de imatinib frente a interacciones de conformación 3C identificadas en el gen cKIT. Los datos muestran una acción terapéutica directa de imatinib sobre las conformaciones 3C.

Además, GSK864 (un inhibidor mutante de IDH1) muestra una acción directa sobre el número de copias de MMP1 3C a 1  $\mu$ m en las estirpes celulares DBTRG y HL60.

## REIVINDICACIONES

- 1.- Un método para determinar el diagnóstico o pronóstico de un cáncer que comprende detectar la presencia o ausencia de al menos una interacción cromosómica epigenética en un cromosoma que se refiere a uno o más de los siguientes genes: FIP1L1, PDGFRA, Flt3, ABL1, FGFR1, cKIT y tirosina quinasa de Bruton, en donde dicha interacción cromosómica está seleccionada entre una interacción definida por:

(i) cualquiera de las sondas siguientes:

- (a) 5'-CTTATAGCCTGTCTCTCTTGCTGATCGAGGTTGCAACGAGCTGAGATTGC-3'  
 (b) 5'-GGGCCAAGTGTGACTCTCAGGTTTTGACCTGCCTCAGCCTCCCAAAGT-3'  
 (c) 5'-TTGAGGACAAGGACCTCGAGATACTGCCGAGAAATCC-3'  
 (d) 5'-TCACCCAGAATAAGGCTTCTCGATTCTAAGTTCTACAAGA-3'  
 (e) 5'-AAACCCCGAGCAGCCCTGCCCAAGTCGAGGGAGCAGCTCCCCACCCAGCCC-3'  
 (f) 5'-TATGCTTGTGGGACATCGACAAAAGCAATTATGC-3'  
 (g) 5'-CCTCGGCGACAGAGTGAGACTCTGTCGATCTCATTCTTTGGTTTCTGAAC-3'  
 (h) 5'-ACTTTACTGTGTCCCCATCAGTGTCGACAGAGTCTCACTCTGTGCGCCGA-3'  
 (i) 5'-GAGTTCAGAAACCAAGAATGAGATCGAGTGATTGCTAGGTGATGGACCT-3'  
 (j) 5'-GCTGGAGGATTGCTTGAGCTTGGGAGGTCGATCTCATTCTTTGGTTTCTGAAC-3'  
 (k) 5'-TCAGCAAGGACCTCGAAAAGATAAAACA-3'  
 (l) 5'-ATTTATTCGAAAACCTGGGACCC-3'

o

- (ii) uno cualquiera de los siguientes pares de cebadores que amplifican el producto ligado preparado en el proceso de detección de la interacción cromosómica:

- (m) TTCCACGTGGCCTACCACAG and CAGCTGCGAGGTTTTCTTT  
 (n) TGGGAGTGGGTGGAGTGAGA and ACGGACTGACATCTTTAGCTTCC  
 (o) GCAGCTGCGAGGTTTTCTTT and GCCATGTGGCTTGGGCATAC  
 (p) AGTACTTCCTCTCCCTCCCA and GCAGCTGCGAGGTTTTCTTT  
 (q) CCTCTCCACACAAACCTGCTA and ACATGGAGCACACATACAAGCTAC  
 (r) AAATGATGAGGCACGGGTGAA and CATGGAGCACACATACAAGCTAC,

- en donde la detección de la interacción cromosómica definida por cualquiera de las sondas (a) a (e) y (g) a (j) es para determinar el diagnóstico de cáncer, y

en donde la detección de las interacciones cromosómicas definidas por cualquiera de las sondas (f), (k) y (l) o cualquiera de los pares de cebadores definidos por (m) a (r) es para determinar el pronóstico de cáncer.

- 2.- Un método según la reivindicación 1, en donde dicha detección de la presencia o ausencia de la interacción cromosómica se lleva a cabo mediante un método que comprende las etapas de:

(i) reticular in vitro de regiones cromosómicas que se han unido en una interacción cromosómica para formar ácido nucleico reticulado;

(ii) someter dicho ácido reticulado a escisión por digestión de restricción con una enzima; y

(iii) unir de dichos extremos de ácido nucleico escindidos reticulados para formar el producto ligado; y

(iv) la detección del producto ligado para determinar así la presencia o ausencia de la interacción cromosómica.

3.- Un método según la reivindicación 2, en donde dicho producto ligado comprende secuencias de ácido nucleico de longitud:

- de 10 a 1000 bases nucleotídicas,

- de 10 a 800 bases nucleotídicas,

- 10 a 500 bases nucleotídicas,
- 10 a 100 bases nucleotídicas,
- 5 - de 10 a 400 bases nucleotídicas,
- 10 a 500 bases nucleotídicas,
- 200 a 600 bases nucleotídicas,
- 10 - 200 a 800 bases nucleotídicas, o
- de 200 a 1000 bases nucleotídicas,
- 15 4.- Un método según la reivindicación 2 o 3, en donde dicha detección comprende la detección específica del producto ligado por PCR cuantitativa (qPCR) que utiliza cebadores capaces de amplificar el producto ligado y una sonda que se une al sitio de enlace durante la reacción de PCR, en donde dicha sonda comprende una secuencia que es complementaria a la secuencia de cada una de las regiones cromosómicas que se han unido en la interacción cromosómica.
- 20 5.- Un método según la reivindicación 4, en donde dicha sonda comprende:  
un oligonucleótido que se une específicamente a dicho producto ligado, y/o
- 25 un fluoróforo unido covalentemente al extremo 5' del oligonucleótido, y/o  
un agente de inactivación unido covalentemente al extremo 3' del oligonucleótido, y  
opcionalmente dicho fluoróforo está seleccionado entre HEX, Texas Red y FAM.
- 30 6.- Un método según la reivindicación 4 o 5, en donde dicha sonda comprende una secuencia de ácido nucleico de 10 a 40 bases nucleotídicas de longitud, preferentemente una longitud de 20 a 30 bases nucleotídicas.
- 35 7.- Un método según cualquiera de las reivindicaciones anteriores, en donde el cáncer es uno cualquiera de los siguientes:  
(a) síndrome hipereosinofílico idiopático (iHES)
- 40 (b) leucemias eosinofílicas crónicas y agudas.
- (c) leucemia mieloide aguda (AML)
- (d) carcinoma de pulmón no microcítico positivo para PDGFRA.
- 45 (e) glioblastoma, incluidos glioblastoma multiforme y astrocitoma.
- (f) cáncer de próstata.
- (g) cáncer de ovario avanzado.
- 50 (h) tumor del estroma gastrointestinal (GIST).
- 8.- Un método según la reivindicación 4, en donde hay detección cuantitativa del producto ligado mediante el uso de una sonda que se puede detectar tras la activación durante la reacción de PCR,
- 55 en donde dicho método comprende poner en contacto la secuencia ligada con la sonda durante una reacción de PCR, y detectar el grado de activación de la sonda, y  
en donde dicha sonda se une al sitio de unión.
- 60 9.- Un método según la reivindicación 8, en donde la sonda es como se define en la reivindicación 5 o 6.
- 10.- Un método según la reivindicación 8 o 9, en donde la longitud del producto ligado es como se define en la reivindicación 3.
- 65 11.- Un método según la reivindicación 8, 9 o 10, en donde el sitio de unión de dicho producto ligado comprende

la secuencia de reconocimiento de enzima de restricción de la enzima de restricción usada para cortar el ácido nucleico reticulado, y preferentemente dicha enzima de restricción es Taq1.

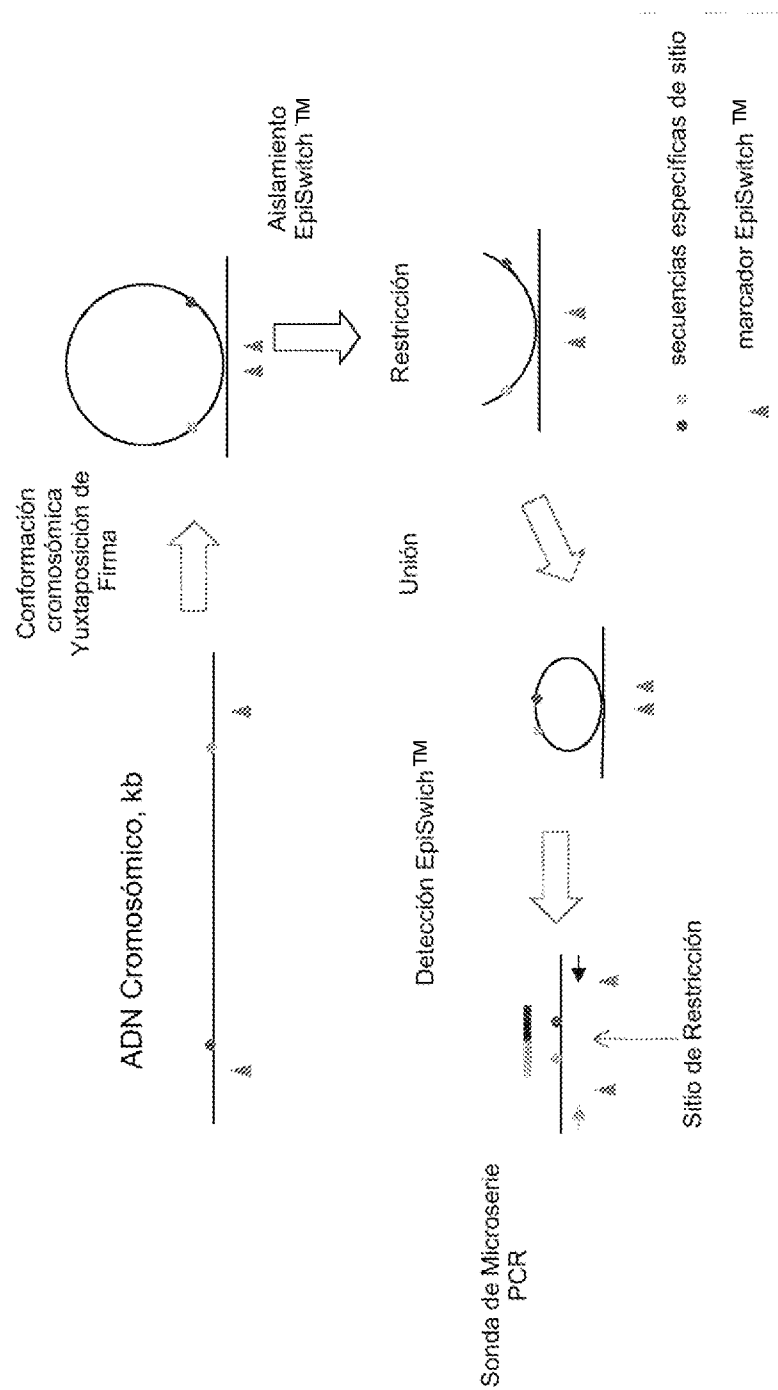


Figura 1

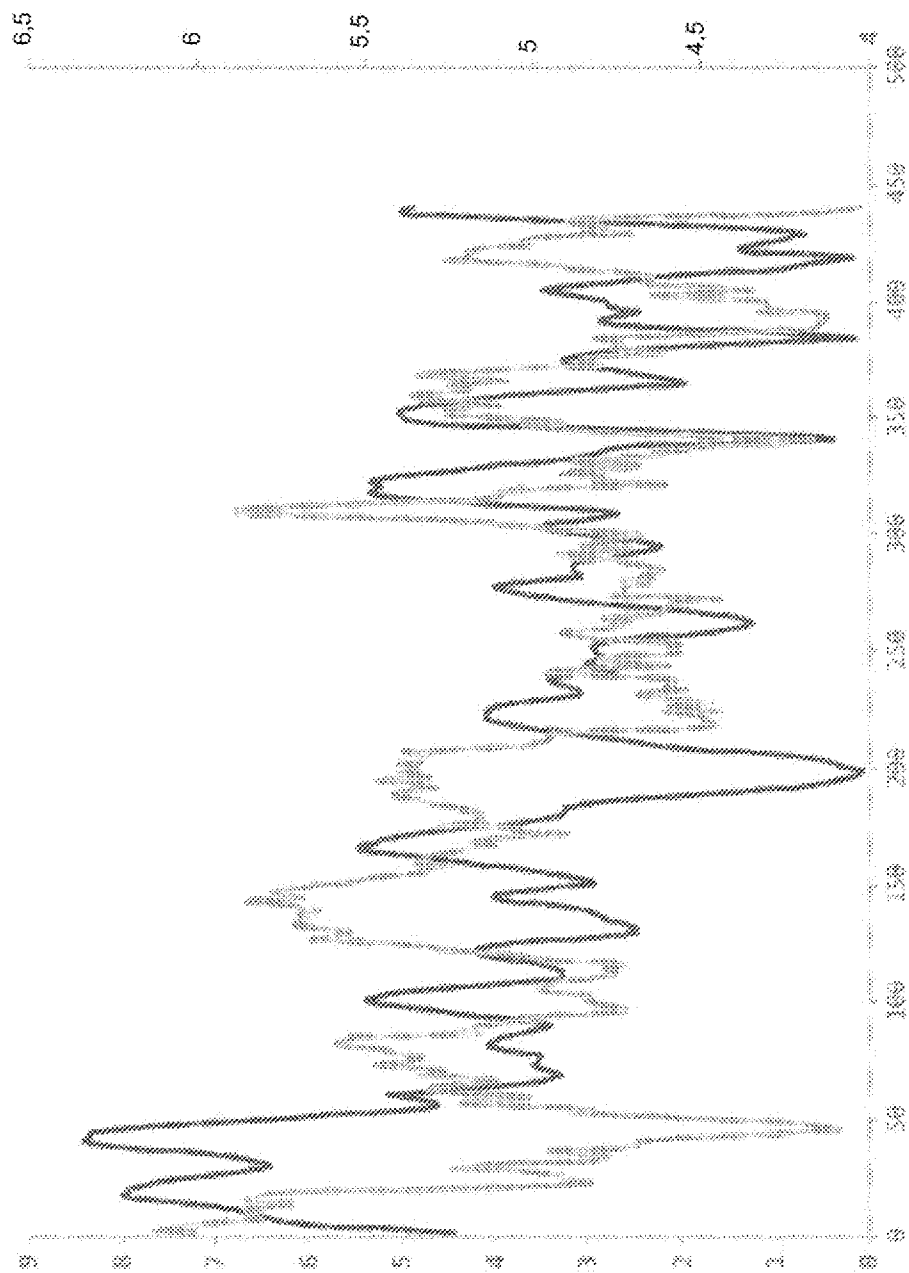


Figura 2

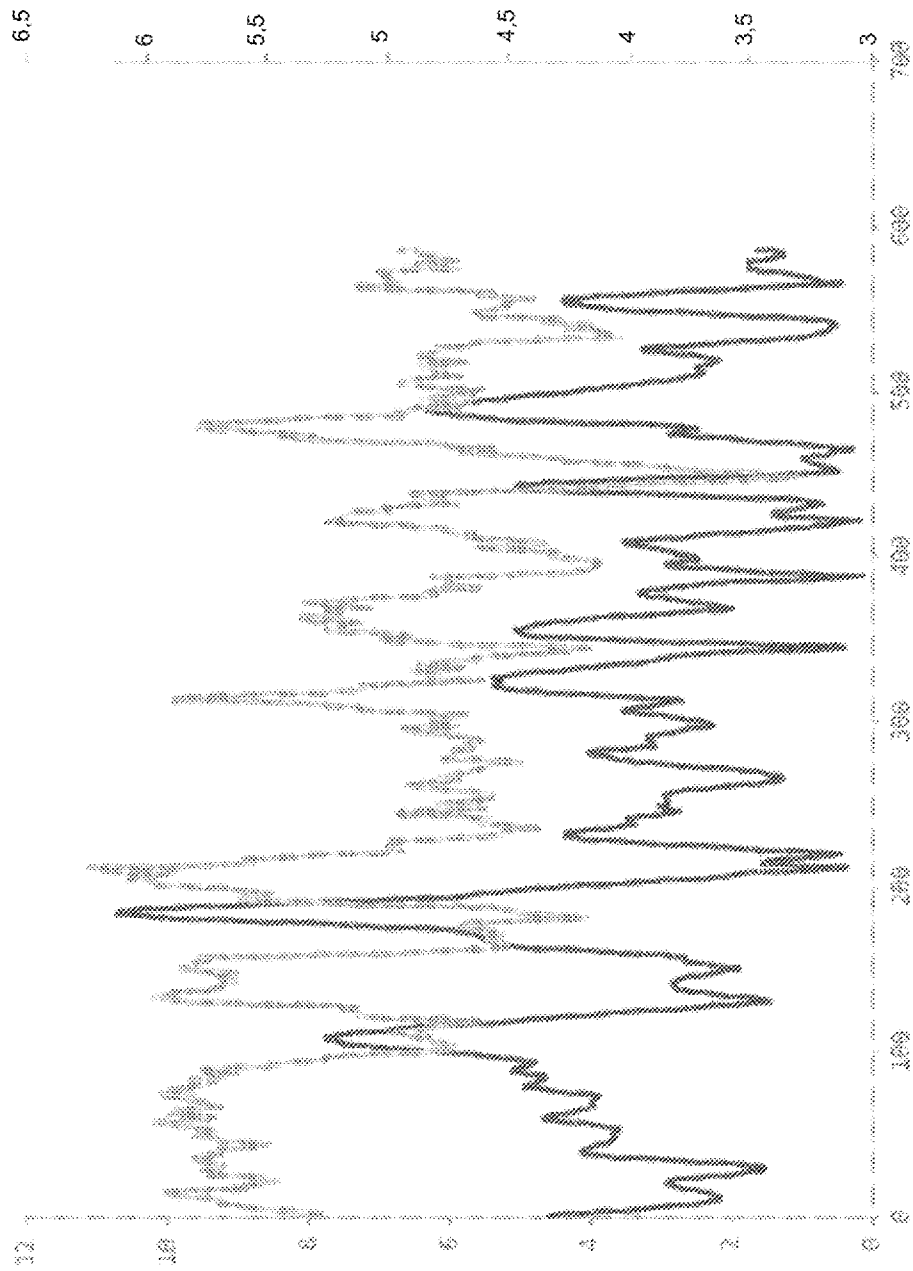


Figura 3



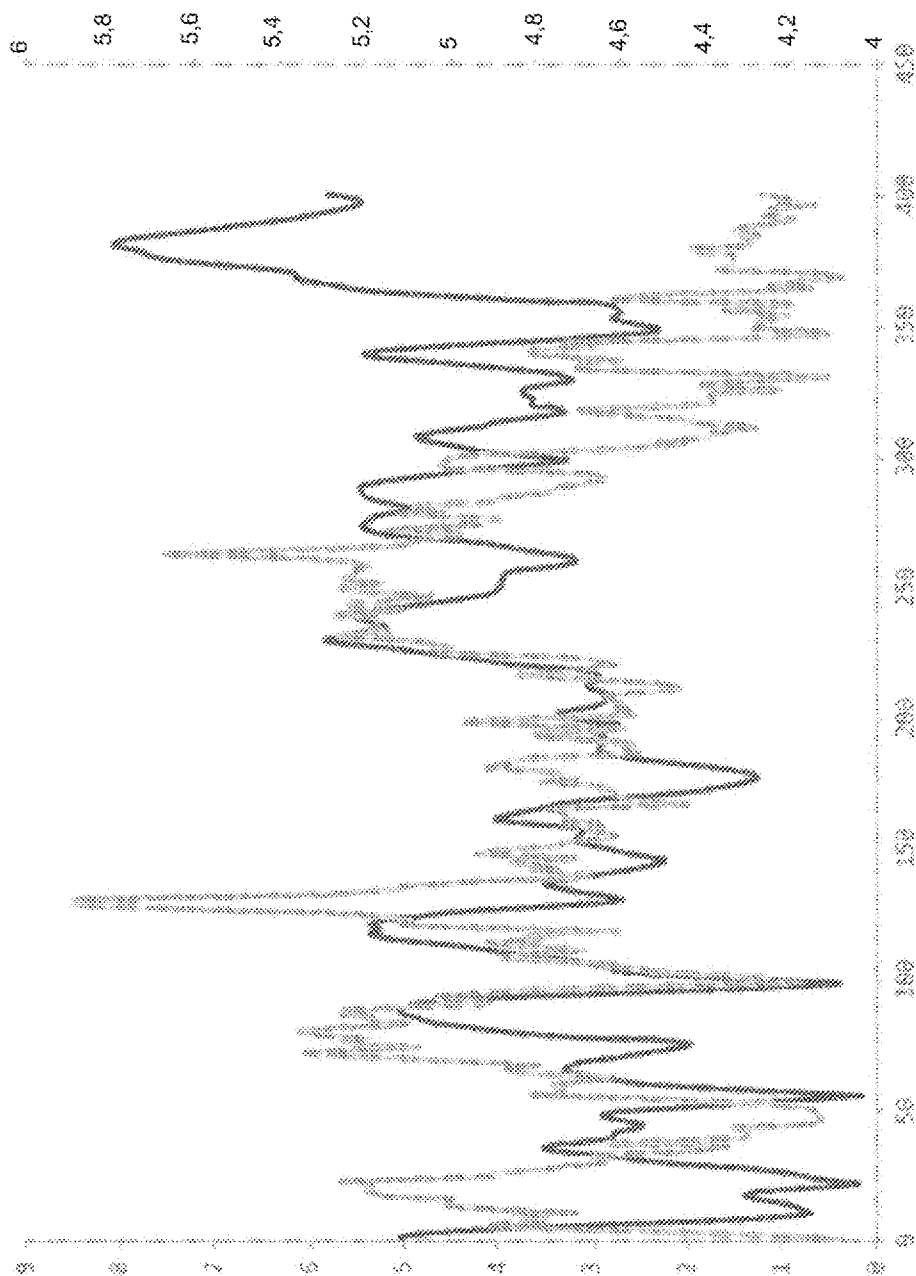


Figura 4

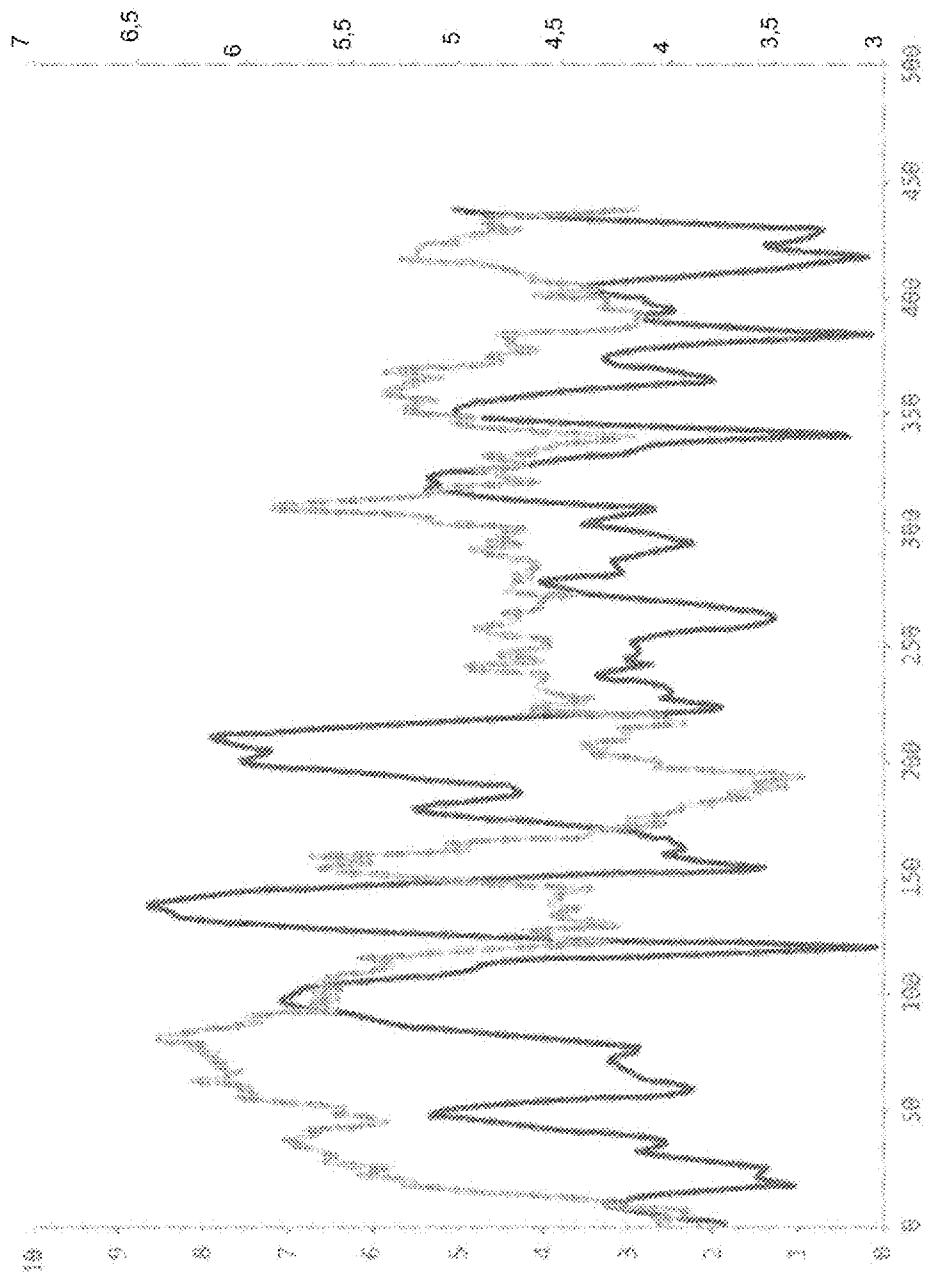


Figura 5

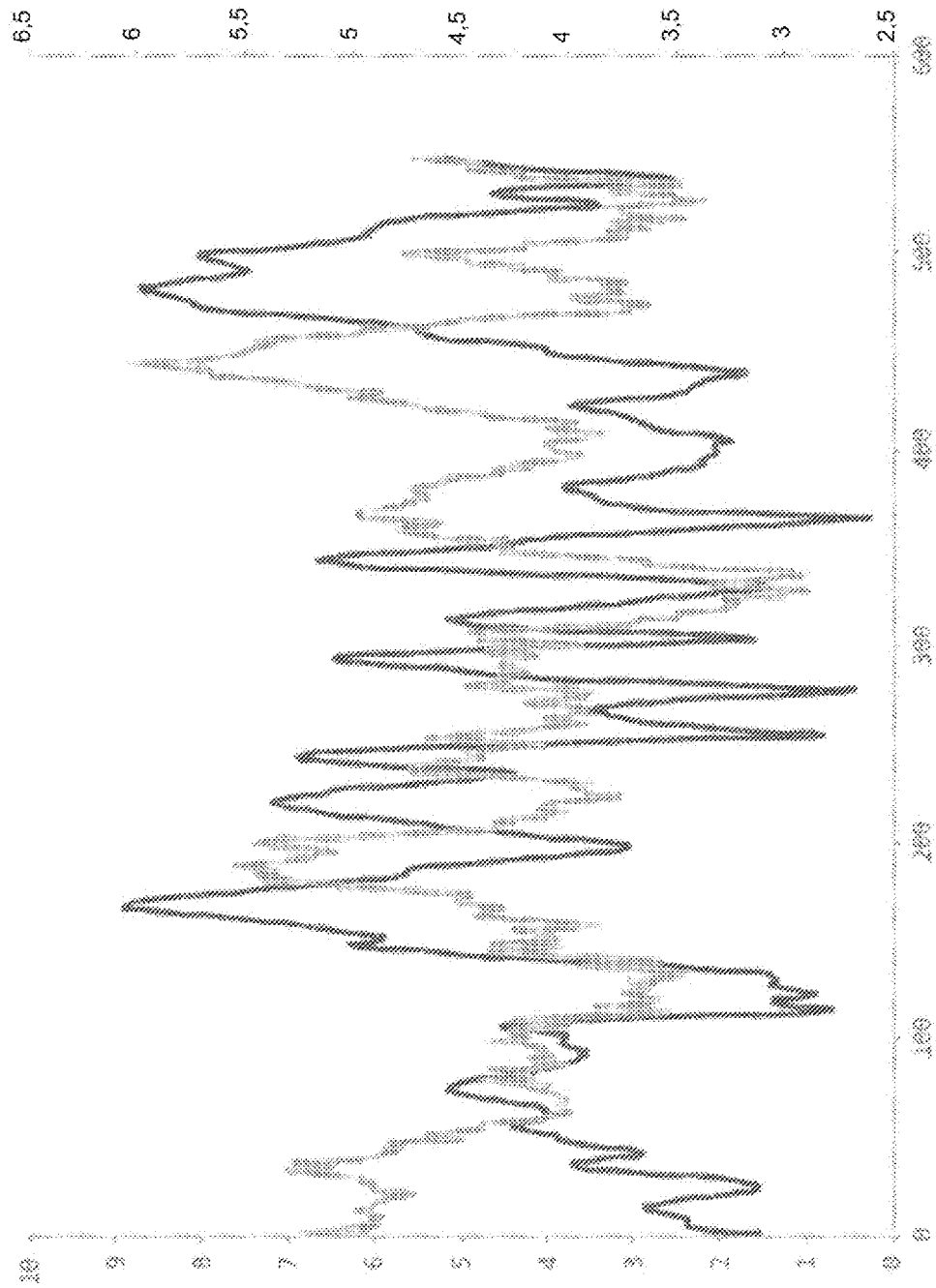


Figura 6

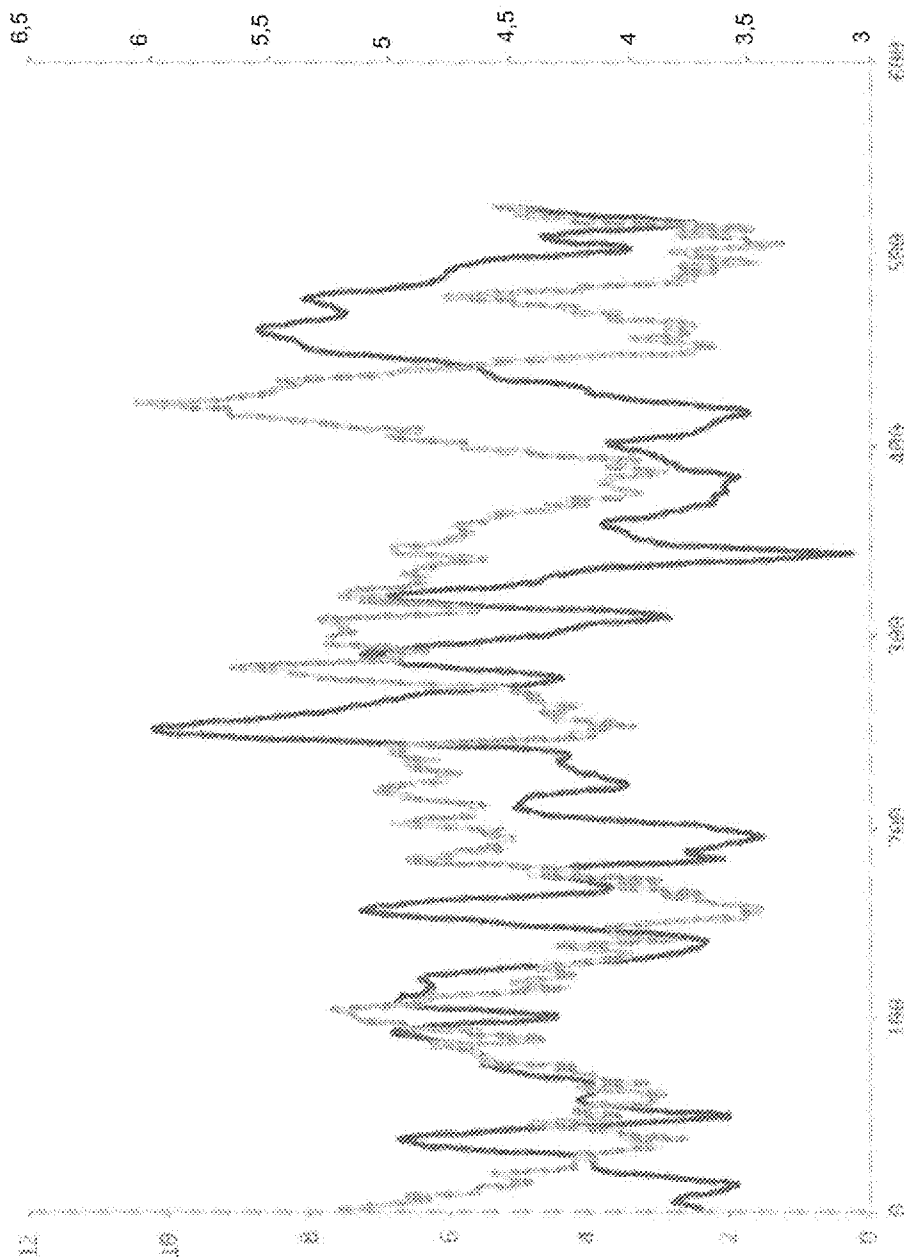


Figura 7

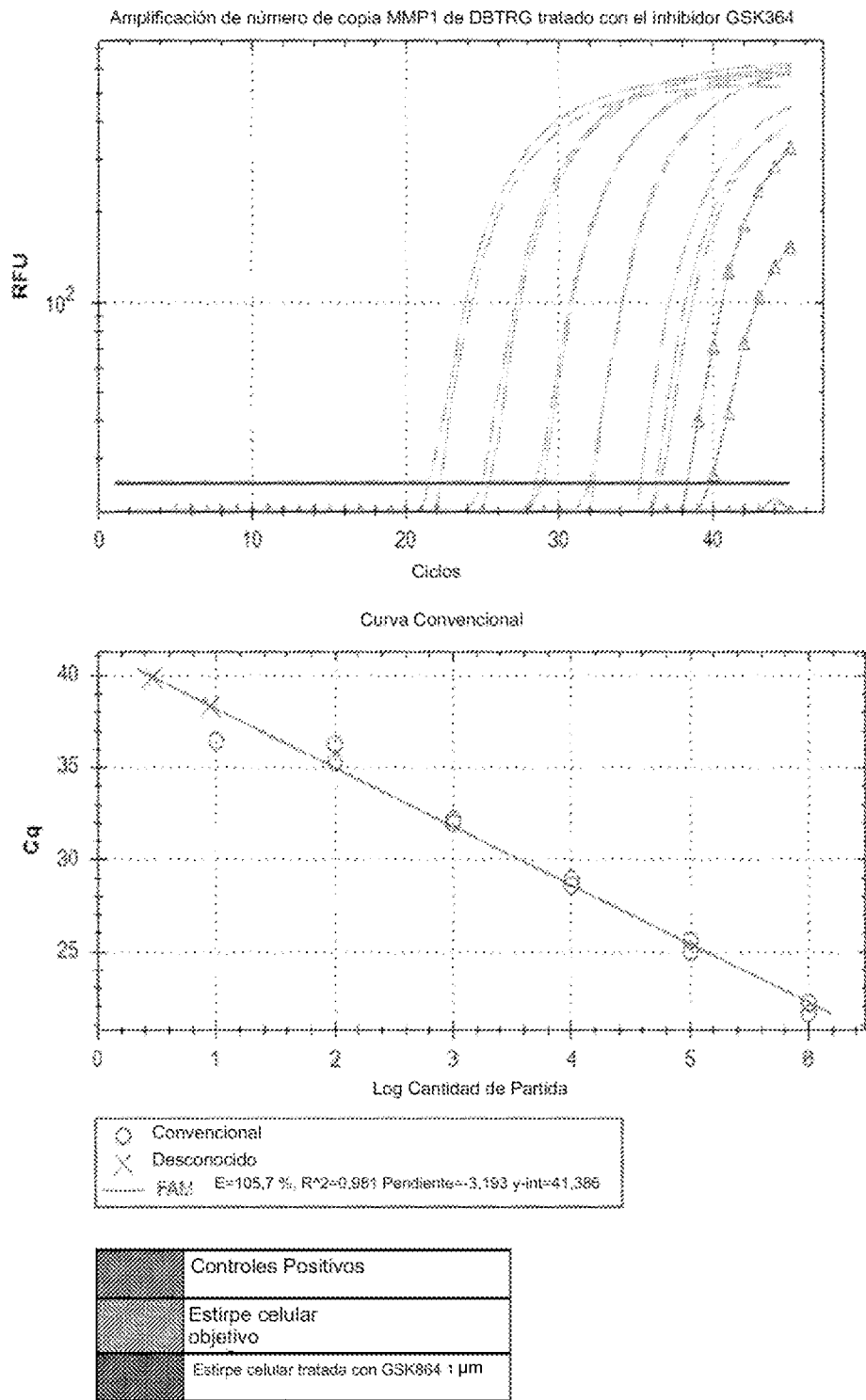


Figura 8

Amplificación de número de copia MMP1 en HL-60 tratado con GSK864

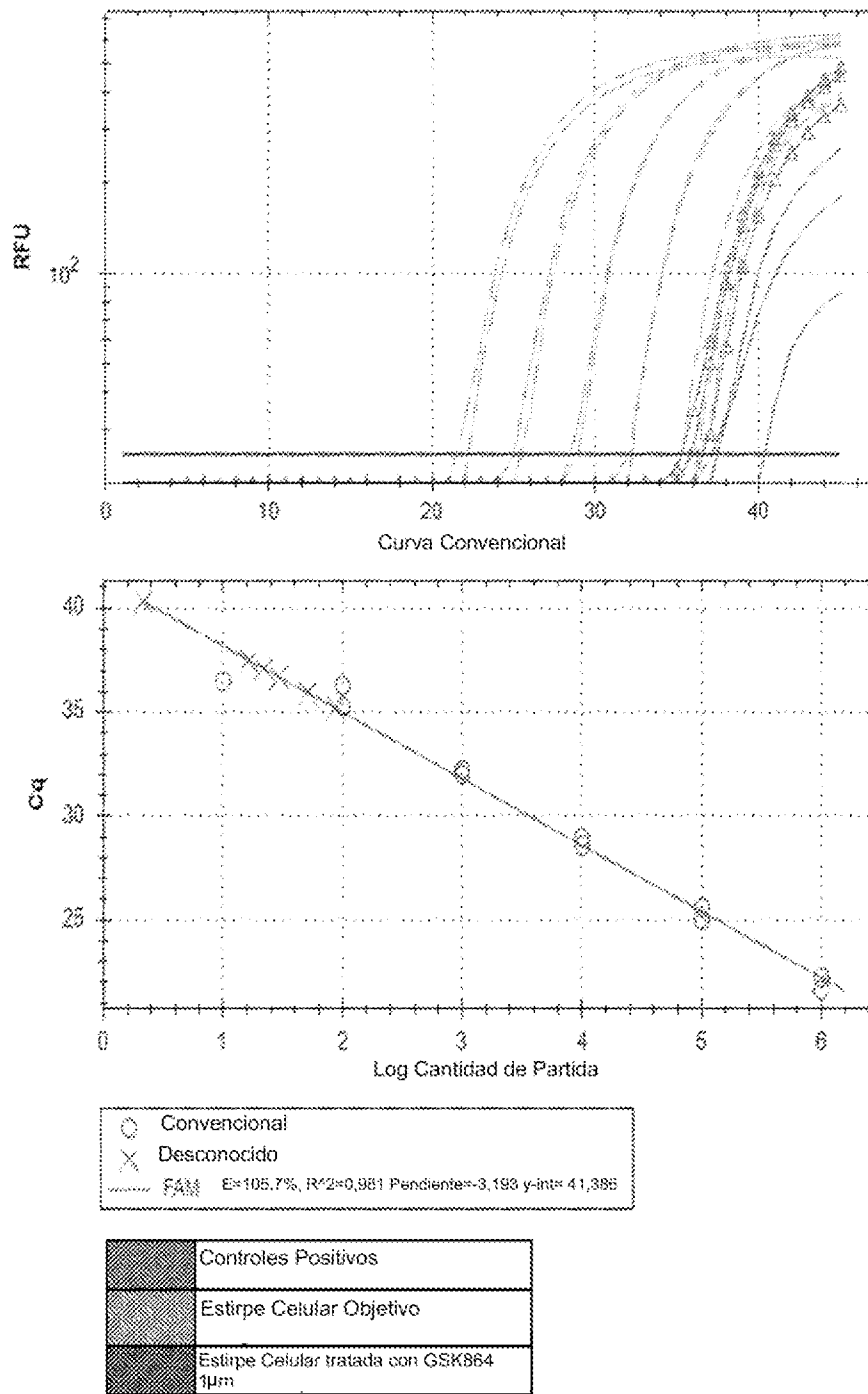


Figura 9