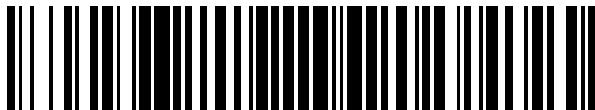


(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 923 602**

(51) Int. Cl.:

**C12Q 1/68** (2008.01)

**C12N 15/11** (2006.01)

**C12Q 1/6869** (2008.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **28.12.2015 E 20179648 (9)**

(97) Fecha y número de publicación de la concesión europea: **01.06.2022 EP 3766986**

---

(54) Título: **Detección y tratamiento de enfermedades que muestran heterogeneidad celular de enfermedad y sistemas y métodos para comunicar los resultados de las pruebas**

(30) Prioridad:

**31.12.2014 US 201462098426 P  
01.05.2015 US 201562155763 P**

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

**28.09.2022**

(73) Titular/es:

**GUARDANT HEALTH, INC. (100.0%)  
505 Penobscot Drive  
Redwood City, CA 94063, US**

(72) Inventor/es:

**ELTOUKHY, HELMY;  
TALASAZ, AMIRALI;  
KERMANI, BAHRAM, GHAFFARZADEH y  
IHUEGBU, NNAMDI**

(74) Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

**ES 2 923 602 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

Detección y tratamiento de enfermedades que muestran heterogeneidad celular de enfermedad y sistemas y métodos para comunicar los resultados de las pruebas

5 **ANTECEDENTES**

La atención de salud recién ahora está empezando a utilizar eficazmente la información del genoma humano para diagnosticar y tratar la enfermedad. En ninguna parte es esto más crucial que en el tratamiento del cáncer, del cual mueren 7,6 millones de personas en los EE.UU. cada año, y para el cual EE.UU. gasta 87 mil millones de dólares al año en tratamiento. El cáncer se refiere a cualquier trastorno de diversas neoplasias malignas caracterizadas por la proliferación de células anaplásicas que tienden a invadir el tejido circundante y metastatizar a nuevos sitios corporales y las condiciones patológicas caracterizadas por tales crecimientos.

10 Una de las razones que el cáncer es difícil de tratar es que los métodos de prueba actuales no pueden ayudar a que los médicos hagan coincidir cánceres específicos con tratamientos farmacológicos eficaces. Y es un objetivo en movimiento: las células cancerosas cambian y mutan constantemente. Los cánceres pueden acumular variantes genéticas mediante, por ejemplo, mutación de células somáticas. Tales variantes incluyen, por ejemplo, variantes de secuencia y variantes de número de copia. El análisis de tumores ha indicado que diferentes células de un tumor pueden tener diferentes variantes genéticas. Tal diferenciación entre células tumorales se ha denominado heterogeneidad tumoral.

15 Los cánceres pueden evolucionar con el tiempo, llegando a ser resistentes a una intervención terapéutica. Se sabe que ciertas variantes se correlacionan con la capacidad de respuesta o la resistencia a intervenciones terapéuticas específicas. Serían beneficiosos tratamientos más eficaces para los cánceres que presentan heterogeneidad tumoral. Dichos cánceres pueden tratarse con una segunda intervención terapéutica diferente a la que responde el cáncer.

20 Métodos de secuenciación de ADN permiten la detección de variantes genéticas en el ADN de las células tumorales. Los tumores cancerosos arrojan continuamente su material genómico único al torrente sanguíneo. Desafortunadamente, estas "señales" genómicas reveladoras son tan débiles que las tecnologías de análisis genómico actuales, incluida la secuenciación de próxima generación, solo pueden detectar tales señales esporádicamente o en pacientes con una carga tumoral terminal alta. La razón principal de esto es que estas tecnologías están plagadas de tasas de error y sesgos que pueden ser órdenes de magnitud más altas que las que se requieren para detectar de manera confiable las alteraciones genómicas de novo asociadas con el cáncer.

25 En una tendencia paralela, para entender el significado clínico de una prueba genética, el tratamiento de los profesionales debe tener un conocimiento práctico de los principios básicos de la herencia genética y facilidades razonables, con la interpretación de probabilísticas de datos. Algunos estudios sugieren que muchos profesionales tratantes no están adecuadamente preparados para interpretar las pruebas genéticas de susceptibilidad a enfermedades. Algunos médicos tienen dificultades para interpretar datos probabilísticos relacionados con la utilidad clínica de las pruebas de diagnóstico, como el valor predictivo positivo o negativo de una prueba de laboratorio.

30 Las tasas de error y el sesgo en la detección de alteraciones novo genómicas asociadas con el cáncer, junto con inadecuada explicación o las implicaciones de las pruebas genéticas para el cáncer, han reducido la calidad del cuidado de pacientes con cáncer. Las sociedades profesionales, como el College of American Pathologists (CAP) y el American College of Medical Genetics (ACMG), han publicado estándares o pautas para los laboratorios que brindan pruebas genéticas, que requieren que los informes que contienen información genética incluyan contenido interpretativo que sea comprensible para médicos generalistas.

35 La WO 2014/121117 (Univ Leland Stanford Junior) se refiere a la "identificación y uso de marcadores de tumores de ácidos nucleicos circulantes" (título). La WO 2012/024543 ((aris Life Sciences Luxembourg Holdings) se refiere a "biomarcadores circulantes para enfermedad" (título).

40 **55 SUMARIO**

45 La invención proporciona un método que comprende usar una base de datos informática para identificar una o más intervenciones terapéuticamente eficaces para un sujeto que tiene cáncer, en donde la base de datos informática incluye, para cada uno de la pluralidad de sujetos que tiene cáncer: (i) datos de pruebas genómicas de tumores, incluyendo alteraciones somáticas, recogidas en dos o más intervalos de tiempo por sujeto mediante biopsia en serie de ADN libre de células; (ii) una o más intervenciones terapéuticas administradas a cada uno de los sujetos en uno o más momentos; y (iii) eficacia de las intervenciones terapéuticas.

50 A continuación se exponen detalles adicionales del método, que incluyen métodos para obtener los datos incluidos en una base de datos informática, por ejemplo, datos de pruebas genómicas de tumores y de eficacia de

tratamiento.

En algunas realizaciones, la determinación de una intervención terapéutica tiene en cuenta las frecuencias relativas de las alteraciones genéticas relacionadas con el tumor. En algunas realizaciones, la intervención terapéutica comprende administrar, en combinación o en serie, una pluralidad de fármacos, en donde cada fármaco es relativamente más eficaz contra un cáncer que se presenta con una diferente de las mutaciones somáticas que ocurren con diferente frecuencia relativa. En algunas realizaciones, se administra en mayor cantidad un fármaco que es relativamente más eficaz contra un cáncer que se presenta con una mutación somática que se produce con una frecuencia relativa más alta. En algunas realizaciones, los fármacos se administran en dosis que se estratifican para reflejar las cantidades relativas de las variantes en el ADN. En algunas realizaciones, los cánceres que se presentan con al menos una de las variantes genéticas son resistentes a al menos uno de los fármacos. En algunas realizaciones, la determinación de una intervención terapéutica tiene en cuenta el tejido de origen del cáncer. En algunas realizaciones, la intervención terapéutica se determina basándose en una base de datos de intervenciones que han demostrado ser terapéuticas para cánceres que tienen heterogeneidad tumoral caracterizada por cada una de las mutaciones somáticas.

En algunas realizaciones, el cáncer se selecciona de carcinomas, sarcomas, leucemias, linfomas, mielomas y cánceres del sistema nervioso central (por ejemplo, cáncer de mama, cáncer de próstata, cáncer colorrectal, cáncer de cerebro, cáncer del esófago, cáncer de cabeza y cuello, cáncer de la vejiga, cáncer ginecológico, liposarcoma y mieloma múltiple). En algunas realizaciones, las células cancerosas del tumor se derivan de una célula patológica parental común. En algunas realizaciones, las células cancerosas del tumor se derivan de diferentes células cancerosas parentales del mismo o diferente tipo de cáncer. En algunas realizaciones, el método comprende además determinar una medida de las mutaciones somáticas en una o más referencias de control para determinar la cantidad relativa.

La invención reivindicada puede usarse como parte de un método que comprende: (a) realizar análisis biomolecular de polímeros biomoleculares de células de enfermedad (por ejemplo, células de enfermedad espacialmente distintas) de un sujeto; (b) identificar y cuantificar variantes biomoleculares en las macromoléculas biomoleculares; (c) desarrollar un perfil de heterogeneidad de células de enfermedad en el sujeto que indique la presencia y cantidad relativa de una pluralidad de variantes en las macromoléculas biomoleculares, en donde diferentes cantidades relativas indican heterogeneidad de células de enfermedad; y (d) determinar una intervención terapéutica para una enfermedad que muestra la heterogeneidad de las células de enfermedad, en donde la intervención terapéutica es eficaz contra una enfermedad que tiene el perfil determinado de heterogeneidad de las células de enfermedad, y en donde la intervención terapéutica se determina de acuerdo con la invención reivindicada, en base a una base de datos de intervenciones que han demostrado ser terapéuticas para cánceres que tienen heterogeneidad tumoral caracterizada por cada una de las mutaciones somáticas.

De acuerdo con la invención reivindicada, puede proporcionarse una intervención terapéutica en donde la intervención terapéutica se determina a partir de un perfil de heterogeneidad de células enfermas en el sujeto, en donde el perfil indica la presencia y la cantidad relativa de una pluralidad de las mutaciones somáticas en los polinucleótidos, en donde las cantidades relativas diferentes indican heterogeneidad de células de enfermedad; y en donde la intervención terapéutica es eficaz, según se determina por la invención reivindicada, contra una enfermedad que tiene el perfil de heterogeneidad de células de enfermedad determinado, por ejemplo, más eficaz contra una enfermedad que presenta la pluralidad de mutaciones somáticas que contra una enfermedad con cualquiera, pero no todas, de las mutaciones somáticas.

Una base de datos informática como se define en las reivindicaciones puede compilarse. La base de datos es útil para inferir la eficacia de las intervenciones terapéuticas en sujetos con un perfil genómico tumoral. En algunas realizaciones, la pluralidad de sujetos es por lo menos 50, por lo menos 500 o por lo menos 5000. En algunas realizaciones, los datos de pruebas genómicas tumorales se recopilan mediante biopsia en serie, ADN libre de células, ARN libre de células o células tumorales circulantes. En algunas realizaciones, se usan frecuencias relativas de variantes genéticas detectadas para clasificar la eficacia del tratamiento. En algunas realizaciones, se usa información adicional para ayudar a clasificar la eficacia del tratamiento incluyendo, pero no limitada a, el peso, los efectos adversos del tratamiento, las pruebas histológicas, las pruebas de sangre, la información radiográfica, los tratamientos anteriores y el tipo de cáncer. En algunas realizaciones, se recopila y clasifica cuantitativamente mediante pruebas adicionales la respuesta al tratamiento por paciente. En algunas realizaciones, la prueba adicional es una prueba basada en sangre u orina.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

- La Figura 1** muestra un diagrama de flujo de un método ejemplar de determinación y uso de una intervención terapéutica.  
**La Figura 2** muestra un diagrama de flujo de un método ejemplar para determinar la frecuencia de variantes en una muestra corregida en base a CNV en un locus.  
**La Figura 3** muestra un diagrama de flujo de un método ejemplar para proporcionar ciclos de terapia pulsada

que pueden retrasar la resistencia a los fármacos.

**La Figura 4** muestra un diagrama de flujo de un método ejemplar para detectar la carga tumoral usando CNV en los orígenes de la replicación para detectar el ADN de las células en división.

**La Figura 5** muestra un sistema informático ejemplar.

**La Figura 6** muestra un escaneo ejemplar de CNV a través de una región de un genoma de muestras que contienen células en estado de reposo y en estado de división celular. No se observa NVC genómica en los loci a y b, pero el locus c muestra duplicación de genes. En las células en estado de reposo, el número de copias es relativamente igual en todos los intervalos de la región, excepto aquellos intervalos que se superponen al locus de la duplicación de genes. En la muestra que contiene ADN de células tumorales, que están experimentando división celular, el número de copias parece aumentar inmediatamente después de los orígenes de la replicación, lo que proporciona una variación en el CNV en la región. La desviación es particularmente dramática en un locus que exhibe NVC en un origen de replicación (c).

**La Figura 7** muestra un curso ejemplar de seguimiento y tratamiento de una enfermedad en un sujeto.

**La Figura 8** muestra un panel ejemplar de 70 genes que exhiben variación genética en el cáncer.

**La Figura 9A** muestra un sistema ejemplar para comunicar los resultados de las pruebas de cáncer.

**La Figura 9B** muestra un proceso ejemplar para reducir las tasas de error y el sesgo en las lecturas de secuencias de ADN y generar informes genéticos para los usuarios.

**Las Figuras 10A-10C** muestran procesos ejemplares para informar los resultados de las pruebas genéticas a los usuarios.

**La Figura 10D-10I-2** muestra páginas de un informe de prueba genética ejemplar.

**La Figura 10J-10P** muestra varios gráficos de flujo modificados ejemplares.

**La Figura 11A-11B** muestra procesos ejemplares para detectar mutaciones y notificar los resultados de las pruebas a los usuarios.

## 25 DESCRIPCIÓN DETALLADA

Puede detectarse en una muestra biológica el mosaicismo biomolecular (por ejemplo, mosaicismo genético), como una población genómica heterogénea de células de ácido desoxirribonucleico (ADN). El mosaicismo genético puede existir a nivel de organismo. Por ejemplo, las variantes genéticas que surgen temprano en el desarrollo pueden dar como resultado que diferentes células somáticas tengan diferentes genomas. Un individuo puede ser una quimera, por ejemplo, producida por la fusión de dos cigotos. El trasplante de órganos de un donante alogénico puede dar como resultados mosaicos genéticos, que también pueden detectarse examinando polinucleótidos vertidos en la sangre del órgano transplantado. La heterogeneidad de las células enfermas, en donde las células enfermas tienen diferentes variantes genéticas, es otra forma de mosaicismo genético. Puede detectar el mosaicismo y, en el caso de enfermedad, puede proporcionarse una intervención terapéutica. La realización de perfiles en todo el cuerpo del mosaicismo biomolecular puede realizarse mediante el uso de polinucleótidos circulantes, que pueden derivarse u originarse de otro modo a partir de células en diversas ubicaciones del cuerpo de un sujeto.

Las células enfermas, tales como tumores, pueden evolucionar con el tiempo, resultando en diferentes subpoblaciones clonales que tienen nuevas características genéticas y fenotípicas. Esto puede resultar de mutaciones naturales a medida que las células se dividen, o pueden ser impulsadas por tratamientos que se dirigen a ciertas subpoblaciones clonales, permitiendo que los clones más resistentes al tratamiento proliferen por selección negativa. La existencia de subpoblaciones de células enfermas que tienen diferentes características genotípicas o fenotípicas se denomina aquí heterogeneidad de células enfermas o, en el caso de cáncer, heterogeneidad tumoral.

En la actualidad, los cánceres se tratan en base a formas mutantes encontradas en una biopsia de cáncer.

Por ejemplo, el hallazgo de Her2+ incluso en pequeñas cantidades de células de cáncer de mama puede ser indicativo de cáncer de mama, que puede seguirse con un tratamiento que utilice una terapia anti-Her2+. Como otro ejemplo, un cáncer colorrectal en donde se encuentra un mutante de KRAS en pequeñas cantidades puede tratarse con una terapia para la que responde KRAS.

Las herramientas para el análisis fino de células enfermas (por ejemplo, tumores) permiten la detección de la heterogeneidad de las células enfermas. Además, el análisis de polinucleótidos procedentes de células enfermas ubicadas en todo el cuerpo permite un perfil de heterogeneidad de células enfermas en todo el cuerpo. El uso de ADN libre de células, o ADN circulante, es particularmente poderoso porque los polinucleótidos en la sangre no provienen de células localizadas físicamente. Más bien, incluyen células de sitios metastásicos en todo el cuerpo. Por ejemplo, el análisis puede mostrar que una población de células de cáncer de mama incluye un 90% que son Her2+ y un 10% que son Her2-. Esto puede determinarse, por ejemplo, cuantificando el ADN para cada forma en una muestra, por ejemplo, ADN libre de células (cfADN), detectando así la heterogeneidad en el tumor.

Esta información puede ser utilizada por un profesional de la salud, por ejemplo, un médico, para desarrollar intervenciones terapéuticas. Por ejemplo, un sujeto que tiene un tumor heterogéneo puede tratarse como si tuviera dos tumores, y una intervención terapéutica puede tratar cada uno de los tumores. La intervención terapéutica podría incluir, por ejemplo, una terapia de combinación que incluye un primer fármaco eficaz contra el primer tipo de tumor y un segundo fármaco eficaz contra el segundo tipo de tumor. Los fármacos se pueden administrar en cantidades que

reflejen las cantidades relativas de las formas mutantes detectadas. Por ejemplo, un fármaco para tratar la forma mutante que se encuentra en cantidades relativas más altas puede administrarse a una dosis mayor que un fármaco para tratar la forma mutante en menor cantidad relativa. O bien, el tratamiento para el mutante en menor cantidad relativa puede retrasarse o escalonarse con respecto al mutante en mayor cantidad.

5 La monitorización de los cambios en el perfil de la heterogeneidad de las células patológicas a lo largo del tiempo permite calibrar la intervención terapéutica para un tumor en evolución. Por ejemplo, el análisis puede mostrar cantidades crecientes de polinucleótidos que portan mutantes de resistencia a fármacos. En este caso, la intervención terapéutica puede modificarse para disminuir la cantidad de fármaco eficaz para tratar un tumor que no porta el mutante de resistencia y aumentar la administración de un fármaco que si trata un tumor que porta el marcador de resistencia.

10 15 Las intervenciones terapéuticas pueden ser determinadas por un profesional médico o por un algoritmo de ordenador, o una combinación de los dos. Una base de datos de acuerdo con la invención reivindicada puede contener los resultados de intervenciones terapéuticas contra enfermedades que tienen varios perfiles de heterogeneidad de células de enfermedad. La base de datos se puede consultar para determinar una intervención terapéutica para una enfermedad con un perfil particular.

20 25 Los métodos de la invención pueden ser útiles para determinar una intervención terapéutica para un sujeto que tienen un cáncer que muestra heterogeneidad de células de enfermedad, por ejemplo, heterogeneidad de tumores. Esto puede implicar analizar macromoléculas biológicas (por ejemplo, polinucleótidos de secuenciación) de células de enfermedad (por ejemplo, células de enfermedad espacialmente distintas) de un sujeto que tiene la enfermedad. Se desarrolla un perfil de heterogeneidad de las células enfermas que indica la existencia de variantes genéticas específicas de las células enfermas y la cantidad de estas variantes entre sí. Esta información, a su vez, se utiliza para determinar una intervención terapéutica que tenga en cuenta el perfil.

### **Células de la enfermedad**

30 Un objeto de los métodos de esta descripción es cualquier organismo multicelular. Más específicamente, el sujeto puede ser una planta o un animal, un vertebrado, un mamífero, un ratón, un primate, un simio o un ser humano. Los animales incluyen, pero no se limitan a, animales de granja, animales deportivos y mascotas. Un sujeto puede ser un paciente, por ejemplo, un sujeto bajo el cuidado de un profesional sanitario.

35 A las células que presentan patología de la enfermedad se hace referencia en la presente como células enfermas.

40 45 En particular, la enfermedad es un cáncer. El cáncer es una afección caracterizada por células anormales que se dividen sin control. Los cánceres incluyen, sin limitación, carcinomas, sarcomas, leucemias, linfomas, mielomas y cánceres del sistema nervioso central. Ejemplos más específicos de cánceres son cáncer de mama, cáncer de próstata, cáncer colorrectal, cáncer de cerebro, cáncer de esófago, cáncer de cabeza y cuello, cáncer de vejiga, cáncer ginecológico, liposarcoma y mieloma múltiple.

50 55 60 Otros tipos de cáncer incluyen, por ejemplo, leucemia linfoblástica aguda (LLA), leucemia mieloide aguda (LMA), carcinoma adrenocortical, Sarcoma de Kaposi, cáncer anal, carcinoma de células basales, cáncer del conducto biliar, cáncer de vejiga, cáncer de hueso, osteosarcoma, maligno histiocitoma fibroso, glioma de tronco encefálico, cáncer de cerebro, craneofaringioma, ependimoblastoma, ependimoma, meduloblastoma, meduloeptilioma, tumor del parénquima pineal, cáncer de mama, tumor bronquial, linfoma de Burkitt, linfoma no Hodgkin, cáncer de cuello uterino, leucemia LLC), leucemia mielógena crónica (LMC), cáncer de colon, cáncer colorrectal, linfoma cutáneo de células T, carcinoma ductal *in situ*, cáncer de endometrio, cáncer de esófago, sarcoma de Ewing, cáncer de ojo, melanoma intraocular, retinoblastoma, histiocitoma fibroso, cáncer de vesícula biliar, cáncer gástrico, glioma, leucemia de células pilosas, cáncer de cabeza y cuello, cáncer de corazón, cáncer hepatocelular (hígado), linfoma de Hodgkin, cáncer de hipofaringe, cáncer de riñón, cáncer de laringe, cáncer de labio, cáncer de cavidad oral, cáncer de pulmón, carcinoma de células microcíticas, carcinoma de células pequeñas, melanoma, cáncer de boca, síndromes mielodisplásicos, mieloma múltiple, meduloblastoma, cáncer de cavidad nasal, cáncer de seno paranasal, neuroblastoma, nasofaríngeo cáncer, cáncer oral, cáncer de orofaringe, osteosarcoma, cáncer de ovario, cáncer de páncreas, papilomatosis, paraganglioma, cáncer de paratiroides, cáncer de pene, cáncer de faringe, tumor pituitario, neoplasia de células plasmáticas, cáncer de próstata, cáncer de recto, cáncer de células renales, rhabdomiosarcoma, glándula salival cáncer, síndrome de Sezary, cáncer de piel, no melanoma, cáncer de intestino delgado, sarcoma de tejidos blandos, carcinoma de células escamosas, cáncer de testículo, cáncer de garganta, timoma, cáncer de tiroides, cáncer de uretra, cáncer de útero, sarcoma de útero, cáncer de vagina, cáncer de vulva, macroglobulinemia de Waldenstrom y/o tumor de Wilms.

65 Un tumor es una colección de células cancerosas (células de la enfermedad de cáncer). Esto incluye, por ejemplo, una colección de células en una sola masa de células (por ejemplo, un tumor sólido), una colección de células de diferentes sitios tumorales metastásicos (tumores metastásicos) y tumores difusos (por ejemplo, células tumorales circulantes). Un tumor puede incluir células de un solo cáncer (p. ej., cáncer colorrectal) o cánceres múltiples (p. ej.,

cáncer colorrectal y cáncer de páncreas). Un tumor puede incluir células que se originan a partir de una única célula somática original o de diferentes células somáticas.

Las células de la enfermedad en el sujeto son espacialmente distintas. Las células de enfermedad son espacialmente distintas si las células están ubicadas al menos 1 cm, al menos 2 cm, al menos 5 cm o al menos 10 cm de distancia en un cuerpo, por ejemplo, en diferentes tejidos u órganos, o en el mismo tejido u órgano. En el caso del cáncer, los ejemplos de células cancerosas espacialmente distintas incluyen células cancerosas de cánceres difusos (como leucemias), células cancerosas en diferentes sitios metastásicos y células cancerosas de la misma masa de células tumorales que están separadas por al menos 1 cm.

La carga de células enfermas (por ejemplo, "carga tumoral") es una medida cuantitativa de la cantidad de células enfermas en un sujeto. Una medida de la carga de células patológicas es la fracción de macromoléculas biológicas totales en una muestra que son macromoléculas biológicas patológicas, por ejemplo, la cantidad relativa de polinucleótidos tumorales en una muestra de polinucleótidos libres de células. Por ejemplo, si el ADNc de un primer sujeto tiene un 10% de polinucleótidos cancerosos, se puede decir que el sujeto tiene una carga tumoral libre de células del 10%. Si el ADNc de un segundo sujeto tiene un 5% de polinucleótidos cancerosos, se puede decir que el segundo sujeto tiene la mitad de la carga tumoral libre de células que el primer sujeto. Estas medidas son mucho más relevantes sobre una base intrasujeto que entre sujetos, ya que la carga tumoral libre de células en un individuo puede ser mucho mayor o menor que en otro individuo a pesar de los diferentes niveles de carga de la enfermedad. Sin embargo, estas medidas pueden usarse de manera bastante efectiva para monitorear la carga de enfermedad dentro de un individuo, por ejemplo, un aumento de un 5% a un 15% de carga tumoral de ADN libre de células puede indicar una progresión significativa de la enfermedad, mientras que una disminución del 10% a 1% puede indicar una respuesta parcial al tratamiento.

Los polinucleótidos que se van a secuenciar pueden obtenerse de sitios espacialmente distintos. Esto incluye polinucleótidos obtenidos de biopsias de diferentes lugares en una sola masa tumoral. También incluye polinucleótidos procedentes de células en diferentes sitios de tumores metastásicos. Las células arrojan polinucleótidos a la sangre donde se pueden detectar como polinucleótidos libres de células (por ejemplo, ADN tumoral circulante). Los polinucleótidos libres de células también se pueden encontrar en otros fluidos corporales como la orina. Por lo tanto, cfADN proporciona un perfil más preciso de heterogeneidad tumoral en toda la población de células de la enfermedad que el ADN procedente de una única ubicación tumoral. El ADN extraído de las células de la población de células de la enfermedad en un cuerpo se denomina "ADN de carga de enfermedad" o, en el caso del cáncer, "ADN de carga de tumor".

Las células enfermas, como los tumores, pueden compartir perfiles biomoleculares iguales o similares. Por ejemplo, los tumores pueden compartir una, dos, tres o más variantes genéticas. Dichas variantes pueden compartir la misma estratificación, por ejemplo, la frecuencia más alta, la segunda frecuencia más alta, etc. Los perfiles también pueden compartir cargas de células de enfermedad similares, por ejemplo, cargas de cfADN, por ejemplo, dentro del 15%, dentro del 10%, dentro del 5% o dentro del 2%.

#### Analitos

Como se usa en el presente documento, una macromolécula es una molécula formada a partir de subunidades monoméricas. Las subunidades monoméricas que forman macromoléculas biológicas incluyen, por ejemplo, nucleótidos, aminoácidos, monosacáridos y ácidos grasos. Las macromoléculas biológicas incluyen, por ejemplo, biopolímeros y macromoléculas no poliméricas.

Un polinucleótido es una macromolécula que comprende un polímero de nucleótidos. Los polinucleótidos incluyen, por ejemplo, polidesoxirribonucleótidos (ADN) y polirribonucleótidos (ARN). Un polipéptido es una macromolécula que comprende un polímero de aminoácidos. Un polisacárido es una macromolécula que comprende un polímero de monosacáridos. Los lípidos son un grupo diverso de compuestos orgánicos que incluyen, por ejemplo, grasas, aceites y hormonas que comparten la característica funcional de no interactuar apreciablemente con el agua. Por ejemplo, un triglicérido es una grasa formada por tres cadenas de ácidos grasos.

Un polinucleótido cáncer (por ejemplo, ADN de cáncer) es un polinucleótido (por ejemplo, ADN) derivado de una célula de cáncer. El ADN y/o ARN del cáncer se pueden extraer de tumores, de células cancerosas aisladas o de fluidos biológicos (por ejemplo, saliva, suero, sangre u orina) en forma de ADN libre de células (cfADN) o ARN libre de células.

ADN libre de células es el ADN situado fuera de una célula en un fluido corporal, por ejemplo, en la sangre o la orina. Los ácidos nucleicos circulantes (ANC) son ácidos nucleicos que se encuentran en el torrente sanguíneo. El ADN libre de células en la sangre es una forma de ácido nucleico circulante. ADN libre de células se cree que surge de las células que mueren que arrojan su ADN en la sangre. Debido a que las células cancerosas espacialmente distintas arrojarán ADN a los fluidos corporales, como la sangre, el ADNc de los sujetos cancerosos generalmente comprende ADN canceroso de células cancerosas espacialmente distintas.

## Muestras biológicas

Los analitos pueden derivar de una muestra biológica, por ejemplo, una muestra que comprende una macromolécula biológica. Una muestra biológica puede derivarse de cualquier órgano, tejido o fluido biológico. Una muestra biológica puede comprender, por ejemplo, un fluido corporal o una muestra de tejido sólido. Un ejemplo de una muestra de tejido sólido es una muestra de tumor, por ejemplo, de una biopsia de tumor sólido. Los fluidos corporales incluyen, por ejemplo, sangre, suero, células tumorales, saliva, orina, fluido linfático, fluido prostático, fluido seminal, leche, esputo, heces y lágrimas. Los fluidos corporales son fuentes particularmente buenas de macromoléculas biológicas a partir de células enfermas espacialmente distintas, ya que tales células de muchos lugares del cuerpo pueden verter estas moléculas en el fluido corporal. Por ejemplo, la sangre y la orina son buenas fuentes de polinucleótidos libres de células. Las macromoléculas de tales fuentes pueden proporcionar un perfil más preciso de las células enfermas que las macromoléculas derivadas de una masa de células enfermas localizadas.

Las cantidades de los polinucleótidos de la enfermedad en una muestra de fluido corporal pueden ser aumentadas. Dichos aumentos pueden aumentar la sensibilidad de detección de polinucleótidos patológicos. Una intervención, como una intervención terapéutica, a un sujeto que provoca la lisis de las células enfermas, vaciando su ADN en el fluido circundante. Tales intervenciones pueden incluir la administración de quimioterapia. También puede incluir la administración de radiación o ultrasonido a todo el cuerpo de un sujeto, o a una parte del cuerpo de un sujeto, como dirigirse a un tumor o un órgano enfermo. Después de la administración de la intervención y cuando aumenta la cantidad de polinucleótidos de la enfermedad en el líquido, se recoge una muestra de líquido para su análisis. El intervalo entre la administración de la intervención y la recolección puede ser lo suficientemente largo para que aumenten los polinucleótidos de la enfermedad, pero no tanto como para que se eliminen del cuerpo. Por ejemplo, se puede administrar una dosis baja de quimioterapia aproximadamente una semana antes de la recolección de la muestra.

## Métodos analíticos

Hay varios tipos de análisis biomolecular incluyendo, por ejemplo, genómico, epigenético (por ejemplo, metilación), expresión de ARN y proteómica. El análisis genómico se puede realizar, por ejemplo, con un analizador genético, por ejemplo, usando secuenciación de ADN. El análisis de metilación se puede realizar, por ejemplo, mediante conversión de bases metiladas seguido de secuenciación de ADN. El análisis de expresión de ARN se puede realizar, por ejemplo, mediante hibridación de matriz de polinucleótidos. El análisis proteómico se puede realizar, por ejemplo, mediante espectrometría de masas.

Como se usa en este documento, el término "analizador genético" se refiere a un sistema que incluye un secuenciador de ADN para generar información de la secuencia de ADN y un ordenador que comprende software que realiza el análisis bioinformático en la secuencia de ADN de la información. El análisis bioinformático puede incluir, sin limitación, ensamblar datos de secuencia, detectar y cuantificar variantes genéticas en una muestra, incluidas variantes de la línea germinal (por ejemplo, heterocigosidad) y variantes de células somáticas (por ejemplo, variantes de células cancerosas).

Métodos analíticos pueden incluir la generación y la captura de la información genética. La información genética puede incluir información de secuencia genética, estados de ploidía, la identidad de una o más variantes genéticas, así como una medida cuantitativa de las variantes. El término "medida cuantitativa" se refiere a cualquier medida de cantidad, incluidas las medidas absolutas y relativas. Una medida cuantitativa puede ser, por ejemplo, un número (por ejemplo, un recuento), un porcentaje, una frecuencia, un grado o una cantidad umbral.

Los polinucleótidos se pueden analizar mediante cualquier método conocido en la técnica. Normalmente, el secuenciador de ADN empleará la secuenciación de próxima generación (por ejemplo, Illumina, 454, Ion torrent, SOLiD). El análisis de secuencia se puede realizar mediante secuenciación masivamente paralela, es decir, secuenciando simultáneamente (o en rápida sucesión) cualquiera de al menos 100.000, 1 millón, 10 millones, 100 millones o mil millones de moléculas de polinucleótidos. Los métodos de secuenciación pueden incluir, pero no se limitan a: secuenciación de alto rendimiento, pirosecuenciación, secuenciación por síntesis, secuenciación de molécula única, secuenciación de nanoporos, secuenciación de semiconductores, secuenciación por ligación, secuenciación por hibridación, ARN-Seq (Illumina), expresión génica digital (Helicos), secuenciación de próxima generación, secuenciación de una sola molécula por síntesis (SMSS) (Helicos), secuenciación paralela masiva, matriz clonal de una sola molécula (Solexa), secuenciación rápida, Maxam-Gilbert o Secuenciación de Sanger, caminata de cebadores, secuenciación usando plataformas PacBio, SOLiD, Ion Torrent, Genius (GenapSys) o Nanopore (por ejemplo, Oxford Nanopore) y cualquier otro método de secuenciación conocido en la técnica.

El secuenciador de ADN puede aplicar el método de secuenciación de Gilbert basado en la modificación química de ADN seguida por la escisión en bases específicas, o se puede aplicar la técnica de Sanger que se basa en la terminación de cadena de didesoxinucleótidos. El método Sanger se hizo popular debido a su mayor eficiencia y baja radiactividad. El secuenciador de ADN puede utilizar técnicas que no requieren amplificación de ADN (reacción

en cadena de la polimerasa - RCP), lo que acelera la preparación de la muestra antes de la secuenciación y reduce los errores. Además, se recopilan datos de secuenciación de las reacciones provocadas por la adición de nucleótidos en la hebra complementaria en tiempo real. Por ejemplo, los secuenciadores de ADN pueden utilizar un método llamado Monolécula en tiempo real (SMRT), donde los datos de secuenciación son producidos por la luz (capturada por una cámara) emitida cuando se agrega un nucleótido a la hebra complementaria por enzimas que contienen colorantes fluorescentes.

La secuenciación del genoma puede ser selectiva, por ejemplo, dirigida a porciones del genoma de interés. Por ejemplo, se sabe que muchos genes (y formas mutantes de estos genes) están asociados con varios cánceres. La secuenciación de genes seleccionados o porciones de genes puede ser suficiente para el análisis deseado. Los polinucleótidos que mapean en loci específicos en el genoma que son el tema de interés pueden aislarse para secuenciarlos, por ejemplo, mediante captura de secuencia o amplificación específica de sitio.

Una secuencia de nucleótidos (por ejemplo, secuencia de ADN) se puede referir a lecturas de secuencias en bruto o lecturas de secuencias procesadas, tales como recuentos moleculares únicos inferidos de lecturas de secuencias en bruto.

Las lecturas de secuencias generadas a partir de la secuenciación son objeto de análisis incluyendo, por ejemplo, la identificación de variantes genéticas. Esto puede incluir identificar variantes de secuencias y cuantificar el número de llamadas de bases en cada locus. La cuantificación puede implicar, por ejemplo, contar el número de lecturas mapeadas a un locus genético particular. Diferentes números de lecturas en diferentes loci pueden indicar variación en el número de copias (CNV).

Los métodos de secuenciación y bioinformática que reducen el ruido y la distorsión son particularmente útiles cuando el número de polinucleótidos diana en una muestra es pequeño en comparación con polinucleótidos no diana. Cuando las moléculas diana son pocas, la señal de la diana puede ser débil. Esto puede ser el caso, por ejemplo, en el caso del ADN libre de células, donde un pequeño número de polinucleótidos tumorales puede mezclarse con un número mucho mayor de polinucleótidos de células sanas. Los métodos de seguimiento molecular pueden ser útiles en tales situaciones. El seguimiento molecular implica el seguimiento de las lecturas de secuencia de un protocolo de secuenciación hasta las moléculas de una muestra original (por ejemplo, antes de la amplificación y/o secuenciación) de las que se derivan las lecturas. Ciertos métodos implican marcar moléculas de tal manera que múltiples lecturas de secuencias producidas a partir de moléculas originales se pueden agrupar en familias de secuencias derivadas de moléculas originales. De esta forma, se pueden filtrar las llamadas de base que representan ruido. Tales métodos se describen con más detalle en, por ejemplo, WO 2013/142389 (Schmitt et al.), US 2014/0227705 (Vogelstein et al.) y WO 2014/149134 (Talasaz et al.). Los métodos de muestreo ascendente también son útiles para determinar con mayor precisión los recuentos de moléculas en una muestra. Los métodos de muestreo superior implican determinar una medida cuantitativa de moléculas de ADN individuales para las que se detectan ambas cadenas (cadenas de Watson y Crick); determinar una medida cuantitativa de moléculas de ADN individuales para las que solo se detecta una de las cadenas de ADN; inferir de estas medidas una medida cuantitativa de moléculas de ADN individuales para las que no se detectó ninguna hebra; y utilizar estas medidas para determinar la medida cuantitativa indicativa de una serie de moléculas de ADN de doble hebra individuales en la muestra. Este método se describe en más detalle en el documento PCT/US2014/072383, presentada el 24 de diciembre de 2014.

### Variantes genéticas

A las variantes genéticas también se hace referencia como "alteraciones genéticas". Las variantes genéticas son formas alternativas en un locus genético. En el genoma humano, aproximadamente el 0,1% de las posiciones de nucleótidos son polimórficas, es decir, existen en una segunda forma genética que se presenta en al menos el 1% de la población. Las mutaciones pueden introducir variantes genéticas en la línea germinal y también en las células enfermas, como el cáncer. Las secuencias de referencia, como hg19 o NCBI Build 37 o Build 38, pretenden representar un genoma de "tipo salvaje" o "normal". Sin embargo, en la medida en que tengan una sola secuencia, no identifican polimorfismos comunes que también pueden considerarse normales.

Las variantes genéticas incluyen las variantes de secuencia, número de copias de las variantes y variantes de modificación de nucleótido. Una variante de secuencia es una variación en una secuencia de nucleótidos genética. Una variante del número de copias es una desviación del tipo salvaje en el número de copias de una parte de un genoma. Las variantes genéticas incluyen, por ejemplo, variaciones de un solo nucleótido (SNP), inserciones, delecciones, inversiones, transversiones, translocaciones, fusiones de genes, fusiones de cromosomas, truncamientos de genes, variaciones en el número de copias (por ejemplo, aneuploidía, aneuploidía parcial, poliploidía, amplificación de genes), cambios anormales en las modificaciones químicas del ácido nucleico, cambios anormales en los patrones epigenéticos y cambios anormales en la metilación del ácido nucleico.

Variantes genéticas pueden ser detectadas mediante la comparación de secuencias de polinucleótidos en una muestra con una referencia, por ejemplo, a una secuencia de referencia del genoma, a un índice o a una base de datos de mutaciones conocidas. La secuencia de referencia es una secuencia de referencia disponible públicamente,

como la secuencia del genoma humano HG-19 o NCBI Build 37. La secuencia de referencia es una secuencia en una base de datos no pública. La secuencia de referencia es una secuencia de línea germinal de un organismo inferida o determinada a partir de la secuenciación de polinucleótidos del organismo.

5 Una mutación somática o alteración somática es una variante genética que surge en una célula somática. Las mutaciones somáticas se distinguen de las mutaciones que surgen en el genoma de una célula de la línea germinal (es decir, un espermatozoide u óvulo) o un cigoto, de un individuo. Las mutaciones somáticas, por ejemplo, las que se encuentran en las células cancerosas, se distinguen del genoma de la línea germinal de un sujeto en donde surgió el cáncer. También se pueden detectar comparando el genoma del cáncer con el genoma de la línea germinal o con un genoma de referencia. También existen variantes genéticas conocidas que son comunes en las células cancerosas. Se puede encontrar una base de datos de SNV en cáncer humano en el sitio web: cancer.sanger.ac.uk/cancergenome/projects/cosmic/.

10 15 La Figura 8 muestra genes conocidos, en cáncer, por exhibir mutaciones puntuales, amplificaciones, fusiones e indeles.

#### **Desviación de CNV en células que se dividen rápidamente**

20 Durante la fase S del ciclo celular, la célula replica el ADN. Una célula diploide que tiene cromosomas 2N con ADN replicado puede corresponder a aproximadamente un contenido de ADN 4X, mientras que una célula diploide que tiene cromosomas 2N sin ADN replicado puede corresponder a un contenido de ADN aproximadamente 2X. La replicación procede de los orígenes de la replicación. En los mamíferos, los orígenes de replicación están espaciados a intervalos de aproximadamente 15 kb a 300 kb. Durante este período, existen porciones del genoma en forma poliploide. Aquellas áreas entre los orígenes de replicación y la posición de la polimerasa están duplicadas, mientras 25 que aquellas áreas más allá de la posición de la polimerasa (o justo antes del origen de replicación) todavía están en número de copia única en la hebra sometida a replicación. Cuando se escanea a través del genoma, el número de copias parece desigual o distorsionado, con regiones que existen en forma de poliploidía y regiones que existen en forma diploide. Tal escaneo parece ruidoso. Esto es cierto incluso para las células que no presentan variaciones en el 30 número de copias en el genoma en estado de reposo. Por el contrario, una exploración de CNV en células en Go muestra un perfil en donde el número de copias es relativamente plano o no está distorsionado en todo el genoma. Debido a que las células cancerosas se dividen rápidamente, su perfil de CNV a través del genoma exhibe distorsión, independientemente de que el genoma también lleve CNV en ciertos loci.

35 Se puede aprovechar este hecho para detectar la carga tumoral en el ADN a partir de muestras que comprenden ADN heterogénea, por ejemplo, una mezcla de ADN de la enfermedad y el ADN saludable, como cfADN. Una manera de detectar la carga tumoral implica determinar la variación del número de copias debido a la proximidad del locus o loci examinados a varios orígenes de replicación. Las regiones que incluyen un origen de replicación tendrán muy cerca de 4 copias de ADN en ese locus (en una célula diploide), mientras que las regiones que están muy alejadas de un origen de replicación tendrán más de 2 copias (en una célula diploide). El locus o loci examinados 40 incluyen, al menos 1 kb, al menos 10 kb, al menos 100 kb, al menos 1 mb, al menos 10 mb, al menos 100 mb, a través de un cromosoma completo o a través de un genoma completo. Se determina una medida del CNV del origen de replicación (ROCNV) en la región. Esto puede ser, por ejemplo, una medida de desviación en el número de copias de un valor de tendencia central. El valor de la tendencia central puede ser, por ejemplo, media, mediana o moda. La medida de la desviación puede ser, por ejemplo, varianza o desviación estándar. Esta medida puede compararse con 45 una medida de ROCNV en la misma región en una muestra de control, por ejemplo, de un individuo sano o células en estado de reposo. Los ROCNV se pueden determinar dividiendo la región o regiones analizadas en particiones no superpuestas de varias longitudes y tomando una medida de CNV en esta partición. Esta medida de CNV puede derivarse del número de lecturas o fragmentos determinados para mapear en esas regiones después de la secuenciación. Las particiones pueden tener varios tamaños, para producir varios niveles de resolución, por ejemplo, 50 un solo nivel de base (base por base), 10 bases, 100 bases, 1 kb, 10 kb o 100 kb. Las desviaciones que son mayores que un control indican la presencia de ADN en replicación, lo que, a su vez, indica malignidad. Cuanto mayor sea el grado de desviación, mayor será la cantidad de ADN de las células sometidas a división celular en la muestra.

55 Las variaciones del número de copias genéticas verdaderas que difieren de origen de replicación basadas en la distorsión pueden calcularse de varias maneras. Por ejemplo, las posiciones de SNP heterocigotos en los loci de CNV afectados se pueden usar para inferir la variación del número de copias calculando la desviación del 50% o el desequilibrio alélico en esos loci. La distorsión debida a la proximidad del origen de la replicación no debería afectar este desequilibrio, ya que ambas copias generalmente se copiarían a intervalos de tiempo similares y, por lo tanto, se normalizarían automáticamente (aunque los cambios alélicos podrían cambiar la replicación del origen entre las dos 60 variantes alélicas). Por ejemplo, la duplicación de un segmento cromosómico que contiene un SNP podría detectarse en aproximadamente el 67% de las lecturas, mientras que la duplicación resultante de ROCNV se detectaría en aproximadamente el 50% de las lecturas. Se utilizan técnicas basadas en recuento que utilizan la densidad de fragmentos detectados o lecturas en un determinado lugar para calcular el número relativo de copias. Estas técnicas están generalmente limitadas por el ruido de Poisson y el sesgo sistemático debido a la preparación de la muestra de ADN y al sesgo de secuenciación. Una combinación de estos métodos también puede ser para obtener una precisión 65

aún mayor.

El ROCNV puede calcularse para una muestra dada y usarse para dar un valor sobre la carga tumoral libre de células a pesar de la falta de detección de variantes somáticas tradicionales, tales como SNV, NVC específicos de genes, reordenamientos genómicos, variantes epigenéticas, pérdida de heterocigosidad, etc. Los ROCNV también se pueden usar para restar la distorsión de una muestra determinada para aumentar la sensibilidad y/o especificidad de un método de estimación/detección de CNV dado al eliminar la variación que está relacionada con la proximidad del origen de replicación en lugar de debido a cambios reales en el número de copias en una célula. Las líneas celulares con cambios de número de copia conocidos o sin cambios sobre una referencia también se pueden utilizar como referencia de los ROCNV para estimar su contribución a una muestra determinada.

Puede determinarse un nivel de referencia de copias de moléculas de ADN en uno o más loci de una o más muestras de control, cada una que contiene ADN de las células sometidas a un nivel predeterminado de células de división, por ejemplo, células en estado de reposo o células tumorales que se dividen rápidamente. También se determina una medida de copias de moléculas de ADN en una muestra de prueba. La medida en las muestras de prueba puede ser de uno o más loci divididos en una o más particiones. En cada caso, una pluralidad de loci incluye cada uno un origen o replicación. La medida de copias de la muestra de prueba puede ser un promedio en todas las particiones o un nivel de varianza entre los loci. Se compara una medida de tendencia central o de variación (por ejemplo, varianza o desviación estándar) en el número de copias en la muestra de prueba con la muestra de control. Una medida que es mayor en una muestra de prueba que en un control de células en estado de reposo, o dividiéndose lentamente, indica que las células que generan el ADN en la muestra de prueba se están dividiendo más rápidamente que las células que proporcionan ADN a la muestra de control, por ejemplo, son cancerosas. De manera similar, las medidas que son similares entre una muestra de prueba y un control de células en estado de división activa indican que las células que generan el ADN en la muestra de prueba se están dividiendo a una velocidad similar a las células que se dividen rápidamente, por ejemplo, son cancerosas.

#### **Heterogeneidad de células enfermas**

Heterogeneidad de células enfermas, por ejemplo, la heterogeneidad tumoral, es la aparición de células enfermas que tienen diferentes variantes genéticas. La heterogeneidad de las células enfermas se puede determinar examinando polinucleótidos aislados de células enfermas y detectando diferencias en sus genomas. La heterogeneidad de las células enfermas también se puede inferir a partir del examen de polinucleótidos de una muestra que contiene polinucleótidos tanto de células enfermas como de células sanas basándose en diferencias en la frecuencia relativa de mutaciones somáticas. Por ejemplo, el cáncer se caracteriza por cambios a nivel genético, por ejemplo, a través de la acumulación de mutaciones somáticas en diferentes grupos clonales de células. Estos cambios pueden contribuir al crecimiento no regulado de las células cancerosas o funcionar como marcadores de capacidad de respuesta o falta de respuesta a diversas intervenciones terapéuticas.

Heterogeneidad tumoral es una condición en donde un tumor caracterizado por las células cancerosas que contienen diferentes combinaciones de variantes genéticas, por ejemplo, diferentes combinaciones de mutaciones somáticas. Es decir, el tumor puede tener diferentes células que contengan alteraciones en diferentes genes o que contengan diferentes alteraciones en el mismo gen. Por ejemplo, una primera célula podría incluir una forma mutante de BRAF, mientras que una segunda célula podría incluir formas mutantes de BRAF y ERBB2. Alternativamente, una primera célula cancerosa podría incluir el polimorfismo de un solo nucleótido EGRF 55249063 G>A, mientras que una segunda célula podría incluir el polimorfismo de un solo nucleótido EGRF 55238874 T>A. (Los números se refieren a la posición del nucleótido en la secuencia de referencia genómico.)

Por ejemplo, una célula de tumor original puede incluir una variante genética en un gen, por ejemplo, un oncogén. A medida que las células continúan dividiéndose, algunas células de la progenie, que portan la mutación original, pueden desarrollar independientemente variantes genéticas en otros genes o en diferentes partes del mismo gen. En divisiones posteriores, las células tumorales pueden acumular aún más variantes genéticas.

#### **Perfil de la heterogeneidad de enfermedad**

Pueden realizarse perfiles cuantitativos así como cualitativos de mosaicismo de enfermedad, por ejemplo, heterogeneidad tumoral. El perfil incluye información de polinucleótidos de células patológicas espacialmente distintas. El perfil es un perfil de cuerpo entero que contiene información de células distribuidas por todo el cuerpo. El análisis de polinucleótidos en cfADN permite el muestreo de ADN en toda la extensión geográfica de un tumor, en contraste con el muestreo de un área localizada de un tumor. En particular, permite el muestreo de tumores difusos y metastásicos. Esto contrasta con los métodos que detectan la mera existencia de heterogeneidad tumoral a través del muestreo localizado de un tumor. El perfil puede indicar la secuencia de nucleótidos exacta de la variante, o puede indicar simplemente un gen que porta la mutación somática.

Un perfil de heterogeneidad de las células de la enfermedad, tales como la heterogeneidad de las células tumorales, puede identificar variaciones genéticas y las cantidades relativas de cada variante. A partir de esta

información, se pueden inferir posibles distribuciones de las variantes en diferentes subpoblaciones de células. Por ejemplo, un cáncer puede comenzar con una célula que porta la mutación somática X. Como resultado de la evolución clonal, alguna progenie de esta célula puede desarrollar la variante Y. Otra progenie puede desarrollar la variante Z. A nivel celular, después del análisis, el tumor puede caracterizarse como 50% X, 35% XY y 15% XZ. A nivel de ADN (y considerando solo el ADN de las células tumorales), el perfil puede indicar 100% X, 35% Y y 15% Z. También se pueden detectar CNV en un primer locus y variantes de secuencia en un segundo locus.

La heterogeneidad tumoral puede ser detectada a partir del análisis de las secuencias de polinucleótidos de cáncer, basándose en la existencia de variaciones genómicas en diferentes loci que se producen a diferentes frecuencias. Por ejemplo, en una muestra de ADN libre de células (que probablemente contenga ADN de la línea germinal así como ADN del cáncer), se puede encontrar que una variante de secuencia de BRAF ocurre con una frecuencia del 17%, ocurre una variante de secuencia de CDKN2A a una frecuencia del 6%, se produce una variante de secuencia de ERBB2 a una frecuencia del 3% y una variante de secuencia de ATM se produce a una frecuencia del 1%. Estas diferentes frecuencias de variantes de secuencia indican heterogeneidad tumoral. De manera similar, las secuencias genéticas que exhiben diferentes cantidades de variación en el número de copias también indican heterogeneidad tumoral. Por ejemplo, el análisis de una muestra puede mostrar diferentes niveles de amplificación para los genes EGFR y CCNE1. Esto también indica heterogeneidad tumoral.

En el caso de ADN libre de células, la detección de mutaciones somáticas se puede realizar mediante la comparación de base de las llamadas en la muestra a una secuencia de referencia o, internamente, como base de las llamadas menos frecuentes, a base de las llamadas más comunes, que se presumen en la secuencia de línea germinal. En cualquier caso, la existencia de formas subdominantes (p. ej., menos del 40% de las llamadas de bases totales) en diferentes loci y con diferente frecuencia indica heterogeneidad de células patológicas.

ADN libre de células comprende típicamente una preponderancia de ADN de las células normales que tienen el genoma de la línea germinal de secuencia y, en el caso de una enfermedad, tal como cáncer, un pequeño porcentaje de ADN de las células de cáncer y que tiene una secuencia de genoma del cáncer. Las secuencias generadas a partir de polinucleótidos en una muestra de cfADN se pueden comparar con una secuencia de referencia para detectar diferencias entre la secuencia de referencia y los polinucleótidos en el cfADN. En cualquier locus, todos o casi todos los polinucleótidos de una muestra de prueba pueden ser idénticos a un nucleótido en la secuencia de referencia. Alternativamente, un nucleótido detectado con una frecuencia cercana al 100% en una muestra puede ser diferente de un nucleótido en la secuencia de referencia. Lo más probable es que esto indique una forma polimórfica normal en este locus. Si se detecta un primer nucleótido que coincide con un nucleótido de referencia en aproximadamente el 50% y se detecta un segundo nucleótido que es diferente de un nucleótido de referencia en aproximadamente el 50%, lo más probable es que esto indique heterocigosis normal. La heterocigosis puede presentarse en proporciones de alelos divergentes de 50:50, por ejemplo, 60:40 o incluso 70:30. Sin embargo, si la muestra comprende un nucleótido detectable por encima del ruido a una frecuencia por debajo (de arriba) de un rango inequívocamente heterocigoto (por ejemplo, menos del 45%, menos del 40%, menos del 30%, menos del 20%, menos del 10% o menos del 5%), esto se puede atribuir a la existencia de mutaciones somáticas en un porcentaje de las células que aportan ADN a la población de cfADN. Estos pueden provenir de células enfermas, por ejemplo, células cancerosas. (El porcentaje exacto es una función de la carga tumoral). Si la frecuencia de mutaciones somáticas en dos loci genéticos diferentes es diferente, por ejemplo, 16% en un locus y 5% en otro locus, esto indica que las células de la enfermedad, por ejemplo, las células cancerosas son heterogéneas.

En el caso de ADN de tumores sólidos, que se espera que comprenden ADN tumoral predominantemente, mutaciones somáticas también pueden ser detectadas por comparación con una secuencia de referencia. Detección de mutaciones somáticas que existen en el 100% de las células tumorales puede requerir referencia a una secuencia estándar o información sobre mutantes conocidos. Sin embargo, la existencia de secuencias subdominantes entre el conjunto de polinucleótidos en diferentes loci y en diferentes frecuencias relativas, indica heterogeneidad tumoral.

El perfil puede incluir variantes genéticas en genes que se sabe que son accionables. El conocimiento de tales variantes puede contribuir a seleccionar intervenciones terapéuticas, ya que las terapias pueden dirigirse a tales variantes. En el caso del cáncer, ya se conocen muchas variantes genéticas procesables.

##### **CNV y SNV en la heterogeneidad de células enfermas**

En general, el estado del número de copias de un gen debe reflejarse en la frecuencia de una forma genética del gen en la muestra. Por ejemplo, una variante de secuencia puede detectarse a una frecuencia consistente con la homocigosis o heterocigosis (por ejemplo, aproximadamente el 100% o aproximadamente el 50%, respectivamente) sin variación en el número de copias. Esto es consistente con un polimorfismo o mutación de la línea germinal. Se puede detectar una variante de secuencia con una frecuencia de aproximadamente el 67% (o, alternativamente, aproximadamente el 33%) de polinucleótidos en un locus, y también en un gen medido con un mayor número de copias (generalmente, n = 2). Esto es consistente con duplicación de genes en la línea germinal. Por ejemplo, una trisomía se presentaría de esta manera. Sin embargo, si se detecta una variante de secuencia a un nivel consistente con homocigosis (por ejemplo, aproximadamente 100%) pero en cantidades consistentes con la

variación del número de copias, es más probable que esto refleje la presencia de polinucleótidos de células enfermas que han sufrido amplificación génica. De manera similar, si se detecta una variante de secuencia a un nivel no incompatible con la heterocigosidad (p. ej., desviándose algo del 50%) pero en cantidades consistentes con la variación del número de copias, esto también es más probable que refleje la presencia de polinucleótidos de células enfermas; los polinucleótidos enfermos crean algún nivel de desequilibrio en la frecuencia de los alelos lejos de 50:50.

5 Esta observación puede utilizarse para inferir si una variante de secuencia es más probable presente en el nivel de línea germinal o el resultado de una mutación de células somáticas, por ejemplo, en una célula de cáncer. Por ejemplo, una variante de secuencia en un gen detectada a niveles posiblemente consistentes con heterocigosidad en la línea germinal es más probablemente el producto de una mutación somática en células enfermas si también se detecta variación en el número de copias en ese gen.

10 15 Además, en la medida en que esperamos que una duplicación de genes en la línea germinal debe soportar una variante consistente con aumento de la dosis genética (por ejemplo, aproximadamente 67% para la trisomía en un locus), la amplificación del gen de detección con una variante de secuencia dosis que se desvíe significativamente de esta cantidad esperada indica que es más probable que el CNV esté presente como resultado de una mutación de células somáticas.

20 25 30 35 El hecho de que las mutaciones somáticas en diferentes loci pueden estar presentes en el número único o múltiple de copias en la misma célula de la enfermedad también puede ser utilizado para inferir la heterogeneidad tumoral. Más específicamente, se puede inferir la heterogeneidad tumoral cuando se detectan dos genes con diferente frecuencia pero su número de copias es relativamente igual. Alternativamente, se puede inferir la homogeneidad del tumor cuando la diferencia de frecuencia entre dos variantes de secuencia es coherente con la diferencia en el número de copias de los dos genes. Por lo tanto, si se detecta una variante de EGFR al 11% y una variante de KRAS al 5%, y no se detecta NVC en estos genes, la diferencia en la frecuencia probablemente refleja la heterogeneidad del tumor (p. ej., todas las células tumorales portan un mutante EGFR y la mitad de las células tumorales también portan un mutante KRAS). Alternativamente, si el gen EGFR que lleva el mutante se detecta en un número de copias aumentado, una interpretación consistente es una población homogénea de células tumorales, cada célula porta un mutante en los genes EGFR y KRAS, pero en donde el gen KRAS está duplicado. Por consiguiente, se pueden determinar tanto la frecuencia de una variante de secuencia como una medida de CNV en el locus de la variante de secuencia en una muestra. La frecuencia puede entonces corregirse para reflejar el número relativo de células que llevan la variante ponderando la frecuencia en base a la dosis por célula determinada a partir de la medida de CNV. Este resultado es ahora más comparable en términos de número de células que llevan la variante a una variante de secuencia que no varía en número de copias.

### **Prueba de comunicación de los resultados**

40 Un informe de los resultados de análisis de variante genética (por ejemplo, variantes de secuencia, CNV, heterogeneidad de las células de la enfermedad, y combinaciones de las mismas) puede ser proporcionado por un generador de informes, por ejemplo a un profesional de la salud, por ejemplo, un médico, para ayudar en la interpretación de los resultados de la prueba (por ejemplo, datos) y la selección de opciones de tratamiento. Un informe generado por un generador de informes puede proporcionar información adicional, como resultados de laboratorio clínico, que puede ser útil para diagnosticar enfermedades y seleccionar opciones de tratamiento.

45 50 55 Con referencia ahora a la Figura 9A, se ilustra esquemáticamente un sistema con un generador de informes 1 para informar, por ejemplo, sobre los resultados de las pruebas de cáncer y las opciones de tratamiento de los mismos. El sistema generador de informes puede ser un sistema de procesamiento de datos central configurado para establecer comunicaciones directamente con: un laboratorio o sitio de datos remoto 2, un consultorio médico/proveedor de atención médica (profesional tratante) 4 y/o un paciente/sujeto 6 a través de enlaces de comunicación. El laboratorio 2 puede ser un laboratorio médico, un laboratorio de diagnóstico, una instalación médica, una práctica médica, un dispositivo de prueba en el punto de atención o cualquier otro sitio de datos remoto capaz de generar información clínica del sujeto. La información clínica del sujeto incluye, entre otros, datos de pruebas de laboratorio, por ejemplo, análisis de variantes genéticas; datos de imágenes y rayos X; resultados de los exámenes; y diagnóstico. El proveedor de atención médica o la práctica 6 puede incluir proveedores de servicios médicos, como médicos, enfermeras, asistentes de salud en el hogar, técnicos y asistentes médicos, y la práctica puede ser cualquier centro de atención médica con proveedores de atención médica. El proveedor de atención médica o la práctica también es un sitio de datos remoto. Cuando el cáncer es una enfermedad a tratar, el sujeto puede padecer cáncer, entre otras posibles enfermedades o trastornos.

60 65 Otra información clínica para un sujeto con cáncer 6 puede incluir los resultados de pruebas de laboratorio, por ejemplo, análisis de variantes genéticas, panel metabólico, hemograma completo, etc.; datos de imágenes médicas; y/o procedimientos médicos dirigidos a diagnosticar la condición, proporcionar un pronóstico, monitorear la progresión de la enfermedad, determinar la recaída o remisión, o combinaciones de las mismas. La lista de fuentes apropiadas de información clínica para el cáncer incluye, entre otros, tomografías computarizadas, resonancias magnéticas, ecografías, gammagrafías óseas, tomografías por emisión de positrones, pruebas de médula ósea,

radiografías con bario, endoscopias, linfangiogramas, UIV (urograma intravenoso) o IVP (pielograma intravenoso), punciones lumbares, cistoscopia, pruebas inmunológicas (detección de anticuerpos anti-malignina) y pruebas de marcadores de cáncer.

5 La información clínica del sujeto 6 puede obtenerse del laboratorio 2 de forma manual o automática. Cuando se desee la simplicidad del sistema, la información se puede obtener automáticamente a intervalos de tiempo predeterminados o regulares. Un intervalo de tiempo regular puede referirse a un intervalo de tiempo en donde la recopilación de datos de laboratorio se lleva a cabo automáticamente en función de una medición de tiempo como horas, días, semanas, meses, años, etc. La recopilación de datos y el procesamiento se lleva a cabo al menos una vez al día. La transferencia y recopilación de datos se lleva a cabo aproximadamente de forma mensual, quincenal, 10 semanal, varias veces a la semana o diariamente. Alternativamente, la recuperación de información puede realizarse a intervalos de tiempo predeterminados, que pueden no ser intervalos de tiempo regulares. Por ejemplo, un primer paso de recuperación puede ocurrir después de una semana y un segundo paso de recuperación puede ocurrir después de un mes. La transferencia y recopilación de datos se puede personalizar de acuerdo con la naturaleza del trastorno que se está manejando y la frecuencia de las pruebas y exámenes médicos requeridos de los sujetos.

15 La FIG. 9B muestra un proceso ejemplar para generar informes genéticos, que incluye un mapa de respuesta tumoral y un resumen asociado de alteraciones. Un mapa de respuesta tumoral es una representación gráfica de información genética que indica cambios a lo largo del tiempo en la información genética de un tumor, por ejemplo, cambios cualitativos y cuantitativos. Tales cambios pueden reflejar la respuesta de un sujeto a una intervención terapéutica. Este proceso puede reducir las tasas de error y el sesgo que pueden ser órdenes de magnitud superiores a lo que se requiere para detectar de manera confiable variantes genéticas de novo asociadas con el cáncer. El 20 proceso puede comprender, en primer lugar, capturar información genética mediante la recolección de muestras de fluidos corporales como fuentes de material genético (por ejemplo, sangre, saliva, sudor, orina, etc.). Entonces, el proceso puede comprender secuenciar los materiales (11). Por ejemplo, los polinucleótidos en una muestra se pueden secuenciar, produciendo una pluralidad de lecturas de secuencia. La carga tumoral en una muestra que comprende polinucleótidos puede estimarse como el número relativo de lecturas de secuencia que llevan una variante del número total de lecturas de secuencia generadas a partir de la muestra. Cuando se analizan variantes del número de copias, la carga tumoral se puede estimar como el exceso relativo (p. ej., en el caso de la duplicación de genes) o el déficit relativo (p. ej., en el caso de eliminación de genes) del número total de lecturas de secuencia en la prueba y loci de control. Por ejemplo, una ejecución puede producir 1000 lecturas mapeadas a un locus oncogénico de las cuales 900 corresponden al tipo salvaje y 100 corresponden a un mutante de cáncer, lo que indica una variante del número de copias en este gen. Más detalles sobre la colección de muestras ejemplares y la secuenciación de los materiales genéticos se discuten a continuación en las FIGS. 10-11.

30 35 A continuación, se puede procesar la información genética (12). Entonces se pueden identificar variantes genéticas. El proceso puede comprender determinar la frecuencia de variantes genéticas en la muestra que contiene el material genético. El proceso puede comprender separar la información del ruido (13) si este proceso es ruidoso.

40 Los métodos de secuenciación para análisis genético pueden tener tasas de error. Por ejemplo, el sistema mySeq de Illumina puede producir tasas de error porcentuales en los dígitos bajos de un solo dígito. Para 1000 lecturas de secuencia mapeadas a un locus, se puede esperar que alrededor de 50 lecturas (alrededor del 5%) incluyan errores. Ciertas metodologías, como las descritas en el documento WO 2014/149134, pueden reducir significativamente la 45 tasa de error. Los errores crean ruido que puede ocultar las señales del cáncer presente en niveles bajos en una muestra. Por ejemplo, si una muestra tiene una carga tumoral en un nivel cercano a la tasa de error del sistema de secuenciación, por ejemplo, alrededor del 0,1% al 5%, puede ser difícil distinguir una señal correspondiente a una variante genética debida al cáncer de una debida al ruido.

50 55 El análisis de variantes genéticas se pueden utilizar para el diagnóstico en presencia de ruido. El análisis puede basarse en la frecuencia de las variantes de secuencia o el nivel de NVC (14) y se puede establecer una indicación de confianza en el diagnóstico o un nivel para detectar variantes genéticas en el intervalo de ruido (15).

A continuación, el procedimiento puede comprender el aumento de la confianza diagnóstica. Esto se puede hacer usando una pluralidad de mediciones para aumentar la confianza del diagnóstico (16), o alternativamente utilizar 60 mediciones en una pluralidad de puntos de tiempo para determinar si el cáncer está avanzando, en remisión o estabilizado (17). La confianza del diagnóstico se puede utilizar para identificar estados de enfermedad. Por ejemplo, los polinucleótidos libres de células tomados de un sujeto pueden incluir polinucleótidos derivados de células normales, así como polinucleótidos derivados de células enfermas, tales como células cancerosas. Los polinucleótidos de las células cancerosas pueden tener variantes genéticas, como mutaciones de células somáticas y variantes del número de copias. Cuando se secuencian polinucleótidos libres de células de una muestra de un sujeto, estos polinucleótidos cancerosos se detectan como variantes de secuencia o como variantes de número de copias.

65 Las mediciones de un parámetro, estén o no en el intervalo de ruido, pueden proporcionarse con un intervalo de confianza. Probado a lo largo del tiempo, se puede determinar si un cáncer está avanzando, estabilizado o en remisión comparando los intervalos de confianza a lo largo del tiempo. Cuando los intervalos de confianza se

superponen, es posible que no se pueda saber si la enfermedad está aumentando o disminuyendo, porque no existe una diferencia estadísticamente significativa entre las medidas. Sin embargo, cuando los intervalos de confianza no se superponen, esto indica la dirección de la enfermedad. Por ejemplo, comparar el punto más bajo en un intervalo de confianza en un punto temporal y el punto más alto en un intervalo de confianza en un segundo punto temporal indica la dirección.

A continuación, el proceso puede comprender generar Informe/Diagnóstico genético. El proceso puede comprender generar un gráfico genético para una pluralidad de mediciones que muestren la tendencia de mutación (18) y generar un informe que muestre los resultados y las opciones del tratamiento (19).

Las FIGS. 10A-10C muestran con más detalle una manera de generar informes genéticos y diagnóstico (por ejemplo, Informe/Diagnóstico). En una implementación, la Figura 10C muestra un pseudocódigo ejemplar ejecutado por el sistema de la Figura 9A para procesar frecuencias de alelos mutantes no informados por CNV. Sin embargo, el sistema también puede procesar frecuencias de alelos mutantes informadas por CNV.

Las muestras que comprenden material genético, tal como cfADN, se pueden recoger de un sujeto a una pluralidad de puntos de tiempo, es decir, en serie. El material genético puede secuenciarse, por ejemplo, usando un sistema de secuenciación de alto rendimiento. La secuenciación puede apuntar a loci de interés para detectar variantes genéticas, tales como genes portadores de mutaciones somáticas, genes que experimentan variación en el número de copias o genes implicados en fusiones de genes, por ejemplo, en el cáncer. En cada momento se puede determinar una medida cuantitativa de las variantes genéticas encontradas. Por ejemplo, en el caso de cfADN, la medida cuantitativa puede ser la frecuencia o el porcentaje de una variante genética entre los polinucleótidos que mapean un locus, o el número absoluto de lecturas de secuencia o polinucleótidos que mapean un locus. Las variantes genéticas que tienen una cantidad distinta de cero en al menos un punto de tiempo pueden representarse gráficamente a través de todos los puntos de tiempo. Por ejemplo, en una colección de 1000 secuencias, la variante 1 puede encontrarse en los puntos temporales 1, 2 y 3 en cantidades de 50, 30 y 0, respectivamente. La variante 2 se puede encontrar en cantidades 0, 10 y 20 en estos momentos. Estos importes se pueden normalizar, para la variante 1, al 5%, 3% y 0%, y, para la variante 2, 0%, 1% y 2%. Una representación gráfica que muestre la unión de todos los resultados distintos de cero puede indicar estas cantidades para ambas variantes en todos los puntos de tiempo. Las cantidades normalizadas se pueden escalar de modo que cada porcentaje esté representado por una capa, por ejemplo, que tenga una altura de 1 mm. Así, por ejemplo, en este caso las alturas estarían en el momento 1: alturas 5 mm (variante 1) y 0 mm (variante 2); en el momento 2: alturas 3 mm (variante 1) y 1 mm (variante 2), en el momento 3: alturas 0 mm (variante 1) y 2 mm (variante 2). La representación gráfica puede tener la forma de un gráfico de áreas apiladas, como un gráfico de flujo. Un punto de tiempo "cero" (antes del primer punto de tiempo) puede ser representado por un punto, con todos los valores en 0. La altura de la cantidad de variantes en la representación gráfica puede ser, por ejemplo, relativa o proporcional entre sí. Por ejemplo, una frecuencia de variante del 5% en un momento determinado podría representarse con una altura del doble de la de una variante con una frecuencia del 2,5% en el mismo momento. El orden de apilamiento se puede elegir para facilitar la comprensión. Por ejemplo, las variantes se pueden apilar en orden de cantidad de mayor a menor, de abajo hacia arriba. O pueden apilarse en un gráfico de flujo con la variante de mayor cantidad inicial en el medio y otras variantes de cantidad decreciente en ambos lados. Las áreas se pueden codificar por colores según la variante. Las variantes del mismo gen se pueden mostrar en diferentes tonalidades del mismo color. Por ejemplo, los mutantes KRAS se pueden mostrar en diferentes tonos de azul, los mutantes EGFR en diferentes tonos de rojo.

Pasando ahora a la Figura 10A, el proceso puede comprender recibir información genética de un secuenciador de ADN (30). Entonces, el proceso puede comprender determinar alteraciones genéticas específicas y cantidades de las mismas (32).

A continuación, se genera un mapa de respuesta tumoral. Para generar el mapa, el proceso puede comprender la normalización de las cantidades para cada alteración genética para la representación en todos los puntos de prueba y luego genera un factor de escala (34). Como se usa en este documento, el término "normalizar" generalmente se refiere a medios que ajustan valores medidos en diferentes escalas a una escala teóricamente común. Por ejemplo, los datos medidos en diferentes puntos se convierten/ajustan para que todos los valores se puedan cambiar de tamaño a una escala común. Como se usa en este documento, el término "factor de escala" generalmente se refiere a un número que escala, o multiplica, alguna cantidad. Por ejemplo, en la ecuación  $y = Cx$ ,  $C$  es el factor de escala para  $x$ .  $C$  es también el coeficiente de  $x$ , y puede llamarse constante de proporcionalidad de  $y$  a  $x$ . Los valores están normalizados para permitir el trazado en una escala común que sea visualmente amigable. Y el factor de escala se usa para conocer las alturas exactas que corresponden a los valores que se trazarán (por ejemplo, el 10% de la frecuencia de alelos mutantes puede representar 1 cm en el informe donde la altura total es de 10 cm). El factor de escala se aplica a todos los puntos de prueba y, por lo tanto, se considera un factor de escala universal. Para cada punto de prueba, el proceso puede comprender la presentación de información en un mapa de respuesta tumoral (36). En la operación 36, el proceso puede comprender renderizar alteraciones y alturas relativas utilizando el factor de escala determinado (38) y asigna un indicador visual único para cada alteración (40). Además del mapa de respuesta, el proceso puede comprender generar un resumen de alteraciones y opciones de tratamiento (42). Además, se presenta información de ensayos clínicos que puede ayudar a las alteraciones genéticas particulares y otras

sugerencias de tratamiento útiles, junto con explicaciones de terminología, metodología de prueba y otra información que se agrega al informe y se presenta al usuario.

- 5 La variación del número de copia puede ser reportada como gráfico, indicando varias posiciones en el genoma y un correspondiente aumento o disminución o mantenimiento de la variación del número de copias en cada posición respectiva. Además, la variación del número de copias puede usarse para informar una puntuación porcentual que indique cuánto material patológico (o ácidos nucleicos que tienen una variación del número de copias) existe en la muestra de polinucleótido libre de células.
- 10 El informe incluye anotaciones para ayudar a los médicos a interpretar los resultados y recomendar opciones de tratamiento. La anotación puede incluir la anotación de un informe para una afección en las NCCN Clinical Practice Guidelines in Oncology™ o las pautas de práctica clínica de la American Society of Clinical Oncology (ASCO). La anotación puede incluir una lista de uno o más medicamentos aprobados por la FDA para uso no indicado en la etiqueta, uno o más medicamentos enumerados en un compendio de tratamiento contra el cáncer de los Centros de Servicios de Medicare y Medicaid (CMS) y/o uno o más medicamentos experimentales que se encuentran en literatura científica, en el informe. La anotación puede incluir conectar una opción de tratamiento farmacológico enumerada con una referencia que contenga información científica sobre la opción de tratamiento farmacológico. La información científica puede provenir de un artículo revisado por pares de una revista médica. La anotación puede incluir proporcionar un enlace a información sobre un ensayo clínico para una opción de tratamiento farmacológico en el informe. La anotación puede incluir la presentación de información en un cuadro emergente o en un cuadro flotante cerca de las opciones de tratamiento farmacológico proporcionadas en un informe electrónico. La anotación puede incluir agregar información a un informe seleccionado del grupo que consta de una o más opciones de tratamiento farmacológico, información científica sobre una o más opciones de tratamiento farmacológico, uno o más enlaces a información científica sobre una o más opciones de tratamiento farmacológico, uno o más enlaces a citas para obtener información científica sobre una o más opciones de tratamiento farmacológico e información sobre ensayos clínicos sobre una o más opciones de tratamiento farmacológico.
- 15
- 20
- 25

FIG. 10B muestra un proceso ejemplar para generar una ruta de mapa de respuesta tumoral que puede ser utilizada por un médico, por ejemplo, un médico, por ejemplo, para tomar decisiones sobre el cuidado del paciente. El proceso puede comprender primero determinar un factor de escala global (43). Para todas las frecuencias de alelos mutantes notificadas que no son CNV (variación del número de copias), el proceso puede comprender transformar el valor absoluto en una métrica/escala relativa que puede ser más fácil de graficar (por ejemplo, multiplicar la frecuencia de alelos mutantes por 100 y tomar log de ese valor) y determinar un factor de escala global utilizando el valor máximo observado. Luego, el proceso implica la visualización de información del primer conjunto de datos de prueba (44). La visualización puede comprender representar gráficamente la información en una interfaz de usuario (por ejemplo, una pantalla de computadora) o en forma tangible (por ejemplo, en una hoja de papel). Para cada alteración que no sea de CNV, el proceso puede comprender multiplicar el factor de escala por un valor transformado para cada gen y usarlo como indicador de cantidad para graficar esa variante, y luego asignar un color/indicador visual único para cada alteración. Luego, el proceso puede comprender la visualización de información para los puntos de prueba posteriores (45) utilizando el siguiente pseudocódigo:

- 30 Si la composición de los resultados de la prueba no ha cambiado, continuar visualizando la fecha del panel anterior en el nuevo panel  
35 Si las alteraciones siguen siendo las mismas, pero las cantidades han cambiado  
40 Vuelva a calcular el indicador de cantidad para trazar esa variante y volver a trazar todos los valores actualizados en los paneles existentes y el nuevo panel para la última fecha de prueba.  
45 Si se agregan nuevas alteraciones
- 45 Agregue las alteraciones a la parte superior de todas las alteraciones existentes  
50 Calcule los valores de transformación  
55 Vuelva a calcular el factor de escala  
Vuelva a dibujar el mapa de respuesta, volviendo a trazar las alteraciones en la fecha de prueba anterior que aún se detectan en la fecha de prueba actual así como las nuevas alteraciones emergentes
- 55 Si la alteración existente anterior no se encuentra entre el conjunto de alteraciones detectadas  
Utilice una altura de cero y grafique la cantidad de alteración para todas las fechas de prueba subsiguientes  
Incluir aún el color es un conjunto de colores no disponibles.
- 60 Cada panel posterior que denota una fecha de prueba también puede incluir información adicional del paciente o la intervención que puede correlacionarse con los cambios de alteración que se ven en el resto del mapa. También se pueden implementar escalas, gráficos y transformaciones similares en CNV y otros tipos de alteraciones del ADN (por ejemplo, metilación) para mostrar estas cantidades en gráficos separados o combinados. Estas anotaciones adicionales también pueden ser cuantificables y trazarse de manera similar en el mapa.
- 65

[0126] El proceso puede entonces comprender determinar un resumen de las alteraciones y las opciones de tratamiento (46). Para la alteración con las frecuencias máximas de los alelos mutantes, se realizan las siguientes acciones:

- 5        Informar todas las alteraciones para ese gen en orden decreciente de frecuencia de alelos mutantes de las alteraciones que no son de CNV  
Informar todas las alteraciones de CNV para ese gen en orden decreciente de valor de CNV  
Repetir para el siguiente gen con la siguiente frecuencia de alelo mutante no CNV más alta que aún no se ha informado.
- 10      Para cada alteración informada, el proceso puede comprender la inclusión de un indicador de tendencia para esa alteración en los diferentes puntos de fecha de prueba.

La agrupación de máximas frecuencias de los alelos mutantes también se puede extender más allá de los genes que se albergaron en mayores anotaciones de encapsulación tales como vías biológicas, nivel de evidencia, etc.

15      Las FIGS. 10D-10I muestran un informe ejemplar generado por el sistema de la FIG. 9A. En la FIG. 10D, una sección 52 de identificación del paciente proporciona información del paciente, fecha de informe e información de contacto del médico. Un mapa 54 de respuesta tumoral incluye un diagrama de flujo 56 modificado que muestra las actividades tumorales con colores únicos para cada gen mutante. El gráfico 56 tiene un cuadro de texto de explicación resumida adjunto 58. Se proporcionan más detalles en la sección 60 de un resumen de las alteraciones y opciones de tratamiento. Las alteraciones 62 y 64 se presentan en la sección 60, junto con la tendencia de mutación, la frecuencia del alelo mutante, la amplificación sin células, Indicación de medicamentos aprobados por la FDA, medicamentos aprobados por la FDA con otras indicaciones e información sobre ensayos clínicos de medicamentos. Figura 10D-1, 10D-2 y 10D-3 proporcionan vistas ampliadas de la Figura 10D.

20      La FIG. 10E muestra una sección de informe ejemplar que proporciona definiciones, comentarios e interpretación de las pruebas. FIGS. 10E-1 y 10E-2 proporcionan vistas ampliadas de la FIG. 10E. La FIG. 10F muestra una parte de resultado de terapia detallada ejemplar del informe. Las FIGS. 10F-1 y 10F-2 proporcionan vistas ampliadas de la FIG. 10F. La FIG. 10G muestra una discusión ejemplar de la relevancia clínica de las alteraciones detectadas. Las FIGS. 10G-1 y 10G-2 proporcionan vistas ampliadas de la FIG. 10G. La FIG. 10H muestra los medicamentos potencialmente disponibles que están pasando por ensayos clínicos. La FIG. 10I muestra los métodos de prueba y las limitaciones de los mismos. Las FIGS. 10I-1 y 10I-2 proporcionan vistas ampliadas de la FIG. 10I.

25      La FIG. 10J-10P muestra varios ejemplos de gráficos de flujo modificados 56. Un gráfico de flujo, o gráfico de flujo, es un tipo de gráfico de áreas apiladas que se desplaza alrededor de un eje central, dando como resultado una forma orgánica fluida. Los gráficos de flujo son una generalización de los gráficos de áreas apiladas donde la línea de base es gratuita. Al cambiar la línea de base, es posible minimizar el cambio en pendiente (o "meneo") en series individuales, lo que facilita la percepción del grosor de cualquier capa dada a través de los datos.

30      Por ejemplo, programas de la FIG. 10J muestran siete capas que representan al menos 8 mutantes más de tres períodos de tiempo, y un punto de tiempo "0" (todos los valores "0"). La FIG. 10K muestra un solo mutante durante 4 períodos de tiempo. No se detectan mutantes en el segundo, tercer y cuarto punto de tiempo. La FIG. 10L indica la frecuencia del alelo dominante en cada momento. La FIG. 10M muestra un único punto de tiempo con un total de cuatro mutantes en dos genes. Los mutantes se identifican por aminoácidos en una posición cambiada (es decir, EGFR T790M).

35      Puede renderizarse un gráfico de flujo para que no sea reflectante del eje x. El gráfico modificado aplica una escala única para denotar atributos proporcionales. El gráfico puede indicar la adición de nuevos atributos a lo largo del tiempo. La presencia o ausencia de una mutación puede reflejarse en forma gráfica, indicando varias posiciones en el genoma y un correspondiente aumento o disminución o mantenimiento de una frecuencia de mutación en cada posición respectiva. Además, pueden usarse mutaciones para informar una puntuación porcentual que indique cuánto material patológico existe en la muestra de polinucleótido libre de células. Una puntuación de confianza puede acompañar a cada mutación detectada, dadas las estadísticas conocidas de variaciones típicas en posiciones informadas en secuencias de referencia que no son enfermedades. Las mutaciones también pueden clasificarse en orden de abundancia en el sujeto o clasificarse por importancia clínicamente procesable.

40      El mapeo de las posiciones del genoma y la variación del número de copias para el sujeto con cáncer puede indicar que un cáncer particular es agresivo y resistente al tratamiento. El sujeto puede ser monitoreado por un período y reevaluado. Si al final del período, el perfil de variación del número de copias, por ejemplo, como se muestra en un mapa de respuesta tumoral, comienza a aumentar dramáticamente, esto puede indicar que el tratamiento actual no está funcionando. También se puede hacer una comparación con perfiles genéticos de otros sujetos. Por ejemplo, si se determina que este aumento en la variación del número de copias indica que el cáncer está avanzando, entonces el régimen de tratamiento original prescrito ya no trata el cáncer y se prescribe un nuevo tratamiento.

Estos informes se pueden enviar y acceder a ellos electrónicamente a través de Internet. El análisis de los datos de la secuencia puede ocurrir en un sitio que no sea la ubicación del sujeto. El informe se puede generar y transmitir a la ubicación del sujeto. A través de una computadora con acceso a Internet, el sujeto puede acceder a los informes que reflejan su carga tumoral.

A continuación, se describen detalles de los procesos ejemplares de prueba de genes. Pasando ahora a la FIG. 11A, un proceso ejemplar recibe materiales genéticos de una muestra de sangre u otras muestras corporales (1102). El proceso puede comprender convertir los polinucleótidos de los materiales genéticos en nucleótidos parentales marcados (1104). Los nucleótidos parentales marcados se amplifican para producir polinucleótidos de la progenie amplificados (1106). Un subconjunto de los polinucleótidos amplificados se secuencia para producir lecturas de secuencia (1108), que se agrupan en familias, cada una generada a partir de un nucleótido original marcado único (1110). En un locus seleccionado, el proceso puede comprender asignar a cada familia una puntuación de confianza para cada familia (1112). A continuación, se determina un consenso utilizando lecturas anteriores. Esto se hace revisando la puntuación de confianza previa para cada familia y, si existen puntuaciones de confianza previas consistentes, se aumenta la puntuación de confianza actual (1114). Si hay puntuaciones de confianza previas, pero son inconsistentes, la puntuación de confianza actual no se modifica en una realización (1116). La puntuación de confianza se ajusta de una manera predeterminada para las puntuaciones de confianza previas inconsistentes. Si es la primera vez que se detecta la familia, se puede reducir la puntuación de confianza actual, ya que puede ser una lectura falsa (1118). El proceso puede comprender inferir la frecuencia de la familia en el locus en el conjunto de polinucleótidos parentales marcados basándose en la puntuación de confianza. Luego, se generan informes de pruebas genéticas como se discutió anteriormente (1120).

Mientras que la información temporal se ha utilizado en las FIGS. 11A-11B para mejorar la información para la detección de mutaciones o variaciones en el número de copias, se pueden aplicar otros métodos de consenso. La comparación histórica se puede usar junto con otras secuencias de consenso que mapean una secuencia de referencia particular para detectar casos de variación genética. El mapeo de secuencias de consenso a secuencias de referencia particulares se puede medir y normalizar frente a muestras de control. Las medidas de las moléculas que mapean las secuencias de referencia se pueden comparar en un genoma para identificar áreas en el genoma en las que el número de copias varía o se pierde la heterocigosidad. Los métodos de consenso incluyen, por ejemplo, métodos lineales o no lineales para construir secuencias de consenso (p. ej., votación, promediado, estadístico, detección de máxima probabilidad a posteriori o máxima, programación dinámica, métodos Bayesanos, Markov ocultos o de máquinas de vectores de soporte, etc.) derivado de la teoría de la comunicación digital, la teoría de la información o la bioinformática. Una vez que se ha determinado la cobertura de lectura de la secuencia, se aplica un algoritmo de modelado estocástico para convertir la cobertura de lectura de la secuencia de ácido nucleico normalizada para cada región de la ventana en los estados del número de copias discretas. En algunos casos, este algoritmo puede comprender uno o más de los siguientes: modelo de Markov oculto, programación dinámica, máquina de vectores de soporte, red bayesiana, decodificación de trellis, decodificación de Viterbi, maximización de expectativas, metodologías de filtrado de Kalman y redes neuronales.

Como se muestra en la FIG. 11B, una comparación de la cobertura de la secuencia con una muestra de control o una secuencia de referencia puede ayudar en la normalización entre ventanas. Aquí, los ADN libres de células se extraen y aíslan de un fluido corporal fácilmente accesible, como sangre, sudor, saliva, orina, etc. Por ejemplo, los ADN libres de células se pueden extraer usando una variedad de métodos conocidos en la técnica, que incluyen pero sin limitarse a precipitación con isopropanol y/o purificación basada en sílice. Los ADN libres de células se pueden extraer de cualquier número de sujetos, tales como sujetos sin cáncer, sujetos con riesgo de cáncer o sujetos que se sabe que tienen cáncer (por ejemplo, por otros medios).

Después de la etapa de aislamiento/extracción, cualquiera de un número de diferentes operaciones de secuenciación se puede realizar en la muestra de polinucleótido libre de células. Las muestras se pueden procesar antes de secuenciarlas con uno o más reactivos (por ejemplo, enzimas, identificadores únicos (por ejemplo, códigos de barras), sondas, etc.). En algunos casos, si la muestra se procesa con un identificador único, como un código de barras, las muestras o fragmentos de muestras pueden etiquetarse individualmente o en subgrupos con el identificador único. La muestra etiquetada se puede usar luego en una aplicación posterior, como una reacción de secuenciación, y las moléculas individuales se pueden rastrear hasta las moléculas originales.

Los polinucleótidos libres de células pueden marcarse o rastrearse para permitir la posterior identificación y origen del polinucleótido particular. La asignación de un identificador a un individuo o subgrupos de polinucleótidos puede permitir que se asigne una identidad única a secuencias individuales o fragmentos de secuencias. Esto puede permitir la adquisición de datos de muestras individuales y no se limita a promedios de muestras. Los ácidos nucleicos u otras moléculas derivadas de una sola hebra pueden compartir una etiqueta o identificador común y, por lo tanto, pueden identificarse posteriormente como derivados de esa hebra. De manera similar, todos los fragmentos de una sola hebra de ácido nucleico pueden etiquetarse con el mismo identificador o etiqueta, permitiendo así la identificación posterior de fragmentos de la hebra parental. En otros casos, los productos de expresión génica (p. ej., ARNm) pueden marcarse para cuantificar la expresión. Se puede contar un código de barras o un código de barras en combinación con la secuencia a la que se adjunta. En otros casos más, los sistemas y métodos pueden usarse como control de

amplificación por RCP. En tales casos, múltiples productos de amplificación de una reacción de RCP pueden etiquetarse con la misma etiqueta o identificador. Si los productos se secuencian posteriormente y muestran diferencias de secuencia, las diferencias entre productos con el mismo identificador pueden atribuirse a un error de RCP. Además, las secuencias individuales pueden identificarse basándose en las características de los datos de secuencia para la propia lectura. Por ejemplo, la detección de datos de secuencia única al principio (inicio) y final (final) de las lecturas de secuenciación individuales se puede utilizar, solo o en combinación, con la longitud o el número de pares de bases de cada lectura de secuencia para asignar identidades únicas a moléculas individuales. Los fragmentos de una sola hebra de ácido nucleico, a los que se les ha asignado una identidad única, pueden permitir así la identificación posterior de fragmentos de la hebra parental. Esto se puede utilizar junto con el cuello de botella del material genético de partida inicial para limitar la diversidad.

Además, el uso de datos de secuencias únicas en las porciones de comienzo (inicio) y final (parada) de lecturas de secuenciación individual y la longitud de lectura de secuenciación se pueden usar, solos o combinados, con el uso de códigos de barras. En algunos casos, los códigos de barras pueden ser únicos. En otros casos, los códigos de barras en sí mismos pueden no ser únicos. En este caso, el uso de códigos de barras no únicos, en combinación con datos de secuencia al principio (inicio) y parada (final) de las lecturas de secuenciación individuales y la longitud de lectura de secuenciación puede permitir la asignación de una identidad única a secuencias individuales. De manera similar, los fragmentos de una sola hebra de ácido nucleico a los que se les ha asignado una identidad única pueden permitir así la identificación posterior de los fragmentos de la hebra parental.

En general, pueden prepararse secuencias de polinucleótidos libres de células para una reacción de secuenciación de aplicación corriente abajo. A menudo, un método de secuenciación es la secuenciación clásica de Sanger. Los métodos de secuenciación pueden incluir, pero no se limitan a: secuenciación de alto rendimiento, pirosecuenciación, secuenciación por síntesis, secuenciación de una sola molécula, secuenciación de nanoporos, secuenciación de semiconductores, secuenciación por ligación, secuenciación por hibridación, ARN-Seq (Illumina), Expresión génica digital (Helicos), secuenciación de próxima generación, secuenciación de molécula única por síntesis (SMSS) (Helicos), secuenciación masivamente paralela, matriz de molécula única clonal (Solexa), secuenciación rápida, secuenciación de Maxim-Gilbert, marcha de cebadores y cualquier otro método de secuenciación conocido en la técnica.

Los métodos de secuenciación implican típicamente la preparación de la muestra, la secuenciación de polinucleótidos en la muestra preparada para producir lecturas de secuencia y la manipulación bioinformática de las lecturas de secuencia para producir información genética cuantitativa y/o cualitativa sobre la muestra. La preparación de muestras típicamente implica convertir polinucleótidos en una muestra en una forma compatible con la plataforma de secuenciación utilizada. Esta conversión puede implicar marcar polinucleótidos. Las etiquetas pueden comprender etiquetas de secuencia de polinucleótidos. Las metodologías de conversión utilizadas en la secuenciación pueden no ser 100% eficientes. Por ejemplo, no es raro convertir polinucleótidos en una muestra con una eficiencia de conversión de aproximadamente 1-5%, es decir, aproximadamente 1-5% de los polinucleótidos en una muestra se convierten en polinucleótidos marcados. Los polinucleótidos que no se convierten en moléculas etiquetadas no se representan en una biblioteca etiquetada para secuenciar. Por consiguiente, los polinucleótidos que tienen variantes genéticas representadas con baja frecuencia en el material genético inicial pueden no estar representados en la biblioteca etiquetada y, por lo tanto, pueden no ser secuenciados o detectados. Al aumentar la eficiencia de conversión, aumenta la probabilidad de que un polinucleótido en el material genético inicial esté representado en la biblioteca etiquetada y, en consecuencia, se detecte mediante secuenciación. Además, en lugar de abordar directamente el problema de la baja eficiencia de conversión de la preparación de bibliotecas, la mayoría de los protocolos hasta la fecha requieren más de 1 microgramo de ADN como material de entrada. Sin embargo, cuando el material de muestra de entrada es limitado o se desea la detección de polinucleótidos con baja representación, una alta eficiencia de conversión puede secuenciar eficientemente la muestra y/o detectar adecuadamente tales polinucleótidos.

Generalmente, la detección de mutaciones se puede realizar en regiones enriquecidas selectivamente del genoma o transcriptoma purificado y aislado (1302). Las regiones específicas, que pueden incluir pero no se limitan a genes, oncogenes, genes supresores de tumores, promotores, elementos de secuencia reguladora, regiones no codificantes, miARN, snARN y similares, pueden amplificarse selectivamente de una población total de polinucleótidos libres de células. Esto se puede realizar como se describe en este documento. Por ejemplo, puede usarse secuenciación multiplex, con o sin etiquetas de código de barras para secuencias de polinucleótidos individuales. En otros ejemplos, la secuenciación se puede realizar usando cualquier plataforma de secuenciación de ácidos nucleicos conocida en la técnica. Este paso genera una pluralidad de lecturas de secuencias de fragmentos genómicos (1304). Además, se obtiene una secuencia de referencia de una muestra de control, tomada de otro sujeto. En algunos casos, el sujeto de control puede ser un sujeto que se sabe que no tiene aberraciones o enfermedades genéticas conocidas. En algunos casos, estas lecturas de secuencia pueden contener información de códigos de barras. En otros ejemplos, no se utilizan códigos de barras.

Después de la secuenciación, se puede asignar una puntuación de calidad a las lecturas. Una puntuación de calidad puede ser una representación de lecturas que indica si esas lecturas pueden ser útiles en análisis posteriores basados en un umbral. En algunos casos, algunas lecturas no son de suficiente calidad o longitud para realizar el

siguiente paso de mapeo. Las lecturas de secuenciación con una puntuación de calidad de al menos el 90%, 95%, 99%, 99,9%, 99,99% o 99,999% pueden filtrarse fuera del conjunto de datos. En otros casos, las lecturas de secuenciación asignadas a una calidad puntuada al menos 90%, 95%, 99%, 99,9%, 99,99% o 99,999% se pueden filtrar fuera del conjunto de datos. En el paso 1306, las lecturas del fragmento genómico que cumplen con un umbral

5 de puntuación de calidad especificado se mapean en un genoma de referencia, o una secuencia de referencia que se sabe que no contiene mutaciones. Después de mapear la alineación, las lecturas de secuencia reciben una puntuación de mapeo. Una puntuación de mapeo puede ser una representación o lecturas mapeadas de nuevo a la secuencia de referencia que indica si cada posición es o no mapeable de forma única. Las lecturas pueden ser secuencias no relacionadas con el análisis de mutaciones. Por ejemplo, algunas lecturas de secuencia pueden originarse a partir de polinucleótidos contaminantes. Las lecturas de secuenciación con una puntuación de mapeo de al menos 90%, 95%, 99%, 99,9%, 99,99% o 99,999% pueden filtrarse del conjunto de datos. En otros casos, las lecturas de secuenciación asignadas una puntuación de mapeo por debajo del 90%, 95%, 99%, 99,9%, 99,99% o 99,999% se pueden filtrar fuera del conjunto de datos.

10 15 Para cada base mapeable, las bases que no cumplen el umbral mínimo de mapeo, o bases de baja calidad, pueden ser reemplazadas por las bases correspondientes que se encuentran en la secuencia de referencia.

20 La frecuencia de bases variante puede ser calculada como el número de lecturas que contiene la variante dividida por el número total de lecturas 1308 después de comprobar la cobertura de lecturas e identificar bases de variantes de la identificación con respecto a la secuencia de control en cada lectura. Esto puede expresarse como una proporción para cada posición mapeable en el genoma.

25 30 Para cada posición de base, las frecuencias de los cuatro nucleótidos, citosina, guanina, timina, adenina pueden analizarse en comparación con la secuencia de referencia. Se puede aplicar un algoritmo de modelado estocástico o estadístico para convertir las relaciones normalizadas para cada posición mapeable para reflejar los estados de frecuencia para cada variante base. En algunos casos, este algoritmo puede comprender uno o más de los siguientes: modelo de Markov oculto, programación dinámica, máquina de vectores de soporte, modelado bayesiano o probabilístico, decodificación de trellis, decodificación de Viterbi, maximización de expectativas, metodologías de filtrado de Kalman y redes neuronales.

35 40 Los estados de mutación discretos de cada posición de la base se pueden utilizar para identificar una variante de base con alta frecuencia de la varianza en comparación con la línea base de la secuencia de referencia. En algunos casos, la línea de base puede representar una frecuencia de al menos 0,0001%, 0,001%, 0,01%, 0,1%, 1,0%, 2,0%, 3,0%, 4,0% 5,0%, 10% o 25%. En otros casos, la línea de base podría representar una frecuencia de al menos 0,0001%, 0,001%, 0,01%, 0,1%, 1,0%, 2,0%, 3,0%, 4,0% 5,0%, 10% o 25%. En algunos casos, todas las posiciones de bases adyacentes con la variante de base o la mutación pueden fusionarse en un segmento para informar la presencia o ausencia de una mutación. En algunos casos, se pueden filtrar varias posiciones antes de fusionarlas con otros segmentos.

45 50 Después del cálculo de las frecuencias de varianza para cada posición de base, la variante con una desviación más grande para una posición específica en la secuencia derivada del sujeto en comparación con la secuencia de referencia puede ser identificada como una mutación. En algunos casos, una mutación puede ser una mutación cancerosa. En otros casos, una mutación puede estar correlacionada con un estado de enfermedad.

55 60 [0150] Una mutación o variante pueden comprender una aberración genética que incluye, pero no se limita a una única base de sustitución, o pequeños indeles, transversiones, translocaciones, inversión, delecciones, truncamientos o truncamientos de genes. En algunos casos, una mutación puede tener como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 o 20 nucleótidos de longitud. En otros casos, una mutación puede tener al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 o 20 nucleótidos de longitud.

65 70 A continuación, un consenso se determina usando las lecturas previas. Esto se hace revisando el puntaje de confianza previo para las bases correspondientes, y si existen puntajes de confianza previos consistentes, entonces el puntaje de confianza actual aumenta (1314). Si hay puntuaciones de confianza previas, pero son inconsistentes, la puntuación de confianza actual no se modifica (1316). En otros casos, la puntuación de confianza se ajusta de una manera predeterminada para las puntuaciones de confianza previas inconsistentes. Si es la primera vez que se detecta la familia, la puntuación de confianza actual puede reducirse ya que puede ser una lectura falsa (1318). El proceso puede comprender luego convertir la frecuencia de variación por cada base en estados variantes discretos para cada posición base (1320).

75 80 Pueden detectarse numerosos cánceres. Las células cancerosas, como la mayoría de las células, pueden caracterizarse por una tasa de renovación, en donde las células viejas mueren y son reemplazadas por células más nuevas. Generalmente, las células muertas, en contacto con la vasculatura de un sujeto determinado, pueden liberar ADN o fragmentos de ADN al torrente sanguíneo. Esto también se aplica a las células cancerosas durante varias etapas de la enfermedad. Las células cancerosas también pueden caracterizarse, dependiendo de la etapa de la enfermedad, por diversas aberraciones genéticas, como la variación del número de copias y las mutaciones. Este

fenómeno puede usarse para detectar la presencia o ausencia de cánceres en individuos usando los métodos y sistemas descritos en este documento.

5 Por ejemplo, la sangre de los sujetos con riesgo de cáncer puede ser extraída y preparada para generar una población de polinucleótidos libres de células. Pueden detectarse mutaciones o variaciones en el número de copias que pueden existir en ciertos cánceres presentes. Puede detectarse la presencia de células cancerosas en el cuerpo, a pesar de la ausencia de síntomas u otras características de la enfermedad.

10 Los tipos y número de cánceres que se pueden detectar pueden incluir, pero no se limitan a cánceres de la sangre, cánceres del cerebro, cánceres de pulmón, cáncer de piel, cáncer de la nariz, cánceres de garganta, cáncer de hígado, cánceres de hueso, linfomas, cánceres pancreáticos, cánceres de la piel, cánceres del intestino, cánceres de recto, cánceres de tiroides, cánceres de vejiga, cánceres de riñón, cánceres de boca, cánceres de estómago, tumores de estado sólido, tumores heterogéneos, tumores homogéneos y similares.

15 Puede detectarse cualquier número de aberraciones genéticas que pueden causar o resultar de cánceres. Estos pueden incluir, entre otros, mutaciones, mutaciones, indeles, variaciones en el número de copias, transversiones, translocaciones, inversión, delecciones, aneuploidía, aneuploidía parcial, poliploidía, inestabilidad cromosómica, alteraciones de la estructura cromosómica, fusiones de genes, fusiones de cromosomas, truncamientos de genes, amplificación, duplicaciones de genes, lesiones cromosómicas, lesiones del ADN, cambios anormales en las modificaciones químicas del ácido nucleico, cambios anormales en los patrones epigenéticos, cambios anormales en la infección por metilación del ácido nucleico y cáncer.

20 Además, pueden caracterizarse ciertos tipos de cáncer. Los datos genéticos pueden permitir a los médicos ayudar a caracterizar mejor una forma específica de cáncer. A menudo, los cánceres son heterogéneos tanto en composición como en estadios. Los datos del perfil genético pueden permitir la caracterización de subtipos específicos de cáncer que pueden ser importantes en el diagnóstico o tratamiento de ese subtipo específico. Esta información también puede proporcionar pistas al sujeto o al médico con respecto al pronóstico de un tipo específico de cáncer.

25 30 Pueden monitorizarse cánceres ya conocidos, u otras enfermedades en un sujeto particular. Esto puede permitir que el sujeto o el médico adapten las opciones de tratamiento de acuerdo con el progreso de la enfermedad. Pueden construirse perfiles genéticos de un sujeto particular del curso de la enfermedad. Los cánceres pueden progresar, volviéndose más agresivos y genéticamente inestables. En otros ejemplos, los cánceres pueden permanecer benignos, inactivos o latentes. Puede determinarse la progresión de la enfermedad.

35 40 Además, puede determinarse de la eficacia de una determinada opción de tratamiento. En un ejemplo, las opciones de tratamiento exitosas pueden aumentar la cantidad de variación en el número de copias o mutaciones detectadas en la sangre del sujeto si el tratamiento es exitoso, ya que más cánceres pueden morir y perder ADN. En otros ejemplos, esto puede no ocurrir. En otro ejemplo, quizás ciertas opciones de tratamiento pueden estar correlacionadas con perfiles genéticos de cánceres a lo largo del tiempo. Esta correlación puede resultar útil para seleccionar una terapia. Además, si se observa que un cáncer está en remisión después del tratamiento, puede monitorizarse la enfermedad residual o la recurrencia de la enfermedad.

45 50 |||Varias otras enfermedades e infecciones pueden dar lugar a otros tipos de afecciones que pueden ser adecuadas para la detección y el seguimiento tempranos. Por ejemplo, en determinados casos, los trastornos genéticos o las enfermedades infecciosas pueden provocar un cierto mosaicismo genético dentro de un sujeto. Este mosaicismo genético puede causar variaciones en el número de copias y mutaciones que podrían observarse. También pueden monitorizarse los genomas de las células inmunes dentro del cuerpo. Las células inmunes, como las células B, pueden experimentar una rápida expansión clonal ante la presencia de ciertas enfermedades. Las expansiones clonales pueden monitorearse usando la detección de variación del número de copias y pueden monitorearse ciertos estados inmunes. En este ejemplo, se puede realizar un análisis de variación del número de copias a lo largo del tiempo para producir un perfil de cómo puede estar progresando una enfermedad en particular.

55 60 Además, también pueden monitorizarse infecciones sistémicas, como pueden causarse por un patógeno tal como una bacteria o virus. Se puede utilizar la variación del número de copias o incluso la detección de mutaciones para determinar cómo está cambiando una población de patógenos durante el curso de la infección. Esto puede ser particularmente importante durante infecciones crónicas, como las infecciones por VIH/SIDA o hepatitis, en las que los virus pueden cambiar el estado del ciclo de vida y/o mutar en formas más virulentas durante el curso de la infección.

65 También puede monitorizarse a los sujetos con trasplantes. Generalmente, el tejido transplantado sufre un cierto grado de rechazo por parte del cuerpo tras el trasplante. Pueden determinarse o realizarse perfiles de las actividades de rechazo del cuerpo huésped, cuando las células inmunes intentan destruir el tejido transplantado. Esto puede ser útil para controlar el estado del tejido transplantado, así como para alterar el curso del tratamiento o la prevención del rechazo.

65 Además, puede caracterizarse la heterogeneidad de una condición anormal en un sujeto, comprendiendo la

- generación de un perfil genético de polinucleótidos extracelulares en el sujeto, en donde el perfil genético comprende una pluralidad de datos que resultan de análisis de variación y mutación del número de copias. En algunos casos, incluidos, entre otros, el cáncer, una enfermedad puede ser heterogénea. Las células de la enfermedad pueden no ser idénticas. En el ejemplo del cáncer, se sabe que algunos tumores comprenden diferentes tipos de células tumorales, algunas células en diferentes etapas del cáncer. En otros ejemplos, la heterogeneidad puede comprender múltiples focos de enfermedad. Nuevamente, en el ejemplo del cáncer, puede haber múltiples focos tumorales, quizás donde uno o más focos son el resultado de metástasis que se han diseminado desde un sitio primario.
- Puede generarse un perfil, huella o conjunto de datos que es una suma de la información genética procedente de diferentes células en una enfermedad heterogénea. Este conjunto de datos puede comprender la variación del número de copias y los análisis de mutaciones solos o en combinación.
- Además, pueden diagnosticarse, pronosticarse, monitorizarse u observarse cánceres u otras enfermedades de origen fetal. Es decir, estas metodologías pueden emplearse en una paciente embarazada para diagnosticar, pronosticar, monitorear u observar cánceres u otras enfermedades en un sujeto por nacer cuyo ADN y otros polinucleótidos pueden co-circular con moléculas maternas.
- Además, estos informes se envían y se accede a ellos electrónicamente a través de Internet. El análisis de los datos de la secuencia se produce en un sitio diferente al del sujeto. El informe se genera y se transmite a la ubicación del sujeto. A través de una computadora con acceso a Internet, el sujeto accede a los informes que reflejan su carga tumoral.
- La información anotada puede ser utilizada por un proveedor de atención médica para seleccionar otras opciones de tratamiento con medicamentos y/o proporcionar información sobre opciones de tratamiento con medicamentos a una compañía de seguros. Pueden anotarse las opciones de tratamiento farmacológico para una afección en, por ejemplo, las NCCN Clinical Practice Guidelines in Oncology™ o las pautas de práctica clínica de la Sociedad Estadounidense de Oncología Clínica (ASCO).
- Las opciones de tratamiento de drogas que se estratifican en un informe se pueden anotar en el informe haciendo una lista de opciones de tratamiento de fármacos adicionales. Un tratamiento farmacológico adicional puede ser un medicamento aprobado por la FDA para un uso no indicado en la etiqueta. Una disposición de la 1993 Omnibus Budget Reconciliation Act (OBRA) requiere que Medicare cubra los usos no aprobados de medicamentos contra el cáncer que están incluidos en los compendios médicos estándar. Los medicamentos utilizados para anotar listas se pueden encontrar en los compendios aprobados por CMS, incluido el National Comprehensive Cancer Network (NCCN) Drugs and Biologics Compendium™, Thomson Micromedex DrugDex®, el compendio de Elsevier Gold Standard's Clinical Pharmacology y el American Hospital Formulary Service-Drug Information Compendium®.
- Las opciones de tratamiento de drogas se pueden anotar enumerando un fármaco experimental que puede ser útil en el tratamiento de un cáncer con uno o más marcadores moleculares de un estado particular. El fármaco experimental puede ser un fármaco para el que se dispone de datos in vitro, datos in vivo, datos de modelos animales, datos de ensayos preclínicos o datos de ensayos clínicos. Los datos se pueden publicar en literatura médica revisada por pares que se encuentra en revistas incluidas en el CMS Medicare Benefit Policy Manual, que incluyen, por ejemplo, American Journal of Medicine, Annals of Internal Medicine, Annals of Oncology, Annals of Surgical Oncology, Biology of Blood y trasplante de médula ósea, sangre, trasplante de médula ósea, British Journal of Cancer, British Journal of Hematology, British Medical Journal, Cancer, Clinical Cancer Research, Drugs, European Journal of Cancer (anteriormente European Journal of Cancer and Clinical Oncology), Gynecologic Oncology, International Journal of Radiation, Oncology, Biology, and Physics, The Journal of the American Medical Association, Journal of Clinical Oncology, Journal of the National Cancer Institute, Journal of the National Comprehensive Cancer Network (NCCN), Journal of Urology, Lancet, Lancet Oncology, Leukemia, The New England Journal of Medicine, y Radiation Oncology.
- Las opciones de tratamiento farmacológico se pueden anotar proporcionando un enlace en un informe electrónico que conecte un medicamento listado con información científica sobre el medicamento. Por ejemplo, se puede proporcionar un enlace a información sobre un ensayo clínico de un medicamento ([clinicaltrials.gov](http://clinicaltrials.gov)). Si el informe se proporciona a través de una computadora o un sitio web informático, el enlace puede ser una nota al pie de página, un hipervínculo a un sitio web, un cuadro emergente o un cuadro desplegable con información, etc. El informe y la información anotada pueden ser proporcionados en un formulario impreso, y las anotaciones pueden ser, por ejemplo, una nota al pie de una referencia.
- La información para anotar una o más opciones de tratamiento farmacológico en un informe puede ser proporcionada por una entidad comercial que almacena información científica. Un proveedor de atención médica puede tratar a un sujeto, como un paciente con cáncer, con un fármaco experimental incluido en la información anotada, y el proveedor de atención médica puede acceder a la opción de tratamiento farmacológico anotado, recuperar la información científica (p. ej., imprimir un artículo de revista médica) y enviarlo (por ejemplo, un artículo de una revista impresa) a una compañía de seguros junto con una solicitud de reembolso por proporcionar el tratamiento para la drogadicción. Los médicos pueden utilizar cualquiera de una variedad de códigos de grupo relacionados con

el diagnóstico (DRG) para permitir el reembolso.

Una opción de tratamiento de drogas en un informe también puede ser anotado con información con respecto a otros componentes moleculares en una vía que afecta un fármaco (por ejemplo, la información sobre un fármaco que se dirige una quinasa corriente abajo de un receptor de superficie celular que es un objetivo de drogas). La opción de tratamiento farmacológico se puede anotar con información sobre medicamentos que se dirigen a uno o más componentes de la ruta molecular. La identificación y/o anotación de información relacionada con las rutas puede ser subcontratada a otra empresa.

La información anotada puede ser, por ejemplo, el nombre de un medicamento (por ejemplo, un medicamento aprobado por la FDA para uso no indicado en la etiqueta; un medicamento que se encuentra en un compendio aprobado por CMS y/o un medicamento descrito en un artículo de una revista científica (médica)), información científica sobre una o más opciones de tratamiento farmacológico, uno o más enlaces a información científica sobre uno o más medicamentos, información sobre ensayos clínicos sobre uno o más medicamentos (por ejemplo, información de Clinicaltrials.gov/), uno o más enlaces a citas de información científica respecto a las drogas, etc.

La información anotada se puede insertar en cualquier lugar en un informe. La información anotada se puede insertar en varias ubicaciones en un informe. Se puede insertar información anotada en un informe cerca de una sección sobre opciones de tratamiento farmacológico estratificado. La información anotada se puede insertar en un informe en una página separada de las opciones estratificadas de tratamiento farmacológico. Un informe que no contenga opciones estratificadas de tratamiento farmacológico se puede anotar con información.

**[0174]** También pueden incluirse informes sobre los efectos de los fármacos sobre la muestra (por ejemplo, células tumorales) aislados de un sujeto (por ejemplo, paciente de cáncer). Puede establecerse un cultivo in vitro usando un tumor de un paciente con cáncer usando técnicas conocidas por los expertos en la técnica. Puede incluirse un cribado de alto rendimiento de fármacos no aprobados por la FDA o fármacos experimentales utilizando dicho cultivo in vitro y/o modelo de xenoinjerto. El sistema también puede incluir la monitorización del antígeno tumoral para la detección de recurrencia.

Puede proporcionarse acceso habilitado por Internet de informes de un sujeto con cáncer. El sistema puede utilizar un secuenciador de ADN portátil o un secuenciador de ADN de escritorio. El secuenciador de ADN es un instrumento científico que se utiliza para automatizar el proceso de secuenciación del ADN. Dado una muestra de ADN, se utiliza un secuenciador de ADN para determinar el orden de las cuatro bases: adenina, guanina, citosina y timina. El orden de las bases de ADN se informa como una cadena de texto, llamada lectura. Algunos secuenciadores de ADN también pueden considerarse instrumentos ópticos, ya que analizan señales de luz que se originan en fluorocromos unidos a nucleótidos.

Los secuenciadores de ADN envían los datos a través de una conexión directa o por Internet a una computadora para su procesamiento. Los aspectos de procesamiento de datos del sistema se pueden implementar en circuitos electrónicos digitales, o en hardware, firmware, software o en combinaciones de ellos. El aparato de procesamiento de datos se puede implementar en un producto de programa informático incorporado de manera tangible en un dispositivo de almacenamiento legible por máquina para su ejecución mediante un procesador programable; y los pasos del método de procesamiento de datos de la invención pueden ser realizados por un procesador programable que ejecuta un programa de instrucciones para realizar funciones de la invención operando con datos de entrada y generando salida. Los datos de procesamiento de aspectos pueden implementarse ventajosamente en uno o más programas de ordenador que son ejecutables en un sistema programable que incluye al menos un procesador programable acoplado para recibir datos e instrucciones desde y a los datos de transmisión y las instrucciones para un sistema de almacenamiento de datos, al menos un dispositivo de entrada y al menos un dispositivo de salida. Cada programa de computadora puede implementarse en un lenguaje de programación orientado a objetos o de procedimiento de alto nivel, o en lenguaje ensamblador o de máquina, si se desea; y, en todo caso, el lenguaje puede ser un lenguaje compilado o interpretado. Los procesadores adecuados incluyen, a modo de ejemplo, microprocesadores de uso general y especial. Generalmente, un procesador recibirá instrucciones y datos de una memoria de solo lectura y/o una memoria de acceso aleatorio. Los dispositivos de almacenamiento adecuados para incorporar de forma tangible instrucciones y datos de programas informáticos incluyen todas las formas de memoria no volátil, incluyendo, a modo de ejemplo, dispositivos de memoria semiconductores, tales como EPROM, EEPROM y dispositivos de memoria flash; discos magnéticos como discos duros internos y discos extraíbles; discos magneto-ópticos; y discos CD-ROM. Cualquiera de los anteriores se puede complementar o incorporar en ASIC (circuitos integrados específicos de la aplicación).

Para proporcionar interacción con un usuario, puede usarse un sistema informático que tiene un dispositivo de visualización como un monitor o pantalla LCD (pantalla de cristal líquido) para la visualización de información al usuario y dispositivos de entrada mediante los cuales el usuario puede proporcionar información al sistema informático, como un teclado, un dispositivo señalador bidimensional como un ratón o una bola de seguimiento, o un dispositivo señalador tridimensional, como un guante de datos o un ratón giroscópico. El sistema informático se puede programar para proporcionar una interfaz gráfica de usuario a través de la cual los programas informáticos interactúan con los

usuarios. El sistema informático se puede programar para proporcionar una interfaz de visualización tridimensional de realidad virtual.

### Intervención terapéutica

5 Los métodos de esta invención permiten que se proporciona intervenciones terapéuticas dirigidas con más precisión a la forma de una enfermedad en un sujeto, y para calibrar estas intervenciones terapéuticas en el tiempo. Esta precisión refleja, en parte, la precisión con la que se puede perfilar el estado tumoral de todo el cuerpo de un sujeto, tal como se refleja en la heterogeneidad del tumor. Por tanto, la intervención terapéutica es más eficaz contra 10 cánceres con este perfil que contra cánceres con cualquiera de estas variantes.

15 Una intervención terapéutica es una intervención que produce un efecto terapéutico (por ejemplo, es terapéuticamente eficaz). Las intervenciones terapéuticamente eficaces previenen, ralentizan la progresión, mejoran el estado de (p. ej., provoca la remisión) o curan una enfermedad, como el cáncer. Una intervención terapéutica puede incluir, por ejemplo, la administración de un tratamiento, como quimioterapia, radioterapia, cirugía, inmunoterapia, administración de un fármaco o nutracéutico, o un cambio de comportamiento, como la dieta. Una medida de la eficacia terapéutica es la eficacia para al menos el 90% de los sujetos que se someten a la intervención en al menos 100 sujetos.

20 Dianas de fármacos en el cáncer y fármacos eficaces contra estas dianas se exponen en las Tablas 1 y 2 (tomadas de Bailey et al., Discovery Medicine, v. 18 nº 92, 2/7/14).

<b>Tabla 1. Ejemplos seleccionados de pruebas de diagnóstico disponibles comercialmente, implicaciones terapéuticas asociadas y tipo de cáncer relevante.</b>				
	<b>Implicaciones de la terapia</b>	<b>Prueba</b>	<b>Tipo de cáncer</b>	<b>Asociación Clínica de Medicamentos y Biomarcadores</b>
<b>Ensayos IHC</b>				
	Cetuximab; Panitumumab	EGFR	CRC	Establecido
30	Imatinib	C-KIT	GIST	Establecido
	Trastuzumab	HER2	Cáncer de mama; Cáncer gástrico	Establecido

(Continuación)

<b>Tabla 1. Ejemplos seleccionados de pruebas de diagnóstico disponibles comercialmente, implicaciones terapéuticas asociadas y tipo de cáncer relevante.</b>				
	<b>Implicaciones de la terapia</b>	<b>Prueba</b>	<b>Tipo de cáncer</b>	<b>Asociación Clínica de Medicamentos y Biomarcadores</b>
<b>Ensayos IHC</b>				
5	Resistencia a inhibidores de PI3K, AKT y MEK	LKB1	NSCLC	En investigación (Mahoney <i>et al.</i> , 2009)
10	Crizotinib	C-MET	NSCLC	En investigación (Sadiq & Salgia, 2013)
15	Inhibidores de Akt/mTOR; resistencia a las terapias anti-EGFR	PTEN	CRC, NSCLC	En investigación (Di Nicolantonio <i>et al.</i> , 2010; Sos <i>et al.</i> , 2009; Wang <i>et al.</i> , 2012)
<b>Ensayos de hibridación <i>in situ</i></b>				
20	Crizotinib	ALK Fusion FISH	NSCLC	Establecido
25	Trastuzumab; Pertuzumab	FISH HER2	Cáncer de mama, cáncer gástrico	Establecido
30	Trastuzumab	HER2 CISH	Cáncer de mama	Establecido
35	Trastuzumab	HER2 ISH	Cáncer de mama	Establecido
<b>Ensayos de mutación</b>				
40	Cetuximab, Panitumumab	KRAS	CRC, NSCLC, cáncer de páncreas	Establecido
45	Erlotinib, Gefitinib	EGFR	NSCLC, CCR	Establecido
50	Vemurafenib, Trametenib, Dabrafenib, Resistencia a terapias anti EGFR	BRAF	CRC, cáncer de tiroides, melanoma	Establecido
55	Imatinib; TKI de segunda generación	BCR-ABL	CML, Ph + LMA	Establecido
60	Crizotinib	ALK	NSCLC	Establecido
	Inhibidores de RAF y MEK, resistencia a anti-EGFR	NRAS	Melanoma, CCR, NSCLC	En investigación (Ascierto <i>et al.</i> , 2013; De Mattos-Arruda et al., 2011; De Roock <i>et al.</i> , 2010; Huang <i>et al.</i> , 2013)
	terapias			
	Imatinib	PDGFR A	ESENCIA	Establecido
	Inhibidores de PI3K/mTOR	PIK3CA	Cáncer de mama, CCR, cáncer de pulmón	En investigación (Di Nicolantonio <i>et al.</i> , 2010; Janku <i>et al.</i> , 2013)
	Inhibidores de Akt/mTOR; resistencia a las terapias anti-EGFR	PTEN	CRC, NSCLC, mama	En investigación (Di Nicolantonio <i>et al.</i> , 2010, Jerusalem <i>et al.</i> , 2013; Sos <i>et al.</i> , 2009; Wang <i>et al.</i> , 2012)
	Resistencia a inhibidores de PI3K, AKT y MEK	LKB1	NSCLC	Investigación (Averette-Byers <i>et al.</i> , 2012)

(Continuación)

<b>Tabla 1. Ejemplos seleccionados de pruebas de diagnóstico disponibles comercialmente, implicaciones terapéuticas asociadas y tipo de cáncer relevante.</b>				
5	Implicaciones de la terapia	Prueba	Tipo de cáncer	Asociación Clínica de Medicamentos y Biomarcadores
Otro				
10	Imatinib	Análisis cuantitativo de transcripciones BCR-ABL1	CML, Ph + LMA	Establecido
15	Resistencia a Imatinib	Número de copia BCR-ABL1	CML, Ph + LMA	Investigación (Hochhaus et al., 2002)
20	Inhibidores de PI3K Erlotinib; Gefitinib; Cetuximab; Panitumumab	Amplificación PIK3CA Amplificación EGFR	Múltiples tipos de cáncer NSCLC, CCR	En investigación (Rodon et al., 2013) En investigación (Gupta et al., 2009)
<p><i>Nota:</i> Las asociaciones clínicas de medicamentos y biomarcadores denominadas "Establecidas" reflejan indicaciones de medicamentos bien conocidas de la FDA. Las denominadas "En investigación" son asociaciones hipotetizadas y demostradas por la literatura científica.</p>				

<b>Tabla 2. Indicaciones y terapias de T diana aprobadas por la FDA de EE.UU.</b>				
Agente	Nombre comercial	Diana(s)	Indicación(es) aprobadas por la FDA	Empresa
<b>Anticuerpos monoclonicos</b>				
Ado-trastuzumab emtansina (T-DM1)*	Kadcyla	HER2	Cáncer de mama (HER2+)*	Genentech
Bevacizumab	Avastin	VEGF	CRC	Genentech
			GBM	
			NCLC	
			RCC	
Cetuximab*	Erbitux	EGFR	CRC (tipo salvaje KRAS) *	Eli Lilly
			HNSCC	
Ipilimumab	Yervoy	CTLA-4	Melanoma	Bristol-Myers Squibb
Obinutuzumab	Gazyva	CD-20	CLL	Genentech
Panitumumab *	Vectibix	EGFR	CRC (tipo salvaje KRAS)*	Amgen
Pertuzumab	Perjeta	HER2	Cáncer de mama (HER2+)*	Genentech
Trastuzumab*	Herceptin	HER2	Cáncer de mama (HER2+)*	Genentech
			Cáncer gástrico (HER2+)*	
<b>Inhibidores de moléculas pequeñas</b>				
Afatinib *	Gilotrif	EGFR, HER2	NSCLC (con delecciones del exón 19 de EGFR o sustitución de L858R)*	Boehringer Ingelheim
Axitinib	Inlyta	KIT, PDGFR $\beta$ , VEGFR <sub>1/2/3</sub>	RCC	Pfizer
Bosutinib *	Bosulif	ABL	LMC (cromosoma Filadelfia positivo)*	Pfizer
Cabozantinib	Cometriq	FLT3, KIT, MET, RET, VEGFR <sub>2</sub>	Cáncer de tiroides medular	Exelixis
Crizotinib *	Xalkori	ALK, MET	NSCLC (con fusión ALK)*	Pfizer
Dabrafenib *	Tafinlar	BRAF	Melanoma (con mutación BRAF V600E)*	GlaxoSmith Kline
Dasatinib*	Sprycel	ABL	LMC (cromosoma Filadelfia positivo)*	Bristol-Myers Squibb
			LLA (cromosoma Filadelfia positivo)*	
Denosumab	Xgeva	RANKL	Tumor de células gigantes de hueso	Amgen

(Continuación)

**Tabla 2. Indicaciones y terapias de T diana aprobadas por la FDA de EE.UU.**

Agente	Nombre comercial	Diana(s)	Indicación(es) aprobadas por la FDA	Empresa
<b>Inhibidores de moléculas pequeñas</b>				
Erlotinib*	Tarceva	EGFR	NSCLC (con delecciones o sustituciones del exón 19)*	Genentech y OSI
			Cáncer de páncreas	
Everolimus*	Afinitor	mTOR	Tumor neuroendocrino pancreático	Novrtis
			RCC	
			Cáncer de mama (ER/PR +) en combinación con exemestano *	
			Astrocitoma subependimario de células gigantes no resecable asociado con esclerosis tuberosa	
Gefitinib	Iressa	EGFR	NSCLC con beneficio previo conocido de gefitinib (aprobación limitada)	AstraZeneca
Ibrutinib	Imbruvica	BTK	Linfoma de células del manto	Farmacyclics
Imatinib*	Gleevec	KIT, ABL	Tumor del estroma gastrointestinal	Novartis
			Dermatofibrosarcoma protuberans	
			Varias neoplasias hematológicas, incluidas LLA y LMC con cromosoma Filadelfia positivo *	
Lapatinib *	Tykerb	HER2, EGFR	Cáncer de mama (HER2+)*	GlaxoSmithKline
Nilotinib *	Tasigna	ABL	LMC (cromosoma Filadelfia positivo)*	Novartis
Pazopanib	Votrient	VEGFR, PDGFR, KIT	RCC	GlaxoSmithKline
Regorafenib	Stivarga	KIT, PDGFR $\beta$ , RAF, RET, VEGFR1/2/3	Sarcoma de tejido blando	Bayer
			CRC	
Ruxolitinib	Jakafi	JAK1/2	Tumores del estroma gastrointestinal	Bayer
			Mielofibrosis	
Sorafenib	Nexavar	VEGFR, PDGFR, KIT, RAF	Carcinoma hepatocelular	Bayer
			RCC	
Sunitinib	Sutent	VEGFR, PDGFR, KIT, RET	GIST	Pfizer
			Tumor neuroendocrino pancreático	
			RCC	
Temsirolimus	Torisel	mTOR	RCC	Wyeth
Trametinib *	Mekinista	MEK	Melanoma (con mutaciones BRAF V600E o V600K)*	GlaxoSmithKline
Vandetanib	Caprelsa	EGFR RET, VEGFR2	Cáncer de tiroides medular	AstraZeneca
Vemurafenib *	Zelboraf	BRAF	Melanoma (con mutación BRAF V600)*	Roche
Nota: LLA, leucemia linfoblástica aguda; LMC, leucemia mieloide crónica; Tumor de estroma gastrointestinal GIST; RE, receptor de estrógeno, RP, receptor de progesterona; NSCLC, cáncer de pulmón de células microcíticas; CRC, cáncer colorrectal; GBM, glioblastoma; RCC, carcinoma de células renales, HNSCC, carcinoma de células escamosas de cabeza y cuello; CLL, leucemia linfoblástica crónica; BTK, tirosina quinasa de Bruton. * Terapia dirigida que se asocia con una alteración del subtipo de cáncer molecular específica. Hay aproximadamente 17 terapias dirigidas que están asociadas con 10 subtipos moleculares específicos de cáncer.				

En una realización, basándose en el perfil de heterogeneidad de la enfermedad, se determina una intervención terapéutica que tiene en cuenta tanto el tipo de variantes genéticas encontradas en las células de la enfermedad como sus cantidades relativas (por ejemplo, proporción). La intervención terapéutica puede tratar al sujeto como si cada variante clonal fuera un cáncer diferente a tratar de forma independiente. En algunos casos, cuando se detectan una o más variantes genéticas en cantidades inferiores a las subclínicas, por ejemplo, al menos 5 veces más bajas, al menos 10 veces más bajas o al menos 100 veces más bajas que los clones detectados dominantes, estas variantes pueden quedar fuera de la intervención terapéutica hasta que alcancen un umbral clínico o una frecuencia relativa significativa (por ejemplo, mayor que el umbral indicado anteriormente).

5 Cuando una pluralidad de diferentes variantes genéticas se encuentra en diferentes cantidades, por ejemplo, diferentes números o diferentes cantidades relativas, una intervención terapéutica puede incluir tratamientos eficaces contra las enfermedades con cada uno de los genéticos variantes. Por ejemplo, en el caso del cáncer, pueden detectarse variantes genéticas, tales como formas mutantes de un gen o amplificación de genes, en varios genes (por ejemplo, un clon principal y un clon menor). Cada una de estas formas puede ser procesable, es decir, puede conocerse un tratamiento para el que responden los cánceres con la variante particular. Sin embargo, el perfil de heterogeneidad tumoral puede indicar que una de las variantes está presente en los polinucleótidos, por ejemplo, cinco veces el nivel de cada una de las otras dos variantes. Puede determinarse una intervención terapéutica que implica administrar tres fármacos diferentes al sujeto, cada fármaco relativamente más eficaz contra los cánceres que portan cada una de las variantes. Los medicamentos se pueden administrar como cóctel o secuencialmente.

10 15 Los fármacos se pueden administrar en dosis estratificadas para reflejar las cantidades relativas de las variantes en el ADN. Por ejemplo, un fármaco eficaz contra la variante más común puede administrarse en mayor cantidad que los fármacos eficaces contra las dos variantes menos comunes.

20 25 Alternativamente, el perfil de la heterogeneidad tumoral puede mostrar la presencia de una subpoblación de células cancerosas que llevan una variante genética que es resistente a un fármaco a la que la enfermedad responde típicamente. En este caso, la intervención terapéutica puede implicar la inclusión tanto de un primer fármaco eficaz contra las células tumorales sin la variante de resistencia como de un segundo fármaco eficaz contra las células tumorales con la variante resistente. Nuevamente, las dosis se pueden estratificar para reflejar las cantidades relativas de cada variante detectada en el perfil.

30 35 Los cambios en el perfil de la heterogeneidad tumoral se examinan con el tiempo, e intervenciones terapéuticas se han desarrollado para tratar el tumor cambiante. Por ejemplo, la heterogeneidad de la enfermedad se puede determinar en una pluralidad de momentos diferentes. Usando los métodos de elaboración de perfiles de esta divulgación, se pueden hacer inferencias más precisas sobre la evolución del tumor. Esto permite al médico controlar la evolución de la enfermedad, en particular a medida que surgen nuevas subpoblaciones clonales después de la remisión efectuada por una primera ola de terapia. En este caso, las intervenciones terapéuticas se pueden calibrar con el tiempo para tratar el tumor cambiante. Por ejemplo, un perfil puede mostrar que un cáncer tiene una forma que responde a un determinado tratamiento. Se administra el tratamiento y se observa que la carga tumoral disminuye con el tiempo. En algún momento, se encuentra una variante genética en el tumor que indica la presencia de una población de células cancerosas que no responde al tratamiento. Se determina una nueva intervención terapéutica que se dirige a las células que portan el marcador de falta de respuesta.

40 45 50 En respuesta a la quimioterapia, una forma de tumor dominante puede eventualmente dar paso a través de la selección darwiniana a células cancerosas que portan mutantes que hacen que el cáncer no responda al régimen de terapia. La aparición de estos mutantes de resistencia puede retrasarse. Por ejemplo, un sujeto puede someterse a uno o más ciclos de terapia pulsada, cada ciclo de terapia pulsada comprende un primer período durante el cual se administra un fármaco en una primera cantidad y un segundo ciclo durante el cual el fármaco se administra en una segunda cantidad reducida. El primer período se caracteriza por una carga tumoral detectada por encima de un primer nivel clínico. El segundo período se caracteriza por una carga tumoral detectada por debajo de un segundo nivel clínico. Los niveles clínicos primero y segundo pueden ser diferentes en diferentes ciclos de terapia pulsada. Entonces, por ejemplo, el primer nivel clínico puede ser menor en los ciclos sucesivos. Una pluralidad de ciclos puede incluir al menos 2, 3, 4, 5, 6, 7, 8 o más ciclos. Por ejemplo, el V600E mutante de BRAF puede detectarse en polinucleótidos de células enfermas en una cantidad que indica una carga tumoral del 5% en cfADN. La quimioterapia puede comenzar con dabrafenib. Las pruebas posteriores pueden mostrar que la cantidad del mutante BRAF en el cfADN cae por debajo del 0,5% o hasta niveles indetectables. En este punto, la terapia con dabrafenib puede interrumpirse o reducirse significativamente. Pruebas posteriores posteriores pueden encontrar que el ADN que lleva la mutación BRAF ha aumentado al 2,5% de los polinucleótidos en el ADNc. En este punto, la terapia con dabrafenib se vuelve a iniciar, por ejemplo, al mismo nivel que el tratamiento inicial. Las pruebas posteriores pueden encontrar que el ADN que lleva la mutación BRAF ha disminuido al 0,5% de los polinucleótidos en el cfADN. Nuevamente, la terapia con dabrafenib se detiene o se reduce. El ciclo se puede repetir varias veces.

55 60 La Figura 7 muestra un curso ejemplar de seguimiento y tratamiento de una enfermedad en un sujeto. Un sujeto analizado en el momento de la extracción de sangre 1 tiene una carga tumoral del 1,4% y presenta alteraciones genéticas en los genes 1, 2 y 3. El sujeto es tratado con el fármaco A. Después de un tiempo, se interrumpe el tratamiento. En un segundo momento posterior, una segunda extracción de sangre muestra que el cáncer está en remisión. En un tercer momento posterior, una tercera extracción de sangre indica que el cáncer ha reaparecido, en este caso, presentando una variante genética en el Gen 4. El sujeto ahora recibe un tratamiento con el Medicamento B, al cual responden los cánceres que tienen esta variante.

65 Una intervención terapéutica se puede cambiar después de la detección de la aparición de una forma mutante resistente a un medicamento original. Por ejemplo, los cánceres con la mutación EGFR L858R responden a la terapia con erlotinib. Sin embargo, los cánceres con la mutación EGFR T790M son resistentes a erlotinib. Sin embargo,

responden al ruxolitinib. Un método de esta divulgación implica controlar los cambios en el perfil del tumor y cambiar una intervención terapéutica cuando una variante genética asociada con la resistencia a los fármacos se eleva a un nivel clínico predeterminado.

## 5 Base de datos

Se construye una base de datos como se define en las reivindicaciones en donde se registra la información genética a partir de muestras de serie recogidas de pacientes con cáncer. Esta base de datos también puede contener tratamientos intermedios y otra información clínicamente relevante, como peso, efectos adversos, pruebas histológicas, análisis de sangre, información radiográfica, tratamientos previos, tipo de cáncer, etc. Los resultados de las pruebas seriadas se pueden utilizar para inferir la eficacia del tratamiento, especialmente cuando se usa con muestras de sangre, lo que puede dar una estimación más imparcial de la carga tumoral que la autoevaluación o el reporte radiográfico de un médico. La eficacia del tratamiento se puede agrupar por aquellos con perfiles genómicos similares y viceversa. Los perfiles genómicos pueden organizarse en torno, por ejemplo, a una alteración genética primaria, una alteración o alteraciones genéticas secundarias, cantidades relativas de estas alteraciones genéticas y carga tumoral. Esta base de datos se puede utilizar para apoyar la decisión de pacientes posteriores. Tanto la línea germinal como alteraciones somáticas se pueden utilizar para determinar la eficacia del tratamiento también. Alteraciones de resistencia adquiridas que también se pueden inferir de la base de datos cuando los tratamientos que inicialmente fueron efectivos comienzan a fallar. Este fallo puede detectarse mediante radiografías, sangre u otros medios. Los datos primarios utilizados para la inferencia de los mecanismos de resistencia adquiridos son los perfiles de tumores genómicos recopilados después del tratamiento por paciente. Estos datos también se pueden utilizar para establecer límites cuantitativos en la respuesta probable al tratamiento, así como para predecir el tiempo hasta el fracaso del tratamiento. Con base en las probables alteraciones de resistencia adquiridas para un tratamiento dado y un perfil genómico tumoral, se puede modificar un régimen de tratamiento para suprimir la adquisición de las alteraciones de resistencia más probables.

## Sistemas de computación

Los métodos de la presente divulgación se pueden implementar utilizando, o con la ayuda de, los sistemas informáticos. FIG. 5 muestra un sistema informático 1501 que está programado o configurado de otro modo para implementar los métodos de la presente divulgación. El sistema informático 1501 incluye una unidad central de procesamiento (CPU, también "procesador" y "procesador informático" en este documento) 1505. El sistema informático 1501 también incluye memoria o ubicación de memoria 1510 (por ejemplo, memoria de acceso aleatorio, memoria de sólo lectura, memoria flash), unidad de almacenamiento electrónico 1515 (por ejemplo, disco duro), interfaz de comunicación 1520 (por ejemplo, adaptador de red) para comunicarse con uno o más de otros sistemas, y dispositivos periféricos 1525, como caché, otra memoria, almacenamiento de datos y/o adaptadores de pantalla. La memoria 1510, la unidad de almacenamiento 1515, la interfaz 1520 y los dispositivos periféricos 1525 están en comunicación con la CPU 1505 a través de un bus de comunicación (líneas continuas). La unidad de almacenamiento 1515 puede ser una unidad de almacenamiento de datos (o depósito de datos) para almacenar datos. El sistema informático 1501 puede acoplarse operativamente a una red informática ("red") 1530 con la ayuda de la interfaz de comunicación 1520. La red 1530 puede ser Internet, un internet y/o extranet, o una intranet y/o extranet que está en comunicación con Internet. La red 1530 en algunos casos es una red de telecomunicaciones y/o datos. La red 1530 puede incluir uno o más servidores informáticos, que pueden permitir la informática distribuida, como la informática en la nube. La CPU 1505 puede ejecutar una secuencia de instrucciones legibles por máquina, que se pueden incorporar en un programa o software. Las instrucciones pueden almacenarse en una ubicación de memoria, como la memoria 1510. La unidad de almacenamiento 1515 puede almacenar archivos, como controladores, bibliotecas y programas guardados. El sistema informático 1501 puede comunicarse con uno o más sistemas informáticos remotos a través de la red 1530. Los métodos de la invención pueden implementarse mediante un código ejecutable de máquina (por ejemplo, procesador informático) almacenado en una ubicación de almacenamiento electrónico del sistema informático 1501, tal como, por ejemplo, en la memoria 1510 o la unidad de almacenamiento electrónica 1515. El código ejecutable o legible por máquina puede proporcionarse en forma de software. En la programación pueden realizarse varios aspectos como el sistema informático 1501. Varios aspectos de la tecnología se pueden considerar como "productos" o "artículos manufacturados" típicamente en forma de código ejecutable de máquina (o procesador) y/o datos asociados que se llevan o incorporan en un tipo de medio legible por máquina. El código ejecutable por máquina se puede almacenar en una unidad de almacenamiento electrónico, tal memoria (por ejemplo, la memoria de sólo lectura, memoria de acceso aleatorio, memoria flash) o un disco duro. Los medios de tipo "almacenamiento" pueden incluir parte o toda la memoria tangible de los ordenadores, procesadores o similares, o módulos asociados de los mismos, como varias memorias de semiconductores, unidades de cinta, unidades de disco y similares, que pueden proporcionar almacenamiento no transitorio en cualquier momento para la programación del software. En ocasiones, todo el software o partes del mismo pueden comunicarse a través de Internet o de otras redes de telecomunicaciones. El sistema informático 1501 puede incluir o estar en comunicación con una pantalla electrónica que comprende una interfaz de usuario (IU) para proporcionar, por ejemplo, uno o más resultados del análisis de muestras.

**EJEMPLOS**

Posiciones de nucleótidos (por ejemplo, loci) en el genoma pueden ser designadas por el número, tal como se representa en la Figura 2. Las posiciones en donde aproximadamente el 100% de las llamadas de base son idénticas a la secuencia de referencia o en donde aproximadamente el 100% de las llamadas de bases son diferentes a las de la secuencia de referencia y se infiere que representan la homocigosidad del cfADN (se presume normal). Se infiere que las posiciones en las que aproximadamente el 50% de las llamadas de bases son idénticas a la secuencia de referencia representan la heterocigosidad del cfADN (también se presume normal). Se infiere que las posiciones en las que el porcentaje de llamadas de bases en un locus están sustancialmente por debajo del 50% y por encima del límite de detección del sistema de llamadas de bases representan variantes genéticas asociadas a tumores.

10    **Ejemplo 1. Métodos para la variación del número de copias de detección de Recogida de sangre**

15    Muestras de sangre de 10-30 ml se recogen a temperatura ambiente. Las muestras se centrifugan para eliminar las células. El plasma se recoge después de la centrifugación.

**Extracción de cfADN**

20    La muestra se somete a digestión con proteinasa K. El ADN se precipita con isopropanol. El ADN se captura en una columna de purificación de ADN (por ejemplo, un QIAamp DNA Blood Mini Kit) y se eluye en 100 µl de solución. Los ADN por debajo de 500 pb se seleccionan con captura de cuentas magnéticas Ampure SPRI (PEG/sal). La producción resultante se suspende en 30 µL H<sub>2</sub>O. La distribución de tamaño se comprueba (pico principal = 166 nucleótidos; pico menor = 330 nucleótidos) y se cuantifica. 5 ng de ADN extraído contienen aproximadamente 1700 equivalentes de genoma haploide ("HGE"). La correlación general entre la cantidad de ADN y HGE es la siguiente: 3 pg de ADN = 1 HGE; 3 ng de ADN = 1K HGE; 3 µg de ADN = 1M HGE; 10 pg de ADN = 3 HGE; 10 ng de ADN = 3K HGE; 10 µg de ADN = 3M HGE.

**Prep de biblioteca de "molécula única"**

30    Etiquetado de ADN de alta eficiencia (>80%) se realiza mediante la reparación final, cola A y ligación de extremos pegajosos con 2 diferentes octómeros (es decir, 4 combinaciones) con adaptadores de horquilla sobrecargados. Se utilizan 2,5 ng de ADN (es decir, aproximadamente 800 HGE) como material de partida. Cada adaptador de horquilla comprende una secuencia aleatoria en su parte no complementaria. Ambos extremos de cada fragmento de ADN están unidos con adaptadores de horquilla. Cada fragmento etiquetado puede identificarse mediante una combinación de la secuencia octomérica en los adaptadores de horquilla y porciones endógenas de la secuencia del inserto.

40    El ADN marcado se amplifica mediante 12 ciclos de RCP para producir aproximadamente 1-7 µg de ADN que contiene aproximadamente 500 copias de cada uno de los 800 HGE en el material de partida.

45    Optimización de tampón, optimización de la polimerasa y la reducción de ciclo se puede realizar para optimizar las reacciones de RCP. El sesgo de amplificación, por ejemplo, el sesgo no específico, el sesgo de GC y/o el sesgo de tamaño también se reducen mediante la optimización. Los ruidos (p. ej., errores introducidos por la polimerasa) se reducen utilizando polimerasas de alta fidelidad.

50    Las secuencias se pueden enriquecer de la siguiente manera: los ADN con regiones de interés (ROI) se capturan usando una perla marcada con biotina con sonda para ROI. Las ROI se amplifican con 12 ciclos de RCP para generar una amplificación de 2000 veces.

55    **Secuenciación masivamente paralela**

0,1 a 1% de la muestra (aproximadamente 100 pg) se utilizan para la secuenciación. A continuación, el ADN resultante se desnaturaliza y se diluye a 8 pM y se carga en un secuenciador Illumina.

55    **Bioinformática digital**

60    Las lecturas de secuencia se agrupan en familias, con aproximadamente 10 lecturas de secuencia en cada familia. Las familias se colapsan en secuencias de consenso mediante votación (por ejemplo, votación sesgada) cada posición en una familia. Se pide una base para la secuencia de consenso si 8 o 9 miembros están de acuerdo. No se requiere una base para la secuencia de consenso si no más del 60% de los miembros están de acuerdo.

65    Las secuencias de consenso resultantes se asignan a un genoma de referencia, tales como hg19. Cada base en una secuencia de consenso está cubierta por aproximadamente 3000 familias diferentes. Se calcula una puntuación de calidad para cada secuencia y las secuencias se filtran en función de sus puntuaciones de calidad. Las llamadas de base en cada posición en una secuencia de consenso se comparan con la secuencia de referencia de HG-19. En

cada posición en donde una llamada de base difiere de la secuencia de referencia, se determina e informa la identidad de las diferentes base o bases y su porcentaje en función de las llamadas de bases totales en el locus.

La variación de secuencia se detecta contando la distribución de bases en cada locus. Si el 98% de las lecturas tienen la misma base (homocigotas) y el 2% tiene una base diferente, es probable que el locus tenga una variante de secuencia, presumiblemente del ADN del cáncer. El CNV se detecta contando el número total de secuencias (bases) mapeadas a un locus y comparándolas con un locus de control. Para aumentar la detección de CNV, el análisis de CNV se realiza en regiones específicas, incluidas regiones en ALK, APC, BRAF, CDKN2A, EGFR, ERBB2, FBXW7, KRAS, MYC, NOTCH1, NRAS, PIK3CA, PTEN, RB1, TP53, MET, AR, ABL1, AKT1, ATM, CDH1, CSF1R, CTNNB1, ERBB4, EZH2, FGFR1, FGFR2, FGFR3, FLT3, GNA11, GNAQ, GNAS, HNF1A, HRAS, IDH1, IDH2, JAK2, JAK3, KDR, KITL, MLH1, PDGFRA, PROC, PTPN11, RET, SMAD4, SMARCB1, SMO, SRC, STK11, VHL, TERT, CCND1, CDK4, CDKN2B, RAF1, BRCA1, CCND2, CDK6, NF1, TP53, ARID1A, BRCA2, CCNE1T1, ESR, GATA3, MAP2K1, RHEB, ROS1, ARAF, MAP2K2, NFE2L2, RHOA o NTRK1.

**15 Ejemplo 2. Método para corregir base llamando por determinar el número total de moléculas no vistas en una muestra**

Después de fragmentos se amplifican y las secuencias de los fragmentos amplificados se leen y se alinean, los fragmentos se someten a la llamada de base. Las variaciones en el número de fragmentos amplificados y fragmentos amplificados no vistos pueden introducir errores en la llamada de bases. Estas variaciones se corrigen calculando el número de fragmentos amplificados no vistos.

Al llamar de base para locus A (un locus arbitrario), se supone primero que hay N fragmentos amplificados. Las lecturas de secuencia pueden provenir de dos tipos de fragmentos: fragmentos bicatenarios y fragmentos monocatenarios. El siguiente es un ejemplo teórico de cómo calcular el número total de moléculas no vistas en una muestra.

N es el número total de moléculas en la muestra.

Suponiendo que 1000 es el número de dúplex detectados.

30 Suponiendo que 500 es el número de moléculas monocatenarias detectadas.

P es la probabilidad de ver una hebra.

Q es la probabilidad de no detectar una hebra.

$$35 \quad \text{Al ser } Q = 1 - P.$$

$$1000 = NP(2).$$

$$500 = N2PQ.$$

$$40 \quad 1000 / P(2) = N.$$

$$500 / 2 PQ = N.$$

$$45 \quad 1000 / P(2) = 500 / 2PQ.$$

$$1000 * 2 PQ = 500 P(2).$$

$$50 \quad 2000 PQ = 500 P(2).$$

$$2000 Q = 500 P.$$

$$55 \quad 2000 (1-P) = 500P$$

$$2000 - 2000 P = 500P.$$

$$60 \quad 2000 = 500P + 2000 P.$$

$$2000 = 2500 P.$$

$$65 \quad 2000 / 2500 = P.$$

0,8 = P.

5                    $1000 / P(2) = N.$

$1000 \div 0.64 = N.$

10                 1562 = N.  
Número de fragmentos no vistos = 62.

**Ejemplo 3. Identificación de variantes genéticas en variantes somáticas asociadas al cáncer en un paciente**

15                 Se utiliza un ensayo para analizar un panel de genes para identificar variantes genéticas en variantes somáticas asociadas al cáncer con alta sensibilidad.

20                 El ADN libre de células se extrae del plasma de un paciente y se amplifica mediante RCP. Las variantes genéticas se analizan mediante secuenciación masivamente paralela de los genes diana amplificados. Para un conjunto de genes, todos los exones se secuencian, ya que se ha demostrado que la cobertura de secuenciación tiene utilidad clínica (Tabla 3). Para otro conjunto de genes, la cobertura de secuenciación incluyó aquellos exones con una mutación somática informada previamente (Tabla 4). El alelo mutante mínimo detectable (límite de detección) depende de la concentración de ADN libre de células de la muestra del paciente, que varió de menos de 10 a más de 1000 equivalentes genómicos por ml de sangre periférica. Es posible que la amplificación no se detecte en muestras con cantidades más bajas de ADN libre de células y/o amplificación de copias de genes de bajo nivel. Ciertas características de la muestra o variantes dieron como resultado una sensibilidad analítica reducida, como una muestra de baja calidad o una recolección incorrecta.

30                 El porcentaje de variantes genéticas encontradas en el ADN libre de células que circulan en la sangre está relacionada con la biología única tumoral de este paciente. Los factores que afectaron a la cantidad/porcentajes de variantes genéticas detectadas en el ADN libre de células circulante en la sangre incluyen el crecimiento, recambio, tamaño, heterogeneidad, vascularización, progresión de la enfermedad o tratamiento del tumor. La Tabla 5 anota el porcentaje, o la frecuencia alélica, de ADN libre de células circulante alterado (% cfADN) detectado en este paciente. Algunas de las variantes genéticas detectadas se enumeran en orden descendente por % cfADN.

35                 Las variantes genéticas se detectan en el ADN libre de células circulante aislado de muestra de sangre de este paciente. Estas variantes genéticas son variantes somáticas asociadas al cáncer, algunas de las cuales se han asociado con una respuesta clínica aumentada o reducida a un tratamiento específico. Las "alteraciones menores" se definen como aquellas alteraciones detectadas a menos del 10% de la frecuencia alélica de las "alteraciones mayores". Una alteración mayor es la alteración predominante en un locus. Se anotan las frecuencias alélicas detectadas de estas alteraciones (Tabla 5) y los tratamientos asociados para este paciente.

40                 Todos los genes enumerados en las Tablas 3 y 4 se analizan como parte de la prueba. No se detecta amplificación para *ERBB2*, *EGFR* o *MET* en el ADN libre de células circulante aislado de la muestra de sangre de este paciente.

45                 Resultados de las pruebas de paciente que comprenden las variantes genéticas se enumeran en la Tabla 6.

50                 En referencia a la Tabla 4, en 13 posiciones, un nucleótido detectado en al menos 98,8% de frecuencia en la muestra es diferente de un nucleótido en la secuencia de referencia, indicando homocigosis en estos loci. Por ejemplo, en el gen KRAS, en la posición 25346462, se detectó T en lugar del nucleótido C de referencia en el 100% de los casos.

55                 En 35 posiciones, un nucleótido detectado en entre 41,4% y 55% de frecuencia en la muestra es diferente de un nucleótido en la secuencia de referencia, lo que indica heterocigosis en estos loci. Por ejemplo, en el gen ALK, en la posición 29455267, se detectó G en lugar del nucleótido A de referencia en el 50% de los casos.

60                 A las 3 posiciones un nucleótido detectado en menos de 9% de frecuencia es diferente de un nucleótido en la referencia de secuencia. Estos incluyen variantes en BRAF (140453136 A>T, 8,9%), NRAS (115256530 G>T 2,6%) y JAK2 (5073770 G>T 1,5%). Se presume que son mutaciones somáticas del ADN del cáncer.

65                 Se calculan las cantidades relativas de variantes genéticas asociadas a tumores. La proporción de cantidades de BRAF:NRAS:JAK2 es 8,9: 2,6: 1,5 o 1: 0,29: 0,17. De este resultado se puede inferir la presencia de heterogeneidad tumoral. Por ejemplo, una posible interpretación es que el 100% de las células tumorales contienen una variante en BRAF, el 83% contienen variantes en BRAF y NRAS, y el 17% contienen variantes en BRAF, NRAS y JAK2. Sin embargo, el análisis de CNV puede mostrar amplificación de BRAF, en cuyo caso el 100% de las células tumorales

# ES 2 923 602 T3

pueden tener variantes tanto en BRAF como en NRAS.

Tabla 3. Genes en los que se secuencian todos los exones

<b>GENES EN LOS QUE TODOS LOS EXONES SON SECUENCIADOS</b>				
ALK	< 0,1%	APC	< 0,1%	
AR	< 0,1%	BRAF	< 0,1%	
CDKN2A	< 0,1%	EGFR	< 0,1%	
ERBB2	< 0,1%	FBXW7	< 0,1%	
KRAS	< 0,1%	MET	< 0,1%	
MYC	< 0,1%	NOTCH1	< 0,1%	
NRAS	< 0,1%	PIK3CA	< 0,1%	
PTEN	< 0,1%	PROC	< 0,1%	
RB1	< 0,1%	TP53	< 0,1%	

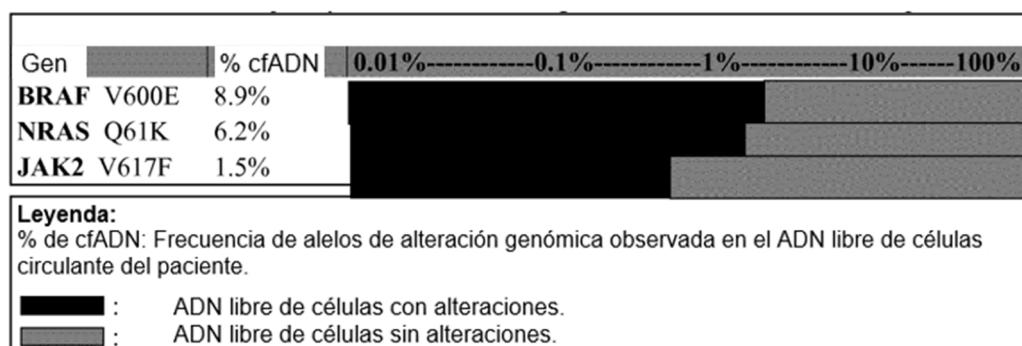
LOD: Límite de detección. La frecuencia mínima de alelos mutantes detectables para esta muestra en donde se detecta el 80% de las variantes somáticas.

20 Tabla 4. Genes en los que se encuentran exones con una mutación somática informada previamente secuenciado

<b>GENES EN LOS QUE SE SECUENCIAN EXONES CON UNA MUTACIÓN SOMÁTICA REPORTADA ANTERIORMENTE</b>				
ABL1	< 0,1%	AKT1	< 0,1%	
ATM	< 0,1%	CDH1	< 0,1%	
CSF1R	< 0,1%	CTNNB1	< 0,1%	
ERBB4	< 0,1%	EZH2	< 0,1%	
FGFR1	< 0,1%	FGFR2	< 0,1%	
FGFR3	< 0,1%	FLT3	< 0,1%	
GNA11	< 0,1%	GNAQ	< 0,1%	
GNAS	< 0,1%	HNF1A	< 0,1%	
HRAS	< 0,1%	IDH1	< 0,1%	
IDH2	< 0,1%	JAK2	< 0,1%	
JAK3	< 0,1%	KDR	< 0,1%	
KIT	< 0,1%	MLH1	< 0,1%	
MPL	< 0,1%	NPM1	< 0,1%	
PDGFRA	< 0,1%	PTPN11	< 0,1%	
RET	< 0,1%	SMAD4	< 0,1%	
SMARCB1	< 0,1%	SMO	< 0,1%	
SRC	< 0,1%	STK11	< 0,1%	
TERT	< 0,1%	VHL	< 0,1%	

LOD: Límite de detección. La frecuencia mínima de alelos mutantes detectables para esta muestra en donde se detecta el 80% de las variantes somáticas.

45 Tabla 5. Frecuencia de alelos de ADN libre de células circulante detectado en este paciente



**Tabla 6. Alteraciones genómicas detectadas en genes seleccionados**  
**Detectadas: 51 alteraciones genómicas**

<b>Gen</b>	<b>Cromosoma</b>	<b>Posición</b>	<b>Mutación (nt)</b>	<b>Mutación (AA)</b>	<b>Porcentaje</b>	<b>ID cósmico</b>	<b>DBSNP ID</b>
KRAS	12	25368462	C>T		100,0%		rs4362222
ALK	2	29416572	T>C	11461V	100,0%		rs1670283
ALK	2	29444095	C>T		100,0%		rs1569156
ALK	2	29543663	T>C	Q500Q	100,0%		rs2293564
ALK	2	29940529	A>T	P234P	100,0%		rs2246745
APC	5	112176756	T>A	V1822D	100,0%		rs459552
CDKN2A	9	21968199	C>G		100,0%	COSM14251	rs11515
FGFR3	4	1807894	G>A	T651T	100,0%		rs7688609
NOTCH1	9	139410424	A>G		100,0%		rs3125006
PDGFRA	4	55141055	A>G	P567P	100,0%		rs1873778
HRAS	11	534242	A>G	H27H	100,0%	COSM249860	rs12628
EGFR	7	55214348	C>T	N158N	99,9%	COSM42978	rs2072454
TP53	17	7579472	G>C	P72R	99,8%		rs1042522
APC	5	112162854	T>C	Y486Y	55,0%		rs2229992
APC	5	112177171	G>A	P1960P	53,8%		rs465899
EGFR	7	55266417	T>C	T903T	53,6%		rs1140475
APC	5	112176325	G>A	G1678G	53,2%		rs42427
APC	5	112176559	T>G	S1756S	53,0%		rs866006
EGFR	7	55229255	G>A	R521K	53,0%		
MET	7	116397572	A>G	Q648Q	52,7%		
APC	5	112175770	G>A	T1493T	52,7%		rs41115
EGFR	7	55249063	G>A	Q787Q	52,6%		rs1050171
NOTCH1	9	139411714	T>C		52,4%		rs11145767
EGFR	7	55238874	T>A	T629T	52,0%		rs2227984

(Continuación)

Detectadas: 51 alteraciones genómicas							
Gen	Cromosoma	Posición	Mutación (nt)	Mutación (AA)	Porcentaje	ID cósmico	DBSNP ID
ERBB2	17	37879588	A>G	I655V	51,6%		rs1136201
NOTCH1	9	139397707	G>A	D1698D	51,3%	COSM33747	rs10521
ALK	2	30143499	G>C	L9L	51,0%		rs4358080
APC	5	112164561	G>A	A545A	51,0%		rs351771
FLT3	13	28610183	A>G		50,8%		rs2491231
NOTCH1	9	139418260	A>G	N104N	50,5%		rs4489420
ALK	2	29444076	G>T		50,4%		rs1534545
PIK3CA	3	178917005	A>G		50,3%		rs3729674
NOTCH1	9	139412197	G>A		50,2%		rs9411208
ALK	2	29455267	A>G	G845G	50,0%	COSM148825	rs2256740
KIT	4	55593464	A>C	M541L	49,9%	COSM128026	
NOTCH1	9	139391636	G>A	D2185D	48,9%		rs2229974
PDGFRA	4	55152040	C>T	V824V	48,9%	COSM22413	rs2228230
ALK	2	29416481	T>C	K1491R	48,9%	COSM1130802	rs1881420
ALK	2	29445458	G>T	G1125G	48,6%		rs3795850
NOTCH1	9	139410177	T>C		48,5%	rs3124603	
RET	10	43613843	G>T	L769L	48,2%		rs1800861
EGFR	7	55214443	G>A		48,0%		rs7801956
ALK	2	29416366	G>C	D1529E	47,2%		rs1881421
EGFR	7	55238087	C>T		45,5%		rs10258429
RET	10	43615633	C>G	S904S	44,8%		rs1800863
BRAF	7	140453136	A>T	V600E	8,9%	COSM476	
NRAS	1	115256530	G>T	Q61K	6,2%	COSM580	rs121913254
JAK2	9	5073770	G>T	V617F	1,5%	COSM12600	rs77375493

**Ejemplo 4. Determinación de los límites específicos del paciente de la detección de los genes analizados mediante ensayos**

Usando el método del Ejemplo 3, se detectan alteraciones genéticas en el ADN libre de células de un paciente. Las lecturas de secuencias de estos genes incluyen secuencias de exones y/o intrones.

**Ejemplo 5. Corrección de errores de secuencia que comparan de secuencias de Watson y Crick**

ADN libre de células de hebra doble está aislado del plasma de un paciente. Los fragmentos de ADN libres de células se etiquetan utilizando 16 adaptadores diferentes que contienen burbujas, cada uno de los cuales comprende un código de barras distintivo. Los adaptadores que contienen burbujas se unen a ambos extremos de cada fragmento de ADN libre de células mediante ligación. Después de la ligación, cada uno de los fragmentos de ADN libre de células se puede identificar de forma distinta mediante la secuencia de códigos de barras distintos y dos secuencias endógenas de 20 pb en cada extremo del fragmento de ADN libre de células.

Los fragmentos de ADN libres de células etiquetadas se amplifican por RCP. Los fragmentos amplificados se enriquecen usando cuentas que comprenden sondas de oligonucleótidos que se unen específicamente a un grupo de genes asociados al cáncer. Por tanto, los fragmentos de ADN libres de células del grupo de genes asociados al cáncer se enriquecen selectivamente.

Los adaptadores de secuenciación, cada uno de los cuales comprende un sitio de unión del cebador de secuenciación, un código de barras de muestra y una secuencia de flujo celular, se unen a las moléculas de ADN enriquecidas. Las moléculas resultantes se amplifican mediante RCP.

Las dos hebras de los fragmentos amplificados se secuencian. Debido a que cada adaptador que contiene burbujas comprende una parte no complementaria (por ejemplo, la burbuja), la secuencia de una hebra del adaptador que contiene burbujas es diferente de la secuencia de la otra hebra (complemento). Por lo tanto, las lecturas de secuencia de amplicones derivados de la hebra de Watson de un ADN libre de células original se pueden distinguir de los amplicones de la hebra de Crick del ADN libre de células original por las secuencias adaptadoras que contienen burbujas adjuntas.

Las lecturas de secuencias de una hebra de un fragmento de ADN libre de células originales se comparan con las lecturas de secuencias de la otra hebra del fragmento de ADN libre de células originales. Si una variante ocurre solo en la secuencia que se lee de una hebra, pero no de otra hebra, del fragmento de ADN libre de células original, esta variante se identificará como un error (p. ej., resultado de RCP y/o amplificación), en lugar de una verdadera variante genética.

Las lecturas de secuencia se agrupan en familias. Se corrigen los errores en la secuencia de lecturas. La secuencia de consenso de cada familia se genera colapsando.

**Ejemplo 6. Intervención terapéutica**

Se determina una intervención terapéutica para tratar el cáncer. Los cánceres con mutantes BRAF responden al tratamiento con vemurafenib, regorafenib, tranetinib y dabrafenib. Los cánceres con mutantes NRAS responden al tratamiento con trametinib. Los cánceres con mutantes JAK2 responden al tratamiento con ruxolitinib. Se determina que una intervención terapéutica que incluye la administración de trametinib y ruxolitinib es más eficaz contra este cáncer que el tratamiento con cualquiera de los fármacos mencionados anteriormente solo. El sujeto se trata con una combinación de trametinib y ruxolitinib en una proporción de dosis de 5:1.

Después de varias rondas de tratamiento, el cfADN del sujeto se prueba de nuevo por la presencia de la heterogeneidad tumoral. Los resultados muestran que la proporción de BRAF:NRAS:JAK2 es ahora de aproximadamente 4:2:1.5. Esto indica que la intervención terapéutica ha reducido el número de células con mutantes BRAF y NRAS, y ha detenido el crecimiento de células con mutantes JAK2. Se determina una segunda intervención terapéutica en donde se determina que trametinib y ruxolitinib son eficaces en una proporción de dosis de 1:1. El sujeto recibe un curso de quimioterapia en cantidades en esta proporción. Las pruebas posteriores muestran que los mutantes BRAF, NRAS y JAK2 están presentes en el cfADN en cantidades inferiores al 1%.

**Ejemplo 7. Intervención terapéutica**

Una muestra de sangre se obtiene de un individuo con melanoma pre-tratamiento y se determina que el paciente tiene una mutación BRAF V600E a una concentración de 2,8% y no hay mutaciones NRAS detectables utilizando análisis de ADN libre de células. El paciente recibe una terapia anti-BRAF (dabrafenib). Después de 3 semanas, se recolecta y analiza otra muestra de sangre. Se determina que el nivel de BRAF V600E ha bajado al 0,1%. Se interrumpe la terapia y se repite la prueba cada 2 semanas. El nivel de BRAF V600E aumenta de nuevo y la terapia se reinicia cuando el nivel de BRAF V600E aumenta al 1,5%. La terapia se detiene de nuevo cuando el nivel vuelve a

bajar al 0,1%. Este ciclo se repite.

#### **Ejemplo 8. Corrección de CNV basado en mediciones ROCNV**

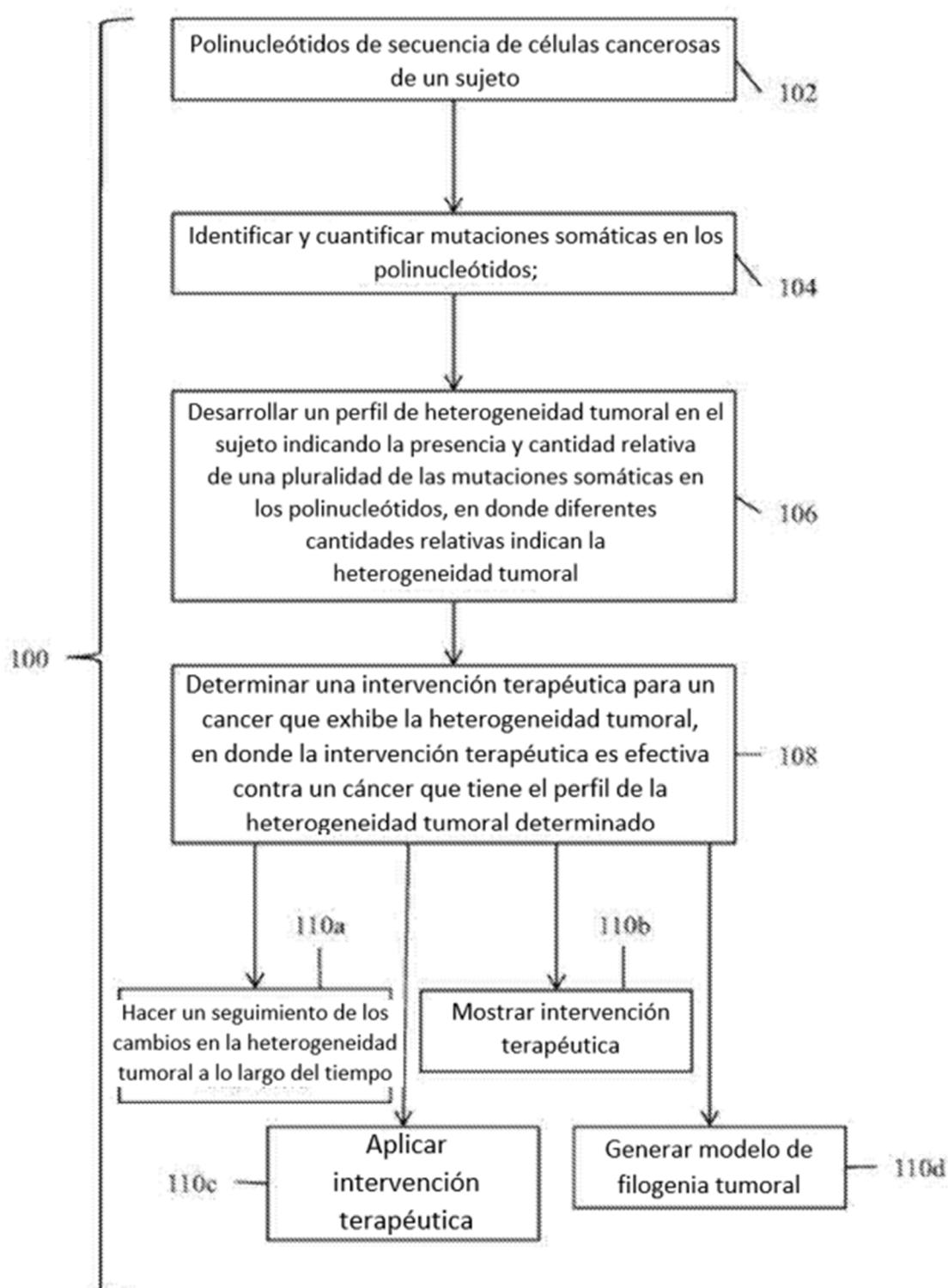
5 Se determinan variaciones de número de copias en una muestra del paciente. Los métodos para determinar  
pueden incluir seguimiento molecular y muestreo superior, como se describió anteriormente. Se utiliza un modelo de  
markov oculto basado en ubicaciones esperadas de los orígenes de replicación para eliminar el efecto de la proximidad  
10 del origen de replicación de las variaciones estimadas del número de copias en la muestra del paciente. La desviación  
estándar de las variaciones del número de copias para cada gen se reduce posteriormente en un 40%. El modelo de  
proximidad del origen de replicación también se utiliza para inferir la carga tumoral libre de células en el paciente.

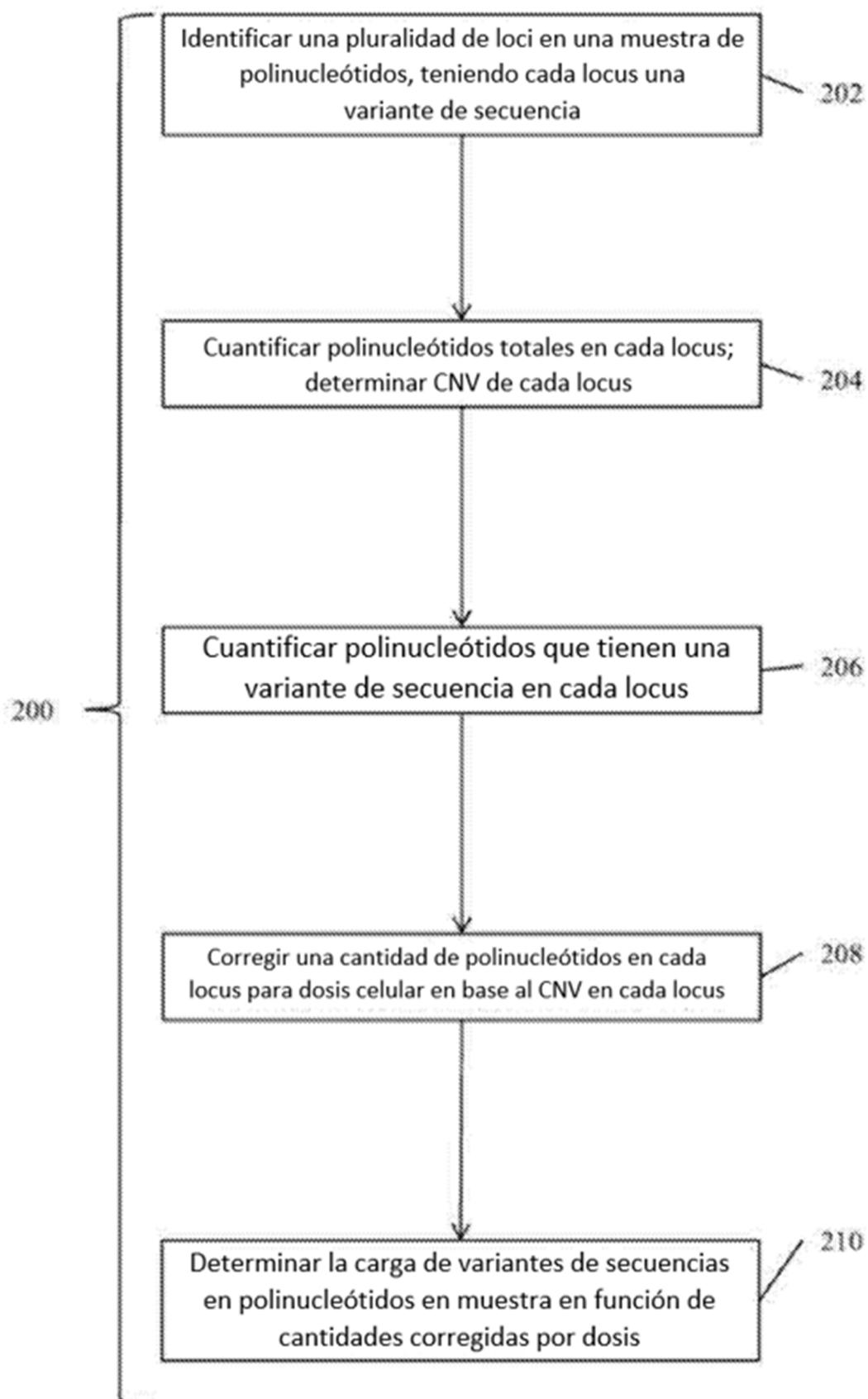
15 En muchos casos, el nivel de tumor sin células derivado puede ser bajo o por debajo del límite de detección  
de una tecnología particular. Este puede ser el caso cuando el número de equivalentes de genoma humano de ADN  
derivado de tumor en plasma es inferior a 1 copia por 5 ml. Se ha demostrado que la radiación y las quimioterapias  
afectan a las células que se dividen rápidamente más que a las células sanas estables, de ahí su eficacia en el  
tratamiento de pacientes con cáncer avanzado. Por tanto, se administra un procedimiento con efectos adversos  
20 mínimos a un paciente antes de la extracción de sangre para aumentar preferentemente la fracción de ADN derivado  
del tumor recolectado. Por ejemplo, se podría administrar una dosis baja de quimioterapia al paciente y se podría  
recolectar una muestra de sangre dentro de las 24 horas, 48 horas, 72 horas o menos de 1 semana. Para  
quimioterapias efectivas, esta muestra de sangre contiene concentraciones más altas de ADN derivado de tumores  
25 sin células debido a tasas potencialmente más altas de muerte celular de las células cancerosas. Alternativamente, la  
radioterapia de dosis baja se aplica a través de un instrumento radiográfico de cuerpo entero o localmente a las  
regiones afectadas en lugar de la quimioterapia de dosis baja. Se prevén otros procedimientos, que incluyen someter  
a un paciente a ultrasonidos, ondas sonoras, ejercicio, estrés, etc.

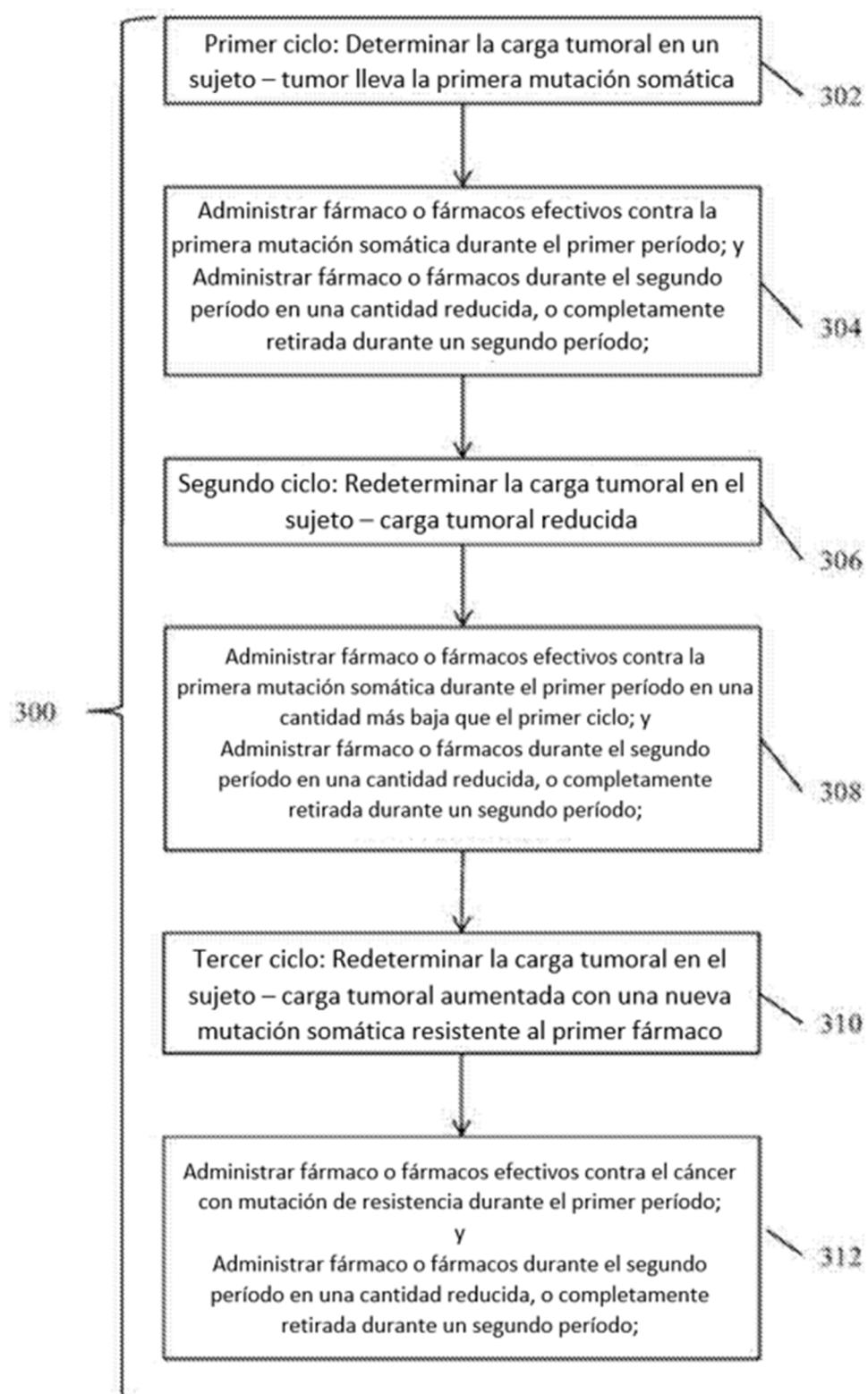
25

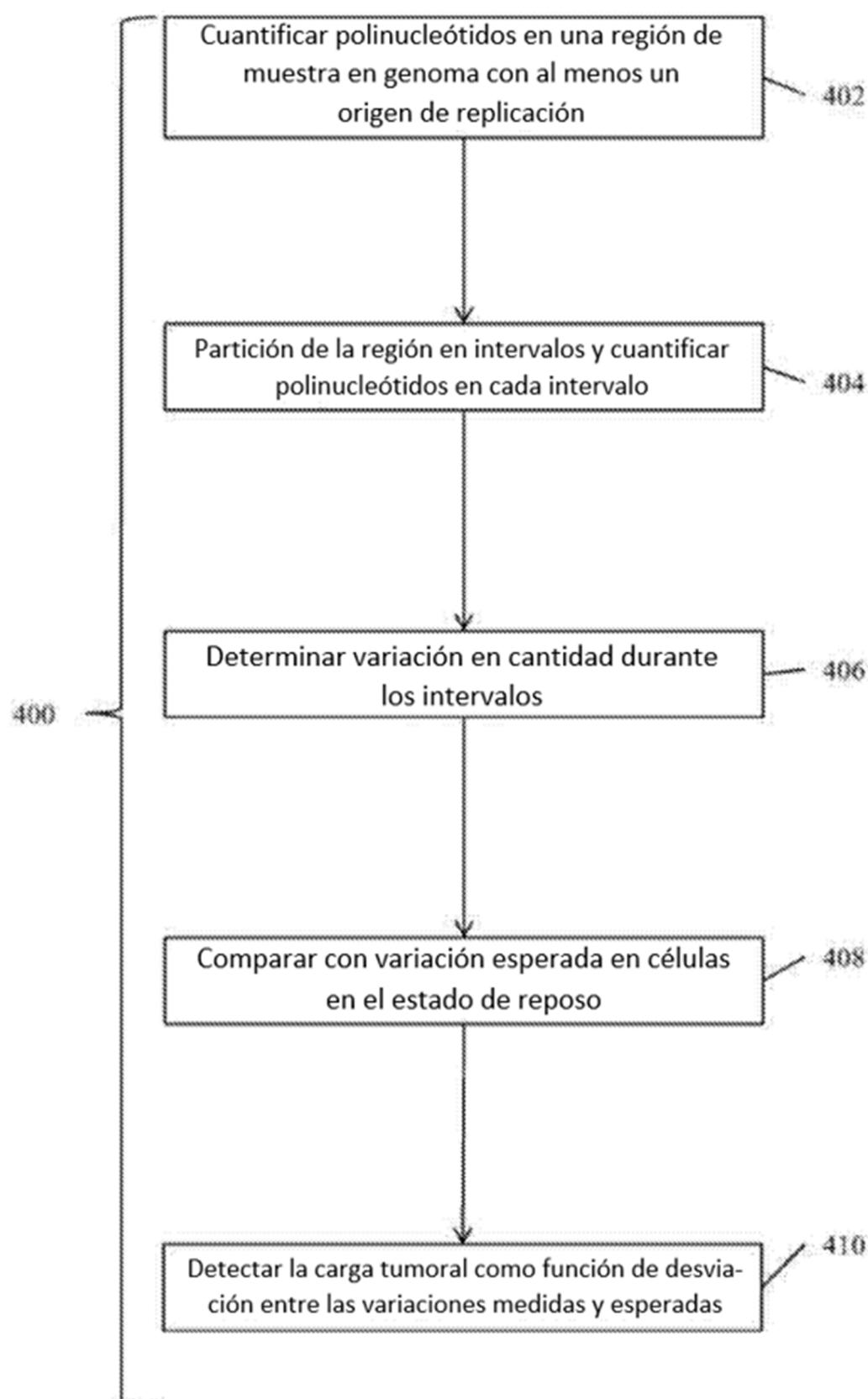
## REIVINDICACIONES

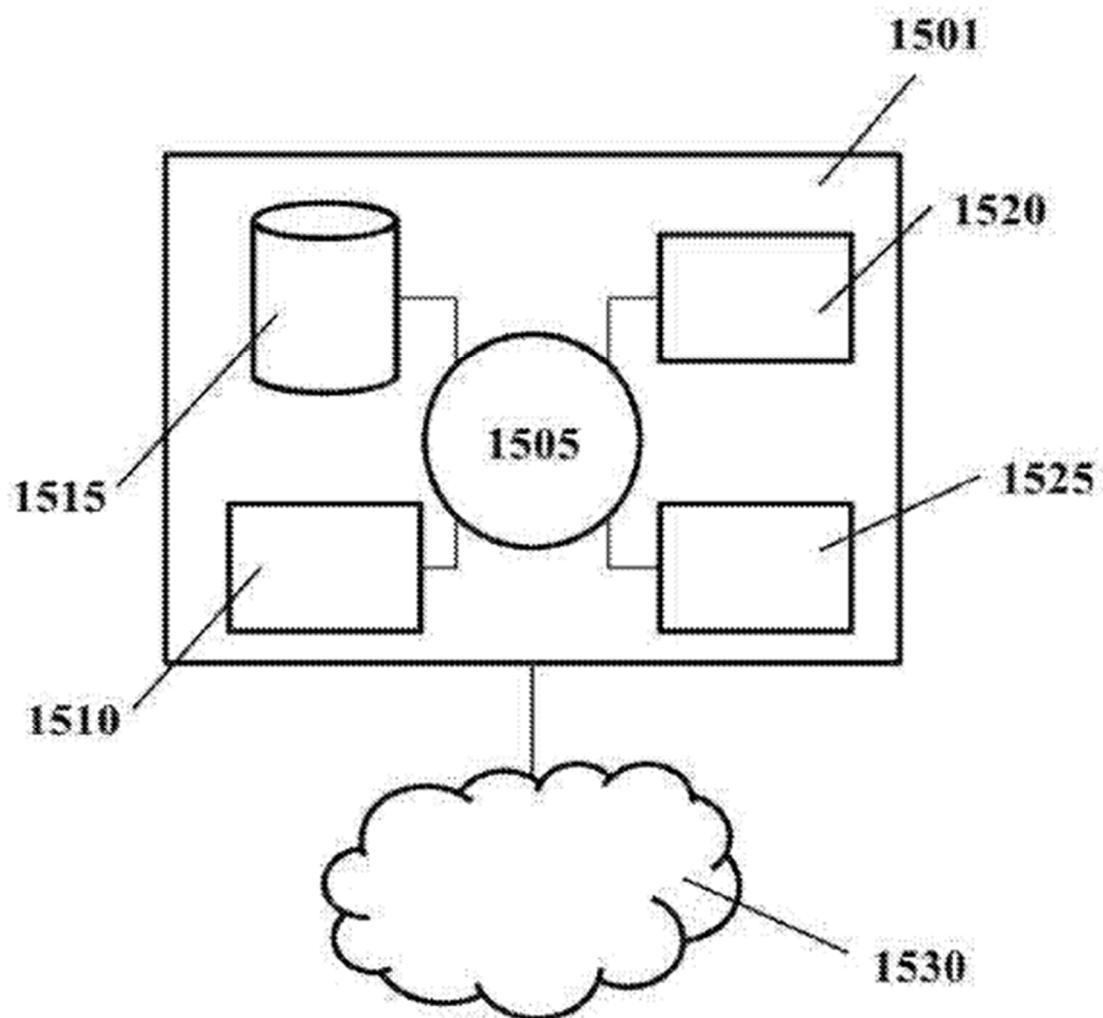
- 5        1. Un método implementado por ordenador que comprende el uso de una base de datos informática para identificar una o más intervenciones terapéuticas eficaces para un sujeto que tiene cáncer, en donde la base de datos informática incluye, para cada uno de una pluralidad de sujetos que tienen cáncer:
- 10      (i) datos de pruebas genómicas tumorales, incluyendo las alteraciones somáticas, recopilados en dos o más intervalos de tiempo por sujeto mediante biopsias en serie de ADN libre de células;  
           (ii) una o más intervenciones terapéuticas administradas a cada uno de los sujetos en uno o más momentos; y  
           (iii) eficacia de las intervenciones terapéuticas.
- 15      2. El método de la reivindicación 1, en el que las intervenciones terapéuticas identificadas se estratifican por eficacia.
- 20      3. El método de la reivindicación 2, en el que se informa de los límites cuantitativos sobre la eficacia prevista de la intervención terapéutica o la ausencia de la misma.
- 25      4. El método de cualquiera de las reivindicaciones 1-3, en el que las intervenciones terapéuticas usan información de la evolución genómica tumoral prevista o mecanismos de resistencia adquiridos en sujetos similares en respuesta al tratamiento.
- 30      5. El método de una cualquiera de las reivindicaciones 1-4, en el que la pluralidad de sujetos es de por lo menos 50, por lo menos 500 o por lo menos 5000 sujetos.
- 35      6. El método de una cualquiera de las reivindicaciones 1-5, en el que para clasificar la eficacia del tratamiento se usan frecuencias relativas de variantes genéticas detectadas.
- 40      7. Los métodos de cualquiera de las reivindicaciones 1-6, en los que se usan el peso, los efectos adversos del tratamiento, las pruebas histológicas, los análisis de sangre, la información radiográfica, los tratamientos anteriores y/o el tipo de cáncer para ayudar a clasificar la eficacia del tratamiento.
- 45      8. Los métodos de cualquiera de las reivindicaciones 1-7, en los que la respuesta al tratamiento por sujeto se recopila y clasifica cuantitativamente a través de pruebas adicionales, en donde las pruebas adicionales son pruebas basadas en sangre u orina.
- 50      9. El método de cualquiera de las reivindicaciones 1-8, en donde el método comprende clasificar la eficacia del tratamiento usando un algoritmo de clasificación, como:
- 55      (i) procesos de regresión lineal como regresión lineal múltiple, mínimos cuadrados parciales, regresión y regresión de componentes principales;  
           (ii) árboles de decisión binarios, como procesos de partición recursivos como árboles de clasificación y regresión;  
           (iii) redes neurales artificiales como redes de retropropagación;  
           (iv) análisis discriminantes como el clasificador bayesiano o el análisis de Fischer;  
           (v) clasificadores logísticos; y/o  
           (vi) clasificadores de vectores de soporte, como máquinas de vectores de soporte.
- 60      10. El método de cualquiera de las reivindicaciones 1-9, en el que se usan alteraciones tanto de la línea germinal como somáticas para determinar la eficacia del tratamiento.
- 65      11. El método de una cualquiera de las reivindicaciones 1-10, en el que las alteraciones de resistencia adquiridas se infieren de la base de datos cuando los tratamientos que fueron eficaces en la pluralidad de sujetos inicialmente comenzaron a fallar.
- 70      12. El método de cualquiera de las reivindicaciones 1-11, en el que los mecanismos de resistencia adquiridos se infieren usando perfiles genómicos tumorales recopilados después del tratamiento por sujeto.
- 75      13. El método de cualquiera de las reivindicaciones 1-12, en el que los datos de las pruebas genómicas tumorales son datos de secuenciación de ADN.
- 80      14. El método de la reivindicación 13, en el que los datos de secuenciación de ADN incluyen el mapeo de polinucleótidos en loci específicos en el genoma que son la materia de interés, y han sido aislados para la secuenciación por captura de secuencias o amplificación específica del sitio.
- 85      15. El método de cualquiera de las reivindicaciones 1-14, en el que el ADN libre de células ha sido marcado o rastreado para permitir la posterior identificación y el origen del polinucleótido particular.

**FIG. 1**

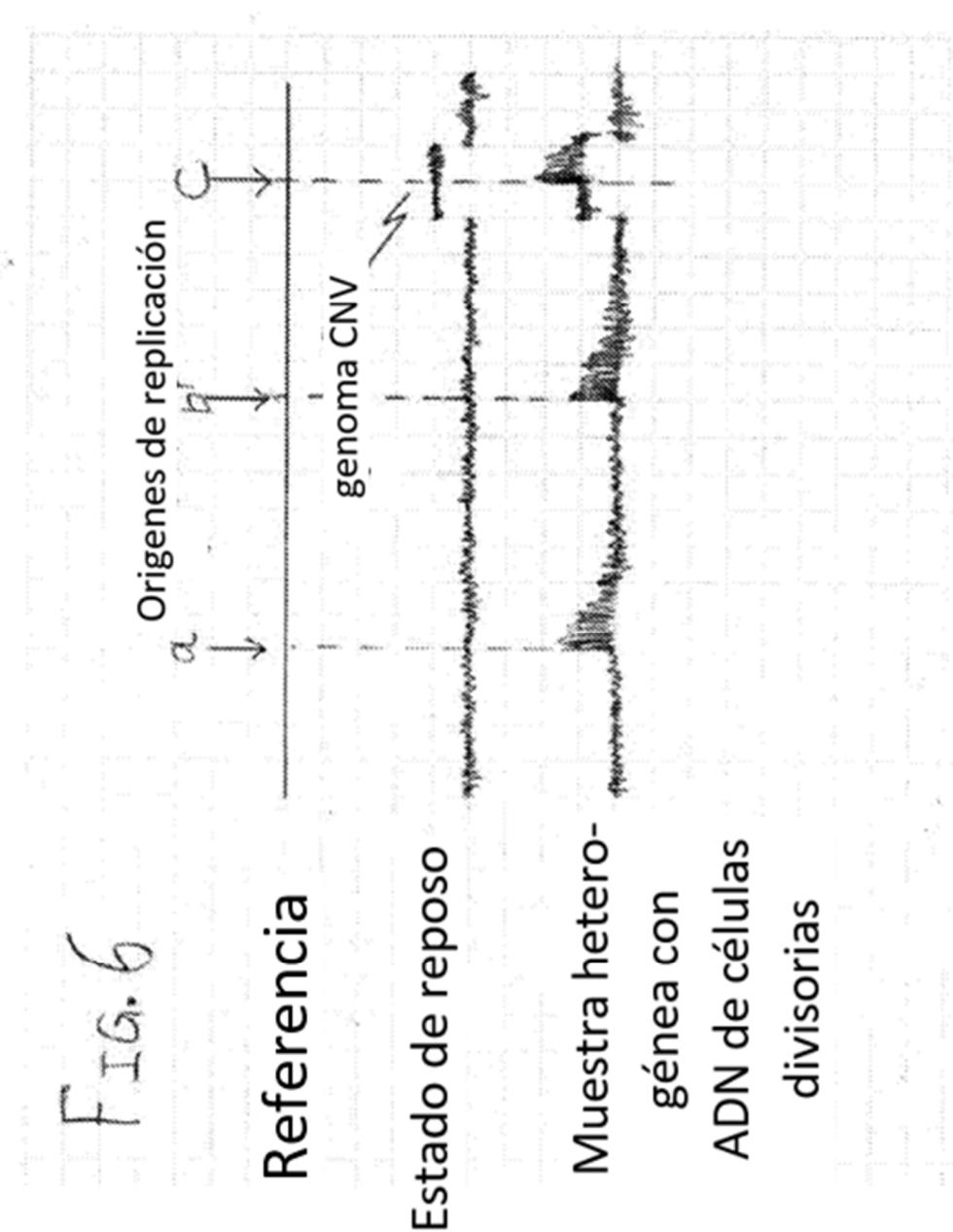
**FIG. 2**

**FIG. 3**

**FIG. 4**



***FIG. 5***

**FIG. 6**

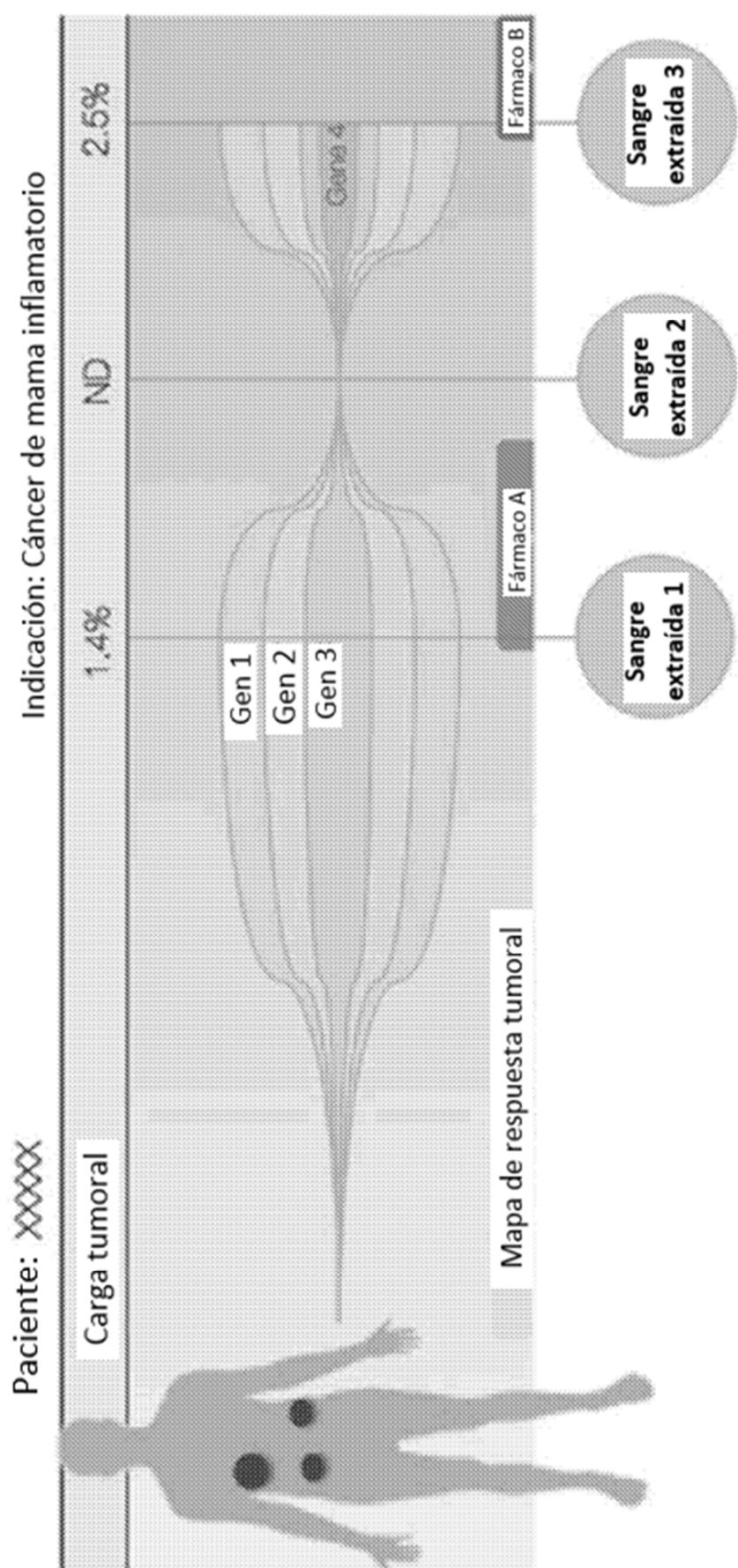


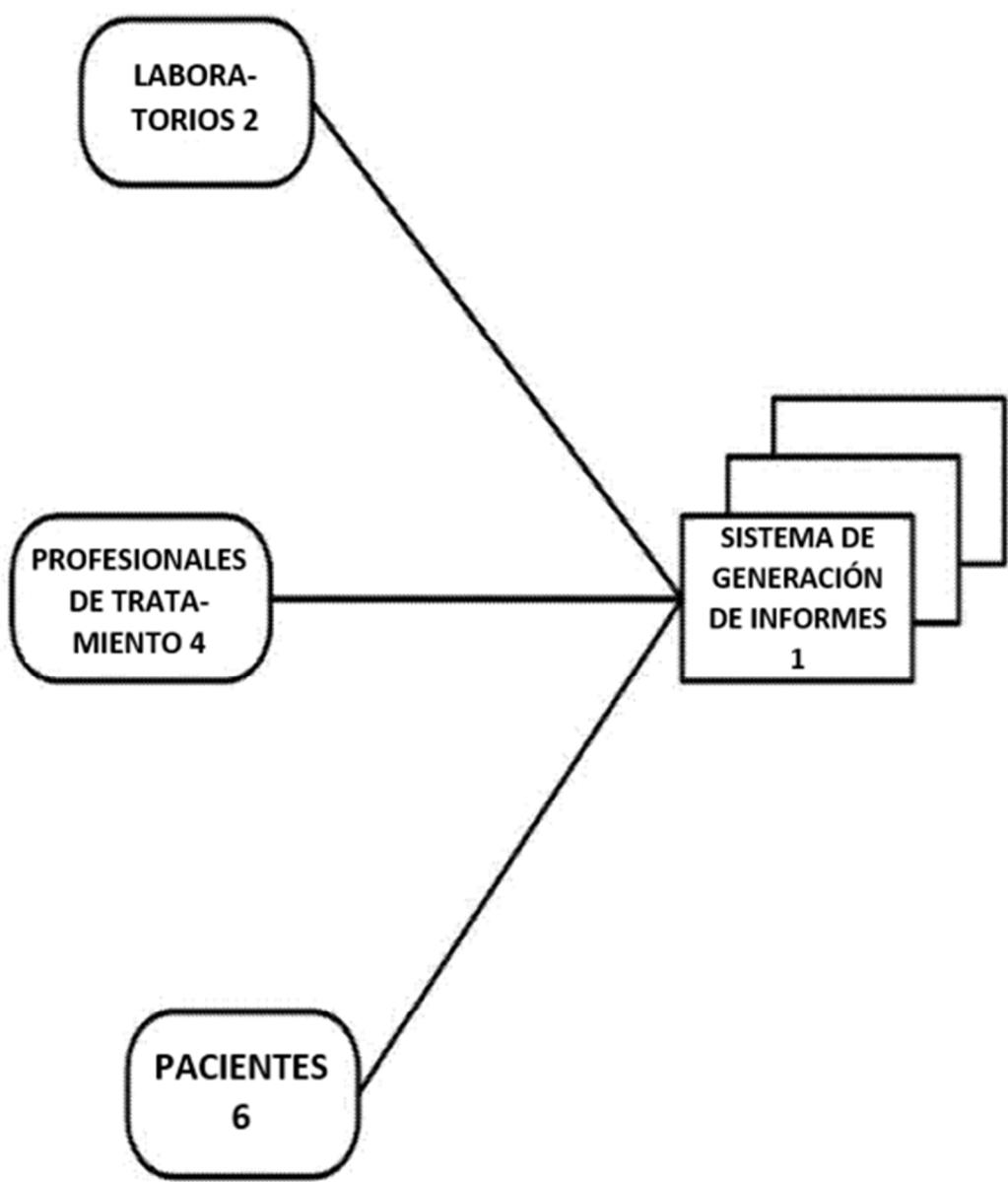
FIG. 7

### Secuenciación completa de exones

Mutaciones de punto (SNVs) (70 genes)		Amplificaciones (CNVs) (16 genes)		Fusiones (6 genes)		Indelos (3 genes)	
<i>AKT1</i>	<i>ALK</i>	<i>APC</i>	<i>AR</i>	<i>ARID1A</i>	<i>BRAF</i>	<i>ALK</i>	<i>EGFR</i>
<i>ATM</i>	<i>BRAF</i>	<i>BRCA1</i>	<i>BRCA2</i>	<i>CCND1</i>	<i>CDK4</i>	<i>FGFR2</i>	<i>ERBB2*</i>
<i>CCNE1</i>	<i>CDH1</i>	<i>CDK4</i>	<i>CDK6</i>	<i>CDKN2A</i>	<i>CDKN2B</i>	<i>FGFR</i>	<i>MET**</i>
<i>CTNNA1</i>	<i>EGFR</i>	<i>ERBB2</i>	<i>ESR1</i>	<i>EZH2</i>	<i>FBXW7</i>	<i>FGFR1</i>	<i>NTRK1</i>
<i>FGFR1</i>	<i>FGFR2</i>	<i>FGFR3</i>	<i>GATA3</i>	<i>GNA11</i>	<i>GNAQ</i>	<i>FGFR2</i>	<i>RET</i>
<i>GNAS</i>	<i>HNF1A</i>	<i>HRAS</i>	<i>IDH1</i>	<i>IDH2</i>	<i>JAK2</i>	<i>KRAS</i>	<i>ROS1</i>
<i>JAK3</i>	<i>KIT</i>	<i>KRAS</i>	<i>MAP2K1</i>	<i>MAP2K2</i>	<i>MET</i>	<i>MYC</i>	<i>PDGFRA</i>
<i>MLL1</i>	<i>MPL</i>	<i>MYC</i>	<i>NF1</i>	<i>NFE2L2</i>	<i>NOTCH1</i>	<i>PIK3CA</i>	<i>RAF1</i>
<i>NPM1</i>	<i>NRAS</i>	<i>NTRK1</i>	<i>POGFR</i>	<i>PIK3CA</i>	<i>PTEN</i>		
<i>PTPN11</i>	<i>RAF1</i>	<i>RET</i>		<i>RHEB</i>	<i>RHOA</i>		
<i>RIT1</i>	<i>ROS1</i>	<i>SMA4</i>	<i>SMO</i>	<i>SRC</i>	<i>STK11</i>		
<i>TERC</i>	<i>TP53</i>	<i>TSC1</i>	<i>VHL</i>				

\*exones 19 & 20  
\*\*exones 19&20  
\*\*\*exon 14 saltar

**FIG. 8**



**FIG. 9A**

**I. Información genética**

- A. Recoger fuentes de material genético y secuenciación de los materiales (11)
- B. Procesar información genética (12)
  - 1. Variantes genéticas
    - a) Variantes de secuencias
    - b) Variantes de número de copias
  - 2. Frecuencia de variantes genéticas en una muestra que contiene material genético
- C. Separar información del ruido (13)

**II. Diagnóstico del cáncer**

- A. En base a la frecuencia de las variantes de secuencias o nivel de CNV (14)
- B. Fiabilidad de detección de variantes genéticas en el intervalo de ruido (15)

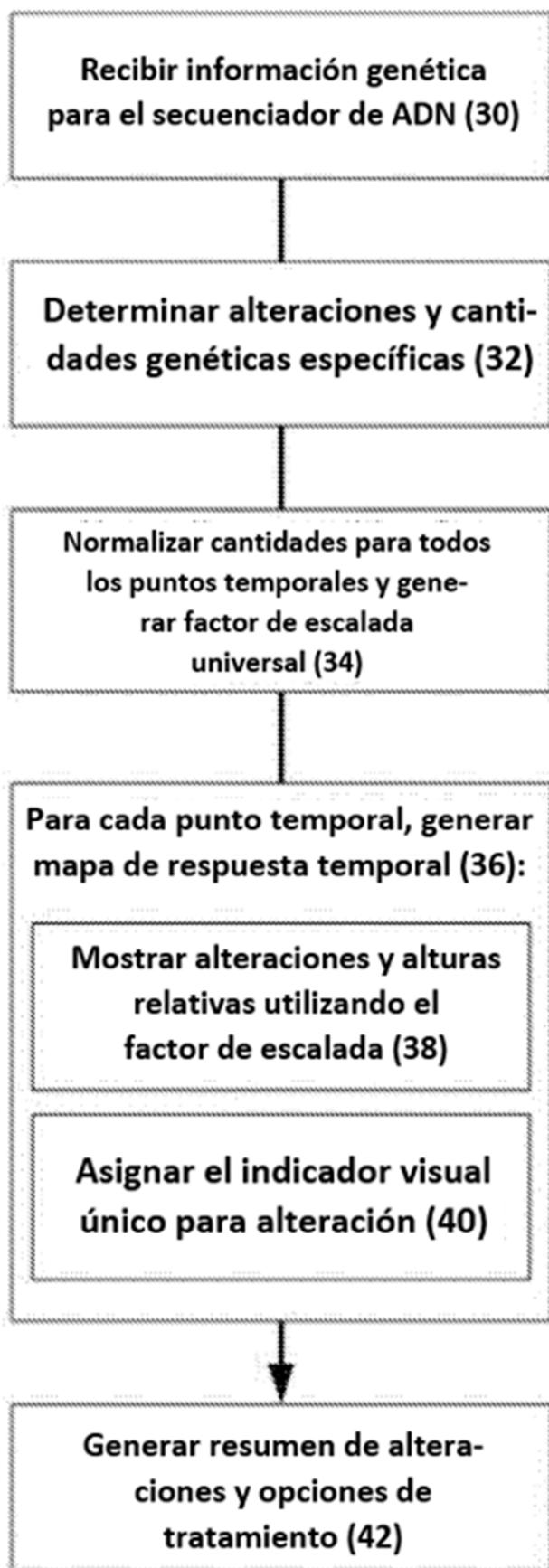
**III. Aumentar fiabilidad de diagnóstico**

- A. Utilizar una pluralidad de mediciones para aumentar la fiabilidad del diagnóstico (16)
  - B. Utilizar mediciones en una pluralidad de puntos temporales para determinar si hay progreso en el cáncer o remisión o estabilización (17)

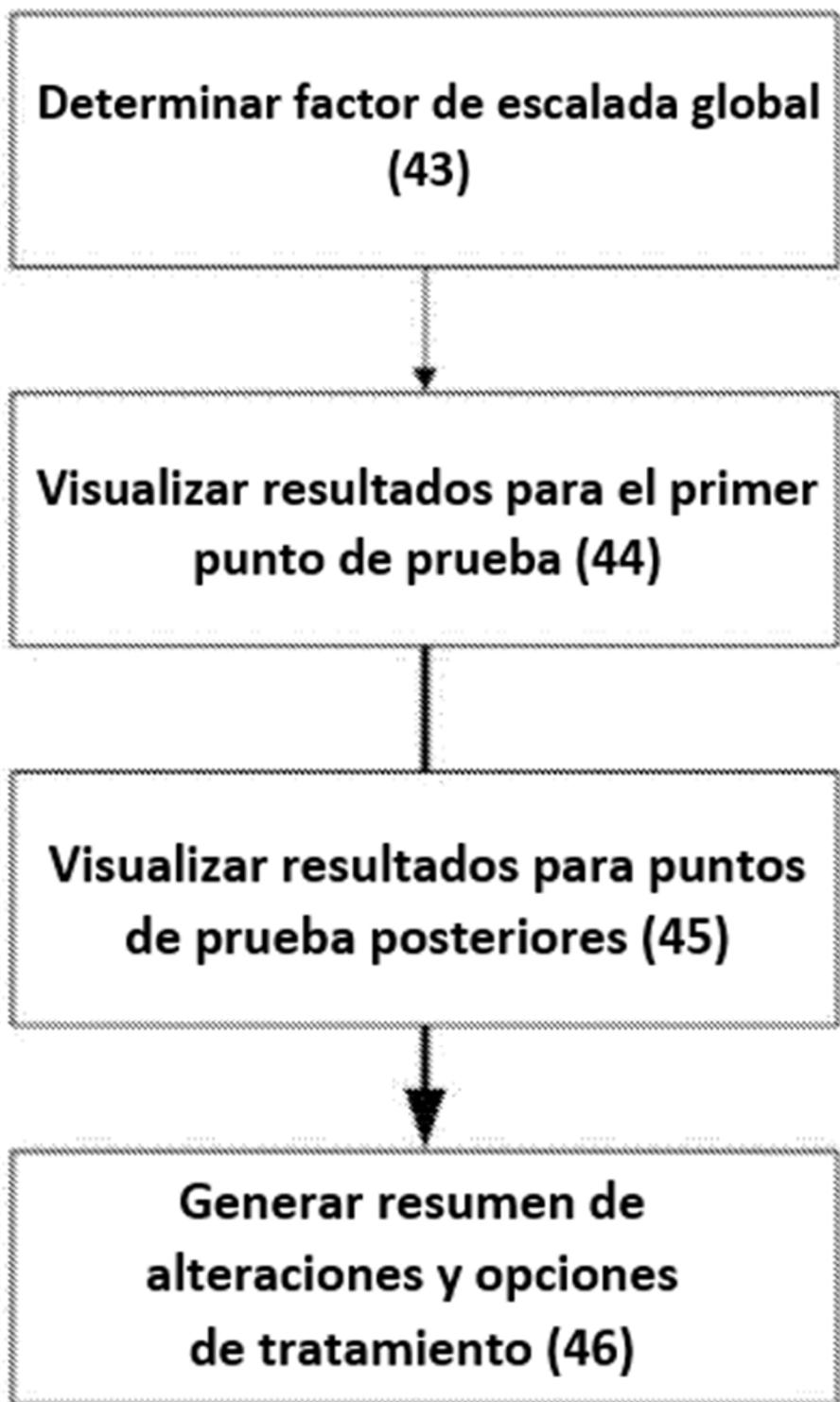
**IV. Generar informe/diagnóstico**

- A. Generar gráfico genético para una pluralidad de mediciones que muestran la tendencia de la mutación (18)
- B. Generar informe que muestra los resultados y opciones del tratamiento (19)

***FIG. 9B***



**FIG. 10A**



**FIG. 10B**

**Para todas las frecuencias de alelos mutantes informadas no de CNV (variación de número de copias)**

Transformar el valor absoluto en una métrica/escala relativa que es más fácil de trazar (p. ej. Multiplicar frecuencia de alelos mutante por 100 y registrar ese valor)

Computar el factor de escala utilizando el valor observado máximo

Utilizar conjunto de datos de prueba más temprano

Para cada alteración no CNV

Multiplicar el factor de escala por el valor transformado para cada gen y utilizar como un indicador de cantidad para trazar esa variante

Asignar un indicador de color/único para cada alteración

Para conjunto de datos para fechas de prueba posteriores

Si hay composición no cambiada de resultados de prueba, continuar vista de fecha de panel anterior en nuevo panel

Si las alteraciones permanecen las mismas, pero han cambiado las cantidades

Recomputar el indicador de cantidades para trazar esa variante y retrazar todos los valores actualizados en panel(es) y nuevo panel para la última fecha de prueba

Si nueva adición de alteraciones

Añadir las alteraciones a todas las alteraciones existentes

Computar los valores de transformación

Recomputar el factor de escala

Retrazar el mapa de respuesta, retrazar alteraciones en la fecha de prueba anterior que todavía están detectadas en la fecha de prueba actual, así como alteraciones de aparición nueva

Si la alteración de existencia anterior no está entre el conjunto de alteraciones detectadas

Utilizar una altura de cero y trazar la cantidad de una alteración para todas las fechas de prueba posteriores

Generar resumen de alteraciones y opciones de tratamiento

Agrupar alteraciones en base a frecuencias de alelos mutantes máximas

**FIG. 10C**

DNI: Johnxxxxxxxxxxxxxxxxxxxxx (XXXXXX)

Fecha nacimiento XX/XX/XXXX | Número de prueba 3

Paciente	Informe	Médico
52 2345678	FECHA DE INFORME I TIEMPO 03/02/2015 a 10:33 h	NOMBRE Johnxxxxxxxxxxxxxxxxxxxxx, MD
GÉNERO Varón	FECHA DE RECEPCIÓN 02/20	CUENTA Pleasantville Oncologyxxxxxxxxxxxxxxxxxxxxx Filadelfia PA 12345
DIAGNÓSTICO CLÍNICO Cáncer de mama	FECHA DE RECOGIDA 02/19	TELÉFONO (888) 888-8888 FAX (888) 888-8888
	MUESTRA Sangre	NOMBRES ADICIONALES Jane Smith, MD John Smith, Title

## Mapa de respuesta tumoral 54

Notas: Response Map ilustra los cambios genéticos detectados en diferentes puntos de la muestra de sangre.



NOTA: Somatic alterations may be present that are below the limit of detection of this test. Certain samples or certain characteristics may result in reduced analytic sensitivity, such as poor sample quality or incomplete resection. Corrected alterations in a tumor may be present, but also not detected in circulating cell-free DNA from this blood specimen with this test.

## Resumen de alteraciones y opciones de tratamiento asociadas 60

The percentage of genomic alterations found in cell-free DNA circulating in blood is related to the clinical status/biology of this patient. Factors that may affect the percentage of detected genomic alterations in circulating cell-free DNA in blood include tumor growth, age over 60, heterogeneous cancer variant, disease progression, or treatment.

Alteración	Tendencia de mutación,	Frecuencia de alelos mutante	Amplificación libre de células	Aprobado en indicación	Disponible para uso en otras indicaciones	Ensayos de dosis clínica
ALK 62	NP	2%				
PROC 64	N139S	20%				
R81	D878N	20%				
APF	G329E	20%				
FBXW7	P319Q	20%				

FIG. 10D

XXXX, John Smith, MD

Fecha nacimiento XX/XX/XXXX | Número de prueba 3

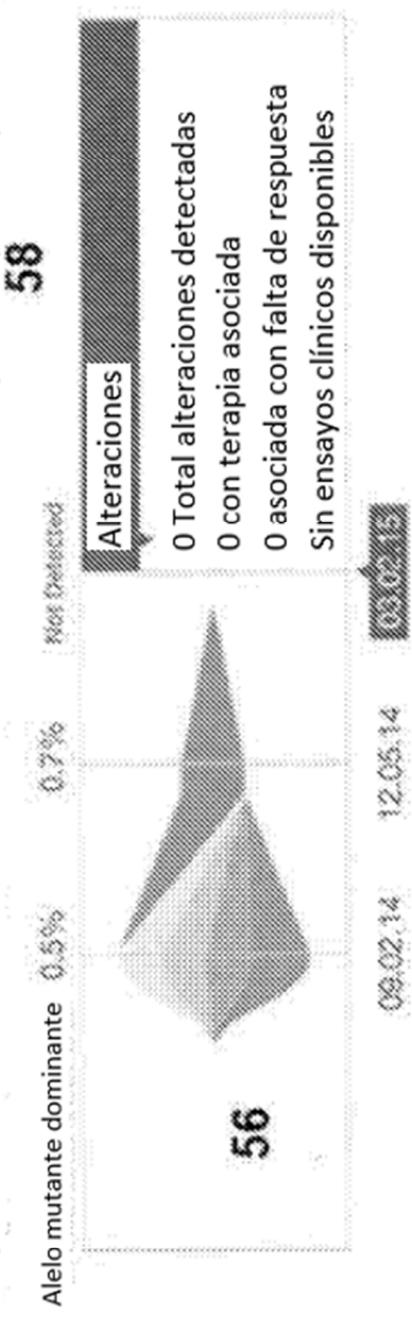
**Paciente**52  
2345678GÉNERO  
VarónDIAGNÓSTICO CLÍNICO  
Cáncer de mama**Informe**FECHA DE INFORME I TIEMPO  
03/02/2015 a 10:33 hFECHA DE RECEPCIÓN  
02/20FECHA DE RECOGIDA  
02/19MUESTRA  
Sangre**Médico**NOMBRE  
John Smith, MDCUENTA  
Philadelphia Oncology XXXXXXXXFiladelfia, PA, 12345  
TELÉFONO  
(888) 888-8888

NOMBRES ADICIONALES

Jane Smith, MD  
Josh Smith, TUE**FIG. 10D-1**

## Mapa de respuesta tumoral 54

El mapa de respuesta tumoral ilustra los cambios relativos de ctADN observado en diferentes puntos temporales de presentación de muestras:



Nota: Alteraciones genómicas pueden estar presentes que están por debajo del límite de detección de esta prueba. Ciertas características de muestras o variantes pueden resultar en sensibilidad analítica reducida, tal como mala calidad de muestra o mala recogida. Pueden estar presentes alteraciones genómicas en un tumor, pero no se detectan en ADN libre de células circulante de la muestra de sangre con esta prueba.

**FIG. 10D-2**

**60****62****64****66****68****70****72****74****76****78****80****82****84****86****88****90****92****94****96****98****100****102****104****106****108****110****112****114****116****118****120****122****124****126****128****130****132****134****136****138****140****142****144****146****148****150****152****154****156****158****160****162****164****166****168****170****172****174****176****178****180****182****184****186****188****190****192****194****196****198****200****202****204****206****208****210****212****214****216****218****220****222****224****226****228****230****232****234****236****238****240****242****244****246****248****250****252****254****256****258****260****262****264****266****268****270****272****274****276****278****280****282****284****286****288****290****292****294****296****298****300****302****304****306****308****310****312****314****316****318****320****322****324****326****328****330****332****334****336****338****340****342****344****346****348****350****352****354****356****358****360****362****364****366****368****370****372****374****376****378****380****382****384****386****388****390****392****394****396****398****400****402****404****406****408****410****412****414****416****418****420****422****424****426****428****430****432****434****436****438****440****442****444****446****448****450****452****454****456****458****460****462****464****466****468****470****472****474****476****478****480****482****484****486****488****490****492****494****496****498****500****502****504****506****508****510****512****514****516****518****520****522****524****526****528****530****532****534****536****538****540****542****544****546****548****550****552****554****556****558****560****562****564****566****568****570****572****574****576****578****580****582****584****586****588****590****592****594****596****598****600****602****604****606****608****610****612****614****616****618****620****622****624****626****628****630****632****634****636****638****640****642****644****646****648****650****652****654****656****658****660****662****664****666****668****670****672****674****676****678**

Doe, John ██████████ (AC000001)

Fecha nacimiento: XX/XX/XXXX | Nº Prueba 3 | Estado: CORREGIDO

Secuenciación de tumores libres de biopsia

El gráfico anterior apunta el porcentaje de la frecuencia de alelos de ADN libre de células circulante alterado (% cfADN) detectado en este paciente. Las alteraciones genómicas detectadas están enumeradas en orden descendente por % cfADN por gen.

Las columnas "indicación aprobada" y "aprobado en otra indicación" describen fármacos asociados con alteraciones genómicas específicas. Se basa en información públicamente disponible como se ha descrito en el apéndice de este informe.

#### Definiciones

**Número de copias:** La amplificación fue detectada para este gen en el ADN libre de células circulante del paciente. La prueba solo detecta amplificación en ciertos genes con asterisco en la Tabla 3. Esta prueba no comprueba otras alteraciones de número de copias en otros genes.

**Delección (del):** La prueba detecta cortas delecciones en exones 19 y 20 del gen EGFR. Esta prueba no comprueba delecciones en otros genes.

**Fusión:** La prueba detecta fusiones de los genes ALX, NTRK3, RET y ROS1 y todos los genes compañeros. Esta prueba no comprueba fusiones de otros genes.

**^Denota mutaciones genómicas sinónimas que no resultan en cambios en aminoácidos**

**Comentarios** EGFR 4.747t: Esta alteración genómica codifica una sustitución sinónima que no resulta en un cambio en el aminoácido en esta posición. Aunque no se espera que se altere la estructura de proteína, no se conocen las consecuencias funcionales y la significación clínica exacta de esta variante. No se conoce la relevancia de las terapias disponibles y ensayos clínicos que se dirigen a esta alteración.

#### Firmado electrónicamente

Arthur Garcia, MD PhD, 05/27/2016 14:53

#### Interpretación

Las alteraciones genómicas fueron detectadas en el ADN libre de células circulante aislado de la muestra de sangre del paciente. Estas alteraciones genómicas son variantes somáticas asociadas con cáncer, algunas de las cuales se han asociado con la respuesta clínica aumentada o reducida a tratamientos específicos.

Todos los genes enumerados en la Tabla X fueron analizados como parte de la prueba y alteraciones genómicas fueron detectadas solo en los genes enumerados en la primera página de este informe. Las alteraciones genómicas no fueron detectadas en otros genes enumerados en la Tabla X.

La amplificación fue detectada en el ADN libre de células circulante aislado de la sangre de este paciente para el (los) gen(es) anotado(s). La prueba solo detecta la amplificación en los genes con asterisco en la Tabla X por análisis de los datos de secuenciación de próxima generación en el ADN libre de células circulante. Esta prueba no comprueba otras alteraciones de número de copias en otros genes. Esta prueba no sustituye otros métodos basados en tejido establecidos que detectan la amplificación de genes, por ejemplo, por inmunohistoquímica (IHQ) o FISH. A diferencia de las pruebas de amplificaciones basadas en tejido (IHQ o FISH), una prueba positiva representa la amplificación media para el gen interrogada en todo el ADN libre de células circulante presente en la muestra de sangre del paciente. Por ejemplo, una prueba positiva podría representar una pequeña población de células con niveles extremadamente altos de la amplificación de genes detectada. Alternativamente, podría representar una población grande de células con niveles bajos a medios de la amplificación de genes. La exacta correlación entre la amplificación obtenida por la prueba en comparación con el IHQ o FISH, y cómo cada prueba guía la gestión del paciente diferencialmente, es un área de investigación activa.

**FIG. 10E**

Doc#\_Johnxxxxxxxxxxxxxx (A0000001)

**Fecha nacimiento:** xx/xx/xxxx | **Nº Prueba** 3 | **Estado:** CORREGIDO

**El gráfico anterior apunta el porcentaje de la frecuencia de alelos de ADN libre de células circulante alterado (% cfADN) detectado en este paciente. Las alteraciones genómicas detectadas están enumeradas en orden descendiente por % cfADN por gen.**

**Las columnas “indicación aprobada” y “aprobado en otra indicación” describen fármacos asociados con alteraciones genómicas específicas. Se basa en información públicamente disponible como se ha descrito en el apéndice de este informe.**

Definiciones

**Número de copias:** La amplificación fue detectada para este gen en el ADN libre de células circulante del paciente. La prueba solo detecta amplificación en ciertos genes con asterisco en la Tabla 3. Esta prueba no comprueba otras alteraciones de número de copias en otros genes.

**Delección (del):** La prueba detecta cortas delecciones en exones 19 y 20 del gen EGFR. Esta prueba no comprueba delecciones en otros genes.

**Fusión:** La prueba detecta fusiones de los genes ALX, NTRK3, RET y ROS1 y todos los genes compañeros. Esta prueba no comprueba fusiones de otros genes.

<sup>a</sup>Denota mutaciones sénómicas sinónimas que no resultan en cambios en aminoácidos.

FIG. 10E-1

**Comentarios**

**EGFR 4.747t:** Esta alteración genómica codifica una sustitución sinónima que no resulta en un cambio en el aminoácido en esta posición. Aunque no se espera que se altere la estructura de proteína, no se conocen las consecuencias funcionales y la significación clínica exacta de esta variante. No se conoce la relevancia de las terapias disponibles y ensayos clínicos que se dirigen a esta alteración.

**Firmado electrónicamente****Arthur Bacá, MD PhD, 01/21/2015 14:53****Interpretación**

Las alteraciones genómicas fueron detectadas en el ADN libre de células circulante aislado de la muestra de sangre del paciente. Estas alteraciones genómicas son variantes somáticas asociadas con cáncer, algunas de las cuales se han asociado con la respuesta clínica aumentada o reducida a tratamientos específicos.

Todos los genes enumerados en la Tabla X fueron analizados como parte de la prueba y alteraciones genómicas fueron detectadas solo en los genes enumerados en la primera página de este informe. Las alteraciones genómicas no fueron detectadas en otros genes enumerados en la Tabla X.

La amplificación fue detectada en el ADN libre de células circulante aislado de la sangre de este paciente para el (los) gen(es) anotado(s). La prueba solo detecta la amplificación en los genes con asterisco en la Tabla X por análisis de los datos de secuenciación de próxima generación en el ADN libre de células circulante. Esta prueba no comprueba otras alteraciones de número de copias en otros genes. Esta prueba no sustituye otros métodos basados en tejido establecidos que detectan la amplificación de genes, por ejemplo, por inmunohistoquímica (IHQ) o FISH. A diferencia de las pruebas de amplificaciones basadas en tejido (IHQ o FISH), una prueba positiva representa la amplificación media para el gen interrogada en todo el ADN libre de células circulante presente en la muestra de sangre del paciente. Por ejemplo, una prueba positiva podría representar una pequeña población de células con niveles extremadamente altos de la amplificación de genes detectada. Alternativamente, podría representar una población grande de células con niveles bajos a medios de la amplificación de genes. La exacta correlación entre la amplificación obtenida por la prueba en comparación con el IHQ o FISH, y cómo cada prueba guía la gestión del paciente diferencialmente, es un área de investigación activa.

**FIG. 10E-2**

Día, 30 hora 00:00:00:00:00:00 (A000001)  
 Fecha nacimiento: XX/XX/XXXX | Nº Prueba 3 | Estado: CORREGIDO

## Resultados de terapia detallados

Alteración	Fármaco	Nombre comercial	Diana	Estado actual
EGFR T790M Amplificación	Neratinib	Egfr/Her2/Erbb4 Inhibidor		Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 3 (cáncer de mama)
	CO-1626	Egr-T790M Inhibidor		Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 2 (cáncer de pulmóns)
	Paratumumab	EGFR Anticuerpo inhibidor		Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Aprobado por FDA en otras indicaciones (carcinoma colorectal)
	AZD9331	Egr-T790M Inhibidor		Fase 3 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 3 (cáncer de pulmóns)
	Atazan	Gefitinib	EGFR/ERBB1/ERBB2/ERBB3/ERBB4 Inhibidor	Aprobado por FDA en la indicación (carcinoma de pulmóns de células no pequeñas (NSCLC)) Aprobado por FDA en la indicación (cáncer pulmonar)
	Lapatinib	Tkisib	Egr-Anticuerpo inhibidor	Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Aprobado por FDA en las indicaciones (cáncer de mama)
	Dacomitinib		Egr-T790M Inhibidor	Fase 3 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 3 (cáncer de pulmóns)
	Nimotuzumab	Therizide	EGFR/Anticuerpo inhibidor	Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 3 (carcinoma gástrico, adeno carcinoma de la nasofaringe, carcinoma de células escamosas de cabeza y cuello (HNSCC). Carcinoma de unión gastroesofágica. Carcinoma esofágico. Carcinoma de células escamosas de la lengua).
	Cetuximab	Egrtuz	EGFR/Anticuerpo inhibidor	Fase 3 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Aprobado por FDA en otras indicaciones. Carcinoma de células escamosas de cabeza y cuello (HNSCC) Carcinoma de células escamosas de la lengua. Carcinoma colorectal
	Osimertinib	OSIRIB	ALK/HER1/HER2 Inhibidor	Aprobado por FDA en las indicaciones. (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 2 (cáncer gástrico, carcinoma celular renal).
MET Amplificación	LY2609623		Inhibidor de multi-quinasa dirigido a Met, Ros1, Axl, Flt3 y otras proteínas	Fase 1 (tumor sólido)
	BMS-777507		Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR1,2,3/Fit3/Tie2/TrkB/Axl	Fase 2 (cáncer gástrico, carcinoma de próstata. Carcinoma celular renal. Carcinoma de células escamosas de cabeza y cuello (HNSCC) Carcinoma de células escamosas cutáneas. Carcinoma esofágico
	Cabozantinib	Comosig	Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR1,2,3/Fit3/Tie2/TrkB/Axl	Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Aprobado por FDA en otras indicaciones. (carcinoma medular de la tiroides).
	Ponatinib		Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR2	Fase 2 (carcinoma de pulmóns de células no pequeñas (NSCLC)) Fase 2 (cáncer gástrico, carcinoma hepatocelular. Carcinoma celular renal. Carcinoma de células escamosas de cabeza y cuello

FIG. 10F

## Resultados de terapia detallados

Alteración	Fármaco	Nombre comercial	Diana	Estado actual
<b>EGFR T790M</b> Amplificación	<b>Neratinib</b>		<i>Egfr/Her2/ErbB4</i> Inhibidor	Fase 2 (carcinoma pulmonar de células no pequeñas (NSCLC)) Fase 3 (cáncer de mama)
	<b>CO-1686</b>		<i>EGFR T790M</i> Inhibidor	Fase 2 (carcinoma pulmonar de células no pequeñas (NSCLC)) Fase 2 (cáncer de los pulmones)
	<b>Panitumumab</b>	<b>Vectibix</b>	<i>EGFR</i> Anticuerpo inhibitorio	Fase 2 (carcinoma pulmonar de células no pequeñas (NSCLC)) Aprobado por la FDA en otras indicaciones (carcina colorectal)
	<b>AZD9291</b>		<i>EGFR T790M</i> Inhibidor	Fase 3 (carcinoma pulmonar de células no pequeñas (NSCLC)) Fase 3 (cáncer de los pulmones)
	<b>Afatinib</b>		<b>Gilotrif</b> <b>Inhibidor de quinasa pan-ErbB irreversible</b>	Aprobado por la FDA en esta indicación (carcina pulmonar de células no pequeñas (NSCLC)) Aprobado por la FDA en otras indicaciones (cáncer de los pulmones)
	<b>Lapatinib</b>	<b>Tykerb</b>	<i>EGFR</i> Anticuerpo inhibitorio	Fase 2 (carcinoma pulmonar de células no pequeñas (NSCLC)) Aprobado por la FDA en otras indicaciones (cáncer de mama)
	<b>Dacomitinib</b>			Fase 3 (carcinoma pulmonar de células no pequeñas (NSCLC)) Fase 3 (cáncer de los pulmones)
	<b>Nimotuzumab</b>	<b>Theraloc</b>		Fase 2 (carcinoma pulmonar de células no pequeñas (NSCLC)) Fase 3 (carcinoma gástrico, Adeno carcinoma de la nasofaringe, glioblastoma, carcinoma de células escamosas de cabeza y cuello (HNSCC), carcinoma de unión gastroesofágico, carcinoma esofágico, carcinoma de céulas escamosas de la lengua)
	<b>Cetuximab</b>		<i>EGFR</i> Anticuerpo inhibitorio	Fase 3 (carcinoma pulmonar de células no pequeñas (NSCLC)) Aprobado por la FDA en otras indicaciones (carcina de células escamosas de cabeza y cuello (HNSCC), carcinoma de céulas escamosas de la lengua), carcinoma colorectal
	<b>Erbitux</b>			

**FIG. 10F-1**

MET Amplificación	Crizotinib	Xalkori	ALK/Met Inhibidor de quinasa	Aprobado por la FDA en esta indicación (carcinoma de pulmones de células no pequeñas (NSCLC)) Fase 2 (carcinaoma gástrico, carcinoma celular renal)
	LY2801653	Inhibidor de multi-quinasa dirigido a Met, Ros1, Axl, Flt3 y otras proteínas	Fase 1 (tumor sólido)	
BMS-777607		Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR1,2,3/Flt3/Tie2/TrkB/Axl	Fase 2 (carcinaoma gástrico, carcinoma de próstata, carcinoma celular renal, carcinoma de células escamosas de cabeza y cuello (HNSSC), carcinoma de células escamosas cutáneas, carcinoma esofágico)	
	Cometriq	Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR1,2,3/Flt3/Tie2/TrkB/Axl	Fase 2 (carcinoma de pulmón de células no pequeñas (NSCLC)) Aprobado por FDA en otras indicaciones (carcinaoma medular de la tiroides)	
Cabozantinib		Inhibidor de quinasa de molécula pequeña de Met/Ret/VEGFR2	Fase 2 (carcinoma de pulmón de células no pequeñas (NSCLC)) Fase 2 (carcinaoma gástrico, carcinoma hepatocelular, carcinoma celular renal, carcinoma de células escamosas de cabeza y cuello)	
	Faretinib			Continuado en la próxima página

FIG. 10F-2

Dosis, Jó hxxxxxxxxxxxxxx (3000001)

Fecha nacimiento: XX/XX/XXXX | Nº Prueba 3 | Estado: CORREGIDO

## Relevancia clínica de alteraciones detectadas

Alteración	Papel en enfermedad	Efecto en sensibilidad farmacéutica	Efecto en resistencia a fármacos
<b>EGFR</b> T790M Amplificación		18784101, Van Cutsem et al., 2007, 17470858). En base a la evidencia acumulada, la American Society for Clinical Oncology (ASCO) ha emitido una Opinión Clínica Provisional recomendando análisis mutacional EGFR para pacientes NSCLC para predecir el beneficio de Egfr TKIs (Keedy et al., 2011; 21482992).	
<b>MET</b> Amplificación	Activación o sobreexpresión de proteína Met fomenta angiogenesis, resistencia a apoptosis, proliferación e invasión de células cancerosas (Appleman, 2011; 22042966, Jung et al., 2012; 22553051, Gherardi et al., 2012, 22270953, Tsuchi et al. 2003; 12684423). Expresión de proteína Met en NSCLC se ha asociado con una predisposición al desarrollo de metástasis cerebral (Benadettini et al., 2020; 20489150).	Expresión MET aumentada posiblemente como resultado de mutación MET o amplificación, puede llevar a activación MET mejorada y por tanto puede otorgar sensibilidad a inhibidores MET (Cecchi et al., 2010, 20303741). Crizotinib y cabozantinib se dirigen a múltiples quinasas, incluyendo Met. Crizotinib ha sido aprobado por la FDA para uso en cáncer pulmonar de células no pequeñas positivo de EML4-ALK y cabozantinib ha sido aprobado por la FDA para el cáncer medular de la tiroídeos (Camidge et al., 2011; ASCO 2011, Abstract 2501, Bang et al., 2010; ASCO 2010, Abstract 3, Traynor, 2013; 23292257, Hart y De Boer, 2013, 23319867). Informes en múltiples tipos de tumores sugieren que la amplificación de MET puede predecir la sensibilidad a crizotinib (Lennerz et al., 2011; 22042942, Chi et al., 2012 22162573, Ou et al., 2011; 21623265), inhibidores específicos a Met están en desarrollo clínico, incluyendo Tivantinib (ARQ 197), un inhibidor MET competitivo selectivo no ATP y MET dirigido a anticuerpo monoclonal de onartuzumab (Adjet et al., 2011; 21632449, Bendell et al., 2013; 23810377; Eng et al., 2013; ASCO 2013, Abstract 3508).	Amplificación Met o expresión elevada de MET ha sido implicada en resistencia adquirida a inhibidores Egfr en algunos tipos de cáncer; los estudios actualmente investigan la terapia en combinación con inhibidores de MET e inhibidores de Egfr en este contexto (Krumbach et al., 2011; 21273060; Engelman et al., 2007; 17463250, Chan et al, 2013; 23527257); activación de Met ha sido implicada como un mecanismo clave de resistencia a terapia dirigida a Egfr en NSCLC (Engelman, et al., 2007; 17463250, La et al., 2013, 24167634, Benedettini et al.)
<b>NOTCH1</b> P24086	Dependiendo del contexto celular, NOTCH1 puede actuar como supresor tumoral u oncogen (Wang et al., 2011; 22006338, Klinakis et al., 2011, 21562564). La actividad de NOTCH1 ha sido informada para ser requerida para tumorigenesis pulmonar en modelos NSCLC impulsados por K-Ras, y ha sido relacionado con el fenotipo de células madre (Licitra et al., 2013; 23943799, Baumgart et al., 2013; 24509876, Hassan et al., 2013; 23444212, Allen et al., 2011; 21803744).	Mutaciones activadoras de NOTCH1 estabilizan la proteína intracellular Notch y llevan al aumento de señales Notch1 (Weng et al, 2004; 15472075). Inhibidores Notch1 (es decir, inhibidores de gamma-secretasa, que evitan escisión del dominio intracellular) pueden ser un enfoque terapéutico potencial en el caso de mutaciones activadoras de NOTCH1 y estas están en ensayos clínicos para varios cánceres (Fouad et al., 2011; 21825264, Groth y Fortini, 2012; 22309842). Inhibidores de gamma-secretasa han mostrado eficacia en modelos tumorales preclínicos con activación de Notch (Westhort et al., 2009, 20007775, Arcani et al., 2013; 23868008). Algunos estudios sugieren que los inhibidores HDAC, incluyendo LBH589 y ácido valproico pueden reactivar señales de vía Notch en algunos tumores que han perdido la expresión de Notch (Egloff y Grandis, 2012; 22773529, Platta et al., 2008; 18570928). Sin embargo, este enfoque es relevante en el contexto de silenciamiento oligonéutico y no se espera que sea relevante en el caso de mutación inactivadora de NOTCH1. En el caso de esta variante no caracterizada, no se conoce la relevancia de cualesquier enfoques terapéuticos disponibles.	En un estudio preclínico utilizando líneas celulares de NSCLC y xenoinjertos, se ha sugerido que Notch1 tenga un papel en la resistencia adquirida a inhibidores de tirosina quinasa como gefitinib (Xia et al., 2013; 23916913, Xia et al., 2012; 22173954). Adicionalmente, las líneas celulares NSCLC con Notch1 alto o activado han sido informadas a ser resistentes a quimioterapias, incluyendo cisplatino y docetaxel (Hassan et al., 2013; 23444212). Liu et al., 2013; 20136908

**FIG. 10G**

## Relevancia clínica de alteraciones detectadas

Alteración	Papel en enfermedad	Efecto en sensibilidad farmacéutica	Efecto en resistencia a fármacos
<b>EGFR</b> <b>T790M</b> Amplificación		18784101, Van Cutsem et al., 2007, 17470858). En base a la evidencia acumulada, la American Society for Clinical Oncology (ASCO) ha emitido una Opinión Clínica Provisional recomendando análisis mutacional EGFR para pacientes NSCLC para predecir el beneficio de Egfr TKIs (Keedy et al., 2011; 21482992).	
<b>MET</b> Amplificación	Activación o sobreexpresión de proteína Met fomenta angiogenesis, resistencia a apoptosis, proliferación e invasión de células cancerosas (Appleman, 2011; 22042966, Jung et al., 2012; 22553051, Gherardi et al., 2012, 22270953, Taksuchi et al. 2003; 12684423). Expresión de proteína Met en NSCLC se ha asociado con una predisposición al desarrollo de metástasis cerebral (Benadettini et al., 2020; 20489150).	Expresión MET aumentada posiblemente como resultado de mutación MET o amplificación, puede llevar a activación MET mejorada y por tanto puede otorgar sensibilidad a inhibidores MET (Cecchi et al., 2010, 20303741). Crizotinib y cabozantinib se dirigen a múltiples quinasas, incluyendo Met. Crizotinib ha sido aprobado por la FDA para uso en cáncer pulmonar de células no pequeñas positivo de EML4-ALK y cabozantinib ha sido aprobado por la FDA para el cáncer medular de la tiroides (Camidge et al., 2011; ASCO 2011, Abstract 2501, Bang et al., 2010, ASCO 2010, Abstract 3, Traynor, 2013; 23292257, Hart y De Boer, 2013, 23319867). Informes en múltiples tipos de tumores sugieren que la amplificación de MET puede predecir la sensibilidad a crizotinib (Lennerz et al., 2011; 22042942, Chi et al., 2012 22162573, Ou et al., 2011; 21623265), inhibidores específicos a Met están en desarrollo clínico, incluyendo Tivantinib (ARQ 197), un inhibidor MET competitivo selectivo no ATP y MET dirigido a anticuerpo monoclonal de onartuzumab (Adjet et al., 2011; 21632449, Bendell et al., 2013; 23810377; Eng et al., 2013; ASCO 2013, Abstract 3508).	Amplificación Met o expresión elevada de MET ha sido implicada en resistencia adquirida a inhibidores Egfr en algunos tipos de cáncer; los estudios actualmente investigan la terapia en combinación con inhibidores de MET e inhibidores de Egfr en este contexto (Krumbach et al., 2011; 21273060; Engelman et al., 2007; 17463250, Chan et al, 2013; 23527257); activación de Met ha sido implicada como un mecanismo clave de resistencia a terapia dirigida a Egfr en NSCLC (Engelman, et al., 2007; 17463250, La et al., 2013, 24167634, Benedettini et al.)

**FIG. 10G-1**

**NOTCH1  
P2008**

Dependiendo del contexto celular, NOTCH1 puede actuar como supresor tumoral u oncogen (Wang et al., 2011; 22006338, Klinakis et al., 2011, 21562564). La actividad de NOTCH1 ha sido informada para ser requerida para tumorigenesis pulmonar en modelos NSCLC impulsados por K-Ras, y ha sido relacionado con el fenotipo de células madre (Licciulli et al., 2013; 23943799, Baumgart et al., 2013; 24509876, Hassan et al., 2013; 23444212, Allen et al., 2011; 21803744).

**Mutaciones activadoras de NOTCH1** estabilizan la proteína intracelular Notch1 y llevan al aumento de señales Notch1 (Weng et al., 2004; 15472075). Inhibidores Notch1 (es decir, inhibidores de gamma-secretasa, que evitan escisión del dominio intracelular) pueden ser un enfoque terapéutico potencial en el caso de mutaciones activadoras de NOTCH1 y estas están en ensayos clínicos para varios cánceres (Fouladi et al., 2011; 21825264, Groth y Fortini, 2012; 22309842). Inhibidores de gamma-secretasa han mostrado eficacia en modelos tumorales preclínicos con activación de Notch (Westhoff et al., 2009, 20007775, Arcaroli et al., 2013; 23868008). Algunos estudios sugieren que los inhibidores HDAC, incluyendo LBH589 y ácido valproico pueden reactivar señales de vía Notch en algunos tumores que han perdido la expresión de Notch (Egloff y Grandis, 2012; 22773529, Platta et al., 2008; 18570928). Sin embargo, este enfoque es relevante en el contexto de silenciamiento epigenético y no se espera que sea relevante en el caso de mutación inactivadora de NOTCH1. En el caso de esta variante no caracterizada, no se conoce la relevancia de cualesquier enfoques terapéuticos disponibles.

En un estudio preclínico utilizando líneas celulares de NSCLC y xeno-injertos, se ha sugerido que Notch1 tenga un papel en la resistencia adquirida a inhibidores de tirosina quinasa como gefitinib (Xia et al., 2013; 23916913, Xia et al., 2012; 22173954). Adicionalmente, las líneas celulares NSCLC con Notch1 alto o activado han sido informadas a ser resistentes a quimioterapias, incluyendo cisplatino y docetaxel (Hassan et al., 2013; 23444212).

**FIG. 10G-2**

Dos, Johnxxxxxxxxxxxxx (A000001)  
 Fecha nacimiento: XX/XX/XXXX | Nº Prueba 3 | Estado: CORREGIDO

### Ensayos clínicos disponibles

Alteración	Ensayo ID	Título	Fase	Sitio
EGFR T790M Amplificación	NCT01647711	Un estudio de Afatinib a dosis alta intermitente para determinar la dosis tolerada máxima y evaluar la actividad de esta dosis contra cáncer de pulmón de células no pequeñas con mutaciones T790M	Fase 1	Aurora, CO; Boston, MA {2}
	NCT01526928	Un estudio para evaluar la seguridad, farmacocinética y eficacia de rociletinib (CD-1686) en receptor de factor de crecimiento epidérmico mutante tratado previamente (EGFR) en pacientes del cáncer pulmonar de células no pequeñas (NSCLC)	Fase 1/Fase 2	Duarte, CA; Fountain Valley, CA; Los Angeles, CA (2); Orange, CA; Sacramento, CA; Santa Monica, CA; Stanford, CA; Whittier, CA (2); Aurora, OD; Washington, DC; Miami, FL; Orlando, FL; Atenas, GA; Boston, MA (2); Ann Arbor, MI; Detroit, MI; Morristown; Buffalo, NY; Lake Success, NY; Nueva York, NY; Cincinnati, OH; Columbus, OH; Tulsa, OH; Portland; OR; Filadelfia, PA; Nashville, TN; Houston, TX.
MET Amplificación	NCT01324479	Estudio de INC280 en pacientes con tumores sólidos avanzados dependientes de c-MET	Fase 1	Fayetteville, AK; Los Angeles, CA; Chicago, IL; Rockville, MD; Detroit, MI; Buffalo, NY; Nashville, TN; Dallas, TX; Houston, TX; Salt Lake City UT; Australia (2); Canadá (2); Francia (3); Alemania (6); Hong Kong (2); Israel (4); Italia (6); Corea (5); Países Bajos (3); Noruega; Singapur; Granada; España (5); Taiwán (2); Tailandia (3)
	NCT01712217	Un estudio de AT13387 en pacientes con cáncer pulmonar de células no pequeñas (NSCLC) solo y en combinación con Crizotinib	Fase 1/Fase 2	Scottsdale, AZ; La Jolla, CA; Los Ángeles, CA; San Diego, CA; Santa Monica, CA; Whittier, CA; Aurora, CO; New Haven, CT; Newark, DE; Orlando, FL; Tampa, FL; Chicago, IL (2); Ann Arbor, MI; Detroit, MI; Rochester, MN; St Louis, MO; Omaha, NE; Lebanon, NH; Bronx, NY; Nueva York, NY; Greensboro, NC; Winston Salem, NC; Cincinnati, OH (2); Cleveland, OH; Columbus, OH; Portland, OR; Hershey, PA; Filadelfia, PA; Memphis, TN; Nashville, TN; Dallas, TX; Fairfax, VA; Seattle, WA (2); Madison, WI; Canadá (6); Francia (10); Corea (6); España (5)

**FIG. 10H**

Dos, Jo http://xxxxxxxxxxxxxx@xxxxxxxxx | Fecha nacimiento: xx/xx/XXXX | Nº Prueba 3 | Estado: CORREGIDO

## Métodos y limitaciones

La prueba ensaya un panel de 68 genes para identificar alteraciones genómicas en variantes somáticas asociadas con cáncer alta sensibilidad. El ADN libre de células se extrae de plasma y alteraciones genómicas se analizan por secuenciación masivamente paralela de genes diana amplificados. Para un conjunto de genes, todos los exones se secuencian como tal cobertura de secuenciación se ha mostrado tener utilidad clínica. Para otro conjunto de genes, la cobertura de secuenciación incluye los exones con mutaciones somáticas informadas previamente. Sustituciones sinónimas son informadas si presentes, aunque la exacta significación clínica y relevancia de cualesquier terapias disponibles dirigidas a tales variantes no se definen precisamente. El alelo mutante detectable mínimo (límite de detección) depende de la concentración de ADN libre de células de la muestra del paciente, la cual puede variar de menos de 10 a más de 1000 equivalentes genómicos por mL de sangre periférica. La prueba es validada para solo detectar amplificaciones genéticas en los genes con asterisco en la Tabla X en muestras con al menos 2% de ADN libre de células circulante derivado de tumores que contiene amplificación de 8 – 10 genes; la amplificación puede no ser detectada en muestras con cantidades más bajas de ADN libre de células y/o amplificación de copias de genes de bajo nivel. Los tipos de alteraciones genómicas que pueden ser detectadas por la prueba incluyen variaciones de un solo nucleótido, amplificaciones, fusiones de ALX, NTRK3, RET y ROS1 y cortas inserciones/delecciones presentes en exones 19 y 20 del gen EGFR. Esta versión de la prueba no es validada para la detección de otros tipos de alteraciones genómicas (por ejemplo, delecciones genéticas, marcos de lectura, variantes de escisión o indeles). Ciertas características de muestras o variantes pueden resultar en sensibilidad analítica reducida, tal como baja calidad de muestra o recogida no adecuada.

La Tabla X anota el límite específico a pacientes de detección (LOD) para cada uno de los genes analizados, el cual es dependiente de la concentración de ADN libre de células y cobertura de secuenciación para ese gen.

**Genes en negrita** son los que también se analizan para variaciones de número de copias (CNVs).

\*EGFR indoleos en exons 18 y 20

\* Incluye región promotora

## Reestructuraciones

GEN LOG GEN LOG GEN LOG GEN LOG

FIG. 10I

Doe, Johnxxxxxxxxxxxxxx (A000001)

Fecha nacimiento: XX/XX/XXXX | N° Prueba 3 | Estado: CORREGIDO

### Métodos y limitaciones

La prueba evalúa un panel de 68 genes para identificar alteraciones genómicas en variantes somáticas asociadas con cáncer con alta sensibilidad. ADN libre de células se extrae de plasma y alteraciones genómicas son analizadas por secuenciación masivamente paralela de genes diana amplificados. Para un conjunto de genes, todos los exones son secuenciados, al haber mostrado dicha cobertura de secuenciación tener utilidad clínica. Para otro conjunto de genes, la cobertura de secuenciación incluye todos los exones con una mutación somática informada anteriormente. Las sustituciones sinónimas son informadas si presentes, aunque la significación clínica exacta y relevancia de cualesquiera terapias disponibles no son precisamente definidas. El alelo mutante detectable mínimo (límite de detección) es dependiente de la concentración de ADN libre de células de muestra del paciente, la cual puede variar de menos de 10 a más de 1000 equivalentes genómicos por mL de sangre periférica. La prueba es validada para solo detectar amplificaciones genéticas en los genes con asterisco en Tabla X en muestras con al menos 2% de ADN libre de células circulante derivado de tumores que contienen 8-10x amplificación de genes; amplificación puede no ser detectada en muestras con cantidades más bajas de ADN libre de células y/o amplificación de copia de genes de nivel bajo. Los tipos de alteraciones genómicas que pueden ser detectadas por la prueba incluyen variaciones de nucleótido único, amplificaciones, fusiones ALK, NTRK1, RET y ROS1 y cortas inserciones/delecciones presentes en exones 19 y 20 del gen EGFR. Esta versión de la prueba no es validada para la detección de otros tipos de alteraciones genómicas (por ejemplo, delecciones genéticas, marcos de lectura, variantes de escisión o indeles). Ciertas características de muestras o variantes pueden resultar en sensibilidad analítica reducida, tal como baja calidad de muestra o recogida no adecuada.

### ***FIG. 10I-1***

**Tabla X** anota el límite específico al paciente de detección (LOD) para cada uno de los genes analizados, el cual es dependiente de la concentración de ADN libre de células y cobertura de secuenciación para ese gen

Genes con cobertura completa de exones y parcial de intrones

GEN	LOD	GEN	LOD	GEN	LOD	GEN	LOD
APC	0.1%	AR	0.2%	ARID1A	0.2%	BRAF	0.1%
BRCA1	0.2%	BRCA2	0.2%	CCND1	0.1%	CCND2	0.1%
CCNE1	0.1%	CDK4	0.1%	CDK6	0.2%	CDKN2A	0.1%
CDKN2B	0.1%	EGFR*+	0.1%	ERBB2	<0.1%	FGFR1	0.1%
FGFR2	0.1%	HRAS	0.1%	KIT	0.1%	KRAS	0.2%
MET	0.1%	MYC	<0.1%	NF1	0.2%	NRAS	0.2%
PDGFRA	0.1%	PIK3CA	0.2%	PTEN	0.2%	RAF1	0.1%
TP53	0.2%						

Genes en negrita son los que también se analizan para variaciones de número de copias (CNVs).

\*+EGFR indeles en exon 19 y 20

Genes con cobertura crítica de exones

GEN	LOD	GEN	LOD	GEN	LOD	GEN	LOD
AKT1	<0.1%	ALK	0.1%	ARAF	0.2%	ATM	0.2%
CDH1	0.2%	CTNNB1	<0.1%	ESR1	0.1%	EZH2	0.1%
FBXW7	0.1%	FGFR3	0.1%	GATA3	0.1%	GNA11	0.1%
GNAQ	0.1%	GNAS	0.1%	HNF1A	0.1%	IDH1	0.1%
IDH2	0.1%	JAK2	0.2%	JAK3	0.1%	MAP2K1	0.1%
MAP2K2	0.1%	MLH1	0.2%	MPL	0.1%	NFE2L2	0.1%
NOTCH1	0.1%	NPM1	0.2%	NTRK1	<0.1%	PTPN11	0.1%
RET	<0.1%	RHEB	0.2%	RHOA	0.1%	RIT1	0.3%
ROS1	0.2%	SMAD4	0.2%	SMO	0.1%	SRC	<0.1%
STK11	0.1%	TERT*	0.1%	VHL	0.2%		


\* Incluye región promotora

Reestructuraciones


GEN	LOD	GEN	LOD	GEN	LOD	GEN	LOD
ALK	<0.1%	NTRK1	<0.1%	RET	<0.1%	ROS1	<0.2%

**FIG. 10I-2**

# ES 2 923 602 T3



**FIG. 10J**



**FIG. 10K**





FIG. 10L


01/07/2014

Alelo mutante dominante al 12.5%




**FIG. 10M**

# ES 2 923 602 T3




**FIG. 10N**

# ES 2 923 602 T3




**FIG. 100**


# ES 2 923 602 T3



**FIG. 10P**



**FIG. 11A**



**FIG. 11B**