(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **3 017 068**

(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 22176398 (0)**

(97) Fecha y número de publicación de la concesión europea: **05.02.2025 EP 4123032**

(54) Título: **Detección y tratamiento de enfermedades que presentan heterogeneidad de células pacientes y sistemas y métodos paracomunicar 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:
12.05.2025

(73) Titular/es:

GUARDANT HEALTH, INC. (100.00%)
3100 Hanover Street
Palo Alto, CA 94304, US

(72) Inventor/es:

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

(74) Agente/Representante:

IZQUIERDO BLANCO, María Alicia

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 presentan heterogeneidad de células pacientes y sistemas y métodos para comunicar los resultados de las pruebas

ANTECEDENTES

La atención sanitaria recién ahora está comenzando a utilizar de manera efectiva la información del genoma humano es fundamental para diagnosticar y tratar enfermedades. En ningún ámbito es más importante que en el tratamiento del cáncer, del que mueren 7,6 millones de personas cada año en Estados Unidos y para el que el país gasta 87 mil millones de dólares al año en tratamientos. El cáncer se refiere a cualquier trastorno de diversas neoplasias malignas que se caracteriza por la proliferación de células anaplásicas que tienden a invadir el tejido circundante y hacer metástasis en nuevas zonas del cuerpo y las condiciones patológicas caracterizadas por tales crecimientos.

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

Los cánceres pueden evolucionar con el tiempo y volverse resistentes a una intervención terapéutica. Se sabe que ciertas variantes se correlacionan con la capacidad de respuesta o la resistencia a intervenciones terapéuticas específicas. Sería beneficioso contar con tratamientos más eficaces para los cánceres que presentan heterogeneidad tumoral. Dichos cánceres pueden tratarse con una segunda intervención terapéutica diferente a la que responda el cáncer.

Los métodos de secuenciación de ADN permiten la detección de variantes genéticas en el ADN de las células tumorales. Los tumores cancerosos liberan continuamente su material genómico único en el torrente sanguíneo. Desafortunadamente, estas "señales" genómicas reveladoras son tan débiles que las tecnologías actuales de análisis genómico, incluida la secuenciación de próxima generación, solo pueden detectar dichas señales esporádicamente o en pacientes con una carga tumoral terminalmente alta. La razón principal de esto es que dichas tecnologías están plagadas de tasas de error y sesgo que pueden ser órdenes de magnitud superiores a lo que se requiere para detectar de manera confiable alteraciones genómicas de novo asociadas con el cáncer.

En una tendencia paralela, para comprender la importancia clínica de una prueba genética, los profesionales que tratan deben tener un conocimiento práctico de los principios básicos de la herencia genética y una facilidad razonable con la interpretación de datos probabilísticos. Algunos estudios sugieren que muchos profesionales que tratan no están adecuadamente preparados para interpretar pruebas genéticas para la susceptibilidad a la enfermedad. Algunos médicos tienen dificultad para interpretar datos probabilísticos relacionados con la utilidad clínica de las pruebas diagnósticas, como el valor predictivo positivo o negativo de una prueba de laboratorio.

Las tasas de error y el sesgo en la detección de alteraciones genómicas de novo asociadas con el cáncer, junto con una explicación inadecuada de las implicaciones de las pruebas genéticas para el cáncer, han reducido la calidad de la atención a los pacientes con cáncer. Las sociedades profesionales, como el Colegio Americano de Patólogos (CAP) y el Colegio Americano de Genética Médica (ACMG), han publicado normas o directrices para los laboratorios que proporcionan pruebas genéticas, que exigen que los informes que contienen información genética incluyan un contenido interpretativo que sea comprensible para los médicos generalistas.

El documento WO 2014/165549 (University of Florida Research Foundation, Incorporated) describe la "*determination of methylation state and chromatin structure of target genetic loci*" (título); Sharma et al. (Cell (2010) 141(1):69-80) describen "*a chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations*" (título); el documento WO 2014/043763 (The Chinese University of Hong Kong) describe la "*non-invasive determination of methylome of fetus or tumor from plasma*" (título); y Chan et al. (PNAS (2013) 110(47):18761-8) describen la "*Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing*" (título).

RESUMEN

La invención se expone en el conjunto de reivindicaciones adjunto. En algunas formas de realización, el método comprende, además: (e) controlar los cambios en la heterogeneidad tumoral en el sujeto a lo largo del tiempo y determinar diferentes intervenciones terapéuticas a lo largo del tiempo en función de los cambios. En algunas formas de realización, el método comprende, además: (e) mostrar la intervención terapéutica. En algunas formas de realización, el método comprende, además: (e) generar una filogenia de la evolución tumoral en función del perfil tumoral; en donde la determinación de la intervención terapéutica tiene en cuenta la filogenia.

En algunas formas de realización, la determinación se realiza con la ayuda de un algoritmo ejecutado por computadora. En algunas formas de realización, las lecturas de secuencia generadas por secuenciación se someten a reducción de ruido antes de la identificación y cuantificación. En algunas formas de realización, la reducción de ruido comprende el seguimiento molecular de secuencias generadas a partir de un único polinucleótido en la muestra.

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

En algunas formas de realización, el ADNlc proviene de una muestra de sangre. En algunas formas de realización, los polinucleótidos comprenden polinucleótidos de diferentes sitios de tumores metastásicos.

En algunas formas de realización, la identificación comprende generar una pluralidad de lecturas de secuencia para polinucleótidos parentales de la muestra y colapsar las lecturas de secuencia para generar llamadas de consenso para bases en cada polinucleótido parental. En algunas formas de realización, la cuantificación comprende determinar la frecuencia con la que se detectan las mutaciones somáticas en la población de polinucleótidos de la muestra biológica. En algunas formas de realización, la muestra biológica comprende moléculas biológicas de células no patológicas. En algunas formas de realización, la muestra biológica comprende moléculas biológicas de una pluralidad de tejidos diferentes. En algunas formas de realización, las biomoléculas están comprendidas en una muestra biológica. En algunas formas de realización, las biomoléculas están comprendidas en una pluralidad de muestras biológicas. En algunas formas de realización, la pluralidad de muestras biológicas son tumores de una pluralidad de metástasis.

En algunas formas de realización, la secuenciación comprende la secuenciación de todo o parte de un subconjunto de genes en el genoma del sujeto. En algunas formas de realización, las mutaciones somáticas se seleccionan de entre variaciones de un solo nucleótido (VNS), inserciones, deleciones, inversiones, transversiones, translocaciones, variaciones del número de copias (VNC) (por ejemplo, aneuploidía, aneuploidía parcial, poliploidía), inestabilidad cromosómica, alteraciones de la estructura cromosómica, fusiones genéticas, fusiones cromosómicas, truncamientos genéticos, amplificación genética, duplicaciones genéticas, lesiones cromosómicas, lesiones del ADN, cambios anormales en las modificaciones químicas de los ácidos nucleicos, cambios anormales en los patrones epigenéticos y cambios anormales en la metilación de los ácidos nucleicos. En algunas formas de realización, los loci genéticos se seleccionan de entre nucleótidos individuales, genes y cromosomas.

En algunas formas de realización, 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 cerebral, cáncer de esófago, cáncer de cabeza y cuello, cáncer de vejiga, cáncer ginecológico, liposarcoma y mieloma múltiple). En algunas formas de realización, las células cancerosas del tumor se derivan de una célula patológica parental común. En algunas formas de realización, las células cancerosas del tumor se derivan de diferentes células cancerosas parentales del mismo tipo de cáncer o de diferentes tipos de cáncer. En algunas formas de realización, el método comprende, además determinar una medida de las mutaciones somáticas con respecto a una o más referencias de control para determinar la cantidad relativa.

En algunas formas de realización, los polinucleótidos también se obtienen de una biopsia de tumor sólido. En algunas formas de realización, se desarrollan perfiles por separado para polinucleótidos obtenidos de polinucleótidos de cáncer circulantes y de la biopsia de tumor sólido.

En algunas formas de realización, el perfil tumoral indica la frecuencia relativa de una pluralidad de mutaciones somáticas más. En algunas formas de realización, el método comprende, además monitorear los cambios en las frecuencias relativas en el sujeto a lo largo del tiempo y determinar diferentes intervenciones terapéuticas a lo largo del tiempo en función de los cambios. En algunas formas de realización, la intervención terapéutica es más eficaz contra un cáncer que se presenta con cada una de las mutaciones somáticas que contra un cáncer que se presenta con cualquiera, pero no todas, las mutaciones somáticas. En algunas formas de realización, 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 mutación somática diferente que se produce a una frecuencia relativa diferente. En algunas formas de realización, un fármaco que es relativamente más eficaz contra un cáncer que se presenta con una mutación somática que se produce a una frecuencia relativa más alta se administra en mayor cantidad. En algunas formas de

realización, los fármacos se administran en dosis que se estratifican para reflejar las cantidades relativas de las variantes en el ADN. En algunas formas de realización, los cánceres que presentan al menos una de las variantes genéticas son resistentes a al menos uno de los fármacos. En algunas formas de realización, el cáncer se selecciona entre 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 de esófago, cáncer de cabeza y cuello, cáncer de vejiga, cáncer ginecológico, liposarcoma y mieloma múltiple).

En algunas formas de realización, 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 formas de realización, detectar la heterogeneidad tumoral en un sujeto comprende: a) cuantificar polinucleótidos que portan una variante de secuencia en cada uno de una pluralidad de loci genéticos en el ADN libre de células de una muestra del sujeto, en donde la muestra comprende polinucleótidos de células somáticas y de células enfermas; b) determinar para cada locus una medida de variación del número de copias (VNC) para polinucleótidos que portan la variante de secuencia; c) determinar para cada locus una medida ponderada de la cantidad de polinucleótidos que portan una variante de secuencia en el locus en función de VNC en el locus; y d) comparar las medidas ponderadas en cada uno de la pluralidad de loci, en donde diferentes medidas ponderadas indican heterogeneidad de células enfermas.

En algunas formas de realización, el método comprende: a) determinar una medida de desviación de un valor de tendencia central (por ejemplo, desviación estándar, varianza) del número de copias en polinucleótidos en una muestra a lo largo de una región de al menos 1 kb, al menos 10 kb, al menos 100 kb, al menos 1 mb, al menos 10 mb o al menos 100 mb de un genoma; b) inferir una medida de carga de ADN de células que experimentan división celular en la muestra basándose en la medida de desviación. En algunas formas de realización, el valor de tendencia central es media, mediana o moda. En algunas formas de realización, la determinación comprende dividir la región en una pluralidad de intervalos no superpuestos, determinar una medida del número de copias en cada intervalo y determinar la medida de desviación basándose en medidas del número de copias en cada intervalo. En algunas formas de realización, el intervalo no es mayor que cualquiera de 1 base, 10 bases, 100 bases, 1 kb o 10 kb.

En algunas formas de realización, el método comprende inferir una medida de carga de ADN de células que experimentan división celular en una muestra que comprende medir la variación del número de copias inducida por la proximidad de uno o más loci genómicos a los orígenes de replicación de las células, en donde el aumento de VNC indica células que experimentan división celular. De acuerdo con la invención, la carga se mide en ADN libre de células. En algunas formas de realización, la medida de carga se relaciona con la fracción de células tumorales o equivalentes genómicos de ADN de células tumorales en la muestra. En algunas formas de realización, se infiere VNC debido a la proximidad a los orígenes de replicación a partir de un conjunto de muestras de control o líneas celulares. En algunas formas de realización, se utiliza un modelo de Markov oculto, un modelo de regresión, un modelo basado en análisis de componentes principales o un modelo modificado por genotipo para aproximar las variaciones debidas a los orígenes de replicación. En algunas formas de realización, la medida de la carga es la presencia o ausencia de células que experimentan división celular. En algunas formas de realización, la proximidad está dentro de 1 kb de un origen de replicación.

En algunas formas de realización, el método comprende aumentar la sensibilidad y/o especificidad de la determinación de variaciones en el número de copias relacionadas con genes mejorando el efecto de las variaciones debidas a la proximidad a los orígenes de replicación. En algunas formas de realización, el método comprende medir la VNC en un locus, determinar la cantidad de VNC debida a la proximidad del locus a un origen de replicación y corregir la VNC medida para reflejar la VNC genómica, por ejemplo, restando la cantidad de VNC atribuible a la división celular. De acuerdo con la invención, los datos genómicos se obtienen a partir de ADN libre de células. En algunas formas de realización, la medida de la carga se relaciona con la fracción de células tumorales o equivalentes genómicos de ADN en una muestra. En algunas formas de realización, las variaciones debidas a los orígenes de replicación se infieren a partir de un conjunto de muestras de control o líneas celulares. En algunas formas de realización, se utiliza un modelo de Markov oculto, un modelo de regresión, un modelo basado en análisis de componentes principales o un modelo modificado por genotipo para aproximar las variaciones debidas a los orígenes de replicación.

En algunas formas de realización, el método comprende: a) determinar una medida de referencia de copias de moléculas de ADN en uno o más loci de una o más muestras de control, en donde uno o más de los loci incluyen un origen de replicación, cada uno conteniendo ADN de células que experimentan un nivel predeterminado de división celular; b) determinar una medida de prueba de moléculas de ADN en una muestra de prueba; en donde la medida en la muestra de prueba es de uno o más loci divididos en una o más particiones y en donde uno o más de los loci incluyen un origen de replicación; c) comparar la medida de prueba y la medida de referencia, en donde una medida de prueba por encima de una medida de referencia indica ADN en la muestra de prueba de células que se dividen a una velocidad más rápida que las células que proporcionan ADN a la muestra de control. En algunas formas de realización, la medida se selecciona de recuento de moléculas, una medida de tendencia central del recuento de moléculas a lo largo de las particiones o una medida de variación del recuento de moléculas a lo largo de las particiones.

En algunas formas de realización, el método comprende compilar una base de datos, en donde la base de datos incluye, para cada uno de una pluralidad de sujetos que tienen cáncer, datos de pruebas genómicas de tumores, incluyendo alteraciones somáticas, recopilados en dos o más intervalos de tiempo por sujeto, una o más intervenciones terapéuticas administradas a cada uno de los sujetos en uno o más momentos y eficacia de las intervenciones terapéuticas, en donde la base de datos es útil para inferir la eficacia de las intervenciones terapéuticas en sujetos con un perfil genómico tumoral. En algunas formas de realización, la pluralidad es de al menos 50, al menos 500 o al menos 5000. En algunas formas de realización, 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 formas de realización, se utilizan frecuencias relativas de variantes genéticas detectadas para clasificar la eficacia del tratamiento. En algunas formas de realización, se utiliza información adicional para ayudar a clasificar la eficacia del tratamiento, que incluye, entre otros, peso, efectos adversos del tratamiento, pruebas histológicas, análisis de sangre, información radiográfica, tratamientos previos y tipo de cáncer. En algunas formas de realización, la respuesta al tratamiento por paciente se recopila y clasifica cuantitativamente mediante pruebas adicionales. En algunas formas de realización, las pruebas adicionales son pruebas basadas en sangre u orina.

En algunas formas de realización, el método comprende el uso de una base de datos para identificar una o más intervenciones terapéuticas eficaces para un sujeto que tiene cáncer, en donde la base de datos incluye, para cada uno de una pluralidad de sujetos que tienen cáncer, datos de pruebas genómicas de tumores, incluyendo alteraciones somáticas, recopilados en dos o más intervalos de tiempo por sujeto, una o más intervenciones terapéuticas administradas a cada uno de los sujetos en uno o más momentos y la eficacia de las intervenciones terapéuticas. En algunas formas de realización, las intervenciones terapéuticas identificadas se estratifican por eficacia. En algunas formas de realización, se informan límites cuantitativos sobre la eficacia o falta de eficacia de las intervenciones terapéuticas previstas. En algunas formas de realización, las intervenciones terapéuticas utilizan información de la evolución genómica de tumores prevista o mecanismos de resistencia adquiridos en pacientes similares en respuesta al tratamiento.

En algunas formas de realización, el método comprende clasificar la efectividad del tratamiento utilizando un algoritmo de clasificación, por ejemplo, procesos de regresión lineal (por ejemplo, regresión lineal múltiple (RLM), regresión de mínimos cuadrados parciales (MCP) y regresión de componentes principales (RCP)), árboles de decisión binarios (por ejemplo, procesos de partición recursiva como CART - árboles de clasificación y regresión), redes neuronales artificiales como redes de retropropagación, análisis discriminantes (por ejemplo, clasificador bayesiano o análisis de Fischer), clasificadores logísticos y clasificadores de vectores de soporte (por ejemplo, máquinas de vectores de soporte).

En algunas formas de realización, el método comprende informar los resultados de una o más pruebas genéticas que comprenden: capturar información genética que incluye variantes genéticas y medidas cuantitativas de la misma en uno o más puntos de prueba utilizando un analizador genético; normalizar las medidas cuantitativas para su representación con uno o más puntos de prueba y generar un factor de escala; aplicar el factor de escala para representar un mapa de respuesta tumoral; y generar un resumen de variantes genéticas. En algunas formas de realización, el método comprende analizar frecuencias de alelos mutantes que no son VNC (variación del número de copias). En algunas formas de realización, el método comprende transformar un valor absoluto en una métrica relativa para representar el mapa de respuesta tumoral. En algunas formas de realización, el método comprende multiplicar una frecuencia de alelo mutante por un valor predeterminado y tomar un logaritmo del mismo. En algunas formas de realización, el método comprende: multiplicar el factor de escala por un valor transformado para cada gen para determinar un indicador de cantidad que se representará en el mapa de respuesta tumoral; y asignar un indicador visual único para cada alteración en un panel visual. En algunas formas de realización, el método comprende centrar en Y o centrar verticalmente el indicador de cantidad en un panel colocado de forma contigua que indica continuidad. En algunas formas de realización, la asignación comprende, además proporcionar un color único para cada alteración.

En algunas formas de realización, el método comprende analizar información genética de otro punto de prueba o tiempo de prueba. En algunas formas de realización, en las que un nuevo resultado de prueba no difiere de un resultado de prueba anterior, el método comprende reproducir el panel visual anterior. En algunas formas de realización, en las que, si las alteraciones permanecen iguales, pero las cantidades han cambiado, el método comprende: mantener el orden y el indicador visual único para cada alteración; y determinar un nuevo indicador de cantidad y generar un nuevo panel visual para todos los puntos de prueba. En algunas formas de realización, el método comprende determinar una nueva alteración en la información genética y agregar la alteración a la parte superior de las alteraciones existentes. En algunas formas de realización, el método comprende determinar una nueva alteración en la información genética y determinar nuevos valores de transformación y factor de escala y asignar un indicador visual único para cada nueva alteración. En algunas formas de realización, el método comprende determinar una nueva alteración en la información genética y volver a generar el mapa de respuesta tumoral que incluye alteraciones de un punto de prueba anterior que aún se detectan en el punto de prueba actual y la nueva alteración. En algunas formas de realización, el método comprende determinar si una alteración anterior ya no está presente y, de ser así, comprende usar una altura de cero al reproducir la cantidad de la alteración de la alteración anterior para puntos de prueba posteriores. En algunas formas de realización, el método comprende determinar si una alteración anterior ya no está presente y, de ser así, reservar el indicador visual único asociado con la alteración anterior para uso futuro.

En algunas formas de realización, el método comprende analizar las frecuencias de alelo mutante de VNC y frecuencias de alelos mutantes de metilación. En algunas formas de realización, el método comprende la agrupación de frecuencias máximas de alelos mutantes para representarlas primero en el mapa de respuesta tumoral. En algunas formas de realización, el método comprende representar alteraciones para el gen en orden decreciente de frecuencia de alelos mutantes de alteraciones. En algunas formas de realización, el método comprende representar alteraciones para el gen en un orden decreciente. En algunas formas de realización, el método comprende seleccionar un siguiente gen con la siguiente frecuencia de alelos mutantes más alta.

En algunas formas de realización, para cada alteración notificada, el método comprende generar un indicador de tendencia para la alteración en los diferentes puntos de prueba. En algunas formas de realización, el método comprende generar un resumen de alteraciones. En algunas formas de realización, el método comprende generar un resumen de opciones de tratamiento. En algunas formas de realización, el método comprende generar un resumen de frecuencia de alelos mutantes, amplificación libre de células, indicación de aprobación clínica y ensayo clínico. En algunas formas de realización, el método comprende generar un panel basado en una vía biológica. En algunas formas de realización, el método comprende generar un panel basado en un nivel de evidencia. En algunas formas de realización, la información genética incluye una o más de variaciones de un solo nucleótido, variaciones del número de copias, inserciones y deleciones y reordenamientos genéticos. En algunas formas de realización, el método comprende generar un informe de relevancia clínica sobre alteraciones detectadas. En algunas formas de realización, el método comprende generar un resumen de resultados de terapia.

En algunas formas de realización, el método comprende generar un informe genético que comprende: generar datos de variación de número de copias no (VNC) utilizando un analizador genético; determinar un factor de escala para cada frecuencia de alelo mutante no VNC; para una primera prueba, generar un panel visual de cada alteración no VNC utilizando el factor de escala; y para cada prueba posterior, generar cambios en la alteración no VNC para el panel visual utilizando el factor de escala.

En algunas formas de realización, el método comprende transformar un valor absoluto en una métrica relativa para la representación. En algunas formas de realización, el método comprende multiplicar una frecuencia de alelo mutante por un valor predeterminado y tomar un logaritmo del valor predeterminado. En algunas formas de realización, el método comprende determinar un factor de escala utilizando un valor máximo observado. En algunas formas de realización, para cada alteración no VNC, el método comprende multiplicar un factor de escala por un valor transformado para cada variante genética como un indicador de cantidad para visualizar la variante genética.

En algunas formas de realización, el método comprende asignar un indicador visual único para cada alteración. En algunas formas de realización, para la prueba posterior, el método comprende utilizar el panel visual si el resultado de la prueba no cambia. En algunas formas de realización, si las alteraciones permanecen iguales en la prueba posterior, el método comprende mantener el orden y el indicador visual único para cada alteración; y recalcular un indicador de cantidad para visualizar esa variante y volver a representar los valores actualizados en los paneles existentes y en el nuevo panel para la última prueba. En algunas formas de realización, si se encuentra una nueva alteración en la prueba posterior, el método comprende añadir las alteraciones a la parte superior de todas las alteraciones existentes; calcular los valores de transformación y el factor de escala; y asignar un indicador visual único para cada nueva alteración.

En algunas formas de realización, el método comprende: volver a renderizar las alteraciones en el punto de prueba anterior y la nueva alteración; y centrar verticalmente una imagen de las alteraciones en un panel colocado de manera contigua que indica continuidad. En algunas formas de realización, si una alteración anterior no está presente en una prueba posterior, el método comprende utilizar una altura de cero como la cantidad de la alteración para una renderización posterior. En algunas formas de realización, el método comprende la presentación de información del sujeto o de la intervención asociada con cambios de alteración. En algunas formas de realización, el método comprende la identificación de una alteración con la frecuencia máxima de alelos mutantes.

En algunas formas de realización, el método comprende: informar alteraciones para ese gen en orden decreciente de frecuencia de alelos mutantes de alteraciones no VNC; e informar alteraciones VNC para ese gen en orden decreciente de valor VNC. En algunas formas de realización, el método comprende seleccionar el siguiente gen con la siguiente frecuencia más alta de alelos mutantes no VNC e informar alteraciones para ese gen en orden decreciente de frecuencia de alelos mutantes de alteraciones no VNC; e informar alteraciones VNC para ese gen en orden decreciente de valor VNC.

En algunas formas de realización, el método comprende generar un indicador de tendencia para una alteración a lo largo de diferentes fechas de prueba. En algunas formas de realización, el método comprende agrupar frecuencias máximas de alelos mutantes y generar anotaciones que incluyen vías biológicas o nivel de evidencia. En algunas formas de realización, el método comprende generar un panel basado en un nivel de evidencia. En algunas formas de realización, el método comprende generar un panel basado en una vía biológica. En algunas formas de realización, la información genética incluye una o más de variaciones de un solo nucleótido, variaciones en el número de copias, inserciones y deleciones y reordenamientos genéticos.

En algunas formas de realización, el método comprende: a) proporcionar una pluralidad de muestras de ácido nucleico de un sujeto, siendo las muestras recolectadas en puntos de tiempo seriales; b) secuenciar ADN libre de células de las muestras para generar secuencias; c) determinar una medida cuantitativa de cada una de una pluralidad de variantes genéticas entre los polinucleótidos en cada muestra; d) representar gráficamente por computadora cantidades relativas de variantes genéticas en cada punto de tiempo serial para aquellas mutaciones somáticas presentes en una cantidad distinta de cero en al menos uno de los puntos de tiempo seriales. En algunas formas de realización, la medida cuantitativa es la frecuencia de la variante genética entre todas las secuencias que se asignan al mismo locus genético. En algunas formas de realización, las cantidades relativas se representan como un gráfico de áreas apiladas. En algunas formas de realización, las cantidades relativas se apilan, en el punto temporal más temprano, de mayor a menor desde la parte inferior a la superior del gráfico, y en donde una variante genética que aparece por primera vez en una cantidad distinta de cero en un punto temporal posterior se apila en la parte superior del gráfico. En algunas formas de realización, las áreas se representan mediante diferentes colores. En algunas formas de realización, la representación gráfica indica, además, para cada punto temporal, la medida cuantitativa de la variante genética predominante. En algunas formas de realización, la representación gráfica incluye además una clave que identifica las variantes genéticas representadas en el gráfico. En algunas formas de realización, la representación gráfica comprende normalizar y escalar las medidas cuantitativas.

En algunas formas de realización, los loci se encuentran en oncogenes. En algunas formas de realización, la pluralidad de variantes genéticas se asigna a un gen diferente en el genoma. En algunas formas de realización, la pluralidad de variantes genéticas se asigna al mismo gen en el genoma. En algunas formas de realización, se secuencian al menos 10 oncogenes diferentes.

En algunas formas de realización, la determinación comprende recibir las secuencias en la memoria de la computadora y utilizar un procesador de computadora para ejecutar un software para determinar la medición cuantitativa. En algunas formas de realización, la representación gráfica comprende utilizar un procesador de computadora para ejecutar un software que transforma las mediciones cuantitativas en un formato gráfico y representar el formato gráfico en una interfaz gráfica de usuario electrónica, por ejemplo, una pantalla de visualización.

En algunas formas de realización, el método comprende generar un informe de prueba de paciente en papel o electrónico a partir de datos generados por un analizador genético que comprende: a) resumir datos de dos o más puntos de tiempo de prueba, por lo que se informa una unión de todos los resultados de prueba distintos de cero en cada punto de prueba posterior después de la primera prueba; y b) presentar los resultados de prueba en el informe de prueba de paciente en papel o electrónico. En algunas formas de realización, el resumen y la presentación se realizan en una computadora ejecutando código con un procesador de computadora para (i) identificar todos los resultados de prueba distintos de cero, (ii) generar el informe de prueba y (iii) mostrar el informe de prueba en una interfaz gráfica de usuario.

En algunas formas de realización, el método comprende representar gráficamente la evolución de variantes genéticas de un tumor en un sujeto a partir de datos generados por un analizador genético que comprende: a) generar por computadora una representación apilada de variantes genéticas detectadas en cada uno de una pluralidad de puntos de tiempo en el sujeto, en donde una altura o ancho de cada capa en la pila que corresponde a una variante genética representa una contribución cuantitativa de la variante genética a la cantidad total de variantes genéticas en cada punto temporal; y b) mostrar la representación apilada en un monitor de computadora o un informe en papel. En algunas formas de realización, el método comprende, además el uso de una combinación de una magnitud de variantes genéticas detectadas en una prueba basada en fluidos corporales para inferir una carga de enfermedad. En algunas formas de realización, el método comprende, además el uso de fracciones de alelos de mutaciones detectadas, desequilibrios alélicos, cobertura específica de genes para inferir la carga de enfermedad.

En algunas formas de realización, la altura total de la pila es representativa de la carga total de la enfermedad o de una puntuación de la carga de la enfermedad en el sujeto. En algunas formas de realización, se utiliza un color distinto para representar cada variante genética. En algunas formas de realización, solo se representa gráficamente un subconjunto de las variantes genéticas detectadas. En algunas formas de realización, el subconjunto se elige en función de la probabilidad de que sea una alteración impulsora o una asociación con una respuesta aumentada o reducida al tratamiento.

En algunas formas de realización, el método comprende producir un informe de prueba para una prueba genómica. En algunas formas de realización, se utiliza una escala no lineal para representar las alturas o los anchos de cada variante genética representada. En algunas formas de realización, se representa un gráfico de puntos de prueba anteriores en el informe. En algunas formas de realización, el método comprende estimar una progresión o remisión de la enfermedad en función de la tasa de cambio y/o la precisión cuantitativa de cada resultado de prueba. En algunas formas de realización, el método comprende mostrar una intervención terapéutica entre puntos de prueba intermedios. En algunas formas de realización, la visualización comprende: a) recibir datos que representan las variantes genéticas tumorales detectadas en la memoria de la computadora; b) ejecutar código con un procesador de computadora para representar gráficamente la contribución cuantitativa de cada variante genética en un punto de tiempo como una línea o área proporcional a la contribución relativa; y c) mostrar la representación gráfica en una interfaz gráfica de usuario.

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 función de VNC 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 medicamentos.

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

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

La **Figura 6** muestra un escaneo ejemplar de VNC 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 VNC genómico en los loci a y b, pero el locus c muestra duplicación génica. En las células en estado de reposo, el número de copias es relativamente igual en todos los intervalos de la región, excepto en aquellos intervalos que se superponen al locus de duplicación génica. 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 replicación, lo que proporciona una variación en VNC en la región. La desviación es particularmente dramática en un locus que muestra VNC en un origen de replicación (c).

La **Figura 7** muestra un curso ejemplar de monitoreo 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 pruebas de cáncer.

La **Figura 9B** muestra un proceso ejemplar para reducir las tasas de error y 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 pruebas genéticas a los usuarios.

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

Las **Figuras 10J-10P** muestran varios ejemplos de gráficos de flujo modificados.

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

DESCRIPCION DETALLADA

El mosaicismo biomolecular (por ejemplo, mosaicismo genético) en una muestra biológica puede ser detectado, como una población genómica heterogénea de células o 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 diferentes células somáticas que tienen 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 resultado mosaicos genéticos, que también se pueden detectar examinando los polinucleótidos vertidos en la sangre desde el órgano trasplantado. La heterogeneidad de células enfermas, en la que las células enfermas tienen diferentes variantes genéticas, es otra forma de mosaicismo genético. Los métodos de acuerdo con la invención detectan el mosaicismo y proporcionan una intervención terapéutica. Ciertas formas de realización comprenden la forma de realización de un perfil de todo el cuerpo de mosaicismo biomolecular mediante el uso de polinucleótidos circulantes, que pueden derivar o tener su origen en células de diversas partes del cuerpo de un sujeto.

Las células enfermas, como los tumores, pueden evolucionar con el tiempo, dando lugar a diferentes subpoblaciones clonales que tienen nuevas características genéticas y fenotípicas. Esto puede ser el resultado de mutaciones naturales a medida que las células se dividen, o puede ser impulsado por tratamientos que se dirigen a ciertas subpoblaciones clonales, lo que permite 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 en el presente documento heterogeneidad de células enfermas o, en el caso del cáncer, heterogeneidad tumoral.

En la actualidad, los cánceres se tratan basándose en 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 ser seguido por un tratamiento utilizando una terapia anti-Her2+. Como otro ejemplo, un

cáncer colorrectal en el que se encuentra un mutante de KRAS en pequeñas cantidades puede ser tratado con una terapia para la que KRAS responde.

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 provenientes de células enfermas ubicadas en todo el cuerpo permite obtener un perfil corporal completo de la heterogeneidad de las células enfermas. 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. En cambio, 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 se puede determinar, por ejemplo, cuantificando el ADN para cada forma en una muestra, por ejemplo, ADN libre de células (ADNlc), detectando así la heterogeneidad en el tumor.

Esta información puede ser utilizada por un proveedor de atención médica, por ejemplo, un médico, para desarrollar intervenciones terapéuticas. Por ejemplo, un sujeto que tiene un tumor heterogéneo puede ser tratado como si tuviera dos tumores, y una intervención terapéutica puede tratar cada uno de los tumores. La intervención terapéutica según la invención incluye 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 pueden administrarse en cantidades que reflejan 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 en una dosis mayor que un fármaco para tratar la forma mutante en una cantidad relativa menor. O bien, el tratamiento para el mutante en la cantidad relativa menor puede retrasarse o escalonarse con respecto al mutante en una cantidad mayor.

El seguimiento de los cambios en el perfil de heterogeneidad de las células de la enfermedad 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 sí trata un tumor que porta el marcador de resistencia.

Las intervenciones terapéuticas pueden ser determinadas por un proveedor de atención médica o por un algoritmo informático, o una combinación de ambos. Una base de datos puede contener los resultados de intervenciones terapéuticas contra enfermedades que tienen diversos perfiles de heterogeneidad de células patológicas. La base de datos puede ser consultada para determinar una intervención terapéutica para una enfermedad con un perfil particular.

Células de enfermedad

El sujeto puede tener una condición patológica (enfermedad). Las células que presentan patología de enfermedad se denominan en el presente documento células de enfermedad.

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 cerebral, cáncer de esófago, cáncer de cabeza y cuello, cáncer de vejiga, cáncer ginecológico, liposarcoma y mieloma múltiple.

Otros cánceres 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 huesos, osteosarcoma, histiocitoma fibroso maligno, glioma del tronco encefálico, cáncer cerebral, craneofaringioma, ependimoblastoma, ependimoma, meduloblastoma, meduloepitelioma, tumor del parénquima pineal, cáncer de mama, tumor bronquial, linfoma de Burkitt, linfoma no Hodgkin, tumor carcinoide, cáncer de cuello uterino, cordoma, leucemia linfocítica crónica (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 no pequeñas, carcinoma de células pequeñas, melanoma, cáncer de boca, síndromes mielodisplásicos, mieloma múltiple, meduloblastoma, cáncer de cavidad nasal, cáncer de seno paranasal, neuroblastoma, cáncer nasofaríngeo, cáncer oral, cáncer orofaríngeo, 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 rectal, cáncer de células renales, rabdomiosarcoma, cáncer de glándula salival, síndrome de Sézary, cáncer de piel, no melanoma, cáncer de intestino delgado, sarcoma de tejidos blandos, carcinoma de células escamosas, cáncer testicular, cáncer de garganta, timoma, cáncer de tiroides, cáncer uretral, cáncer de útero, sarcoma uterino, cáncer vaginal, cáncer vulvar, macroglobulinemia de Waldenström y/o tumor de Wilms.

Un tumor es una colección de células cancerosas (células de la enfermedad cancerosa). 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 de tumores metastásicos (tumores metastásicos) y tumores difusos (por ejemplo, células tumorales circulantes). Un tumor puede incluir células de un solo cáncer (por ejemplo, cáncer colorrectal) o de múltiples cánceres (por ejemplo, cáncer colorrectal y cáncer de páncreas). Un tumor puede incluir células que se originan a partir de una sola célula somática original o de diferentes células somáticas.

Las células de la enfermedad son espacialmente distintas si las células están ubicadas al menos a 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 de la enfermedad (por ejemplo, "carga tumoral") es una medida cuantitativa de la cantidad de células de la enfermedad en un sujeto. Una medida de la carga de células de la enfermedad es la fracción de macromoléculas biológicas totales en una muestra que son macromoléculas biológicas de la enfermedad, por ejemplo, la cantidad relativa de polinucleótidos tumorales en una muestra de polinucleótidos libres de células. Por ejemplo, si el ADNlc 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 ADNlc 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 del primer sujeto. Estas medidas son mucho más relevantes en una base intrasujeto que en una base intersujeto, ya que las cargas tumorales libres de células en un individuo pueden ser mucho más altas o más bajas que en otro individuo a pesar de los diferentes niveles de carga de enfermedad. Sin embargo, estas medidas se pueden utilizar de manera bastante efectiva para monitorear la carga de enfermedad dentro de un individuo, por ejemplo, un aumento de una carga tumoral de ADN libre de células del 5 % al 15 % puede indicar una progresión significativa de la enfermedad, mientras que una disminución del 10 % al 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 ubicaciones en una única masa tumoral. También incluye polinucleótidos obtenidos de células en diferentes sitios tumorales metastásicos. Las células liberan polinucleótidos en la sangre, donde son detectables 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, el ADN libre de células proporciona un perfil más preciso de la heterogeneidad tumoral en toda la población de células de la enfermedad que el ADN obtenido de una única ubicación tumoral. El ADN extraído de las células de toda la población de células enfermas de un cuerpo se denomina "ADN de carga de enfermedad" o, en el caso del cáncer, "ADN de carga tumoral".

Las células de enfermedades, 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 ADNlc, por ejemplo, dentro del 15 %, dentro del 10 %, dentro del 5 % o dentro del 2 %.

Análisis

Tal como se utiliza 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 a partir de tres cadenas de ácidos grasos.

Un polinucleótido de cáncer (por ejemplo, ADN de cáncer) es un polinucleótido (por ejemplo, ADN) derivado de una célula cancerosa. El ADN y/o ARN canceroso se puede 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 (ADNlc) o ARN libre de células.

El ADN libre de células es ADN ubicado 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. Se cree que el ADN libre de células surge de células moribundas que liberan su ADN en la sangre. Debido a que las células cancerosas espacialmente distintas liberan ADN en fluidos corporales, como la sangre, el ADNlc de sujetos con cáncer típicamente comprende ADN canceroso de células cancerosas espacialmente distintas.

Muestras biológicas

Los analitos de acuerdo con la invención se derivan de un fluido corporal. 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, líquido linfático, líquido prostático, líquido seminal, leche, esputo, heces y lágrimas. Los fluidos corporales son fuentes particularmente buenas de macromoléculas biológicas de células patológicas espacialmente distintas, ya que dichas células de muchas ubicaciones en un cuerpo pueden liberar estas moléculas en el fluido corporal. Por ejemplo, la sangre y la orina son buenas fuentes de polinucleótidos libres de células. Las macromoléculas de dichas fuentes pueden proporcionar un perfil más preciso de las células enfermas que las macromoléculas derivadas de una masa localizada de células patológicas.

Se pueden aumentar las cantidades de polinucleótidos de la enfermedad en una muestra de fluido corporal. Dichos aumentos pueden aumentar la sensibilidad de detección de polinucleótidos de la enfermedad. Se puede administrar una intervención, tal como una intervención terapéutica, a un sujeto que haga que las células de la enfermedad se lisen, vaciando su ADN en el fluido circundante. Dichas intervenciones pueden incluir la administración de quimioterapia. También puede incluir la administración de radiación o ultrasonidos a todo el cuerpo de un sujeto, o a una parte del cuerpo de un sujeto, tal como ser dirigido a un tumor o un órgano enfermo. Después de la administración de la intervención y cuando la cantidad de polinucleótidos de la enfermedad en el fluido aumenta, se recoge una muestra de fluido para su análisis. El intervalo entre la administración de la intervención y la recogida puede ser lo suficientemente largo para que los polinucleótidos de la enfermedad aumenten, pero no tan largo como para que se eliminen del cuerpo. Por ejemplo, se puede administrar una dosis baja de quimioterapia aproximadamente una semana antes de la recogida de la muestra.

Métodos analíticos

Existen varios tipos de análisis biomolecular, incluidos, por ejemplo, genómico, epigenético (por ejemplo, metilación), expresión de ARN y proteómico. El análisis genómico se puede realizar, por ejemplo, mediante un analizador genético, por ejemplo, utilizando secuenciación de ADN. El análisis de metilación se puede realizar, por ejemplo, mediante conversión de bases seguida de secuenciación de ADN. El análisis de la expresión del ARN se puede realizar, por ejemplo, mediante hibridación de matrices de polinucleótidos. El análisis proteómico se puede realizar, por ejemplo, mediante espectrometría de masas.

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

Los métodos analíticos pueden incluir la generación y captura de 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 pueden analizarse mediante cualquier método conocido en la técnica. Normalmente, el secuenciador de ADN empleará secuenciación de próxima generación (por ejemplo, Illumina, 454, Ion torrent, SOLiD). El análisis de secuencias puede realizarse mediante secuenciación masiva en paralelo, es decir, secuenciando simultáneamente (o en rápida sucesión) cualquiera de al menos 100.000, 1 millón, 10 millones, 100 millones o 1.000 millones de moléculas de polinucleótidos. Los métodos de secuenciación pueden incluir, entre otros: 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, RNA-Seq (Illumina), expresión génica digital (Helicos), secuenciación de próxima generación, secuenciación de una sola molécula por síntesis (SMSS) (Helicos), secuenciación masiva paralela, matriz de una sola molécula clonal (Solexa), secuenciación shotgun, secuenciación Maxam-Gilbert o Sanger, primer walking, secuenciación utilizando 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 del ADN seguida de la escisión en bases específicas, o puede aplicar la técnica de Sanger que se basa en la terminación de la cadena de didesoxinucleótidos. El método de 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), que acelera la preparación de la muestra antes de la secuenciación y reduce los errores. Además, los datos de secuenciación se recopilan a partir de las reacciones provocadas por la adición de nucleótidos en la cadena complementaria en tiempo real. Por ejemplo, los secuenciadores de ADN pueden utilizar un método llamado Single-

molecule real-time (SMRT), donde los datos de secuenciación se producen mediante luz (capturada por una cámara) emitida cuando se agrega un nucleótido a la cadena complementaria mediante 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 diversos tipos de cáncer. La secuenciación de genes seleccionados, o porciones de genes, puede ser suficiente para el análisis deseado. Los polinucleótidos que se asignan a loci específicos en el genoma que son el tema de interés pueden aislarse para la secuenciación mediante, por ejemplo, captura de secuencia o amplificación específica del sitio.

Una secuencia de nucleótidos (por ejemplo, secuencia de ADN) puede referirse a lecturas de secuencias sin procesar o lecturas de secuencias procesadas, tales como recuentos moleculares únicos inferidos a partir de lecturas de secuencias sin procesar.

Las lecturas de secuencias generadas a partir de la secuenciación están sujetas a análisis que incluyen, por ejemplo, la identificación de variantes genéticas. Esto puede incluir la identificación de variantes de secuencia y la cuantificación de números de llamadas de bases en cada locus. La cuantificación puede implicar, por ejemplo, contar el número de lecturas que se asignan a un locus genético particular. Diferentes números de lecturas en diferentes loci pueden indicar variación del número de copias (VNC).

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 los polinucleótidos no diana. Cuando las moléculas diana son pocas en número, la señal de la diana puede ser débil. Este puede ser el caso, por ejemplo, en el caso del ADN libre de células, donde una pequeña cantidad de polinucleótidos tumorales puede mezclarse con una cantidad mucho mayor de polinucleótidos de células sanas. Los métodos de rastreo molecular pueden ser útiles en tales situaciones. El rastreo molecular implica rastrear lecturas de secuencias desde un protocolo de secuenciación hasta moléculas en una muestra original (por ejemplo, antes de la amplificación y/o secuenciación) de la que se derivan las lecturas. Ciertos métodos implican etiquetar moléculas de tal manera que múltiples lecturas de secuencias producidas a partir de moléculas originales se puedan agrupar en familias de secuencias derivadas de moléculas originales. De esta manera, se pueden filtrar las llamadas de bases que representan ruido. Dichos 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. En algunas formas de realización, los métodos de muestreo ascendente implican determinar una medida cuantitativa de moléculas de ADN individuales para las que se detectan ambas cadenas (cadenas 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 a partir de estas medidas una medida cuantitativa de moléculas de ADN individuales para las que no se detectó ninguna cadena; y utilizar estas medidas para determinar la medida cuantitativa indicativa de una cantidad de moléculas de ADN bicatenarias individuales en la muestra. Este método se describe con más detalle en PCT/US2014/072383, presentada el 24 de diciembre de 2014.

Variantes genéticas

Los métodos de la presente divulgación se pueden utilizar en la detección de variantes genéticas (también denominadas "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 células de enfermedades, 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 tienen una sola secuencia, no identifican polimorfismos comunes que también pueden considerarse normales.

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

Las variantes genéticas se pueden detectar comparando secuencias de polinucleótidos en una muestra con una referencia, por ejemplo, a una secuencia de genoma de referencia, con un índice o con una base de datos de mutaciones conocidas. En una forma de realización, la secuencia de referencia es una secuencia de referencia disponible públicamente, como la secuencia del genoma humano HG-19 o NCBI Build 37. En otra forma de realización, la secuencia de referencia es una secuencia en una base de datos no pública. En otra forma de realización, 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.

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, espermatozoide u óvulo) o un cigoto, de un individuo. Las mutaciones somáticas, por ejemplo, las que se encuentran en células cancerosas, se pueden distinguir del genoma de la línea germinal de un sujeto en el que 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 se conocen variantes genéticas que son comunes en las células cancerosas. Se puede encontrar una base de datos de VNS en el cáncer humano en el sitio web: cancer.sanger.ac.uk/cancergenome/projects/cosmic/.

La Figura 8 muestra genes que se sabe que, en el cáncer, presentan mutaciones puntuales, amplificaciones, fusiones y indeles.

Desviación de la VNC en células que se dividen rápidamente

Durante la fase S del ciclo celular, la célula replica ADN. Una célula diploide que tiene 2N cromosomas con ADN replicado puede corresponder a aproximadamente 4X de contenido de ADN, mientras que una célula diploide que tiene 2N cromosomas sin ADN replicado puede corresponder a aproximadamente 2X de contenido de ADN. La replicación procede de los orígenes de replicación. En los mamíferos, los orígenes de replicación están espaciados a intervalos de aproximadamente 15 kb a 300 kb. Durante este período, partes del genoma existen en forma poliploide. Aquellas áreas entre los orígenes de replicación y la posición de la polimerasa están duplicadas, mientras 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 cadena que experimenta replicación. Cuando se escanea a través del genoma, el número de copias parece desigual o distorsionado, con regiones que existen en forma poliploide y regiones que existen en forma diploide. Tal escaneo parece ruidoso. Esto es cierto incluso para las células que no presentan variaciones en el número de copias en el genoma en estado de reposo. Por el contrario, una exploración de VNC en células en G_0 muestra un perfil en el que 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 VNC en todo el genoma muestra distorsión, independientemente de que el genoma también contenga VNC en determinados loci.

Se puede aprovechar este hecho para detectar la carga tumoral en el ADN de muestras que comprenden ADN heterogéneo, por ejemplo, una mezcla de ADN patológico y ADN sano, como ADNc. Un método para 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 cerca de 2 copias (en una célula diploide). En ciertas formas de realización, el locus o loci examinados 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 lo largo de un cromosoma completo o a lo largo de un genoma completo. Se determina una medida de VNC del origen de replicación (VNCOR) a lo largo de la región. Esto puede ser, por ejemplo, una medida de desviación en el número de copias con respecto a un valor de tendencia central. El valor de tendencia central puede ser, por ejemplo, media, mediana o moda. La medida de desviación puede ser, por ejemplo, varianza o desviación estándar. Esta medida puede compararse con una medida de VNCOR en la misma región en una muestra de control, por ejemplo, de un individuo sano o células en estado de reposo. Los VNCOR pueden determinarse dividiendo la región o regiones analizadas en particiones no superpuestas de varias longitudes y tomando una medida de VNC en esta partición. Esta medida de VNC puede derivarse del número de lecturas o fragmentos determinados para mapear esas regiones después de la secuenciación. Las particiones pueden tener varios tamaños, para producir varios niveles de resolución, por ejemplo, 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 células en proceso de división celular en la muestra.

Se pueden utilizar varios métodos para calcular las variaciones reales del número de copias genéticas que difieren de la distorsión basada en el origen de replicación. Por ejemplo, las posiciones de SNP heterocigóticas en los loci VNC afectados se pueden utilizar 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 replicación no debería afectar a este desequilibrio, ya que ambas copias generalmente se copiarían en intervalos de tiempo similares y, por lo tanto, se autonormalizarían (aunque los cambios alélicos podrían posiblemente cambiar la replicación del origen entre las dos variantes alélicas). Por ejemplo, la duplicación de un segmento cromosómico que contiene un SNP podría detectarse en alrededor del 67 % de las lecturas, mientras que la duplicación resultante de VNCOR se detectaría en aproximadamente el 50 % de las lecturas. En otro método, se utilizan técnicas basadas en el recuento que utilizan la densidad de fragmentos o lecturas detectados en un locus determinado para calcular el número relativo de copias. Estas técnicas generalmente están 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 útil para obtener una precisión aún mayor.

Se puede calcular VNCOR para una muestra dada y se puede utilizar para dar un valor sobre la carga tumoral libre de células a pesar de la falta de detección de variantes somáticas tradicionales, tales como VNS, VNC específicas de genes, reordenamientos genómicos, variantes epigenéticas, pérdida de heterocigosidad, etc. Los VNCOR también se pueden utilizar para restar la distorsión para una muestra dada para aumentar la sensibilidad y/o especificidad de un método de detección/estimación de VNC dado eliminando la variación que está relacionada con la proximidad del origen de replicación en lugar de deberse a cambios reales en el número de copias en una célula. Las líneas celulares con cambios conocidos o nulos en el número de copias sobre una referencia también se pueden utilizar como una referencia de VNCOR para su uso en la estimación de su contribución a una muestra dada.

En una forma de realización, el método implica determinar un nivel de referencia de copias de moléculas de ADN en uno o más loci de una o más muestras de control, cada una de las cuales contiene ADN de células que experimentan un nivel predeterminado de división celular, 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 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 a lo largo de todas las particiones, o un nivel de varianza a lo largo de los loci. Una medida de tendencia central o de variación (por ejemplo, varianza o desviación estándar) en el número de copias en la muestra de prueba se compara 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 que se dividen 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 patológicas

La heterogeneidad de células enfermas, por ejemplo, la heterogeneidad tumoral, es la aparición de células enfermas que tienen diferentes variantes genéticas. La heterogeneidad de células enfermas se puede determinar mediante el examen de polinucleótidos aislados de células enfermas y la detección de diferencias en sus genomas. La heterogeneidad de células enfermas también se puede inferir a partir del examen de polinucleótidos de una muestra que contiene polinucleótidos de células enfermas y sanas en función de las 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 descontrolado de las células cancerosas o funcionar como marcadores de la respuesta o no respuesta a diversas intervenciones terapéuticas.

La heterogeneidad tumoral es una condición en la que un tumor se caracteriza por 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 contienen alteraciones en diferentes genes, o que contienen 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 tanto de BRAF como de 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 genómica).

Por ejemplo, una célula tumoral 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 de forma independiente 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 heterogeneidad de enfermedades

Los métodos de esta divulgación permiten la elaboración de perfiles cuantitativos y cualitativos de la heterogeneidad tumoral. En una forma de realización, el perfil es un perfil de cuerpo entero que contiene información de células distribuidas por todo el cuerpo. El análisis de polinucleótidos en ADNlc 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.

En una forma de realización de un perfil de heterogeneidad de células tumorales, el perfil identifica variaciones genéticas y las cantidades relativas de cada una variante. A partir de esta información, se pueden inferir posibles distribuciones de las variantes en diferentes subpoblaciones celulares. Por ejemplo, un cáncer puede comenzar con una célula portadora de una 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 células tumorales), el perfil puede indicar 100 % X, 35 % Y y 15 % Z. También se pueden detectar tanto VNC en un primer locus como variantes de secuencia en un segundo locus.

La heterogeneidad tumoral se puede detectar a partir del análisis de 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 es probable que contenga ADN de línea germinal, así como ADN de cáncer), se puede encontrar que una variante de secuencia de BRAF se produce a una frecuencia del 17 %, una variante de secuencia de CDKN2A se produce a una frecuencia del 6 %, una variante de secuencia de ERBB2 se produce 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 muestran 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 del ADN libre de células, la detección de mutaciones somáticas se puede realizar comparando las llamadas de bases en la muestra con una secuencia de referencia o, internamente, como llamadas de bases menos frecuentes con llamadas de bases más comunes, que se supone que están en la secuencia de la línea germinal. En cualquier caso, la existencia de formas subdominantes (por ejemplo, menos del 40 % de las llamadas de bases totales) en diferentes loci y con diferente frecuencia indica heterogeneidad de células patológicas.

El ADN libre de células comprende típicamente una preponderancia de ADN de células normales que tienen la secuencia del genoma de la línea germinal y, en el caso de una enfermedad, como el cáncer, un pequeño porcentaje de ADN de células cancerosas y que tienen una secuencia del genoma del cáncer. Las secuencias generadas a partir de polinucleótidos en una muestra de ADNc se pueden comparar con una secuencia de referencia para detectar diferencias entre la secuencia de referencia y los polinucleótidos en el ADNc. 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. Esto probablemente indica 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 un segundo nucleótido diferente de un nucleótido de referencia se detecta en aproximadamente el 50 %, esto probablemente indica heterocigosidad normal. La heterocigosidad 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 (o por encima) de un rango de heterocigotos inequívoco (por ejemplo, menos de aproximadamente el 45 %, menos del 40 %, menos del 30 %, menos del 20 %, menos del 10 % o menos del 5 %), esto puede atribuirse a la existencia de mutaciones somáticas en un porcentaje de las células que aportan ADN a la población de ADNc. Estas 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 enfermas, por ejemplo, las células cancerosas, son heterogéneas.

En el caso del ADN de tumores sólidos, que se espera que comprenda predominantemente ADN tumoral, también se pueden detectar mutaciones somáticas mediante comparación con una secuencia de referencia. La 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 susceptibles de tratamiento. El conocimiento de dichas variantes puede contribuir a la selección de intervenciones terapéuticas, ya que las terapias pueden dirigirse a dichas variantes. En el caso del cáncer, ya se conocen muchas variantes genéticas susceptibles de tratamiento.

VNC y VNS en la heterogeneidad de células patológicas

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 homocigosidad o heterocigosidad (por ejemplo, aproximadamente el 100 % o aproximadamente el 50 %, respectivamente) sin variación del número de copias. Esto es consistente con un polimorfismo o mutación de la línea germinal. Una variante de secuencia puede detectarse a una frecuencia de aproximadamente el 67 % (o, alternativamente, aproximadamente el 33 %) de polinucleótidos en un locus, y también en un gen medido a un número de copias aumentado (generalmente, n = 2). Esto es consistente con la duplicación génica 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 la homocigosidad (por ejemplo, aproximadamente el 100 %) pero en cantidades consistentes con la variación del número de copias, es más probable que esto refleje la presencia de polinucleótidos de células de la enfermedad que han experimentado una amplificación génica. De manera similar, si una variante de secuencia es detectada en un nivel no incompatible con la heterocigosidad (por ejemplo, desviándose un poco 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 alélica alejada del 50:50.

Esta observación se puede utilizar para inferir si es más probable que una variante de secuencia esté presente en el nivel de la línea germinal o que haya sido resultado de una mutación de células somáticas, por ejemplo, en una célula cancerosa. Por ejemplo, podría decirse que una variante de secuencia en un gen detectado a niveles es consistente con la heterocigosidad en la línea germinal, pero es más probable que sea el producto de una mutación somática en las células enfermas si también se detecta variación en el número de copias en ese gen.

Además, en la medida en que esperamos que una duplicación de genes en la línea germinal deba llevar una variante consistente con una dosis genética aumentada (por ejemplo, alrededor del 67 % para la trisomía en un locus), la amplificación del gen de detección con una dosis de variante de secuencia que se desvía significativamente de esta cantidad esperada indica que es más probable que la VNC esté presente como resultado de una mutación de células somáticas.

El hecho de que las mutaciones somáticas en diferentes loci puedan estar presentes en un número de copia único o múltiple en la misma célula de la enfermedad también se puede utilizar para inferir la heterogeneidad tumoral. Más específicamente, la heterogeneidad tumoral se puede inferir cuando se detectan dos genes en una frecuencia diferente pero su número de copias es relativamente igual. Alternativamente, la homogeneidad tumoral se puede inferir cuando la diferencia en la frecuencia entre dos variantes de secuencia es consistente con la diferencia en el número de copias para los dos genes. Por lo tanto, si se detecta una variante de EGFR en un 11 % y una variante de KRAS en un 5 %, y no se detecta ninguna VNC en estos genes, la diferencia en la frecuencia probablemente refleja la heterogeneidad tumoral (por ejemplo, todas las células tumorales portan un mutante de EGFR y la mitad de las células tumorales también portan un mutante de KRAS). Alternativamente, si el gen EGFR que porta el mutante se detecta en un número de copias aumentado, una interpretación consistente es una población homogénea de células tumorales, cada célula portando un mutante en los genes EGFR y KRAS, pero en la que el gen KRAS está duplicado. En consecuencia, se puede determinar tanto la frecuencia de una variante de secuencia como una medida de VNC en el locus de la variante de secuencia en una muestra. La frecuencia se puede corregir entonces para reflejar el número relativo de células que portan la variante ponderando la frecuencia en función de la dosis por célula determinada a partir de la medida de VNC. Este resultado es ahora más comparable en términos de número de células que portan la variante con una variante de secuencia que no varía en número de copias.

Comunicación de los resultados de las pruebas

Un generador de informes puede proporcionar un informe de los resultados del análisis de variantes genéticas (por ejemplo, variantes de secuencia, VNC, heterogeneidad de células de la enfermedad y combinaciones de las mismas), por ejemplo, a un profesional de la salud, por ejemplo, un médico, para ayudar a la interpretación de los resultados de las pruebas (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 pueden ser útiles para diagnosticar enfermedades y seleccionar opciones de tratamiento.

Con referencia ahora a la FIG. 9A, se ilustra esquemáticamente un sistema con un generador de informes 1 para informar, por ejemplo, sobre los resultados de pruebas de cáncer y las opciones de tratamiento a partir de los mismos. El sistema generador de informes puede ser un sistema de procesamiento de datos central configurado para establecer comunicaciones directamente con: un sitio de datos remoto o laboratorio 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, un consultorio médico, 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 exámenes; y diagnóstico. El proveedor o consultorio de atención médica 6 puede incluir proveedores de servicios médicos, como médicos, enfermeras, asistentes de atención médica a domicilio, técnicos y asistentes médicos, y el consultorio puede ser cualquier instalación de atención médica dotada de personal de proveedores de atención médica. En algunos casos, el proveedor o consultorio de atención médica también es un sitio de datos remoto. Cuando el cáncer es una enfermedad que se debe tratar, el sujeto puede padecer cáncer, entre otras posibles enfermedades o trastornos.

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 afección, proporcionar un pronóstico, monitorear la progresión de la enfermedad, determinar la recaída o remisión, o combinaciones de los mismos. La lista de fuentes apropiadas de información clínica para el cáncer incluye, pero no se limita a, tomografías computarizadas, resonancias magnéticas, ecografías, gammagrafías óseas, tomografías PET, análisis de médula ósea, rayos X con bario, endoscopias, linfangiogramas, IVU (urograma intravenoso) o IVP (pielograma intravenoso), punciones lumbares, cistoscopia, pruebas inmunológicas (examen de anticuerpos anti-malignina) y pruebas de marcadores de cáncer.

La información clínica del sujeto 6 puede obtenerse del laboratorio 2 de forma manual o automática. Cuando se desea simplicidad del sistema, la información puede obtenerse automáticamente a intervalos de tiempo predeterminados o regulares. Un intervalo de tiempo regular puede referirse a un intervalo de tiempo en el que la recogida de los datos de laboratorio se lleva a cabo automáticamente por los métodos y sistemas descritos en el presente documento basándose en una medición del tiempo, como horas, días, semanas, meses, años, etc. En una forma de realización, la recogida de datos y el procesamiento se llevan a cabo al menos una vez al día. En una forma de realización, la transferencia y recogida de datos se lleva a cabo aproximadamente de forma mensual, quincenal, semanal, varias veces a la semana o diaria. Alternativamente, la recuperación de información puede llevarse a cabo en intervalos de tiempo predeterminados, que pueden no ser intervalos de tiempo regulares. Por ejemplo, un primer paso de recuperación puede producirse después de una semana y un segundo paso de recuperación puede producirse después de un mes. La transferencia y recogida de datos se pueden personalizar según la naturaleza del trastorno que se está tratando y la frecuencia de las pruebas y los exámenes médicos requeridos de los sujetos.

La FIG. 9B muestra un proceso ejemplar para generar informes genéticos, incluyendo 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. Dichos 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 proceso puede comprender primero capturar información genética mediante la recolección de muestras de fluidos corporales como fuentes de material genético (por ejemplo, sangre, saliva, sudor, orina, etc.). A continuación, el proceso puede comprender la secuenciación de los materiales (11). Por ejemplo, los polinucleótidos en una muestra pueden secuenciarse, 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 tienen una variante con respecto al número total de lecturas de secuencia generadas a partir de la muestra. Cuando se analizan variantes en el número de copias, la carga tumoral puede estimarse como el exceso relativo (por ejemplo, en el caso de duplicación génica) o el déficit relativo (por ejemplo, en el caso de eliminación génica) del número total de lecturas de secuencia en los loci de prueba y control. Por ejemplo, una serie puede producir 1000 lecturas que se asignan a un locus de oncogén, de las cuales 900 corresponden al tipo salvaje y 100 corresponden a un mutante de cáncer, lo que indica una variante en el número de copias en este gen. Más detalles sobre la recolección de muestras ejemplares y la secuenciación de los materiales genéticos se discuten a continuación en las FIGS. 10-11.

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

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 simples bajos. Para 1000 lecturas de secuencia asignadas a un locus, se puede esperar que aproximadamente 50 lecturas (aproximadamente el 5 %) incluyan errores. Ciertas metodologías, como las descritas en WO 2014/149134, pueden reducir significativamente la tasa de error. Los errores crean ruido que puede oscurecer las señales del cáncer presente en niveles bajos en una muestra. Por ejemplo, si una muestra tiene una carga tumoral a un nivel cercano a la tasa de error del sistema de secuenciación, por ejemplo, alrededor del 0,1 %-5 %, puede ser difícil distinguir una señal correspondiente a una variante genética debida al cáncer de una debida al ruido.

El análisis de variantes genéticas se puede utilizar para diagnosticar en presencia de ruido. El análisis puede basarse en la frecuencia de variantes de secuencia o en el nivel de VNC (14) y se puede establecer una indicación de confianza diagnóstica o un nivel para detectar variantes genéticas en el rango de ruido (15).

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

Las mediciones de un parámetro, ya sea que se encuentren o no en el rango de ruido, pueden proporcionarse con un intervalo de confianza. Al realizar pruebas a lo largo del tiempo, se puede determinar si un cáncer está avanzando, estabilizándose 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 determinar si la enfermedad está aumentando o disminuyendo, porque no hay una diferencia estadísticamente significativa entre las mediciones. Sin embargo, cuando los intervalos de

confianza no se superponen, esto indica la dirección de la enfermedad. Por ejemplo, comparar el punto más bajo de un intervalo de confianza en un momento dado y el punto más alto de un intervalo de confianza en un segundo momento indica la dirección.

A continuación, el proceso puede comprender la generación de un informe/diagnóstico genético. El proceso puede comprender la generación de un gráfico genético para una pluralidad de mediciones que muestran la tendencia de mutación (18) y la generación de un informe que muestra los resultados y las opciones del tratamiento (19).

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

Las muestras que comprenden material genético, como ADNc, se pueden recolectar de un sujeto en una pluralidad de puntos temporales, es decir, en serie. El material genético se puede secuenciar, por ejemplo, utilizando un sistema de secuenciación de alto rendimiento. La secuenciación puede dirigirse a loci de interés para detectar variantes genéticas, como genes que presentan mutaciones somáticas, genes que experimentan variación en el número de copias o genes involucrados en fusiones genéticas, por ejemplo, en el cáncer. En cada punto temporal, se puede determinar una medida cuantitativa de las variantes genéticas encontradas. Por ejemplo, en el caso de ADNc, la medida cuantitativa puede ser la frecuencia o el porcentaje de una variante genética entre los polinucleótidos que se asignan a un locus, o el número absoluto de lecturas de secuencia o polinucleótidos que se asignan a un locus. Las variantes genéticas que tienen una cantidad distinta de cero en al menos un punto temporal se pueden representar gráficamente a través de todos los puntos temporales. 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 puede encontrarse en cantidades de 0, 10 y 20 en estos puntos temporales. Estas cantidades pueden normalizarse, para la variante 1, a 5 %, 3 % y 0 %, y, para la variante 2, a 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 temporales. Las cantidades normalizadas pueden escalarse 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 serían en el punto temporal 1: alturas de 5 mm (variante 1) y 0 mm (variante 2); en el punto temporal 2: alturas de 3 mm (variante 1) y 1 mm (variante 2), en el punto temporal 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 temporal "cero" (antes del primer punto temporal) puede representarse mediante un punto, con todos los valores en 0. La altura de la cantidad de las variantes en la representación gráfica puede ser, por ejemplo, relativa o proporcional entre sí. Por ejemplo, una frecuencia de variante del 5 % en un punto de tiempo podría representarse con una altura del doble de la de una variante con una frecuencia del 2,5 % en el mismo punto de tiempo. El orden de apilamiento puede elegirse para facilitar la comprensión. Por ejemplo, las variantes pueden apilarse en orden de cantidad de mayor a menor de abajo a arriba. O bien, pueden apilarse en un gráfico de flujo con la variante de mayor cantidad inicial en el medio y otras variantes de cantidad decreciente a cada lado. En ciertas formas de realización, las áreas se pueden codificar por colores según la variante. Las variantes del mismo gen se pueden mostrar en diferentes tonos del mismo color. Por ejemplo, los mutantes de KRAS se pueden mostrar en diferentes tonos de azul, los mutantes de EGFR en diferentes tonos de rojo.

Volviendo ahora a la FIG. 10A, el proceso puede comprender recibir información genética de un secuenciador de ADN (30). El proceso puede comprender luego 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 su representación en todos los puntos de prueba y luego generar un factor de escala (34). Como se utiliza en el presente documento, el término "normalizar" se refiere generalmente a medios para ajustar valores medidos en diferentes escalas a una escala nocionalmente común. Por ejemplo, los datos medidos en diferentes puntos se convierten/ajustan de modo que todos los valores se puedan redimensionar a una escala común. Como se utiliza en el presente documento, el término "factor de escala" se refiere generalmente 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 denominarse la constante de proporcionalidad de y a x. Los valores se normalizan para permitir la representación gráfica en una escala común que sea visualmente fácil de entender. Y el factor de escala se utiliza para conocer las alturas exactas que corresponden a los valores que se van a representar gráficamente (por ejemplo, la frecuencia de alelos mutantes 10 % puede representar 1 cm en el informe en el que la altura total es 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 representación de información en un mapa de respuesta tumoral (36). En la operación 36, el proceso puede comprender la representación de 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 la generación de 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 representa para el usuario.

En una implementación, la variación del número de copias puede informarse como un gráfico, que indica varias posiciones en el genoma y un aumento o disminución correspondiente o un mantenimiento de la variación del número de copias encada posición respectiva. Además, la variación del número de copias puede usarse para informar una puntuación porcentual que indica cuánto material patológico (o ácidos nucleicos que tienen una variación del número de copias) existe en la muestra de polinucleótidos sin células.

En otra forma de realización, 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 to Oncology™ o las pautas de práctica clínica de la Sociedad Americana de Oncología Clínica (ASCO). La anotación puede incluir la inclusión de una o más drogas aprobadas por la FDA para uso no indicado en la etiqueta, una o más drogas incluidas en compendios de tratamiento contra el cáncer de los Centros de Servicios de Medicare y Medicaid (CMS) y/o una o más drogas experimentales encontradas en la literatura científica, en el informe. La anotación puede incluir la conexión de una opción de tratamiento farmacológico incluida en la lista con una referencia que contiene información científica sobre la opción de tratamiento farmacológico. La información científica puede ser de un artículo revisado por pares de una revista médica. La anotación puede incluir la provisión de 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 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 consiste en 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 de información científica sobre una o más opciones de tratamiento farmacológico e información de ensayos clínicos sobre una o más opciones de tratamiento farmacológico.

La FIG. 10B muestra un proceso ejemplar para generar una vía de mapa de respuesta tumoral que puede ser utilizada por un profesional de la salud, por ejemplo, un médico, por ejemplo, para tomar decisiones sobre el cuidado del paciente. En esta forma de realización , el proceso puede comprender primero determinar un factor de escala global (43). En una forma de realización, para todas las frecuencias de alelos mutantes informadas que no son VNC (variación del número de copias), el proceso puede comprender transformar el valor absoluto en una métrica/escala relativa que puede ser más adecuada para representar gráficamente (por ejemplo, multiplicar la frecuencia de alelos mutantes por 100 y tomar el logaritmo de ese valor) y determinar un factor de escala global utilizando el valor máximo observado. El proceso luego implica visualizar información del conjunto de datos de prueba más antiguo (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 VNC, el proceso puede comprender la multiplicación del factor de escala por un valor transformado para cada gen y su uso como indicador de cantidad para representar gráficamente esa variante, y luego asignar un color/indicador visual único para cada alteración. A continuación, el proceso puede comprender la visualización de la información para los puntos de prueba posteriores (45) utilizando el siguiente pseudocódigo:

Si la composición de los resultados de la prueba no cambia, continúe con la fecha del panel anterior visualizándola en el nuevo panel.
Si las modificaciones siguen siendo las mismas, pero las cantidades han cambiado
Volver a calcular el indicador de cantidad para trazar esa variante y vuelva a trazar todos los valores actualizados en los paneles existentes y en el nuevo panel para la última fecha de prueba.
Si se añaden nuevas alteraciones

Añadir las modificaciones en la parte superior de todas las modificaciones existentes
Calcular los valores de transformación
Recalcular el factor de escala
Volver 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 alteraciones emergentes.

Si la alteración existente previa no se encuentra entre el conjunto de alteraciones detectadas

Utilizar una altura de cero y trace gráficamente la cantidad de alteración para todas las fechas de prueba posteriores.
Aún se incluye el color en el conjunto de colores no disponibles

Cada panel posterior que indica una fecha de prueba también puede incluir información adicional sobre el paciente o la intervención que puede correlacionarse con los cambios de alteración observados en el resto del mapa. También se pueden implementar escalas, representaciones gráficas y transformaciones similares en VNC 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 graficarse de manera similar en el mapa.

El proceso puede entonces comprender la determinación de un resumen de alteraciones y opciones de tratamiento (46). En una forma de realización, para la alteración con las frecuencias máximas de alelos mutantes, se realizan las siguientes acciones:

Informar todas las alteraciones de ese gen en orden decreciente de frecuencia de alelos mutantes de alteraciones no VNC
Informar todas las alteraciones de VNC para ese gen en orden decreciente de valor de VNC
Repetir para el próximo gen con la siguiente frecuencia de alelo mutante no VNC más alta aún no reportado
Para cada alteración informada, el proceso puede incluir la inclusión de un indicador de tendencia para esa alteración a lo largo de los diferentes puntos de fecha de prueba.
La agrupación de frecuencias máximas de alelos mutantes también puede extenderse más allá de los genes en los que se encuentran hacia anotaciones más encapsulantes, como vías biológicas, nivel de evidencia, etc.

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

La FIG. 10E muestra una sección de informe ejemplar que proporciona definiciones, comentarios e interpretación de las pruebas. Las FIGS. 10E-1 y 10E-2 proporcionan vistas ampliadas de la FIG. 10E. La FIG. 10F muestra una parte ejemplar del informe que contiene el resultado de la terapia detallada. 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 medicamentos potencialmente disponibles que se encuentran en fase de ensayos clínicos. La FIG. 10I-1 muestra los métodos de ensayo y las limitaciones de los mismos. Las FIGs. 101-1 y 101-2 proporcionan vistas ampliadas de la FIG. 10I.

La FIG. 10J-10P muestra varios gráficos de flujo modificados ejemplares 56. Un gráfico de flujo, o gráfico de corriente, es un tipo de gráfico de área apilada que se desplaza alrededor de un eje central, lo que da como resultado una forma orgánica fluida. Los gráficos de flujo son una generalización de los gráficos de área apilados donde la línea base es libre. Al desplazar la línea base, es posible minimizar el cambio en la pendiente (o "ondulación") en series individuales, lo que hace que sea más fácil percibir el espesor de cualquier capa dada a lo largo de los datos.

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

Una forma de realización presenta un gráfico de flujo de modo que no sea reflectante en el eje x. El gráfico modificado aplica una escala única para indicar 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 aumento o disminución correspondiente o mantenimiento de una frecuencia de mutación en cada posición respectiva. Además, las mutaciones se pueden utilizar para informar una puntuación porcentual que indica cuánto material de enfermedad existe en la muestra de polinucleótido libre de células. Una puntuación de confianza puede acompañar a cada mutación detectada, dadas las estadísticas conocidas de las variaciones típicas en las posiciones informadas en las secuencias de referencia sin enfermedad. Las mutaciones también se pueden clasificar en orden de abundancia en el sujeto o se pueden clasificar por importancia clínicamente procesable.

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 durante un período y volver a realizar la prueba. Si al final del período, el perfil de variación del número de copias, por ejemplo, como se representa en un mapa de respuesta tumoral, comienza a aumentar drásticamente, esto puede indicar que el tratamiento actual no está funcionando. También se puede realizar una comparación con los 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 según lo prescrito ya no trata el cáncer y se prescribe un nuevo tratamiento.

Estos informes pueden presentarse y accederse electrónicamente a través de Internet. El análisis de los datos de secuencia puede realizarse en un sitio distinto de la ubicación del sujeto. El informe puede generarse y transmitirse 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 procesos de prueba de genes ejemplares. Volviendo 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 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 parental 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 anterior para cada familia y, si existen puntuaciones de confianza anteriores consistentes, se aumenta la puntuación de confianza actual (1114). Si existen puntuaciones de confianza anteriores, pero son inconsistentes, la puntuación de confianza actual no se modifica en una forma de realización (1116). En otras formas de realización, la puntuación de confianza se ajusta de una manera predeterminada para puntuaciones de confianza anteriores inconsistentes. Si es la primera vez que se detecta la familia, la puntuación de confianza actual se puede reducir, 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 en función de la puntuación de confianza. A continuación, se generan los informes de pruebas genéticas como se explicó anteriormente (1120).

Aunque se ha utilizado información temporal 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. En otras formas de realización, la comparación histórica se puede utilizar junto con otras secuencias de consenso que se asignan a una secuencia de referencia particular para detectar casos de variación genética. Las secuencias de consenso que se asignan a secuencias de referencia particulares se pueden medir y normalizar frente a muestras de control. Las medidas de las moléculas que se asignan a secuencias de referencia se pueden comparar en un genoma para identificar áreas en el genoma en las que varía el número de copias o se pierde la heterocigosidad. Los métodos de consenso incluyen, por ejemplo, métodos lineales o no lineales de construcción de secuencias de consenso (por ejemplo, votación, promediado, estadístico, detección de máxima verosimilitud o máxima a posteriori, programación dinámica, métodos bayesianos, de Markov oculto o de máquina de vectores de soporte, etc.) derivados 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 ventana en estados de número de copias discretos. En algunos casos, este algoritmo puede comprender uno o más de los siguientes: modelo oculto de Markov, programación dinámica, máquina de vectores de soporte, red bayesiana, decodificación de enrejado, 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 secuencia con una muestra de control o secuencia de referencia puede ayudar en la normalización a través de ventanas. En esta forma de realización, 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 utilizando una variedad de métodos conocidos en la técnica, que incluyen, entre otros, precipitación con isopropanol y/o purificación a base de sílice. Los ADN libres de células se pueden extraer de cualquier número de sujetos, como sujetos sin cáncer, sujetos en riesgo de cáncer o sujetos que se sabe que tienen cáncer (por ejemplo, a través de otros medios).

Después del paso de aislamiento/extracción, se puede realizar cualquiera de varias operaciones de secuenciación diferentes en la muestra de polinucleótido libre de células. Las muestras se pueden procesar antes de la secuenciación 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 utilizar en una aplicación posterior, como una reacción de secuenciación, y se pueden rastrear moléculas individuales hasta las moléculas originales.

Los polinucleótidos libres de células pueden etiquetarse o rastrearse para permitir la identificación posterior y el 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. En algunos ejemplos, 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, lo que permite la identificación posterior de fragmentos de la hebra parental. En otros casos, los productos de expresión génica (por ejemplo, ARNm) pueden etiquetarse para cuantificar la expresión. Se puede contar un código de barras o un código de barras en combinación con la secuencia a la que está unido. En otros casos más, los sistemas y métodos pueden usarse como un control de amplificación por RCP. En tales casos, se pueden etiquetar múltiples productos de amplificación de una reacción de RCP 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 se pueden atribuir a un error de RCP. Además, se pueden identificar secuencias individuales basándose en las características de los datos de secuencia de la lectura en sí. Por ejemplo, se puede utilizar la detección de datos de secuencia únicos en las partes de inicio (inicio) y final (fin) de lecturas de secuenciación individuales, sola 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 cadena de ácido nucleico, a los que se les ha asignado una identidad única, pueden permitir así la identificación posterior de fragmentos de la cadena original. Esto se puede utilizar junto con el bloqueo del material genético de partida inicial para limitar la diversidad.

Además, el uso de datos de secuencia únicos en las partes de inicio (inicio) y final (fin) de lecturas de secuenciación individuales y la longitud de lectura de secuenciación se pueden utilizar, solos o en combinación, con el uso de códigos de barras. En algunos casos, los códigos de barras pueden ser únicos como se describe en este documento. En otros casos, los propios códigos de barras pueden no ser únicos. En este caso, el uso de códigos de barras no únicos, en combinación con datos de secuencia en las partes de inicio (inicio) y final (de terminación) de lecturas de secuenciación individuales y la longitud de las lecturas de secuenciación pueden permitir la asignación de una identidad única a secuencias individuales. De manera similar, los fragmentos de una sola cadena de ácido nucleico a los que se les ha asignado una identidad única pueden permitir la identificación posterior de fragmentos de la cadena original.

En general, los métodos y sistemas proporcionados en el presente documento son útiles para la preparación de secuencias de polinucleótidos libres de células para una reacción de secuenciación de aplicación posterior. A menudo, un método de secuenciación es la secuenciación clásica de Sanger. Los métodos de secuenciación pueden incluir, entre otros: 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, RNA-Seq (Illumina), expresión génica digital (Helicos), secuenciación de próxima generación, secuenciación de una sola molécula por síntesis (SMSS) (Helicos), secuenciación masiva en paralelo, matriz de una sola molécula clonal (Solexa), secuenciación shotgun, secuenciación Maxim-Gilbert, primer walking y cualquier otro método de secuenciación conocido en la técnica.

Los métodos de secuenciación típicamente implican 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 la muestra típicamente implica la conversión de polinucleótidos en una muestra en una forma compatible con la plataforma de secuenciación utilizada. Esta conversión puede implicar el etiquetado de polinucleótidos. En ciertas formas de realización de esta invención, las etiquetas comprenden 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 etiquetados. Los polinucleótidos que no se convierten en moléculas etiquetadas no están representados en una biblioteca etiquetada para secuenciación. En consecuencia, los polinucleótidos que tienen variantes genéticas representadas a 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, sea detectado por secuenciación. Además, en lugar de abordar directamente el problema de la baja eficiencia de conversión de la preparación de la biblioteca, 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 quiere la detección de polinucleótidos con baja representación, la alta eficiencia de conversión puede secuenciar eficientemente la muestra y/o detectar de manera adecuada tales polinucleótidos.

Generalmente, la detección de mutación puede realizarse en regiones selectivamente enriquecidas del genoma o transcriptoma purificado y aislado (1302). Como se describe en el presente documento, regiones específicas, que pueden incluir, entre otras, genes, oncogenes, genes supresores de tumores, promotores, elementos de secuencia reguladora, regiones no codificantes, miARN, ARNsn y similares, se pueden amplificar selectivamente a partir de una población total de polinucleótidos libres de células. Esto se puede realizar como se describe en el presente documento. En un ejemplo, se puede utilizar la secuenciación multiplex, con o sin etiquetas de código de barras para secuencias de polinucleótidos individuales. En otros ejemplos, la secuenciación se puede realizar utilizando 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 a partir de una muestra de control, tomada de otro sujeto. En algunos casos, el sujeto de control puede ser un sujeto que no padece aberraciones genéticas ni enfermedades conocidas. En algunos casos, estas lecturas de secuencias pueden contener información de código de barras. En otros ejemplos, no se utilizan códigos de barras.

Después de la secuenciación, a las lecturas se les puede asignar una puntuación de calidad. Una puntuación de calidad puede ser una representación de lecturas que indica si esas lecturas pueden ser útiles en un análisis posterior en función de un umbral. En algunos casos, algunas lecturas no tienen la calidad o la longitud suficientes para realizar el paso de mapeo posterior. Las lecturas de secuenciación con una puntuación de calidad de al menos el 90 %, 95 %, 99 %, 99,9 %, 99,99 % o 99,999 % se pueden filtrar del conjunto de datos. En otros casos, las lecturas de secuenciación a las que se

les ha asignado una puntuación de calidad de al menos el 90 %, 95 %, 99 %, 99,9 %, 99,99 % o 99,999 % se pueden filtrar del conjunto de datos. En el paso 1306, las lecturas de fragmentos genómicos que cumplen un umbral de puntuación de calidad especificado se asignan a un genoma de referencia o a una secuencia de referencia que se sabe que no contiene mutaciones. Después de la alineación de asignación, a las lecturas de secuencia se les asigna una puntuación de asignación. Una puntuación de asignación puede ser una representación o lecturas asignadas de nuevo a la secuencia de referencia que indica si cada posición es o no asignable de forma única. En algunos casos, 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 asignación de al menos el 90 %, 95 %, 99 %, 99,9 %, 99,99 % o 99,999 % se pueden filtrar del conjunto de datos. En otros casos, las lecturas de secuenciación a las que se les asigne un puntaje de mapeo menor a 90 %, 95 %, 99 %, 99,9 %, 99,99 % o 99,999 % pueden filtrarse del conjunto de datos.

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

La frecuencia de bases variantes se puede calcular como el número de lecturas que contienen la variante dividido por el número total de lecturas 1308 después de determinar la cobertura de lectura e identificar bases variantes en relación con la secuencia de control en cada lectura. Esto se puede expresar como una relación para cada posición mapeable en el genoma.

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

Los estados de mutación discretos de cada posición de base se pueden utilizar para identificar una variante de base con alta frecuencia de variación en comparación con la línea base de la secuencia de referencia. En algunos casos, la línea 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 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 base adyacentes con la variante o mutación de base se pueden fusionar 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.

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

Una mutación o variante puede comprender una aberración genética que incluye, pero no se limita a, una única sustitución de bases, o pequeñas indeles, transversiones, translocaciones, inversiones, deleciones, 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.

A continuación, se determina un consenso utilizando lecturas anteriores. Esto se hace revisando la puntuación de confianza anterior para las bases correspondientes y, si existen puntuaciones de confianza anteriores consistentes, entonces se aumenta la puntuación de confianza actual (1314). Si existen puntuaciones de confianza anteriores, pero son inconsistentes, la puntuación de confianza actual no se modifica en una forma de realización (1316). En otras formas de realización, la puntuación de confianza se ajusta de una manera predeterminada para puntuaciones de confianza anteriores inconsistentes. Si es la primera vez que se detecta la familia, la puntuación de confianza actual se puede reducir ya que puede ser una lectura falsa (1318). El proceso puede comprender luego convertir la frecuencia de varianza por cada base en estados de variante discretos para cada posición de base (1320).

Se pueden detectar numerosos cánceres utilizando los métodos y sistemas descritos en este documento. Las células cancerosas, como la mayoría de las células, se pueden caracterizar por una tasa de recambio, en la que las células viejas mueren y son reemplazadas por células más nuevas. Generalmente, las células muertas, en contacto con la vasculatura en un sujeto determinado, pueden liberar ADN o fragmentos de ADN en el torrente sanguíneo. Esto también es cierto para las células cancerosas durante varias etapas de la enfermedad. Las células cancerosas también se pueden caracterizar, dependiendo de la etapa de la enfermedad, por varias aberraciones genéticas, como la variación del número de copias, así como las mutaciones. Este fenómeno se puede utilizar para detectar la presencia o ausencia de individuos con cáncer utilizando los métodos y sistemas descritos en este documento.

Por ejemplo, se puede extraer sangre de sujetos en riesgo de cáncer y prepararla como se describe en este documento para generar una población de polinucleótidos libres de células. En un ejemplo, esto podría ser ADN libre de células. Los sistemas y métodos de la divulgación se pueden emplear para detectar mutaciones o variaciones del número de copias que pueden existir en ciertos cánceres presentes. El método puede ayudar a detectar la presencia de células cancerosas en el cuerpo, a pesar de la ausencia de síntomas u otras características de la enfermedad.

Los tipos y el número de cánceres que se pueden detectar pueden incluir, entre otros, cánceres de sangre, cánceres cerebrales, cánceres de pulmón, cánceres de piel, cánceres de nariz, cánceres de garganta, cánceres de hígado, cánceres de huesos, linfomas, cánceres de páncreas, cánceres de piel, cánceres de intestino, cánceres de recto, cánceres de tiroides, cánceres de vejiga, cánceres de riñón, cánceres de boca, cánceres de estómago, tumores en estado sólido, tumores heterogéneos, tumores homogéneos y similares.

El sistema y los métodos se pueden utilizar para detectar cualquier número de aberraciones genéticas que pueden causar o resultar de cánceres. Estas pueden incluir, entre otros, mutaciones, mutaciones, indels, variaciones del número de copias, transversiones, translocaciones, inversiones, deleciones, aneuploidía, aneuploidía parcial, poliploidía, inestabilidad cromosómica, alteraciones de la estructura cromosómica, fusiones genéticas, fusiones cromosómicas, truncamientos genéticos, amplificación genética, duplicaciones genéticas, lesiones cromosómicas, lesiones del ADN, cambios anormales en las modificaciones químicas de los ácidos nucleicos, cambios anormales en los patrones epigenéticos, cambios anormales en la metilación de los ácidos nucleicos, infección y cáncer.

Además, los sistemas y métodos descritos en el presente documento también pueden utilizarse para ayudar a caracterizar ciertos tipos de cáncer. Los datos genéticos producidos a partir del sistema y los métodos de esta divulgación 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 estadificación. 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 a un sujeto o médico pistas sobre el pronóstico de un tipo específico de cáncer.

Los sistemas y métodos proporcionados en este documento pueden utilizarse para controlar cánceres ya conocidos u otras enfermedades en un sujeto en particular. Esto puede permitir que un sujeto o un médico adapten las opciones de tratamiento de acuerdo con el progreso de la enfermedad. En este ejemplo, los sistemas y métodos descritos en este documento pueden utilizarse para construir perfiles genéticos de un sujeto en particular del curso de la enfermedad. En algunos casos, los cánceres pueden progresar, volviéndose más agresivos y genéticamente inestables. En otros ejemplos, los cánceres pueden permanecer benignos, inactivos o latentes. El sistema y los métodos de esta divulgación pueden ser útiles para determinar la progresión de la enfermedad.

Además, los sistemas y métodos descritos en el presente documento pueden ser útiles para determinar la eficacia de una opción de tratamiento particular. En un ejemplo, las opciones de tratamiento exitosas pueden en realidad aumentar la cantidad de variación del número de copias o mutaciones detectadas en la sangre del sujeto si el tratamiento es exitoso, ya que más cánceres pueden morir y desprender ADN. En otros ejemplos, esto puede no ocurrir. En otro ejemplo, tal vez ciertas opciones de tratamiento pueden correlacionarse con perfiles genéticos de cánceres a lo largo del tiempo. Esta correlación puede ser útil para seleccionar una terapia. Además, si se observa que un cáncer está en remisión después del tratamiento, los sistemas y métodos descritos en el presente documento pueden ser útiles para monitorear la enfermedad residual o la recurrencia de la enfermedad.

Diversas 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 ciertos casos, los trastornos genéticos o las enfermedades infecciosas pueden provocar un cierto mosaicismo genético en un sujeto. Este mosaicismo genético puede provocar variaciones en el número de copias y mutaciones que podrían observarse. También se pueden controlar los genomas de las células inmunitarias del organismo. Las células inmunitarias, como las células B, pueden experimentar una rápida expansión clonal ante la presencia de determinadas enfermedades. Las expansiones clonales se pueden controlar mediante la detección de variaciones en el número de copias y se pueden controlar determinados estados inmunitarios. En este ejemplo, se puede realizar un análisis de las variaciones en el número de copias a lo largo del tiempo para generar un perfil de cómo puede estar progresando una enfermedad en particular.

También se pueden controlar las infecciones sistémicas, ya que pueden estar causadas por un patógeno como una bacteria o un virus. La variación del número de copias o incluso la detección de mutaciones se pueden utilizar para determinar cómo cambia 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.

Otro ejemplo más es el seguimiento de sujetos trasplantados. Generalmente, el tejido trasplantado sufre un cierto grado de rechazo por parte del cuerpo tras el trasplante. Las actividades de rechazo del cuerpo huésped pueden determinarse o perfilarse, a medida que las células inmunitarias intentan destruir el tejido trasplantado. Esto puede ser útil para controlar el estado del tejido trasplantado, así como para modificar el curso del tratamiento o la prevención del rechazo.

Un perfil genético puede comprender una pluralidad de datos resultantes de la variación del número de copias y de los análisis de mutación. 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 propagado desde un sitio primario.

Los métodos de esta divulgación se pueden utilizar para generar o perfilar, identificar o crear un conjunto de datos que sea una suma de información genética derivada de diferentes células en una enfermedad heterogénea. Este conjunto de datos puede comprender análisis de variación del número de copias y de mutación solos o en combinación.

Adicionalmente, se pueden diagnosticar, pronosticar, monitorear u observar cánceres u otras enfermedades de origen fetal. Es decir, estas metodologías se pueden emplear en una embarazada para diagnosticar, pronosticar, monitorear u observar cánceres u otras enfermedades en un feto 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 secuencia se produce en un sitio distinto de la ubicación 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 farmacológico y/o proporcionar información sobre las opciones de tratamiento farmacológico a una compañía de seguros. El método puede incluir la anotación de 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 farmacológico que se estratifican en un informe se pueden anotar en el informe enumerando opciones de tratamiento farmacológico adicionales. Un tratamiento farmacológico adicional puede ser un fármaco aprobado por la FDA para un uso no indicado en la etiqueta. Una disposición de la Ley Ómnibus de Reconciliación Presupuestaria (OBRA) de 1993 exige que Medicare cubra los usos no indicados en la etiqueta de los fármacos contra el cáncer que se incluyen en los compendios médicos estándar. Los fármacos utilizados para anotar las listas se pueden encontrar en los compendios aprobados por CMS, incluidos el Compendio de Medicamentos y Productos Biológicos™ de la Red Nacional Integral del Cáncer (NCCN), DrugDex® de Thomson Micromedex, el compendio de Farmacología Clínica de Elsevier Gold Standard y el Compendio de Información sobre Medicamentos del Servicio de Formularios de los Hospitales Americanos®.

Las opciones de tratamiento farmacológico pueden anotarse enumerando un fármaco experimental que puede ser útil para tratar un cáncer con uno o más marcadores moleculares de un estado particular. El fármaco experimental puede ser un fármaco para el que se encuentran disponibles datos in vitro, datos in vivo, datos de modelos animales, datos de ensayos preclínicos o datos de ensayos clínicos. Los datos pueden publicarse en la literatura médica revisada por pares que se encuentra en las revistas enumeradas en el Manual de Políticas de Beneficios de Medicare de CMS, incluyendo, por ejemplo, American Journal of Medicine, Annals of Internal Medicine, Annals of Oncology, Annals of Surgical Oncology, Biology of Blood and Marrow Transplantation, Blood, Bone Marrow Transplantation, 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, and Radiation Oncology.

Las opciones de tratamiento farmacológico pueden anotarse proporcionando un enlace en un informe electrónico que conecte un fármaco enumerado con información científica sobre el fármaco. Por ejemplo, puede proporcionarse un enlace a información sobre un ensayo clínico para un fármaco (clinicaltrials.gov). Si el informe se proporciona a través de una computadora o un sitio web de computadora, el enlace puede ser una nota al pie, un hipervínculo a un sitio web, un cuadro emergente o un cuadro desplegable con información, etc. El informe y la información anotada pueden proporcionarse en un formato 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 enumerado en la información anotada, y el proveedor de atención médica puede acceder a la opción de tratamiento farmacológico anotada, recuperar la información científica (por ejemplo, imprimir un artículo de una revista médica) y enviarla (por ejemplo, un artículo de revista impreso) a una compañía

de seguros junto con una solicitud de reembolso por proporcionar el tratamiento farmacológico. Los médicos pueden utilizar cualquiera de una variedad de códigos de grupo relacionado con el diagnóstico (DRG) para permitir el reembolso.

Una opción de tratamiento farmacológico en un informe también puede incluir información sobre otros componentes moleculares en una vía a la que afecta un fármaco (por ejemplo, información sobre un fármaco que se dirige a una quinasa aguas abajo de un receptor de la superficie celular que es un objetivo farmacológico). La opción de tratamiento farmacológico puede incluir información sobre fármacos que se dirigen a uno o más componentes de otras vías moleculares. La identificación y/o anotación de información relacionada con vías puede externalizarse o subcontratarse a otra empresa.

La información anotada puede ser, por ejemplo, un nombre de fármaco (por ejemplo, un fármaco aprobado por la FDA para uso no indicado en la etiqueta; un fármaco que se encuentra en un compendio aprobado por CMS y/o un fármaco 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 fármacos, información de ensayos clínicos sobre uno o más fármacos (por ejemplo, información de clinicaltrials.gov/), uno o más enlaces a citas de información científica sobre medicamentos, etc.

La información anotada se puede insertar en cualquier ubicación de un informe. La información anotada se puede insertar en múltiples ubicaciones de un informe. La información anotada se puede insertar en un informe cerca de una sección sobre opciones de tratamiento farmacológico estratificadas. La información anotada se puede insertar en un informe en una página separada de las opciones de tratamiento farmacológico estratificadas. Un informe que no contiene opciones de tratamiento farmacológico estratificadas se puede anotar con información.

El sistema también puede incluir informes sobre los efectos de los fármacos en una muestra (por ejemplo, células tumorales) aislada de un sujeto (por ejemplo, un paciente con cáncer). Se puede establecer un cultivo in vitro utilizando un tumor de un paciente con cáncer utilizando técnicas conocidas por los expertos en la materia. El sistema también puede incluir un cribado de alto rendimiento de fármacos no aprobados por la FDA o fármacos experimentales utilizando dicho modelo de cultivo in vitro y/o xenoinjerto. El sistema también puede incluir el control del antígeno tumoral para la detección de recurrencia.

El sistema puede proporcionar acceso a través de Internet a 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 utilizado para automatizar el proceso de secuenciación de ADN. Dada 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, denominada lectura. Algunos secuenciadores de ADN también pueden considerarse instrumentos ópticos, ya que analizan señales de luz que se originan a partir de fluorocromos unidos a nucleótidos.

Los datos son enviados por los secuenciadores de ADN a través de una conexión directa o a través de Internet a una computadora para su procesamiento. Los aspectos de procesamiento de datos del sistema pueden implementarse en circuitos electrónicos digitales, o en hardware, firmware, software de computadora, o en combinaciones de ellos. El aparato de procesamiento de datos puede implementarse en un producto de programa de computadora incorporado de manera tangible en un dispositivo de almacenamiento legible por máquina para su ejecución por 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 sobre datos de entrada y generando una salida. Los aspectos de procesamiento de datos de la invención pueden implementarse de manera ventajosa en uno o más programas de computadora que son ejecutables en un sistema programable que incluye al menos un procesador programable acoplado para recibir datos e instrucciones desde y para transmitir datos e instrucciones a un sistema de almacenamiento de datos, al menos un dispositivo de entrada y al menos un dispositivo de salida. Cada programa informático puede implementarse en un lenguaje de programación de alto nivel, orientado a objetos o de procedimiento, o en lenguaje ensamblador o de máquina, si se desea; y, en cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado. Los procesadores adecuados incluyen, a modo de ejemplo, microprocesadores de propósito 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 manera tangible instrucciones y datos de programas informáticos incluyen todas las formas de memoria no volátil, incluyendo, a modo de ejemplo, dispositivos de memoria semiconductora, tales como EPROM, EEPROM y dispositivos de memoria flash; discos magnéticos tales como discos duros internos y discos extraíbles; discos magneto-ópticos; y discos CD-ROM. Cualquiera de los anteriores puede complementarse con, o incorporarse en, ASIC (circuitos integrados específicos de la aplicación).

Para proporcionar interacción con un usuario, la invención puede implementarse utilizando un sistema informático que tenga un dispositivo de visualización, como un monitor o una pantalla LCD (pantalla de cristal líquido) para mostrar 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 un trackball, o un dispositivo señalador tridimensional, como un guante de datos o un ratón giroscópico. El sistema informático puede programarse para proporcionar una interfaz gráfica de usuario a través de la cual los programas informáticos interactúan con los

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

Intervención terapéutica

Los métodos de esta divulgación permiten proporcionar intervenciones terapéuticas dirigidas con mayor precisión a la forma de una enfermedad en un sujeto, y calibrar estas intervenciones terapéuticas a lo largo del 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 tumoral. Por lo tanto, la intervención terapéutica es más eficaz contra los cánceres con este perfil que contra los cánceres con cualquiera de estas variantes.

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

Los objetivos farmacológicos en el cáncer y los fármacos eficaces contra estos objetivos se establecen en las Tablas 1 y 2 (tomados de Bailey et al., Discovery Medicine, v. 18 #92, 7/2/14).

Tabla 1. Ejemplos seleccionados de pruebas diagnósticas comercialmente disponibles, implicación de terapia asociada y el tipo de cáncer relevante.

Implicaciones de terapia	Prueba	Tipo de cáncer	Asociación clínica de fármaco-biomarcador
Ensayos IHQ			
Cetuximab; Panitumumab	EGFR	CRC	Establecido
Imatinib	C-KIT	TEGI	Establecido
Trastuzumab	HER2	Cáncer de mama; cáncer gástrico	Establecido
Resistencia a PI3K, AKT e inhibidores MEK	LKB1	CPCNP	Investigacional (Mahoney et al., 2009)
Crizotinib	C-MET	CPCNP	Investigacional (Sadiq & Salgia, 2013)
Inhibidores Akt/mTOR; resistencia a terapias de anti-EGFR	PTEN	CRC, CPCNP	Investigacional (Di Nicolantonio et al., 2010; Sos et al., 2009; Wang et al., 2012)
Ensayos de hibridación in situ			
Crizotinib	Fusión ALK FISH	CPCNP	Establecido
Trastuzumab; Pertuzumab	HER2 FISH	Cáncer de mama; cáncer gástrico	Establecido
Trastuzumab	HER2 CISH	Cáncer de mama	Establecido
Trastuzumab	HER2 ISH	Cáncer de mama	Establecido
Ensayos de mutación			
Cetuximab, Panitumumab	KRAS	CRC, CPCNP, cáncer pancreático	Establecido
Erlotinib, Gefitinib	EGFR	CPCNP, CRC	Establecido
Vemurafenib, Trametenib, Dabrafenib, Resistencia a terapias anti-EGFR	BRAF	CRC, cáncer del tiroides, melanoma	Establecido
Imatinib; TKI de 2ª generación	BCR-ABL	LMC, Ph+ LMA	Establecido
Crizotinib	ALK	CPCNP	Establecido
Inhibidores de RAF y MEK, resistencia a terapias de anti-EGFR	NRAS	Melanoma, CRC, CPCNP	Investigacional (Ascierto et al., 2013; De Mattos-Arruda et al., 2011; De Roock et al., 2010; Huang et al., 2013)
Imatinib	PDGFRA	TEGI	Establecido
Inhibidores PI3K/mTOR	PIK3CA	Cáncer de mama, CRC, cáncer pulmonar	Investigacional (Di Nicolantonio et al., 2010; Janku et al., 2013)

(Continuación)

Tabla 1. Ejemplos seleccionados de pruebas diagnósticas comercialmente disponibles, implicación de terapia asociada y el tipo de cáncer relevante.			
Implicaciones de terapia	**Prueba**	**Tipo de cáncer**	**Asociación clínica de fármaco-biomarcador**
Ensayos de mutación			
Inhibidores de Akt/mTOR; resistencia a terapias de anti-EGFR	PTEN	CRC, CPCNP, mama	Investigacional (Di Nicolantonio et al., 2010; Jerusalem et al., 2013; Sos et al., 2009; Wang et al., 2012)
Resistencia a inhibidores de PI3K, AKT y MEK	LKB1	CPCNP	Investigacional (Averette-Byers et al., 2012)
Otros			
Imatinib	Análisis de transcripción cuantitativa de BCR-ABL1	LMC, Ph+ LMA	Establecido
Resistencia a Imatinib	Nº de copias BCR-ABL1	LMC, Ph+ LMA	Investigacional (Hochhaus et al., 2002)
Inhibidores de PI3K	Amplificación de PIK3CA	Tipos múltiples de cáncer	Investigacional (Rodon et al., 2013)
Erlotinib; Getfitnib; Cetuximab; Panitumumab	Amplificación de EGFR	CPCNP, CRC	Investigacional (Gupta et al., 2009)
Nota: Las asociaciones clínicas de fármaco-biomarcador denominadas "Establecidas" reflejan evidencias bien conocidas de fármacos de la FDA. Las indicadas “investigacional” son asociaciones que están hipotetizadas y demostradas por la literatura científica.			

Tabla 2. Terapias e indicaciones objetivo aprobadas por la FDA de los EE. UU.				
Agente	**Denominación social**	**Objetivo(s)**	**Indicación(es) aprobada(s) por la FDA**	**Empresa**
Anticuerpos monoclonales				
Ado-trastuzumab emtansina (T-DM1)*	Kadcyla	HER2	Cáncer de mama (HER2+)*	Genentech
Bevacizumab	Avastin	VEGF	CRC	Genentech
			GBM	
			NCLC	
			CCR	
Cetuximab*	Erbitux	EGFR	CRC (tipo silvestre KRAS)	Eli Lilly
			CCECC	
Ipilimumab	Yervoy	CTLA-4	Melanoma	Bristol-Myers Squibb
Obinutuzumab	Gazyva	CD-20	LLC	Genentech
Panitumumab*	Vectibix	EGFR	CRC (tipo silvestre KRAS)*	Amgen
Pertuzumab	Perjeta	HER2	Cáncer de mama (HER2+)*	Genentech
Trastuzumab*	Herceptin	HER2	Cáncer de mama (HER2+)*	Genentech
			Cáncer gástrico (HER2+)*	

(Continuación)

Tabla 2. Terapias e indicaciones objetivo aprobadas por la FDA de los EE. UU.				
Agente	**Denominación social**	**Objetivo(s)**	**Indicación(es) aprobada(s) por la FDA**	**Empresa**
Inhibidores de molécula pequeña				
Afatinib*	Gilotrif	EGFR, HER2	CPCNP (con deleciones EGFR exón 19 o sustitución L858R)*	Boehringer Ingelheim
Axitinib	Inlyta	KIT, PDGFRβ, VEGFR1/2/3	CCR	Pfizer
Bosutinib*	Bosulif	ABL	LMC (positivo por cromosoma Filadelfia)*	Pfizer
Cabozantinib	Cometriq	FLT3, KIT, MET, RET, VEGFR2	Cáncer de tiroides medular	Exelixis
Crizotinib*	Xalkori	ALK, MET	CPCNP (con fusión ALK)*	Pfizer
Dabrafenib*	Tafinlar	BRAF	Melanoma (con mutación BRAF V600E)*	GlaxoSmithKline
Dasatinib*	Sprycel	ABL	LMC (positivo por cromosoma Filadelfia)*	Bristol-Myers Squibb
			LLA (positivo por cromosoma Filadelfia)*	
Denosumab	Xgeva	RANKL	Tumor de células gigantes de hueso	Amgen
Erlotinib*	Tarceva	EGFR	CPCNP (con deleciones de exón 19 o sustituciones L858R)*	Genentech & OSI
			Cáncer pancreático	

(Continuación)

Tabla 2. Terapias e indicaciones objetivo aprobadas por la FDA de los EE. UU.				
Agente	**Denominación social**	**Objetivo(s)**	**Indicación(es) aprobada(s) por la FDA**	**Empresa**
Inhibidores de molécula pequeña				
Everolimus*	Afinitor	mTOR	Tumor neuroendocrino pancreático	Novartis
			CCR	
			Cáncer de mama (ER/PR+) en combinación con exemestano*	
			Astrocitoma de células gigantes subependimales no reseccionables con esclerosis tuberosa	
Gefitinib	Iressa	EGFR	CPCNP con beneficio previo conocido de gefitinib (aprobación limitada)	AstraZeneca
Ibrutinib	Imbruvica	TQB	Linfoma de mastocito	Pharmacyclics
Imatinib*	Gleevec	KIT, PDGFR, ABL	Tumor estromal GI	Novartis
			Dermatofibrosarcoma protuberans	
			Neoplasmias malignas hematológicas múltiples que incluyen LLA y LMC positivos por cromosoma Filadelfia*	
Lepatinib*	Tykerb	HER2, EGFR	Cáncer de mama (HER2+)*	GlaxoSmithKline
Nilotinib*	Tasigna	ABL	LMC (positivo por cromosoma Filadelfia)*	Novartis
Pazopanib	Votrient	VEGFR, PDGFR, KIT	CCR	GlaxoSmithKline
			Sarcoma de tejido blando	
Regorafenib	Stivarga	KIT, PDGFRβ, RAF, RET, VEGFR1/2/3	CRC	Bayer
			Tumores estromales gastrointestinales	
Ruxolitinib	Jakafi	JAK1/2	Mielofibrosis	Incyte
Sorafenib	Nexavar	VEGFR, PDGFR, KIT, RAF	Carcinoma hepatocelular	Bayer
			CCR	

(Continuación)

Tabla 2. Terapias e indicaciones objetivo aprobadas por la FDA de los EE. UU.				
Agente	**Denominación social**	**Objetivo(s)**	**Indicación(es) aprobada(s) por la FDA**	**Empresa**
Inhibidores de molécula pequeña				
Sunitinib	Sutent	VEGFR,	TEGI	Pfizer
		PDGFR, KIT, RET	Tumor neuroendocrino pancreático	
			CCR	
Temsirolimus	Torisel	mTOR	CCR	Wyeth
Trametinib*	Mekinist	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; TEGI, tumor del estroma gastrointestinal; RE, receptor de estrógeno; RP receptor de progesterona; CPCNP, cáncer de pulmón de células no pequeñas; CC, cáncer colorrectal; GBM, glioblastoma; CCR, carcinoma de células renales; CCECC, carcinoma de células escamosas de cabeza y cuello; LLC, leucemia linfoblástica crónica; TQB: tirosina quinasa de Bruton. *Terapia dirigida que se asocia con una alteración de substancia de cáncer molecularmente específica. Hay aproximadamente 17 terapias dirigidas que están asociadas con 10 subtipos moleculares específicos de cáncer.				

En una forma de realización, en función del 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 que se trataría 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 inferiores, al menos 10 veces inferiores o al menos 100 veces inferiores a las de los clones dominantes detectados, estas variantes pueden dejarse fuera de la intervención terapéutica hasta que alcancen un umbral clínico o una frecuencia relativa significativa (por ejemplo, superior al umbral indicado anteriormente).

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

En una forma de realización adicional, 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 medicamentos eficaces contra las dos variantes menos comunes.

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

En otra forma de realización, se examinan los cambios en el perfil de heterogeneidad tumoral a lo largo del tiempo, y se desarrollan intervenciones terapéuticas para tratar el tumor cambiante. Por ejemplo, la heterogeneidad de la enfermedad se puede determinar en una pluralidad de momentos diferentes. Utilizando 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 a lo largo del 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 tienen el marcador de falta de respuesta.

En respuesta a la quimioterapia, una forma tumoral 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 resistentes puede retrasarse. Por ejemplo, un sujeto puede ser sometido a uno o más ciclos de terapia pulsátil, comprendiendo cada ciclo de terapia pulsátil un primer período durante el cual se administra un fármaco en una primera cantidad y un segundo ciclo durante el cual se administra el fármaco 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. El primer y segundo nivel clínico pueden ser diferentes en diferentes ciclos de terapia pulsátil. Por ejemplo, el primer nivel clínico puede ser más bajo en los ciclos siguientes. Una pluralidad de ciclos puede incluir al menos 2, 3, 4, 5, 6, 7, 8 o más ciclos. Por ejemplo, el mutante BRAF V600E puede detectarse en polinucleótidos de células de la enfermedad en una cantidad que indica una carga tumoral del 5 % en el ADNlc. La quimioterapia puede iniciarse el tratamiento con dabrafenib. Las pruebas posteriores pueden mostrar que la cantidad de la mutación BRAF en el ADNc cae por debajo del 0,5 % o a niveles indetectables. En este punto, se puede suspender o reducir significativamente la terapia con dabrafenib. Las pruebas posteriores pueden mostrar que el ADN que lleva la mutación BRAF ha aumentado al 2,5 % de los polinucleótidos en el ADNc. En este punto, se reinicia la terapia con dabrafenib, por ejemplo, al mismo nivel que el tratamiento inicial. Las pruebas posteriores pueden mostrar que el ADN que lleva la mutación BRAF ha disminuido al 0,5 % de los polinucleótidos en el ADNc. Nuevamente, se suspende o reduce la terapia con dabrafenib. El ciclo se puede repetir varias veces.

La Figura 7 muestra un curso ejemplar de monitoreo y tratamiento de la enfermedad en un sujeto. Un sujeto examinado 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, el tratamiento se interrumpe. 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, presentándose con una variante genética en el Gen 4. El sujeto ahora es sometido a un tratamiento con el Fármaco B, al que responden los cánceres que tienen esta variante.

En otra forma de realización, se puede cambiar una intervención terapéutica tras la detección del surgimiento de una forma mutante resistente a un fármaco 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 a ruxolitinib. Un método de esta divulgación implica monitorear los cambios en el perfil tumoral y cambiar una intervención terapéutica cuando una variante genética asociada con la resistencia a fármacos aumenta hasta un nivel clínico predeterminado.

Base de datos

En otra forma de realización, se construye una base de datos en la que se registra la información genética de muestras seriadas recogidas de pacientes con cáncer. Esta base de datos también puede contener el tratamiento intermedio y otra información clínicamente relevante, como peso, efectos adversos, pruebas histológicas, análisis de sangre, información radiográfica, tratamientos anteriores, tipo de cáncer, etc. Los resultados de las pruebas seriadas se pueden utilizar para inferir la eficacia del tratamiento, especialmente cuando se utilizan con muestras de sangre, que pueden dar una estimación más imparcial de la carga tumoral que la autodeclaración o la declaración radiográfica de un médico. La eficacia del tratamiento se puede agrupar por aquellos con perfiles genómicos similares y viceversa. Los perfiles genómicos se pueden organizar en torno a, por ejemplo, alteración genética primaria, alteración(es) genética(s) secundaria(s), cantidades relativas de estas alteraciones genéticas y carga tumoral. Esta base de datos se puede utilizar para apoyo a la toma de decisiones para pacientes posteriores. Tanto las alteraciones de la línea germinal como las somáticas se pueden utilizar para determinar la eficacia del tratamiento. Las alteraciones de la resistencia adquirida también se pueden inferir de la base de datos cuando los tratamientos que fueron efectivos inicialmente comienzan a fallar. Este fracaso se puede detectar mediante radiografías, sangre u otros medios. Los datos primarios utilizados para inferir los mecanismos de resistencia adquiridos son los perfiles genómicos tumorales recopilados después del tratamiento por paciente. Estos datos también se pueden utilizar para establecer límites cuantitativos sobre la probable respuesta al tratamiento, así como para predecir el tiempo hasta el fracaso del tratamiento. probables alteraciones de resistencia adquiridas para un tratamiento determinado y el perfil genómico del tumor, se puede modificar un régimen de tratamiento para suprimir la adquisición de las alteraciones de resistencia más probables.

Sistemas informáticos

Los métodos de la presente divulgación se pueden implementar utilizando, o con la ayuda de, sistemas informáticos. La 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 el presente documento) 1505. El sistema informático 1501 también incluye una memoria o ubicación de memoria 1510 (por ejemplo, memoria de acceso aleatorio, memoria de solo lectura, memoria flash), una unidad de almacenamiento electrónico 1515 (por ejemplo, disco duro), una interfaz de comunicación

1520 (por ejemplo, adaptador de red) para comunicarse con uno o más sistemas diferentes, y dispositivos periféricos 1525, como caché, otra memoria, almacenamiento de datos y/o adaptadores de pantalla electrónica. 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 repositorio de datos) para almacenar datos. El sistema informático 1501 puede estar acoplado operativamente a una red informática ("red") 1530 con la ayuda de la interfaz de comunicación 1520. La red 1530 puede ser Internet, una red de Internet y/o extranet, o una intranet y/o extranet que esté en comunicación con Internet. La red 1530 en algunos casos es una red de telecomunicaciones y/o de datos. La red 1530 puede incluir uno o más servidores informáticos, que pueden permitir la computación distribuida, como la computación en la nube. La CPU 1505 puede ejecutar una secuencia de instrucciones legibles por máquina, que pueden estar incorporadas 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 como los descritos en este documento se pueden implementar por medio de un código ejecutable de máquina (por ejemplo, procesador de computadora) almacenado en una ubicación del sistema informático 1501, como, por ejemplo, en la memoria 1510 o la unidad de almacenamiento electrónico 1515. El código ejecutable por máquina o legible por máquina se puede proporcionar en forma de software. Los aspectos de los sistemas y métodos proporcionados en este documento, como el sistema informático 1501, se pueden incorporar en la programación. Varios aspectos de la tecnología se pueden considerar como "productos" o "artículos de fabricación" típicamente en forma de código ejecutable por máquina (o procesador) y/o datos asociados que se transportan o incorporan en un tipo de medio legible por máquina. El código ejecutable por máquina se puede almacenar en una unidad de almacenamiento electrónico, como una memoria (por ejemplo, memoria de solo lectura, memoria de acceso aleatorio, memoria flash) o un disco duro. Los medios de tipo "almacenamiento" pueden incluir cualquiera o toda la memoria tangible de las computadoras, procesadores o similares, o módulos asociados a las mismas, tales 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. Todo el software o partes del mismo pueden comunicarse en ocasiones a través de Internet o de varias 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 de análisis de muestras.

EJEMPLOS

Las posiciones de nucleótidos (por ejemplo, loci) en el genoma se pueden designar por número, como se muestra en la Figura 2. Las posiciones en las que aproximadamente el 100 % de las denominaciones de bases son idénticas a la secuencia de referencia o en las que aproximadamente el 100 % de las denominaciones de bases son diferentes a la secuencia de referencia se infieren para representar la homocigosidad del ADNc (presuntamente normal). Las posiciones en las que aproximadamente el 50 % de las denominaciones de bases son idénticas a la secuencia de referencia se infieren para representar la heterocigosidad del ADNc (también presuntamente normal). Las posiciones en las que el porcentaje de denominaciones de bases en un locus está sustancialmente por debajo del 50 % y por encima del límite de detección del sistema de denominación de bases se infieren para representar variantes genéticas asociadas a tumores.

Ejemplo 1. Métodos para la detección de variaciones en el número de copias

Extracción de sangre

Se recogen muestras de sangre de 10 a 30 mL a temperatura ambiente. Las muestras se centrifugan para eliminar las células. El plasma se recoge después de la centrifugación.

Extracción de ADNc

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 kit QIAamp DNA Blood Mini) y se eluye en una solución de 100 µl. Los ADN por debajo de 500 pb se seleccionan con captura de perlas magnéticas Ampure SPRI (PEG/sal). La producción resultante se suspende en 30 µl de H_2O. Se comprueba la distribución del tamaño (pico mayor = 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 = 3 K HGE; 10 µg de ADN = 3 M HGE.

Preparación de la biblioteca de "Molécula Única"

El marcado de ADN de alta eficiencia (>80 %) se realiza mediante reparación de extremos, unión de extremos pegajosos y colas A con 2 octómeros diferentes (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 porción no complementaria. Ambos extremos de cada fragmento de ADN están unidos con

adaptadores de horquilla. Cada fragmento marcado se puede identificar mediante una combinación de la secuencia de octómero en los adaptadores de horquilla y porciones endógenas de la secuencia de inserción.

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

Se puede realizar una optimización del tampón, una optimización de la polimerasa y una reducción del ciclo 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. El ruido (por ejemplo, los errores introducidos por la polimerasa) se reducen mediante el uso de polimerasas de alta fidelidad.

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

Secuenciación masiva paralela

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

Bioinformática digital

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

Las secuencias de consenso resultantes se asignan a un genoma de referencia, como hg19. Cada base en una secuencia de consenso está cubierta por aproximadamente 3000 familias diferentes. Se calcula el puntaje de calidad de cada secuencia y las secuencias se filtran en función de sus puntajes de calidad. Las llamadas de bases en cada posición de una secuencia de consenso se comparan con la secuencia de referencia HG-19. En cada posición en la que una llamada de bases difiere de la secuencia de referencia, se determina y se informa la identidad de la base o bases diferentes y su porcentaje en función del total de llamadas de bases 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 (homocigotos) y el 2 % tienen una base diferente, es probable que el locus tenga una variante de secuencia, presumiblemente de ADN de cáncer.

La VNC se detecta contando el número total de secuencias (bases) asignadas a un locus y comparándolas con un locus de control. Para aumentar la detección de VNC, el análisis de VNC se realiza en regiones específicas, incluidas las regiones en genes ALK, APC, BRAF, CDKN2A, EGFR, ERBB2, FBXW7, KRAS, MYC, NOTCH1, NRAS, PIK3CA, PTEN, RBI, TP53, MET, AR, ABL1, AKT1, ATM, CDHI, CSFIR, CTNNBI, ERBB4, EZH2, FGFR1, FGFR2, FGFR3, FLT3, GNA11, GNAQ, GNAS, HNFIA, HRAS, IDH1, IDH2, JAK2, JAK3, KDR, KIT, MLH1, MPL, NPM1, PDGFRA, PROC, PTPN11, RET, SMAD4, SMARCB1, SMO, SRC, STK11, VHL, TERT, CCND1, CDK4, CDKN2B, RAF1, BRCA1, CCND2, CDK6, NF1, TP53, ARID1A, BRCA2, CCNE1, ESR1, RIT1, GATA3, MAP2K1, RHEB, ROS1, ARAF, MAP2K2, NFE2L2, RHOA o NTRK1.

Ejemplo 2. Método para corregir la determinación de bases mediante la determinación del número total de moléculas no observadas en una muestra

Después de que los fragmentos se amplifican y las secuencias de los fragmentos amplificados se leen y alinean, los fragmentos se someten a una llamada de bases. 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 observados.

Cuando se realiza una búsqueda de bases para el locus A (un locus arbitrario), primero se supone que hay N fragmentos amplificados. Las lecturas de secuencia pueden proceder de dos tipos de fragmentos: fragmentos de doble cadena y fragmentos de cadena sencilla. El siguiente es un ejemplo teórico de cálculo del número total de moléculas no observadas en una muestra.

N es el número total de moléculas en la muestra.

Suponiendo que 1000 es el número de dúplex detectados.

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.

Dado que Q = 1 -P.

1000 = NP(2).
500 = N2PQ.
1000 / P(2) = N.
500÷2 PQ = N.
1000 / P(2) = 500÷2PQ.
1000 * 2 PQ = 500 P(2).
2000 PQ = 500 P(2).
2000 Q = 500 P.
2000 (1-P) = 500P
2000-2000 P = 500P.
2000 = 500P + 2000 P.
2000 = 2500 P.
2000÷2500 = P.
0,8 = P.
1000/ P(2) = N.
1000÷0,64 = N.
1562 = N.
Número de fragmentos no vistos = 62.

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

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

El ADN libre de células se extrae del plasma de un paciente y se amplifica por RCP. Las variantes genéticas se analizan mediante secuenciación masiva paralela de los genes diana amplificados. Para un conjunto de genes, todos los exones se secuencian ya que dicha cobertura de secuenciación ha demostrado tener 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. La amplificación puede no detectarse en muestras con cantidades inferiores de ADN libre de células y/o amplificación de copias de genes de bajo nivel. Ciertas características de la muestra o variante dieron como resultado una sensibilidad analítica reducida, como una calidad de muestra baja o una recolección inadecuada.

El porcentaje de variantes genéticas encontradas en el ADN libre de células que circula en la sangre está relacionado con la biología tumoral única de este paciente. Los factores que afectaron la cantidad/porcentajes de variantes genéticas detectadas en el ADN circulante de células libres en la sangre incluyen el crecimiento del tumor, la extensión, el tamaño, la heterogeneidad, la vascularización, la progresión de la enfermedad o el tratamiento. La Tabla 5 indica el porcentaje o la frecuencia de alelos de ADN libre de células circulantes alteradas (% ADNcl) detectado en este paciente. Algunas de las variantes genéticas detectadas se enumeran en orden descendente por % ADNcl.

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

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

Los resultados de las pruebas de pacientes que comprenden las variantes genéticas se enumeran en la Tabla 6.

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

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

En 3 posiciones, un nucleótido detectado con una frecuencia inferior al 9 % es diferente de un nucleótido en la secuencia de referencia. Estas 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.

Se calculan las cantidades relativas de variantes genéticas asociadas a tumores. La relación de cantidades de BRAF:NRAS:JAK2 es 8,9: 2,6: 1,5, o 1: 0,29: 0,17. A partir 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 VNC puede mostrar amplificación de BRAF, en cuyo caso el 100 % de las células tumorales pueden tener variantes tanto en BRAF como en NRAS.

Tabla 3. Genes en los que están secuenciados todos los exones

GENES EN LOS QUE ESTÁN SECUENCIADOS TODOS LOS EXONES			
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. Frecuencia mínima detectable de alelos mutantes para esta muestra en la que se detecta el 80 % de variantes somáticas.

Tabla 4. Genes en los que se secuencian exones con una mutación somática previamente reportada

GENES EN LOS QUE SE SECUENCIAN EXONES CON UNA MUTACIÓN SOMÁTICA PREVIAMENTE REPORTADA			
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. Frecuencia mínima detectable de alelos mutantes para esta muestra en la que se detecta el 80 % de variantes somáticas.

Tabla 5. Frecuencia de alelos del ADN libre circulante alterado detectado en este paciente

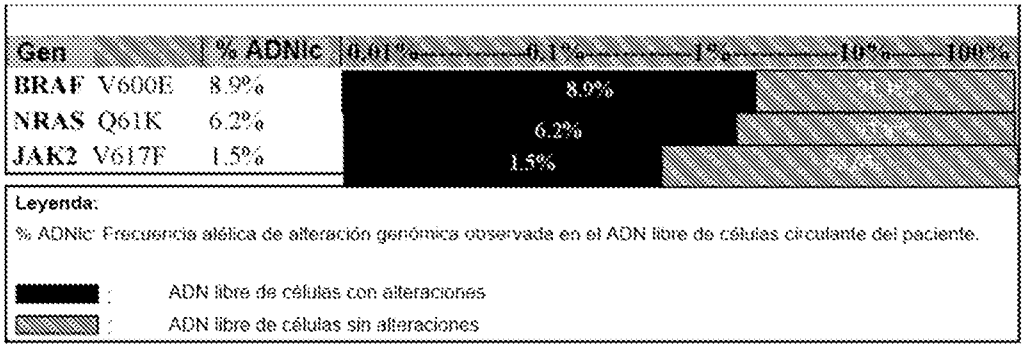


Tabla 6. Alteraciones genómicas detectadas en genes seleccionados

Detectado: 51 Alteraciones Genómicas

Gen	Cromosoma	Posición	Mutación (nt)	Mutación (AA)	Porcentaje	ID cósmico	DBSNP ID
KRAS	12	25368462	C>T		100.0%		rs4362222
ALK	2	29416572	T>C	I1461V	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	55206417	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
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%	COSM28026	
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 límites de detección específicos del paciente para genes analizados mediante ensayos

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

Ejemplo 5. Corrección de errores de secuencia comparando secuencias Watson y Crick

Se aísla ADN libre de células de doble cadena del plasma de un paciente. Los fragmentos de ADN libre 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 ligadura. Después de la ligadura, cada uno de los fragmentos de ADN libre de células se puede identificar de forma distintiva mediante la secuencia de los códigos de barras distintivos y dos secuencias endógenas de 20 pb en cada extremo del fragmento de ADN libre de células.

Los fragmentos de ADN libres de células marcados se amplifican mediante RCP. Los fragmentos amplificados se enriquecen utilizando perlas que comprenden sondas de oligonucleótidos que se unen específicamente a un grupo de genes asociados al cáncer. Por lo 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 de 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.

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

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

Las lecturas de secuencia se agrupan en familias. Los errores en las lecturas de secuencia se corrigen. La secuencia de consenso de cada familia se genera mediante colapso.

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 de NRAS responden al tratamiento con trametinib. Los cánceres con mutantes de JAK2 responden al tratamiento con ruxolitinib. Se ha determinado que una intervención terapéutica que incluya la administración de trametinib y ruxolitinib es más eficaz contra este cáncer que el tratamiento con cualquiera de los fármacos mencionados anteriormente por separado. El sujeto recibe tratamiento con una combinación de trametinib y ruxolitinib en una proporción de dosis de 5:1.

Después de varias rondas de tratamiento, se vuelve a analizar el ADNc del sujeto para detectar la presencia de 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 la que se determina que trametinib y ruxolitinib son eficaces en una proporción de dosis de 1:1. Se administra al sujeto un ciclo de

quimioterapia en cantidades en esta proporción. Las pruebas posteriores muestran que los mutantes BRAF, NRAS y JAK2 están presentes en el ADNc en cantidades inferiores al 1 %.

Ejemplo 7. Intervención terapéutica

Se toma una muestra de sangre de un individuo con melanoma antes del tratamiento y se determina que el paciente tiene una mutación BRAF V600E a una concentración del 2,8 % y ninguna mutación NRAS detectable mediante un análisis de ADN libre de células. Se le administra al paciente una terapia anti BRAF (dabrafenib). Después de 3 semanas, se toma otra muestra de sangre y se analiza. Se determina que el nivel de BRAF V600E ha descendido al 0,1 %. Se interrumpe el tratamiento y se repite la prueba cada 2 semanas. El nivel de BRAF V600E vuelve a aumentar y se reinicia el tratamiento cuando el nivel de BRAF V600E aumenta al 1,5 %. Se interrumpe nuevamente el tratamiento cuando el nivel desciende al 0,1 % nuevamente. Este ciclo se repite.

Ejemplo 8. Corrección de VNC en función de las mediciones de VNCOR

Se determinan las variaciones del número de copias en una muestra de paciente. Los métodos para determinar pueden incluir el seguimiento molecular y el sobremuestreo, como se describió anteriormente. Se utiliza un modelo de Markov oculto basado en las ubicaciones esperadas de los orígenes de replicación para eliminar el efecto de la proximidad del origen de replicación a partir 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.

En muchos casos, el nivel de células libres derivadas de tumores puede ser bajo o estar por debajo del límite de detección de una tecnología particular. Este puede ser el caso cuando el número de equivalentes del genoma humano de ADN derivado de tumores 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 lo tanto, se administra un procedimiento al paciente una dosis mínima de quimioterapia con efectos adversos antes de la extracción de sangre para aumentar preferentemente la fracción de ADN derivado del tumor que se obtiene. Por ejemplo, se podría administrar al paciente una dosis baja de quimioterapia y se podría extraer una muestra de sangre en un plazo de 24, 48, 72 horas o menos de una semana. Para lograr una quimioterapia eficaz, esta muestra de sangre contiene concentraciones más altas de ADN derivado del tumor libre de células debido a las tasas potencialmente más altas de muerte celular de las células cancerosas. Como alternativa, se aplica radioterapia de dosis baja a través de un instrumento radiográfico de cuerpo entero o de forma local en las regiones afectadas en lugar de quimioterapia de dosis baja. Se prevén otros procedimientos, como someter al paciente a ultrasonidos, ondas sonoras, ejercicio, estrés, etc.

REIVINDICACIONES

1. Un método que comprende:

(a) realizar análisis genómico y análisis epigenético de ADN libre de células de células cancerosas espacialmente distintas de un sujeto, en donde

(i) el análisis genómico se realiza mediante secuenciación de ADN, y
(ii) el ADN libre de células procede de una muestra de fluido corporal del sujeto;

(b) identificar y cuantificar variantes biomoleculares en el ADN libre de células;
(c) desarrollar un perfil de heterogeneidad tumoral en el sujeto que indique la presencia y cantidad relativa de una pluralidad de variantes en el ADN libre de células, en donde diferentes cantidades relativas indican heterogeneidad tumoral; y
(d) determinar una intervención terapéutica para un cáncer que presenta la heterogeneidad tumoral, en donde la intervención terapéutica es efectiva contra un cáncer que tiene el perfil de heterogeneidad tumoral determinado e incluye una terapia de combinación que comprende un primer fármaco efectivo contra un primer tipo de tumor y un segundo fármaco efectivo contra un segundo tipo de tumor.

2. El método de la reivindicación 1, en el que el análisis epigenético es un análisis de metilación.

3. El método de la reivindicación 1 o 2, en el que el método comprende un seguimiento molecular, por ejemplo, que comprende el etiquetado de polímeros biomoleculares 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, comprendiendo además opcionalmente el método el filtrado de llamadas de bases que representan ruido.

4. El método de cualquier reivindicación anterior, en el que el método comprende monitorear cambios en el perfil de heterogeneidad del tumor a lo largo del tiempo.

5. El método de cualquier reivindicación anterior, en el que el perfil de heterogeneidad del tumor se determina en una pluralidad de momentos diferentes, y el método comprende, además calibrar las intervenciones terapéuticas a lo largo del tiempo.

6. El método de cualquier reivindicación precedente, en el que el método comprende el uso de una base de datos para identificar la intervención terapéutica eficaz para un sujeto que tiene cáncer, en el que la base de datos incluye, para cada uno de una pluralidad de sujetos que tienen cáncer, datos de pruebas genómicas del tumor, incluyendo alteraciones somáticas, recopilados en dos o más intervalos de tiempo por sujeto, una o más intervenciones terapéuticas administradas a cada uno de los sujetos en uno o más momentos y la eficacia de las intervenciones terapéuticas.

7. El método de cualquier reivindicación precedente, en el que el método comprende, además generar y proporcionar (A) un informe de la heterogeneidad del tumor y/o (B) un informe de los resultados de las pruebas de cáncer y las opciones de tratamiento, por ejemplo, proporcionando el informe a un profesional sanitario tal como un médico.

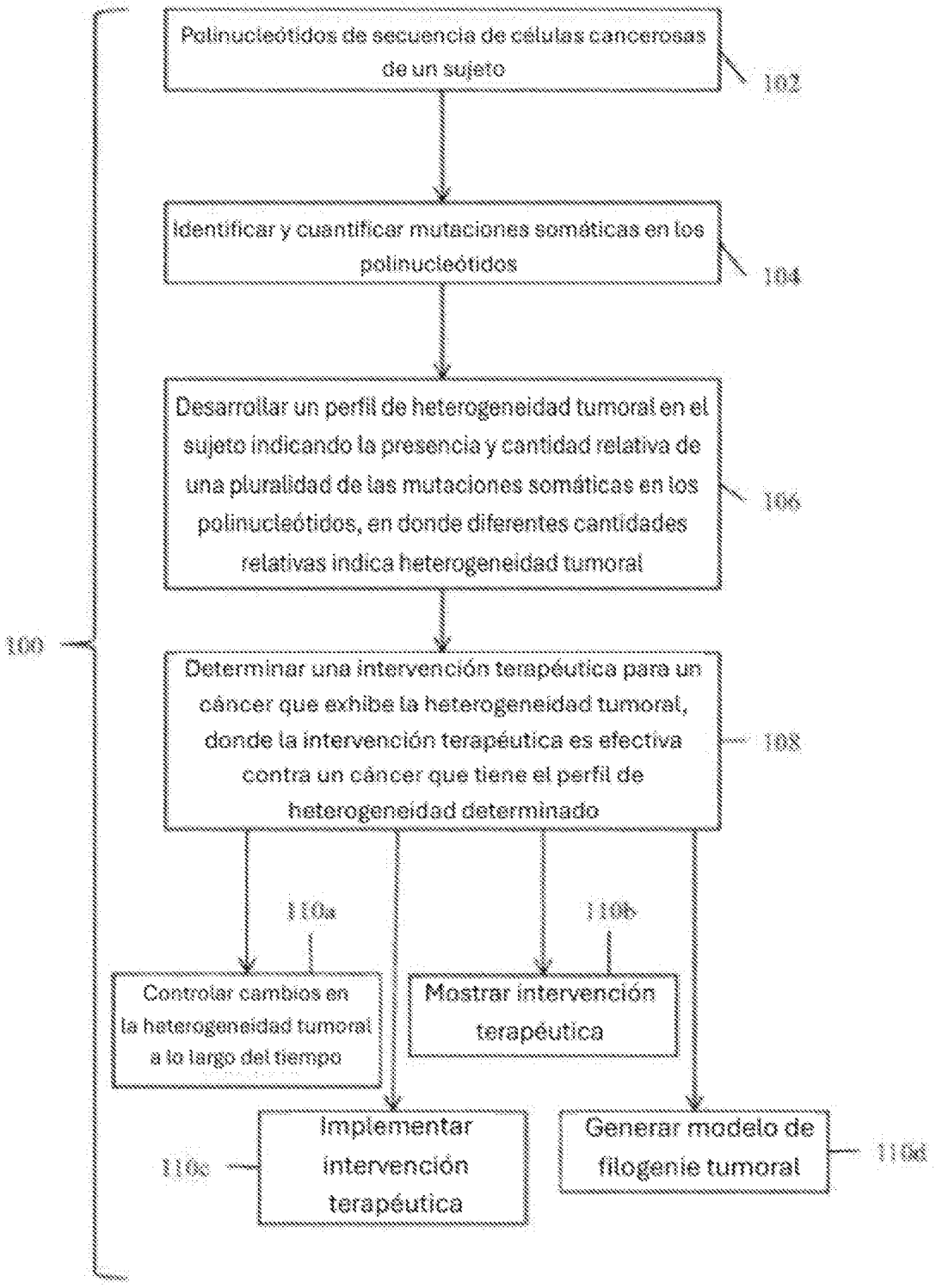

Polinucleótidos de secuencia de células cancerosas de un sujeto
102
Identificar y cuantificar mutaciones somáticas en los polinucleótidos
104
Desarrollar un perfil de heterogeneidad tumoral en el sujeto indicando la presencia y cantidad relativa de una pluralidad de las mutaciones somáticas en los polinucleótidos, en donde diferentes cantidades relativas indica heterogeneidad tumoral
106
100
Determinar una intervención terapéutica para un cáncer que exhibe la heterogeneidad tumoral, donde la intervención terapéutica es efectiva contra un cáncer que tiene el perfil de heterogeneidad determinado
108
110a
110b
Controlar cambios en la heterogeneidad tumoral a lo largo del tiempo
Mostrar intervención terapéutica
110c
Implementar intervención terapéutica
Generar modelo de filogenie tumoral
110d

FIG. 1

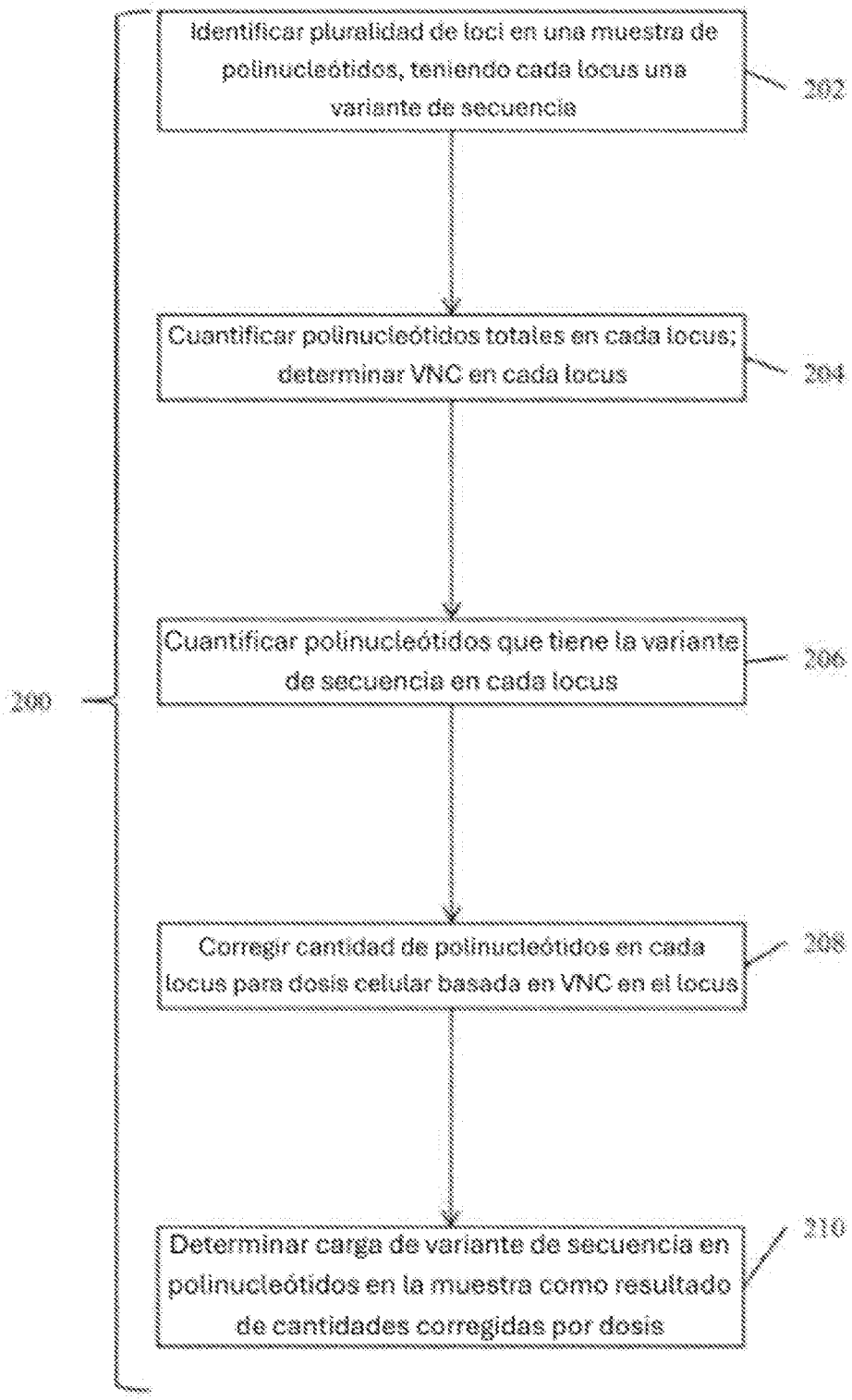

Identificar pluralidad de loci en una muestra de polinucleótidos, teniendo cada locus una variante de secuencia
202
Cuantificar polinucleótidos totales en cada locus; determinar VNC en cada locus
204
Cuantificar polinucleótidos que tiene la variante de secuencia en cada locus
206
200
Corregir cantidad de polinucleótidos en cada locus para dosis celular basada en VNC en el locus
208
Determinar carga de variante de secuencia en polinucleótidos en la muestra como resultado de cantidades corregidas por dosis
210

FIG. 2

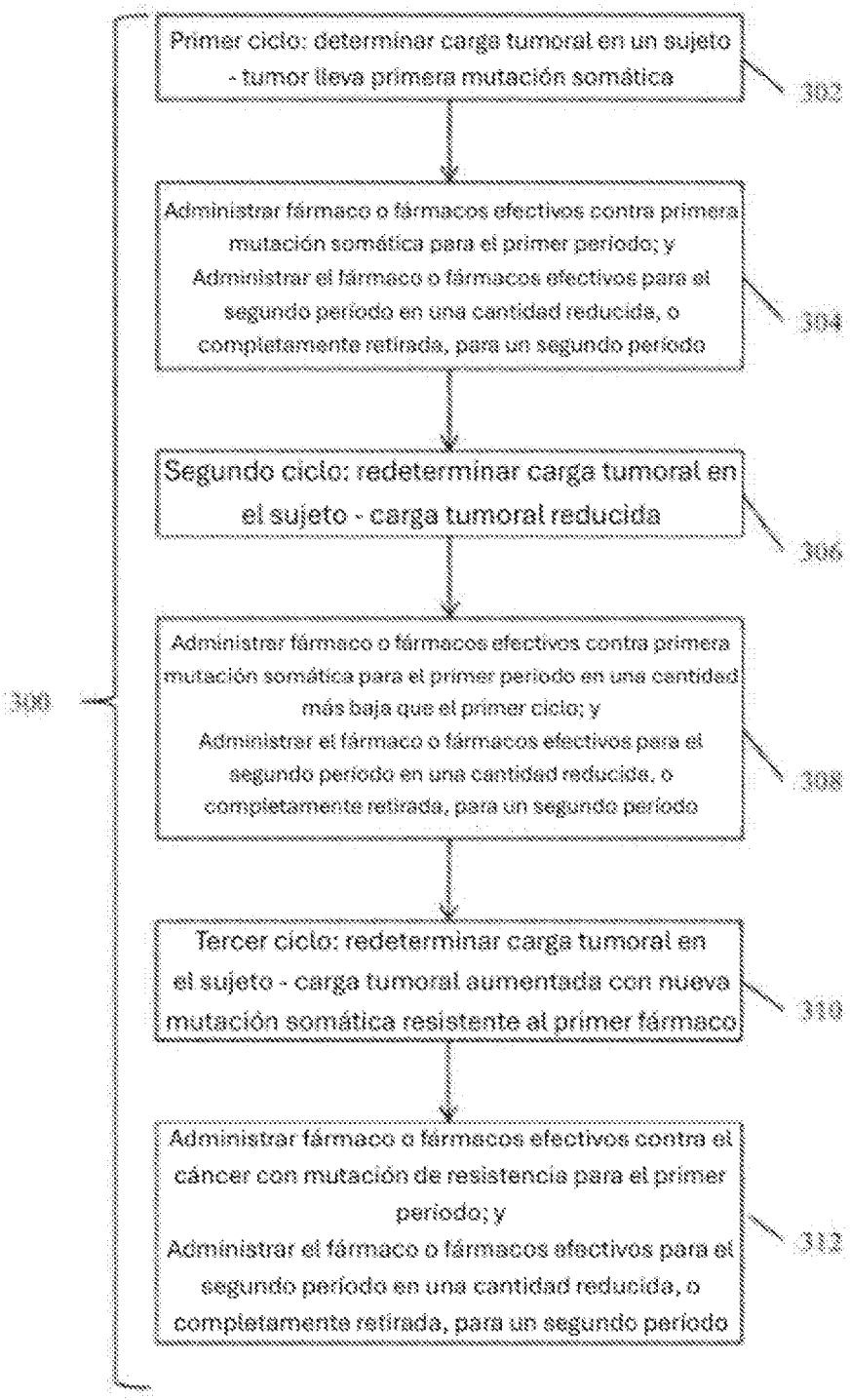

Primer ciclo: determinar carga tumoral en un sujeto - tumor lleva primera mutación somática
302
Administrar fármaco o fármacos efectivos contra primera mutación somática para el primer periodo; y
Administrar el fármaco o fármacos efectivos para el segundo periodo en una cantidad reducida, o completamente retirada, para un segundo periodo
304
Segundo ciclo: redeterminar carga tumoral en el sujeto - carga tumoral reducida
306
Administrar fármaco o fármacos efectivos contra primera mutación somática para el primer periodo en una cantidad más baja que el primer ciclo; y
Administrar el fármaco o fármacos efectivos para el segundo periodo en una cantidad reducida, o completamente retirada, para un segundo periodo
308
Tercer ciclo: redeterminar carga tumoral en el sujeto - carga tumoral aumentada con nueva mutación somática resistente al primer fármaco
310
Administrar fármaco o fármacos efectivos contra el cáncer con mutación de resistencia para el primer periodo; y
Administrar el fármaco o fármacos efectivos para el segundo periodo en una cantidad reducida, o completamente retirada, para un segundo periodo
312
300

FIG. 3

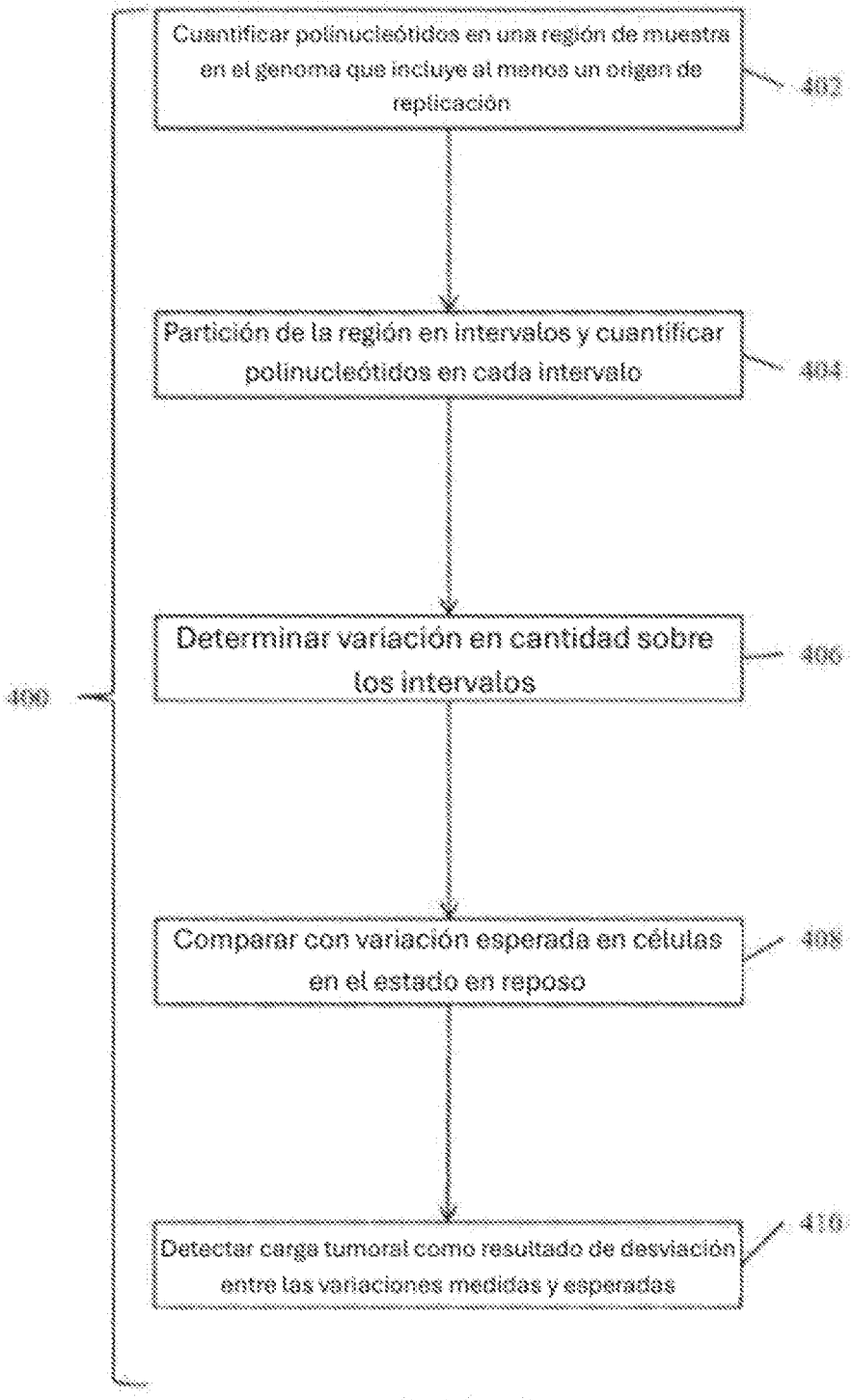

Cuantificar polinucleótidos en una región de muestra en el genoma que incluye al menos un origen de replicación
402
Partición de la región en intervalos y cuantificar polinucleótidos en cada intervalo
404
Determinar variación en cantidad sobre los intervalos
406
408
Comparar con variación esperada en células en el estado en reposo
408
Detectar carga tumoral como resultado de desviación entre las variaciones medidas y esperadas
410

FIG. 4

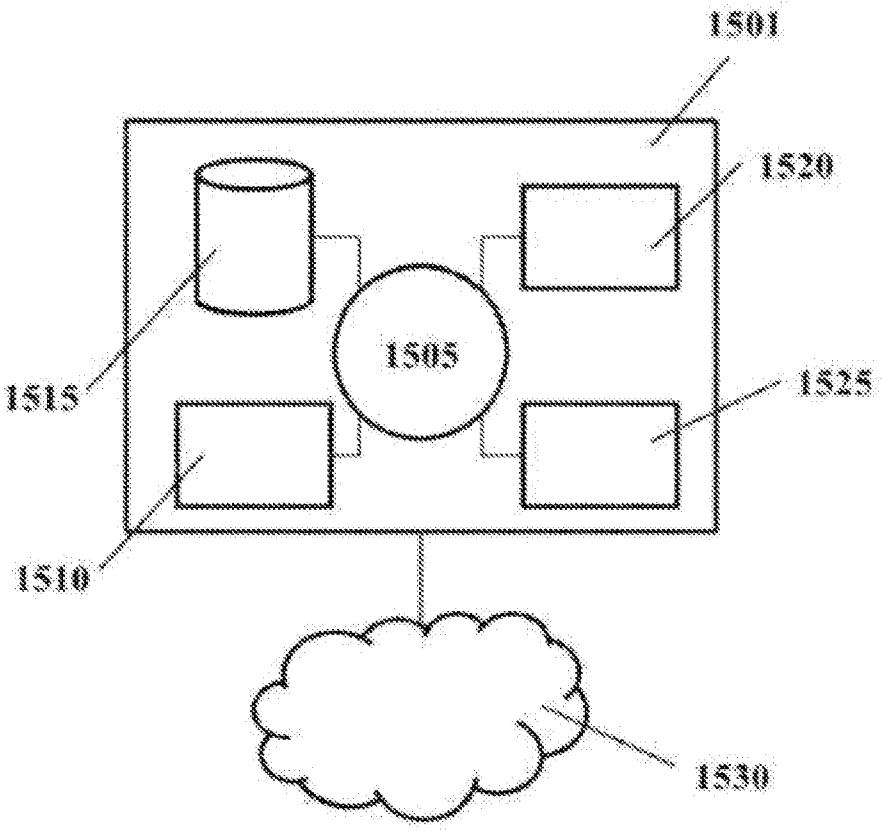


FIG. 5

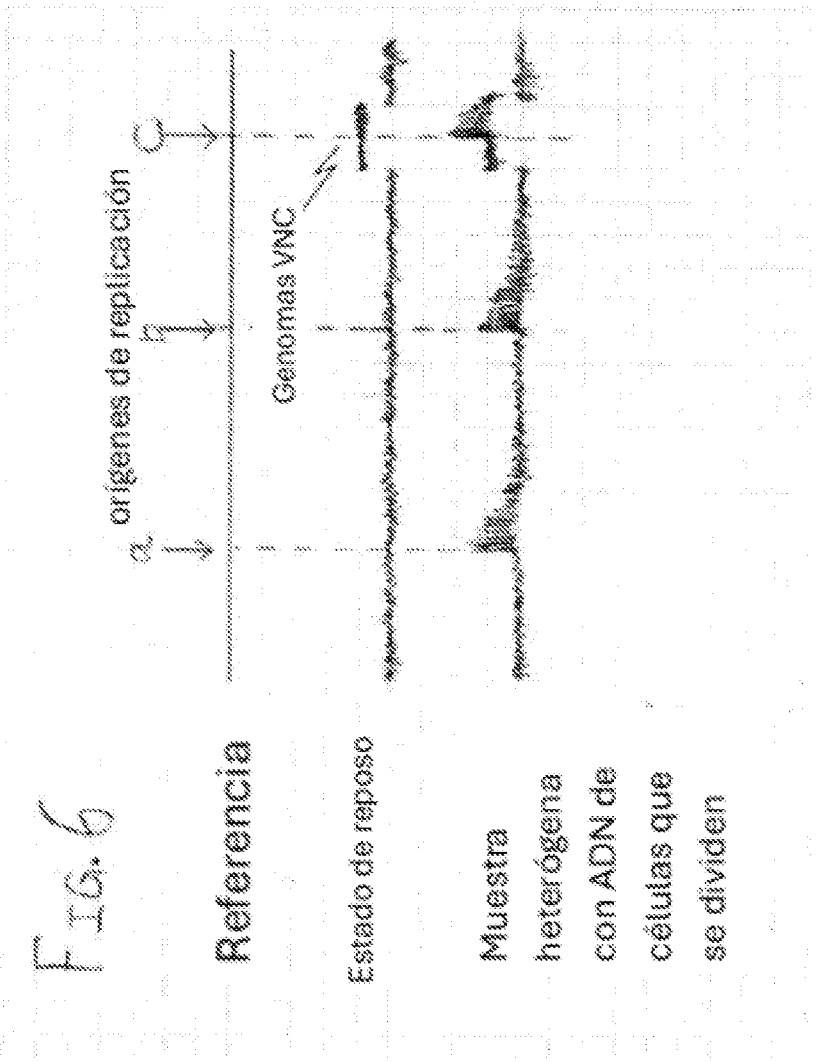

FIG. 6
Referencia
Estado de reposo
Muestra heterógena con ADN de células que se dividen
orígenes de replicación
a
b
c
Genomas VNC

FIG. 6

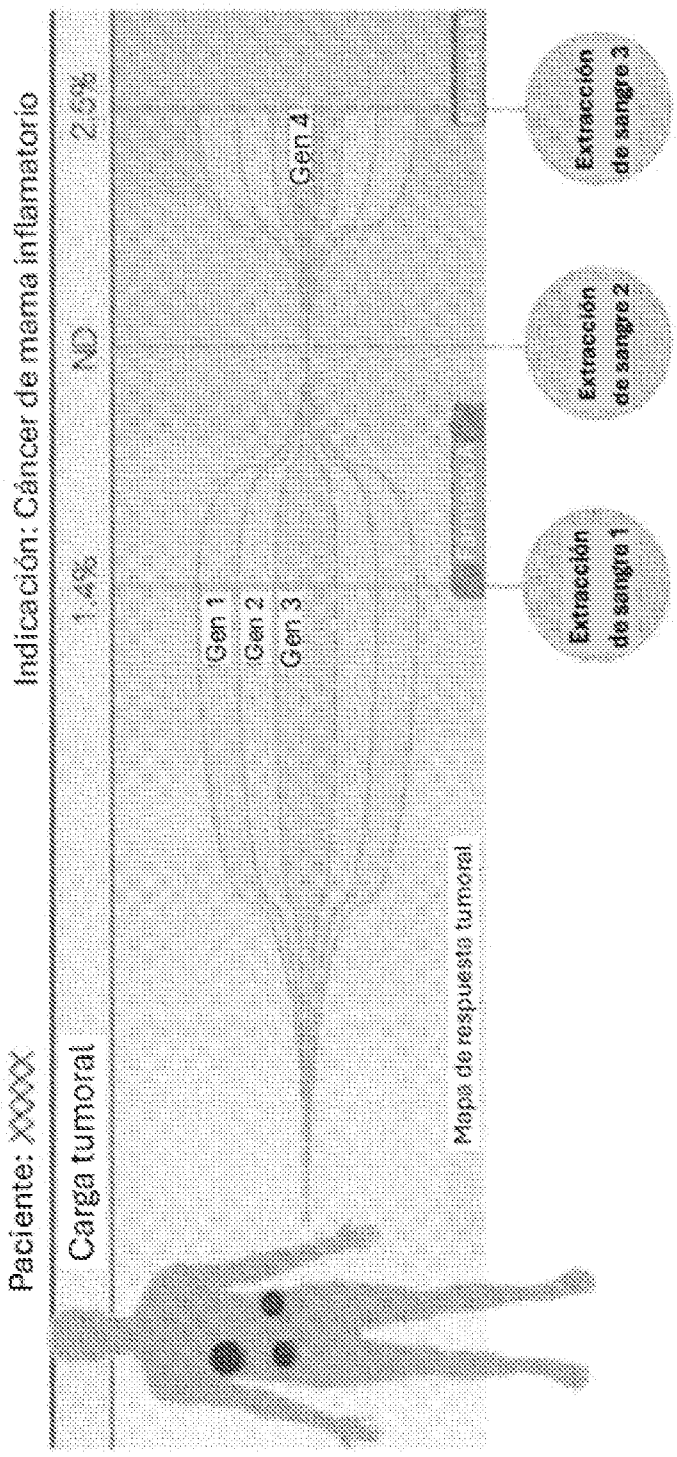

Paciente: XXXX
Indicación: Cáncer de mama inflamatorio
Carga tumoral
1.4%
MD
2.8%
Gen 1
Gen 2
Gen 3
Gen 4
Mapa de respuesta tumoral
Extracción de sangre 1
Extracción de sangre 2
Extracción de sangre 3

FIG. 7

Secuenciación de exón completa
Mutaciones de punto/INDEL
(70 genes)
AKT1 ALK APC AR ARAF ARID1A AR BRAF AJK
ATM BRAF BRCA1 BRCA2 CCND1 CCND2 CCNE1 CDK4 FGFR2
CCND1 CDH1 CDK4 CDK8 CDKN2A, CDKN2B CDK6 EGFR FGFR3
CTNNB1 EGFR ERBB2 ESR1 ESR2 FBXW7 ERBB2 FGFR3 KIT1
FGFR1 FGFR2 FGFR3 GATA3 GNA11 GNAQ PDGFRA KIT RET
GNAS HNF1A KRAS IDH1 IDH2 JAK2 KRAS MET ROS1
JAK3 KIT KRAS MAP2K1 MAP2K2 MET MYD PDGFRA
NOTCH1 NPM1 PIK3CA PTEN RB1 SMAD4 SMO NOTCH1 PIK3CA RAF1
NPM1 NRAS NTRK1 PDGFRA PIK3CA PTEN
PTPN11 RAF1 RB1 RET RHEB RHOA
RET ROS1 SMAD4 SMO SRC STK11
TERT TP53 TSC1 VHL
*exones 18 y 20
*exones 19 y 20
**exón 14 salto

FIG. 8

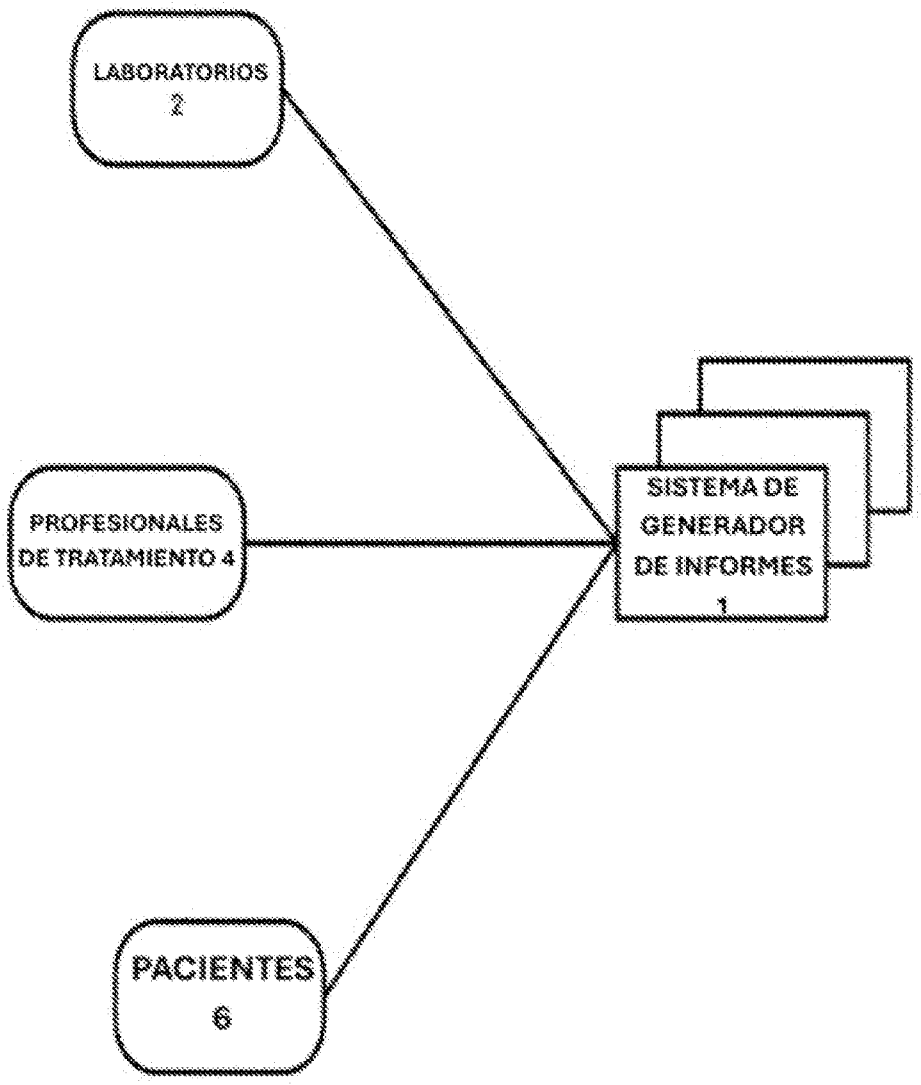

LABORATORIOS
2
PROFESIONALES
DE TRATAMIENTO 4
PACIENTES
6
SISTEMA DE
GENERADOR
DE INFORMES
1

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 secuencia

b) Variantes de número de copia

2. Frecuencia de variantes genéticas en muestra que contiene material genético

C. Separar información de ruido (13)

II. Diagnóstico de cáncer

A. Basado en frecuencia de variantes de secuencia o nivel de VNC (14)

B. Seguridad de detectar variantes genéticas en el rango de ruido (15)

III. Aumentar seguridad de diagnóstico

A. Utilizar una pluralidad de mediciones para aumentar seguridad de diagnóstico (16)

B. Utilizar mediciones en una pluralidad de puntos temporales para determinar si el cáncer avanza, en remisión o estabilizado (17)

IV. Generar informe/diagnóstico

A. Generar gráfico genético para una pluralidad de mediciones mostrando tendencia de mutación (18)

B. Generar informe mostrando resultados y opciones de tratamiento (19)

FIG. 9B

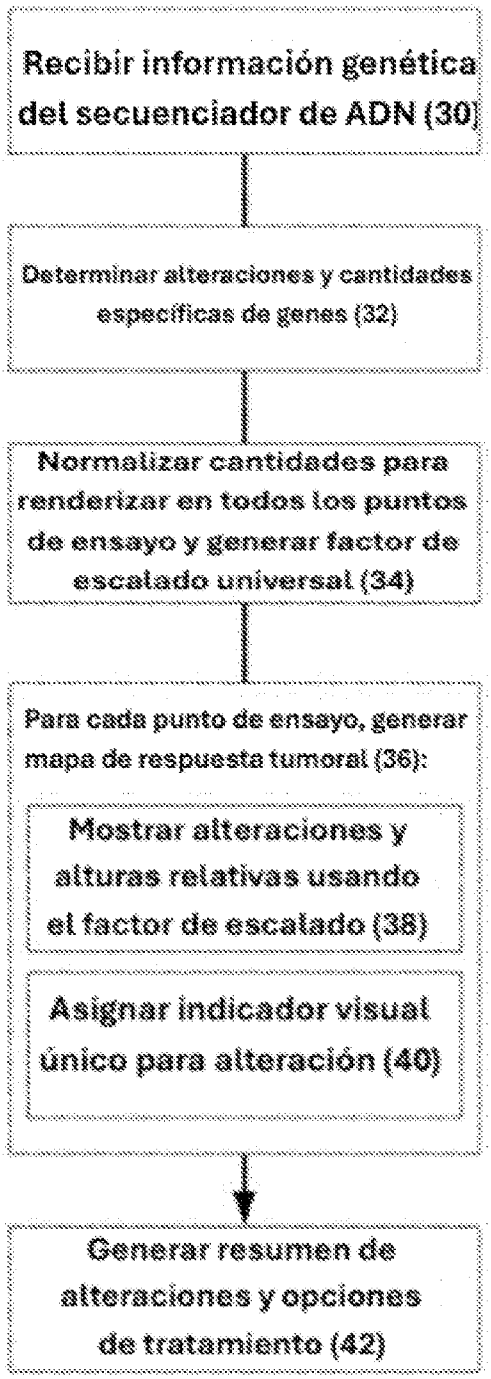

Recibir información genética del secuenciador de ADN (30)
Determinar alteraciones y cantidades específicas de genes (32)
Normalizar cantidades para renderizar en todos los puntos de ensayo y generar factor de escalado universal (34)
Para cada punto de ensayo, generar mapa de respuesta tumoral (36):
Mostrar alteraciones y alturas relativas usando el factor de escalado (38)
Asignar indicador visual único para alteración (40)
Generar resumen de alteraciones y opciones de tratamiento (42)

FIG. 10A

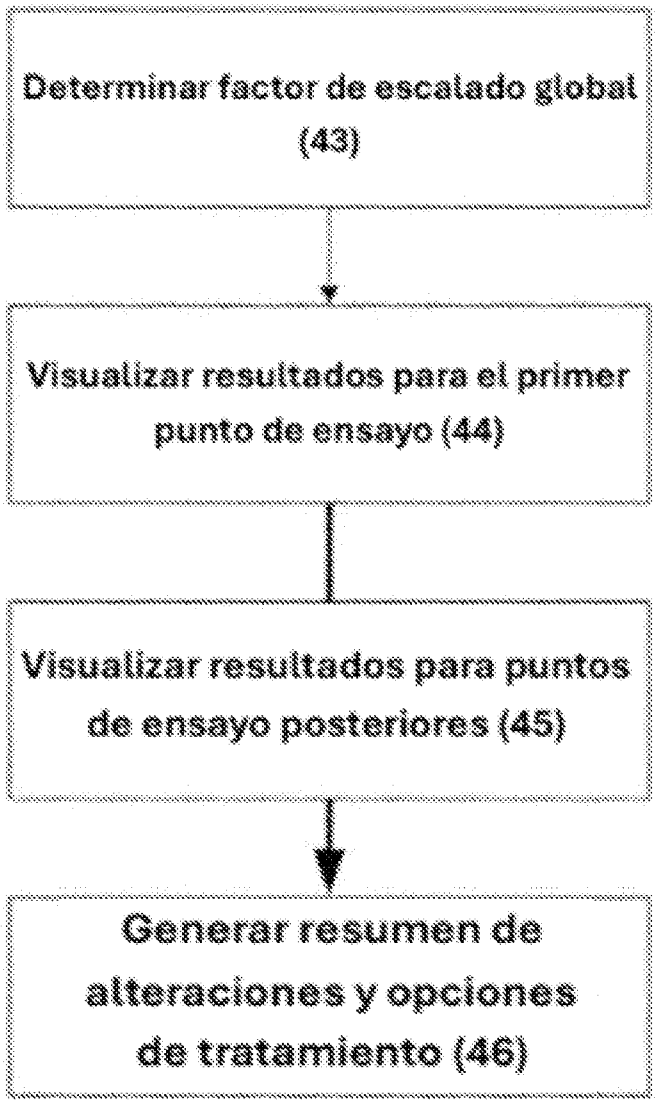

Determinar factor de escalado global (43)
Visualizar resultados para el primer punto de ensayo (44)
Visualizar resultados para puntos de ensayo posteriores (45)
Generar resumen de alteraciones y opciones de tratamiento (46)

FIG. 10B

Para todas las frecuencias alélicas mutantes informadas no de VNC (variación de número de copia)

Transformar el valor absoluto en una métrica/escala relativa que se puede trazar (p. ej., multiplicar la frecuencia alélica por 100 y restar log de ese valor)

Computar factor de escalado utilizando valor observado máximo

Utilizar conjunto de datos de ensayo más antiguo

Para cada alteración no de VNC

Multiplicar el factor de escalado por el valor transformado para cada gen y utilizarlo como indicador de cantidad al trazar esa variante

Asignar un indicador visual de color/único para alteración

Para conjunto de datos para fechas de ensayo posteriores

Si la composición de los resultados de ensayo permanece sin cambiar, continuar visualización de fechas de panel anterior en panel nuevo

Si las alteraciones permanecen iguales, pero las cantidades han cambiado

Recomputar el indicador de cantidad para trazar esa variante y retrazar todos los valores actualizados en panel(es) existente(s) y panel nuevo para la fecha de ensayo última.

Si se añaden nuevas alteraciones

Añadir las alteraciones al punto máximo de todas las alteraciones existentes

Computar valores de transformación

Recomputar factor de escalado

Redibujar el mapa de respuesta, retrazar las alteraciones en la fecha de ensayo anterior que todavía se detectan en la fecha de ensayo actual, así como alteraciones de aparición nueva

Si alteración existente anterior no está entre el conjunto de alteraciones detectadas

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

Generar resumen de alteraciones y opciones de tratamiento

Agrupar alteraciones basadas en las frecuencias alélicas mutantes máximas

FIG. 10C

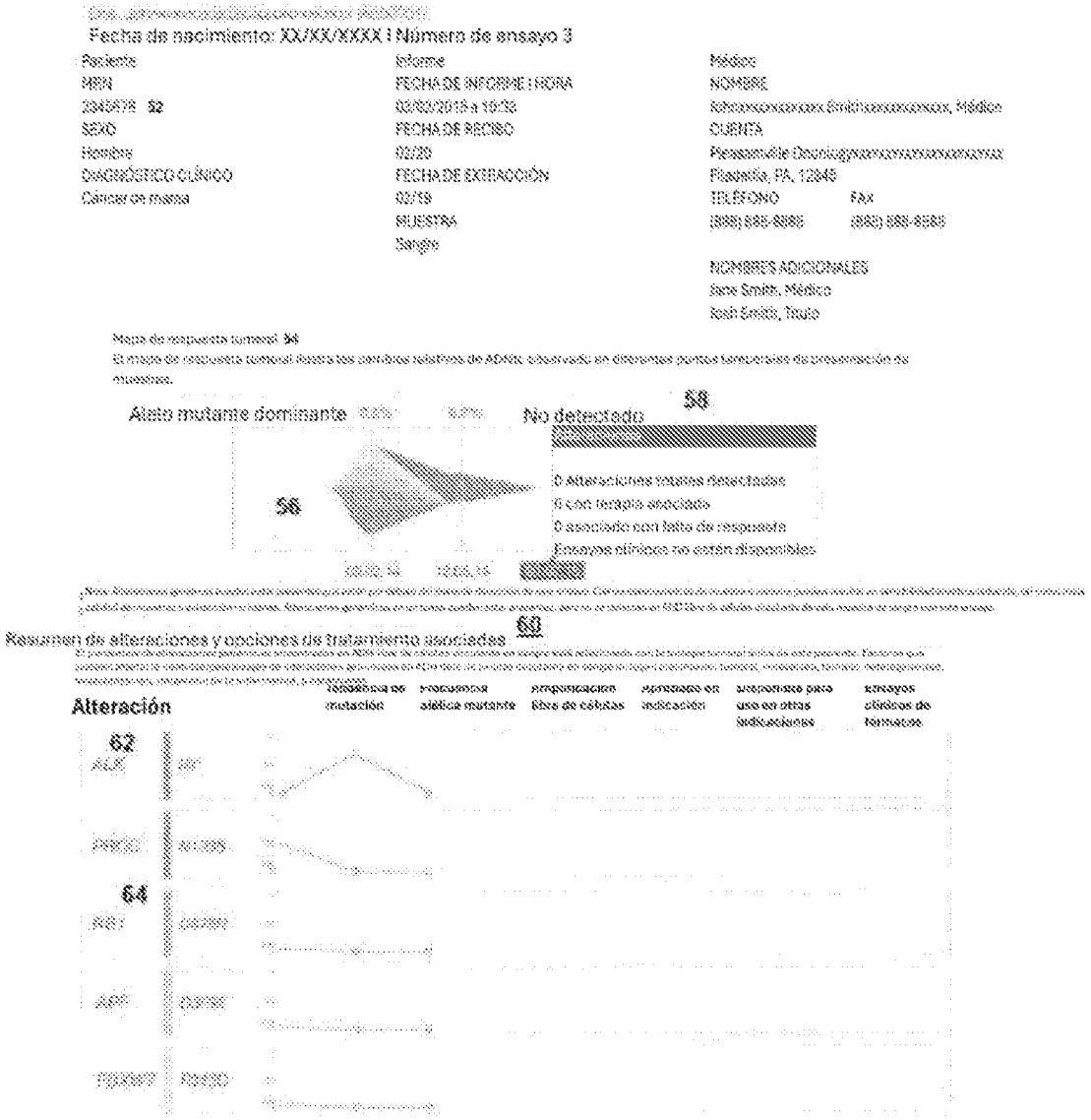

Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3
Paciente
MRN
SEXO
Hombre
DIAGNÓSTICO CLÍNICO
Cáncer de mama
Informe
FECHA DE INFORME | HORA
FECHA DE RECIBO
FECHA DE EXTRACCIÓN
MUESTRA
Sangre
Médico
NOMBRE
CUENTA
TELÉFONO
FAX
NOMBRES ADICIONALES
Jane Smith, Médico
Josh Smith, Tutor
Mapa de respuesta tumoral 66
Alelo mutante dominante
No detectado
58
56
0 Alteraciones mutantes detectadas
0 con terapia asociada
0 asociado con falta de respuesta
Ensayos clínicos no están disponibles
60
Resumen de alteraciones y opciones de tratamiento asociadas
Alteración
Frecuencia de mutación
Frecuencia alélica mutante
Amplificación libre de células
Aprobados en indicación
Disponibles para uso en otras indicaciones
Ensayos clínicos de fármacos
62
64

FIG. 10D

Dra. Johnxxxxxxxxxxxxxxxxxxxx (XXXXX)

Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3

Paciente	Informe	Médico
MRN	FECHA DE INFORME	NOMBRE
2345678 **52**	03/02/2015 a 10:33	Johnxxxxxxxxxx Smithxxxxxxxxxx, Médico
SEXO	FECHA DE RECIBO	CUENTA
Hombre	02/20	Pleasantville Oncologyxxxxxxxxxxxxxxxxxx
DIAGNÓSTICO CLÍNICO	FECHA DE EXTRACCIÓN	Filadelfia, PA, 12345
Cáncer de mama	02/19	TELÉFONO FAX
	MUESTRA	(555) 888-8888 (888) 888-8888
	Sangre	

NOMBRES ADICIONALES
Jane Smith, Médico
Josh Smith, Tirulo

FIG. 10D-1

Mapa de respuesta tumoral 54
El mapa de respuesta tumoral muestra los cambios relativos de ADNtc observado en diferentes puntos temporales de presentación de muestras.
Alelo mutante dominante
No detectado
58
56
0 Alteraciones totales detectadas
0 con terapia asociada
0 asociado con falta de respuesta
Ensayos clínicos no están disponibles
Nota: Alteraciones genéticas pueden estar presentes que están por debajo del límite de detección de este ensayo. Ciertas características de muestra o variante pueden resultar en sensibilidad analítica reducida, tal como mala calidad de muestras o extracción no idónea. Alteraciones genómicas en un tumor pueden estar presentes, pero no se detectan en ADN libre de células circulante de esta muestra de sangre con este ensayo.

FIG. 10D-2

60
Resumen de alteraciones y opciones de tratamiento asociadas
El porcentaje de alteraciones genómicas encontradas en ADN libre de células circulante en sangre está relacionado con la biología tumoral única de este paciente. Factores que pueden afectar la cantidad/porcentajes de alteraciones genómicas en ADN libre de células circulante en sangre incluyen crecimiento tumoral, renovación, tamaño, heterogeneidad, vascularización, progresión de la enfermedad, o tratamiento.
Alteración
Tendencia de mutación
Frecuencia alélica mutante
Amplificación libre de células
Aprobado en indicación
Disponible para uso en otras indicaciones
Ensayos clínicos de fármacos
62
ALK
MET
PIK3CA
AKT1/E17K
64
RB1
DAY10M
APC
G12D/E
FBXW7
R310Q

FIG. 10D-3

Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Secuenciación de tumor libre de biops

El gráfico anterior anota el porcentaje de frecuencia alélica de ADN libre de células circulante alterado (% ADNlc) detectado en este paciente. Las alteraciones genómicas detectadas se enumeran en orden descendente por % ADNlc por gen.

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

Definiciones

Número de copia: La amplificación fue detectada para este gen en el ADN libre de células circulante de este paciente. El ensayo únicamente detecta la amplificación en determinados genes con asterisco en la Tabla 3. Este ensayo no examina otras alteraciones de número de copias en otros genes.

Deleción (Del): El ensayo detecta deleciones cortas en exones 18 y 20 del gen *EGFR*. Este ensayo no examina deleciones en otros genes.

Fusión: El ensayo detecta fusiones de los genes *ALK*, *NTRK3* y *ROS1* y todos los genes compañeros. Este ensayo no examina fusiones de otros genes.

*Detecta mutaciones genómicas sinónimas que no resultan en cambios de aminoácidos.

Comentarios

EGFR-L747L: Esta alteración genómica codifica una sustitución sinónima que no resulta en un cambio en el aminoácido en esta posición. Aunque no se espera que la estructura proteínica se altere, las consecuencias funcionales e importancia clínica exacta de esta variante permanecen desconocidas. No se conoce la relevancia de terapias disponibles y ensayos clínicos que se dirigen a esta alteración.

Firmado electrónicamente

Interpretación

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

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

La amplificación fue detectada en el ADN libre de células circulante aislado de la muestra de sangre de este paciente para los gen(es) anotado(s). El ensayo únicamente detecta la amplificación en los genes con asterisco en la Tabla X por análisis de datos de secuenciación de próxima generación en el ADN libre de células circulante. Este ensayo no examina otras alteraciones de número de copias en otros genes. Este ensayo no es un sustituto para otros métodos basados en tejido establecidos que detectan amplificación genética, por ejemplo, por inmunohistoquímica (IHQ) o FISH. A diferencia de ensayos de amplificación basados en tejido (IHQ o FISH), un ensayo positivo representa la amplificación promedio para el gen interrogado en el ADN libre de células circulante presente en la muestra de sangre del paciente. Por ejemplo, un ensayo positivo podría representar una población pequeña 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 detectada. La correlación exacta entre la amplificación obtenida por el ensayo comparada a IHQ o FISH, y cómo cada ensayo guía la gestión de pacientes de manera distinta, es un área de investigación activa.

FIG. 10E

Doe, Johxxxxxxxxxxxxxxxxxxxxxxxxxx (A000001)
Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Secuenciación de tumor libre de biopsia

El gráfico anterior anota el porcentaje de frecuencia alélica de ADN libre de células circulante alterado (% ADNlc) detectado en este paciente. Las alteraciones genómicas detectadas se enumeran en orden descendente por % ADNlc por gen.

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

Definiciones

Número de copia: La amplificación fue detectada para este gen en el ADN libre de células circulante de este paciente. El ensayo únicamente detecta la amplificación en determinados genes con asterisco en la Tabla 3. Este ensayo no examina otras alteraciones de número de copia en otros genes.

Deleción (Del): El ensayo detecta deleciones cortas en exones 19 y 20 del gen *EGFR*. Este ensayo no examina deleciones en otros genes.

Fusión: El ensayo detecta fusiones de los genes *ALK*, *NTRK1*, *RET* y *ROS1* y todos los genes compañeros. Este ensayo no examina fusiones de otros genes.

*Denota mutaciones genómicas sinónimas que no resultan en cambios de aminoácidos.

FIG. 10E-1

Comentarios

EGFR-L747L: Esta alteración genómica codifica una sustitución sinónima que no resulta en un cambio en el aminoácido en esta posición. Aunque no se espera que la estructura proteínica se altere, las consecuencias funcionales e importancia clínica exacta de esta variante permanecen desconocidas. No se conoce la relevancia de terapias disponibles y ensayos clínicos que se dirigen a esta alteración.

Firmado electrónicamente
Arthur Baca, MD PhD, 01/27/2015 14:53

Interpretación

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

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

La amplificación fue detectada en el ADN libre de células circulante aislado de la muestra de sangre de este paciente para los gen(es) anotado(s). El ensayo únicamente detecta la amplificación en los genes con asterisco en la Tabla X por análisis de datos de secuenciación de próxima generación en el ADN libre de células circulante. Este ensayo no examina otras alteraciones de número de copias en otros genes. Este ensayo no es un sustituto para otros métodos basados en tejido establecidos que detectan amplificación genética, por ejemplo, por inmunohistoquímica (IHQ) o FISH. A diferencia de ensayos de amplificación basados en tejido (IHQ o FISH), un ensayo positivo representa la amplificación promedio para el gen interrogado en el ADN libre de células circulante presente en la muestra de sangre del paciente. Por ejemplo, un ensayo positivo podría representar una población pequeña 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 detectada. La correlación exacta entre la amplificación obtenida por el ensayo comparada a IHQ o FISH, y como cada ensayo guía la gestión de pacientes de manera distinta, es un área de investigación activa.

FIG. 10F-2

Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Resultados de terapia en detalle

Alteración	Fármaco	Nombre comercial	Objetivo	Estado actual
Amplificación EGFR T790M	Necitumab		Inhibidor EGFR/Her2/ErbB4	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Carcinoma de mama)
	CO-1686		Inhibidor EGFR T790M	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Cáncer pulmonar)
	Panitumumab	Vectibix	Anticuerpo antiproteína EGFR	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma colorrectal)
	AZD9291		Inhibidor EGFR T790M	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Cáncer pulmonar)
	Afatinib	Gilotrif	Inhibidor de quinasa pan-ErbB irreversible	Aprobado por FDA en esta indicación (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Cáncer pulmonar)
	Lapatinib	Tykerb	Anticuerpo inhibidor EGFR	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma de mama)
	Dacomitinib		Inhibidor EGFR T790M	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Cáncer pulmonar)
	Nimotuzumab	Theraloc	Anticuerpo inhibidor EGFR	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Carcinoma glioma, Astrocitoma de astrocitoma, Glioblastoma, Carcinoma de células escamosas de cabeza y cuello (CCECC), Carcinoma de colangiocarcinoma, Carcinoma gastrointestinal, Carcinoma de células escamosas de la lengua)
	Cetuximab	Erbitux	Anticuerpo inhibidor EGFR	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma de células escamosas de cabeza y cuello), Carcinoma de células escamosas de la lengua, Carcinoma colorrectal)
Amplificación MET	Crizotinib	Xalkori	Inhibidor de quinasa ALK/Met	Aprobado por FDA en esta indicación (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Carcinoma glioma, Carcinoma de células renal)
	EXEL-2880 (XL-880)		Inhibidor de multiquinasa que se dirige a Met, Ros1, Axl, Ret3 y otras proteínas	Fase 1 (Tumor sólido)
	SGX-523		Inhibidor de quinasa de molécula pequeña Met/Ret/VEGFR1,2,3/Flt3/Tie-2/Trk/Axl	Fase 2 (Carcinoma glioma, Carcinoma de próstata, Carcinoma de células renal, Carcinoma de células escamosas de cabeza y cuello (CCECC), Carcinoma de células escamosas cutáneo, Carcinoma colorrectal)
	Cabozantinib	Cometriq	Inhibidor de quinasa de molécula pequeña Met/Ret/VEGFR1,2,3/Kit2/Tie-2/Trk/Axl	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma medular tiroideo)
	Foretinib		Inhibidor de quinasa de molécula pequeña Met/VEGFR2	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Carcinoma glioma, Carcinoma hepatocelular, Carcinoma de células renales, Carcinoma de células escamosas de cabeza y cuello)

Continúa en la siguiente página

FIG. 10F

Resultados de terapia en detalle

Alteración	Fármaco	Nombre comercial	Objetivo	Estado actual
Amplificación EGFR	Neratinib		Inhibidor Egfr/Her2/Erbb4	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Carcinoma de mama)
T790M	CO-1686		Inhibidor EGFR T790M	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Cáncer pulmonar)
	Panitumumab	Vectibix	Anticuerpo inhibitorio EGFR	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma colorrectal)
	AZD9291		Inhibidor EGFR T790M	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Cáncer pulmonar)
	Afatinib	Gilotrif	Inhibidor de quinasa pan-ErbB irreversible	Aprobado por FDA en esta indicación (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Cáncer pulmonar)
	Lapatinib	Tykerb	Anticuerpo inhibitorio EGFR	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma de mama)
	Dacomitinib		Inhibidor EGFR T790M	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Cáncer pulmonar)
	Nimotuzumab	Theraloc	Anticuerpo inhibitorio EGFR	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 3 (Carcinoma gástrico, Adenocarcinoma del esófago, Glioblastoma, Carcinoma de células escamosas de cabeza y cuello (SCCHN), Carcinoma de unión gastroesofágica, Carcinoma gastroesofágico, Carcinoma de célula escamosa de la lengua)
	Cetuximab	Erbitux	Anticuerpo inhibitorio EGFR	Fase 3 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma de célula escamosa de cabeza y cuello), Carcinoma de célula escamosa de la lengua, Carcinoma colorrectal

FIG. 10F-1

Amplificación MET	Crizotinib	Xalkori	Inhibidor de quinasa ALK/Met	Aprobado por FDA en esta indicación (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Carcinoma gástrico, Carcinoma de célula renal)
	LY2801653		Inhibidor de multiquinasa que se dirige a Met, Ros 1, Axl, Rt3 y otras proteínas	Fase 1 (Tumor sólido)
	BMS-777607		Inhibidor de quinasa de molécula pequeña Met/Ret/VEGFR1,2,3/Rt3/ Tie-2/Tyro/Axl	Fase 2 (Carcinoma gástrico, Carcinoma de próstata, Carcinoma de célula renal, Carcinoma de célula escamosa de cabeza y cuello (CCECC), Carcinoma de célula escamosa cutánea, Carcinoma esofágico)
	Cabozantinib	Cometriq	Inhibidor de quinasa de molécula pequeña Met/Ret/VEGFR1,2,3/Rt3/ Tie-2/Tyro/Axl	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Aprobado por FDA en otras indicaciones (Carcinoma medular tiroideo)
	foretinib		Inhibidor de quinasa de molécula pequeña Met/VEGFR2	Fase 2 (Carcinoma pulmonar de células no pequeñas (CPCNP)) Fase 2 (Carcinoma gástrico, Carcinoma hepatocelular, Carcinoma de células renales, Carcinoma de célula escamosa de cabeza y cuello)

Continuado en la siguiente página

FIG. 10F-2

Sexo, [illegible] (XXYXXY)
Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Relevancia clínica de alteraciones detectadas

Alteración	Papel en enfermedad	Efecto en sensibilidad al fármaco	Efecto en resistencia al fármaco
Amplificación EGFR 7792M		16784701, Yan Cataxon et al., 2007, 17479883), Basado en recomendación consistente, la American Society for Clinical Oncology (ASCO) ha publicado un Documento Clínico Provisional recomendando análisis mutacional de EGFR para pacientes con CPCNP para considerar un beneficio de TKI de Egfr (Keedy et al., 2011, 21482982).	
Amplificación MET	Activación de proteína Met e sobreexpresión promueve angiogénesis, resistencia a apoptosis, proliferación e invasión de células cancerosas (Acquaviva, 2011; 23042886, Jung et al., 2012; 22370863; Gherardi et al., 2012; 22270863; Takeuchi et al., 2002; 12084429). La expresión de proteína Met en CPCNP se ha asociado con una predisposición al desarrollo de metástasis cerebral (Benestetinti et al., 2010; 21083159).	Pacientes de NSCLC cancerosa, positivamente correlacionado de mutaciones amplificación de MET puede causar activación de MET oncogénica y promover sensibilidad a inhibidores de PSC (Cancer et al., 2010; 20097187). Clínicamente y radiográficamente respondedores epidermoides, incluyendo Met. Crizotinib ha sido aprobado por la FDA para cáncer de pulmón de células no pequeñas positivo por ALK-R, y actualmente se está aprobando por la FDA para el cáncer de pulmón [Camidge et al., 2011; ASCO 2011, Abstract 7521, Bang et al., 2010; ASCO 2010, Abstract 3, Pfizer, 2010; [illegible], Iafrate y Bang, 2011; [illegible], información adicional se puede registrar que la amplificación de MET puede producir sensibilidad a crizotinib (Camidge et al., 2011; [illegible], Cho et al., 2012, 22760202, Cho et al., 2011, 21913000, Inhibidores específicos a Met como el desarrollo clínico, incluyendo Tivantinib (ARQ 197), un inhibidor MET no competitivo por ATP selectivo, y onartuzumab un anticuerpo monovalente que se dirige a MET (Rajkumar et al., 2011; 21913444, Basilico et al., 2013; 23307437; Eng et al., 2011; ASCO 2011, Abstract 7508).	Amplificación de Met o expresión elevada de Met se ha implicado en resistencia adquirida a inhibidores Egfr en algunos tipos de cáncer; actualmente se está investigando terapia de combinación con inhibidores Egfr en esta configuración [Pao/Miller, et al., 2011; 17803386, Chan et al., 2010; 20527157]; Activación Met se ha visto implicado como un mecanismo clave de resistencia a terapia dirigida a Egfr en CPCNP [Engelman, et al., 2007, 17463250]; La et al., 2013; 24167838, Bachleitner, et al.,
NOTCH1 or PSEN1	Dependiendo del contexto celular NOTCH1 puede operar como supresor tumoral o oncogen (Wang et al., 2011, 22086338, Klinakis et al., 2011, 21562484). Actividad de NOTCH1 ha sido informada a requerirse para tumorigénesis pulmonar en modelos CLCNP impulsados por K-Ras, y ha sido enlazado a un fenotipo de célula madre (Lobry et al., 2013, 23943799, Baumgart, et al., 2013; 24569976, Hassan et al., 2013; 23446212, Allen et al., 2011, 21803744).	Activos mutaciones NOTCH1 contribuyen la proteína transmembrana Notch1 y forman o reclutamiento Notch1 aumentando Wang et al., 2009; 19470835), incluyendo Notch1 los casos, translocaciones genómicas aberrantes que implican receptor o dominios intracelulares truncados que carecen de secuencias PEST [illegible] y como tales en los ensayos clínicos para varios cánceres (Radtke et al., 2011; 21430364, Grego y Fenton, 2012; 22456420). Inhibidores de Gamma secretasa se han investigado ampliamente en modelos tumorales preclínicos con actividades específicas (Knoechel et al., 2006; 16677779, Arcaroli et al., 2011; 20980566). Algunos estudios sugieren que las mutaciones NOTCH1, independiente Lealidad y dosis relativas, pueden relacionar la sensibilidad a vía Notch1 en algunos tumores que han perdido supresores Notch (Egloff y Grandis, 2012; 22773430, Rizzo et al., 2008; 18579808). Sin embargo, esta enfoque es relevante en el desarrollo de [illegible] epigenéticos y no se sugiere que se relacionen en el caso de mutación relacionada de NOTCH1. En el caso de esta variante no caracterizada, la relevancia de enfoques terapéuticos disponibles no es conocida.	En un estudio preclínico utilizando líneas celulares CPCNP y xenoinjertos, se ha sugerido que Notch1-2 juega un papel en mecanismos adquiridos a inhibidores de quinasas de Receptor Egfr, tal como gefitinib (Xie et al., 2010; 20971898, Xie et al., 2012; 22713998-81. Además, se ha informado que los tumores exitosos CPCNP con Notch1 alto o activado son resistentes a quimio-terapias, radioterapias convencionales y [illegible] (Sanchez et al., 2013; 23584621). Lee et al., 2013; 23563268

FIG. 10G

Relevancia clínica de alteraciones detectadas

Alteración	Papel en enfermedad	Efecto en sensibilidad al fármaco	Efecto en resistencia al fármaco
Amplificación EGFR T790M		18784107, Van Cutsem et al., 2007, 17878356). Basado en evidencia acumulada, la American Society for Clinical Oncology (ASCO) ha publicado un Consenso Clínico Provisional recomendando análisis mutacional de EGFR para pacientes con CPCNP para predecir un beneficio de TKI de Egfr (Keedy et al., 2011, 21482982)	
Amplificación MET	Activación de proteína Met o sobreexpresión promueve angiogenesis, resistencia a apoptosis, proliferación e invasión de células cancerosas (Appleman, 2011; 22542856, Kong et al., 2012; 22553081, Cleveland et al., 2012; 22278538, Tolcher et al., 2012; 19584328). La expresión de proteína Met en CPCNP se ha asociado con una predisposición al desarrollo de metastasis cerebral (Benedettini et al., 2010; 20488158).	Expresión de MET aumentada genéticamente como resultado de mutación o amplificación de MET puede causar activación de MET mejorada y potencial sensibilidad a inhibidores de MET (Cecchi et al., 2010; 20302741). Crizotinib y cabozantinib se dirigen múltiples quinasas, incluyendo Met. Crizotinib ha sido aprobado por la FDA para uso en cáncer pulmonar de células no pequeñas positivo por EML4-ALK, y cabozantinib ha sido aprobado por la FDA para el cáncer tiroideo medular (Camidge et al., 2011; ASCO 2011, Abstract 2001, Sang et al., 2010; ASCO 2010, Abstract 3, Treyns, 2013; 23292237, Hunt y De Boer, 2013; 23318867). Informes sobre múltiples tipos de tumor sugieren que la amplificación de MET puede predecir sensibilidad a concomitante (Lennerz et al., 2011; 22042394; Cihi et al., 2012; 22162373; Ou et al., 2011; 21822805). Inhibidores específicos a Met están en desarrollo clínico, incluyendo tivantinib (ARQ 197), un inhibidor MET no competitivo por ATP selectivo, y onartuzumab un anticuerpo monoclonal que se dirige a MET (Adjei et al., 2011; 21632448; Benvenuti et al., 2010; 20810077; Eng et al., 2013; ASCO 2013, Abstract 3008).	Amplificación de Met o expresión elevada de Met se ha implicado en resistencia adquirida a inhibidores Egfr en algunos tipos de cáncer, actualmente se está investigando terapia de combinación con inhibidores Egfr en esta configuración (Bean Bosch, et al., 2011; 21460356; Chan et al., 2013; 23537287). Activación Met se ha visto implicado como un mecanismo clave de resistencia a terapia dirigia a Egfr en CPCNP (Engelman, et al., 2007; 17463250; La et al., 2013; 24107834; Benedettini, et al.;

FIG. 10G-1

[illegible]	Dependiendo del contexto celular NOTCH1 puede operar como supresor tumoral o oncogén (Weng et al, 2011, 22086338, Klinakis et al, 2011, 21685884). Actividad de NOTCH1 ha sido informada a requerirse para tumorigenesis pulmonar en modelos CPCNP impulsados por K-Ras, y ha sido enlazado a un fenotipo de célula madre (Licciulli et al., 2013, 23643388, Baumgart, et al., 2012, 24509876, Hassan et al., 2013; 23444212, Allen et al., 2011, 21602744).	Activar mutaciones NOTCH1 estratificas la gestión intracelular Notch1 y llevan a señalización Notch1 aumentada (Weng et al., 2004, 15272075). Inhibidores Notch1 (anticuerpos, pequeñas moléculas gamma-secretasa, que impiden escisión o dominio extracelular) pueden ser un enfoque terapéutico potencial en el caso de mutaciones de activación NOTCH1, y estos están en los ensayos clínicos para varios cánceres (Roti et al., 2011; 21825384, Groth y Fortini, 2012; 22389842). Inhibidores de Gamma-secretasa se han mostrado eficaces en modelos tumorales preclínicos con activación Notch (Weerkamp et al., 2006, 16501775, Arcaroli et al., 2013, 23588309). Algunos estudios sugieren que los inhibidores HDAC, incluyendo LBH589 y ácido valproico, pueden mediante la señalización de vía Notch en algunos tumores que han perdido expresión Notch (Egloff y Grandis, 2012, 22773520, Sierra et al., 2006, 16675626). Sin embargo, este enfoque es relevante en el contexto de alteraciones epigenéticas y no se espera que es relevante en el caso de mutación inactivadora de NOTCH1. En el caso de esta variante no caracterizada, la relevancia de enfoques terapéuticos disponibles no se conoce.	En un estudio preclínico utilizando líneas celulares CPCNP y xenoinjertos, se ha sugerido que Notch1 juega un papel en resistencia adquirida a inhibidores de quinasa de tirosina Egfr, tal como gefitinib (Xie et al., 2013; 23818813, Xie et al., 2012; 22173854). Además, se ha informado que las líneas celulares CPCNP con Notch1 silio o activado son resistentes a químio-terapias, incluyendo cisplatino y docetaxel (Hassan et al., 2013; 23444212).

FIG. 10G-2

[redacted] (ACRONIMO 1)
Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Ensayos clínicos disponibles

Alteración	ID Ensayo	Título	Fase	Sitio
Amplificación EGFR T790M	NCT01647711	Un estudio de Afatinib intermitente de dosis alta para determinar la dosis tolerada máxima y evaluar actividad de esta dosis contra cáncer pulmonar de células no pequeñas con mutaciones T790M	Fase 1	Aurora, CO; Boston, MA (2)
	NCT01526928	Estudio para evaluar seguridad, farmacocinética y eficacia de Rociletinib (CO-1686) en receptor de factor de crecimiento epidérmico mutante tratado previamente (EGFR) en pacientes con cáncer pulmonar de células no pequeñas (CPCNP)	Fase 1/Fase 2	Aurora, CO; Fountain Valley, CA; Los Angeles, CA (2); Orange, CA; Sacramento, CA; Santa Monica, CA; Stanford, CA; Whittier, CA (2); Aurora, CO; Washington, DC; Miami, FL; Sarasota, FL; Atlanta, GA; Boston, MA (2); Ann Arbor, MI; Detroit, MI; Minneapolis, MN; Buffalo, NY; New Hyde Park, NY; Nueva York, NY; Cincinnati, OH; Columbus, OH; Toledo, OH; Portland, OR; Filadelfia, PA; Pittsburgh, PA; Memphis, TN; Houston, TX
Amplificación MET	NCT01324479	Estudio de INC280 en pacientes con tumores sólidos avanzados dependientes c-MET	Fase 1	Filadelfia, PA; Los Angeles, CA; Chicago, IL; Rochester, MN; Detroit, MI; Buffalo, NY; Nashville, TN; Dallas, TX; Houston, TX; Salt Lake City, UT; Australia (2), Canadá (2), Francia (3), Alemania (2), Hong Kong (2), Israel (4), Italia (6), Corea (4), Países Bajos (3), Noruega, Singapur, Granada, España (3), Taiwán (2), Tailandia (3)
	NCT01712217	Un estudio de AP13287 en pacientes con cáncer pulmonar de células no pequeñas (CPCNP) solo y en combinación con Crizotinib	Fase 1/Fase 2	Scottsdale, AZ; La Jolla, CA; Los Angeles, CA; San Diego, CA; Santa Monica, CA; Whittier, CA; Aurora, CO; New Haven, CT; Detroit, MI; Orlando, FL; Tampa, FL; Chicago, IL (2); Ann Arbor, MI; Detroit, MI; Rochester, MN; St. Louis, MO; Omaha, NE; Lebanon, NH; Reno, NV; 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; Houston, TX; Fairfax, VA; Seattle, WA (2); Madison, WI; Canadá (5); Francia (10); Corea (6); España (5)

FIG. 10H

Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Métodos y limitaciones

Genes en el panel y LOD asociado

Genes con cobertura de exón completa y de intrón parcial

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

*Indels +EGFR en exón 19 y 20

Genes con cobertura de exón crítica

Incluye región promotora

Reestructuraciones

FIG. 101

Dra. Johanna xxxxxxxxxxxxxxxx (MD001)
Fecha de nacimiento: XX/XX/XXXX | Número de ensayo 3 | Estado: CORREGIDO

Métodos y limitaciones

La prueba ensaya un panel de 68 genes para identificar alteraciones genómicas en variantes somáticas asociadas al cáncer con alta sensibilidad. ADN libre de células se extrae del plasma y alteraciones genómicas se analizan por secuenciación paralela masiva de genes objetivo amplificados. Para un conjunto de genes, todos los exones se secuencian al haberse mostrado que tal cobertura de secuenciación tiene utilidad clínica. Para otro conjunto de genes, la cobertura de secuenciación incluye todos aquellos exones con una mutación somática informada previamente. Las sustituciones sinónimas se informan si presentes, aunque la exacta importancia clínica y relevancia de terapias disponibles que se dirigen a tales variantes no se definen precisamente. El alelo mutante detectable mínimo (límite de detección) es dependiente de la concentración de ADN libre de células de muestra de paciente, que puede variar de menos de 10 a más de 1000 equivalentes genómicos por mL de sangre periférica. La prueba se valida para solo detectar amplificaciones de genes en los genes con asterisco en Tabla X en muestras con tumor de al menos 2 % derivado de ADN libre de células circulante que contienen 8-10x amplificación de genes; amplificación que puede no detectarse en muestras con cantidades más bajas de ADN libre de células y/o amplificación de copia de genes de bajo nivel. Los tipos de alteraciones genómicas que pueden detectarse por la prueba incluyen variaciones de nucleotido único, amplificaciones, fusiones ALK, NTRK1, RET y ROS1 e inserciones/deleciones cortas presentes en exones 19 y 20 del gen EGFR. Esta versión de la prueba no se valida para la detección de otros tipos de alteraciones genómicas (por ejemplo, deleciones de genes, cambios de marco, variantes de empalme o indeles). Ciertas características de muestra o variante pueden resultar en sensibilidad analítica reducida, tal como baja calidad de las muestras o extracción no correcta.

FIG. 10I-1

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

Genes con cobertura de exón completa y de intrón parcial

GEN	LOD	GEN	LOD	GEN	LOD	GEN	LOD
APC	0.5%	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%	EGFA*	0.1%	ERBB2	<0.1%	FGFR1	0.1%
FGFR2	0.5%	HRAS	0.1%	KIT	0.1%	KRAS	0.2%
MET	0.5%	MYC	<0.1%	NF1	0.2%	NRAS	0.2%
PDGFRA	0.5%	PIK3CA	0.2%	PTEN	0.2%	RAF1	0.1%
TP53	0.2%						

Genes en negrita son aquellos que también se analizan para variaciones de número de copias (VNC).
*Indeles +EGFR en exón 18 y 20

Genes con cobertura de exón crítica

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	0.5%	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%	RHOB	0.2%	SMCA	0.1%	RB1	0.1%
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. 101-2

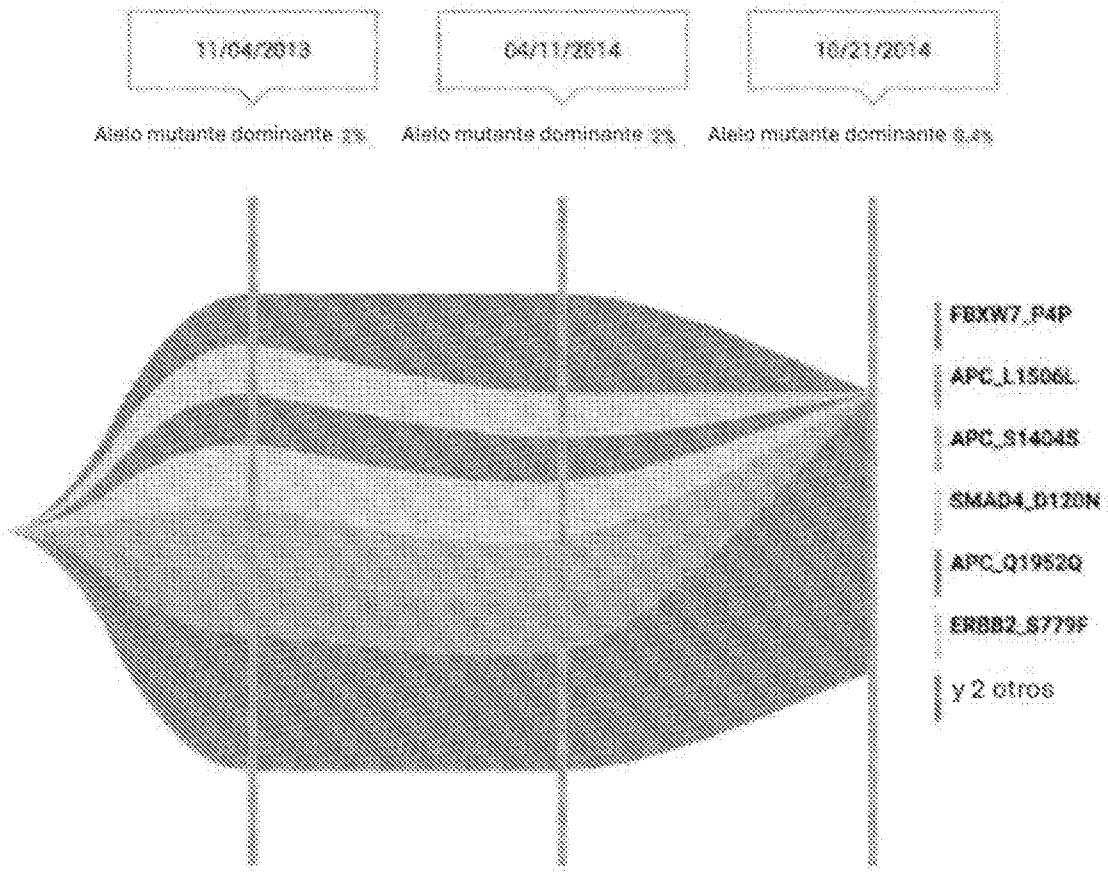

11/04/2013
04/11/2014
10/21/2014
Alelo mutante dominante 3%
Alelo mutante dominante 2%
Alelo mutante dominante 8.4%
FBXW7_P4P
APC_L1508L
APC_S1404S
SMAD4_D132N
APC_Q1952Q
ERBB2_S779F
y 2 otros

FIG. 10J

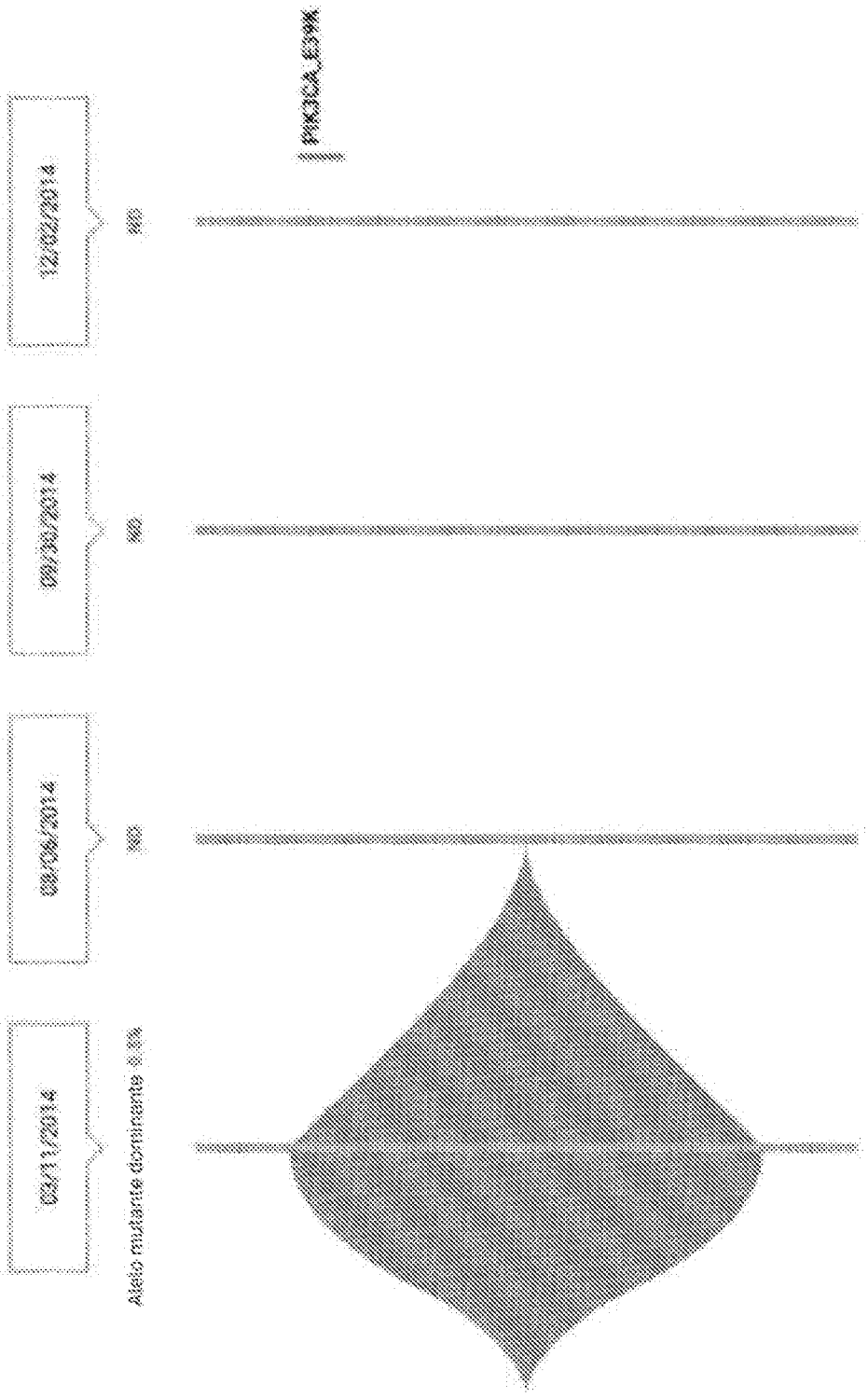

PIK3CA_E39K
12/02/2014
#0
09/30/2014
#0
08/06/2014
#0
03/11/2014
Alelo mutante dominante 0.3%

FIG. 10K

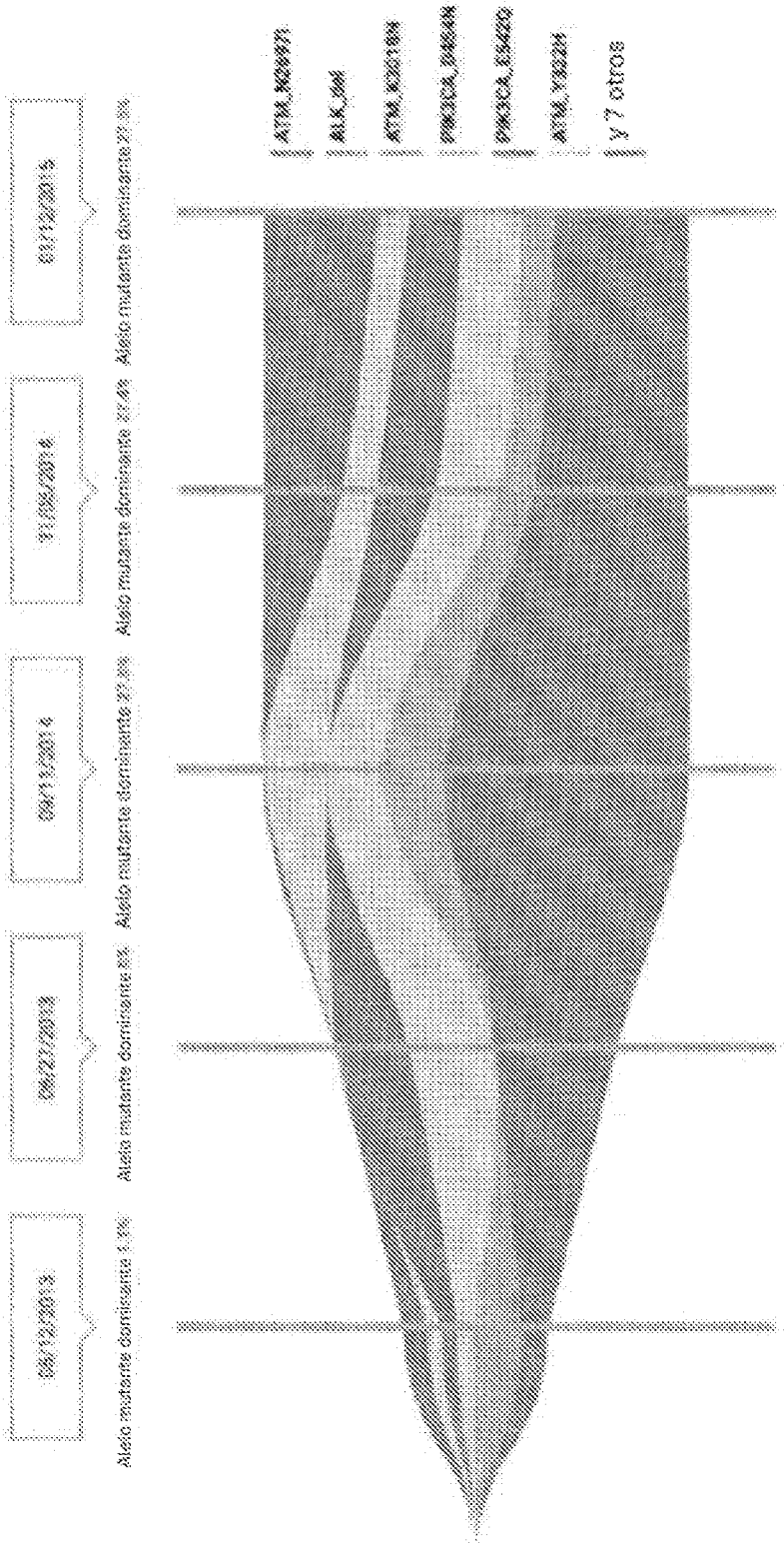

05/12/2013
06/27/2013
09/11/2014
31/05/2014
01/12/2015
Alelo mutante dominante 4.3%
Alelo mutante dominante 8%
Alelo mutante dominante 27.8%
Alelo mutante dominante 27.4%
Alelo mutante dominante 23.3%
ATM_M2691T
ALK_I3M
ATM_K2019R
PIK3CA_E545K
PIK3CA_E542K
ATM_Y1023H
y 7 otros

FIG. 10L

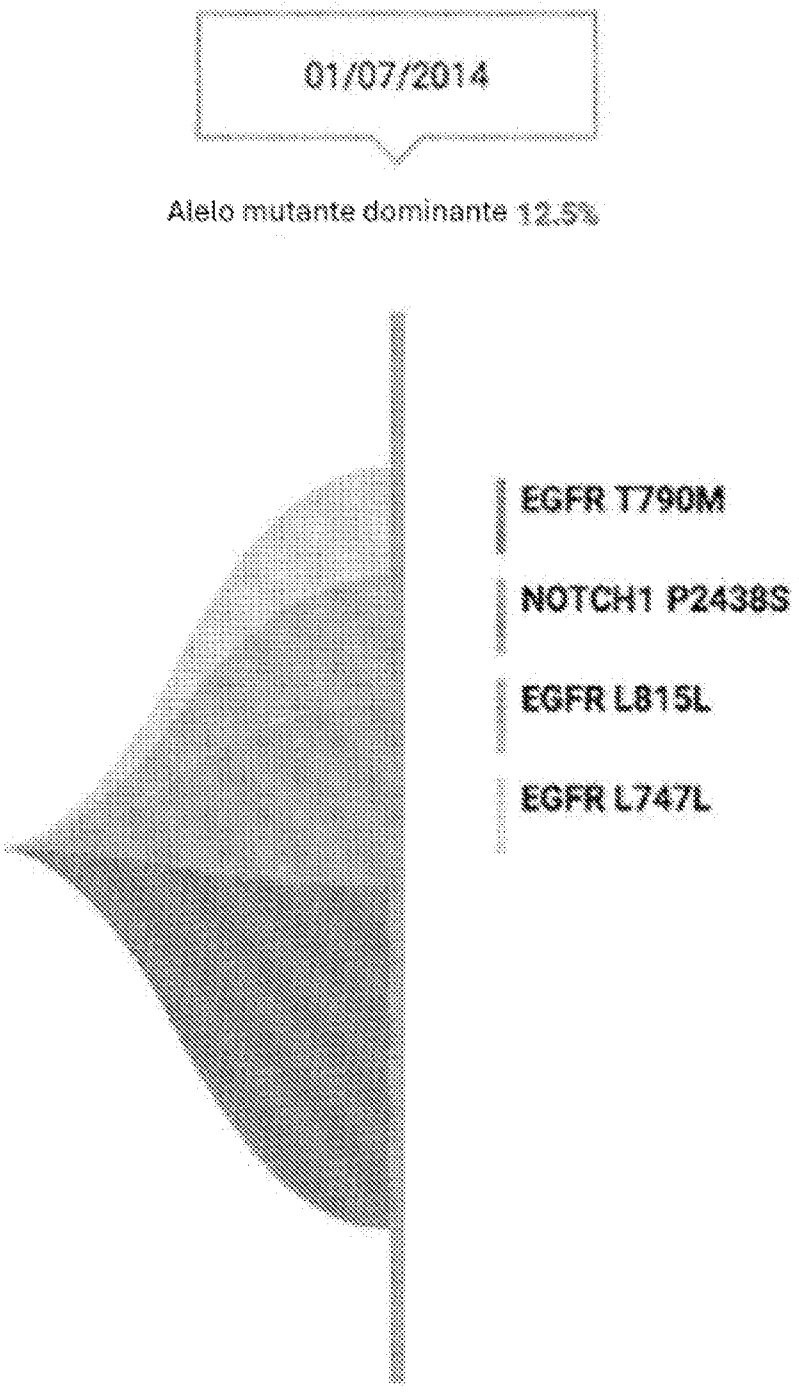

01/07/2014
Alelo mutante dominante 12,5%
EGFR T790M
NOTCH1 P2438S
EGFR L815L
EGFR L747L

FIG. 10M

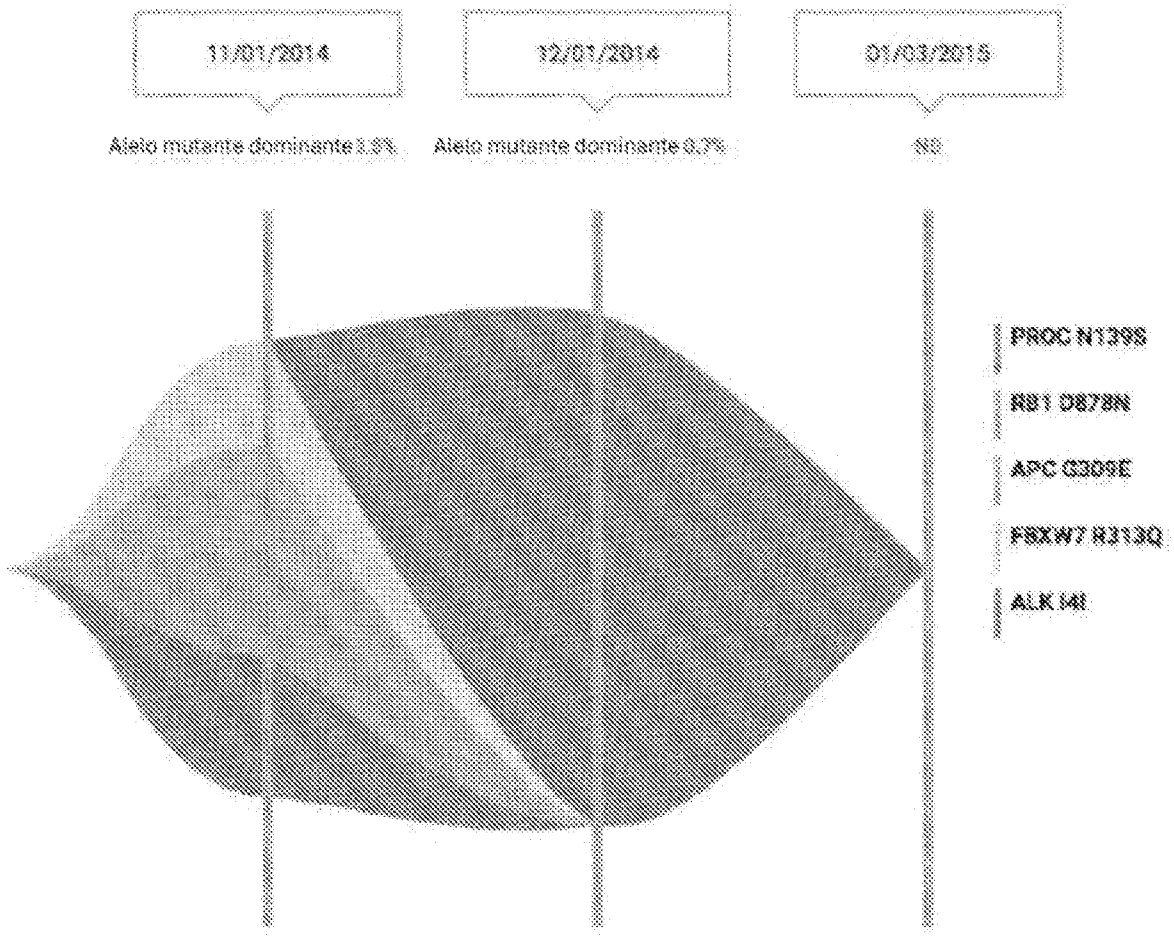

11/01/2014
12/01/2014
01/03/2015
Alelo mutante dominante 1.8%
Alelo mutante dominante 0.7%
ND
PIK3CA N1399
RB1 D878N
APC G309E
FBXW7 R313Q
ALK I4I

FIG. 10N

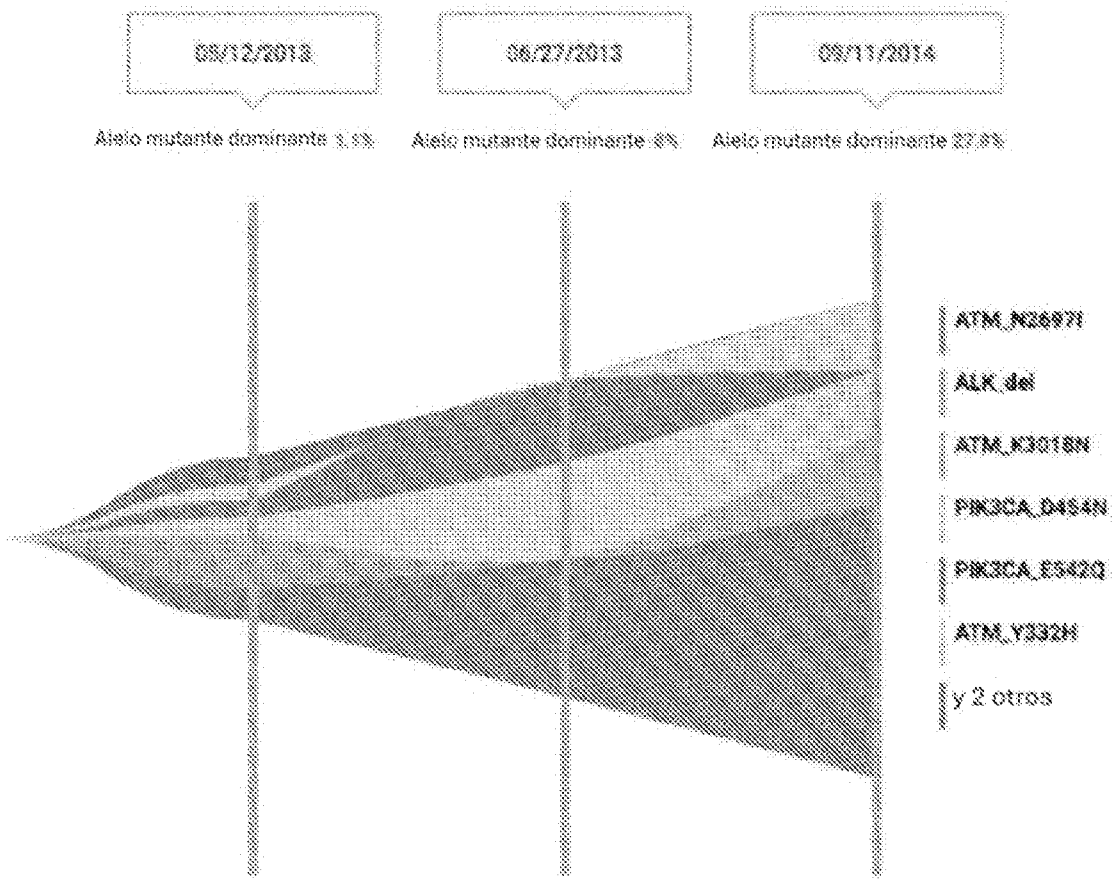

05/12/2013
06/27/2013
09/11/2014
Alelo mutante dominante 1.1%
Alelo mutante dominante 6%
Alelo mutante dominante 27.8%
ATM_N2697I
ALK_del
ATM_K3018N
PIK3CA_D454N
PIK3CA_E542Q
ATM_Y332H
y 2 otros

FIG. 10O

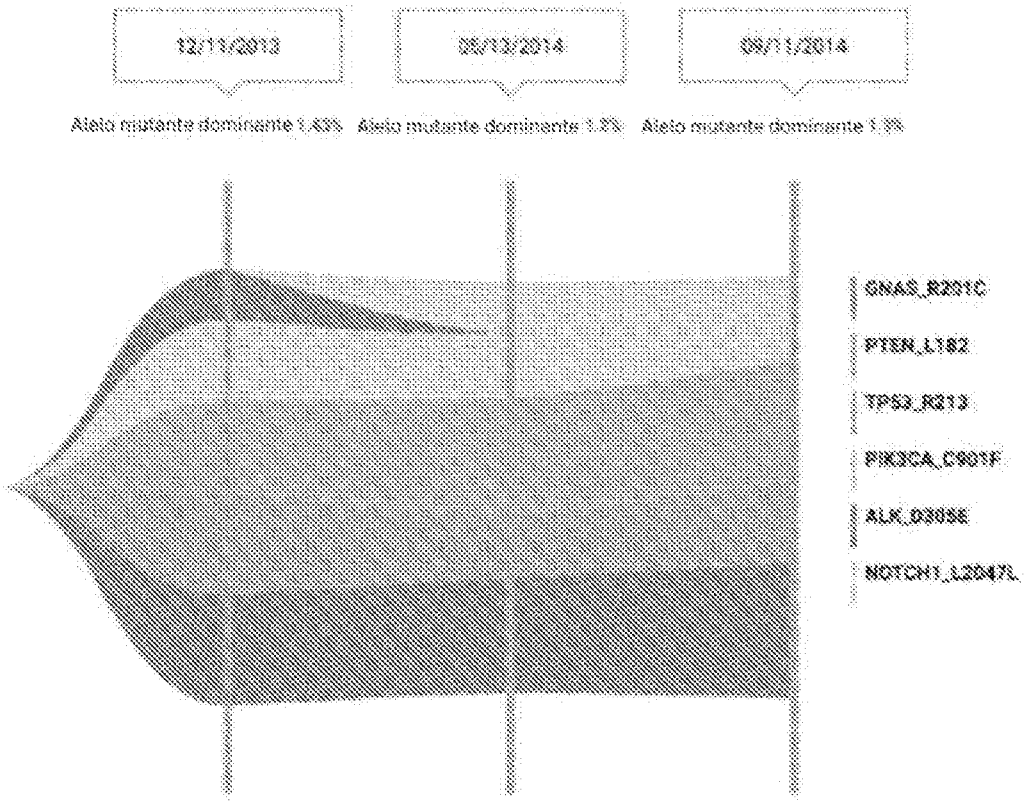

12/11/2013
05/13/2014
09/11/2014
Alelo mutante dominante 1.43%
Alelo mutante dominante 1.2%
Alelo mutante dominante 1.9%
GNAS_R201C
PTEN_L182
TP53_R213
PIK3CA_C901F
ALK_D305E
NOTCH1_L2047L

FIG. 10P

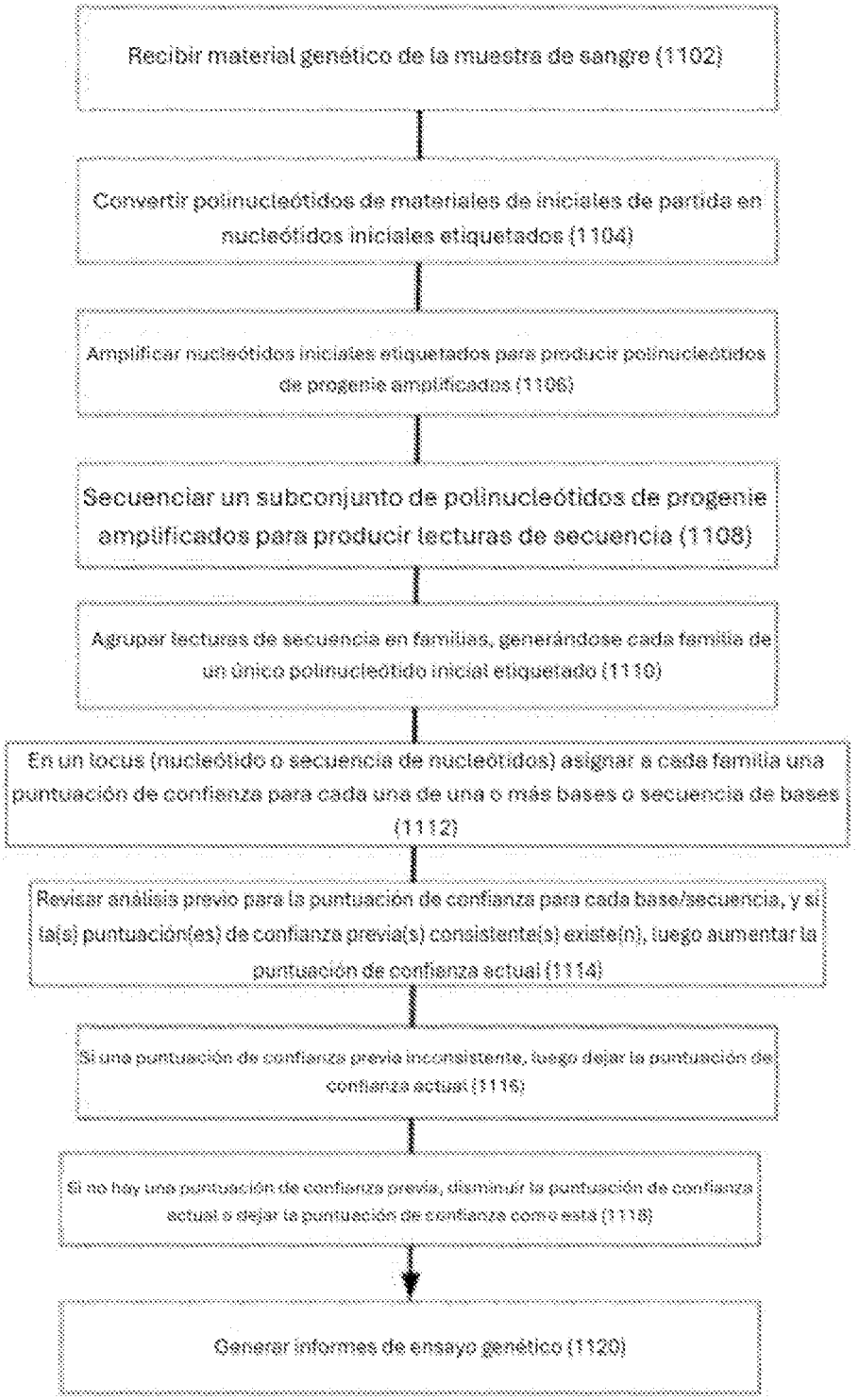

Recibir material genético de la muestra de sangre (1102)
Convertir polinucleótidos de materiales de iniciales de partida en nucleótidos iniciales etiquetados (1104)
Amplificar nucleótidos iniciales etiquetados para producir polinucleótidos de progenie amplificados (1106)
Secuenciar un subconjunto de polinucleótidos de progenie amplificados para producir lecturas de secuencia (1108)
Agrupar lecturas de secuencia en familias, generándose cada familia de un único polinucleótido inicial etiquetado (1110)
En un locus (nucleótido o secuencia de nucleótidos) asignar a cada familia una puntuación de confianza para cada una de una o más bases o secuencia de bases (1112)
Revisar análisis previo para la puntuación de confianza para cada base/secuencia, y si la(s) puntuación(es) de confianza previa(s) consistente(s) existe(n), luego aumentar la puntuación de confianza actual (1114)
Si una puntuación de confianza previa inconsistente, luego dejar la puntuación de confianza actual (1116)
Si no hay una puntuación de confianza previa, disminuir la puntuación de confianza actual o dejar la puntuación de confianza como esta (1118)
Generar informes de ensayo genético (1120)

FIG. 11A

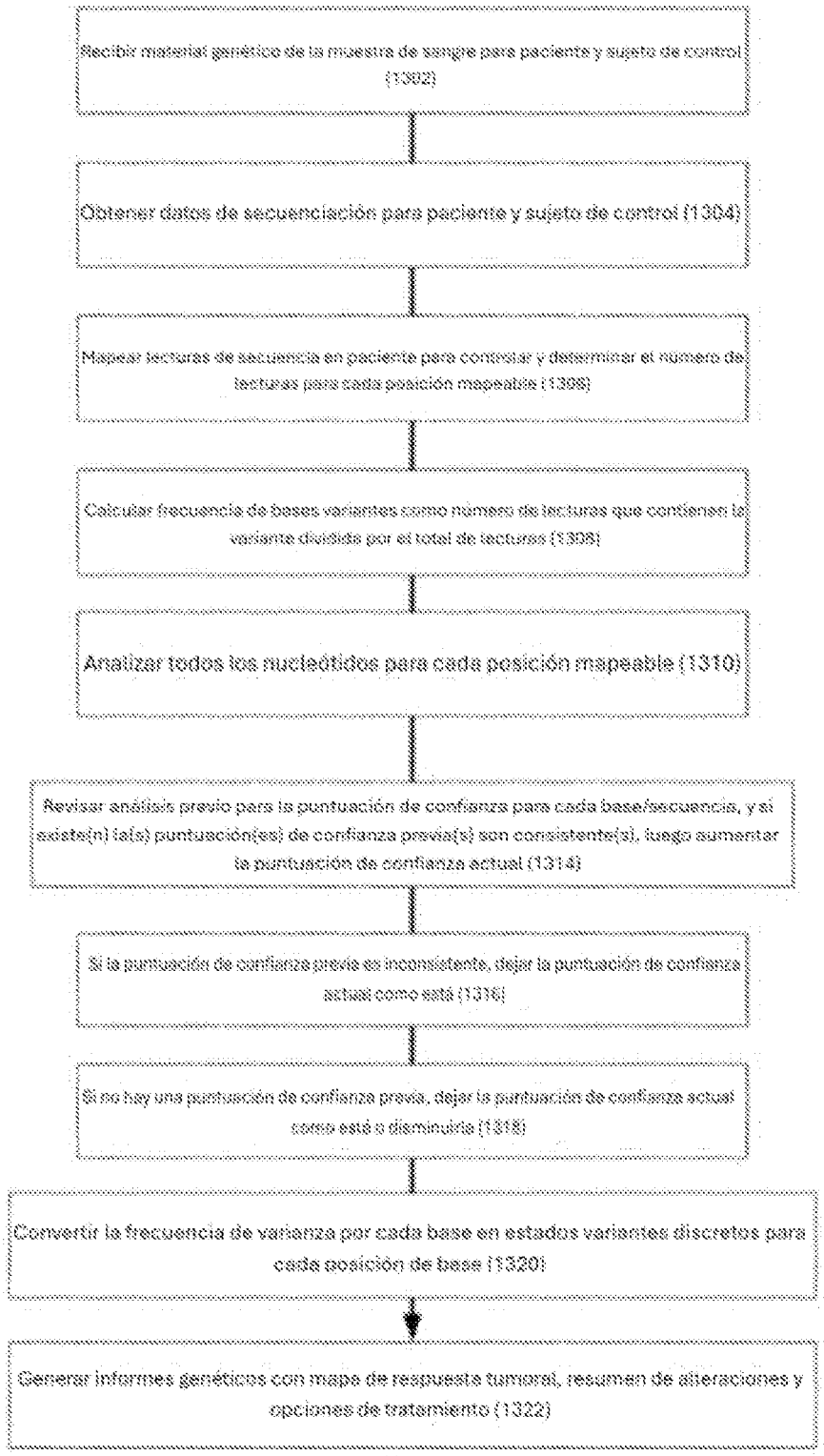

Recibir material genético de la muestra de sangre para paciente y sujeto de control (1302)
Obtener datos de secuenciación para paciente y sujeto de control (1304)
Mapear lecturas de secuencia en paciente para controlar y determinar el número de lecturas para cada posición mapeable (1306)
Calcular frecuencia de bases variantes como número de lecturas que contienen la variante dividido por el total de lecturas (1308)
Analizar todos los nucleótidos para cada posición mapeable (1310)
Revisar análisis previo para la puntuación de confianza para cada base/secuencia, y si existe(n) la(s) puntuación(es) de confianza previa(s) son consistente(s), luego aumentar la puntuación de confianza actual (1314)
Si la puntuación de confianza previa es inconsistente, dejar la puntuación de confianza actual como está (1316)
Si no hay una puntuación de confianza previa, dejar la puntuación de confianza actual como está o disminuirla (1318)
Convertir la frecuencia de varianza por cada base en estados variantes discretos para cada posición de base (1320)
Generar informes genéticos con mapa de respuesta tumoral, resumen de alteraciones y opciones de tratamiento (1322)

FIG. 11B