

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 704 303**

51 Int. Cl.:

**C12Q 1/6886** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.08.2011** **E 16166825 (6)**

97 Fecha y número de publicación de la concesión europea: **26.12.2018** **EP 3109325**

54 Título: **Procedimientos para la predicción de una respuesta contra el cáncer**

30 Prioridad:

**24.08.2010 US 402116 P**

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

**15.03.2019**

73 Titular/es:

**DANA-FARBER CANCER INSTITUTE, INC.**

**(25.0%)**

**450 Brookline Avenue**

**Boston, MA 02215, US;**

**THE BRIGHAM AND WOMEN'S HOSPITAL, INC.**

**(25.0%);**

**CHILDREN'S MEDICAL CENTER CORPORATION**

**(25.0%) y**

**THE TECHNICAL UNIVERSITY OF DENMARK**

**(25.0%)**

72 Inventor/es:

**RICHARDSON, ANDREA L.;**

**WANG, ZHIGANG C.;**

**SILVER, DANIEL P.;**

**SZALLASI, ZOLTAN;**

**BIRKBAK, NICOLAI JUUL y**

**EKLUND, ARON C.**

74 Agente/Representante:

**SALVÀ FERRER, Joan**

**ES 2 704 303 T3**

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

## DESCRIPCIÓN

Procedimientos para la predicción de una respuesta contra el cáncer

5 **[0001]** Los oncólogos médicos se han beneficiado en gran medida de los esfuerzos relativamente recientes para diseccionar y comprender los elementos genéticos subyacentes del cáncer en mamíferos. La identificación de predisposiciones genéticas específicas, tales como mutaciones en BRCA-1, BRCA2, y HER2, ha proporcionado información clave en los mecanismos subyacentes de la tumorigénesis y ha demostrado ser útil para el diseño de nuevas generaciones de enfoques dirigidos para la intervención clínica. Con la determinación de la secuencia del genoma humano y las mejoras en las tecnologías de secuenciación y bioinformática, se han convertido en posibles los análisis sistemáticos de alteraciones genéticas en cánceres humanos. Sin embargo, las intervenciones clínicas en base a esta información se han visto gravemente obstaculizadas por el hecho de que a menudo sólo un porcentaje de los pacientes responderá favorablemente a un tratamiento contra un cáncer determinado. Actualmente, los oncólogos médicos, en general, no pueden predecir qué pacientes responderán o no a un tratamiento quimioterapéutico propuesto. En consecuencia, hay una gran necesidad en la técnica para identificar la capacidad de respuesta del paciente a determinadas terapias contra el cáncer.

10 **[0002]** El documento WO 2004/042032 describe anomalías citogenéticas que son predictivas de la respuesta al tratamiento de la leucemia linfocítica crónica. El documento WO 2006/098978 describe un procedimiento de diagnóstico para la identificación de pacientes candidatos para el tratamiento con trastuzumab. El documento WO 2009/148528 describe la evaluación de alteraciones cromosómicas en loci genéticos específicos para predecir el resultado clínico del tratamiento con bortezomib.

15 **[0003]** Ferreira et al., Oncogene, vol. 27 (14), 27 de marzo de 2008, páginas 2084-2090, describe que la matriz ("array") de CGH y elaboración de perfiles de expresión genética revelan distintos patrones de inestabilidad genómica asociada con los mecanismos de reparación del ADN y los puntos de control del ciclo celular en el sarcoma de Ewing.

20 **[0004]** Etemadmoghadam et al., Clinical Cancer Research, vol. 15 (4), 15 de febrero de 2009, páginas 1417-1427, describe que el número de copias de ADN integrado en todo el genoma y el análisis de expresión identifican mecanismos distintos de quimioresistencia primaria en carcinomas de ovario.

25 **[0005]** La presente invención es tal como se define en las reivindicaciones adjuntas. La presente invención se basa, al menos en parte, en el descubrimiento de que ciertos patrones de aberraciones de ADN descritas en este documento son predictivos de la respuesta contra el cáncer de las células que albergan tales aberraciones de ADN a las terapias contra el cáncer. Por consiguiente, en un aspecto, la presente descripción presenta un procedimiento para predecir el resultado del tratamiento contra el cáncer de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberraciones cromosómicas globales (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si una pluralidad de regiones cromosómicas que muestran una aberración cromosómica existen dentro de una pluralidad de loci cromosómicos, en el que dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico (AI), pérdida de heterocigosidad (LOH), aberraciones en el número de copias (CNA), ganancia del número de copias (CNG), disminución del número de copias (CND) y combinaciones de los mismos, en relación a un control, y en el que la presencia de una pluralidad de regiones cromosómicas que muestran dichas aberraciones cromosómicas predice el resultado del tratamiento contra el cáncer del sujeto. El sujeto puede ser un mamífero, tal como un humano.

30 **[0006]** En un aspecto, el tratamiento contra el cáncer es el tratamiento por quimioterapia. En otra realización, el tratamiento contra el cáncer comprende agentes quimioterapéuticos basados en platino (por ejemplo, cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino).

35 **[0007]** En otro aspecto, el trastorno hiperproliferativo celular se selecciona del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de vejiga de células transicionales, cáncer de pulmón broncogénico, cáncer de tiroides, cáncer de páncreas, cáncer de próstata, cáncer de útero, cáncer testicular, cáncer gástrico, sarcomas de tejidos blandos y osteogénicos, neuroblastoma, tumor de Wilms, linfoma maligno (de Hodgkin y no Hodgkin), leucemia mieloblástica aguda, leucemia linfoblástica aguda, sarcoma de Kaposi, tumor de Ewing, mieloma múltiple refractario, y carcinomas de células escamosas de la cabeza, el cuello, el cuello del útero y la vagina.

40 **[0008]** En aún otro aspecto, la muestra biológica se selecciona del grupo que consiste en células, líneas celulares, cortes histológicos, biopsias centrales congeladas, tejidos embebidos en parafina, tejidos fijados con formalina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspadura bucal, saliva, líquido cefalorraquídeo, orina, heces, y médula ósea. En una realización, la muestra biológica se enriquece por la presencia de células hiperproliferativas hasta por lo menos el 75% de la población total de células. En otra realización, el enriquecimiento se realiza según al menos una técnica seleccionada del grupo que consiste en microdissección con aguja, microdissección por láser, clasificación de células activadas por fluorescencia y clasificación de células inmunológicas. En aún otra realización, una máquina automatizada realiza dicha al menos una técnica para transformar así la muestra biológica en una forma purificada enriquecida por la presencia de células

hiperproliferativas. En aún otra realización, la muestra biológica se obtiene antes de que el sujeto haya recibido quimioterapia adyuvante. Alternativamente, la muestra biológica se obtiene después de que el sujeto haya recibido quimioterapia adyuvante.

5 [0009] En aún otro aspecto, el control se determina a partir de una muestra de células no hiperproliferativas del paciente o miembro de la misma especie a la que pertenece el paciente. En una realización, el control se determina a partir de la frecuencia media de aparición del locus genómico de regiones cromosómicas del mismo grupo étnico dentro de la especie a la que pertenece el paciente. En otra realización, el control proviene de tejido no canceroso que es el mismo tipo de tejido que dicho tejido canceroso del sujeto. En aún otra realización, el control proviene de  
10 tejido no canceroso que no es el mismo tipo de tejido que dicho tejido canceroso del sujeto.

[0010] En otro aspecto, el AI se determina usando una proporción mayoritaria de copias (MCP). En una realización, el IA para una región genómica determinada se cuenta cuando MCP es mayor que 0,70.

15 [0011] En aún otro aspecto, la pluralidad de loci cromosómicos están distribuidos al azar en todo el genoma al menos cada 100 Kb de ADN. En una realización, la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada uno de los 23 pares de cromosomas humanos. En otra realización, la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada brazo de cada uno de los 23 pares de cromosomas humanos. En aún otra realización, la pluralidad de loci cromosómicos comprenden al menos un locus cromosómico en al menos un telómero de cada uno de los 23 pares de cromosomas humanos. En aún otra  
20 realización, la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada telómero de cada uno de los 23 pares de cromosomas humanos.

[0012] En aún otro aspecto, las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 1 Mb. En una realización, las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 12 Mb.

[0013] En otro aspecto, la pluralidad de aberraciones cromosómicas comprende al menos 5 aberraciones cromosómicas. En una realización, la pluralidad de aberraciones cromosómicas comprende al menos 13 aberraciones cromosómicas.

30 [0014] En aún otro aspecto, los loci cromosómicos se seleccionan del grupo que consiste en polimorfismos de nucleótido único (SNP), polimorfismos de longitud de fragmento de restricción (RFLP), y repeticiones en tándem simples (STR).

35 [0015] En aún otro aspecto, los loci cromosómicos se analizan usando al menos una técnica seleccionada del grupo que consiste en sonda de inversión molecular (MIP), matriz de polimorfismos de nucleótido único (SNP), hibridación in situ, transferencia Southern, matriz de hibridación genómica comparada (aCGH) y la secuenciación de próxima generación.

40 [0016] En otro aspecto, el resultado del tratamiento se mide mediante al menos un criterio seleccionado del grupo que consiste en la supervivencia hasta la mortalidad, respuesta patológica completa, mediciones semicuantitativas de la respuesta patológica, remisión clínica completa, remisión clínica parcial, enfermedad clínica estable, supervivencia libre de recurrencia, supervivencia libre de metástasis, supervivencia libre de enfermedad, disminución de células tumorales circulantes, respuesta del marcador circulante y criterios RECIST.

45 [0017] En aún otro aspecto, el procedimiento comprende además la determinación de un régimen de tratamiento adecuado para el sujeto. En una realización, el régimen de tratamiento adecuado comprende al menos un agente quimioterapéutico basado en platino cuando se determina una pluralidad de aberraciones cromosómicas genómicas o no comprende al menos un agente quimioterapéutico basado en platino cuando no se determina una pluralidad de  
50 aberraciones cromosómicas genómicas.

**Breve descripción de los dibujos**

[0018]  
55 Las **figuras 1A-1C** muestran la correlación entre las regiones de desequilibrio alélico (AI) y la sensibilidad al cisplatino in vitro. La **figura 1A** muestra curvas dosis-respuesta de seis líneas celulares TNBC tal como se determina por un ensayo de proliferación después de 48 horas de exposición a cisplatino. Las curvas para células con valores de IC<sub>50</sub> inferiores (mayor sensibilidad) se muestran en azul; la línea celular con la mayor IC<sub>50</sub> (mayor resistencia) se muestra en rojo; las líneas celulares con una sensibilidad intermedia se muestran en gris. La **Figura 1B** muestra el efecto del umbral de tamaño de segmento de AI sobre la correlación entre el número de regiones de AI telomérico y la sensibilidad a cisplatino en las seis líneas celulares. Cada punto representa un valor R<sup>2</sup> basado en la regresión lineal entre el número de regiones CNA de un tamaño mínimo indicado en el eje X, y IC<sub>50</sub> de cisplatino en un panel de 6 líneas celulares de 6 TNBC (BT20, BT-549, HCC1187, HCC38, MDA-MB-231, MDA-MB-468). El umbral de tamaño de segmento mínimo óptimo se indica por la línea de puntos. La **Figura 1C** muestra una comparación entre  
60 el número de regiones de AI telomérico (N<sub>AI,12</sub>) y la sensibilidad a cisplatino en el umbral óptimo seleccionado de 12 Mb. Las líneas celulares se indican de la siguiente manera: 1, BT-20; 2, BT-549; 3, HCC1187; 4, HCC38; 5, MDA-

MB-231; 6, MDA-MB-468.

Las **figuras 2A-2C** muestran que el análisis de la proporción mayoritaria de copias (MCP) identifica el desequilibrio alélico en muestras de biopsia de tumores con diferentes grados de pureza de las células tumorales. La **Figura 2A** muestra la fórmula para el cálculo de MCP, así como los cromosomas bialélicos normales y tres maneras diferentes en las que se puede producir el desequilibrio alélico de una región cromosómica y el cálculo de MCP correspondiente. Las Figuras **2B y 2C** muestran diagramas que representan la observación de la pérdida de heterocigosidad (LOH), AI determinado por MCP, y el análisis del número absoluto de copias en dos muestras de tumores con diferentes grados de contaminación celular normal: T7 con un contenido de células tumorales > 95% (Figura 2B) y T5 con un contenido de tumor de aproximadamente el 80% (Figura 2C). Los cromosomas se indican a lo largo de la parte izquierda. Las primeras columnas para cada tumor muestran las células para LOH (azul) y la retención de la heterocigosidad (amarillo) en cada posición de cromosoma. La segunda columna muestra los niveles de MCP (entre 0,5 y 1,0) en cada posición cromosómica. El corte de MCP de 0,7 se indica con líneas rojas. El AI se pronostica para las regiones con MCP superiores a 0,7. La tercera y cuarta columnas muestran el número absoluto de copias de ADN en cada posición, indicando el blanco el diploide, indicando los tonos de rojo la ganancia de copias e indicando los tonos de azul la pérdida de copias. Los niveles del número de copias se muestran en los paneles de más a la derecha. La muestra de tumor con mayor pureza (T7 en la Figura 2B), muestra una concordancia entre la LOH y las células con AI determinado por MCP. En la muestra de tumor con sólo un 80% de células tumorales, se pierde la señal LOH, pero AI todavía se puede estimar mediante MCP con un umbral de 0,70.

Las **figuras 3A-3D** muestran la asociación entre la sensibilidad a cisplatino y el número de alteraciones genómicas en un panel de líneas celulares TNBC. La **figura 3A** muestra la  $IC_{50}$  de cisplatino frente al número de regiones AI teloméricas de al menos 1 Mb de largo con AI definido por MCP > 0,7. La **Figura 3B** muestra  $IC_{50}$  de cisplatino frente al recuento de las regiones con aberración en el número de copias, incluyendo las ganancias y las pérdidas, de al menos 1 Mb de largo. La figura 3C muestra  $IC_{50}$  de cisplatino frente al recuento de las regiones con una ganancia del número de copias, de al menos 1 Mb de largo. La **Figura 3D** muestra  $IC_{50}$  de cisplatino frente al recuento de las regiones con una pérdida del número de copias, de al menos 1 Mb largos. Las líneas de células se indican en cada figura y son las mismas que en la figura 1.

Las **figuras 4A-4B** muestran la asociación entre la sensibilidad a cisplatino y el recuento de regiones de AI telomérico o intersticial en un panel de líneas celulares TNBC. La **figura 4A** muestra  $IC_{50}$  de cisplatino frente al número de regiones de AI telomérico de al menos 1 Mb de largo con AI definido por MCP > 0,7. La figura 4B muestra  $IC_{50}$  de cisplatino frente a varias regiones de AI intersticiales de al menos 1 Mb de largo con AI definido por MCP > 0,7. Las líneas celulares se indican en cada figura y son las mismas que en la figura 1.

Las figuras **5A-5F** muestran la asociación entre las aberraciones del número de copias (CNA) enumeradas y la sensibilidad al cisplatino in vitro. Las **figuras 5A-5C** muestran la determinación del tamaño de segmento mínimo que demuestra la mejor correlación con la sensibilidad a cisplatino para el número aberraciones del número de copias (NCNA; Figura 5A), el número de regiones con ganancia del número de copias (NCNA, ganancia; figura 5B), y el número de regiones con pérdida del número de copias (NCNA, pérdida; Figura 5C). Cada punto representa un valor  $R^2$  basado en la regresión lineal entre el recuento de regiones de CNA de un tamaño mínimo indicado en el eje X, y  $IC_{50}$  de cisplatino en un panel de 6 líneas celulares TNBC (BT20, BT-549, HCC1187, HCC38, MDA-MB-231, MDA-MB-468). El tamaño mínimo óptimo de regiones de CNA se indica por la línea de puntos. Las **Figuras 5D-5F** muestran representaciones de los valores de  $IC_{50}$  de cisplatino ( $\mu M$ , eje X) frente al número de regiones de CNA con tamaños de segmento mínimo óptimos (eje Y) tal como se indica a continuación: NCNA de al menos 9 Mb de largo (figura 5D), NCNA, ganancia, de al menos 9 Mb de largo (Figura 5E), y NCNA, pérdida, de al menos 5 Mb de largo, en 6 líneas celulares TNBC (Figura 5F), tal como se indica.

Las figuras **6A-6C** muestran regiones AI y respuesta a cisplatino en cáncer de mama. La respuesta patológica a cisplatino se evaluó mediante la puntuación de Miller-Payne (MP), que puede variar de 0 (progresión) a 5 (respuesta patológica completa, pCR). Higo. La **figura 6A** muestra representaciones de genomas de tumores individuales dispuestos en orden creciente de puntuación de MP. Se indican las regiones de AI telomérico (azul oscuro) y AI intersticial (azul claro), con finas líneas blancas delimitando los cromosomas individuales. La **figura 6B** muestra la asociación entre la puntuación MP y  $N_{IAI,12}$ . La **Figura 6C** muestra una curva de características de operación del receptor (ROC) para evaluar el rendimiento de  $N_{IAI,12}$  para predecir la pCR al tratamiento con cisplatino (pCR, n = 4; no pCR, n = 20).

La **figura 7** muestra el desequilibrio alélico del cromosoma completo (isodisomía) y la sensibilidad a cisplatino en los cánceres de mama. Las regiones de AI del cromosoma completo se indican en rojo para cada localización cromosómica. Cada fila definida por las líneas blancas finas representa un cromosoma diferente y los números de cromosoma se indican a lo largo de la parte izquierda. Cada columna representa una muestra de tumor individual. La puntuación de respuesta patológica de Miller-Payne (MP) para cada tumor se indica en la parte inferior. Los casos están dispuestos en orden creciente de respuesta patológica al cisplatino (0 = progresión, 5 = respuesta patológica completa (pCR)).

Las figuras **8A-8B** muestran regiones AI y el tiempo hasta la recaída en el cáncer de ovario seroso tratado con terapia basada en platino. La **figura 8A** muestra un rango de individuos según el aumento de  $N_{IAI,12}$ . Los que recayeron en menos de un año se indican con círculos sólidos y los que no experimentaron recaída en menos de un año se indican con círculos no sólidos. El valor de corte de  $N_{IAI,12} = 13$ , basado en el análisis ROC de TNBC para la predicción de la respuesta patológica completa (pCR) a cisplatino, está indicado por la línea de puntos. La **Figura 8B** muestra las curvas de supervivencia de Kaplan-Meier para el tiempo hasta la recaída en individuos clasificados como  $N_{IAI,12}$  alto (regiones de  $N_{IAI,12}$  13 o mayor, azul) o  $N_{IAI,12}$  bajo (regiones de  $N_{IAI,12}$  de menos de 13, rojo).

La **figura 9** muestra un modelo que relaciona la reparación del ADN a la acumulación de AI y la respuesta a los

agentes de platino. Diversas lesiones genéticas pueden dar lugar a defectos en los mecanismos comunes de la reparación del ADN, lo que lleva primero a la reparación de roturas espontáneas de ADN, a continuación, la recombinación ilegítima de cromosomas y la formación de cromosoma cuadrirradial aberrante, y, finalmente, a altos niveles de desequilibrio alélico telomérico. En paralelo, el mecanismo de reparación del ADN defectuoso también puede dar lugar a la incapacidad de la célula tumoral de reparar el daño del ADN inducido por fármacos, lo que conduce a la sensibilidad del tumor a fármacos, tales como las sales de platino. Por lo tanto, el nivel de AI telomérico en un tumor sirve como un indicador de la reparación del ADN defectuoso y predice la sensibilidad al tratamiento con agentes genotóxicos.

## 10 **Descripción detallada de la invención**

[0019] La presente descripción se refiere a procedimientos para predecir la respuesta de un cáncer en un sujeto a las terapias contra el cáncer basadas en la determinación y el análisis de una puntuación global de aberración cromosómica (GCAS).

15 I. Determinación de la puntuación global de aberración cromosómica (GCAS)

[0020] Según un aspecto de la descripción, la GCAS es una medición predictiva de la capacidad de respuesta a las terapias contra el cáncer de un cáncer en un sujeto. Esta utilidad de GCAS se basa en el hallazgo novedoso de que la suma de las aberraciones cromosómicas individuales puede predecir la capacidad de respuesta de un cáncer en un sujeto a los agentes contra el cáncer independientemente de identificar aberraciones cromosómicas específicas. Se utilizan loci informativos de interés (por ejemplo, polimorfismos de nucleótido único (SNP), polimorfismos de longitud de fragmento de restricción (RFLP), repeticiones en tándem simples (STRs), etc.), para determinar GCAS, ya que son útiles para detectar y/o distinguir aberraciones cromosómicas (por ejemplo, desequilibrio alélico, pérdida de heterocigosidad, cambio del número total de copias, ganancia del número de copias y pérdida del número de copias).

[0021] La GCAS se determina mediante la determinación de una pluralidad o el número total de regiones cromosómicas que presentan desequilibrio alélico (NAI), la pérdida de heterocigosidad (LOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG), y/o la pérdida del número de copias (NCND), tal como se describe adicionalmente en este documento y según procedimientos bien conocidos en la técnica. Una GCAS de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, o 60 o más es predictivo de la respuesta a la terapia contra el cáncer de la célula cancerosa de la que se derivaba el ácido nucleico ensayado.

[0022] En una realización, el análisis se basa en los ácidos nucleicos obtenidos de un sujeto y/o muestra de control. Dichas muestras pueden incluir "fluidos corporales", que se refieren a los fluidos que se excretan o secretan del cuerpo, así como los fluidos que normalmente no lo hacen (por ejemplo, líquido amniótico, humor acuoso, bilis, sangre y plasma sanguíneo, líquido cefalorraquídeo, cerumen y cera de los oídos, fluido de Cowper o pre-eyaculatorio, quilo, quimo, heces, eyaculado femenino, líquido intersticial, líquido intracelular, linfa, menstruación, leche materna, moco, líquido pleural, pus, saliva, sebo, semen, suero, sudor, líquido sinovial, lágrimas, orina, lubricación vaginal, humor vítreo, vómito). En una realización preferida, el sujeto y/o muestra de control se seleccionan del grupo que consiste en células, líneas celulares, cortes histológicos, tejidos embebidos en parafina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspado bucal, saliva, líquido cefalorraquídeo, orina, heces y médula ósea.

[0023] En una realización, se utilizan SNP en la determinación de GCAS para predecir la capacidad de respuesta de un cáncer a una terapia contra el cáncer. Hay seis tipos posibles de SNP, ya sean transiciones (A <> T o G <> C) o transversiones (A <> G, A <> C, G <> T o C <> T). Los SNP son ventajosos en que se pueden identificar grandes cantidades.

[0024] En algunas realizaciones, los SNP u otros loci genómicos se puede puntuar para detectar anomalías en el número de copias. En tales casos, dichos loci genómicos no necesitan ser informativos en términos de genotipo, ya que el número de copias se determina mediante las intensidades de hibridación y no depende del genotipo. Además, las anomalías del número de copias se pueden detectar usando procedimientos que no utilizan SNP, tales como, por ejemplo, una matriz de CGH utilizando BAC, cDNA y/o grupos de oligonucleótidos.

[0025] Por ejemplo, los procedimientos para evaluar el número de copias de ácido nucleico en una muestra incluyen, pero no se limitan a, ensayos basados en hibridación. Un procedimiento para evaluar el número de copias del ácido nucleico codificante en una muestra implica una transferencia Southern. En una transferencia Southern, el ADN genómico (habitualmente fragmentado y separado en un gel electroforético) se hibrida con una sonda específica para la región diana. La comparación de la intensidad de la señal de hibridación de la sonda para la región diana con la señal de sonda de control a partir del análisis de ADN genómico normal (por ejemplo, una parte no amplificada del mismo o célula, tejido, órgano, etc. relacionados) proporciona una estimación del número relativo de copias del ácido nucleico diana. Alternativamente, se puede utilizar una transferencia Northern para evaluar el número de copias del

ácido nucleico codificante en una muestra. En una transferencia Northern, el ARNm se hibrida con una sonda específica para la región diana. La comparación de la intensidad de la señal de hibridación de la sonda para la región diana con la señal de sonda de control a partir del análisis de ARNm normal (por ejemplo, una parte no amplificada del mismo o célula, tejido, órgano, etc. relacionados) proporciona una estimación del número relativo de copias del ácido nucleico diana. Se pueden realizar procedimientos similares para determinar el número de copias utilizando matrices de transcripción que son bien conocidos en la técnica.

**[0026]** Un medio alternativo para la determinación del número de copias es la hibridación in situ (por ejemplo, Angerer (1987) Meth Enzymol 152: 649). En general, la hibridación in situ comprende las siguientes etapas: (1) fijación del tejido o de la estructura biológica a analizar; (2) tratamiento previo a la hibridación de la estructura biológica para aumentar la accesibilidad del ADN diana, y para reducir la unión no específica; (3) hibridación de la mezcla de ácidos nucleicos con el ácido nucleico en la estructura o tejido biológico; (4) lavados posteriores a la hibridación para eliminar fragmentos de ácido nucleico no unidos en la hibridación y (5) detección de los fragmentos de ácido nucleico hibridados. El reactivo utilizado en cada una de estas etapas y las condiciones de uso varían dependiendo de la aplicación particular.

**[0027]** Los ensayos basados en hibridación preferidos incluyen, pero sin limitarse a, procedimientos tradicionales de "sonda directa", tales como transferencias Southern o hibridación in situ (por ejemplo, FISH y FISH más SKY), y los procedimientos de "sonda comparativa", tales como hibridación genómica comparada (CGH), por ejemplo, CGH basada en ADNc o basada en oligonucleótidos. Los procedimientos se pueden utilizar en una amplia variedad de formatos, incluyendo, pero sin limitarse a, procedimientos de unión a sustrato (por ejemplo, membrana o vidrio) o estrategias basadas en matrices.

**[0028]** En un ensayo de hibridación in situ típico, las células se fijan a un soporte sólido, habitualmente un portaobjetos de vidrio. Si se debe sondear un ácido nucleico, las células habitualmente se desnaturalizan con calor o álcali. Las células se ponen en contacto a continuación con una solución de hibridación a una temperatura moderada para permitir la hibridación de sondas marcadas específicas a la secuencia de ácido nucleico que codifica la proteína. Las dianas (por ejemplo, células) a continuación habitualmente se lavan a una rigurosidad predeterminada o a una rigurosidad creciente hasta que se obtiene una relación de señal a ruido apropiada.

**[0029]** Las sondas se marcan habitualmente, por ejemplo, con radioisótopos o indicadores fluorescentes. Las sondas preferidas son suficientemente largas para hibridarse específicamente con el ácido o ácidos nucleicos diana en condiciones rigurosas. El intervalo de tamaño preferido es de aproximadamente 200 bases a aproximadamente 1000 bases.

**[0030]** En algunas aplicaciones, es necesario bloquear la capacidad de hibridación de secuencias repetitivas. Por lo tanto, en algunas realizaciones, ARNt, ADN genómico humano, o ADN Cot-I se utilizan para bloquear la hibridación no específica.

**[0031]** En los procedimientos de CGH, se marca una primera colección de ácidos nucleicos (por ejemplo, de una muestra, por ejemplo, un posible tumor) con un primer marcador, mientras que una segunda colección de ácidos nucleicos (por ejemplo, un control, por ejemplo, una célula/tejido sano) se marca con un segundo marcador. Se determina la proporción de hibridación de los ácidos nucleicos mediante la proporción de los dos marcadores (primero y segundo) que se unen a cada fibra en la matriz. Donde hay deleciones o multiplicaciones cromosómicas, se detectarán diferencias en la proporción de las señales de los dos marcadores y la proporción proporcionará una medida del número de copias. La CGH basada en una matriz también se puede realizar con marcaje de un solo color (en lugar de marcar el control y la posible muestra de tumor con dos colorantes diferentes y mezclarlos antes de la hibridación, lo que producirá una proporción debido a la hibridación competitiva de sondas en las matrices). En CGH de un solo color, el control se marca y se hibrida con una matriz y se leen las señales absolutas, y la posible muestra de tumor se marca y se hibrida con una segunda matriz (con idéntico contenido) y se leen las señales absolutas. Se calcula la diferencia en el número de copias en base a las señales absolutas de las dos matrices. Se describen protocolos de hibridación adecuados para su uso con los procedimientos de la invención, por ejemplo, en Albertson (1984) EMBO J. 3: 1227-1234; Pinkel (1988) Proc. Natl. Acad. Sci. USA. 85: 9138-9142; EPO Pub. N° 430.402; Methods in Molecular Biology, Vol. 33: In situ Hybridization Protocols, Choo, ed, Humana Press, Totowa, NJ (1994), etc. En una realización, se utiliza el protocolo de hibridación de Pinkel, et al. (1998) Nature Genetics 20: 207-211, o de Kallioniemi (1992) Proc. Natl Acad Sci USA. 89: 5321-5325 (1992).

**[0032]** Los procedimientos de la invención están particularmente bien adaptados a los formatos de hibridación basados en matrices. La CGH basada en una matriz se describe en la patente de Estados Unidos N° 6.455.258.

**[0033]** En aún otra realización, se pueden utilizar ensayos basados en la amplificación para medir el número de copias. En tales ensayos basados en la amplificación, las secuencias de ácidos nucleicos actúan como una plantilla en una reacción de amplificación (por ejemplo, reacción en cadena de la polimerasa (PCR)). En una amplificación cuantitativa, la cantidad de producto de amplificación será proporcional a la cantidad de plantilla en la muestra original. La comparación con los controles adecuados, por ejemplo, tejido sano, proporciona una medida del número de copias.

**[0034]** Los procedimientos de amplificación "cuantitativa" son bien conocidos por los expertos en la materia. Por ejemplo, una PCR cuantitativa implica la coamplificación simultánea de una cantidad conocida de una secuencia de control usando los mismos cebadores. Esto proporciona un patrón interno que puede usarse para calibrar la reacción de PCR. Los protocolos detallados para la PCR cuantitativa se proporcionan en Innis, et al. (1990) PRC Protocol, A guide to methods and application, Academic Press, Inc. NY). La medición del número de copias de ADN en los loci de microsatélites utilizando análisis de PCR cuantitativa se describe en Ginzonger, et al. (2000) Cancer Research 60: 5405-5409. La secuencia nucleica conocida para los genes es suficiente para permitir a un experto en la materia seleccionar de forma rutinaria cebadores para amplificar cualquier parte del gen. También se puede usar PCR cuantitativa fluorogénica en los procedimientos de la invención. En la PCR cuantitativa fluorogénica, la cuantificación se basa en la cantidad de señales de fluorescencia, por ejemplo, TaqMan y verde sybr.

**[0035]** Otros procedimientos de amplificación adecuados incluyen, pero no se limitan a, reacción en cadena de la ligasa (LCR) (véase Wu y Wallace (1989) Genomics 4: 560, Landegren, et al (1988) Science 241: 1077, y Barringer et al. (1990) Gene 89: 117), amplificación de la transcripción (Kwoh, et al. (1989). Proc Natl Acad Sci USA. 86: 1173), replicación de secuencia autosostenida (Guatelli, et al. (1990), Proc Nat Acad Sci USA. 87: 1874), PCR de puntos y PCR con adaptador enlazador, etc.

**[0036]** En aún otras realizaciones de los procedimientos proporcionados en este documento, se realiza la secuenciación por separado de las moléculas nucleicas individuales (o sus productos de amplificación), como una alternativa a los ensayos basados en hibridación, usando técnicas de secuenciación de ácido nucleico. En una realización, se utiliza una técnica de secuenciación en paralelo de alto rendimiento que aísla moléculas de ácido nucleico individuales de una población de moléculas de ácido nucleico antes de la secuenciación. Dichas estrategias utilizadas se llaman "sistemas de secuenciación de próxima generación", incluyendo, sin limitación, máquinas de secuenciación y/o estrategias bien conocidas en la técnica, tales como las desarrollados por Illumina/Solexa (el analizador Genoma; Bennett et al (2005) Pharmacogenomics, 6: 373-382), por Applied Biosystems, Inc. (el secuenciador SOLiD; solid.appliedbiosystems.com), por Roche (por ejemplo, el secuenciador 454 GS FLX; Margulies et al (2005). Nature, 437: 376-380; patente de Estados Unidos N° 6.274.320; 6.258.568; 6.210.891), mediante el sistema Heliscope® de Helicos Biosciences (véase, por ejemplo, publicación de la solicitud de patente de Estados Unidos N° 2007/0070349), y por otros. Todas las estrategias de determinación del número de copias descritas en el presente documento se pueden aplicar de manera similar a cualquiera de otros análisis basados en ácidos nucleicos descritos en este documento, tales como para loci informativo y similares descritos más adelante.

**[0037]** En otras realizaciones, los SNP se puede valorar para heterocigosidad o ausencia de heterocigosidad. Técnicas, como el análisis de la proporción mayoritaria de copias, utilizan la información del desequilibrio alélico y el número de copias para ampliar los análisis que se pueden realizar con el número de copias de sucesos LOH solos, ya que pueden implicar sucesos de eliminación, ganancia o ninguna variación del número de copias eliminación.

**[0038]** En otras realizaciones, para determinar la GCAS de un cáncer en un sujeto, se identifican los SNP heterocigotos localizados en todo el genoma usando muestras de ácido nucleico derivadas de tejido no canceroso del sujeto o una población de sujetos de una misma especie, y se determina el número de los SNP heterocigotos identificados que mantienen la heterocigosidad (o, alternativamente, no exhiben heterocigosidad, es decir, han perdido la heterocigosidad) en una muestra de ácido nucleico de, o derivado de, ADN genómico de tejido canceroso del sujeto. Una muestra de ácido nucleico "derivado de" ADN genómico incluye, pero no se limita a, pre-ARN mensajero (que contiene intrones), productos de amplificación de ADN genómico o pre-ARN mensajero, fragmentos de ADN genómico opcionalmente con oligonucleótidos adaptadores ligados al mismo o presentes en vectores de clonación u otros vectores, etc. (los intrones y regiones no codificantes no deben ser eliminados de forma selectiva).

**[0039]** Todos los SNP conocidos que muestran heterocigosidad en la especie a la que el sujeto con cáncer pertenece, no deben incluirse en el número de SNP heterocigotos utilizados. En algunas realizaciones, se utilizan al menos 250, 500, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 21.000, 22.000, 23.000, 24.000, 25.000, 26.000, 27.000, 28.000, 29.000, 30.000, 31.000, 32.000, 33.000, 34.000, 35.000, 36.000, 37.000, 38.000, 39.000, 40.000, 41.000, 42.000, 43.000, 44.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 150.000, 200.000, 250.000, 300.000, 350.000, 400.000, 450.000, 500.000, 550.000, 600.000, 650.000, 700.000, 750.000, 800.000, 850.000, 900.000, 950.000, 1.000.000 SNPs o más o en cualquier intervalo intermedio o en otros loci informativos de interés (por ejemplo, RFLP, STR, etc.). Preferiblemente, dichos SNP están en el genoma humano. En una realización, la pluralidad de SNP heterocigotos están distribuidos al azar en todo el genoma al menos cada 1, 5, 10, 50, 100, 250, 500, 1.000, 1.500, 2.000, 2.500, 3.000, 5.000, 10.000 kb o más, o cualquier intervalo intermedio. Por "distribuido al azar", como se usa anteriormente, se entiende que los SNP de la pluralidad no se selecciona por predisposición hacia cualquier locus o loci cromosómicos específicos; sin embargo, se pueden utilizar otras preferencias (por ejemplo, evitar secuencias de ADN repetitivas) en la selección de los SNP. En otras realizaciones, la pluralidad de SNP heterocigotos no están distribuidos al azar en todo el genoma (es decir, distribuidos dentro de al menos 250, 500, 1.000, 1.500, 2.000, 2.500, 3.000, 5.000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 21.000, 22.000, 23.000, 24.000, o 25.000 kb = 25 Mb). Dichas regiones pueden además estar predispuestas, en algunas realizaciones, a regiones cromosómicas específicas, tales como telómeros,

definidas como regiones que se extienden hacia el telómero pero que no cruzan el centrómero. En una realización, el tamaño de segmento telomérico con desequilibrio alélico es de al menos 1 Mb, 2 Mb, 3 Mb, 4 Mb, 5 Mb, 6 Mb, 7 Mb, 8 Mb, 9 Mb, 10 Mb, 11 Mb, 12 Mb, 13 Mb, 14 Mb, 15 Mb, 16 Mb, 17 Mb, 18 Mb, 19 Mb, 20 Mb, 21 Mb, 22 Mb, 23 Mb, 24 Mb, 25 Mb o más, o cualquier intervalo intermedio, como por ejemplo entre 5 y 25 Mb. En otra realización, el tamaño de segmento telomérico con desequilibrio alélico es de 12 Mb. Por el contrario, las regiones intersticiales no implican el telómero. Las regiones intersticiales se definen en el presente documento como regiones de desequilibrio alélico que comienzan en dirección 3' del telómero, lo que significa que hay al menos una parte del cromosoma sin desequilibrio alélico entre el telómero y la región de desequilibrio alélico. En una realización, la pluralidad de SNP heterocigotos no se encuentra en regiones de ADN genómico que son repetitivas. En otra realización, la pluralidad de SNP heterocigotos comprende SNP situados en el genoma en diferentes loci cromosómicos, en el que los diferentes loci cromosómicos comprenden loci en cada uno de los cromosomas de la especie, o en cada brazo de cada cromosoma de la especie (por ejemplo, región telomérica del mismo).

**[0040]** Los SNP heterocigotos que se puede utilizar en los procedimientos de la invención para determinar el fenotipo de un cáncer son informativos, lo que significa que se observa heterocigosidad en la muestra de ácido nucleico de tejidos y/o células no cancerosas de un sujeto. Según los procedimientos de la invención, estos SNP informativos se examinan en la muestra de ácido nucleico de un tejido y/o células cancerosas de un sujeto para determinar la GCAS.

**[0041]** En una realización adicional, las muestras de ácido nucleico usadas para determinar el número de SNP heterocigotos en la pluralidad de SNP, que muestran heterocigosidad en el ADN genómico de tejido no canceroso de la especie a la que pertenece el paciente de cáncer, se toman de al menos 1, 2, 5, 10, 20, 30, 40, 50, 100, o 250 organismos diferentes de esa especie. Un experto en la materia podrá entender que se pueden determinar controles adecuados en base a la frecuencia promedio de alelos SNP que existen dentro del mismo grupo étnico de la especie a la que pertenece el paciente.

**[0042]** En ciertas realizaciones, los SNP informativos utilizados en los procedimientos de la invención para determinar y/o predecir el fenotipo de un cáncer comprenden al menos un SNP en cada cromosoma de un sujeto (por ejemplo, una región telomérica de cada cromosoma). En una realización relacionada, los SNP informativos utilizados en los procedimientos de la invención para determinar y/o predecir el fenotipo de un cáncer comprenden al menos un SNP en cada brazo de cada cromosoma de un sujeto (por ejemplo, una región telomérica de cada brazo de cada cromosoma).

## II. Predicción de la respuesta a la terapia

**[0043]** En ciertas realizaciones, la descripción proporciona procedimientos para determinar el fenotipo de un cáncer, en el que el fenotipo es la respuesta a la terapia. La terapia puede ser cualquier terapia contra el cáncer, incluyendo, pero no limitado a, quimioterapia, radioterapia, inmunoterapia, inhibidores de moléculas pequeñas, shARN, hormonal, y combinaciones de los mismos.

**[0044]** Cuando GCAS representa deleciones de copias, ganancia de copias, pérdidas de cromosomas enteros, ganancias de cromosomas enteros y/o pérdida de heterocigosidad, los sujetos cuyo tejido canceroso muestra una GCAS por debajo de un valor umbral se predice que tienen una peor respuesta al tratamiento (por ejemplo, radiación o quimioterapia) que aquellos con una GCAS alta (por encima del valor umbral). Cuando la GCAS representa la falta de copias o cambios en el número de cromosomas y/o retención de heterocigosidad, los sujetos cuyo tejido canceroso muestra una GCAS por encima de un valor umbral se predice que tienen una peor respuesta al tratamiento (por ejemplo, radiación o quimioterapia) que aquellos con una GCAS baja (por debajo del valor umbral).

**[0045]** A modo de explicación, pero sin pretender imponer ninguna teoría, se cree que cuando el valor de GCAS representa la pérdida de heterocigosidad o desequilibrio alélico, se identifican células que albergan roturas de la doble cadena de ADN cromosómico reparado incorrectamente y el recuento en todo el genoma de estos reordenamientos cromosómicos en un tumor específico indica el grado de incompetencia en la reparación del ADN, independientemente del defecto causante específico de la reparación del ADN. En estos sujetos, el número total de reordenamientos cromosómicos en un tumor indica la incapacidad de reparar el daño en el ADN inducido por las terapias contra el cáncer, y en consecuencia predice la sensibilidad a este tipo de terapias contra el cáncer. También a modo de explicación y sin pretender imponer ninguna teoría, se cree que la GCAS que representa las ganancias de copias puede indicar defectos genéticos distintos o adicionalmente a los defectos de reparación del ADN y que la GCAS que representa la pérdida o ganancia del cromosoma entero puede indicar defectos de verificación mitótica o defectos en la segregación de cromosomas, y similares. Se ha determinado que tales aberraciones en la reparación fiel de ADN, la segregación, el control del punto de verificación, etc. son predictivas de las células que albergan tales aberraciones al tratamiento con terapias contra el cáncer (por ejemplo, agentes quimioterapéuticos) en sujetos.

**[0046]** La respuesta a las terapias contra el cáncer se refiere a cualquier respuesta del tumor a la quimioterapia, preferiblemente a un cambio en la masa y/o el volumen del tumor después de la iniciación de la quimioterapia neoadyuvante o adyuvante. La respuesta del tumor puede evaluarse en un entorno neoadyuvante o adyuvante en la que se puede comparar el tamaño de un tumor después de una intervención sistémica con el tamaño y las

dimensiones iniciales medidas por CT, PET, mamografía, ecografía o palpación y se puede estimar la celularidad de un tumor histológicamente y compararse con la celularidad de una biopsia de tumor antes de iniciar el tratamiento. La respuesta también puede evaluarse por medición del calibre o examen patológico del tumor después de la biopsia o resección quirúrgica. La respuesta puede registrarse de forma cuantitativa como el cambio porcentual en el volumen del tumor o la celularidad o utilizando un sistema de puntuación semicuantitativo, tal como la carga de cáncer residual (Symmans et al., J. Clin Oncol (2007) 25: 4.414 a 4.422) o la puntuación de Miller-Payne (Ogston et al, Breast (Edimburgo, Escocia) (2003). 12: 320-327) de una manera cualitativa como la "respuesta patológica completa" (pCR), "remisión clínica completa" (cCR), " remisión clínica parcial "(cPR)," enfermedad clínica estable" (cSD)," enfermedad clínica progresiva" (cPD) u otros criterios cualitativos. La evaluación de la respuesta del tumor puede llevar a cabo poco después de la aparición de la terapia neoadyuvante o adyuvante, por ejemplo, después de unas horas, días, semanas o preferiblemente después de unos meses. Un punto final típico para la evaluación de la respuesta es tras la terminación de la quimioterapia neoadyuvante o tras la extirpación quirúrgica de las células tumorales residuales y/o el lecho del tumor.

**[0047]** Los criterios adicionales para evaluar la respuesta a las terapias contra el cáncer están relacionados con la "supervivencia", que incluye todas las características siguientes: la supervivencia hasta la mortalidad, también conocida como la supervivencia global (en la que dicha mortalidad puede ser independiente de la causa o relacionado con el tumor); "supervivencia libre de recurrencia" (en la que el término recurrencia incluirá tanto la recurrencia localizada como distante); supervivencia libre de metástasis; supervivencia libre de enfermedad (en la que el término enfermedad incluirá el cáncer y las enfermedades asociadas con el mismo). El tiempo de dicha supervivencia puede calcularse en referencia a un punto de inicio definido (por ejemplo, tiempo de diagnóstico o inicio del tratamiento) y un punto final (por ejemplo, la muerte, la recurrencia o metástasis). Además, los criterios de eficacia del tratamiento se pueden ampliar para incluir la respuesta a la quimioterapia, la probabilidad de supervivencia, la probabilidad de metástasis dentro de un período de tiempo determinado, y la probabilidad de recurrencia del tumor.

**[0048]** Por ejemplo, con el fin de determinar los valores umbral apropiados, se puede administrar un régimen terapéutico contra el cáncer particular a una población de sujetos y el resultado puede correlacionarse con la de GCAS que se determinaron antes de la administración de cualquier terapia contra el cáncer. La medición de los resultados puede ser la respuesta patológica al tratamiento proporcionado en el entorno neoadyuvante. Alternativamente, las mediciones del resultado, tales como la supervivencia global y la supervivencia libre de enfermedad, pueden monitorizarse durante un periodo de tiempo en los sujetos después de la terapia contra el cáncer para los que se conocen los valores de GCAS. En ciertas realizaciones, se administran las mismas dosis de agentes contra el cáncer a cada sujeto. En realizaciones relacionadas, las dosis administradas son dosis estándar conocidas en la técnica para agentes contra el cáncer. El período de tiempo durante el cual se supervisan los sujetos puede variar. Por ejemplo, los sujetos se pueden monitorizar durante al menos 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 55, o 60 meses. Los valores umbral de GCAS que se correlacionan con el resultado de una terapia contra el cáncer se pueden determinar usando procedimientos tales como los descritos en la sección de Ejemplos.

### III. Agentes terapéuticos contra el cáncer

**[0049]** La eficacia de las terapias contra el cáncer que dañan el ADN, así como agentes que se aprovechan de los defectos de reparación del ADN, pero no dañan el ADN por sí mismos, tales como inhibidores de la poli ADP ribosa polimerasa (PARP), así como la quimioterapia o terapia de radiación, se prevé según el nivel de GCAS de un cáncer en un sujeto según los procedimientos descritos en el presente documento.

**[0050]** En una realización, se predice la eficacia de la quimioterapia. La quimioterapia incluye la administración de un agente quimioterapéutico. Dicho agente quimioterapéutico puede ser, pero no se limita a, los seleccionados entre los siguientes grupos de compuestos: compuestos de platino, antibióticos citotóxicos, antimetabolitos, agentes antimitóticos, agentes alquilantes, compuestos de arsénico, inhibidores de la ADN topoisomerasa, taxanos, análogos de nucleósidos, alcaloides de plantas, y toxinas; y derivados sintéticos de los mismos. Los compuestos de ejemplo incluyen, pero no se limitan a, agentes alquilantes: cisplatino, treosulfan, y trofosfamida; alcaloides de plantas: vinblastina, paclitaxel, docetaxol; inhibidores de la ADN topoisomerasa: tenipósido, crisnatol, y mitomicina; antifolatos: metotrexato, ácido micofenólico, e hidroxurea; análogos de pirimidina: 5-fluorouracilo, doxifluridina, y arabinósido de citosina; análogos de purina: mercaptopurina y tioguanina; antimetabolitos de ADN: 2'-desoxi-5-fluorouridina, glicinato de afidicolina, y pirazoloimidazol; y agentes antimitóticos: halicondrina, colchicina, y rizoxina. También se pueden utilizar composiciones que comprenden uno o más agentes quimioterapéuticos (por ejemplo, FLAG, CHOP). FLAG comprende fludarabina, arabinósido de citosina (Ara-C) y G-CSF. CHOP comprende ciclofosfamida, vincristina, doxorubicina y prednisona. En otras realizaciones, se utilizan inhibidores de PARP (por ejemplo, PARP-1 y/o PARP-2) y dichos inhibidores son bien conocidos en la técnica (por ejemplo, olaparib, ABT-888, BSI-201, BGP-15 (N-Gene Research Laboratories, Inc.); INO-1001 (Inotek Pharmaceuticals Inc.); PJ34 (Soriano et al, 2001; Pacher et al, 2002b); 3-aminobenzamida (Trevigen); 4-amino-1,8-naftalimida; (Trevigen); 6(5H)-fenantridinona (Trevigen); benzamida (patente de Estados Unidos Re 36.397), y NU1025 (Bowman et al). Los ejemplos anteriores de agentes quimioterapéuticos son ilustrativos, y no se pretende que sean limitativos.

**[0051]** En una realización preferida, los agentes quimioterapéuticos son compuestos de platino, tales como cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino. Otros compuestos de coordinación de platino antineoplásicos son bien conocidos en la técnica, y se pueden modificar según procedimientos bien conocidos en la técnica, e incluyen los compuestos descritos en las patentes de los Estados Unidos. Nos. 4.996.337, 4.946.954, 5.091.521, 5.434.256, 5.527.905, y 5.633.243.

**[0052]** En otra realización, la GCAS predice la eficacia de la radioterapia. La radiación utilizada en radioterapia puede ser radiación ionizante. La radioterapia también puede ser rayos gamma, rayos X o haces de protones. Los ejemplos de radioterapia incluyen, pero no se limitan a, radioterapia de haz externo, implantación intersticial de radioisótopos (I-125, paladio, iridio), radioisótopos tales como el estroncio-89, radioterapia torácica, radioterapia P-32 intraperitoneal, y/o radioterapia abdominal y pélvica total. Para una visión general de la radioterapia, véase Hellman, Capítulo 16: Principles of Cancer Management: Radiation Therapy, 6ª edición, 2001, DeVita et al, eds, J.B. Lippencott Company, Philadelphia. La radioterapia se puede administrar en forma de radiación de haz externo o teleterapia, en la que la radiación se dirige desde una fuente remota. El tratamiento con radiación también se puede administrar como una terapia interna o braquiterapia, en la que se coloca una fuente radiactiva dentro del cuerpo cerca de las células cancerosas o una masa tumoral. También se abarca el uso de terapia fotodinámica que comprende la administración de fotosensibilizadores, tales como hematoporfirina y sus derivados, Vertoporfina (BPD-MA), ftalocianina, fotosensibilizador Pc4, desmetoxi-hipocrellina A; y 2BA-2-DMHA.

**[0053]** Las terapias contra el cáncer que dañan el ADN en menor medida que la quimioterapia o la radioterapia pueden tener eficacia en sujetos que tienen determinaciones de GCAS relativamente más bajas o más altas utilizando los procedimientos de la descripción para determinar el fenotipo de un cáncer. Los ejemplos de dichas terapias incluyen la inmunoterapia, terapia hormonal y terapia génica. Dichas terapias incluyen, pero no se limitan a, el uso de polinucleótidos antisentido, ribozimas, moléculas de ARN de interferencia, polinucleótidos de triple hélice y similares, donde la secuencia de nucleótidos de dichos compuestos está relacionada con las secuencias de nucleótidos de ADN y/o ARN de genes que están relacionados con la iniciación, progresión y/o patología de un tumor o cáncer. Por ejemplo, se pueden utilizar oncogenes, genes de factores de crecimiento, genes de receptores de factores de crecimiento, genes del ciclo celular, genes de reparación del ADN, y otros, en este tipo de terapias.

**[0054]** La inmunoterapia puede comprender, por ejemplo, el uso de vacunas contra el cáncer y/o células presentadoras de antígenos sensibilizadas. La inmunoterapia puede implicar la inmunidad pasiva para la protección a corto plazo de un huésped, lograda mediante la administración de anticuerpo preformado dirigido contra un antígeno de cáncer o antígeno de enfermedad (por ejemplo, la administración de un anticuerpo monoclonal, opcionalmente ligado a un agente quimioterapéutico o una toxina, a un antígeno tumoral). La inmunoterapia también puede centrarse en el uso de los epítomos reconocidos por linfocitos citotóxicos de líneas celulares de cáncer.

**[0055]** Los tratamientos terapéuticos hormonales pueden comprender, por ejemplo, agonistas hormonales, antagonistas hormonales (por ejemplo, flutamida, bicalutamida, tamoxifeno, raloxifeno, acetato de leuprolide (LUPRON), antagonistas de LH-RH), inhibidores de la biosíntesis y procesamiento de hormonas, y esteroides (por ejemplo, dexametasona, retinoides, deltoides, betametasona, cortisol, cortisona, prednisona, deshidrotestosterona, glucocorticoides, mineralocorticoides, estrógeno, testosterona, progestinas), derivados de vitamina A (por ejemplo, ácido retinoico todo trans (ATRA)); análogos de la vitamina D3; antigestágenos (por ejemplo, mifepristona, onapristona), o antiandrógenos (por ejemplo, acetato de ciproterona).

**[0056]** En una realización, la terapia contra el cáncer utilizada para los cánceres cuyo fenotipo se determina mediante los procedimientos de la descripción pueden comprender uno o más tipos de terapias descritas en el presente documento, incluyendo, pero no limitado a, agentes quimioterapéuticos, agentes inmunoterapéuticos, agentes antiangiogénicos, citoquinas, hormonas, anticuerpos, polinucleótidos, radiación y agentes terapéuticos fotodinámicos. Por ejemplo, las terapias de combinación pueden comprender uno o más agentes quimioterapéuticos y radiación, uno o más agentes quimioterapéuticos e inmunoterapia, o uno o más agentes quimioterapéuticos, radiación y quimioterapia.

**[0057]** La duración y/o la dosis de tratamiento con terapias contra el cáncer pueden variar según el agente contra el cáncer particular o combinación de los mismos. El experto en la materia entenderá el tiempo de tratamiento apropiado para un agente terapéutico particular contra el cáncer será apreciado. La descripción contempla la evaluación continua de las pautas de tratamiento óptimas para cada agente terapéutico contra el cáncer, en la que el fenotipo del cáncer del sujeto, tal como se determina por los procedimientos de descripción, es un factor en la determinación de las dosis y pautas de tratamiento óptimas.

IV. Cánceres para los cuales se puede determinar el fenotipo

**[0058]** Los procedimientos descritos en el presente documento se pueden utilizar para determinar el fenotipo de muchos tipos de cánceres diferentes. Los ejemplos específicos de los tipos de cáncer para los que se puede determinar el fenotipo por los procedimientos comprendidos por la descripción incluyen, pero no se limitan a, sarcomas y carcinomas humanos, por ejemplo, fibrosarcoma, mixosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, cordoma, angiosarcoma, endoteliosarcoma, linfangiosarcoma, linfangioendoteliosarcoma, sinovioma,

mesotelioma, tumor de Ewing, leiomiocarcinoma, rhabdomiocarcinoma, carcinoma de colon, cáncer colorrectal, cáncer de páncreas, cáncer de mama, cáncer de ovario, cáncer de próstata, carcinoma de células escamosas, carcinoma de células basales, adenocarcinoma, carcinoma de las glándulas sudoríparas, carcinoma de glándulas sebáceas, carcinoma papilar, adenocarcinomas papilares, cistadenocarcinoma, carcinoma medular, carcinoma broncogénico, carcinoma de células renales, hepatoma, carcinoma del conducto biliar, cáncer de hígado, coriocarcinoma, seminoma, carcinoma embrionario, tumor de Wilms, cáncer cervical, cáncer de hueso, tumor cerebral, cáncer testicular, carcinoma de pulmón, carcinoma de pulmón de células pequeñas, carcinoma de vejiga, carcinoma epitelial, glioma, astrocitoma, meduloblastoma, craneofaringioma, ependimoma, pinealoma, hemangioblastoma, neuroma acústico, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma; leucemias, por ejemplo, leucemia linfocítica aguda y leucemia mielocítica aguda (mieloblástica, promielocítica, mielomonocítica, monocítica y eritroleucemia); leucemia crónica (leucemia mielocítica (granulocítica) crónica y leucemia linfocítica crónica); y policitemia vera, linfoma (enfermedad de Hodgkin y enfermedad no Hodgkin), mieloma múltiple, macroglobulinemia de Waldenström y enfermedad de las cadenas pesadas.

**[0059]** En algunas realizaciones, el cáncer cuyo fenotipo se determina mediante el procedimiento de la descripción es un cáncer epitelial, tal como, pero no limitado a, cáncer de vejiga, cáncer de mama, cáncer cervical, cáncer de colon, cáncer ginecológico, cáncer renal, cáncer de laringe, cáncer de pulmón, cáncer de boca, cáncer de cabeza y cuello, cáncer de ovario, cáncer de páncreas, cáncer de próstata o cáncer de piel. En otras realizaciones, el cáncer es cáncer de mama, cáncer de próstata, cáncer de pulmón o cáncer de colon. En todavía otras realizaciones, el cáncer epitelial es el cáncer no microcítico de pulmón, carcinoma de células renales no papilar, carcinoma cervical, carcinoma de ovario (por ejemplo, carcinoma de ovario seroso), o carcinoma de mama. Los cánceres epiteliales pueden caracterizarse de varias otras maneras incluyendo, pero no limitadas a, seroso, endometriode, mucinoso, célula de clara, Brenner, o indiferenciada.

#### V. Sujetos

**[0060]** En una realización, el sujeto para el cual se determina la eficacia predicha de una terapia contra el cáncer es un mamífero (por ejemplo, ratón, rata, primate, mamífero no humano, animal doméstico, tal como perro, gato, vaca, caballo), y es preferiblemente un ser humano.

**[0061]** En otra realización de los procedimientos de la descripción, el sujeto no se ha sometido a quimioterapia o radioterapia. En realizaciones alternativas, el sujeto se ha sometido a quimioterapia o radioterapia (por ejemplo, tal como con cisplatino, carboplatino, y/o taxano). En realizaciones relacionadas, el sujeto no ha estado expuesto a niveles de radiación o agentes quimiotóxicos por encima de los que se encuentran en general o en un promedio por los sujetos de una especie.

**[0062]** En ciertas realizaciones, el sujeto se ha sometido a cirugía para extirpar el tejido canceroso o precanceroso. En otras realizaciones, el tejido canceroso no se ha extraído, por ejemplo, el tejido canceroso puede estar situado en una región inoperable del cuerpo, tal como en un tejido que es esencial para la vida, o en una región en la que un procedimiento quirúrgico podría causar considerable un riesgo de daño para el paciente.

**[0063]** Según un aspecto de la descripción, la GCAS se puede utilizar para determinar el fenotipo, es decir la capacidad de respuesta a la terapia de un cáncer en un sujeto, donde el sujeto se ha sometido previamente a quimioterapia, radioterapia, o ha estado expuesto a la radiación, o un agente quimiotóxico. Dicha terapia o exposición podrían potencialmente dañar el ADN y alterar el número de SNP heterocigotos informativos en un sujeto. El número alterado de SNP heterocigotos informativos, a su vez, alteraría la GCAS de un sujeto. Dado que las muestras de ADN no cancerosas mostrarían mayor o menor número de SNP heterocigotos, el intervalo de GCAS se alteraría para una población de sujetos. En ciertas realizaciones, el daño del ADN de la terapia o la exposición en un sujeto o población de sujetos se produce alrededor de 1 mes, 2 meses, 3 meses, 4 meses, 5 meses, 6 meses, 7 meses, 8 meses, 9 meses, 10 meses, 11 meses, 1 año, 1,5 años, 2 años o más antes de la determinación de GCAS.

**[0064]** Para determinar los valores umbral de GCAS para los sujetos que presentan daños en el ADN de la terapia o la exposición, se controla una población de sujetos que han recibido quimioterapia o radioterapia, preferiblemente a través de regímenes de tratamiento idénticos o similares, incluyendo dosis y frecuencia, para dichos sujetos.

#### VI. Preparación de la muestra de ácido nucleico

##### A. Aislamiento de ácido nucleico

**[0065]** Las muestras de ácidos nucleicos derivados de células cancerosas y no cancerosas de un sujeto que se pueden utilizar en los procedimientos de la invención para determinar el fenotipo de un cáncer se pueden preparar por medios bien conocidos en la técnica. Por ejemplo, se pueden utilizar procedimientos quirúrgicos o biopsia por aspiración con aguja para recoger muestras cancerosas de un sujeto. En algunas realizaciones, es importante enriquecer y/o purificar el tejido canceroso y/o muestras de células del tejido no canceroso y/o muestras de células. En otras realizaciones, las muestras de tejido y/o de células cancerosas se pueden a continuación microdisecionar para reducir la cantidad de contaminación de tejidos normales antes de la extracción de ácido nucleico genómico o

pre-ARN para su uso en los procedimientos de la invención. En todavía otra realización, las muestras de tejido y/o de células cancerosas están enriquecidas en células cancerosas en por lo menos 50%, 55%, 60%, 65%, 70%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, o cualquier intervalo intermedio, en el contenido de células cancerosas. Dicho enriquecimiento se puede realizar según procedimientos bien conocidos en la técnica, tales como la microdissección con aguja, microdissección por láser, clasificación de células activadas por fluorescencia, y clasificación de células inmunológicas. En una realización, una máquina automática realiza el enriquecimiento de células hiperproliferativas para transformar así la muestra biológica en una forma purificada enriquecida en la presencia de células hiperproliferativas.

[0066] La recogida de muestras de ácidos nucleicos a partir de células no cancerosas de un sujeto también se puede realizar con cirugía o aspiración. En los procedimientos quirúrgicos en los que se extirpa tejido canceroso, los cirujanos a menudo extraen muestras de tejido y/o de células no cancerosas del mismo tipo de tejido del paciente con cáncer para la comparación. Las muestras de ácidos nucleicos pueden aislarse a partir de dicho tejido no canceroso del sujeto para su uso en los procedimientos de la invención.

[0067] En ciertas realizaciones de los procedimientos de la invención, las muestras de ácidos nucleicos de tejidos no cancerosos se no derivan del mismo tipo de tejido que el tejido y/o células cancerosas de la muestra, y/o no se derivan del paciente con cáncer. Las muestras de ácidos nucleicos de tejidos no cancerosos pueden derivar de cualquier tejido y/o células no cancerosas y/o libres de enfermedad. Dichas muestras no cancerosas pueden recogerse mediante procedimientos quirúrgicos o no quirúrgicos. En ciertas realizaciones, las muestras de ácido nucleico no cancerosas derivan de tejidos libres de tumores. Por ejemplo, las muestras no cancerosas se pueden recoger de ganglios linfáticos, linfocitos de sangre periférica y/o células mononucleares de sangre, o cualquier subpoblación de las mismas. En una realización preferida, el tejido no canceroso no es tejido precanceroso, por ejemplo, no presenta ningún indicio de una afección preneoplásica, tal como hiperplasia, metaplasia, o displasia.

[0068] En una realización, las muestras de ácidos nucleicos utilizadas para calcular la GCAS (por ejemplo, el número de SNP heterocigotos en la pluralidad de SNP totales que muestran heterocigosidad en el ADN genómico de tejido no canceroso de la especie a la que pertenece el paciente de cáncer) se toman al menos de 1, 2, 5, 10, 20, 30, 40, 50, 100, o 200 organismos diferentes de esa especie.

[0069] Según ciertos aspectos de la invención, el ácido nucleico "derivado de" ADN genómico, tal como se usa en los procedimientos de la invención, por ejemplo, en experimentos de hibridación para determinar la heterocigosidad de los SNP, puede ser fragmentos de ácido nucleico genómico generado por digestión con enzimas de restricción y/o ligadura a otro ácido nucleico, y/o productos de amplificación de ácidos nucleicos genómicos, o pre-ARN mensajero (pre-ARNm), productos de amplificación de pre-ARNm, o fragmentos de ADN genómico desarrollados en vectores de clonación generados, por ejemplo, mediante procedimientos de clonación "shotgun". En ciertas realizaciones, las muestras de ácido nucleico genómico se digirieron con enzimas de restricción.

#### B. Amplificación de ácidos nucleicos

[0070] Aunque la muestra de ácido nucleico no necesita comprender ácido nucleico amplificado, en algunas realizaciones, los ácidos nucleicos aislados se pueden procesar de maneras que requieren y/o se aprovechan de la amplificación. Las muestras de ADN genómico de un sujeto, opcionalmente, pueden fragmentarse utilizando endonucleasas de restricción y/o se amplifican antes de determinar la GCAS. En una realización, los fragmentos de ADN se amplifican usando la reacción en cadena de la polimerasa (PCR). Los procedimientos para llevar a la práctica la PCR son bien conocidos por los expertos en la técnica. Una ventaja de la PCR es que se pueden utilizar pequeñas cantidades de ADN. Por ejemplo, el ADN genómico de un sujeto puede ser aproximadamente 150 ng, 175 ng, 200 ng, 225 ng, 250 ng, 275 ng, o 300 ng de ADN.

[0071] En ciertas realizaciones de los procedimientos de la invención, el ácido nucleico de un sujeto se amplifica usando un solo par de cebadores. Por ejemplo, se pueden digerir muestras de ADN genómico con endonucleasas de restricción para generar fragmentos de ADN genómico que a continuación se ligan a una secuencia de ADN adaptadora que el par de cebadores reconoce. En otras realizaciones de los procedimientos de la invención, el ácido nucleico de un sujeto se amplifica usando conjuntos de pares de cebadores específicos para loci de interés (por ejemplo, RFLP, STR, SNP, etc.) situados por todo el genoma. Dichos conjuntos de pares de cebadores reconocen cada uno secuencias de ADN genómico que flanquean loci particulares de interés (por ejemplo, SNP, RFLP, STR, etc.). Se puede obtener una muestra de ADN adecuada para la hibridación, por ejemplo, mediante amplificación con la reacción en cadena de la polimerasa (PCR) e ADN genómico, fragmentos de ADN genómico, fragmentos de ADN genómico ligados secuencias adaptadoras o secuencias clonadas. Se pueden utilizar programas informáticos que son bien conocidos en la técnica en el diseño de cebadores con la especificidad deseada y propiedades de amplificación óptimas, tales como Oligo versión 5.0 (National Biosciences). Los procedimientos de PCR son bien conocidos en la técnica, y se describen, por ejemplo, en Innis et al, eds, 1990, PCR Protocols: A guide to Methods and Applications, Academic Press Inc., San Diego, Calif. Será evidente para un experto en la técnica que los sistemas robóticos controlados son útiles para aislar y amplificar los ácidos nucleicos y se pueden utilizar.

**[0072]** En otras realizaciones, cuando el ADN genómico de un sujeto se fragmenta utilizando endonucleasas de restricción y se amplifica antes de determinar la GCAS, la amplificación puede comprender regiones de clonación de ADN genómico del sujeto. En dichos procedimientos, la amplificación de las regiones de ADN se logra mediante el proceso de clonación. Por ejemplo, los vectores de expresión pueden modificarse genéticamente para expresar grandes cantidades de fragmentos particulares de ADN genómico del sujeto (Sambrook, J. et al, eds, 1989, Molecular Cloning: A Laboratory Manual, 2ª Edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, en las págs. 9.47 a 9.51).

**[0073]** En aún otras realizaciones, cuando el ADN de un sujeto se fragmenta utilizando endonucleasas de restricción y se amplifica antes de determinar la GCAS, la amplificación comprende expresar un ácido nucleico que codifica un gen, o un gen y regiones genómicas flanqueantes de ácidos nucleicos, del sujeto. A continuación, el ARN (pre-ARN mensajero) que comprende la totalidad de la transcripción, incluyendo intrones, se aísla y se utiliza en los procedimientos de la descripción para determinar la GCAS y el fenotipo de un cáncer.

**[0074]** En ciertas realizaciones, no se requiere amplificación. En dichas realizaciones, el ADN genómico, o preARN, de un sujeto pueden fragmentarse utilizando endonucleasas de restricción u otros procedimientos. Los fragmentos resultantes se pueden hibridar con sondas de SNP. Habitualmente, se necesitan aislar mayores cantidades de ADN en comparación con la cantidad de ADN o pre-ARN necesaria cuando se amplifican fragmentos. Por ejemplo, cuando no se amplifica el ácido nucleico de un sujeto, una muestra de ADN de un sujeto para su uso en la hibridación puede ser de aproximadamente 400 ng, 500 ng, 600 ng, 700 ng, 800 ng, 900 ng, o 1000 ng de ADN o mayor. Alternativamente, en otras realizaciones, se utilizan procedimientos que requieren cantidades muy pequeñas de ácidos nucleicos para el análisis, tales como menos de 400 ng, 300 ng, 200 ng, 100 ng, 90 ng, 85 ng, 80 ng, 75 ng, 70 ng, 65 ng, 60 ng, 55 ng, 50 ng, o menos, tal como se utiliza para los ensayos de sonda de inversión molecular (MIP). Estas técnicas son particularmente útiles para el análisis de muestras clínicas, tal como material fijado con formalina embebido en parafina o biopsias centrales pequeñas con aguja, caracterizadas por ser de fácil acceso, pero en general por tener una calidad de ADN reducida (por ejemplo, ADN pequeño, fragmentado) y/o no proporcionar grandes cantidades de ácidos nucleicos.

#### C. Hibridación

**[0075]** Las muestras de ácido nucleico derivadas de un sujeto utilizadas en los procedimientos de la invención se pueden hibridar a matrices que comprenden sondas (por ejemplo, sondas de oligonucleótidos) con el fin de identificar loci informativo de interés (por ejemplo, SNP, RFLP, STR, etc.). La hibridación también puede utilizarse para determinar si los loci informativo de interés (por ejemplo, SNP, RFLP, STR, etc.) identificados muestran aberraciones cromosómicas (por ejemplo, desequilibrio alélico, pérdida de heterocigosidad, cambio en el número total de copias, ganancia del número de copias y pérdida del número de copias) en muestras de ácidos nucleicos de tejidos y/o células cancerosas del sujeto. En realizaciones preferidas, las sondas utilizadas en los procedimientos de la invención comprenden una matriz de sondas colocadas en un chip de ADN (por ejemplo, sondas de oligonucleótidos de SNP). En algunas realizaciones, la heterocigosidad de un locus SNP se determina mediante un procedimiento que no comprende la detección de un cambio en el tamaño de fragmentos de ácido nucleico digeridos por enzimas de restricción. En otras realizaciones, se analizan SNP para identificar el desequilibrio alélico.

**[0076]** Las condiciones de hibridación y condiciones de lavado utilizadas en los procedimientos de la invención se eligen de modo que las muestras de ácido nucleico a analizar por la invención se unen específicamente o se hibridan específicamente con las secuencias de oligonucleótidos complementarias de la matriz, preferiblemente a un sitio específico de la matriz, en el que se encuentra su ADN complementario. En algunas realizaciones, el ADN complementario puede aparearse completamente o desaparecerse en algún grado tal como se utiliza, por ejemplo, en matrices de oligonucleótidos Affymetrix, tales como los utilizados para analizar SNP en ensayos MIP.

**[0077]** Las sondas de ADN de ácido oligodesoxirribonucleico sintético de una cadena de una matriz pueden necesitar la desnaturalización antes del contacto con las muestras de ácido nucleico de un sujeto, por ejemplo, para eliminar horquillas o dímeros que forman las correspondientes secuencias complementarias o autocomplementarias.

**[0078]** Las condiciones de hibridación óptimas dependerán de la longitud de las sondas y el tipo de muestras de ácido nucleico de un sujeto. Los parámetros generales para condiciones de hibridación específicas (es decir, rigurosas) para ácidos nucleicos se describen en Sambrook, J. et al, eds, 1989, Molecular Cloning: A Laboratory Manual, Segunda Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, págs 9.47 a 9.51 y de 11.55 a 11.61; Ausubel et al., Eds., 1989, Current Protocols in Molecules Biology, Vol. 1, Green Publishing Associates, Inc., John Wiley & Sons, Inc., Nueva York, págs. 2.10.1-2.10.16. Las condiciones de hibridación útiles de ejemplo se proporcionan en, por ejemplo, Tijessen, 1993, Hybridization with Nucleic Acid Probes, Elsevier Science Publishers B. V. y Kricka 1992, Nonisotopic DNA Probe Techniques, Academic Press, San Diego, California.

#### D. Matrices de ácidos nucleicos oligonucleótidos

**[0079]** En algunas realizaciones de los procedimientos de la presente invención, se pueden utilizar matrices de ADN para determinar si las muestras de ácido nucleico presentan desequilibrio alélico midiendo el nivel de hibridación de

la secuencia de ácidos nucleicos a sondas de oligonucleótidos que comprenden secuencias complementarias. La hibridación se puede utilizar para determinar la presencia o ausencia de heterocigosidad. Varios formatos de matrices de ADN que emplean "sondas" de oligonucleótidos, (es decir, moléculas de ácido nucleico que tienen secuencias definidas) son bien conocidos por los expertos en la técnica.

5 [0080] Habitualmente, se inmoviliza un conjunto de sondas de ácido nucleico, cada una de los cuales, con una secuencia definida, sobre un soporte sólido de manera que cada sonda diferente está inmovilizada a una región predeterminada. En ciertas realizaciones, el conjunto de sondas forma una matriz de sitios de unión (por ejemplo, la hibridación) posicionalmente direccionables sobre un soporte. Cada uno de dichos sitios de unión comprende una pluralidad de moléculas de oligonucleótidos de una sonda unida a la región predeterminada sobre el soporte. Más específicamente, cada sonda de la matriz está situada preferiblemente en una posición predeterminada conocida sobre el soporte sólido, de modo que se puede determinar la identidad (es decir, la secuencia) de cada sonda a partir de su posición en la matriz (es decir, sobre el soporte o superficie). Las micromatrices se pueden fabricar de varias maneras, de las cuales varias están descritas en este documento. Sin embargo, las micromatrices producidas comparten ciertas características, son reproducibles, permiten la producción de múltiples copias de una matriz determinada y se comparan fácilmente con las demás.

[0081] Preferiblemente, las micromatrices se fabrican de materiales que son estables en las condiciones de unión (por ejemplo, hibridación de ácidos nucleicos). Las micromatrices son preferiblemente pequeñas, por ejemplo, entre aproximadamente 1 cm<sup>2</sup> y 25 cm<sup>2</sup>, preferiblemente de aproximadamente 1 a 3 cm<sup>2</sup>. Sin embargo, también se contemplan matrices más grandes y más pequeñas y pueden ser preferibles, por ejemplo, para evaluar simultáneamente un gran número de diferentes sondas.

[0082] Las sondas de oligonucleótidos se pueden sintetizar directamente sobre un soporte para formar la matriz. Las sondas se pueden unir a un soporte sólido o superficie, que puede estar fabricada, por ejemplo, de vidrio, plástico (por ejemplo, polipropileno, nylon), poliacrilamida, nitrocelulosa, gel, u otro material poroso o no poroso. El conjunto de sondas inmovilizadas o la matriz de sondas inmovilizadas se pone en contacto con una muestra que contiene especies de ácidos nucleicos marcados de modo que los ácidos nucleicos que tienen secuencias complementarias a una sonda inmovilizada se hibridan o se unen a la sonda. Después de la separación, por ejemplo, mediante lavado, de cualquier material no unido, las, se detectan y se miden las secuencias marcadas unidas. La medición se efectúa habitualmente con ayuda de un ordenador. Utilizando ensayos de matriz de ADN, se pueden analizar mezclas complejas de ácidos nucleicos marcados, por ejemplo, fragmentos de ácido nucleico derivados de una digestión de restricción de ADN genómico de tejido no canceroso. Las tecnologías de matriz de ADN han hecho posible determinar la heterocigosidad de un gran número de loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) en todo el genoma.

[0083] En ciertas realizaciones, se utilizan matrices de oligonucleótidos de alta densidad en los procedimientos de la invención. Estas matrices que contienen miles de oligonucleótidos complementarios a secuencias definidas, en sitios definidos en una superficie, se pueden sintetizar in situ en la superficie, por ejemplo, mediante técnicas fotolitográficas (véase, por ejemplo, Fodor et al, 1991, Science 251: 767-773; Pease et al, 1994, Proc Natl Acad Sci USA. 91: 5.022 a 5.026; Lockhart et al, 1996, Nature Biotechnology 14: 1675; patentes de Estados Unidos nº 5.578.832; 5.556.752; 5.510.270; 5.445.934; 5.744.305 y 6.040.138). Los procedimientos para generar matrices usando la tecnología de inyección de tinta para la síntesis de oligonucleótidos in situ también son conocidos en la técnica (véase, por ejemplo, Blanchard, publicación de patente internacional WO 98/41531, publicada el 24 de septiembre 1998; Blanchard et al., 1996, Biosensors and Bioelectrónica 11: 687-690; Blanchard, 1998, en Synthetic DNA Arrays in Genetic Engineering, Vol 20, J.K. Setlow, Ed, Plenum Press, Nueva York, en las páginas 111-123). Otro procedimiento para la fijación de los ácidos nucleicos a una superficie es mediante la impresión sobre placas de vidrio, tal como se describe en general por Schena et al. (1995, Science 270: 467-470). También se pueden utilizar otros procedimientos para la fabricación de micromatrices, por ejemplo, mediante el enmascaramiento (Mask y del Sur, 1992, Nucl Acids Res. 20: 1679-1684). Cuando se utilizan estos procedimientos, los oligonucleótidos (por ejemplo, de 15 a 60 unidades) de secuencia conocida se sintetizan directamente sobre una superficie, tal como un portaobjetos de vidrio derivatizado. La matriz producida puede ser redundante, con varias moléculas de oligonucleótidos que corresponden a cada locus informativo de interés (por ejemplo, SNP, RFLP, STR, etc.).

[0084] Un medio de ejemplo para generar las sondas de oligonucleótidos de la matriz de ADN es mediante síntesis de polinucleótidos u oligonucleótidos sintéticos, por ejemplo, utilizando químicas de N-fosfonato o fosforamida (Froehler et al, 1986, Nucleic Acid Res. 14: 5399-5407; McBride et al, 1983, Tetrahedron Lett 24: 246-248). Las secuencias sintéticas tienen habitualmente entre aproximadamente 15 y aproximadamente 600 bases de longitud, más habitualmente entre aproximadamente 20 y aproximadamente 100 bases, lo más preferiblemente entre aproximadamente 40 y aproximadamente 70 bases de longitud. En algunas realizaciones, los ácidos nucleicos sintéticos incluyen bases no naturales, tales como, pero de ninguna manera se limita a, inosina. Como se señaló anteriormente, los análogos de ácido nucleico pueden utilizarse como sitios de unión para la hibridación. Un ejemplo de un análogo de ácido nucleico adecuado es el ácido nucleico peptídico (véase, por ejemplo, Egholm et al, 1993, Nature 363: 566 a 568; patente de Estados Unidos No. 5.539.083). En realizaciones alternativas, los sitios de hibridación (es decir, las sondas) están formados de clones de plásmido o de fago de las regiones de ADN genómico correspondientes a SNP o el complemento de las mismas.

**[0085]** El tamaño de las sondas de oligonucleótidos usadas en los procedimientos de la invención puede ser de al menos 10, 20, 25, 30, 35, 40, 45, o 50 nucleótidos de longitud. Es bien conocido en la técnica que, aunque la hibridación es selectiva para secuencias complementarias, otras secuencias que no son perfectamente complementarios también pueden hibridarse con una sonda determinada en un cierto nivel. Por lo tanto, se pueden utilizar múltiples sondas de oligonucleótidos con ligeras variaciones, para optimizar la hibridación de muestras. Para optimizar adicionalmente la hibridación, se pueden alterar las condiciones de rigurosidad de la hibridación, por ejemplo, la temperatura de hibridación y las concentraciones de sal, mediante procedimientos que son bien conocidos en la técnica.

**[0086]** En realizaciones preferidas, las matrices de oligonucleótidos de alta densidad usadas en los procedimientos de la invención comprenden oligonucleótidos correspondientes a loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.). Las sondas de oligonucleótidos pueden comprender ADN o "miméticos" de ADN (por ejemplo, derivados y análogos) que corresponden a una parte de cada locus informativo de interés (por ejemplo, SNP, RFLP, STR, etc.) en el genoma de un sujeto. Las sondas de oligonucleótidos se pueden modificar en el resto de base, en el resto de azúcar, o en la cadena principal de fosfato. Ejemplos de miméticos de ADN incluyen, por ejemplo, fosforotioatos. Para cada locus de SNP, se puede utilizar una pluralidad de diferentes oligonucleótidos que son complementarios a las secuencias de los ácidos nucleicos de la muestra. Por ejemplo, para un único locus informativo de interés (por ejemplo, SNP, RFLP, STR, etc.) se pueden utilizar aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, o más oligonucleótidos diferentes. Cada uno de los oligonucleótidos para un locus informativo particular de interés puede tener una ligera variación en apareamientos perfectos, desapareamientos, y la secuencia de flanqueo alrededor del SNP. En ciertas realizaciones, las sondas se generan de manera que las sondas para un locus informativo particular de interés comprenden secuencias solapantes y/o solapantes sucesivas que abarcan o se sitúan a lo largo de una región genómica que contiene el sitio diana, donde todas las sondas contienen el sitio diana. A modo de ejemplo, las secuencias de sondas solapantes se pueden colocar en las etapas de intervalos de bases predeterminadas, por ejemplo en las etapas de 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 intervalos de bases.

**[0087]** En ciertas realizaciones, los ensayos se pueden realizar utilizando matrices adecuadas para su uso con protocolos de sonda de inversión molecular, tales como los descritos por Wang et al. (2007) Genome Biol. 8, R246.

**[0088]** Para las sondas de oligonucleótidos dirigidas a las especies de ácidos nucleicos de secuencias estrechamente parecidas (es decir, homólogas), la "hibridación cruzada" entre sondas similares puede contaminar de manera significativa y confundir los resultados de las mediciones de la hibridación. Hibridación cruzada es una preocupación especialmente significativa en la detección de SNP, ya que la secuencia a detectar (es decir, el SNP en particular) debe distinguirse de otras secuencias que difieren sólo en un único nucleótido. La hibridación cruzada puede minimizarse mediante la regulación de las condiciones rigurosas de hibridación y/o durante los lavados post-hibridación. Las condiciones altamente rigurosas permiten la detección de variantes alélicas de una secuencia de nucleótidos, por ejemplo, aproximadamente 1 desapareamiento por cada 10-30 nucleótidos.

**[0089]** No hay una única condición de hibridación o de lavado que sea óptima para todas las diferentes secuencias de ácidos nucleicos. Para determinadas matrices de loci informativo de interés, estas condiciones pueden ser idénticas a las sugeridas por el fabricante o pueden ajustarse por un experto en la técnica.

**[0090]** En realizaciones preferidas, las sondas utilizadas en los procedimientos de la invención se inmovilizan (es decir, se colocan) en un portaobjetos de vidrio llamado chip. Por ejemplo, una micromatriz de ADN puede comprender un chip en el que los oligonucleótidos (secuencias de ADN de cadena simple purificadas en solución) han sido impresas por un robot en una matriz (aproximadamente) rectangular, donde cada punto de la matriz corresponde a una única muestra de ADN que codifica un oligonucleótido. En resumen, el procedimiento comprende, inundar el chip de micromatrices de ADN con una muestra marcada en condiciones adecuadas para que se produzca la hibridación entre las secuencias del portaobjetos y la muestra marcada, a continuación la matriz se lava y se seca, y la matriz se escanea con un microscopio láser para detectar la hibridación. En ciertas realizaciones hay por lo menos 250, 500, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 21.000, 22.000, 23.000, 24.000, 25.000, 26.000, 27.000, 28.000, 29.000, 30.000, 31.000, 32.000, 33.000, 34.000, 35.000, 36.000, 37.000, 38.000, 39.000, 40.000, 41.000, 42.000, 43.000, 44.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000 o más, o cualquier intervalo intermedio, de loci informativos de interés para los que aparecen sondas en la matriz (con sondas con desapareamientos/apareamientos para un único locus de interés o sondas colocadas a lo largo de un solo locus que cuenta como un locus de interés). El número máximo de loci informativo de interés que se sonda por matriz se determina por el tamaño del genoma y la diversidad genética de la especie del sujeto. Los chips de ADN son bien conocidos en la técnica y se pueden adquirir en forma prefabricada con secuencias específicas para especies particulares. En algunas realizaciones, en los procedimientos de la invención se utilizan las matrices del Genome-Wide Human SNP Array 6,0® y/o 50K Xbal (Affymetrix, Santa Clara, Calif.). En otras realizaciones, pueden detectarse y cuantificarse SNP y/o número de copias de ADN usando procedimientos de secuenciación, tales como "procedimientos de secuenciación de próxima generación", tal como se describe anteriormente.

## E. Detección de la señal

**[0091]** En algunas realizaciones, las muestras de ácido nucleico derivadas de un sujeto se hibridan con los sitios de unión de una matriz descrita en este documento. En ciertas realizaciones, las muestras de ácido nucleico derivadas de cada uno de los dos tipos de muestras de un sujeto (es decir, cancerosas y no cancerosas) se hibridan con matrices separadas, aunque idénticas. En ciertas realizaciones, las muestras de ácido nucleico derivadas de uno de los dos tipos de muestras de un sujeto (es decir, cancerosas y no cancerosas) se hibridan a dicha matriz, a continuación, después de la detección de la señal, el chip se lava para eliminar la primera muestra marcada y se vuelven a utilizar para hibridar la muestra restante. En otras realizaciones, la matriz no se vuelve a utilizar más de una vez. En ciertas realizaciones, las muestras de ácidos nucleicos derivadas de cada uno de los dos tipos de muestras de un sujeto (es decir, cancerosas y no cancerosas) se marcan de forma diferente de modo que puedan distinguirse. Cuando las dos muestras se mezclan y se hibridan a la misma matriz, la intensidad relativa de la señal de cada muestra se determina para cada sitio en la matriz, y se detecta cualquier diferencia relativa en la abundancia de un alelo de loci informativos de interés.

**[0092]** Las señales se pueden registrar y, en algunas realizaciones, analizarse por ordenador. En una realización, en una imagen escaneada se extraen los puntos usando un programa de gráficos (por ejemplo, HiJaak Graphics Suite) y a continuación se analizan usando un programa de cuadrícula de imágenes que crea una hoja de cálculo de la hibridación promedio en cada longitud de onda en cada sitio. Si es necesario, se puede realizar una corrección determinada experimentalmente para la "intercomunicación" (o superposición) entre los canales para los dos compuestos fluorescentes. Para cualquier sitio de hibridación particular en la matriz, se puede calcular una relación de la emisión de los dos fluoróforos, lo cual puede ayudar en la eliminación de las señales de hibridación cruzadas para determinar con mayor precisión si un locus de SNP particular es homocigoto o heterocigoto.

## F. Marcaje

**[0093]** En algunas realizaciones, las muestras de ácidos nucleicos, fragmentos de los mismos, o fragmentos de los mismos ligados a regiones adaptadoras utilizadas en los procedimientos de la invención se marcan de forma detectable. Por ejemplo, el marcador detectable puede ser un marcador fluorescente, por ejemplo, mediante la incorporación de análogos de nucleótidos. Otros marcadores adecuados para su uso en la presente invención incluyen, pero no se limitan a, biotina, iminobiotina, antígenos, cofactores, dinitrofenol, ácido lipoico, compuestos olefinicos, polipéptidos detectables, moléculas ricas de electrones, enzimas capaces de generar una señal detectable por acción sobre un sustrato, e isótopos radiactivos.

**[0094]** Los isótopos radiactivos incluyen los que se pueden utilizar conjuntamente con los procedimientos de la invención, pero no se limitan a,  $^{32}\text{P}$  y  $^{14}\text{C}$ . Las moléculas fluorescentes adecuadas para la presente invención incluyen, pero no se limitan a, fluoresceína y sus derivados, rodamina y sus derivados, rojo texas, 5'-carboxi-fluoresceína ("FAM"), 2',7'-dimetoxi-4',5'-dicloro-6-carboxi-fluoresceína ("JOE"), N,N,N',N'-tetrametil-6-carboxi-rodamina ("TAM-RA"), 6-carboxi-X-rodamina ("ROX"), HEX, TET, IRD40 e IRD41.

**[0095]** Las moléculas fluorescentes que son adecuados para usar según la invención incluyen además: colorantes de cianina, incluyendo pero no limitado a Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7 y FLUORX; colorantes BODIPY incluyendo, pero no limitado a, BODIPY-FL, BODIPY-TR, BODIPY-TMR, BODIPY-630/650, y BODIPY-650/670; y colorantes ALEXA, incluyendo pero no limitado a, ALEXA-488, ALEXA -532, ALEXA -546, ALEXA -568 y ALEXA -594; así como otros colorantes fluorescentes, que serán conocidos para los expertos en la técnica. Entre las moléculas indicadoras ricas en electrones adecuadas para la presente invención se incluyen, pero no se limitan a, ferritina, hemocianina, y oro coloidal.

**[0096]** También se pueden utilizar esquemas de marcaje y detección con fluorescencia de dos colores (Shena y otros, 1995, Science 270: 467-470). El uso de dos o más marcadores puede ser útil en la detección de variaciones debido a las diferencias menores en las condiciones experimentales (por ejemplo, condiciones de hibridación). En algunas realizaciones de la invención, se pueden utilizar al menos 5, 10, 20, o 100 colorantes de diferentes colores para el marcaje. Dicho marcaje también permitiría el análisis de múltiples muestras simultáneamente, lo cual está comprendido por la invención.

**[0097]** Las muestras de ácidos nucleicos marcados, fragmentos de los mismos, o fragmentos de los mismos ligados a regiones adaptadoras que se pueden utilizar en los procedimientos de la invención se ponen en contacto con una pluralidad de sondas de oligonucleótidos en condiciones que permiten a los ácidos nucleicos de muestra que tiene secuencias complementarias a las sondas hibridarse a las mismas.

**[0098]** Dependiendo del tipo de marcador usado, las señales de hibridación se pueden detectar usando procedimientos bien conocidos por los expertos en la técnica incluyendo, pero no limitado a, película de rayos X, obtención de imágenes fosforescentes o cámara CCD. Cuando se utilizan sondas marcadas con fluorescencia, las emisiones de fluorescencia en cada sitio de una matriz de transcripción pueden detectarse, preferiblemente, mediante microscopía confocal láser de barrido. En una realización, se lleva a cabo un barrido separado, utilizando

la línea de excitación apropiada, para cada uno de los dos fluoróforos utilizados. Alternativamente, puede usarse un láser que permita la iluminación de muestras simultáneas a longitudes de onda específicas para los dos fluoróforos y se pueden analizar simultáneamente las emisiones de los dos fluoróforos (ver Shalon et al. (1996) Genome Res. 6,639-645). En una realización preferida, las matrices se escanean con un escáner de fluorescencia láser con una etapa X-Y controlada por ordenador y un objetivo de microscopio. La excitación secuencial de los dos fluoróforos se consigue con un láser de gas mixto multilínea y la luz emitida se divide por la longitud de onda y se detecta con dos tubos fotomultiplicadores. Dichos dispositivos de barrido por láser de fluorescencia se describen, por ejemplo, en Schena et al. (1996) Genome Res. 6, 639-645. Alternativamente, se puede utilizar un haz de fibra óptica, tal como el descrito por Ferguson et al. (1996) Nat. Biotech. 14, 1681-1684. Las señales resultantes se pueden analizar a continuación para determinar la presencia o ausencia de heterocigosidad o homocigosidad para los loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) utilizando software informático.

G. Algoritmos para el análisis de loci informativos de interés

**[0099]** Una vez que se ha detectado la señal de hibridación, los datos resultantes se pueden analizar utilizando algoritmos. En ciertas realizaciones, el algoritmo para determinar la heterocigosidad en loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) se basa en procedimientos bien conocidos que requieren desequilibrio alélico (AI), pérdida de heterocigosidad (LOH), aberraciones en el número de copias (CNA), ganancia del número de copias (CNG), y disminución del número de copias (CND). Por ejemplo, AI se puede determinar utilizando la proporción mayoritaria de copias (MCP) a la que se requiere AI para un determinado SNP, cuando el valor de MCP es mayor que 0,60, 0,61, 0,62, 0,63, 0,64, 0,65, 0,66, 0,67, 0,68, 0,69, 0,70, 0,71, 0,72, 0,73, 0,74, 0,75, 0,76, 0,77, 0,78, 0,79, 0,80, 0,81, 0,82, 0,83, 0,84, 0,85, 0,86, 0,87, 0,88, 0,89, 0,90, 0,91, 0,92, 0,92, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98, o 0,99. Una vez que se determina el requerimiento, se pueden aplicar adicionalmente procedimientos de enumeración. Por ejemplo, se puede determinar la GCAS, por ejemplo, mediante: 1) el recuento del número total de SNP afectados por el AI o la ganancia de copias o LOH, 2) el recuento del número de regiones afectadas por AI (por ejemplo,  $N_{AI}$ , tal como describe adicionalmente en los Ejemplos, una región individual se define como una cadena de SNP vecinos que muestran todos AI limitado en al menos un lado por SNPs que no muestran AI/retención de heterocigosidad. El tamaño de la región se define por la longitud del cromosoma representado por la cadena de SNP con AI); 3) el recuento del número de cromosomas con la pérdida de todo el cromosoma, o 4) el recuento del número de regiones cromosómicas con la CNA, CNG, CND, etc. Ilustraciones representativas adicionales de dichos algoritmos conocidos se proporcionan en la sección de ejemplos a continuación.

H. Sistemas y procedimientos de implementación por ordenador

**[0100]** En ciertas realizaciones, los procedimientos de la invención implementan un programa informático para calcular la GCAS. Por ejemplo, se puede utilizar un programa informático para realizar los algoritmos descritos en este documento. Un sistema informático también puede almacenar y manipular los datos generados por los procedimientos de la presente invención, que comprende una pluralidad de cambios/perfiles de la señal de hibridación durante la aproximación al equilibrio en diferentes mediciones de hibridación y que pueden ser utilizados por un sistema informático en la implementación de los procedimientos de esta invención. En ciertas realizaciones, un sistema informático recibe datos de la hibridación de las sondas; (i) almacena datos de la hibridación de las sondas; y (iii) compara datos de la hibridación de las sondas para determinar el estado de los loci informativos de interés en dicha muestra de ácido nucleico de tejido canceroso o precanceroso. A continuación, se calcula la GCAS. En otras realizaciones, un sistema informático (i) compara la GCAS determinada con un valor umbral; y (ii) da salida a una indicación de si dicha GCAS está por encima o por debajo de un valor umbral, o un fenotipo en base a dicha indicación. Dichos sistemas informáticos también se consideran parte de la presente descripción.

**[0101]** Se pueden utilizar numerosos tipos de sistemas informáticos para implementar los procedimientos analíticos de esta invención según el conocimiento que tiene un experto en bioinformática y/o arte digital.

**[0102]** Se pueden cargar varios componentes de software en la memoria durante el funcionamiento de dicho sistema informático. Los componentes de software pueden comprender componentes de software que son estándar en la técnica y componentes que son especiales para la presente invención (por ejemplo, software dCHIP descrito en Lin et al (2004) Bioinformatics 20, 1233-1240; software CRLMM descrito en Silver et al., (2007) Cell 128, 991-1002; software Aroma Affymetrix descrito en Richardson et al (2006) Cancer Cell 9, 121-132. Los procedimientos de la invención también pueden programarse o modelarse en paquetes de software matemáticos que permiten la entrada simbólica de ecuaciones y la especificación de alto nivel de procesamiento, incluyendo algoritmos específicos a utilizar, liberando así al usuario de la necesidad de programa procesalmente ecuaciones y algoritmos individuales. Tales paquetes incluyen, por ejemplo, Matlab de Mathworks (Natick, Mass.), Mathematica de Wolfram Research (Champaign, Illinois.) o S-Plus de MathSoft (Seattle, Wash.).

**[0103]** En ciertas realizaciones, el ordenador comprende una base de datos para el almacenamiento de los perfiles de la señal de hibridación. Dichos perfiles almacenados pueden ser accesibles y utilizarse para calcular SAAG. Por ejemplo, el perfil de la señal de hibridación de una muestra derivada de tejido no canceroso de un sujeto y/o perfiles generados a partir de distribuciones basadas en la población de loci informativos de interés en las poblaciones

relevantes de la misma especie almacenados, podrían entonces compararse con el perfil de la señal de hibridación de una muestra derivada de tejido canceroso del sujeto.

5 [0104] Además de las estructuras de los programas de ejemplo y los sistemas informáticos descritos en el presente documento, otras estructuras de programas alternativas y sistemas informáticos serán fácilmente evidentes para el experto en la materia. Dichos sistemas alternativos, que no se apartan del sistema informático y las estructuras de programas descritos anteriormente ni en el espíritu ni en su alcance, por lo tanto, se pretende que estén comprendidas dentro de las reivindicaciones adjuntas.

10 [0105] Otras realizaciones de la presente invención se describen en los siguientes Ejemplos. La presente descripción se ilustra adicionalmente mediante los siguientes ejemplos que no deben interpretarse como limitantes adicionales.

## 15 EJEMPLOS

### Ejemplo 1: Materiales y Procedimientos para el Ejemplo 2

[0106] Se midió la respuesta patológica después de la terapia neoadyuvante con cisplatino en la cohorte de TNBC utilizando la escala semicuantitativa de Miller-Payne tal como se ha descrito (Silver et al (2010) J. Clin Oncol 28, 1145-1153; Ogston et al. (2003) Breast 12, 320-327). Se realizó el genotipado de MIP tal como se ha descrito (Wang et al. (2007) Genome Biol. 8, R246). Se procesaron la intensidad de señal de alelos y los genotipos de MIP o los análisis de matrices de SNP públicos por el algoritmo CRLMM (Lin et al. (2008) Genome Biol. 9, R63) tal como se aplica en el "oligo" del paquete R. Se determinó el número de copias de ADN usando el paquete R "AromaAffymetrix" (Bengtsson et al. (2008) Bioinformatics 24, 759-767). Los datos del genotipo procesados se exportaron a dChip (disponible en la World Wide Web en <http://biosun1.harvard.edu/complab/dchip/>) para la determinación de la proporción mayoritaria de copias (MCP), que se define como la relación del número de copias mayoritario con respecto al número de copias mayoritario + menor (Li et al. (2008) Bioinformatics 9, 204). Se realizó una estimación del nivel de contaminación de ADN normal a partir de la curva de MCP genómica tal como se ha descrito (Li et al. (2008) Bioinformatics 9, 204). En los análisis se incluyeron los casos de mama u ovario que se estimó que tenía un 75% o más de contenido tumoral. El desequilibrio alélico (AI) para fines específicos de los ejemplos descritos en el presente documento se define como  $MCP > 0,7$  y las regiones de AI definidas como más de 10 sondas consecutivas con AI. El AI telomérico para fines específicos de los ejemplos descritos en el presente documento se define como regiones de AI que se extienden hasta el telómero y no cruzan el centrómero. Se estimó la asociación entre  $N_{tAI,12}$  y la respuesta a cisplatino en sujetos con TNBC mediante el área bajo la curva (AUC) de la curva característica de operación del receptor (ROC); valor p es de la prueba de rangos de Wilcoxon de dos colas. Se estimó la asociación entre AI telomérico y el tiempo hasta la recurrencia del cáncer de ovario después de la terapia con platino mediante el análisis de Kaplan-Meier utilizando un punto de corte de 13 para definir un grupo de  $t_{tAI,12}$  alto; valor p se basa en la prueba de rangos logarítmicos. Una lista completa de los materiales y procedimientos es la siguiente:

#### 40 A. Líneas celulares y ensayos de sensibilidad a fármacos

[0107] Se mantuvieron las líneas celulares de cáncer de mama triple negativa BT20, BT549, HCC1187, HCC38, MDA-MB231 y MDA-MB468 37°C con 5% de CO<sub>2</sub> en medio RPMI 1640 y/o medio MEM suplementado con FBS al 10% u otros suplementos tal como se recomienda por la ATCC para cada línea celular. Para probar la sensibilidad a los fármacos, las células se expusieron a una serie de concentraciones de cisplatino durante 48 horas. Se cuantificó el número de células viables utilizando el ensayo CellTiter 96 Aqueous One Solution Cell Proliferation Assay según las instrucciones del fabricante (Promega). Los resultados se presentan como el porcentaje de células viables en los pocillos tratados con fármaco frente a los pocillos de control tratados con medio y se representan como curvas de supervivencia celular dependiente de la dosis del fármaco (Figura 1A). La sensibilidad a los fármacos se cuantificó como la dosis de fármaco que causa una reducción del 50% del crecimiento (IC<sub>50</sub>). Estos datos se generaron originalmente para un estudio separado en el que se informó como "datos no presentados" en Li et al. (2010) Nat. Medicina. 16, 214-218.

#### 55 B. Cohorte de cáncer de mama

[0108] Se trataron un total de 28 pacientes con TNBC principalmente esporádico con monoterapia de cisplatino en un entorno neoadyuvante (Silver et al. (2010) J. Clin. Oncol. 28, 1145-1153). La respuesta del cisplatino se midió utilizando la puntuación semicuantitativa de Miller-Payne mediante la evaluación patológica de muestras quirúrgicas después de la terapia (Ogston et al. (2003) Breast 12, 320-327). La respuesta patológica completa es equivalente a la puntuación de Miller-Payne de 5 y se define como un carcinoma no invasivo residual de mama o ganglios linfáticos.

#### 65 C. Preparación de las muestras de cáncer de mama

**[0109]** Se obtuvo una biopsia central congelada del tumor antes de iniciar el tratamiento. El tejido tumoral estaba disponible en la biopsia central congelada para 24 de los 28 casos y en muestras de biopsias centrales diagnósticas embebidas en parafina fijadas con formalina de otros 3 casos adicionales. Las células tumorales se enriquecieron mediante microdissección con aguja para extraer el estroma de secciones de tejido teñidos con hematoxilina y eosina (H y E). El tejido restante en el portaobjetos se examinó por microscopía para la estimación del enriquecimiento. Se extrajo el ADN de células tumorales enriquecidas mediante digestiones con proteinasa K y ARNasa A, extracción con fenol/cloroformo seguido de precipitación con etanol. Se obtuvo ADN adecuado para el análisis del genotipado de MIP (mínimo 80 ng) de los 27 casos para los que se disponía de tejido tumoral. Se obtuvo el ADN normal emparejado de cada paciente de linfocitos de sangre periférica.

#### D. Análisis de genotipado de la sonda de inversión molecular (MIP)

**[0110]** Se envió ADN de muestras de biopsia de tumor de mama a Affymetrix, Inc. (Santa Clara, CA) para el análisis de genotipado dirigido a MIP que generó una intensidad de la señal de alelo y genotipos para 42.000 polimorfismos de nucleótido único (SNP) individuales. El conjunto de datos del genotipo completo de MIP está disponible en la base de datos NCBI GEO.

#### E. Conjuntos de datos públicos

**[0111]** Se adquirieron los perfiles genómicos con Affymetrix SNP 6.0 de seis líneas celulares triple negativas de cáncer de mama, BT20, BT549, HCC1187, HCC38, MDA-MB231 y MDA-MB468, del Wellcome Trust Sanger Institute (información disponible en la World Wide Web en <http://www.sanger.ac.uk/>).

**[0112]** Los datos de SNP que representan 118 tumores de carcinoma de ovario dispuestos en la plataforma Affymetrix 50K XbaI se adquirieron de los ómnibus de expresión génica (GEO, GSE13813; Etemadmoghadam et al (2009) Clin Cancer Res 15, 1417-1427.). De estos, 38 tumores eran del subtipo seroso, tenía tumor residual después de la citorreducción quirúrgica de menos de 1 cm, y habían recibido cisplatino adyuvante o tratamiento con carboplatino. La mayoría de los pacientes (35 de 38) también habían recibido tratamiento con taxanos.

#### F. Genotipo y análisis del número de copias

**[0113]** Se procesaron la intensidad de señal de alelos y los genotipos de del genotipado de MIP o los análisis de matrices de SNP mediante el algoritmo CRLMM (Lin et al. (2008) Genome Biol. 9, R63) tal como se aplica en el "oligo" del paquete R. Se determinó el número de copias de ADN usando el paquete R "AromaAffymetrix" (Bengtsson et al. (2008) Bioinformatics 24, 759-767). Los datos del genotipo procesados se exportaron a dChip (disponible en la World Wide Web en <http://biosun1.harvard.edu/complab/dchip/>) para la determinación de la proporción mayoritaria de copias (MCP), que se define como la relación del número de copias mayoritario con respecto al número de copias mayoritario + menor (Li et al. (2008) Bioinformatics 9, 204). Se realizó una estimación del grado de contaminación de células normales mediante el grado de desplazamiento en la curva de MCP de la mayoría de regiones que muestran desequilibrio alélico a lo largo del genoma, con exclusión de todas las regiones de ganancia del número de copias (Figuras 2B y 2C). El desplazamiento observado en las curvas de MCP genómicas en experimentos de mezcla de líneas celulares tumorales y normales emparejadas se utilizó como referencia para estimar la contaminación normal, tal como se ha descrito (Waddell et al (2009) Breast Cancer Res Treat (4 Dec; publicadao online)). Por consiguiente, 21 de las 27 muestras de tumores de mama y 33 de 38 de los casos de cáncer de ovario se estimó que tenían un 25% o menos de la contaminación de ADN normal (contenido de tumor  $\geq$  75%) y se consideraron aceptables para su posterior análisis.

**[0114]** El desequilibrio alélico (AI) se definió para los propósitos de los ejemplos descritos en el presente documento como  $MCP > 0,70$ , lo que permite la detección de la mayoría de sucesos de pérdidas de heterocigosidad (LOH) y de amplificaciones monoalélicas de copias elevadas en muestras con 25% o menos de contaminación o heterogeneidad, pero que también excluye las ganancias de copia de bajo nivel (ganancias de 4 copias o menos). Las regiones de AI se definieron para los propósitos de los ejemplos descritos en el presente documento como más de 10 sondas consecutivas que muestran AI. En el conjunto de datos de TNBC, las regiones de AI definidas por estos criterios incluían todas las regiones LOH requeribles determinadas a partir de la comparación de genotipo convencional. El número total de copias (la combinación de ambos alelos) se segmentó mediante el algoritmo de segmentación binaria circular. El ochenta y cinco por ciento de las regiones AI tenía un número de copias total casi diploide o por debajo, el 9% de las regiones de AI mostró ganancia de copias totales de 3 y un 6% con una ganancia de copias totales  $\geq$  4. Por lo tanto, las regiones de AI identificadas representan predominantemente LOH o delección cromosómica uniparental.

#### G. Asociación entre el número de aberraciones genómicas y la sensibilidad al platino in vitro

**[0115]** El número de regiones de AI o regiones con aberración en el número de copias se compararon con valores  $IC_{50}$  específicos de líneas celulares después de aplicar un filtro de 1 Mb de tamaño mínimo para eliminar regiones muy pequeñas que podrían ser causadas por el ruido en los datos de SNP 6.0 (Figura 3). Para la comparación de las regiones de AI telomérico e intersticial, el AI telomérico se definió para los propósitos de los ejemplos descritos

en el presente documento como el AI que se extiende hasta el telómero, pero que no cruza el centrómero. Por el contrario, el AI intersticial se definió para los propósitos de los ejemplos descritos en el presente documento como regiones de AI que no implican el telómero. Para investigar si había un tamaño mínimo óptimo de AI telomérico o segmentos de alteración del número de copias que mostraban una correlación superior a la IC<sub>50</sub> del cisplatino, se utilizó la regresión lineal para comparar los valores de IC<sub>50</sub> con el número total de segmentos más grandes que un cierto umbral, que se incrementó en intervalos de 1 Mb entre 0 y 100 Mb (figura 5).

H. Asociación entre el número de regiones de AI telomérico y la sensibilidad al platino en tumores

**[0116]** Se determinó el número total de regiones de AI telomérico para cada caso de TNBC con un contenido de tumor de al menos un 75%. Se aplicó el umbral del tamaño de segmento mínimo óptimo de AI telomérico de 12 Mb hallado en las líneas celulares y se contaron N<sub>tAI,12</sub> para cada sujeto. Se realizó un análisis de la curva ROC (Receiver Operating Characteristic) para evaluar la capacidad del número total de segmentos de AI telomérico para predecir la pCR (puntuación de Miller-Payne de 5) para el tratamiento con cisplatino. La asociación de N<sub>tAI,12</sub> con pCR al cisplatino se estimó mediante el área bajo la curva (AUC); el valor p correspondiente es la prueba de rangos de Wilcoxon de dos colas. Basándose en el análisis ROC, un N<sub>tAI,12</sub> de 13 dio lugar a una sensibilidad del 100% para la predicción de pCR en la cohorte con TNBC tratada con cisplatino.

**[0117]** La asociación entre N<sub>tAI,12</sub> y el tiempo hasta la recurrencia después de la terapia basada en platino en la cohorte de cáncer de ovario se estimó mediante el análisis de Kaplan-Meier con el grupo de "N<sub>tAI,12</sub> alto" definido como al menos 13 regiones de N<sub>tAI,12</sub>. El valor p se basa en una prueba de rangos logarítmicos.

**Ejemplo 2: El número total de reordenamientos cromosómicos es predictivo de la sensibilidad a fármacos quimioterapéuticos**

**[0118]** Sin pretender imponer ninguna teoría, se cree que la pérdida intracromosómica heterocigidad (LOH) o desequilibrio alélico (AI) resulta de una reparación incorrecta de las roturas de las dobles cadenas de ADN cromosómico que el recuento en todo el genoma de estos reordenamientos cromosómicos en un tumor específico puede indicar el grado de incompetencia en la reparación del ADN, independientemente del defecto específico de reparación del ADN causante. Por lo tanto, el número total de reordenamientos cromosómicos en un tumor refleja la incapacidad para reparar el daño del ADN inducido por fármacos, tales como cisplatino, y por lo tanto, predice la sensibilidad a estos agentes. La sensibilidad al cisplatino de seis líneas celulares de TNBC para las que los datos de matrices de SNP estaban disponibles en Wellcome Trust Sanger Institute, Reino Unido, se determinó de esta manera (Figura 1A). AI se determinó por análisis de la proporción mayoritaria de copias (MCP), un procedimiento menos sensible a la contaminación normal en muestras de tumores heterogéneos (Li et al. (2008) Bioinformatics 9, 204). La MCP es el número de alelos de copia mayoritaria en un locus dividido por la suma de los alelos de copia mayoritaria más menores (Figura 2). Las ganancias o reducciones en el número total de copias de ADN en cada región cromosómica se dedujeron utilizando software dChip (Lin et al. (2004) Bioinformatics 20, 1233-1240).

**[0119]** La lesión o lesiones de reparación del ADN que hacen las células sensibles al cisplatino pueden inducir preferencialmente alteraciones cromosómicas de un tipo específico o con un intervalo de tamaños específicos. En las seis líneas celulares, se analizó la asociación entre la sensibilidad al cisplatino y cada una de las cuatro mediciones de alteraciones cromosómicas. Las cuatro mediciones fueron (1) el número de regiones cromosómicas con AI (N<sub>AI</sub>), (2) el número de aberraciones en el número de copias (N<sub>CNA</sub>), (3) el número de regiones con ganancia del número de copias, y (4) el número de regiones con disminución del número de copias (figura 3). De estas cuatro mediciones, la N<sub>AI</sub> se correlacionó más fuertemente con la sensibilidad a cisplatino (R<sup>2</sup> = 0,5).

**[0120]** Los defectos conocidos en la reparación de roturas del ADN de doble cadena, incluyendo la pérdida de BRCA1 o mutaciones en la helicasa Bloom, causa la formación espontánea de estructuras cromosómicas triradiales y cuadriradiales, que son indicios citológicos de recombinación aberrante (Silver et al. (2007) Cell 128, 991-1002; Luo et al (2000) Nat Genet 26, 424-429; Xu et al (1999) Mol Cell 3, 389-395). La resolución de estos reordenamientos cromosómicos en la mitosis puede dar lugar a la pérdida de fragmentos cromosómicos distales (teloméricos) y grandes regiones de AI (Luo et al (2000) Nat Genet 26, 424-429; Vrieling (2001) Nat Genet. 28, 101-102). Por lo tanto, las regiones de AI telomérico e intersticial (no telomérico) se compararon y se encontró que la correlación entre la sensibilidad al cisplatino y el AI era más fuerte cuando se limita a regiones de AI que implican telómeros, mientras que sólo se observó una asociación débil entre la sensibilidad a cisplatino y el número de regiones de AI intersticial (figura 4).

**[0121]** A continuación, se determinó si las correlaciones podrían mejorarse entre la sensibilidad a cisplatino y las mediciones de las aberraciones genómicas mediante el análisis de un intervalo de tamaños de segmento mínimos, en líneas celulares de TNBC (Figura 1B y figuras 5A-5C). Se observó una correlación significativa con la sensibilidad a cisplatino usando puntos de corte del tamaño mínimo de segmento de AI telomérico entre 5 y 25 Mb observándose el nivel de correlación más alto para el número total de segmentos con AI telomérico (N<sub>tAI</sub>) de al menos 12 MB (R<sup>2</sup> = 0,8; P = 0,016; Figura 1C). El análisis del tamaño de segmento mínimo óptimo no mejoró sensiblemente la correlación entre la sensibilidad a cisplatino y las mediciones de las aberraciones del número de copias, que permaneció como no significativa (Figuras 5D-5F).

**[0122]** También se investigó si misma asociación entre  $N_{tAI}$  y la sensibilidad a cisplatino estaba presente en muestras de tumores clínicos utilizando el punto de corte óptimo del tamaño de segmento de 12 MB ( $N_{tAI,12}$ ). Se comparó el  $N_{tAI,12}$  con la respuesta a quimioterapia en pacientes con TNBC tratados con monoterapia de cisplatino preoperatoria (Silver et al. (2010) J. Clin. Oncol. 28, 1145-1153). Las secciones de tejido de criostato de biopsias centrales previas al tratamiento se enriquecieron en células tumorales mediante microdissección con aguja, y se extrajo el ADN para el genotipado. Los genotipos de 42.000 SNP se determinaron con el sistema de genotipado dirigido a la Sonda de Inversión Molecular (MIP) (Affymetrix, Inc.) (Wang et al. (2007) Genome Biol. 8, R246). El grado de contaminación de células normal se estimó a partir de los datos de genotipo de MIP tal como se ha descrito (Figura 2B; Li et al (2008) Bioinformatics 9, 204). No se observó ninguna asociación entre el grado de contaminación normal y la respuesta a cisplatino ( $R^2 = 0,004$ ,  $P = 0,75$ ).

**[0123]** Se evaluaron los datos del genotipo de MIP de 21 casos con un contenido de células tumorales de al menos el 75% mediante el análisis de MCP para definir las regiones de AI telomérico, intersticial, o todo el cromosoma en todo el genoma (Figura 6A y Figura 7). Se observó una correlación entre el  $N_{tAI,12}$  y la tasa de respuesta, tal como se cuantifica por la puntuación de Miller-Payne ( $R^2 = 0,5$ ;  $P = 0,00032$ ; Figura 6B; Ogston et al (2003) Breast 12, 320-327), con un mayor número de regiones tAI asociadas con una mayor sensibilidad a cisplatino. El análisis de la curva característica de operación del receptor (ROC) reveló que el  $N_{tAI,12}$  se asociaba significativamente con la respuesta patológica completa a cisplatino (Miller-Payne 5) mediante el área bajo la curva (AUC = 0,85;  $P = 0,017$ ; Figura 6C). No hubo asociación aparente entre el número de segmentos de AI intersticial (Figura 6 A) o el nivel de AI de todo el cromosoma (Figura 7) y la respuesta al cisplatino.

**[0124]** El carcinoma de ovario seroso se trata a menudo con terapias a base de platino. Se investigó un conjunto de datos de SNP disponibles públicamente gama de carcinomas de ovario tratados con cisplatino o carboplatino más un taxano (Etemadmoghadam et al., (2009) Clin Cancer Res. 15, 1417-1427) y se identificaron 33 casos del subtipo seroso tratados después óptima de la citorreducción quirúrgica (tumor residual <1 cm) y pureza de tumor razonable (> 75%, estimada a partir de datos de SNP). El  $N_{tAI,12}$  se determinó mediante análisis de MCP. En estos casos de cáncer de ovario tratado con platino, se encontró una asociación entre los niveles más altos de AI telomérico en tumores y la ausencia de recaída en menos de un año (Figura 8A). Se utilizó el análisis ROC en la cohorte de TNBC para definir un punto de corte de  $N_{tAI,12}$  de al menos 13 sucesos, lo que proporcionó la mayor sensibilidad para la clasificación de pCR a la terapia de platino en la cohorte de TNBC. Este punto de corte se utilizó para clasificar la cohorte del cáncer ovárico en grupos de  $N_{tAI,12}$  altos y bajos y se encontró una mayor supervivencia libre de enfermedad, un indicador indirecto de la mayor sensibilidad al platino, en el grupo de  $N_{tAI,12}$  alto (Figura 8B).

**[0125]** Por lo tanto, la inestabilidad cromosómica, que se manifiesta por altos niveles de AI telomérico, caracterizan subconjuntos de TNBC y cáncer de ovario, y, además, los niveles más altos de estos cambios predicen vulnerabilidades terapéuticas específicas. Aunque el TNBC esporádico parece similar al cáncer de mama asociado a BRCA1 en los patrones de alteraciones cromosómicas y otros inmunofenotipos y características histológicas, el defecto o defectos moleculares precisos en el mantenimiento de la estabilidad cromosómica en estos tumores son desconocidos. Los resultados de los ejemplos descritos en este documento indican que la carga de reordenamientos cromosómicos resultantes de las roturas de la cadena de ADN reparadas incorrectamente son indicadores de los defectos de reparación del ADN que sensibilizan las células a ciertas quimioterapias (Figura 9). Por tanto, los niveles de desequilibrio alélica proporcionan un biomarcador preciso para predecir la sensibilidad del tumor al tratamiento con agentes genotóxicos, independientemente del conocimiento de la lesión causante de la reparación del ADN.

### Equivalentes

**[0126]** La presente invención es tal como se define en las reivindicaciones. Los expertos en la materia entenderán o serán capaces de determinar sin utilizar más que la experimentación de rutina, muchos equivalentes a las realizaciones de la presente invención descritas en el presente documento.

**[0127]** También se describen los siguientes procedimientos (cláusulas):

1. Un procedimiento para predecir el resultado del tratamiento contra el cáncer de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación global de aberraciones cromosómicas (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si una pluralidad de regiones cromosómicas que muestran una aberración cromosómica existen dentro de una pluralidad de loci cromosómicos, en el que dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico ( $N_{AI}$ ), pérdida de heterocigosidad ( $N_{LOH}$ ), aberraciones en el número de copias ( $N_{CNA}$ ), ganancia del número de copias ( $N_{CNG}$ ), disminución del número de copias ( $N_{CND}$ ) y combinaciones de los mismos, en relación a un control, y en el que la presencia de una pluralidad de regiones cromosómicas que muestran dichas aberraciones cromosómicas predice el resultado del tratamiento contra el cáncer del sujeto.

2. El procedimiento de la cláusula 1, en el que el tratamiento contra el cáncer es un tratamiento con quimioterapia.

3. El procedimiento de la cláusula 2, en el que el tratamiento con quimioterapia comprende agentes

quimioterapéuticos basados en platino.

- 5 4. El procedimiento de la cláusula 3, en el que los agentes quimioterapéuticos basados en platino se seleccionan del grupo que consiste en cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino
- 5 5. El procedimiento de la cláusula 1, en el que el sujeto es un humano.
- 10 6. El procedimiento de la cláusula 1, en el que el trastorno hiperproliferativo celular se selecciona del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de vejiga de células transicionales, cáncer de pulmón broncogénico, cáncer de tiroides, cáncer de páncreas, cáncer de próstata, cáncer de útero, cáncer testicular, cáncer gástrico, sarcomas de tejidos blandos y osteogénicos, neuroblastoma, tumor de Wilms, linfoma maligno (de Hodgkin y no Hodgkin), leucemia mieloblástica aguda, leucemia linfoblástica aguda, sarcoma de Kaposi, tumor de Ewing, mieloma múltiple refractario, carcinomas de células escamosas de la cabeza, cuello, cuello del útero, cáncer de colon, melanoma y vagina,
- 15 7. El procedimiento de la cláusula 1, en el que la muestra biológica se selecciona del grupo que consiste en células, líneas celulares, cortes histológicos, biopsias centrales congeladas, tejidos embebidos en parafina, tejidos fijados a formalina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspadura bucal, saliva, líquido cefalorraquídeo, orina, heces y médula ósea.
- 20 8. El procedimiento de la cláusula 1, en el que la muestra biológica se enriquece en la presencia de células hiperproliferativas hasta al menos el 75% de la población total de células.
- 25 9. El procedimiento de la cláusula 8, en el que el enriquecimiento se realiza según al menos una técnica seleccionada del grupo que consiste en microdissección con aguja, microdissección por láser, clasificación de células activadas por fluorescencia y clasificación de células inmunológicas.
- 30 10. El procedimiento de la cláusula 8 ó 9, en el que una máquina automatizada realiza dicha al menos una técnica para transformar así la muestra biológica en una forma purificada enriquecida en la presencia de células hiperproliferativas.
- 35 11. El procedimiento de la cláusula 1, en el que la muestra biológica se obtiene antes de que el sujeto haya recibido quimioterapia adyuvante.
- 40 12. El procedimiento de la cláusula 1, en el que la muestra biológica se obtiene después de que el sujeto haya recibido quimioterapia adyuvante.
- 45 13. El procedimiento de la cláusula 1, en el que el control se determina a partir de una muestra de células no hiperproliferativas del paciente o miembro de la misma especie a la que pertenece el paciente.
- 50 14. El procedimiento de la cláusula 1, en el que el control se determina a partir de la frecuencia media de aparición del locus genómico de regiones cromosómicas del mismo grupo étnico dentro de la especie a la que pertenece el paciente.
- 55 15. El procedimiento de la cláusula 13 ó 14, en el que el control proviene de tejido no canceroso que es el mismo tipo de tejido que dicho tejido canceroso del sujeto.
16. El procedimiento de la cláusula 13 ó 14, en el que el control proviene de tejido no canceroso que no es el mismo tipo de tejido que dicho tejido canceroso del sujeto.
17. El procedimiento de la cláusula 1, en el que  $N_{AI}$  se determina utilizando la proporción mayoritaria de copias (MCP).
18. El procedimiento de la cláusula 17, en el que  $N_{AI}$  para una región genómica determinada se cuenta cuando MCP es superior a 0,70.
19. El procedimiento de la cláusula 1, en el que la pluralidad de loci cromosómicos se distribuye al azar a lo largo de todo el genoma al menos cada 100 kb de ADN.
- 60 20. El procedimiento de la cláusula 1, en el que la pluralidad de loci cromosómicos comprenden al menos un locus cromosómico en cada uno de los 23 pares de cromosomas humanos.
- 65 21. El procedimiento de la cláusula 1, en el que la pluralidad de loci cromosómicos comprenden al menos un locus cromosómico en cada uno de los brazos de cada uno de los 23 pares de cromosomas humanos.
22. El procedimiento de la cláusula 21, en el que la pluralidad de loci cromosómicos comprenden al menos un locus

cromosómico en al menos un telómero de cada uno de los 23 pares de cromosomas humanos.

23. El procedimiento de la cláusula 22, en el que la pluralidad de loci cromosómicos comprenden al menos un locus cromosómico en cada uno de los telómeros de cada uno de los 23 pares de cromosomas humanos.

5 24. El procedimiento de la cláusula 1 ó 22, en el que las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 1 Mb.

10 25. El procedimiento de la cláusula 24, en el que las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 12 Mb.

26. El procedimiento de la cláusula 1, en el que la pluralidad de aberraciones cromosómicas comprende al menos 5 aberraciones cromosómicas.

15 27. El procedimiento de la cláusula 26, en el que la pluralidad de aberraciones cromosómicas comprende al menos 13 aberraciones cromosómicas.

20 28. El procedimiento de la cláusula 1, en el que los loci cromosómicos se seleccionan del grupo que consiste en polimorfismos de nucleótido único (SNP), polimorfismos de longitud de fragmento de restricción (RFLP) y repeticiones en tándem simples (STR).

25 29. El procedimiento de la cláusula 1, en el que los loci cromosómicos se analizan usando al menos una técnica seleccionada del grupo que consiste en sonda de inversión molecular (MIP), matriz de polimorfismos de nucleótido único (SNP), hibridación *in situ*, transferencia Southern, matrices de transcripción, matriz de hibridación genómica comparada (aCGH) y secuenciación de próxima generación.

30 30. El procedimiento de la cláusula 1, en el que el resultado del tratamiento se mide mediante al menos un criterio seleccionado del grupo que consiste en la supervivencia hasta la mortalidad, respuesta patológica completa, mediciones semicuantitativas de la respuesta patológica, remisión clínica completa, remisión clínica parcial, enfermedad clínica estable, supervivencia libre de recurrencia, supervivencia libre de metástasis, supervivencia libre de enfermedad, disminución de células tumorales circulantes, respuesta del marcador circulante y criterios RECIST.

35 31. El procedimiento de la cláusula 1, que comprende además la determinación de un régimen de tratamiento adecuado para el sujeto.

40 32. El procedimiento de la cláusula 31, en el que dicho régimen de tratamiento adecuado comprende al menos un agente quimioterapéutico basado en platino cuando se determina una pluralidad de aberraciones cromosómicas genómicas o no comprende al menos un agente quimioterapéutico basado en platino cuando no se determina una pluralidad de aberraciones cromosómicas genómicas.

**REIVINDICACIONES**

1. Procedimiento para determinar un régimen de tratamiento adecuado para un sujeto con cáncer, que comprende:  
 5 determinar en una muestra biológica que comprende células tumorales de un sujeto, el número de desequilibrios alélicos presentes en una pluralidad de regiones cromosómicas que se extienden hacia y que comprenden al menos un telómero, pero que no cruzan el centrómero,  
 en el que, cuando hay un mayor número de desequilibrios alélicos en relación a un control, se determina que un régimen de tratamiento que comprende al menos un agente quimioterapéutico basado en platino o al menos un inhibidor de poli(ADP ribosa) polimerasa (PARP), es un régimen de tratamiento adecuado, y que cuando no hay un  
 10 mayor número de desequilibrios alélicos en relación a un control, se determina que un régimen de tratamiento que no comprende un agente quimioterapéutico basado en platino ni un inhibidor de poli(ADP ribosa) polimerasa (PARP), es un régimen de tratamiento adecuado.
2. Procedimiento, según la reivindicación 1, en el que el número de desequilibrios alélicos se determina detectando  
 15 una pluralidad de SNP heterocigotos en la pluralidad de regiones cromosómicas que se extienden hacia y que comprenden al menos un telómero, pero que no cruzan el centrómero.
3. Procedimiento, según la reivindicación 1, en el que al menos una de las regiones cromosómicas que se extiende hacia y que comprende el telómero, pero que no cruza el centrómero, tiene al menos 5 Mb.  
 20
4. Procedimiento, según la reivindicación 1, en el que al menos una de las regiones cromosómicas que se extiende hacia y que comprende el telómero, pero que no cruza el centrómero, tiene al menos 10 Mb.
5. Procedimiento, según la reivindicación 1, en el que dicho al menos un agente quimioterapéutico que contiene platino se selecciona del grupo que consiste en cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino  
 25
6. Procedimiento, según la reivindicación 1, en el dicho al menos un inhibidor de poli(ADP ribosa) polimerasa (PARP) se selecciona del grupo que consiste en: Olaparib, ABT-888, BSI-201, INO-1001, PJ34, 3-aminobenzamida, 4-amino-1,8-naftalimida, 6(5H)-fenantridinona, benzamida, NU1025 y BGP-15.  
 30
7. Procedimiento, según la reivindicación 1, en el que la determinación del número de desequilibrios alélicos comprende al menos una técnica de secuenciación de ácidos nucleicos.
8. Procedimiento, según la reivindicación 7, en el que dicha al menos una técnica de secuenciación de ácidos nucleicos comprende sistemas de secuenciación de próxima generación.  
 35
9. Procedimiento, según la reivindicación 1, en el que la pluralidad de desequilibrios alélicos comprende al menos 5 desequilibrios alélicos.
- 40 10. Procedimiento, según la reivindicación 1, en el que el cáncer se selecciona del grupo que consiste en: cáncer de mama, cáncer de ovario, cáncer de vejiga de células transicionales, cáncer de pulmón broncogénico, cáncer de tiroides, cáncer de páncreas, cáncer de próstata, cáncer de útero, cáncer testicular, cáncer gástrico, sarcomas de tejidos blandos y osteogénicos, neuroblastoma, tumor de Wilms, linfoma maligno (de Hodgkin y no Hodgkin), leucemia mieloblástica aguda, leucemia linfoblástica aguda, sarcoma de Kaposi, tumor de Ewing, cáncer de colon, melanoma, tal como mieloma múltiple refractario, y carcinomas de células escamosas de la cabeza, cuello, cuello  
 45 del útero y vagina.

Figura 1

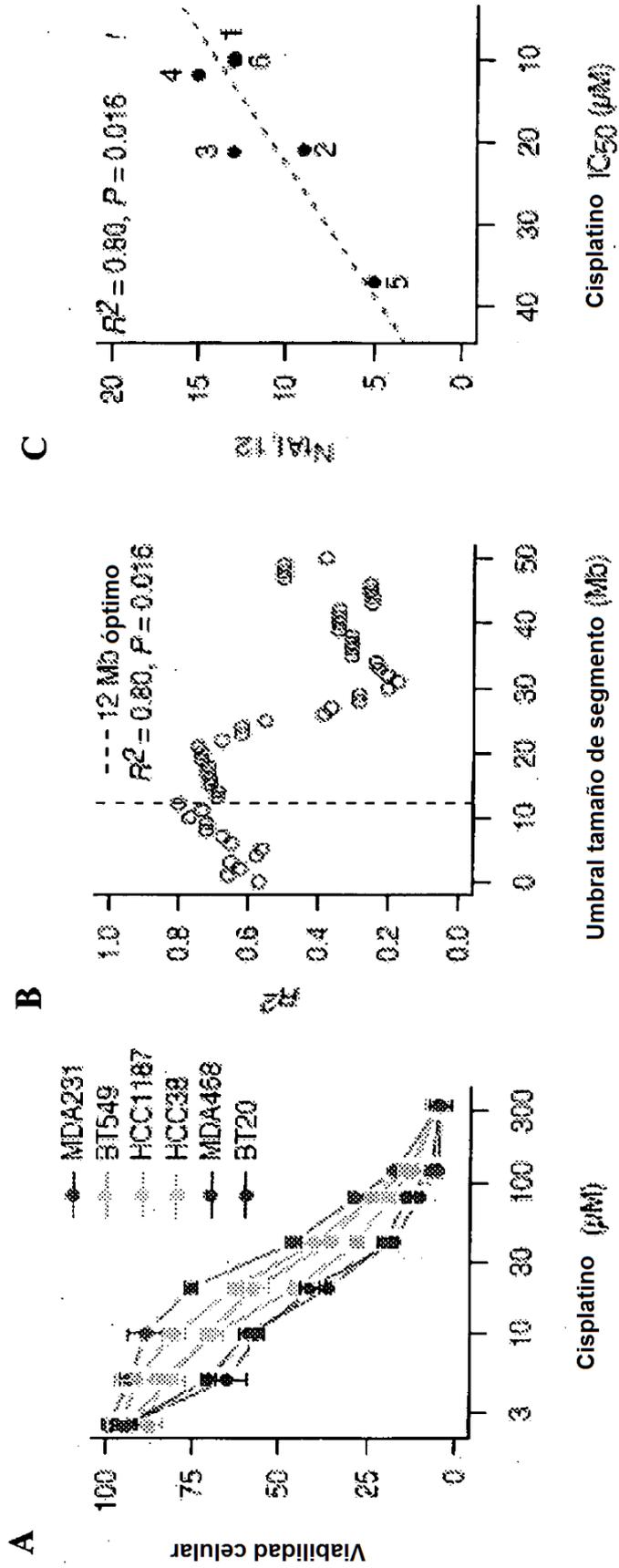


Figura 2

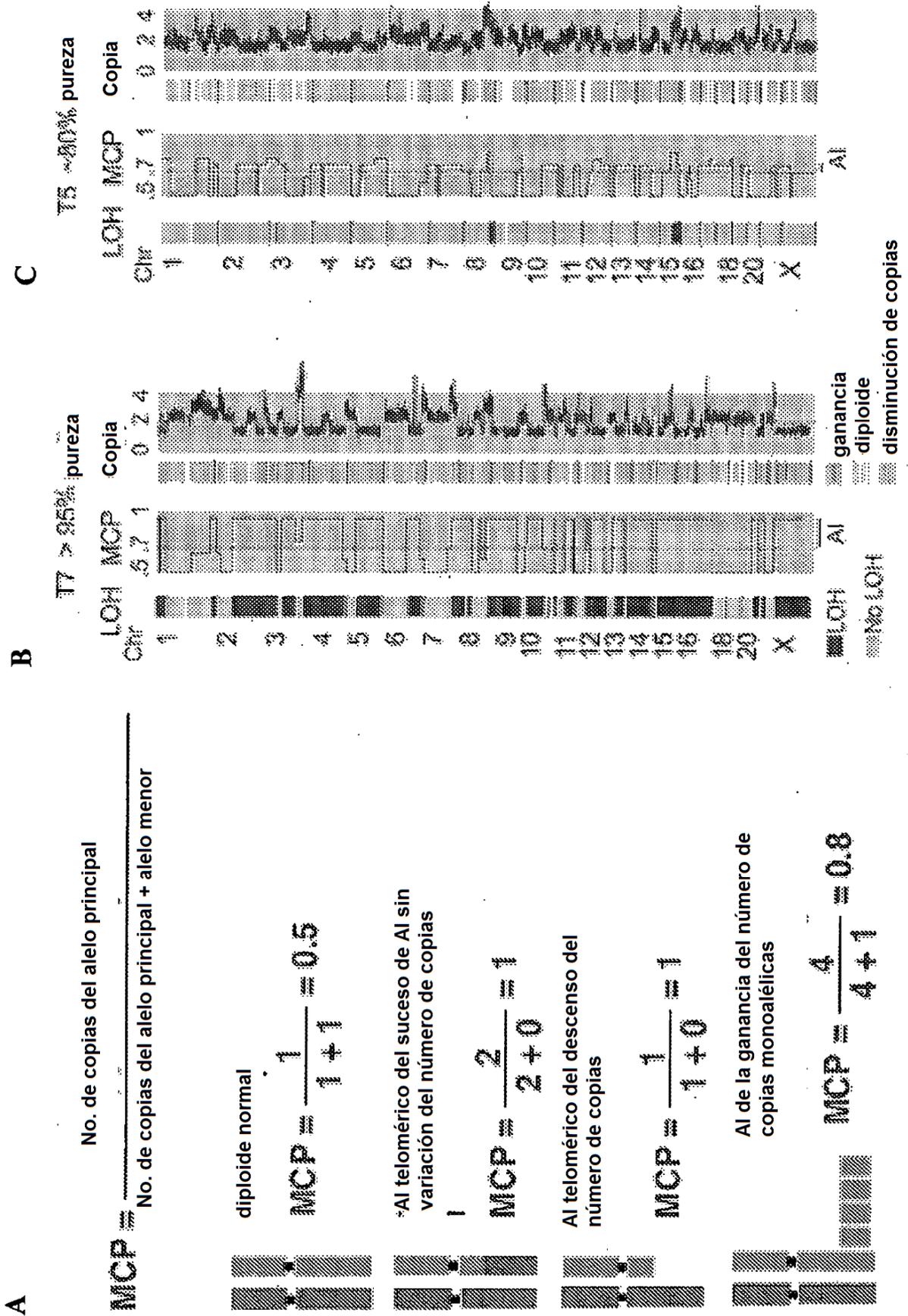


Figura 3

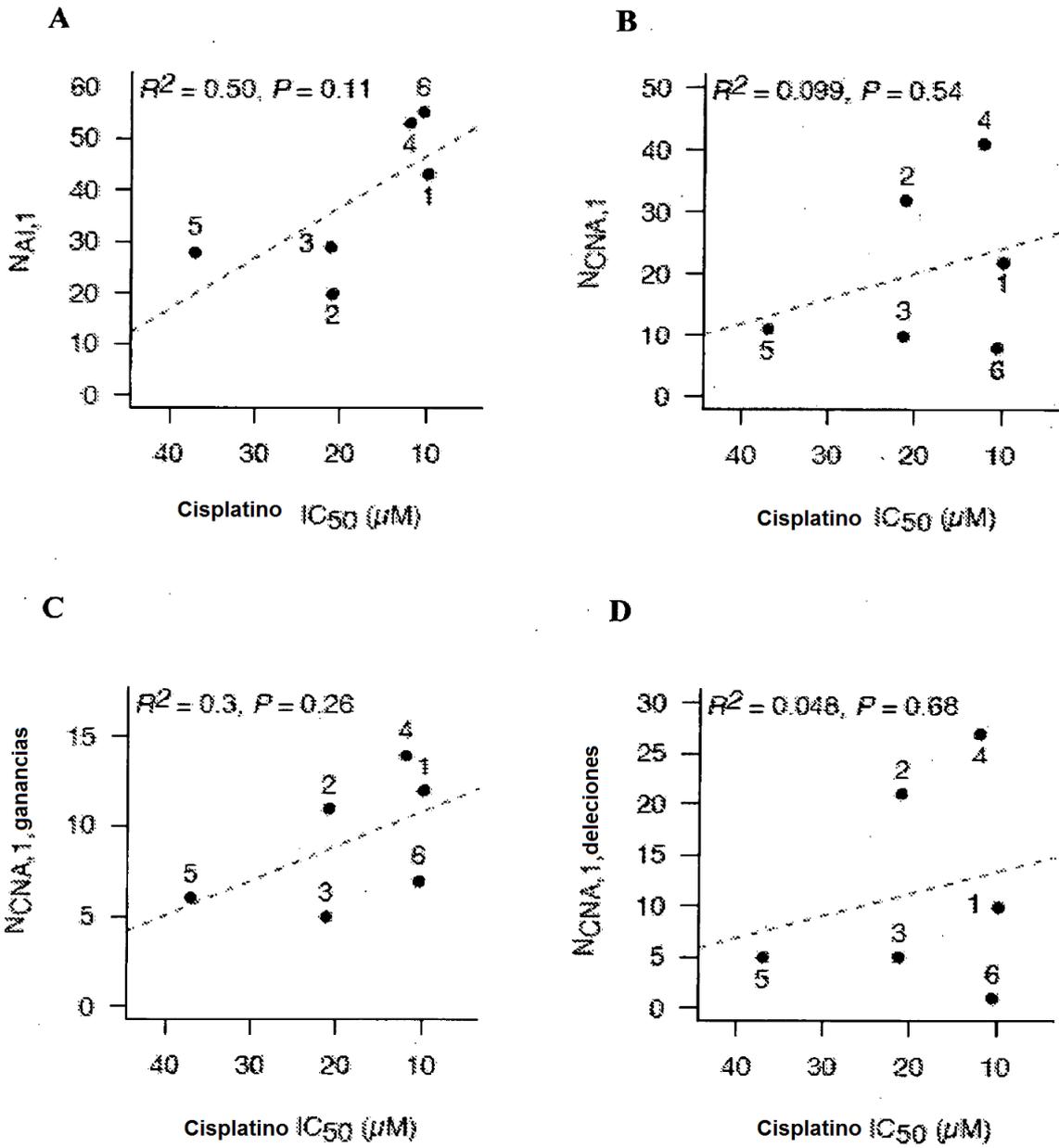


Figura 4

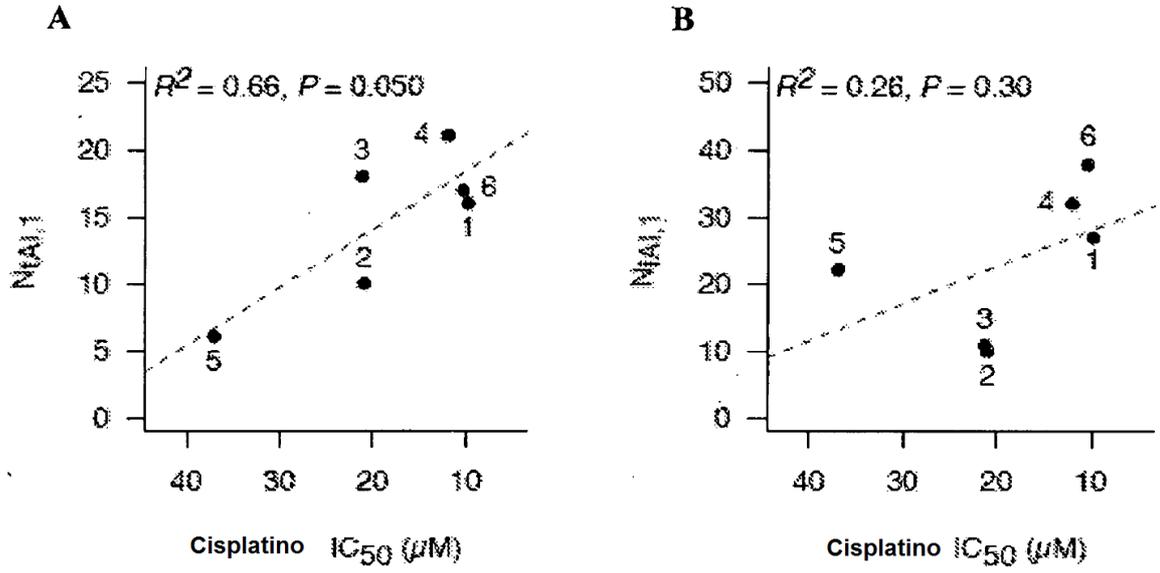


Figura 5

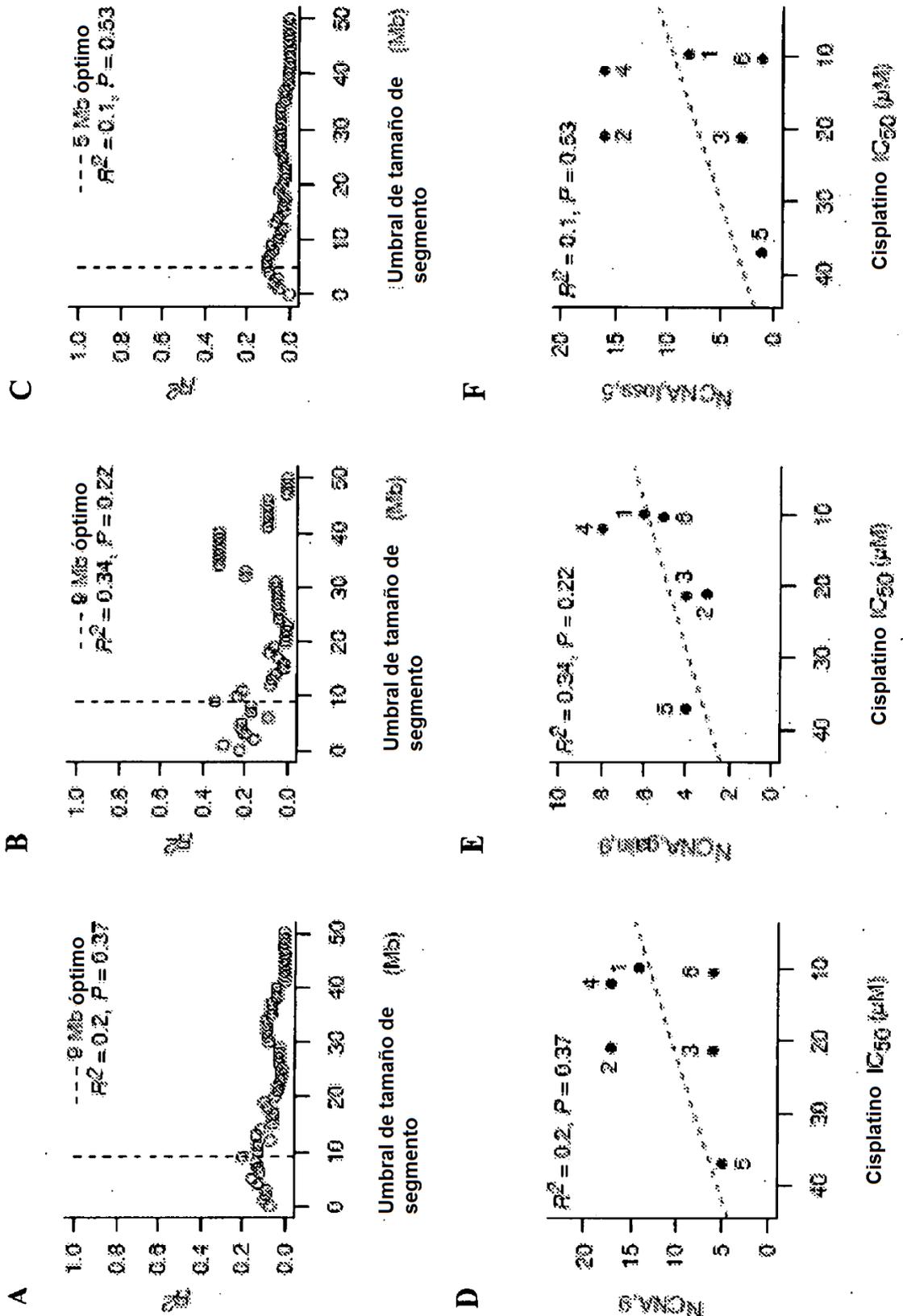


Figura 6

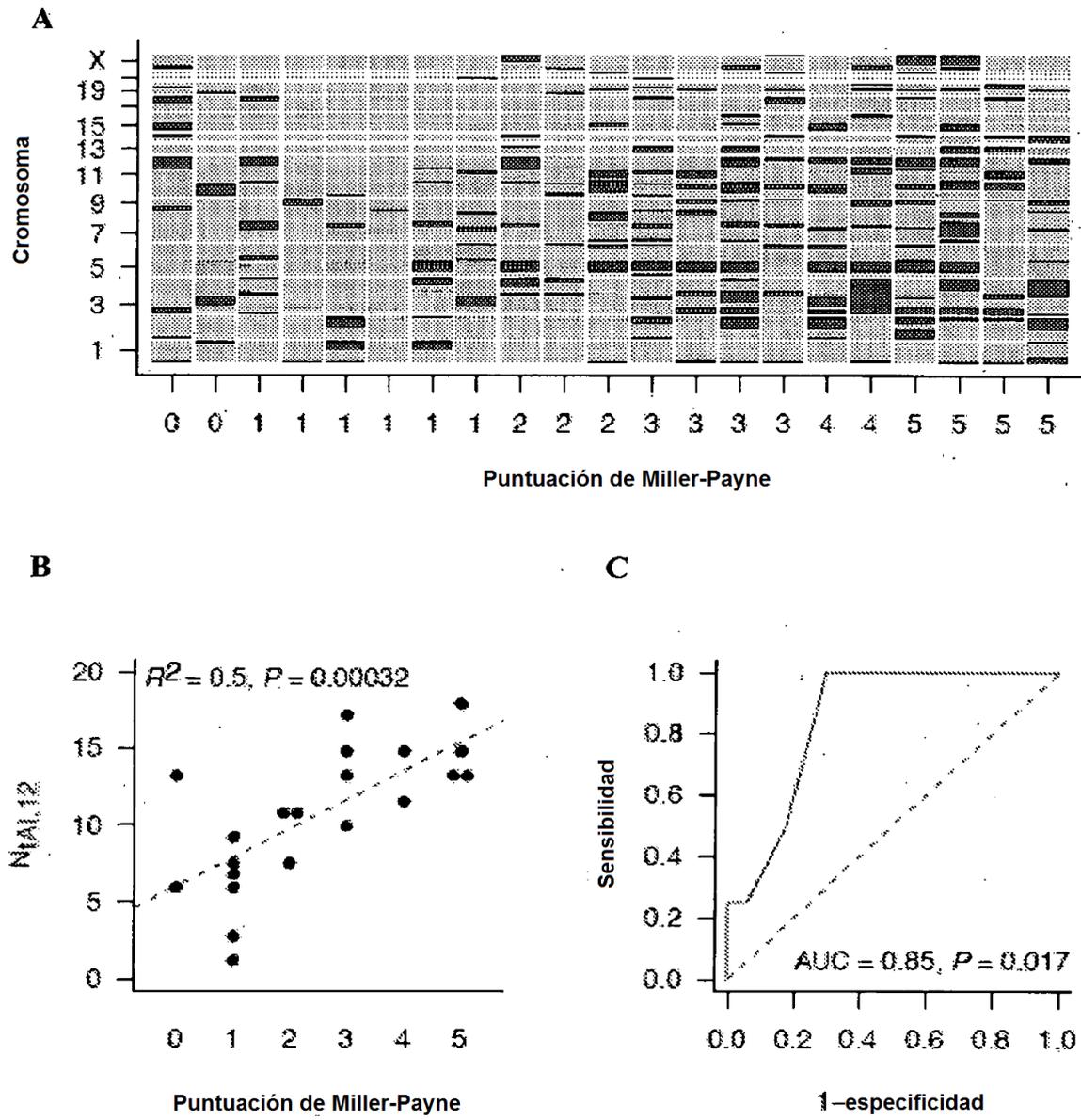


Figura 7

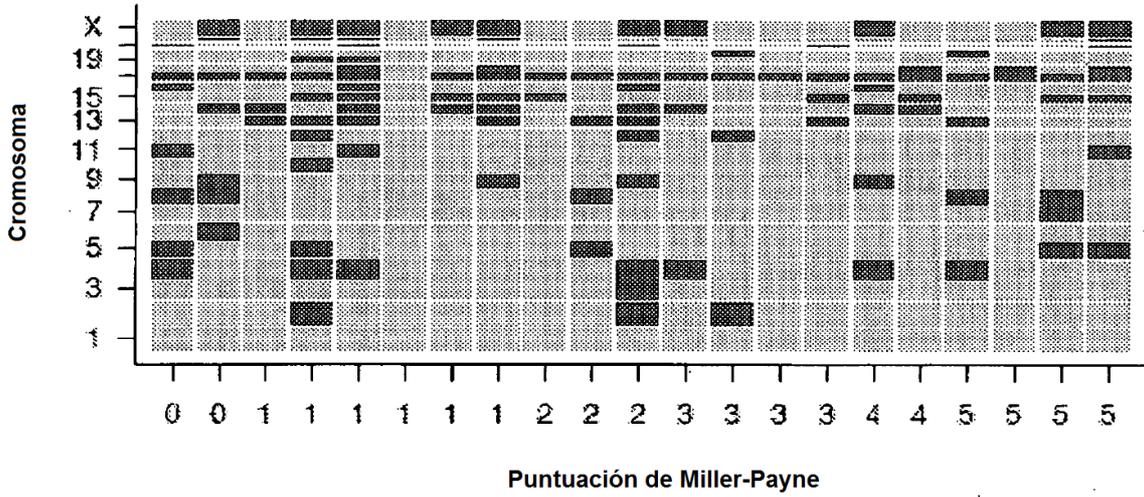


Figura 8

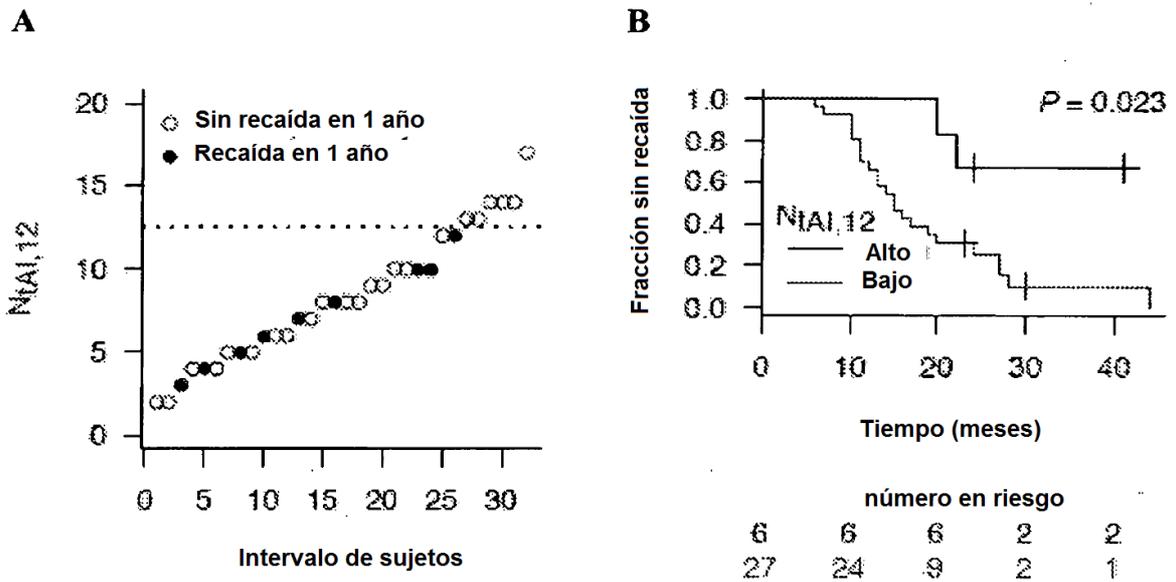


Figura 9

