

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 687 024**

51 Int. Cl.:

**G01N 33/574** (2006.01)

**C12Q 1/68** (2008.01)

**G01N 33/50** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.02.2013 PCT/US2013/027295**

87 Fecha y número de publicación internacional: **06.09.2013 WO13130347**

96 Fecha de presentación y número de la solicitud europea: **22.02.2013 E 13755251 (9)**

97 Fecha y número de publicación de la concesión europea: **11.07.2018 EP 2817630**

54 Título: **Procedimientos para predecir la respuesta contra el cáncer**

30 Prioridad:

**23.02.2012 US 201261602460 P**  
**29.02.2012 US 201261604810 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**23.10.2018**

73 Titular/es:

**CHILDREN'S MEDICAL CENTER CORPORATION**  
**(25.0%)**  
**55 Shattuck Street**  
**Boston, Massachusetts 02115, US;**  
**THE TECHNICAL UNIVERSITY OF DENMARK**  
**(25.0%);**  
**DANA-FARBER CANCER INSTITUTE, INC.**  
**(25.0%) y**  
**THE BRIGHAM AND WOMEN'S HOSPITAL, INC.**  
**(25.0%)**

72 Inventor/es:

**SILVER, DANIEL;**  
**WANG, ZHIGANG;**  
**SZALLASI, ZOLTAN;**  
**RICHARDSON, ANDREA;**  
**BIRKBAK, NICOLAI JUUL y**  
**EKLUND, ARON**

74 Agente/Representante:

**SALVÀ FERRER, Joan**

ES 2 687 024 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 predecir la respuesta contra el cáncer

5 **[0001]** Esta solicitud reivindica el beneficio según 35 USC 119 (e) de las solicitudes provisionales de los EE. UU. n.º 61/602.460 presentada el 23 de febrero de 2012 y 61/604.810 presentada el 29 de febrero de 2012.

**[0002]** Esta invención se hizo con ayuda del gobierno bajo el número de concesión CA089393097193 otorgado por los Institutos Nacionales de Salud. El gobierno tiene ciertos derechos en la invención.

10

### ANTECEDENTES DE LA INVENCION

**[0003]** Los oncólogos médicos se han beneficiado enormemente de los esfuerzos relativamente recientes para diseccionar y comprender los elementos genéticos que subyacen al cáncer de mamíferos. La identificación de 15 predisposiciones genéticas específicas, como las mutaciones en BRCA-1, BRCA2 y HER2, ha proporcionado información clave sobre los mecanismos subyacentes a la tumorigénesis y ha demostrado ser útil para el diseño de nuevas generaciones de enfoques específicos 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 hecho posibles los análisis sistemáticos de las alteraciones genéticas en los cánceres humanos. El documento WO 2011/160063 20 describe procedimientos y materiales para evaluar la pérdida de heterocigosidad. N. Juul et al., Cancer Research, vol. 69 (24), 15 de diciembre de 2009, páginas 509S-510S, informa que una medida resumida derivada del perfil genómico de los puntos de corte cromosómicos predice la respuesta al tratamiento con cisplatino, agente dañino para el ADN. Ott Katya et al., Clinical Cancer Research, vol. 9 (6), 1 de junio de 2003, páginas 2307-2315, informa que la inestabilidad cromosómica, en lugar de la mutación p53, se asocia con la respuesta a la quimioterapia 25 neoadyuvante basada en cisplatino en el carcinoma gástrico.

**[0004]** Sin embargo, las intervenciones clínicas basadas en esta información se han visto gravemente obstaculizadas por el hecho de que a menudo solo un porcentaje de pacientes responderá favorablemente a un 30 tratamiento anticanceroso en particular. Los oncólogos médicos actualmente no pueden predecir qué pacientes responderán o no a un tratamiento quimioterapéutico propuesto.

**[0005]** Por consiguiente, existe una gran necesidad en la técnica de identificar la capacidad de respuesta del paciente a terapias anticancerosas particulares.

### 35 RESUMEN DE LA INVENCION

**[0006]** La presente invención se basa, al menos en parte, en el descubrimiento de que ciertos patrones de aberraciones de ADN descritos en este documento son predictivos de respuesta anticancerosa de las células que albergan dichas aberraciones de ADN a terapias contra el cáncer.

40

**[0007]** Por consiguiente, en un aspecto, la presente descripción presenta un procedimiento para predecir el resultado del tratamiento anticancerígeno de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberración cromosómica global (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si existe una pluralidad de regiones cromosómicas que exhiben una aberración 45 cromosómica 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 del número de copias (CNA), ganancia de número de copias (CNG), disminución del número de copias (CND) y combinaciones de las mismas, con respecto 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 50 anticancerígeno del sujeto. El sujeto puede ser un mamífero, como un ser humano.

**[0008]** Por ejemplo, las mutaciones en BRCA1 o BRCA2 causan defectos en la reparación del ADN que predicen la sensibilidad a las sales de platino en cáncer de mama y de ovario; sin embargo, algunos pacientes sin mutaciones en BRCA también se benefician de estos agentes. Este estudio muestra que los defectos en la 55 reparación del ADN que causan la sensibilidad al platino se pueden deducir del número de desequilibrio alélico (AI) o del número de desequilibrio alélico telomérico (NtAI), una medida de la aberración genómica en los tumores. El NtAI puede identificar pacientes con cáncer sin mutaciones en BRCA que puedan beneficiarse de la terapia basada en platino.

**[0009]** En un aspecto, el tratamiento contra el cáncer es el tratamiento de quimioterapia. En otra realización, el tratamiento anticanceroso comprende agentes quimioterapéuticos basados en platino (por ejemplo, cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino).

5 **[0010]** 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 uterino, cáncer testicular, cáncer gástrico, sarcomas de tejido blando 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  
10 y carcinomas de células escamosas de cabeza, cuello, cuello uterino y vagina

**[0011]** En otro aspecto más, la muestra biológica se selecciona del grupo que consiste en células, líneas celulares, portaobjetos histológicos, biopsias con núcleo congelado, tejidos embebidos en parafina, tejidos fijados con formalina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspado bucal, saliva, líquido  
15 cefalorraquídeo, orina, heces y médula ósea. En una realización, la muestra biológica está enriquecida por la presencia de células hiperproliferativas a al menos el 75 % de la población total de células. En otra realización, el enriquecimiento se realiza de acuerdo con al menos una técnica seleccionada del grupo que consiste en microdissección con aguja, microdissección con láser, clasificación de células activadas por fluorescencia y  
20 clasificación inmunológica de células. En otra realización más, una máquina automática realiza la al menos una técnica para transformar de ese modo 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. Como alternativa, la muestra biológica se obtiene después de que el sujeto haya recibido quimioterapia adyuvante.

25 **[0012]** En otro aspecto más, el control se determina a partir de una muestra de células hiperproliferativas no celulares 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 promedio de aparición del locus genómico de las regiones cromosómicas del mismo grupo étnico dentro de la especie a la que pertenece el paciente. En otra realización, el control es de tejido no canceroso que es del mismo tipo de tejido que dicho tejido canceroso del sujeto. En otra realización más, el control  
30 es de tejido no canceroso que no es del mismo tipo de tejido que dicho tejido canceroso del sujeto.

**[0013]** En otro aspecto, el AI se determina usando una proporción de copia mayor (MCP). En una realización, se cuenta el AI para una región genómica dada cuando MCP es mayor que 0,70.

35 **[0014]** En otro aspecto más, la pluralidad de loci cromosómicos se distribuye aleatoriamente por 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 otra realización más, la pluralidad de loci cromosómicos comprende al menos un locus  
40 cromosómico en al menos un telómero de cada uno de los 23 pares de cromosomas humanos. En otra realización más, la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada telómero de cada uno de los 23 pares de cromosomas humanos.

45 **[0015]** En otro aspecto más, 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.

**[0016]** 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  
50 aberraciones cromosómicas.

**[0017]** En otro aspecto más, los loci cromosómicos se seleccionan del grupo que consiste en polimorfismos de un solo nucleótido (SNP), polimorfismos de longitud de fragmentos de restricción (RFLP) y repeticiones en tándem simples (STR).  
55

**[0018]** En otro aspecto más, los loci cromosómicos se analizan usando al menos una técnica seleccionada del grupo que consiste en la sonda de inversión molecular (MIP), matriz de polimorfismo de un solo nucleótido (SNP), hibridación *in situ*, transferencia de Southern, hibridación genómica comparativa de matriz (aCGH) y secuenciación de próxima generación.

**[0019]** En otro aspecto, el resultado del tratamiento se mide por al menos un criterio seleccionado del grupo que consiste en supervivencia hasta mortalidad, respuesta patológica completa, medidas semicuantitativas de respuesta patológica, remisión completa clínica, remisión parcial clínica, enfermedad estable clínica, supervivencia libre de recurrencia, supervivencia libre de metástasis, supervivencia libre de enfermedad, disminución de células tumorales circulantes, respuesta de marcador circulante y criterios RECIST.

**[0020]** En otro aspecto más, el procedimiento además comprende determinar 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 genómicas cromosómicas o no comprende al menos un agente quimioterapéutico basado en platino cuando no se determina una pluralidad de aberraciones genómicas cromosómicas.

**[0021]** La invención proporciona un procedimiento in vitro como se define en las reivindicaciones para seleccionar una terapia que comprende platino para un sujeto humano que tiene cáncer, el procedimiento que comprende: someter una muestra biológica que comprende células tumorales tomadas del ser humano a análisis de desequilibrio alélico telomérico (tAI); detectar el número de desequilibrio alélico telomérico (NtAI) y seleccionar una terapia que comprenda platino para el sujeto cuando se detecta que el NtAI está por encima de un valor de referencia, basado en el reconocimiento de que la terapia que comprende platino es efectiva en pacientes que tienen un NtAI por encima del valor de referencia; en el que el valor de referencia se determina a partir del NtAI obtenido de cánceres similares que se sabe que son resistentes al platino. En una realización, se selecciona una terapia contra el cáncer que no comprende platino para el sujeto cuando se detecta que el NtAI está por debajo de un valor de referencia, basándose en el reconocimiento de que la terapia contra el cáncer que comprende platino no es efectiva en pacientes que tienen el NtAI por debajo de un valor de referencia. La terapia seleccionada puede administrarse opcionalmente al sujeto.

**[0022]** Un ensayo o un procedimiento para seleccionar una terapia que comprende platino para un sujeto que tiene cáncer que comprende: someter una muestra biológica tomada del sujeto a un análisis de desequilibrio alélico (AI); detectar el número de AI; y seleccionar una terapia contra el cáncer que comprende platino para el sujeto cuando el número de AI está por encima de un valor de referencia basado en el reconocimiento de que la terapia contra el cáncer que comprende platino es eficaz en pacientes que tienen el número de AI por encima de un valor de referencia y opcionalmente administrar la terapia contra el cáncer que comprende platino, si se selecciona.

**[0023]** Los ensayos pueden comprender opcionalmente las etapas de obtener una muestra que comprende células cancerosas o ADN derivado de células cancerosas del sujeto, sometiendo la muestra a manipulaciones, tales como purificación, amplificación de ADN, contacto de la muestra con una sonda, marcaje y otras etapas similares que son necesarias en el análisis del NtAI o NAI. Además, el ensayo y el análisis pueden ser realizados por una máquina no humana ejecutando un algoritmo y determinando automáticamente si la muestra comprende las condiciones para seleccionar una terapia contra el cáncer que comprende platino o terapia contra el cáncer que no comprende platino para el sujeto basándose en el análisis del NAI o el NtAI.

**[0024]** El cáncer puede ser cualquier cáncer. En algunos aspectos de todas las realizaciones de la invención, el cáncer se selecciona entre cánceres de mama y de ovario. En algunos aspectos de todas las realizaciones de la invención, el sujeto es negativo para las bien conocidas mutaciones en BRCA1 y/o BRCA2. En algunos aspectos de todas las realizaciones, el sujeto tiene disminución o aumento en el ARNm de BRCA1 y/o BRCA2, que se puede determinar opcionalmente junto con el ensayo o antes o después de realizar el ensayo, y que puede ayudar a determinar si el cáncer será receptivo o resistente al tratamiento con terapia contra el cáncer que comprenda platino.

**[0025]** También proporcionamos un procedimiento para predecir el resultado del tratamiento anticancerígeno de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberración cromosómica global (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si existe una pluralidad de regiones cromosómicas que muestran una aberración cromosómica dentro de una pluralidad de loci cromosómicos, donde dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico (NAI), pérdida de heterocigosidad (NLOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG), disminución del número de copias (NCND) y combinaciones de los mismos, con respecto 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 anticancerígeno del sujeto.

**[0026]** También proporcionamos un procedimiento para predecir el resultado del tratamiento anticancerígeno

de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberración cromosómica global (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si existe una pluralidad de regiones cromosómicas que muestran una aberración cromosómica dentro de una pluralidad de loci cromosómicos, donde dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico (NAI), pérdida de heterocigosidad (NLOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG), disminución del número de copias (NCND) y combinaciones de los mismos, con respecto 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 anticancerígeno del sujeto.

10 **[0027]** Además proporcionamos un procedimiento para determinar el pronóstico en un paciente que comprende: (a) determinar si el paciente comprende células cancerosas que tienen una firma de LOH, en el que la presencia de más de un número de referencia de regiones de LOH en al menos un par de cromosomas humanos de células cancerosas del paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH indican que las células cancerosas tienen la firma de LOH, donde al menos un par de cromosomas humanos no es un par de cromosomas sexuales humanos X/Y, en el que la primera longitud es de aproximadamente 1,5 o más megabases, y (b) (1) determinar, basándose al menos en parte en la presencia de la firma de LOH, que el paciente tiene un pronóstico relativamente bueno o (b) (2) determinar, basándose al menos en parte en la ausencia de la firma de LOH, que el paciente tiene un pronóstico relativamente malo.

20 **[0028]** Proporcionamos una composición que comprende un agente terapéutico seleccionado del grupo que consiste en agente dañino para el ADN, antraciclina, inhibidor de la topoisomerasa I e inhibidor de PARP para usar en el tratamiento de un cáncer seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de esófago, cáncer de pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, 25 cáncer de recto, cáncer colorrectal y cáncer de páncreas en un paciente con más de un número de referencia de regiones de LOH en al menos un par de cromosomas humanos de una célula cancerosa del paciente que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH, donde al menos un par de cromosomas humanos no es un par de cromosomas sexuales humanos X/Y, en el que la primera longitud es de aproximadamente 1,5 o más megabases.

30 **[0029]** Además, proporcionamos un procedimiento para tratar el cáncer en un paciente, que comprende: (a) determinar en una muestra de dicho paciente que el número de regiones de LOH en al menos un par de cromosomas humanos de una célula cancerosa del paciente con cáncer son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH indica que las células cancerosas tienen la firma de LOH, en la que el al menos un par de cromosomas humanos no es un par de cromosomas sexuales humanos X/Y, en el que la primera longitud es de aproximadamente 1,5 o más megabases; (b) proporcionar un valor de prueba derivado del número de dichas regiones de LOH; (c) comparar dicho valor de prueba con uno o más valores de referencia derivados del número de dichas regiones de LOH en una población de referencia (por ejemplo, media, mediana, terciles, cuartiles, quintiles, etc.); y (d) administrar a dicho paciente un fármaco anticanceroso, o recomendar o prescribir o iniciar un régimen de tratamiento que comprende quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revela que el valor de prueba es mayor (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor; al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 desviaciones típicas mayor) que al menos uno de dichos valores de referencia; o (e) recomendar o prescribir o iniciar un régimen de tratamiento que no comprende quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revela que el valor de la prueba no es mayor (por ejemplo, no más de 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor, no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor) que al menos uno de dichos valores de referencia.

50 **[0030]** En una realización, dicho agente que daña el ADN es cisplatino, carboplatino, oxalapatino o picoplatino, dicha antraciclina es epirubicina o doxorubicina, dicho inhibidor de la topoisomerasa I es camptotecina, topotecán o irinotecán, y/o dicho inhibidor de PARP es iniparib, olaparib o velapirib.

**[0031]** Proporcionamos una composición que comprende un agente terapéutico seleccionado del grupo que consiste en terapia contra el cáncer que comprende platino y antraciclina para usar en el tratamiento de un cáncer 55 seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de esófago, cáncer de pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, cáncer de recto, cáncer colorrectal y cáncer de páncreas en un paciente con desequilibrio alélico aumentado. La invención proporciona una composición que comprende una terapia contra el cáncer que comprende platino para tratar un cáncer seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de esófago, cáncer de

pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, cáncer de recto, cáncer colorrectal y cáncer de páncreas en un paciente; en el que el paciente tiene un número aumentado de desequilibrio alélico telomérico (NtAI) en comparación con un valor de referencia, en el que el valor de referencia es un NtAI de al menos 20, 22, 24 o 27; y en donde además se determina que el paciente no porta una mutación en BRCA1 y/o BRCA2, o en donde se determina adicionalmente que el paciente tiene una cantidad disminuida de ARNm de BRCA1 en la célula cancerosa y/o se determina adicionalmente que tiene una metilación aumentada de la región promotora de BRCA1, en comparación con un control BRCA1 de tipo silvestre.

10 **[0032]** Proporcionamos un procedimiento para predecir el resultado del tratamiento anticancerígeno de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberración cromosómica global (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si existe una pluralidad de regiones cromosómicas que muestra una aberración cromosómica dentro de una pluralidad de loci cromosómicos, donde dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico (NAI), pérdida de heterocigosidad (NLOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG), disminución del número copias (NCND) y combinaciones de los mismos, con respecto 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 anticancerígeno del sujeto. En algunos aspectos de todas las realizaciones de la invención, el tratamiento anticancerígeno es un tratamiento de quimioterapia, que también puede ser agentes quimioterapéuticos basados en platino, por ejemplo, cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatino.

25 **[0033]** En algunos aspectos de todas las realizaciones de la invención, el trastorno hiperproliferativo celular puede seleccionarse 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 tejido blando 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 carcinoma de células escamosas de la cabeza, cuello, cuello uterino, cáncer de colon, melanoma y vagina.

30 **[0034]** La muestra biológica puede seleccionarse del grupo que consiste en células, líneas celulares, portaobjetos histológicos, biopsias con núcleo congelado, tejidos embebidos en parafina, tejidos fijados con formalina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspado bucal, saliva, líquido cefalorraquídeo, orina, heces y médula ósea, en el que la muestra comprende células cancerosas.

35 **[0035]** En algún aspecto de todas las realizaciones de la invención, que incluyen los ensayos y procedimientos, las células cancerosas en la muestra pueden enriquecerse usando, por ejemplo, microdissección con aguja, microdissección con láser, clasificación de células activadas por fluorescencia y clasificación inmunológica de células.

40 **[0036]** En algunos aspectos de todas las realizaciones de la invención, una máquina automática realiza la al menos una técnica para transformar de ese modo la muestra biológica en una forma purificada enriquecida por la presencia de células hiperproliferativas.

45 **[0037]** En algunos aspectos de todas las realizaciones de la invención, la muestra o muestra biológica se obtiene antes de que el sujeto haya recibido quimioterapia adyuvante, o después de que el sujeto haya recibido quimioterapia adyuvante.

50 **[0038]** En algunos aspectos de todas las realizaciones de la invención, el control se determina a partir de la frecuencia promedio de aparición del locus genómico de las regiones cromosómicas del mismo grupo étnico dentro de la especie a la que pertenece el paciente. El control también puede ser de tejido no canceroso que es del mismo tipo de tejido que dicho tejido canceroso del sujeto, o de tejido no canceroso que no es del mismo tipo de tejido que dicho tejido canceroso del sujeto.

#### Breve descripción de los dibujos

55 **[0039]**

Las **Figuras 1A-1C** muestran la correlación entre las regiones del desequilibrio alélico (AI) y la sensibilidad al cisplatino *in vitro*. La **Figura 1A** muestra una curva de respuesta a la dosis de seis líneas celulares de TNBC como

se determina mediante un ensayo de proliferación después de 48 horas de exposición al cisplatino. Las curvas para las células con valores de  $CI_{50}$  más bajos (mayor sensibilidad) se muestran en azul; la línea celular con  $CI_{50}$  más alta (resistencia máxima) se muestra en rojo; las líneas celulares con sensibilidad intermedia se muestran en gris. La **Figura 1B** muestra el efecto del umbral del tamaño del segmento 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  $R^2$  basado en la regresión lineal entre el recuento de las regiones CNA de un tamaño mínimo indicado en el eje X, y la  $CI_{50}$  del cisplatino en un panel de 6 líneas celulares TNBC (BT20, BT-549, HCC1187, HCC38, MDA-MB-231, MDA-MB-468). El umbral de tamaño de segmento mínimo óptimo se indica mediante la línea punteada. La **Figura 1C** muestra una comparación entre el número de regiones de AI telomérico (NtAI, 12) y la sensibilidad al 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.

La **Figura 2** muestra que el análisis de la proporción de copias principales (MCP) identifica un desequilibrio alélico en muestras de biopsia tumoral con diferentes grados de pureza de la célula tumoral. La Figura 2 muestra la fórmula para el cálculo de MCP, así como los cromosomas bialélicos normales y tres formas diferentes en que se puede producir el desequilibrio alélico de una región cromosómica y el cálculo de MCP correspondiente. También preparamos diagramas que representan la pantalla de pérdida de heterocigosidad (LOH), AI determinado por MCP y análisis del número de copias absoluto en dos muestras tumorales con diferentes grados de contaminación celular normal: T7 con > 95 % de contenido de células tumorales y T5 con aproximadamente el 80 % de contenido tumoral. Los cromosomas están indicados en el lado izquierdo. Las primeras columnas para cada tumor muestran las células para LOH (azul) y la retención de heterocigosidad (amarillo) en cada posición del cromosoma. Las segundas columnas muestran 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 llama para regiones con MCP mayor que 0,7. Las columnas tercera y cuarta muestran el número de copias de ADN absoluto en cada posición, el blanco indica diploide, los tonos de rojo indican la ganancia de copia y los tonos de azul indican la pérdida de copia. Los niveles del número de copias se muestran en los paneles del extremo derecho. La muestra de tumor con mayor pureza (T7) muestra acuerdo entre LOH y llamadas de AI determinadas por MCP. En la muestra tumoral con solo un 80 % de células tumorales, la señal de LOH se pierde, pero AI todavía se puede estimar mediante MCP con un umbral de 0,70. La muestra de tumor con mayor pureza (T7) muestra acuerdo entre LOH y llamadas de AI determinadas por MCP. En la muestra tumoral con solo un 80 % de células tumorales, la señal de LOH se pierde, pero todavía se puede estimar el AI mediante MCP con un umbral de 0,70.

Las **Figuras 3A-3D** muestran la asociación entre la sensibilidad al cisplatino y el número de anomalías genómicas en un panel de líneas celulares TNBC. La **Figura 3A** muestra la  $CI_{50}$  del cisplatino frente al número de regiones de AI telomérico de al menos 1 Mb de largo con un AI definido por MCP > 0,7. La **Figura 3B** muestra la  $CI_{50}$  del cisplatino frente al recuento de regiones con aberración del número de copias, incluidas las ganancias y pérdidas, de al menos 1 Mb de longitud. La **Figura 3C** muestra la  $CI_{50}$  del cisplatino frente al recuento de regiones con ganancia en el número de copias, de al menos 1 Mb de longitud. La **Figura 3D** muestra la  $CI_{50}$  del cisplatino frente al recuento de regiones con pérdida de número de copias, de al menos 1 Mb de longitud. Las líneas celulares están indicadas en cada figura y son las mismas que en la Figura 1.

Las **Figuras 4A-4B** muestran la asociación entre la sensibilidad al cisplatino y el recuento de regiones de AI telomérico o intersticiales en un panel de líneas celulares TNBC. La **Figura 4A** muestra la  $CI_{50}$  del cisplatino frente al número de regiones de AI telomérico de al menos 1 Mb de largo con un AI definido por MCP > 0,7. La **Figura 4B** muestra la  $CI_{50}$  del cisplatino frente al número de regiones de AI intersticiales de al menos 1 Mb de longitud con un AI definido por MCP > 0,7. Las líneas celulares están indicadas en cada figura y son las mismas que en la Figura 1.

Las **Figuras 5A-5F** muestran la asociación entre las aberraciones enumeradas del número de copias (CNA) y la sensibilidad al cisplatino *in vitro*. Las **Figuras 5A-5C** muestran la determinación del tamaño mínimo del segmento que demuestra la mejor correlación con la sensibilidad al cisplatino para el número de aberraciones del número de copias (NCNA; Figura 5A), número de regiones con ganancia de número de copias (NCNA, ganancia; Figura 5B), y número de regiones con pérdida de número de copias (NCNA, pérdida; Figura 5C). Cada punto representa un valor  $R^2$  basado en la regresión lineal entre el recuento de las regiones CNA de un tamaño mínimo indicado en el eje X, y la  $CI_{50}$  del 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 las regiones CNA está indicado por la línea punteada. Las **Figuras 5D-5F** muestran gráficas de los valores de la  $CI_{50}$  del cisplatino ( $\mu M$ , eje X) frente al número de regiones CNA con tamaños de segmentos mínimos óptimos (eje Y) de la siguiente manera: NCNA de al menos 9 Mb de longitud (Figura 5D), NCNA, ganancia de al menos 9 Mb de longitud (Figura 5E), y NCNA, pérdida de al menos 5 Mb de longitud, en 6 líneas celulares TNBC (Figura 5F), como se indica.

Las **Figuras 6A-6C** muestran las regiones de AI y la respuesta al cisplatino en el cáncer de mama. La respuesta patológica al cisplatino se evaluó mediante la puntuación de Miller-Payne (MP), que puede oscilar entre 0 (progresión) y 5 (respuesta patológica completa, pCR). La **Figura 6A** muestra representaciones de genomas tumorales individuales ordenados para aumentar la puntuación MP. Se indican las regiones de AI telomérico (azul

oscuro) e AI intersticial (azul claro), con líneas blancas delgadas que demarcan cromosomas individuales. La **Figura 6B** muestra la asociación entre la puntuación MP y el NtAI, 12. La **Figura 6C** muestra una curva de características de funcionamiento del receptor (ROC) que evalúa el rendimiento de NtAI, 12 para predecir la terapia pCR a cisplatino (pCR,  $n = 4$ ; sin pCR,  $n = 20$ ).

5 La **Figura 7** muestra el desequilibrio alélico de todo el cromosoma (isodisomía) y la sensibilidad al cisplatino en los cánceres de mama. Las regiones de AI de todo el cromosoma están indicadas en rojo para cada ubicación cromosómica. Cada fila definida por líneas blancas finas representa un cromosoma diferente y los números de cromosomas están indicados a lo largo del lado izquierdo. Cada columna representa una muestra de tumor individual. La puntuación de la respuesta patológica de Miller-Payne (MP) para cada tumor se indica en la parte inferior. Los casos se ordenan para aumentar la respuesta patológica al cisplatino (0 = progresión, 5 = respuesta patológica completa (pCR)).

10 Las **Figuras 8A-8B** muestran las regiones de AI y el tiempo de 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 NtAI, 12. Aquellos que recayeron dentro de un año están indicados por círculos cerrados y los que no tienen recaída dentro de un año están indicados por círculos abiertos. Un valor de corte de NtAI, 12 = 13, basado en el análisis de ROC TNBC para la predicción de la respuesta patológica completa (pCR) a cisplatino, se indica mediante la línea de puntos. La **Figura 8B** muestra las curvas de supervivencia de Kaplan-Meier para el tiempo de recaída en individuos clasificados como alto NtAI, 12 (13 o más NtAI, 12 regiones, azul) o bajo NtAI, 12 (menos de 13 NtAI, 12 regiones, rojo).

15 La **Figura 9** muestra un modelo que relaciona la reparación del ADN con la acumulación de AI y la respuesta a los agentes de platino. Diversas lesiones genéticas pueden producir defectos en las vías comunes de reparación del ADN, lo que conduce primero a la reparación anormal de las roturas espontáneas del ADN, a la recombinación ilegítima de cromosomas y a la formación de cromosomas cuadrirradiales aberrantes y finalmente a altos niveles de desequilibrio alélico telomérico. En paralelo, la vía de reparación del ADN defectuosa también puede dar como resultado la incapacidad de la célula tumoral para reparar el daño del ADN inducido por el fármaco, lo que conduce a

20 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 indicador de una reparación defectuosa del ADN y predice la sensibilidad al tratamiento con agentes genotóxicos.

25 Las **Figuras 10A-10C** muestran aberraciones cromosómicas y la sensibilidad al cisplatino in vitro. La relación entre las regiones de AI y la sensibilidad al cisplatino se analizó en 10 líneas celulares de cáncer de mama: 1: CAMA-1, 2: HCC1954, 3: MDA-MB-231, 4: MDA-MB-361, 5: HCC1187, 6: BT -549, 7: HCC1143, 8: MDA-MB-468, 9: BT-20, 10: T47D. La **Fig. 10A** muestra los valores de  $Cl_{50}$  para cada una de las líneas celulares 10. Se utilizó un ensayo de proliferación para evaluar la viabilidad después de 48 horas de exposición a cisplatino y se determinó la  $Cl_{50}$  a partir de las curvas de respuesta a la dosis. La **Fig. 10B** muestra una comparación entre el número de regiones con desequilibrio alélico telomérico (NtAI) y la sensibilidad al cisplatino. El subtipo de cáncer de mama se indica de la siguiente manera: TN, rojo; HER2+, verde, ER+ HER2-, azul. La **Fig. 10C** muestra la comparación entre (NtAI) y la sensibilidad al cisplatino según lo determinado por  $GI_{50}$  en líneas celulares de cáncer de mama de Heiser et al. (18). El subtipo de transcripción informado se indica de la siguiente manera: basal, rojo; claudin-bajo, rosa; ERBB2Amp, verde; luminal, azul.

30 Las **Figuras 11A-11D** muestran un NtAI y una respuesta al cisplatino en el cáncer de mama. En dos ensayos clínicos, los pacientes con TNBC recibieron cisplatino preoperatorio (Cisplatino-1, Fig. 11A-11B) o cisplatino y bevacizumab (Cisplatino-2, Fig. 11C-11D). Los tumores sensibles al cisplatino están indicados en rojo, los tumores insensibles al cisplatino están indicados en negro. Los tumores con mutaciones en la línea germinal en BRCA1/2 están indicados con triángulos. La **Fig. 11A** y la **Fig. 11C** muestran diagramas de caja que muestran la distribución de NtAI en tumores sensibles y resistentes al cisplatino. La **Fig. 11B** y la **Fig. 11D** muestran las curvas características operativas del receptor que muestran la capacidad de NtAI para predecir la sensibilidad al cisplatino.

35 La **Figura 12** muestra el NtAI y la respuesta al cisplatino en el cáncer de ovario seroso. Diagramas de caja que muestran la distribución de NtAI en tumores resistentes y sensibles al platino en cánceres sin mutaciones en BRCA1 o BRCA2 (wt-BRCA) y para cánceres con mutación germinal o somática en BRCA1 (mBRCA1) o en BRCA2 (mBRCA2). El rojo indica muestras sensibles, los triángulos indican muestras con mutaciones en la línea germinal o somática en BRCA1 o BRCA2. Se indican diferencias significativas entre los grupos wt-BRCA resistentes y sensibles. Además, se encontraron diferencias significativas entre wt-BRCA sensible y mBRCA2 sensible ( $P = 0,047$ ), y wt-BRCA sensible y mBRCA1 sensible ( $P = 0,014$ ).

40 Las **Figuras 13A-13B** muestran el enriquecimiento de CNV comunes en los puntos de corte cromosómicos tAI de TNBC. Asociación de puntos de corte tAI con loci CNV comunes basados en simulaciones computacionales que compararon el número esperado de puntos de corte que contienen CNV con el número observado en total de casos en Cisplatino-1 (Fig. 13A) y Cisplatino-2 (Fig. 13B).

45 Las **Figuras 14A-14C** muestran la asociación entre la expresión de BRCA1, NtAI y la metilación del promotor de BRCA1. El rojo indica tumores sensibles al cisplatino. Los tumores con una mutación germinal en BRCA1 o BRCA2 se excluyen en la Fig. 14A y 14B, pero están incluidos en la Fig. 14C, representados como triángulos. La **Fig. 14A**

muestra que la expresión de BRCA1 medida por qPCR es significativamente más baja en tumores sensibles en la cohorte Cisplatino-2. La **Fig. 14B** muestra que la expresión de BRCA1 es menor en las muestras que presentan la metilación de la región promotora de BRCA1 en las cohortes combinadas de Cisplatino-1 y Cisplatino-2. La **Fig. 14C** muestra que la expresión de BRCA1 medida por qPCR presenta una correlación negativa con NtAI en las cohortes combinadas de Cisplatino-1 y Cisplatino-2.

Las **Figuras 15A-15D** muestran un modelo que relaciona la reparación del ADN con la acumulación de AI telomérico y la respuesta a agentes de platino. La **Fig. 15A** muestra en las células competentes para la reparación del ADN, que las roturas del ADN se reparan usando recombinación homóloga libre de errores empleando la cromátida hermana idéntica como molde, lo que da como resultado la ausencia de AI. La **Fig. 15B** y la **Fig. 15C** muestran que la reparación del ADN comprometida favorece el uso de vías de reparación propensas a errores, lo que resulta en reordenamientos cromosómicos y la formación aberrante de cromosomas radiales. Después de la división mitótica, las células hijas tendrán un desequilibrio en la contribución parental de los segmentos teloméricos de los cromosomas (AI telomérico). La **Fig. 15B** muestra que la unión final no homóloga es un mecanismo propenso a errores que une una cromátida quebrada de un cromosoma (azul oscuro) con la cromátida de otro cromosoma, generalmente no homólogo (blanco). La segregación mitótica da como resultado células con AI telomérico debido al cambio monoalélico en el número de copias de ADN de la región telomérica afectada. La **Fig. 15C** muestra que la recombinación mitótica puede dar lugar a reordenamientos entre cromosomas homólogos (azul oscuro y azul claro). La segregación mitótica da como resultado células con AI debido a una LOH neutra. Se espera que la replicación inducida por rotura dé lugar a un resultado similar. La **Fig. 15D** muestra el mismo compromiso en la reparación del ADN que causa el AI telomérico, que también puede dar como resultado la incapacidad de la célula tumoral para reparar el daño inducido por el ADN, lo que lleva a la sensibilidad del tumor a fármacos como las sales de platino.

La **Figura 16** muestra una definición de ejemplo de desequilibrio alélico. El diagrama muestra cromosomas bialélicos normales y tres formas diferentes en las que puede producirse un desequilibrio alélico de una región cromosómica.

Las **Figuras 17A-17C** muestran la asociación entre la sensibilidad al cisplatino y las medidas de anomalías genómicas en un panel de líneas celulares de cáncer de mama.  $CI_{50}$  del cisplatino versus: **Fig. 17A**, número total de regiones AI; **Fig. 17B**, número total de regiones de ganancia de número de copias; y la **Fig. 17C**, número total de regiones de pérdida de número de copias. Los números representan las mismas líneas celulares que en la Figura 1.

Las **Figuras 18A-18D** muestran una asociación entre la sensibilidad al cisplatino y las ganancias y pérdidas teloméricas/intersticiales.  $CI_{50}$  del cisplatino versus: **Fig. 18A**, el número de regiones de ganancia de números de copias teloméricas; **Fig. 18B**, el número de regiones de pérdida de números de copias teloméricas; **Fig. 18C**, el número de regiones de ganancia de número de copias intersticiales; y la **Fig. 18D**, el número de regiones de pérdida de número de copias intersticiales. Los números representan las mismas líneas celulares que en la Figura 10 (1).

La **Figura 19** muestra la curva característica operativa del receptor que muestra la capacidad de NtAI para predecir la sensibilidad a la terapia basada en platino en el cáncer de ovario seroso de wt-BRCA.

La **Figura 20** muestra la distribución de las roturas de ADNdc que da como resultado un desequilibrio alélico telomérico y la asociación con CNV comunes según la respuesta al cisplatino. Los cuadrados indican la ubicación cromosómica inferida de las roturas de ADNdc que dan como resultado el tAI, agrupados en ambos ensayos. Los cuadrados apilados representan tumores múltiples con roturas de ADNdc en la misma posición.

Las **Figuras 21A-21B** muestran la expresión de BRCA1 y NtAI frente a la respuesta al cisplatino en Cisplatino-1 y Cisplatino-2 combinados. La **Fig. 21A** muestra la identificación del punto de corte óptimo para NtAI (negro) y el ARNm de BRCA1 (azul) para predecir la respuesta al cisplatino por separado. Los círculos rellenos representan los puntos de corte óptimos. La **Fig. 21B** muestra cómo la combinación de la expresión de BRCA1 y NtAI puede mejorar la predicción de la respuesta al cisplatino. El rojo indica muestras sensibles al cisplatino. Las líneas representan el corte óptimo para la predicción de respuesta basada en NtAI y ARNm de BRCA1, como se determina en la Fig. 21A.

"Sens" representa el número de casos sensibles por el total que se muestran en cada cuadrante, definidos por los valores límite del NtAI y ARNm de BRCA1. La tabla muestra la precisión de predicción basada en los límites definidos para NtAI solo, ARNm de BRCA1 solo y las dos medidas combinadas. ACC: precisión. PPV: valor predictivo positivo. VAN: valor predictivo negativo. SENS: sensibilidad. SPEC: especificidad. P: valor p basado en la prueba exacta de Fisher. Esta tabla se basa solo en las muestras que se muestran en la Fig. 21B.

Las **Figuras 22A-22C** muestran la expresión de BRCA1 por micromatrices de expresión génica en cohortes TCGA. Fig. 22A. La expresión de ARNm de BRCA1 frente a NtAI en los cánceres de mama TCGA ER-/HER2- (n = 78). La Fig. 22B muestra la expresión de ARNm de BRCA1 frente a NtAI en los cánceres de ovario serosos TCGA wt-BRCA (n = 165). La Fig. 22C muestra la expresión de ARNm de BRCA1 frente a la respuesta al tratamiento en los cánceres de ovario serosos TCGA wt-BRCA.

La **Figura 23** muestra un proceso ejemplar mediante el cual un sistema informático puede determinar una puntuación de aberración cromosómica.

La **Figura 24** es un diagrama de un ejemplo de un dispositivo informático 1400 y un dispositivo informático móvil 1450, que se puede usar con las técnicas descritas en este documento.

**DESCRIPCIÓN DETALLADA DE LA INVENCION**

**[0040]** La presente descripción se refiere a procedimientos para predecir la respuesta de un cáncer en un sujeto a terapias contra el cáncer basándose en una determinación y análisis de una puntuación de aberración cromosómica, tal como el número de desequilibrios alélicos o el número de desequilibrios alélicos teloméricos en los cromosomas del genoma humano

**[0041]** De acuerdo con un aspecto divulgado en este documento, Puntuación de Aberración Cromosómica Global (GCAS) es una medida predictiva de la capacidad de respuesta a terapias contra el cáncer de un cáncer en un sujeto. Esta utilidad de GCAS se basa en el nuevo hallazgo de que la suma de las aberraciones cromosómicas individuales puede predecir la capacidad de respuesta de un cáncer en un sujeto a agentes anticancerosos independientemente de la identificación de aberraciones cromosómicas específicas. Los loci informativos de interés (por ejemplo, polimorfismos de un solo nucleótido (SNP), polimorfismos de longitud de fragmentos de restricción (RFLP), repeticiones en tándem simples (STR), etc.), se utilizan para determinar la GCAS ya que son útiles para detectar y/o distinguir aberraciones cromosómicas. Como se usa en el presente documento, "aberración cromosómica" significa desequilibrio alélico (AI), pérdida de heterocigosidad (LOH), aberraciones del número de copias (CNA), ganancia del número de copias (CNG), disminución del número de copias (CND) y combinaciones de los mismos. La GCAS es un tipo de puntuación de aberración cromosómica, de las cuales otros tipos incluyen puntuación de aberración telomérica, puntuación de desequilibrio alélico telomérico, etc. Por lo tanto, a menos que se indique explícitamente lo contrario o salvo que el contexto indique claramente lo contrario, las referencias a la GCAS pueden aplicarse en algunas realizaciones a puntuaciones de aberración cromosómica (por ejemplo, puntuación de aberración telomérica, puntuación de desequilibrio alélico telomérico, etc.).

**[0042]** La GCAS se determina determinando una pluralidad o el número total de regiones cromosómicas que muestran desequilibrio alélico (NAI), pérdida de heterocigosidad (LOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG) y/o disminución del número de copias (NCND), como se describe adicionalmente en este documento y de acuerdo con 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 es predictiva de la respuesta a la terapia contra el cáncer de la célula cancerosa a partir de la cual se derivó el ácido nucleico analizado.

**[0043]** En una realización, el análisis se basa en 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 a los líquidos que normalmente no son excretados o secretados (por ejemplo, líquido amniótico, humor acuoso, bilis, sangre y plasma sanguíneo, líquido cefalorraquídeo, cerumen y cera de los oídos, líquido de Cowper o fluido preeyaculatorio, quilo, quimo, heces, eyaculación femenina, 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 selecciona del grupo que consiste en células, líneas celulares, portaobjetos histológicos, tejidos embebidos en parafina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspado bucal, saliva, fluido cerebroespinal, orina, heces y médula ósea.

**[0044]** En una realización, los SNP se usan para determinar la GCAS, para predecir la capacidad de respuesta de un cáncer a una terapia contra el cáncer. Hay seis posibles tipos de SNP, ya sea transiciones (A <> T o G <> C) o transversiones (A <> G, A <> C, G <> T o C <> T). Los SNP son ventajosos porque se pueden identificar números grandes.

**[0045]** En algunas realizaciones, los SNP u otros loci genómicos pueden puntuarse para detectar anomalías del número de copias. En tales casos, dichos loci genómicos no necesitan ser informativos en términos de genotipo ya que el número de copias está determinado por intensidades de hibridación y no depende del genotipo. Además, las anomalías del número de copias se pueden detectar usando procedimientos que no usan SNP, tales como, por ejemplo, matriz CGH usando matrices de BAC, ADNc y/u oligonucleótidos; marcadores de microsatélites; STRs, RFLPS; etc.

**[0046]** 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 de ácidos nucleicos codificantes en una muestra implica una transferencia de Southern. En una transferencia de Southern, el ADN genómico (normalmente 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 del análisis de ADN genómico normal (por ejemplo, una porción no amplificada de la misma célula, tejido, órgano o una célula, tejido, órgano relacionado, etc.) proporciona una estimación del número de copias relativo del ácido nucleico diana. Como alternativa, se puede utilizar una transferencia de Northern para evaluar el número de copias del ácido nucleico codificador en una muestra. En una transferencia de 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 del análisis de ARNm normal (por ejemplo, una porción no amplificada de la misma célula, tejido, órgano o una célula, tejido, órgano relacionado, etc.) proporciona una estimación del número de copias relativo del ácido nucleico diana. Se pueden realizar procedimientos similares para determinar el número de copias usando matrices transcripcionales, que son bien conocidas en la técnica.

**[0047]** Un medio alternativo para determinar el 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 estructura biológica a analizar; (2) tratamiento de prehibridación de la estructura biológica para aumentar la accesibilidad del ADN diana y 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 los 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 según la aplicación particular.

**[0048]** Ensayos basados en hibridación preferidos incluyen, pero no se limitan a, procedimientos tradicionales de "sonda directa" tales como transferencias de Southern o hibridación *in situ* (por ejemplo, FISH y FISH más SKY), y procedimientos de "sonda comparativa" tales como hibridación genómica comparativa (CGH), por ejemplo, CGH basado en ADNc u oligonucleótidos. Los procedimientos se pueden usar en una amplia variedad de formatos que incluyen, pero no se limitan a, procedimientos unidos al sustrato (por ejemplo, membrana o vidrio) o enfoques basados en matrices.

**[0049]** En un ensayo de hibridación *in situ* típico, las células se fijan a un soporte sólido, normalmente un portaobjetos de vidrio. Si se va a probar un ácido nucleico, las células normalmente se desnaturalizan con calor o álcali. A continuación, las células se ponen en contacto con una solución de hibridación a una temperatura moderada para permitir la hibridación de las sondas marcadas específicas de la secuencia de ácido nucleico que codifica la proteína. Los objetivos (por ejemplo, las células) normalmente luego se lavan a una restricción predeterminada o a una rigurosidad creciente hasta que se obtiene una relación señal/ruido apropiada.

**[0050]** Las sondas normalmente se marcan, por ejemplo, con radioisótopos o informadores fluorescentes. Las sondas preferidas son suficientemente largas para hibridar específicamente con el ácido o ácidos nucleicos objetivo en condiciones rigurosas. El intervalo de tamaño preferido es de aproximadamente 200 bases a aproximadamente 1000 bases.

**[0051]** En algunas aplicaciones, es necesario bloquear la capacidad de hibridación de secuencias repetitivas. Por lo tanto, en algunas realizaciones, se usa ARNt, ADN genómico humano o ADN de Cot-I para bloquear la hibridación no específica.

**[0052]** En los procedimientos de CGH, una primera colección de ácidos nucleicos (por ejemplo, de una muestra, por ejemplo, un posible tumor) se marca con una primera etiqueta, mientras que una segunda colección de ácidos nucleicos (por ejemplo, un control, por ejemplo, de una célula/tejido sano) se marca con una segunda etiqueta. La relación de hibridación de los ácidos nucleicos se determina por la relación de los dos (primero y segundo) marcadores que se unen a cada fibra en la matriz. Cuando hay supresiones o multiplicaciones cromosómicas, se detectarán diferencias en la proporción de las señales de las dos etiquetas y la relación proporcionará una medida del número de copias. El CGH basado en matriz también puede realizarse con marcaje de un solo color (en oposición al marcaje del control y la posible muestra tumoral con dos colorantes diferentes y mezclarlos antes de la hibridación, lo que arrojará una relación debido a la hibridación competitiva de las sondas en las matrices). En CGH de un solo color, el control se marca e hibrida con una matriz y se leen las señales absolutas, y la posible muestra de tumor se marca e hibrida con una segunda matriz (con contenido idéntico) y se leen las señales absolutas. La diferencia de número de copias se calcula en función de 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. No. 430.402; Methods in Molecular Biology, vol. 33: In-situ Hybridization Protocols, Choo, ed., Humana Press,

Totowa, NJ (1994), etc. En una realización, se usa 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).

**[0053]** Los procedimientos de la invención son particularmente adecuados para formatos de hibridación basados en matriz. El CGH basado en matriz se describe en la Patente de los Estados Unidos n.º 6.455.258. En otra realización más, los ensayos basados en amplificación se pueden usar para medir el número de copias. En dichos ensayos basados en amplificación, las secuencias de ácido nucleico actúan como molde 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 molde en la muestra original. La comparación con los controles apropiados, por ejemplo, tejido sano, proporciona una medida del número de copias.

**[0054]** Los procedimientos de amplificación "cuantitativa" son bien conocidos por los expertos en la materia. Por ejemplo, la 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) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. NY. La medición del número de copias de ADN en loci de microsatélites utilizando análisis de PCR cuantitativo se describe en Ginzonger, et al. (2000) *Cancer Research* 60: 5405-5409. La secuencia de ácido nucleico conocida para los genes es suficiente para permitir a un experto en la materia seleccionar rutinariamente cebadores para amplificar cualquier porción del gen. La PCR cuantitativa fluorogénica también puede usarse en los procedimientos de la invención. En la PCR fluorogénica cuantitativa, la cuantificación se basa en la cantidad de señales de fluorescencia, por ejemplo, TaqMan y sybr verde.

**[0055]** Otros procedimientos de amplificación adecuados incluyen, sin limitación, la 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 por puntos, y PCR de adaptador de enlazador, etc.

**[0056]** En otras realizaciones más de los procedimientos proporcionados en este documento, se realiza la secuenciación de 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 puede usar 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 pueden usar los denominados "sistemas de secuenciación de próxima generación" que incluyen, sin limitación, máquinas y/o estrategias de secuenciación bien conocidas en la técnica, tales como las desarrolladas por Illumina/Solexa (el Genome Analyzer, Bennett et al., (2005) *Pharmacogenomics*, 6: 373-20 382), por Applied Biosystems, Inc. (SOLiD Sequencer; solid.appliedbiosystems.com), por Roche (por ejemplo, el secuenciador 454 GS FLX; Margulies et al. (2005) *Nature*, 437: 376-380; Patentes de Estados Unidos n.º 6.274.320; 6.258.568; 6.210.891), mediante el sistema HELISCOPE™ de Helicos Biosciences (ver, por ejemplo, Solicitud de Patente de los Estados Unidos n.º Pub. 2007/0070349) y por otros. También se pueden usar otras estrategias de secuenciación, como la secuenciación estocástica (por ejemplo, desarrollada por Oxford Nanopore), por ejemplo, tal como se describe en la Solicitud Internacional n.º PCT/GB2009/001690 publicada como WO2010/004273. Todas las estrategias de determinación del número de copias descritas en la presente memoria se pueden aplicar de manera similar a cualquiera de los análisis basados en ácidos nucleicos descritos en la presente memoria, tales como loci informativos y similares descritos más adelante.

**[0057]** En otras realizaciones, los SNP pueden puntuarse para heterocigosidad o ausencia de heterocigosidad. Las técnicas como el análisis de proporción de copias principales utilizan el desequilibrio alélico y la información del número de copias para ampliar los análisis que se pueden realizar con el número de copias de los eventos de LOH solo ya que pueden implicar eventos de eliminación, neutros o de ganancia de número de copias. En otras realizaciones, para determinar la GCAS de un cáncer en un sujeto, se identifican SNP heterocigóticos ubicados en todo el genoma usando muestras de ácido nucleico derivadas de tejido no canceroso del sujeto o una población de sujetos de una sola especie, y se determina el número de aquellos SNP heterocigóticos identificados que mantienen heterocigosidad (o alternativamente no exhiben heterocigosidad, es decir, han perdido la heterocigosidad) en una muestra de ácido nucleico, o derivada de ADN genómico de tejido canceroso del sujeto. Una muestra de ácido nucleico "derivada de" ADN genómico incluye, sin limitación, ARN pre-mensajero (que contiene intrones), productos de amplificación de ADN genómico o ARN pre-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 las regiones no codificadoras no deben eliminarse selectivamente).

**[0058]** No es necesario incluir todos los SNP que exhiben heterocigosidad en la especie a la que pertenece el sujeto con cáncer en la cantidad de SNP heterocigóticos utilizados o analizados. En algunas realizaciones, al menos 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 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 SNP o más, o cualquier rango intermedio, u 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 heterocigóticos se distribuye aleatoriamente en todo el genoma al menos cada 1, 5, 10, 50, 100, 250, 500, 1000, 1.500, 2000, 2500, 3000, 5000, 10.000 kb o más, o cualquier rango entre medias. Por "distribuido aleatoriamente", como se usa anteriormente, se entiende que los SNP de la pluralidad no se seleccionan por sesgo hacia ningún locus o loci cromosómicos específicos; sin embargo, se pueden usar otros sesgos (por ejemplo, la evitación de secuencias de ADN repetitivas) en la selección de los SNP. En otras realizaciones, la pluralidad de SNP heterocigóticos no se distribuye aleatoriamente en todo el genoma (es decir, se distribuye en al menos 250, 500, 1000, 1.500, 2000, 2500, 3000, 5000, 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 además pueden estar sesgadas, en algunas realizaciones, a regiones cromosómicas específicas tales como telómeros (a veces denominadas en este documento "regiones teloméricas" o "segmentos teloméricos") definidas como regiones que se extienden hacia el telómero pero que no cruzan el centrómero. En una realización, el tamaño del segmento de desequilibrio alélico telomérico 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 rango intermedio, tal como entre 5 y 25 Mb. En otra realización, el tamaño del segmento de desequilibrio alélico telomérico es de 12 Mb. Por el contrario, las regiones intersticiales no involucran al telómero. Las regiones intersticiales se definen en este documento como regiones de desequilibrio alélico que comienzan aguas abajo del telómero lo que significa que hay al menos una parte del cromosoma sin desequilibrio alélico entre el telómero y la región del desequilibrio alélico. En una realización, la pluralidad de SNP heterocigóticos no se encuentran en regiones de ADN genómico que son repetitivas. En otra realización, la pluralidad de SNP heterocigóticos comprende SNP localizados 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, su región telomérica).

**[0059]** Con muchas técnicas modernas de alto rendimiento (incluidas las tratadas en este documento), es posible determinar el genotipo, número de copias, proporción de copias, etc., para decenas, cientos, miles, millones o incluso miles de millones de loci genómicos (por ejemplo, todos los SNP heterocigóticos conocidos en una especie particular, secuenciación del genoma completo, etc.). Una vez que se ha realizado un ensayo global (por ejemplo, analizando todos o sustancialmente todos los SNP heterocigóticos conocidos), uno puede analizar informáticamente uno o más subconjuntos de loci (es decir, paneles de loci de prueba o, como se usa en este documento algunas veces, pluralidades de loci de prueba). Por lo tanto, en algunas realizaciones, después de analizar el desequilibrio alélico en cientos de loci o más en una muestra (o después de recibir los datos de dicho ensayo), se puede analizar (por ejemplo, informáticamente) un panel o pluralidad de loci de prueba de acuerdo con la presente invención (por ejemplo, SNP completa o principalmente teloméricos) combinando los datos relacionados con los loci de prueba individuales para obtener un valor de prueba indicativo del nivel general, naturaleza, etc. del desequilibrio alélico en el grupo deseado de loci de prueba.

**[0060]** Por lo tanto, en un aspecto la descripción proporciona un procedimiento para derivar una puntuación de aberración cromosómica (por ejemplo, la GCAS, puntuación de aberración telomérica, puntuación de desequilibrio alélico telomérico, etc.) que comprende: determinar si una muestra tiene una aberración cromosómica (por ejemplo, de desequilibrio alélico, pérdida de heterocigosidad, aberraciones del número de copias, ganancia del número de copias, disminución del número de copias) en una pluralidad de loci de prueba (por ejemplo, genómicos); analizar una pluralidad de loci de prueba dentro de dicha pluralidad de loci de prueba para aberraciones cromosómicas; combinar los datos de (2) para derivar una puntuación que refleje la extensión global de la aberración cromosómica en dicha pluralidad de loci de prueba, derivando así una puntuación de aberración cromosómica.

**[0061]** En algunas realizaciones, determinar si una muestra tiene una aberración cromosómica en la pluralidad de loci de prueba comprende analizar una muestra de tejido (por ejemplo, procesar físicamente una

muestra tangible de un paciente para derivar datos de la misma) y analizar los datos (por ejemplo, datos de genotipo de SNP) derivados de dicho ensayo. En algunas realizaciones, determinar si una muestra tiene una aberración cromosómica en la pluralidad de loci de prueba comprende analizar datos derivados de un ensayo en una muestra de tejido.

5

**[0062]** En algunas realizaciones, los loci de prueba representan todos los loci analizados en el ensayo relevante (por ejemplo, todos los SNP heterocigóticos representados en la matriz de SNP particular, todos los nucleótidos en un ensayo de secuenciación). En algunas realizaciones, los loci de prueba representan loci particulares analizados en el ensayo relevante (por ejemplo, ciertos nucleótidos, tales como SNP, en un ensayo de secuenciación). En algunas realizaciones, todos los loci de prueba son loci de prueba. En algunas realizaciones, los loci de prueba representan al menos algún porcentaje (por ejemplo, el 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 100 %) de los loci de prueba. En algunas realizaciones, al menos algún porcentaje (por ejemplo, el 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 100 %) de los loci de prueba son loci teloméricos. En algunas realizaciones, al menos algún porcentaje (por ejemplo, el 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 100 %) de los loci de prueba son loci teloméricos. Así, en algunas realizaciones, la descripción proporciona un procedimiento para derivar una puntuación de aberración cromosómica (por ejemplo, la GCAS, puntuación de aberración telomérica, puntuación de desequilibrio alélico telomérico, etc.) que comprende: determinar si una muestra tiene una aberración cromosómica (por ejemplo, de desequilibrio alélico, pérdida de heterocigosis, aberraciones del número de copias, ganancia del número de copias, disminución del número de copias) en una pluralidad de loci de prueba (por ejemplo, genómicos); analizar una pluralidad de loci de prueba dentro de dicha pluralidad de loci de prueba para aberraciones cromosómicas, donde al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, u 80 %, o 90 %, o 95 %, o 100 %) de los loci de prueba son loci teloméricos; combinar los datos de (2) para obtener una puntuación que refleje la extensión global de la aberración cromosómica en dicha pluralidad de loci de prueba, derivando así una puntuación de aberraciones cromosómicas.

**[0063]** En algunas realizaciones, a cada locus de prueba se le asigna un peso particular en el cálculo de la puntuación de aberración cromosómica. En algunas realizaciones, a los loci de prueba se les asigna un peso por el que cada uno recibe un coeficiente particular en una fórmula (final o intermedia) utilizada para calcular la puntuación de aberración cromosómica. En algunas realizaciones, los loci de prueba teloméricos (todos o al menos el 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % de loci de prueba teloméricos) se ponderan de manera que contribuyan al menos con algún porcentaje (el 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 100 %) a la puntuación de aberración cromosómica. Por lo tanto, en algunas realizaciones, la descripción proporciona un procedimiento para derivar una puntuación de aberración cromosómica (por ejemplo, la GCAS, puntuación de aberración telomérica, puntuación de desequilibrio alélico telomérico, etc.) que comprende: determinar si una muestra tiene una aberración cromosómica (por ejemplo, de desequilibrio alélico, pérdida de heterocigosis, aberraciones del número de copias, ganancia del número de copias, disminución del número de copias) en una pluralidad de loci de prueba (por ejemplo, genómicos); analizar una pluralidad de loci de prueba dentro de dicha pluralidad de loci de prueba para aberraciones cromosómicas; combinar los datos de (2) para derivar una puntuación que refleje la extensión global de la aberración cromosómica en dicha pluralidad de loci de prueba, asignando a cada locus de prueba un coeficiente de ponderación que determina su contribución a la puntuación de aberración cromosómica y en el que se ponderan los loci teloméricos que contribuyen al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) a la puntuación de aberración cromosómica, derivando así una puntuación de aberración cromosómica.

45

**[0064]** "Locus telomérico" como se usa en la presente memoria significa un locus dentro de un telómero o dentro de una distancia definida a lo largo del cromosoma desde el telómero. En algunas realizaciones, un locus telomérico está dentro de 1 Kb, 2 Kb, 3 Kb, 4 Kb, 5 Kb, 6 Kb, 7 Kb, 8 Kb, 9 Kb, 10 Kb, 15 Kb, 20 Kb, 25 Kb, 30 Kb, 35 Kb, 40 Kb, 45 Kb, 50 Kb, 100 Kb, 200 Kb, 300 Kb, 400 Kb, 500 Kb, 750 Kb, 1 Mb, 2 Mb, 3 Mb, 4 Mb, 5 Mb, 6 Mb, 7 Mb, 8 Mb, 9 Mb, 10 Mb, 15 Mb, 20 Mb, 25 Mb, 30 Mb, 35 Mb, 40 Mb, 45 Mb, 50 Mb, 60 Mb, 70 Mb, 80 Mb, 90 Mb o 100 Mb o menos del telómero. En algunas realizaciones, la distancia entre el locus telomérico y el telómero es menor que el 1 %, 2 %, 3 %, 4 %, 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 % o 90 % de la longitud del brazo cromosómico completo.

**[0065]** Por lo tanto, en algunas realizaciones, una región telomérica o segmento telomérico es una región cromosómica que abarca al menos un cierto número de loci teloméricos (por ejemplo, al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 7500, o 10.000 o más loci teloméricos). En algunas realizaciones, una región telomérica o segmento telomérico es una región cromosómica que abarca al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000,

1500, 2000, 2500, 3000, 4000, 5000, 7500, o 10.000 o más loci teloméricos, en el que tales loci teloméricos están dentro de 50 Kb, 100 Kb, 200 Kb, 300 Kb, 400 Kb, 500 Kb, 750 Kb, 1 Mb, 2 Mb, 3 Mb, 4 Mb, 5 Mb, 6 Mb, 7 Mb, 8 Mb, 9 Mb, o 10 Mb del telómero (o cualquier combinación de ese tipo de loci teloméricos y la distancia de dichos loci desde el telómero). En algunas realizaciones, las regiones teloméricas no cruzan el centrómero.

5

**[0066]** La capacidad de reparación del ADN es un factor determinante de la sensibilidad a ciertos medicamentos de quimioterapia, como el cisplatino. Las células cancerosas con reparación intacta de ADN pueden evitar la acumulación de daño en el genoma durante el crecimiento y también pueden reparar el daño del ADN inducido por platino. Se buscaron firmas genómicas indicativas de una reparación defectuosa del ADN en líneas  
10 celulares y tumores, y se correlacionaron estas firmas con la sensibilidad al platino. El número de regiones subcromosómicas con desequilibrio alélico que se extiende a los telómeros (NtAI) predijo la sensibilidad al cisplatino in vitro y la respuesta patológica al tratamiento preoperatorio con cisplatino en pacientes con cáncer de mama triple negativo (TNBC). En el cáncer de ovario seroso tratado con quimioterapia basada en platino, un mayor NtAI pronostica una mejor respuesta inicial. Encontramos una relación inversa entre la expresión de BRCA1 y el número  
15 de regiones de tAI (NtAI) en TNBC esporádicos y cánceres de ovario serosos sin la mutación en BRCA1 o BRCA2. Por lo tanto, la acumulación de tAI es un marcador de sensibilidad al cisplatino y sugiere una reparación deteriorada del ADN, y el NtAI puede ser útil para predecir la respuesta a tratamientos dirigidos a la reparación defectuosa del ADN.

20 **[0067]** Las mutaciones en BRCA1 o BRCA2 causan defectos en la reparación del ADN que predicen la sensibilidad a las sales de platino en cáncer de mama y de ovario; sin embargo, algunos pacientes sin mutaciones en BRCA también se benefician de estos agentes. Este estudio muestra que los defectos en la reparación del ADN que causan la sensibilidad al platino se pueden inferir del número de desequilibrio alélico (NAI), por ejemplo, el número de desequilibrio telomérico (NtAI), una medida de la aberración genómica en los tumores. Hemos  
25 demostrado que NAI y/o NtAI pueden identificar pacientes con cáncer sin mutaciones en BRCA que probablemente se beneficien de la terapia basada en platino, como pacientes con cáncer de mama triple negativo o cáncer de ovario triple negativo.

**[0068]** Las líneas celulares portadoras de mutaciones en BRCA1 o BRCA2 son más sensibles a la muerte por  
30 las sales de platino cisplatino y carboplatino que las células silvestres (Samouelian, et al. Chemosensitivity and radiosensitivity profiles of four new human epithelial ovarian cancer cell lines exhibiting genetic alterations in BRCA2, TGFbeta-RII, KRAS2, TP53 and/or CDNK2A. Cancer Chemother Pharmacol 2004; 54: 497-504, Tassone, et al. BRCA1 expression modulates chemosensitivity of BRCA1-defective HCC1937 human breast cancer cells. Br J Cancer 2003; 88: 1285-1291). Los cánceres de mama y de ovario en pacientes portadores de mutaciones en  
35 BRCA1 o BRCA2 también son sensibles a la quimioterapia basada en platino (Byrski, T., et al. Response to neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients. Breast Cancer Res Treat 2009; 115: 359-363, Cass, I., et al. Improved survival in women with BRCA-associated ovarian carcinoma. Cancer 2003; 97: 2187-2195). La mayoría de los cánceres de mama que surgen en mujeres con una mutación germinal en BRCA1 carecen de expresión de receptores de estrógeno y progesterona o de amplificación del gen HER2-neu ("triple negativo"). Los cánceres de mama relacionados con BRCA1 comparten varias características fenotípicas con el  
40 cáncer de mama triple negativo (TNBC) esporádico (Turner, N. C., et al. BRCA1 dysfunction in sporadic basal-like breast cancer. Oncogene 2007; 26: 2126-2132; Sorlie, T., et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003; 100: 8418-8423; Lakhani, S. R., et al. Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer  
45 Res 2005; 11: 5175-5180). Ambos tipos de tumores comparten un patrón común de anomalías genómicas y tienen altos niveles globales de aberraciones cromosómicas incluyendo desequilibrio alélico (AI), la contribución desigual de las secuencias de ADN materno y paterno con o sin cambios en el número total de copias de ADN (Wang, Z. C., et al. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. Cancer Research 2004; 64: 64-71; Richardson, A. L., et al. X chromosomal abnormalities in basal-like human breast  
50 cancer. Cancer Cell 2006; 9: 121-132; Van Loo, P., et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 2010; 107: 16910-16915). Dado que tienen en común las aberraciones genómicas que sugieren una lesión compartida en el control de la integridad genómica, es razonable postular que el TNBC esporádico que ha acumulado niveles elevados de AI podría compartir la sensibilidad a la quimioterapia basada en platino que caracteriza el cáncer asociado a BRCA1.

55

**[0069]** Se realizó un ensayo clínico, Cisplatino-1, en el que 28 pacientes con TNBC operable se trataron preoperatoriamente con cisplatino en monoterapia. El tratamiento preoperatorio en Cisplatino-1 dio como resultado una reducción del tumor superior al 90 % en 10 de 28 (36 %) pacientes, incluida la respuesta patológica completa (pCR) en 6 mujeres, 2 de las cuales tenían cánceres asociados con BRCA1 (Silver, D. P., et al. Efficacy of

neoadjuvant Cisplatin in triple-negative breast cancer. *J Clin Oncol* 2010; 28: 1145-1153). Un segundo ensayo, Cisplatino-2, acumuló 51 pacientes con TNBC que recibieron el mismo régimen de cisplatino preoperatorio que Cisplatino-1, pero en combinación con el inhibidor de la angiogénesis bevacizumab (Ryan, P. D., et al. Neoadjuvant cisplatin and bevacizumab in triple negative breast cancer (TNBC): Safety and Efficacy. *J Clin Oncol* 2009; 27: 551).

5 La tasa de respuesta en Cisplatino-2 fue similar a Cisplatino-1. En el segundo ensayo, se observó una reducción del tumor superior al 90 % en 17 de las 44 mujeres (39 %) que completaron el tratamiento. En Cisplatino-2, 8 pacientes portaban una mutación en la línea germinal en BRCA1 o BRCA2, de las cuales 4 pacientes lograron una pCR o casi pCR en el régimen de Cisplatino-bevacizumab. En ambos ensayos, todas las pacientes se habían sometido a secuenciación de investigación para determinar su estado germinal de BRCA1 y BRCA2. Por lo tanto, en algunos

10 aspectos de todas las realizaciones de la invención, el estado de BRCA1 y BRCA2 puede determinarse simultáneamente con o antes del análisis de desequilibrio alélico o de desequilibrio alélico telomérico. En algunos aspectos, solo las pacientes sin mutaciones en BRCA1 o BRCA2 se someten a los ensayos o análisis de NtAI (número de desequilibrio alélico telomérico). El cisplatino se utiliza como ejemplo de terapias contra el cáncer que comprenden platino.

15

**[0070]** Comparamos el número de diversas anomalías cromosómicas, incluido el AI presente en las biopsias tumorales obtenidas antes del tratamiento, con la respuesta tumoral determinada por patología al cisplatino, solo o en combinación con bevacizumab, evaluada mediante el examen de la muestra quirúrgica posterior al tratamiento.

20 **[0071]** Sin desear estar sujetos a la teoría, las anomalías cromosómicas, como las regiones de desequilibrio alélico, distintas de las que resultan de la ganancia o pérdida de cromosomas, pueden ser el resultado de una reparación incorrecta de las roturas de la doble cadena del ADN durante el desarrollo tumoral. Si es así, un recuento genómico de regiones cromosómicas anormales en tumores puede indicar el grado de incompetencia de reparación del ADN, independientemente del conocimiento de cualquier defecto específico de reparación del ADN causante.

25 Nuestra hipótesis fue que el número de regiones cromosómicas de AI en tumores predeciría la sensibilidad a fármacos que inducen enlaces cruzados de ADN como el cisplatino.

**[0072]** Se buscaron asociaciones entre diversas medidas de anomalías subcromosómicas y la sensibilidad al cisplatino en líneas celulares de cáncer de mama y se encontró el predictor más preciso de AI que se extiende hasta

30 el extremo telomérico del cromosoma (NtAI). Finalmente, probamos si NtAI se asociaba a la respuesta al tratamiento en muestras tumorales de pacientes en los ensayos Cisplatino-1 y Cisplatino-2 de TNBC y en el conjunto de datos públicos de cáncer de ovario seroso del Cancer Genome Atlas (TCGA), un cáncer tratado rutinariamente con terapia basada en platino. En un esfuerzo por comprender más acerca de los procesos que dan lugar a desequilibrios alélicos teloméricos, mapeamos la ubicación de sus puntos críticos y observamos una asociación sorprendente de

35 estos puntos de corte con regiones del genoma que son difíciles de replicar, las variantes comunes de número de copias (CNV). Además, un subconjunto de tumores con alto NtAI muestra bajos niveles de ARNm de BRCA1. Estas observaciones comienzan a sugerir modelos de cómo puede ocurrir el tAI.

**[0073]** Mostramos que la sensibilidad al cisplatino se correlaciona con la carga, es decir, aumenta el número de desequilibrios alélicos teloméricos en comparación con las células normales, por ejemplo, en líneas celulares de

40 cáncer de mama. Se obtuvieron datos de matriz de genotipo de polimorfismo de nucleótido único (SNP) del Wellcome Trust Sanger Institute para un conjunto de líneas celulares de cáncer de mama de tipo silvestre BRCA1 establecidas para las que se había determinado la sensibilidad al cisplatino (Li, Y Li, Y., et al. Amplification of LPTM4B and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer. *Nat Med* 2010; 16:

45 214-218) (Fig. 10A). El número de copias de alelos se determinó a partir de los datos de la matriz de SNP y AI se detectó usando el análisis ASCAT (10) (Fig. 16), aunque también se puede usar cualquier otro análisis de desequilibrio alélico. Se probó la asociación entre los valores de  $CI_{50}$  para el cisplatino y cada una de tres medidas de resumen de alteración cromosómica: el número de regiones cromosómicas con AI (NAI, Fig. 17A), el número de regiones con ganancia de número de copias (NGain, Fig. 17B), y el número de regiones con pérdida de número de

50 copias (NLoss, Fig. 17C). Ninguna de estas medidas se correlacionó con la sensibilidad al cisplatino en las líneas celulares.

**[0074]** Los defectos conocidos en la reparación de roturas de la doble cadena del ADN, incluida la pérdida de BRCA1, provocan la formación espontánea de estructuras trirradiales y cuadrirradiales cromosómicas, que son

55 indicaciones citológicas de recombinación cromosómica aberrante (Silver, D. P., et al. Further evidence for BRCA1 communication with the inactive X chromosome. *Cell* 2007; 128: 991-1002; Luo, G., et al. Cancer predisposition caused by elevated mitotic recombination in Bloom mice. *Nat Genet* 2000; 26: 424-429; Xu, X., et al., Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell* 1999; 3: 389-395). La resolución de estos reordenamientos cromosómicos en la mitosis

puede dar lugar a grandes regiones de AI y/o cambios en el número de copias que se extienden desde el cruce al telómero (Luo, G., et al. Cancer predisposition caused by elevated mitotic recombination in Bloom mice. *Nat Genet* 2000; 26: 424-429; Vrieling, H. Mitotic maneuvers in the light. *Nat Genet* 2001; 28: 101-102). Más generalmente, varios procesos de reparación propensos a errores potencialmente empleados por células con reparación defectuosa de ADN causan eventos cruzados cromosómicos o de elección de copias que dan como resultado pérdida alélica o cambio del número de copias que se extiende desde el sitio de daño del ADN al telómero.

**[0075]** Se buscó una asociación entre la sensibilidad al cisplatino y el número de regiones contiguas de AI, ganancia de copia o pérdida de copia que se extendió a un telómero y no cruzó el centrómero (regiones teloméricas) o no se extendió a un telómero (regiones intersticiales) (Fig. 16, Fig. 10B y Fig. 18). El número de regiones de AI telomérico (NtAI) fue la única medida genómica de resumen que se asoció significativamente con la sensibilidad al cisplatino en las líneas celulares de cáncer de mama ( $r = 0,76$   $P = 0,011$ , Fig. 10B); la correlación entre el NtAI y la sensibilidad al cisplatino fue más fuerte cuando el análisis se restringió a las líneas de cáncer de mama triple negativo (Fig. 1B, círculos rojos;  $r = 0,82$   $P = 0,0499$ ). Se observó una relación similar entre el NtAI y la sensibilidad al cisplatino medida por GI50 en un estudio recientemente publicado de líneas celulares de cáncer de mama ( $r = 0,57$   $P = 0,0018$ , Fig. 10C) (18). De todos los fármacos probados en este estudio, el NtAI fue el que más se correlacionó con la sensibilidad al cisplatino.

**[0076]** Mostramos que los tumores sensibles a la quimioterapia basada en cisplatino tienen niveles más altos de desequilibrio alélico telomérico. Investigamos si la asociación entre el NtAI en muestras clínicas de tumores y la sensibilidad al cisplatino estaba presente en el ensayo Cisplatino-1. La sensibilidad se midió por la respuesta patológica determinada después del tratamiento preoperatorio (Silver, D. P., et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. *J Clin Oncol* 2010; 28: 1145-1153). Los datos del genotipo de SNP de la sonda de inversión molecular de muestras de tumores de pretratamiento ( $n = 27$ ) se evaluaron mediante análisis ASCAT para determinar el NtAI. Comparamos tumores con una reducción de al menos un 90 % en el contenido de células malignas (sensibles al cisplatino) a tumores con respuesta limitada o nula al cisplatino (resistentes al cisplatino, definido por una reducción tumoral de menos del 90 %).

**[0077]** Mostramos que los tumores sensibles al cisplatino tenían significativamente mayor NtAI (mediana de 24 frente a 17,5,  $P = 0,047$ , Fig. 11A) Probamos la capacidad de NtAI para predecir la respuesta al cisplatino mediante el cálculo del área bajo la curva (AUC) característica operativa del receptor (ROC). El análisis ROC mostró que un mayor NtAI se asoció con la sensibilidad al cisplatino (AUC = 0,74, CI 0,50-0,90, Fig. 11B).

**[0078]** En el ensayo Cisplatino-2, los tumores sensibles al cisplatino ( $n = 9$ ) tuvieron un NtAI significativamente más alto que los tumores resistentes ( $n = 17$ , mediana de 27 frente a 20,  $P = 0,019$ , Fig. 11C). NtAI también se asoció con la respuesta a cisplatino y bevacizumab por análisis ROC (AUC = 0,79, CI 0,55-0,93, Fig. 11D). La asociación entre el NtAI y la sensibilidad al cisplatino siguió siendo significativa cuando se excluyeron los casos con mutación en BRCA1 o BRCA2 y solo se analizaron los casos normales de BRCA ( $P = 0,030$  y  $P = 0,023$  en Cisplatino-1 y Cisplatino-2, respectivamente). Por lo tanto, en dos ensayos preoperatorios separados en cáncer de mama, en los que se evaluó la sensibilidad del tratamiento mediante una medida cuantitativa de la respuesta patológica, el NtAI pronosticó con fiabilidad la respuesta al tratamiento basado en cisplatino.

**[0079]** Para evaluar si la métrica de NtAI indica sensibilidad al platino en cánceres distintos al de mama, determinamos la asociación entre el NtAI y la respuesta de tratamiento inicial en la cohorte de pacientes con cáncer de ovario seroso del The Cancer Genome Atlas (TCGA) que recibieron quimioterapia adyuvante con platino y taxano (Bell, D., et al., Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; 474: 609-615). Nuevamente, entre los cánceres de ovario sin mutación en BRCA1 o BRCA2 (wt-BRCA), los tumores sensibles al platino tuvieron un NtAI significativamente más alto que los cánceres resistentes al platino (mediana 22 frente a 20,  $P = 0,036$ , Fig. 12), y fueron predictivos de la respuesta al tratamiento por análisis ROC (AUC = 0,63, CI 0,50-0,76, Fig. 19). Los cánceres de ovario con mutación somática o de la línea germinal en BRCA1 o BRCA2 que eran sensibles a la terapia de platino tenían incluso mayor NtAI (mediana = 26,  $P = 0,0017$  y mediana de 23,5,  $P = 0,037$  frente al wt-BRCA resistente, respectivamente, Fig. 12). Todos los cánceres mutados en BRCA2 eran sensibles al platino; sin embargo, 5 tumores mutados en BRCA1 fueron resistentes a la terapia con platino y parecían tener niveles relativamente altos de NtAI. Por lo tanto, un NtAI alto es característico del cáncer de ovario seroso con mutación conocida en BRCA1 o BRCA2; el alto NtAI también se encuentra en un subconjunto de cánceres esporádicos sin mutaciones en BRCA donde es predictivo de sensibilidad al platino.

**[0080]** Por consiguiente, proporcionamos un procedimiento para seleccionar una terapia para un paciente humano con cáncer, el procedimiento que comprende analizar una muestra que comprende células tumorales

tomadas del paciente humano con cáncer por el número de desequilibrio alélico, por ejemplo, desequilibrio alélico telomérico, y seleccionar, y opcionalmente administrar una terapia contra el cáncer que comprende platino al paciente humano con cáncer si aumenta el número de desequilibrios alélicos en comparación con un valor de referencia. El valor de referencia para el número de desequilibrio alélico, como el desequilibrio alélico telomérico, puede ser, por ejemplo, de al menos 20, al menos 21, al menos 22, al menos 23, al menos 23,5, al menos 24, al menos 25, a al menos 26, al menos 27, al menos 28 al menos 29, o al menos 30. El valor de referencia se puede determinar para cada tumor, por ejemplo, a partir del número de desequilibrios alélicos recogidos de cánceres similares que son resistentes al platino. Entonces, por ejemplo, en un cáncer de pulmón, las muestras de células de cáncer de pulmón de cánceres resistentes al platino pueden proporcionar un número medio de desequilibrio alélico para que el cáncer se use como valor de referencia para muestras que no responden.

**[0081]** Además, demostramos que las ubicaciones de roturas cromosómicas asociadas al NtAI no son aleatorias. Para comprender mejor los procesos que dan lugar al tAI, mapeamos la ubicación de los puntos de corte cromosómicos que definen el límite de las regiones de tAI. Observamos que muchos puntos de rotura se ubicaron muy cerca el uno del otro (Fig. 20), lo que sugiere una distribución no aleatoria de las roturas del ADN que causan un desequilibrio alélico telomérico.

**[0082]** Sin desear estar ligados por una teoría, los puntos de rotura de translocación cromosómica recurrentes se pueden asociar con regiones de secuencias de ADN repetidas que pueden provocar horquillas de replicación estancadas, una mayor frecuencia de roturas de ADN y posterior reordenamiento por recombinación homóloga no alélica u otros mecanismos similares (Kolomietz, E., et al., The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. *Genes Chromosomes Cancer* 2002; 35: 97-112; Hastings, P. J., Ira, G., and Lupski, J. R. A microhomology-mediated break- induced replication model for the origin of human copy number variation. *PLoS Genet* 2009; 5: e1000327).

**[0083]** Las variantes de número de copias (CNV) son secuencias de ADN altamente homólogas para las cuales el número de copias de la línea germinal varía entre individuos sanos (Iafrate, A. J., et al. Detection of large-scale variation in the human genome. *Nat Genet* 2004; 36: 949-951; Sebat, J., et al., Large-scale copy number polymorphism in the human genome. *Science* 2004;305: 525-528). Se han propuesto CNV para facilitar la generación de alteraciones cromosómicas, similares a sitios frágiles (Hastings, P. J., Ira, G., and Lupski, J. R. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* 2009; 5: e1000327; Stankiewicz, P., et al. Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet* 2003; 72: 1101-1116; Hastings, P. J., et al., Mechanisms of change in gene copy number. *Nat Rev Genet* 2009; 10: 551-564). Comparamos el número de interrupciones observadas dentro de los 25 kB de una CNV con la frecuencia esperada solo por casualidad, en base a los datos permutados. En la cohorte Cisplatino-1, de 517 puntos de corte de NtAI, 255 (49 %) se asociaron con CNV superpuestas. De manera similar, en la cohorte de Cisplatino-2, de 599 puntos de corte de NtAI, 340 (57 %) se asociaron con CNV. En ambos ensayos, el número observado de roturas de NtAI asociadas con CNV fue significativamente mayor de lo esperado por casualidad (Fig. 13A-13B). Por lo tanto, muchos de los puntos de rotura que dan lugar al AI telomérico en TNBC ocurren cerca de las CNV, lo que sugiere que las horquillas de replicación estancadas, el estrés de replicación u otros mecanismos asociados a CNV pueden estar implicados en la génesis del AI telomérico.

**[0084]** Por consiguiente, en algunos aspectos de todas las realizaciones de la descripción, proporcionamos un procedimiento o un ensayo para determinar si un paciente responde a la terapia que comprende platino analizando el número de desequilibrio alélico, tal como desequilibrio alélico telomérico, en el que el AI se asocia con variaciones en el número de copias (CNV). Si el aumento de CNV se asocia a NAI, tal como NtAI, entonces se determina que el paciente responde al tratamiento contra el cáncer que comprende platino y, opcionalmente, la terapia que comprende platino se administra al paciente con cáncer. Si no se detecta un aumento en el NAI asociado a CNV, como NtAI, entonces se determina que el paciente con cáncer no responde a la terapia contra el cáncer que comprende platino y opcionalmente administra al paciente con cáncer un tratamiento contra el cáncer que no contiene platino.

**[0085]** Demostramos que el ARNm bajo de BRCA1 se asocia con un alto NtAI y sensibilidad al cisplatino.

**[0086]** Por consiguiente, en algunos aspectos de todas las realizaciones de la descripción, proporcionamos un ensayo o un procedimiento para determinar la capacidad de respuesta de un paciente con cáncer a una terapia contra el cáncer que comprende platino, el ensayo o procedimiento que comprende analizar en una célula cancerosa que comprende una muestra tomada del paciente con cáncer la cantidad de desequilibrio alélico y/o la cantidad de ARNm de BRCA1, y si aumenta el número de desequilibrio alélico y/o disminuye la cantidad de ARNm de BRCA1,

entonces se selecciona y opcionalmente se administra la terapia contra el cáncer que comprende platino al paciente con cáncer. Si, por otro lado, no se detecta un aumento en el número de desequilibrio alélico y/o no disminuye la cantidad de ARNm de BRCA1, entonces se selecciona, y opcionalmente se administra a dicho paciente con cáncer un tratamiento contra el cáncer que no contiene platino.

5

**[0087]** En nuestro ensayo Cisplatino-1, encontramos una asociación entre bajos niveles de transcripción de BRCA1 y una mejor respuesta a cisplatino. En el ensayo Cisplatino-2, los niveles de transcripción de BRCA1 medidos, por ejemplo, por qPCR también se asocian con la respuesta al cisplatino ( $P = 0,015$ , Fig. 14A). En un análisis combinado de los datos de ambos ensayos, los niveles más bajos de transcripción de BRCA1 están asociados con la metilación del promotor de BRCA1 ( $P = 0,027$ , Fig. 14B), aunque la metilación del promotor de BRCA1 en sí no está significativamente asociada con la respuesta al cisplatino ( $P = 0,25$ , prueba exacta de Fisher). Los niveles de ARNm de BRCA1 están inversamente asociados con el NtAI en los dos ensayos de cisplatino ( $r = -0,50$ ,  $P = 0,0053$ , Fig. 14C). Este hallazgo sugiere que la disfunción de un proceso dependiente de BRCA1 u otra anomalía que causa ARNm de BRCA1 bajo puede ser responsable del alto nivel de desequilibrio alélico telomérico y también de la sensibilidad al cisplatino en muchos de estos TNBC.

**[0088]** En algunos aspectos de todas las realizaciones descritas en este documento, los ensayos y procedimientos comprenden analizar el estado de metilación de BRCA1, donde el aumento en la metilación de la región promotora de BRCA1 está asociado con una respuesta a terapia contra el cáncer que comprende platino y ningún aumento en la metilación de la región promotora de BRCA1 está asociado con la resistencia a la terapia contra el cáncer que comprende platino. Si se detecta una mayor metilación de la región promotora de BRCA1, entonces se selecciona, y opcionalmente se administra, una terapia que comprende platino para el paciente con cáncer, y si no se detecta un aumento en la metilación de la región promotora de BRCA1, se selecciona y opcionalmente se administra una terapia contra el cáncer que no comprende platino para el paciente con cáncer. En algunos aspectos de esta realización, el cáncer es cáncer de mama.

**[0089]** El análisis ROC de los ensayos combinados de TNBC sugiere que el nivel de expresión de BRCA1 o NtAI puede proporcionar una precisión predictiva similar para la sensibilidad al cisplatino (Fig. 21A). Cuando se combinan un alto NtAI y una baja expresión de BRCA1 en un modelo predictivo, el valor predictivo positivo y la especificidad de la predicción mejoraron considerablemente, pero la sensibilidad disminuyó en relación con el NtAI solo (Fig. 21B), lo que sugiere que la baja expresión de BRCA1 no representa todos los tumores sensibles al cisplatino.

**[0090]** En los ensayos de TNBC, observamos algunos tumores sensibles al cisplatino con altos niveles de NtAI pero con alto ARNm de BRCA1, lo que sugiere que, en algunos tumores, mecanismos alternativos pueden impulsar la generación de tAI. El análisis de los datos de TCGA de cáncer de mama ER-/HER2- y cáncer de ovario seroso de wt-BRCA demuestra una correlación inversa entre la expresión de NtAI y BRCA1. Sin embargo, en ambas cohortes hubo un subconjunto considerable de tumores con alto NtAI y alta expresión de BRCA1 (Fig. 22A, 22B). A diferencia de NtAI, la expresión de BRCA1 no fue aparentemente diferente entre los cánceres de ovario serosos wt-BRCA sensibles y resistentes (Fig. 22C). Estos hallazgos sugieren un modelo mediante el cual un NtAI elevado puede representar una lectura de la deficiencia de reparación del ADN que resulta de la expresión baja de BRCA1 o de otros mecanismos conocidos o desconocidos (Fig. 15).

**[0091]** En algunos aspectos de todas las realizaciones de la invención, el análisis de NtAI se realiza solo sin detectar o determinar por separado el estado de BRCA1 y/o BRCA2, tal como si la célula tumoral portase una mutación en BRCA1 y/o BRCA2 o si la expresión de BRCA1 o BRCA2 se encuentra disminuida o no, o si la metilación del promotor de BRCA1 y/o BRCA2 está aumentada o no. En algunos aspectos de todas las realizaciones de la invención, el análisis de NtAI se realiza en combinación con la detección o determinación del estado de BRCA1 y/o BRCA2.

50

**[0092]** Varias realizaciones de la invención descrita en este documento implican una etapa de correlación del número de tAI según la presente invención (por ejemplo, el número total de regiones de tAI en al menos un par de cromosomas humanos de dicha célula cancerosa que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de tAI, donde dicho al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que dicha primera longitud es de aproximadamente 1,5 o más megabases) para una característica clínica particular (por ejemplo, una mayor probabilidad de una deficiencia en el gen BRCA1 o BRCA2, una mayor probabilidad de deficiencia de HDR, una mayor probabilidad de respuesta a un régimen de tratamiento que comprende un agente dañino para el ADN, una antraciclina, un inhibidor de la topoisomerasa I, radiación y/o un inhibidor de PARP, etc.) si el número es mayor que

55

alguna referencia (u opcionalmente a otra característica si el número es menor que alguna referencia). A lo largo de este documento, dondequiera que se describa dicha realización, otra realización de la invención puede implicar, además o en lugar de una etapa correlativa, una o ambas de las siguientes etapas: (a) concluir que el paciente tiene la característica clínica basada al menos en parte en el aumento o no del número de tAI; o (b) comunicar que el paciente tiene la característica clínica basada al menos en parte en el aumento de NtAI.

**[0093]** A modo de ilustración, pero no de limitación, una realización descrita en este documento es un procedimiento para predecir la respuesta de un paciente de cáncer a un régimen de tratamiento contra el cáncer que comprende un agente que daña el ADN, una antraciclina, un inhibidor de la topoisomerasa I, radiación y/o un inhibidor PARP, dicho procedimiento que comprende: (1) determinar, en una célula cancerosa de dicho paciente con cáncer, el número de regiones de LOH/AI/tAI en al menos un par de cromosomas humanos de una célula cancerosa de dicho paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH/AI/tAI, donde dicho al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, donde dicha primera longitud es de aproximadamente 1,5 o más megabases; y (2) correlacionar dicho número total que es mayor que un número de referencia con una mayor probabilidad de que dicho paciente con cáncer responda a dicho régimen de tratamiento contra el cáncer. Según el párrafo anterior, se entiende que esta descripción de esta realización incluye una descripción de dos realizaciones relacionadas, es decir, un procedimiento para predecir la respuesta de un paciente de cáncer a un régimen de tratamiento contra el cáncer que comprende un agente que daña el ADN, una antraciclina, un inhibidor de la topoisomerasa I, radiación, y/o un inhibidor de PARP, dicho procedimiento que comprende: (1) determinar, en una célula cancerosa de dicho paciente con cáncer, el número de regiones de LOH/AI/tAI en al menos un par de cromosomas humanos de una célula cancerosa de dicho paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud de todo el cromosoma que contiene la región de LOH/AI/tAI, en el que dicho al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que dicha primera longitud es de aproximadamente 1,5 o más megabases; y (2)(a) concluir que dicho paciente tiene una mayor probabilidad de que dicho paciente con cáncer responderá a dicho régimen de tratamiento contra el cáncer basándose, al menos en parte, en un número total que es mayor que un número de referencia; o (2)(b) comunicar que dicho paciente tiene una mayor probabilidad de que dicho paciente con cáncer responda a dicho régimen de tratamiento contra el cáncer basándose al menos en parte en un número total que es mayor que un número de referencia.

**[0094]** En cada realización descrita en este documento, se correlaciona un ensayo o resultado de análisis particular (por ejemplo, número total de regiones de LOH/AI/tAI mayores que un número de referencia, etc.) con alguna probabilidad (por ejemplo, aumentada, no aumentada, disminuida, etc.) de alguna característica clínica (por ejemplo, respuesta a un tratamiento particular, muerte específica por cáncer, etc.), o adicional o alternativamente concluir o comunicar dicha característica clínica basada al menos en parte en dicho ensayo o resultado de análisis particular, dicha correlación, conclusión o comunicación que puede comprender asignar un riesgo o probabilidad de que la característica clínica se produzca basándose, al menos en parte, en el ensayo o resultado de análisis particular. En algunas realizaciones, dicho riesgo es un porcentaje de probabilidad de que se produzca el evento o resultado. En algunas realizaciones, el paciente se asigna a un grupo de riesgo (por ejemplo, bajo riesgo, riesgo intermedio, alto riesgo, etc.). En algunas realizaciones, "bajo riesgo" es cualquier porcentaje de probabilidad por debajo del 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 % o 50 %. En algunas realizaciones, "riesgo intermedio" es cualquier porcentaje de probabilidad superior al 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 % o 50 % e inferior al 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 % o 75 %. En algunas realizaciones, "alto riesgo" es cualquier porcentaje de probabilidad superior al 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 %.

**[0095]** Tal como se usa en el presente documento, "comunicar" una determinada información significa hacer conocer dicha información a otra persona o transferir dicha información a una cosa (por ejemplo, un ordenador). En algunos procedimientos de la invención, se comunica el pronóstico de un paciente o la probabilidad de respuesta a un tratamiento particular. En algunas realizaciones, se comunica la información utilizada para llegar a dicho pronóstico o predicción de respuesta (por ejemplo, NtAI según la presente invención, etc.). Esta comunicación puede ser auditiva (por ejemplo, verbal), visual (por ejemplo, escrita), electrónica (por ejemplo, datos transferidos de un sistema informático a otro), etc. En algunas realizaciones, comunicar una clasificación de cáncer (por ejemplo, pronóstico, probabilidad de respuesta, tratamiento apropiado, etc.) comprende generar un informe que comunique la clasificación del cáncer. En algunas realizaciones, el informe es un informe impreso, un informe auditivo o un registro electrónico. En algunas realizaciones, el informe se visualiza y/o almacena en un dispositivo informático (por ejemplo, dispositivo portátil, ordenador de escritorio, dispositivo inteligente, sitio web, etc.). En algunas realizaciones, la clasificación del cáncer se comunica a un médico (por ejemplo, un informe que comunica la clasificación se

proporciona al médico). En algunas realizaciones, la clasificación del cáncer se comunica a un paciente (por ejemplo, un informe que comunica la clasificación se proporciona al paciente). La comunicación de una clasificación de cáncer también se puede lograr transfiriendo información (por ejemplo, datos) que incorporan la clasificación a un ordenador servidor y permitiendo que un intermediario o usuario final acceda a dicha información (por ejemplo, mediante la visualización de la información como se muestra desde el servidor, mediante la descarga de la información en forma de uno o más archivos transferidos desde el servidor al intermediario o dispositivo del usuario final, etc.).

**[0096]** Siempre que una realización de la invención comprenda concluir algún hecho (por ejemplo, el pronóstico de un paciente o la probabilidad de respuesta del paciente a un régimen de tratamiento particular), esto puede incluir en algunas realizaciones un programa informático que concluya dicho hecho, normalmente después de realizar un algoritmo que aplique información a regiones de tAI de acuerdo con la presente invención.

**[0097]** En cada realización descrita en este documento que implica varias regiones de LOH (por ejemplo, regiones indicadoras de LOH) o una longitud total combinada de dichas regiones de LOH, la presente descripción abarca una realización relacionada que implica un valor o puntuación de prueba (por ejemplo, puntuación HRD, puntuación LOH, NAI, NtAI, etc.) derivada de, que incorpora y/o, al menos hasta cierto punto, refleja dicho número o longitud. En otras palabras, los números o longitudes de la región de tAI desnudo no tienen por qué usarse en los diversos procedimientos, sistemas, etc. de la invención; se puede usar un valor de prueba o puntuación derivado de dichos números o longitudes. Por ejemplo, una realización de la descripción proporciona un procedimiento para tratar el cáncer en un paciente, que comprende: (1) determinar en una muestra de dicho paciente el número de regiones de LOH/Al/tAI en al menos un par de cromosomas humanos de una célula cancerosa del paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud total del cromosoma que contiene la región de LOH/Al/tAI indica que las células cancerosas tienen la firma de LOH o el número de Al o tAI, donde el al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que la primera longitud es de aproximadamente 1,5 o más megabases; (2) proporcionar un valor de prueba derivado del número de dichas regiones de LOH/Al/tAI; (3) comparar dicho valor de prueba con uno o más valores de referencia derivados del número de dichas regiones de LOH/Al/tAI en una población de referencia (por ejemplo, media, mediana, terciles, cuartiles, quintiles, etc.); y (4)(a) administrar a dicho paciente un medicamento contra el cáncer, o recomendar o prescribir o iniciar un régimen de tratamiento que comprende quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revela que el valor de la prueba es mayor (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9, o 10 veces mayor, al menos 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor) que al menos uno de dicho valor de referencia; o (4)(b) recomendar o prescribir o iniciar un régimen de tratamiento que no comprende quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revela que el valor de la prueba no es mayor (por ejemplo, no más de 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor; no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor) que al menos uno de dichos valores de referencia. La invención abarca, mutatis mutandis, realizaciones correspondientes donde el valor de prueba o puntuación se usa para determinar el pronóstico del paciente, la probabilidad de respuesta del paciente a un régimen de tratamiento particular, la probabilidad del paciente o de la muestra del paciente de tener una deficiencia de BRCA1, BRCA2, RAD51C o HDR, etc.

**[0098]** En un aspecto, la descripción proporciona un kit que comprende, en un recipiente, reactivos adecuados para determinar el desequilibrio alélico en al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 7500, o 10.000 o más loci teloméricos (por ejemplo, loci dentro de 50 Kb, 100 Kb, 200 Kb, 300 Kb, 400 Kb, 500 Kb, 750 Kb, 1 Mb, 2 Mb, 3 Mb, 4 Mb, 5 Mb, 6 Mb, 7 Mb, 8 Mb, 9 Mb o 10 Mb del telómero). En algunas realizaciones, el kit comprende reactivos para determinar el desequilibrio alélico en no más de 10.000, 7500, 5000, 4000, 3000, 2000, 1000, 750, 500, 400, 300, 200, 150, 100, 90, 80, 70, 60, o 50 loci totales. En algunas realizaciones, los loci teloméricos comprenden al menos el 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 99 %, o 100 % de los loci para los cuales el kit contiene reactivos para determinar el desequilibrio alélico.

**[0099]** Como se ha descrito anteriormente, el desequilibrio alélico es un tipo de aberración cromosómica. Como se usa en el presente documento, "desequilibrio alélico" significa un cambio en el número y/o tipo de alelos de un cromosoma en un tejido somático en comparación con la línea germinal. En algunas realizaciones, el desequilibrio alélico es la pérdida de heterocigosidad ("LOH"). Esto puede ser un número de copias neutro, como cuando uno de los alelos parentales heterocigóticos se pierde y el otro alelo se duplica como "reemplazo". La LOH también puede ocurrir en una forma sin un número de copias neutro, donde un alelo parental simplemente se pierde. En algunas realizaciones, el desequilibrio alélico es la duplicación de un alelo sobre otro (AA/B somático de la línea germinal AA/B) o mayor duplicación de un alelo en comparación con otro (por ejemplo, AAAA/BB somático de la

línea germinal A/B). En algunas realizaciones, una región tiene desequilibrio alélico si los loci en esa región muestran MCP (como se analiza en mayor detalle en la Sección IV.H. a continuación) que es mayor que 0,51, 0,52, 0,53, 0,54, 0,55, 0,56, 0,57, 0,58, 0,59, 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, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98, o 0,99 (incluyendo MCP de 1).

**[0100]** Por lo tanto, se puede analizar un número predefinido de cromosomas para determinar el número total de regiones de LOH indicadoras, preferiblemente el número total de regiones de LOH de una longitud mayor que 9 megabases, 10 megabases, 12 megabases, 14 megabases, más preferiblemente mayores que 15 megabases. Adicional o alternativamente, los tamaños de todas las Regiones de LOH Indicadoras identificadas se pueden sumar para obtener una longitud total de las Regiones de LOH Indicadoras.

**[0101]** Para la clasificación del estado positivo de firma de LOH, el número de referencia discutido anteriormente para el número total de Regiones de LOH Indicadoras puede ser de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20 o superior, preferiblemente 5, preferiblemente 8, más preferiblemente 9 o 10, lo más preferiblemente 10. El número de referencia para la longitud total (por ejemplo, combinada) de las regiones de LOH indicadoras puede ser de aproximadamente 75, 90, 105, 120, 130, 135, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 megabases o más, preferiblemente de aproximadamente 75 megabases o más, preferiblemente de aproximadamente 90 o 105 megabases o más, más preferiblemente de aproximadamente 120 o 130 megabases o más, y más preferiblemente de aproximadamente 135 megabases o más, y lo más preferiblemente de aproximadamente 150 megabases o más.

**[0102]** En algunas realizaciones específicas, se determina el número total de regiones de LOH de una longitud mayor que aproximadamente 14 o 15 megabases y se compara con un número de referencia de aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19 o 20. Adicional o alternativamente, la longitud total de las regiones de LOH de una longitud mayor que 14 o 15 megabases se determina y compara con un número de referencia de aproximadamente 75, 90, 105, 120, 130, 135, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475 o 500 megabases.

**[0103]** En algunas realizaciones, el número de regiones de LOH (o la longitud combinada, o un valor de prueba o puntuación derivado de cualquiera de ellos) en una muestra de paciente se considera "mayor" que una referencia si es al menos 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor que la referencia, mientras que en algunas realizaciones, se considera "mayor" si es al menos 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor que la referencia. Por el contrario, en algunas realizaciones, el número de regiones de LOH (o la longitud combinada, o un valor de prueba o puntuación derivado de cualquiera de ellos) en una muestra de paciente se considera "no mayor" que una referencia si no es más que 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor que la referencia, mientras que en algunas realizaciones, se considera "no mayor" si no es más de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor que la referencia.

**[0104]** En algunas realizaciones, el número de referencia (o longitud, valor o puntuación) se deriva de una población de referencia relevante. Dichas poblaciones de referencia pueden incluir pacientes (a) con el mismo cáncer que el paciente analizado, (b) con el mismo subtipo de cáncer, (c) con cáncer que tiene características genéticas u otras características clínicas o moleculares similares, (d) que respondieron a un tratamiento en particular, (e) que no respondieron a un tratamiento en particular, (f) que aparentemente están sanos (por ejemplo, no tienen cáncer o al menos no tienen el cáncer del paciente probado), etc. El número de referencia (o longitud, valor o puntuación) puede ser (a) representativo del número (o longitud, valor o puntuación) encontrado en la población de referencia como un todo, (b) un promedio (media, mediana, etc.) del número (o longitud, valor o puntuación) encontrado en la población de referencia como un todo o una subpoblación en particular, (c) representativo del número (o longitud, valor o puntuación) (por ejemplo, un promedio tal como media o mediana) encontrado en terciles, cuartiles, quintiles, etc. de la población de referencia clasificada por (i) su número respectivo (o longitud, valor o puntuación) o (ii) la característica clínica que se encontró que tenían (por ejemplo, fuerza de respuesta, pronóstico (incluido el tiempo hasta la muerte específica del cáncer), etc.).

**[0105]** Como se describe en este documento, los pacientes que tienen células cancerosas identificadas con un estado positivo de firma de LOH o aumento en el NAI o el NtAI pueden clasificarse, basándose al menos en parte en un estado positivo de firma de LOH o aumento en el NAI o el NtAI, como probable de responder a un régimen de tratamiento de cáncer en particular. Por ejemplo, los pacientes que tienen células cancerosas con un genoma que contiene una firma de LOH o aumento en el NAI o el NtAI pueden clasificarse, basándose al menos en parte en un estado positivo de firma de LOH o aumento en el NAI o el NtAI, como probable de responder a un régimen de

tratamiento contra el cáncer que incluye el uso de un agente que daña el ADN, un agente de letalidad sintético (por ejemplo, un inhibidor de PARP), radiación o una combinación de los mismos. Preferiblemente, los pacientes son pacientes sin tratamiento previo.

5 **[0106]** Los ejemplos de agentes que dañan el ADN incluyen, sin limitación, fármacos de quimioterapia basados en platino (por ejemplo, cisplatino, carboplatino, oxaliplatino y picoplatino), antraciclinas (por ejemplo, epirrubicina y doxorubicina), inhibidores de la topoisomerasa I (por ejemplo, campotecina, topotecán e irinotecán), reticulantes de ADN tales como mitomicina C y compuestos de triazeno (por ejemplo, dacarbazina y temozolomida). Los enfoques terapéuticos de letalidad sintética normalmente implican la administración de un agente que inhibe al  
 10 menos un componente crítico de una vía biológica que es especialmente importante para la supervivencia de una célula tumoral particular. Por ejemplo, cuando una célula tumoral tiene una vía de reparación homóloga deficiente (por ejemplo, según se determina de acuerdo con la presente invención), inhibidores de poli ADP ribosa polimerasa (o fármacos de platino, inhibidores de la reparación de la rotura de la doble cadena, etc.) pueden ser especialmente potentes contra dichos tumores debido a que se obstruyen dos vías críticas para la supervivencia (una  
 15 biológicamente, por ejemplo, mediante la mutación en BRCA1, y la otra sintéticamente, por ejemplo, mediante la administración de un fármaco para la vía). Los enfoques de letalidad sintética para la terapia contra el cáncer se describen en, por ejemplo, O'Brien et al., Converting cancer mutations into therapeutic opportunities, EMBO MOL. MED. (2009) 1:297-299.

20 **[0107]** Los ejemplos de agentes de letalidad sintética incluyen, sin limitación, inhibidores de PARP o inhibidores de la reparación de la rotura de la doble cadena en células tumorales deficientes en reparación homóloga, inhibidores de PARP en células tumorales deficientes en PTEN, metotrexato en células tumorales deficientes en MSH2, etc. Ejemplos de inhibidores de PARP incluyen, sin limitación, olaparib, iniparib y veliparib. Los ejemplos de inhibidores de la reparación de la rotura de la doble cadena incluyen, sin limitación, KU55933 (inhibidor  
 25 de ATM) y NU7441 (inhibidor de ADN-PKcs). Los ejemplos de información que pueden usarse además de un estado positivo de firma de LOH para basar una clasificación de probabilidad de responder a un régimen de tratamiento de cáncer particular incluyen, sin limitación, resultados de tratamiento previos, mutaciones de ADN germinal o somático, perfiles de expresión de genes o proteínas (por ejemplo, estado de ER/PR/HER2, niveles de PSA), histología tumoral (por ejemplo, adenocarcinoma, carcinoma de células escamosas, carcinoma papilar seroso, carcinoma  
 30 mucinoso, carcinoma ductal invasivo, carcinoma ductal *in situ* (no invasivo), etc.), etapa de la enfermedad, tumor o grado de cáncer (por ejemplo, bien, moderadamente, o pobremente diferenciado (por ejemplo, Gleason, Bloom Richardson modificado), etc.), el número de ciclos de tratamiento anteriores, etc.

**[0108]** Además de predecir la posible respuesta al tratamiento o seleccionar regímenes de tratamiento  
 35 deseables, se puede utilizar una firma de LOH o un aumento en el NAI o el NtAI para determinar el pronóstico de un paciente. Hemos demostrado que los pacientes cuyos tumores tienen una firma de LOH o aumentan el NAI o el NtAI muestran una supervivencia significativamente mejor que los pacientes cuyos tumores no tienen dicha firma de LOH o aumento en el NAI o el NtAI. Por lo tanto, en un aspecto, este documento presenta un procedimiento para determinar el pronóstico de un paciente basado al menos en parte en la detección de la presencia o ausencia de una  
 40 firma de LOH o un aumento en el NAI o el NtAI en una muestra del paciente. El procedimiento comprende, o consiste esencialmente en, (a) determinar si el paciente comprende células cancerosas que tienen una firma de LOH o un aumento en el NAI o el NtAI como se describe en este documento (por ejemplo, en el que la presencia de más de un número de referencia de regiones de LOH en al menos un par de cromosomas humanos de una célula cancerosa del paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud del  
 45 cromosoma completo que contiene la región de LOH o la región AI o tAI indica que las células cancerosas tienen la firma de LOH o un aumento en el NAI o el NtAI, donde el al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que la primera longitud es de aproximadamente 1,5 o más megabases), y (b)(1) determinando, basado al menos en parte en la presencia de la firma de LOH o aumento en el NAI o el NtAI, que el paciente tiene un pronóstico relativamente bueno, o (b)(2) determinando, basado al menos en parte en la  
 50 ausencia de la firma de LOH o aumento en el NAI o NtAI, que el paciente tiene un pronóstico relativamente malo.

**[0109]** El pronóstico puede incluir la probabilidad de supervivencia del paciente (por ejemplo, supervivencia libre de progresión, supervivencia global), donde un pronóstico relativamente bueno incluiría una mayor probabilidad de supervivencia en comparación con una población de referencia (por ejemplo, paciente promedio con este  
 55 tipo/subtipo de cáncer, paciente promedio que no tiene una firma de LOH o aumento en el NAI o el NtAI, etc.). Por el contrario, un pronóstico relativamente malo en términos de supervivencia incluiría una menor probabilidad de supervivencia en comparación con alguna población de referencia (por ejemplo, paciente promedio con el tipo/subtipo de cáncer de este paciente, paciente promedio con una firma de LOH o aumento en el NAI o el NtAI, etc.).

**[0110]** "Desequilibrio alélico telomérico" significa un desequilibrio alélico en una región o segmento telomérico. El "desequilibrio alélico" en una región o una "región de desequilibrio alélico" significa un desequilibrio alélico en al menos un cierto número de loci que definen (en todo o en parte) dicha región. Estos generalmente se deben distinguir de loci aislados de desequilibrio alélico. Por lo tanto, en algunas realizaciones las regiones de desequilibrio alélico se definen como al menos 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500 o más sondas consecutivas que muestran un desequilibrio alélico.

**[0111]** El número o proporción de regiones teloméricas que tienen desequilibrio alélico se puede utilizar para derivar una puntuación de desequilibrio alélico telomérico (a veces denominado en este documento NtAl), que es análogo a la puntuación de aberración cromosómica descrita anteriormente (incluido la GCAS) y particularmente a continuación, aunque centrado en regiones teloméricas. Así, en algunas realizaciones, la descripción proporciona un procedimiento para derivar una puntuación de desequilibrio alélico telomérico que comprende: determinar si una muestra tiene una aberración cromosómica (por ejemplo, de desequilibrio alélico, pérdida de heterocigosidad, aberraciones del número de copias, ganancia del número de copias, disminución del número de copias) en una pluralidad de loci de prueba (por ejemplo, genómicos); analizar una pluralidad de loci de prueba dentro de dicha pluralidad de loci de prueba para aberraciones cromosómicas; combinar los datos de (2) para derivar una puntuación que refleje la extensión global de la aberración cromosómica en dicha pluralidad de loci de prueba, derivando así una puntuación de aberración cromosómica.

**[0112]** En algunas realizaciones, los datos se combinan en (3) de tal manera que a cada locus de prueba se le asigna un coeficiente de ponderación que determina su contribución a la puntuación de aberración cromosómica y los loci teloméricos se ponderan de modo que contribuyan al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) a la puntuación de la aberración cromosómica. En algunas realizaciones, al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) de los loci de prueba en (2) son loci teloméricos.

**[0113]** En algunas realizaciones, la puntuación del desequilibrio alélico telomérico contará todas las regiones teloméricas que muestren un desequilibrio alélico. En algunas realizaciones, esto incluirá regiones de desequilibrio alélico que abarcan un cromosoma completo. En algunas realizaciones, la puntuación del desequilibrio alélico telomérico contará todas las regiones teloméricas de al menos un tamaño mínimo (por ejemplo, 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 rango intermedio, tal como entre 5 y 25 Mb) que muestra un desequilibrio alélico.

**[0114]** En algunas realizaciones, al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75 o 100 o más regiones teloméricas con desequilibrio alélico (por ejemplo, una puntuación de desequilibrio alélico telomérico alto (por ejemplo, puntuación de al menos 1, 2, 3, 4, 5, 10, 15, 20, 22, 23, 235, 24, 25, 26, 27, 50, 75, o 100 o más)) indica una mayor probabilidad de respuesta a terapia que comprende una modalidad particular (por ejemplo, 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; cisplatino, treosulfán y trofosfamida; alcaloides vegetales: vinblastina, paclitaxel, docetaxol; inhibidores de la ADN topoisomerasa: tenipósido, crinatól y mitomicina; anti-folatos: 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: halichondrina, colchicina y rizoxina; y derivados sintéticos de los mismos).

**[0115]** Por lo tanto, en algunas realizaciones la descripción proporciona un procedimiento para predecir si un paciente responderá a un tratamiento particular que comprende: determinar si una muestra tiene una aberración cromosómica (por ejemplo, de desequilibrio alélico, pérdida de heterocigosidad, aberraciones del número de copias, ganancia del número de copias) disminución del número de copias) en una pluralidad de loci de prueba (por ejemplo, genómicos); analizar una pluralidad de loci de prueba dentro de dicha pluralidad de loci de prueba para aberraciones cromosómicas; combinar los datos de (2) para derivar una puntuación de aberración cromosómica que refleja la extensión global de la aberración cromosómica en dicha pluralidad de lugares de prueba; correlacionar una puntuación de aberración cromosómica alta con una mayor probabilidad de respuesta a un tratamiento particular.

**[0116]** En algunas realizaciones, los datos se combinan en (3) de tal manera que a cada locus de prueba se le asigna un coeficiente de ponderación que determina su contribución a la puntuación de aberración cromosómica y los loci teloméricos se ponderan de modo que contribuyan al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) a la puntuación de la aberración cromosómica.

o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) a la puntuación de la aberración cromosómica. En algunas realizaciones, al menos el 5 % (o 10 %, o 15 %, o 20 %, o 30 %, o 40 %, o 50 %, o 60 %, o 70 %, o 80 %, o 90 %, o 95 %, o 100 %) de los loci de prueba en (2) son loci teloméricos. En algunas realizaciones, la puntuación del desequilibrio alélico telomérico contará todas las regiones teloméricas que muestren un desequilibrio alélico. En algunas realizaciones, la puntuación del desequilibrio alélico telomérico contará todas las regiones teloméricas de al menos un tamaño mínimo (por ejemplo, 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 rango intermedio, tal como entre 5 y 25 Mb) que muestra un desequilibrio alélico. En algunas realizaciones, una puntuación de desequilibrio alélico telomérico es alta si al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75 o 100 regiones teloméricas tienen desequilibrio alélico.

**[0117]** Los SNP heterocigóticos que pueden usarse en los procedimientos descritos en este documento para determinar el fenotipo de un cáncer son informativos, lo que significa que se observa heterocigosidad en la muestra de ácido nucleico de tejido y/o células no cancerosas de un sujeto. De acuerdo con los procedimientos de la descripción, 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. En una realización adicional, las muestras de ácido nucleico usadas para determinar el número de SNP heterocigóticos en la pluralidad de SNP, que exhiben 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 entenderá que los controles apropiados se pueden determinar basándose en la frecuencia promedio de los alelos SNP que existen dentro del mismo grupo étnico de la especie a la que pertenece el paciente. En determinadas realizaciones, los SNP informativos usados en los procedimientos de la descripción 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 usados en los procedimientos de la descripción 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).

**[0118]** 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 que incluye, pero no se limita a, quimioterapia, radioterapia, inmunoterapia, inhibidores de moléculas pequeñas, shRNA, hormonas y combinaciones de los mismos. Cuando la GCAS representa supresiones de copia, ganancias de copias, pérdidas de cromosomas completos, ganancias de cromosomas completos y/o pérdida de heterocigosidad, se predice que los sujetos cuyo tejido canceroso muestra una GCAS por debajo de un valor umbral tienen una peor respuesta al tratamiento (por ejemplo, radiación o quimioterapia) que aquellos con una GCAS alta (por encima del valor umbral). Cuando GCAS representa la falta de cambios en el número de copias o cromosomas y/o la retención de heterocigosidad, se predice que los sujetos cuyo tejido canceroso exhibe una GCAS por encima de un valor umbral tienen una peor respuesta a la terapia (por ejemplo, radiación o quimioterapia) que aquellos con una baja GCAS (por debajo del valor umbral).

**[0119]** A modo de explicación, pero sin estar sujetos a ninguna teoría, se cree que cuando el valor de la GCAS representa pérdida de heterocigosidad o desequilibrio alélico, identifica células que albergan roturas de la doble cadena de ADN cromosómico incorrectamente reparadas y el recuento genómico completo de estos reordenamientos cromosómicos en un tumor específico indica el grado de incompetencia de la reparación del ADN, independientemente del defecto de reparación del ADN causal específico. En dichos sujetos, el número total de reordenamientos cromosómicos en un tumor indica la incapacidad para reparar el daño en el ADN inducido por terapias contra el cáncer y, en consecuencia, predice la sensibilidad a tales terapias contra el cáncer. También a modo de explicación y sin estar ligados por la teoría, se cree que la GCAS que representa ganancias de copias puede indicar defectos genéticos distintos o adicionales a los defectos de reparación del ADN y que la GCAS que representa la pérdida o ganancia del cromosoma completo puede indicar defectos del punto de control mitótico o defectos de segregación cromosómica, y similares. Se ha determinado que dichas aberraciones en la reparación fidedigna del ADN, la segregación, el control del punto de control, etc. son predictivas de las células que albergan dichas aberraciones al tratamiento con terapias contra el cáncer (por ejemplo, quimioterapéuticos) en sujetos.

**[0120]** 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 tumoral después del inicio de la quimioterapia neoadyuvante o adyuvante. La respuesta tumoral puede evaluarse en una situación neoadyuvante o adyuvante donde el tamaño de un tumor después de la intervención sistémica se puede comparar con el tamaño y las dimensiones iniciales medidos por TC, PET, mamografía, ultrasonido o palpación y se puede estimar la celularidad

de un tumor histológicamente y en comparación con la celularidad de una biopsia tumoral tomada antes del inicio del tratamiento. La respuesta también se puede evaluar mediante la medición con un calibre o el examen patológico del tumor después de una biopsia o resección quirúrgica. La respuesta se puede registrar de forma cuantitativa como un cambio porcentual en el volumen tumoral o celularidad o usando un sistema de puntuación semicuantitativo tal como carga residual de cáncer (Symmans et al., J. Clin. Oncol. (2007) 25: 4414-5 4422 ) o puntuación de Miller-Payne (Ogston et al., Breast (Edimburgo, Escocia) (2003) 12: 320-327) en una forma cualitativa como "respuesta patológica completa" (pCR), "remisión clínica completa" (cCR), "remisión parcial clínica" (cPR), "enfermedad estable clínica" (cSD), "enfermedad progresiva clínica" (cPD) u otros criterios cualitativos. La evaluación de la respuesta tumoral se puede realizar temprano después del inicio de la terapia neoadyuvante o adyuvante, por ejemplo, después de unas pocas horas, días, semanas o preferiblemente después de algunos meses. Un criterio de valoración típico para la evaluación de la respuesta es tras la finalización de la quimioterapia neoadyuvante o tras la extirpación quirúrgica de las células tumorales residuales y/o el lecho tumoral.

**[0121]** Los criterios adicionales para evaluar la respuesta a terapias contra el cáncer están relacionados con la "supervivencia", que incluye todo lo siguiente: supervivencia hasta la muerte, también conocida como supervivencia global (en la que dicha mortalidad puede ser independiente de la causa o relacionada con el tumor); "supervivencia libre de recurrencia" (en la que el término recurrencia incluirá recurrencia tanto localizada como distante); supervivencia libre de metástasis; supervivencia libre de enfermedad (en la que el término enfermedad incluirá cáncer y enfermedades asociadas a él). La duración de dicha supervivencia se puede calcular por referencia a un punto de inicio definido (por ejemplo, tiempo de diagnóstico o inicio del tratamiento) y punto final (por ejemplo, muerte, recurrencia o metástasis). Además, los criterios para la 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 dado, y la probabilidad de recurrencia tumoral.

**[0122]** Por ejemplo, para determinar los valores umbral apropiados, se puede administrar un régimen terapéutico anticanceroso particular a una población de sujetos y el resultado se puede correlacionar con las GCAS que se determinaron antes de la administración de cualquier terapia contra el cáncer. La medida de resultado puede ser la respuesta patológica a la terapia administrada en el entorno neoadyuvante. Como alternativa, las medidas de resultado, como la supervivencia global y la supervivencia libre de enfermedad se pueden controlar durante un período de tiempo para los sujetos que siguen una terapia contra el cáncer para quienes se conocen los valores de GCAS. En ciertas realizaciones, se administran las mismas dosis de agentes anticancerosos a cada sujeto. En realizaciones relacionadas, las dosis administradas son dosis convencionales conocidas en la técnica para agentes anticancerosos. El período de tiempo durante el cual se supervisan los sujetos puede variar. Por ejemplo, los sujetos pueden controlarse 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 Ejemplo.

**[0123]** En algunas realizaciones, el valor de prueba que representa la puntuación de aberración cromosómica se compara con uno o más valores de referencia (o valores de índice), y opcionalmente se correlaciona con una mayor (o no) probabilidad de respuesta a un tratamiento particular. Por ejemplo, el valor del índice puede representar la puntuación de aberración cromosómica promedio para un conjunto de individuos de una población de cáncer diversa o un subconjunto de la población. Por ejemplo, uno puede determinar la puntuación de aberración cromosómica promedio en un muestreo aleatorio de pacientes con cáncer (o un cáncer particular). Esta puntuación de aberración cromosómica promedio puede denominarse "valor de índice de umbral", y los pacientes que tienen una puntuación de aberración cromosómica superior a este valor tienen una probabilidad de respuesta mayor que los que tienen una puntuación de aberración cromosómica inferior a este valor. En algunas realizaciones, el valor de prueba se correlaciona con una mayor probabilidad de respuesta a un tratamiento particular si el valor de prueba excede el valor de referencia en al menos alguna cantidad (por ejemplo, al menos 0,5, 0,75, 0,85, 0,90, 0,95, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 o más veces o desviaciones típicas).

**[0124]** Como alternativa, el valor del índice puede representar la puntuación de aberración cromosómica promedio en una pluralidad de pacientes de entrenamiento con resultados similares cuyos datos clínicos y de seguimiento están disponibles y son suficientes para definir y categorizar a los pacientes por resultado, por ejemplo, respuesta a un tratamiento particular. Ver, por ejemplo, Ejemplos, infra. Por ejemplo, se puede generar un "valor de índice de respuesta" a partir de una pluralidad de pacientes de cáncer de entrenamiento caracterizados por tener "respuesta" al tratamiento particular. Puede generarse un "valor de índice sin respuesta (o pobre)" a partir de una pluralidad de pacientes de entrenamiento con cáncer definidos como "sin respuesta (o pobre)" al tratamiento particular. Por lo tanto, un valor de índice de pronóstico de respuesta puede representar la puntuación promedio de aberración cromosómica en pacientes que tienen "respuesta", mientras que un valor de índice sin respuesta (o

pobre) representa la puntuación de aberración cromosómica promedio en pacientes que tienen una respuesta nula o pobre. Por lo tanto, cuando la puntuación de aberración cromosómica determinada está más cerca del valor del índice de respuesta que del valor del índice sin respuesta, se puede concluir que es más probable que el paciente responda. Por otro lado, si la puntuación de aberración cromosómica 1 determinada está más cerca del valor del índice sin respuesta que del valor del índice de respuesta, entonces se puede concluir que el paciente no tiene una mayor probabilidad de respuesta.

**[0125]** El pronóstico puede incluir la probabilidad de supervivencia del paciente (por ejemplo, supervivencia libre de progresión, supervivencia global), donde un pronóstico relativamente bueno incluiría una mayor probabilidad de supervivencia en comparación con una población de referencia (por ejemplo, paciente promedio con este tipo/subtipo de cáncer, paciente promedio que no tiene una firma de LOH, paciente que no tiene un número mayor de AI o tAI, etc.). Por el contrario, un pronóstico relativamente malo en términos de supervivencia incluiría una menor probabilidad de supervivencia en comparación con alguna población de referencia (por ejemplo, paciente promedio con el tipo/subtipo de cáncer de este paciente, paciente promedio con una firma de LOH, paciente sin mayor número de AI o tAI, etc.).

#### Agentes terapéuticos contra el cáncer

**[0126]** La eficacia de las terapias contra el cáncer que dañan el ADN, así como los agentes que aprovechan los defectos de la reparación del ADN pero que no dañan el ADN, como los inhibidores de la poli ADP ribosa polimerasa (PARP), así como la quimioterapia o la radioterapia, se predice según el nivel de GCAS de un cáncer en un sujeto de acuerdo con los procedimientos descritos en este documento. En una realización, se predice la eficacia de las quimioterapias. 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 antimetabólicos, agentes alquilantes, compuestos de arsénico, inhibidores de la ADN topoisomerasa, taxanos, análogos de nucleósidos, alcaloides vegetales y toxinas; y derivados sintéticos de los mismos. Los ejemplos de compuestos incluyen, pero no se limitan a, agentes alquilantes: cisplatino, treosulfán y trofosfamida; alcaloides vegetales: vinblastina, paclitaxel, docetaxol; inhibidores de la ADN topoisomerasa: tenipósido, crinátol y mitomicina; anti-folatos: metotrexato, ácido micofenólico e hidroxiurea; análogos de pirimidina: 5-fluorouracilo, doxifluridina y citosina arabinósido; análogos de purina: mercaptopurina y tioguanina; antimetabolitos de ADN: 2'-desoxi-5-fluorouridina, glicinato de afidicolina y pirazoloimidazol; y agentes antimetabólicos: halicondrina, colchicina y rizoxina. También se pueden usar composiciones que comprenden uno o más agentes quimioterapéuticos (por ejemplo, FLAG, CHOP). FLAG comprende fludarabina, citosina arabinósido (Ara-C) y G-CSF. CHOP comprende ciclofosfamida, vincristina, doxorubicina y prednisona. En otras realizaciones, se usan 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 los Estados Unidos Re. 36.397) y NU1025 (Bowman et al.). Los ejemplos anteriores de agentes quimioterapéuticos son ilustrativos, y no pretenden ser limitantes.

**[0127]** En una realización preferida, los agentes quimioterapéuticos son compuestos de platino o terapias contra el cáncer que comprenden platino, tales como cisplatino, carboplatino, oxaliplatino, nedaplatino e iroplatino. Otros compuestos de coordinación de platino antineoplásicos son bien conocidos en la técnica, pueden modificarse según procedimientos bien conocidos en la técnica e incluyen los compuestos descritos en las Patentes de Estados Unidos n.º 4.996.337, 4.946.954, 5.091.521, 5.434.256, 5.527.905 y 5.633.243. En otra realización, la GCAS predice la eficacia de la radioterapia. La radiación utilizada en la radioterapia puede ser radiación ionizante. La radioterapia también puede ser rayos gamma, rayos X o haces de protones. Los ejemplos de radioterapia incluyen, entre otros, radioterapia de haz externo, implantación intersticial de radioisótopos (I-125, paladio, iridio), radioisótopos como estroncio-89, radioterapia torácica, radioterapia intraperitoneal con P-32 y/o radioterapia total abdominal y pélvica. Para una descripción general de la radioterapia, ver Hellman, Capítulo 16: Principles of Cancer Management: Radiation Therapy, 6th edition, 2001, DeVita et al., eds., J. B. Lippencott Company, Philadelphia. La radioterapia se puede administrar como radiación de haz externo o teleterapia en la que la radiación se dirige desde una fuente remota. El tratamiento de radiación también se puede administrar como 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 incluye el uso de la terapia fotodinámica que comprende la administración de fotosensibilizadores, tales como hematoporfirina y sus derivados, Vertoporfirina (BPD-MA), ftalocianina, fotosensibilizador Pc4, demetoxi-hipocrelina A; y 2BA-2-DMHA.

**[0128]** Las terapias anticancerosas que dañan el ADN en menor medida que la quimioterapia o la radioterapia pueden tener eficacia en sujetos que se determina 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 inmunoterapia, terapia hormonal y terapia génica. Dichas terapias incluyen, pero no se limitan a, el uso de polinucleótidos antisentido, ribozimas, moléculas de interferencia de ARN, 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 vinculados a la iniciación, progresión y/o patología de un tumor o cáncer. Por ejemplo, en dichas terapias se pueden usar oncogenes, genes de factores de crecimiento, genes receptores del factor de crecimiento, genes del ciclo celular, genes de reparación del ADN y otros.

**[0129]** La inmunoterapia puede comprender, por ejemplo, el uso de vacunas contra el cáncer y/o células presentadoras de antígenos sensibilizadas. La inmunoterapia puede implicar inmunidad pasiva para la protección a corto plazo de un huésped, lograda mediante la administración de anticuerpos preformados dirigidos contra un antígeno cancerígeno o antígeno de enfermedad (por ejemplo, la administración de un anticuerpo monoclonal, opcionalmente unido a un agente quimioterapéutico o toxina, a un antígeno tumoral). La inmunoterapia también puede enfocarse en el uso de epítomos citotóxicos reconocidos por linfocitos de líneas celulares de cáncer.

**[0130]** Los tratamientos terapéuticos hormonales pueden comprender, por ejemplo, agonistas hormonales, antagonistas hormonales (por ejemplo, flutamida, bicalutamida, tamoxifeno, raloxifeno, acetato de leuprolida (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 la vitamina A (por ejemplo, ácido retinoico todo trans (ATRA)); análogos de vitamina D3; antigestágenos (por ejemplo, mifepristona, onapristona) o antiandrógenos (por ejemplo, acetato de ciproterona).

**[0131]** En una realización, la terapia contra el cáncer usada para cánceres cuyo fenotipo se determina por los procedimientos de la descripción puede comprender uno o más tipos de terapias descritas en este documento que incluyen, pero no se limitan a, agentes quimioterapéuticos, 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.

**[0132]** La duración y/o dosis del tratamiento con terapias contra el cáncer puede variar de acuerdo con el agente anticanceroso particular o una combinación de los mismos. El experto en la materia apreciará un tiempo de tratamiento apropiado para un agente terapéutico contra el cáncer en particular. La descripción contempla la evaluación continua de programas de tratamiento óptimos para cada agente terapéutico contra el cáncer, donde el fenotipo del cáncer del sujeto como se determina por los procedimientos descritos en la presente memoria es un factor en la determinación de dosis y programas de tratamiento óptimos.

Cánceres para los que se puede determinar el fenotipo

**[0133]** Los procedimientos de la invención se pueden usar para determinar el fenotipo de muchos cánceres diferentes. Los ejemplos específicos de tipos de cánceres para los que se puede determinar el fenotipo mediante los procedimientos abarcados por la invención incluyen, pero no se limitan a, sarcomas y carcinomas humanos, por ejemplo, fibrosarcoma, miosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, cordoma, angiosarcoma, endoteliosarcoma, linfangiosarcoma, linfangioendoteliosarcoma, sinovioma, mesotelioma, tumor de Ewing, leiomiomasarcoma, rhabdomiomasarcoma, 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 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 de cuello uterino, 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 crónica (granulocítica) y leucemia linfocítica crónica); y policitemia vera, linfoma (enfermedad de Hodgkin y enfermedad no Hodgkin), mieloma múltiple, macroglobulinemia

de Waldenstrom y enfermedad de la cadena pesada.

**[0134]** En algunas realizaciones, las células cancerosas son células cancerosas primarias o metastásicas de cáncer de ovario, cáncer de mama, cáncer de pulmón o cáncer de esófago.

5

**[0135]** En algunas realizaciones, el cáncer cuyo fenotipo se determina por el procedimiento de la invención es un cáncer epitelial tal como, pero no limitado a, cáncer de vejiga, cáncer de mama, cáncer de cuello uterino, cáncer de colon, cáncer ginecológico, cáncer renal, cáncer de laringe, pulmón cáncer, cáncer oral, 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 otras formas de realización más, el cáncer epitelial es cáncer de pulmón de células no pequeñas, carcinoma de células renales no papilar, carcinoma de cuello uterino, carcinoma de ovario (por ejemplo, carcinoma de ovario seroso) o carcinoma de mama. Los cánceres epiteliales se pueden caracterizar de diversas formas que incluyen, pero no se limitan a, serosa, endometrioide, mucinosa, de células claras, brenner o indiferenciada.

10

15

Sujetos

**[0136]** En todas las realizaciones de la invención, el sujeto para el que se determina la eficacia prevista de una terapia contra el cáncer es un ser humano. En una realización de los procedimientos de la invención, 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 cisplatino, carboplatino y/o taxano). En realizaciones relacionadas, el sujeto no ha sido expuesto a niveles de radiación o agentes quimiotóxicos por encima de los encontrados generalmente o en promedio por los sujetos de una especie. En ciertas realizaciones, el sujeto ha tenido cirugía para eliminar tejido canceroso o precanceroso. En otras realizaciones, el tejido canceroso no se ha eliminado, por ejemplo, el tejido canceroso puede estar localizado en una región inoperable del cuerpo, tal como en un tejido que es esencial para la vida, o en una región donde un procedimiento quirúrgico causaría un riesgo considerable de daño al paciente.

20

25

**[0137]** En algunas realizaciones, los pacientes son pacientes sin tratamiento previo.

30

**[0138]** Según un aspecto de la descripción, la GCAS se puede usar para determinar el fenotipo, es decir, la capacidad de respuesta a la terapia de un cáncer en un sujeto, en el que el sujeto se haya sometido previamente a quimioterapia, radioterapia o haya estado expuesto a radiación o un agente quimiotóxico. Dicha terapia o exposición podría dañar el ADN y alterar el número de SNP heterocigóticos informativos en un sujeto. El número alterado a SNP heterocigóticos informativos alteraría la GCAS de un sujeto. Debido a que las muestras de ADN no cancerosas exhibirían mayores o menores SNP heterocigóticos, el rango de GCAS se alteraría para una población de sujetos. En ciertas realizaciones, el daño en el ADN por terapia o exposición en un sujeto o población de sujetos ocurre aproximadamente 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 la GCAS. Para determinar los valores umbrales de GCAS para sujetos que exhiben daño en el ADN por terapia o 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.

35

40

Preparación de muestra de ácido nucleico

45

*Aislamiento de ácidos nucleicos*

**[0139]** Las muestras de ácido nucleico derivadas de células cancerosas y no cancerosas de un sujeto que pueden usarse 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, los procedimientos quirúrgicos o la aspiración de biopsia con aguja pueden usarse para recolectar muestras cancerosas de un sujeto. En algunas realizaciones, es importante enriquecer y/o purificar las muestras de tejido y/o células cancerosas a partir de las muestras de tejido y/o células no cancerosas. En otras realizaciones, las muestras de tejido y/o células cancerosas pueden microdisccionarse a continuación para reducir la cantidad de contaminación tisular normal antes de la extracción de ácido nucleico genómico o pre-ARN para usar en los procedimientos de la invención. Todavía en otra realización, las muestras de tejido y/o células cancerosas están enriquecidas para células cancerosas en al menos el 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 puede realizarse de acuerdo con los procedimientos bien conocidos en la

50

55

técnica, tales como microdissección con aguja, microdissección con láser, clasificación de células activadas por fluorescencia y clasificación inmunológica de células. En una realización, una máquina automática realiza el enriquecimiento celular hiperproliferativo para transformar de ese modo la muestra biológica en una forma purificada enriquecida por la presencia de células hiperproliferativas.

5

**[0140]** La recolección de muestras de ácido nucleico de células no cancerosas de un sujeto también puede realizarse con cirugía o aspiración. En procedimientos quirúrgicos en los que se extirpa tejido canceroso, los cirujanos suelen extraer muestras de tejido y/o células no cancerosas del mismo tipo de tejido del paciente con cáncer para compararlas. Las muestras de ácido nucleico se pueden aislar a partir de dicho tejido no canceroso del sujeto para su uso en los procedimientos de la invención. En ciertas realizaciones de los procedimientos de la invención, las muestras de ácido nucleico de tejidos no cancerosos no se derivan del mismo tipo de tejido que el tejido canceroso y/o células muestreadas, y/o no se derivan del paciente con cáncer. Las muestras de ácido nucleico de tejidos no cancerosos pueden derivarse de cualquier tejido y/o célula no cancerosa y/o libre de enfermedad. Dichas muestras no cancerosas se pueden recoger mediante procedimientos quirúrgicos o no quirúrgicos. En ciertas realizaciones, las muestras de ácido nucleico no cancerosas se derivan de tejidos libres de tumor. Por ejemplo, pueden recogerse muestras no cancerosas de ganglios linfáticos, linfocitos de sangre periférica y/o células sanguíneas mononucleares, o cualquier subpoblación de las mismas. En una realización preferida, el tejido no canceroso no es tejido precanceroso, por ejemplo, no muestra ningún indicio de un estado preneoplásico tal como hiperplasia, metaplasia o displasia.

20

**[0141]** En una realización, las muestras de ácido nucleico usadas para calcular la GCAS (por ejemplo, el número de SNP heterocigóticos en la pluralidad de SNP totales que exhiben heterocigosidad en 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 200 organismos diferentes de esa especie. De acuerdo con ciertos aspectos de la invención, el ácido nucleico "derivado de" ADN genómico, como se usa en los procedimientos de la invención, por ejemplo, en experimentos de hibridación para determinar la heterocigosidad de SNP, pueden ser fragmentos de ácido nucleico genómico generado por digestión con enzimas de restricción y/o ligación a otros ácidos nucleicos, y/o productos de amplificación de ácidos nucleicos genómicos, o ARN pre-mensajero (pre-ARNm), productos de amplificación de pre-ARNm, o fragmentos de ADN genómico crecidos en vectores de clonación generados, por ejemplo, mediante procedimientos de clonación con "pistola". En ciertas realizaciones, las muestras de ácido nucleico genómico se digieren con enzimas de restricción.

30

#### *Amplificación de ácidos nucleicos*

**[0142]** Aunque la muestra de ácido nucleico no necesita comprender ácido nucleico amplificado, en algunas realizaciones, los ácidos nucleicos aislados pueden procesarse de maneras que requieren y/o se aprovechan de la amplificación. Las muestras de ADN genómico de un sujeto opcionalmente pueden fragmentarse usando endonucleasas de restricción y/o amplificarse 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 practicar la PCR son bien conocidos por los expertos en la materia. Una ventaja de la PCR es que pueden usarse pequeñas cantidades de ADN. Por ejemplo, el ADN genómico de un sujeto puede ser de aproximadamente 150 ng, 175 ng, 200 ng, 225 ng, 250 ng, 275 ng o 300 ng de ADN.

40

**[0143]** En ciertas realizaciones de los procedimientos de la invención, el ácido nucleico de un sujeto se amplifica utilizando un solo par de cebadores. Por ejemplo, las muestras de ADN genómico se pueden digerir con endonucleasas de restricción para generar fragmentos de ADN genómico que luego se ligan a una secuencia de ADN adaptadora que reconoce el par de cebadores. En otras realizaciones de los procedimientos de la invención, el ácido nucleico de un sujeto se amplifica utilizando conjuntos de pares de cebadores específicos de los loci de interés (por ejemplo, RFLP, STR, SNP, etc.) ubicados en todo el genoma. Dichos conjuntos de pares de cebadores reconocen cada uno secuencias de ADN genómicas que flanquean lugares de interés particulares (por ejemplo, SNP, RFLP, STR, etc.). Se puede obtener una muestra de ADN adecuada para la hibridación, por ejemplo, mediante la reacción en cadena de la polimerasa (PCR) de ADN genómico, fragmentos de ADN genómico, fragmentos de ADN genómico-ligados a secuencias adaptadoras o secuencias clonadas. Se pueden usar programas informáticos que son bien conocidos en la técnica en el diseño de cebadores con la especificidad deseada y las 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 materia que los sistemas robóticos controlados son útiles y que pueden usarse para aislar y amplificar ácidos nucleicos.

55

**[0144]** En otras realizaciones, cuando el ADN genómico de un sujeto se fragmenta usando 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 a través del 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, NY, en pp. 9.47-9.51).

**[0145]** En otras realizaciones más, cuando el ADN de un sujeto se fragmenta usando 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. El ARN (ARN mensajero) que comprende el transcrito completo que incluye intrones se aísla entonces y se usa en los procedimientos de la invención para determinar la GCAS y el fenotipo de un cáncer. En ciertas realizaciones, no se requiere amplificación. En dichas realizaciones, el ADN genómico, o pre-ARN, de un sujeto se puede fragmentar usando endonucleasas de restricción u otros procedimientos. Los fragmentos resultantes pueden hibridarse con sondas SNP. Normalmente, se necesitan cantidades mayores de ADN para aislar en comparación con la cantidad de ADN o pre-ARN que se necesita cuando los fragmentos se amplifican. Por ejemplo, cuando el ácido nucleico de un sujeto no se amplifica, una muestra de ADN de un sujeto para su uso en hibridación puede ser de aproximadamente 400 ng, 500 ng, 600 ng, 700 ng, 800 ng, 900 ng o 1000 ng de ADN o mayor. Como alternativa, en otras realizaciones, se usan procedimientos que requieren cantidades muy pequeñas de ácidos nucleicos para su 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 usa para ensayos de sonda de inversión molecular (MIP). Estas técnicas son particularmente útiles para analizar muestras clínicas, tales como material fijado con formalina embebido en parafina o biopsias de aguja con núcleo pequeño, caracterizado porque está fácilmente disponible pero generalmente tiene una calidad de ADN reducida (por ejemplo, ADN fragmentado pequeño) y/o no proporciona grandes cantidades de ácidos nucleicos.

#### *Hibridación*

**[0146]** Las muestras de ácido nucleico derivadas de un sujeto usadas en los procedimientos de la invención se pueden hibridar con matrices que comprenden sondas (por ejemplo, sondas de oligonucleótidos) para identificar loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.). La hibridación también se puede usar para determinar si los loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) identifican aberraciones cromosómicas (por ejemplo, desequilibrio alélico, pérdida de heterocigosidad, cambio total del número de copias, ganancia del número de copias y pérdida del número de copias) en muestras de ácido nucleico de tejidos cancerosos y/o células del sujeto. En realizaciones preferidas, las sondas usadas en los procedimientos de la invención comprenden una matriz de sondas que se pueden fijar en un chip de ADN (por ejemplo, sondas de oligonucleótidos de SNP). En algunas realizaciones, la heterocigosidad de un locus de SNP se determina mediante un procedimiento que no comprende detectar un cambio en el tamaño de los fragmentos de ácido nucleico digeridos con enzimas de restricción. En otras realizaciones, los SNP se analizan para identificar un desequilibrio alélico. Las condiciones de hibridación y lavado usadas en los procedimientos de la invención se eligen para que las muestras de ácido nucleico a analizar por la invención se unan específicamente o se hibriden específicamente con las secuencias de oligonucleótidos complementarias de la matriz, preferiblemente en un sitio de matriz específico, en el que se encuentra su ADN complementario. En algunas realizaciones, el ADN complementario puede coincidir completamente o no coincidir en cierto grado como se usa, por ejemplo, en matrices de oligonucleótidos Affymetrix tales como las usadas para analizar SNP en ensayos MIP. Las sondas de ADN de ácido oligodesoxirribonucleico sintético monocatenario de una matriz pueden tener que desnaturalizarse antes del contacto con las muestras de ácido nucleico de un sujeto, por ejemplo, para eliminar horquillas o dímeros que se forman debido a secuencias autocomplementarias.

**[0147]** Las condiciones óptimas de hibridación dependerán de la longitud de las sondas y del 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*, 2ª edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, en pp. 9.47-9.51 y Ausubel et al., Eds., 1989, *Current Protocols in Molecular Biology*, vol. 1, Green Publishing Associates, Inc., John Wiley & Sons, Inc., Nueva York, pp. 2.10.1-2.10.16. Se proporcionan condiciones de hibridación útiles ejemplares en, por ejemplo, Tijessen, 1993, *hibridación con sondas de ácido nucleico*, Elsevier Science Publishers BV y Kricka, 1992, *Nonisotopic DNA Probe Techniques*, Academic Press, San Diego, California.

*Matrices oligonucleótidos de ácidos nucleicos de*

**[0148]** En algunas realizaciones de los procedimientos de la presente invención, las matrices de ADN pueden usarse para determinar si las muestras de ácido nucleico exhiben aberraciones cromosómicas (por ejemplo, 5 desequilibrio alélico, pérdida de heterocigosidad, cambio total del número de copias, ganancia del número de copias y pérdida del número de copias) midiendo el nivel de hibridación de la secuencia de ácido nucleico a sondas de oligonucleótidos que comprenden secuencias complementarias. La hibridación se puede usar para determinar la presencia o ausencia de heterocigosidad. Varios formatos de matrices de ADN que emplean "sondas" de oligonucleótidos (es decir, las moléculas de ácido nucleico que tienen secuencias definidas) son bien conocidas por 10 los expertos en la materia. Normalmente, un conjunto de sondas de ácido nucleico, cada una de las cuales tiene una secuencia definida, se inmoviliza sobre un soporte sólido de tal manera que cada sonda diferente se inmoviliza en una región predeterminada. En ciertas realizaciones, el conjunto de sondas forma una matriz de sitios de unión direccionable por posición (por ejemplo, hibridación) en un soporte. Cada uno de dichos sitios de unión comprende una pluralidad de moléculas de oligonucleótidos de una sonda unida a la región predeterminada en el soporte. Más 15 específicamente, cada sonda de la matriz se ubica preferiblemente en una posición conocida, predeterminada en el soporte sólido de modo que la identidad (es decir, la secuencia) de cada sonda se puede determinar a partir de su posición en la matriz (es decir, sobre el soporte o la superficie). Las micromatrices pueden prepararse de varias maneras, varias de las cuales se describen en el presente documento. Independientemente de lo que se produzca, las micromatrices comparten ciertas características, son reproducibles, permitiendo que se produzcan y se puedan 20 comparar fácilmente múltiples copias de un conjunto determinado.

**[0149]** En la técnica se conocen numerosas variaciones sobre matrices de ácidos nucleicos útiles en la invención. Estas incluyen matriz Affymetrix 500K GeneChip; Affymetrix OncoScan™ FFPE Express 2.0 Services (anteriormente MIP CN Services), y similares. 25

**[0150]** Preferiblemente, las micromatrices están hechas de materiales que son estables en condiciones de unión (por ejemplo, hibridación de ácidos nucleicos). Las micromatrices preferiblemente son 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 mayores y menores y pueden ser preferibles, por ejemplo, para evaluar simultáneamente un 30 gran número de sondas diferentes. Las sondas de oligonucleótidos se pueden sintetizar directamente en un soporte para formar la matriz. Las sondas se pueden unir a un soporte o superficie sólido, que puede estar hecho, por ejemplo, de vidrio, plástico (por ejemplo, polipropileno, nailon), 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 ácido nucleico marcadas de modo que los ácidos nucleicos que tienen 35 secuencias complementarias a una sonda inmovilizada se hibriden o se unan a la sonda. Después de separar, por ejemplo, mediante lavado, cualquier material no unido, se detectan y se miden las secuencias marcadas y unidas. La medición generalmente se realiza con asistencia informática. Usando 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 40 permitido determinar la heterocigosidad de un gran número de loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) en todo el genoma.

**[0151]** En ciertas realizaciones, se usan matrices de oligonucleótidos de alta densidad en los procedimientos de la invención. Estas matrices que contienen miles de oligonucleótidos complementarios a secuencias definidas, en 45 ubicaciones definidas en una superficie pueden sintetizarse *in situ* en la superficie mediante, por ejemplo, técnicas fotolitográficas (véase, por ejemplo, Fodor et al., 1991, Science 251: 767-773; Pease et al., 1994, Proc. Natl. Acad. Sci. USA 91: 5022-5026; Lockhart et al., 1996, Nature Biotechnology 14: 1675; Patentes de Estados Unidos n.º 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 se conocen en la 50 técnica (véase, por ejemplo, Blanchard, publicación de patente internacional WO 98/41531, publicada el 24 de septiembre de 1998; Blanchard et al., 1996, Biosensors and Bioelectronics 11: 687-690; Blanchard, 1998, en Synthetic DNA Arrays in Genetic Engineering, vol. 20, JK Setlow, Ed., Plenum Press, Nueva York en las páginas 111-123). Otro procedimiento para unir los ácidos nucleicos a una superficie es mediante impresión sobre placas de vidrio, como se describe generalmente por Schena et al. (1995, Science 270: 467-470). También se pueden usar 55 otros procedimientos para preparar micromatrices, por ejemplo, mediante enmascaramiento (Maskos y Southern, 1992, Nucl. Acids. Res. 20: 1679-1684). Cuando se usan estos procedimientos, los oligonucleótidos (por ejemplo, de 15 a 60-meros) 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 correspondientes a cada locus informativo de interés (por ejemplo, SNP, RFLP, STR, etc.).

**[0152]** Un ejemplo de medio para generar las sondas de oligonucleótidos de la matriz de ADN es por síntesis de polinucleótidos u oligonucleótidos sintéticos, por ejemplo, usando las 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 5 secuencias sintéticas normalmente tienen entre aproximadamente 15 y aproximadamente 600 bases de longitud, más normalmente 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 limitadas a, inosina. Como se ha indicado anteriormente, los análogos de ácido nucleico se pueden usar como sitios de unión para la hibridación. Un 10 ejemplo de un análogo de ácido nucleico adecuado es un ácido nucleico peptídico (véase, por ejemplo, Egholm et al., 1993, *Nature* 363: 566-568; Patente de los Estados Unidos n.º 5.539.083). En realizaciones alternativas, los sitios de hibridación (es decir, las sondas) están hechos a partir de clones de plásmidos o fagos de regiones de ADN genómico-correspondientes a SNP o el complemento de los mismos. 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 15 longitud. Es bien sabido en la técnica que aunque la hibridación es selectiva para secuencias complementarias, otras secuencias que no son perfectamente complementarias también pueden hibridarse en cierto grado a una sonda dada. Por lo tanto, pueden usarse múltiples sondas de oligonucleótidos con ligeras variaciones para optimizar la hibridación de las muestras. Para optimizar aún más la hibridación, pueden alterarse las condiciones de rigurosidad de hibridación, por ejemplo, la temperatura de hibridación y las concentraciones de sal, por 20 procedimientos que son bien conocidos en la técnica.

**[0153]** En algunas realizaciones, 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 "miméticos" de ADN o ADN 25 (por ejemplo, derivados y análogos) correspondientes a una porción de cada locus informativo de interés (por ejemplo, SNP, RFLP, STR, etc.) en el genoma de un sujeto. Las sondas de oligonucleótidos pueden modificarse 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 usar una pluralidad de oligonucleótidos diferentes que son complementarios a las secuencias de ácidos nucleicos de muestra. Por ejemplo, para un solo locus de 30 interés informativo (por ejemplo, SNP, RFLP, STR, etc.) pueden usarse 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 coincidencias perfectas, discordancias y la secuencia flanqueante alrededor del SNP. En ciertas realizaciones, las sondas se generan de manera que las 35 sondas para un locus informativo particular de interés comprendan secuencias superpuestas y/o secuencias superpuestas sucesivas que se extienden o están fijadas a través de una región genómica que contiene el sitio objetivo, donde todas las sondas contienen el sitio diana. A modo de ejemplo, las secuencias de sonda superpuestas se pueden fijar en etapas de intervalos de base predeterminados, por ejemplo, en las etapas de intervalos de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 bases. En ciertas realizaciones, los ensayos pueden realizarse usando 40 matrices adecuadas para su uso con protocolos de sonda de inversión molecular tales como los descritos por Wang et al. (2007) *Genoma Biol.* 8, R246. Para sondas de oligonucleótidos dirigidas a especies de ácido nucleico de secuencias estrechamente semejantes (es decir, homólogas), la "hibridación cruzada" entre sondas similares puede contaminar y confundir significativamente los resultados de las mediciones de hibridación. La hibridación cruzada es una preocupación particularmente importante en la detección de SNP ya que la secuencia a detectar (es decir, el 45 SNP particular) debe distinguirse de otras secuencias que difieren en un solo nucleótido. La hibridación cruzada se puede minimizar regulando las condiciones de rigurosidad de la hibridación y/o durante los lavados posteriores a la hibridación. Las condiciones altamente rigurosas permiten la detección de variantes alélicas de una secuencia de nucleótidos, por ejemplo, aproximadamente 1 desemparejamiento por 10-30 nucleótidos. No existe una única condición de hibridación o lavado que sea óptima para todas las diferentes secuencias de ácidos nucleicos. Para 50 matrices particulares de loci informativos de interés, estas condiciones pueden ser idénticas a las sugeridas por el fabricante o pueden ser ajustadas por un experto en la materia. En realizaciones preferidas, las sondas usadas en los procedimientos de la invención están inmovilizadas (es decir, fijadas) en un portaobjetos de vidrio llamado chip. Por ejemplo, una micromatriz de ADN puede comprender un chip sobre el que se han impreso robóticamente oligonucleótidos (secuencias de ADN monocatenario purificadas en solución) en una matriz (aproximadamente) 55 rectangular con cada punto en la matriz correspondiente a una muestra de ADN que codifica un oligonucleótido. En resumen, el proceso 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, luego se lava y seca la matriz y se explora con un microscopio láser para detectar hibridación. En ciertas realizaciones hay al menos 250, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 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 rango intermedio, de loci informativos de interés para los cuales las sondas aparecen en la matriz (con sondas de coincidencia/desajuste para un solo locus de interés o sondas fijadas en un solo locus de interés que cuenta como un locus de interés). El número máximo de loci informativos de interés sondeados por matriz está determinado por el tamaño del genoma y la diversidad genética de las especies del sujeto. Los chips de ADN son bien conocidos en la técnica y pueden comprarse en forma fabricada pre-5 con secuencias específicas para especies particulares. En algunas realizaciones, se usan la matriz Genome-Wide Human SNP Array 6.0™ y/o las matrices 50K XbaI (Affymetrix, Santa Clara, CA) en los procedimientos de la invención. En otras realizaciones, los SNP y/o el número de copias de ADN pueden detectarse y cuantificarse usando procedimientos de secuenciación, tales como "procedimientos de secuenciación de próxima generación" como se describe adicionalmente más arriba.

#### *Detección de señal*

15

**[0154]** En algunas realizaciones, las muestras de ácidos nucleicos 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 muestra 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 muestra de un sujeto (es decir, cancerosas y no cancerosas) se hibridan con dicha matriz, luego, después de la detección de la señal se lava el chip para eliminar la primera muestra marcada y se reutiliza para hibridar la muestra restante. En otras realizaciones, la matriz no se reutiliza más de una vez. En ciertas realizaciones, las muestras de ácido nucleico derivadas de cada uno de los dos tipos de muestra de un sujeto (es decir, cancerosas y no cancerosas) se marcan de manera diferente para que puedan distinguirse. Cuando las dos muestras se mezclan e hibridan con la misma matriz, se determina la intensidad relativa de la señal de cada muestra para cada sitio en la matriz, y se detecta cualquier diferencia relativa en la abundancia de un alelo de loci informativos de interés. Las señales pueden grabarse y, en algunas realizaciones, analizarse por ordenador. En una realización, la imagen escaneada se limpia utilizando un programa de gráficos (por ejemplo, Hijaak Graphics Suite) y luego se analiza usando un programa de cuadrículado 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 hacer una corrección determinada experimentalmente para "interferencia" (o superposición) entre los canales para los dos fluoros. Para cualquier sitio de hibridación particular en la matriz, puede calcularse una relación de la emisión de los dos fluoróforos, que puede ayudar a eliminar las señales de hibridación cruzada para determinar con mayor precisión si un locus de SNP particular es heterocigótico u homocigótico.

35

#### Marcaje

**[0155]** En algunas realizaciones, las muestras de ácidos nucleicos, fragmentos de las mismas, o fragmentos de las mismas ligadas a regiones adaptadoras usadas 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, entre otros, biotina, iminobiotina, antígenos, cofactores, dinitrofenol, ácido lipoico, compuestos olefínicos, polipéptidos detectables, moléculas ricas en electrones, enzimas capaces de generar una señal detectable por acción sobre un sustrato, e isótopos radiactivos.

45

**[0156]** Los isótopos radiactivos incluyen los que se pueden usar junto con los procedimientos de la invención, pero no están limitados a, <sup>32</sup>P y <sup>14</sup>C. Las moléculas fluorescentes adecuadas para la presente invención incluyen, sin limitación, fluoresceína y sus derivados, rodamina y sus derivados, texas red, 5'-carboxi-fluoresceína ("FAM"), 2',7'-dimetoxi-4',5'-dicloro-6-carboxi-fluoresceína ("JOE"), N,N,N',N'-tetrametil-6-carboxi-rodamina ("TAMRA"), 6-carboxi-X-rodamina ("ROX"), HEX, TET, IRD40 e IRD41.

50

**[0157]** Las moléculas fluorescentes que son adecuadas para su uso de acuerdo con la invención incluyen además: colorantes de cianina, que incluyen, pero no se limitan a, Cy2, Cy3, Cy3,5, Cy5, Cy5,5, Cy7 y FLUORX; colorantes BODIPY que incluyen pero no se limitan a BODIPY-FL, BODIPY-TR, BODIPY-TMR, BODIPY-630/650 y BODIPY-650/670; y colorantes ALEXA, que incluyen, entre otros, ALEXA-488, ALEXA-532, ALEXA-546, ALEXA-568 y ALEXA-594; así como otros colorantes fluorescentes que serán conocidos por los expertos en la materia. Las moléculas indicadoras ricas en electrones adecuadas para la presente invención incluyen, sin limitación, ferritina, hemocianina y oro coloidal.

55

**[0158]** También se pueden usar esquemas de detección y marcaje de fluorescencia bicolor (Shena et al., 1995, Science 270: 467-470). El uso de dos o más marcadores puede ser útil para detectar variaciones debido a pequeñas diferencias en las condiciones experimentales (por ejemplo, condiciones de hibridación). En algunas realizaciones de la invención, se pueden usar al menos 5, 10, 20 o 100 colorantes de diferentes colores para el marcaje. Dicho marcaje también permitiría el análisis simultáneo de múltiples muestras que abarca la invención.

**[0159]** Las muestras de ácido nucleico marcadas, fragmentos de las mismas o fragmentos de las mismas ligadas a regiones adaptadoras que pueden usarse en los procedimientos de la invención se ponen en contacto con una pluralidad de sondas de oligonucleótidos en condiciones que permiten que los ácidos nucleicos de muestra que tienen secuencias complementarias a las sondas se hibriden. Dependiendo del tipo de marcador utilizado, las señales de hibridación pueden detectarse usando procedimientos bien conocidos por los expertos en la materia, que incluyen, pero no se limitan a, película de rayos X, cámara de imágenes de fósforo o cámara CCD. Cuando se usan sondas marcadas fluorescentemente, las emisiones de fluorescencia en cada sitio de una matriz de transcritos, preferiblemente, pueden detectarse mediante exploración con microscopio láser confocal. En una realización, se lleva a cabo una exploración separada, usando la línea de excitación apropiada para cada uno de los dos fluoróforos usados. Como alternativa, se puede usar un láser que permite la iluminación simultánea de especímenes en longitudes de onda específicas para los dos fluoróforos y las emisiones de los dos fluoróforos se pueden analizar simultáneamente (ver Shalon et al. (1996) Genome Res. 6, 639-645). En una realización preferida, las matrices se exploran 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 logra con un láser de gas mixto de varias líneas, y la luz emitida se divide por la longitud de onda y se detecta con dos tubos fotomultiplicadores. Dichos dispositivos de exploración con láser de fluorescencia se describen, por ejemplo, en Schena et al. (1996) Genome Res. 6, 639-645. Como alternativa, se puede usar un haz de fibra óptica tal como el descrito por Ferguson et al. (1996) Nat. Biotech. 14, 1681-1684. Las señales resultantes pueden analizarse entonces para determinar la presencia o ausencia de heterocigosidad u homocigosidad para loci informativos de interés (por ejemplo, SNP, RFLP, STR, etc.) usando un programa informático.

#### *Algoritmos para analizar loci informativos de interés*

**[0160]** 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 para llamar al desequilibrio alélico (AI), pérdida de heterocigosidad (LOH), aberraciones del número de copias (CNA), ganancia del número de copias (CNG) y disminución del número de copias (CND). Por ejemplo, el AI puede determinarse usando la proporción de copias principales (MCP) en la que se llama el AI para un SNP dado, 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 la llamada, pueden aplicarse aún más procedimientos de enumeración. Por ejemplo, la GCAS se puede determinar, por ejemplo, por: 1) el recuento del número total de SNP afectados por AI o ganancia de copias o LOH, 2) el recuento de la cantidad de regiones afectadas por AI (por ejemplo, NAI como se describe adicionalmente en los Ejemplos; una región única se define como una cadena de SNP vecinos que muestran un AI limitado en al menos un lado por SNP 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 pérdida cromosómica completa, o 4) el recuento del número de regiones cromosómicas con CNA, CNG, CND, etc. Se proporcionan ilustraciones representativas adicionales de dichos algoritmos bien conocidos en la sección de Ejemplos a continuación.

#### *Sistemas y procedimientos de implementación informática*

**[0161]** En ciertas realizaciones, los procedimientos de la invención implementan un programa informático para calcular una puntuación de desequilibrio alélico telomérico. Por ejemplo, se puede usar un programa de ordenador para realizar los algoritmos descritos en este documento. Un sistema informático también puede almacenar y manipular datos generados por los procedimientos de la presente invención que comprenden una pluralidad de cambios/perfiles de señal de hibridación durante el acercamiento al equilibrio en diferentes mediciones de hibridación y que pueden ser utilizados por un sistema informático para implementar los procedimientos de esta invención. En ciertas realizaciones, un sistema informático recibe datos de hibridación de la sonda; (ii) almacena datos de hibridación de la sonda; y (iii) compara los datos de hibridación de la sonda para determinar el estado de los loci informativos de interés en dicha muestra de ácido nucleico de tejido canceroso o precanceroso. La GCAS se calcula. En algunas realizaciones, un sistema informático (i) compara la GCAS determinada con un valor umbral; y

(ii) produce una indicación de si dicha GCAS está por encima o por debajo de un valor umbral, o un fenotipo basado en dicha indicación. Dichos sistemas informáticos también se describen en el presente documento.

**[0162]** Se pueden usar numerosos tipos de sistemas informáticos para implementar los procedimientos analíticos de esta invención de acuerdo con el conocimiento que posea un experto en bioinformática y/o en materia de computación.

**[0163]** Varios componentes de software se pueden cargar en la memoria durante el funcionamiento de dicho sistema informático. Los componentes de software pueden comprender tanto componentes de software que son convencionales en la técnica como 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 se pueden programar o modelar en paquetes de software matemático 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 programar proceduralmente ecuaciones y algoritmos individuales. Dichos paquetes incluyen, por ejemplo, Matlab de Mathworks (Natick, MA), Mathematica de Wolfram Research (Champaign, ILL) o S-Plus de MathSoft (Seattle, WA). En ciertas realizaciones, el ordenador comprende una base de datos para el almacenamiento de perfiles de señal de hibridación. Se puede acceder a estos perfiles almacenados y usarlos para calcular la GCAS. Por ejemplo, del perfil de la señal de hibridación de una muestra derivada del tejido no canceroso de un sujeto y/o perfiles generados a partir de distribuciones basadas en la población de loci informativos de interés en poblaciones relevantes de la misma especie que estuvieran almacenados, entonces se podría comparar con el perfil de señal de hibridación de una muestra derivada del tejido canceroso del sujeto.

**[0164]** Además de las estructuras de programas y sistemas informáticos ejemplares descritos en la presente memoria, otras estructuras alternativas de programas y sistemas informáticos serán fácilmente evidentes para el experto en la materia.

**[0165]** Una vez que un técnico de laboratorio o profesional de laboratorio o un grupo de técnicos de laboratorio o profesionales de laboratorio determina si una muestra tiene una aberración cromosómica en una pluralidad de loci de prueba como se ha descrito anteriormente (por ejemplo, la etapa (1) en muchos de los procedimientos anteriores), el mismo o un técnico de laboratorio o un profesional de laboratorio (o grupo) diferente puede analizar una pluralidad de loci de prueba para determinar si tienen una aberración cromosómica (por ejemplo, la etapa (2) en muchos de los procedimientos anteriores). Luego, el mismo o un técnico de laboratorio o profesional de laboratorio (o grupo) diferente puede combinar los datos de aberración cromosómica de los loci de prueba para derivar una puntuación de aberración cromosómica (por ejemplo, la etapa (3) en muchos de los procedimientos anteriores). Opcionalmente, el mismo o un técnico de laboratorio o profesional de laboratorio (o grupo) diferente puede correlacionar una puntuación de aberración cromosómica alta con una mayor probabilidad de respuesta a un tratamiento en particular (por ejemplo, los mencionados anteriormente). Por ejemplo, uno o más técnicos de laboratorio o profesionales de laboratorio pueden identificar a un paciente con células cancerosas que se detectó que tiene una puntuación de aberración cromosómica alta asociando esa puntuación de aberración cromosómica alta o el resultado (o resultados o un resumen de resultados) del análisis de diagnóstico con el nombre del paciente correspondiente, registro médico, identificador simbólico/numérico, o una combinación de los mismos. Dicha identificación puede basarse únicamente en la detección de la presencia de una puntuación de aberración cromosómica alta o puede basarse, al menos en parte, en la detección de la presencia de una puntuación de aberración cromosómica alta. Por ejemplo, un técnico de laboratorio o un profesional de laboratorio puede identificar a un paciente que tiene células cancerosas que tienen una puntuación de aberración cromosómica alta como células cancerosas con una mayor probabilidad de respuesta a una terapia particular basada en una combinación de una puntuación de aberración cromosómica alta y los resultados de otras pruebas genéticas y bioquímicas realizadas en el laboratorio de pruebas.

**[0166]** La Figura 23 muestra un proceso ejemplar mediante el cual un sistema informático puede determinar una puntuación de aberración cromosómica. El proceso comienza en el recuadro 300, donde los datos con respecto al genotipo (por ejemplo, número de copias relativo o absoluto, homocigótico, heterocigótico) de una pluralidad de loci a lo largo de un cromosoma son recogidos por el sistema informático. Como se describe en este documento, se puede usar cualquier ensayo apropiado tal como un ensayo basado en matriz de SNP o ensayo basado en secuenciación para evaluar loci a lo largo de un cromosoma para el genotipo. En algunos casos, se puede usar un sistema que incluye un detector de señal y un ordenador para recopilar datos (por ejemplo, señales fluorescentes o resultados de secuenciación) con respecto al genotipo de la pluralidad de loci. En el recuadro 310, los datos con

respecto al genotipo de una pluralidad de loci, así como la ubicación o relación espacial de cada locus, son evaluados por el sistema informático para determinar, por ejemplo, la longitud de cualquier región de aberración cromosómica (por ejemplo, desequilibrio alélico) presente a lo largo de un cromosoma o el número de regiones de aberración telomérica (por ejemplo, desequilibrio alélico). En el recuadro 320, los datos con respecto al número de regiones de aberración cromosómica detectadas y opcionalmente la longitud o la ubicación de cada región de aberración cromosómica detectada se evalúan mediante el sistema informático para determinar el número de regiones de aberración cromosómica que son regiones teloméricas. En el recuadro 330, el sistema informático formatea una salida que proporciona una indicación de la presencia o ausencia de una puntuación de aberración cromosómica alta. Una vez formateado, el sistema informático puede presentar el resultado a un usuario (por ejemplo, un técnico de laboratorio, un médico o un profesional médico). Como se describe en este documento, la presencia o ausencia de una puntuación de aberración cromosómica alta puede usarse para proporcionar una indicación sobre posibles regímenes de tratamiento del cáncer.

**[0167]** La Figura 24 es un diagrama de un ejemplo de un dispositivo informático 1400 y un dispositivo informático móvil 1450, que se puede usar con las técnicas descritas en este documento. El dispositivo 1400 de computación está destinado a representar diversas formas de ordenadores digitales, tales como ordenadores portátiles, ordenadores de escritorio, estaciones de trabajo, asistentes digitales personales, servidores, servidores blade, mainframes y otros ordenadores apropiados. El dispositivo de cálculo 1450 pretende representar diversas formas de dispositivos móviles, tales como asistentes digitales personales, teléfonos móviles, teléfonos inteligentes y otros dispositivos informáticos similares. Los componentes mostrados en este documento, sus conexiones y relaciones, y sus funciones, están destinados a ser solo ejemplares, y no están destinados a limitar las implementaciones de las invenciones descritas y/o reivindicadas en este documento.

**[0168]** El dispositivo informático 1400 incluye un procesador 1402, memoria 1404, un dispositivo de almacenamiento 1406, una interfaz de alta velocidad 1408 que conecta con la memoria 1404 y puertos de expansión de alta velocidad 1410, y una interfaz de baja velocidad 1415 que conecta con el bus de baja velocidad 1414 y el dispositivo de almacenamiento 1406. Cada uno de los componentes 1402, 1404, 1406, 1408, 1410 y 1415 están interconectados usando varios buses, y pueden montarse en una placa base común o de otras maneras, según corresponda. El procesador 1402 puede procesar instrucciones para su ejecución dentro del dispositivo informático 1400, incluidas instrucciones almacenadas en la memoria 1404 o en el dispositivo de almacenamiento 1406 para mostrar información gráfica para una GUI en un dispositivo de entrada/salida externo, como la pantalla 1416 acoplada a la interfaz de alta velocidad 1408. En otras implementaciones, pueden usarse múltiples procesadores y/o múltiples buses, según corresponda, junto con múltiples memorias y tipos de memoria. Además, se pueden conectar múltiples dispositivos informáticos 1400, proporcionando cada dispositivo porciones de las operaciones necesarias (por ejemplo, como un banco servidor, un grupo de servidores blade, o un sistema multiprocesador).

**[0169]** La memoria 1404 almacena información dentro del dispositivo informático 1400. En una implementación, la memoria 1404 es una unidad o unidades de memoria volátil. En otra implementación, la memoria 1404 es una unidad o unidades de memoria no volátil. La memoria 1404 también puede ser otra forma de medio legible por ordenador, tal como un disco magnético u óptico.

**[0170]** El dispositivo de almacenamiento 1406 es capaz de proporcionar almacenamiento masivo para el dispositivo informático 1400. En una implementación, el dispositivo de almacenamiento 1406 puede ser o contener un medio legible por ordenador, como un dispositivo de disquete, un dispositivo de disco duro, un dispositivo de disco óptico, o un dispositivo de cinta, una memoria flash u otro dispositivo de memoria de estado sólido similar, o una matriz de dispositivos, incluidos los dispositivos en una red de área de almacenamiento u otras configuraciones. Un producto de programa informático puede incorporarse de forma tangible en un soporte de información. El producto de programa informático también puede contener instrucciones que, cuando se ejecutan, realizan uno o más procedimientos, tales como los descritos en este documento. El soporte de información es un medio legible por ordenador o máquina, tal como la memoria 1404, el dispositivo de almacenamiento 1406, la memoria sobre el procesador 1402, o una señal propagada.

**[0171]** El controlador de alta velocidad 1408 gestiona operaciones de ancho de banda intensivo para el dispositivo informático 1400, mientras que el controlador de baja velocidad 1415 gestiona operaciones de un ancho de banda intensivo menor. Dicha asignación de funciones es solo ejemplar. En una implementación, el controlador de alta velocidad 1408 está acoplado a la memoria 1404, pantalla 1416 (por ejemplo, a través de un procesador o acelerador de gráficos) y a puertos de expansión de alta velocidad 1410, que pueden aceptar varias tarjetas de expansión (no mostradas). En la implementación, el controlador de baja velocidad 1415 está acoplado al dispositivo de almacenamiento 1406 y al puerto de expansión de baja velocidad 1414. El puerto de expansión de baja

velocidad, que puede incluir varios puertos de comunicación (por ejemplo, USB, Bluetooth, Ethernet o Ethernet inalámbrico) puede estar acoplado a uno o más dispositivos de entrada/salida, como un teclado, un dispositivo señalador, un escáner, un lector óptico, un detector de señal fluorescente, o un dispositivo de red tal como un conmutador o enrutador, por ejemplo, a través de un adaptador de red.

5

**[0172]** El dispositivo informático 1400 puede implementarse en una variedad de formas diferentes, como se muestra en la figura. Por ejemplo, puede implementarse como un servidor convencional 1420, o múltiples veces en un grupo de dichos servidores. También se puede implementar como parte de un sistema servidor en rack 1424. Además, se puede implementar en un ordenador personal tal como un ordenador portátil 1422. Como alternativa, los componentes del dispositivo informático 1400 se pueden combinar con otros componentes en un dispositivo móvil (no mostrado), tal como el dispositivo 1450. Cada uno de dichos dispositivos puede contener uno o más dispositivos informáticos 1400, 1450, y un sistema completo puede estar compuesto por múltiples dispositivos informáticos 1400, 1450 que se comunican entre sí.

10

**[0173]** El dispositivo de cálculo 1450 incluye un procesador 1452, memoria 1464, un dispositivo de entrada/salida tal como una pantalla 1454, una interfaz de comunicación 1466 y un transceptor 1468, entre otros componentes (por ejemplo, un escáner, un lector óptico, un detector de señal fluorescente). El dispositivo 1450 también puede estar provisto de un dispositivo de almacenamiento, tal como un microdispositivo u otro dispositivo, para proporcionar almacenamiento adicional. Cada uno de los componentes 1450, 1452, 1464, 1454, 1466 y 1468 están interconectados usando varios buses, y varios de los componentes pueden montarse en una placa base común o de otras maneras, según corresponda.

20

**[0174]** El procesador 1452 puede ejecutar instrucciones dentro del dispositivo informático 1450, incluidas las instrucciones almacenadas en la memoria 1464. El procesador puede implementarse como un chipset de chips que incluyen procesadores analógicos y digitales separados y múltiples. El procesador puede proporcionar, por ejemplo, la coordinación de los otros componentes del dispositivo 1450, tal como el control de las interfaces de usuario, las aplicaciones ejecutadas por el dispositivo 1450, y la comunicación inalámbrica por el dispositivo 1450.

25

**[0175]** El procesador 1452 puede comunicarse con un usuario a través de la interfaz de control 1458 y la interfaz de visualización 1456 acoplada a una pantalla 1454. La pantalla 1454 puede ser, por ejemplo, una pantalla de LCD TFT (pantalla de cristal líquido de transistores de película delgada) o una pantalla de OLED (diodo orgánico de emisión de luz) u otra tecnología de visualización apropiada. La interfaz de visualización 1456 puede comprender una circuitería apropiada para hacer que la pantalla 1454 presente información gráfica y otra información a un usuario. La interfaz de control 1458 puede recibir comandos de un usuario y convertirlos para enviarlos al procesador 1452. Además, se puede proporcionar una interfaz externa 1462 en comunicación con el procesador 1452, para permitir la comunicación de área cercana del dispositivo 1450 con otros dispositivos. La interfaz externa 1462 puede proporcionar, por ejemplo, comunicaciones por cable en algunas implementaciones, o comunicación inalámbrica en otras implementaciones, y también se pueden usar múltiples interfaces.

30

35

**[0176]** La memoria 1464 almacena información dentro del dispositivo informático 1450. La memoria 1464 puede implementarse como uno o más medios legibles por ordenador, unidad o unidades de memoria volátil, o unidad o unidades de memoria no volátil. La memoria de expansión 1474 también se puede proporcionar y conectar al dispositivo 1450 a través de la interfaz de expansión 1472, que puede incluir, por ejemplo, una interfaz de tarjeta SIMM (módulo de memoria en línea individual). Dicha memoria de expansión 1474 puede proporcionar espacio de almacenamiento adicional para el dispositivo 1450, o también puede almacenar aplicaciones u otra información para el dispositivo 1450. Por ejemplo, la memoria de expansión 1474 puede incluir instrucciones para llevar a cabo o complementar los procesos descritos en este documento, y también puede incluir salvar información. Por lo tanto, por ejemplo, la memoria de expansión 1474 se puede proporcionar como módulo de seguridad para el dispositivo 1450, y se puede programar con instrucciones que permitan el uso seguro del dispositivo 1450. Además, se pueden proporcionar aplicaciones seguras a través de las tarjetas SIMM, junto con información adicional, como colocar información de identificación en la tarjeta SIMM de una manera no pirateable.

40

45

50

**[0177]** La memoria puede incluir, por ejemplo, memoria flash y/o memoria NVRAM, como se describe a continuación. En una implementación, un producto de programa informático se materializa de forma tangible en un soporte de información. El producto del programa informático contiene instrucciones que, cuando se ejecutan, realizan uno o más procedimientos, como los descritos en este documento. El soporte de información es un medio legible por ordenador o máquina, tal como la memoria 1464, la memoria de expansión 1474, la memoria sobre el procesador 1452, o una señal propagada que puede recibirse, por ejemplo, sobre el transceptor 1468 o la interfaz externa 1462.

55

**[0178]** El dispositivo 1450 puede comunicarse de forma inalámbrica a través de la interfaz de comunicación 1466, que puede incluir un circuito de procesamiento de una señal digital cuando sea necesario. La interfaz de comunicación 1466 puede proporcionar comunicaciones bajo diversos modos o protocolos, tales como llamadas de voz GSM, SMS, EMS o mensajería MMS, CDMA, TDMA, PDC, WCDMA, CDMA2000 o GPRS, entre otros. Dicha comunicación puede ocurrir, por ejemplo, a través del transceptor 1468 de radiofrecuencia. Además, puede producirse una comunicación de corto alcance, tal como el uso de un transceptor Bluetooth, WiFi u otro transceptor de este tipo (no mostrado). Además, el módulo receptor 1470 de GPS (sistema de posicionamiento global) puede proporcionar datos inalámbricos adicionales relacionados con la navegación y la ubicación al dispositivo 1450, que pueden usarse según sea apropiado por las aplicaciones que se ejecutan en el dispositivo 1450.

**[0179]** El dispositivo 1450 también puede comunicarse de manera audible usando el códec de audio 1460, que puede recibir información hablada de un usuario y convertirla en información digital utilizable. El códec de audio 1460 también puede generar sonido audible para un usuario, tal como a través de un altavoz, por ejemplo, en un auricular del dispositivo 1450. Dicho sonido puede incluir sonido de llamadas telefónicas de voz, puede incluir sonido grabado (por ejemplo, mensajes de voz, archivos de música, etc.) y también puede incluir sonido generado por aplicaciones que operan en el dispositivo 1450.

**[0180]** El dispositivo informático 1450 puede implementarse en una variedad de formas diferentes, como se muestra en la figura. Por ejemplo, puede implementarse como un teléfono móvil 1480. También se puede implementar como parte de un teléfono inteligente 1482, asistente digital personal u otro dispositivo móvil similar.

**[0181]** Diversas implementaciones de los sistemas y técnicas descritos en este documento pueden realizarse en circuitos electrónicos digitales, circuitos integrados, ASIC especialmente diseñados (circuitos integrados específicos de la aplicación), hardware informático, firmware, software y/o combinaciones de los mismos. Estas diversas implementaciones pueden incluir la implementación en uno o más programas informáticos que son ejecutables y/o interpretables en un sistema programable que incluye al menos un procesador programable, que puede ser de propósito especial o de propósito general, para recibir datos e instrucciones de, y para transmitir datos e instrucciones para, un sistema de almacenamiento, al menos un dispositivo de entrada, y al menos un dispositivo de salida.

**[0182]** Estos programas informáticos (también conocidos como programas, software, aplicaciones de software o código) incluyen instrucciones de la máquina para un procesador programable, y pueden implementarse en un lenguaje de programación orientado a objetos y/o de procedimiento de alto nivel, y/o en lenguaje ensamblador/código máquina. Como se usa en el presente documento, los términos "medio legible por máquina" y "medio legible por ordenador" se refieren a cualquier producto, aparato y/o dispositivo de programa informático (por ejemplo, discos magnéticos, discos ópticos, memoria y dispositivos lógicos programables (PLD)) utilizado para proporcionar instrucciones y/o datos de la máquina a un procesador programable, que incluye un medio legible por máquina que recibe instrucciones de la máquina como una señal legible por la máquina. El término "señal legible por máquina" se refiere a cualquier señal utilizada para proporcionar instrucciones y/o datos de la máquina a un procesador programable.

**[0183]** Para proporcionar interacción con un usuario, los sistemas y técnicas descritos en este documento pueden implementarse en un ordenador que tiene un dispositivo de visualización (por ejemplo, un monitor CRT (tubo de rayos catódicos) o LCD (pantalla de cristal líquido) para mostrar información al usuario y un teclado y un dispositivo señalador (por ejemplo, un ratón o una bola de seguimiento) mediante los cuales el usuario puede proporcionar información al ordenador. También se pueden usar otros tipos de dispositivos para proporcionar interacción con un usuario; por ejemplo, la retroalimentación proporcionada al usuario puede ser cualquier forma de retroalimentación sensorial (por ejemplo, retroalimentación visual, retroalimentación auditiva o retroalimentación táctil); y la entrada del usuario se puede recibir en cualquier forma, incluida la entrada acústica, de voz o táctil.

**[0184]** Los sistemas y técnicas descritos en este documento pueden implementarse en un sistema informático que incluye un componente de back-end (por ejemplo, como servidor de datos) o que incluye un componente de middleware (por ejemplo, un servidor de aplicaciones) o que incluye un componente de front-end (por ejemplo, un ordenador cliente que tiene una interfaz gráfica de usuario o un navegador web a través del cual un usuario puede interactuar con una implementación de los sistemas y técnicas descritas en este documento), o cualquier combinación de dichos componentes de back-end, middleware o front-end. Los componentes del sistema se pueden interconectar mediante cualquier forma o medio de comunicación de datos digitales (por ejemplo, una red de comunicación). Los ejemplos de redes de comunicación incluyen una red de área local ("LAN"), una red de área

amplia ("WAN") e Internet.

**[0185]** El sistema informático puede incluir clientes y servidores. Un cliente y un servidor generalmente son remotos entre sí y suelen interactuar a través de una red de comunicación. La relación del cliente y el servidor surge en virtud de los programas de ordenador que se ejecutan en los ordenadores respectivos y que tienen una relación cliente-servidor entre sí.

**[0186]** Los resultados de cualquier análisis de acuerdo con la invención a menudo serán comunicados a médicos, asesores genéticos y/o pacientes (u otras partes interesadas, tales como investigadores) en una forma transmisible que puede ser comunicada o transmitida a cualquiera de las partes mencionadas. Dicha forma puede variar y puede ser tangible o intangible. Los resultados pueden incorporarse en declaraciones descriptivas, diagramas, fotografías, gráficos, imágenes o cualquier otra forma visual. Por ejemplo, se pueden usar gráficos o diagramas que muestran información de genotipo o LOH (o estado de HRD) para explicar los resultados. Las declaraciones y las formas visuales se pueden grabar en un medio tangible, como documentos, medios legibles por ordenador, como disquetes, discos compactos, memoria flash, etc., o en un medio intangible, por ejemplo, un medio electrónico en forma de correo electrónico o sitio web en internet o intranet. Además, los resultados también pueden grabarse en una forma de sonido y transmitirse a través de cualquier medio adecuado, por ejemplo, líneas de cable analógico o digital, cables de fibra óptica, etc., a través de teléfono, fax, teléfono móvil inalámbrico, teléfono de Internet y similares.

**[0187]** Por lo tanto, la información y los datos sobre el resultado de una prueba pueden producirse en cualquier parte del mundo y transmitirse a una ubicación diferente. Como ejemplo ilustrativo, cuando una prueba se realiza fuera de los Estados Unidos, la información y los datos sobre un resultado de prueba pueden generarse, enviarse en una forma transmisible como se describe anteriormente, y luego importarse a los Estados Unidos. Por consiguiente, la presente descripción también abarca un procedimiento para producir una forma de información transmisible en una firma de LOH para al menos una muestra de paciente. El procedimiento comprende las etapas de (1) determinar una firma de LOH según los procedimientos de la presente descripción; y (2) incorporar el resultado de la etapa de determinación en una forma transmisible. La forma transmisible es un producto de dicho procedimiento.

**[0188]** En algunos casos, un sistema informático proporcionado en este documento puede configurarse para incluir uno o más analizadores de muestras. Un analizador de muestra puede configurarse para producir una pluralidad de señales sobre ADN genómico de al menos un par de cromosomas humanos de una célula cancerosa. Por ejemplo, un analizador de muestras puede producir señales que pueden interpretarse de manera que identifique la naturaleza homocigótica o heterocigótica de los loci a lo largo de un cromosoma. En algunos casos, un analizador de muestras puede configurarse para llevar a cabo una o más etapas de un ensayo basado en matriz de SNP o un ensayo basado en secuenciación y puede configurarse para producir y/o capturar señales de dichos ensayos. En algunos casos, un sistema informático proporcionado en este documento se puede configurar para incluir un dispositivo informático. En tales casos, el dispositivo informático se puede configurar para recibir señales de un analizador de muestras. El dispositivo informático puede incluir instrucciones ejecutables por ordenador o un programa informático (por ejemplo, software) que contiene instrucciones ejecutables por ordenador para llevar a cabo uno o más de los procedimientos o etapas descritos en el presente documento. En algunos casos, dichas instrucciones ejecutables por ordenador pueden instruir a un dispositivo informático para que analice señales de un analizador de muestras, de otro dispositivo informático, de un ensayo basado en matrices de SNP, o de un ensayo basado en secuenciación. El análisis de dichas señales puede llevarse a cabo para determinar genotipos, aberraciones cromosómicas en ciertos loci, regiones de aberración cromosómica, el número de regiones de aberración cromosómica, para determinar la ubicación de las regiones de aberración cromosómica (por ejemplo, teloméricas), para determinar el número de regiones de aberración cromosómica que tienen una ubicación particular (por ejemplo, telomérica), determinar si una muestra es positiva o no para una puntuación de aberración cromosómica alta, para determinar la probabilidad de que un paciente con cáncer responda a un régimen de tratamiento del cáncer en particular (por ejemplo, un régimen como el descrito anteriormente) o para determinar una combinación de estos elementos.

**[0189]** En algunos casos, un sistema informático proporcionado en este documento puede incluir instrucciones ejecutables por ordenador o un programa informático (por ejemplo, software) que contiene instrucciones ejecutables por ordenador para formatear un resultado que proporciona una indicación sobre el número de regiones de aberración cromosómica, la ubicación de regiones de aberración cromosómicas (por ejemplo, telomérica), el número de regiones de LOH que tienen una ubicación particular (por ejemplo, telomérica), si una muestra es positiva para una puntuación de aberración cromosómica alta o probabilidad de que un paciente con

cáncer responda a un régimen de tratamiento de cáncer particular (por ejemplo, un régimen como se ha descrito anteriormente), o una combinación de estos elementos. En algunos casos, un sistema informático proporcionado en este documento puede incluir instrucciones ejecutables por ordenador o un programa informático (por ejemplo, software) que contiene instrucciones ejecutables por ordenador para determinar un régimen de tratamiento de 5 cáncer deseado para un paciente particular basándose, al menos en parte, en la presencia o ausencia de una puntuación de aberración cromosómica alta.

**[0190]** En algunos casos, un sistema informático proporcionado en este documento puede incluir un dispositivo de procesamiento previo configurado para procesar una muestra (por ejemplo, células cancerosas) de 10 manera que se pueda realizar un ensayo basado en un conjunto de SNP o un ensayo basado en secuenciación. Los ejemplos de dispositivos de preprocesamiento incluyen, sin limitación, dispositivos configurados para enriquecer las poblaciones de células cancerosas en oposición a células no cancerosas, dispositivos configurados para lisar células y/o extraer ácido nucleico genómico, y dispositivos configurados para enriquecer una muestra para fragmentos de ADN genómico particular.

**[0191]** En general, un aspecto de esta descripción presenta un procedimiento para evaluar LOH en una 15 célula cancerosa o ADN genómico de la misma. En algunas realizaciones, el procedimiento comprende, o consiste esencialmente en, (a) detectar, en una célula cancerosa o ADN genómico derivado de la misma, regiones de LOH en al menos un par de cromosomas humanos de la célula cancerosa (por ejemplo, cualquier par de cromosomas 20 humanos distinto de un par de cromosomas sexuales X/Y humanos); y (b) determinar el número y tamaño (por ejemplo, longitud) de dichas regiones de LOH. En algunas realizaciones, las regiones de LOH se analizan en varios pares de cromosomas que son representativos del genoma completo (por ejemplo, se analizan suficientes cromosomas de modo que se espera que el número y el tamaño de las regiones de LOH sean representativos del número y tamaño de las regiones de LOH a través del genoma). En algunas realizaciones, el procedimiento 25 comprende además determinar el número total de regiones de LOH que son más largas que aproximadamente 1,5, 5, 12, 13, 14, 15, 16, 17 megabases o más (preferiblemente 14, 15, 16 megabases o más, más preferiblemente 15 megabases o más) pero más cortas que toda la longitud del cromosoma respectivo en el que se encuentra la región de LOH (Regiones de LOH Indicadoras). Alternativa o adicionalmente, se determina la longitud combinada total de dichas regiones de LOH indicadoras. En algunas realizaciones específicas, si el número total de regiones de LOH 30 indicadoras o la longitud total combinada de regiones de LOH indicadoras es igual o mayor que un número de referencia predeterminado, entonces dicha célula cancerosa o ADN genómico o un paciente que tiene dicha célula cancerosa o ADN genómico es identificado como que tiene una firma de LOH con deficiencia de HDR.

**[0192]** Otras realizaciones de la presente descripción se describen en los siguientes ejemplos. La presente 35 descripción se ilustra adicionalmente mediante los siguientes ejemplos que no deben interpretarse como una limitación adicional.

**[0193]** Los siguientes párrafos definen la descripción con más detalle

- 40 1. Un ensayo para seleccionar la terapia para un sujeto que tiene cáncer, el ensayo que comprende someter una muestra biológica que comprende una célula cancerosa o ácido nucleico de una célula cancerosa tomada del sujeto a un análisis de desequilibrio alélico telomérico (tAI); detectar el número de desequilibrio alélico telomérico (NtAI) en la célula cancerosa o ácido nucleico de la célula cancerosa, y
- 45 seleccionar una terapia que comprenda platino para el sujeto cuando se detecta que el NtAI está por encima de un valor de referencia basado en el reconocimiento de que la terapia que comprende platino es efectiva en pacientes que tienen el NtAI por encima del valor de referencia; y
- seleccionar una terapia contra el cáncer que no comprende platino para el sujeto cuando se detecta que el NtAI está por debajo de un valor de referencia basado en el reconocimiento de que la terapia contra el cáncer que comprende 50 platino no es efectiva en pacientes que tienen el NtAI por debajo de un valor de referencia.
2. El ensayo del párrafo 1 que comprende además la etapa de tratar al sujeto con la terapia seleccionada.
3. El ensayo de cualquiera de los párrafos anteriores, en el que el cáncer es cáncer de mama o cáncer de ovario.
4. El ensayo de cualquiera de los párrafos anteriores, en el que el valor de referencia es 22.
5. El ensayo de cualquiera de los párrafos anteriores, en el que el valor de referencia es 24.
- 55 6. El ensayo de cualquiera de los párrafos anteriores, en el que el valor de referencia es 27.
7. El ensayo de cualquiera de los párrafos anteriores, en el que la célula cancerosa no tiene mutaciones en el gen BRCA1 y/o BRCA2.
8. El ensayo de cualquiera de los párrafos anteriores que comprende además una etapa de ensayo para la expresión de ARNm de BRCA1 o el estado de metilación del promotor de BRCA1, detectando la cantidad de

expresión de ARNm de BRCA1 o la cantidad de metilación del promotor de BRCA1, donde se selecciona la terapia que comprende platino cuando se detecta una expresión disminuida de BRCA1 o metilación aumentada del promotor de BRCA1.

9. Un procedimiento para seleccionar una terapia que comprende platino para un sujeto que tiene cáncer que  
5 comprende  
someter una muestra biológica tomada del sujeto a un análisis de desequilibrio alélico (AI);  
detectar el número de AI; y  
seleccionar una terapia contra el cáncer que comprende platino para el sujeto cuando el número de AI está por  
encima de un valor de referencia basado en el reconocimiento de que la terapia contra el cáncer que comprende  
10 platino es eficaz en pacientes que tienen el número de AI por encima de un valor de referencia.
10. El procedimiento de cualquiera de los párrafos anteriores que comprende además la etapa de tratar al sujeto con  
la terapia contra el cáncer que comprende platino cuando se selecciona la terapia contra el cáncer que comprende  
platino.
11. El procedimiento de cualquiera de los párrafos anteriores, en el que el cáncer se selecciona entre cáncer de  
15 mama y cáncer de ovario.
12. El procedimiento de cualquiera de los párrafos anteriores, en el que el cáncer de mama no tiene mutaciones en  
BRCA1.
13. El procedimiento de cualquiera de los párrafos anteriores, en el que el desequilibrio alélico está dentro de  
aproximadamente 25 kB de una variación del número de copias (CNV).
- 20 14. El procedimiento de cualquiera de los párrafos anteriores, en el que la CNV es CNV pericentromérica o  
subtelomérica.
15. El procedimiento de cualquiera de los párrafos anteriores, en el que el desequilibrio alélico es desequilibrio  
alélico telomérico.
16. Un procedimiento que comprende:  
25 detectar, en una célula cancerosa o ADN genómico derivado de la misma, un desequilibrio alélico en un número  
representativo de pares de cromosomas humanos de la célula cancerosa; y  
determinar el número de desequilibrio alélico.
17. El procedimiento del párrafo 16, dicho número representativo de pares de cromosomas humanos que es  
representativo de todo el genoma.
- 30 18. El procedimiento del párrafo 16-17, que comprende además correlacionar un mayor número de regiones de  
desequilibrio alélico con una mayor probabilidad de deficiencia en HDR.
19. El procedimiento del párrafo 16-18, que comprende además correlacionar un mayor número de regiones de  
desequilibrio alélico con una mayor probabilidad de que dicha célula cancerosa responda a la terapia contra el  
cáncer que comprende platino.
- 35 20. El procedimiento del párrafo 16-19, que comprende además correlacionar un número no aumentado de regiones  
de desequilibrio alélico con una menor probabilidad de que dicha célula cancerosa responda a la terapia contra el  
cáncer que comprende platino.
21. El procedimiento del párrafo 16-21, en el que la terapia contra el cáncer que comprende platino comprende  
cisplatino, carboplatino, oxalapatino o picoplatino.
- 40 22. Un procedimiento que comprende:  
a) detectar, en una célula cancerosa o ADN genómico derivado de la misma, regiones de LOH en un número  
representativo de pares de cromosomas humanos de la célula cancerosa; y  
b) determinar el número y tamaño de dichas regiones de LOH.
23. El procedimiento del párrafo 22, dicho número representativo de pares de cromosomas humanos que es  
45 representativo de todo el genoma.
24. El procedimiento del párrafo 22-23, que comprende además correlacionar un mayor número de regiones de LOH  
de un tamaño particular con una mayor probabilidad de deficiencia en HDR.
25. El procedimiento del párrafo 22-24, en el que dicho tamaño particular es más largo que aproximadamente 1,5, 2,  
2,5, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 75 o 100 megabases y menor  
50 que la longitud del cromosoma completo que contiene la región de LOH.
26. El procedimiento de cualquiera de los párrafos 22-25, en el que 6, 7, 8, 9, 10, 11, 12 o 13 o más regiones de  
LOH de dicho tamaño particular se correlacionan con una mayor probabilidad de deficiencia en HDR.
27. Un procedimiento para determinar el pronóstico en un paciente que comprende:  
a) determinar si el paciente comprende células cancerosas que tienen una firma de LOH, en el que la presencia de  
55 más de un número de referencia de regiones de LOH en al menos un par de cromosomas humanos de una célula  
cancerosa del paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud del  
cromosoma completo que contiene la región de LOH indica que las células cancerosas tienen la firma de LOH,  
donde el al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que  
la primera longitud es de aproximadamente 1,5 o más megabases, y

- b)(1) determinar, basándose al menos en parte en la presencia de la firma de LOH, que el paciente tiene un pronóstico relativamente bueno, o b)(2) determinar, basándose al menos en parte en la ausencia de la firma de LOH, que el paciente tiene un pronóstico relativamente malo.
28. Una composición que comprende un agente terapéutico seleccionado del grupo que consiste en agente dañino para el ADN, antraciclina, inhibidor de la topoisomerasa I e inhibidor de PARP para usar en el tratamiento de un cáncer seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de esófago, cáncer de pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, cáncer de recto, cáncer colorrectal y cáncer de páncreas en un paciente con más de un número de referencia de regiones de LOH en al menos un par de cromosomas humanos de una célula cancerosa del paciente que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH, donde al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que la primera longitud es de aproximadamente 1,5 o más megabases.
29. La composición de cualquiera de los párrafos anteriores, en la que dichas regiones de LOH se determinan en al menos dos, cinco, diez o 21 pares de cromosomas humanos.
30. La composición de cualquiera de los párrafos anteriores, en la que el número total de dichas regiones de LOH es 9, 15, 20 o más.
31. La composición de cualquiera de los párrafos anteriores, en la que dicha primera longitud es de aproximadamente 6, 12 o 15 megabases o más.
32. La composición de cualquiera de los párrafos anteriores, en la que dicho número de referencia es 6, 7, 8, 9, 10, 11, 12 o 13 o superior.
33. Un procedimiento para tratar el cáncer en un paciente, que comprende:
- determinar en una muestra de dicho paciente el número de regiones de LOH en al menos un par de cromosomas humanos de una célula cancerosa del paciente con cáncer que son más largas que una primera longitud pero más cortas que la longitud del cromosoma completo que contiene la región de LOH indica que las células cancerosas tienen la firma de LOH, en la que el al menos un par de cromosomas humanos no es un par de cromosomas sexuales X/Y humanos, en el que la primera longitud es de aproximadamente 1,5 o más megabases;
  - proporcionar un valor de prueba derivado del número de dichas regiones de LOH;
  - comparar dicho valor de prueba con uno o más valores de referencia derivados del número de dichas regiones de LOH en una población de referencia (por ejemplo, media, mediana, terciles, cuartiles, quintiles, etc.); y
  - administrar a dicho paciente un fármaco anticanceroso, o recomendar o prescribir o iniciar un régimen de tratamiento que comprende quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revela que el valor de prueba es mayor (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor; al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 desviaciones típicas mayor) que al menos uno de dichos valores de referencia; o
  - recomendar o prescribir o iniciar un régimen de tratamiento que no incluya quimioterapia y/o un agente de letalidad sintético basado al menos en parte en dicha etapa de comparación que revele que el valor de prueba no es mayor (por ejemplo, no más de 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor; no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 desviaciones típicas mayor) que al menos uno de dichos valores de referencia.
34. El procedimiento del párrafo 33, en el que dichas regiones de LOH se determinan en al menos dos, cinco, diez o 21 pares de cromosomas humanos.
35. El procedimiento del párrafo 33-34, en el que el número total de dichas regiones de LOH es 9, 15, 20 o más.
36. El procedimiento del párrafo 33-35, en el que dicha primera longitud es de aproximadamente 6, 12 o 15 megabases o más.
37. El procedimiento del párrafo 33-36, en el que dicho número de referencia es 6, 7, 8, 9, 10, 11, 12 o 13 o superior.
38. El procedimiento del párrafo 33-37, en el que dicha quimioterapia se selecciona del grupo que consiste en un agente que daña el ADN, una antraciclina y un inhibidor de la topoisomerasa I y/o en el que dicho agente de letalidad sintético es un fármaco inhibidor de PARP.
39. El procedimiento del párrafo 33-38, en el que dicho agente que daña el ADN es cisplatino, carboplatino, oxalapatino o picoplatino, dicha antraciclina es epirubicina o doxorubicina, dicho inhibidor de la topoisomerasa I es campotecina, topotecán o irinotecán, y/o dicho inhibidor de PARP es iniparib, olaparib o velaparib.
40. Una composición que comprende un agente terapéutico seleccionado del grupo que consiste en terapia contra el cáncer que comprende platino y antraciclina para usar en el tratamiento de un cáncer seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de esófago, cáncer de pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, cáncer de recto, cáncer colorrectal y cáncer de páncreas en un paciente con desequilibrio alélico aumentado.
41. La composición del párrafo 40, en el que el desequilibrio alélico es un desequilibrio alélico telomérico.
42. La composición del párrafo 40-41, en el que el desequilibrio alélico está dentro de aproximadamente 25 kB de una variación del número de copias (CNV).
43. La composición del párrafo 40-42, en la que se determina además que el paciente no porta una mutación en

BRCA1 y/o BRCA2.

44. La composición del párrafo 40-43, en la que se determina además que el paciente tiene una cantidad disminuida de ARNm de BRCA1 en la célula cancerosa y/o se determina adicionalmente que tiene una metilación aumentada de la región promotora de BRCA1.
- 5 45. Un procedimiento para predecir el resultado del tratamiento anticanceroso de un sujeto con un trastorno hiperproliferativo celular, que comprende determinar una puntuación de aberración cromosómica global (GCAS), que comprende obtener una muestra biológica del sujeto y determinar si existe una pluralidad de regiones cromosómicas que muestra una aberración cromosómica dentro de una pluralidad de loci cromosómicos, donde dichas aberraciones cromosómicas se seleccionan del grupo que consiste en desequilibrio alélico (NAI), pérdida de  
10 heterocigosidad (NLOH), aberraciones del número de copias (NCNA), ganancia del número de copias (NCNG), disminución del número copias (NCND) y combinaciones de los mismos, con respecto 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 anticancerígeno del sujeto.
46. El procedimiento del párrafo 45, en el que el tratamiento anticanceroso es un tratamiento de quimioterapia.
- 15 47. El procedimiento del párrafo 45-46, en el que el tratamiento de quimioterapia comprende agentes quimioterapéuticos basados en platino.
48. El procedimiento del párrafo 45-47, en el que los agentes quimioterapéuticos basados en platino se seleccionan del grupo que consiste en cisplatino, carboplatino, oxaliplatino, nedaplatino e iproplatinato.
49. El procedimiento del párrafo 45-48, en el que el sujeto es un ser humano.
- 20 50. El procedimiento del párrafo 45-49, 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 de testículo, cáncer gástrico, sarcomas de tejido blando 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,  
25 mieloma múltiple refractario y carcinomas de células escamosas de la cabeza, cuello, cuello uterino, cáncer de colon, melanoma y vagina.
51. El procedimiento del párrafo 45-50, en el que la muestra biológica se selecciona del grupo que consiste en células, líneas celulares, portaobjetos histológicos, biopsias con núcleo congelado, tejidos embebidos en parafina, tejidos fijados con formalina, biopsias, sangre completa, aspirado de pezón, suero, plasma, raspado bucal, saliva,  
30 líquido cefalorraquídeo, orina, heces y médula ósea.
52. El procedimiento del párrafo 45-51, en el que la muestra biológica está enriquecida por la presencia de células hiperproliferativas a al menos el 75 % de la población total de células.
53. El procedimiento del párrafo 45-52, en el que el enriquecimiento se realiza según al menos una técnica seleccionada del grupo que consiste en microdissección con aguja, microdissección con láser, clasificación de células  
35 activadas por fluorescencia y clasificación inmunológica de células.
54. El procedimiento del párrafo 45-53, en el que una máquina automática realiza la al menos una técnica para transformar de ese modo la muestra biológica en una forma purificada enriquecida por la presencia de células hiperproliferativas.
55. El procedimiento del párrafo 45-54, en el que la muestra biológica se obtiene antes de que el sujeto haya  
40 recibido quimioterapia adyuvante.
56. El procedimiento del párrafo 45-55, en el que la muestra biológica se obtiene después de que el sujeto haya recibido quimioterapia adyuvante.
57. El procedimiento del párrafo 45-56, en el que el control se determina a partir de una muestra de células hiperproliferativas no celulares del paciente o miembro de la misma especie a la que pertenece el paciente.
- 45 58. El procedimiento del párrafo 45-58, en el que el control se determina a partir de la frecuencia promedio del aspecto del locus genómico de las regiones cromosómicas del mismo grupo étnico dentro de la especie a la que pertenece el paciente.
59. El procedimiento del párrafo 45-58, en el que el control es de tejido no canceroso que es del mismo tipo de tejido que dicho tejido canceroso del sujeto.
- 50 60. El procedimiento del párrafo 45-59, en el que el control es de tejido no canceroso que no es del mismo tipo de tejido que dicho tejido canceroso del sujeto.
61. El procedimiento del párrafo 45-60, en el que el NAI se determina usando la proporción de copia mayor (MCP).
62. El procedimiento del párrafo 45-61, en el que el NAI para una región genómica dada se cuenta cuando MCP es mayor que 0,70.
- 55 63. El procedimiento del párrafo 45-62, en el que la pluralidad de loci cromosómicos se distribuye aleatoriamente en todo el genoma al menos cada 100 Kb de ADN.
64. El procedimiento del párrafo 45-63, en el que la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada uno de los 23 pares de cromosomas humanos.
65. El procedimiento del párrafo 45-64, en el que la pluralidad de loci cromosómicos comprende al menos un locus

cromosómico en cada brazo de cada uno de los 23 pares de cromosomas humanos.

66. El procedimiento del párrafo 45-65, en el que la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en al menos un telómero de cada uno de los 23 pares de cromosomas humanos.

67. El procedimiento del párrafo 45-66, en el que la pluralidad de loci cromosómicos comprende al menos un locus cromosómico en cada telómero de cada uno de los 23 pares de cromosomas humanos.

68. El procedimiento del párrafo 45-67, en el que las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 1 Mb.

69. El procedimiento del párrafo 45-68, en el que las aberraciones cromosómicas tienen un tamaño de segmento mínimo de al menos 12 Mb.

70. El procedimiento del párrafo 45-69, en el que la pluralidad de aberraciones cromosómicas comprende al menos 5 aberraciones cromosómicas.

71. El procedimiento del párrafo 45-70, en el que la pluralidad de aberraciones cromosómicas comprende al menos 13 aberraciones cromosómicas.

72. El procedimiento del párrafo 45-71, en el que los loci cromosómicos se seleccionan del grupo que consiste en polimorfismos de un solo nucleótido (SNP), polimorfismos de longitud de fragmentos de restricción (RFLP) y repeticiones en tándem simples (STR).

73. El procedimiento del párrafo 45-72, en el que los loci cromosómicos se analizan usando al menos una técnica seleccionada del grupo que consiste en la sonda de inversión molecular (MIP), matriz de polimorfismo de nucleótido único (SNP), hibridación *in situ*, transferencia de Southern, matrices transcripcionales, hibridación genómica comparativa de matriz (aCGH) y secuenciación de próxima generación.

74. El procedimiento del párrafo 45-73, en el que el resultado del tratamiento se mide por al menos un criterio seleccionado del grupo que consiste en supervivencia hasta mortalidad, respuesta patológica completa, medidas semicuantitativas de respuesta patológica, remisión completa clínica, remisión parcial clínica, enfermedad estable clínica, supervivencia libre de recurrencia, supervivencia libre de metástasis, supervivencia libre de enfermedad, disminución de células tumorales circulantes, respuesta de marcador circulante y criterios RECIST.

75. El procedimiento del párrafo 45-74, que comprende además determinar un régimen de tratamiento adecuado para el sujeto.

76. El procedimiento del párrafo 45-75, en el que dicho régimen de tratamiento adecuado comprende al menos un agente quimioterapéutico basado en platino cuando se determina una pluralidad de aberraciones genómicas cromosómicas o no comprende al menos un agente quimioterapéutico basado en platino cuando no se determina una pluralidad de aberraciones cromosómicas genómicas.

## EJEMPLOS

### 35 Ejemplo 1: Materiales y procedimientos para el ejemplo 2

[0194] La respuesta patológica después del tratamiento neoadyuvante con cisplatino en la cohorte de TNBC se midió usando la escala semicuantitativa de Miller-Payne como se describe (Silver et al. (2010) J. Clin. Oncol. 28, 1145-1153; Ogston et al. (2003) Breast 12, 320-327). El genotipado de la MIP se realizó como se describe (Wang et al. (2007) Genome Biol. 8, R246). La intensidad de la señal del alelo y los genotipos del genotipado de la MIP o los análisis de la matriz de SNP pública se procesaron mediante el algoritmo CRLMM (Lin et al. (2008) Genome Biol. 9, R63) tal como se implementó en el paquete R "oligo". El número de copias de ADN se determinó usando el paquete R "AromaAffymetrix" (Bengtsson et al. (2008) Bioinformatics 24, 759-767). Los datos del genotipo procesado se exportaron a dChip (disponible en Internet en <http://biosun1.harvard.edu/complab/dchip/>) para la determinación de la proporción de copias principales (MCP), definida como la relación entre el número de copias principales y el número de copias principales + secundarias (Li et al. (2008) Bioinformatics 9, 204). Se hizo una estimación del nivel de contaminación normal de ADN a partir de la curva de MCP genómica como se describe (Li et al. (2008) Bioinformatics 9, 204). Se incluyeron casos de mamas u ovarios con un contenido tumoral del 75 % o más en los análisis. El desequilibrio alélico (AI) para fines específicos de algunos ejemplos descritos en este documento se definió como  $MCP > 0,7$  y regiones de AI definidas como más de 10 sondas consecutivas con AI. La AI telomérica para fines específicos de algunos Ejemplos descritos en este documento se definió como regiones de AI que se extienden a los telómeros y no cruzan el centrómero. La asociación entre el NtAI, 12 y la respuesta a cisplatino en sujetos con TNBC se estimó mediante el área bajo la curva (AUC) de la curva característica del operador receptor (ROC); el valor de p es de la prueba de rango de Wilcoxon de dos lados. La asociación entre el AI telomérico y el tiempo hasta la recurrencia del cáncer de ovario después de la terapia con platino se estimó mediante el análisis de Kaplan Meier utilizando un punto de corte de 13 para definir un alto NtAI, grupo 12; el valor de p se basa en la prueba de log-rank. Una lista completa de materiales y procedimientos es la siguiente:

*Ensayos de líneas celulares y sensibilidad a los medicamentos*

**[0195]** Las líneas celulares de cáncer de mama triple-negativo BT20, BT549, HCC1187, HCC38, MDA-MB231 y MDA-MB468 se mantuvieron a 37 °C con el 5 % de CO<sub>2</sub> en medio RPMI 1640 y/o medio MEM suplementado con FBS al 10 % u otros suplementos según lo recomendado por la ATCC para cada línea celular.

5 Para probar la sensibilidad del medicamento, las células se expusieron a una serie de concentraciones de cisplatino durante 48 horas. El número de células viables se cuantificó usando el ensayo CellTiter 96 Aqueous One Solution Cell Proliferation Assay de acuerdo con las instrucciones del fabricante (Promega). Los resultados se presentan como el porcentaje de células viables en pocillos tratados con fármaco frente a pocillos de control tratados con medio y representados como curvas de supervivencia celular dependientes de fármaco (Figura 1A). La sensibilidad  
10 del fármaco se cuantificó como la dosis de fármaco que causa una reducción del 50 % del crecimiento (CI<sub>50</sub>). Estos datos se generaron originalmente para un estudio separado en el que se informó como "datos no mostrados" en Li et al. (2010) Nat. Med. 16, 214-218.

#### *Cohorte de cáncer de mama*

15

**[0196]** Un total de 28 pacientes con TNBC principalmente esporádico se trataron con monoterapia con cisplatino en el entorno neoadyuvante (Silver et al. (2010) J. Clin. Oncol., 28, 1145-1153). La respuesta al cisplatino se midió utilizando la puntuación semicuantitativa de Miller Payne mediante evaluación patológica de muestras quirúrgicas después de la terapia (Ogston et al. (2003) Breast 12, 320-327). La respuesta completa patológica es  
20 equivalente a la puntuación de Miller-Payne y se define como carcinoma invasivo residual en mama o en los ganglios linfáticos.

#### *Preparación de muestras de cáncer de mama*

25 **[0197]** Se obtuvo una biopsia de núcleos congelados del tumor antes de comenzar el tratamiento. El tejido tumoral estaba disponible en la biopsia de núcleos congelados para 24 de 28 casos y en muestras de biopsia de núcleos diagnósticos embebidas en parafina y fijadas con formalina de 3 casos adicionales. Las células tumorales se enriquecieron mediante microdissección con aguja para eliminar el estroma de las secciones de tejido teñidas con hematoxilina y eosina (H & E). El tejido restante en los portaobjetos se examinó por microscopía para la estimación  
30 del enriquecimiento. El ADN se extrajo de células tumorales enriquecidas mediante digestiones con proteinasa K y RNasa A, extracción con fenol/cloroformo seguido de precipitación con etanol. El ADN adecuado para el análisis de genotipado de la MIP (mínimo 80 ng) se obtuvo de los 27 casos para los que el tejido tumoral estaba disponible. El ADN normal apareado de cada paciente se obtuvo a partir de linfocitos de sangre periférica.

#### 35 *Análisis de genotipado de la sonda de inversión molecular (MIP)*

**[0198]** El ADN de las muestras de biopsia tumoral de mama se envió a Affymetrix, Inc. (Santa Clara, CA) para el análisis de genotipado dirigido a MIP que generó la intensidad de señal de los alelos y genotipos para 42.000 polimorfismos individuales de nucleótido único (SNP). El conjunto completo de datos del genotipo de MIP está  
40 disponible en la base de datos NCBI GEO.

#### *Conjuntos de datos públicos*

45 **[0199]** Los perfiles genómicos de Affymetrix SNP 6.0 de seis líneas celulares de cáncer de mama triple negativo, BT20, BT549, HCC1187, HCC38, MDA-MB231 y MDA-MB468, se adquirieron del Wellcome Trust Sanger Institute (información disponible en Internet en <http://www.sanger.ac.uk/>).

50 **[0200]** Los datos de SNP que representan 118 tumores de carcinoma de ovario distribuidos en la plataforma Affymetrix 50K XbaI se adquirieron del omnibus de expresión génica (GEO, GSE13813, Etemadmoghadam et al. (2009) Clin. Cancer Res. 15, 1417-1427). De estos, 38 tumores eran del subtipo seroso, tenían un tumor residual después de una citorreducción quirúrgica de menos de 1 cm y habían recibido tratamiento con cisplatino adyuvante o con carboplatino. La mayoría de los pacientes (35 de 38) también habían recibido tratamiento con taxano.

#### *Genotipo y análisis de número de copias*

55

**[0201]** La intensidad de la señal del alelo y los genotipos del genotipado de la MIP o los análisis de la matriz de SNP se procesaron mediante el algoritmo CRLMM (Lin et al. (2008) Genome Biol. 9, R63) tal como se implementó en el paquete R "oligo". El número de copias de ADN se determinó usando el paquete R "AromaAffymetrix" (Bengtsson y otros (2008) Bioinformatics 24, 759-767). Los datos del genotipo procesado se

exportaron a dChip (disponible en Internet en <http://biosun1.harvard.edu/complab/dchip/>) para la determinación de la proporción de copias principales (MCP).

5 [0202] La MCP se define como la relación entre el número de copias de alelo principal y el número de copias de alelo principal + secundario (Li et al. (2008) *Bioinformatics* 9, 204). El grado de contaminación celular normal se estimó por el grado de desplazamiento en la curva MCP de la mayoría de las regiones que muestran un desequilibrio alélico en el genoma, excluyendo todas las regiones de ganancia del número de copias (El desplazamiento observado en las curvas de MCP genómicas en los experimentos de la mezcla de líneas celulares normales y tumorales se utilizó como referencia para estimar la contaminación normal como se describe (Waddell et al., (2009) *Breast Cancer Res. Treat.* (4 de diciembre; publicación electrónica)). En consecuencia, 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 contaminación normal de ADN ( $\geq 75$  % de contenido tumoral) y se consideraron aceptables para análisis posteriores.

15 [0203] El desequilibrio alélico (AI) se definió para algunos ejemplos descritos en este documento como MCP  $> 0,70$ , que permite la detección de la mayoría de eventos de pérdida de heterocigosidad (LOH) y de amplificaciones monoalélicas de alto número de copias en muestras con el 25 % o menos de contaminación o heterogeneidad, pero también excluye ganancias de copias de bajo nivel (ganancias de 4 copias o menos). Las regiones de AI se definieron para los fines de algunos ejemplos descritos en este 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 de LOH que se pueden llamar según se determina a partir de la comparación de genotipos convencionales. Los números de copia totales (que combinan ambos alelos) se segmentaron mediante el algoritmo de segmentación binaria circular. El ochenta y cinco por ciento de las regiones de AI tenían un número total de copias cercano a diploide o inferior, el 9 % de las regiones de AI mostraron una ganancia total de copia de 3, y el 6 % con una ganancia de copia total  $\geq 4$ . Por lo tanto, las regiones de AI identificadas representan predominantemente LOH o supresión cromosómica uniparental.

*Asociación entre el número de aberraciones genómicas y la sensibilidad al platino in vitro*

30 [0204] Los números de regiones de AI o regiones con aberración del número de copias se compararon con valores de  $CI_{50}$  específicos de línea celular después de aplicar un filtro de tamaño mínimo de 1 Mb para eliminar regiones muy pequeñas que podrían ser causadas por ruido en los datos de SNP 6.0 (Figura 3). Para la comparación de las regiones de AI telomérico e intersticiales, se definió el AI telomérico para los fines de los Ejemplos descritos en la presente memoria como el AI que se extiende al telómero pero que no cruza el centrómero. Por el contrario, el AI intersticial se definió para los fines de los Ejemplos descritos en la presente memoria como regiones de AI que no implican al 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 con la  $CI_{50}$  del cisplatino, se utilizó la regresión lineal para comparar los valores de  $CI_{50}$  con el número total de segmentos mayores que un determinado umbral, que se incrementó en intervalos de 1 Mb entre 0 y 100 Mb (Figura 5).

40 *Asociación entre el número de regiones de AI teloméricas y la sensibilidad al platino en tumores*

[0205] Se determinó el número total de regiones de AI telomérico para cada caso de TNBC con al menos un 75 % de contenido tumoral. Se aplicó el umbral de tamaño de segmento de AI telomérico mínimo óptimo de 12 Mb encontrado en las líneas celulares, y se contaron los NtAI, 12 para cada sujeto. Se realizó un análisis de curva ROC (Característica operativa del receptor) para evaluar la capacidad del número total de segmentos de AI teloméricos para predecir la pCR (puntuación 5 de Miller-Payne) para el tratamiento con cisplatino.

50 [0206] La asociación de NtAI, 12 con pCR a cisplatino se estimó por el área bajo la curva (AUC); el valor de p correspondiente es de la prueba de rango de Wilcoxon de dos lados. Con base en el análisis ROC, un NtAI.12 de 13 dio como resultado un 100 % de sensibilidad para la predicción de pCR en la cohorte de TNBC tratada con cisplatino.

[0207] La asociación entre NtAI, 12 y el tiempo hasta la recurrencia después del tratamiento basado en platino en la cohorte de cáncer de ovario se estimó mediante análisis de Kaplan-Meier con el grupo "alto NtAI, 12" definido como al menos 13 regiones de NtAI, 12. El valor de P se basa en una prueba de log-rank.

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

**[0208]** Sin estar limitados por la teoría, se cree que la pérdida de heterocigosis (LOH) intracromosómica o el desequilibrio alélico (AI) resulta de una reparación incorrecta de las roturas de la doble cadena del ADN cromosómico y que el recuento genómico completo de estos reordenamientos cromosómicos en un tumor específico puede indicar el grado de incompetencia de reparación del ADN, independientemente del defecto de reparación causante del ADN específico. Por lo tanto, el número total de reordenamientos cromosómicos en un tumor refleja la incapacidad de reparar el daño en el ADN inducido por fármacos como el cisplatino y, en consecuencia, predice la sensibilidad a estos agentes. Se determinó así la sensibilidad al cisplatino de seis líneas celulares de TNBC para las cuales los datos de la matriz de SNP estaban disponibles en Wellcome Trust Sanger Institute, Reino Unido (Figura 1A). La AI se determinó por análisis de la proporción de copias principales (MCP), un procedimiento menos sensible a la contaminación normal en muestras de tumores heterogéneos (Li et al. (2008) *Bioinformatics* 9, 204).

**[0209]** El MCP es el número de alelos de copias principales en un locus dividido por la suma del número de alelos de copias principales y secundarias (Figura 2). Las ganancias o reducciones en el número total de copias de ADN en cada región cromosómica se infirieron usando el software dChip (Lin et al. (2004) *Bioinformatics* 20, 1233-1240).

**[0210]** La(s) lesión(es) de reparación del ADN que hace(n) que las células sean sensibles al cisplatino puede(n) inducir preferentemente alteraciones cromosómicas de un tipo específico o con un rango de tamaño específico. En las seis líneas celulares, se probó la asociación entre la sensibilidad al cisplatino y cada una de las cuatro medidas de alteraciones cromosómicas. Las cuatro medidas fueron (1) el número de regiones cromosómicas con AI (NAI), (2) el número de aberraciones del número de copias (NCNA), (3) el número de regiones con ganancia de número de copias, y (4) el número de regiones con disminución del número de copias (Figura 3). De estas cuatro medidas, el NAI se correlacionó más fuertemente con la sensibilidad al cisplatino ( $R^2 = 0,5$ ).

**[0211]** Los defectos conocidos en la reparación de la rotura de la doble cadena de ADN, incluida la pérdida de BRCA1 o mutaciones en la helicasa de Bloom, provocan la formación espontánea de estructuras trirradiales y cuadrirradiales cromosómicas, que son indicaciones citológicas 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 como resultado 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, se compararon las regiones de AI telomérico e intersticiales (no teloméricas) y se encontró que la correlación entre la sensibilidad al cisplatino y el AI era más fuerte cuando se limitaba a regiones de AI que implican telómeros, mientras que solo se observaba una asociación débil entre la sensibilidad al cisplatino y el número de regiones de AI intersticiales (Figura 4).

**[0212]** A continuación, se determinó si podrían mejorarse las correlaciones entre la sensibilidad al cisplatino y las medidas de aberraciones genómicas mediante la prueba de un rango de tamaños de segmento mínimo, en líneas celulares TNBC (Figura 1B y Figuras 5A-5C). Se observó una correlación significativa con la sensibilidad al cisplatino utilizando cortes mínimos de tamaño del segmento de AI telomérico entre 5 y 25 Mb con el nivel más alto de correlación visto para el número total de segmentos con AI telomérico (NtAI) de al menos 12 MB ( $R^2 = 0,8$ ;  $P = 0,016$ ; Figura 1C). La prueba para el tamaño óptimo del segmento mínimo no mejoró apreciablemente la correlación entre la sensibilidad al cisplatino y las medidas de las aberraciones del número de copias, que no se mantuvieron significativas (Figuras 5D-5F).

**[0213]** También se investigó si la misma asociación entre el NtAI y la sensibilidad al cisplatino estaba presente en muestras de tumores clínicos utilizando el corte de tamaño de segmento óptimo de 12 MB (NtAI, 12). El NtAI, 12 se comparó con la respuesta a la quimioterapia en sujetos con TNBC tratados con monoterapia con cisplatino preoperatorio (Silver et al. (2010) *J. Clin. Oncol.* 28, 1145-1153). Secciones de tejido criostático de biopsias de núcleo antes del tratamiento se enriquecieron para células tumorales mediante microdissección con aguja, y se extrajo ADN para genotipado. Los genotipos de 42.000 SNP se determinaron con el sistema de genotipado dirigido de la Sonda de Inversión Molecular (MIP) (Affymetrix, Inc.) (Wang et al. (2007) *Genome Biol.* 8, R246). El grado de contaminación celular normal se estimó a partir de los datos del genotipo de MIP como se describe (Li et al. (2008) *Bioinformatics* 9, 204). No se observó asociación entre el grado de contaminación normal y la respuesta a cisplatino ( $R^2 = 0,004$ ,  $P = 0,75$ ).

**[0214]** Los datos del genotipo de MIP de 21 casos con al menos un 75 % de contenido de células tumorales se evaluaron mediante análisis MCP para definir las regiones de AI telomérico, intersticial o de cromosoma completo en todo el genoma (Figura 6A y Figura 7). Se observó una correlación entre el NtAI, 12 y la tasa de respuesta, según

se cuantificó mediante 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 de tAI asociadas con una mayor sensibilidad al cisplatino. El análisis de la curva característica operativa del receptor (ROC) reveló que el NtAI, 12 se asoció significativamente con la respuesta patológica completa a cisplatino (Miller-Payne 5) por el área bajo la curva ( $AUC = 0,85$ ;  $P = 0,017$ ; 5 Figura 6C). No hubo asociación aparente entre el número de segmentos de AI intersticiales (Figura 6A) o el nivel de AI del cromosoma entero (Figura 7) y la respuesta a cisplatino.

**[0215]** El carcinoma de ovario seroso a menudo se trata con terapias basadas en platino. Se investigó un conjunto de datos de matriz de SNP disponible públicamente de carcinomas ováricos tratados con cisplatino o 10 carboplatino más un taxano (Etemadmoghadam et al. (2009) Clin. Cancer Res. 15, 1417-1427) y se identificaron 33 casos del subtipo seroso después de citorreducción quirúrgica óptima (tumor residual  $<1$  cm) y una pureza tumoral razonable ( $> 75$  %, estimada a partir de datos de SNP). El NtAI, 12 se determinó mediante análisis MCP. En estos casos de cáncer de ovario tratado con platino, se encontró una asociación entre niveles más altos de AI telomérico en tumores y ausencia de recaída en un año (Figura 8A). El análisis ROC en la cohorte de TNBC se utilizó para 15 definir un valor de corte de NtAI, 12 de al menos 13 eventos, que dieron la mayor sensibilidad para la clasificación de pCR a terapia de platino en la cohorte de TNBC. Este punto de corte se utilizó para clasificar la cohorte de cáncer de ovario en grupos de NtAI, 12 alto y bajo y se encontró una supervivencia libre de enfermedad más prolongada, un indicador sustituto de mayor sensibilidad al platino, en el grupo de NtAI, 12 alto (Figura 8B).

20 **[0216]** Por lo tanto, la inestabilidad cromosómica, que se manifiesta por altos niveles de AI telomérico, caracteriza 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 algunos otros inmuno-fenotipos y características histológicas, se desconoce el defecto o defectos moleculares precisos en el mantenimiento de la estabilidad 25 cromosómica en estos tumores. Los resultados de los ejemplos descritos en este documento indican que la carga de reordenamientos cromosómicos que resultan de roturas de la cadena de ADN reparadas incorrectamente son indicadores de defectos de reparación de ADN que sensibilizan las células a ciertas quimioterapias (Figura 9). Como tal, los niveles de desequilibrio alélico 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 30 del ADN.

### Ejemplo 3

**[0217]** En este estudio, utilizamos dos ensayos clínicos preoperatorios en mujeres con cáncer de mama triple 35 negativo tratado con cisplatino, en el que la respuesta patológica en el momento de la cirugía proporcionó un punto final experimental. Los cánceres de mama triple negativo esporádicos son heterogéneos en sus respuestas a las sales de platino, agentes quimioterapéuticos que dependen en parte de los defectos de reparación del ADN para su actividad citotóxica (Sakai, W., et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. Nature 2008; 451: 1116-1120; Edwards, S. L., et al. Resistance to therapy caused by intragenic deletion in 40 BRCA2. Nature 2008; 451: 1111-1115). Las lesiones en la reparación del ADN causada por la disfunción en BRCA1 o BRCA2 dan lugar a la sensibilidad al platino; razonamos que los tipos de aberraciones cromosómicas que surgen en el contexto de la disfunción de BRCA también podrían estar asociados con la sensibilidad al platino en cánceres wt-BRCA (BRCA de tipo silvestre). En base a los resultados en las líneas celulares, elegimos enumerar una de esas anomalías cromosómicas, el desequilibrio alélico telomérico (NtAI) en los genomas tumorales antes al tratamiento y 45 relacionar esto con la respuesta patológica después del cisplatino, una terapia ejemplar que comprende platino.

**[0218]** El NtAI se asoció con la respuesta al tratamiento con platino en nuestros ensayos de TNBC con cisplatino y en el cáncer de ovario seroso tratado con platino y sugiere que la carga de esta anomalía genómica expone una deficiencia subyacente de reparación del ADN en el subconjunto sensible al platino de estos cánceres. 50 El desequilibrio alélico propagado desde una ubicación cromosómica determinada al telómero sugiere la operación de procesos propensos a errores que dan lugar a eventos de cruce o de cambio de molde anormales, en lugar de reparación de ADN libre de errores.

**[0219]** Encontramos que los puntos de corte de las regiones de tAI no son aleatorios y están enriquecidos 55 para CNV. Este patrón también sugiere una reparación defectuosa del ADN. Las CNV se asocian a otras secuencias repetitivas como las repeticiones Alu, que se concentran en regiones pericentroméricas y subteloméricas y se asocian también con sitios frágiles comunes (McVean, G. What drives recombination hotspots to repeat DNA in humans? Philos Trans R Soc Lond B Biol Sci 2010; 365: 1213-1218; Puliti, A., et al. Low-copy repeats on chromosome 22q11.2 show replication timing switches, DNA flexibility peaks and stress inducible asynchrony,

sharing instability features with fragile sites. *Mutat Res* 2010; 686: 74-83). Se cree que estos elementos repetitivos dan lugar a "zonas lentas" de replicación propensas a estancamiento de replicación y formación de roturas de la doble cadena de ADN (Richard, G. F., Kerrest, A., and Dujon, B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev* 2008; 72: 686-727; Cha, R. S. and Kleckner, N. ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science* 2002; 297: 602-606). Además, la regulación por disminución de Rad51 o la inhibición de BRCA1 aumenta la fragilidad en dichos sitios cuando las células están bajo estrés de replicación (Arlt, M. F., et al., BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function. *Mol Cell Biol* 2004; 24: 6701-6709; Schwartz, M., et al. Homologous recombination and nonhomologous end-joining repair pathways regulate fragile site stability. *Genes Dev* 2005; 19: 2715-2726). La asociación observada de bajos niveles de expresión de BRCA1 en muchos tumores con alto NtAI sugiere una recombinación homóloga deficiente, una función de punto de control S o G2/M alterada, o una combinación de estos factores que subyace a la generación de este tipo de anomalía genómica.

**[0220]** El cisplatino forma enlaces entre intercatenarios en el ADN que dan lugar a horquillas de replicación estancadas y roturas de la doble cadena de ADN que deben repararse para que la célula pueda sobrevivir. Es probable que estas roturas se reparen usando mecanismos similares a los empleados en horquillas de replicación estancadas y roturas de ADN generadas en sitios de CNV. Por lo tanto, el pretratamiento alto de NtAI identifica tumores incapaces de reparar con precisión las roturas y reiniciar las horquillas de replicación estancadas en los sitios de CNV. Estos mismos tumores también son incapaces de contender con las horquillas estancadas en sitios de enlaces cruzados de cisplatino.

**[0221]** Si bien el desequilibrio alélico en los sitios de CNV puede reflejar una reparación ineficiente sin errores, se deben considerar otras explicaciones. Ambas cohortes triples negativas mostraron una relación significativa entre el NtAI y la respuesta patológica a la quimioterapia con cisplatino. Sin embargo, hubo pacientes en ambos ensayos cuyos tumores mostraron una respuesta deficiente a la terapia de cisplatino a pesar de tener un alto NtAI. Del mismo modo, algunos de los cánceres de ovario mutados en BRCA1 tenían un alto NtAI pero eran resistentes a la terapia con platino. Dado que el NtAI es una suma de lesiones de ADN en curso y pasadas, los mecanismos de resistencia adquiridos después de la generación de tAI confundirían la relación entre el NtAI y la respuesta. En los portadores de mutaciones en BRCA1 o BRCA2, algunos tumores que se vuelven resistentes a los agentes de platino tienen una mutación de reversión que restablece parcial o completamente la función BRCA1 o BRCA2 y restaura la recombinación homóloga (Sakai, W., et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* 2008; 451: 1116-1120; Edwards, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* 2008; 451: 1111-1115; Swisher, E. M., et al., Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. *Cancer Res* 2008; 68: 2581-2586). También se ha visto reversión en una línea celular con una mutación en BRCA2 seleccionada para resistencia al inhibidor de PARP (Edwards, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* 2008; 451: 1111-1115). Se observaron mutaciones de reversión y mutaciones compensadoras cis en pacientes con anemia de Fanconi, lo que produjo una mejoría en la función de la médula ósea (Kalb, R., et al., Fanconi anemia: causes and consequences of genetic instability. *Genome Dyn* 2006; 1:218-242). La inactivación de TP53BP1 restablece el equilibrio entre la recombinación homóloga y la unión de extremos no homólogos en células mutadas de BRCA1 y los hace resistentes a inhibidores de PARP (Bouwman, P., et al. 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nat Struct Mol Biol* 2010; 17: 688-695; Bunting, S. F., et al. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* 2010; 141: 243-254). Finalmente, los transportadores de fármacos pueden evitar la acumulación de agentes de platino en las células tumorales (Burger, H., et al., Drug transporters of platinum-based anticancer agents and their clinical significance. *Drug Resist Updat* 2011). Por lo tanto, la reversión o compensación de un defecto de reparación de ADN preexistente puede generar un tumor con alto NtAI pero resistencia al tratamiento con platino; otros mecanismos de resistencia al platino no relacionados con la reparación del ADN tendrían el mismo efecto.

**[0222]** Nuestro análisis sugiere un esquema de la taxonomía molecular del TNBC y el cáncer de ovario con respecto a la reparación del ADN y la sensibilidad a los medicamentos. La mayoría de los cánceres de mama o de ovario resistentes al platino son tumores con capacidad de reparación y niveles bajos de NtAI. Dos subconjuntos de tumores wt-BRCA poseen alto NtAI y son sensibles a los fármacos que contienen platino. En uno de estos subconjuntos, la deficiencia de la reparación puede ser la consecuencia de una baja expresión de BRCA1 y en el otro subconjunto, la reparación puede verse afectada por mecanismos que no dependen de la expresión de BRCA1. Estas observaciones sin duda se refinarán aún más; la inclusión de mutaciones de reversión, compensaciones por otros eventos en las vías de reparación del ADN, otros mecanismos de resistencia a los fármacos y otros factores aún no apreciados pueden ayudar a mejorar nuestra predicción de la sensibilidad a los medicamentos en el futuro.

**[0223]** En conclusión, una medida resumida de las aberraciones cromosómicas teloméricas en el genoma del tumor, NtAI, predice la sensibilidad al tratamiento con platino. Nuestros hallazgos implican el NtAI como marcador de deterioro de la reparación de la rotura de la doble cadena de ADN. Los ensayos para determinar el NtAI son factibles utilizando material tumoral embebido en parafina fijado con formalina y algoritmos recientes como ASCAT permiten la determinación precisa del número de copias y el desequilibrio alélico en la mayoría de las muestras a pesar del bajo contenido de células tumorales. El NtAI puede ser útil para predecir la respuesta a una variedad de estrategias terapéuticas que explotan la reparación defectuosa del ADN.

#### *Materiales y procedimientos*

10

**[0224]** Líneas celulares y ensayos de sensibilidad a fármacos: Mediciones de sensibilidad a fármacos en líneas celulares de cáncer de mama BT20, BT549, HCC1187, HCC1143, MDA-MB-231, MDA-MB-468, HCC38, MDA-MB-453 (triple negativo), CAMA-1, MCF7, T47D (ER positivo), BT474, HCC1954 y MDA-MB-361 (HER2 positivo) se generaron originalmente para un estudio separado en el que se informó como "datos no mostrados" en un manuscrito recientemente publicado (Li, Y., et al. Amplification of LAPT4B and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer. Nat Med 2010; 16: 214-218). Brevemente, las células se expusieron a una serie de concentraciones de diversos agentes quimioterapéuticos durante 48 horas. El número de células viables se cuantificó usando el ensayo CellTiter 96 AQueous One Solution Cell Proliferation Assay de acuerdo con las instrucciones del fabricante (Promega). La sensibilidad del fármaco se cuantificó como la dosis de fármaco que da como resultado una reducción del 50 % del crecimiento ( $Cl_{50}$ ). Encontramos que MCF7 es altamente resistente a todos los agentes quimioterapéuticos probados, consistente con su deficiencia reportada de caspasa-3 y resistencia a la apoptosis inducida por fármacos (Yang, XH, et al., Reconstitution of caspase 3 sensitizes MCF-7 breast cancer cells to doxorubicin- and etoposide-induced apoptosis. Cancer Res 2001; 61: 348-354). En nuestros análisis con medidas de aberración genómica, MCF7 fue el único valor atípico claro y por estas razones, se excluyó de nuestros análisis.

#### *Cohortes de cáncer de mama y evaluación de la respuesta terapéutica*

**[0225]** Para este estudio, los sujetos se incluyeron para el análisis de la respuesta al cisplatino si progresaron en el tratamiento o si recibieron al menos 3 de 4 ciclos del tratamiento planeado con cisplatino, no recibieron ningún otro tratamiento fuera del protocolo antes de la cirugía y si había disponible una cantidad adecuada del tumor a partir de la biopsia previa al tratamiento. La respuesta terapéutica se midió utilizando el sistema semicuantitativo de clasificación de Miller Payne, que estima el porcentaje de reducción del volumen tumoral invasivo y la celularidad basándose en la evaluación patológica de muestras quirúrgicas después del tratamiento (Ogston, K. N., et al., A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. Breast 2003; 12: 320-327). Cisplatino-1 consta de 28 pacientes con TNBC principalmente esporádico tratados con cisplatino preoperatorio en monoterapia, de los cuales 4 progresaron en el tratamiento y 24 completaron 4 ciclos de tratamiento con cisplatino (Silver, D. P., et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. J Clin Oncol 2010; 28: 1145-1153). Cisplatino-2 consiste en 51 pacientes con TNBC tratados con cisplatino preoperatorio y bevacizumab, de los cuales un paciente progresó a terapia y 44 pacientes completaron 4 ciclos de terapia con cisplatino antes de la cirugía (Ryan, P. D., et al. Neoadjuvant cisplatin and bevacizumab in triple negative breast cancer (TNBC): Safety and Efficacy. J Clin Oncol 2009; 27: 551). Dos pacientes incluidos en este estudio fueron llevados a cirugía después de completar 3 ciclos de terapia con cisplatino debido al desarrollo de toxicidad; en ambos casos no hubo una respuesta patológica apreciable en el tumor extirpado después de 3 ciclos de cisplatino.

#### *Preparación de muestras de cáncer de mama*

**[0226]** Para ambos ensayos, se obtuvieron biopsias de tumor antes del inicio del tratamiento. El tumor adecuado para el análisis estaba presente para 27 de 28 sujetos en Cisplatino-1 y 37 de 51 sujetos en Cisplatino-2. Se examinaron microscópicamente las secciones de tejido teñidas con H & E de las biopsias con aguja gruesa previa al tratamiento; para todas las biopsias cuyo enriquecimiento se consideró factible, las secciones se microdisecionaron manualmente con una aguja de calibre 18. El ADN se extrajo mediante digestiones con proteinasa K y RNasa A, extracción con fenol/cloroformo y precipitación con etanol. El ADN normal emparejado de los pacientes se obtuvo de linfocitos de sangre periférica para todos los casos en Cisplatino-1 y de 10 casos en Cisplatino-2.

#### *Cohortes de cáncer de ovario y de mama de TCGA*

**[0227]** Se obtuvieron datos públicos de matriz de SNP, datos de expresión y anotación clínica para cohortes de cáncer de ovario de TCGA (Bell, D., et al., Integrated genomic analyses of ovarian carcinoma. Nature 2011; 474: 609-615) y cohortes de cáncer de mama del sitio web de TCGA (<http://tcga-data.nci.nih.gov/tcga/>). El estado de mutación en BRCA1 y BRCA2 para los cánceres de ovario se obtuvo del portal de datos cBIO (<http://bit.ly/wpwRXd>).

- 5 En la cohorte de cáncer de ovario, identificamos 218 muestras con datos de SNP que pasaron ASCAT, estado de mutación en BRCA y anotaciones clínicas interpretables para el tratamiento y los resultados que indican tratamiento inicial con quimioterapia adyuvante basada en platino, predominantemente la combinación de carboplatino y docetaxel. Clasificamos "tratamiento sensible" como aquellos anotados como respuesta parcial o completa al tratamiento inicial y sin progresión o recurrencia dentro de los 6 meses del tratamiento inicial (n = 187); "resistente al
- 10 tratamiento" fueron aquellos anotados como enfermedad estable o progresiva en la terapia inicial o recurrencia o progresión de la enfermedad dentro de los 6 meses (n = 31). En la cohorte de cáncer de mama, identificamos 78 muestras con expresión génica coincidente y datos de SNP que pasaron ASCAT, que se clasificaron como ER-/HER2- basado en la agrupación del gen ESR1 y ERBB2 (ver los procedimientos complementarios).

#### 15 *Genotipado y análisis del número de copias*

- [0228]** El ADN se envió a Affymetrix, Inc. (Santa Clara, CA) para la determinación de genotipos utilizando el sistema de genotipado basado en la sonda de inversión molecular, OncoScan FFPE Express (Wang, Y., et al. Analysis of molecular inversion probe performance for allele copy number determination. Genome Biol 2007; 8: R246). El análisis comercial, que determina el genotipo de 330.000 SNP, se utilizó para el análisis del ensayo
- 20 Cisplatino-2. Una versión temprana del análisis OncoScan cuyos genotipos de 42.000 SNP se utilizó para el ensayo Cisplatino-1. La intensidad de la señal del alelo y los genotipos del ensayo de genotipado OncoScan fueron procesados y proporcionados por Affymetrix. Los datos del genotipo OncoScan SNP para los ensayos de terapia con cisplatino se envían a la base de datos NCBI GEO con acceso GSE28330. Los datos en bruto de la matriz de SNP
- 25 públicos para las líneas celulares de cáncer de mama se obtuvieron del sitio web del Sanger Institute del sitio web Somatic Mutations In Cancer, "punto" sanger "punto" ac "punto" uk/cosmic (Bamford, S., et. al. Base de datos y sitio web COSMIC (Catálogo de mutaciones somáticas en el cáncer) Br J Cancer 2004; 91: 355-358), datos públicos de la matriz de SNP de un estudio independiente de línea celular de cáncer de mama, Heiser et al. (Heiser, L. M., et al. Subtype and pathway specific responses to anticancer compounds in breast cancer. Proc Natl Acad Sci USA 2011),
- 30 y los datos públicos de la matriz de SNP de las cohortes de cáncer de ovario TCGA (Bell, D., et al., Integrated genomic analyses of ovarian carcinoma. Nature 2011; 474: 609-615) y las cohortes de cáncer de mama fueron preprocesados por los algoritmos AROMAv2 y CalMaTe (Bengtsson, H., Wirapati, P., and Speed, T. P. A single-array preprocessing method for estimating full-resolution raw copy numbers from all Affymetrix genotyping arrays including GenomeWideSNP 5 & 6. Bioinformatics 2009;25: 2149-2156) y, cuando se disponía de muestras normales
- 35 pareadas, TumorBoost (Bengtsson, H., Neuvial, P., and Speed, T. P. TumorBoost: normalization of allele-specific tumor copy numbers from a single pair of tumor-normal genotyping microarrays. BMC Bioinformatics 2010; 11: 245). Los datos de genotipo procesado del genotipado OncoScan y los datos de la matriz de SNP públicos se analizaron para números de copias específicas de alelos y contenido de células tumorales mediante el algoritmo "Análisis de tumores con número de copias de alelos específico", ASCAT (Van Loo, P., et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci USA 2010; 107: 16910-16915). ASCAT está diseñado para corregir la contaminación celular normal y la ploidía de células tumorales, pero ocasionalmente no se ajusta un modelo a una muestra determinada. En este estudio, ASCAT no pudo procesar 3 de 14 líneas celulares de Sanger, 15 de 42 líneas celulares de Heiser et al., y 5 de 37 muestras del ensayo Cisplatino-2. El desequilibrio alélico se definió como cualquier momento en que el número de copias de los dos alelos no fuera igual, y estaba presente al menos un alelo
- 40 (Fig. 16). Para garantizar que todos los casos de prueba fueran comparables, eliminamos los casos estimados por ASCAT de tener menos del 36 % de contenido de células tumorales, el nivel más alto de mezcla celular normal en el ensayo Cisplatino-1, que fue el ensayo con una mayor pureza tumoral general. Por lo tanto, se incluyeron las 27 muestras con los datos de la matriz de SNP del ensayo Cisplatino-1, 26 de 32 muestras con datos de la matriz de SNP que pasaron ASCAT del ensayo Cisplatino-2.

50

- [0229]** Se requirió un número mínimo de sondas consecutivas que mostraban una aberración para llamar a las regiones de AI y CNA con confianza. Para garantizar una detección de aberración similar en las tres plataformas que se utilizaron, se estableció que la cantidad mínima de sondas necesarias para definir una región de aberración era proporcional a la densidad general de SNP de la plataforma. Las densidades de la sonda de las plataformas fueron 42.000/genoma OncoScan (prototipo), 330.000/genoma OncoScan FFPE Express y 900.000/genoma SNP
- 55 6.0 para una relación aproximada de 1:8:20. Los requisitos mínimos de sonda de 25 sondas para el prototipo OncoScan de 42k, 200 sondas para 330k OncoScan FFPE Express y 500 sondas para la plataforma SNP6.0 se eligieron basándose en la optimización de la correlación de la medición de aberraciones en un subconjunto de muestras con datos duplicados generados en ambas versiones de la plataforma OncoScan (Ver también

Procedimientos Suplementarios).

**[0230]** El AI telomérico y la CNA telomérica se definen como regiones que se extienden a uno de los telómeros secundarios pero que no cruzan el centrómero. El número de copias de las regiones de AI telomérico se definió como el número medio de copias de las sondas que mapean en la región. La pérdida de copia se definió como una media de menos de 1,5 copias y la ganancia de copia se definió como una media de más de 2,5 copias. La asociación entre el NtAI y la respuesta a cisplatino en los ensayos clínicos de TNBC se midió por el AUC de la curva ROC para la respuesta binaria. La significancia estadística se evaluó mediante la prueba de suma de rangos de Wilcoxon. Todos los valores de P son bilaterales.

10

*Enriquecimiento de las variantes del número de copias en el sitio de los puntos críticos de ADN*

**[0231]** La ubicación genómica de las variantes comunes de número de copias (CNV) se adquirió de la Base de datos de variantes genómicas (<http://projects.tcag.ca/variation/>). Se adquirió el mapeo para HG17 y HG18 para que coincida con el mapeo de la sonda SNP del prototipo 42K y las plataformas comerciales OncoScan de 330 K, respectivamente. Las CNV se consideraron asociadas a un punto de rotura si se superponían dentro de una ventana de 25 kb a cada lado del punto de corte. Para evaluar el enriquecimiento, realizamos 1000 permutaciones para cada cohorte, donde barajamos aleatoriamente la ubicación de los puntos de corte de ADN en función de la ubicación de las sondas de SNP, y determinamos cuántas se asociaron con CNV.

15

20

*Cuantificación de la transcripción de BRCA1 y análisis de metilación del promotor*

**[0232]** El ensayo de reacción en cadena de la polimerasa cuantitativa del exón 16/17 de BRCA1 y RPLP0 (control) se realizó como se ha descrito anteriormente (Silver, D. P., et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. J Clin Oncol 2010; 28: 1145-1153) usando ADNc de tumor amplificado generado usando el kit Ovation RNA Amplification System V2 (NuGen Technologies, Inc., San Carlos, CA). El ensayo de metilación del promotor de BRCA1 se realizó como se ha descrito previamente (Silver, D. P., et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. J Clin Oncol 2010; 28: 1145-1153).

25

30

*Expresión de BRCA1 en cohortes públicas de TCGA.*

**[0233]** Los datos de expresión genética basados en Agilent normalizados y resumidos públicos se adquirieron de TCGA para todas las muestras de cáncer de mama (nivel 3). Se obtuvieron archivos Raw Affymetrix CEL para muestras de cáncer de ovario (nivel 1). Los datos de expresión para todas las muestras de cáncer de ovario TCGA se normalizaron y resumieron usando RMA, y el conjunto de sonda "204531\_s\_at" se identificó como el conjunto de sonda óptimo para medir la expresión de BRCA1 usando el paquete R "JetSet" (Li, Q., et al., Jetset: selecting the optimal microarray probe set to represent a gene. BMC Bioinformatics 2011; 12: 474).

35

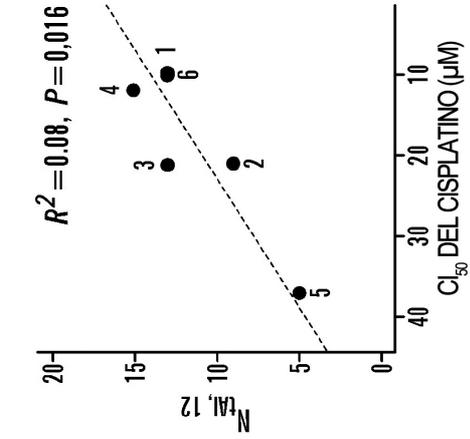
40

#### **Equivalentes**

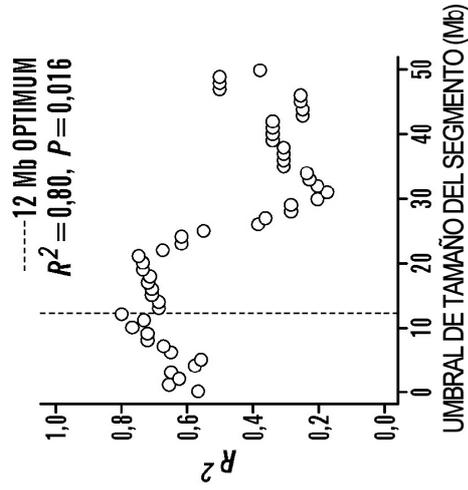
**[0234]** Los expertos en la materia reconocerán, o serán capaces de determinar usando no más que la experimentación de rutina, muchos equivalentes a las realizaciones específicas de la invención descrita en el presente documento.

**REIVINDICACIONES**

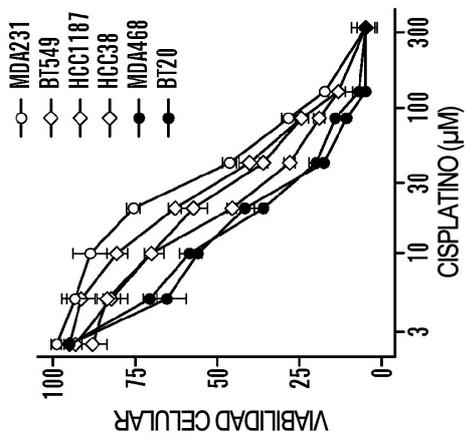
1. Procedimiento *in vitro* para seleccionar una terapia que comprende platino para un sujeto humano que tiene cáncer, que comprende:
- 5 someter una muestra biológica que comprende células tumorales tomadas del ser humano a análisis de desequilibrio alélico telomérico (tAl);  
detectar el número de tAls (NtAl); y  
seleccionar una terapia contra el cáncer que comprende platino para el sujeto cuando el NtAl está por encima de un  
valor de referencia; en el que el valor de referencia se determina a partir del NtAl obtenido de cánceres similares que  
10 se sabe que son resistentes al platino.
2. Procedimiento *in vitro* de acuerdo con la reivindicación 1, que comprende además seleccionar una terapia contra  
el cáncer que no comprende platino para el sujeto cuando se detecta que el NtAl está por debajo del valor de  
referencia.
- 15 3. Procedimiento *in vitro* de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el valor de referencia es al  
menos 20, 22, 24 o 27.
4. Procedimiento *in vitro* de acuerdo con la reivindicación 1 o la reivindicación 2, en el que la muestra biológica  
20 comprende ácido nucleico de una célula cancerosa tomada del sujeto, y en el que el NtAl se detecta en el ácido  
nucleico de la célula cancerosa.
5. Procedimiento *in vitro* de acuerdo con cualquiera de las reivindicaciones 1-4, en el que el cáncer se selecciona  
entre cáncer de mama y cáncer de ovario.
- 25 6. Procedimiento *in vitro* de acuerdo con cualquiera de las reivindicaciones 1-5, que comprende además una etapa  
de ensayo para la expresión de ARNm de BRCA1 o el estado de metilación del promotor de BRCA1, y detectar la  
cantidad de expresión de ARNm de BRCA1 o la cantidad de metilación del promotor de BRCA1.
- 30 7. Procedimiento *in vitro* de acuerdo con cualquiera de las reivindicaciones 1-6, en el que el cáncer no tiene una  
mutación en el gen BRCA1 y/o en el gen BRCA2.
8. Procedimiento *in vitro* de acuerdo con cualquiera de las reivindicaciones 1-7, en el que el cáncer es cáncer de  
mama que no tiene una mutación en BRCA1.
- 35 9. Composición que comprende agentes quimioterapéuticos basados en platino para su uso en el tratamiento de un  
cáncer seleccionado del grupo que consiste en cáncer de mama, cáncer de ovario, cáncer de hígado, cáncer de  
esófago, cáncer de pulmón, cáncer de cabeza y cuello, cáncer de próstata, cáncer de colon, cáncer de recto, cáncer  
colorrectal y cáncer de páncreas en un paciente;
- 40 en el que el paciente tiene un número aumentado de desequilibrio alélico telomérico (NtAl) en comparación con un  
valor de referencia, en el que el valor de referencia es un NtAl de al menos 20, 22, 24 o 27;  
y en el que además se determina que el paciente no porta una mutación en BRCA1 y/o BRCA2, o en el que se  
determina adicionalmente que el paciente tiene una cantidad disminuida de ARNm de BRCA1 en la célula cancerosa  
y/o se determina adicionalmente que tiene una metilación aumentada de la región promotora de BRCA1, en  
45 comparación con un control de BRCA1 de tipo natural.



**FIG. 1C**

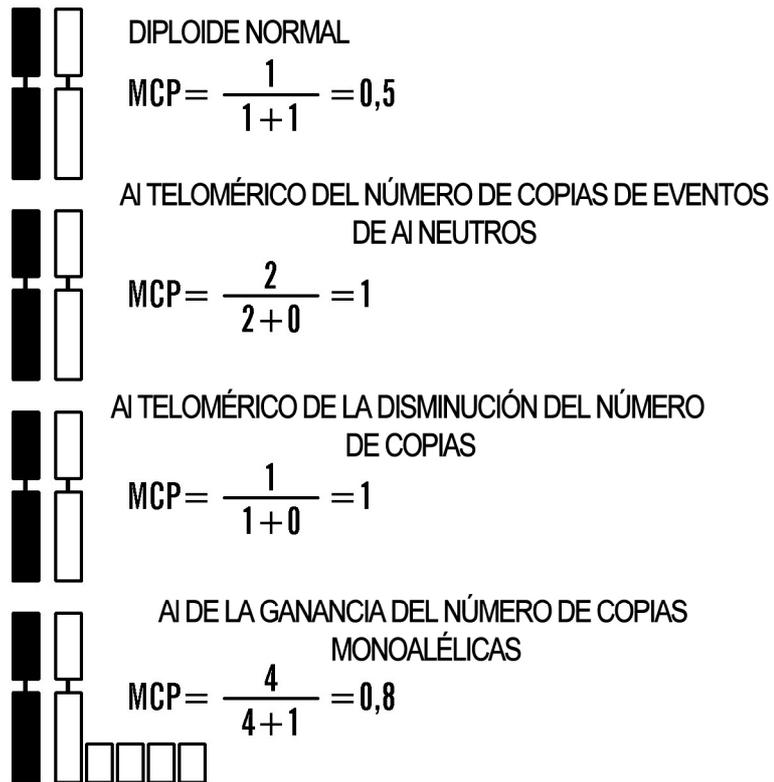


**FIG. 1B**

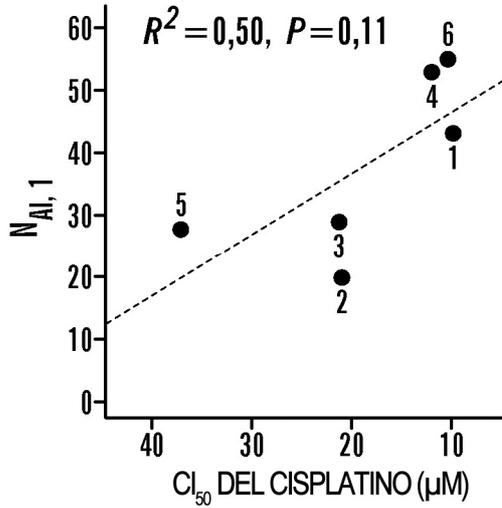


**FIG. 1A**

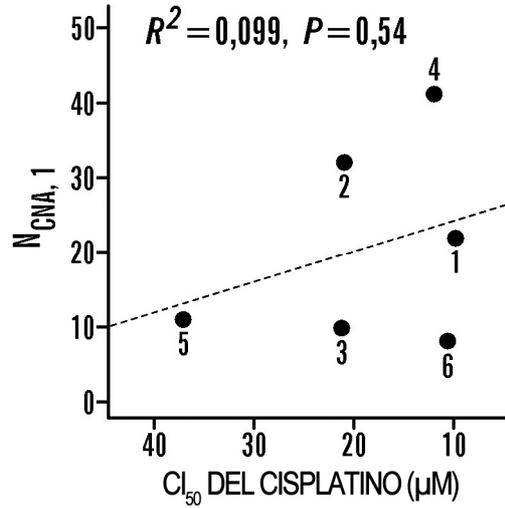
$$\text{MCP} = \frac{\text{NÚMERO DE COPIAS DEL ALELO PRINCIPAL}}{\text{NÚMERO DE COPIAS DEL ALELO PRINCIPAL} + \text{SECUNDARIO}}$$



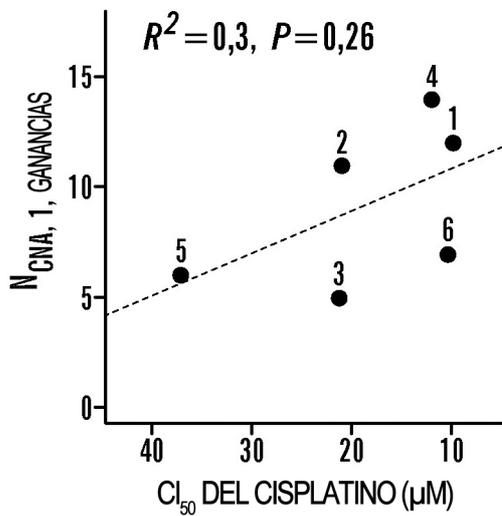
**FIG. 2**



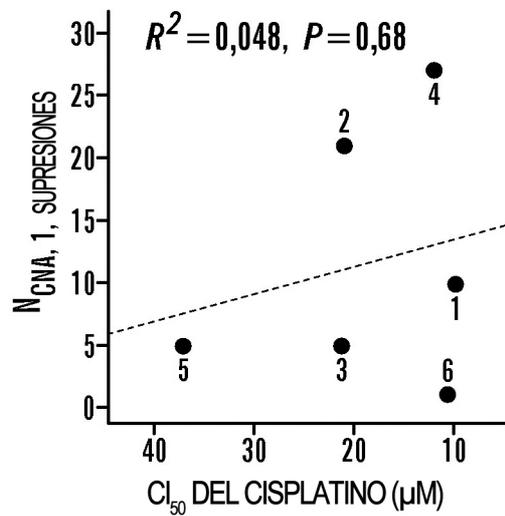
**FIG. 3A**



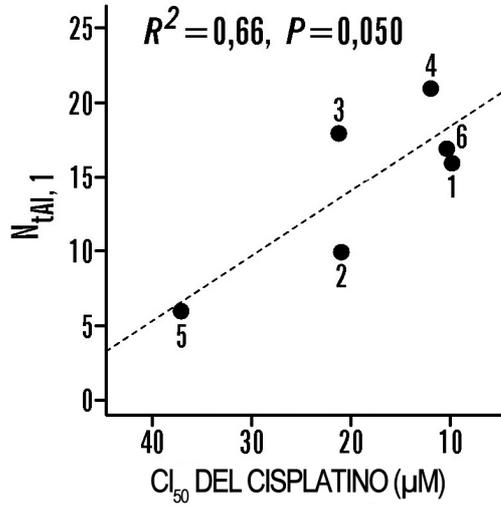
**FIG. 3B**



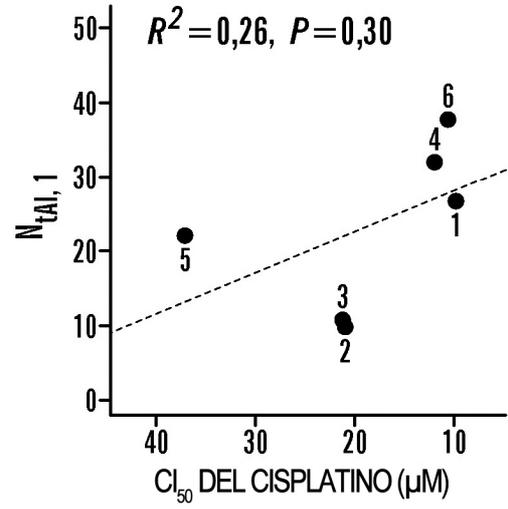
**FIG. 3C**



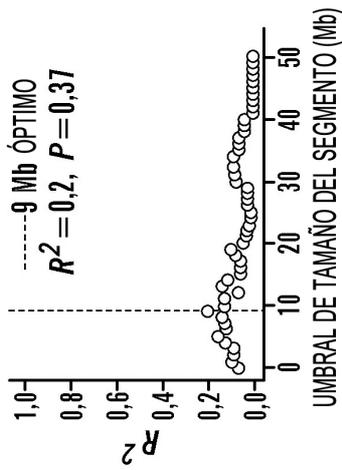
**FIG. 3D**



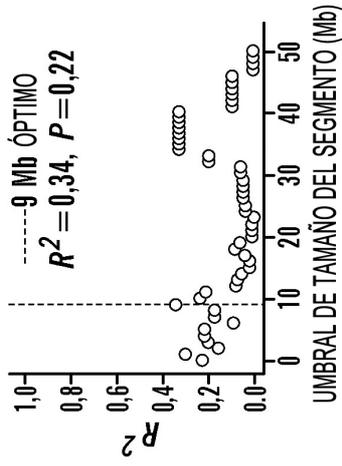
**FIG. 4A**



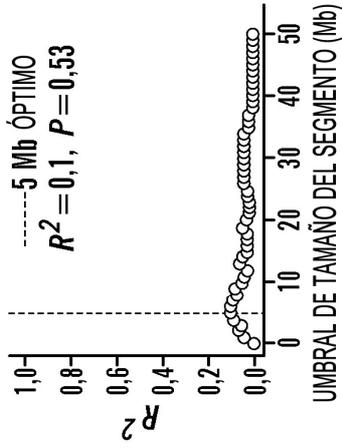
**FIG. 4B**



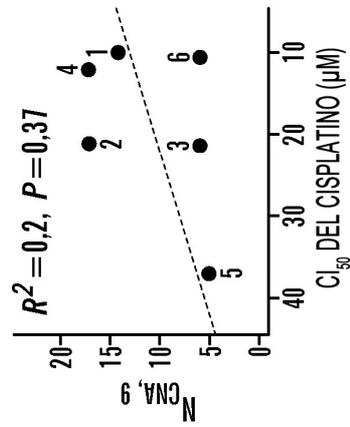
**FIG. 5A**



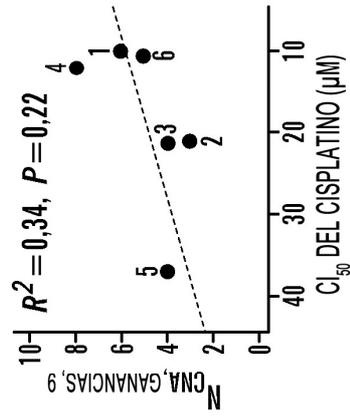
**FIG. 5B**



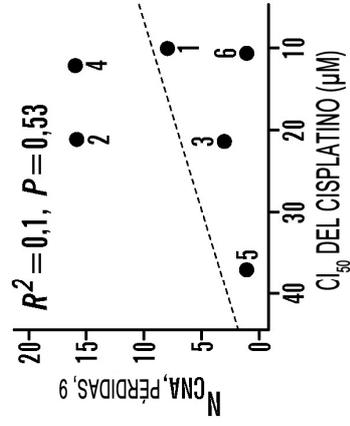
**FIG. 5C**



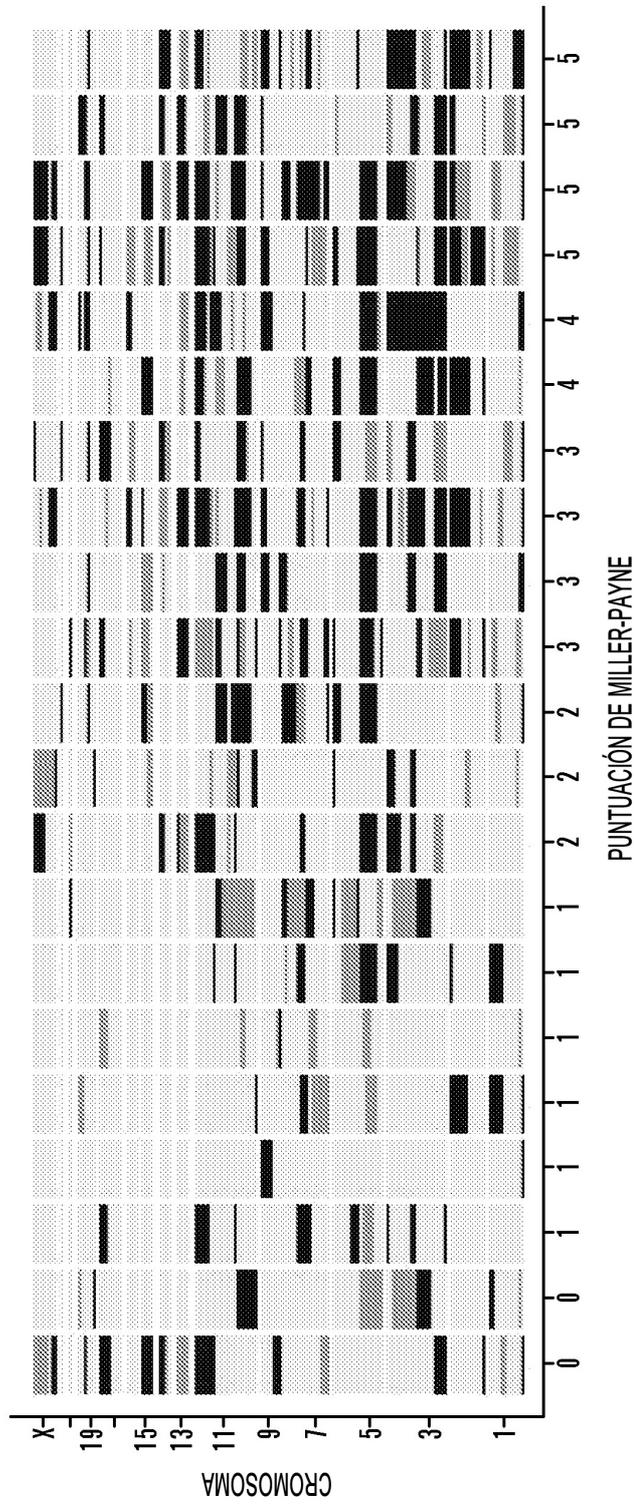
**FIG. 5D**



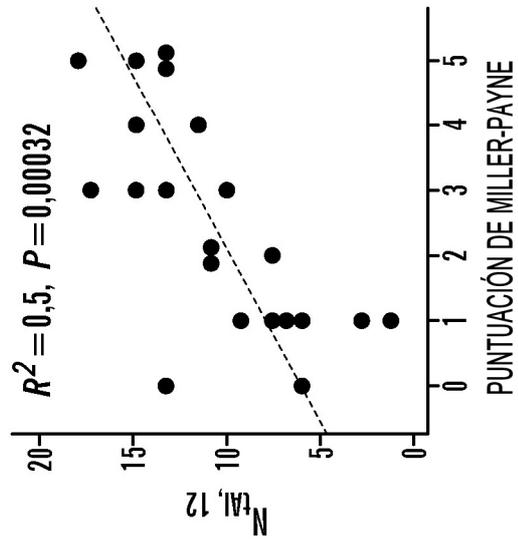
**FIG. 5E**



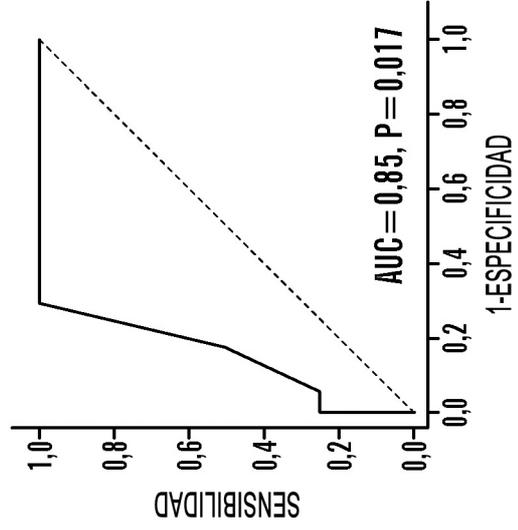
**FIG. 5F**



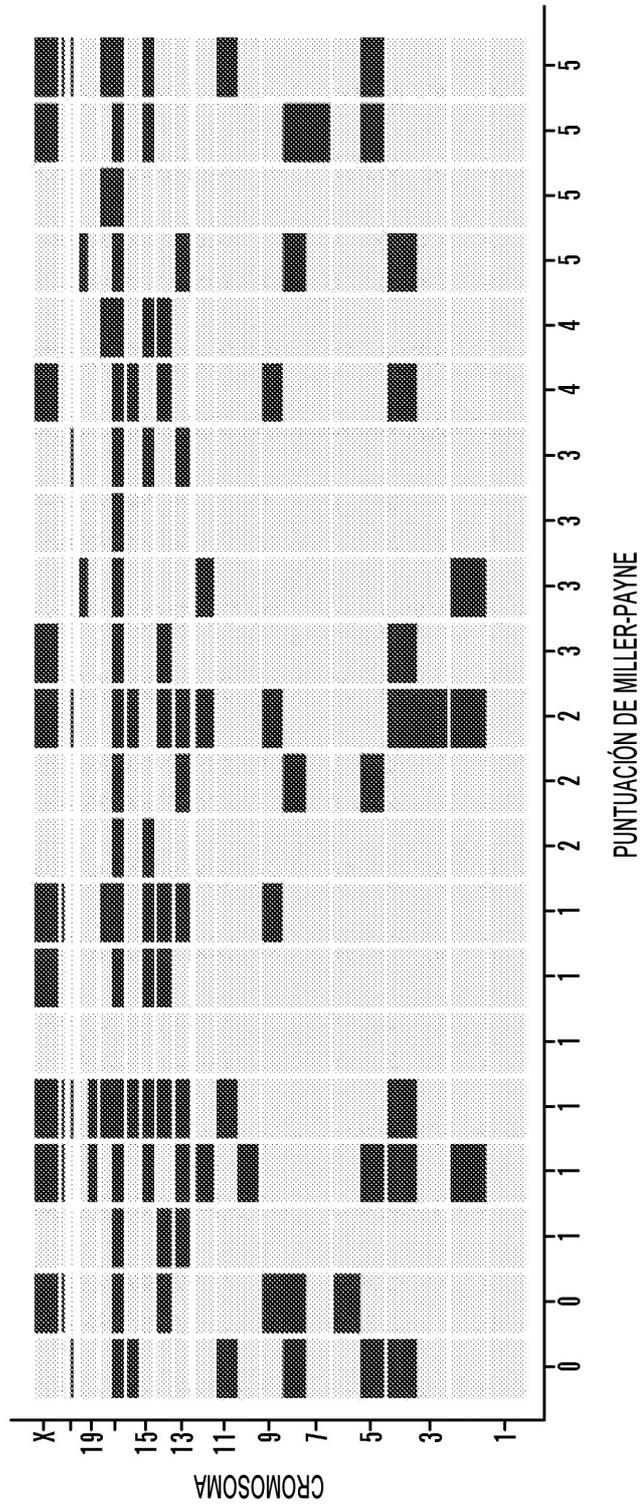
**FIG. 6A**



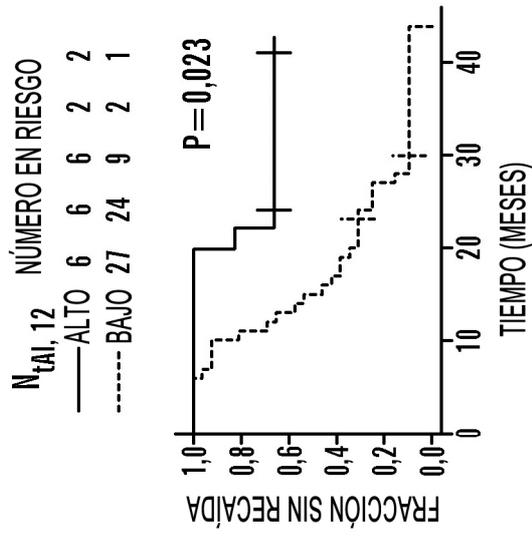
**FIG. 6B**



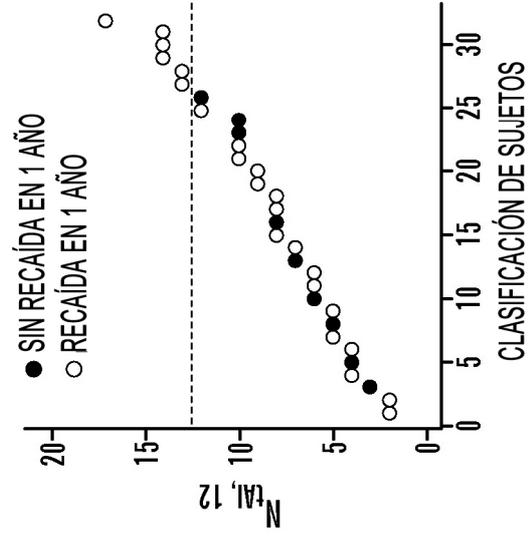
**FIG. 6C**



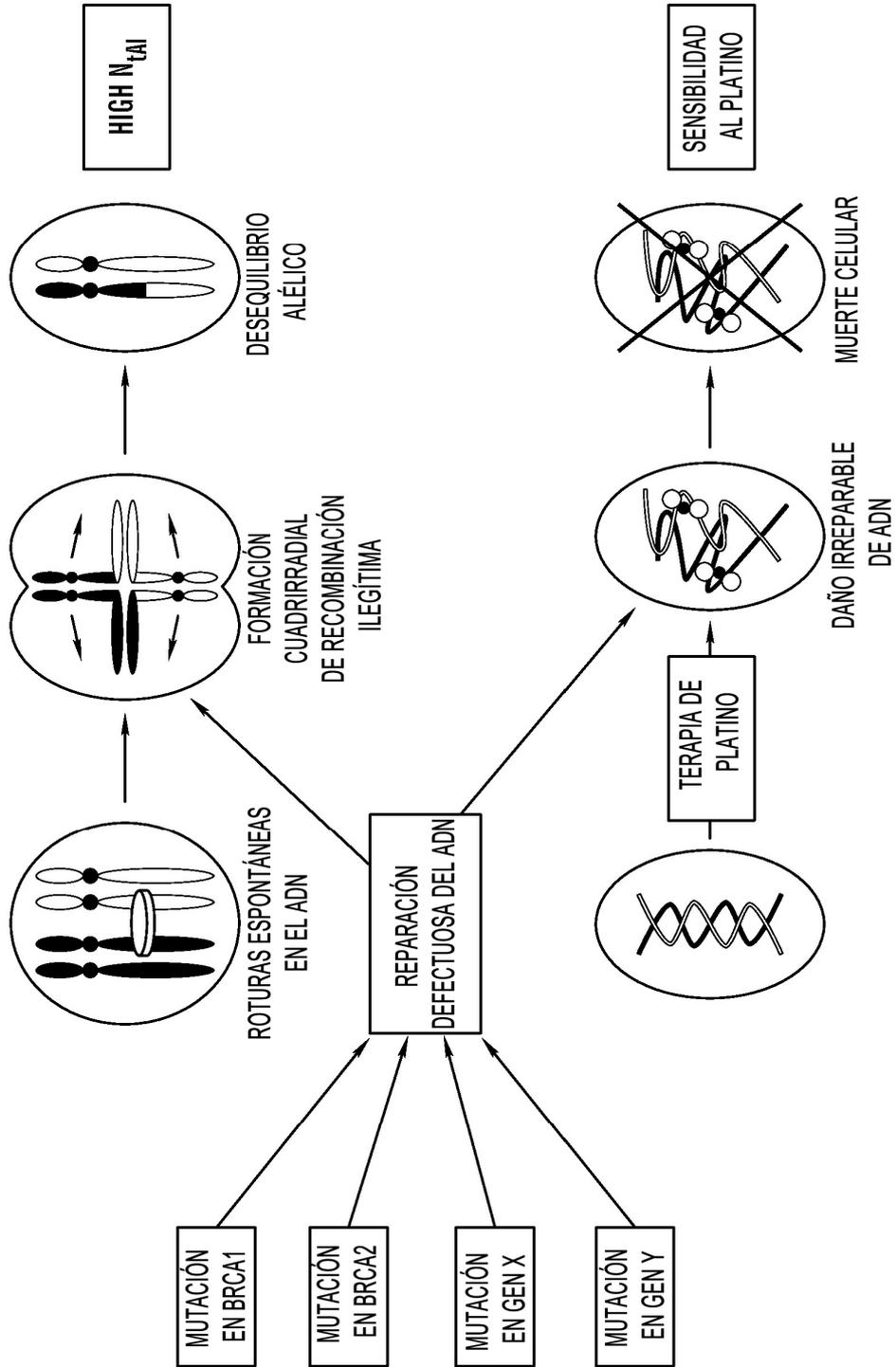
**FIG. 7**



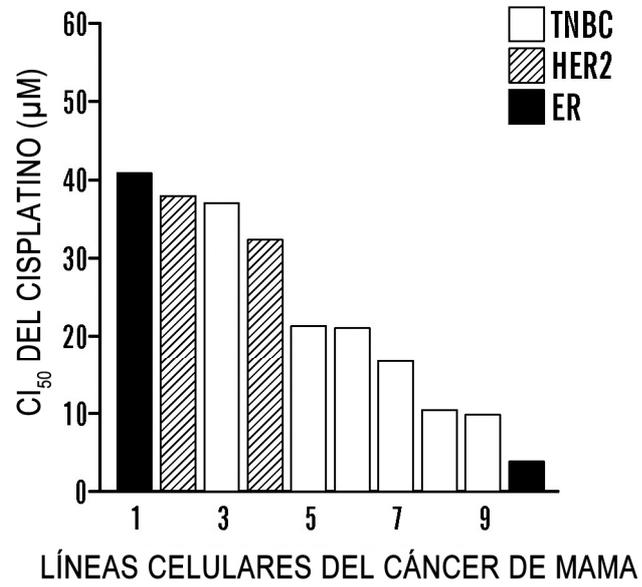
**FIG. 8B**



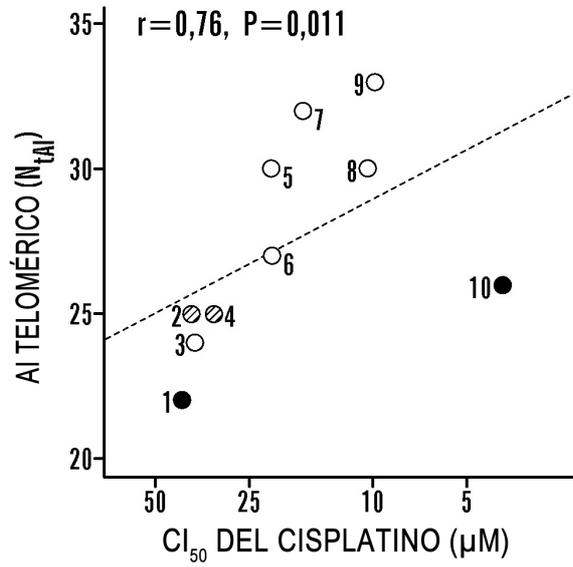
**FIG. 8A**



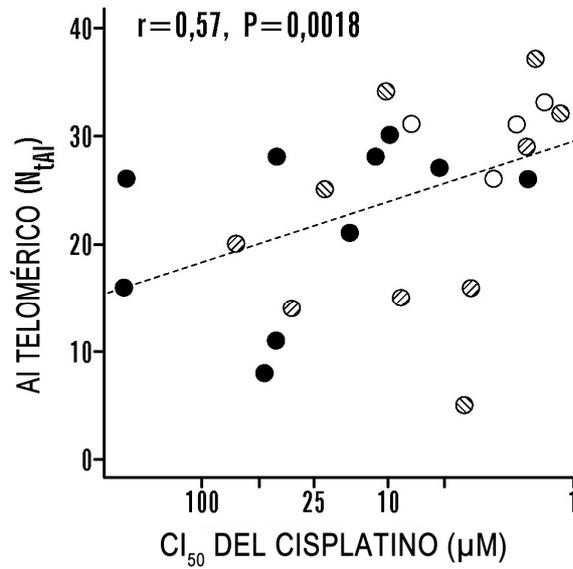
**FIG. 9**



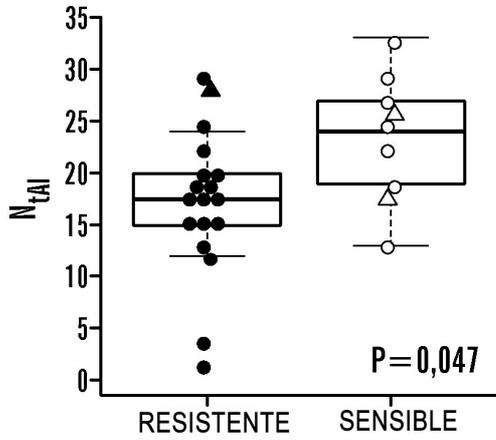
**FIG. 10A**



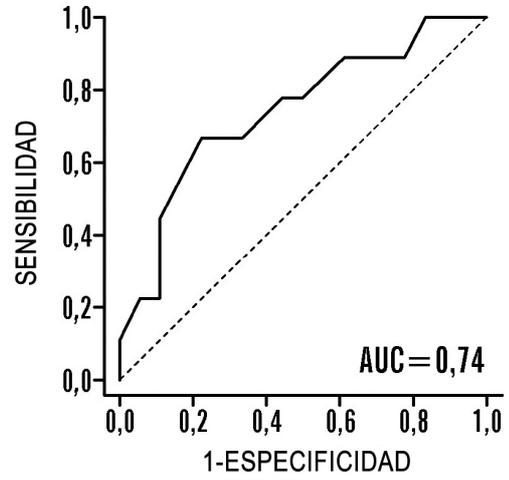
**FIG. 10B**



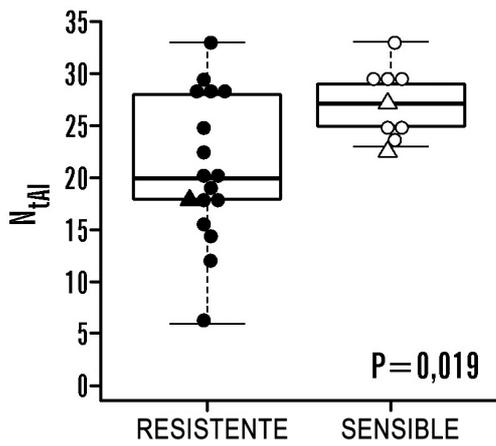
**FIG. 10C**



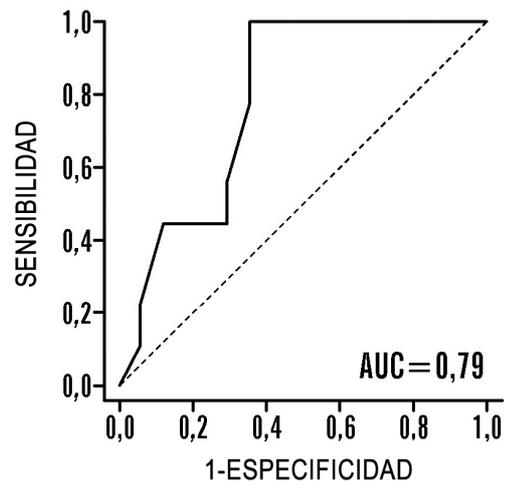
**FIG. 11A**



**FIG. 11B**

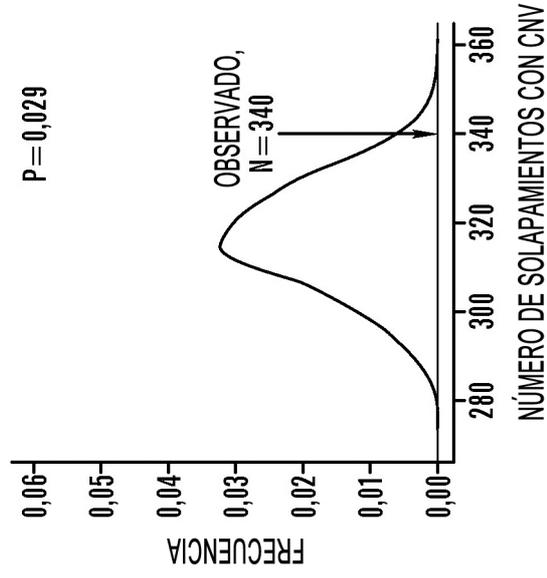


**FIG. 11C**

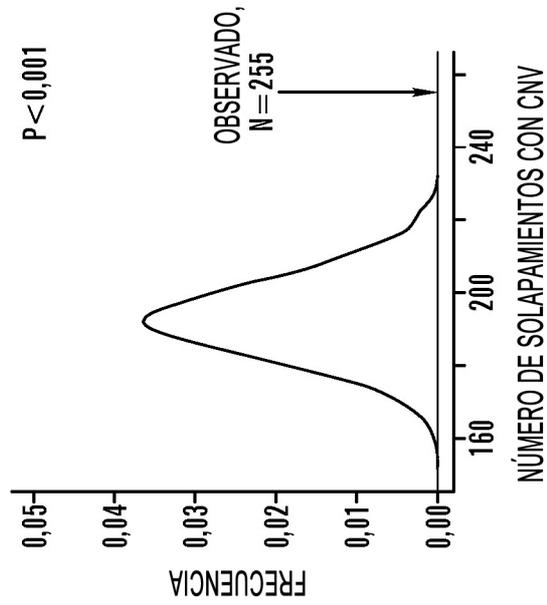


**FIG. 11D**

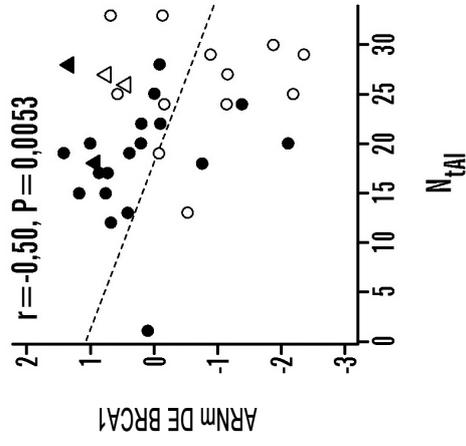




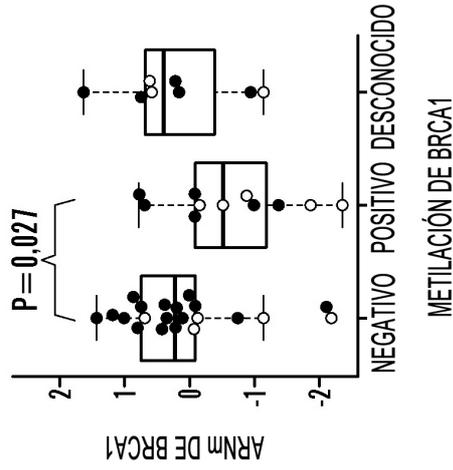
**FIG. 13B**



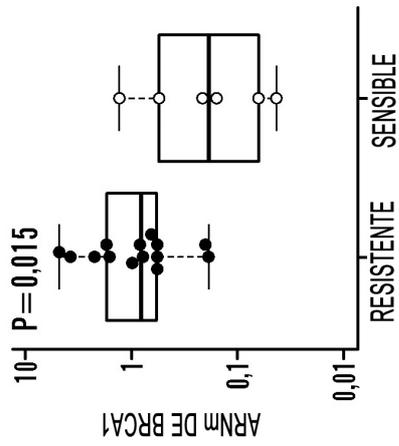
**FIG. 13A**



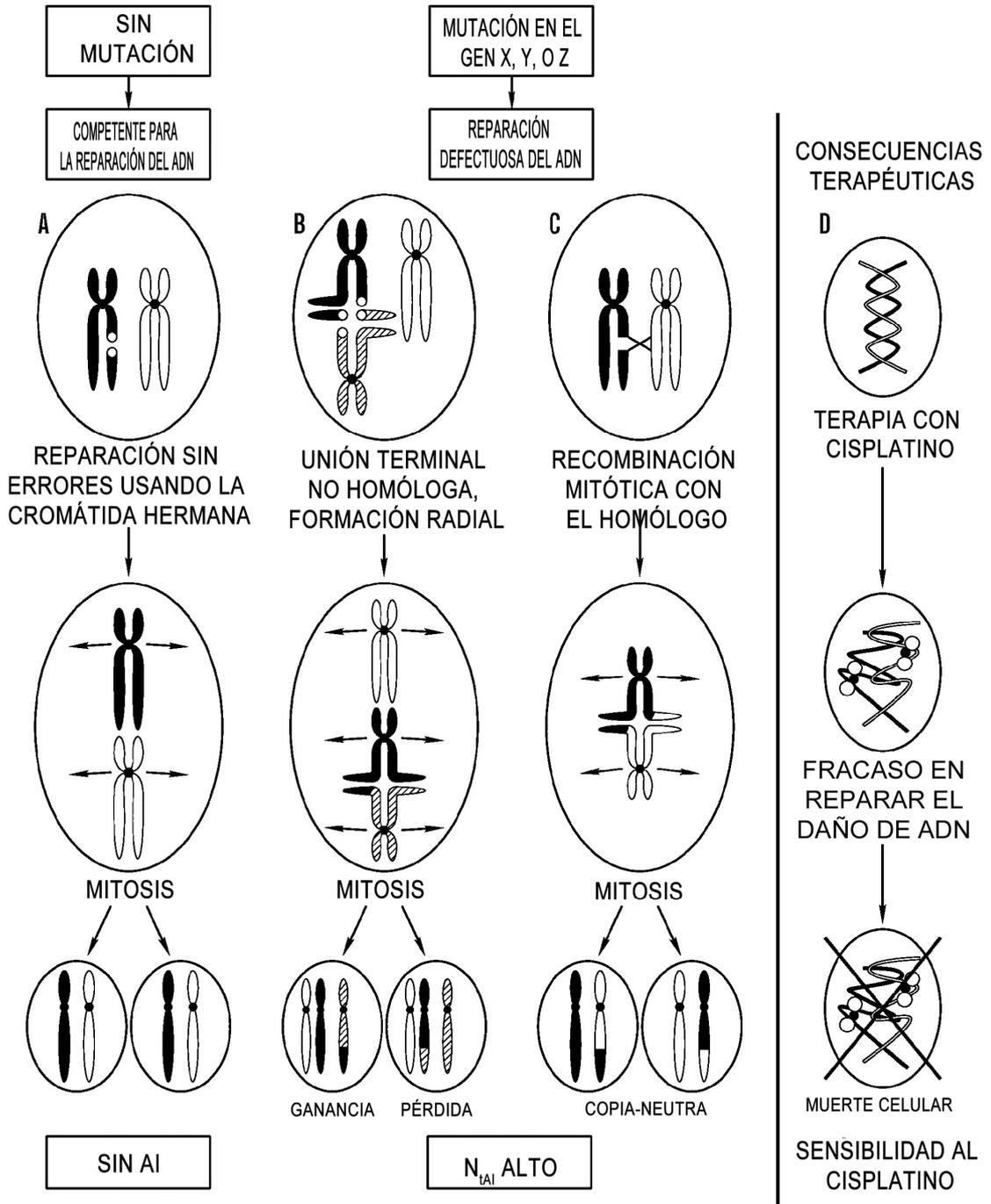
**FIG. 14C**



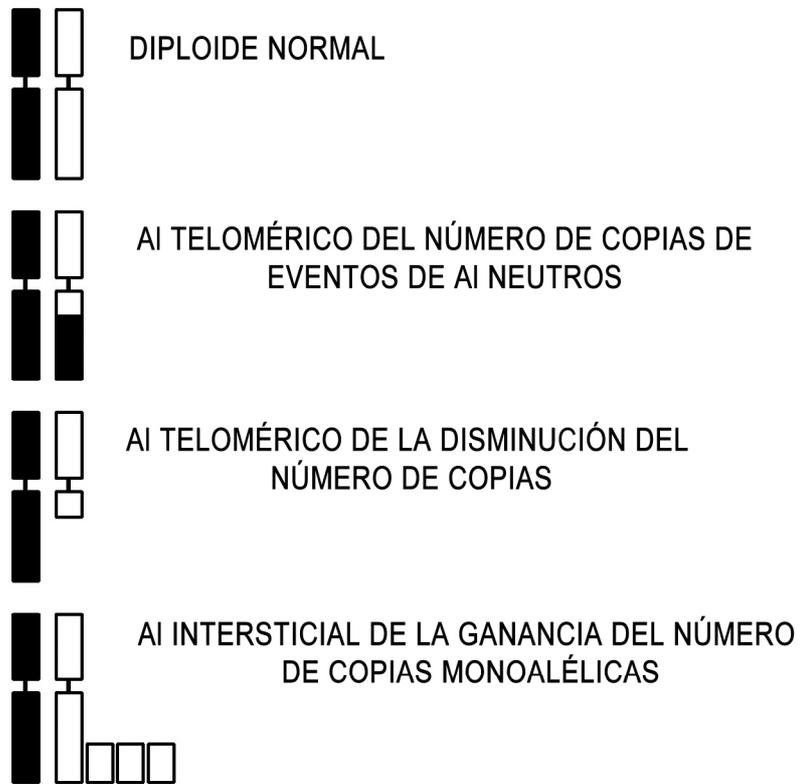
**FIG. 14B**



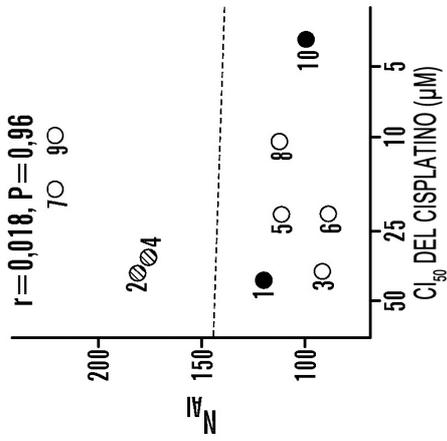
**FIG. 14A**



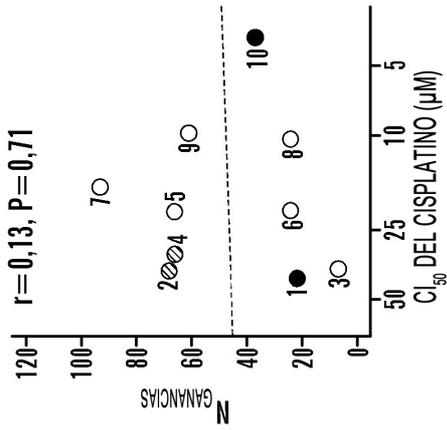
**FIG. 15**



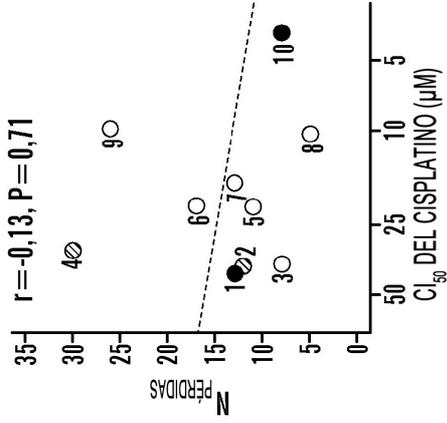
***FIG. 16***



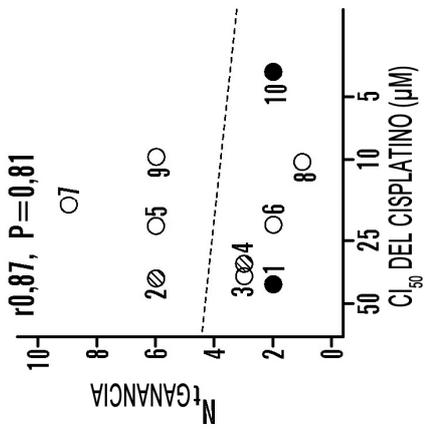
**FIG. 17A**



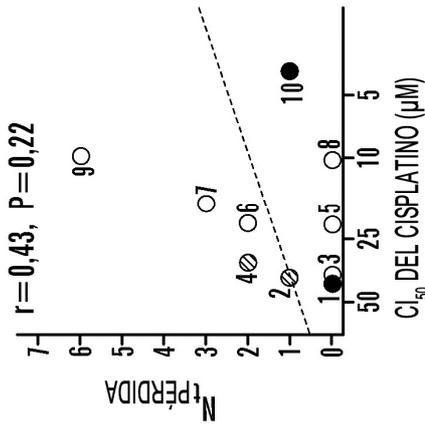
**FIG. 17B**



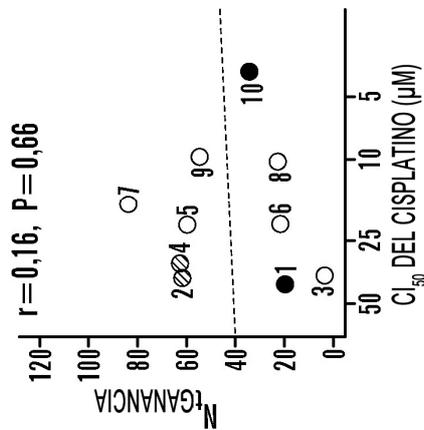
**FIG. 17C**



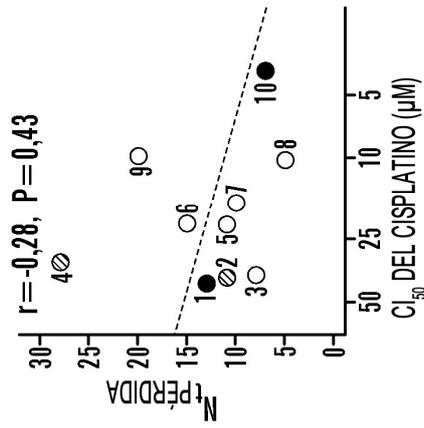
**FIG. 18A**



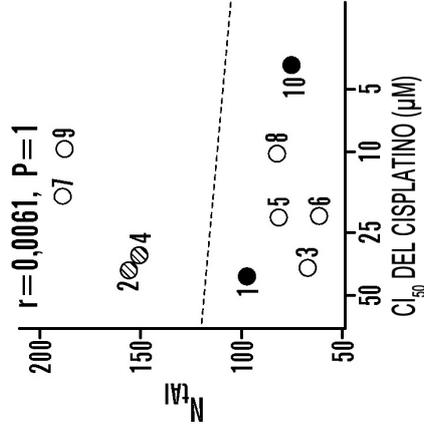
**FIG. 18B**



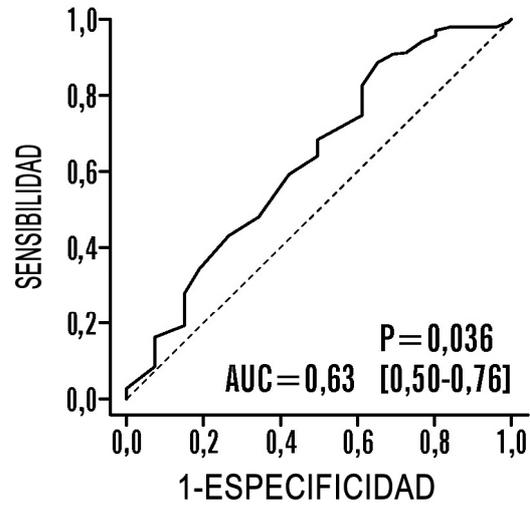
**FIG. 18C**



**FIG. 18D**



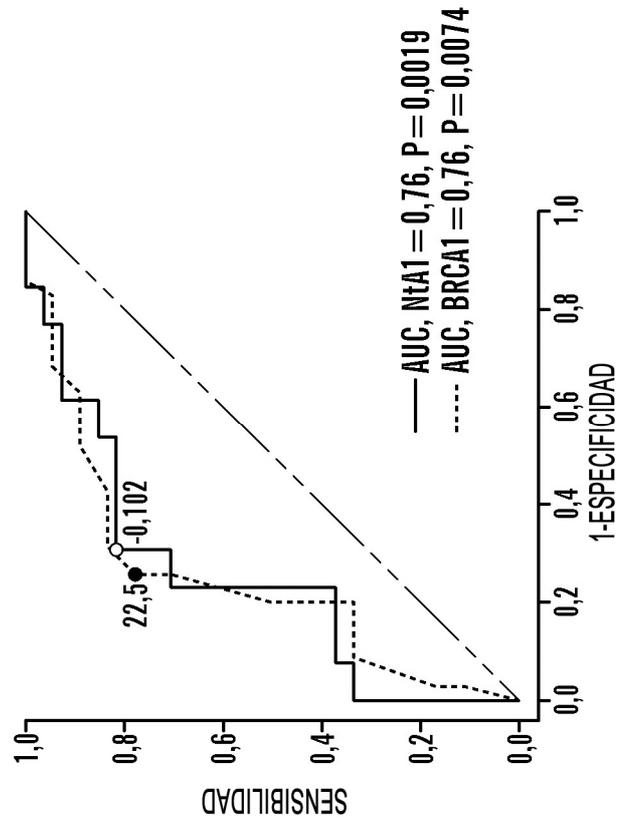
**FIG. 18E**



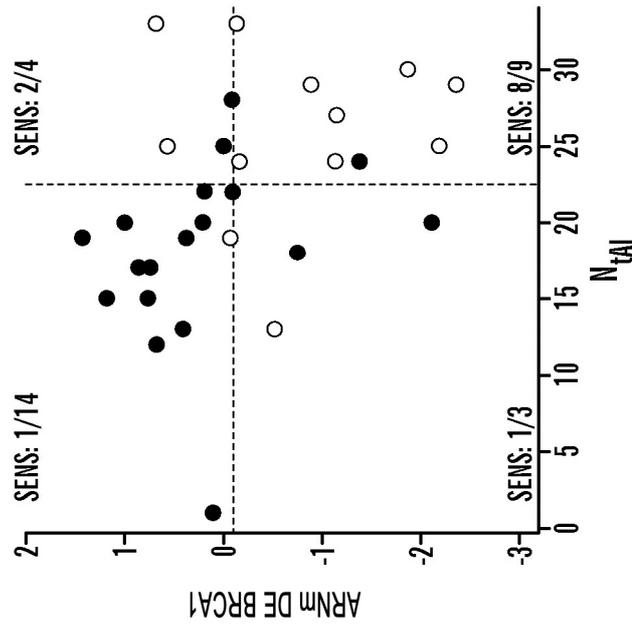
**FIG. 19**



**FIG. 20**

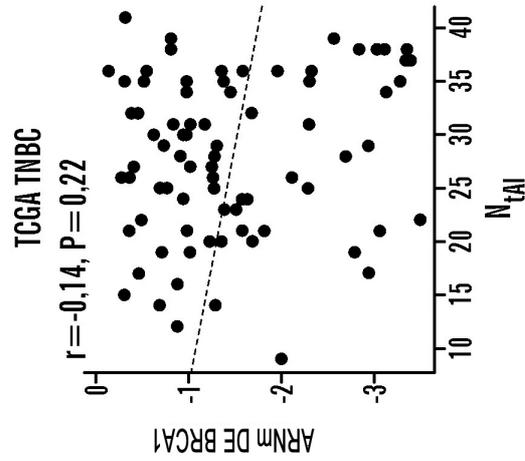


**FIG. 21A**

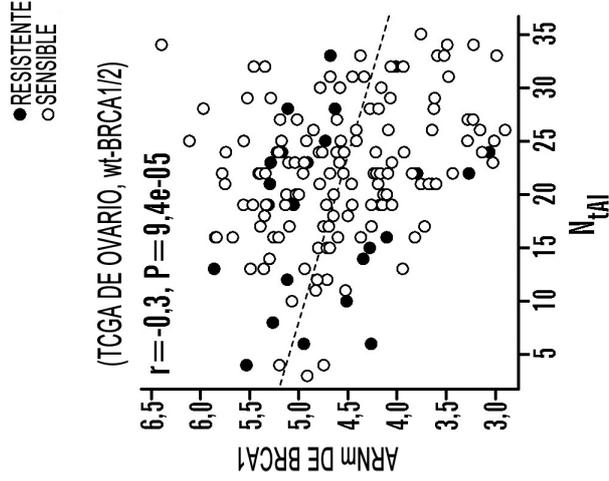


	NTAI	BRCA1	NTAI + BRCA1
ACC	0,83	0,80	0,83
PPV	0,77	0,75	0,89
NPV	0,88	0,83	0,81
SENS	0,83	0,75	0,67
SPEC	0,83	0,83	0,94
P	0,00054	0,0024	0,00064

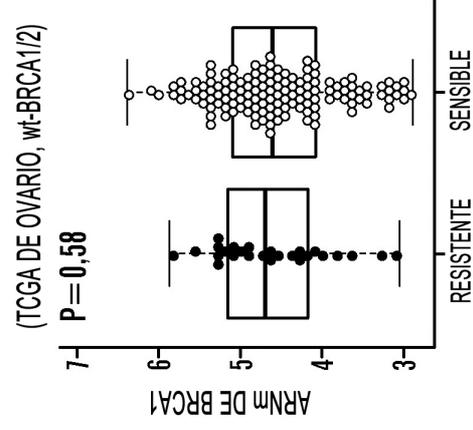
**FIG. 21B**



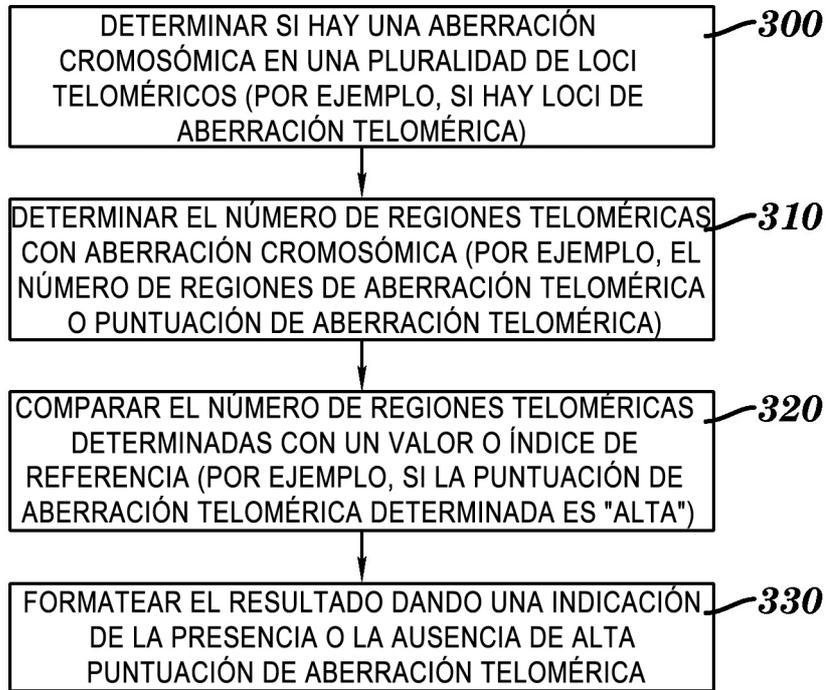
**FIG. 22A**



**FIG. 22B**



**FIG. 22C**



**FIG. 23**

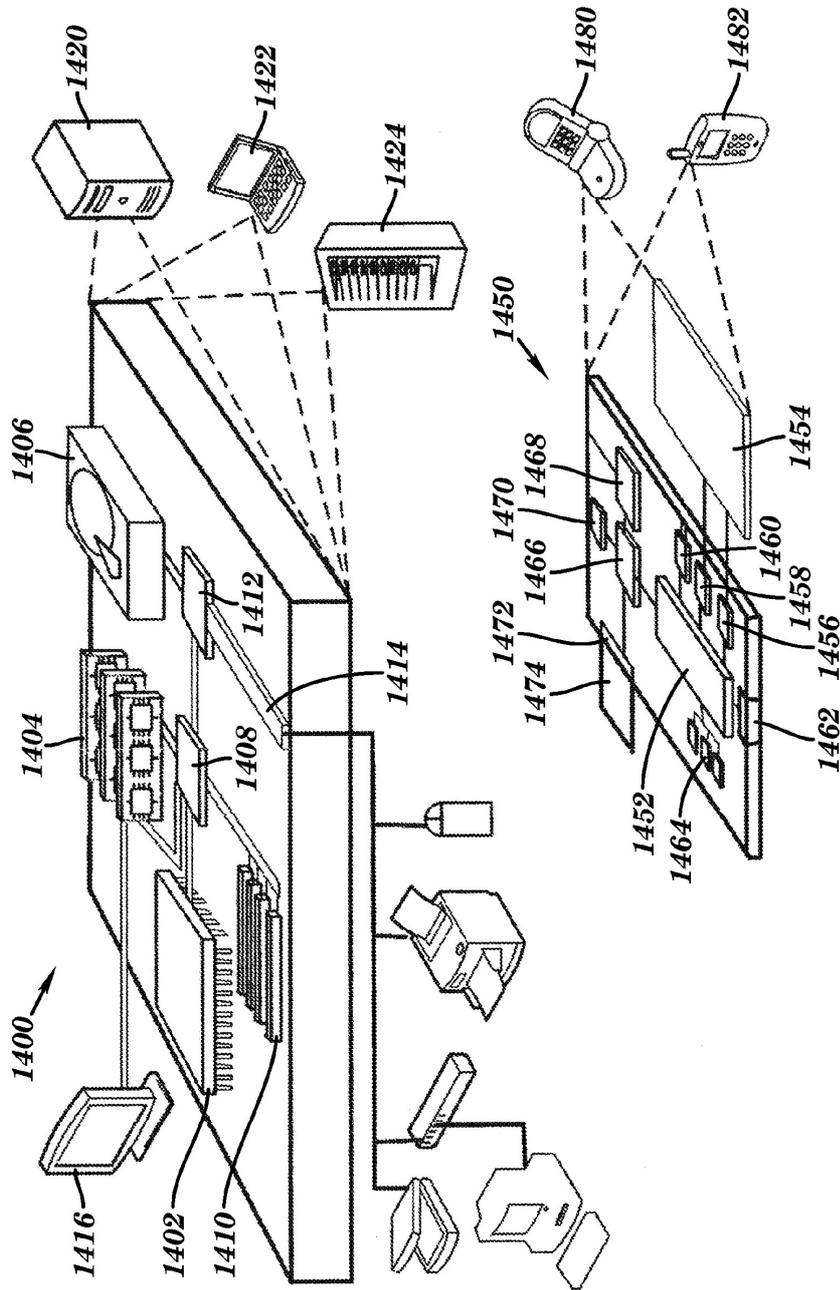


FIG. 24