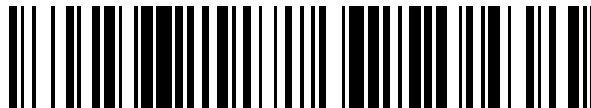


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 405 654**

51 Int. Cl.:

C07H 21/04 (2006.01)

G01N 33/574 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.09.2007 E 07867158 (3)**

97 Fecha y número de publicación de la concesión europea: **20.03.2013 EP 2081950**

54 Título: **Perfiles de expresión asociados con el tratamiento con irinotecan**

30 Prioridad:

21.09.2006 US 846298 P
12.03.2007 US 906438 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
31.05.2013

73 Titular/es:

NUCLEA BIOMARKERS LLC (100.0%)
105 SOUTH STREET
PITTSFIELD, MA 01201, US

72 Inventor/es:

MURACA, PATRICK J.

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 405 654 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

Perfiles de expresión asociados con el tratamiento con irinotecan

5 Solicitudes relacionadas

Esta solicitud reivindica la prioridad bajo el artículo 35 U.S.C. párrafo 119 (e) de la Solicitud provisional de EE.UU. N° de Serie 60/846.298 presentada el 21 de septiembre de 2006 y la solicitud N° de serie 60/906.438 presentada el 12 de marzo de 2007.

10

Antecedentes de la invención

Pacientes a quienes se ha diagnosticado cáncer se enfrentan a opciones de tratamiento costosas y, a menudo, dolorosas. Estos tratamientos pueden no ser eficaces en una sub-población de pacientes y, como resultado de ello, estos pacientes sobrellevan estos tratamientos sin beneficio pequeño alguno o sin ningún beneficio terapéutico. Algunos pacientes pueden reaccionar adversamente a determinados agentes, provocando un sufrimiento adicional y, posiblemente, la muerte.

15

El tratamiento ineficaz es también problemático, dado que el tiempo es una variable clave cuando se trata el cáncer. Un proveedor de tratamiento tiene una posibilidad mucho mayor de contener y controlar la enfermedad si el cáncer es diagnosticado en una fase temprana y es tratado con un agente terapéuticamente eficaz. Un agente puede proporcionar grandes beneficios terapéuticos si se administra en una fase temprana de la enfermedad; sin embargo, con el paso del tiempo, este mismo agente puede dejar de ser eficaz.

20

El cáncer colorrectal es un ejemplo de una afección en la que el diagnóstico temprano es clave para un tratamiento eficaz. El cáncer colorrectal es un cáncer que se desarrolla en el colon o en el recto. Las paredes del colon y del recto tienen varias capas de tejidos. El cáncer colorrectal comienza a menudo en la capa más interna y puede desarrollarse a través de alguna o de la totalidad de las otras capas; la fase (grado de diseminación) de un cáncer colorrectal depende en gran medida de lo profundo que se haya desarrollado en estas capas.

25

30

La quimioterapia se utiliza a menudo para tratar el cáncer colorrectal. Hidrocloruro de irinotecan (CAMPTOSAR®) es un agente quimioterapéutico indicado para la terapia de primera línea de cánceres colorrectales. Al igual que con muchos agentes quimioterapéuticos, la administración de hidrocloruro de irinotecan ("irinotecan") provoca, a menudo, efectos secundarios perjudiciales para el paciente, y algunos pacientes no responden bien al tratamiento. Así, algunos pacientes son sometidos a tratamiento con irinotecan y padecen los efectos secundarios dolorosos únicamente para reconocer posteriormente que el agente no ha sido terapéuticamente beneficioso para su afección. Además de un sufrimiento innecesario, se pierde un tiempo crítico para determinar un tratamiento alternativo.

35

40 Sumario de la invención

La presente descripción proporciona perfiles de expresión de genes y de proteínas y métodos para utilizarlos, para identificar aquellos pacientes que probablemente respondan al tratamiento con irinotecan (a estos pacientes se les alude como "respondedores"), así como aquellos pacientes que no es probable que se beneficien de un tratamiento de este tipo (a estos pacientes se les alude como "no respondedores"). La presente descripción permite que un proveedor de tratamiento identifique a aquellos pacientes que son respondedores al tratamiento con irinotecan, y aquellos que no son respondedores a este tipo de tratamiento, antes de la administración del agente.

45

En un aspecto, la presente descripción comprende perfiles de expresión de genes, a los que se alude también como "firmas de genes", que son indicativos de la tendencia de un paciente que padece cáncer a responder al tratamiento con irinotecan. El perfil de expresión de genes (GEP – siglas en inglés) comprende al menos uno y preferiblemente una pluralidad de genes seleccionados del grupo que consiste en ERBB2, GRB7, quinasa Erk1, quinasa JNK1, BCL2, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, MK167, STK6, MRP14, fosfo-Akt, CD68, BAG1 y GSTM1. La firma de genes puede incluir, además, genes de referencia o control. Los genes de referencia actualmente preferidos son ACTB, GAPD, GUSB, RPLP0 y TFRC. Algunos o la totalidad de estos genes son expresados de modo diferencial (p. ej. supra-regulados o sub-regulados) en pacientes que son respondedores a la terapia con irinotecan. Específicamente, ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1 están supra-regulados (sobre-expresados), y quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5,

55

STK6, MRP14 y GSTM1 están sub-regulados (sub-expresados) en pacientes que son respondedores a irinotecan. Genes de referencia ACTB, GAPD, GUSB, RPLP0 y TFRC están supra-regulados (sobre-expresados).

5 La presente descripción comprende, además, perfiles de expresión de proteínas que son indicativos de la tendencia de un paciente de cáncer a responder al tratamiento con irinotecan. Los perfiles de expresión de proteínas comprenden aquellas proteínas codificadas por los genes del GEP que también están diferencialmente expresadas en cánceres de colon que responden a la terapia con irinotecan. El presente perfil de expresión de proteínas (PEP – siglas en inglés) comprende al menos una y preferiblemente una pluralidad de proteínas codificadas por los genes seleccionados del grupo que consiste en ERBB2, GRB7, quinasa Erk1, quinasa JNK1, BCL2, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, MK167, STK6, MRP14, fosfo-Akt, CD68, BAG1 y GSTM1. El perfil de expresión de proteínas puede incluir, además, proteínas codificadas por genes de referencia. Los genes de referencia actualmente preferidos son ACTB, GAPD, GUSB, RPLP0 y TFRC. Algunas o la totalidad de estas proteínas son expresadas de modo diferencial (p. ej. supra-reguladas o sub-reguladas) en pacientes que son respondedores a la terapia con irinotecan. Específicamente, proteínas codificadas por los siguientes genes están supra-reguladas (sobre-expresadas): ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1; y proteínas codificadas por los siguientes genes están sub-reguladas (sub-expresadas): quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1 en pacientes que son respondedores a irinotecan. Proteínas de referencia ACTB, GAPD, GUSB, RPLP0 y TFRC están supra-reguladas (sobre-expresadas).

20 Los perfiles de expresión de genes y proteínas (a los que se alude en esta memoria como GPEPs) comprenden un grupo de genes y proteínas que están diferencialmente expresados (p. ej. supra-regulados o sub-regulados) en pacientes que son respondedores a una terapia con irinotecan con respecto a la expresión de los mismos genes en pacientes que no son respondedores a esta terapia. Pacientes con tumores que no responden a irinotecan experimentan a menudo una recurrencia de su enfermedad o la muerte relacionada con la enfermedad. Así, los GPEPs se pueden utilizar para predecir no sólo la capacidad de respuesta de un cáncer de colon a una terapia con irinotecan, sino también la probabilidad de recurrencia del cáncer y/o muerte relacionada con la enfermedad.

30 La presente descripción comprende, además, un método para determinar si un paciente es un respondedor o un no respondedor al tratamiento con irinotecan. El método comprende obtener una muestra del tumor del paciente, determinar el perfil de expresión de genes y/o proteínas de la muestra y determinar, a partir del perfil de expresión de genes o proteínas, si al menos uno, preferiblemente al menos 4, más preferiblemente al menos 10 y, lo más preferiblemente al menos 16 de los genes seleccionados del grupo que consiste en ERBB2, GRB7, quinasa Erk1, quinasa JNK1, BCL2, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, MK167, STK6, MRP14, fosfo-Akt, CD68, BAG1 y GSTM1, o al menos una proteína seleccionada de las proteínas codificadas por estos genes, está diferencialmente expresado en la muestra. A partir de esta información, el proveedor de tratamiento puede averiguar si es probable que el paciente se beneficie de la terapia con irinotecan. El presente método también se puede utilizar para predecir una recurrencia tardía y la muerte relacionada con la enfermedad, asociada con la terapia.

40 La presente descripción comprende, además, ensayos para determinar el perfil de expresión de genes y/o proteínas en una muestra de un paciente, e instrucciones para utilizar el ensayo. El ensayo se puede basar en la detección de ácidos nucleicos (p. ej. utilizando sondas de ácidos nucleicos específicas para los ácidos nucleicos de interés) o proteínas o péptidos (p. ej. utilizando anticuerpos específicos para las proteínas/péptidos de interés). En un ejemplo actualmente preferido, el ensayo comprende un ensayo de inmunohistoquímica (IHC – siglas en inglés) en el que muestras de tejidos, preferiblemente dispuestas en una microdisposición ordenada de tejidos (TMA – siglas en inglés) se ponen en contacto con soportes específicos para las proteínas/péptidos identificados en el GPEP como indicativos de una capacidad de respuesta del paciente a irinotecan. Ensayos que comprenden marcadores se describen en el documento US 2006/0166231. El alcance de protección se define por las reivindicaciones.

50 Breve descripción de los dibujos

La FIG. 1 es una gráfica que muestra las tasas de supervivencia para pacientes tratados con irinotecan HCl, correlacionadas con el presente perfil de expresión de genes que predice la capacidad de respuesta a la terapia con irinotecan.

55 Descripción detallada

- La presente descripción proporciona perfiles de expresión de genes y proteínas y su uso para predecir la capacidad de respuesta de un paciente a un tratamiento de cáncer. Más específicamente, los presentes GPEPs son indicativos de si un paciente es un respondedor o un no respondedor al tratamiento con irinotecan. Es probable que aquellos pacientes identificados como respondedores que utilicen el presente GPEP se beneficien de la terapia con irinotecan, mientras que aquellos pacientes identificados como no respondedores pueden evitar un tratamiento innecesario con irinotecan y considerar otras opciones de tratamiento de una manera oportuna. Los presentes GPEPs también se pueden utilizar para predecir la probabilidad de recurrencia de cáncer de colon y muerte relacionada con la enfermedad asociada con la terapia con irinotecan en algunos pacientes.
- Irinotecan es un agente quimioterapéutico que pertenece al grupo de medicinas denominado antineoplásicos. Está indicado, como terapia de primera línea, para tratar cánceres del colon o del recto. Irinotecan interfiere en el crecimiento de células cancerígenas que finalmente son destruidas. Debido a que el crecimiento de células normales puede verse también afectado por la medicina, también se pueden producir otros efectos. Estos otros efectos pueden incluir: sudoración y producción de saliva incrementadas, diarrea, náuseas (encontrarse mal) y vómitos, pérdida del apetito, resistencia reducida a la infección, magulladuras o sangría, anemia, pérdida del pelo, cansancio y una sensación general de debilidad. La presente descripción permite al proveedor de tratamiento determinar de antemano aquellos pacientes que probablemente se beneficien del tratamiento con irinotecan y considerar opciones de tratamiento alternativas para no respondedores. Se entiende que el tratamiento con irinotecan incluye administrar irinotecan solo y en combinación con otros agentes o adyuvantes terapéuticos. Las presentes indicaciones para CAMPTOSAR® incluyen administrar irinotecan HCl en combinación con 5-fluorouracilo (5-FU) y leucovorina (LV) como terapia de primera línea para el cáncer colorrectal metastásico, y solo, como una terapia de segunda línea, para pacientes cuya enfermedad haya vuelto o progresado después de una terapia inicial con 5-FU.
- Los genes que comprenden el presente GEP incluyen: ERBB2, GRB7, quinasa Erk1, quinasa JNK1, BCL2, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, MK167, STK6, MRP14, fosfo-Akt, CD68, BAG1 y GSTM1. En un ejemplo preferido, el presente perfil de expresión de genes incluye, además, los siguientes genes de referencia: ACTB, GAPD, GUSB, RPLP0 y TFRC. El número de acceso NCBI de una variante de cada uno de estos genes se recoge en la Tabla 1; existen otras variantes que pueden fácilmente ser verificadas haciendo referencia a una base de datos apropiada tal como NCBI Entrez (www.ncbi.nlm.nih.gov/qquery). Estos genes están supra- o sub-regulados en pacientes que responden a una terapia con irinotecan y no en pacientes que experimentan una recurrencia tardía de su enfermedad o muerte relacionada con la enfermedad, asociada con la terapia. Por consiguiente, es posible determinar de antemano si es probable que un paciente se beneficie de una terapia con irinotecan para obtener un perfil de expresión de genes a partir del tejido del paciente, y determinar si uno o más de los genes en el presente GEP está supra- o sub-regulado. La Tabla 1 identifica los genes e indica si estos genes están supra- o sub-regulados en pacientes que son respondedores a una terapia con irinotecan.

Tabla 1

	NOMBRE DEL GEN	NOMBRE ALT DEL GEN	SUPRA-o SUB-REGULACIÓN	Nº ACCESO NCBI	SEQ ID Nº
Racimo de Proliferación	ERBB2	HER2	Supra	NM_004448	1
	GRB7		Supra	NM_005310	2
Racimo de Invasión	quinasa Erk 1		Sub	X60188	3
	quinasa JNK1		Supra	NM_002750	4
	BCL2		Supra	NM_000633	5
Racimo de Expresión de ER	GSK-3-beta		Sub	NM_002093	6
	MMP11	STMY3; estromolisina 3	Sub	NM_005940	7
	CTSL2	catepsina L2	Sub	NM_001333	8
	CCNB1	ciclina B1	Sub	NM_031966	9
Racimo de Proliferación	BIRC5	SURV; survivina	Sub	NM_001168	10
	MK167	antígeno Ki-67	Supra	NM_002417	11

STK6	STK15; BTAK	Sub	NM_003600	12
Akt (Ser473)		Supra	NM_005163	13
CD68		Supra	NM_001251	14
BAG1		Supra	NM_004323	15
GSTM1	glutación-s-transferasa M1	Sub	NM_146421	16
MRP14	proteína de unión a calcio S100 A9	Sub	NM_002965	17
ACTB	β - actina	Supra	NM_001101	18
GAPD	GAPDH	Supra	NM_002046	19
GUSB	GUS	Supra	NM_000181	20
RPLP0		Supra	NM_001002	21
TFRC		Supra	NM_003234	22
Genes de Referencia				
Otros Genes				

En un aspecto preferido, el perfil de genes de la presente invención comprende al menos cuatro, preferiblemente entre cuatro y diez, más preferiblemente al menos diez y, lo más preferiblemente, al menos dieciséis de los genes en el presente GEP, supra- o sub-regulados según sea aplicable, junto con uno o más genes de referencia.

- 5 Los perfiles de expresión de genes se pueden utilizar para predecir la capacidad de respuesta de un paciente de cáncer de colon a una terapia con irinotecan. En un aspecto, el presente método comprende (a) obtener un perfil de expresión de genes a partir de una muestra biológica de un paciente que padece cáncer de colon; (b) determinar a partir del perfil de expresión de genes si uno o más de los siguientes genes están supra-regulados (sobre-expresados): ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1; y/o si al menos uno de los siguientes genes está sub-regulado (sub-expresado): quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1. Preferiblemente, también se mide la expresión de al menos dos genes de referencia. El valor predictivo del perfil de genes para determinar la respuesta a irinotecan aumenta con el número de estos genes que se encuentran supra- o sub-regulados. Preferiblemente, al menos aproximadamente cuatro, más preferiblemente al menos aproximadamente diez y lo más preferiblemente al menos aproximadamente dieciséis de los genes en el presente GPEP están expresados diferencialmente.

La presente descripción comprende, además, perfiles de expresión de proteínas que son indicativos de la tendencia de un paciente de cáncer a responder al tratamiento con irinotecan. Los perfiles de expresión de proteínas comprenden al menos una y preferiblemente una pluralidad de proteínas codificadas por los genes seleccionados del grupo que consiste en ERBB2, GRB7, quinasa Erk1, quinasa JNK1, BCL2, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, MK167, STK6, MRP14, fosfo-Akt, CD68, BAG1 y GSTM1. Algunas o la totalidad de estas proteínas son expresadas de modo diferencial (p. ej. supra-reguladas o sub-reguladas) en pacientes que son respondedores a la terapia con irinotecan. Específicamente, las proteínas codificadas por los siguientes genes están supra-reguladas (sobre-expresadas): ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1, y las proteínas codificadas por los siguientes genes están sub-reguladas (sub-expresadas): quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1 en pacientes que son respondedores a irinotecan. Se pueden incluir los siguientes genes de referencia: ACTB, GAPD, GUSB, RPLP0 y TFRC.

La Tabla 2 lista los genes en el presente GPEP y una variante de una proteína codificada con ello. La Tabla 2 indica también si la expresión de la proteína está supra- o sub-regulada en pacientes que responden a una terapia con irinotecan. La Tabla 2 incluye el N° de acceso NCBI de una variante de cada una de las proteínas; existen otras variantes de estas proteínas que pueden ser fácilmente averiguadas haciendo referencia a una base de datos apropiada tal como NCBI Entrez (www.ncbi.nlm.nih.gov/gquery). Nombres alternativos para las proteínas listadas en la Tabla 2 también se pueden determinar a partir del sitio de NCBI.

35

Tabla 2

NOMBRE DEL GEN	NOMBRE(S) DE LAS PROTEÍNAS	SUPRA- O SUB-REGULACIÓN de la proteína	N° de Acceso NCBI de la proteína	SEQ ID N° de la proteína
ERBB2	ErbB-2; HER-2	Supra	NP_004449	23
GRB7	GRB7; proteína 7 unida al receptor del factor de crecimiento	Supra	NP_005301	24
quinasa Erk1	quinasa Erk1; proteína quinasa activada por mitógeno 3	Sub	P 27361	25
quinasa JNK1	quinasa JNK1	Supra	NP_002741	26
BCL2	Bcl-2; proteína 2 del linfoma de células B	Supra	NP_000624	27
GSK-3-beta	Fosfo-GSK-3-beta; glucógeno sintasa quinasa 3 beta	Sub	NP_002084	28
MMP11	STMY3; estromolisina 3; metaloproteinasa de matriz 11	Sub	NP_005931	29
CTSL2	catepsina L2	Sub	NP_001324	30

CCNB1	ciclina B1	Sub	NP_114172	31
BIRC5	BIRC5; survivina	Sub	NP_001159	32
MK167	antígeno	Supra	NP_002408	33
STK6	STK15; BTAK; aurora-A	Sub	NP_003591	34
Akt	Fosfo-akt; oncogen viral del timoma murino v-akt; proteína quinasa RAC-alfa	Supra	NP_005154	35
CD68	Antígeno CD68; macrosialina	Supra	NP_001242	36
BAG1	atanogen asociado a Bcl-2	Supra	NP_004314	37
GSTM1	glutation-s- transferasa M1	Sub	NP_666533	38
ACTB	β -actina	Supra	NP_001092	39
GAPD	GAPD	Supra	NP_002037	40
GUSB	GUS; glucuronidasa beta	Supra	NP_000172	41
RPLP0	proteína ribosomal P0	Supra	NP_000993	42
TFRC	receptor de transferrina	Supra	NP_003225	43
MRP14	proteína de unión a calcio S100 A9	Sub	NP_002956	44

Definiciones:

5 Por conveniencia, se proporciona en lo que sigue el significado de determinados términos y frases empleados en la memoria descriptiva, ejemplos y reivindicaciones anejas. Las definiciones sirven para proporcionar una comprensión más clara de determinados aspectos de la presente invención.

10 El término "genoma" pretende incluir el complemento de ADN completo de un organismo, incluido el componente de ADN nuclear, ADN cromosómico o extracromosómico, así como el dominio citoplásmico (p. ej. ADN mitocondrial).

15 El término "gen" se refiere a una secuencia de ácidos nucleicos que comprende secuencias de control y codificadoras, necesarias para producir un polipéptido o precursor. El polipéptido puede ser codificado por una secuencia codificadora de longitud completa o por cualquier porción de la secuencia codificadora. El gen se puede derivar, en su totalidad o en parte, de cualquier fuente conocida en la técnica, incluida una planta, un hongo, un animal, un genoma bacteriano o episoma, ADN eucariótico, nuclear o de plásmido, ADNc, ADN viral o ADN químicamente sintetizado. Un gen puede contener una o más modificaciones en cualquiera de las regiones codificadoras o las no traducidas que podrían afectar a la actividad biológica o a la estructura química del producto de expresión, a la tasa de expresión o a la manera de controlar la expresión. Modificaciones de este tipo incluyen, pero no se limitan a mutaciones, inserciones, deleciones y sustituciones de uno o más nucleótidos. El gen puede constituir una secuencia codificadora ininterrumpida, o puede incluir uno o más intrones, unidos por las uniones de corte y empalme apropiadas. El término "gen", tal como se utiliza en esta memoria, incluye variantes de los genes identificados en la Tabla 1.

25 La expresión "expresión de genes" se refiere al proceso mediante el cual una secuencia de ácidos nucleicos sufre una transcripción y traducción con éxito de modo que se expresan niveles detectables de la secuencia de nucleótidos.

30 Las expresiones "perfil de expresión de genes" o "firma de genes" se refieren a un grupo de genes expresados por un tipo de célula o tejido particular, en donde la presencia de los genes considerados juntos o la expresión diferencial de dichos genes, es indicativa/predictiva de una determinada afección.

La expresión “ácido nucleico”, tal como se utiliza en esta memoria, se refiere a una molécula constituida por uno o más nucleótidos, es decir, ribonucleótidos, desoxirribonucleótidos, o ambos. La expresión incluye monómeros y polímeros de ribonucleótidos y desoxirribonucleótidos, estando los ribonucleótidos y/o desoxirribonucleótidos unidos entre sí, en el caso de los polímeros, a través de enlaces 5' a 3'. Los polímeros de ribonucleótidos y desoxirribonucleótidos pueden ser de cadena sencilla o doble. Sin embargo, los enlaces pueden incluir cualquiera de los enlaces conocidos en la técnica, incluidos, por ejemplo, ácidos nucleicos que comprenden enlaces 5' a 3'. Los nucleótidos pueden presentarse en la naturaleza o pueden ser análogos producidos de forma sintética que son capaces de formar relaciones de pares de bases con pares de bases que se presentan en la naturaleza. Ejemplos de bases que se presentan de forma no natural, que son capaces de formar relaciones de apareamiento de bases, incluyen, pero no se limitan a análogos de aza y deaza pirimidina, análogos de aza y deaza purina y otros análogos de bases heterocíclicas, en donde uno o más de los átomos de carbono y nitrógeno de los anillos de pirimidina han sido sustituidos por heteroátomos, p. ej., oxígeno, azufre, selenio, fósforo y similares. Además, la expresión “secuencias de ácidos nucleicos” contempla la secuencia complementaria e incluye específicamente cualquier secuencia de ácidos nucleicos que sea esencialmente homóloga tanto a la secuencia de ácidos nucleicos como a su complemento.

Los términos “disposición ordenada” y “microdisposición ordenada” se refieren al tipo de genes o proteínas representado por una disposición ordenada de oligonucleótidos o agentes de captura de proteínas, y en donde el tipo de genes o proteínas representado en la disposición ordenada depende del fin pretendido de la disposición ordenada (p. ej. para vigilar la expresión de genes o proteínas humanos). Los oligonucleótidos o agentes de captura de proteínas en una disposición ordenada dada pueden corresponder al mismo tipo, categoría o grupo de genes o proteínas. Genes o proteínas se pueden considerar del mismo tipo si comparten algunas características comunes tales como la especie de origen (p. ej. ser humano, ratón, rata); estado patológico (p. ej. cáncer); funciones (p. ej. proteína quinasas, supresores de tumores); mismo proceso biológico (p. ej. apoptosis, transducción de señales, regulación, proliferación y diferenciación del ciclo celular). Por ejemplo, un tipo de disposición ordenada puede ser una “disposición ordenada de cáncer” en la que cada uno de los oligonucleótidos o agentes de captura de proteínas de la disposición ordenada corresponde a un gen o proteína asociado con un cáncer. Una “disposición ordenada epitelial” puede ser una disposición ordenada de oligonucleótidos o agentes de captura de proteínas que corresponde a genes o proteínas epiteliales únicas. De manera similar, una “disposición ordenada del ciclo celular” puede ser un tipo de disposición ordenada en la que los oligonucleótidos o agentes de captura de proteínas corresponden a genes o proteínas únicos asociados con el ciclo celular.

La expresión “tipo de célula” se refiere a una célula procedente de una fuente dada (p. ej. un tejido, órgano) o una célula en un estado de diferenciación dado, o una célula asociada con una patología o estructura genética dada.

El término “activación”, tal como se utiliza en esta memoria, se refiere a cualquier alteración de una vía de señalización o respuesta biológica que incluye, por ejemplo, aumentos por encima de los niveles basales, restauración a los niveles basales a partir de un estado inhibido y estimulación de la vía por encima de los niveles basales.

La expresión “expresión diferencial” se refiere tanto a diferencias cuantitativas como cualitativas en los modelos de expresión temporales y de tejidos de un gen o una proteína en tejidos o células enfermos frente a tejido adyacente normal. Por ejemplo, un gen diferencialmente expresado puede tener su expresión activada o inactivada por completo en estados normales frente a estados patológicos, o puede estar supra-regulado (sobre-expresado) o sub-regulado (sub-expresado) en un estado patológico frente a un estado normal. Un gen cualitativamente regulado de este tipo puede exhibir un modelo de expresión dentro de un tejido o tipo de célula dado que sea detectable en estados de control o patológicos, pero que no sea detectable en ambos. Dicho de otra manera, un gen o proteína está diferencialmente expresado cuando la expresión del gen o de la proteína se produce a un nivel superior o inferior en los tejidos o células enfermos de un paciente con respecto al nivel de su expresión en los tejidos o células normales (libres de enfermedades) del paciente y/o tejidos o células control.

El término “detectable” se refiere a un modelo de expresión de ARN que puede ser detectado a través de las técnicas convencionales de la reacción en cadena de la polimerasa (PCR – siglas en inglés), PCR con transcriptasa inversa (RT – siglas en inglés), representación diferencial y análisis Northern, todos los cuales son bien conocidos por los expertos en la técnica. De manera similar, modelos de expresión de proteínas se pueden “detectar” a través de técnicas convencionales tales como transferencias Western.

El término “complementario” se refiere a la compatibilidad o emparejamiento conjunto topológico de las superficies interactuantes de una molécula sonda y su diana. La diana y su sonda se pueden describir como complementarias y, además, las características de la superficie de contacto son complementarias una con otra. La hibridación o el apareamiento de bases entre nucleótidos o ácidos nucleicos tales como, por ejemplo, entre las dos cadenas de una molécula de ADN de doble cadena o entre una sonda de oligonucleótido y una diana son complementarios.

La expresión “muestra biológica” se refiere a una muestra obtenida de un organismo (p. ej. un paciente humano) o de componentes (p. ej. células) de un organismo. La muestra puede ser de cualquier tejido o fluido biológico. La muestra puede ser una “muestra clínica” que es una muestra derivada de un paciente. Muestras de este tipo incluyen, pero no se limitan a esputo, sangre, células de la sangre (p. ej. leucocitos), fluido amniótico, plasma, semen, médula ósea, y muestras de tejido o de biopsia con aguja fina, orina, fluido peritoneal y fluido pleural, o células procedentes de ellos. Muestras biológicas pueden incluir también secciones de tejidos tales como secciones congeladas tomadas para fines histológicos. A una muestra biológica se la puede aludir también como una “muestra de un paciente”.

Una “proteína” significa un polímero de residuos aminoácidos enlazados entre sí por enlaces péptido. El término, tal como se utiliza en esta memoria, se refiere a proteínas, polipéptidos y péptidos de cualquier tamaño, estructura o función. Sin embargo, típicamente, una proteína tendrá una longitud de al menos seis aminoácidos. Si la proteína es un péptido corto, éste será de una longitud de al menos aproximadamente 10 residuos aminoácidos. Una proteína puede ser que se presenta en la naturaleza, recombinante o sintética, o cualquier combinación de éstas. Una proteína puede comprender también un fragmento de una proteína o péptido que se presenta en la naturaleza. Una proteína puede ser una molécula sencilla o puede ser un complejo multi-molecular. El término proteína también se puede aplicar a polímeros de aminoácidos en los que uno o más residuos aminoácidos es un análogo químico artificial de un correspondiente aminoácido que se presenta en la naturaleza.

Un “fragmento de una proteína”, tal como se utiliza en esta memoria, se refiere a una proteína que es una parte de otra proteína. Por ejemplo, fragmentos de proteínas pueden comprender polipéptidos obtenidos al digerir proteína de longitud completa aislada de células cultivadas. Un fragmento de proteína comprende al menos aproximadamente seis aminoácidos. En otro ejemplo, el fragmento comprende al menos aproximadamente diez aminoácidos. Todavía en otro ejemplo, el fragmento de proteína comprende al menos aproximadamente dieciséis aminoácidos.

Tal como se utiliza en esta memoria, un “producto de expresión” es una biomolécula tal como una proteína, que se produce cuando un gen se expresa en un organismo. Un producto de expresión puede comprender modificaciones post-traducción.

La expresión “expresión de proteínas” se refiere al proceso mediante el cual una secuencia de ácidos nucleicos sufre una transcripción y traducción con éxito, de modo que se expresan niveles detectables de la secuencia de aminoácidos o de la proteína.

La expresión “perfil de expresión de proteínas” o “firma de expresión de proteínas” se refiere a un grupo de proteínas expresado por un tipo de célula o tejido particular (p. ej. neurona, endotelio de arteria coronaria o tejido patológico) en donde la presencia de las proteínas consideradas juntas o la expresión diferencial de este tipo de proteínas es indicativo/predictivo de una determinada afección.

El término “anticuerpo” significa una inmunoglobulina, ya sea natural o producida de forma sintética parcial o totalmente. Todos sus derivados que conservan su capacidad de unión específica están incluidos también en el término. El término cubre también cualquier proteína que tenga un dominio de unión que sea homólogo o ampliamente homólogo a un dominio de unión de inmunoglobulina. Un anticuerpo puede ser monoclonal o policlonal. El anticuerpo puede ser un miembro de cualquier clase de inmunoglobulina, incluida cualquiera de las clases humanas: IgG, IgM, IgA, IgD e IgE.

La expresión “fragmento de anticuerpo” se refiere a cualquier derivado de un anticuerpo que tiene una longitud que es menor que la longitud completa. En un aspecto, el fragmento de anticuerpo conserva al menos una porción importante de la capacidad de unión específica del anticuerpo de longitud completa; específicamente en calidad de un participante en la unión. Ejemplos de fragmentos de anticuerpo incluyen, pero no se limitan a fragmentos Fab, Fab', F(ab')₂, scFv, Fv, diacuerpo de dsFv y Fd. El fragmento de anticuerpo se puede producir por cualquier medio. Por ejemplo, el fragmento de anticuerpo puede producirse enzimática o químicamente mediante fragmentación de

un anticuerpo intacto o puede ser producido de forma recombinante a partir de un gen que codifica la secuencia parcial del anticuerpo. Alternativamente, el fragmento de anticuerpo puede ser producido de forma sintética en su totalidad o en parte. El fragmento de anticuerpo puede comprender un fragmento de anticuerpo de cadena sencilla. En otro ejemplo, el fragmento puede comprender cadenas múltiples que están enlazadas entre sí, por ejemplo mediante enlaces disulfuro. El fragmento también puede comprender un complejo multi-molecular. Un fragmento de anticuerpo funcional puede comprender, típicamente, al menos aproximadamente 50 aminoácidos y, más típicamente, comprenderá al menos aproximadamente 200 aminoácidos.

Determinación de los perfiles de expresión de genes

Se describe a continuación el método para identificar y validar los presentes perfiles de expresión de genes, indicativos de si un paciente de cáncer de colon respondería a un tratamiento con irinotecan. Se conocen otros métodos para identificar los perfiles de expresión de genes y/o proteínas; también podrían utilizarse cualquiera de estos métodos alternativos. Véase, *p. ej.*, Chen et al., *NEJM*, 356(1):11-20 (2007); Lu et al., *PLOS Med.*, 3(12):e467 (2006); Golub et al., *Science*, 286:531-537 (1999).

El presente método utiliza un ensayo paralelo en el que, en una pista, se identifican aquellos genes que están sobre-/sub-expresados en comparación con muestras de tejido normal (no canceroso) y, en una segunda pista, se identifican aquellos genes que comprenden inserciones o deleciones cromosómicas en comparación con muestras normales, procedentes de las mismas muestras. Estas dos pistas de análisis producen dos conjuntos de datos. Los datos se analizan utilizando un algoritmo que identifica los genes del perfil de expresión de genes (*es decir*, los genes que son diferencialmente expresados en tejido cancerígeno). Para normalizar los resultados se pueden emplear controles positivos y negativos, incluyendo eliminar aquellos genes y proteínas que también son expresados diferencialmente en tejidos normales procedentes de los mismos pacientes y confirmando que el perfil de expresión de genes es único para el cáncer de interés.

En el presente caso, como una etapa inicial, se adquirieron muestras biológicas de pacientes que padecían cáncer colorrectal. Se utilizaron aproximadamente quinientas (500) muestras de tejido obtenidas de pacientes con cáncer colorrectal, incluido tejido tumoral y tejido de colon normal (no enfermo) adyacente. Las muestras de tejido se obtuvieron de pacientes que padecían diversas fases de cáncer de colon, e incluían las obtenidas de pacientes quienes habían sido tratados con irinotecan. Todos los pacientes eran respondedores a una terapia con irinotecan. En una base de datos se registró la información clínica asociada con cada muestra, incluido el tratamiento con irinotecan y el resultado del tratamiento. La información clínica incluye también información tal como edad, sexo, historial médico, historial del tratamiento, síntomas, historial familiar, recurrencia (sí/no), etc. También se adquirieron de los mismos pacientes muestras control, que incluían muestras de tejido normal (no canceroso). Como controles positivos se utilizaron muestras de tejido de colon no enfermo normal procedente de un conjunto de individuos sanos, y como controles negativos se utilizaron muestras de tumores de colon procedentes de pacientes quienes eran no respondedores a la terapia con irinotecan.

Después se generaron perfiles de expresión de genes (GEPs) a partir de las muestras biológicas en base al ARN total de acuerdo con métodos bien establecidos. En síntesis, un método típico implica aislar ARN total procedente de la muestra biológica, amplificar el ARN, sintetizar ADNc, marcar el ADNc con un marcador detectable, hibridar el ADNc con una disposición ordenada genómica tal como GeneChip® de Affymetrix U133 y determinar la unión del ADNc marcado con la disposición ordenada genómica midiendo la intensidad de la señal procedente del marcador detectable unido a la disposición ordenada. Véanse, *p. ej.*, los métodos descritos en Lu, et al., Chen, et al. y Golub et al., *supra*, y las referencias citadas en los mismos. Los datos de expresión resultantes se introducen en una base de datos.

ARNms en las muestras de tejidos se pueden analizar utilizando sondas comercialmente disponibles o adaptadas según las necesidades del cliente o disposiciones ordenadas de oligonucleótidos tales como disposiciones ordenadas de ADNc o de oligonucleótidos. El uso de estas disposiciones ordenadas permite la medición simultánea de niveles de ARNm de estado estable de miles de genes, representando con ello una herramienta poderosa para identificar efectos tales como el brote, la detención o la modulación de la proliferación incontrolada de células. La hibridación y/o unión de las sondas en las disposiciones ordenadas a los ácidos nucleicos de interés procedentes de las células se puede determinar detectando y/o midiendo la localización e intensidad de la señal recibida de la sonda marcada o utilizada para detectar una secuencia de ADN/ARN procedente de la muestra que se hibrida a una secuencia de ácidos nucleicos en una localización conocida en la microdisposición ordenada. La intensidad de la señal es proporcional a la cantidad de ADNc o de ARNm presente en el tejido de la muestra. Están

disponibles y son útiles numerosas disposiciones ordenadas y técnicas. Métodos para determinar la expresión de genes y/o proteínas en tejidos de muestra se describen, por ejemplo, en la patente de EE.UU. nº 6.271.002; patente de EE.UU. nº 6.218.122; patente de EE.UU. nº 6.218.114; y patente de EE.UU. nº 6.004.755; y en Wang et al., *J. Clin. Oncol.*, 22(9):1564-1671 (2004); Golub et al, (*supra*); y Schena et al., *Science*, 270:467-470 (1995).

5 El aspecto del análisis de genes utilizado en el presente método investiga la expresión de genes así como datos de inserción/delección. Como una primera etapa, el ARN se aisló de las muestras de tejido y se marcó. Se realizaron procesos paralelos en la muestra para desarrollar dos conjuntos de datos: (1) la sobre-/sub-expresión de genes basada en niveles de ARNm; y (2) datos de inserción/delección cromosómica. Estos dos conjuntos de datos se correlacionaron luego por medio de un algoritmo. La sobre-/sub-expresión de los genes en cada una de las muestras de tejido cancerígeno se comparó con la expresión de genes en las muestras normales (no cancerosas), y se identificó un subconjunto de genes que se expresaron diferencialmente en el tejido cancerígeno. Preferiblemente, niveles de supra- y sub-regulación se distinguen en base a cambios múltiples de las mediciones de intensidad de las sondas de la microdisposición ordenada hibridadas. Se prefiere una diferencia de aproximadamente 2,0 veces o mayor para realizar este tipo de distinciones, o un valor p menor que 10 aproximadamente 0,05. Es decir, antes de que se diga que un gen está diferencialmente expresado en células enfermas frente a normales, se encuentra que la célula enferma proporciona una intensidad de expresión al menos aproximadamente 2 veces mayor o menor que las células normales. Generalmente, cuanto mayor sea el múltiplo de diferencia (o menor sea el valor p), tanto más se preferirá el gen para uso como una herramienta de diagnóstico o pronóstico. Genes seleccionados para las firmas de genes tienen niveles de expresión que resultan en la generación de una señal que es distinguible de las de los genes normales o no modulados en una cantidad que excede el fondo utilizando instrumentación de laboratorio clínico.

Se pueden utilizar valores estadísticos para distinguir con seguridad genes modulados de no modulados y el ruido. Ensayos estadísticos pueden identificar los genes expresados diferencialmente de la manera más significativa entre diversos grupos de muestras. El test t de Student es un ejemplo de un test estadístico robusto que puede utilizarse para encontrar diferencias significativas entre dos grupos. Cuando menor sea el valor p, tanto más convincente será la evidencia de que el gen está mostrando una diferencia entre los diferentes grupos. No obstante, dado que las microdisposiciones ordenadas permiten la medición de más de un gen a la vez, se pueden consultar a la vez decenas de miles de test estadísticos. Debido a ello, no es probable observar pequeños valores p al azar, y se pueden realizar ajustes utilizando una corrección de Sidak o etapa similar así como un experimento de aleatoriedad/permutación. Un valor p menor que aproximadamente 0,05 por parte del test t es una evidencia de que el nivel de expresión del gen es significativamente diferente. Una evidencia más convincente es un valor p menor que aproximadamente 0,05 después de haber tenido en cuenta la corrección de Sidak. Para un gran número de muestras en cada grupo, un valor p menor que aproximadamente 0,05 después del test de aleatoriedad/permutación es la evidencia más convincente de una diferencia significativa.

Otro parámetro que se puede utilizar para seleccionar genes que generan una señal que es mayor que la del gen no modulado o ruido es la medición de la diferencia de señales absoluta. Preferiblemente, la señal generada por los genes diferencialmente expresados difiere en al menos aproximadamente 20% de las del gen normal o no modulado (sobre una base absoluta). Se prefiere incluso más que este tipo de genes produzcan modelos de expresión que sean al menos aproximadamente 30% diferentes de los de genes normales o no modulados.

Este análisis de expresión diferencial se puede realizar utilizando disposiciones ordenadas comercialmente disponibles, por ejemplo disposiciones ordenadas GeneChip® de Affymetrix U133 (Affymetrix, Inc.). Estas disposiciones ordenadas tienen conjuntos de sondas para el genoma humano completo inmovilizado en el chip y se pueden utilizar para determinar la supra- y sub-regulación de genes en muestras de ensayo. También se pueden utilizar otros sustratos que tengan fijados sobre los mismos ADN genómico humano o sondas capaces de detectar productos de expresión tales como los disponibles de Affymetrix, Agilent Technologies, Inc. o Illumina, Inc. Microdisposiciones ordenadas de genes actualmente preferidas incluyen disposiciones ordenadas GeneChip® de Affymetrix U133 y microdisposiciones ordenadas de ADN genómico de Agilent Technologies. Instrumentos y reactivos para llevar a cabo el análisis de la expresión de genes están comercialmente disponibles. Véase, *p. ej.*, el sistema GeneChip® de Affymetrix. Los datos de expresión obtenidos del análisis se introducen entonces en la base de datos.

55 En el segundo brazo del presente método se obtuvieron datos de inserción/delección cromosómica para los genes de cada una de las muestras en comparación con muestras de tejido normal. El análisis de inserción/delección se generó utilizando una hibridación genómica comparativa ("CGH" – siglas en inglés) basada en disposiciones

ordenadas. La CGH basada en disposiciones ordenadas mide las variaciones en el número de copias en múltiples loci simultáneamente, proporcionando una importante herramienta para estudiar el cáncer y trastornos de desarrollo y para desarrollar dianas diagnósticas y terapéuticas. Microchips para llevar a cabo la CGH basada en disposiciones ordenadas están comercialmente disponibles, *p. ej.*, de Agilent Technologies. El chip de Agilent es una disposición ordenada cromosómica que muestra la localización de genes en los cromosomas y proporciona datos adicionales para la firma de genes. Los datos de inserción/delección procedentes de este ensayo se introducen en la base de datos.

Los análisis se llevan a cabo en las mismas muestras de los mismos pacientes para generar datos paralelos. Para reducir la variabilidad se utilizan los mismos chips y la misma preparación de muestras.

También se determina la expresión de determinados genes conocidos como “genes de referencia”, “genes control” o “genes de control interno”, preferiblemente al mismo tiempo, como medios para asegurar la veracidad del perfil de expresión. Genes de referencia son genes que se expresan de forma consistente en muchos tipos de tejidos, incluidos tejidos cancerosos y normales y, así, son útiles para normalizar los perfiles de la expresión de genes. Véase, *p. ej.*, Silvia et al., *BMC Cancer*, 6:200 (2006); Lee et al., *Genome Research*, 12(2):292-297 (2002); Zhang et al., *BMC Mol. Biol.*, 6:4 (2005). La determinación de la expresión de genes de referencia en paralelo con los genes en el perfil de expresión de genes único proporciona una garantía adicional de que las técnicas utilizadas para la determinación del perfil de expresión de genes están actuando adecuadamente. Los datos de expresión relacionados con los genes de referencia también se introducen en la base de datos. En una realización actualmente preferida, los siguientes genes se utilizan como genes de referencia: ACTB, GAPD, GUSB, RPLP0 y/o TRFC.

Correlación de datos

Los datos de la expresión diferencial y los datos de inserción/delección en la base de datos se correlacionan con la información de los resultados clínicos asociada con cada muestra de tejido también en la base de datos por medio de un algoritmo para determinar un perfil de expresión de genes para determinar la eficacia terapéutica de irinotecan, así como la recurrencia tardía de la enfermedad y/o la muerte relacionada con la enfermedad asociada con la terapia con irinotecan. Están disponibles diversos algoritmos que son útiles para correlacionar los datos e identificar las firmas de genes predictivas. Por ejemplo, para la práctica de las realizaciones descritas en esta memoria se pueden utilizar algoritmos tales como los identificados en Xu et al., A Smooth Response Surface Algorithm For Constructing A Gene Regulatory Network, *Physiol. Genomics* 11:11-20 (2002).

Otro método para identificar perfiles de expresión de genes es a través del uso de algoritmos de optimización tal como el algoritmo de media-varianza ampliamente utilizado para establecer portafolios de acciones. Un método de este tipo se describe en detalle en la solicitud de patente de EE.UU. con nº de publicación 2003/0194734. En esencia, el método exige el establecimiento de un conjunto de entradas (según se mide por la intensidad) que optimizarán el retorno (señal que se genera) que se recibe para utilizarlo, al tiempo que se minimiza la variabilidad del retorno. También se puede utilizar el algoritmo descrito en Irizarry et al., *Nucleic Acids Res.*, 31:e15 (2003). El algoritmo actualmente preferido es el algoritmo de JMP Genomics disponible de JMP Software.

El proceso para seleccionar perfiles de expresión de genes también puede incluir la aplicación de reglas heurísticas. Tales reglas se formulan en base a la biología y a una comprensión de la tecnología utilizada para producir resultados clínicos, y se aplican para extraer el resultado a partir del método de optimización. Por ejemplo, el método de la media-varianza para la identificación de la firma de genes se puede aplicar a datos de la microdisposición ordenada para un cierto número de genes diferencialmente expresados en sujetos con cáncer colorrectal. El resultado del método sería un conjunto optimizado de genes que podría incluir algunos genes que se expresan en la sangre periférica así como en tejido enfermo. Si muestras utilizadas en el método de ensayo se obtienen de la sangre periférica y determinados genes diferencialmente expresados en casos de cáncer también podrían ser diferencialmente expresados en la sangre periférica, entonces se puede aplicar una regla heurística en la que se selecciona un portafolio de la frontera eficiente que excluye aquellos que son diferencialmente expresados en la sangre periférica. Naturalmente, la regla se puede aplicar antes de la formación de la frontera eficiente, por ejemplo aplicando la regla durante la pre-selección de datos.

Se pueden aplicar otras reglas heurísticas que no estén necesariamente relacionadas con la biología en cuestión. Por ejemplo, se puede aplicar una regla de que sólo pueda ser representado un determinado porcentaje del portafolio por parte de un gen o grupo de genes particular. Software comercialmente disponible tal como el

software de Wagner alberga fácilmente este tipo de heurísticas (Wagner Associates Mean-Variance Optimization Application, www.wagner.com). Esto puede ser útil, por ejemplo, cuando factores distintos de la seguridad y precisión tengan un impacto sobre el deseo de incluir uno o más genes.

5 Como un ejemplo, el algoritmo se puede utilizar para comparar perfiles de expresión de genes para diversos genes (o portafolios) para atribuir pronósticos. Los perfiles de expresión de genes de cada uno de los genes que comprenden el portafolio se fijan en un medio tal como un medio legible por ordenador. Éste puede adoptar un cierto número de formas. Por ejemplo, se puede establecer una tabla en la que se introduzcan el intervalo de señales (p. ej. mediciones de intensidad) indicativas de la enfermedad. Los datos reales del paciente se pueden
10 luego comparar con los valores en la tabla para determinar si las muestras del paciente son normales o enfermas. En una realización más sofisticada, se registran digital o gráficamente modelos de las señales de expresión (p. ej. intensidad fluorescente). Los modelos de expresión de genes procedentes de los portafolios de genes utilizados en unión con las muestras de los pacientes se comparan entonces con los modelos de expresión. Software de comparación de modelos se puede luego utilizar para determinar si las muestras de los pacientes tienen un modelo
15 indicativo de recurrencia de la enfermedad. Naturalmente, estas comparaciones también se pueden utilizar para determinar si no es probable que el paciente experimente recurrencia de la enfermedad. Estos perfiles de expresión de las muestras se comparan entonces con el portafolio de una célula control. Si los modelos de expresión de las muestras son consistentes con los modelos de expresión para la recurrencia de un cáncer colorrectal, entonces (en ausencia de consideraciones médicas compensatorias) el paciente es tratado como se trataría a un paciente de
20 recaída. Si los modelos de expresión de las muestras son consistentes con el modelo de expresión procedente de la célula normal/control, entonces al paciente se le diagnostica negativo para el cáncer colorrectal.

Un método para analizar las firmas de genes de un paciente para determinar el pronóstico de cáncer es a través del uso de un programa de análisis de riesgo Cox. El análisis se puede realizar utilizando el software S-Plus
25 (comercialmente disponible de Insightful Corporation, www.insightful.com). Al utilizar métodos de este tipo, un perfil de expresión de genes se compara con un perfil que representa con seguridad la recaída (*es decir*, los niveles de expresión para la combinación de genes en el perfil es indicativa de la recaída). El modelo de riesgo de Cox con el umbral establecido se utiliza para comparar la similitud de los dos perfiles (recaída conocida frente a paciente) y luego determina si el perfil del paciente excede del umbral. Si lo hace, entonces el paciente se clasifica como
30 alguien que recaerá y se le acuerda tratamiento tal como terapia adyuvante. Si el perfil del paciente no excede del umbral, entonces estos se clasifican como un paciente de no recaída. Otras herramientas analíticas también se pueden utilizar para responder la misma cuestión tales como análisis discriminativo lineal, regresión logística y enfoques de red neuronales. Véase, p. ej., el software disponible de JMP statistical software (www.jmp.com).

35 Están disponibles numerosos otros métodos bien conocidos del reconocimiento de modelos. Las siguientes referencias proporcionan algunos ejemplos:

Votación ponderada: Golub, T R., Slonim, D K., Tamaya, P., Huard, C., Gaasenbeek, M., Mesirov, J P., Coller, H., Loh, L., Downing, J R, Caligiuri, M A., Bloomfield, C D., Lander, E S. Molecular classification of cancer: class
40 discovery and class prediction by gene expression monitoring. *Science* 286:531-537, 1999.

Máquinas de vectores de apoyo: Su, A I., Welsh, J B., Sapinosa, L M., Kern, S G., Dimitrov, P., Lapp, H., Schultz, P G., Powell, S M., Moskaluk, C A., Frierson, H F. Jr., Hampton, G M. Molecular classification of human
45 carcinomas by use of gene expression signatures. *Cancer Research* 61:7388-93, 2001. Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J P., Poggio, T., Gerald, W., Loda, M., Lander, E S., Gould, T R. Multiclass cancer diagnosis using tumor gene expression signatures *Proceedings of the National Academy of Sciences of the USA* 98:15149-15154, 2001.

Vecinos más próximos a K: Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J P., Poggio, T., Gerald, W., Loda, M., Lander, E S., Gould, T R. Multiclass cancer diagnosis using tumor gene expression signatures *Proceedings of the National Academy of
50 Sciences of the USA* 98:15149-15154, 2001.

Coefficientes de correlación: van't Veer L J, Dai H, van de Vijver M J, He Y D, Hart A, Mao M, Peterse H L, van der Kooy K, Marton M J, Witteveen A T, Schreiber G J, Kerkhoven R M, Roberts C, Linsley P S, Bernards R, Friend S
55 H. Gene expression profiling predicts clinical outcome of breast cancer, *Nature*. Enero de 2002 31; 415(6871):530-6.

El análisis de expresión de genes identifica un perfil de expresión de genes (GEP) único para las muestras de cáncer, es decir, los genes que son diferencialmente expresados por las células cancerígenas. Este GEP se valida luego, por ejemplo, utilizando la reacción en cadena de la polimerasa cuantitativa en tiempo real (RT-qPCR), que se puede llevar a cabo utilizando instrumentos y reactivos comercialmente disponibles, tales como los disponibles de Applied Biosystems (www.appliedbiosystems.com).

En el presente caso, los resultados de los análisis de la expresión de genes demostraron que en pacientes de cáncer de colon, quienes respondían al tratamiento con irinotecan, estaban supra-regulados los siguientes genes: ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt y BAG1, y estaban sub-regulados los siguientes genes: quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, en comparación con la expresión de estos genes en muestras de tejido de colon normal procedentes de estos pacientes, y procedentes de los pacientes control negativos, es decir, las muestras de tejido de pacientes que habían experimentado una recurrencia de su cáncer después de tratamiento con irinotecan. Los genes de referencia ACTB, GAPD, GUSB, RPLP0 y TFRC estaban todos supra-regulados.

Determinación de los perfiles de expresión de proteínas

No todos los genes expresados por una célula son traducidos en proteínas y, por lo tanto, una vez que se ha identificado un GEP, es deseable confirmar si las proteínas que corresponden a alguno o a todos los genes diferencialmente expresados en el GEP son también diferencialmente expresadas por parte de las mismas células o tejido. Por lo tanto, se generan perfiles de expresión de proteínas (PEPs) procedentes de los mismos tejidos cancerígenos y de control utilizados para identificar los GEPs. Los PEPs también se utilizan para validar el GEP en otros pacientes de cáncer de colon.

El método preferido para generar PEPs de acuerdo con la presente invención es mediante análisis de inmunohistoquímica (IHC – siglas en inglés). En este método, se utilizan anticuerpos específicos para las proteínas en el PEP para interrogar muestras de tejido procedentes de pacientes de cáncer. Se conocen otros métodos para identificar PEPs, *p. ej.* hibridación *in situ* (ISH – siglas en inglés) utilizando sondas de ácidos nucleicos específicas para proteínas. Véase, *p. ej.* Hofer et al., *Clin. Can. Res.*, 11(16):5722 (2005); Volm et al., *Clin. Exp. Metas.*, 19(5):385 (2002). También podría utilizarse cualquiera de estos métodos alternativos.

En el presente caso, muestras de tejido tumoral de colon y de tejido de colon normal se obtuvieron de pacientes que padecían cáncer de colon y que habían sido sometidos a tratamiento con éxito con irinotecan; estas son las mismas muestras que las utilizadas para identificar el GEP. Las muestras de tejido se dispusieron ordenadamente en microdisposiciones ordenadas de tejidos (TMAs – siglas en inglés) para permitir el análisis simultáneo. Las TMAs consisten en sustratos tales como portaobjetos de vidrio, sobre los cuales se reúnen de una forma en disposición ordenada hasta aproximadamente 1000 muestras de tejido separadas para permitir el análisis histológico simultáneo. Las muestras de tejido pueden comprender tejido obtenido de muestras de biopsia conservadas, *p. ej.* tejidos embebidos en parafina o congelados. Técnicas para producir microdisposiciones ordenadas de tejido son bien conocidas en la técnica. Véase, *p. ej.*, Simon et al., *BioTechniques*, 36(1):98-105 (2004); Kallioniemi et al., documento WO 99/44062; Kononen et al., *Nat. Med.*, 4:844-847 (1998). En el presente caso, se utilizó una aguja hueca para separar núcleos de tejido tan pequeños como de 0,6 mm de diámetro de regiones de interés en tejidos embebidos en parafina. Las “regiones de interés” son aquellas que han sido identificadas por un patólogo como que contienen el tejido enfermo o normal deseado. Estos núcleos de tejido se insertaron entonces en un bloque de parafina receptor en un modelo de disposición ordenada espaciada con precisión. Se cortaron secciones de este bloque utilizando un microtomo, montado sobre un portaobjetos de microscopio, y luego se analizaron mediante análisis histológico estándar. Cada uno de los bloques de disposición ordenada se puede cortar en aproximadamente 100 a aproximadamente 500 secciones que pueden ser sometidas a test independientes.

Las TMAs se prepararon utilizando dos muestras de tejido de cada uno de los pacientes: una de tejido tumoral de colon y otra de tejido de colon normal. También se prepararon disposiciones ordenadas control; en una realización actualmente preferida, se utilizaron las siguientes TMAs de control: una disposición ordenada que contenía muestras de tejido de colon normal procedente de individuos sanos y exentos de cáncer; una disposición ordenada de “controles positivos” que contenían tejidos tumorales procedentes de pacientes de cáncer que padecían cánceres distintos a cáncer de colon, *p. ej.* cáncer de mama, cáncer de pulmón, cáncer de próstata, etc.; y una disposición ordenada de “controles negativos” que contenían muestras de tumores procedentes de pacientes de cáncer de colon que habían experimentado recurrencias del cáncer después del tratamiento con irinotecan – es

decir, pacientes que eran “no respondedores” a la terapia.

Proteínas en las muestras de tejidos se pueden analizar interrogando las TMAs utilizando agentes específicos para proteínas tales como anticuerpos o sondas de ácidos nucleicos tales como aptámeros. Para este fin se prefieren anticuerpos debido a su especificidad y disponibilidad. Los anticuerpos pueden ser anticuerpos monoclonales o policlonales, fragmentos de anticuerpos y/o diversos tipos de anticuerpos sintéticos, incluidos anticuerpos quiméricos o fragmentos de los mismos. Los anticuerpos están comercialmente disponibles de un cierto número de fuentes (*p. ej.* Abcam (www.abcam.com), Cell Signaling Technology (www.cellsignal.com), Santa Cruz Biotechnology (www.santacruz.com)) o se pueden generar utilizando técnicas bien conocidas por los expertos en la técnica. Típicamente, los anticuerpos están equipados con marcadores detectables tales como enzimas, cromógenos o puntos cuánticos que permiten detectar a los anticuerpos. Los anticuerpos se pueden conjugar o marcar directamente con un marcador detectable, o indirectamente con un miembro de un par de unión, de los que el otro miembro contiene un marcador detectable. Sistemas de detección para uso con los mismos se describen, por ejemplo, en la página web de Ventana Medical Systems, Inc. (www.ventanamed.com). Los puntos cuánticos son particularmente útiles como marcadores detectables. El uso de puntos cuánticos se describe, por ejemplo, en las siguientes referencias: Jaiswal et al., *Nat. Biotechnol.*, 21:47-51 (2003); Chan et al., *Curr. Opin Biotechnol.*, 13:40-46 (2002); Chan et al., *Science*, 281:435-446 (1998).

El uso de anticuerpos para identificar proteínas de interés en las células de un tejido, al que se alude como inmunohistoquímica (IHC), está bien establecido. Véase, *p. ej.* Simon et al., *BioTechniques*, 36(1):98 (2004); Haedicke et al., *BioTechniques*, 35(1): 164 (2003). El ensayo de IHC se puede automatizar utilizando instrumentos comercialmente disponibles tales como los instrumentos Benchmark disponibles de Ventana Medical Systems Inc.

En el presente caso, las TMAs se pusieron en contacto con anticuerpos específicos para las proteínas codificadas por los genes identificados en el estudio de expresión de genes como supra- o sub-regulados en pacientes de cáncer de colon quienes eran respondedores a la terapia con irinotecan con el fin de determinar la expresión de estas proteínas en cada uno de los tipos de tejido. Los resultados del ensayo de IHC demostraron que en pacientes de cáncer de colon que respondían al tratamiento con irinotecan, las siguientes proteínas estaban supra-reguladas: ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1, y las siguientes proteínas estaban sub-reguladas: quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, en comparación con una expresión de estas proteínas en tejido de colon normal procedente de estos pacientes y en las muestras control negativo, es decir, las muestras de tumor de colon procedentes de pacientes que habían experimentado una recurrencia de su cáncer después de tratamiento con irinotecan (no respondedores). Adicionalmente, el análisis de IHC demostró que una mayoría de estas proteínas no estaba supra- ni sub-reguladas en las muestras de tejido control positivo. Las proteínas de referencia ACTB, GAPD, GUSB, RPLP0 y TFRC estaban todas supra-reguladas.

Ensayos

La presente descripción comprende, además, métodos y ensayos para determinar si es probable que un paciente con cáncer de colon responda a un tratamiento con irinotecan, y/o para predecir si es probable que el cáncer recurra, o se produzca una muerte relacionada con la enfermedad. De acuerdo con un aspecto, se puede utilizar un ensayo de IHC formateado para determinar si un tumor de cáncer de colon exhibe el presente GPEP. Los ensayos se pueden formular en kits que incluyen la totalidad o algunos de los materiales necesarios para realizar el análisis, incluidos reactivos (anticuerpos, marcadores detectables, etc.) e instrucciones.

El método de ensayo comprende poner en contacto una muestra de tumor de un paciente de cáncer de colon con un grupo de anticuerpos específicos para algunos o la totalidad de genes o proteínas en el presente GPEP, y determinar la ocurrencia de una supra- o sub-regulación de estos genes o proteínas en la muestra. El uso de TMAs permite ensayar simultáneamente numerosas muestras, incluidas muestras control.

En un ejemplo preferido, el método comprende poner en contacto una muestra del tumor procedente de un paciente con cáncer de colon y muestras control con un grupo de anticuerpos específicos para alguna o la totalidad de las proteínas en el presente GPEP, y determinar la ocurrencia de una supra- o sub-regulación de estas proteínas. La supra-regulación de alguna o de la totalidad de las siguientes proteínas: ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1; y la sub-regulación de alguna o la totalidad de las siguientes proteínas: quinasa Erk 1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, es indicativa de la capacidad de respuesta del paciente a irinotecan. Preferiblemente, al menos aproximadamente

cuatro, preferiblemente entre aproximadamente cuatro y diez y, lo más preferiblemente, entre aproximadamente diez y dieciséis (o más) anticuerpos se utilizan en el presente método.

5 Preferiblemente, el método incluye también detectar y/o cuantificar proteínas control o "proteínas de referencia". La detección y/o cuantificación de las proteínas de referencia en las muestras normaliza los resultados y, así, proporciona una garantía adicional de que el ensayo está funcionando adecuadamente. En un ejemplo actualmente preferido, se incluyen anticuerpos específicos para una o más de las siguientes proteínas de referencia: ACTB, GAPD, GUSB, RPLP0 y/o TRFC.

10 La presente descripción comprende, además, un kit que contiene reactivos para llevar a cabo un análisis por IHC de muestras de tejido o células procedentes de pacientes de cáncer de colon, incluidos anticuerpos específicos para al menos aproximadamente cuatro de las proteínas en el GPEP y para cualesquiera proteínas de referencia. Los anticuerpos están preferiblemente marcados con medios para detectar la unión de los anticuerpos a las proteínas de interés, *p. ej.*, marcadores detectables. Marcadores detectables preferidos incluyen compuestos fluorescentes o puntos cuánticos, pero se pueden utilizar otros tipos de marcadores detectables. Marcadores detectables para anticuerpos están comercialmente disponibles, *p. ej.*, de Ventana Medical Systems, Inc.

20 Son bien conocidos métodos inmunohistoquímicos para detectar y cuantificar la expresión de proteínas en muestras de tejidos. Se puede utilizar cualquier método que permita la determinación de la expresión de varias proteínas diferentes. Véase, *p. ej.*, Signoretti *et al.*, "Her-2-neu Expression and Progressión Toward Androgen Independence in Human Prostate Cancer", *J. Natl. Cancer Instit.*, 92(23): 1918-25 (2000); Gu *et al.* "Prostate stem cell antigen (PSCA) expression increases with high gleason score, advanced stage and bone metastasis in prostate cancer", *Oncogene*, 19:1288-96 (2000). Métodos de este tipo se pueden llevar a cabo eficazmente utilizando instrumentos automatizados diseñados para el análisis inmunohistoquímico (IHC). Instrumentos para realizar rápidamente ensayos de este tipo están comercialmente disponibles, *p. ej.* de Ventana Molecular Discovery Systems o Lab Vision Corporation. Métodos de acuerdo con la presente invención que utilizan este tipo de instrumentos se llevan a cabo de acuerdo con las instrucciones del fabricante.

30 Anticuerpos específicos para proteínas para uso en este tipo de métodos o ensayos están disponibles fácilmente o pueden prepararse utilizando técnicas bien establecidas. Anticuerpos específicos para las proteínas en el GPEP descrito en esta memoria se pueden obtener, por ejemplo, de Cell Signaling Technology, Inc., Santa Cruz Biotechnology, Inc. o Abcam.

EJEMPLOS

35 Se sometieron a ensayo una serie de factores de pronóstico con el fin de validar la eficacia del perfil de expresión de genes/proteínas (GPEP) de la presente invención para predecir la respuesta terapéutica de una terapia con irinotecan. Los niveles de expresión de estos factores, que consisten en veintidós (22) proteínas en el presente GPEP listadas en la Tabla 2 (que incluye diecisiete proteínas diferencialmente expresadas y cinco proteínas de referencia) se determinó mediante una metodología inmunohistoquímica en muestras de tejido de biopsia obtenidas de pacientes de cáncer de colon en estado avanzado, cuyo tratamiento con irinotecan había sido un éxito, así como muestras procedentes de pacientes cuyo tratamiento no tuvo éxito, *p. ej.* quienes habían experimentado una recurrencia tardía (LRec) o una muerte relacionada con la enfermedad (DRD) asociada con la terapia. Para fines de seleccionar los pacientes para el estudio, se determinó que la terapia con irinotecan fracasaba si se presentaba una recurrencia en el espacio de tres años del diagnóstico.

50 De acuerdo con la actual información de prescripción para CAMPTOSAR®, irinotecan está actualmente indicado para la terapia de primera línea del cáncer de colon en combinación con 5-fluorouracilo (5-FU) y leucovorina, o después de una terapia inicial con 5-FU en pacientes de cáncer de colon en estado avanzado. Los pacientes en el estudio han sido tratados utilizando la terapia de combinación de acuerdo con la información de prescripción para CAMPTOSAR®.

Perfil de expresión de genes/proteínas (GPEP):

55 En este estudio, se evaluaron probetas de cáncer de colon primario embebidas con parafina fijada con formalina de 280 pacientes (media de edad de 63 años), seguido de un mínimo de 120 meses para el tamaño primario del tumor, la calidad histológica y el estadio de Duke. Estos pacientes incluían sólo aquellos que habían respondido a una terapia con irinotecan. Ninguno de los pacientes recibió un tratamiento con adyuvantes antes del primer

episodio de la recurrencia de la enfermedad. Utilizando las técnicas descritas anteriormente, se generó un GPEP que consistía en los siguientes diecisiete genes y proteínas codificados: ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68, BAG1, quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, y cinco genes y proteínas de referencia: ACTB, GAPD, GUSB, RPLP0 y TRFC.

Se prepararon microdisposiciones ordenadas de tejidos utilizando los adenocarcinomas de colon y tejido de colon normal (no canceroso) de pacientes descritos anteriormente con cánceres en estado avanzado que habían sido tratados con irinotecan. También se prepararon TMAs que contenían muestras control, positivas y negativas. Las TMAs utilizadas en este estudio se describen en la Tabla A:

Tabla A. Microdisposiciones ordenadas de tejidos

Disposición ordenada de rastreo normal	Esta disposición ordenada contenía muestras de tejido de colon normal (no canceroso) procedente de 200 pacientes (2 muestras por paciente)
Irinotecan para tratamiento del colon	Esta disposición ordenada contenía 280 muestras de pacientes obtenidas de los pacientes que padecían adenocarcinoma de colon en fase avanzada quienes habían sido tratados con CAMPTOSAR®, junto con muestras de tejido de colon normal procedentes de cada uno de los pacientes.
Disposición ordenada de supervivencia de rastreo de cáncer (disposición ordenada de control positivo)	Esta disposición ordenada contenía 200 muestras de tumores para cánceres distintos de cáncer de colon, que incluyen cáncer de mama, cáncer pancreático, cáncer de próstata, cáncer de ovarios, cáncer de las glándulas salivares, cáncer de pulmón y cáncer de cerebro.
Progreso del cáncer de colon (disposición ordenada de control negativo – disposición ordenada TE30)	Esta disposición ordenada contenía muestras de tejido de cáncer de colon procedente de treinta pacientes quienes habían progresado a la siguiente fase del cáncer o habían experimentado una recurrencia del cáncer después de tratamiento con CAMPTOSAR®.

Las TMAs se construyeron de acuerdo con el siguiente proceso:

Construcción: se utilizó un instrumento para crear agujeros en un bloque de parafina receptora que luego se llenaron con núcleo de tejido adquirido de un bloque de donantes seleccionado. Estos núcleos de tejido se troquelaron con una broca afilada de paredes delgadas. Una guía de precisión X-Y permitió la colocación ordenada de estas muestras de tejido en un formato de disposición ordenada.

Presentación: Secciones de TMA se cortaron con espesores de 4 micras y se montaron en micro-portaobjetos de vidrio cargados positivamente. Los elementos individuales tenían un diámetro de 0,6 mm y estaban separados 0,2 mm.

Elementos: Además de TMAs que contenían las muestras de cáncer de colon, se produjeron disposiciones ordenadas de rastreo y control que contenían las muestras de tejido descritas en la Tabla A.

Especificidad: Las TMAs se diseñaron para uso con los métodos de tinción especializada e inmunohistoquímicos para fines de rastreo de la expresión de genes utilizando anticuerpos monoclonales y policlonales a lo largo de una amplia gama de tipos de tejidos caracterizados.

Acompañando a cada disposición ordenada se encontraba un mapa localizador de disposiciones ordenadas y una hoja de cálculo que contenía el diagnóstico del paciente, datos histológicos y demográficos para cada uno de los elementos.

Se utilizaron técnicas de tinción inmunohistoquímica para la visualización de proteínas (células) de tejidos presentes en las muestras de tejidos en las TMAs. Estas técnicas se basaron en la inmunoreactividad de anticuerpos y en las propiedades químicas de enzimas o complejos enzimáticos que reaccionan con cromógenos

de sustrato incoloros para producir un producto final coloreado. Manchas inmunoenzimáticas iniciales utilizaban el método directo, el cual conjugaba directamente a un anticuerpo con especificidad antigénica conocida (anticuerpo primario).

5 Se empleó una técnica de avidina-biotina marcada modificada, en la que un anticuerpo secundario biotinilado formaba un complejo con moléculas de estreptavidina conjugadas con peroxidasa. La actividad de peroxidasa endógena se mitigó mediante la adición de peróxido de hidrógeno al 3%. Las probetas se incubaron después con los anticuerpos primarios, seguido de incubaciones secuenciales con el anticuerpo de enlace secundario biotinilado (que contiene inmunoglobulinas anti-conejo o anti-ratón) y estreptavidina marcada con peroxidasa. El anticuerpo primario, el anticuerpo secundario y el complejo avidina-enzima se visualizan después utilizando un cromógeno del sustrato que produce un pigmento pardo en el lugar del antígeno que es visible mediante microscopía óptica. Los anticuerpos utilizados en este estudio eran anticuerpos específicos para las proteínas en el presente perfil de expresión de proteínas, es decir, ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68, BAG1, quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, y las proteínas de referencia ACTB, GAPD, GUSB, RPLP0 y TRFC. Todos los anticuerpos se obtuvieron de Cell Signaling Technology, Inc. y Abcam.

Proceso de tinción IHC automatizado:

- 20 1. La recuperación del epítipo inducida por calor (HIER) utilizando disolución de tampón citrato 10 mM, pH 6,0, se realizó como sigue:
 - a. Secciones desparafinadas y rehidratadas se colocaron en un bastidor para la tinción de portaobjetos.
 - 25 b. El bastidor se colocó en una estufa a presión que funcionaba con microondas; para cubrir los portaobjetos se añadieron 750 ml de tampón citrato 10 mM, pH 6,0.
 - c. La estufa a presión cubierta se colocó en el microondas en elevada potencia durante 15 minutos.
 - d. La estufa a presión se retiró del microondas y se enfrió hasta que el indicador de la presión había caído y la tapa podía ser retirada con seguridad.
 - 30 e. Se dejó que los portaobjetos se enfriaran hasta la temperatura ambiente (TA) y se llevó a cabo una tinción IHC.
2. Los portaobjetos se trataron con H₂O₂ al 3% durante 10 min a TA para mitigar la actividad de peroxidasa endógena.
- 35 3. Los portaobjetos se aclararon suavemente con solución salina tamponada con fosfato (PBS – siglas en inglés).
4. Los anticuerpos primarios se aplicaron a la dilución predeterminada (de acuerdo con las especificaciones de Cell Signaling Technology) durante 30 min a la TA. Suero normal de ratón o conejo en la dilución 1:750 se aplicó a portaobjetos control negativos.
- 40 5. Los portaobjetos se aclararon con solución salina tamponada con fosfato (PBS).
- 45 6. Anticuerpos de enlace biotinilados secundarios (anticuerpos secundarios: inmunoglobulinas anti-pollo y anti-ratón biotiniladas en solución salina tamponada con fosfato (PBS) que contenía proteína soporte y azida de sodio 15 mM) se aplicaron durante 30 min a la TA.
7. Los portaobjetos se aclararon con solución salina tamponada con fosfato (PBS).
- 50 8. Los portaobjetos se trataron con estreptavidina-HRP (estreptavidina conjugada a peroxidasa de rábano picante en PBS que contenía proteína soporte y agentes antimicrobianos procedentes de Ventana) durante 30 min a la temperatura ambiente.
- 55 9. Los portaobjetos se aclararon con solución salina tamponada con fosfato (PBS).
10. Los portaobjetos se trataron con sustrato/cromógeno (en tampón sustrato-imidazol-HCl pH 7,5 que contenía H₂O₂ y agentes antimicrobianos; DAB-3,3'-diaminobenzidina en disolución cromógena

procedente de Ventana) durante 10 min a la temperatura ambiente.

11. Los portaobjetos se aclararon con agua destilada.

5 12. Una contratinción en hematoxilina se aplicó durante 1 min.

13. Los portaobjetos se lavaron en agua corriente durante 2 min.

14. Los portaobjetos se deshidrataron después, se aclararon y se montó el cubreobjetos.

10 Todos los anticuerpos primarios se titularon a diluciones de acuerdo con las especificaciones del fabricante. La tinción de portaobjetos de disposiciones ordenadas de ensayo TE30 (descritos más abajo) se realizó con y sin la recuperación del epítipo (HIER). Los portaobjetos se rastrearon por parte de un patólogo para determinar la dilución de funcionamiento óptima. El tratamiento previo con HIER proporcionó una fuerte tinción específica con un pequeño fondo o sin fondo. El proceso de IHC anteriormente descrito se llevó a cabo utilizando un instrumento Benchmark de Ventana Medical Systems, Inc.

Criterios de puntuación:

20 La tinción se puntuó por parte de un patólogo en una escala de 0-3+ con 0 = sin tinción y siendo traza menor que 1+ pero mayor que 0. Los procesos de puntuación se describen en Signoretti et al., *J. Nat. Cancer Inst.*, Vol. 92, (23):1918 (diciembre 2000) y Gu et al., *Oncogene*, vol. 19, pág. 1288 (2000). Las categorías de 1+ a 3+ representan una intensidad incrementada de tinción, siendo 3+ una tinción parda oscura intensa. Los criterios de puntuación se basaron también en el porcentaje total de tinción 0 = 0%, 1 = menos de 25%, 2 = 25-50% y 3 = mayor que 50%. Se analizaron la positividad en porcentaje y la intensidad de tinción tanto para componentes nucleares como citoplásmicos así como sub-celulares. Las puntuaciones tanto de la intensidad como de porcentaje positivas se multiplicaron para producir un número del 0-9. La tinción 3+ se determinó a partir de la expresión conocida del antígeno procedente de los controles positivos.

30 Controles de tejido positivos se definieron a través de análisis de transferencia Western. Este experimento se realizó para confirmar el nivel de expresión de proteínas en cada uno de los tejidos controles. Los controles negativos se definieron también por la misma metodología. Los controles positivos consistían en muestras de adenocarcinomas de mama, próstata, pulmón, glándulas salivares, páncreas y ovarios y muestras de tejido tumoral de cerebro no relacionadas con los pacientes de cáncer de colon quienes eran los sujetos del estudio.

35 Muestras de de tejido de cáncer de colon procedentes de pacientes que no respondían a una terapia con irinotecan (*es decir*, quienes experimentaban una recurrencia de la enfermedad o la muerte por la enfermedad después de tratamiento) se utilizaron como controles negativos.

40 La expresión positiva se evaluó también utilizando una disposición ordenada de xeno-injertos. A ratones SCID se les inyectaron células tumorales derivadas de pacientes quienes respondían a un tratamiento con irinotecan y se dejó que los tumores se desarrollaran en los ratones. Una vez que se establecieron los tumores, a los ratones se les inyectaron 200 mg/kg de irinotecan, y los ratones fueron vigilados para observar la capacidad de respuesta al fármaco. Como resultado del tratamiento con irinotecan, los tumores formados en los ratones SCID se redujeron o eliminaron. Antes del tratamiento con el fármaco, las muestras de los tumores se extrajeron de los ratones y se utilizaron para producir una TMA. El ensayo de IHC de la TMA que contenía el tejido tumoral de xeno-injerto de ratón demostró que los tumores de xeno-injerto tienen el mismo GPEP que el identificado en pacientes humanos quienes respondían a una terapia con irinotecan.

50 Todas las tandas se agruparon por disposiciones ordenadas de anticuerpos y tejidos que aseguraban que las tandas fueran normalizadas, significando que la totalidad de las disposiciones ordenadas de tejidos se tiñeron bajo las mismas condiciones con el mismo anticuerpo en la misma tanda. La reproducibilidad se comparó y validó.

Resultados:

55 El análisis univariable de perfiles GPEP de las muestras de pacientes descrito anteriormente predecía de forma independiente y fiable la respuesta a una terapia con irinotecan ($p < 0,0001$); recurrencia tardía (LRec, $p < 0,0005$); y muerte relacionada con la enfermedad (DRD, $p < 0,0001$). Cuando se estratifica en grupos GPEP-negativos, límite-positivos y altamente positivos, pacientes con tumores altamente positivos tenían un riesgo relativo (riesgo relativo

ajustado) de respuesta a la terapia con irinotecan de 8,3 (intervalo 2,1 – 32,4); LRec de 4,3 (intervalo 1,7 – 11,0) y DRD de 11,0 (intervalo 3,0 – 40,7). El tamaño del tumor y la calidad histológica no predecían una respuesta a la terapia con irinotecan, LRec ni DRD.

- 5 Los resultados de este estudio demuestran que en pacientes de cáncer de colon en fase avanzada, la positividad de GPEP mediante inmunohistoquímica predecía de forma fiable una respuesta a la terapia con irinotecan, recurrencia de la enfermedad tardía y muerte relacionada con la enfermedad, independientemente del tamaño del tumor, calidad y estadio de Duke. El ensayo detectó de forma fiable noventa y dos por ciento (92%) de no respondedores a una terapia con irinotecan (tasa de error o de clasificación errónea menor que 1,5%). Para
- 10 determinar la capacidad de respuesta a irinotecan, se determinó que la tasa de sensibilidad al ensayo era de aproximadamente noventa y seis por ciento (96%) y que la tasa de especificidad del ensayo era de aproximadamente noventa y ocho por ciento (98%). “Positividad de GPEP” significa que en las muestras de tumor procedentes de pacientes quienes eran respondedores a la terapia con irinotecan, las siguientes proteínas estaban supra-reguladas: ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD-68 y BAG1, y los siguientes genes y proteínas codificadas estaban sub-regulados: quinasa Erk1, fosfo-GSK-3beta, MMP11, CTSL2, CCNB1, BIRC5, STK6, MRP14 y GSTM1, en comparación con la expresión de estos genes y proteínas en tejido de colon normal
- 15 procedente de estos pacientes y el tejido de colon normal y tejidos de cáncer no de colon de otros pacientes. En todos los tejidos estaban supra-reguladas las proteínas de referencia ACTB, GAPD, GUSB, RPLP0 y TFRC.
- 20 Los resultados de este estudio se ilustran en la Figura 1. La Figura 1 es una gráfica que muestra las tasas de supervivencia de pacientes con cáncer colorrectal tratados con irinotecan representada frente a la presencia de un GPEP de la invención. Tal como se muestra en la Figura 1, pacientes con tumores con un perfil de expresión de genes en los que al menos dieciséis de los genes en el presente GPEP eran diferencialmente expresados y tenían las tasas de supervivencia más prolongadas después de tratamiento con irinotecan. Pacientes cuyo perfil de
- 25 expresión de genes mostraron que diez o más de los genes en el GPEP tenían las tasas de supervivencia siguientes más prolongadas. Las tasas de supervivencia de pacientes cuyos perfiles de expresión de genes indicaban que cuatro o menos de estos genes estaban diferencialmente expresados tenían las tasas de supervivencia menores después de la terapia con irinotecan.
- 30 Los veinte genes señalados en la leyenda a la Figura 1 incluyen cinco proteínas de referencia.

Estudios de validación

- 35 Se realizaron estudios utilizando tejidos de biopsia de cáncer de colon adicionales (procedentes de pacientes distintos a los utilizados en el estudio arriba descrito) para validar adicionalmente la utilidad del GPEP de la presente invención para predecir una capacidad de respuesta del paciente a una terapia con irinotecan. En un estudio de este tipo, En este estudio, se evaluaron probetas de cáncer de colon primario embebidas con parafina fijada con formalina de 220 pacientes, seguido de un mínimo de 120 meses para el tamaño primario del tumor, la calidad histológica, el estadio de Duke y la expresión de las proteínas en el presente GPEP. Ninguno de estos
- 40 pacientes recibió un tratamiento con adyuvantes antes del primer episodio de la recurrencia de la enfermedad. El estudio se llevó a cabo utilizando la misma metodología IHC que la descrita en el Ejemplo precedente y utilizando las mismas disposiciones ordenadas de controles negativos y positivos.

- 45 El análisis univariable de perfiles GPEP de las muestras de pacientes descrito anteriormente predecía de forma fiable la respuesta a una terapia con irinotecan ($p < 0,0001$); recurrencia tardía (LRec, $p < 0,0005$); y muerte relacionada con la enfermedad (DRD, $p < 0,0001$). Cuando se estratifica en grupos GPEP-negativos, límite-positivos y altamente positivos, pacientes con tumores altamente positivos tenían un riesgo relativo (riesgo relativo ajustado) de respuesta a la terapia con irinotecan de 6,7 (intervalo 2,1 – 22,4); LRec de 3,3 (intervalo 1,7 – 9,0) y DRD de 7,0 (intervalo 3,0 – 30,0).
- 50

- Los resultados de este estudio validan, además, que en pacientes de cáncer de colon en fase avanzada, la positividad de GPEP predecía de forma fiable una respuesta a la terapia con irinotecan, recurrencia de la enfermedad tardía y muerte relacionada con la enfermedad, independientemente del tamaño del tumor, calidad y
- 55 estadio de Duke. Para determinar la capacidad de respuesta a irinotecan, se determinó que la tasa de sensibilidad al ensayo era de aproximadamente noventa y seis por ciento (96%) y que la tasa de especificidad del ensayo era de aproximadamente noventa y ocho por ciento (98%).

REIVINDICACIONES

- 5 1. Un ensayo para determinar si un paciente es un respondedor a tratamiento con irinotecan, que comprende medios para determinar el nivel de expresión en una célula tumoral o tejido tumoral de cada una de las proteínas codificadas por los genes seleccionados del grupo que consiste en: ERBB2, que codifica la secuencia de aminoácidos de SEQ ID N° 23, GRB7, que codifica la secuencia de aminoácidos de SEQ ID N° 24, quinasa JNK1, que codifica la secuencia de aminoácidos de SEQ ID N° 26, BCL2, que codifica la secuencia de aminoácidos de SEQ ID N° 27, MK167, que codifica la secuencia de aminoácidos de SEQ ID N° 33, fosfo-Akt, que codifica la secuencia de aminoácidos de SEQ ID N° 35, CD68, que codifica la secuencia de aminoácidos de SEQ ID N° 36 y
- 10 BAG1, que codifica la secuencia de aminoácidos de SEQ ID N° 37, en donde los medios para determinar los niveles de expresión comprenden al menos un anticuerpo específico para cada una de las proteínas y que tiene un marcador detectable.
- 15 2. El ensayo de la reivindicación 1, que comprende, además, medios para detectar el nivel de expresión de la menos una proteína de referencia.
3. El ensayo de la reivindicación 2, en donde la proteína de referencia es codificada por el gen GAPD y tiene la secuencia de aminoácidos de SEQ ID N° 40.
- 20 4. El ensayo de la reivindicación 1, en donde los medios para determinar los niveles de expresión se seleccionan del grupo que consiste en: método de ensayo inmunohistoquímico y ensayo de transferencia Western.
- 25 5. El ensayo de la reivindicación 1, en donde el paciente es un respondedor a tratamiento con irinotecan si el nivel de cada una de las proteínas codificadas por los genes ERBB2, GRB7, quinasa JNK1, BCL2, MK167, fosfo-Akt, CD68 y BAG1 está elevado en el tejido o las células en comparación con los niveles correspondientes en tejido o células de colon normales.
6. Un kit que comprende el ensayo de la reivindicación 4.

Supervivencia mediante Tratamiento con Irinotecan de Cáncer de Colon en Estadío D de Duke

