



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 357 658**

51 Int. Cl.:
C12Q 1/68 (2006.01)
C12N 15/12 (2006.01)
C07K 14/47 (2006.01)
C12N 15/11 (2006.01)
A61P 35/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05768832 .7**
96 Fecha de presentación : **29.07.2005**
97 Número de publicación de la solicitud: **1786925**
97 Fecha de publicación de la solicitud: **23.05.2007**

54 Título: **Genes y polipéptidos relacionados con cánceres de mama.**

30 Prioridad: **10.08.2004 US 600146 P**

45 Fecha de publicación de la mención BOPI:
28.04.2011

45 Fecha de la publicación del folleto de la patente:
28.04.2011

73 Titular/es: **ONCOTHERAPY SCIENCE, Inc.**
2-1, Sakado 3-chome
Takatsu-ku, Kawasaki-shi
Kanagawa 213-0012, JP

72 Inventor/es: **Nakamura, Yusuke;**
Katagiri, Toyomasa y
Nakatsuru, Shuichi

74 Agente: **Arias Sanz, Juan**

ES 2 357 658 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).

DESCRIPCION

Genes y polipéptidos relacionados con cánceres de mama.

CAMPO DE LA INVENCION

La presente invención se refiere al campo de la ciencia biológica, más específicamente al campo de la investigación del cáncer. En particular, se dan a conocer genes novedosos *B1194* y *A2282*, implicados en el mecanismo de proliferación del cáncer de mama, así como polipéptidos codificados por los genes. Los genes y polipéptidos dados a conocer pueden usarse, por ejemplo, en el diagnóstico del cáncer de mama, y como moléculas diana para desarrollar fármacos contra el cáncer de mama. Por tanto, la presente invención se refiere a las realizaciones caracterizadas en las reivindicaciones.

ANTECEDENTES DE LA INVENCION

El cáncer de mama, una enfermedad genéticamente heterogénea, es el tumor maligno más común en mujeres. Cada año se notifica una estimación de aproximadamente 800.000 nuevos casos a nivel mundial (Parkin DM, *et al.*, (1999). *CA Cancer J Clin* 49:33-64). La mastectomía es aún la primera opción concurrente para el tratamiento médico. A pesar de la extracción quirúrgica de los tumores primarios, puede producirse recidiva en sitios localizados o distantes debido a micrometástasis indetectable en el momento del diagnóstico (Saphner T, *et al.*, (1996). *J Clin Oncol*, 14, 2738-2746). Habitualmente se administran agentes citotóxicos como terapia auxiliar tras la cirugía, lo que tiene como objetivo destruir las células residuales o premalignas. El tratamiento con agentes quimioterápicos convencionales es a menudo empírico y mayormente se basa en parámetros tumorales histológicos, y, en ausencia de un entendimiento mecanístico específico, los fármacos dirigidos a una diana están convirtiéndose por tanto en el tratamiento base para el cáncer de mama. El tamoxifeno y los inhibidores de la aromataasa, dos representantes de su clase, han demostrado tener grandes respuestas cuando se usan como adyuvante o quimioprevención en pacientes con cáncer de mama con metástasis (Fisher B, *et al.*, (1998). *J Natl Cancer Inst*, 90, 1371-1388; Cuzick J (2002). *Lancet* 360, 817-824). Sin embargo, la desventaja es que sólo los receptores de estrógenos expresados en los pacientes son sensibles a estos fármacos. Recientemente, incluso surgieron preocupaciones en cuanto a sus efectos secundarios, por ejemplo cáncer de endometrio que resulta de tratamiento con tamoxifeno a largo plazo y fracturas óseas que resultan de la terapia con aromataasa en las mujeres posmenopáusicas (Coleman RE (2004). *Oncology*. 18 (5 Supl. 3), 16-20).

A pesar del avance reciente en las estrategias de diagnóstico y terapéuticas, el pronóstico de los pacientes con cánceres avanzados sigue siendo muy pobre. Aunque estudios moleculares han revelado la implicación de alteraciones en genes supresores tumorales y/u oncogenes en la carcinogénesis, el mecanismo preciso queda aún por esclarecer.

Tecnologías de micromatrices de ADNc han permitido la construcción de perfiles completos de expresión génica en células normales y malignas, y la comparación de la expresión génica en células malignas y normales correspondientes (Okabe *et al.*, *Cancer Res* 61:2129-37 (2001); Kitahara *et al.*, *Cancer Res* 61: 3544-9 (2001); Lin *et al.*, *Oncogene* 21: 4120-8 (2002); Hasegawa *et al.*, *Cancer Res* 62:7012-7 (2002)). Este enfoque facilita el entendimiento de la naturaleza compleja de las células cancerosas, y ayuda a esclarecer el mecanismo de la carcinogénesis. La identificación de los genes que se desregulan en tumores puede conducir a un diagnóstico más preciso y exacto de cánceres individuales, y al desarrollo de dianas terapéuticas novedosas (Bienz y Clevers, *Cell* 103:311-20 (2000)). Para dar a conocer los mecanismos que subyacen a los tumores desde un punto de vista para todo el genoma, y descubrir moléculas diana para el diagnóstico y desarrollo de fármacos terapéuticos novedosos, los presentes inventores han analizado los perfiles de expresión de células tumorales usando una micromatriz de ADNc de 23.040 genes (Okabe *et al.*, *Cancer Res* 61:2129-37 (2001); Kitahara *et al.*, *Cancer Res* 61:3544-9 (2001); Lin *et al.*, *Oncogene* 21:4120-8 (2002); Hasegawa *et al.*, *Cancer Res* 62:7012-7 (2002)).

Estudios diseñados para revelar los mecanismos de la carcinogénesis ya han facilitado la identificación de dianas moleculares para agentes antitumorales. Por ejemplo, se ha demostrado que los inhibidores de la farnesiltransferasa (FTI), que se desarrollaron originalmente para inhibir la ruta de señalización del crecimiento relacionada con Ras y cuya activación depende de la farnesilación postraduccional, son eficaces en el tratamiento de tumores dependientes de Ras en modelos con animales (Sun J, *et al.*, *Oncogene*. 1998; 16:1467-73.). Ensayos clínicos en seres humanos, usando una combinación de fármacos anticancerígenos y el anticuerpo monoclonal anti-HER2, trastuzumab, para antagonizar el receptor del protooncogén HER2/neu, han estado logrando una respuesta clínica mejorada y la supervivencia global de los pacientes con cáncer de mama (Molina MA, *et al.*, *Cancer Res*. 2001; 61:4744-9.). Se ha desarrollado un inhibidor de tirosina cinasa, STI-571, que inactiva selectivamente proteínas de fusión bcr-abl, para tratar leucemias mielógenas crónicas en las que la activación constitutiva de la tirosina cinasa de bcr-abl desempeña un papel crucial en la transformación de leucocitos. Se diseñan agentes de estas clases para suprimir la actividad oncogénica de productos génicos específicos (O'Dwyer ME & Druker BJ, *Curr Opin Oncol*. 2000; 12:594-7.). Por tanto, los productos génicos comúnmente regulados por incremento en células cancerosas pueden servir como dianas potenciales para desarrollar agentes anticancerígenos novedosos.

Se ha demostrado que los linfocitos T citotóxicos CD8⁺ (CTL) reconocen péptidos de epítomos derivados de antígenos asociados a tumores (TAA) presentados en la molécula de clase I del CMH, y células tumorales de lisis. Desde el descubrimiento de la familia MAGE como primer ejemplo de TAA, se han descubierto muchos otros TAA

usando enfoques inmunológicos (Boon, *Int J Cancer* 54: 177-80 (1993); Boon y van der Bruggen, *J Exp Med* 183: 725-9 (1996); van der Bruggen *et al.*, *Science* 254: 1643-7 (1991); Richard *et al.*, *J Exp Med* 178: 489-95 (1993); Kawakami *et al.*, *J Exp Med* 180: 347-52 (1994)). Algunos de los TAA descubiertos se encuentran ahora en la fase de desarrollo clínico como dianas de inmunoterapia. Los TAA descubiertos hasta la fecha incluyen MAGE (van der Bruggen *et al.*, *Science* 254: 1643-7 (1991)), gp100 (Kawakami *et al.*, *J Exp Med* 180: 347-52 (1994)), SART (Shichijo *et al.*, *J Exp Med* 187: 277-88 (1998)), y NY-ESO-1 (Chen *et al.*, *Proc Natl Acad Sci USA* 94: 1914-8 (1997)). Por otra parte, se ha mostrado que productos génicos, que se había demostrado que se sobreexpresaban específicamente en células tumorales, se reconocen como dianas que inducen respuestas inmunitarias celulares. Tales productos génicos incluyen p53 (Umano *et al.*, *Brit J Cancer* 84: 1052-7 (2001)), HER2/neu (Tanaka *et al.*, *Brit J Cancer* 84: 94-9 (2001)), CEA (Nukaya *et al.*, *Int J Cancer* 80: 92-7 (1999)), y similares.

A pesar del avance significativo en la investigación básica y clínica sobre los TAA (Rosenberg *et al.*, *Nature Med* 4: 321-7 (1998); Mukherji *et al.*, *Proc Natl Acad Sci USA* 92: 8078-82 (1995); Hu *et al.*, *Cancer Res* 56: 2479-83 (1996)), sólo un número limitado de TAA candidatos para el tratamiento de adenocarcinomas, incluyendo cáncer de mama, se encuentran actualmente disponibles. Los TAA expresados abundantemente en células cancerosas, y al mismo tiempo cuya expresión se limita a células cancerosas, serían candidatos prometedores como dianas inmunoterapéuticas. Además, se espera que la identificación de nuevos TAA que inducen respuestas inmunitarias antitumorales potentes y específicas fomente el uso clínico de estrategias de vacunación con péptidos en diversos tipos de cáncer (Boon y van der Bruggen, *J Exp Med* 183: 725-9 (1996); van der Bruggen *et al.*, *Science* 254: 1643-7 (1991); Brichard *et al.*, *J Exp Med* 178: 489-95 (1993); Kawakami *et al.*, *J Exp Med* 180: 347-52 (1994); Shichijo *et al.*, *J Exp Med* 187: 277-88 (1998); Chen *et al.*, *Proc Natl Acad Sci USA* 94: 1914-8 (1997); Harris, *J Natl Cancer Inst* 88: 1442-55 (1996); Butterfield *et al.*, *Cancer Res* 59: 3134-42 (1999); Vissers *et al.*, *Cancer Res* 59: 5554-9 (1999); van der Burg *et al.*, *J Immunol* 156: 3308-14 (1996); Tanaka *et al.*, *Cancer Res* 57: 4465-8 (1997); Fujie *et al.*, *Int J Cancer* 80: 169-72 (1999); Kikuchi *et al.*, *Int J Cancer* 81: 459-66 (1999); Oiso *et al.*, *Int J Cancer* 81: 387-94 (1999).

Se ha notificado repetidamente que las células mononucleares de sangre periférica (PBMC) estimuladas con péptido de determinados donantes sanos producen niveles significativos de IFN- γ en respuesta al péptido, pero pocas veces ejercen citotoxicidad contra células tumorales de una manera limitada para HLA-A24 o -A0201 en ensayos de liberación de ^{51}Cr (Kawano *et al.*, *Cancer Res* 60: 3550-8 (2000); Nishizaka *et al.*, *Cancer Res* 60: 4830-7 (2000); Tamura *et al.*, *Jpn J Cancer Res* 92: 762-7 (2001)). Sin embargo, tanto HLA-A24 como HLA-A0201 son alelos de HLA populares en las poblaciones japonesa, así como en las caucásicas (Date *et al.*, *Tissue Antigens* 47: 93-101 (1996); Kondo *et al.*, *J Immunol* 155: 4307-12 (1995); Kubo *et al.*, *J Immunol* 152: 3913-24 (1994); Imanishi *et al.*, *Proceeding of the eleventh International Histocompatibility Workshop and Conference Oxford University Press, Oxford, 1065 (1992); Williams et al.*, *Tissue Antigen* 49: 129 (1997)). Por tanto, los péptidos antigénicos de cánceres presentados por estos HLA pueden ser especialmente útiles para el tratamiento de cánceres entre las poblaciones japonesa y caucásicas. Además, se conoce que la inducción de CTL de baja afinidad *in vitro* habitualmente resulta del uso de un péptido en una alta concentración, lo que genera un alto nivel de complejos péptido/CMH específicos en células presentadoras de antígenos (APC), que activará eficazmente estos CTL (Alexander-Miller *et al.*, *Proc Natl Acad Sci USA* 93: 4102-7 (1996)).

Anteriormente, mediante el análisis del perfil de la expresión génica usando una micromatriz de ADN para todo el genoma, se ha mostrado que un polinucleótido (Gen 2089) que codifica para un polipéptido de secuencia SEQ ID NO:8 (que corresponde a SEQ ID NO: 4 de la presente solicitud) se regula por incremento en líneas celulares de cáncer de mama entre numerosos genes distintos (documento WO03/065006). Además, se ha mostrado mediante la micromatriz de ADN que la expresión de las secuencias de SEQ ID NO: 111-112 (correspondientes a las secuencias SEQ ID NO: 3 a 4 de la presente solicitud) se potencia en células de cáncer de mama (documento US2002/156263).

Para dar a conocer el mecanismo de la carcinogénesis de la mama e identificar marcadores de diagnóstico y/o dianas farmacológicas novedosos para el tratamiento de estos tumores, los presentes inventores analizaron los perfiles de expresión de genes en la carcinogénesis de mama usando una micromatriz de ADNc para todo el genoma que contenía 27.648 genes. Desde el punto de vista farmacológico, la supresión de señales oncogénicas es más fácil en la práctica que la activación de efectos supresores tumorales. Por tanto, los presentes inventores buscaron genes que se regularan por incremento durante la carcinogénesis de mama.

50 SUMARIO DE LA INVENCION

Por consiguiente, en un esfuerzo para entender los mecanismos carcinogénicos asociados con el cáncer e identificar dianas potenciales para desarrollar agentes anticancerígenos novedosos, se realizaron análisis a gran escala de patrones de expresión génica en poblaciones purificadas de células de cáncer de mama usando una micromatriz de ADNc que representaba 27.648 genes. Más particularmente, para aislar dianas moleculares novedosas para los tratamientos de cáncer de mama, usando una combinación de micromatriz de ADNc y microdissección por rayo láser, se examinaron perfiles de expresión para todo el genoma de 77 tumores de mama, incluyendo 8 carcinomas ductales *in situ* (CDIS) y 69 carcinomas ductales invasivos (CDI).

Entre los genes regulados por incremento, se identificó B1194, designado proteína hipotética FLJ 10252, que se sobreexpresaba más de dos veces en 24 de los 41 (59%) casos de cáncer de mama para los que se encontraban

5 disponibles datos de expresión, especialmente en 20 de los 36 (56%) casos con muestras de cáncer de mama de tipo bien diferenciado. También se identificó A2282, designado la cinasa de cremallera de leucina embrionaria materna (MELK), que se sobreexpresaba más de tres veces en 25 de los 33 (76%) casos de cáncer de mama para los que se encontraban disponibles datos de expresión, especialmente en 10 de los 14 (71%) casos con muestras de cáncer de mama de tipo moderadamente diferenciado. La RT-PCR semicuantitativa posterior confirmó que B1194 y A2282 se regulaban por incremento en muestras de cáncer de mama clínicas y líneas celulares de cáncer de mama en comparación con tejidos humanos normales, incluyendo células ductales de mama o mama normal. El análisis por transferencia de tipo Northern también reveló que el transcrito de aproximadamente 2,4 kb de B1194 y A2282 se expresaban exclusivamente en líneas celulares de de cáncer de mama (B1194 y A2282) y testículos (B1194). La tinción inmunocitoquímica indicó que B1194 exógeno se localizaba en el aparato del núcleo en células COS7. El tratamiento de células de cáncer de mama con ARN de interferencia pequeños (ARNip) inhibía eficazmente la expresión de B1194 y A2282, y suprimía el crecimiento celular/tumoral de la línea celular de cáncer de mama T47D y/o MCF7, lo que sugiere que estos genes desempeñan un papel clave en la proliferación del crecimiento celular. Estos hallazgos sugieren que la sobreexpresión de B1194 y A2282 puede estar implicada en la tumorigénesis de mama y puede proporcionar estrategias prometedoras para el tratamiento específico de pacientes con cáncer de mama.

10 Por tanto, se aislaron genes novedosos B1194 y A2282 que se sobreexpresaban significativamente en células de cáncer de mama, y se confirmó mediante RT-PCR semicuantitativa y análisis por transferencia de tipo Northern que el patrón de expresión de B1194 y aquél entre las variantes de A2282, V1, V2 y V3, se sobreexpresaban específicamente en células de cáncer de mama. Se notificó anteriormente que las EST de tanto B1194 como A2282 se regulaban por incremento en un cáncer de pulmón de células no pequeñas (documento WO 2004/031413). Sin embargo, anteriormente se desconocía la relación de estos genes con el cáncer de mama. Además, se proporciona la secuencia de nucleótidos de longitud completa de estos genes.

15 Por consiguiente, se dan a conocer proteínas novedosas implicadas en el mecanismo de proliferación de células de cáncer de mama y los genes que codifican para tales proteínas, así como métodos para producir y usar las mismas en el diagnóstico y usar las mismas en el diagnóstico y tratamiento del cáncer de mama.

20 Entre los transcritos que comúnmente se regulaban por incremento en cánceres de mama, se identificaron los genes humanos novedosos *B1194*, *A2282V1*, *A2282V2* y *A2282V3* en la banda de cromosoma 1q41 y 9p13.1, respectivamente. La transferencia génica de *B1194*, *A2282V1*, *A2282V2* y *A2282V3* promovía la proliferación de células. Además, la reducción de la expresión de *B1194*, *A2282V1*, *A2282V2* y *A2282V3* mediante la transfección de sus S-oligonucleótidos antisentido específicos o ARN de interferencia pequeños inhibía el crecimiento de las células de cáncer de mama. Muchos fármacos anticancerígenos, tales como inhibidores de la síntesis de ADN y/o ARN, supresores metabólicos, e intercaladores de ADN, no sólo son tóxicos para las células cancerosas sino también para células en crecimiento normal. Sin embargo, los agentes que suprimen la expresión de B1194 pueden no afectar de manera adversa a otros órganos debido al hecho de que la expresión normal del gen se limita a los testículos, y por tanto puede ser de gran importancia para tratar el cáncer.

30 Por tanto, se proporcionan genes aislados, *B1194*, *A2282V1*, *A2282V2* y *A2282V3*, que son candidatos como marcadores de diagnóstico así como dianas potenciales prometedoras para desarrollar nuevas estrategias para el diagnóstico y agentes anticancerígenos eficaces. Además, la presente invención proporciona polipéptidos codificados por estos genes, así como la producción y el uso de los mismos. Más específicamente, la presente invención proporciona lo siguiente:

35 La presente descripción también proporciona polipéptidos humanos novedosos, B1194, A2282V1, A2282V2 y A2282V3, o equivalentes funcionales de los mismos, que promueven la proliferación celular y se regulan por incremento en enfermedades proliferativas células, tales como cáncer de mama.

40 En una realización preferida, el polipéptido B1194 incluye una proteína de 528 aminoácidos putativa. B1194 se codifica por el marco de lectura abierto de SEQ ID NO: 1. El polipéptido B1194 incluye preferiblemente la secuencia de aminoácidos expuesta en SEQ ID NO: 2. La presente descripción también proporciona una proteína aislada codificada a partir de al menos una parte de la secuencia de polinucleótido de *B1194*, o secuencias de polinucleótido de al menos el 30%, y más preferiblemente al menos el 40% complementarias a la secuencia expuesta en SEQ ID NO: 1.

45 En una realización preferida, los polipéptidos A2282V1, A2282V2 y A2282V3 incluyen una proteína de 651, 619 y 580 aminoácidos putativa codificada por el marco de lectura abierto de SEQ ID NO: 3, 5 y 7, respectivamente. Los polipéptidos A2282V1, A2282V2 y A2282V3 incluyen preferiblemente las secuencias de aminoácidos expuestas en SEQ ID NO: 4, 6 y 8, respectivamente. La presente descripción también proporciona una proteína aislada codificada a partir de al menos una parte de las secuencias de polinucleótido de *A2282V1*, *A2282V2*, y *A2282V3*, o secuencias de polinucleótido de al menos el 15%, y más preferiblemente al menos el 25% complementarias a la secuencia expuesta en SEQ ID NO: 3, 5 y 7, respectivamente.

50 La presente descripción proporciona además genes humanos novedosos, *B1194*, *A2282V1*, *A2282V2* y *A2282V3*, cuyas expresiones se elevan notablemente en una gran mayoría de cánceres de mama en comparación con los tejidos no cancerosos correspondientes. El gen *B1194* aislado incluye una secuencia de polinucleótido tal como se describe en SEQ ID NO: 1. En particular, el ADNc de *B1194* incluye 2338 nucleótidos que contienen un marco de

lectura abierto de 1587 nucleótidos (SEQ ID NO: 1). La presente descripción abarca además polinucleótidos que se hibridan con y que son al menos el 30%, y más preferiblemente al menos el 40% complementarios a la secuencia de polinucleótido expuesta en SEQ ID NO: 1, en la medida de que codifican para una proteína B1194 o un equivalente funcional de la misma. Ejemplos de tales polinucleótidos son mutantes alélicos y degenerados de SEQ ID NO: 1. Por otra parte, los genes *A2282V1*, *A2282V2*, y *A2282V3* aislados incluyen una secuencia de polinucleótido tal como se describe en SEQ ID NO: 3, 5 y 7, respectivamente. En particular, los ADNc de *A2282V1*, *A2282V2*, y *A2282V3* incluyen 2501, 2368 y 2251 nucleótidos que contienen un marco de lectura abierto de 1956, 1860 y 1743 nucleótidos, respectivamente (SEQ ID NO: 3, 5 y 7, respectivamente). La presente descripción abarca además polinucleótidos que se hibridan con y que son al menos el 15%, y más preferiblemente al menos el 25% complementarios a las secuencias de polinucleótido expuestas en SEQ ID NO: 3, 5 y 7, respectivamente, en la medida de que codifican una proteína *A2282V1*, *A2282V2* o *A2282V3* o un equivalente funcional de la misma. Ejemplos de tales polinucleótidos son mutantes alélicos y degenerados de SEQ ID NO: 3, 5 y 7.

Tal como se usa en el presente documento, un gen aislado es un polinucleótido cuya estructura no es idéntica a la de ningún polinucleótido que se produce de manera natural o a la de ningún fragmento de un polinucleótido genómico que se produce de manera natural que abarca más de tres genes separados. El término, por tanto incluye, por ejemplo, (a) un ADN que tiene la secuencia de parte de una molécula de ADN genómico que se produce de manera natural en el genoma del organismo en el que ésta se produce de manera natural; (b) un polinucleótido incorporado en un vector o en el ADN genómico de un procarionte o eucariote de una manera tal que la molécula resultante no es idéntica a ningún vector o ADN genómico que se produce de manera natural; (c) una molécula separada, tal como un ADNc, un fragmento genómico, un fragmento producido mediante reacción en cadena de la polimerasa (PCR), o un fragmento de restricción; y (d) una secuencia de nucleótidos recombinante que es parte de un gen híbrido, es decir, un gen que codifica para un polipéptido de fusión.

Por consiguiente, en un aspecto, la descripción proporciona un polinucleótido aislado que codifica para un polipéptido descrito en el presente documento o un fragmento del mismo. Preferiblemente, el polipéptido aislado incluye una secuencia de nucleótidos que es al menos el 60% idéntica a la secuencia de nucleótidos mostrada en SEQ ID NO: 1, 3, 5 ó 7. Más preferiblemente, la molécula de ácido nucleico aislada es al menos el 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, idéntica a la secuencia de nucleótidos mostrada en SEQ ID NO: 1, 3, 5 ó 7. En el caso de un polinucleótido aislado que es más largo que o equivalente en longitud a la secuencia de referencia, por ejemplo, SEQ ID NO: 1, 3, 5 ó 7, se realiza la comparación con la longitud completa de la secuencia de referencia. Cuando el polinucleótido aislado es más corto que la secuencia de referencia, por ejemplo, más corto que la SEQ ID NO: 1, 3, 5 ó 7, se realiza la comparación con el segmento de la secuencia de referencia de la misma longitud (excluyendo cualquier bucle requerido por el cálculo de homología).

La presente descripción también proporciona un método de producción de una proteína transfectando o transformando una célula huésped con una secuencia de polinucleótido que codifica para una proteína B1194, *A2282V1*, *A2282V2* o *A2282V3*, y que expresa la secuencia de polinucleótido. Además, la presente descripción proporciona vectores que comprenden una secuencia de nucleótidos que codifica para una proteína B1194, *A2282V1*, *A2282V2* o *A2282V3*, y células huésped que albergan un polinucleótido que codifica para una proteína B1194, *A2282V1*, *A2282V2* o *A2282V3*. Tales vectores y células huésped pueden usarse para producir las proteínas B1194, *A2282V1*, *A2282V2* y *A2282V3*.

También se proporciona un anticuerpo que reconoce una proteína B1194, *A2282V1*, *A2282V2* o *A2282V3*. En parte, también se proporciona un polinucleótido antisentido (por ejemplo, ADN antisentido), ribozima, y ARNip (ARN de interferencia pequeño) del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3*.

La presente descripción proporciona además un método para el diagnóstico del cáncer de mama que incluye la etapa de determinar un nivel de expresión de un gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en una muestra biológica de espécimen y comparar el nivel de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* con el de una muestra normal, siendo un alto nivel de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en la muestra indicativo de cáncer de mama.

Además, se proporciona un método de selección de un compuesto útil en el tratamiento del cáncer de mama. El método incluye la etapa de poner en contacto un polipéptido B1194, *A2282V1*, *A2282V2* o *A2282V3* con compuestos de prueba, y seleccionar compuestos de prueba que se unen al polipéptido B1194, *A2282V1*, *A2282V2* o *A2282V3*.

La presente invención proporciona además un método de selección de un útil en el tratamiento del cáncer de mama, incluyendo el método la etapa de poner en contacto un polipéptido B1194, *A2282V1*, *A2282V2* o *A2282V3* con un compuesto de prueba, y seleccionar el compuesto de prueba que suprime la actividad biológica del polipéptido B1194, *A2282V1*, *A2282V2* o *A2282V3*.

La presente descripción también proporciona una composición farmacéutica útil en el tratamiento del cáncer de mama. La composición farmacéutica puede ser, por ejemplo, un agente anticancerígeno. La composición farmacéutica puede describirse como al menos una parte de los S-oligonucleótidos antisentido o ARNip de la secuencia de polinucleótido de *B1194*, *A2282V1*, *A2282V2* o *A2282V3* mostrada y descrita en SEQ ID NO: 1, 3, 5 ó 7 respectivamente. Ejemplos de secuencias diana adecuadas de ARNip incluyen las secuencias de nucleótidos expuestas

5 en SEQ ID NO: 38, 39, 40 y 41. La secuencia diana de ARNip para B1194, incluyendo aquellas que tienen la secuencia de nucleótidos de SEQ ID NO: 38 ó 39, puede usarse adecuadamente para tratar el cáncer de mama; la secuencia diana de ARNip para A2282V1, A2282V2 o A2282V3, incluyendo aquellas que tienen la secuencia de nucleótidos de SEQ ID NO: 40 ó 41, también pueden usarse adecuadamente para tratar el cáncer de mama. Las composiciones farmacéuticas de la presente invención también incluyen aquellos compuestos seleccionados mediante los presentes métodos de selección de compuestos útiles en el tratamiento del cáncer de mama.

El transcurso de acción de la composición farmacéutica es de manera deseable inhibir el crecimiento de las células cancerosas. La composición farmacéutica puede aplicarse a mamíferos incluyendo seres humanos y mamíferos domesticados.

10 La presente invención proporciona además métodos para tratar el cáncer de mama usando la composición farmacéutica proporcionada por la presente invención.

15 Además, la presente descripción proporciona un método para tratar o prevenir el cáncer de mama que comprende la etapa de administrar un polipéptido B1194, A2282V1, A2282V2 o A2282V3. Se espera que la inmunidad antitumoral se indujera mediante la administración de un polipéptido B1194, A2282V1, A2282V2 o A2282V3 de este tipo. Por tanto, la presente descripción proporciona también un método para inducir inmunidad antitumoral que comprende la etapa de administrar el polipéptido B1194, A2282V1, A2282V2 o A2282V3, así como composiciones farmacéuticas para tratar o prevenir el cáncer de mama que comprenden el polipéptido B1194, A2282V1, A2282V2 o A2282V3.

20 Éstos y otros objetos y características de la descripción o la invención resultarán más completamente evidentes cuando se lea la siguiente descripción detallada junto con las figuras y ejemplos adjuntos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

25 La figura 1 está compuesta por una serie de fotografías que muestran los resultados de RT-PCR para validar la sobreexpresión de genes B1194. Las líneas 1 a 12 de la figura 1(a) son muestras clínicas sometidas a una ronda de amplificación basada en T7 antes de la transcripción inversa, y las líneas 1 a 9 de la figura 1(b) son líneas celulares de cáncer de mama. La abreviatura "pre" representa una mezcla de células ductales de mama normales y se usa en el presente documento como control universal en el experimento con micromatrices.

La figura 2 está compuesta por una serie de fotografías que muestran los resultados del análisis por transferencia de tipo Northern, particularmente el patrón de expresión de B1194 en (a) múltiples tejidos normales, (b) líneas celulares de cáncer de mama, respectivamente. "#" indica órganos vitales.

30 La figura 3(a) está compuesta por una serie de fotografías que representan la localización subcelular de la proteína B1194 en células COS7. DAPI (núcleo); B1194 (FITC) y la unión de los mismos se demuestra en la fotografía n.º 1, 2 y 3 respectivamente. La figura 3(b) es una fotografía de análisis por inmunotransferencia de tipo Western de la proteína B1194.

35 La figura 4(a) está compuesta por una serie de fotografías que representan los resultados de la RT-PCR semicuantitativa, que muestra particularmente el efecto de desactivación de B1194 endógeno en líneas celulares T47D. La figura 4(b) es un diagrama de barras que representa los resultados de un ensayo de MTT que muestra baja proliferación en cultivos de si1 y si5.

40 La figura 5 está compuesta por una serie de fotografías que representan los resultados de RT-PCR semicuantitativa, particularmente la expresión de A2282. Las líneas 1 a 12 de la figura 5(a) son muestras clínicas sometidas a una ronda de amplificación basada en T7 antes de la transcripción inversa. Las líneas 1 a 6 de la figura 5 (b) son líneas celulares de cáncer. De nuevo, la abreviatura "pre" representa una mezcla de células ductales de mama y se usa en el presente documento como un control universal para el análisis con micromatrices de ADNc.

45 La figura 6(a) representa la estructura genómica de A2282. Las figuras 6(b) y 6(c) son fotografías que representan los resultados del análisis por transferencia de tipo Northern, particularmente que muestran el patrón de expresión de (b) múltiples tejidos normales, (c) líneas celulares de cáncer de mama y tejidos normales, respectivamente. "#" indica órganos vitales.

50 La figura 7(a) representa la estructura de las variantes de A2282. Se aislaron cinco transcritos de A2282 de la selección de bibliotecas de ADNc. ATG y TAA representan el codón de iniciación y terminación de la traducción, respectivamente. Los bloques en negro y sombreados indican regiones no traducidas y secuencias codificantes. La figura 7(b) es una fotografía que representa los resultados del análisis por transferencia de tipo Northern en líneas celulares de cáncer de mama y tejidos normales.

La figura 8 es una fotografía que muestra la capacidad de traducción de cuatro transcritos de A2282 *in vitro*. El peso molecular de la proteína predicho se indica entre paréntesis al lado de cada variante. N representa el control negativo.

La figura 9(a) es una fotografía que muestra los resultados del análisis por inmunotransferencia de tipo Western, particularmente que muestra la expresión de la proteína A2282. Las abreviaturas “E” y “M” indican la fase sin tratamiento con fármacos (crecimiento exponencial) y la fase mitótica, respectivamente. Se usó β actina (ACTB) como control interno. Se cargaron cantidades iguales de proteína total (10 μ g) en cada línea. La figura 9 (b) está compuesta por una serie de gráficos que muestran los resultados del análisis del ciclo celular. Se examina el efecto de tres transcritos de A2282 en la transición del ciclo celular en fase G1 sincronizada de las células HeLa mediante citometría de flujo. Se usaron células no transfectadas como control.

La figura 10(a) son fotografías de RT-PCR semicuantitativa que muestran el efecto de desactivación de A2282 endógeno en las líneas celulares MCF-7 y T47D. La figura 10(b) está compuesta por una serie de diagramas de barras que representan los resultados de un ensayo de MTT, que muestra particularmente baja proliferación en el cultivo n.º 3 y n.º 4. La figura 10(c) está compuesta por una serie de fotografías que representan los resultados de un ensayo de formación de colonias que demuestra una disminución en la densidad de las colonias en cultivos de desactivación del gen A2282.

La figura 11 demuestra que la proteína A2282 está fosforilada en el dominio cinasa. Específicamente, la figura 11(a) es una representación sistemática de dos proteínas A2282 truncadas y tipo natural. La figura 11 (b) representa los resultados del análisis por inmunotransferencia de tipo Western para los tres transcritos usando anticuerpo anti-HA. La figura 11(c) representa los resultados del ensayo con λ -PPasa que confirmó la fosforilación de la proteína A2282 de tipo natural.

La figura 12 representa los resultados de ensayos de cinasa de inmunocomplejos e inmunoprecipitación. Específicamente, la figura 12 (a) es una representación sistemática de transcritos de tipo natural y mutantes. La figura 12(b) examina el estado de fosforilación de tres mutantes en líneas celulares HEK 293. La figura 12(c) evalúa la actividad cinasa de tres mutantes mediante ensayo de cinasa de inmunocomplejos. Se usó la histona H1 como sustrato *in vitro*.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES PREFERIDAS

1. Visión general

La presente solicitud identifica genes humanos novedosos *B1194*, *A2282V1*, *A2282V2* y *A2282V3* cuya expresión se eleva notablemente en cáncer de mama en comparación con tejidos no cancerosos correspondientes. El ADNc de *B1194* consiste en 2338 nucleótidos que contiene un marco de lectura abierto de 1587 nucleótidos tal como se expone en SEQ ID NO: 1. El marco de lectura abierto codifica para una proteína de 528 aminoácidos putativa. De manera similar, el ADNc de *A2282V1*, *A2282V2* y *A2282V3* incluye 2501, 2368 y 2251 nucleótidos que contienen un marco de lectura abierto de 1956, 1860 y 1743 nucleótidos, respectivamente (SEQ ID NO: 3, 5 y 7, respectivamente). Las secuencias de nucleótidos de SEQ ID NO: 5 y 7 (*A2282V2* y *A2282V3*, respectivamente) se presentaron en el banco de datos de ADN de Japón (DDBJ), y se les asignaron respectivamente los números de registro AB183427 y AB183428. Dado que la expresión de la proteína se reguló por incremento en el cáncer de mama, las proteínas se denominaron *A2282V1*, *A2282V2* y *A2282V3* (reguladas por incremento en el cáncer de mama).

Consecuentemente, la expresión exógena de *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en células confirió un aumento en el crecimiento celular, mientras que la supresión de la expresión con S-oligonucleótidos antisentido o ARN de interferencia pequeño (ARNip) dio como resultado una inhibición del crecimiento significativa de células cancerosas. Estos hallazgos sugieren que *B1194*, *A2282V1*, *A2282V2* y *A2282V3* proporcionan actividades oncogénicas a las células cancerosas, y que la inhibición de la actividad de estas proteínas podría ser una estrategia prometedora para el tratamiento del cáncer.

Más particularmente, en el presente documento se descubrió que *B1194*, designado gen de proteína hipotética FLJ-10252, se regula por incremento significativamente mediante los perfiles de expresión de cáncer de mama. Este hallazgo se confirmó mediante RT-PCR semicuantitativa usando muestras clínicas y análisis por transferencia de tipo Northern. Además, se encontró que la expresión de este gen era específica del cáncer más que un acontecimiento ubicuo. El tratamiento de células de cáncer de mama con ARN de interferencia pequeños (ARNip) inhibió de manera eficaz la expresión de *B1194* y suprimió el crecimiento celular/tumoral de la línea celular de cáncer de mama T47D. Estos hallazgos tomados juntos sugieren que la proteína hipotética FLJ-10252 es un candidato molecular novedoso destacado para el desarrollo de fármacos contra el cáncer de mama.

Además, mediante los perfiles de expresión precisos de cáncer de mama por medio de una micromatriz de ADNc para todo el genoma, se identificó el gen novedoso *A2282* que se sobreexpresaba significativamente en células de cáncer de mama en comparación con tejidos humanos normales. El tratamiento de células de cáncer de mama con ARNip inhibió eficazmente la expresión de *A2282* y suprimió significativamente el crecimiento celular/tumoral del cáncer de mama.

Se seleccionó *A2282*, designado MELK, un nuevo miembro de la familia KIN1/PAR-1/MARK identificado durante el desarrollo de los embriones de xenopus y ratón (Blot J, *et al.*, (2002). *Dev Biol*, 241, 327-338; Heyer BS, *et al.*, (1999). *Dev Dyn*, 215, 344-351), para su estudio debido a su expresión elevada significativa en el cáncer de mama.

Se identificaron cinco variantes del gen de MELK humano, y de entre ellos, los transcritos de aproximadamente 2,4 kb mostraron una expresión específica del cáncer, mientras que otros dos transcritos no pudieron traducirse en casi todos los órganos. El análisis de secuencias reveló deleciones internas en el dominio catalítico en el extremo N-terminal en dos transcritos, que posteriormente se designaron V2 y V3. Estas deleciones provocaron una terminación prematura de la traducción que condujo a la traducción de una proteína más corta con un codón de iniciación alternativo en el mismo marco de lectura, lo que generó codones de iniciación de la traducción novedosos de V2 o V3, y que produjo la parte N-terminal delecionada. Mediante el alineamiento de las secuencias de aminoácidos de las tres variantes, se demostró claramente que V2 y V3 aún retenían el dominio cinasa parcial más no la región transmembrana putativa. Sin embargo, no está claro si esta deleción afecta a la actividad cinética de esta proteína o no.

Con el fin de caracterizar estos transcritos, se examinó el estado de fosforilación de estas variantes a lo largo del ciclo celular. De acuerdo con estudios anteriores (Davezac N, *et al.*, (2002). *Oncogene*, 21, 7630-7641), se mostró que V1 se fosforilaba fuertemente durante la mitosis. Sin embargo, no se observaron ni V2 ni V3 fosforilados en ninguna de las fases del ciclo celular. De manera interesante, la expresión transitoria de estos transcritos en células HeLa sincronizadas tuvo efectos ligeramente diferentes. La expresión de V1 dio como resultado el acortamiento del primer ciclo celular, seguido de una detención en la fase G2/M. Por el contrario, la inducción de V2 y V3 condujo a un primer ciclo celular prolongado en comparación con las células control no transfectadas. A pesar de la discrepancia del resultado inicial, todas las variantes pudieron finalmente detener las células en la fase G2/M. También se han notificado resultados similares (Davezac N, *et al.*, (2002). *Oncogene*, 21, 7630-7641; Vulsteke V, *et al.*, (2004). *J Biol Chem*, 279, 8642-7). Además, la reciente evidencia sugiere que es probable que el dominio C-terminal de MELK sea el inhibidor fuerte de la actividad cinasa de MELK (Vulsteke V, *et al.*, (2004). *J Biol Chem*, 279, 8642-7). Este hallazgo especula que el dominio cinasa de MELK podría contribuir a un acortamiento del ciclo celular en el cáncer de mama, y su actividad podría controlarse estrictamente bajo la regulación negativa de su dominio C-terminal.

En conclusión, estos hallazgos muestran que A2282 es un gen indispensable específico para el cáncer, esencial para el crecimiento de células cancerosas por medio de una ruta de señalización no identificada. Basándose en estos resultados, parece que MELK es una diana molecular prometedora para el tratamiento del cáncer de mama.

II. Definiciones

Las palabras “uno”, “una”, y “el/la/los/las” tal como se usan en el presente documento significan “al menos uno” a menos que se indique específicamente lo contrario.

En el contexto de la presente invención, “inhibición de la unión” entre dos proteínas se refiere a reducir al menos la unión entre las proteínas. Por tanto, en algunos casos, el porcentaje de pares de unión en una muestra estará disminuido en comparación con un control apropiado (por ejemplo, sin tratar con compuesto de prueba o a partir de una muestra sin cáncer, o a partir de una muestra con cáncer). La reducción en la cantidad de proteínas unidas puede ser, por ejemplo, menos del 90%, 80%, 70%, 60%, 50%, 40%, 25%, 10%, 5%, 1% o menos (por ejemplo, el 0%), que los pares unidos en una muestra control.

Una “cantidad farmacéuticamente eficaz” de un compuesto es una cantidad que es suficiente para tratar y/o mejorar el cáncer en un individuo. Un ejemplo de una cantidad farmacéuticamente eficaz puede ser una cantidad necesaria para disminuir la expresión de B1194, A2282V1, A2282V2 o A2282V3 cuando se administra a un animal. La disminución puede ser, por ejemplo, al menos un 5%, 10%, 20%, 30%, 40%, 50%, 75%, 80%, 90%, 95%, 99% o un 100% de cambio en la expresión.

La expresión “vehículo farmacéuticamente aceptable” se refiere a una sustancia inerte usada como diluyente o vehículo para un fármaco.

En la presente descripción, la expresión “funcionalmente equivalente” significa que el polipéptido objeto tiene la actividad de promover la proliferación celular como las proteínas B1194, A2282V1, A2282V2 o A2282V3 y de conferir actividad oncogénica a las células cancerosas. Además, un polipéptido funcionalmente equivalente puede tener la actividad proteína cinasa asociada con las proteínas A2282V1, A2282V2 y A2282V3. Los ensayos para determinar tales actividades se conocen bien en la técnica. Por ejemplo, puede evaluarse si el polipéptido objeto tiene una actividad de proliferación celular o no introduciendo el ADN que codifica para el polipéptido objeto en una célula que expresa el polipéptido respectivo, y detectando la promoción de la proliferación de las células o el aumento en la actividad de formación de colonias. Tales células incluyen, por ejemplo, células COS7 y NIH3T3 para B1194 y A2282V1, A2282V2, A2282V3.

El término “aislado” y la expresión “biológicamente puro” se refieren a un material que está sustancial o esencialmente libre de componentes que normalmente lo acompañan tal como se encuentra en su estado nativo. Sin embargo, el término “aislado” no pretende hacer referencia a los componentes presentes en un gel electroforético u otro medio de separación. Un componente aislado está libre de tales medios de separación y en una forma lista para su uso en otra aplicación o ya en uso en la/el nueva(o) aplicación/medio.

En el contexto de la presente invención, se determina un “porcentaje de identidad de secuencia” comparando dos secuencias alineadas de manera óptima a lo largo de una ventana de comparación, pudiendo comprender la parte

de la secuencia de polinucleótido en la ventana de comparación adiciones o deleciones (es decir, huecos) en comparación con la secuencia de referencia (por ejemplo, un polipéptido de la invención), que no comprende adiciones o deleciones, para el alineamiento óptimo de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las que aparece la base de ácido nucleico o el residuo de aminoácido idénticos en ambas secuencias para dar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para dar el porcentaje de identidad de secuencia.

Los términos “idéntico” o “identidad” en porcentaje, en el contexto de dos o más ácidos nucleicos o secuencias de polipéptido, se refieren a dos o más secuencias o subsecuencias que son las mismas secuencias. Dos secuencias son “sustancialmente idénticas” si las dos secuencias tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son los mismos (es decir, una identidad del 60%, opcionalmente una identidad del 65%, 70%, 75%, 80%, 85%, 90% o del 95% a lo largo de una región especificada, o, cuando no se especifica, a lo largo de toda la secuencia), cuando se comparan o alinean para una correspondencia máxima a lo largo de una ventana de comparación, o región diseñada tal como se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante alineamiento manual e inspección visual. Opcionalmente, la identidad existe a lo largo de una región que es de al menos aproximadamente 50 nucleótidos de longitud, o más preferiblemente a lo largo de una región que es de 100 a 500 ó 1000 o más nucleótidos de longitud.

Para la comparación de secuencias, normalmente una secuencia actúa como secuencia de referencia con la que se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, se introducen las secuencias de prueba y de referencia en un ordenador, se designan coordenadas de subsecuencias, si es necesario, y se designan parámetros del programa de algoritmos de secuencia. Pueden usarse los parámetros del programa por defecto, o pueden designarse parámetros alternativos. Entonces el algoritmo de comparación de secuencias calcula el porcentaje de identidades de secuencia para las secuencias de prueba con respecto a la secuencia de referencia, basándose en los parámetros del programa.

Una “ventana de comparación”, tal como se usa en el presente documento, incluye una referencia de un segmento de una cualquiera del número de posiciones contiguas seleccionadas del grupo que consiste en desde 20 hasta 600, habitualmente de aproximadamente 50 hasta aproximadamente 200, más habitualmente de aproximadamente 100 a aproximadamente 150 en el que puede compararse una secuencia con una secuencia de referencia del mismo número de posiciones contiguas tras alinearse óptimamente las dos secuencias. Los métodos de alineamiento de secuencias para comparación se conocen bien en la técnica. El alineamiento óptimo de las secuencias para su comparación pueden llevarse a cabo, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman (1981) *Adv. Appl. Math.* 2:482-489, mediante el algoritmo de alineamiento por homología de Needleman y Wunsch (1970) *J. Mol. Biol.* 48:443, mediante la búsqueda por el método de similitud de Pearson y Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444, mediante implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineamiento manual e inspección visual (véase, por ejemplo, Ausubel *et al.*, *Current Protocols in Molecular Biology* (suplemento de 1995)).

Dos ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencias y la similitud de secuencias son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1977) *Nuc. Acids Res.* 25:3389-3402, y Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectivamente. El software para realizar análisis por BLAST está disponible al público a través del Centro Nacional para la Información Biotecnológica. Este algoritmo implica en primer lugar identificar pares de secuencia de alta puntuación (HSP) identificando palabras cortas de longitud *W* en la secuencia de consulta, que o bien coinciden o bien satisfacen cierta puntuación umbral de valor positivo *T* cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. *T* se denomina umbral de puntuación de palabra vecina (Altschul *et al.*, citado anteriormente). Estas coincidencias de palabras vecinas iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que las contengan. Las coincidencias de palabras se extienden en ambas direcciones a lo largo de cada secuencia tanto como pueda aumentarse la puntuación de alineamiento acumulativa. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros *M* (puntuación de recompensa para un par de residuos coincidentes; siempre > 0) y *N* (puntuación de penalidad para residuos no coincidentes; siempre < 0). Para las secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las coincidencias de palabras en cada dirección se detiene cuando: la puntuación de alineamiento acumulativa se reduce en la cantidad *X* con respecto a su valor máximo logrado; la puntuación acumulativa tiende a cero o menos, debido a la acumulación de uno o más alineamientos de residuos de puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST *W*, *T* y *X* determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para secuencias de nucleótidos) usa como valores por defecto una longitud de palabra (*W*) de 11, una expectativa (*E*) de 10, *M*=5, *N*=-4 y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa como valores por defecto una longitud de palabra de 3, y una expectativa (*E*) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alineamientos (*B*) de 50, expectativa (*E*) de 10, *M*=5, *N*=-4, y una comparación de ambas cadenas.

El algoritmo de BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-7). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña ($P(N)$), que proporciona una indicación de la probabilidad mediante la que se produciría una coincidencia entre dos secuencias de nucleótidos o aminoácidos al azar. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es menos de aproximadamente 0,2, más preferiblemente menos de aproximadamente 0,01, y lo más preferiblemente menos de aproximadamente de 0,001.

La expresión "oligonucleótidos antisentido" tal como se usan en el presente documento significa, no sólo aquellos en los que los nucleótidos correspondientes a aquellos que constituyen una región especificada de un ADN o ARNm son completamente complementarios, sino también aquellos que tienen un apareamiento erróneo de uno o más nucleótidos, siempre que el ADN o ARNm y el oligonucleótido antisentido pueda hibridarse específicamente con la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7.

Tales polinucleótidos están contenidos como aquellos que tienen, en la "región de secuencia al menos 15 nucleótidos contiguos", una homología de al menos el 70% o superior, preferiblemente al 80% o superior, más preferiblemente el 90% o superior, incluso más preferiblemente el 95% o superior. Puede usarse el algoritmo indicado en el presente documento para determinar la homología. Tales polinucleótidos son útiles como sondas para el aislamiento o la detección del ADN que codifica para el polipéptido de la invención tal como se indica en un ejemplo posterior o como cebador usado para amplificaciones.

Los términos "marcador" y "marcador detectable" se usan en el presente documento para referirse a cualquier composición detectable mediante medios espectroscópicos, fotoquímicos, bioquímicos, inmuoquímicos, eléctricos, ópticos o químicos. Tales marcadores incluyen biotina para teñir con conjugado de estreptavidina marcado, perlas magnéticas (por ejemplo, DYNABEADS™), colorantes fluorescentes (por ejemplo, fluoresceína, rojo Texas, rodamina, proteína fluorescente verde y similares), radiomarcadores (por ejemplo, ^3H , ^{125}I , ^{35}S , ^{14}C , o ^{32}P), enzimas (por ejemplo, peroxidasa del rábano picante, fosfatasa alcalina y otras usadas comúnmente en un ELISA), y marcadores calorimétricos tales como oro coloidal o perlas de vidrio o plástico coloreadas (por ejemplo, poliestireno, polipropileno, látex, etc.). Las patentes que enseñan el uso de tales marcadores incluyen las patentes estadounidenses n.ºs 3.817.837; 3.850.752; 3.939.350; 3.996.345; 4.277.437; 4.275.149; y 4.366.241. Medios de detección de tales marcadores se conocen bien por los expertos en la técnica. Por tanto, por ejemplo, los radiomarcadores pueden detectarse usando una película fotográfica o contadores de centelleo, los marcadores fluorescentes pueden detectarse usando un fotodetector para detectar la luz emitida. Los marcadores enzimáticos se detectan normalmente dotando a la enzima de un sustrato y detectando el producto de reacción producido por la acción de la enzima sobre el sustrato, y los marcadores calorimétricos se detectan visualizando simplemente el marcador coloreado.

El término "anticuerpo" tal como se usa en el presente documento abarca anticuerpos que se producen de manera natural así como anticuerpos que no se producen de manera natural, incluyendo, por ejemplo, anticuerpos de cadena sencilla, anticuerpos quiméricos, bifuncionales y humanizados, así como fragmentos de unión a antígeno de los mismos, (por ejemplo, Fab', F(ab')₂, Fab, Fv y rIgG). Véase también, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL). Véase también, por ejemplo, Kuby, J., Immunology, 3ª Ed., W.H. Freeman & Co., Nueva York (1998). Tales anticuerpos que no se producen de manera natural pueden construirse usando síntesis de péptidos en fase sólida, pueden producirse de manera recombinante o pueden obtenerse, por ejemplo, seleccionando bibliotecas combinatorias que consisten en cadenas pesadas variables y cadenas ligeras variables tal como se describe por Huse *et al.*, Science 246:1275-1281 (1989). Estos y otros métodos de preparación, por ejemplo, de anticuerpos quiméricos, humanizados, con injerto de CDR, de cadena sencilla y bifuncionales se conocen bien por los expertos en la técnica (Winter y Harris, Immunol. Today 14:243-246 (1993); Ward *et al.*, Nature 341:544-546 (1989); Harlow y Lane, Antibodies, 511-52, Cold Spring Harbor Laboratory Publications, Nueva York, 1988; Hilyard *et al.*, Protein Engineering: A practical approach (IRL Press 1992); Borrebaeck, Antibody Engineering, 2ª ed. (Oxford University Press 1995).

El término "anticuerpo" incluye tanto anticuerpos policlonales como monoclonales. El término también incluye formas genéticamente modificadas tales como anticuerpos quiméricos (por ejemplo, anticuerpos murinos humanizados) y anticuerpos heteroconjugados (por ejemplo, anticuerpos biespecíficos). El término también se refiere a fragmentos Fv de cadena simple recombinantes (scFv). El término anticuerpo también incluye moléculas bivalentes o biespecíficas, diacuerpos, triacuerpos y tetracuerpos. Moléculas bivalentes y biespecíficas se describen en, por ejemplo, Kostelny *et al.* (1992) J Immunol 148:1547, Pack y Pluckthun (1992) Biochemistry 31:1579, Holliger *et al.* (1993) Proc Natl Acad Sci U S A. 90:6444, Gruber *et al.* (1994) J Immunol: 5368, Zhu *et al.* (1997) Protein Sci 6:781, Hu *et al.* (1997) Cancer Res. 56:3055, Adams *et al.* (1993) Cancer Res. 53:4026, y McCartney, *et al.* (1995) Protein Eng. 8:301.

Normalmente, un anticuerpo tiene una cadena pesada y ligera. Cada cadena pesada y ligera contiene una región constante y una región variable, (las regiones también se conocen como "dominios"). Las regiones variables de cadena ligera y pesada contienen cuatro regiones de "entramado" interrumpidas por tres regiones hipervariables, también denominadas "regiones determinantes de la complementariedad" o "CDR". Se ha definido la extensión de las regiones de entramado y las CDR. Las secuencias de las regiones de entramado de diferentes cadenas ligeras o pesadas están relativamente conservadas dentro de una especie. La región de entramado de un anticuerpo, es decir,

las regiones de entramado combinadas de las cadenas ligeras y pesadas constituyentes, sirve para ubicar y alinear las CDR en espacios tridimensionales.

Las CDR son principalmente responsables de la unión a un epítipo de un antígeno. Las CDR de cada cadena normalmente se denominan CDR1, CDR2 y CDR3, enumeradas secuencialmente partiendo del extremo N-terminal, y también se identifican normalmente por la cadena en la que se está ubicada la CDR particular. Por tanto, una CDR3 de V_H está ubicada en el dominio variable de la cadena pesada del anticuerpo en el que se encuentra, mientras que una CDR1 de V_L es la CDR1 del dominio variable de la cadena ligera del anticuerpo en el que se encuentra. Las referencias a " V_H " se refieren a la región variable de una cadena pesada de inmunoglobulina de un anticuerpo, incluyendo la cadena pesada de un Fv, scFv o Fab. Las referencias a " V_L " se refieren a la región variable de una cadena ligera de inmunoglobulina, incluyendo la cadena ligera de un Fv, scFv, dsFv o Fab.

La expresión "Fv de cadena sencilla" o "scFv" se refiere a un anticuerpo en el que los dominios variables de la cadena pesada y de la cadena ligera de un anticuerpo de dos cadenas tradicional se han unido para formar una cadena. Normalmente, se inserta un péptido ligador entre las dos cadenas para permitir el plegamiento apropiado y la creación de un sitio de unión activo.

Un "anticuerpo quimérico" es una molécula de inmunoglobulina en la que (a) se altera, sustituye o intercambia la región constante, o parte de la misma, de modo que el sitio de unión a antígeno (región variable) se une a una región constante de una clase, función efectora y/o especie diferente o alterada, o una molécula completamente diferente que confiere nuevas propiedades al anticuerpo quimérico, por ejemplo, una enzima, toxina, hormona, factor de crecimiento, fármaco, etc.; o (b) se altera, sustituye o intercambia la región variable, o una parte de la misma, con una región variable que tiene una especificidad de antígeno diferente o alterada.

Un "anticuerpo humanizado" es una molécula de inmunoglobulina que contiene una secuencia mínima derivada de inmunoglobulina no humana. Los anticuerpos humanizados incluyen inmunoglobulinas humanas (anticuerpo receptor) en las que residuos de una región determinante de la complementariedad (CDR) del receptor se sustituyen por residuos de una CDR de una especie no humana (anticuerpo donante) tal como ratón, rata o conejo que tiene la especificidad, afinidad y capacidad deseadas. En algunos casos, los residuos de entramado de Fv de la inmunoglobulina humana se sustituyen por los residuos no humanos correspondientes. Los anticuerpos humanizados también pueden comprender residuos que se no encuentran ni en el anticuerpo receptor ni en la CDR o las secuencias de entramado importadas. En general, un anticuerpo humanizado comprenderá sustancialmente todos de al menos uno, y normalmente dos, dominios variables, en los que todas o sustancialmente todas las regiones CDR corresponden a las de una inmunoglobulina no humana, y todas o sustancialmente todas las regiones de entramado (FR) son las de una secuencia consenso de inmunoglobulina humana. El anticuerpo humanizado también comprenderá de manera óptima al menos una parte de una región constante de inmunoglobulina (Fc), normalmente la de una inmunoglobulina humana (Jones *et al.*, Nature 321:522-525 (1986); Riechmann *et al.* Nature 332: 323-327 (1988); y Presta, Curr. Op. Struct. Biol. 2:593-596 (1992)). La humanización puede realizarse esencialmente siguiendo el método de Winter y colaboradores (Jones *et al.*, Nature 321:522-525 (1986); Riechmann *et al.*, Nature 332: 323-327 (1988); Verhoeven *et al.*, Science 239:1534-1536 (1988)), sustituyendo las CDR o secuencias de CDR de roedor por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, tales anticuerpos humanizados son anticuerpos quiméricos (patente estadounidense n.º 4.816.567), en los que se ha sustituido sustancialmente menos de un dominio variable humano intacto por la secuencia correspondiente de una especie no humana.

El término "epítipo" y la expresión "determinante antigénico" se refieren a un sitio en un antígeno al que se une un anticuerpo. Los epítopos pueden formarse tanto a partir de aminoácidos contiguos como de aminoácidos no contiguos yuxtapuestos por el plegamiento terciario de una proteína. Los epítopos formados de aminoácidos contiguos normalmente se conservan con la exposición a disolventes desnaturizantes mientras que los epítopos formados por el plegamiento terciario normalmente se pierden con el tratamiento con disolventes desnaturizantes. Un epítipo incluye normalmente al menos 3, y más habitualmente, al menos 5 u 8-10 aminoácidos en una conformación especial única. Los métodos de determinación de la conformación espacial de los epítopos incluyen, por ejemplo, cristalografía de rayos x y resonancia magnética nuclear en 2 dimensiones. Véase, por ejemplo, Epitope Mapping Protocols in Methods in Molecular Biology, Vol. 66, Glenn E. Morris, Ed (1996).

Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente en el presente documento para hacer referencia a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácido es un mimético químico artificial de un aminoácido que se produce de manera natural correspondiente, así como a polímeros de aminoácidos que se producen de manera natural, aquellos que contienen residuos modificados, y polímero de aminoácidos que no se producen de manera natural.

El término "aminoácido" se refiere a aminoácidos que se producen de manera natural y sintéticos, así como análogos de aminoácidos y miméticos de aminoácidos que funcionan de manera similar a los aminoácidos que se producen de manera natural. Aminoácidos que se producen de manera natural son aquellos codificados por el código genético, así como aquellos aminoácidos que se modifican posteriormente, por ejemplo, hidroxiprolina, γ -carboxiglutamato y O-fosfoserina. Análogos de aminoácidos se refiere a compuestos que tienen la misma estructura química básica que un aminoácido que se produce de manera natural, por ejemplo, un carbono α que está unido a un

hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, por ejemplo, homoserina, norleucina, metionina sulfóxido, metionina metilsulfonio. Tales análogos pueden tener grupos R modificados (por ejemplo, norleucina) o estructuras principales peptídicas modificadas, pero conservan la misma estructura química básica que un aminoácido que se produce de manera natural. Miméticos de aminoácidos se refiere a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funciona de manera similar a un aminoácido que se produce de manera natural.

En el presente documento los aminoácidos pueden denominarse mediante sus símbolos de tres letras o mediante los símbolos de una letra comúnmente conocidos recomendados por la Comisión de Nomenclatura Bioquímica IUPAC-IUB. Igualmente, los nucleótidos pueden denominarse mediante sus códigos de una sola letra comúnmente aceptados.

El término "recombinante" cuando se usa con referencia, por ejemplo, a una célula, o ácido nucleico, proteína o vector, indica que la célula, ácido nucleico, proteína o vector, se ha modificado mediante la introducción de un ácido nucleico o proteína heterólogo o la alteración de un ácido nucleico o proteína nativo, o que la célula se deriva de una célula así modificada. Por tanto, por ejemplo, las células recombinantes expresan genes que no se encuentran dentro de la forma nativa (no recombinante) de la célula o expresan genes nativos que de otro modo se expresan de manera anómala, se subexpresan o no se expresan en absoluto. Por la expresión "ácido nucleico recombinante" en el presente documento se entiende ácido nucleico, formado originariamente *in vitro*, en general, mediante la manipulación del ácido nucleico, por ejemplo, usando polimerasas y endonucleasas, en una forma que no se encuentra normalmente en la naturaleza. De esta manera, se logra el ligamiento operativo de diferentes secuencias. Por tanto, un ácido nucleico aislado, en una forma lineal, o un vector de expresión formado *in vitro* mediante moléculas de ADN de ligamiento que normalmente no están unidas, se consideran ambos recombinantes para los fines de esta invención. Se entiende que una vez que se prepara un ácido nucleico recombinante y se introduce nuevamente en una célula u organismo huésped, éste se replicará de manera no recombinante, es decir, usando la maquinaria celular *in vivo* de la célula huésped en vez de manipulaciones *in vitro*; sin embargo, tales ácidos nucleicos, una vez producidos de manera recombinante, aunque posteriormente se repliquen de manera no recombinante, aún se consideran recombinantes para los fines de la invención. De manera similar, una "proteína recombinante" es una proteína preparada usando técnicas recombinantes, es decir, mediante la expresión de un ácido nucleico recombinante tal como se describió anteriormente.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que se entiende comúnmente por un experto habitual en la técnica a la pertenece que esta invención. En caso de conflicto, predominará la presente memoria descriptiva, incluyendo las definiciones.

III. Nucleótidos, polipéptidos, vectores y células huésped novedosos

La presente descripción abarca el gen humano novedoso *B1194*, que incluye una secuencia de polinucleótido tal como se describe en SEQ ID NO: 1, así como mutantes y degenerados del mismo, en la medida de que codifican para una proteína *B1194*, que incluye la secuencia de aminoácidos expuesta en SEQ ID NO: 2 o su equivalente funcional. Los ejemplos de polipéptidos funcionalmente equivalentes a *B1194* incluyen, por ejemplo, proteínas homólogas de otros organismos correspondientes a la proteína *B1194* humana, así como mutantes de las proteínas *B1194* humanas.

La presente descripción también abarca genes humanos novedosos *A2282V1*, *A2282V2* y *A2282V3*, que incluyen secuencias de polinucleótido descritas en SEQ ID NO: 3, 5, 7 respectivamente, así como mutantes y degenerados de los mismos, en la medida de que éstos codifican para una proteína *A2282V1*, *A2282V2* o *A2282V3*, incluyendo la secuencia de aminoácidos expuesta en SEQ ID NO: 4, 6, 8 o su equivalente funcional. Los ejemplos de polipéptidos funcionalmente equivalentes a *A2282V1*, *A2282V2* o *A2282V3* incluyen, por ejemplo, proteínas homólogas de otros organismos correspondientes a la proteína *A2282V1*, *A2282V2* o *A2282V3* humana, así como mutantes de las proteínas *A2282V1*, *A2282V2* o *A2282V3* humanas.

Un experto en la técnica conoce bien métodos para preparar polipéptidos funcionalmente equivalentes a una proteína dada e incluyen métodos conocidos de introducción de mutaciones en la proteína. Por ejemplo, un experto en la técnica puede preparar polipéptidos funcionalmente equivalentes a la proteína *B1194*, *A2282V1*, *A2282V2* o *A2282V3* humana introduciendo una mutación apropiada en la secuencia de aminoácidos de cualquiera de estas proteínas mediante mutagénesis dirigida al sitio (Hashimoto-Gotoh *et al.*, Gene 152:271-5 (1995); Zoller y Smith, Methods Enzymol 100: 468-500 (1983); Kramer *et al.*, Nucleic Acids Res. 12:9441-9456 (1984); Kramer y Fritz, Methods Enzymol 154: 350-67 (1987); Kunkel, Proc Natl Acad Sci USA 82: 488-92 (1985); Kunkel, Methods Enzymol 204: 125-139 (1991)). Las mutaciones de aminoácidos también pueden producirse en la naturaleza. El polipéptido de la presente descripción incluye aquellas proteínas que tienen las secuencias de aminoácidos de la proteína *B1194*, *A2282V1*, *A2282V2* o *A2282V3* humana en la que uno o más aminoácidos están mutados, que proporciona polipéptidos mutados resultantes que son funcionalmente equivalentes a las proteínas *B1194*, *A2282V1*, *A2282V2* o *A2282V3* humanas. El número de aminoácidos que va a mutarse en un mutante de este tipo generalmente es de 10 aminoácidos o menos, preferiblemente 6 aminoácidos o menos, y más preferiblemente 3 aminoácidos o menos.

Se conoce que las proteínas mutadas o modificadas, proteínas que tienen secuencias de aminoácidos modificadas sustituyendo, delecionando, insertando y/o añadiendo uno o más residuos de aminoácido de una determinada secuencia de aminoácidos, conservan la actividad biológica original (Mark *et al.*, Proc Natl Acad Sci USA 81: 5662-6 (1984); Zoller y Smith, Nucleic Acids Res 10: 6487-500 (1982); Dalbadie-McFarland *et al.*, Proc Natl Acad Sci USA 79: 6409-13 (1982)).

El residuo de aminoácido que va a mutarse preferiblemente se muta en un aminoácido diferente en el que se conservan las propiedades de la cadena lateral de aminoácido (un proceso conocido como sustitución de aminoácidos conservativa). Ejemplos de propiedades de las cadenas laterales de aminoácidos son aminoácidos hidrófobos (A, I, L, M, F, P, W, Y, V), aminoácidos hidrófilos (R, D, N, C, E, Q, G, H, K, S, T), y cadenas laterales que tienen los siguientes grupos funcionales o características en común: una cadena lateral alifática (G, A, V, L, I, P); una cadena lateral que contiene grupo hidroxilo (S, T, Y); una cadena lateral que contiene átomo de azufre (C, M); una cadena lateral que contiene ácido carboxílico y amida (D, N, E, Q); una cadena lateral que contiene base (R, K, H); y una cadena lateral que contiene grupo aromático (H, F, Y, W). Obsérvese que las letras entre paréntesis indican códigos de una letra de los aminoácidos.

Las tablas de sustitución conservativa que proporcionan aminoácidos funcionalmente similares se conocen bien en la técnica. Tales variantes modificadas de manera conservativa son además de y no excluyen las variantes polimórficas, homólogos entre especies, y alelos de la invención. Por ejemplo, cada uno de los siguientes ocho grupos contiene aminoácidos que son sustituciones conservativas entre sí:

- 1) Alanina (A), Glicina (G);
- 2) Ácido aspártico (D), ácido glutámico (E);
- 3) Asparagina (N), Glutamina (Q);
- 4) Arginina (R), Lisina (K);
- 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V);
- 6) Fenilalanina (F), Tirosina (Y), Triptófano (W);
- 7) Serina (S), Treonina (T); y
- 8) Cisteína (C), Metionina (M) (véase, por ejemplo, Creighton, Proteins (1984)).

Un ejemplo de un polipéptido al que se le añade uno o más residuos de aminoácidos a la secuencia de aminoácidos de la proteína B1194, A2282V1, A2282V2 o A2282V3 humana es una proteína de fusión que contiene la proteína B1194, A2282V1, A2282V2 o A2282V3 humana. En la presente invención se incluyen proteínas de fusión, fusiones de la proteína B1194, A2282V1, A2282V2 o A2282V3 humana y otros péptidos o proteínas. Las proteínas de fusión pueden prepararse mediante técnicas bien conocidas por un experto en la técnica, tal como ligando un ADN que codifica para una proteína B1194, A2282V1, A2282V2 o A2282V3 humana de la presente invención con ADN que codifica para otro péptido o proteína, de modo que los marcos coinciden, insertando el ADN de fusión en un vector de expresión, y expresándolo en un huésped. No existe ninguna restricción en cuanto a los péptidos o proteínas fusionados a la proteína de la presente invención.

Los péptidos conocidos que pueden usarse como péptidos que se fusionan a la proteína de la presente invención incluyen, por ejemplo, FLAG (Hopp *et al.*, Biotechnology 6: 1204-10 (1988)), 6xHis que contiene seis residuos de His (histidina), 10xHis, aglutinina de *Influenza* (HA), fragmento de c-myc humano, fragmento VSP-GP, fragmento p18HIV, etiqueta de T7, etiqueta de HSV, etiqueta de E, fragmento de antígeno SV40T, etiqueta de Ick, fragmento de β -tubulina, etiqueta de B, fragmento de proteína C y similares. Los ejemplos de proteínas que pueden fusionarse con una proteína de la invención incluyen GST (glutati6n-S-transferasa), aglutinina de *Influenza* (HA), regi6n constante de inmunoglobulina, β -galactosidasa, MBP (proteína de uni6n a maltosa) y similares.

Pueden prepararse proteínas de fusión fusionando ADN comercialmente disponible, que codifica para los péptidos o proteínas de fusión comentados anteriormente, con el ADN que codifica para el polipéptido de la presente invención y expresando el ADN fusionado preparado.

Un método alternativo conocido en la técnica para aislar polipéptidos funcionalmente equivalentes es, por ejemplo, el método que usa una técnica de hibridaci6n (Sambrook *et al.*, Molecular Cloning 2^a ed. 9.47-9.58, Cold Spring Harbor Lab. Press (1989)). Un experto en la técnica puede aislar fácilmente un ADN que tiene alta homología con la totalidad o una parte de la secuencia de ADN que codifica para la proteína B1194, A2282V1, A2282V2, A2282V3 humana (es decir, SEQ ID NO: 1, 3, 5 ó 7), e aislar polipéptidos funcionalmente equivalentes a la proteína B1194, A2282V1, A2282V2 o A2282V3 humana a partir del ADN aislado. Los polipéptidos dados a conocer incluyen aquellos que se codifican por el ADN que se hibrida con la totalidad o una parte de la secuencia de ADN que codifica para la proteína B1194, A2282V1, A2282V2 o A2282V3 humana y son funcionalmente equivalentes a la proteína B1194, A2282V1, A2282V2 o A2282V3 humana. Estos polipéptidos incluyen homólogos de mamífero correspondientes a la

proteína derivada de ser humano (por ejemplo, un polipéptido codificado por un gen de mono, rata, conejo y bovino). Por ejemplo, en el aislamiento de un ADNc altamente homólogo a un ADN que codifica para la proteína B1194 humana de animales, es particularmente preferible usar tejidos de líneas celulares de cáncer de testículos o de mama. Alternativamente, en el aislamiento de un ADNc altamente homólogo a un ADN que codifica para la proteína A2282V1, A2282V2 o A2282V3 humana de animales, es particularmente preferible usar tejidos de línea celular de cáncer de mama.

Un experto en la técnica puede seleccionar rutinariamente la condición de hibridación para aislar un ADN que codifica para un polipéptido funcionalmente equivalente a la proteína B1194, A2282V1, A2282V2 o A2282V3 humana. Por ejemplo, la hibridación puede realizarse llevando a cabo una prehibridación a 68°C durante 30 min. o más, usando “tampón Rapid-hyb” (Amersham LIFE SCIENCE), añadiendo una sonda marcada y calentando a 68°C durante 1 hora o más. Puede llevarse a cabo la siguiente etapa de lavado, por ejemplo, en una condición de baja rigurosidad. Una condición de baja rigurosidad es, por ejemplo, 42°C; SSC 2X, SDS al 0,1%, o preferiblemente 50°C, SSC 2X, SDS al 0,1%. Más preferiblemente, se usan condiciones de alta rigurosidad. Un ejemplo de una condición de alta rigurosidad incluye lavar 3 veces en SSC 2X, SDS al 0,01% a temperatura ambiente durante 20 min., después lavar 3 veces en SSC 1X, SDS al 0,1% a 37°C durante 20 min. y lavar dos veces en SSC 1X, SDS al 0,1% a 50°C durante 20 min. Sin embargo, varios factores, tales como temperatura y concentración salina, pueden influir en la rigurosidad de la hibridación y un experto en la técnica puede seleccionar adecuadamente los factores para lograr la rigurosidad requerida.

En lugar de la hibridación, puede utilizarse un método de amplificación génica, por ejemplo, el método de reacción en cadena de la polimerasa (PCR), para aislar un ADN que codifica para un polipéptido funcionalmente equivalente a la proteína B1194, A2282V1, A2282V2 o A2282V3 humana, usando un cebador sintetizado basado en la información de secuencia del ADN que codifica para la proteína (SEQ ID NO: 1, 3, 5 ó 7).

Los polipéptidos que son funcionalmente equivalentes a la proteína B1194, A2282V1, A2282V2 o A2282V3 humana codificada por el ADN aislado mediante las técnicas de hibridación o técnicas de amplificación génica anteriores, normalmente tienen una alta homología con la secuencia de aminoácidos de la proteína B1194, A2282V1, A2282V2 o A2282V3 humana. “Alta homología” se refiere normalmente a una homología del 40% o superior, preferiblemente del 60% o superior, más preferiblemente del 80% o superior, incluso más preferiblemente del 95% o superior. La homología de un polipéptido puede determinarse siguiendo el algoritmo en “Wilbur y Lipman, Proc Natl Acad Sci USA 80: 726-30 (1983)”.

Un polipéptido tal como se da a conocer puede tener variaciones en la secuencia de aminoácidos, peso molecular, punto isoeléctrico, la presencia o ausencia de cadenas de azúcar, o la forma, dependiendo de la célula o el huésped usado para producirlo o del método de purificación utilizado. No obstante, siempre que tenga una función equivalente a la de la proteína B1194, A2282V1, A2282V2, A2282V3 humana dada a conocer, está dentro del alcance de la presente descripción.

Los polipéptidos dados a conocer pueden prepararse como proteínas recombinantes o proteínas naturales, mediante métodos bien conocidos por los expertos en la técnica. Una proteína recombinante puede prepararse insertando un ADN, que codifica para un polipéptido de la presente invención (por ejemplo, un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7), en un vector de expresión apropiado, introduciendo el vector en una célula huésped apropiada, obteniendo el extracto y purificando el polipéptido sometiendo el extracto a cromatografía, por ejemplo, cromatografía de intercambio iónico, cromatografía de fase inversa, filtración en gel, o cromatografía de afinidad utilizando una columna a la que se fijan anticuerpos frente a la proteína dada a conocer, o combinando más de una de las columnas anteriormente mencionadas.

También cuando se expresa el polipéptido dado a conocer dentro de células huésped (por ejemplo, células animales y *E. coli*) como una proteína de fusión con proteína glutatión-S-transferasa o como una proteína recombinante complementada con múltiples histidinas, puede purificarse la proteína recombinante expresada usando una columna de glutatión o columna de níquel. Alternativamente, cuando se expresa el polipéptido dado a conocer como una proteína etiquetada con una c-myc, múltiples histidinas, o FLAG, ésta puede detectarse y purificarse usando anticuerpos frente a c-myc, His o FLAG, respectivamente.

Tras purificar la proteína de fusión, también es posible excluir regiones distintas del polipéptido objetivo cortando con trombina o factor-Xa tal como se requiera. Puede aislarse una proteína natural mediante métodos conocidos por un experto en la técnica, por ejemplo, poniendo en contacto la columna de afinidad, en la que se unen los anticuerpos que se unen a la proteína B1194, A2282V1, A2282V2, A2282V3 descrita a continuación, expresando el extracto de tejidos o células el polipéptido dado a conocer. Los anticuerpos pueden ser anticuerpos policlonales o anticuerpos monoclonales.

La presente descripción también abarca péptidos parciales del polipéptido de la presente invención. El péptido parcial tiene una secuencia de aminoácidos específica para el polipéptido de la presente invención y consiste en al menos 7 aminoácidos, preferiblemente 8 aminoácidos o más, y más preferiblemente 9 aminoácidos o más. Puede usarse el péptido parcial, por ejemplo, para preparar anticuerpos frente al polipéptido de la presente invención,

seleccionar un compuesto que se une al polipéptido dado a conocer, y seleccionar aceleradores o inhibidores del polipéptido dado a conocer.

Puede producirse un péptido parcial de la descripción mediante ingeniería genética, mediante métodos conocidos de síntesis de péptidos, o digiriendo el polipéptido dado a conocer con una peptidasa apropiada. Para la síntesis de péptidos, por ejemplo, puede usarse síntesis en fase sólida o síntesis en fase líquida.

Además, la presente descripción proporciona polinucleótidos que codifican para un polipéptido de la presente invención. Los polinucleótidos dados a conocer pueden usarse para la producción *in vivo* o *in vitro* de un polipéptido dado a conocer tal como se describió anteriormente. Puede usarse cualquier forma del polinucleótido dado a conocer, siempre que ésta codifique para un polipéptido dado a conocer, incluyendo ARNm, ARN, ADNc, ADN genómico, polinucleótidos sintetizados químicamente. Los polinucleótidos dados a conocer incluyen un ADN que comprende una secuencia de nucleótidos dada así como sus secuencias degeneradas, siempre que el ADN resultante codifique para un polipéptido dado a conocer.

Los polinucleótidos pueden prepararse mediante métodos conocidos por un experto en la técnica. Por ejemplo, el polinucleótido puede proceder de una biblioteca de ADNc de células que expresan un polipéptido de la presente invención, llevado a cabo la hibridación usando una secuencia parcial del ADN de la presente invención (por ejemplo, SEQ ID NO: 1, 3, 5 ó 7) como sonda. Puede prepararse una biblioteca de ADNc, por ejemplo, mediante el método descrito en Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor Laboratory Press (1989); alternativamente, pueden usarse bibliotecas de ADNc comercialmente disponibles. También puede prepararse una biblioteca de ADNc extrayendo ARN de células que expresan el polipéptido dado a conocer, sintetizando como oligo-ADN basados en la secuencia de un ADN dado a conocer (por ejemplo, SEQ ID NO: 1, 3, 5 ó 7), llevando a cabo PCR usando los oligo-ADN como cebadores, y amplificando los ADNc que codifican para la proteína dada a conocer.

Además, secuenciando los nucleótidos del ADNc obtenido, puede determinarse rutinariamente la región de traducción codificada por el ADNc, y puede obtenerse fácilmente la secuencia de aminoácidos del polipéptido dado a conocer. Además, examinando la biblioteca de ADN genómico usando el ADNc obtenido o partes del mismo como sonda, puede aislarse el ADN genómico.

Más específicamente, en primer lugar los ARNm pueden prepararse a partir de una célula, tejido u órgano (por ejemplo, línea celular de cáncer de testículo o de mama para B1194; y línea celular de cáncer de mama para A2282V1, A2282V2 o A2282V3) en las que se expresa un polipéptido objeto. Pueden usarse métodos conocidos para aislar ARNm; por ejemplo, puede prepararse ARN total mediante ultracentrifugación con guanidina (Chirgwin *et al.*, Biochemistry 18:5294-9 (1979)) o método de AGPC (Chomczynski y Sacchi, Anal Biochem 162:156-9 (1987)). Además, puede purificarse el ARNm a partir del ARN total usando el kit de purificación de ARNm (Pharmacia) y similares o, alternativamente, el ARNm puede purificarse directamente mediante el kit de purificación de ARNm QuickPrep (Pharmacia).

El ARNm obtenido se usa para sintetizar ADNc usando transcriptasa inversa. El ADNc puede sintetizarse usando un kit comercialmente disponible, tal como el kit de síntesis de ADNc AMV Reverse Transcriptase First-Strand (Seikagaku Kogyo). Alternativamente, el ADNc puede sintetizarse y amplificarse siguiendo el método 5'-RACE (Frohman *et al.*, Proc Natl Acad Sci USA 85: 8998-9002 (1988); Belyavsky *et al.*, Nucleic Acids Res 17: 2919-32 (1989)), que usa un cebador y similares, descrito en el presente documento, el kit 5'-Ampli FINDER RACE (Clontech), y reacción en cadena de la polimerasa (PCR).

Se prepara un fragmento de ADN deseado a partir de productos de PCR y se ligan con un ADN de vector. Se usan los vectores recombinantes para transformar *E. coli* y similares, y se prepara un vector recombinante deseado a partir de una colonia seleccionada. Puede verificarse la secuencia de nucleótidos del ADN deseado mediante métodos convencionales, tales como terminación de la cadena con didesoxinucleótido.

La secuencia de nucleótidos de un polinucleótido dado a conocer puede diseñarse para que se exprese más eficazmente teniendo en cuenta la frecuencia del uso de codones en el huésped que va usarse para la expresión (Grantham *et al.*, Nucleic Acids Res 9: 43-74 (1981)). Además, puede alterarse la secuencia del polinucleótido dado a conocer mediante un kit comercialmente disponible o un método convencional. Por ejemplo, puede alterarse la secuencia mediante digestión con enzimas de restricción, inserción de un oligonucleótido sintético o un fragmento de polinucleótido apropiado, adición de un ligador, o inserción del codón de iniciación (ATG) y/o el codón de parada (TAA, TGA o TAG).

En una realización particularmente preferida, el polinucleótido de la presente descripción abarca ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7.

Además, la presente descripción proporciona un polinucleótido que se hibrida en condiciones rigurosas con un polinucleótido que tiene una secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7, y codifica para un polipéptido funcionalmente equivalente a la proteína B1194, A2282V1, A2282V2 o A2282V3 descrita anteriormente. Tal como se trató anteriormente, un experto en la técnica puede elegir apropiadamente condiciones rigurosas. Por ejemplo, pueden usarse condiciones de baja rigurosidad. Más preferiblemente, se usan condiciones de alta rigurosidad. Estas

condiciones son tal como se describieron anteriormente. El ADN que se hibrida anterior es preferiblemente un ADNc o un ADN cromosómico.

La presente descripción también proporciona un vector en el que se inserta un polinucleótido dado a conocer. Un vector de la presente descripción es útil para mantener un polinucleótido, especialmente un ADN, de la presente descripción en una célula huésped; para expresar el polipéptido dado a conocer.

Cuando se selecciona *E. coli* como la célula huésped y se amplifica y produce el vector en gran cantidad en *E. coli* (por ejemplo, JM109, DH5 α , HB101 o XL1Blue), el vector debe tener "ori" que va a amplificarse en *E. coli* y un gen marcador para seleccionar la *E. coli* transformada (por ejemplo, un gen de resistencia a fármacos seleccionado mediante un fármaco tal como ampicilina, tetraciclina, kanamicina, cloranfenicol o similares). Por ejemplo, pueden usarse vectores de la serie M13, vectores de la serie pUC, pBR322, pBluescript, pCR-Script, etc. Además, también puede usarse pGEM-T, pDIRECT y pT7 para subclonar y extraer el ADNc así como los vectores descritos anteriormente. Cuando se usa un vector para producir una proteína de la presente invención, es especialmente útil un vector de expresión. Por ejemplo, un vector de expresión que va expresarse en *E. coli* debe tener las características anteriores para que se amplifique en *E. coli*. Cuando se usan *E. coli*, tal como JM109, DH5 α , HB101 o XL1Blue, como célula huésped, el vector debe tener un promotor, por ejemplo, promotor lacZ (Ward *et al.*, Nature 341: 544-6 (1989); FASEB J 6: 2422-7 (1992)), promotor araB (Better *et al.*, Science 240: 1041-3 (1988)), o promotor de T7 o similares, que puede expresar eficazmente el gen deseado en *E. coli*. Con respecto a esto, puede usarse pGEX-SX-1 (Pharmacia), "sistema QIAexpress" (Qiagen), pEGFP y pET (en este caso, el huésped es preferiblemente BL21 que expresa ARN polimerasa de T7), por ejemplo, en vez de los vectores anteriores. Adicionalmente, el vector también puede contener una secuencia señal para la secreción del polipéptido. Una secuencia señal a modo de ejemplo que dirige el polipéptido que va a secretarse hacia el periplasma de la *E. coli* es la secuencia señal pelB (Lei *et al.*, J Bacteriol 169: 4319-83 (1987)). Los medios para introducir los vectores en las células huésped diana incluyen, por ejemplo, el método del cloruro de calcio y el método de electroporación.

Además de *E. coli*, por ejemplo, pueden usarse vectores de expresión derivados de células de mamífero (por ejemplo, pcDNA3 (Invitrogen) y pEGF-BOS (Mizushima S., Nucleic Acids Res 18(17): 5322 (1990)), pEF, pCDM8), vectores de expresión derivados de células de insecto (por ejemplo, "sistema de expresión de baculovirus Bac-to-BAC" (GIBCO BRL), pBacPAK8), vectores de expresión derivados de plantas (por ejemplo, pMH1, pMH2), vectores de expresión derivados de virus de animales (por ejemplo, pHSV, pMV, pAdexLcw), vectores de expresión derivados de retrovirus (por ejemplo, pZlpneo), vector de expresión derivado de levadura (por ejemplo, "Kit de expresión de *Pichia*" (Invitrogen), pNV11, SP-Q01), y vectores de expresión derivados de *Bacillus subtilis* (por ejemplo, pPL608, pKTH50) para producir el polipéptido de la presente invención.

Con el fin de expresar un vector en células animales, tales como células CHO, COS, o NIH3T3, el vector debe tener un promotor necesario para su expresión en tales células, por ejemplo, el promotor de SV40 (Mulligan *et al.*, Nature 277: 108-14 (1979)), el promotor MMLV-LTR, el promotor EF1 α (Mizushima *et al.*, Nucleic Acids Res 18: 5322 (1990)), el promotor de CMV y similares, y preferiblemente un gen marcador para seleccionar transformantes (por ejemplo, un gen de resistencia a fármacos seleccionado mediante un fármaco a (por ejemplo, neomicina, G418)). Los ejemplos de vectores conocidos con estas características incluyen, por ejemplo, pMAM, pDR2, pBK-RSV, pBK-CMV, pOPRSV y pOP13.

Además, pueden usarse métodos para expresar un gen de manera estable y, al mismo tiempo, para amplificar el número de copias del gen en las células. Por ejemplo, puede introducirse un vector que comprende el gen DHFR complementario (por ejemplo, pCHO I) en células CHO en las que se deleciona la ruta de síntesis de ácido nucleico, y entonces se amplifica mediante metotrexato (MTX). Además, en caso de la expresión transitoria de un gen, puede usarse el método en el que se transforma un vector que comprende un origen de replicación de SV40 (pcD, etc.) en células COS que comprenden el gen que expresa el antígeno T de SV40 en el cromosoma.

Un polipéptido de la presente descripción obtenido tal como anteriormente puede aislarse a partir del interior o exterior (tal como el medio) de las células huésped, y purificarse como un polipéptido homogéneo sustancialmente puro. La expresión "sustancialmente puro" tal como se usa en el presente documento haciendo referencia a un polipéptido dado significa que el polipéptido está sustancialmente libre de otras macromoléculas biológicas. El polipéptido sustancialmente puro es al menos el 75% (por ejemplo, al menos el 80, 85, 95 ó 99%) puro en peso seco. La pureza puede medirse mediante cualquier método convencional apropiado, por ejemplo mediante cromatografía en columna, electroforesis en gel de poliacrilamida o análisis por HPLC. El método para el aislamiento y purificación del polipéptido no se limita a ningún método específico; de hecho, puede usarse cualquier método convencional.

Por ejemplo, pueden seleccionarse y combinarse apropiadamente cromatografía en columna, filtración, ultrafiltración, precipitación con sales, precipitación con disolventes, extracción con disolventes, destilación, inmunoprecipitación, electroforesis en gel de SDS-poliacrilamida, electroforesis de punto isoeléctrico, diálisis y recristalización para aislar y purificar el polipéptido.

Los ejemplos de cromatografía incluyen, por ejemplo, cromatografía de afinidad, cromatografía de intercambio iónico, cromatografía hidrófoba, filtración en gel, cromatografía de fase inversa, cromatografía de adsorción y similares

(Strategies for Protein Purification and Characterization: A Laboratory Course Manual. Ed. Daniel R. Marshak *et al.*, Cold Spring Harbor Laboratory Press (1996)). Estas cromatografías pueden realizarse mediante cromatografía líquida, tal como HPLC y FPLC. Por tanto, la presente descripción proporciona polipéptidos altamente purificados preparados mediante los métodos anteriores.

5 Un polipéptido de la presente descripción puede estar opcionalmente modificado o delecionado parcialmente tratándolo con un enzima para modificación de proteínas apropiada o tras purificación. Las enzimas para modificación de proteínas útiles incluyen, pero no se limitan a, tripsina, quimiotripsina, lisilendopeptidasa, proteína cinasa, glucosidasa y similares.

IV. Anticuerpos

10 La presente descripción también proporciona anticuerpos que se unen a un polipéptido dado a conocer. Un anticuerpo de este tipo puede usarse en cualquier forma, tal como anticuerpos monoclonales o policlonales, e incluye antisero obtenido inmunizando un animal tal como un conejo con el polipéptido dado a conocer, todas las clases de anticuerpos policlonales y monoclonales, anticuerpos humanos y anticuerpos humanizados producidos mediante recombinación genética.

15 Un polipéptido de la presente descripción usado como antígeno para obtener un anticuerpo puede derivarse de cualquier especie animal, pero preferiblemente se deriva de un mamífero tal como un ser humano, un ratón o una rata, más preferiblemente de un ser humano. Puede obtenerse un polipéptido derivado de ser humano a partir de las secuencias de nucleótidos o aminoácidos dadas a conocer en el presente documento. Según la presente descripción el polipéptido que va usarse como un antígeno de inmunización puede ser una proteína completa o un péptido parcial de la proteína. Un péptido parcial puede comprender, por ejemplo, el fragmento amino (N)-terminal o carboxilo (C)-terminal de un polipéptido de la presente invención.

20 Puede insertarse un gen que codifica para un polipéptido de la descripción o su fragmento en un vector de expresión conocido, que entonces se usa para transformar una célula huésped tal como se describe en el presente documento. El polipéptido deseado o su fragmento puede recuperarse desde el exterior o el interior de las células huésped mediante cualquier método convencional, y posteriormente puede usarse como antígeno. Alternativamente, pueden usarse células completas que expresan el polipéptido o sus lisados, o un polipéptido químicamente sintetizado como antígeno.

25 Puede inmunizarse cualquier animal mamífero con el antígeno, pero preferiblemente se tiene en cuenta la compatibilidad con las células originales usadas para la fusión celular. En general, se usan animales de los órdenes *Rodentia*, *Lagomorpha* o *Primates*. Los animales del orden *Rodentia* incluyen, por ejemplo, el ratón, rata y hámster. Los animales del orden *Lagomorpha* incluyen, por ejemplo, el conejo. Animales del orden *Primate* incluyen, por ejemplo, un mono de Catarrhini (mono del viejo mundo) tal como *Macaca fascicularis*, mono rhesus, babuino sagrado y chimpancés.

30 En la técnica se conocen métodos para inmunizar animales con antígenos. Por ejemplo, la inyección intraperitoneal o inyección subcutánea de antígenos es un método convencional para la inmunización de mamíferos. Más específicamente, los antígenos pueden diluirse y suspenderse en una cantidad apropiada de solución salina tamponada con fosfato (PBS), solución salina fisiológica, etc. Si se desea, la suspensión de antígenos puede mezclarse con una cantidad apropiada de un adyuvante convencional, tal como adyuvante completo de Freund, prepararse en emulsión, y luego administrarse a los animales mamíferos. Preferiblemente, esto va seguido de varias administraciones de antígeno mezcladas con una cantidad apropiada de adyuvante incompleto de Freund de cada 4 a 21 días. Para la inmunización también puede usarse un vehículo apropiado. Tras la inmunización tal como anteriormente, se examina el suero mediante un método convencional para determinar un aumento en la cantidad de anticuerpos deseados.

35 Los anticuerpos policlonales frente a los polipéptidos de la presente descripción pueden prepararse extrayendo sangre del mamífero inmunizado examinados para detectar el aumento de los anticuerpos deseados en el suero, y separando el suero de la sangre mediante cualquier método convencional. Los anticuerpos policlonales incluyen suero que contiene anticuerpos policlonales, así como fracciones que contienen los anticuerpos policlonales aislados del suero. Puede prepararse inmunoglobulina G o M a partir de una fracción que reconoce sólo el polipéptido de la presente invención usando, por ejemplo, una columna de afinidad acoplada con el polipéptido de la presente descripción, y además purificando esta fracción usando columna de proteína A o proteína G.

40 Para preparar anticuerpos monoclonales, se extraen células inmunitarias del mamífero inmunizado con el antígeno y se examinan para detectar el aumento del nivel de anticuerpos deseados en el suero tal como se describió anteriormente, y se someten a fusión celular. Las células inmunitarias usadas para la fusión celular se obtienen preferiblemente del bazo. Otras células originales preferidas que van a fusionarse con el inmunocito anterior incluyen, por ejemplo, células de mieloma de mamíferos, y más preferiblemente células de mieloma que tienen una propiedad adquirida para la selección de células fusionadas mediante fármacos.

55 El inmunocito anterior y las células de mieloma pueden fusionarse según métodos conocidos, por ejemplo, el método de Milstein *et al.* (Galfre y Milstein, *Methods Enzymol* 73: 3-46 (1981)).

5 Los hibridomas resultantes obtenidos mediante la fusión celular pueden seleccionarse cultivándolos en un medio de selección convencional, tal como medio HAT (medio que contiene hipoxantina, aminopterina y timidina). Normalmente se mantiene el cultivo celular en el medio HAT durante varios días hasta varias semanas, siendo el tiempo suficiente para permitir que mueran todas las otras células, con la excepción del hibridoma deseado (células no fusionadas). Entonces, se realiza la dilución limitante convencional para seleccionar y clonar una célula de hibridoma que produce el anticuerpo deseado.

10 Además del método anterior, en el que se inmuniza un animal no humano con un antígeno para preparar un hibridoma, pueden inmunizarse linfocitos humanos tales como los infectados por virus EB con un polipéptido, células que expresan polipéptido, o sus lisados *in vitro*. Entonces, se fusionan los linfocitos inmunizados con células de mieloma derivadas de ser humano que pueden dividirse indefinidamente, tal como U266, para producir un hibridoma que produce un anticuerpo humano deseado que puede unirse al polipéptido puede obtenerse (solicitud de patente japonesa publicada no examinada n.º (JPA) Sho 63-17688).

15 Posteriormente se trasplantan los hibridomas obtenidos en la cavidad abdominal de un ratón y se extraen las ascitis. Los anticuerpos monoclonales obtenidos pueden purificarse mediante, por ejemplo, precipitación con sulfato de amonio, una columna de proteína A o proteína G, cromatografía de intercambio iónico con DEAE, o una columna de afinidad a la que se acopla el polipéptido de la presente descripción. El anticuerpo de la presente descripción puede usarse no sólo para la purificación y detección del polipéptido de la presente descripción, sino también como un candidato para agonistas y antagonistas del polipéptido de la presente descripción. Además, este anticuerpo puede aplicarse al tratamiento con anticuerpos para enfermedades relacionadas con el polipéptido de la presente descripción. Cuando va a administrarse el anticuerpo obtenido al cuerpo humano (tratamiento con anticuerpos), es preferible un anticuerpo humano o un anticuerpo humanizado para reducir la inmunogenicidad.

20 Por ejemplo, los animales transgénicos que tienen un repertorio de genes de anticuerpos humanos pueden inmunizarse con un antígeno seleccionado de un polipéptido, células que expresan polipéptido o sus lisados. Entonces, se recogen las células productoras de anticuerpos de los animales y se fusionan con células de mieloma para obtener un hibridoma, a partir del cual pueden prepararse anticuerpos humanos frente al polipéptido (véanse los documentos WO92-03918, WO94-02602, WO94-25585, WO96-33735 y WO96-34096).

25 Alternativamente, puede immortalizarse una célula inmunitaria, tal como un linfocito inmunizado, que produce anticuerpos, mediante un oncogén y usarse para preparar anticuerpos monoclonales.

30 Los anticuerpos monoclonales así obtenidos también pueden prepararse de manera recombinante usando técnicas de ingeniería genética (véase, por ejemplo; Borrebaeck y Larrick, Therapeutic Monoclonal Antibodies, publicado en el Reino Unido por MacMillan Publishers LTD (1990)). Por ejemplo, puede clonarse un ADN que codifica para un anticuerpo a partir de una célula inmunitaria, tal como un hibridoma o un linfocito inmunizado que produce el anticuerpo, insertarse en un vector apropiado e introducirse en células huésped para preparar un anticuerpo recombinante. La presente invención también proporciona anticuerpos recombinantes preparados tal como se describió anteriormente.

35 Además, un anticuerpo de la presente descripción puede ser un fragmento de un anticuerpo o anticuerpo modificado, siempre que éste se una a uno o más de los polipéptidos dados a conocer. Por ejemplo, el fragmento de anticuerpo puede ser Fab, F(ab')₂, Fv o Fv de cadena sencilla (scFv), en el que se ligan los fragmentos Fv de las cadenas H y L mediante un ligador apropiado (Huston *et al.*, Proc Natl Acad Sci USA 85: 5879-83 (1988)). Más específicamente, puede generarse un fragmento de anticuerpo tratando un anticuerpo con una enzima, tal como papaína o pepsina. Alternativamente, puede construirse un gen que codifica para el fragmento de anticuerpo, insertarse en un vector de expresión y expresarse en una célula huésped apropiada (véanse, por ejemplo, Co *et al.*, J Immunol 152: 2968-76 (1994); Better y Horwitz, Methods Enzymol 178: 476-96 (1989); Pluckthun y Skerra, Methods Enzymol 178: 497-515 (1989); Lamoyi, Methods Enzymol 121: 652-63 (1986); Rousseaux *et al.*, Methods Enzymol 121: 663-9 (1986); Bird y Walker, Trends Biotechnol 9: 132-7 (1991)).

40 Un anticuerpo puede modificarse mediante conjugación con una variedad de moléculas, tales como polietilenglicol (PEG). La presente descripción proporciona tales anticuerpos modificados. El anticuerpo modificado puede obtenerse modificando químicamente un anticuerpo. Estos métodos de modificación son convencionales en el campo.

45 Alternativamente, pueden obtenerse un anticuerpo de la presente descripción como un anticuerpo quimérico, entre una región variable derivada de anticuerpo no humano y la región constante derivada de anticuerpo humano, o como un anticuerpo humanizado, que comprende la región determinante de la complementariedad (CDR) derivada de anticuerpo no humano, la región de entramado (FR) y la región constante derivada de anticuerpo humano. Tales anticuerpos pueden prepararse usando tecnología conocida.

50 Los anticuerpos obtenidos tal como anteriormente pueden purificarse hasta homogeneidad. Por ejemplo, la separación y purificación del anticuerpo pueden realizarse según métodos de separación y purificación usados para proteínas generales. Por ejemplo, el anticuerpo puede separarse y aislarse mediante el uso combinado y seleccionado apropiadamente de cromatografías en columna, tales como cromatografía de afinidad, filtración, ultrafiltración,

precipitación con sales, diálisis, electroforesis en gel de SDS-poliacrilamida, isoelectroenfoque y otras (Antibodies: A Laboratory Manual. Ed Harlow y David Lane, Cold Spring Harbor Laboratory (1988)), pero sin limitarse a las mismas. Puede usarse una columna de proteína A y de proteína G como columna de afinidad. Las columnas de proteína A a modo de ejemplo que van a usarse incluyen, por ejemplo, Hyper D, POROS, y Sepharose F.F. (Farmacia).

5 Los ejemplos de cromatografía, con la excepción de la de afinidad incluyen cromatografía de intercambio iónico, cromatografía hidrófoba, filtración en gel, cromatografía de fase inversa, cromatografía de adsorción y similares (Strategies for Protein Purification and Characterization: A Laboratory Course Manual. Ed Daniel R. Marshak *et al.*, Cold Spring Harbor Laboratory Press (1996)). Los procedimientos cromatográficos pueden llevarse a cabo mediante cromatografía en fase líquida, tal como HPLC y FPLC.

10 Por ejemplo, pueden usarse ensayos de absorbancia, ensayos de inmunoadsorción ligado a enzimas (ELISA), inmunoensayos enzimáticos (EIA), radioinmunoanálisis (RIA) y/o ensayos de inmunofluorescencia para medir la actividad de unión a antígeno del anticuerpo de la descripción. En el ELISA, se inmoviliza un anticuerpo de la presente descripción en una placa, se aplica un polipéptido de la descripción a la placa, y entonces se aplica una muestra que contiene un anticuerpo deseado, tal como sobrenadante de cultivo de células productoras de anticuerpos o anticuerpos purificados. Entonces, se aplica un anticuerpo secundario que reconoce el anticuerpo primario y está marcado con una enzima, tal como fosfatasa alcalina, y se incuba la placa. A continuación, tras el lavado, se añade un sustrato enzimático, tal como fosfato de p-nitrofenilo a la placa, y se mide la absorbancia para evaluar la actividad de unión a antígeno de la muestra. Puede usarse un fragmento del polipéptido, tal como un fragmento C-terminal o N-terminal, como antígeno para evaluar la actividad de unión del anticuerpo. Puede usarse BIAcore (Farmacia) para evaluar la actividad del anticuerpo según la presente descripción.

Los métodos anteriores permiten la detección o medición del polipéptido dado a conocer, exponiendo el anticuerpo dado a conocer a una muestra que se supone que contiene el polipéptido dado a conocer, y detectando o midiendo el inmunocomplejo formado por el anticuerpo y el polipéptido.

25 Dado que el método de detección o medición del polipéptido según la descripción puede detectar o medir específicamente un polipéptido, el método puede ser útil en una variedad de experimentos en los que se usa el polipéptido.

V. Oligonucleótidos antisentido

30 Tal como se observó anteriormente, la presente descripción también proporciona un polinucleótido que se hibrida con un polinucleótido que codifica para la proteína B1194, A2282V1, A2282V2 o A2282V3 humana (SEQ ID NO: 1, 3, 5 ó 7) o la cadena complementaria del mismo, y que comprende al menos 15 nucleótidos. El polinucleótido de la presente descripción es preferiblemente un polinucleótido que se hibrida específicamente con el ADN que codifica para el polipéptido de la presente invención. La expresión "se hibrida específicamente" tal como se usa en el presente documento, significa que no se produce significativamente hibridación cruzada con ADN que codifica para otras proteínas, bajo las condiciones de hibridación habituales, preferiblemente en condiciones de hibridación rigurosas. Tales polinucleótidos incluyen sondas, cebadores, nucleótidos y derivados de nucleótidos (por ejemplo, oligonucleótidos antisentido y ribozimas), que se hibrida específicamente con ADN que codifica para el polipéptido de la invención o su cadena complementaria. Además, tal polinucleótido puede utilizarse para la preparación de un chip de ADN.

40 Por consiguiente, la presente descripción incluye un oligonucleótido antisentido que se hibrida con cualquier sitio dentro de la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7. Un oligonucleótido antisentido de este tipo está dirigido preferiblemente frente al menos 15 nucleótidos contiguos de la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 ó 7. Incluso se prefiere más el oligonucleótido antisentido mencionado anteriormente que contiene un codón de iniciación en los al menos 15 nucleótidos contiguos mencionados anteriormente.

45 Pueden usarse derivados o productos modificados de oligonucleótidos antisentido como oligonucleótidos antisentido de la presente descripción. Los ejemplos de tales productos modificados incluyen modificaciones de fosfonato de alquilo inferior, tales como modificaciones de de tipo metil-fosfonato o de tipo etil-fosfonato, de fosforotioato y modificaciones de fosforamido.

50 Los derivados de oligonucleótido antisentido de la presente descripción actúan sobre las células que producen el polipéptido dado a conocer uniéndose al ADN o ARNm que codifica para el polipéptido, inhibiendo su transcripción o traducción, promoviendo la degradación del ARNm e inhibiendo la expresión del polipéptido dado a conocer, dando como resultado de ese modo la inhibición de la función del polipéptido.

Puede prepararse un derivado de oligonucleótido antisentido de la presente descripción en una preparación externa, tal como un linimento o un cataplasma, mezclando con un material base adecuado que es inactivo frente a los derivados.

55 También, cuando sea necesario, pueden formularse los derivados en comprimidos, polvos, gránulos, cápsulas, cápsulas de liposomas, inyecciones, disoluciones, gotas para la nariz y agentes de liofilización añadiendo excipientes,

agentes isotónicos, solubilizantes, estabilizantes, conservantes, analgésicos y similares. Estos pueden prepararse siguiendo métodos habituales.

Puede proporcionarse al paciente el derivado de oligonucleótido antisentido aplicando directamente sobre el sitio afectado o inyectando en un vaso sanguíneo de manera que éste alcanzará el sitio de enfermedad. También puede usarse un medio que aumenta el nivel de antisentido para aumentar la durabilidad y permeabilidad de la membrana. Los ejemplos incluyen, pero no se limitan a, liposoma, poli-L-lisina, lípido, colesterol, lipofectina o derivados de estos.

Puede ajustarse adecuadamente la dosificación del derivado de oligonucleótido antisentido de la presente descripción según el estado del paciente y usarse en las cantidades deseadas. Por ejemplo, puede administrarse un intervalo de dosis de 0,1 a 100 mg/kg, preferiblemente de 0,1 a 50 mg/kg.

El término "ARNip" se refiere a una molécula de ARN bicatenaria que evita la traducción de un ARNm diana. Se usan técnicas convencionales para introducir ARNip en las células, que incluyen aquellas en las que se usa el ADN como el molde para transcribir el ARN. Un ARNip de la presente descripción comprende una secuencia de ácido nucleico sentido y una secuencia de ácido nucleico antisentido de un polinucleótido que codifica para proteína B1194, A2282V1, A2282V2 o A2282V3 humana (SEQ ID NO: 1, 3, 5 ó 7). El ARNip se construye de manera que un transcrito individual (ARN bicateario) tiene tanto las secuencias sentido como antisentido complementarias del gen diana, por ejemplo, una horquilla.

La unión del ARNip a un transcrito correspondiente a B1194, A2282V1, A2282V2 o A2282V3 en la célula diana da como resultado una reducción en la producción de proteínas por las células. La longitud del oligonucleótido es de al menos 10 nucleótidos y puede ser tan largo como el transcrito que se produce de manera natural. Preferiblemente, el oligonucleótido es de menos de 75, 50, 25 nucleótidos de longitud. Lo más preferiblemente, el oligonucleótido es de 19-25 nucleótidos de longitud. Los ejemplos de oligonucleótidos de ARNip de B1194, A2282V1, A2282V2 y A2282V3 que inhiben el crecimiento de la célula cancerosa incluyen la secuencia diana que contiene SEQ ID NO:38-41.

Además, con el fin de potenciar la actividad de inhibición del ARNip, puede añadirse el nucleótido "u" al extremo 3' de la cadena antisentido de la secuencia diana. El número de "u" que van a añadirse es al menos 2, generalmente de 2 a 10, preferiblemente de 2 a 5. Las "u" añadidas forman una cadena sencilla en el extremo 3' de la cadena antisentido del ARNip.

Los ARNip de B1194, A2282V1, A2282V2 y A2282V3 pueden introducirse directamente en las células en una forma que puede unirse a los transcritos de ARNm. Alternativamente, el ADN que codifica para el ARNip de B1194, A2282V1, A2282V2 o A2282V3 puede estar contenido en un vector.

Los vectores se producen, por ejemplo, clonando una secuencia diana de B1194, A2282V1, A2282V2 o A2282V3 en un vector de expresión ligado operativamente a secuencias reguladoras que flanquean la secuencia B1194, A2282V1, A2282V2 o A2282V3 de una manera que permite la expresión (mediante la transcripción de la molécula de ADN) de ambas cadenas (Lee, N.S. *et al.*, (2002) *Nature Biotechnology* 20:500-505.). Se transcribe una molécula de ARN que es antisentido con respecto a un ARNm de B1194, A2282V1, A2282V2 o A2282V3 mediante un primer promotor (por ejemplo, una secuencia promotora en 3' del ADN clonado) y se transcribe una molécula de ARN que es la cadena sentido para un ARNm de B1194, A2282V1, A2282V2 o A2282V3 mediante un segundo promotor (por ejemplo, una secuencia promotora en 5' del ADN clonado). Las cadenas sentido y antisentido se hibridan *in vivo* para generar constructos de ARNip para silenciar el gen B1194, A2282V1, A2282V2 o A2282V3. Alternativamente, pueden utilizarse dos constructos para crear las cadenas sentido y antisentido del constructo de ARNip. B1194, A2282V1, A2282V2 o A2282V3 clonados pueden codificar para un constructo que tiene estructura secundaria, por ejemplo, horquillas, en los que un transcrito individual tiene tanto las secuencias sentido como antisentido complementarias del gen diana.

Además, una secuencia de bucle que consiste en una secuencia de nucleótidos arbitraria puede ubicarse entre la secuencia sentido y antisentido con el fin de formar la estructura de bucle en horquilla. Por tanto, la presente descripción también proporciona ARNip que tiene la fórmula general 5'-[A]-[B]-[A']-3', en la que [A] es una secuencia de ribonucleótido que corresponde a una secuencia de nucleótidos SEQ ID NO:38-41, [B] es una secuencia de ribonucleótido que consiste en de 3 a 23 nucleótidos, y [A'] es una secuencia de ribonucleótido que consiste en la secuencia complementaria de [A]. La secuencia de bucle puede consistir en una secuencia arbitraria preferiblemente de 3 a 23 nucleótidos de longitud. La secuencia de bucle, por ejemplo, puede seleccionarse del grupo que consiste en las siguientes secuencias (http://www.ambion.com/techlib/tb/tb_506.html). En el ARNip de la presente descripción, puede añadirse el nucleótido "u" al extremo 3' de [A'], con el fin de potenciar la actividad de inhibición del ARNip. El número de "u" que van a añadirse es al menos 2, generalmente de 2 a 10, preferiblemente de 2 a 5. Además, la secuencia de bucle que consiste en 23 nucleótidos también proporciona ARNip activo (Jacque, J.-M. *et al.*, *Nature* 418: 435-438 (2002).):

- CCC, CCACC o CCACACC: Jacque, J. M. *et al.*, *Nature*, Vol. 418: 435-438 (2002);
- UUCG: Lee, N.S. *et al.*, *Nature Biotechnology* 20: 500-505.; Fruscoloni, P. *et al.*, *Proc. Natl. Acad. Sci. USA* 100(4): 1639-1644 (2003); y

- UUCAAGAGA: Dykxhoorn, D. M. *et al.*, Nature Reviews Molecular Cell Biology 4: 457-467 (2003).

A continuación se muestran ejemplos de ARNip preferidos que tienen estructura de horquilla de la presente invención. En la siguiente estructura, la secuencia de bucle puede seleccionarse del grupo que consiste en CCC, UUCG, CCACC, CCACACC y UUCAAGAGA. Una secuencia de bucle preferida es UUCAAGAGA ("ttcaagaga" en ADN).

- 5 guauaucuugcccucugaa-[B]-uucagagggaagauauac (para la secuencia diana de SEQ ID NO: 38)
 guccgaacacacuuuguu-[B]-aacaagauguguucggac (para la secuencia diana de SEQ ID NO: 39)
 gacaucuaucugcgca-[B]-ugcagcuagauaggauguc (para la secuencia diana de SEQ ID NO: 40)
 aguauauggaacuacaa-[B]-uugguaguucuaugaacu (para la secuencia diana de SEQ ID NO: 41)

- 10 Las secuencias reguladoras que flanquean la secuencia de B1194, A2282V1, A2282V2 o A2282V3 son idénticas o diferentes, de manera que su expresión puede modularse independientemente, o de una manera temporal o espacial. Se transcriben los ARNip intracelularmente clonando los moldes del gen B1194, A2282V1, A2282V2 o A2282V3 en un vector que contiene, por ejemplo, una unidad de transcripción de ARN polimerasa III de U6 de ARN nuclear pequeño (ARNnp) o el promotor de ARN H1 humano. Para introducir el vector en la célula, puede usarse un agente de potenciación de la transfección. FuGENE (Roche Diagnostics), Lipofectamine 2000 (Invitrogen),
 15 Oligofectamine (Invitrogen) y Nucleofector (Wako pure Chemical) son útiles como agente de potenciación de la transfección.

La secuencia de nucleótidos de los ARNip puede diseñarse usando un programa informático de diseño de ARNip disponible del sitio web de Ambion (http://www.ambion.com/techlib/misc/siRNA_finder.html). Se seleccionan las secuencias de nucleótidos para el ARNip mediante el programa informático basado en el siguiente protocolo:

- 20 Selección de sitios diana para ARNip:

1. Comenzando por el codón de iniciación AUG del transcrito objeto, explorar en el sentido de 3' para detectar secuencias de dinucleótido AA. Registrar la aparición de cada AA y los 19 nucleótidos adyacentes en 3' como posibles sitios diana para ARNip. Tuschl, *et al.*, Genes Dev 13(24):3191-7(1999), no se recomienda frente al diseño de ARNip para las regiones no traducidas (UTR) en 5' y 3' y las regiones cerca del codón de iniciación (dentro de las 75 bases) ya que éstas pueden ser más ricas en sitios de unión a proteína reguladora. Las proteínas de unión a UTR y/o los complejos de iniciación de la traducción pueden interferir con la unión del complejo endonucleasa-ARNip.
2. Comparar los posibles sitios diana con la base de datos del genoma humano y eliminar de consideración cualquier secuencia diana con homología significativa a otras secuencias codificantes. La búsqueda de homología puede realizarse usando BLAST (Altschul SF, *et al.*, J Mol Biol. 1990; 215:403-10.; Altschul SF, *et al.*; Nucleic Acids Res. 1997; 25:3389-402.), que puede encontrarse en el servidor del NCBI en: www.ncbi.nlm.nih.gov/BLAST/.
3. Seleccionar las secuencias dianas calificadas para la síntesis. En Ambion, preferiblemente pueden seleccionarse varias secuencias diana a lo largo de la longitud del gen para su evaluación.

- 35 Se sometieron a prueba *in vitro* oligonucleótidos y oligonucleótidos complementarios a diversas partes del ARNm de B1194, A2282V1, A2282V2 y A2282V3 para detectar su capacidad para disminuir la producción de B1194, A2282V1, A2282V2 o A2282V3 en células tumorales (por ejemplo, usando la línea celular de cáncer de mama T47D o MCF7) según métodos convencionales. Puede detectarse una reducción en el producto génico de B1194, A2282V1, A2282V2 o A2282V3 en las células puestas en contacto con la composición de ARNip candidato en comparación con las células cultivadas en ausencia de la composición candidata, usando anticuerpos específicos frente a B1194, A2282V1, A2282V2 o A2282V3 u otras estrategias de detección. Entonces pueden someterse a prueba secuencias que reducen la producción de B1194, A2282V1, A2282V2 o A2282V3 en ensayos basados en células o ensayos libres de células *in vitro* para detectar sus efectos inhibidores sobre el crecimiento celular. Las secuencias que inhiben el crecimiento celular en un ensayo basado en células *in vitro* se someten a prueba en ratas o ratones *in vivo* para confirmar la disminución de la producción de B1194, A2282V1, A2282V2 o A2282V3 y la disminución del crecimiento celular tumoral en animales con neoplasias malignas.

- 45 En la descripción se incluyen también moléculas bicatenarias que incluyen la secuencia de ácido nucleico de secuencias diana, por ejemplo, nucleótidos SEQ ID NO: 38-41. En la presente invención, la molécula bicatenaria que comprende una cadena sentido y una cadena antisentido, en la que la cadena sentido comprende una secuencia de ribonucleótido correspondiente a SEQ ID NO: 38-41, y en la que la cadena antisentido comprende una secuencia de ribonucleótido que es complementaria a dicha cadena sentido, hibridándose dicha cadena sentido y dicha cadena antisentido entre sí para formar dicha molécula bicatenaria, e inhibiendo dicha molécula bicatenaria, cuando se introduce en una célula que expresa el gen B1194, A2282V1, A2282V2 o A2282V3, la expresión de dicho gen. En la presente invención, cuando el ácido nucleico aislado es ARN o derivados del mismo, debe sustituirse la base "t" por "u" en las secuencias de nucleótidos. Tal como se usa en el presente documento, el término "complementario" se refiere un
 55

apareamiento de bases de Watson-Crick o Hoogsteen entre unidades de nucleótidos de una molécula de ácido nucleico, y el término “unión” significa la interacción física o química entre dos polipéptidos o compuestos o polipéptidos o compuestos asociados o combinaciones de los mismos.

5 Las secuencias de ácido nucleico complementarias se hibridan en condiciones apropiadas para formar dúplex estables que contienen pocos o ningún apareamiento erróneo. Además, la cadena sentido y cadena antisentido del nucleótido aislado de la presente invención, pueden formar nucleótido bicatenario o estructura de bucle en horquilla mediante la hibridación. En una realización preferida, tales dúplex contienen no más de 1 apareamiento erróneo por cada 10 apareamientos. En una realización especialmente preferida, en la que las cadenas del dúplex son completamente complementarias, tales dúplex no contienen ningún apareamiento erróneo. Por ejemplo, la molécula de ácido nucleico es de menos de 500, 200 ó 75 nucleótidos de longitud. También se incluye en la invención un vector que contiene uno o más de los ácidos nucleicos descritos en el presente documento, y una célula que contiene los vectores. Los ácidos nucleicos aislados de la presente descripción son útiles para detectar ARNip contra B1194, A2282V1, A2282V2 o A2282V3 o ADN que codifica para el ARNip. Cuando se usan los ácidos nucleicos para ARNip o ADN codificante del mismo, la cadena sentido es preferiblemente más larga de 19 nucleótidos, y más preferiblemente más larga de 21 nucleótidos. La presente invención específicamente se refiere a los ARNip caracterizados en las reivindicaciones.

VI. Diagnóstico del cáncer de mama

20 Un oligonucleótido antisentido o ARNip de la presente descripción inhibe la expresión de un polipéptido de la descripción y por tanto es útil para suprimir la actividad biológica del polipéptido de la descripción. También, los inhibidores de la expresión que comprenden el oligonucleótido antisentido o ARNip de la descripción, son útiles hasta el punto de que pueden inhibir la actividad biológica del polipéptido de la descripción. Por tanto, una composición que comprende el oligonucleótido antisentido o ARNip de la presente invención es útil en el tratamiento del cáncer de mama. Además, la presente descripción proporciona un método para diagnosticar cáncer de mama usando el nivel de expresión de los polipéptidos de la presente descripción como marcador de diagnóstico.

25 El método de diagnóstico de la presente descripción comprende preferiblemente las etapas de: (a) detectar el nivel de expresión del gen *B1194*, *A2282V1*, *A2282Y2* o *A2282V3* de la presente descripción, y (b) relacionar una elevación en el nivel de expresión con el cáncer de mama.

30 Puede estimarse el nivel de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en un espécimen particular cuantificando el ARNm correspondiente a o la proteína codificada por el gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3*. Los expertos en la técnica conocen métodos de cuantificación para ARNm. Por ejemplo, pueden estimarse los niveles de ARNm correspondientes al gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* mediante transferencia de tipo Northern o RT-PCR. Dado que las secuencias de nucleótidos de longitud completa de los genes *B1194*, *A2282V1*, *A2282Y2* y *A2282V3* se muestran en SEQ ID NO: 1, 3, 5 ó 7, cualquier experto en la técnica puede diseñar las secuencias de nucleótidos para sondas o cebadores para cuantificar el gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3*.

35 También puede analizarse el nivel de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* basándose en la actividad o cantidad de proteína codificada por el gen. A continuación se muestra un método para determinar la cantidad de la proteína *B1194*, *A2282V1*, *A2282V2* o *A2282V3*. Por ejemplo, un método de inmunoensayo es útil para determinar proteínas en materiales biológicos. Puede usarse cualquier material biológico para la determinación de la proteína o su actividad. Por ejemplo, puede analizarse una muestra de sangre para la estimación de la proteína codificada por un marcador sérico. Por otra parte, puede seleccionarse un método adecuado para la determinación de la actividad de una proteína codificada por el gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* según la actividad de cada proteína que va a analizarse.

45 Según los métodos de la presente descripción, se estiman los niveles de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en un espécimen (muestra de prueba) y se comparan con los de una muestra normal. Cuando una comparación de este tipo muestra que el nivel de expresión del gen diana es superior que el de la muestra normal, se evalúa que el sujeto va a verse afectado por cáncer de mama. El nivel de expresión del gen *B1194*, *A2282V1*, *A2282V2* o *A2282V3* en los especímenes de la muestra normal y del sujeto puede determinarse al mismo tiempo. Alternativamente, pueden determinarse intervalos normales de los niveles de expresión mediante un método estadístico basado en los resultados obtenidos del análisis de especímenes previamente recogidos de un grupo control. Se compara un resultado obtenido de una muestra del sujeto con el intervalo normal, cuando el resultado no se encuentra dentro del intervalo normal, se evalúa que el sujeto va a verse afectado por cáncer de mama.

55 En la presente descripción, también se proporciona un agente de diagnóstico para diagnosticar cáncer de mama. El agente de diagnóstico comprende un compuesto que se une a un polinucleótido o un polipéptido de la presente descripción. Preferiblemente, puede usarse un oligonucleótido que se hibrida con el polinucleótido dado a conocer, o un anticuerpo que se une al polipéptido dado a conocer, como compuesto de este tipo.

VII. Monitorización del tratamiento del cáncer de mama

Los niveles de expresión de los genes *B1194*, *A2282V1*, *A2282Y2* y *A2282V3* también permiten que el ciclo del tratamiento del cáncer de mama se monitorice. En este método, se proporciona una población celular de prueba de un sujeto que se somete a tratamiento para cáncer de mama. Si se desea, se obtienen poblaciones celulares de prueba del sujeto a diversos puntos de tiempo, antes, durante y/o tras el tratamiento. Entonces, se determina la expresión de uno o más de los genes *B1194*, *A2282V1*, *A2282V2* y *A2282V3* en la población celular y se compara con una población celular de referencia, que incluye células cuyo estado de cáncer de mama se conoce. En el contexto de la presente descripción, no deben haberse expuesto las células de referencia al tratamiento de interés.

Si la población celular de referencia no contiene ninguna célula de cáncer de mama, una similitud en la expresión de uno o más de los genes *B1194*, *A2282V1*, *A2282V2* y *A2282V3* en la población celular de prueba y la población celular de referencia indica que el tratamiento de interés es eficaz. Sin embargo, una diferencia en la expresión de estos genes en la población de prueba y una población celular de referencia de control normal indica un desenlace o pronóstico clínico menos favorable. De manera similar, si la población celular de referencia contienen células de cáncer de mama, una diferencia entre la expresión de uno o más de los genes de la presente descripción en la población celular de prueba y la población celular de referencia indica que el tratamiento de interés es eficaz, mientras una similitud en la expresión de tales genes en la población de prueba y una población celular de referencia indica un desenlace o pronóstico menos favorable.

Adicionalmente, puede compararse el nivel de expresión de los genes de la presente descripción determinado en una muestra biológica derivado del sujeto obtenida tras el tratamiento (es decir, niveles tras el tratamiento), con el nivel de expresión de uno o más de los genes *B1194*, *A2282V1*, *A2282V2*, y *A2282V3* determinado en una muestra biológica derivada del sujeto obtenida antes del inicio del tratamiento (es decir, niveles antes del tratamiento). Una disminución en el nivel de expresión en una muestra tras el tratamiento indica que el tratamiento de interés es eficaz, mientras que un aumento o mantenimiento en el nivel de expresión en la muestra tras el tratamiento indica un desenlace o pronóstico clínico menos favorable.

Tal como se usa en el presente documento, el término "eficaz" indica que el tratamiento conduce a una reducción en la expresión de un gen regulado por incremento de manera patológica, un aumento en la expresión de un gen regulado por disminución de manera patológica o una disminución en tamaño, prevalencia o potencial metastásico de carcinoma ductal de mama en un sujeto. Cuando se aplica un tratamiento de interés de manera profiláctica, el término "eficaz" significa que el tratamiento retarda o previene que se forme un tumor de mama o retarda, previene o alivia un síntoma del cáncer de mama clínico. La evaluación de los tumores de mama puede realizarse usando protocolos clínicos convencionales.

Además, la eficacia puede determinarse en asociación con cualquier método conocido para diagnosticar o tratar el cáncer de mama. El cáncer de mama puede diagnosticarse, por ejemplo, identificando anomalías sintomáticas, por ejemplo, pérdida de peso, dolor abdominal, dolor de espalda, anorexia, náuseas, vómitos y malestar generalizado, debilidad e ictericia.

VIII. Tratamiento del cáncer de mama

La presente invención proporciona además un método de selección de un compuesto útil en el tratamiento del cáncer de mama usando un polipéptido de la presente invención. Una realización de un método de selección de este tipo comprende las etapas de: (a) poner en contacto un compuesto de prueba con un polipéptido de la presente invención, (b) detectar la actividad de unión entre el polipéptido de la presente invención y el compuesto de prueba, y (c) seleccionar el compuesto de prueba que se une al polipéptido de la presente invención e inhibir su actividad biológica de potenciación de la proliferación celular y la actividad cinasa.

Un polipéptido de la presente invención que va usarse para la selección puede ser un polipéptido recombinante o una proteína derivada de la naturaleza, o un péptido parcial de los mismos. Puede usarse cualquier compuesto de prueba, por ejemplo, extractos celulares, sobrenadante de cultivo celular, productos de microorganismos fermentadores, extractos de organismos marinos, extractos de plantas, proteínas brutas o purificadas, péptidos, compuestos no peptídicos, compuestos micromoleculares sintéticos y compuestos naturales. El polipéptido de la presente invención que va ponerse en contacto con un compuesto de prueba puede ser, por ejemplo, un polipéptido purificado, una proteína soluble, una forma unida a un portador o una proteína de fusión fusionada con otros polipéptidos.

Como método de selección de proteínas, por ejemplo, que se unen a un polipéptido de la presente invención usando un polipéptido de la presente invención, pueden usarse muchos métodos bien conocidos por un experto en la técnica. Puede llevarse a cabo una selección de este tipo usando, por ejemplo, un método de inmunoprecipitación, específicamente, de la siguiente manera. Se expresa un gen que codifica para un polipéptido de la presente invención en células animales insertando el gen en un vector de expresión para genes foráneos, tal como pSV2neo, pcDNA I y pCD8. El promotor que va usarse para la expresión puede ser cualquier promotor que se usa comúnmente e incluye, por ejemplo, el promotor temprano de SV40 (Rigby en Williamson (ed.), Genetic Engineering, vol. 3. Academic Press, Londres, 83-141 (1982)), el promotor EF-1 α (Kim *et al.*, Gene 91: 217-23 (1990)), el promotor CAG (Niwa *et al.*, Gene 108: 193-9 (1991)), el promotor LTR de VSR (Cullen, Methods in Enzymology 152: 684-704 (1987)), el promotor SR α (Takebe *et al.*, Mol Cell Biol 8: 466-72 (1988)), el promotor temprano inmediato de CMV (Seed y Aruffo, Proc Natl Acad Sci USA 84: 3365-9 (1987)), el promotor tardío de SV40 (Gheysen y Fiers, J Mol Appl Genet 1: 385-94 (1982)), el

promotor tardío de adenovirus (Kaufman *et al.*, Mol Cell Biol 9: 946-58 (1989)), el promotor TK de VHS, etc. La introducción del gen en células animales para expresar un gen foráneo puede realizarse según cualquier método, por ejemplo, el método de electroporación (Chu *et al.*, Nucleic Acids Res 15: 1311-26 (1987)), el método de fosfato de calcio (Chen y Okayama, Mol Cell Biol 7: 2745-52 (1987)), el método de DEAE-dextrano (Lopata *et al.*, Nucleic Acids Res 12: 5707-17 (1984); Sussman y Milman, Mol Cell Biol 4: 1641-3 (1984)), el método de Lipofectin (Derijard B, Cell 76: 1025-37 (1994); Lamb *et al.*, Nature Genetics 5: 22-30 (1993); Rabindran *et al.*, Science 259: 230-4 (1993)), etc. El polipéptido de la presente invención puede expresarse como una proteína de fusión que comprende un sitio de reconocimiento (epítipo) de un anticuerpo monoclonal introduciendo el epítipo del anticuerpo monoclonal, cuya especificidad se ha revelado, en el extremo N- o C-terminal del polipéptido de la presente invención. Puede usarse un sistema de epítipo-anticuerpo comercialmente disponible (Experimental Medicine 13: 85-90 (1995)). Están comercialmente disponibles vectores que pueden expresar una proteína de fusión con, por ejemplo, β -galactosidasa, proteína de unión a maltosa, glutatión S-transferasa, proteína fluorescente verde (GFP) etc. mediante el uso de sus múltiples sitios de clonación.

También se notifica una proteína de fusión preparada introduciendo sólo epítipos pequeños que consisten en de varios a una docena de aminoácidos de modo que no se cambia la propiedad del polipéptido de la presente invención mediante la fusión. Pueden usarse epítipos, tales como polihistidina (etiqueta de His), HA de agregado de influenza, c-myc humano, FLAG, glicoproteína del virus de la estomatitis vesicular (VSV-GP), proteína del gen 10 de T7 (etiqueta de T7), glicoproteína del virus del herpes simple humano (etiqueta de VHS), etiqueta de E (un epítipo en fago monoclonal), y similares, y pueden usarse anticuerpos monoclonales que los reconocen como sistema de epítipo-anticuerpo para seleccionar proteínas que se unen al polipéptido de la presente invención (Experimental Medicine 13: 85-90 (1995)).

En la inmunoprecipitación, se forma un inmunocomplejo añadiendo estos anticuerpos al lisado celular preparado usando un detergente apropiado. El inmunocomplejo consiste en un polipéptido de la presente invención, un polipéptido que tiene una afinidad de unión para el polipéptido, y un anticuerpo. La inmunoprecipitación también puede llevarse a cabo usando anticuerpos contra un polipéptido de la presente invención, además de usar anticuerpos contra los epítipos anteriores, anticuerpos que pueden prepararse tal como se describió anteriormente.

Puede precipitarse un inmunocomplejo, por ejemplo, mediante proteína A sefarosa o proteína G sefarosa cuando el anticuerpo es un anticuerpo IgG de ratón. Si se prepara el polipéptido de la presente invención como una proteína de fusión con un epítipo, tal como GST, puede formarse un inmunocomplejo de la misma manera que en el uso del anticuerpo contra el polipéptido de la presente invención, usando una sustancia que se une específicamente a estos epítipos, tal como glutatión-sefarosa 4B.

La inmunoprecipitación puede realizarse siguiendo o según, por ejemplo, los métodos en la bibliografía (Harlow y Lane, Antibodies, 511-52, publicaciones de Cold Spring Harbor Laboratory, Nueva York (1988)).

Se usa comúnmente SDS-PAGE para el análisis de proteínas inmunoprecipitadas y la proteína unida puede analizarse mediante el peso molecular de la proteína usando geles con una concentración apropiada. Dado que la proteína unida al polipéptido de la presente invención es difícil de detectar mediante un método de tinción común, tal como tinción con Coomassie o tinción con plata, puede mejorarse la sensibilidad de la detección para la proteína cultivando células en medio de cultivo que contiene isótopo radioactivo, ³⁵S-metionina o ³⁵S-cisteína, marcando las proteínas en las células y detectando las proteínas. Puede purificarse la proteína diana directamente del gel de SDS-poliacrilamida y puede determinarse su secuencia, cuando se ha revelado el peso molecular de una proteína.

Como método para seleccionar proteínas que se unen al polipéptido usando el polipéptido, por ejemplo, puede usarse el análisis de inmunotransferencia de tipo West-Western (Skolnik *et al.*, Cell 65: 83-90 (1991)). Específicamente, puede obtenerse una proteína que se une al polipéptido preparando una biblioteca de ADNc de células, tejidos, órganos (por ejemplo, líneas celulares de cáncer de mama para seleccionar proteínas que se unen a A2282V1, A2282V2, A2282V3), o células cultivadas que se espera que expresen una proteína que se une al polipéptido de la presente invención usando un vector de fago (por ejemplo, ZAP), expresando la proteína en LB-agarosa, fijando la proteína expresada en un filtro, haciendo reaccionar el polipéptido purificado y marcado de la presente invención con el filtro anterior y detectando las placas que expresan proteínas unidas al polipéptido de la presente invención según el marcador. El polipéptido puede marcarse utilizando la unión entre biotina y avidina, o utilizando un anticuerpo que se une específicamente al polipéptido, o a un péptido o polipéptido (por ejemplo, GST) que se fusiona al polipéptido. También pueden usarse métodos que usan radioisótopos o fluorescencia.

Alternativamente, en otra realización del método de selección de la presente invención, puede usarse un sistema doble híbrido que utiliza células ("sistema doble híbrido MATCHMAKER", "kit de ensayo de doble híbrido MATCHMAKER para mamíferos", "sistema monohíbrido MATCHMAKER" (Clontech); "Sistema de vector doble híbrido HybriZAP" (Stratagene); las referencias "Dalton y Treisman, Cell 68: 597-612 (1992)", "Fields y Sternglanz, Trends Genet 10: 286-92 (1994)").

En el sistema doble híbrido, se fusiona el polipéptido a la región de unión a SRF o a la región de unión a GAL4 y se expresa en células de levadura. Se prepara una biblioteca de ADNc de células que se espera que expresen una proteína que se une al polipéptido, de manera que la biblioteca, cuando se expresa, se fusiona a la región de activación de la transcripción VP16 o GAL4. Entonces se introduce la biblioteca de ADNc en las células de levadura anteriores y se

aísla el ADNc derivado de la biblioteca de los clones positivos detectados (cuando se expresa una proteína que se une al polipéptido en las células de levadura, la unión de las dos activa un gen indicador, lo que hace que los clones positivos puedan detectarse). Puede prepararse una proteína codificada por el ADNc introduciendo el ADNc aislado anteriormente en *E. coli* y expresando la proteína.

5 Como gen indicador pueden usarse, por ejemplo, el gen Ade2, gen lacZ, gen CAT, gen de la luciferasa y similares además del gen HIS3.

10 Un compuesto que se une a un polipéptido también puede seleccionarse usando cromatografía de afinidad. Por ejemplo, el polipéptido puede inmovilizarse en un portador de una columna de afinidad, y se aplica un compuesto de prueba, que contiene una proteína que puede unirse al polipéptido, a la columna. Un compuesto de prueba en el presente documento puede ser, por ejemplo, extractos celulares, lisados celulares, etc. Tras cargar el compuesto de prueba, se lava la columna y pueden prepararse compuestos unidos al polipéptido.

15 Cuando el compuesto de prueba es una proteína, se analiza la secuencia de aminoácidos de la proteína obtenida, se sintetiza un oligo-ADN basándose en la secuencia y se seleccionan bibliotecas de ADN usando el oligo-ADN como sonda para obtener un ADN que codifica para la proteína.

20 Puede usarse un biosensor que usa el fenómeno de resonancia de plasmón de superficie como medio para detectar o cuantificar el compuesto unido en la presente invención. Cuando se usa un biosensor de este tipo, la interacción entre el polipéptido y un compuesto de prueba puede observarse en tiempo real como una señal de resonancia de plasmón de superficie, usando sólo una cantidad diminuta de polipéptido y sin marcar (por ejemplo, BIAcore, Pharmacia). Por tanto, es posible evaluar la unión entre el polipéptido y un compuesto de prueba usando un biosensor tal como BIAcore.

25 Un experto en la técnica conoce bien los métodos de selección de moléculas que se unen cuando se expone el polipéptido inmovilizado a compuestos químicos sintéticos, o bancos de sustancias naturales, o una biblioteca de presentación de péptidos en fagos al azar, y los métodos de selección que usan alto rendimiento basados en técnicas de química combinatoria (Wrighton *et al.*, Science 273: 458-63 (1996); Verdine, Nature 384: 11-13 (1996)) para aislar no sólo proteínas sino compuestos químicos que se unen a la proteína (incluyendo agonistas y antagonistas).

Alternativamente, el método de selección de la presente invención puede comprender las siguientes etapas:

30 (a) poner en contacto un compuesto candidato con una célula en la que se ha introducido un vector que comprende la región reguladora de la transcripción de uno o más genes marcadores y un gen indicador que se expresa bajo el control de la región reguladora de la transcripción, en el que se seleccionan uno o más genes marcadores del grupo que consiste en A2282V1, A2282V2 y A2282V3

(b) medir la expresión o actividad de dicho gen indicador; y

(c) seleccionar un compuesto que reduce el nivel de expresión o actividad de dicho gen indicador en comparación con el nivel de expresión o actividad de dicho gen indicador detectado en ausencia del compuesto de prueba.

35 En la técnica se conocen bien células huésped y genes indicadores adecuados. El constructo indicador requerido para la selección puede prepararse usando la región reguladora de la transcripción de un gen marcador. Cuando los expertos en la técnica conocen la región reguladora de la transcripción de un gen marcador, puede prepararse un constructo indicador usando la información de secuencia previa. Cuando la región reguladora de la transcripción de un gen marcador sigue sin identificarse, puede aislarse un segmento de nucleótidos que contiene la región reguladora de la transcripción de una biblioteca de genoma basándose en la información de la secuencia de nucleótidos del gen marcador.

40 Un compuesto aislado mediante la selección es un candidato para fármacos que inhiben la actividad de un polipéptido que, a su vez, puede usarse para tratar o prevenir el cáncer de mama. Un compuesto en el que una parte de la estructura del compuesto obtenido por el presente método de selección que tiene la actividad de unión a un polipéptido se convierte mediante adición, delección y/o sustitución, se incluye en los compuestos obtenidos mediante el método de selección de la presente invención.

45 En una realización adicional, la presente descripción proporciona métodos para seleccionar agentes candidatos que son dianas potenciales en el tratamiento del cáncer de mama. Tal como se trató en detalle anteriormente, controlando el nivel de expresión de la proteína B1194, A2282V1, A2282V2 o A2282V3, puede controlarse el inicio y la evolución del cáncer de mama. Por tanto, pueden identificarse agentes candidatos, que son dianas potenciales en el tratamiento del cáncer de mama, mediante selecciones que usan los niveles de expresión y las actividades de A2282V1, A2282V2, A2282V3 como índices. En el contexto de la presente invención, tal selección puede comprender, por ejemplo, las siguientes etapas:

55 (a) poner en contacto un compuesto candidato con una célula que expresa la proteína A2282V1, A2282V2 o A2282V3 y

(b) seleccionar un compuesto que reduce el nivel de expresión de A2282V1, A2282V2 o A2282V3 en comparación con el nivel de expresión detectado en ausencia del compuesto de prueba.

Las células que expresan al menos una de las proteínas A2282V1, A2282V2 o A2282V3 incluyen, por ejemplo, líneas celulares establecidas de cáncer de mama; tales células pueden usarse para la selección anterior de la presente invención. Puede estimarse el nivel de expresión mediante métodos bien conocidos por el experto en la técnica. En el método de selección, puede seleccionarse un compuesto que reduce el nivel de expresión de al menos una de A2282V1, A2282V2 o A2282V3 como agentes candidatos.

En otra realización del método para seleccionar un compuesto útil en el tratamiento del cáncer de mama de la presente invención, el método utiliza la actividad biológica de un polipéptido como índice; es decir la potenciación de la proliferación celular y la actividad cinasa. Dado que las proteínas A2282V1, A2282V2 y A2282V3 tienen la actividad de promover la proliferación celular, puede seleccionarse un compuesto que inhibe la actividad de una de estas proteínas usando esta actividad como índice. Además, en la presente invención, se confirmó que las proteínas A2282V1, A2282V2 y A2282V3 tienen actividad proteína cinasa. Por tanto, un compuesto que inhibe la actividad proteína cinasa de una de las proteínas A2282V1, A2282V2 o A2282V3 puede seleccionarse usando tal actividad como índice. Este método de selección incluye las etapas de: (a) poner en contacto un compuesto de prueba con el polipéptido; (b) detectar la actividad biológica del polipéptido de la etapa (a), es decir la potenciación de la proliferación celular y la actividad cinasa; y (c) seleccionar un compuesto que suprime dicha actividad biológica del polipéptido en comparación con la actividad biológica detectada en ausencia del compuesto de prueba.

Puede usarse cualquier polipéptido para seleccionar siempre que comprendan la actividad biológica de la proteína A2282V1, A2282V2 o A2282V3. Tal actividad biológica incluye la actividad de proliferación celular de la proteína A2282V1, A2282V2, A2282V3 humana o la actividad proteína cinasa de la proteína A2282V1, A2282V2 o A2282V3. Por ejemplo, puede usarse una proteína A2282V1, A2282V2, A2282V3 humana y también pueden usarse polipéptidos funcionalmente equivalentes a estas proteínas. Tales polipéptidos pueden expresarse de manera endógena o exógena por las células.

En la presente invención, la actividad biológica de la proteína A2282V1, A2282V2 o A2282V3 es preferiblemente actividad proteína cinasa. El experto puede estimar la actividad proteína cinasa. Por ejemplo, puede ponerse en contacto una célula que expresa al menos una de las proteínas A2282V1, A2282V2 o A2282V3 con un compuesto de prueba en presencia de [γ -³²P]-ATP. A continuación, pueden determinarse las proteínas fosforiladas mediante la actividad proteína cinasa de la proteína A2282V1, A2282V2 o A2282V3. Para la detección de la proteína fosforilada, puede usarse SDS-PAGE o inmunoprecipitación. Además, puede usarse un anticuerpo que reconoce residuo de tirosina fosforilado para determinar el nivel de proteína fosforilada.

Puede usarse cualquier compuesto de prueba, por ejemplo, extractos celulares, sobrenadante de cultivo celular, productos de microorganismos fermentadores, extractos de organismos marinos, extractos de plantas, proteínas brutas o purificadas, péptidos, compuestos no péptidos, compuestos micromoleculares sintéticos, compuestos naturales.

El compuesto aislado mediante esta selección es un candidato para antagonistas del polipéptido. Asimismo, el término "antagonista" se refiere a moléculas que inhiben la función del polipéptido uniéndose al mismo. Además, un compuesto aislado mediante esta selección es un candidato para compuestos que inhiben la interacción *in vivo* del polipéptido con moléculas (incluyendo ADN y proteínas).

Cuando la actividad biológica que va a detectarse en el presente método es proliferación celular, puede detectarse, por ejemplo, preparando células que expresan el polipéptido, cultivando las células en presencia de un compuesto de prueba y determinando la velocidad de proliferación celular, midiendo el ciclo celular y similares, así como midiendo la actividad formadora de colonias tal como se describe en los ejemplos.

IX. Compuestos aislados y composiciones farmacéuticas

Un compuesto aislado mediante las selecciones anteriores es un candidato para fármacos que inhiben la actividad del polipéptido y puede aplicarse al tratamiento del cáncer de mama. Más particularmente, cuando se usa la actividad biológica de la proteína B1194, A2282V1, A2282V2 o A2282V3 como índice, los compuestos seleccionados mediante el presente método sirven como candidato para fármacos para el tratamiento del cáncer de mama.

Además, también se incluyen en los compuestos que pueden obtenerse mediante el método de selección de la presente invención compuestos en los que una parte de la estructura del compuesto que inhibe la actividad de la proteína B1194, A2282V1, A2282V2 o A2282V3 se convierte mediante adición, delección y/o sustitución.

Cuando se administra un compuesto aislado mediante los métodos de la invención como agente farmacéutico para seres humanos y otros mamíferos, tales como ratones, ratas, cobayas, conejos, gatos, perros, ovejas, cerdos, ganado, monos, babuinos, chimpancés, para tratar el cáncer de mama, el compuesto aislado puede administrarse directamente o puede formularse en una forma farmacéutica usando métodos de preparación farmacéuticos conocidos. Por ejemplo, según la necesidad, los fármacos pueden tomarse por vía oral, como comprimidos recubiertos con azúcar,

5 cápsulas, elixires y microcápsulas, o por vía no oral, en forma de inyecciones de disoluciones o suspensiones estériles con agua o cualquier otro líquido farmacéuticamente aceptable. Por ejemplo, los compuestos pueden mezclarse con medio o portadores farmacológicamente aceptables, específicamente, agua esterilizada, solución salina fisiológica, aceite vegetal, emulsionantes, agentes de suspensión, tensioactivos, estabilizantes, agentes aromatizantes, excipientes, vehículos, conservantes, aglutinantes y similares, en una forma de dosis unitaria requerida para la implementación de fármacos generalmente aceptada. La cantidad de principios activos en estas preparaciones produce una dosificación adecuada dentro del intervalo indicado obtenible.

10 Ejemplos de aditivos que pueden mezclarse para preparar comprimidos y cápsulas son aglutinantes tales como gelatina, almidón de maíz, goma tragacanto y goma arábica; excipientes tales como celulosa cristalina; agentes de hinchamiento tales como almidón de maíz, gelatina y ácido alginico; lubricantes tales como estearato de magnesio; edulcorantes tales como sacarosa, lactosa o sacarina; agentes aromatizantes tales como menta, aceite de *Gaultheria adenothrix* y cereza. Cuando la forma farmacéutica unitaria es una cápsula, puede incluirse también adicionalmente un portador líquido, tal como aceite, en los componentes anteriores. Pueden formularse materiales compuestos estériles para inyecciones siguiendo implementaciones de fármacos normales usando vehículos tales como agua destilada usada para inyecciones.

15 Puede usarse solución salina fisiológica, glucosa y otros líquidos isotónicos que incluyen adyuvantes, tales como D-sorbitol, D-manosa, D-manitol y cloruro de sodio, como disoluciones acuosas para inyecciones. Estas pueden usarse conjuntamente con solubilizantes adecuados, tales como alcohol, específicamente etanol, polialcoholes tales como propilenglicol y polietilenglicol, tensioactivos no iónicos, tales como polisorbato 80 (TM) y HCO-50.

20 Puede usarse aceite de sésamo o aceite de soja como líquido oleaginoso y puede usarse conjuntamente con benzoato de bencilo o alcohol bencilico como solubilizante y puede formularse con un tampón, tal como tampón fosfato y tampón acetato de sodio; un analgésico, tal como clorhidrato de procaína; un estabilizante, tal como alcohol bencilico, fenol; y un antioxidante. Puede llenarse una ampolla adecuada con la inyección preparada.

25 Pueden usarse métodos bien conocidos por el experto en la técnica para administrar el compuesto farmacéutico de la invención a pacientes, por ejemplo como inyecciones intraarteriales, intravenosas, percutáneas y también como administraciones intranasales, transbronquiales, intramusculares u orales. La dosificación y el método de administración varían según el peso corporal y la edad de un paciente y el método de administración; sin embargo, un experto en la técnica puede seleccionarlos de manera rutinaria. Si dicho compuesto puede codificarse por un ADN, el ADN puede insertarse en un vector para terapia génica y administrarse el vector para realizar la terapia. La dosificación y el método de administración varían según el peso corporal, la edad y los síntomas de un paciente pero un experto en la técnica puede seleccionarlos adecuadamente.

30 Por ejemplo, aunque existen algunas diferencias según los síntomas, la dosis de un compuesto que se une con el polipéptido de la presente invención y regula su actividad es de aproximadamente 0,1 mg a aproximadamente 100 mg al día, preferiblemente de aproximadamente 1,0 mg a aproximadamente 50 mg al día y más preferiblemente de aproximadamente 1,0 mg a aproximadamente 20 mg al día, cuando se administra por vía oral a un adulto normal (peso de 60 kg).

35 Cuando se administra por vía parenteral, en forma de una inyección a un adulto normal (peso de 60 kg), aunque existen algunas diferencias según el paciente, el órgano diana, los síntomas y el método de administración, es conveniente inyectar por vía intravenosa una dosis de aproximadamente 0,01 mg a aproximadamente 30 mg al día, preferiblemente de aproximadamente 0,1 a aproximadamente 20 mg al día y más preferiblemente de aproximadamente 0,1 a aproximadamente 10 mg al día. Además, en el caso de otros animales también, es posible administrar una cantidad convertida a 60 kg de peso corporal.

X. Métodos de inducción de inmunidad antitumoral y vacunas tumorales

45 La presente descripción también se refiere a un método de inducción de inmunidad antitumoral que comprende la etapa de administrar una proteína B1194, A2282V1, A2282V2 o A2282V3 o un fragmento inmunológicamente activo de la misma, o un polinucleótido que codifica para la proteína o fragmentos del mismo. Las proteínas B1194, A2282V1, A2282V2 y A2282V3 o los fragmentos inmunológicamente activos de las mismas son útiles como vacunas contra el cáncer de mama. En algunos casos, pueden administrarse las proteínas o los fragmentos de las mismas en una forma unida al receptor de células T (TCR) o presentarse mediante una célula presentadora de antígenos (APC), tal como macrófago, célula dendrítica (CD) o células B. Debido a la fuerte capacidad de presentación de antígenos de CD, el uso de CD es lo más preferible entre las APC. En la presente descripción, una vacuna contra el cáncer de mama se refiere a una sustancia que tiene la función de inducir inmunidad antitumoral tras su inoculación en animales. En general, la inmunidad antitumoral incluye respuestas inmunitarias tales como las siguientes:

- inducción de linfocitos citotóxicos contra cáncer de mama,
- 55 - inducción de anticuerpos que reconocen cáncer de mama, e
- inducción de la producción de citocinas antitumorales.

Por tanto, cuando una determinada proteína induce una cualquiera de estas respuestas inmunitarias tras su inoculación en un animal, se considera que la proteína tiene un efecto inductor de inmunidad antitumoral. La inducción de la inmunidad antitumoral mediante una proteína puede detectarse observando *in vivo* o *in vitro* la respuesta del sistema inmunitario en el huésped contra la proteína.

5 Por ejemplo, se conoce bien un método para detectar la inducción de linfocitos T citotóxicos. Se presenta una sustancia foránea que ingresa en el organismo vivo a células T y células B mediante la acción de las células presentadoras de antígenos (APC). Las células T que responden al antígeno presentado por la APC de una manera específica de antígeno, se diferencian en células T citotóxicas (o linfocitos T citotóxicos; CTL) debido a la estimulación por el antígeno, y entonces proliferan (esto se denomina activación de células T). Por tanto, la inducción de CTL mediante un determinado péptido puede evaluarse presentando el péptido a la célula T mediante la APC, y detectando la inducción de CTL. Además, la APC tiene el efecto de activación de células T CD4+ y células T CD8+, macrófagos, eosinófilos y células NK. Dado que las células T CD4+ y las células T CD8+ también son importantes en la inmunidad antitumoral, la acción de inducción de inmunidad antitumoral del péptido puede evaluarse usando el efecto de activación de estas células como indicadores.

15 En la técnica se conoce bien un método para evaluar la acción de inducción de CTL usando células dendríticas (CD) como APC. La CD es una APC representativa que tiene la acción de inducción de CTL más fuerte entre las APC. En este método, el polipéptido de prueba se pone en contacto inicialmente con la CD, y luego esta CD se pone en contacto con células T. La detección de células T que tienen efectos citotóxicos contra las células de interés tras el contacto con CD muestra que el polipéptido de prueba tiene una actividad de inducción de las células T citotóxicas. Puede detectarse actividad de CTL contra tumores, por ejemplo, usando la lisis de células tumorales marcadas con ⁵¹Cr como indicador. Alternativamente, también se conoce bien el método de evaluación del grado de daño de células tumorales usando la actividad de captación de ³H-timidina o liberación de LDH (lactato deshidrogenasa) como indicador.

25 Aparte de las CD, también pueden usarse células mononucleares de sangre periférica (PBMC) como la APC. Se notifica que la inducción de CTL puede potenciarse cultivando PBMC en presencia de GM-CSF e IL-4. De manera similar, se ha demostrado que pueden inducirse CTL cultivando PBMC en presencia de hemocianina de lapa californiana (KLH) e IL-7.

30 Los polipéptidos de prueba que se confirma que poseen actividad de inducción de CTL mediante estos métodos son polipéptidos que tienen un efecto de activación de CD y una actividad de inducción de CTL posterior. Por tanto, los polipéptidos que inducen CTL contra células tumorales son útiles como vacunas contra tumores. Además, las APC que adquieren la capacidad para inducir CTL contra tumores poniéndose en contacto con los polipéptidos son útiles como vacunas contra tumores. Además, también puede usarse CTL que adquirieron citotoxicidad debido a la presentación de los antígenos del polipéptido por la APC como vacunas contra tumores. Tales métodos terapéuticos para tumores usando la inmunidad antitumoral debida a APC y CTL se denominan inmunoterapia celular.

35 Generalmente, cuando se usa un polipéptido para inmunoterapia celular, se sabe que la eficiencia de la inducción de CTL aumenta combinando una pluralidad de polipéptidos que tienen diferentes estructuras y poniéndolos en contacto con CD. Por tanto, cuando se estimulan CD con fragmentos de proteínas, es ventajoso usar una mezcla de múltiples tipos de fragmentos.

40 Alternativamente, la inducción de inmunidad antitumoral mediante un polipéptido puede confirmarse observando la inducción de la producción de anticuerpos contra tumores. Por ejemplo, cuando se inducen anticuerpos contra un polipéptido en un animal de laboratorio inmunizado con el polipéptido, y cuando se suprime el crecimiento de células tumorales mediante esos anticuerpos, puede determinarse que el polipéptido tiene una capacidad para inducir inmunidad antitumoral.

45 Se induce inmunidad antitumoral administrando la vacuna dada a conocer y la inducción de inmunidad antitumoral permite el tratamiento y la prevención del cáncer de mama. La terapia contra el cáncer o la prevención del inicio del cáncer incluye cualquiera de las etapas, tales como inhibición del crecimiento de células cancerosas, involución del cáncer y supresión de la aparición del cáncer. La disminución en la mortalidad de individuos que tienen cáncer, la disminución de marcadores tumorales en la sangre, el alivio de síntomas detectables que acompañan al cáncer y similares también se incluyen en la terapia o la prevención del cáncer. Tales efectos terapéuticos y de prevención son preferiblemente estadísticamente significativos. Por ejemplo, en la observación, a un nivel de significación del 5% o menos, en la que el efecto terapéutico o de prevención de una vacuna contra el cáncer de mama se compara con un control sin administración de vacunas. Por ejemplo, pueden usarse la prueba de la t de Student, la prueba U de Mann-Whitney o ANOVA para los análisis estadísticos.

50 Puede combinarse la proteína anteriormente mencionada que tiene actividad inmunológica o un vector que codifica para la proteína con un adyuvante. Un adyuvante se refiere a un compuesto que potencia la respuesta inmunitaria contra la proteína cuando se administran junto (o de manera sucesiva) con la proteína que tiene actividad inmunológica. Los ejemplos de adyuvantes incluyen toxina del cólera, toxina de *Salmonella*, alumbre y similares, pero sin limitarse a los mismos. Además, puede combinarse apropiadamente la vacuna dada a conocer con un portador farmacéuticamente aceptable. Ejemplos de tales portadores son agua esterilizada, solución salina fisiológica, tampón fosfato, líquido de cultivo y similares. Además, la vacuna puede contener, si es necesario, estabilizantes suspensiones,

conservantes, tensioactivos y similares. La vacuna se administra de manera sistémica o local. La administración de la vacuna puede realizarse mediante administración única, o reforzarse mediante múltiples administraciones.

5 Cuando se usan APC o CTL como la vacuna de esta descripción, pueden tratarse o prevenirse los tumores, por ejemplo, mediante el método *ex vivo*. Más específicamente, se recogen PBMC del sujeto que recibe el tratamiento o la prevención, se ponen en contacto las células con el polipéptido *ex vivo* y, tras la inducción de APC o CTL, las células pueden administrarse al sujeto. También pueden inducirse APC introduciendo un vector que codifica para el polipéptido en PBMC *ex vivo*. Pueden clonarse las APC o CTL inducidas *in vitro* antes de la administración. Clonando y haciendo crecer las células que tienen alta actividad de daño de células diana, puede realizarse más eficazmente la inmunoterapia celular. Además, pueden usarse APC y CTL aislados de esta manera para inmunoterapia celular no sólo contra individuos de quienes se derivan las células, sino también contra tipos similares de tumores de otros individuos.

10 Además, se proporciona una composición farmacéutica para tratar o prevenir el cáncer de mama, que comprende una cantidad farmacéuticamente eficaz del polipéptido de la presente descripción. La composición farmacéutica puede usarse para generar inmunidad antitumoral. La expresión normal de B1194 se restringe a los testículos; no se observa expresión de A2282V1, A2282V2 y A2282V3 en órganos normales. Por tanto, la supresión de estos genes puede no afectar de manera adversa a otros órganos. Por tanto, los polipéptidos de B1194, A2282V1, A2282V2 y A2282V3 son preferibles para tratar el cáncer de mama. En la presente descripción, se administra el polipéptido o fragmento del mismo a una dosificación suficiente para inducir inmunidad antitumoral, que está en el intervalo de 0,1 mg a 10 mg, preferiblemente de 0,3 mg a 5 mg, más preferiblemente de 0,8 mg a 1,5 mg. Las administraciones se repiten. Por ejemplo, puede administrarse 1 mg del péptido o fragmento del mismo 4 veces cada dos semanas para inducir la actividad antitumoral.

15 A continuación en la presente invención, se describe la presente invención en más detalle mediante referencia a los ejemplos.

EJEMPLOS

20 Tal como puede apreciarse a partir de la descripción proporcionada anteriormente, la presente invención tiene una gran variedad de aplicaciones.

La presente invención se ilustra en detalle mediante los siguientes ejemplos.

Ejemplo 1 - Materiales y Métodos

(1) Líneas celulares y materiales clínicos

30 Se adquirieron líneas celulares de cáncer de mama humano HBL100, HCC1937, MCF7, MDA-MB-435S, YMB1, SKBR3, T47D, líneas celulares HeLa y COS-7 de adenocarcinoma cervical de la Colección Americana de Cultivos Tipo (ATCC) y se cultivaron según las recomendaciones del depositante respectivo. El Dr. Yamori de Farmacología Molecular, Centro de Quimioterapia del Cáncer de la Fundación Japonesa para la Investigación del Cáncer ("*Molecular Pharmacology, Cancer Chemotherapy Center of the Japanese Foundation for Cancer Research*") donó gentilmente las líneas celulares HBC4, HBC5 y MDA-MB-231.

35 Se cultivaron todas las células en medios apropiados; es decir RPMI-1640 (Sigma, St. Louis, MO) para HBC4, HBC5, SKBR3, T47D, YMB1 y HCC1937 (con L-glutamina 2 mM); medio de Eagle modificado por Dulbecco (Invitrogen, Carlsbad, CA) para HBL100, COS7; EMEM (Sigma) con aminoácido esencial 0,1 mM (Roche), piruvato de sodio 1 mM (Roche), insulina 0,01 mg/ml (Sigma) para MCF-7; L-15 (Roche) para MDA-MB-231 y MDA-MB-435S. Se complementó cada medio de cultivo con suero bovino fetal al 10% (Cansera) y disolución de antibiótico/antimicótico al 1% (Sigma). Se mantuvieron las células MDA-MB-231 y MDA-MB-435S a 37°C en una atmósfera de aire humidificado sin CO₂. Se mantuvieron otras líneas celulares a 37°C en una atmósfera de aire humedecido con CO₂ al 5%. Se obtuvieron muestras clínicas (cáncer de mama y conducto de mama normal) de especímenes quirúrgicos, sobre lo cual todos los pacientes habían dado su consentimiento informado.

(2) Aislamiento de genes humanos novedosos en la micromatriz de ADNc

45 La fabricación de los portaobjetos con micromatrices de ADNc se ha descrito en otra parte (Ono K, *et al.*, (2000). *Cancer Res*, 60, 5007-5011). Para cada análisis de perfiles de expresión, se prepararon conjuntos duplicados de portaobjetos que contenían 27.648 manchas de ADNc para reducir la fluctuación experimental. En resumen, se purificaron los ARN totales de cada muestra de células microdisecionadas con láser y se llevó a cabo la amplificación del ARN basada en T7 para obtener cantidades adecuadas de ARN para los experimentos con micromatrices. Se marcaron alícuotas de ARN amplificado de células de cáncer de mama y células del conducto de mama normales mediante transcripción inversa con Cy5-dCTP y Cy3-dCTP, respectivamente (Amersham Biosciences, Buckinghamshire, RU). Se llevaron a cabo la hibridación, el lavado y la detección tal como se describió anteriormente (Ono K, *et al.*, (2000). *Cancer Res*, 60, 5007-5011). Para detectar genes que comúnmente estaban regulados por incremento en el cáncer de mama, se examinaron los patrones de expresión globales de los 27.648 genes en la micromatriz para seleccionar aquellos con razones de expresión > 2,0 que estuvieran presentes en >50% de todos los

77 casos de cáncer de mama premenopáusicos. Finalmente, se identificó un total de 468 genes que parecía que estaban regulados por incremento en células tumorales.

Para detectar genes que comúnmente estaban regulados por incremento en el cáncer de mama, se examinaron los patrones de expresión globales de los 27.648 genes en la micromatriz para seleccionar aquellos con razones de expresión > 3,0 que estuvieran presentes en >50% de i) todos los 77 casos de cáncer de mama premenopáusicos, ii) 69 carcinomas ductales invasivos, iii) 31 lesiones bien diferenciadas, iv) 14 moderadamente diferenciadas, o v) 24 escasamente diferenciadas, respectivamente. Finalmente, se seleccionó el total de 493 genes que parecía que estaban regulados por incremento en las células tumorales.

(3) Análisis de RT-PCR semicuantitativa

Se extrajo el ARN total de cada población de células capturas por láser y entonces se realizó una amplificación basada en T7 y una transcripción inversa tal como se describió anteriormente (Kitahara O, *et al.*, Cancer Res 61, 3544-3549 (2001)). Se prepararon diluciones adecuadas de cada ADNc monocatenario para la posterior amplificación por PCR monitorizando la gliceraldehído-3-fosfato deshidrogenasa (GAPDH) como control interno cuantitativo. Las secuencias de los cebadores de PCR fueron 5'-CGACCACTTTGTCAAGCTCA-3' (SEQ ID NO.9) y 5'-GGTTGAGCACAGGGTACTTTATT-3'(SEQ ID NO.10) para GAPDH; y 5'-TGGGTAACAAGAGAATGGTTCA-3'(SEQ ID NO.11) y 5'-ATCCAAGTCCTAATCCCTTTGG-3' (SEQ ID NO.12) para B1194; y 5'-GCTGCAAGGTATAATTGATGGA-3' (SEQ ID NO.13) y 5'-CAGTAACATAATGACAGATGGGC-3' (SEQ ID NO. 14) para A2282.

(4) Análisis de transferencia de tipo Northern

Se extrajo los ARN totales de todas las líneas celulares de cáncer de mama usando el kit RNeasy (QIAGEN) según las instrucciones del fabricante. Tras el tratamiento con ADNsa I (Nippon Gene, Osaka, Japón), se aisló el ARNm con el kit de purificación de ARNm (Amersham Biosciences) siguiendo las instrucciones del fabricante. Se separó una alícuota de 1 µg de cada ARNm, junto con ARN poliA(+) aislados de pulmón, corazón, hígado, riñón, médula ósea, cerebro de seres humanos adultos normales (BD, Clontech, Palo Alto, CA), en geles de agarosa desnaturalizantes al 1% y se transfirieron a membranas de nailon (transferencias de tipo Northern de cáncer de mama). Se hibridaron las transferencias de tipo Northern de múltiples tejido humanos (Clontech, Palo Alto, CA) y la transferencia de tipo Northern de cáncer de mama con productos de PCR marcados con [α^{32} P]-dCTP de B1194 y A2282 preparados mediante RT-PCR (véase a continuación). Se realizaron la prehibridación, la hibridación y el lavado según las recomendaciones del proveedor. Las transferencias se sometieron a autorradiografía con pantallas intensificadoras a -80°C durante 14 días. Se prepararon sondas específicas para B1194 (411 pb) mediante PCR utilizando un conjunto de cebadores; 5'-TGGGTAACAAGAGAATGGTTCA-3'(SEQ ID NO.11) y 5'-ATCCAAGTCCTAATCCCTTTGG-3'(SEQ ID NO.12); para las variantes 1, 2 y 3 de A2282 (554 pb) y para las variantes 1 y 2 de A2282 (170 pb) se prepararon mediante PCR usando los siguientes conjuntos de cebadores; 554 pb: 5'-TTATCACTGTGCTCACCAGGAG-3' (SEQ ID NO.15) y 5'-CAGTAACATAATGACAGATGGGC-3' (SEQ ID NO.14); 170 pb 5'-AAACTTGCCTGCCATATCCTTA-3' (SEQ ID NO.16) y 5'-ATTTTGTGGCTGTCTCTAGCA-3' (SEQ ID NO.17), y se marcaron de manera radioactiva con un sistema de marcaje de ADN Megaprime (Amersham bioscience).

(5) Selección de la biblioteca de ADNc

Se construyó una biblioteca de ADNc usando un sistema de plásmidos Superscript™ con tecnología Gateway™ para la síntesis de ADNc y el kit de clonación (Invitrogen), y se obtuvo el ARN poli(A)+ de la línea celular de cáncer de mama T47D, y se seleccionaron 3 X 10⁶ clones independientes de esta biblioteca con una sonda de ADNc correspondiente al nucleótido 1-1112 de la variante V1.

(6) Ensayo de traducción *in vitro*

Se clonaron las cuatro variantes (V1, V2, V3 y V4) de A2282 en el vector de expresión pSPORT-1 usado construyendo una biblioteca de ADNc como anteriormente y entonces se usaron como moldes para experimentos de transcripción/traducción *in vitro*. Se transcribieron y tradujeron los plásmidos (1 µg) usando sistemas de lisados de reticulocitos acoplados a TNT (Promega, Madison, WI) en presencia de ARNt de lisina biotinilado ϵ -marcado según las instrucciones del fabricante. Se separaron los productos de proteína mediante electroforesis en geles de SDS-poliacrilamida en gradiente del 5-20%. Tras la electrotransferencia, se visualizan las proteínas biotiniladas mediante su unión a estreptavidina-peroxidasa del rábano, seguido de detección por quimioluminiscencia (Amersham Biosciences).

(7) Construcción de vectores de expresión

Para la construcción de un vector de expresión de mamíferos, se amplificó toda la secuencia codificante del ADNc de B1194 mediante PCR usando ADN polimerasa KOD-Plus (Toyobo, Osaka, Japón) con conjuntos de cebadores; 5'-AAAGAATTCGGGTGTCGTTAATGTTCCGGG-3' (SEQ ID NO.18); y 5'-AAAGCGCCGCTTAGCGGATTTTCCTGCA-3' (SEQ ID NO.19). Se insertaron los productos de PCR en los sitios EcoRI y Not I del vector de expresión pCMV-N-myc (Clontech).

5 Para construir las variantes V1, V2 y V3 de los vectores de expresión de A2282, se amplificó toda la secuencia codificante de cada variante de ADNc de A2282 mediante PCR usando ADN polimerasa KOD-Plus (Toyobo, Osaka, Japón) con los siguientes conjuntos de cebadores; 5'-CGGAATTCACCTATGAAAGATTATGATGAAC-3' (SEQ ID NO.20); y 5'-AAACTCGAGTACCTTGCAGCTAGATAGGAT-3' (SEQ ID NO.21) porque todas las variantes contienen la misma secuencia en 5' de ORF. Se insertaron los productos de PCR en los sitios EcoRI y Xho I del vector de expresión pCAGGS-HA. Se confirmaron estos constructos, pCMV-myc-B1194 y pCAGGS-A2282-HA, mediante secuenciación del ADN.

(8) Tinción inmunocitoquímica

10 Para examinar la localización subcelular de B1194, se sembraron células COS7 transfectadas con B1194 a 5 x 10⁴ células por pocillo en el sistema de portaobjetos con cámara Lab-Tek® II (Nalgen Nunc International), y a esto le siguió la fijación con paraformaldehído al 4% en PBS y permeabilización con Triton X-100 al 0,1% en PBS durante 3 min. a 4°C. Tras bloquear con BSA al 3% en PBS durante 1 hora a temperatura ambiente, se incubaron las células con anticuerpos monoclonales 9E10 de ratón antimyc (0,2 µg/ml, Santa Cruz Biotechnology) durante 1 hora a temperatura ambiente. Posteriormente, se tiñeron las células con un anticuerpo secundario de cabra anti-ratón conjugado con FITC antes de la visualización con un microscopio TCS SP2 AOBS (Leica, Tokio, Japón).

(9) Análisis de inmunotransferencia de tipo Western

20 Para examinar la expresión de proteína B1194 y A2282 exógena, se transfectaron de manera transitoria plásmido que expresa de B1194, pCAGGSA2282-HA o pCMV-N-myc (simulado) como control negativo en células COS7 o células HeLa, respectivamente. Se separaron los lisados celulares en geles de SDS-poliacrilamida al 5%-10% y se transfirieron a una membrana de nitrocelulosa, seguido de bloqueo con polvo BlockAce™ (Dainippon Seiyaku) y se trataron con anticuerpos monoclonales 9E10 de ratón anti-myc o anticuerpos de ratón anti-HA (0,4 µg/ml, Santa Cruz Biotechnology) o anticuerpo frente a β-actina monoclonal que sirvió como control de carga para proteínas con una dilución 1:1000 (clon AC-15, Sigma-Aldrich, MO). Tras el lavado, se trataron las transferencias con IgG de asno anti-ratón conjugada con peroxidasa del rábano para el anticuerpo frente a β-actina (Amersham Biosciences) y se visualizaron las proteínas mediante la unión de la peroxidasa del rábano, seguido de detección por quimioluminiscencia (Amersham Biosciences).

(10) Sincronización y análisis por citometría de flujo

30 Se transfectaron células HeLa (1 x 10⁶) con 8 µg del vector de expresión pCAGGS-A2282-HA usando FuGENE6 (Roche) según el protocolo del proveedor. Se detuvieron las células en fase G1 24 horas tras la transfección con afidicolina 1 (µg/ml) durante otras 16 horas. Se liberó el ciclo celular lavando tres veces con medio nuevo y se recogieron las células a los puntos de tiempo indicados. Para detener las células en fase mitótica, se incubaron las células con nocodazol (250 ng/ml) 16 horas antes de la recogida.

35 Para el análisis de FACS, se combinó una alícuota de 400 µl de células desprendidas y adherentes sincronizadas y se fijaron con etanol al 70% a 4°C. Tras el lavado con PBS (-) dos veces, se incubaron las células durante 30 min. con 1 µl de PBS que contenía 1 µg de ARNasa I a 37°C. Entonces se tiñeron las células en 1 ml de PBS que contenía 50 µg de yoduro de propidio (PI). Se determinaron los porcentajes de cada fracción de las fases del ciclo celular de al menos 10000 células en un citómetro de flujo (FACScalibur; Becton Dickinson, San Diego, CA).

(11) Construcción de vectores de expresión de ARNip específicos para B1194 o A2282

40 Se estableció un sistema RNAi basado en vectores usando un vector de expresión de ARNip psiH1BX3.0 según el informe anterior (Shimokawa T, *et al.*, (2003). Cancer Res, 63, 6116-6120). Se preparó un vector de expresión de ARNip contra B1194 (psiH1BX-B1194 Si-1 y Si-5) y A2282 (psiH1BX-A2282 Si-3 y Si-4) clonando oligonucleótidos bicatenarios en la tabla 1 en el sitio BbsI en el vector psiH1BX. Se preparó un plásmido de control, psiH1BX-SC y psiH1BX-LUC, clonando oligonucleótidos bicatenarios de

5'

TCCCGCGCGCTTTGTAGGATTCGTTCAAGAGACGAATCCTACAAAGCGCGC -
 3' (SEQ ID NO.22) y
 5'-
 AAAAGCGCGCTTTGTAGGATTCGTCTCTTGAACGAATCCTACAAAGCGCGC-3'
 (SEQ ID NO.23) para SC (Control de mezcla); y
 5'-TCCCGGTACGCGGAATACTTCGATTCAAGAGATCGAAGTATTCCGCGTACG
 -3'(SEQ ID NO.24) y
 5'-AAAACGTACGCGGAATACTTCGATCTCTTGAATCGAAGTATTCCGCGTACG
 -3' (SEQ ID NO.25)

para LUC (control de luciferasa) en el sitio BbsI en el vector psiH1BX3.0, respectivamente.

Tabla 1 Secuencias de oligonucleótidos bicatenarios específicos insertados en el vector de expresión de ARNip

			SEQ ID No.	
B1194	Si-1	F	5'-TCCCGTATATCTTGCCCTCTGAATTC AAGAGATTCAGAGGGCAAGATATAC-3'	30
		R	5'-AAAAGTATATCTTGCCCTCTGAATCT CTTGAATTCAGAGGGCAAGATATAC-3'	31
	Si-5	F	5'-TCCCGTCCGAACACATCTTTGTTTC AAGAGAAACAAGATGTGTTCCGGAC-3'	32
		R	5'-AAAAGTCCGAACACATCTTTGTTTCT CTTGAAAACAAGATGTGTTCCGGAC-3'	33
A2282	Si-3	F	5'-TCCCGACATCCTATCTAGCTGCATTC AAGAGATGCAGCTAGATAGGATGTC-3'	34
		R	5'-AAAAGACATCCTATCTAGCTGCATCT CTTGAATGCAGCTAGATAGGATGTC-3'	35
	Si-4	F	5'-TCCCAGTTCATTGGAAC TACCAATTC AAGAGATTGGTAGTTCCAATGAACT-3'	36
		R	5'-AAAAAGTTCATTGGAAC TACCAATCT CTTGAATTGGTAGTTCCAATGAACT-3'	37

5 Las secuencias dianas de cada ARNip se muestran en la tabla 2.

Tabla 2

		Secuencia diana	SEQ ID NO.
B1194	Si-1	GTATATCTTGCCCTCTGAA	38
	Si-5	GTCCGAACACATCTTTGTT	39

A2282	Si-3	GACATCCTATCTAGCTGCA	40
	Si-4	AGTTCATTGGAACACCAA	41

(12) Efecto de silenciamiento génico de B1194 o A2282

Se sembraron en placa líneas celulares de cáncer de mama humano, T47D o MCF-7, en placas de 15 cm (4 x 10⁶ células/placa) y se transfectaron con 16 µg de cada psiH1BX-LUC (control de luciferasa), psiH1BX-SC (control mezclado) como controles negativos, psiH1BX-A2282 y psiH1BX-B1194 usando un reactivo FuGENE6 según las recomendaciones del proveedor (Roche). 24 horas tras la transfección, se sembraron de nuevo las células para el ensayo de formación de colonias (3 x 10⁶ células/placa de 10 cm), RT-PCR (1 x 10⁶ células/placa de 10 cm) y ensayo MTT (5 x 10⁵ células/pocillo). Se seleccionaron las células que introdujeron B1194 o A2282 con medio que contenía 0,7 ó 0,6 mg/ml de neomicina (Geneticin, Gibco) en células T47D o MCF-7, respectivamente. Después de eso, se cambió el medio cada dos días durante 3 semanas. Para evaluar el funcionamiento del ARNip, se extrajo el ARN total de las células a los 4 días tras la selección con neomicina, y entonces se confirmó el efecto de desactivación de los ARNip mediante RT-PCR semicuantitativa usando cebadores específicos para B1194 o A2282 y para β2MG; 5'-TTAGCTGTGCTCGCGCTACT-3' (SEQ ID NO.26) y 5'-TCACATGGTTCACACGGCAG-3' (SEQ ID NO.27) para β2MG como control interno; 5'-TTAAGTGAAGGCTCTGATTCTAGTT-3' (SEQ ID NO.28) y 5'-GTCCTTATTGGCTGGTTCGTT-3' (SEQ ID NO.29) para B1194; y 5'-TTATCACTGTGCTCACCAGGAG-3 (SEQ ID NO.15) y 5'-CAGTAACATAATGACAGATGGGC-3' (SEQ ID NO.14) para A2282.

Además, se hicieron crecer transfectantes que expresaban ARNip usando células T47D o MCF-7 durante 28 días en medios selectivos que contenían 0,7 mg/ml de neomicina. Tras la fijación con paraformaldehído al 4% se tiñeron las células transfectadas con solución de Giemsa para evaluar la formación de colonias. Se realizaron ensayos MTT para cuantificar la viabilidad celular. Tras 7 días de cultivo en medio que contenía neomicina, se añadió disolución de MTT (bromuro de 3-(4,5-dimetiltiazol-2-il)-2,5-difenil-tetrazolio) (Sigma) a una concentración de 0,5 mg/ml. Tras la incubación a 37°C durante 2,5 horas, se añadió ácido-SDS (HCl 0,01 N/SDS al 10%); se mezcló la suspensión vigorosamente y entonces se incubó durante la noche a 37°C para disolver los cristales de color azul oscuro. Se midió la absorbancia a 570 nm con un lector de microplacas 550 (BioRad). Se realizó cada experimento por triplicado.

(13) Construcción de la proteína A2282 truncada usando el vector pCAGGS-HA

Se prepararon todos los constructos mediante reacción en cadena de la polimerasa (PCR) según los siguientes conjuntos de cebadores. (Directo de longitud completa: 5'-CGGAATTCATGAAAGATTATGATGAAC-3' (SEQ ID NO; 20); directo n.º 1 truncado: 5'-ACGGAATTCATCATGCAAGATTACAACATATCC-3' (SEQ ID NO; 42); directo n.º 2 truncado: 5'-GACGGAATTCATATGGAGGAGACTCCAAAAG-3' (SEQ ID NO; 43) y cebador inverso común: 5'-CCCTCGAGTACCTTGCAGCTAGATAGGATG-3' (SEQ ID NO; 44)). Entonces, se ligaron los ORF en los sitios de enzimas de restricción EcoRI y Xho I del vector pCAGGS-HA en marco con etiqueta de hemaglutinina (HA).

(14) Cultivo celular para la identificación de posibles sitios de fosforilación

Se cultivaron células HBC5 en una placa de 10 cm hasta que alcanzaron un 80% de confluencia antes de la transfección. Se transfectaron 8 µg de cada constructo en células HBC5 según la recomendación del fabricante (FuGENE 6). 36 horas tras la transfección, se trataron las células con afidicolina 1 µg/ml (Sigma A-0781) y nocodazol 0,5 µg/ml (Sigma M-1404) durante otras 18 horas.

(15) Ensayo con fosfatasa lambda

Se llevó a cabo el ensayo con fosfatasa tal como lo describe el fabricante (New England). En resumen, se incubó el lisado celular con 1 µl de fosfatasa lambda a 30 grados durante 1 hora.

(16) Generación de los vectores de expresión de A2282 mutantes para el ensayo de cinasa *in vitro*

Se generaron los constructos mutantes, D150A, T167A y T478A mediante el kit de mutagénesis dirigida al sitio QuickChange (cat. de Stratagene 200518-5) usando pCAGGS-A2282-HA como molde según la recomendación del fabricante.

(17) Transfección y cultivo celular para inmunoprecipitación

Se transfectaron células HEK 293 con 16 µg de cada constructo. 48 tras la transfección, se llevó a cabo la inmunoprecipitación usando tampón de lisis (Tris-HCl 50 mM (pH 7,5), NaCl 150 mM, NP-40 al 1%, NAF 40 mM, β-glicerofosfato 40 mM, cóctel inhibidor de proteasas) y anticuerpo de rata anti-HA (Roche). Se lavaron perlas de sefarosa 4B-proteína G-REC unidas a proteína (ZYMED) dos veces en tampón de lisis y una vez en tampón de cinasa (Tris-HCl 50 mM, (pH 7,5), MgCl 10 mM, NaCl 25 mM, DTT 1 mM).

(18) Ensayo de cinasa *in vitro*

Se incubaron 20 μ l de producto IP con 5 μ g de histona H1 (UPSTATE, Lake Placid, NY 12946) en presencia de 30 μ l de tampón de cinasa que contenía ATP 50 μ M y 10 ci de [32 P]-ATP a 30°C durante 30 minutos. Se detuvo la reacción añadiendo 10 μ l de tampón de muestra SDS y se sometió a ebullición durante 3 min.

Ejemplo 2 - B1194(1) Identificación de B1194 como un gen regulado por incremento en el cáncer de mama

Mediante los criterios de selección para el gen candidato a partir de los perfiles de expresión génica de células cancerosas de 77 pacientes con cáncer de mama premenopáusicas usando una micromatriz de ADNc que representa 27.648 seres humanos, se identificaron 468 genes que comúnmente estaban regulados por incremento al menos dos veces en células de cáncer de mama, en comparación con genes de células del conducto de mama normales (véase Materiales y métodos). De entre ellos, se seleccionó B1194, que se designó FLJ-10252, (registro de Genbank NM_018040). La expresión del gen B1194 estaba elevada en 24 de los 41 casos de cáncer de mama informados en la micromatriz. La posterior RT-PCR semicuantitativa confirmó que B1194 estaba significativamente regulado por incremento en 8 de los 12 especímenes clínicos (de tipo bien diferenciado) (figura 1a) y en casi todas las 9 líneas celulares de cáncer de mama (figura 1b) en comparación con células de conducto de mama normales y otros tejidos normales, aunque se observó una débil expresión de B1194 en la glándula mamaria y el corazón. Para examinar adicionalmente los patrones de expresión del B1194, se realizó un análisis de transferencia de tipo Northern con múltiples tejidos humanos y líneas celulares de cáncer de mama usando las 411 pb del fragmento de ADNc como sonda (véase Material y método). Como resultado, B1194 se expresaba exclusivamente en testículos (figura 2a), y se sobreexpresaba específicamente en todas las líneas celulares de cáncer de mama (figura 2b), lo que sugiere que B1194 podría ser un buen candidato a diana para el desarrollo de fármacos anticancerígenos. B1194 consiste en 10 exones, designada proteína hipotética FLJ-10252, ubicada en el cromosoma 1q41. La secuencia de ADNc de longitud completa de B1194 contenía 2338 nucleótidos, y codifica para 528 aminoácidos (58 kDa). El programa informático SMART predijo que este producto génico tiene un parche G altamente conservado, dominio de unión a ácido nucleico rico en glicina en sus extremos carboxilo, lo que sugiere que se predijo que tenía una función de unión a ARN. El marco de lectura abierto (ORF) comienza en el exón 1, y termina en el exón 10.

(2) Localización subcelular de B1194

El programa informático PSORTII predijo que el producto génico de B1194 se localiza principalmente en el núcleo. Para examinar adicionalmente la caracterización de B1194, se investigó la localización subcelular de este producto génico en células de mamíferos. Cuando se transfirió de manera transitoria un plásmido que expresaba una proteína B1194 (pCMV(+)-myc-B1194) en células COS7, la tinción inmunocitoquímica reveló que B1194 exógena se localizaba en todo el núcleo en células COS7 (figura 3a) con un peso molecular de 58 kDa (figura 3b).

(3) Efectos inhibidores del crecimiento de ARNip contra B1194

Para evaluar el papel de promoción del crecimiento del B1194, se desactivó la expresión de B1194 endógena en la línea celular de cáncer de mama T47D, una línea celular que ha mostrado la sobreexpresión de B1194 (véanse la figura 1b y 2b), mediante la técnica de interferencia por ARN (RNAi) basada en vector de mamífero (véase materiales y métodos). Se examinaron los niveles de expresión de B1194 mediante experimentos de RT-PCR semicuantitativa. Tal como se muestra en la figura 4a, entre los dos constructos de ARNip del gen examinado, los ARNip específicos para B1194 (si1 y si5) suprimieron significativamente la expresión de B1194, en comparación con un constructo de ARNip de control (psiH1BX-SC). Para confirmar la inhibición del crecimiento celular con ARNip específicos para B1194, se realizaron ensayos MTT. Como resultado, la introducción de constructos de ARNip específicos para B1194 (si1 y si5) suprimió el crecimiento de células T47D (figura 4b), lo que concuerda con el resultado de la expresión reducida anterior. Se verificó cada resultado mediante tres experimentos independientes. Por tanto, los presentes hallazgos sugieren que B1194 tiene una función significativa en el crecimiento celular del cáncer de mama.

Ejemplo 3 - A2282(1) Identificación de A2282 como un gen regulado por incremento en cáncer de mama

Cuando se analizaron los perfiles de expresión génica de células cancerosas de 77 pacientes con cáncer de mama premenopáusicas usando una micromatriz de ADNc que representaba 27.648 genes humanos, se identificaron 493 genes que comúnmente estaban regulados por incremento en células de cáncer de mama. De entre ellos, se seleccionó A2282, designado leucina cinasa embrionaria materna. MELK (registro de Genbank NM014791) se ubica en el cromosoma 9p13.1 con un transcrito de ARNm de 2501 bases de longitud que consiste en 18 exones. La expresión de A2282 estaba elevada en 25 de los 33 (76%) de los casos de cáncer de mama de los que pudieron obtenerse datos de expresión, especialmente, en 10 de los 14 (71%) casos con especímenes de cáncer de mama de tipo moderadamente diferenciado. De manera intrigante, A2282 se expresa principalmente en pacientes cuyo estado del receptor de estrógeno y progesterona es negativo. Para confirmar el patrón de expresión de este gen en cánceres de mama, se realizó análisis de RT-PCR semicuantitativa usando líneas celulares de cáncer de mama y tejidos humanos

normales que incluyen células de mama normales. Como resultado, se descubrió que A2282, cuya expresión mostraba expresión elevada en 11 de los 12 especímenes de cáncer de mama clínicos (de tipo moderadamente diferenciado) en comparación con células del conducto de mama normales y otros tejidos vitales normales (figura 5a), también se sobreexpresaba en las 6 líneas celulares de cáncer de mama (figura 5b), aunque se observó la expresión en médula ósea. Para examinar adicionalmente el patrón de expresión de este gen, se realizó análisis de transferencia de tipo Northern con múltiples tejidos humanos y líneas celulares de cáncer de mama usando un fragmento de ADNc (554 pb) ubicado dentro de la UTR en 3' de A2282 como sonda (figura 6a). De manera inesperada, se observó que dos transcritos aparentes (de aproximadamente 1,4 kb y 0,5 kb) se expresaban de forma ubicua en todos los tejidos humanos normales. Particularmente, el transcrito de 0,5 kb mostró una expresión superior en casi todos los tejidos normales que el transcrito de 1,4 kb (figura 6b). Por el contrario, un transcrito de aproximadamente 2,4 kb encontrado con análisis de transferencia de tipo Northern se sobreexpresaba específicamente en líneas celulares de cáncer de mama (figura 6c).

(2) Aislamiento de transcrito expresado específico para cáncer de mama de A2282

Para aislar variantes expresadas específicas para cáncer de mama de A2282, se examinó una biblioteca de ADNc construida usando un ARN poli(A)+ obtenido de líneas celulares de cáncer de mama T47D. Se aislaron cinco variantes diferentes (figura 7a). Entre ellas, tres transcritos tenían un tamaño similar de aproximadamente 2,4 kb, designados V1, V2 y V3 según la longitud completa, la delección de 133 de bases y la delección de 250 bases en el transcrito. Los otros dos transcritos V4 y V5, son de 1,23 kb y 0,5 kb, respectivamente. Para investigar qué transcrito se sobreexpresaba en células de cáncer de mama en comparación con mama normal, se realizó análisis de transferencia de tipo Northern usando una secuencia específica para V1 y V2 como sonda (nucleótidos 214-383). Como era de esperar, los transcritos V1 y V2 del gen A2282 se sobreexpresan específicamente en todas las líneas celulares de cáncer de mama (figura 7b). Por tanto, se seleccionaron los transcritos V1, V2 y V3 de A2282. A2282V1, designado MELK, codifica para una proteína de 75 kDa con un dominio catalítico de serina treonina cinasa en el extremo N-terminal y un dominio asociado a cinasa (KA1) cerca al extremo C-terminal. La proteína MELK humana está conservada evolutivamente entre especies que comparten el 65% (Heyer BS, *et al.*, Dev Dyn. 1999; 215:344-51) y el 29,91% (Gilardi-Hebenstreit, P. *et al.*, (1992) Oncogene 7(12),2499-2506) de identidad con la proteína MELK de xenopus y ratón, respectivamente, lo que sugiere que la proteína MELK humana puede tener parejas de unión o funciones similares *in vivo*. Se ha notificado la cinasa MELK de xenopus como el nuevo miembro de la familia KIN1/PAR-1/MARK que está implicada en el establecimiento de la polaridad celular y tanto la organización dinámica de los microtúbulos como del citoesqueleto. Lo más importantemente, se cree que desempeña un papel importante en la regulación del ciclo celular durante la embriogénesis de xenopus (Blot J, *et al.*, (2002) Dev Biol 241, 327-338). También se ha sugerido que la proteína MELK humana participa en la evolución del ciclo celular (Davezac N, *et al.*, (2002). Oncogene, 21, 7630-7641). Sin embargo, aún falta por aclararse las rutas y maquinarias moleculares exactas.

(3) Traducción *in vitro*

Para examinar adicionalmente si estas cinco variantes diferentes podrían traducirse potencialmente en proteínas funcionales, se realizaron experimentos de traducción *in vitro*. Se confirmó que los transcritos V1, V2 y V3 pueden traducirse *in vitro* con el peso molecular pronosticado de la proteína de 75, 71 y 66 kDa respectivamente (figura 8). Sin embargo, no se detectó ninguna banda para V4 en líneas celulares de cáncer de mama. Estos hallazgos sugieren que los transcritos V1, V2 y V3 de A2282 son buenos candidatos como dianas moleculares para el desarrollo de fármacos anticancerígenos novedosos.

(4) Expresión de V1, V2 y V3 de la proteína A2282 en cualquier fase del ciclo celular

Dado que se notificó que MELK humano estaba implicado en la regulación del ciclo celular, y se observó que su actividad cinasa y estado de fosforilación alcanzaban un nivel máximo durante la fase mitótica (Davezac N, *et al.*, (2002). Oncogene, 21, 7630-7641), se realizaron análisis de inmunotransferencia de tipo Western y de citometría de flujo para examinar si V2 y V3 también poseen características de V1. Por consiguiente, se observó una banda que migraba más lentamente extra de la proteína V1 durante la fase mitótica en células HeLa sincronizadas (figura 9a, b); sin embargo, no se observó ninguna banda extra en las proteínas V2 y V3, lo que sugiere que puede no haber ninguna fosforilación de las proteínas V2 y V3 en ninguna fase del ciclo celular.

(5) Efectos inhibidores del crecimiento de ARNip contra A2282

Para evaluar el papel de promoción del crecimiento de A2282, se desactivó la expresión de A2282 endógeno en las líneas celulares de cáncer de mama T47D y MCF-7, cada una de las cuales se ha demostrado que sobreexpresa A2282, mediante la técnica de interferencia por ARN (RNAi) basada en vector de mamíferos (véase Materiales y métodos) (figura 10). Se examinaron los niveles de expresión de A2282 mediante experimentos de RT-PCR semicuantitativa. Tal como se muestra en la figura 10a, ARNip específicos para A2282 (si3 y si4) suprimieron significativamente la expresión, en comparación con constructos de ARNip de control (psiH1BX-LUC o -SC). Para confirmar la inhibición del crecimiento celular con ARNip específicos para A2282, se realizaron ensayos de MTT y de formación de colonias, respectivamente (figura 10b, c). Como resultado, la introducción de constructos de ARNip específicos para A2282 suprimió el crecimiento de células T47D y MCF-7, lo que concuerda con el resultado de la expresión reducida anterior de este gen. Se verificó cada resultado mediante tres experimentos independientes. Por

tanto, los presentes hallazgos sugieren que A2282 tiene una función significativa en el crecimiento celular del cáncer de mama.

(6) Proteína A2282 fosforilada en el dominio cinasa

5 Para identificar los posibles sitios de fosforilación, se examinaron proteínas de tipo natural (WT) y dos proteínas
 10 delecionadas en el dominio cinasa de A2282 para determinar la fosforilación (figura 11a). Tal como se muestra en la
 15 figura 11b, la proteína WT migró como una banda difusa en todas las fases del ciclo celular, particularmente en la fase
 20 de mitosis; sin embargo, no hubo ninguna banda difusa en los constructos truncados (TC1 y TC2). Para investigar si una
 25 banda difusa de proteína WT está fosforilada, se realizó un ensayo con fosfatasa lambda (figura 11c). El tratamiento con
 fosfatasa lambda redujo la banda difusa a una única banda, lo que sugiere que la proteína WT estaba exhaustivamente
 fosforilada en la fase M. En cambio, no se observó ni una banda doble ni difusa en las otras dos proteínas cinasa
 truncadas. Estos datos indicaron que era probable que el(los) sitio(s) de fosforilación estuviese(n) ubicado(s) en la
 región cinasa de la proteína A2282.

(7) Regulación de la actividad cinasa de A2282 mediante fosforilación

15 Se ha notificado que la activación de cinasas relacionadas con MARK y AMPK se regula mediante la
 fosforilación de su residuo de treonina del bucle T (Drewes G y Nurse P, FEBS Lett. 2003; 554:45-9; Spicer J, *et al.*,
 Oncogene. 2003; 22:4752-6). Por consiguiente, se construyeron varios mutantes sustituidos y delecionados (véase
 20 Materiales y métodos). Se realizó inmunoprecipitación en células de mamíferos y se examinaron las actividades cinasa
 de estos constructos usando histona H1, el sustrato usado en informes anteriores. Se sustituyeron el paquete de unión a
 25 ATP (residuos D150) pronosticado y la posible treonina fosforilada (residuos T167) por residuo de alanina tal como se
 describió en Materiales y métodos. También se mutó un sitio de fosforilación notificado (Vulsteke V, *et al.*, J Biol Chem.
 2004; 279(10):8642-7) Thr 478 como control (figura 12a). También se descubrió que todos los transcritos estaban
 fosforilados en la línea celular HEK293 tal como se observa en la inmunotransferencia tipo Western (figura 12b). Con
 respecto a la actividad cinasa, se observó que las proteínas de tipo natural y T478A que poseen un dominio cinasa
 intacto fosforilaban histona H1 *in vitro* (figura 12c). Esta actividad estaba gravemente comprometida en el mutante
 T167A y completamente suprimida en la proteína D150A. Además, se observó una banda ubicada a los 75 kDa en WT,
 T167A y T478A pero no en D150A, lo que indica la posibilidad de autofosforilación de la proteína A2282.

Aplicabilidad industrial

30 La expresión de los genes humanos novedosos *B1194*, *A2282V1*, *A2282V2* y *A2282V3* está notablemente
 elevada en cáncer de mama en comparación con tejidos humanos no cancerosos. Por consiguiente, estos genes
 pueden servir como marcadores de diagnóstico de cáncer y las proteínas codificadas por los mismos pueden usarse en
 ensayos de diagnóstico del cáncer.

35 En el presente documento, se mostró que la expresión de las proteínas novedosas *B1194*, *A2282V1*, *A2282V2*
 y *A2282V3* promovía el crecimiento celular mientras que se suprimía el crecimiento celular mediante oligonucleótidos
 antisentido o ARN de interferencia pequeños correspondientes a los genes *B1194*, *A2282V1*, *A2282V2* y *A2282V3*.
 40 Estos hallazgos sugieren que cada una de las proteínas *B1194*, *A2282V1*, *A2282V2* y *A2282V3* estimula la actividad
 oncogénica. Por tanto, cada una de estas oncoproteínas novedosas es una diana útil para el desarrollo de agentes
 farmacéuticos anticancerígenos. Por ejemplo, agentes que bloquean la expresión de *B1194*, *A2282V1*, *A2282V2*, o
A2282V3 o evitan su actividad pueden encontrar utilidad terapéutica como agentes anticancerígenos, en particular
 agentes anticancerígenos para el tratamiento del cáncer de mama. Ejemplos de tales agentes incluyen oligonucleótidos
 antisentido, ARN de interferencia pequeños y anticuerpos que reconocen *B1194*, *A2282V1*, *A2282V2* o *A2282V3*.

LISTA DE SECUENCIAS

<110> ONCOTHERAPY SCIENCE, INC.

THE UNIVERSITY OF TOKYO

<120> GENES Y POLIPÉPTIDOS RELACIONADOS CON CÁNCERES DE MAMA

<130> ONC-A0409P

<150> US 60/600 146

<151> 10-08-2004

<160> 44

<170> PatentIn versión 3.3

<210> 1

<211> 2338

<212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

<222> (99)..(1685)

<400> 1

aaaaaatgct gaactgctct ttggaagtcg ccggtgctgt tglagttaga gtcgtttcac	60
ggcctgagc ttcgaggcca ggctcccggg tgtcgta atg ttc ggg gcc gcc ggg	116
Met Phe Gly Ala Ala Gly	
1 5	
cgc caa ccg atc gga gct cca gca gcc ggg aac agc tgg cat ttc agt	164
Arg Gln Pro Ile Gly Ala Pro Ala Ala Gly Asn Ser Trp His Phe Ser	
10 15 20	
aga acc atg gag gag ctg gtt cat gac ctt gtc tca gca ttg gaa gag	212
Arg Thr Met Glu Glu Leu Val His Asp Leu Val Ser Ala Leu Glu Glu	
25 30 35	
agc tca gag caa gct cga ggt gga ttt gct gaa aca gga gac cat tct	260
Ser Ser Glu Gln Ala Arg Gly Gly Phe Ala Glu Thr Gly Asp His Ser	
40 45 50	
cga agt ata tct tgc cct ctg aaa cgc cag gca agg aaa agg aga ggg	308
Arg Ser Ile Ser Cys Pro Leu Lys Arg Gln Ala Arg Lys Arg Arg Gly	
55 60 65 70	
aga aaa cgg agg tcg tat aat gtg cat cac ccg tgg gag act ggt cac	356
Arg Lys Arg Arg Ser Tyr Asn Val His His Pro Trp Glu Thr Gly His	
75 80 85	

ES 2 357 658 T3

tgc tta agt gaa ggc tct gat tct agt tta gaa gaa cca agc aag gac	404
Cys Leu Ser Glu Gly Ser Asp Ser Ser Leu Glu Glu Pro Ser Lys Asp	
90 95 100	
tat aga gag aat cac aat aat aat aaa aaa gat cac agt gac tct gat	452
Tyr Arg Glu Asn His Asn Asn Asn Lys Lys Asp His Ser Asp Ser Asp	
105 110 115	
gac caa atg tta gta gca aag cgc agg ccg tca tca aac tta aat aat	500
Asp Gln Met Leu Val Ala Lys Arg Arg Pro Ser Ser Asn Leu Asn Asn	
120 125 130	
aat gtt cga ggg aaa aga cct cta tgg cat gag tct gat ttt gct gtg	548
Asn Val Arg Gly Lys Arg Pro Leu Trp His Glu Ser Asp Phe Ala Val	
135 140 145 150	
gac aat gtt ggg aat aga act ctg cgc agg agg aga aag gta aaa cgc	596
Asp Asn Val Gly Asn Arg Thr Leu Arg Arg Arg Arg Lys Val Lys Arg	
155 160 165	
atg gca gta gat ctc cca cag gac atc tct aac aaa cgg aca atg acc	644
Met Ala Val Asp Leu Pro Gln Asp Ile Ser Asn Lys Arg Thr Met Thr	
170 175 180	
cag cca cct gag ggt tgt aga gat cag gac atg gac agt gat aga gcc	692

ES 2 357 658 T3

Gln Pro Pro Glu Gly Cys Arg Asp Gln Asp Met Asp Ser Asp Arg Ala	
185	190
195	
tac cag tat caa gaa ttt acc aag aac aaa gtc aaa aaa aga aag ttg	740
Tyr Gln Tyr Gln Glu Phe Thr Lys Asn Lys Val Lys Lys Arg Lys Leu	
200	205
210	
aaa ata atc aga caa gga cca aaa atc caa gat gaa gga gta gtt tta	788
Lys Ile Ile Arg Gln Gly Pro Lys Ile Gln Asp Glu Gly Val Val Leu	
215	220
225	230
gaa agt gag gaa acg aac cag acc aat aag gac aaa atg gaa tgt gaa	836
Glu Ser Glu Glu Thr Asn Gln Thr Asn Lys Asp Lys Met Glu Cys Glu	
235	240
245	
gag caa aaa gtc tca gat gag ctc atg agt gaa agt gat tcc agc agt	884
Glu Gln Lys Val Ser Asp Glu Leu Met Ser Glu Ser Asp Ser Ser Ser	
250	255
260	
ctc agc agc act gat gct gga ttg ttt acc aat gat gag gga aga caa	932
Leu Ser Ser Thr Asp Ala Gly Leu Phe Thr Asn Asp Glu Gly Arg Gln	
265	270
275	
ggt gat gat gaa cag agt gac tgg ttc tac gaa aag gaa tca ggt gga	980
Gly Asp Asp Glu Gln Ser Asp Trp Phe Tyr Glu Lys Glu Ser Gly Gly	
280	285
290	

ES 2 357 658 T3

cca aag gga tta gga cit gga ttt cct cta cca aaa agt act tcc gca	1652
Pro Lys Gly Leu Gly Leu Gly Phe Pro Leu Pro Lys Ser Thr Ser Ala	
505 510 515	
act act acc ccc aat gca gga aaa tcc gcc taa gaaaagcaaa gaagaaatgt	1705
Thr Thr Thr Pro Asn Ala Gly Lys Ser Ala	
520 525	
tttacagact ttattcacta tgtccattg ttctaaaatg ataacatgac ttctgttttt	1765
gaagcaaaaa tctacattgc ctcaaacaca tcaactctagc ttccttactg catacagtcc	1825
tgccatagtg agagaaatgg gatttcatca caattcatgg tgctaaaatg aaaacctctg	1885
cactttaatt ttttcagta atttccagct atttctaggt ataaagagca gctcgtttct	1945
cttattttatt ttagtctcat gtgtcaaac tttccgatgc ttigtctaat tcatgtatgt	2005
gtgcagtgct gcaatgccca gacaaacgtg agcacaccca ccagtttcta aaatggaata	2065
gacaggaaaa gatttggttt tataatcatcc ctatctattg taacccaaaa gacctacat	2125
cgcacatcagtg aagtcogaac acatctttgt ttgaaaggct tgtcaatttc atattccttg	2185
aattggcttc ttggtagga tttctgaca gagtgatacc catcaatfff ctatccttag	2245
acaatgtagt gtaagtcca cagttgacaa acaacaatta atgtttccct tggatgtttt	2305
gacaaaaata aaccicatcg ttgttatcac cag	2338

<210> 2

<211> 528

<212> PRT

<213> *Homo sapiens*

<400> 2

Met Phe Gly Ala Ala Gly Arg Gln Pro Ile Gly Ala Pro Ala Ala Gly
1 5 10 15

Asn Ser Trp His Phe Ser Arg Thr Met Glu Glu Leu Val His Asp Leu
 20 25 30

Val Ser Ala Leu Glu Glu Ser Ser Glu Gln Ala Arg Gly Gly Phe Ala
 35 40 45

Glu Thr Gly Asp His Ser Arg Ser Ile Ser Cys Pro Leu Lys Arg Gln

Arg Arg Lys Val Lys Arg Met Ala Val Asp Leu Pro Gln Asp Ile Ser

165

170

175

Asn Lys Arg Thr Met Thr Gln Pro Pro Glu Gly Cys Arg Asp Gln Asp

180

185

190

Met Asp Ser Asp Arg Ala Tyr Gln Tyr Gln Glu Phe Thr Lys Asn Lys

195

200

205

Val Lys Lys Arg Lys Leu Lys Ile Ile Arg Gln Gly Pro Lys Ile Gln

210

215

220

Asp Glu Gly Val Val Leu Glu Ser Glu Glu Thr Asn Gln Thr Asn Lys

225

230

235

240

Asp Lys Met Glu Cys Glu Glu Gln Lys Val Ser Asp Glu Leu Met Ser

245

250

255

Glu Ser Asp Ser Ser Ser Leu Ser Ser Thr Asp Ala Gly Leu Phe Thr

Ile Pro Gly Pro Val Gly Asn Lys Arg Met Val His Phe Ser Pro Asp
 370 375 380

Ser His His His Asp His Trp Phe Ser Pro Gly Ala Arg Thr Glu His
 385 390 395 400

Asp Gln His Gln Leu Leu Arg Asp Asn Arg Ala Glu Arg Gly His Lys
 405 410 415

Lys Asn Cys Ser Val Arg Thr Ala Ser Arg Gln Thr Ser Met His Leu
 420 425 430

Gly Ser Leu Cys Thr Gly Asp Ile Lys Arg Arg Arg Lys Ala Ala Pro
 435 440 445

Leu Pro Gly Pro Thr Thr Ala Gly Phe Val Gly Glu Asn Ala Gln Pro
 450 455 460

Ile Leu Glu Asn Asn Ile Gly Asn Arg Met Leu Gln Asn Met Gly Trp

465

470

475

480

Thr Pro Gly Ser Gly Leu Gly Arg Asp Gly Lys Gly Ile Ser Glu Pro

485

490

495

Ile Gln Ala Met Gln Arg Pro Lys Gly Leu Gly Leu Gly Phe Pro Leu

500

505

510

Pro Lys Ser Thr Ser Ala Thr Thr Thr Pro Asn Ala Gly Lys Ser Ala

515

520

525

<210> 3

<211> 2501

<212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

<222> (139)..(2094)

<400> 3

ES 2 357 658 T3

cgaaaagatt cttaggaacg ccgtaccagc cgcgtctctc aggacagcag gccccctgtcc 60

ttctgtcggg cgccgctcag ccgtgccctc cgcccctcag gttctttttc taattccaaa 120

taaacttgca agaggact atg aaa gat tat gat gaa ctt ctc aaa tat tat 171

Met Lys Asp Tyr Asp Glu Leu Leu Lys Tyr Tyr

1 5 10

gaa tta cat gaa act att ggg aca ggt ggc ttt gca aag gtc aaa ctt 219

Glu Leu His Glu Thr Ile Gly Thr Gly Gly Phe Ala Lys Val Lys Leu

15 20 25

gcc tgc cat atc ctt act gga gag atg gta gct ata aaa atc atg gat 267

Ala Cys His Ile Leu Thr Gly Glu Met Val Ala Ile Lys Ile Met Asp

30 35 40

aaa aac aca cta ggg agt gat ttg ccc cgg atc aaa acg gag att gag 315

Lys Asn Thr Leu Gly Ser Asp Leu Pro Arg Ile Lys Thr Glu Ile Glu

45 50 55

gcc ttg aag aac ctg aga cat cag cat ata tgt caa ctc tac cat gtg 363

Ala Leu Lys Asn Leu Arg His Gln His Ile Cys Gln Leu Tyr His Val

60 65 70 75

cta gag aca gcc aac aaa ata ttc atg gtt ctt gag tac tgc cct gga 411

Leu Glu Thr Ala Asn Lys Ile Phe Met Val Leu Glu Tyr Cys Pro Gly

ES 2 357 658 T3

80	85	90	
gga gag ctg ttt gac tat ata att tcc cag gat cgc ctg tca gaa gag			459
Gly Glu Leu Phe Asp Tyr Ile Ile Ser Gln Asp Arg Leu Ser Glu Glu			
95	100	105	
gag acc cgg gtt gtc ttc cgt cag ata gta tct gct gtt gct tat gtg			507
Glu Thr Arg Val Val Phe Arg Gln Ile Val Ser Ala Val Ala Tyr Val			
110	115	120	
cac agc cag ggc tat gct cac agg gac ctc aag cca gaa aat ttg ctg			555
His Ser Gln Gly Tyr Ala His Arg Asp Leu Lys Pro Glu Asn Leu Leu			
125	130	135	
ttt gat gaa tat cat aaa tta aag ctg att gac ttt ggt ctc tgt gca			603
Phe Asp Glu Tyr His Lys Leu Lys Leu Ile Asp Phe Gly Leu Cys Ala			
140	145	150	155
aaa ccc aag ggt aac aag gat tac cat cta cag aca tgc tgt ggg agt			651
Lys Pro Lys Gly Asn Lys Asp Tyr His Leu Gln Thr Cys Cys Gly Ser			
160	165	170	
ctg gct tat gca gca cct gag tta ata caa ggc aaa tca tat ctt gga			699
Leu Ala Tyr Ala Ala Pro Glu Leu Ile Gln Gly Lys Ser Tyr Leu Gly			
175	180	185	

tca gag gca gat gtt tgg agc atg ggc ata ctg ita tat gtt ctt atg 747
 Ser Glu Ala Asp Val Trp Ser Met Gly Ile Leu Leu Tyr Val Leu Met
 190 195 200

tgt gga ttt cta cca ttt gat gat gat aat gta atg gct tta tac aag 795
 Cys Gly Phe Leu Pro Phe Asp Asp Asp Asn Val Met Ala Leu Tyr Lys
 205 210 215

aag att atg aga gga aaa tat gat gtt ccc aag tgg ctc tct ccc agt 843
 Lys Ile Met Arg Gly Lys Tyr Asp Val Pro Lys Trp Leu Ser Pro Ser
 220 225 230 235

agc att ctg ctt ctt caa caa atg ctg cag gtg gac cca aag aaa cgg 891
 Ser Ile Leu Leu Leu Gln Gln Met Leu Gln Val Asp Pro Lys Lys Arg
 240 245 250

att tct atg aaa aat cta ttg aac cat ccc tgg atc atg caa gat tac 939
 Ile Ser Met Lys Asn Leu Leu Asn His Pro Trp Ile Met Gln Asp Tyr
 255 260 265

aac tat cct gtt gag tgg caa agc aag aat cct ttt att cac ctc gat 987
 Asn Tyr Pro Val Glu Trp Gln Ser Lys Asn Pro Phe Ile His Leu Asp
 270 275 280

gat gat tgc gta aca gaa ctt tct gta cat cac aga aac aac agg caa 1035
 Asp Asp Cys Val Thr Glu Leu Ser Val His His Arg Asn Asn Arg Gln

ES 2 357 658 T3

285	290	295	
aca atg gag gat tta att tca ctg tgg cag tat gat cac ctc acg gct			1083
Thr Met Glu Asp Leu Ile Ser Leu Trp Gln Tyr Asp His Leu Thr Ala			
300	305	310	315
acc tat ctt ctg ctt cta gcc aag aag gct cgg gga aaa cca gtt cgt			1131
Thr Tyr Leu Leu Leu Leu Ala Lys Lys Ala Arg Gly Lys Pro Val Arg			
	320	325	330
tta agg ctt tct tct ttc tcc tgt gga caa gcc agt gct acc cca ttc			1179
Leu Arg Leu Ser Ser Phe Ser Cys Gly Gln Ala Ser Ala Thr Pro Phe			
	335	340	345
aca gac atc aag tca aat aat tgg agt ctg gaa gat gtg acc gca agt			1227
Thr Asp Ile Lys Ser Asn Asn Trp Ser Leu Glu Asp Val Thr Ala Ser			
350	355	360	
gat aaa aat tat gtg gcg gga tta ata gac tat gat tgg tgt gaa gat			1275
Asp Lys Asn Tyr Val Ala Gly Leu Ile Asp Tyr Asp Trp Cys Glu Asp			
365	370	375	
gat tta tca aca ggt gct gct act ccc cga aca tca cag ttt acc aag			1323
Asp Leu Ser Thr Gly Ala Ala Thr Pro Arg Thr Ser Gln Phe Thr Lys			
380	385	390	395

tac tgg aca gaa tca aat ggg gtg gaa tct aaa tca tta act cca gcc 1371
 Tyr Trp Thr Glu Ser Asn Gly Val Glu Ser Lys Ser Leu Thr Pro Ala
 400 405 410

tta tgc aga aca cct gca aat aaa tta aag aac aaa gaa aat gta tat 1419
 Leu Cys Arg Thr Pro Ala Asn Lys Leu Lys Asn Lys Glu Asn Val Tyr
 415 420 425

act cct aag tct gct gta aag aat gaa gag tac ttt atg ttt cct gag 1467
 Thr Pro Lys Ser Ala Val Lys Asn Glu Glu Tyr Phe Met Phe Pro Glu
 430 435 440

cca aag act cca gtt aat aag aac cag cat aag aga gaa ata ctc act 1515
 Pro Lys Thr Pro Val Asn Lys Asn Gln His Lys Arg Glu Ile Leu Thr
 445 450 455

acg cca aat cgt tac act aca ccc tca aaa gct aga aac cag tgc ctg 1563
 Thr Pro Asn Arg Tyr Thr Thr Pro Ser Lys Ala Arg Asn Gln Cys Leu
 460 465 470 475

aaa gaa act cca att aaa ata cca gta aat tca aca gga aca gac aag 1611
 Lys Glu Thr Pro Ile Lys Ile Pro Val Asn Ser Thr Gly Thr Asp Lys
 480 485 490

tta atg aca ggt gtc att agc cct gag agg cgg tgc cgc tca gtg gaa 1659
 Leu Met Thr Gly Val Ile Ser Pro Glu Arg Arg Cys Arg Ser Val Glu

ES 2 357 658 T3

495	500	505	
ttg gat ctc aac caa gca cat atg gag gag act cca aaa aga aag gga			1707
Leu Asp Leu Asn Gln Ala His Met Glu Glu Thr Pro Lys Arg Lys Gly			
510	515	520	
gcc aaa gtg ttt ggg agc ctt gaa agg ggg ttg gat aag gtt atc act			1755
Ala Lys Val Phe Gly Ser Leu Glu Arg Gly Leu Asp Lys Val Ile Thr			
525	530	535	
gtg ctc acc agg agc aaa agg aag ggt tct gcc aga gac ggg ccc aga			1803
Val Leu Thr Arg Ser Lys Arg Lys Gly Ser Ala Arg Asp Gly Pro Arg			
540	545	550	555
aga cta aag ctt cac tat aat gtg act aca act aga tta gtg aat cca			1851
Arg Leu Lys Leu His Tyr Asn Val Thr Thr Thr Arg Leu Val Asn Pro			
560	565	570	
gat caa ctg ttg aat gaa ata atg tct att ctt cca aag aag cat gtt			1899
Asp Gln Leu Leu Asn Glu Ile Met Ser Ile Leu Pro Lys Lys His Val			
575	580	585	
gac ttt gta caa aag ggt tat aca ctg aag tgt caa aca cag tca gat			1947
Asp Phe Val Gln Lys Gly Tyr Thr Leu Lys Cys Gln Thr Gln Ser Asp			
590	595	600	

<400> 4

Met Lys Asp Tyr Asp Glu Leu Leu Lys Tyr Tyr Glu Leu His Glu Thr
 1 5 10 15

Ile Gly Thr Gly Gly Phe Ala Lys Val Lys Leu Ala Cys His Ile Leu
 20 25 30

Thr Gly Glu Met Val Ala Ile Lys Ile Met Asp Lys Asn Thr Leu Gly
 35 40 45

Ser Asp Leu Pro Arg Ile Lys Thr Glu Ile Glu Ala Leu Lys Asn Leu
 50 55 60

Arg His Gln His Ile Cys Gln Leu Tyr His Val Leu Glu Thr Ala Asn
 65 70 75 80

Lys Ile Phe Met Val Leu Glu Tyr Cys Pro Gly Gly Glu Leu Phe Asp
 85 90 95

Tyr Ile Ile Ser Gln Asp Arg Leu Ser Glu Glu Glu Thr Arg Val Val
 100 105 110

Phe Arg Gln Ile Val Ser Ala Val Ala Tyr Val His Ser Gln Gly Tyr
 115 120 125

Ala His Arg Asp Leu Lys Pro Glu Asn Leu Leu Phe Asp Glu Tyr His
 130 135 140

Lys Leu Lys Leu Ile Asp Phe Gly Leu Cys Ala Lys Pro Lys Gly Asn
 145 150 155 160

Lys Asp Tyr His Leu Gln Thr Cys Cys Gly Ser Leu Ala Tyr Ala Ala
 165 170 175

Pro Glu Leu Ile Gln Gly Lys Ser Tyr Leu Gly Ser Glu Ala Asp Val
 180 185 190

Trp Ser Met Gly Ile Leu Leu Tyr Val Leu Met Cys Gly Phe Leu Pro
 195 200 205

Phe Asp Asp Asp Asn Val Met Ala Leu Tyr Lys Lys Ile Met Arg Gly
 210 215 220

Lys Tyr Asp Val Pro Lys Trp Leu Ser Pro Ser Ser Ile Leu Leu Leu
 225 230 235 240

Gln Gln Met Leu Gln Val Asp Pro Lys Lys Arg Ile Ser Met Lys Asn
 245 250 255

Leu Leu Asn His Pro Trp Ile Met Gln Asp Tyr Asn Tyr Pro Val Glu
 260 265 270

Trp Gln Ser Lys Asn Pro Phe Ile His Leu Asp Asp Asp Cys Val Thr
 275 280 285

Glu Leu Ser Val His His Arg Asn Asn Arg Gln Thr Met Glu Asp Leu
 290 295 300

Ile Ser Leu Trp Gln Tyr Asp His Leu Thr Ala Thr Tyr Leu Leu Leu
 305 310 315 320

Leu Ala Lys Lys Ala Arg Gly Lys Pro Val Arg Leu Arg Leu Ser Ser
 325 330 335

Phe Ser Cys Gly Gln Ala Ser Ala Thr Pro Phe Thr Asp Ile Lys Ser
 340 345 350

Asn Asn Trp Ser Leu Glu Asp Val Thr Ala Ser Asp Lys Asn Tyr Val
 355 360 365

Ala Gly Leu Ile Asp Tyr Asp Trp Cys Glu Asp Asp Leu Ser Thr Gly
 370 375 380

Ile Ser Pro Glu Arg Arg Cys Arg Ser Val Glu Leu Asp Leu Asn Gln
 500 505 510

Ala His Met Glu Glu Thr Pro Lys Arg Lys Gly Ala Lys Val Phe Gly
 515 520 525

Ser Leu Glu Arg Gly Leu Asp Lys Val Ile Thr Val Leu Thr Arg Ser
 530 535 540

Lys Arg Lys Gly Ser Ala Arg Asp Gly Pro Arg Arg Leu Lys Leu His
 545 550 555 560

Tyr Asn Val Thr Thr Thr Arg Leu Val Asn Pro Asp Gln Leu Leu Asn
 565 570 575

Glu Ile Met Ser Ile Leu Pro Lys Lys His Val Asp Phe Val Gln Lys
 580 585 590

Gly Tyr Thr Leu Lys Cys Gln Thr Gln Ser Asp Phe Gly Lys Val Thr
 595 600 605

Met Gln Phe Glu Leu Glu Val Cys Gln Leu Gln Lys Pro Asp Val Val
 610 615 620

Gly Ile Arg Arg Gln Arg Leu Lys Gly Asp Ala Trp Val Tyr Lys Arg
 625 630 635 640

Leu Val Glu Asp Ile Leu Ser Ser Cys Lys Val
 645 650

<210> 5

<211> 2368

<212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

<222> (146)..(2002)

<400> 5

aaagattctt aggaacgccg taccagccgc gtctctcagg acagcaggcc cctgtccttc	60
tgtcggggcg cgctcagccg tgccctccgc cctcagggt ctttttctaa ttccaaataa	120
acttgcaaga ggactatgaa agatt atg atg aac ttc tca aat att atg aat	172
Met Met Asn Phe Ser Asn Ile Met Asn	
1 5	
tac atg aaa cta ttg gga cag agt gat ttg ccc cgg atc aaa acg gag	220
Tyr Met Lys Leu Leu Gly Gln Ser Asp Leu Pro Arg Ile Lys Thr Glu	
10 15 20 25	
att gag gcc ttg aag aac ctg aga cat cag cat ata tgt caa ctc tac	268
Ile Glu Ala Leu Lys Asn Leu Arg His Gln His Ile Cys Gln Leu Tyr	
30 35 40	
cat gtg cta gag aca gcc aac aaa ata ttc atg gtt ctt gag tac tgc	316
His Val Leu Glu Thr Ala Asn Lys Ile Phe Met Val Leu Glu Tyr Cys	
45 50 55	
cct gga gga gag ctg ttt gac tat ata att tcc cag gat cgc ctg tca	364
Pro Gly Gly Glu Leu Phe Asp Tyr Ile Ile Ser Gln Asp Arg Leu Ser	
60 65 70	
gaa gag gag acc cgg gtt gtc ttc cgt cag ata gta tct gct gtt gct	412

ES 2 357 658 T3

Glu Glu Glu Thr Arg Val Val Phe Arg Gln Ile Val Ser Ala Val Ala	
75	80
85	
tat gtg cac agc cag ggc tat gct cac agg gac ctc aag cca gaa aat	460
Tyr Val His Ser Gln Gly Tyr Ala His Arg Asp Leu Lys Pro Glu Asn	
90	95
100	105
ttg ctg ttt gat gaa tat cat aaa tta aag ctg att gac ttt ggt ctc	508
Leu Leu Phe Asp Glu Tyr His Lys Leu Lys Leu Ile Asp Phe Gly Leu	
110	115
120	
tgt gca aaa ccc aag ggt aac aag gat tac cat cta cag aca tgc tgt	556
Cys Ala Lys Pro Lys Gly Asn Lys Asp Tyr His Leu Gln Thr Cys Cys	
125	130
135	
ggg agt ctg gct tat gca gca cct gag tta ata caa ggc aaa tca tat	604
Gly Ser Leu Ala Tyr Ala Ala Pro Glu Leu Ile Gln Gly Lys Ser Tyr	
140	145
150	
cit gga tca gag gca gat gtt tgg agc atg ggc ata ctg tta tat gtt	652
Leu Gly Ser Glu Ala Asp Val Trp Ser Met Gly Ile Leu Leu Tyr Val	
155	160
165	
cit atg tgt gga ttt cta cca ttt gat gat gat aat gta atg gct tta	700
Leu Met Cys Gly Phe Leu Pro Phe Asp Asp Asp Asn Val Met Ala Leu	
170	175
180	185

ES 2 357 658 T3

tac aag aag att atg aga gga aaa tat gat gtt ccc aag tgg ctc tct	748
Tyr Lys Lys Ile Met Arg Gly Lys Tyr Asp Val Pro Lys Trp Leu Ser	
190 195 200	
ccc agt agc att ctg ctt ctt caa caa atg ctg cag gtg gac cca aag	796
Pro Ser Ser Ile Leu Leu Leu Gln Gln Met Leu Gln Val Asp Pro Lys	
205 210 215	
aaa cgg att tct atg aaa aat cta ttg aac cat ccc tgg atc atg caa	844
Lys Arg Ile Ser Met Lys Asn Leu Leu Asn His Pro Trp Ile Met Gln	
220 225 230	
gat tac aac tat cct gtt gag tgg caa agc aag aat cct ttt att cac	892
Asp Tyr Asn Tyr Pro Val Glu Trp Gln Ser Lys Asn Pro Phe Ile His	
235 240 245	
ctc gat gat gat tgc gta aca gaa ctt tct gta cat cac aga aac aac	940
Leu Asp Asp Asp Cys Val Thr Glu Leu Ser Val His His Arg Asn Asn	
250 255 260 265	
agg caa aca atg gag gat tta att tca ctg tgg cag tat gat cac ctc	988
Arg Gln Thr Met Glu Asp Leu Ile Ser Leu Trp Gln Tyr Asp His Leu	
270 275 280	
acg gct acc tat ctt ctg ctt cta gcc aag aag gct cgg gga aaa cca	1036

Thr Ala Thr Tyr Leu Leu Leu Leu Ala Lys Lys Ala Arg Gly Lys Pro	
285	290
295	
gtt cgt tta agg ctt tct tct ttc tcc tgt gga caa gcc agt gct acc	1084
Val Arg Leu Arg Leu Ser Ser Phe Ser Cys Gly Gln Ala Ser Ala Thr	
300	305
310	
cca ttc aca gac atc aag tca aat aat tgg agt ctg gaa gat gtg acc	1132
Pro Phe Thr Asp Ile Lys Ser Asn Asn Trp Ser Leu Glu Asp Val Thr	
315	320
325	
gca agt gat aaa aat tat gtg gcg gga tta ata gac tat gat tgg tgt	1180
Ala Ser Asp Lys Asn Tyr Val Ala Gly Leu Ile Asp Tyr Asp Trp Cys	
330	335
340	345
gaa gat gat tta tca aca ggt gct gct act ccc cga aca tca cag ttt	1228
Glu Asp Asp Leu Ser Thr Gly Ala Ala Thr Pro Arg Thr Ser Gln Phe	
350	355
360	
acc aag tac tgg aca gaa tca aat ggg gtg gaa tct aaa tca tta act	1276
Thr Lys Tyr Trp Thr Glu Ser Asn Gly Val Glu Ser Lys Ser Leu Thr	
365	370
375	
cca gcc tta tgc aga aca cct gca aat aaa tta aag aac aaa gaa aat	1324
Pro Ala Leu Cys Arg Thr Pro Ala Asn Lys Leu Lys Asn Lys Glu Asn	
380	385
390	

ES 2 357 658 T3

gta tat act cct aag tct gct gta aag aat gaa gag tac ttt atg ttt	1372
Val Tyr Thr Pro Lys Ser Ala Val Lys Asn Glu Glu Tyr Phe Met Phe	
395 400 405	
cct gag cca aag act cca gtt aat aag aac cag cat aag aga gaa ata	1420
Pro Glu Pro Lys Thr Pro Val Asn Lys Asn Gln His Lys Arg Glu Ile	
410 415 420 425	
ctc act acg cca aat cgt tac act aca ccc tca aaa gct aga aac cag	1468
Leu Thr Thr Pro Asn Arg Tyr Thr Thr Pro Ser Lys Ala Arg Asn Gln	
430 435 440	
tgc ctg aaa gaa act cca att aaa ata cca gta aat tca aca gga aca	1516
Cys Leu Lys Glu Thr Pro Ile Lys Ile Pro Val Asn Ser Thr Gly Thr	
445 450 455	
gac aag tta atg aca ggt gtc att agc cct gag agg cgg tgc cgc tca	1564
Asp Lys Leu Met Thr Gly Val Ile Ser Pro Glu Arg Arg Cys Arg Ser	
460 465 470	
gtg gaa ttg gat ctc aac caa gca cat atg gag gag act cca aaa aga	1612
Val Glu Leu Asp Leu Asn Gln Ala His Met Glu Glu Thr Pro Lys Arg	
475 480 485	
aag gga gcc aaa gtg ttt ggg agc ctt gaa agg ggg ttg gat aag gtt	1660

Lys Gly Ala Lys Val Phe Gly Ser Leu Glu Arg Gly Leu Asp Lys Val	
490	505
atc act gtg ctc acc agg agc aaa agg aag ggt tct gcc aga gac ggg	1708
Ile Thr Val Leu Thr Arg Ser Lys Arg Lys Gly Ser Ala Arg Asp Gly	
510	520
ccc aga aga cta aag ctt cac taj aat gtg act aca act aga tta gtg	1756
Pro Arg Arg Leu Lys Leu His Tyr Asn Val Thr Thr Thr Arg Leu Val	
525	535
aat cca gat caa ctg ttg aat gaa ata atg tct att ctt cca aag aag	1804
Asn Pro Asp Gln Leu Leu Asn Glu Ile Met Ser Ile Leu Pro Lys Lys	
540	550
cat gtt gac ttt gta caa aag ggt tat aca ctg aag tgt caa aca cag	1852
His Val Asp Phe Val Gln Lys Gly Tyr Thr Leu Lys Cys Gln Thr Gln	
555	565
tca gat ttt ggg aaa gtg aca atg caa ttt gaa tta gaa gtg tgc cag	1900
Ser Asp Phe Gly Lys Val Thr Met Gln Phe Glu Leu Glu Val Cys Gln	
570	585
ctt caa aaa ccc gat gtg gtg ggt atc agg agg cag cgg ctt aag ggc	1948
Leu Gln Lys Pro Asp Val Val Gly Ile Arg Arg Gln Arg Leu Lys Gly	
590	600

ES 2 357 658 T3

gat gcc tgg gtt tac aaa aga tta gtg gaa gac atc cta tct agc tgc 1996
 Asp Ala Trp Val Tyr Lys Arg Leu Val Glu Asp Ile Leu Ser Ser Cys
 605 610 615

aag gta taattgatgg attcttccat cctgccggat gagtgggggt gtgatacagc 2052
 Lys Val

ctacataaag actgttatga tcgctttgat tttaaagttc atiggaacta ccaacttgtt 2112

tctaaagagc tactttaaga ccaatatctc ttgttttta aacaaaagat attattttgt 2172

gtatgaatct aaatcaagcc catctgtcat taigtctacg tcttttttaa tcatgtggtt 2232

ttgtatatta ataattgttg actttcttag attcacttcc atatgtgaat gtaagctctt 2292

aactatgtct ctttgaatg tgtaatttct ttctgaaata aaaccatttg tgaatataaa 2352

aaaaaaaaaa aaaaaa 2368

<210> 6

<211> 619

<212> PRT

<213> *Homo sapiens*

<400> 6

Met Met Asn Phe Ser Asn Ile Met Asn Tyr Met Lys Leu Leu Gly Gln
 1 5 10 15

Ser Asp Leu Pro Arg Ile Lys Thr Glu Ile Glu Ala Leu Lys Asn Leu
 20 25 30

Arg His Gln His Ile Cys Gln Leu Tyr His Val Leu Glu Thr Ala Asn
 35 40 45

Lys Ile Phe Met Val Leu Glu Tyr Cys Pro Gly Gly Glu Leu Phe Asp
 50 55 60

Tyr Ile Ile Ser Gln Asp Arg Leu Ser Glu Glu Glu Thr Arg Val Val
 65 70 75 80

Phe Arg Gln Ile Val Ser Ala Val Ala Tyr Val His Ser Gln Gly Tyr
 85 90 95

Ala His Arg Asp Leu Lys Pro Glu Asn Leu Leu Phe Asp Glu Tyr His
 100 105 110

Lys Leu Lys Leu Ile Asp Phe Gly Leu Cys Ala Lys Pro Lys Gly Asn
 115 120 125

Lys Asp Tyr His Leu Gln Thr Cys Cys Gly Ser Leu Ala Tyr Ala Ala
 130 135 140

Pro Glu Leu Ile Gln Gly Lys Ser Tyr Leu Gly Ser Glu Ala Asp Val
 145 150 155 160

Trp Ser Met Gly Ile Leu Leu Tyr Val Leu Met Cys Gly Phe Leu Pro
 165 170 175

Phe Asp Asp Asp Asn Val Met Ala Leu Tyr Lys Lys Ile Met Arg Gly
 180 185 190

Lys Tyr Asp Val Pro Lys Trp Leu Ser Pro Ser Ser Ile Leu Leu Leu

ES 2 357 658 T3

195 200 205

Gln Gln Met Leu Gln Val Asp Pro Lys Lys Arg Ile Ser Met Lys Asn
 210 215 220

Leu Leu Asn His Pro Trp Ile Met Gln Asp Tyr Asn Tyr Pro Val Glu
 225 230 235 240

Trp Gln Ser Lys Asn Pro Phe Ile His Leu Asp Asp Asp Cys Val Thr
 245 250 255

Glu Leu Ser Val His His Arg Asn Asn Arg Gln Thr Met Glu Asp Leu
 260 265 270

Ile Ser Leu Trp Gln Tyr Asp His Leu Thr Ala Thr Tyr Leu Leu Leu
 275 280 285

Leu Ala Lys Lys Ala Arg Gly Lys Pro Val Arg Leu Arg Leu Ser Ser
 290 295 300

405 410 415

Asn Lys Asn Gln His Lys Arg Glu Ile Leu Thr Thr Pro Asn Arg Tyr
 420 425 430

Thr Thr Pro Ser Lys Ala Arg Asn Gln Cys Leu Lys Glu Thr Pro Ile
 435 440 445

Lys Ile Pro Val Asn Ser Thr Gly Thr Asp Lys Leu Met Thr Gly Val
 450 455 460

Ile Ser Pro Glu Arg Arg Cys Arg Ser Val Glu Leu Asp Leu Asn Gln
 465 470 475 480

Ala His Met Glu Glu Thr Pro Lys Arg Lys Gly Ala Lys Val Phe Gly
 485 490 495

Ser Leu Glu Arg Gly Leu Asp Lys Val Ile Thr Val Leu Thr Arg Ser
 500 505 510

Lys Arg Lys Gly Ser Ala Arg Asp Gly Pro Arg Arg Leu Lys Leu His
 515 520 525

Tyr Asn Val Thr Thr Thr Arg Leu Val Asn Pro Asp Gln Leu Leu Asn
 530 535 540

Glu Ile Met Ser Ile Leu Pro Lys Lys His Val Asp Phe Val Gln Lys
 545 550 555 560

Gly Tyr Thr Leu Lys Cys Gln Thr Gln Ser Asp Phe Gly Lys Val Thr
 565 570 575

Met Gln Phe Glu Leu Glu Val Cys Gln Leu Gln Lys Pro Asp Val Val
 580 585 590

Gly Ile Arg Arg Gln Arg Leu Lys Gly Asp Ala Trp Val Tyr Lys Arg
 595 600 605

Leu Val Glu Asp Ile Leu Ser Ser Cys Lys Val

610

615

<210> 7

<211> 2251

<212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

<222> (148)..(1887)

<400> 7

gaaaagattc ttaggaacgc cgtaccagcc gcgtctctca ggacagcagg ccctgtcct 60

tctgtcgggc gccgctcagc cgtgccctcc gccctcagg ttcttttct aattccaaat 120

aaacttgcaa gaggactatg aaagatt atg atg aac ttc tca aat att atg aat 174

Met Met Asn Phe Ser Asn Ile Met Asn

1 5

tac atg aaa cta ttg gga cag tac tgc cct gga gga gag ctg ttt gac 222

Tyr Met Lys Leu Leu Gly Gln Tyr Cys Pro Gly Gly Glu Leu Phe Asp

10 15 20 25

tat ata att tcc cag gat cgc ctg tca gaa gag gag acc cgg gtt gtc	270
Tyr Ile Ile Ser Gln Asp Arg Leu Ser Glu Glu Glu Thr Arg Val Val	
30 35 40	
ttc cgt cag ata gta tct gct gtt gct tat gtg cac agc cag ggc tat	318
Phe Arg Gln Ile Val Ser Ala Val Ala Tyr Val His Ser Gln Gly Tyr	
45 50 55	
gct cac agg gac ctc aag cca gaa aat ttg ctg ttt gat gaa tat cat	366
Ala His Arg Asp Leu Lys Pro Glu Asn Leu Leu Phe Asp Glu Tyr His	
60 65 70	
aaa tta aag ctg att gac ttt ggt ctc tgt gca aaa ccc aag ggt aac	414
Lys Leu Lys Leu Ile Asp Phe Gly Leu Cys Ala Lys Pro Lys Gly Asn	
75 80 85	
aag gat tac cat cta cag aca tgc tgt ggg agt ctg gct tat gca gca	462
Lys Asp Tyr His Leu Gln Thr Cys Cys Gly Ser Leu Ala Tyr Ala Ala	
90 95 100 105	
cct gag tta ata caa ggc aaa tca tat ctt gga tca gag gca gat gtt	510
Pro Glu Leu Ile Gln Gly Lys Ser Tyr Leu Gly Ser Glu Ala Asp Val	
110 115 120	
tgg agc atg ggc ata ctg tta tat gtt ctt atg tgt gga ttt cta cca	558
Trp Ser Met Gly Ile Leu Leu Tyr Val Leu Met Cys Gly Phe Leu Pro	

ES 2 357 658 T3

125	130	135	
ttt gat gat gat aat gta atg gct tta tac aag aag att atg aga gga			606
Phe Asp Asp Asp Asn Val Met Ala Leu Tyr Lys Lys Ile Met Arg Gly			
140	145	150	
aaa tat gat gtt ccc aag tgg ctc tct ccc agt agc att ctg ctt ctt			654
Lys Tyr Asp Val Pro Lys Trp Leu Ser Pro Ser Ser Ile Leu Leu Leu			
155	160	165	
caa caa atg ctg cag gtg gac cca aag aaa cgg att tct atg aaa aat			702
Gln Gln Met Leu Gln Val Asp Pro Lys Lys Arg Ile Ser Met Lys Asn			
170	175	180	185
cta ttg aac cat ccc tgg atc atg caa gat tac aac tat cct gtt gag			750
Leu Leu Asn His Pro Trp Ile Met Gln Asp Tyr Asn Tyr Pro Val Glu			
190	195	200	
tgg caa agc aag aat cct ttt att cac ctc gat gat gat tgc gta aca			798
Trp Gln Ser Lys Asn Pro Phe Ile His Leu Asp Asp Asp Cys Val Thr			
205	210	215	
gaa ctt tct gta cat cac aga aac aac agg caa aca atg gag gat tta			846
Glu Leu Ser Val His His Arg Asn Asn Arg Gln Thr Met Glu Asp Leu			
220	225	230	

att tca ctg tgg cag tat gat cac ctc acg gct acc tat ctt ctg ctt 894
 Ile Ser Leu Trp Gln Tyr Asp His Leu Thr Ala Thr Tyr Leu Leu Leu
 235 240 245

cta gcc aag aag gct cgg gga aaa cca gtt cgt tta agg ctt tct tct 942
 Leu Ala Lys Lys Ala Arg Gly Lys Pro Val Arg Leu Arg Leu Ser Ser
 250 255 260 265

ttc tcc tgt gga caa gcc agt gct acc cca ttc aca gac atc aag tca 990
 Phe Ser Cys Gly Gln Ala Ser Ala Thr Pro Phe Thr Asp Ile Lys Ser
 270 275 280

aat aat tgg agt ctg gaa gat gtg acc gca agt gat aaa aat tat gtg 1038
 Asn Asn Trp Ser Leu Glu Asp Val Thr Ala Ser Asp Lys Asn Tyr Val
 285 290 295

gcg gga tta ata gac tat gat tgg tgt gaa gat gat tta tca aca ggt 1086
 Ala Gly Leu Ile Asp Tyr Asp Trp Cys Glu Asp Asp Leu Ser Thr Gly
 300 305 310

gct gct act ccc cga aca tca cag ttt acc aag tac tgg aca gaa tca 1134
 Ala Ala Thr Pro Arg Thr Ser Gln Phe Thr Lys Tyr Trp Thr Glu Ser
 315 320 325

aat ggg gtg gaa tct aaa tca tta act cca gcc tta tgc aga aca cct 1182
 Asn Gly Val Glu Ser Lys Ser Leu Thr Pro Ala Leu Cys Arg Thr Pro

ES 2 357 658 T3

330	335	340	345	
gca aat aaa tta aag aac aaa gaa aat gta tat act cct aag tct gct				1230
Ala Asn Lys Leu Lys Asn Lys Glu Asn Val Tyr Thr Pro Lys Ser Ala				
	350	355	360	
gta aag aat gaa gag tac ttt atg ttt cct gag cca aag act cca gtt				1278
Val Lys Asn Glu Glu Tyr Phe Met Phe Pro Glu Pro Lys Thr Pro Val				
	365	370	375	
aat aag aac cag cat aag aga gaa ata ctc act acg cca aat cgt tac				1326
Asn Lys Asn Gln His Lys Arg Glu Ile Leu Thr Thr Pro Asn Arg Tyr				
	380	385	390	
act aca ccc tca aaa gct aga aac cag tgc ctg aaa gaa act cca att				1374
Thr Thr Pro Ser Lys Ala Arg Asn Gln Cys Leu Lys Glu Thr Pro Ile				
	395	400	405	
aaa ata cca gta aat tca aca gga aca gac aag tta atg aca ggt gtc				1422
Lys Ile Pro Val Asn Ser Thr Gly Thr Asp Lys Leu Met Thr Gly Val				
	410	415	420	425
att agc cct gag agg cgg tgc cgc tca gtg gaa ttg gat ctc aac caa				1470
Ile Ser Pro Glu Arg Arg Cys Arg Ser Val Glu Leu Asp Leu Asn Gln				
	430	435	440	

gca cat atg gag gag act cca aaa aga aag gga gcc aaa gtg ttt ggg	1518
Ala His Met Glu Glu Thr Pro Lys Arg Lys Gly Ala Lys Val Phe Gly	
445 450 455	
agc ctt gaa agg ggg ttg gat aag gtt atc act gtg ctc acc agg agc	1566
Ser Leu Glu Arg Gly Leu Asp Lys Val Ile Thr Val Leu Thr Arg Ser	
460 465 470	
aaa agg aag ggt tct gcc aga gac ggg ccc aga aga cta aag ctt cac	1614
Lys Arg Lys Gly Ser Ala Arg Asp Gly Pro Arg Arg Leu Lys Leu His	
475 480 485	
tat aat gtg act aca act aga tta gtg aat cca gat caa ctg ttg aat	1662
Tyr Asn Val Thr Thr Thr Arg Leu Val Asn Pro Asp Gln Leu Leu Asn	
490 495 500 505	
gaa ata atg tct att ctt cca aag aag cat gtt gac ttt gta caa aag	1710
Glu Ile Met Ser Ile Leu Pro Lys Lys His Val Asp Phe Val Gln Lys	
510 515 520	
ggt tat aca ctg aag tgt caa aca cag tca gat ttt ggg aaa gtg aca	1758
Gly Tyr Thr Leu Lys Cys Gln Thr Gln Ser Asp Phe Gly Lys Val Thr	
525 530 535	
atg caa ttt gaa tta gaa gtg tgc cag ctt caa aaa ccc gat gtg gtg	1806
Met Gln Phe Glu Leu Glu Val Cys Gln Leu Gln Lys Pro Asp Val Val	

ES 2 357 658 T3

540	545	550	
ggt atc agg agg cag cgg ctt aag ggc gat gcc tgg gtt tac aaa aga			1854
Gly Ile Arg Arg Gln Arg Leu Lys Gly Asp Ala Trp Val Tyr Lys Arg			
555	560	565	
tta gtg gaa gac atc cta tct agc tgc aag gta taattgatgg attcttccat			1907
Leu Val Glu Asp Ile Leu Ser Ser Cys Lys Val			
570	575	580	
cctgccgat gagtgtgggt gtgatacagc ctacataaag actgttatga tcgctttgat			1967
tttaaagttc attggaacta ccaacttgtt tctaaagagc tatcttaaga ccaatactc			2027
tttgttttta aacaaaagat attattttgt gtatgaatct aaatcaagcc catctgtcat			2087
tatgttactg tcttttttaa tcatgtgggt ttgtatatta ataattgttg actttcttag			2147
attcacttcc atatgtgaat gtaagctctt aactatgtct ctttgtaatg tgtaatttct			2207
ttctgaaata aaaccatttg tgaatataaa aaaaaaaaaa aaaa			2251

<210> 8

<211> 580

<212> PRT

<213> *Homo sapiens*

<400> 8

Met Met Asn Phe Ser Asn Ile Met Asn Tyr Met Lys Leu Leu Gly Gln
 1 5 10 15

Tyr Cys Pro Gly Gly Glu Leu Phe Asp Tyr Ile Ile Ser Gln Asp Arg
 20 25 30

Leu Ser Glu Glu Glu Thr Arg Val Val Phe Arg Gln Ile Val Ser Ala
 35 40 45

Val Ala Tyr Val His Ser Gln Gly Tyr Ala His Arg Asp Leu Lys Pro
 50 55 60

Glu Asn Leu Leu Phe Asp Glu Tyr His Lys Leu Lys Leu Ile Asp Phe
 65 70 75 80

Gly Leu Cys Ala Lys Pro Lys Gly Asn Lys Asp Tyr His Leu Gln Thr
 85 90 95

Cys Cys Gly Ser Leu Ala Tyr Ala Ala Pro Glu Leu Ile Gln Gly Lys
 100 105 110

Ser Tyr Leu Gly Ser Glu Ala Asp Val Trp Ser Met Gly Ile Leu Leu
 115 120 125

Tyr Val Leu Met Cys Gly Phe Leu Pro Phe Asp Asp Asp Asn Val Met
 130 135 140

Ala Leu Tyr Lys Lys Ile Met Arg Gly Lys Tyr Asp Val Pro Lys Trp
 145 150 155 160

Leu Ser Pro Ser Ser Ile Leu Leu Leu Gln Gln Met Leu Gln Val Asp
 165 170 175

Pro Lys Lys Arg Ile Ser Met Lys Asn Leu Leu Asn His Pro Trp Ile
 180 185 190

Met Gln Asp Tyr Asn Tyr Pro Val Glu Trp Gln Ser Lys Asn Pro Phe
 195 200 205

Ile His Leu Asp Asp Asp Cys Val Thr Glu Leu Ser Val His His Arg
 210 215 220

Asn Asn Arg Gln Thr Met Glu Asp Leu Ile Ser Leu Trp Gln Tyr Asp
 225 230 235 240

His Leu Thr Ala Thr Tyr Leu Leu Leu Leu Ala Lys Lys Ala Arg Gly
 245 250 255

Lys Pro Val Arg Leu Arg Leu Ser Ser Phe Ser Cys Gly Gln Ala Ser
 260 265 270

Ala Thr Pro Phe Thr Asp Ile Lys Ser Asn Asn Trp Ser Leu Glu Asp
 275 280 285

Val Thr Ala Ser Asp Lys Asn Tyr Val Ala Gly Leu Ile Asp Tyr Asp
 290 295 300

Asn Gln Cys Leu Lys Glu Thr Pro Ile Lys Ile Pro Val Asn Ser Thr
 405 410 415

Gly Thr Asp Lys Leu Met Thr Gly Val Ile Ser Pro Glu Arg Arg Cys
 420 425 430

Arg Ser Val Glu Leu Asp Leu Asn Gln Ala His Met Glu Glu Thr Pro
 435 440 445

Lys Arg Lys Gly Ala Lys Val Phe Gly Ser Leu Glu Arg Gly Leu Asp
 450 455 460

Lys Val Ile Thr Val Leu Thr Arg Ser Lys Arg Lys Gly Ser Ala Arg
 465 470 475 480

Asp Gly Pro Arg Arg Leu Lys Leu His Tyr Asn Val Thr Thr Thr Arg
 485 490 495

Leu Val Asn Pro Asp Gln Leu Leu Asn Glu Ile Met Ser Ile Leu Pro
 500 505 510

Lys Lys His Val Asp Phe Val Gln Lys Gly Tyr Thr Leu Lys Cys Gln

515

520

525

Thr Gln Ser Asp Phe Gly Lys Val Thr Met Gln Phe Glu Leu Glu Val

530

535

540

Cys Gln Leu Gln Lys Pro Asp Val Val Gly Ile Arg Arg Gln Arg Leu

545

550

555

560

Lys Gly Asp Ala Trp Val Tyr Lys Arg Leu Val Glu Asp Ile Leu Ser

565

570

575

Ser Cys Lys Val

580

<210> 9

<211> 20

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 9

cgaccacttt gtcaagctca

20

<210> 10

<211> 23

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 10

ggttgagcac aggtacttt att

23

<210> 11

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 11

tgggtaacaa gagaatggtt ca

22

<210> 12

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 12

atccaagtcc taatcccttt gg

22

<210> 13

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 13

gctgcaaggt ataattgatg ga

22

<210> 14

<211> 23

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 14

cagtaacata atgacagatg ggc

23

<210> 15

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 15

ttatcactgt gctcaccagg ag

22

<210> 16

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 16

aaacttgccct gccatatacct ta

22

<210> 17

<211> 22

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 17

atdddgttgg ctgtctctag ca

22

<210> 18

<211> 30

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 18

aaagaattcg ggtgtcgtta atgttcgggg

30

<210> 19

<211> 30

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 19

aaagcggccg cttaggcgga ttttcctgca

30

<210> 20

<211> 30

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 20

cggaattcac tatgaaagat tatgatgaac

30

<210> 21

<211> 30

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 21

aaactcgagt accttcagc tagataggat

30

<210> 22

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 22

tcccgcgcgc ttgtaggat tcgttcaaga gacgaatcct acaaagcgcg c 51

<210> 23

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para ARNip

<400> 23

aaaagcgcgc ttgtaggat tcgtctcttg aacgaatcct acaaagcgcg c 51

<210> 24

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 24

tccccgtacg cggaatactt cgattcaaga gatcgaagta ttccgcgtac g 51

<210> 25

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 25

aaaacgtacg cggaaactt cgaatcttg aatcgaagta ttccgcgtac g 51

<210> 26

<211> 20

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 26

ttagctgtgc tcgcgtact 20

<210> 27

<211> 20

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 27

tcacatggtt cacacggcag 20

<210> 28

<211> 25

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 28

ttaagtgaag gctctgattc tagtt

25

<210> 29

<211> 21

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para RT-PCR

<400> 29

gtccttattg gctggttcgt t

21

<210> 30

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 30

tcccgtatat cttgccctct gaattcaaga gattcagagg gcaagatata c

51

<210> 31

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 31

aaaagtatat ctigccctct gaatctcttg aattcagagg gcaagatata c 51

<210> 32

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 32

tcccgiccga acacatcttt gtttcaaga gaaacaaaga tggttcggga c 51

<210> 33

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 33

aaaagtcgga acacatcttt gtttctcttg aaaacaaaga tggttcggga c 51

<210> 34

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 34

tcccgacatc ctatctagct gcattcaaga gatgcagcta gataggatgt c 51

<210> 35

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 35

aaaagacatc ctatctagct gcatctcttg aatgcagcta gataggatgt c 51

<210> 36

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 36

tcccagtca ttggaactac caattcaaga gattgtagt tccaatgaac t 51

<210> 37

<211> 51

<212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintetizado de manera artificial para ARNip

<400> 37

aaaaagtcca ttggaactac caatctcttg aattggtagt tccaatgaac t 51

<210> 38

<211> 19

<212> ADN

<213> Artificial

<220>

<223> Secuencia diana para ARNip

<400> 38

gtatatcttg ccctctgaa 19

<210> 39

<211> 19

<212> ADN

<213> Artificial

<220>

<223> Secuencia diana para ARNip

<400> 39

gtccgaacac atctttgtt 19

<210> 40

<211> 19

<212> ADN

<213> Artificial

<220>

<223> Secuencia diana para ARNip

<400> 40

gacatcctat ctagctgca

19

<210> 41

<211> 19

<212> ADN

<213> Artificial

<220>

<223> Secuencia diana para ARNip

<400> 41

agttcattgg aactaccaa

19

<210> 42

<211> 32

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 42

acggaattca tcatgcaaga ttacaactat cc

32

<210> 43

<211> 33

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 43

gacggaattc aatatggagg agactccaaa aag

33

<210> 44

<211> 30

<212> ADN

<213> Artificial

<220>

<223> Cebador sintetizado de manera artificial para PCR

<400> 44

ccctcgagta ccttgcagct agataggatg

30

REIVINDICACIONES

1. ARN de interferencia pequeño, en el que la cadena sentido del mismo se selecciona del grupo que consiste en las secuencias de nucleótidos de SEQ ID NO: 40 y 41 como secuencia diana.
2. Método de selección de un compuesto útil en el tratamiento del cáncer de mama, comprendiendo dicho método las etapas de:
 - (a) poner en contacto un compuesto de prueba con un polipéptido seleccionado del grupo que consiste en:
 - (1) un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 4, 6 u 8;
 - (2) un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 4, 6 u 8 en la que uno o más aminoácidos están sustituidos, delecionados, insertados y/o añadidos y que tiene una identidad del 95% con SEQ ID NO: 4, 6 u 8, y una actividad biológica de potenciación de la proliferación celular y la actividad cinasa; y
 - (3) un polipéptido codificado por un polinucleótido que se hibrida en condiciones rigurosas con un polinucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO: 3, 5 ó 7, en el que el polipéptido tiene una actividad biológica de potenciación de la proliferación celular y la actividad cinasa
 - (b) detectar la actividad de unión entre el polipéptido y el compuesto de prueba; y
 - (c) seleccionar el compuesto de prueba que se une al polipéptido e inhibe su actividad biológica de potenciación de la proliferación celular y la actividad cinasa.
3. Método de selección de un compuesto útil en el tratamiento del cáncer de mama, comprendiendo dicho método las etapas de:
 - (a) poner en contacto un compuesto de prueba con una célula que expresa uno o más polinucleótidos que comprenden la secuencia de nucleótidos de SEQ ID NO: 3, 5 ó 7; y
 - (b) seleccionar un compuesto que reduce el nivel de expresión de uno o más polinucleótidos que comprenden la secuencia de nucleótidos de SEQ ID NO: 3, 5 ó 7 en comparación con el nivel de expresión detectado en ausencia del compuesto de prueba.
4. Método de selección de un compuesto útil en el tratamiento del cáncer de mama, comprendiendo dicho método las etapas de:
 - (a) poner en contacto un compuesto de prueba con un polipéptido seleccionado del grupo que consiste en:
 - (1) un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 4, 6 u 8;
 - (2) un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 4, 6 u 8 en el que uno o más aminoácidos están sustituidos, delecionados, insertados y/o añadidos y que tiene una identidad del 95% con SEQ ID NO: 4, 6 u 8, y una actividad biológica de potenciación de la proliferación celular y la actividad cinasa; y
 - (3) un polipéptido codificado por un polinucleótido que se hibrida en condiciones rigurosas con un polinucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO: 3, 5 ó 7, en el que el polipéptido tiene una actividad biológica de potenciación de la proliferación celular y la actividad cinasa;
 - (b) detectar la actividad biológica de potenciación de la proliferación celular y la actividad cinasa del polipéptido de la etapa (a); y
 - (c) seleccionar un compuesto que suprime la actividad biológica del polipéptido en comparación con la actividad biológica detectada en ausencia del compuesto de prueba.
5. Método de selección de un compuesto útil en el tratamiento del cáncer de mama, comprendiendo dicho método las etapas de:
 - (a) poner en contacto un compuesto de prueba con una célula en la que se ha introducido un vector que comprende la región reguladora de la transcripción de un gen marcador y un gen indicador que se expresa bajo el control de la región reguladora de la transcripción, en el que el gen marcador comprende una cualquiera de las secuencias de nucleótidos seleccionadas del grupo que consiste en SEQ ID:NO 3, 5 y 7,

- (b) medir la expresión o actividad de dicho gen indicador; y
 - (c) seleccionar el compuesto que reduce el nivel de expresión o actividad de dicho gen indicador en comparación con el nivel de expresión o actividad de dicho gen indicador detectado en ausencia del compuesto de prueba.
6. Composición farmacéutica que comprende una cantidad farmacéuticamente eficaz de un polinucleótido antisentido o ARN de interferencia pequeño frente a un polinucleótido que consiste en la secuencia de nucleótidos SEQ ID NO: 3, 5 ó 7 como principio activo, y un vehículo farmacéuticamente aceptable para su uso en el tratamiento del cáncer de mama.
7. Uso de una cantidad farmacéuticamente eficaz de un polinucleótido antisentido o ARN de interferencia pequeño frente a un polinucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO: 3, 5 ó 7 para la preparación de una composición farmacéutica para tratar el cáncer de mama .

Fig. 1

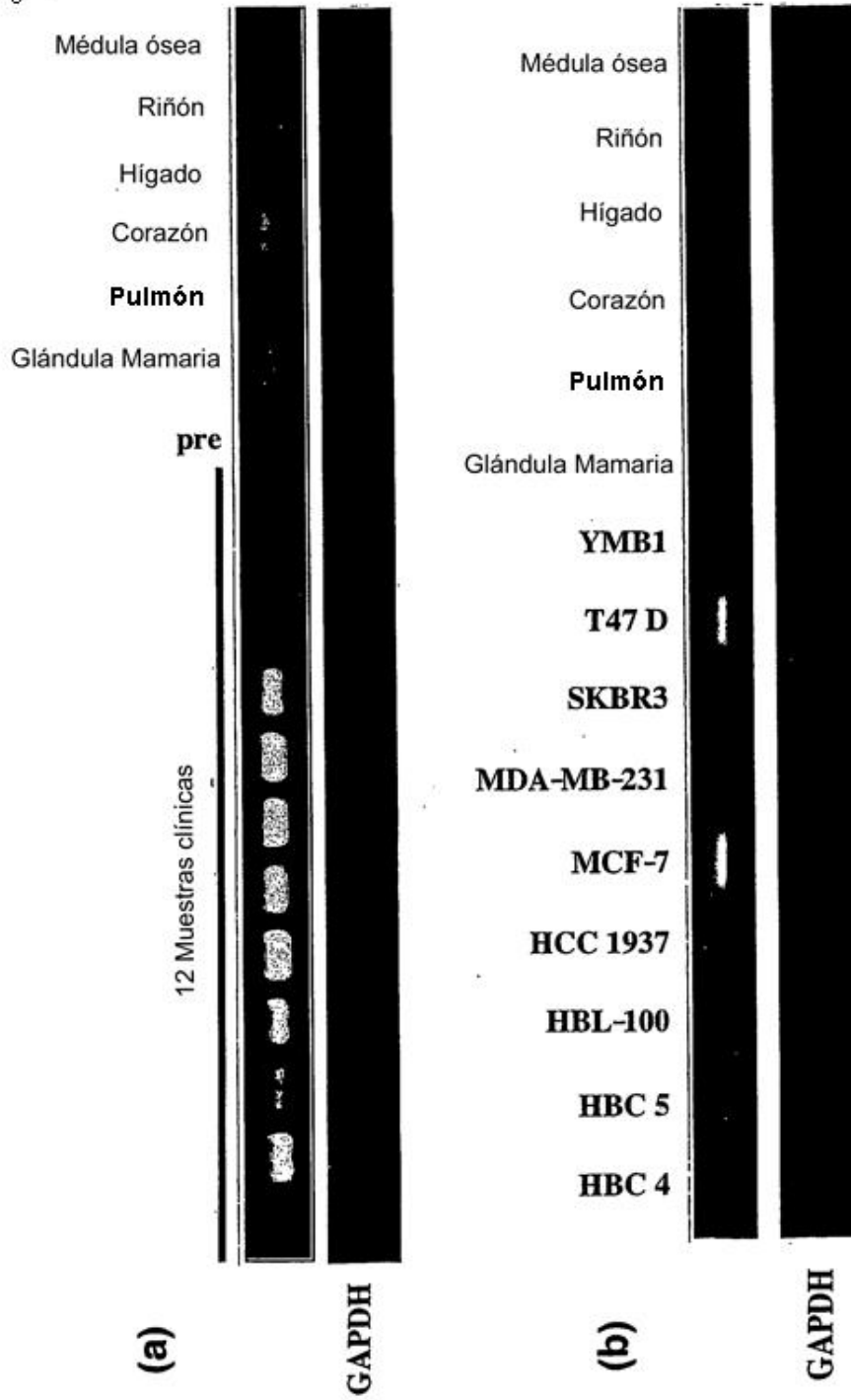
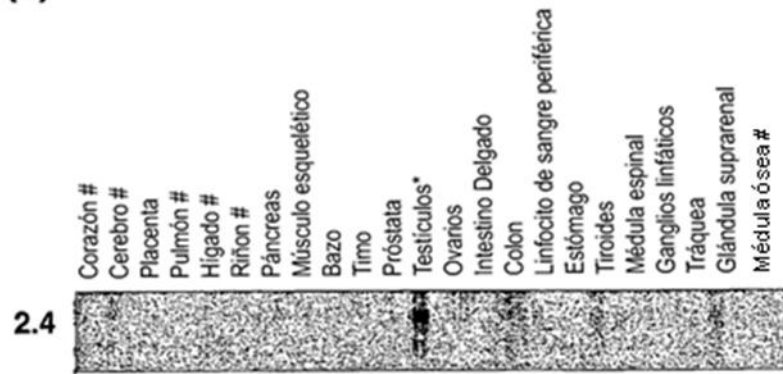


Fig. 2

(a)



(b)

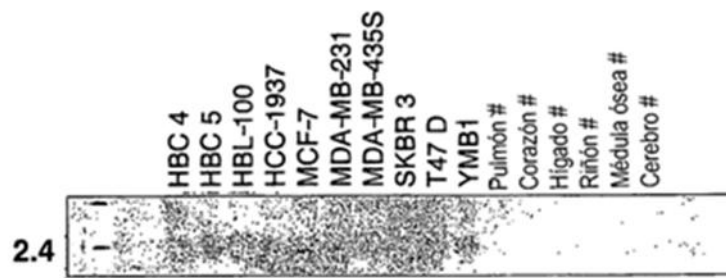
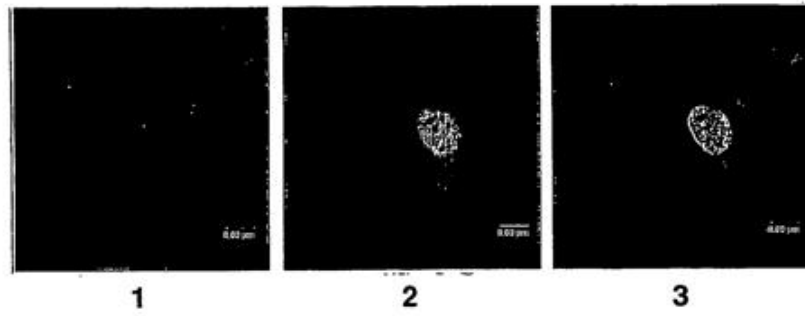


Fig 3

(a)



(b)

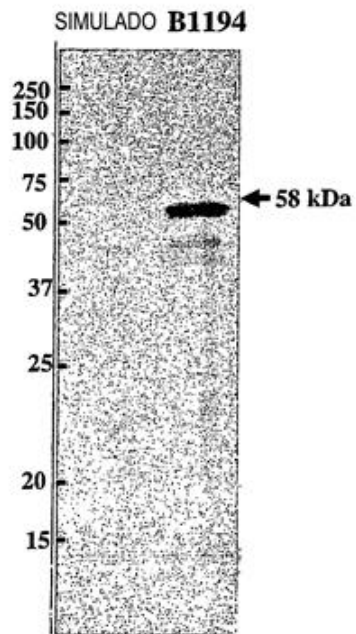


Fig 4

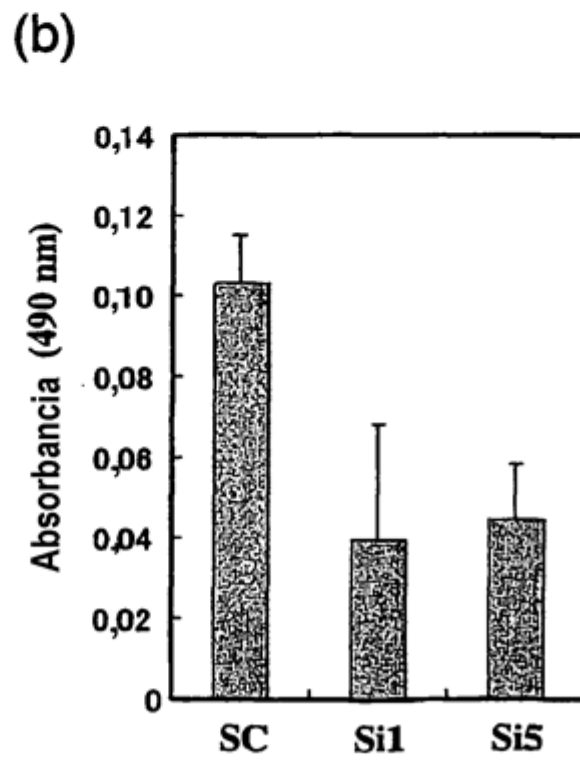
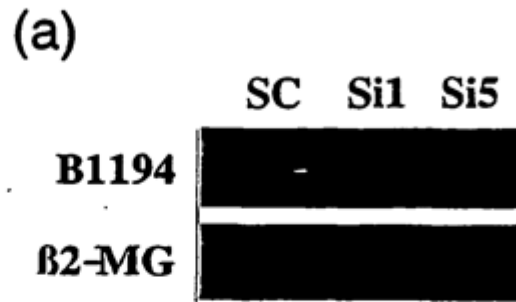


Fig. 5

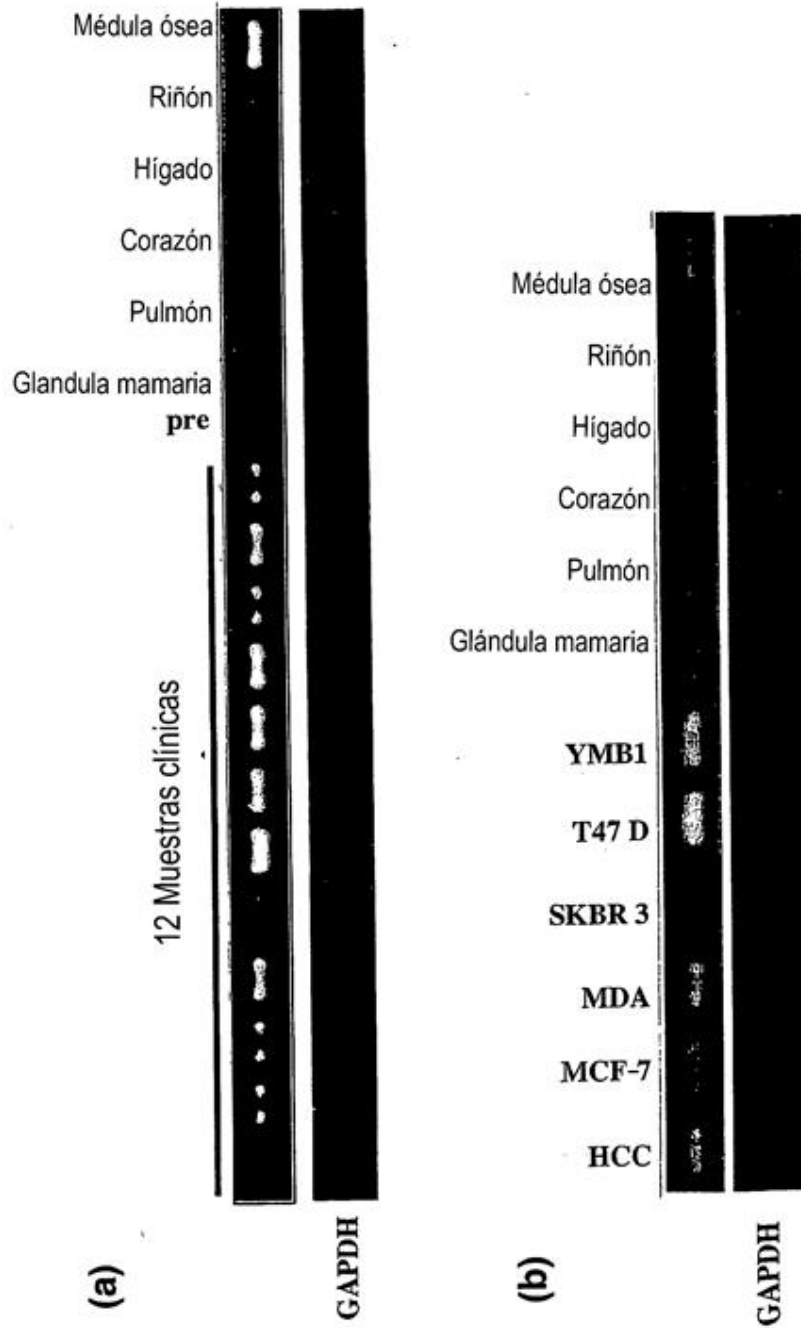


Fig. 6

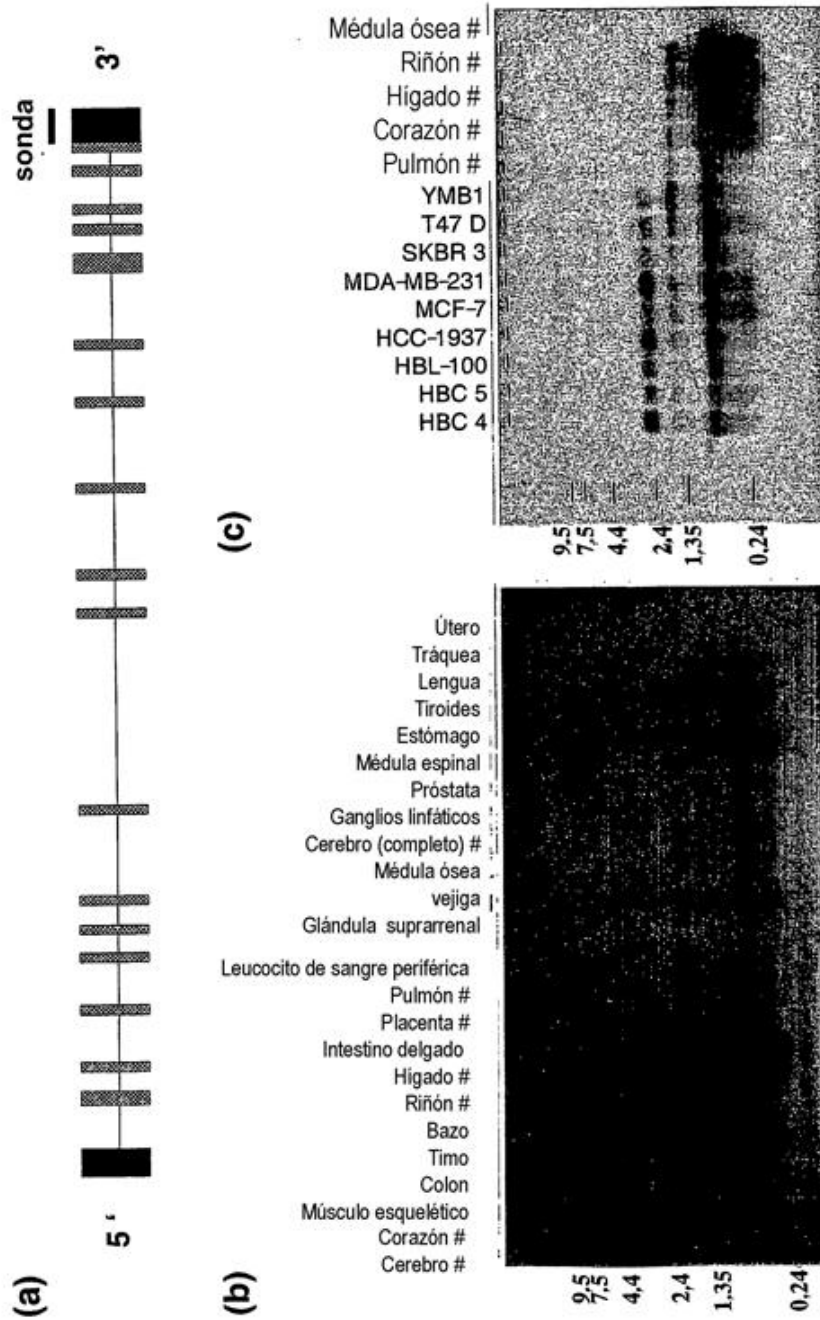


Fig 7

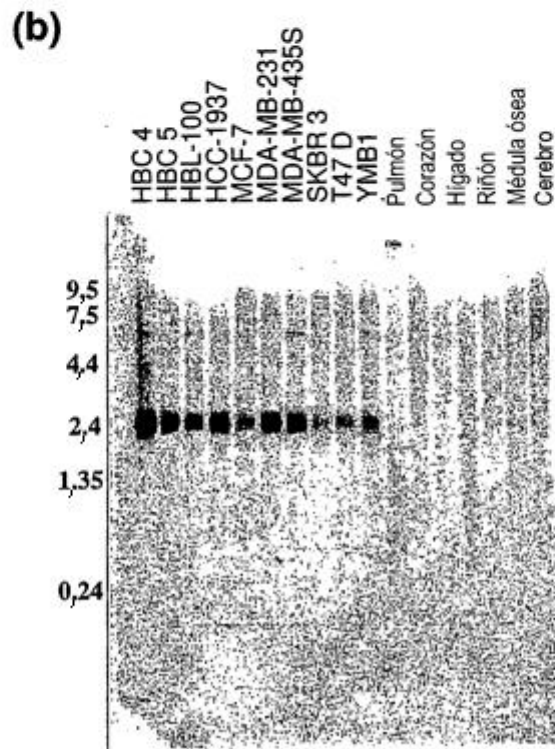
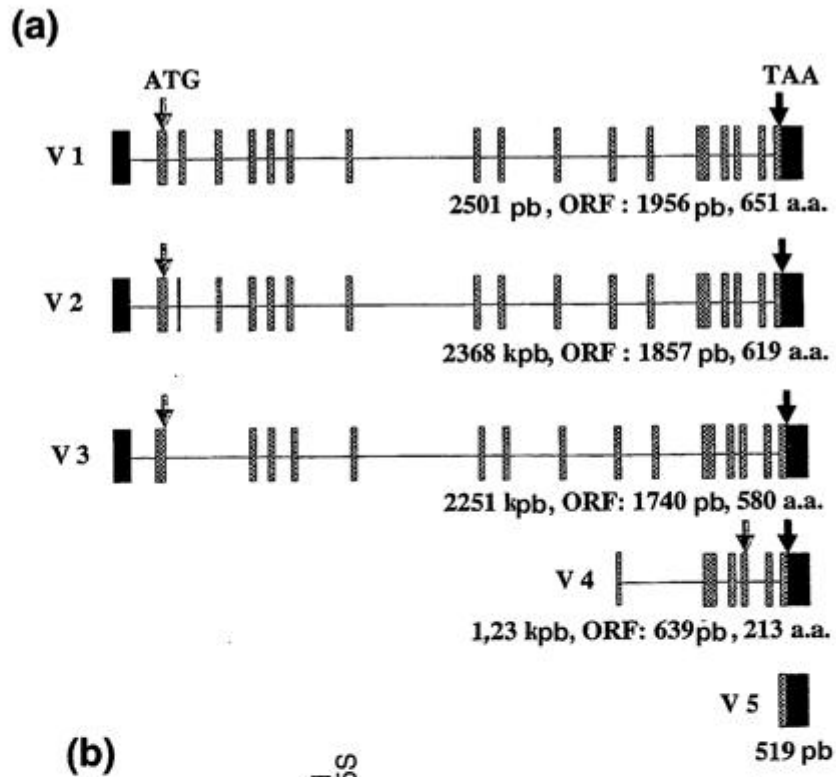


Fig 8

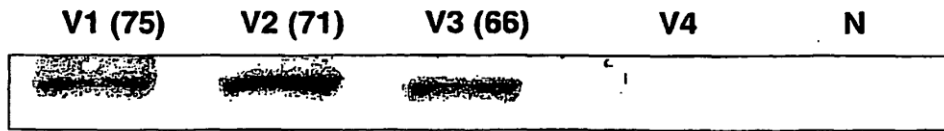
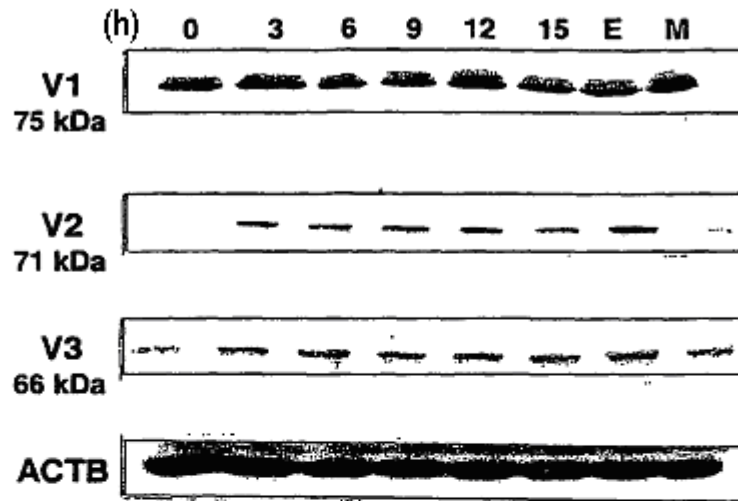


Fig 9

(a)



(b)

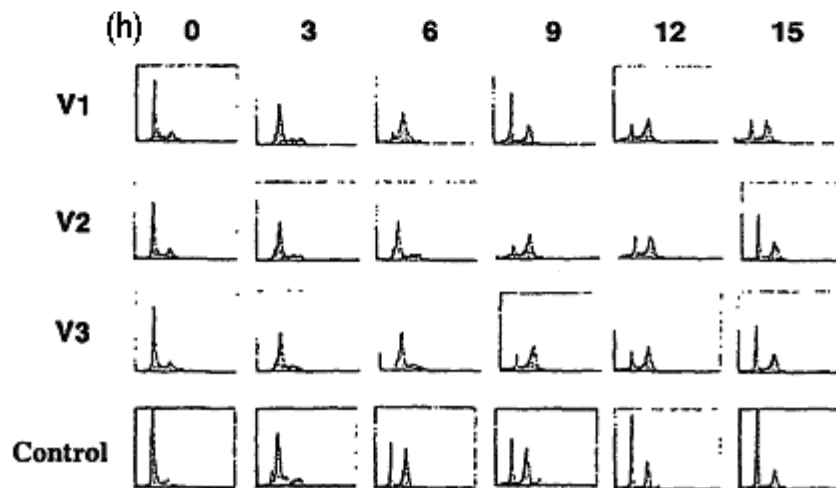
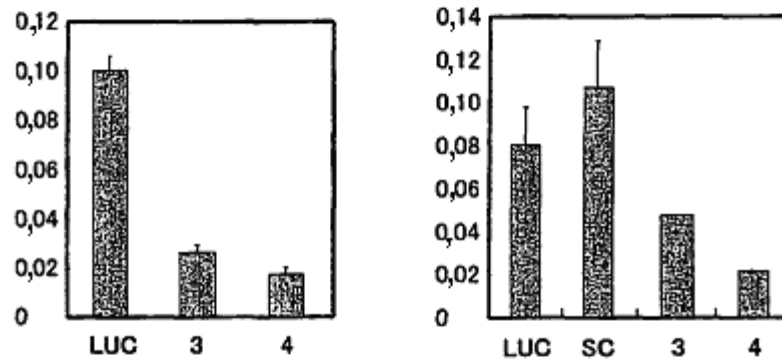


Fig 10

(a)



(b)



(c)

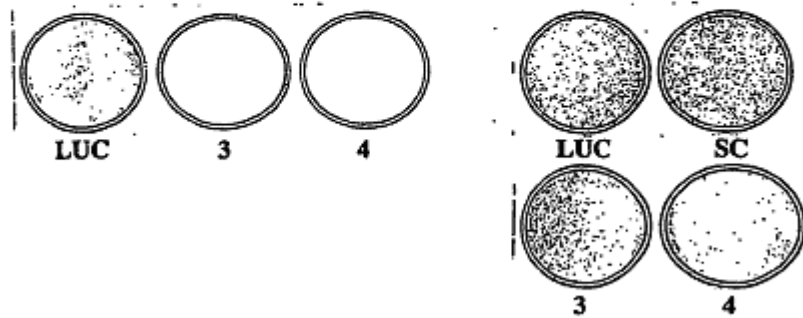


Fig. 11

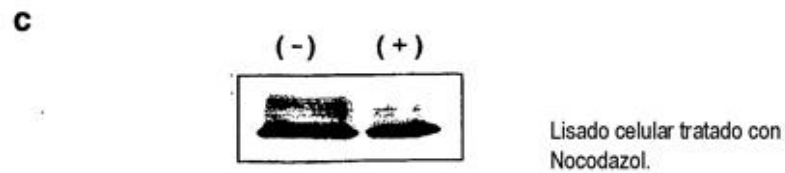
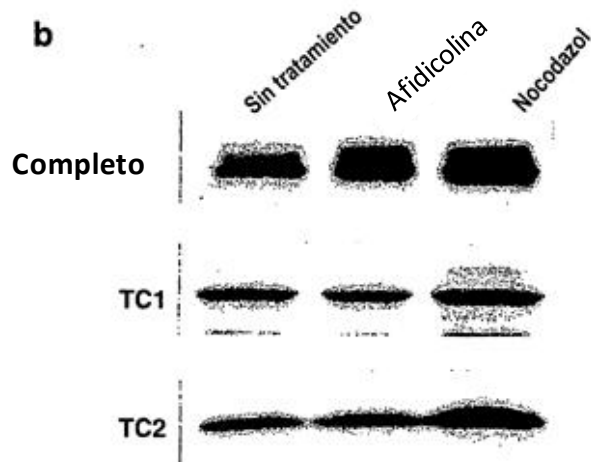
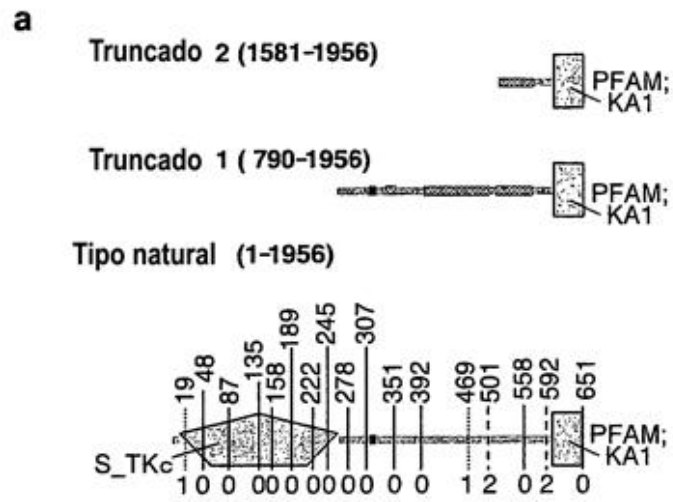
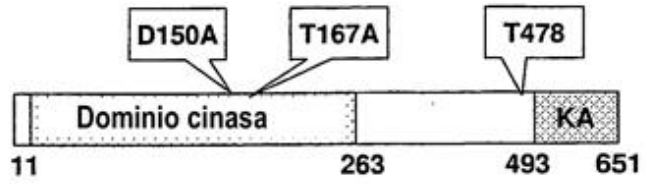
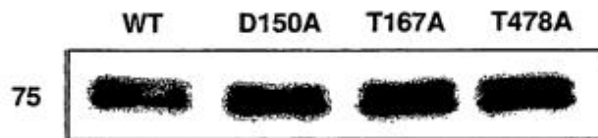


Fig. 12

a



b



c

