



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 279 473**

51 Int. Cl.:
C07K 14/47 (2006.01)
C12Q 1/68 (2006.01)
G01N 33/574 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **05018356 .5**
86 Fecha de presentación : **11.02.2000**
87 Número de publicación de la solicitud: **1607402**
87 Fecha de publicación de la solicitud: **21.12.2005**

54 Título: **Procedimiento de diagnosis de tumores.**

30 Prioridad: **08.03.1999 WO PCT/US99/05028**
11.03.1999 US 123972 P
11.05.1999 US 133459 P
02.06.1999 WO PCT/US99/12252
22.06.1999 US 140650 P
22.06.1999 US 140653 P
20.07.1999 US 144758 P
26.07.1999 US 145698 P
28.07.1999 US 146222 P
17.08.1999 US 149395 P
31.08.1999 US 151689 P
01.09.1999 WO PCT/US99/20111
15.09.1999 WO PCT/US99/21090
30.11.1999 WO PCT/US99/28313
01.12.1999 WO PCT/US99/28301
01.12.1999 WO PCT/US99/28634
05.01.2000 WO PCT/US00/00219

45 Fecha de publicación de la mención BOPI:
16.08.2007

45 Fecha de la publicación del folleto de la patente:
16.08.2007

73 Titular/es: **GENENTECH, Inc.**
1 DNA Way
South San Francisco, California 94080-4990, US

72 Inventor/es: **Ashkenazi, Avi J.;**
Goddard, Audrey;
Godowski, Paul J.;
Gurney, Austin L.;
Hillan, Kenneth J.;
Masters, Scot A.;
Pan, James;
Pitti, Robert M.;
Roy, Margaret Ann;
Smith, Victoria;
Stone, Donna M.;
Watanabe, Colin K. y
Wood, William I.

74 Agente: **Ponti Sales, Adelaida**

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

DESCRIPCIÓN

Procedimiento de diagnóstico de tumores.

5 **Campo de la invención**

La presente invención hace referencia a las composiciones y procedimientos para el diagnóstico y tratamiento de tumores.

10 **Antecedentes de la invención**

Los tumores malignos (cáncer) son la segunda causa de muerte en los Estados Unidos, después de las enfermedades del corazón (Boring y col., CA Cancer J Clin., 43:7 [(198)]).

15 El cáncer está caracterizado por un aumento en el número de células anómalas o neoplásicas derivadas de un tejido normal que prolifera hasta formar una masa tumoral, por la invasión de los tejidos adyacentes por estas células neoplásicas tumorales, y por la generación de células malignas que eventualmente se propagan por la sangre o el sistema linfático a los nódulos linfáticos regionales y los sitios distantes (metástasis). En un estado canceroso, una célula prolifera en condiciones tales, en las que una célula normal no crecería. El cáncer se manifiesta en una gran
20 variedad de formas que se caracterizan por grados distintos de invasión y agresividad.

La alteración en la expresión génica está íntimamente relacionada con el crecimiento celular incontrolado y la re-
diferenciación que es una característica común de todos los cánceres. Los genomas de ciertos tumores bien estudiados
25 han demostrado presentar una expresión disminuida de genes recesivos, referidos generalmente como genes supresores de tumores, que normalmente funcionan para prevenir el crecimiento de las células malignas, y/o la sobreexpresión de ciertos genes dominantes, como los oncogenes, que actúan para promover el crecimiento maligno. Cada uno de estos cambios genéticos parece ser responsable de aportar algunas de estas características que, en conjunto, representan el fenotipo neoplásico completo (Hunter, Cell, 64:1129 [1991] y Bishop, Cell, 64:235-248 [1991]).

30 Un mecanismo bien conocido de la sobreexpresión (por ejemplo de un oncogen) en las células cancerosas es la amplificación génica. Este es un proceso que implica la replicación no programada de la región del cromosoma que comprende el gen, seguido de la recombinación de los segmentos replicados en el cromosoma (Alitalo y col., Adv. Cancer Res., 47:235-281 [1986]). Se cree que la sobreexpresión de genes es paralela a la amplificación génica, es decir, es proporcional al número de copias realizadas.

35 Se han identificado protooncogenes que codifican factores de crecimiento y receptores de los factores de crecimiento, los cuales desempeñan un papel importante en la patogénesis de diversos procesos malignos, incluyendo el cáncer de mama. Por ejemplo, se ha hallado que el gen ErbB2 (erbB2, también denominado her2, o c-erbB-2) que codifica un receptor glicoproteína transmembrana de 185 KD (p185HER2; HER) y que está relacionado con el receptor del factor de crecimiento epitelial EGFR se sobreexpresa en aproximadamente el 25%-30% de cánceres de mama (Slamon y col., Science, 235:177-182 [1987]; Slamon y col., Science, 244:707-712 [1989]).

45 Se ha publicado que la amplificación génica de un proto-oncogen es un evento típicamente implicado en las formas más malignas de cáncer, y podría actuar como un predictor del resultado clínico (Schwab y col., Genes Chromosomes Cancer, 1:181-193 [1990]; Alitalo y col., *supra*). Por ello, la sobreexpresión de erbB2 se considera generalmente como un indicador de mal pronóstico, especialmente en pacientes con enfermedad primaria que implica los nódulos linfáticos axilares (Slamon y col., [1987] y [1989], *supra*; Ravdin y Chamness, Gene, 159:19-27 [1995]; y Hynes y Stern, Biochim. Biophys. Acta, 1198:165-184 [1994]), y que se ha relacionado con la sensibilidad y/o resistencia a la terapia hormonal y los regímenes de terapia y quimioterapia, incluyendo el CMF (ciclofosfamida, metotrexato y fluorouracil) y las antraciclinas (Baselga y col., Oncology, 11 (3, Suppl. 43-48 [1997])). Sin embargo, a pesar de la asociación de la sobreexpresión de erbB2 con un pronóstico pobre, la probabilidad de mejoría en los pacientes
50 HER2 positivos que responden clínicamente al tratamiento con taxanos fue tres veces superior a la de los pacientes HER-2 negativos (*ibid*). Un anticuerpo monoclonal anti-ErbB2 (anti-HER) humanizado recombinante (una versión humanizada del anticuerpo anti-ErbB2 murino, 4D5, referido como rhuMAb HER2 o Herceptina™) ha sido activo clínicamente en pacientes con cáncer de mama metastásico que sobreexpresan ErbB2 y que habían recibido con anterioridad una terapia anticancerosa extensa. (Baselga y col., J. Clin. Oncol., 14:737-744 [1996]).

60 A la luz de los apartados anteriores, existe un interés obvio en la identificación de nuevos procedimientos y composiciones que sean de utilidad para el diagnóstico y tratamiento de tumores, que estén asociados con la amplificación génica.

Resumen de la invenciónA. *Realizaciones*

65 La presente invención hace referencia a las composiciones y procedimientos para el diagnóstico del crecimiento de células neoplásicas y la proliferación en mamíferos, incluyendo el hombre. La presente invención se basa en la identificación de genes que están amplificados en el genoma de las células tumorales. Dicha amplificación génica

se espera que esté asociada con la sobreexpresión del producto génico y que contribuya a la tumorigénesis. Según ello, las proteínas codificadas por los genes amplificados se cree que son dianas de utilidad para el diagnóstico y/o el tratamiento (incluyendo la prevención) de ciertos cánceres y pueden, por tanto, actuar como factores de predicción del pronóstico del tratamiento del tumor.

5

En una realización, la presente invención concierne un anticuerpo aislado que se une a un polipéptido denominado aquí como polipéptido PRO5775. En un aspecto, el anticuerpo aislado específicamente se une al PRO5775. A menudo, una célula que expresa el polipéptido PRO5775 es una célula tumoral que sobreexpresa el polipéptido en comparación con una célula normal del mismo tejido. En otro aspecto, el anticuerpo es un anticuerpo monoclonal, que preferentemente no tiene residuos de la región determinante de la complementariedad (CDR), ni residuos de la región de entramado (FR). El anticuerpo puede marcarse e inmovilizarse en un soporte sólido. En otro aspecto, el anticuerpo es un fragmento de anticuerpo, un anticuerpo de una sola cadena, o un anticuerpo humanizado que se une, preferentemente y de forma específica, con un polipéptido PRO5775.

En otra realización, la invención se refiere a una composición de una sustancia que comprende un anticuerpo, que se une, preferentemente y de forma específica a un polipéptido PRO5775 en una mezcla con un vehículo farmacéuticamente aceptable. En un aspecto, la composición de la sustancia comprende una cantidad efectiva del anticuerpo. En otro aspecto, la composición comprende además un ingrediente activo, el cual puede, por ejemplo, ser un anticuerpo o un agente citotóxico o quimioterapéutico. Preferentemente, la composición será estéril.

20

La invención puede utilizar moléculas de ácido nucleico aisladas que codifican anticuerpos anti-PRO5775 y vectores y células huésped recombinantes que comprenden dichas moléculas de ácido nucleico.

Se describen aquí los procedimientos para la producción de un anticuerpo anti-PRO5775, en donde el procedimiento comprende el cultivo de una célula huésped transformada con una molécula de ácido nucleico que codifica el anticuerpo en condiciones suficientes para permitir la expresión del anticuerpo y la recuperación del mismo a partir del cultivo celular.

En otra realización, la invención hace referencia a un procedimiento para la determinación de la presencia de un polipéptido PRO5775 en una muestra sospechosa de contener un polipéptido PRO5775, en donde el procedimiento comprende la exposición de la muestra a un anticuerpo anti-PRO5775 y la determinación de la unión del anticuerpo con un polipéptido PRO5775 en la muestra. En otra realización, la invención se refiere a un procedimiento para la determinación de la presencia de un polipéptido en una célula, en donde el procedimiento comprende la exposición de la célula a un anticuerpo anti-PRO5775 y la determinación de la unión del anticuerpo a la célula.

35

En otra realización, la presente invención se refiere a un procedimiento de diagnóstico de un tumor en un mamífero, que comprende la detección del nivel de expresión de un gen que codifica un polipéptido (a) en una muestra a analizar de células del tejido obtenidas del mamífero, y (b) en una muestra control de células de tejido normal del mismo tipo celular, en donde un nivel de expresión superior en la muestra a analizar respecto a la muestra control es indicativo de la presencia del tumor en el mamífero del cual se obtienen las células procedentes del tejido determinado.

40

En otra realización, la presente invención se refiere a un procedimiento de diagnóstico tumoral en un mamífero, y comprende (a) el contacto de un anticuerpo anti-PRO5775 con una muestra a analizar de las células del tejido obtenido del mamífero, y (b) la detección de la formación de un complejo entre el anticuerpo anti-PRO5775 y un polipéptido PRO5775 en la muestra a analizar, en donde la formación de un complejo es indicativo de la presencia de un tumor en dicho animal. La detección puede ser cualitativa o cuantitativa, y puede realizarse por comparación con la monitorización de la formación del complejo en una muestra control de células de tejido normal conocido del mismo tipo celular. Una gran cantidad de complejos formados en la muestra a analizar indica la presencia del tumor en el mamífero del cual se han obtenido las células del tejido a analizar. El anticuerpo tiene preferentemente un marcaje detectable. La formación del complejo puede monitorizarse, por ejemplo, mediante microscopía, citometría de flujo, fluorimetría, u otras técnicas conocidas en la materia.

50

La muestra a analizar se obtiene generalmente de un individuo sospechoso de presentar un crecimiento o proliferación celular neoplásicos (por ejemplo, células cancerosas).

55

En otra realización, la presente invención hace referencia a un equipo de diagnóstico para el cáncer que comprende un anticuerpo anti-PRO5775 y un vehículo (por ejemplo un tampón) empaquetados de forma adecuada. El equipo contiene preferentemente las instrucciones para utilizar el anticuerpo con el fin de detectar la presencia de un polipéptido PRO5775 en una muestra sospechosa de contener el mismo.

60

B. Realizaciones adicionales

En otras realizaciones de la presente invención, la invención utiliza una molécula de ácido nucleico aislado que comprende una secuencia nucleotídica que codifica un polipéptido PRO5775.

65

En un aspecto, la molécula de ácido nucleico comprende una secuencia nucleotídica que tiene por lo menos el 80% de identidad de secuencia, preferentemente al menos un 81% de identidad de secuencia, más preferentemente al menos un 82% de identidad de secuencia, aún más preferentemente al menos un 83% de identidad de secuencia,

ES 2 279 473 T3

5 todavía más preferentemente al menos un 84% de identidad de secuencia, todavía más preferentemente al menos un 85% de identidad de secuencia, todavía más preferentemente al menos un 86% de identidad de secuencia, todavía más preferentemente al menos un 87% de identidad de secuencia, todavía más preferentemente al menos un 88% de identidad de secuencia, todavía más preferentemente al menos un 89% de identidad de secuencia, todavía más preferentemente al menos un 90% de identidad de secuencia, todavía más preferentemente al menos un 91% de identidad de secuencia, todavía más preferentemente al menos un 92% de identidad de secuencia, todavía más preferentemente al menos un 93% de identidad de secuencia, todavía más preferentemente al menos un 94% de identidad de secuencia, todavía más preferentemente al menos un 95% de identidad de secuencia, todavía más preferentemente al menos un 96% de identidad de secuencia, todavía más preferentemente al menos un 97% de identidad de secuencia, todavía más preferentemente al menos un 98% de identidad de secuencia, todavía más preferentemente al menos un 99% de identidad de secuencia con (a) una molécula de ADN que codifica un polipéptido PRO5775 con una secuencia aminoacídica completa tal como se describe aquí, una secuencia aminoacídica carente del péptido señal tal como se describe aquí, un dominio extracelular de una proteína transmembrana, con o sin el péptido señal, tal como se describe aquí o cualquier otro fragmento definido específicamente de la secuencia aminoacídica completa tal como se describe aquí, o (b) el complemento de la molécula de ADN de (a).

20 En otros aspectos, la molécula de ácido nucleico aislada comprende una secuencia nucleotídica con al menos un 80% de identidad de secuencia, preferentemente con al menos un 81% de identidad de secuencia, más preferentemente con al menos un 82% de identidad de secuencia, todavía más preferentemente con al menos un 83% de identidad de secuencia, todavía más preferentemente con al menos un 84% de identidad de secuencia, todavía más preferentemente con al menos un 85% de identidad de secuencia, todavía más preferentemente con al menos un 86% de identidad de secuencia, todavía más preferentemente con al menos un 87% de identidad de secuencia, todavía más preferentemente con al menos un 88% de identidad de secuencia, todavía más preferentemente con al menos un 89% de identidad de secuencia, todavía más preferentemente con al menos un 90% de identidad de secuencia, todavía más preferentemente con al menos un 91% de identidad de secuencia, todavía más preferentemente con al menos un 92% de identidad de secuencia, todavía más preferentemente con al menos un 93% de identidad de secuencia, todavía más preferentemente con al menos un 94% de identidad de secuencia, todavía más preferentemente con al menos un 95% de identidad de secuencia, todavía más preferentemente con al menos un 96% de identidad de secuencia, todavía más preferentemente con al menos un 97% de identidad de secuencia, todavía más preferentemente con al menos un 98% de identidad de secuencia, todavía más preferentemente con al menos un 99% de identidad de secuencia con (a) una molécula de ADN que comprende la secuencia codificante de un cADN de un polipéptido de tamaño completo PRO5775 tal como se describe aquí, la secuencia codificante de un polipéptido PRO5775 carente del péptido señal tal como se describe aquí, la secuencia codificante de un dominio extracelular de un polipéptido transmembrana PRO5775, con o sin el péptido señal, tal como se describe aquí o la secuencia codificante de cualquier otro fragmento específicamente definido de la secuencia aminoacídica de tamaño completo, tal como se describe aquí, o (b) la complementaria de la molécula de ADN de (a).

40 En otro aspecto, la invención puede utilizar una molécula de ácido nucleico aislado que comprende una secuencia nucleotídica con al menos el 80% de identidad de secuencia, preferentemente con al menos un 81% de identidad de secuencia, más preferentemente con al menos un 82% de identidad de secuencia, todavía más preferentemente con al menos un 83% de identidad de secuencia, todavía más preferentemente con al menos un 84% de identidad de secuencia, todavía más preferentemente con al menos un 85% de identidad de secuencia, todavía más preferentemente con al menos un 86% de identidad de secuencia, todavía más preferentemente con al menos un 87% de identidad de secuencia, todavía más preferentemente con al menos un 88% de identidad de secuencia, todavía más preferentemente con al menos un 89% de identidad de secuencia, todavía más preferentemente con al menos un 90% de identidad de secuencia, todavía más preferentemente con al menos un 91% de identidad de secuencia, todavía más preferentemente con al menos un 92% de identidad de secuencia, todavía más preferentemente con al menos un 93% de identidad de secuencia, todavía más preferentemente con al menos un 94% de identidad de secuencia, todavía más preferentemente con al menos un 95% de identidad de secuencia, todavía más preferentemente con al menos un 96% de identidad de secuencia, todavía más preferentemente con al menos un 97% de identidad de secuencia, todavía más preferentemente con al menos un 98% de identidad de secuencia, todavía más preferentemente con al menos un 99% de identidad de secuencia con (a) una molécula de ADN que codifica el mismo polipéptido maduro codificado por el cADN de la proteína humana depositado en el ATCC tal como se describe aquí, o (b) el complemento de la molécula de ADN de (a).

55 La invención puede utilizar una molécula de ácido nucleico aislado, que comprende una secuencia nucleotídica que codifica un polipéptido PRO5775, el cual es un dominio transmembrana delecionado o un dominio transmembrana inactivado, o es complementario con la mencionada secuencia nucleotídica codificante, en donde el dominio(s) transmembrana de dicho polipéptido se describe aquí. Por tanto, se contemplan en la presente invención los dominios extracelulares solubles de los polipéptidos PRO5775 aquí descritos.

60 Los fragmentos de una secuencia codificante del polipéptido PRO5775, o de la complementaria pueden hallar una utilización como, por ejemplo, sondas de hibridación, para los fragmentos codificantes de un polipéptido PRO5775 que opcionalmente codifican un polipéptido que comprende un sitio de unión para un anticuerpo anti-PRO5775 o sondas oligonucleotídicas anti-sentido. Dichos fragmentos de ácido nucleico son generalmente de al menos unos 20 nucleótidos de longitud, preferentemente de al menos unos 30 nucleótidos de longitud, más preferentemente de al menos unos 40 nucleótidos de longitud, todavía más preferentemente de al menos unos 50 nucleótidos de longitud, todavía más preferentemente de al menos unos 60 nucleótidos de longitud, todavía más preferentemente de al menos

unos 70 nucleótidos de longitud, todavía más preferentemente de al menos unos 80 nucleótidos de longitud, todavía más preferentemente de al menos unos 90 nucleótidos de longitud, todavía más preferentemente de al menos unos 100 nucleótidos de longitud, todavía más preferentemente de al menos unos 110 nucleótidos de longitud, todavía más preferentemente de al menos unos 120 nucleótidos de longitud, todavía más preferentemente de al menos unos 130 nucleótidos de longitud, todavía más preferentemente de al menos unos 140 nucleótidos de longitud, todavía más preferentemente de al menos unos 150 nucleótidos de longitud, todavía más preferentemente de al menos unos 160 nucleótidos de longitud, todavía más preferentemente de al menos unos 170 nucleótidos de longitud, todavía más preferentemente de al menos unos 180 nucleótidos de longitud, todavía más preferentemente de al menos unos 190 nucleótidos de longitud, todavía más preferentemente de al menos unos 200 nucleótidos de longitud, todavía más preferentemente de al menos unos 250 nucleótidos de longitud, todavía más preferentemente de al menos unos 300 nucleótidos de longitud, todavía más preferentemente de al menos unos 350 nucleótidos de longitud, todavía más preferentemente de al menos unos 400 nucleótidos de longitud, todavía más preferentemente de al menos unos 450 nucleótidos de longitud, todavía más preferentemente de al menos unos 500 nucleótidos de longitud, todavía más preferentemente de al menos unos 600 nucleótidos de longitud, todavía más preferentemente de al menos unos 700 nucleótidos de longitud, todavía más preferentemente de al menos unos 800 nucleótidos de longitud, todavía más preferentemente de al menos unos 900 nucleótidos de longitud, todavía más preferentemente de al menos unos 1000 nucleótidos de longitud, en donde en este contexto el término “aproximadamente” significa que la longitud de la secuencia nucleotídica es más o menos el 10% de la referenciada. Los fragmentos de una secuencia nucleotídica que codifica el polipéptido PRO5775 pueden determinarse de forma rutinaria mediante alineamiento de la secuencia nucleotídica que codifica el polipéptido Pro5775 con otras secuencias nucleotídicas mediante la utilización de cualquiera de los programas de alineamiento conocidos y la determinación de qué fragmentos de secuencia nucleotídica codificante del polipéptido PRO5775 son nuevos.

En otra realización, la invención hace referencia a un polipéptido PRO5775 aislado codificado por cualquiera de las secuencias de ácido nucleico aisladas identificadas más arriba.

En un determinado aspecto, la invención hace referencia a un polipéptido PRO5775 aislado, que comprende una secuencia aminoacídica que tiene al menos un 80% de identidad de secuencia, preferentemente con al menos un 81% de identidad de secuencia, más preferentemente con al menos un 82% de identidad de secuencia, todavía más preferentemente con al menos un 83% de identidad de secuencia, todavía más preferentemente con al menos un 84% de identidad de secuencia, todavía más preferentemente con al menos un 85% de identidad de secuencia, todavía más preferentemente con al menos un 86% de identidad de secuencia, todavía más preferentemente con al menos un 87% de identidad de secuencia, todavía más preferentemente con al menos un 88% de identidad de secuencia, todavía más preferentemente con al menos un 89% de identidad de secuencia, todavía más preferentemente con al menos un 90% de identidad de secuencia, todavía más preferentemente con al menos un 91% de identidad de secuencia, todavía más preferentemente con al menos un 92% de identidad de secuencia, todavía más preferentemente con al menos un 93% de identidad de secuencia, todavía más preferentemente con al menos un 94% de identidad de secuencia, todavía más preferentemente con al menos un 95% de identidad de secuencia, todavía más preferentemente con al menos un 96% de identidad de secuencia, todavía más preferentemente con al menos un 97% de identidad de secuencia, todavía más preferentemente con al menos un 98% de identidad de secuencia, todavía más preferentemente con al menos un 99% de identidad de secuencia con un polipéptido PRO5775 que tiene una secuencia aminoacídica de longitud completa tal como se describe aquí, una secuencia aminoacídica carente del péptido señal tal como se describe aquí, un dominio extracelular de una proteína transmembrana, con o sin el péptido señal, tal como se describe aquí o cualquier otro fragmento definido específicamente de secuencia aminoacídica de longitud completa tal como se describe aquí.

En otro aspecto, la invención hace referencia a un polipéptido PRO5775 aislado que comprende una secuencia aminoacídica que tiene al menos un 80% de identidad de secuencia, preferentemente con al menos un 81% de identidad de secuencia, más preferentemente con al menos un 82% de identidad de secuencia, todavía más preferentemente con al menos un 83% de identidad de secuencia, todavía más preferentemente con al menos un 84% de identidad de secuencia, todavía más preferentemente con al menos un 85% de identidad de secuencia, todavía más preferentemente con al menos un 86% de identidad de secuencia, todavía más preferentemente con al menos un 87% de identidad de secuencia, todavía más preferentemente con al menos un 88% de identidad de secuencia, todavía más preferentemente con al menos un 89% de identidad de secuencia, todavía más preferentemente con al menos un 90% de identidad de secuencia, todavía más preferentemente con al menos un 91% de identidad de secuencia, todavía más preferentemente con al menos un 92% de identidad de secuencia, todavía más preferentemente con al menos un 93% de identidad de secuencia, todavía más preferentemente con al menos un 94% de identidad de secuencia, todavía más preferentemente con al menos un 95% de identidad de secuencia, todavía más preferentemente con al menos un 96% de identidad de secuencia, todavía más preferentemente con al menos un 97% de identidad de secuencia, todavía más preferentemente con al menos un 98% de identidad de secuencia, todavía más preferentemente con al menos un 99% de identidad de secuencia con una secuencia aminoacídica codificada por cualquiera de los cADNs de proteínas humanas depositados en el ATCC tal como se describe aquí.

En otro aspecto, la invención hace referencia a un polipéptido PRO5775 aislado que comprende una secuencia aminoacídica que presenta una valoración de al menos un 80% de valores positivos, preferentemente al menos un 81% de positivos, más preferentemente al menos un 82% de positivos, todavía más preferentemente al menos un 83% de positivos, todavía más preferentemente al menos un 84% de positivos, todavía más preferentemente al menos un 85% de positivos, todavía más preferentemente al menos un 86% de positivos, todavía más preferentemente al menos un 87% de positivos, todavía más preferentemente al menos un 88% de positivos, todavía más preferentemente al menos un

89% de positivos, todavía más preferentemente al menos un 90% de positivos, todavía más preferentemente al menos un 91% de positivos, todavía más preferentemente al menos un 92% de positivos, todavía más preferentemente al menos un 93% de positivos, todavía más preferentemente al menos un 94% de positivos, todavía más preferentemente al menos un 95% de positivos, todavía más preferentemente al menos un 96% de positivos, todavía más preferentemente al menos un 97% de positivos, todavía más preferentemente al menos un 98% de positivos, todavía más preferentemente al menos un 99% de positivos cuando se compara la secuencia aminoacídica de un polipéptido PRO5775 que tiene una secuencia aminoacídica de tamaño completo tal como se describe aquí, una secuencia aminoacídica carente del péptido señal tal como se describe aquí, un dominio extracelular de una proteína transmembrana, con o sin un péptido señal, tal como se describe aquí o cualquier otro fragmento específicamente definido de la secuencia aminoacídica de tamaño completo tal como se describe aquí.

En un aspecto específico, la invención se refiere a un polipéptido PRO5775 aislado sin la secuencia señal N-terminal y/o la metionina del inicio, el cual está codificado por una secuencia nucleotídica que codifica dicha secuencia aminoacídica tal como se ha descrito más arriba. Los procesos para la producción del mismo también se describen aquí. Dichos procesos comprenden el cultivo de una célula huésped que comprende un vector el cual contiene la molécula de ácido nucleico codificante en las condiciones adecuadas para la expresión del polipéptido PRO5775 y la recuperación del polipéptido PRO5775 a partir del cultivo celular.

En otro aspecto de la invención se hace referencia a un polipéptido PRO5775 aislado que o bien tiene un dominio transmembrana deletado o un dominio transmembrana inactivado. Los procesos para la producción del mismo también se describen aquí, en donde dichos procesos comprenden el cultivo de una célula huésped que comprende un vector el cual contiene la molécula de ácido nucleico codificante en condiciones adecuadas para la expresión del polipéptido PRO5775 y la recuperación del polipéptido PRO5775 del cultivo celular.

En otra realización, la invención se refiere a una composición que comprende un anticuerpo anti-PRO5775 en combinación con un vehículo. Opcionalmente, el vehículo es un vehículo farmacéuticamente aceptable.

Se describen aquí los vectores que comprenden el ADN que codifica cualquiera de los polipéptidos aquí descritos. También se proporcionan las células huésped que comprenden dichos vectores. A modo de ejemplo, las células huésped pueden ser células CHO, células de *E. coli*, levaduras, o células de insecto infectadas con *Baculovirus*. Se describe además un proceso para la producción de cualquiera de los polipéptidos aquí descritos y comprende el cultivo de células huésped en condiciones adecuadas para la expresión del polipéptido deseado y la recuperación del polipéptido deseado del cultivo celular.

También se describen moléculas quiméricas que comprenden cualquiera de los polipéptidos descritos aquí fusionados a un polipéptido o secuencia aminoacídica heterólogos. Un ejemplo de dichas moléculas quiméricas comprende cualquiera de los polipéptidos aquí descritos fusionados con una secuencia etiquetada con un epitopo o una región Fc de una inmunoglobulina.

En otra realización, la invención proporciona un anticuerpo que se une específicamente con cualquiera de los polipéptidos descritos más arriba o más adelante. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un anticuerpo humanizado, un fragmento de anticuerpo o un anticuerpo de cadena sencilla.

En otra de las realizaciones, la invención utiliza sondas oligonucleotídicas de utilidad para el aislamiento genómico y de secuencias nucleotídicas de cADN o como sondas antisentido, en donde dichas sondas pueden derivarse de cualquiera de las secuencias nucleotídicas descritas más arriba o más adelante.

Descripción resumida de las figuras

La Figura 1 muestra la secuencia nucleotídica (SEC ID NO:1) de un cADN que contiene una secuencia nucleotídica que codifica una secuencia nativa PRO5775, en donde la secuencia nucleotídica (SEC ID NO:1) es un clon denominado aquí como ADN96869-2673. También se han representado en negrita y subrayadas las posiciones de los codones respectivos de inicio y de parada.

La Figura 2 muestra la secuencia aminoacídica (SEC ID NO:2) de un polipéptido PRO5775 de secuencia nativa derivado de la secuencia codificante SED ID NO:1 que se muestra en la Figura 1.

Descripción detallada de la invención

I. Definiciones

Los términos “amplificación génica” y “duplicación génica” se utilizan de modo intercambiable y hacen referencia a un proceso mediante el cual se forman copias múltiples de un gen o de un fragmento génico en una célula o línea celular determinada. La región duplicada (un fragmento del ADN amplificado) es referida a menudo como “amplicón”. Generalmente, la cantidad de ARN mensajero (mARN) producido, es decir, el nivel de expresión génica, también aumenta en proporción con el número de copias generadas del gen determinado expresado.

ES 2 279 473 T3

“Tumor”, tal como se utiliza aquí, hace referencia al crecimiento de células neoplásicas y a la proliferación, bien sea maligno o benigno, y a todas las células y tejidos precancerosos y cancerosos.

Los términos “cáncer” y “canceroso” hacen referencia a, o describen la condición fisiológica en los mamíferos que se caracteriza por un crecimiento celular no regulado. Ejemplos de cáncer incluyen, aunque no se limitan a éstos, el carcinoma, el linfoma, el blastoma, el sarcoma, y la leucemia. Ejemplos más particulares de dichos cánceres incluyen el cáncer de mama, de próstata, de colon, el cáncer de células escamosas, de células pequeñas de pulmón, de células no pequeñas de pulmón, el cáncer gastrointestinal, el pancreático, el glioblastoma, el cervical, el ovárico, el hepático, el renal, el hepatoma, el colorectal, el carcinoma endometrial, el carcinoma de glándulas salivares, el cáncer de riñón, el cáncer de de riñón, el cáncer de vulva, el de tiroides, el carcinoma hepático, y diversos tipos de cáncer de cabeza y cuello.

El término “tratamiento” es una intervención realizada con la intención de prevenir el desarrollo o la alteración de la patología de una enfermedad. Según ello, “tratamiento” hace referencia tanto al tratamiento terapéutico como al profiláctico o a sus medidas preventivas. Aquellos que necesiten un tratamiento incluyen los que ya han desarrollado el trastorno y en los que se prevé su aparición. En el tratamiento de un tumor (por ejemplo un cáncer), un agente terapéutico puede disminuir directamente la patología de las células tumorales, o convertir el tumor en más susceptible al tratamiento por otros agentes terapéuticos, por ejemplo, la radiación y/o la quimioterapia.

La “patología” del cáncer incluye todos los fenómenos que comprometen el bienestar del paciente. Ello incluye, sin limitación, el crecimiento anómalo o incontrolado de la célula, la metástasis, la interferencia con el funcionamiento normal de las células vecinas, la liberación de citocinas u otros productos secretores a niveles anómalos, la supresión o la agravación de la respuesta inflamatoria o inmunológica, etc.

El término “mamífero” con el propósito de tratamiento hace referencia a cualquier animal clasificado como un mamífero, incluyendo el hombre, los animales domésticos y de granja, del zoo, los animales deportivos o los animales de compañía, como los perros, caballos, gatos, vacas, cerdos, ovejas, etc. Preferentemente, el mamífero es el hombre.

“Vehículos” tal como se utiliza aquí, incluyen los vehículos farmacéuticamente aceptables, los excipientes, o los estabilizantes que no son tóxicos a la célula o al mamífero que se expone a dicho vehículo a las dosificaciones y concentraciones utilizadas. A menudo, el vehículo fisiológicamente aceptable es una solución acuosa de pH tamponado. Ejemplos de vehículos aceptables fisiológicamente incluyen tampones como los fosfatos, el citrato, y otros ácido orgánicos; los antioxidantes incluyendo el ácido ascórbico; polipéptidos de bajo peso molecular (inferiores a 10 residuos); proteínas, como la albúmina sérica, la gelatina o las inmunoglobulinas; polímeros hidrofílicos como la polivinilpirrolidona; aminoácidos como la glicina, la glutamina, la asparragina, la arginina o la lisina; monosacáridos, disacáridos y otros carbohidratos incluyendo la glucosa, la manosa, o las dextrinas; los agentes quelantes como el EDTA; los alcoholes de azúcar como el manitol o el sorbitol; los iones contrarrestadores formadores de sales como el sodio; y/o los tensoactivos no iónicos como el TWEENTM, el polietilén glicol (PEG) y el PLURONICSTM.

La administración “en combinación con” uno o más agentes terapéuticos incluyen la administración simultánea (a la vez) y consecutiva en cualquier orden.

El término “agente citotóxico” tal como se utiliza aquí hace referencia a una sustancia que inhibe o previene la función de las células y/o causa la destrucción de las células. El término pretende incluir isótopos radioactivos (por ejemplo, I¹³¹, I¹²⁵, Y⁹⁰ y Re¹⁸⁶), agentes quimioterapéuticos, y toxinas como las toxinas activas enzimáticamente derivadas de bacterias, hongos, plantas o animales o fragmentos de lo mismo.

Un “agente quimioterapéutico” es un compuesto químico útil en el tratamiento del cáncer. Ejemplos de agentes quimioterapéuticos incluyen la adriamicina, la doxorubicina, el 5-fluorouracilo, la citosina arabinósido (“Ara-C”), la ciclofosfamida, el tiotepa, el busulfán, el citoxín, los taxoides, por ejemplo, el paclitaxel (Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ), y el doxetaxel (Taxotere, Rhone-PoulencRorer, Antony, Francia), el toxotere, el metotrexato, la cisplatina, el melfalan, la vinblastina, la bleomicina, el etoposido, la ifosfamida, la mitomicina C, el mitoxantreno, la vincristina, la vinorelbina, la carboplatina, el tenipósido, la dauromicina, la carminomicina, la aminopterina, la dactinomicina, las mitomicinas, las esperamicinas (véase la patente americana 4.675.187), 5-FU, 6-tioguanina, 6-mercaptopurina, la actinomicina D, VP-16, el clorambucil, el melfalan y otras mostazas nitrogenadas. También se incluyen en esta definición, los agentes hormonales que actúan para regular o inhibir la acción de la hormona sobre los tumores como el tamoxifeno o la onapristona.

Un “agente inhibidor del crecimiento” cuando se utiliza aquí hace referencia a un compuesto o composición que inhibe el crecimiento de una célula, especialmente de células cancerosas que sobreexpresan cualquiera de los genes aquí identificados, bien *in vivo* o *in vitro*. Por tanto, un agente inhibidor del crecimiento es aquel que reduce el porcentaje de las células que sobreexpresan dichos genes en la fase S. Ejemplos de agentes inhibidores del crecimiento incluyen los agentes que bloquean la progresión del ciclo celular (en un momento distinto a la fase S), como los agentes que inducen un paro en fase G1 y fase M. Los bloqueantes típicos de fase M incluyen las vincas (vincristina y vinblastina), el taxol, y los inhibidores de la topo II como la doxorubicina, epirubicina, daunorubicina, etoposido y bleomicina. Estos agentes que bloquean en fase G1 también prolongan su acción provocando un paro en la fase S, por ejemplo, los agentes alquilantes del ADN como el tamoxifeno, la prednisona, la dacarbazina, la mecloretamina, la cisplatina, el metotrexato, el 5-fluorouracil, y la ara-C. Información adicional puede hallarse en The Molecular Basis

of Cancer, Mendelsohn e Israel, eds., Capítulo 1, titulado “Cell Cycle Regulation, oncogens and antineoplastic drugs” por Murakami y col., (WB Saunders: Philadelphia, 1995), especialmente en la página 13.

La “doxorubicina” es un antibiótico antraciclina. El nombre químico completo es la doxorubicina (8S-cis)-10-[(3-amino-2,3,6-tridesoxi- α -L-lixo-hexapiranosil)oxi]-7,8,9,10-tetrahidro-6,8,11-tihidroxi-8-(hidroxiacetil)-1-metoxi-5,12-naftacenediona.

El término “citocina” es un término genérico para las proteínas liberadas por una población celular que actúan sobre otra célula como mediadores intercelulares. Ejemplos de dichas citocinas son las linfocinas, las monocinas y las hormonas polipeptídicas tradicionales. Se incluyen entre las citocinas, las hormonas del crecimiento como la hormona del crecimiento humano, la hormona N-metionil del crecimiento humano y la hormona del crecimiento bovino; la hormona paratiroidea; la tiroxina; la insulina; la relaxina; la prorelaxina; las hormonas glicoproteicas como la hormona estimulante del folículo (HSF), la hormona estimulante del tiroides (HST), y la hormona luteinizante (HL); y el factor de crecimiento hepático, el factor de crecimiento de fibroblastos; la prolactina; el lactógeno placentar; el factor de crecimiento de fibroblastos; el factor- α y el β de necrosis tumoral; la sustancia inhibidora mülleriana; el péptido asociado a la gonadotropina de ratón; la inhibina; la activina; el factor endotelial vascular; la integrina; la trombopoyetina (TPO); los factores de crecimiento nervioso como el NGF- β ; el factor de crecimiento de plaquetas; los factores de crecimiento transformantes (TGFs) como el TGF- α y el TGF- β ; el factor de crecimiento de tipo insulina-I y -II; la eritropoyetina (EPO); los factores osteoinductivos; los interferones como el interferón α , β y γ , los factores estimuladores de colonias (CSFs) como el CSF de macrófagos (M-CSF); el CSF de granulocitos-macrófagos (GM-CSF); y el CSF de granulocitos (G-CSF); las interleucinas (ILs) como la IL-1, la IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; un factor de necrosis tumoral como el TNF- α o el TNF- β ; y otros factores polipeptídicos incluyendo el LIF y el ligando del equipo (KL). Tal como se utiliza aquí, el término citocina incluye proteínas de fuentes naturales o del cultivo de células recombinantes y los equivalentes biológicamente activos de las citocinas de secuencia nativa.

El término “profármaco”, tal como se utiliza aquí en esta aplicación, hace referencia a un precursor o derivado de una sustancia farmacéuticamente activa que es menos citotóxica para las células tumorales comparadas con el fármaco parental y es capaz de ser activado enzimáticamente o convertido en la forma parental más activa, véase, por ejemplo, Willman, Prodrugs in Cancer Chemotherapy, Biochemical Society Transactions, 14:375-382, 615th Meeting, Belfast (1986), y Stella y col., Prodrugs: A Chemical approach to targeted drug delivery, Directed Drug Delivery, Borchardt y col., (ed.) pp. 147-267, Humana Press (1985). Los profármacos de la presente invención incluyen, pero no se limitan a, profármacos que contienen fosfato, tiosfosfato, sulfato, o péptidos. Los pro-fármacos modificados por D-aminoácidos, los profármacos glicosilados, los que contienen β -lactama, opcionalmente los sustituidos con fenoxiacetamidas u opcionalmente fenilacetamidas, los profármacos 5-fluorocitosina y otros 5-fluorouridina que pueden convertirse en un fármaco libre más citotóxico y activo. Ejemplos de fármacos citotóxicos que pueden derivarse en profármacos para su utilización en la presente invención, pero no se limitan a ellos, son los agentes quimioterapéuticos descritos anteriormente.

Una “cantidad efectiva” de un polipéptido descrito aquí o un antagonista de lo mismo, respecto al crecimiento de células neoplásicas, el crecimiento tumoral o el crecimiento de una célula cancerosa, es una cantidad capaz de inhibir, en cierta magnitud, el crecimiento de células diana. El término incluye una cantidad capaz de invocar un efecto inhibidor, citostático y/o citotóxico y/o apoptótico del crecimiento de las células diana. Una “cantidad efectiva” de un antagonista del polipéptido PRO para los propósitos de inhibir el crecimiento de células neoplásicas, el crecimiento del tumor o el crecimiento de células cancerosas, puede determinarse empíricamente y de forma rutinaria.

Una “cantidad terapéuticamente efectiva”, respecto al tratamiento del tumor, se refiere a una cantidad capaz de provocar uno o más de los efectos siguientes: (1) inhibición, en cierta magnitud, del crecimiento tumoral, incluyendo, la ralentización y el completo paro del crecimiento; (2) la reducción en el número de células tumorales; (3) la reducción en el tamaño del tumor; (4) la inhibición (es decir, la reducción, la ralentización o el completo paro) de la infiltración de células tumorales en los órganos periféricos; (5) la inhibición (es decir, la reducción, la ralentización o el completo paro del crecimiento) de metástasis; (6) aumento de la respuesta inmune anti-tumoral, lo cual puede, aunque no es requisito, dar como resultado la regresión o el rechazo del tumor; y/o (7) alivio, en cierta magnitud, de uno o más síntomas asociados con el trastorno. Una “cantidad terapéuticamente efectiva” de un antagonista del polipéptido PRO5775 para propósitos de tratamiento del tumor puede determinarse empíricamente y de forma rutinaria.

Una “cantidad inhibidora del crecimiento” de un antagonista de PRO es una cantidad capaz de inhibir el crecimiento de una célula, por ejemplo, célula cancerosa, bien *in vivo* o *in vitro*. Una “cantidad inhibidora del crecimiento” de un antagonista de PRO con el propósito de inhibir el crecimiento de una célula neoplásica puede determinarse empíricamente y de forma rutinaria.

Una “cantidad citotóxica” de un antagonista de PRO es una cantidad capaz de causar la destrucción de una célula, especialmente tumoral, por ejemplo las células cancerosas, ya sea *in vitro* o *in vivo*. Una “cantidad citotóxica” de un antagonista de PRO con el propósito de inhibir el crecimiento de una célula neoplásica puede determinarse empíricamente y de forma rutinaria.

Los términos “polipéptido PRO” y “PRO” tal como se utilizan aquí y cuando se prosiguen inmediatamente por una denominación numérica hacen referencia a diversos polipéptidos, en donde la denominación completa (es decir,

PRO/número) hace referencia a secuencias polipeptídicas específicas tal como se describe aquí. Los términos “polipéptido PRO/número” y “PRO/número” en donde el término “número”, que se proporciona como una denominación numérica real tal como se utiliza aquí, abarca los polipéptidos de secuencia nativa y las variantes polipeptídicas (que se describen más adelante en la presente invención). Los polipéptidos PRO descritos aquí pueden aislarse de diversas fuentes, tales como tejidos humanos varios o de otras fuentes, o prepararse mediante procedimientos recombinantes o sintéticos.

Un “polipéptido PRO de secuencia nativa” comprende un polipéptido que tiene la misma secuencia aminoacídica que el correspondiente polipéptido PRO derivado de la naturaleza. Dichos polipéptidos PRO de secuencia nativa pueden aislarse de la naturaleza o pueden producirse mediante medios recombinantes o sintéticos. El término “polipéptido PRO de secuencia nativa” abarca específicamente las formas truncadas o secretadas de la naturaleza del polipéptido PRO específico (por ejemplo, una secuencia de dominio extracelular), formas variantes de la naturaleza (por ejemplo, formas ajustadas alternativamente) y variantes alélicas que ocurren de forma natural. En varias realizaciones de la invención, los polipéptidos PRO de secuencia nativa descritos aquí son polipéptidos de secuencia madura o nativa de longitud completa que comprenden las secuencias aminoacídicas de longitud completa que se muestran en las figuras acompañantes. Los codones de inicio y parada se muestran en negrita y en subrayado en las figuras. Sin embargo, mientras que el polipéptido PRO descrito en las figuras acompañantes se inicia en los residuos metionina denominados aquí como aminoácido en la posición 1 en las figuras, es posible que otros residuos metionina localizados cadena arriba o cadena abajo de la posición 1 aminoacídica en las figuras puedan utilizarse como el residuos aminoacídico de los polipéptidos PRO.

El “dominio extracelular” del polipéptido PRO o “variante extracelular” o “ECD” hace referencia a una forma del polipéptido PRO que carece de los dominios transmembrana y citoplasmático. Generalmente, un ECD del polipéptido PRO tendrá menos del 1% de dichos dominios transmembrana y/o citoplasmático y preferentemente, tendrá menos del 0,5% de dichos dominios. Se entenderá que cualquier dominio transmembrana identificado para los polipéptidos PRO de la presente invención se identifican siguiendo los criterios utilizados habitualmente en la materia para la identificación del tipo de dominio hidrofóbico. Los límites exactos de un dominio transmembrana pueden variar, pero en no más de 5 aminoácidos por cada extremo del dominio tal como inicialmente se identificó aquí. Opcionalmente, un dominio extracelular de un polipéptido PRO puede contener desde aproximadamente 5 o menos aminoácidos en cada extremo de los límites del dominio transmembrana/dominio extracelular tal como se identifica en los Ejemplos o especificaciones y dichos polipéptidos, con o sin el péptido señal asociado, y el ácido nucleico que los codifica, están contemplados por la presente invención.

La localización aproximada del “péptido señal” de diversos polipéptidos PRO descritos aquí se muestran en la presente especificación y/o las figuras acompañantes. Se ha de remarcar, sin embargo, que el límite C-terminal de un péptido señal puede variar, pero en no más de 5 aminoácidos en cualquiera de los extremos del C-terminal del péptido señal tal como se identificó inicialmente aquí, en donde el límite C-terminal del péptido señal puede identificarse según un criterio utilizado de forma rutinaria en la materia para identificar el tipo de elemento de secuencia aminoacídica (por ejemplo, Nielsen y col., Prot Eng, 10:1-6 (1997) y von Heinje y col., Nucl Acids Res, 14:4683-4690 (1986)). Además, también se ha reconocido, que en algunos casos, el corte de una secuencia señal de un polipéptido secretado no es totalmente uniforme, dando como resultado más de una especie secretada. Estos polipéptidos maduros, cuando su péptido señal se corta en no más de 5 aminoácidos por cualquier lado del límite C-terminal del péptido señal identificado aquí, y los polinucleótidos que los codifican, están contemplados por la presente invención.

“Variante de polipéptido PRO” significa un polipéptido PRO activo tal como se define más arriba o más adelante que tiene por lo menos el 80% de identidad de secuencia con una secuencia nativa de longitud completa tal como se describe aquí, una secuencia de polipéptido PRO carente de la secuencia señal tal como se describe aquí, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, tal como se describe aquí, un dominio extracelular de un polipéptido PRO tal como se describe aquí. Dichas variantes de polipéptido PRO incluyen, por ejemplo, los polipéptidos PRO en donde uno o más residuos aminoacídicos se añaden, o delecionan, en el extremo N- o C-terminal de la secuencia aminoacídica nativa de longitud completa. Generalmente, una variante de polipéptido PRO tendrá al menos un 80% de identidad de secuencia aminoacídica, preferentemente al menos un 81% de identidad de secuencia aminoacídica, más preferentemente al menos un 82% de identidad de secuencia aminoacídica, más preferentemente al menos un 83% de identidad de secuencia aminoacídica, más preferentemente al menos un 84% de identidad de secuencia aminoacídica, más preferentemente al menos un 85% de identidad de secuencia aminoacídica, más preferentemente al menos un 86% de identidad de secuencia aminoacídica, más preferentemente al menos un 87% de identidad de secuencia aminoacídica, más preferentemente al menos un 88% de identidad de secuencia aminoacídica, más preferentemente al menos un 89% de identidad de secuencia aminoacídica, más preferentemente al menos un 90% de identidad de secuencia aminoacídica, más preferentemente al menos un 91% de identidad de secuencia aminoacídica, más preferentemente al menos un 92% de identidad de secuencia aminoacídica, más preferentemente al menos un 93% de identidad de secuencia aminoacídica, más preferentemente al menos un 94% de identidad de secuencia aminoacídica, más preferentemente al menos un 95% de identidad de secuencia aminoacídica, más preferentemente al menos un 96% de identidad de secuencia aminoacídica, más preferentemente al menos un 97% de identidad de secuencia aminoacídica, más preferentemente al menos un 98% de identidad de secuencia aminoacídica, más preferentemente al menos un 99% de identidad de secuencia aminoacídica con una secuencia de un polipéptido PRO de secuencia nativa de longitud completa tal como se describe aquí, una secuencia de polipéptido PRO carente de un péptido señal tal como se describe aquí, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, tal como se describe aquí o cualquier otro fragmento definido específicamente de una secuencia de polipéptido PRO

ES 2 279 473 T3

de longitud completa tal como se describe aquí. Generalmente, los polipéptidos variantes de PRO tienen una longitud de al menos 10 aminoácidos, a menudo de por lo menos 20 aminoácidos de longitud, más a menudo de por lo menos 30 aminoácidos de longitud, más a menudo de por lo menos 40 aminoácidos de longitud, más a menudo de por lo menos 50 aminoácidos de longitud, más a menudo de por lo menos 60 aminoácidos de longitud, más a menudo de por lo menos 70 aminoácidos de longitud, más a menudo de por lo menos 80 aminoácidos de longitud, más a menudo de por lo menos 90 aminoácidos de longitud, más a menudo de por lo menos 100 aminoácidos de longitud, más a menudo de por lo menos 150 aminoácidos de longitud, más a menudo de por lo menos 200 aminoácidos de longitud, más a menudo de por lo menos 300 aminoácidos de longitud, o más.

Tal como se muestra más adelante, la Tabla 1 proporciona el código de fuente completo para el programa informático de comparación de secuencias por ALIGN-2. Este código de fuente puede ser recopilado para utilizarse en un sistema operativo UNIX para proporcionar el programa informático de comparación de secuencias ALIGN-2.

Además, las tablas 2A-2D muestran las ejemplificaciones hipotéticas para utilizar el procedimiento descrito más abajo para la determinación del % de identidad de secuencia aminoacídica (tablas 2A-2B) y el % de identidad de secuencia de ácido nucleico (Tablas 2C-2D) mediante la utilización del programa informático de comparación de secuencias ALIGN-2, en donde "PRO" representa la secuencia aminoacídica de un polipéptido hipotético PRO de interés. "Proteína de comparación" representa la secuencia aminoacídica de un polipéptido con la que se compara el polipéptido "PRO" de interés. "PRO-ADN" representa una secuencia hipotética de ácido nucleico codificante de PRO de interés. "ADN de comparación" representa la secuencia nucleotídica de una molécula de ácido nucleico con la que se compara la molécula de ácido nucleico "PRO-ADN" de interés, las letras "X", "Y" y "Z" representan cada una los residuos aminoacídicos hipotéticos distintos y "N", "L" y "V" los nucleótidos hipotéticos distintos.

TABLA 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int
_day[26][26] = {
/*  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ {_M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

ES 2 279 473 T3

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4      /* save if there's at least MX-1 bases since last jmp */

#define DMAT        3      /* value of matching bases */
#define DMIS        0      /* penalty for mismatched bases */
#define DINSO       8      /* penalty for a gap */
#define DINSI       1      /* penalty per base */
#define PINSO       8      /* penalty for a gap */
#define PINSI       4      /* penalty per residue */

struct jmp {
    short          n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 -1 */

struct diag {
    int            score;     /* score at last jmp */
    long           offset;    /* offset of prev block */
    short          jmp;      /* current jmp index */
    struct jmp     jp;        /* list of jmps */
};

struct path {
    int            spc;       /* number of leading spaces */
    short          n[JMPS]; /* size of jmp (gap) */
    int            x[JMPS]; /* loc of jmp (last elem before gap) */
};

char             *ofile;     /* output file name */
char             *namex[2]; /* seq names: getseqs() */
char             *prog;     /* prog name for err msgs */
char             *seqx[2];  /* seqs: getseqs() */
int              dmax;      /* best diag: nw() */
int              dmax0;     /* final diag */
int              dna;       /* set if dna: main() */
int              endgaps;   /* set if penalizing end gaps */
int              gapx, gapy; /* total gaps in seqs */
int              len0, len1; /* seq lens */
int              ngapx, ngapy; /* total size of gaps */
int              smax;      /* max score: nw() */
int              *xbm;      /* bitmap for matching */
long             offset;    /* current offset in jmp file */
struct           diag      *dx; /* holds diagonals */
struct           path      pp[2]; /* holds path for seqs */

char             *calloc(), *malloc(), *index(), *strcpy();
char             *getseq(), *g_calloc();

```

ES 2 279 473 T3

```

5  /* Needleman-Wunsch alignment program
   *
   * usage: progs file1 file2
   * where file1 and file2 are two dna or two protein sequences.
   * The sequences can be in upper- or lower-case and may contain ambiguity
   * Any lines beginning with ';', '>' or '<' are ignored
   * Max file length is 65535 (limited by unsigned short x in the jmp struct)
10  * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
   * Output is in the file "align.out"
   *
   * The program may create a tmp file in /tmp to hold info about traceback.
   * Original version developed under BSD 4.3 on a vax 8650
   */
15  #include "nw.h"
   #include "day.h"

   static  _dbval[26] = {
20         1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
   };

   static  _pbval[26] = {
25         1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
         128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
         1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
         1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
   };

   main(ac, av)                                main
30         int      ac;
           char    *av[];
   {
           prog = av[0];
           if (ac != 3) {
35                 fprintf(stderr, "usage: %s file1 file2\n", prog);
                 fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
                 fprintf(stderr, "The sequences can be in upper- or lower-case\n");
                 fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
                 fprintf(stderr, "Output is in the file \"align.out\"\n");
                 exit(1);
           }
40           namex[0] = av[1];
           namex[1] = av[2];
           seqx[0] = getseq(namex[0], &len0);
           seqx[1] = getseq(namex[1], &len1);
           xbm = (dna)? _dbval : _pbval;

45           endgaps = 0;                        /* 1 to penalize endgaps */
           ofile = "align.out";                 /* output file */

           nw();                                /* fill in the matrix, get the possible jmps */
           readjmps();                           /* get the actual jmps */
50           print();                             /* print stats, alignment */

           cleanup();                            /* unlink any tmp files */
   }
55
60
65

```

ES 2 279 473 T3

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()
{
    char          *px, *py;          /* seqs and ptrs */
    int           *ndely, *dely;     /* keep track of dely */
    int           ndelx, delx;       /* keep track of delx */
    int           *tmp;              /* for swapping row0, row1 */
    int           mis;               /* score for each type */
    int           ins0, ins1;        /* insertion penalties */
    register      id;                /* diagonal index */
    register      ij;                /* jmp index */
    register      *col0, *col1;      /* score for curr, last row */
    register      xx, yy;            /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

...NW

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
  mis = col0[yy-1];
  if (dna)
    mis += (xbrn[*px-'A']&xbrn[*py-'A'])? DMAT : DMIS;
  else
    mis += _day[*px-'A'][*py-'A'];

  /* update penalty for del in x seq;
   * favor new del over ongong del
   * ignore MAXGAP if weighting endgaps
   */
  if (endgaps || ndely[yy] < MAXGAP) {
    if (col0[yy] - ins0 >= dely[yy]) {
      dely[yy] = col0[yy] - (ins0+ins1);
      ndely[yy] = 1;
    } else {
      dely[yy] -= ins1;
      ndely[yy]++;
    }
  } else {
    if (col0[yy] - (ins0+ins1) >= dely[yy]) {
      dely[yy] = col0[yy] - (ins0+ins1);
      ndely[yy] = 1;
    } else
      ndely[yy]++;
  }

  /* update penalty for del in y seq;
   * favor new del over ongong del
   */
  if (endgaps || ndelx < MAXGAP) {
    if (col1[yy-1] - ins0 >= delx) {
      delx = col1[yy-1] - (ins0+ins1);
      ndelx = 1;
    } else {
      delx -= ins1;
      ndelx++;
    }
  } else {
    if (col1[yy-1] - (ins0+ins1) >= delx) {
      delx = col1[yy-1] - (ins0+ins1);
      ndelx = 1;
    } else
      ndelx++;
  }

  /* pick the maximum score; we're favoring
   * mis over any del and delx over dely
   */

```

50

55

60

65

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
5   coll[yy] = mis;
else if (delx >= dely[yy]) {
    coll[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10   && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
15   }
        }
    dx[id].jp.n[ij] = ndelx;
    dx[id].jp.x[ij] = xx;
    dx[id].score = delx;
}
20   else {
        coll[yy] = dely[yy];
        ij = dx[id].ijmp;
if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
25   && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
30   }
        }
    dx[id].jp.n[ij] = -ndely[yy];
    dx[id].jp.x[ij] = xx;
    dx[id].score = dely[yy];
}
35   if (xx == len0 && yy < len1) {
        /* last col
        */
        if (endgaps)
            coll[yy] -= ins0+ins1*(len1-yy);
        if (coll[yy] > smax) {
            smax = coll[yy];
            dmax = id;
40   }
        }
    }
}
if (endgaps && xx < len0)
45   coll[yy-1] -= ins0+ins1*(len0-xx);
if (coll[yy-1] > smax) {
    smax = coll[yy-1];
    dmax = id;
}
50   tmp = col0; col0 = coll; coll = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)coll);
55   }
}

```

60

65

```

/*
 *
5  * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
10 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
*/

15 #include "nw.h"

#define SPC      3
#define P_LINE  256 /* maximum output line */
#define P_SPC   3 /* space between name or num and seq */

20 extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

25 print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

55
60
65

```

print

ES 2 279 473 T3

```

/*
 * trace back the best path, count matches
 */
static
5 getmat(lx, ly, firstgap, lastgap)                                getmat
    int    lx, ly;                /* "core" (minus endgaps) */
    int    firstgap, lastgap;     /* leading trailing overlap */
{
    int          nm, i0, i1, siz0, siz1;
    char         outx[32];
    double       pct;
    register     n0, n1;
    register char *p0, *p1;

15     /* get total matches, score
        */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

20     nm = 0;
        while ( *p0 && *p1 ) {
            if (siz0) {
25                 p1++;
                    n1++;
                    siz0--;
            }
            else if (siz1) {
30                 p0++;
                    n0++;
                    siz1--;
            }
            else {
35                 if (xbm[*p0-'A']&xbm[*p1-'A'])
                            nm++;
                    if (n0++ == pp[0].x[i0])
                            siz0 = pp[0].n[i0++];
                    if (n1++ == pp[1].x[i1])
                            siz1 = pp[1].n[i1++];
                    p0++;
                    p1++;
40                 }
            }
        }

45     /* pct homology:
        * if penalizing endgaps, base is the shorter seq
        * else, knock off overhangs and take shorter core
        */
        if (endgaps)
            lx = (len0 < len1)? len0 : len1;
        else
            lx = (lx < ly)? lx : ly;
50     pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "\n");
        fprintf(fx, "< %d match%$ in an overlap of %d: %.2f percent similarity\n",
            nm, (nm == 1)? "" : "es", lx, pct);
55
60
65

```

```

5      fprintf(fx, "< gaps in first sequence: %d", gapx);
      if (gapx) {
          (void) sprintf(outx, "(%d %s%s)",
              ngapx, (dna)? "base":"residue", (ngapx == 1)? "" : "s");
          fprintf(fx, "%s", outx);
10     fprintf(fx, ", gaps in second sequence: %d", gapy);
      if (gapy) {
          (void) sprintf(outx, "(%d %s%s)",
              ngapy, (dna)? "base":"residue", (ngapy == 1)? "" : "s");
          fprintf(fx, "%s", outx);
15     }
      if (dna)
          fprintf(fx,
              "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
              smax, DMAT, DMIS, DINSO, DINS1);
      else
20     fprintf(fx,
              "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
              smax, PINSO, PINS1);
      if (endgaps)
          fprintf(fx,
25     "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
              firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
              lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
      else
          fprintf(fx, "< endgaps not penalized\n");
    }
30     static      nm;          /* matches in core -- for checking */
    static      lmax;        /* lengths of stripped file names */
    static      ij[2];       /* jmp index for a path */
    static      nc[2];       /* number at start of current line */
    static      ni[2];       /* current elem number -- for gapping */
35     static      siz[2];
    static char *ps[2];      /* ptr to current element */
    static char *po[2];      /* ptr to next output char slot */
    static char out[2][P_LINE]; /* output line */
    static char star[P_LINE]; /* set by stars() */
40     /*
     * print alignment of described in struct path pp[]
     */
    static
45     pr_align0
    {
        int      nn;          /* char count */
        int      more;
        register i;

        for (i = 0, lmax = 0; i < 2; i++) {
50             nn = stripname(namex[i]);
            if (nn > lmax)
                lmax = nn;

            nc[i] = 1;
55             ni[i] = 1;
            siz[i] = ij[i] = 0;
            ps[i] = seqx[i];
            po[i] = out[i];
        }
60
65

```

...getmat

pr_align

```

5      for (nn = nm = 0, more = 1; more; ) {
        for (i = more = 0; i < 2; i++) {
            /*
            * do we have more of this sequence?
            */
            if (!*ps[i])
                continue;

10         more++;

            if (pp[i].spc) { /* leading space */
                *po[i]++ = ' ';
                pp[i].spc--;
            }
            else if (siz[i]) { /* in a gap */
                *po[i]++ = '.';
                siz[i]--;
            }
            else { /* we're putting a seq element
            */
                *po[i] = *ps[i];
                if (istlower(*ps[i]))
                    *ps[i] = toupper(*ps[i]);
                po[i]++;
                ps[i]++;

25         /*
            * are we at next gap for this seq?
            */
            if (ni[i] == pp[i].x[ij(i)]) {
                /*
                * we need to merge all gaps
                * at this location
                */
                siz[i] = pp[i].n[ij(i) + +];
                while (ni[i] == pp[i].x[ij(i)])
                    siz[i] += pp[i].n[ij(i) + +];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nm = 0;
    }
}

45 /*
    * dump a block of lines, including numbers, stars: pr_align()
    */
    static
    dumpblock()
    {
        register i;

        for (i = 0; i < 2; i++)
            *po[i]-- = '\0';
55
60
65

```

...pr_align

dumpblock

...dumpblock

```

5      (void)putc('\n', fx);
      for (i = 0; i < 2; i++) {
          IF (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
              if (i == 0)
                  nums(i);
              if (i == 0 && *out[1])
                  stars();
10         putline(i);
              if (i == 0 && *out[1])
                  fprintf(fx, star);
              if (i == 1)
                  nums(i);
          }
      }
15  }
}

/*
 * put out a number line: dumpblock()
 */
static
20  nums(ix)                                nums
    int    ix;    /* index in out[] holding seq line */
    {
        char    nline[P_LINE];
        register    l, j;
        register char    *pn, *px, *py;

        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
            *pn = ' ';
30     for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
            if (*py == ' ' || *py == '-')
                *pn = ' ';
            else {
                if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                    j = (i < 0)? -i : i;
                    for (px = pn; j /= 10, px--)
                        *px = j%10 + '0';
                    if (i < 0)
                        *px = '-';
35                 }
                else
                    *pn = ' ';
                i++;
            }
        }
        *pn = '\0';
        nc[ix] = i;
        for (pn = nline; *pn; pn++)
            (void)putc(*pn, fx);
        (void)putc('\n', fx);
45  }

/*
 * put out a line (name, {num}, seq, {num}): dumpblock()
 */
static
50  putline(ix)                                putline
    int    ix;
    {
55  {
60
65

```

...putline

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
5   (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
   (void) putc(' ', fx);

/* these count from 1:
   * ni[] is current element (from 1)
   * nc[] is number at start of current line
   */
for (px = out[ix]; *px; px++)
10   (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
15 }

```

```

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
20 {
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
30     *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
35                 cx = '*';
                    nm++;
            }
            else if (ldna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
40             else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
45     }
    *px++ = '\n';
    *px = '\0';
}
/*

```

stars

```

/* strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
50 char *pn; /* file name (may be path) */
{
    register char *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
60         py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
65 }

```

stripname

ES 2 279 473 T3

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
5  * g_malloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
10 #include <sys/file.h>

char *jname = "/tmp/homgXXXXXXXXX";      /* tmp file for jumps */
FILE *fj;

int cleanup();                          /* cleanup tmp file */
15 long lseek();

/*
 * remove any tmp file if we blow
 */
20 cleanup(i)
    int i;
    {
        if (fj)
            (void) unlink(jname);
        exit(i);
    }
25

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
30 char *
getseq(file, len)
    char *file; /* file name */
    int *len; /* seq len */
35 {
    char line[1024], *pseq;
    register char *px, *py;
    int natgc, den;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
55
60
65

```

```

py = pseq + 4;
*ten = den;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
    *py++ = '\0';
    *py = '\0';
    (void) fclose(fp);
    dna = natgc > (ten/3);
    return(pseq + 4);
}

char *
g_alloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;         /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;
        }
    }
}

```

...getseq

g_alloc

readjmps

...readjumps

```

5         if (j < 0 && dx[dmax].offset && fj) {
            (void) lseek(fd, dx[dmax].offset, 0);
            (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
            (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
            dx[dmax].ijmp = MAXJMP-1;
        }
        else
            break;
10    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
    if (j >= 0) {
15        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
20            pp[1].n[i1] = -siz;
            xx += siz;

            /* id = xx - yy + len1 - 1
             */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
        }
        else if (siz > 0) { /* gap in first seq */
30            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
        }
        }
        else
            break;
40    }

    /* reverse the order of jumps
     */
    for (j = 0, i0--; j < i0; j++, i0--) {
45        i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
        i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
    }
    for (j = 0, i1--; j < i1; j++, i1--) {
50        i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
        i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
    }
    if (fd >= 0)
        (void) close(fd);
    if (fj) {
55        (void) unlink(jname);
        fj = 0;
        offset = 0;
    }
}

```

60

65

ES 2 279 473 T3

```

5  /*
   * write a filled jmp struct offset of the prev one (if any): nw()
   */
writejumps(ix)                                writejumps
{
    int    ix;
10     char  *mktemp();

    if (!f) {
        if (mktemp(jname) < 0) {
15             fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
                fprintf(stderr, "%s: can't write %s\n", prog, jname);
                exit(1);
        }
20     }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

25

TABLA 2A

30 PRO XXXXXXXXXXXXXXXXX (longitud = 15 aminoácidos)
 Proteína de comparación XXXXXYYYYYYY (Longitud = 12
 aminoácidos)
 35 % identidad de secuencia aminoacídica =
 (número de residuos aminoacídicos completamente
 concordantes entre las dos secuencias polipeptídicas tal
 como se determina por ALIGN-2) dividido por (número total
 40 de residuos aminoacídicos del polipéptidos PRO) =
 5 dividido por 15 = 33,3%

45

TABLA 2B

50 PRO XXXXXXXXXXXX (longitud = 10 aminoácidos)
 Proteína de comparación XXXXXYYYYYYZZYZ (longitud = 15
 aminoácidos)
 55 % identidad de secuencia aminoacídica =
 (número de residuos aminoacídicos completamente
 concordantes entre las dos secuencias polipeptídicas tal
 como se determina por ALIGN-2) dividido por (número total
 60 de residuos aminoacídicos del polipéptidos PRO) =
 5 dividido por 10 = 50%

65

ES 2 279 473 T3

TABLA 2C

5 PRO-ADN NNNNNNNNNNNNNN (longitud = 14 nucleótidos)
ADN de comparación NNNNNNLLLLLLLLLLLL (longitud = 16
nucleótidos)
10 % identidad de secuencia de ácidos nucleicos =
(número de nucleótidos completamente concordantes entre
las dos secuencias de ácidos nucleicos tal como se
15 determina por ALIGN-2) dividido por (número total de
nucleótidos del ácido nucleico PRO-ADN) =
6 dividido por 14 = 42,9%

TABLA 2D

25 PRO-ADN NNNNNNNNNNNNNN (longitud = 12 nucleótidos)
ADN de comparación NNNNLLLVV (longitud = 9 nucleótidos)
% identidad de secuencia de ácido nucleico =
30 (número de nucleótidos completamente concordantes entre
las dos secuencias ácidos nucleicos tal como se determina
por ALIGN-2) dividido por (número total de nucleótidos del
35 ácido nucleico PRO-ADN) =
4 dividido por 12 = 33,3%

40 Porcentaje (%) de "identidad" de secuencia aminoacídica" respecto a las secuencias polipéptidicas PRO identifica-
das aquí se define como el porcentaje de residuos aminoacídicos en una secuencia candidata que son idénticos con
los residuos de una secuencia PRO, después del alineamiento de las secuencias y la introducción de huecos, si fuera
necesario, para lograr el máximo porcentaje de identidad de secuencia, y sin considerar ninguna de las sustituciones
45 conservadoras como parte de la identidad de secuencia. El alineamiento con el propósito de determinar el porcentaje
de identidad de secuencia aminoacídica puede lograrse de diversas maneras bien conocidas por los expertos en la
materia, por ejemplo, mediante la utilización de un programa informático de disponibilidad pública como el BLAST,
BLAST-2, ALIGN, ALIGN-2, o el programa Megalign (ADNSTAR). Los expertos en la materia pueden determinar
los parámetros adecuados para cuantificar el alineamiento incluyendo los algoritmos que se requieran para lograr el
50 máximo alineamiento a lo largo de las secuencias de longitud completa que se compararan. Para los propósitos de la
presente invención, los valores de identidad de secuencia aminoacídica en % se obtienen tal como se describe más ade-
lante mediante la utilización de un programa informático de comparación de secuencias ALIGN-2, en donde el código
fuente completo para el programa ALIGN-2 se proporciona en la Tabla-1. El programa informático de comparación de
secuencias ALIGN-2 fue realizado por Genentech, Inc., y el código fuente que se muestra en la Tabla 1 se ha rellenado
55 con la documentación del usuario en la Oficina de Derechos de Autor de U.S., Washington D.C., 20559, en donde se
ha registrado con el No. TXU510087. El programa ALIGN-2 es de disponibilidad pública gracias a Genentech, Inc.,
South San Francisco, California o puede editarse a partir del código fuente proporcionado en la Tabla 1. El programa
ALIGN-2 debería editarse para utilizarse en un sistema operativo UNIX, preferentemente UNIX V4.0D digital. Todos
los parámetros de comparación de secuencias están determinados por el programa ALIGN-2 y no varían.

60 Para los propósitos de la presente invención, el % de identidad de secuencia aminoacídica de una secuencia ami-
noacídica dada A respecto a o, o con, una secuencia aminoacídica dada B (que alternativamente puede formularse
como una secuencia aminoacídica dada A que tiene o comprende un cierto % de identidad de secuencia aminoacídica
respecto a, con, una secuencia aminoacídica dada B) se calcula de la manera siguiente:

65 100 veces la fracción X/Y

ES 2 279 473 T3

Donde X es el número de residuos aminoacídicos valorados con una concordancia completa por el programa ALIGN-2 de alineamiento de secuencias, en donde se alinean A y B e Y es el número total de residuos aminoacídicos en B. Se apreciará que si la longitud de la secuencia aminoacídica A no es igual a la de B, el % de identidad de secuencia aminoacídica de A respecto a B no será igual al % de identidad de secuencia aminoacídica de B respecto a A. Como ejemplos de % cálculos de identidad de secuencia aminoacídica, en las Tablas 2A-2B se demuestra cómo calcular el % de identidad de secuencia aminoacídica de la secuencia aminoacídica denominada “Proteína de Comparación” respecto a la secuencia aminoacídica denominada “PRO”.

Salvo que se especifique lo contrario, todos los valores en % de identidad de secuencia aminoacídica utilizados aquí se obtuvieron tal como se describió anteriormente mediante la utilización del programa informático de comparación de secuencias ALIGN-2. Sin embargo, el % de identidad de secuencia aminoacídica también puede determinarse mediante la utilización de un programa de comparación NCBI-BLAST2 (Altschul y col., Nucleic Acids Res., 25:3389-3402 (1997)). El programa de comparación de secuencias NCBI-NLAST2 puede descargarse de <http://www.ncbi.nlm.nih.gov>. NCBI-BLAS2 utiliza diversos parámetros de búsqueda, donde todos estos parámetros de búsqueda están determinados a ciertos valores por defecto que incluyen, por ejemplo, no enmascarar (*unmask*) = si, cadena (*strand*) = todas, ocurrencias esperadas (*occurrences expected*) = 10, longitud mínima de complejidad baja (*minimum low complexity length*) = 15/5, multipases valor-e (*multi-pass e-value*) = 0,01, constante para multipase (*multi-pass*) = 25, declinar gradualmente (*dropoff for final gapped alignment*) para el alineamiento con huecos final = 25 y matriz de valoración (*scoring matrix*) = BLOSUM62.

En situaciones en donde se utilice NCBI-BLAST2 para la comparación de secuencias aminoacídicas, el % de identidad de secuencia aminoacídica de una secuencia aminoacídica A respecto, con, una secuencia aminoacídica dada B (la cual puede alternativamente formularse como una secuencia aminoacídica dada A que tiene o comprende un cierto % de identidad de secuencia aminoacídica respecto a, con, una secuencia aminoacídica dada B) se calcula de la manera siguiente:

$$100 \text{ veces la fracción } X/Y$$

Donde X es el número de residuos aminoacídicos valorados con una concordancia completa por el programa de alineamiento de secuencias NCBI-BLAST2 en donde se alinean las secuencias A y B, y en donde Y es el número total de residuos aminoacídicos en B. Se apreciará que si la longitud aminoacídica de A no es igual a la de la secuencia aminoacídica de B, el % de identidad de secuencia aminoacídica de A respecto a B no será igual al % de identidad de secuencia aminoacídica de B respecto a A.

Además, el % de identidad de secuencia aminoacídica también puede determinarse mediante la utilización del programa informático WU-BLAST-2 (Altschul y col., Methods in Enzymology, 266:460-480 (1996)). La mayoría de los parámetros de búsqueda de WU-BLAST-2 se ajustan a los valores por defecto. Los no ajustados a los valores por defecto, es decir, los parámetros ajustables, se determinan con los valores siguientes: extensión del solapamiento (*overlap span*) = 1, fracción de solapamiento (*overlap fraction*) = 0,125, letra del umbral (*word threshold*) (T) = 11, y matriz de valoración (*scoring matrix*) = BLOSUM62. Para los propósitos de la presente invención, un valor de % de identidad de secuencia aminoacídica se determina dividiendo (a) por el número de residuos aminoacídicos concordantes entre la secuencia aminoacídica del polipéptido PRO de interés con una secuencia derivada del polipéptido PRO nativo y la comparación de la secuencia aminoacídica de interés (es decir, la secuencia contra quien se compara el polipéptido PRO de interés que puede ser un polipéptido variante de PRO) tal como se determina por WU-BLAST-2 por (b) el número total de residuos aminoacídicos del polipéptido PRO de interés. Por ejemplo, en la presente invención “un polipéptido que comprende una secuencia aminoacídica A que tiene, o tiene al menos un 80% de identidad de secuencia aminoacídica con la secuencia aminoacídica B”, la secuencia aminoacídica A es la secuencia aminoacídica de comparación de interés y la secuencia aminoacídica B es la secuencia aminoacídica del polipéptido PRO de interés.

“Polipéptido variante PRO” o “secuencia de ácido nucleico variante de PRO” quiere decir una molécula de ácido nucleico que codifica una polipéptido PRO activo tal como se define más adelante y que tiene por lo menos un 80% de identidad de secuencia de ácido nucleico con una secuencia de ácido nucleico que codifica un polipéptido PRO de secuencia nativa de longitud completa tal como se describe aquí, una secuencia de polipéptido PRO de secuencia nativa y longitud completa carente del péptido señal tal como se describe aquí, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, tal como se describe aquí o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa tal como se describe aquí. Generalmente, un polinucleótido variante de PRO tendrá por lo menos un 80% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 81% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 82% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 83% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 84% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 85% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 86% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 87% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 88% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 89% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 90% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 91% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 92% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 93% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 94% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 95% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 96% de identidad de secuencia de ácido

nucleico, más preferentemente al menos un 97% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 98% de identidad de secuencia de ácido nucleico, más preferentemente al menos un 99% de identidad de secuencia de ácido nucleico con la secuencia de ácido nucleico que codifica una secuencia de polipéptido PRO de secuencia nativa y longitud completa tal como se describe aquí, una secuencia de polipéptido PRO de secuencia nativa de tamaño completo carente del péptido señal tal como se describe aquí, un dominio extracelular de un polipéptido PRO, con o sin la secuencia señal, tal como se describe aquí o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa tal como se describe aquí. Las variantes no abarcan la secuencia nucleotídica nativa.

Generalmente, los polinucleótidos variantes de PRO son de al menos 30 nucleótidos de longitud, a menudo de aproximadamente 60 nucleótidos de longitud, más a menudo de al menos 90 nucleótidos de longitud, más a menudo de al menos 120 nucleótidos de longitud, más a menudo de al menos 150 nucleótidos de longitud, más a menudo de al menos 180 nucleótidos de longitud, más a menudo de al menos 210 nucleótidos de longitud, más a menudo de al menos 240 nucleótidos de longitud, más a menudo de al menos 270 nucleótidos de longitud, más a menudo de al menos 300 nucleótidos de longitud, más a menudo de al menos 450 nucleótidos de longitud, más a menudo de al menos 600 nucleótidos de longitud, más a menudo de al menos 900 nucleótidos de longitud, o más.

“Porcentaje (%) de identidad de secuencia de ácido nucleico” respecto a la secuencia de ácido nucleico que codifica el polipéptido PRO identificada aquí se define como el porcentaje de nucleótidos en una secuencia candidata que son idénticos con los nucleótidos en una secuencia de ácido nucleico que codifica un polipéptido PRO, después de alinear las secuencias e introducir los huecos, si fuera necesario, para lograr el % máximo de identidad de secuencia. El alineamiento con el propósito de determinar el porcentaje de identidad de secuencia de ácido nucleico puede lograrse de diversas formas que se hallan dentro del ámbito de la materia, por ejemplo, mediante la utilización de un programa informático disponible públicamente como BLAST, BLAST-2, ALIGN, ALIGN-2, o el programa Megalign (ADNSTAR). Los expertos en la materia pueden determinar los parámetros adecuados para la cuantificación del alineamiento, incluyendo cualquiera de los algoritmos que se requieren para lograr un alineamiento máximo en toda la longitud de la secuencia a comparar. Sin embargo, para los propósitos de la presente invención, los valores en % de identidad de secuencia de ácido nucleico se obtienen tal como se describe más adelante mediante la utilización del programa informático de comparación de secuencias ALIGN-2, en donde el código fuente completo para el programa ALIGN-2 se proporciona en la Tabla 1. El programa informático de comparación de secuencias ALIGN-2 se generó por Genentech Inc., y el código fuente que se muestra en la Tabla 1 se ha rellenado con la documentación del usuario en la Oficina de Derechos de Autor de U.S., Washington D.C., 20559, donde se registró con el N° TXU510087. El programa ALIGN-2 está disponible públicamente gracias a Genentech Inc., South San Francisco, California, o puede ser editado a partir del código fuente proporcionado en la Tabla 1. El programa ALIGN-2 debería editarse para su utilización en un sistema operativo UNIX, preferentemente UNIX V4.0D digital. Todos los parámetros de comparación de secuencias están determinados por el programa ALIGN-2 y no varían.

Para los propósitos de la presente invención, el % de identidad de secuencia de ácido nucleico de una secuencia de ácido nucleico dada C respecto a, con, una secuencia de ácido nucleico dada D (la cual puede formularse alternativamente como una secuencia de ácido nucleico dada C que tiene o comprende un cierto % de identidad de secuencia de ácido nucleico con, respecto a, una secuencia de ácido nucleico dada D) se calcula de la manera siguiente:

$$100 \text{ veces el cociente } W/Z$$

En donde W es el número de nucleótidos valorados como completamente concordantes por el programa de alineamiento de secuencias ALIGN-2, en donde se alinean las secuencias C y D, y Z es el número total de nucleótidos en D. Se apreciará que si la longitud de la secuencia de ácido nucleico C no es igual a la longitud de la secuencia de ácido nucleico D, el % de identidad de secuencia de ácido nucleico de C respecto a D no será igual al % de identidad de secuencia de ácido nucleico de D respecto a C. En las tablas 2C-2D se muestran ejemplos del cálculo del % de identidad de secuencia de ácido nucleico, y se demuestra cómo calcular el % de identidad de secuencia de ácido nucleico de la secuencia de ácido nucleico denominada “ADN de comparación” respecto a la secuencia de ácido nucleico denominada “PRO-ADN”.

Salvo que se especifique lo contrario, todos los valores de % de identidad de secuencia de ácido nucleico que se utilizan aquí se obtienen tal como se describió utilizando el programa informático de comparación de secuencias ALIGN-2. Sin embargo, el % de identidad de secuencia de ácido nucleico también puede determinarse mediante la utilización del programa de comparación de secuencias NCBI-BLAST2 (Altschul y col., *Nucleic Acids Res.*, 25:3389-3402 (1997)). El programa de comparación de secuencias NCBI-BLAST2 puede descargarse de <http://www.ncbi.nlm.nih.gov>. NCBI-BLAST2 utiliza diversos parámetros de búsqueda, donde todos estos parámetros de búsqueda están determinados a ciertos valores por defecto que incluyen, por ejemplo, no enmascarar (*unmask*) = sí, cadena (*strand*) = todas, ocurrencias esperadas (*expected occurrences*) = 10, longitud mínima de complejidad baja (*minimum low complexity length*) = 15/5, multipases valor-e = 0,01, constante para multipase (*constant for multi-pass*) = 25, declinar gradualmente para el alineamiento con huecos final (*dropoff for final gapped alignment*) = 25 y matriz de valoración (*scoring matrix*) = BLOSUM62.

En situaciones donde NCBI-BLAST2 se utiliza para la comparación de secuencias, el % de identidad de secuencia de ácido nucleico de una secuencia de ácido nucleico dada C respecto a, con, una secuencia de ácido nucleico dada D (que alternativamente puede enunciarse como una secuencia de ácido nucleico dada C que tiene o comprende un

ES 2 279 473 T3

cierto % de identidad de secuencia con, respecto a, una secuencia de ácido nucleico dada D) se calcula de la manera siguiente:

$$100 \text{ veces el cociente } W/Z$$

5

En donde W es el número de nucleótidos valorados como completamente concordantes por el programa de alineamiento de secuencias NCBI-BLAST2, en donde se alinean las secuencias C y D, y Z es el número total de nucleótidos en D. Se apreciará que si la longitud de la secuencia de ácido nucleico C no es igual a la longitud de la secuencia de ácido nucleico D, el % de identidad de secuencia de ácido nucleico de C respecto a D no será igual al % de identidad de secuencia de ácido nucleico de D respecto a C.

10

Además, los valores del % de identidad de secuencia de ácido nucleico también pueden generarse utilizando el programa informático WU-BLAST-2 (Altschul y col., *Methods in Enzymology*, 266:460-480) (1996)). La mayoría de los parámetros de búsqueda de WU-BLAST-2 están ajustados a los valores por defecto. Los no ajustados a los valores por defecto, es decir, los parámetros ajustables, se determinan con los valores siguientes: extensión del solapamiento (*overlap span*) = 1, fracción de solapamiento (*overlap fraction*) = 0,125, letra del umbral (*word threshold*) (T) = 11, y matriz de valoración (*scoring matrix*) =BLOSUM62. Para los propósitos de la presente invención, un valor de % de identidad de secuencia de ácido nucleico se determina dividiendo (a) por el número de nucleótidos concordantes entre la secuencia de ácido nucleico codificante del polipéptido PRO de interés que tiene una secuencia derivada del ácido nucleico codificante del polipéptido PRO y la molécula de ácido nucleico de interés de comparación (es decir, la secuencia contra quien se compara la molécula de ácido nucleico codificante del polipéptido PRO de interés, la cual puede ser un polinucleótido PRO variante), tal como se determina por WU-BLAST-2, por (b) el número total de nucleótidos de la molécula de ácido nucleico codificante del polipéptido PRO de interés. Por ejemplo, en la presente invención “una molécula de ácido nucleico aislada que comprende una secuencia de ácido nucleico A, la cual tiene o comprende al menos un 80% de identidad de secuencia de ácido nucleico con la secuencia de ácido nucleico B”, la secuencia de ácido nucleico A es la secuencia de ácido nucleico de comparación de interés y la secuencia de ácido nucleico B es la secuencia de ácido nucleico de la molécula de ácido nucleico codificante del polipéptido PRO de interés.

15

20

25

En otras realizaciones, los polinucleótidos variantes PRO son moléculas de ácido nucleico que codifican un polipéptido PRO activo y que son capaces de hibridar, preferentemente en condiciones astringentes de hibridación y lavado, con las secuencias nucleotídicas que codifican el polipéptido PRO de longitud completa que se muestra en la Figura 2 (SEC ID NO:2). Los polipéptidos variantes de PRO pueden ser aquellos codificados por un polinucleótido variante PRO.

30

El término “positivos”, en el contexto de comparaciones de identidad de secuencia aminoacídica realizadas tal como se describe más arriba, incluye residuos aminoacídicos en las secuencias comparadas que no solamente son idénticos, sino que además tienen propiedades similares. Los residuos aminoacídicos que presentan un valor positivo respecto a un residuo aminoacídico de interés son aquellos que son idénticos al residuo aminoacídico de interés o son una sustitución preferida (tal como se define en la Tabla 3 de más adelante) del residuo aminoacídico de interés.

35

Para el propósito de la presente invención, el % de positivos de una secuencia aminoacídica dada A respecto a, con, un cierto % de valores positivos respecto a, con, una secuencia aminoacídica dada B) se calcula de la manera siguiente:

40

$$100 \text{ veces la fracción } X/Y$$

En donde X es el número de residuos aminoacídicos con una valoración positiva tal como se definió más arriba por el programa de alineamiento de secuencias ALIGN-2 en donde se alinean las secuencias A y B, e Y es el número total de residuos aminoacídicos en B. Se apreciará que si la longitud de la secuencia aminoacídica A no es igual a la longitud de la secuencia aminoacídica B, el % de positivos de A respecto a B no será igual al % de positivos de B respecto a A.

45

El término “aislado” cuando se utiliza para describir los diversos polipéptidos aquí descritos, significa que el polipéptido se ha identificado y separado y/o recuperado de un componente de su entorno natural. Preferentemente, el polipéptido aislado no presenta asociación con ninguno de los componentes normalmente asociados en la naturaleza. Los componentes contaminantes de su entorno natural son materiales que interfieren típicamente con las utilidades diagnósticas o terapéuticas para el polipéptido, y pueden incluir enzimas, hormonas, y otros solutos proteicos o no proteicos. En realizaciones preferidas, el polipéptido se purificará (1) a un grado suficiente para obtener por lo menos 15 residuos del extremo N-terminal o de la secuencia interna aminoacídica mediante la utilización de un secuenciador de taza giratoria, o (2) a homogeneidad mediante SDS-PAGE en condiciones reductoras o no-reductoras utilizando azul de Coomassie o, preferentemente, tinción de plata. El polipéptido aislado incluye polipéptidos *in situ* dentro de las células recombinantes, ya que al menos un componente del entorno natural de PRO no estará presente. Por lo general, sin embargo, el polipéptido aislado se preparará mediante al menos una etapa de purificación.

50

Una molécula de “ácido nucleico” que codifica un polipéptido PRO o un ácido nucleico “aislado” que codifica un anticuerpo PRO, es una molécula de ácido nucleico que se identifica y separa de por lo menos una molécula de ácido nucleico contaminante con la que normalmente está asociada en la fuente natural del ácido que codifica PRO

55

o del ácido nucleico que codifica anti-PRO. Preferentemente, el ácido nucleico aislado está libre de toda asociación con cualquier componente de los que normalmente está asociado en la naturaleza. Una molécula de ácido nucleico que codifica PRO o una molécula de ácido nucleico que codifica anti-PRO son distintas tanto en la forma como en su disposición de la que se halla en la naturaleza. Las moléculas de ácido nucleico aisladas se diferencian en consecuencia de la molécula de ácido nucleico que codifica PRO, o de la molécula de ácido nucleico que codifica anti-PRO tal como existen en la naturaleza. Sin embargo, una molécula de ácido nucleico aislado, que codifica un polipéptido PRO o un anticuerpo anti-PRO incluye moléculas de ácido nucleico y moléculas de ácido nucleico anti-PRO contenidas en células que ordinariamente expresan polipéptidos PRO o que expresan anticuerpos anti-PRO, si, por ejemplo, la molécula de ácido nucleico se halla en una localización cromosómica distinta a la de las células naturales.

El término “secuencias control” hace referencia a secuencias de ADN necesarias para la expresión de una secuencia codificante unida operativamente en un organismo huésped. Las secuencias control que son adecuadas para procariotas, incluyen por ejemplo un promotor, opcionalmente una secuencia operadora, y un sitio de unión a ribosomas. Las células eucariotas son conocidas por utilizar promotores, señales de poliadenilación e intensificadores.

El ácido nucleico está “unido operativamente” cuando se le coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN para una presecuencia o líder secretor está unido operativamente al ADN de un polipéptido si éste se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o un intensificador está unido operativamente a una secuencia codificante si afecta la transcripción de la secuencia; o un sitio de unión a ribosomas está unido operativamente a una secuencia codificante si está localizado para facilitar la traducción. Por lo general, “unido operativamente” significa que las secuencias de ADN que se unen son contiguas y, en el caso de un líder secretor, contiguas y en la misma fase de lectura. Sin embargo, los intensificadores no tienen porque ser contiguos. La unión se logra mediante ligación en las dianas de restricción convenientes. Si dichas dianas no existieran, se utilizarían adaptadores oligonucleotídicos sintéticos o engarces de acuerdo con la práctica convencional.

El término “anticuerpo” se utiliza en el sentido amplio y abarca específicamente, por ejemplo, anticuerpos monoclonales (incluyendo antagonistas, y anticuerpos neutralizantes), composiciones de anticuerpos anti-PRO5775 con especificidad poliepítópica, anticuerpos anti-PRO5775 de cadena sencilla, y fragmentos de anticuerpos anti-PRO5775 (véase más adelante). El término “anticuerpo monoclonal” tal como se utiliza aquí hace referencia a un anticuerpo obtenido de una población de anticuerpos esencialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos salvo por mutaciones posibles que ocurren en la naturaleza y que pueden estar presentes en cantidades menores.

La “astringencia” de las reacciones de hibridación se determina fácilmente por el experto en la materia, y en general es un cálculo aproximado dependiente de la longitud de la sonda, temperatura de lavado, y concentración salina. Por lo general, sondas más largas requieren temperaturas más elevadas para una hibridación adecuada, mientras que sondas más cortas necesitan temperaturas más bajas. La hibridación depende generalmente de la capacidad del ADN desnaturalizado para renaturalizarse cuando se hallan presentes cadenas complementarias en un entorno por debajo de la temperatura de fusión. Cuanto mayor es el grado de homología deseado entre la sonda y la secuencia de hibridación, mayor es la temperatura relativa que puede utilizarse. Como resultado, temperaturas relativamente superiores favorecerán que las condiciones de reacción sean más astringentes, mientras que temperaturas inferiores favorecerán condiciones de reacción menos astringentes. Para detalles adicionales y una explicación de la astringencia de las reacciones de hibridación, véase Ausubel y col., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

“Condiciones astringentes” o “condiciones de astringencia elevada”, tal como se define aquí, pueden identificarse por: (1) utilizar una temperatura elevada y una fuerza iónica baja para el lavado, por ejemplo, cloruro sódico 0,015M/citrato sódico 0,0015M/sodio dodecil sulfato al 0,1% a 50°C; (2) utilizar un agente desnaturalizante durante la hibridación, como la formamida, por ejemplo, formamida (v/v) al 50% con albúmina sérica bovina al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42°C; o (3) utilizar formamida al 50%, 5XSSC (NaCl 0,75M, citrato sódico 0,075M), fosfato sódico 50 mM (pH6,8), pirofosfato sódico al 0,1%, 5X solución de Denhardt, ADN de esperma de salmón sonicado (50 µg/ml), SDS al 0,1% y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2X SSC (cloruro sódico/citrato sódico) y formamida al 50%, seguido de un lavado a astringencia elevada que consiste en 0,1XSSC con EDTA a 55°C.

“Condiciones de astringencia moderada” pueden identificarse tal como se describe por Sambrook y col., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989, e incluyen la utilización de una solución de lavado y condiciones de hibridación (por ejemplo, la temperatura, la fuerza iónica y el % de SDS) menos astringentes que las descritas más arriba. Un ejemplo de condiciones de astringencia moderada es la incubación durante toda la noche a 37°C en una solución que comprende: formamida al 29%, 5XSSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), solución de 5X Denhardt, sulfato de dextran al 10% y ADN de esperma de salmón desmenuzado y desnaturalizado a 20 mg/ml, seguido de lavados de los filtros en 1XSSC a aproximadamente 35-50°C. El experto en la materia reconocerá cómo ajustar la temperatura, fuerza iónica, etc., tanto como sea necesario para adaptar factores como la longitud de la sonda y otros.

El término “etiquetado epítópicamente” cuando se utiliza aquí hace referencia a un polipéptido quimérico que comprende un polipéptido PRO5775 fusionado con un “polipéptido etiqueta”. El polipéptido etiqueta tiene suficientes residuos para proporcionar un epitopo contra el cual se genera un anticuerpo, y es lo suficientemente corto para no

ES 2 279 473 T3

interferir con la actividad del polipéptido con el que se fusiona. El polipéptido etiqueta también es preferentemente único, de modo que el anticuerpo no reacciona de forma cruzada con otros epitopos. Los polipéptidos etiqueta adecuados tienen por lo menos 6 residuos aminoacídicos y generalmente se hallan entre 8 y 50 residuos aminoacídicos (preferentemente entre 10 y 20 residuos aminoacídicos).

5 “Activo” o “actividad” para el propósito de la presente invención hace referencia a la forma(s) de los polipéptidos PRO5775 que retienen una actividad/propiedad biológica y/o inmunológica de un polipéptido PRO5775 nativo o que ocurre en la naturaleza, en donde “actividad” biológica hace referencia a una función (inhibidora o estimuladora) causada por un polipéptido PRO5775 nativo o que ocurre de forma natural distinta a la capacidad para inducir la
10 producción de un anticuerpo contra un epitopo antigénico que forma parte de un polipéptido PRO5775 nativo o que ocurre de forma natural y una actividad “inmunológica” hace referencia a la capacidad para inducir un anticuerpo contra un epitopo antigénico contenido por un polipéptido PRO5775 que ocurre de forma natural.

15 “Actividad biológica” en el contexto de un anticuerpo u otra molécula antagonista que puede identificarse mediante ensayos de cribado descritos aquí (por ejemplo, una molécula pequeña orgánica o inorgánica, péptido, etc.) se utiliza para referirse a la capacidad de dichas moléculas para unirse o formar complejos con los polipéptidos codificados por los genes amplificados identificados aquí, o que interfieren con la interacción de los polipéptidos codificados con otras proteínas celulares o que interfieren con la transcripción o traducción de un polipéptido PRO5775. Una actividad biológica preferida es la inhibición del crecimiento de una célula tumoral. Otra actividad biológica preferida es la
20 actividad citotóxica que resulta en la muerte de las células tumorales.

El término “actividad biológica” en el contexto de un polipéptido PRO5775 hace referencia a la capacidad de un polipéptido PRO5775 para inducir el crecimiento de células neoplásicas o del crecimiento celular incontrolado.

25 El término “actividad inmunológica” hace referencia a la reactividad cruzada inmunológica con por lo menos un epitopo de un polipéptido PRO5775.

El término “reactividad inmunológica cruzada” tal como se utiliza aquí hace referencia a que el polipéptido candidato es capaz de inhibir competitivamente la actividad biológica cualitativa de un polipéptido PRO5775 que tiene esta
30 actividad con el antisuero policlonal generado contra el polipéptido PRO5775 activo conocido. Dicho antisuero se prepara de manera convencional mediante inyección de cabras o conejos, por ejemplo, subcutáneamente con el análogo activo conocido en adyuvante completo de Freund, seguido de una inyección de recuerdo intraperitoneal o subcutánea en adyuvante incompleto de Freund. La reactividad cruzada es preferentemente “específica”, lo que hace referencia a que la afinidad de unión de la molécula identificada que reacciona inmunológicamente de forma cruzada (por ejemplo el anticuerpo), con respecto al polipéptido PRO5775 correspondiente es significativamente superior (preferentemente
35 de por lo menos 2 veces, más preferentemente de al menos 4 veces, aún más preferentemente de al menos 8 veces, más preferentemente de al menos 10 veces superior) que la afinidad de unión de dicha molécula con cualquier otro polipéptido nativo conocido.

40 El término “antagonista” se utiliza en el sentido amplio, e incluye cualquier molécula que bloquea parcial o completamente, inhibe o neutraliza una actividad biológica de un polipéptido PRO5775 descrito aquí, o su transcripción o traducción. Moléculas antagonistas adecuadas incluyen específicamente los anticuerpos o los fragmentos de anticuerpos, los péptidos, las moléculas orgánicas pequeñas, los ácidos nucleicos anti-sentido, etc.

45 Una “molécula pequeña” se define aquí con un peso molecular de aproximadamente 500 Daltons.

“Anticuerpos” (Abs) e “inmunoglobulinas” (Igs) son glicoproteínas que tienen las mismas características. Mientras que los anticuerpos presentan una especificidad de unión con un antígeno específico, las inmunoglobulinas incluyen
50 ambos anticuerpos y otras moléculas de tipo anticuerpo que carecen de la especificidad del antígeno. Los polipéptidos del último tipo son, por ejemplo, producidos a niveles bajos por el sistema linfático y a niveles aumentados por mielomas. El término “anticuerpo” se utiliza en el sentido más amplio y abarca específicamente, sin limitación, anticuerpos monoclonales intactos, anticuerpos policlonales, anticuerpos multiespecíficos (por ejemplo, anticuerpo biespecíficos) formados a partir de dos anticuerpos intactos por lo menos, y fragmentos de anticuerpos siempre que presenten la actividad biológica deseada.

55 “Anticuerpos nativos” e “inmunoglobulinas nativas” son glicoproteínas heterotetraméricas de aproximadamente 150.000 daltons, compuestas de dos cadenas ligeras idénticas (L) y dos cadenas pesadas idénticas (H). Cada cadena ligera está unida a una cadena pesada por enlace disulfuro covalente, mientras que el número de uniones disulfuro varía entre las cadenas pesadas de isotopos distintos de inmunoglobulinas. Cada cadena ligera y pesada también
60 tiene puentes disulfuro intracatenarios espaciados regularmente. Cada cadena pesada tiene en un extremo un dominio variable (V_H) seguido por un número de dominios constantes. Cada cadena ligera tiene un dominio variable en un extremo (V_L) y un dominio constante en su otro extremo; el dominio constante de la cadena ligera está alineado con el primer dominio constante, y el dominio variable de cadena ligera está alineado con el dominio variable de la cadena pesada. Los residuos aminoacídicos particulares se cree que forman una interfaz entre los dominios variables de cadena
65 ligera y pesada.

El término “variable” hace referencia al hecho de que ciertas porciones de los dominios variables difieren extensamente en la secuencia entre los anticuerpos y se utilizan en la especificidad de unión de cada anticuerpo particular para

su antígeno de interés. Sin embargo, la variabilidad no está distribuida uniformemente entre los dominios variables de los anticuerpos. Está concentrada en tres segmentos denominados regiones determinantes de la complementariedad (CDRs) o regiones hipervariables tanto en los dominios variables de cadena ligera como de cadena pesada. Las porciones más altamente conservadas de los dominios variables se denominan las regiones de entramado (FR). Los dominios variables de las cadenas pesada y ligera nativas comprenden cada una 4 regiones FR, que adoptan una configuración en hoja β , conectada por los tres CDRs, que forman bucles de conexión y en algunos casos forman parte de la estructura en hoja β . Las CDRs en cada cadena se mantienen juntas en estrecha proximidad por medio de las regiones FR y, con las CDRs de la otra cadena, contribuyen a la formación del sitio de unión al antígeno de los anticuerpos (véase Kabat y col., NIH Publ. No. 91-3242, vol. I, páginas 647-669 (1991)). Los dominios constantes no están implicados directamente en la unión del anticuerpo con el antígeno, pero presentan diversas funciones, como la participación del anticuerpo en la toxicidad celular dependiente del anticuerpo.

El término “región hipervariable” cuando se utiliza aquí hace referencia a los residuos aminoacídicos de un anticuerpo que son responsables de la unión al antígeno. La región hipervariable comprende residuos aminoacídicos de una “región determinante de la complementariedad” o “CDR” (es decir, residuos 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el dominio variable de cadena ligera y 31-35 (H1), 50-65 (H2) y 95-102 (H3) en el dominio variable de cadena pesada; Kabat y col., *Sequences of Proteins of Immunological Interest*, 5ª ed., Public Health Service, National Institute of Health, Bethesda, MD. [1991]) y/o los residuos de un “bucle hipervariable” (es decir, residuos 26-32 (L1), 50-52 (L2) y 91-96 (L3) en el dominio variable de cadena ligera y 26-32 (H1), 53-55 (H2) y 96-101 (H3) en el dominio variable de cadena pesada; Clothia y Lesk, *J. Mole. Biol.* 196:901-917 [1987]). Los residuos de “entramado” o “FR” son aquellos residuos de dominio variable distintos a los residuos de la región hipervariable tal como se define aquí.

Los “fragmentos de anticuerpo” comprenden una porción de un anticuerpo intacto, preferentemente la región de unión el antígeno o región variable del anticuerpo intacto. Ejemplos de fragmentos de anticuerpo incluyen el Fab, Fab', F(ab')₂ y los fragmentos Fv; los diacuerpos; los anticuerpos lineales (Zapata y col., *Protein Eng.*, 8(10):1057-1062 [1995]); moléculas de anticuerpo de una sola cadena; y anticuerpos multiespecíficos formados por fragmentos de anticuerpos.

La digestión con papaína de los anticuerpos produce dos fragmentos de unión idénticos, denominados fragmentos “Fab”, cada uno con un sitio de unión antigénico único, y un fragmento “Fc” residual, cuyo nombre refleja su capacidad para cristalizar fácilmente. El tratamiento con pepsina produce un fragmento (Fab')₂ que tiene dos sitios de combinación antigénica y es capaz de unirse de forma cruzada con el antígeno.

“Fv” es el fragmento de anticuerpo mínimo que contiene un sitio de reconocimiento y un sitio de unión completos. Esta región consiste de un dímero de un dominio variable de una cadena pesada-una cadena ligera en estrecha asociación no covalente. Es en esta configuración que las tres CDRs de cada dominio variable interactúan para definir un sitio de unión antigénico en la superficie del dímero VH-VL. Colectivamente, las seis CDRs confieren especificidad de unión antigénica al anticuerpo. Sin embargo, incluso un dominio variable único (o la mitad de un Fv que comprende sólo tres CDRs específicas para el antígeno) tiene la capacidad para reconocer y unirse al antígeno, aunque con una afinidad menor que el sitio de unión entero.

El fragmento Fab también contiene el dominio constante de la cadena ligera y el dominio constante (CH1) de la cadena pesada. Los fragmentos Fab difieren de los fragmentos Fab' mediante la adición de unos pocos residuos en el extremo carboxi-terminal del dominio CH1 de cadena pesada incluyendo una o más cisteínas de la región bisagra del anticuerpo, Fab'-SH es la denominación para Fab' en la que los residuos cisteína de los dominios constantes contienen un tercer grupo tiol libre, los fragmentos de anticuerpo F(ab')₂ se producen originalmente como parejas de fragmentos Fab' con una bisagra de cisteínas entre ambos. Otras uniones químicas de fragmentos de anticuerpos se muestran también aquí.

Las “cadenas ligeras” de anticuerpos (inmunoglobulinas) de cualquier especie de vertebrados pueden asignarse a uno de los dos tipos claramente distintos, denominados kappa (κ) y lambda (λ), según las secuencias aminoacídicas de sus dominios constantes.

Las inmunoglobulinas pueden agruparse en cinco clases diferentes en función de la secuencia de aminoácidos del dominio constante de sus cadenas pesadas. Las cinco clases principales de inmunoglobulinas son: IgA, IgD, IgE, IgG, e IgM y algunas de ellas pueden a su vez dividirse en subclases (o isotipos), por ejemplo, IgG1, IgG2, IgG3, IgG4, IgA e IgA2. Los dominios constantes de la cadena pesada que corresponden a las diferentes clases de inmunoglobulinas se denominan respectivamente α , δ , ϵ , γ y μ . Las estructuras de las subunidades y las configuraciones tridimensionales de las diferentes clases de inmunoglobulinas son bien conocidas.

El término “anticuerpo monoclonal” se usa aquí para denominar un anticuerpo obtenido de una población de anticuerpos esencialmente homogéneos, por ejemplo los anticuerpos individuales de los que consta una población son idénticos excepto si existen posibles mutaciones naturales que se presentan en una cantidad mínima. Los anticuerpos monoclonales son altamente específicos, ya que están dirigidos contra un único lugar antigénico. Además, en contraste con las preparaciones de anticuerpos convencionales (policlonales), que incluyen anticuerpos dirigidos contra diferentes determinantes (epitopos), cada anticuerpo monoclonal se dirige contra un único determinante antigénico. Además de su especificidad, los anticuerpos monoclonales presentan la ventaja de ser sintetizados en un cultivo de hibridoma, sin contaminación por otras inmunoglobulinas. El calificativo “monoclonal” indica así el carácter del anticuerpo

obtenido de una población esencialmente homogénea de anticuerpos, y no requiere un procedimiento particular para su producción. Por ejemplo, los anticuerpos monoclonales que se usan de acuerdo con la presente invención, pueden obtenerse por el procedimiento del hibridoma descrito primeramente por Kohler y col., *Nature*, 256:495 [1975], o por procedimientos de ADN recombinante (véase, por ejemplo, la patente americana U.S. No. 4.816.567). Los “anticuerpos monoclonales” también pueden aislarse, por ejemplo, de librerías de anticuerpos en fagos según las técnicas descritas en Clackson y col., *Nature*, 352:624-628 [1991] y Marks y col., *J. Mol. Biol.*, 222:581-597 (1991).

Los anticuerpos monoclonales descritos aquí, específicamente incluyen anticuerpos “quiméricos” (inmunoglobulinas) en los que una porción de la cadena pesada o ligera es idéntica a, o, homóloga a las secuencias correspondientes de anticuerpos derivados de una especie particular, o pertenecen a una clase o subclase particular de anticuerpo, mientras el resto de la cadena(s) es idéntica a, u, homóloga a las secuencias correspondientes en anticuerpos derivados de otras especies o pertenecientes a otras clases de anticuerpos o subclases, así como los fragmentos de estos anticuerpos, siempre que exhiban la actividad biológica deseada (Patente U.S. No. 4,816,567; Morrison y col., *Proc. Natl. Acad. Sci. USA*: 6851-6855 [1984]).

Las formas “humanizadas” de anticuerpos no humanos (por ejemplo los murinos) son inmunoglobulinas quiméricas, cadenas de inmunoglobulinas o fragmentos de éstas (como Fv, Fab, Fab', F(ab')₂) u otras subsecuencias de unión al antígeno de los anticuerpos) que contienen secuencias mínimas derivadas de inmunoglobulinas no humanas. La mayor parte de los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpos receptores) en los que los residuos de un CDR del receptor es reemplazado por los residuos de un CDR de una especie no humana (anticuerpo donante) por ejemplo de ratón, rata o conejo, con la especificidad, afinidad y capacidad deseadas. En algunos casos, los residuos Fv FR residuos de la inmunoglobulina humana se reemplazan por los correspondientes residuos no humanos. Además, los anticuerpos humanizados pueden incluir residuos que no se encuentran ni en el anticuerpo receptor ni en la CDR importada ni en las secuencias de entramado. Estas modificaciones se hacen para realizar un mejor ajuste y maximizar el comportamiento del anticuerpo. En general, el anticuerpo humanizado puede incluir al menos uno, y típicamente dos dominios variables, en los que todas o esencialmente todas las regiones CDR corresponden a las de inmunoglobulinas no humanas y todas o esencialmente todas las regiones FR a las secuencias de inmunoglobulinas humanas. Para ser óptimo, el anticuerpo humanizado también debe incluir al menos una porción de una región constante de una inmunoglobulina (Fc), típicamente de una inmunoglobulina humana. Para más detalles, véase, Jones y col., *Nature*, 321:522-525 (1986); Reichmann y col., *Nature*, 332: 323-329 [1988]; y Presta, *Curr. Op. Struct. Biol.*, 2: 593-596 (1992). El anticuerpo humanizado incluye un anticuerpo PRIMATIZADO™ en el que la región de unión al antígeno deriva de un anticuerpo producido por inmunización de monos macacos con el antígeno de interés.

Los fragmentos de anticuerpo “Fv de una sola cadena” o “sFv” incluyen los dominios V_H y V_L del anticuerpo, dominios que se presentan como cadena polipeptídica única. Preferentemente, el polipéptido Fv además incluye un engarce polipeptídico entre los dominios V_H y V_L que permite al sFv formar la estructura deseada para la unión del antígeno. Para una revisión sobre sFv véase Pluckthun en *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

El término “diacuerpos” se refiere a fragmentos pequeños de anticuerpos que contienen dos lugares de unión antigénica y que poseen un dominio variable de cadena pesada (V_H) conectado a un dominio variable de cadena ligera (V_L) en la misma cadena polipeptídica (VH-VL). Utilizando un engarce suficientemente corto, que permita el emparejamiento entre los dos dominios de la misma cadena, se fuerza al emparejamiento con el dominio complementario en la otra cadena y se crean así dos lugares de unión al antígeno. Los diacuerpos se describen más extensamente en, por ejemplo, la patente europea EP 404,097; la patente internacional WO 93/11161; y Hollinger y col., *Proc. Natl. Acad. Sci. USA*, 90:6444-6448 (1993).

Un anticuerpo “aislado” es aquel que ha sido identificado y separado y/o recuperado de los componentes de su ambiente natural. Los componentes contaminantes de su ambiente natural son materiales que podrían interferir en su posible uso diagnóstico y terapéutico y puede tratarse de enzimas, hormonas u otros solutos proteicos o no proteicos. En realizaciones preferentes, el anticuerpo se purifica (1) en más del 95% en peso según se determina por el procedimiento de Lowry y (2) preferiblemente en más del 99% en peso a un grado suficiente para obtener al menos 15 residuos del N-terminal o de la secuencia interna de aminoácidos que permita el uso de un secuenciador de taza rotatoria o, (3) a homogeneidad mediante separación en SDS-PAGE en condiciones reductoras o no reductoras y tinción con azul de Coomassie o preferentemente con tinción de plata. El anticuerpo aislado incluye el anticuerpo *in situ* sin células recombinantes y sin que se encuentre presente ninguno de los componentes del ambiente natural. No obstante, en general, para preparar anticuerpo aislado, es necesario al menos un paso de purificación.

La palabra “marcador”, cuando se usa aquí, se refiere a un compuesto o composición detectable que se conjuga directa o indirectamente con el anticuerpo para generar un anticuerpo “marcado”. El marcador puede ser detectable por sí mismo (por ejemplo marcadores radioisotópicos o marcadores fluorescentes) o, en el caso del marcador enzimático, pueden catalizar la alteración química de un compuesto o composición detectables que actúan como sustrato. Entre los radionucleidos, que son marcadores detectables, se incluyen por ejemplo, I-131, I-123, I-125, Y-90, Re-188, Re-186, At-211, Cu-67, Bi-212 y Pd-109. Así mismo el marcador puede ser una entidad no detectable como es el caso de una toxina.

Por “fase sólida” se entiende una matriz no acuosa a la que se puede adherir el anticuerpo de la presente invención. Ejemplos de fases sólidas incluidas aquí son aquellas entera o parcialmente compuestas de vidrio (por ejemplo, vidrio

de poro controlado), de polisacáridos (por ejemplo, agarosa), de poliácridamidas, de poliestireno, de polivinil alcohol y de siliconas. En ciertas realizaciones, dependiendo del contexto, la fase sólida puede referirse al pocillo de una placa de ensayo; en otras, a la purificación en columna (por ejemplo, una columna de cromatografía por afinidad). Este término también incluye las fases sólidas discontinuas de partículas discretas tal como las descritas en la patente americana U.S. 4.275.149.

Un "liposoma" es una pequeña vesícula compuesta de varios tipos de lípidos, fosfolípidos y/o tensoactivos que resulta útil para la administración de un compuesto a un mamífero (tal como el polipéptido PRO5775 o el anticuerpo contra éste, y opcionalmente, un agente quimioterapéutico). Los componentes del liposoma están dispuestos generalmente en una formación de bicapa, de manera similar a las membranas biológicas.

El término "inmuno adhesina", tal como se utiliza aquí, designa moléculas de tipo anticuerpo que combinan la especificidad de unión de una proteína heteróloga (una "adhesina") con las funciones efectoras de los dominios constantes de la inmunoglobulina. Estructuralmente, las inmuno adhesinas se componen de la fusión de una secuencia de aminoácidos con la especificidad de unión deseada y que es distinta del sitio de reconocimiento antigénico y del lugar de unión de un anticuerpo (es decir es "heteróloga"), y de una secuencia del dominio constante de la inmunoglobulina. La parte de adhesina de una molécula de inmuno adhesina típicamente es una secuencia contigua de aminoácidos que incluye al menos el sitio de unión de un receptor a un ligando. La secuencia del dominio constante de la inmunoglobulina en la inmuno adhesina puede obtenerse a partir de cualquier inmunoglobulina, tal como los subtipos IgG-1, IgG-2, IgG-3 o IgG-4, IgA (incluidos IgA-1 y IgA-2), IgE, IgD o IgM.

II. Composiciones y Procedimientos de la Invención

I. Polipéptidos PRO5775 de longitud completa

La presente invención se refiere a secuencias de nucleótidos que codifican los polipéptidos a los que se refiere la presente solicitud como PRO5775. El cADN que codifica los polipéptidos PRO5775 se detalla en los Ejemplos de más abajo. Se ha de remarcar que las proteínas producidas en rondas separadas de expresión pueden dar diferentes números de PRO, pero un número único UNQ para un determinado ADN y proteína codificada. Sin embargo, para simplificar las proteínas codificadas por las secuencias de ácidos nucleicos descritas aquí, así como las homólogas nativas y variantes incluidas en la anterior definición de PRO5775, serán referidas como a "PRO5775" con independencia de su origen y forma de preparación.

Tal como se describe en los Ejemplos de más adelante, los clones de cADN se han depositado en el ATCC. La secuencia nucleotídica real de los clones puede determinarse por un técnico en la materia mediante la secuenciación del clon depositado y siguiendo los procedimientos de rutina propios de la materia. Para los polipéptidos PRO5775 y los ácidos nucleicos que los codifican, descritos aquí, los solicitantes han identificado las que se cree son las mejores pautas de lectura identificables con la información de secuencia disponible hasta el momento.

II. Variantes de PRO5775

Además de los polipéptidos PRO5775 de la secuencia nativa completa descritos aquí, se contempla la preparación de las variantes de PRO5775. Las variantes de PRO5775 se pueden preparar mediante la introducción de cambios apropiados en el ADN de PRO5775 y/o por la síntesis del polipéptido de PRO5775 deseado. Los expertos en la materia sabrán que los cambios de aminoácidos pueden alterar el proceso post-traducciona del PRO5775 como por ejemplo cambios en el número y la posición de los sitios de glicosilación o la alteración de sus características de anclaje en la membrana.

Las variaciones en la secuencia nativa de PRO5775 de longitud completa o en varios de los dominios descritos aquí, pueden hacerse, por ejemplo, mediante alguna de las técnicas y guías para mutaciones conservadoras y no conservadoras que se encuentran, por ejemplo, en la patente americana U.S. 5.364.934. Las variaciones pueden generarse por sustitución, deleción o inserción de uno o más codones que codifican PRO5775, lo que resulta en cambios en la secuencia aminoacídica de PRO5775 comparada con la secuencia nativa de PRO5775. Opcionalmente la variante se genera por sustitución de al menos uno de los aminoácidos por otro aminoácido cualquiera en uno o más de los dominios de PRO5775. Una ayuda para la determinación de cuál es el residuo aminoacídico que debe insertarse, substituirse o deleccionarse sin afectar de forma adversa la actividad deseada, puede hallarse en la comparación de la secuencia de PRO5775 con la de moléculas de proteicas homólogas conocidas y minimizando el número de cambios en la secuencia de aminoácidos a realizar en regiones con alta homología. Las substituciones de aminoácidos pueden ser el resultado de reemplazar un aminoácido por otro aminoácido que tenga propiedades similares estructurales o químicas, tal como el reemplazo de una leucina por una serina, es decir, reemplazos de aminoácidos conservados. Las inserciones o deleciones pueden ser opcionalmente de entre 1 y 5 aminoácidos. La variación permitida se puede determinar mediante inserciones, deleciones o substituciones de aminoácidos realizadas sistemáticamente en la secuencia y examinando las variantes resultantes por la actividad que exhiban respecto a la secuencia nativa madura o de longitud completa.

Se proporcionan aquí fragmentos del polipéptido PRO5775. Estos fragmentos, por ejemplo, pueden estar truncados en su extremo N-terminal o C-terminal, o pueden carecer de residuos internos, cuando se comparan con la secuencia nativa madura o de longitud completa. Algunos fragmentos pierden residuos de aminoácidos que no son esenciales para una actividad biológica deseada del polipéptido PRO5775.

ES 2 279 473 T3

Los fragmentos de PRO5775 pueden prepararse mediante cualquiera de las técnicas convencionales. Así, los fragmentos peptídicos deseados pueden ser sintetizados químicamente. Un enfoque alternativo implica la generación de fragmentos PRO5775 por digestión enzimática, por ejemplo, tratando la proteína con una enzima conocida que corta las proteínas en lugares definidos por residuos aminoácidos particulares, o por digestión del ADN con enzimas de restricción adecuadas y aislamiento del fragmento deseado. Aún más, otra técnica adecuada, incluye el aislamiento y la amplificación de un fragmento de ADN que codifica para el fragmento polipeptídico deseado mediante reacción en cadena de la polimerasa (PCR). Los oligonucleótidos que definen las terminaciones deseadas del fragmento de ADN se emplean en los extremos 5' y 3' de los cebadores en la PCR. Preferentemente, los fragmentos del polipéptido PRO5775 comparten al menos una actividad biológica o inmunológica con el polipéptido nativo PRO5775.

En realizaciones particulares, las sustituciones conservadoras de interés se muestran en la Tabla 3 bajo el título correspondiente de sustituciones preferentes. Como resultado de estas sustituciones se produce un cambio en la actividad biológica, se introducen más cambios esenciales denominados ejemplos de sustituciones en la Tabla 3, o tal como se describe más abajo en referencia a las clases de aminoácidos; y posteriormente se criban dichos productos.

TABLA 3

Residuo Original	Ejemplos de sustituciones	Sustituciones Preferidas
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucina	leu
Leu (L)	norleucina; ile; val; Met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr leu	
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; Ala; norleucina	leu

Las modificaciones esenciales de la función o de la identidad inmunológica del polipéptido se consiguen por medio de la selección de sustituciones que difieren significativamente en su efecto sobre (a) el mantenimiento de la estructura del esqueleto polipeptídico en la zona de la sustitución, por ejemplo, en la conformación helicoidal o en hoja; (b) la carga o hidrofobicidad de la molécula; (c) el conjunto de la cadena lateral. Los residuos que se encuentran de forma natural se dividen en grupos en base a las propiedades comunes de sus cadenas laterales:

ES 2 279 473 T3

- (1) hidrofóbicos: norleucina, met, ala, val, leu, ile;
- (2) hidrofílicos neutros: cys, ser, thr,
- 5 (3) acídicos: asp, glu;
- (4) básico: asn, gln, his, lys, arg;
- (5) residuos que influyen la orientación de la cadena: gly, pro, y
- 10 (6) aromáticos: trp, tyr, phe.

Las sustituciones no conservadoras suponen el intercambio de un miembro de una de estas clases por otra. Estas sustituciones se pueden introducir en sitios de sustitución conservadores o más preferentemente, en el resto de sitios no conservadores.

Las variantes se pueden hacer utilizando procedimientos conocidos en la materia tales como la mutagénesis mediada por oligonucleótidos (de sitio-dirigida), el cribado de alaninas, y la mutagénesis por PCR. La mutagénesis lugar-dirigida [Carter *et al.*, Nucl. Acid. Res., 13:4331 (1986); Zoller y col., Nucl. Acids Res., 10:6487 (1987)], mutagénesis en casete [Wells *et al.*, Gene, 34:315 (1985)], la mutagénesis de selección por restricción [Wells y col., Philos. Trans. R. Soc. SerA, 317:415 (1986)] u otras técnicas conocidas se realizan en el ADN clonado para producir las variante del ADN de PRO5775.

El análisis de aminoácidos por cribado también se puede utilizar para identificar uno o más aminoácidos a lo largo de una secuencia continua. Para este tipo de análisis, son preferibles los aminoácidos pequeños y los neutros. Entre estos aminoácidos se incluyen la alanina, la glicina, la serina, y la cisteína. La alanina es típicamente el aminoácido preferido para este análisis debido a que elimina la cadena lateral entre el carbono beta y se altera menos la conformación principal de la variante [Cunningham y Wells, Science, 244:1081-1085 (1989)]. La alanina también es el aminoácido preferido porque es el aminoácido más común. Además, es posibles encontrarlo frecuentemente en posiciones tanto protegidas como expuestas [Creighton, The Proteins, (W.H. Freeman & Co., N. Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. Si la sustitución por alaninas no produce cantidades adecuadas de la variante, se puede utilizar también un aminoácido isostérico.

III. Modificaciones de PRO5775

Las modificaciones covalentes de PRO5775 se incluyen en el ámbito de la presente invención. Un tipo de modificación covalente consiste en la reacción de los residuos aminoácídicos diana de un polipéptido PRO5775 con un agente derivatizante orgánico que es capaz de reaccionar con una cadena lateral seleccionada o con los residuos N- o C-terminales de PRO5775. La derivatización con agentes bifuncionales es útil, por ejemplo, para el entrecruzamiento de PRO5775 con una matriz soportada insoluble en agua o una superficie para su uso en la purificación de anticuerpos anti-PRO5775 y *vice-versa*. Entre los agentes entrecruzadores más comunes, se incluyen, por ejemplo el 1,1-bis (diazacetil)-2-feniletano, el glutaraldehído, los ésteres de N-hidroxisuccinimida, los ésteres con 4-ácido salicílico, imidoésteres homobifuncionales, incluyendo los ésteres de disuccinimidil como el 3'-3-dithiobis (succinimidilpropionato), las maleimidias bifuncionales como el bis-N.-maleimido-1,8-octono y agentes como el metil-3-[(p-azidofenil) dithio]propioimidato.

Otras posibles modificaciones incluyen la desamidación de residuos glutaminil y asparaguinil hasta los correspondientes residuos glutamil y aspartil, respectivamente, y la hidroxilación de la prolina y la lisina, la fosforilación de los grupos hidroxilo de los residuos seril o threonil, la metilación de grupos α -amino de lisina, arginina y cadenas laterales de histidina [T. Creighton, Proteínas: Estructura y Propiedades Moleculares Moleculares, W.H. Freeman & Co., San Francisco, pp, 79-88 (1983)], la acetilación de la amina N-terminal, y la amidación de cualquier grupo carboxilo N-terminal.

Otro tipo de modificación covalente del polipéptido PRO5775 comprendida en el ámbito de la presente invención, incluye la "alteración del modelo de glicosilación nativa" del polipéptido. La alteración del patrón de glicosilación nativa se prevé para propósitos de delección de una o más porciones de carbohidratos que se encuentran en la secuencia nativa de PRO5775 (tanto por la eliminación de sitios de glicosilación subyacentes como por la delección de la glicosilación por medios químicos y/o enzimáticos), y/o la adición de uno o más sitios de glicosilación que no estén presentes en la secuencia nativa de PRO5775. Además, se pueden incluir cambios cualitativos en la glicosilación de las proteínas nativas, que impliquen cambios en la naturaleza y en la proporción de las diversas porciones de carbohidratos presentes.

También puede conseguirse la adición de sitios de glicosilación en el polipéptido PRO5775 por alteración de la secuencia de aminoácidos. La alteración se realiza, por ejemplo, por la adición de, o sustitución por uno o más residuos serina o treonina en la secuencia nativa de PRO5775, (para obtener sitios de O-glicosilación). Así mismo, la secuencia de aminoácidos PRO5775 puede opcionalmente alterarse por cambios a nivel del ADN, particularmente por la mutación del ADN que codifica el polipéptido PRO5775 en bases preseleccionadas y así los codones generados se traducirán en los aminoácidos deseados.

Otra manera de aumentar el número de porciones de carbohidratos en el polipéptido PRO5775 es por unión química o enzimática de glicósidos con el polipéptido. Estos procedimientos se describen en la materia, por ejemplo en la patente internacional WO 87/05.330 publicada el 11 Septiembre de 1987, y en Aplin y Wriston, CRC Crit Rev. Biochem., pp. 259-306 (1981).

5 También pueden extraerse las porciones de carbohidratos presentes en el polipéptido PRO5775 por tratamiento químico o enzimático o por substitución mutacional de codones que codifican los residuos de aminoácidos que sean dianas para la glicosilación. Las técnicas de desglicosilación química se conocen en la materia y se han describen, por ejemplo, por Hakimuddin, y col., Arch. Biochem. Biophys. 259:52 (1987) y por Edge y col., Anal. Biochem., 118:131
10 (1981). La rotura enzimática de porciones de carbohidratos de los polipéptidos se consigue mediante una variedad de endo- y exo-glicosilasas tal como describió Thotakura y col., Meth. Enzymol., 138:350 (1987).

Otro tipo de modificación covalente de PRO5775 comprende la unión del polipéptido PRO5775 a una variedad de polímeros no proteicos, por ejemplo, el polietilén glicol (PEG), el polipropilén glicol, los orpolioxaquilenos, en la
15 forma descrita en las patentes americanas U.S. No. 4.640.835; 4.301.144; 4.791.192 o 4.179.337.

También otra modificación posible de PRO5775 en la presente invención consiste en formar una molécula quimérica que comprende un PRO5775 fusionado a un polipéptido o una secuencia aminoácídica heterólogos.

20 La molécula quimérica puede a su vez incluir una fusión de PRO5775 con una etiqueta polipeptídica que proporciona un epitopo al que se une un anticuerpo anti-etiqueta. La etiqueta epitópica generalmente está colocada en el extremo amino o carboxilo de PRO5775 y la presencia de formas de PRO5775 etiquetado epitópicamente puede detectarse mediante un anticuerpo contra dicha etiqueta. Así mismo, la presencia de etiqueta epitópica permite purificar el PRO5775 por afinidad con el anticuerpo anti-etiqueta u otro tipo de matriz de afinidad que se una a la etiqueta
25 epitópica. Los diferentes polipéptidos etiquetados epitópicamente y sus anticuerpos respectivos son bien conocidos en la materia. Como ejemplos se incluyen las etiquetas de poli-histidina (poli-His) o poli-histidina-glicina (poli-His-gly); el polipéptido etiqueta del virus de la gripe HA y su anticuerpo 12CA5 [Field y col., Mol. Cell. Biol., 8:2159-2165 (1988)]; la etiqueta de c-myc y los anticuerpos contra ella, 8F9, 3C7, 6E10, E4, B7 y 9E10 [Evan y col., Molecular and Cellular Biology, 5:3610-3616 (1985)]; y la etiqueta de la glicoproteína D (gD) del virus Herpes simplex y su anticuerpo [Paborsky y col., Protein Engineering, 36(6):547-553 (1990)]. Otros polipéptidos etiqueta incluyen el péptido-Flag [Hopp y col., Bio-Technology, 6:1204 (1988)]; el péptido epitópico KT3 [Martin y col., Science, 255:192-194 (1992)]; y un epitopo del péptido de la α -tubulina [Skinner y col., J. Biol. Chem., 266:15163-15166 (1991)]; y la etiqueta
30 peptídica de la proteína 10 del gen T7 [Lutz-Freyermuth y col., Proc. Natl. Acad. USA, 87:6393-6397 (1990)].

35 La molécula quimérica puede alternativamente incluir una fusión de PRO5775 o con una inmunoglobulina o con una región particular de una inmunoglobulina. Para una forma bivalente de la molécula quimérica (también referida como "inmuno adhesina"), la fusión puede realizarse con una región FC de una molécula de IgG. Las fusiones Ig preferentemente incluyen la substitución de una forma soluble (un dominio transmembrana deleccionado o inactivado) de un polipéptido PRO5775 en lugar de al menos una región variable de la molécula de Ig. En una realización particularmente preferente, la fusión de la inmunoglobulina incluye la bisagra, y las regiones CH2 y CH3, o la bisagra y CH1, CH2 y CH3 de una molécula de IgG1. Para la producción de las fusiones de inmunoglobulina véase también, la patente americana U.S. 5.428.130 publicada en Junio 27 de 1995.

45 D. Preparación de Polipéptidos PRO5775

La descripción de más abajo se refiere a la producción primaria de PRO5775 por medio de un cultivo de célula transformadas o transfectadas con un vector que contiene el ácido nucleico que codifica PRO5775. Se contempla también, que puedan emplearse procedimientos alternativos bien conocidos en la materia, para la preparación de PRO5775. Por ejemplo, la secuencia de PRO5775 o porciones de la misma, pueden producirse por síntesis directa del
50 péptido mediante técnicas de fase sólida [véase, por ejemplo, Stewart y col., Solid-Phase Peptide Synthesis W.H. Freeman Co., San Francisco, Ca (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. La síntesis de proteínas *in vitro* puede realizarse mediante técnicas manuales o automáticas. La síntesis automática puede realizarse, por ejemplo, mediante un sintetizador de péptidos de Applied Biosystems (Foster City, CA) según las instrucciones del fabricante. Varias porciones de PRO5775 se sintetizan por separado químicamente y se combinan mediante procedimientos químicos o enzimáticos para producir el PRO5775 de longitud completa.

a. Aislamiento del ADN que codifica el Polipéptido PRO5775

60 El ADN que codifica PRO5775 puede obtenerse a partir de una librería de cADN preparada de un tejido que exprese el mRNA de PRO5775 a un nivel detectable. Por consiguiente el ADN del PRO5775 humano, se obtiene de una librería de cADN preparada de tejido humano, tal como se describe en los Ejemplos. También se obtiene PRO5775 de una librería genómica o mediante la síntesis de oligonucleótidos.

65 Las librerías puede cribarse con sondas (como por ejemplo, anticuerpos contra el polipéptido PRO5775, u oligonucleótidos de al menos 20-80 bases) diseñadas para identificar genes de interés o proteínas codificadas por éstos. El cribado del cADN o de la librería genómica con la sonda seleccionada se realiza mediante procedimientos estándares, como los descritos en Sambrook y col., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). Una manera alternativa de aislar el gen que codifica para PRO5775 es usar la metodología

ES 2 279 473 T3

de la PCR [Sambrook y col., *supra*; Dieffenbach y col., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Los ejemplos de más adelante, describen técnicas de cribado de librerías de cADN. Las secuencias de oligonucleótidos seleccionadas como sondas deben tener la suficiente longitud y carecer de ambigüedades con tal de minimizar los falsos positivos. Preferentemente, los oligonucleótidos se marcan para que poder detectar su hibridación con el ADN de la librería y realizar el cribado. Los procedimientos de marcado son bien conocidos en la materia e incluyen el uso de radiomarcadores como el ATP marcado con P³², la biotilación o el marcaje enzimático. Las condiciones de hibridación, incluyendo la astringencia moderada y alta, se proporcionan en Sambrook y col., *supra*.

Mediante los procedimientos de cribado, se identifican en la librería, secuencias que pueden ser comparadas y alineadas con otras secuencias que están depositadas y disponibles en los bancos de datos públicos como el GenBank u otras bases de datos de secuencias privadas. La identidad de la secuencia (tanto a nivel de aminoácido como del nucleótido) en las regiones definidas de la molécula o a lo largo de la secuencia de longitud completa pueden determinarse mediante procedimientos bien conocidos en la materia, y tal como se describe aquí.

El ácido nucleico que tiene la secuencia que codifica la proteína puede obtenerse por cribado de librerías de cADN o genómicas a partir de la secuencia aminoacídica deducida y descrita aquí por primera vez, y si es necesario, mediante cebadores convencionales en procedimientos de extensión, tal como se describe en Sambrook y col., *supra*, para detectar precursores e intermediarios del mRNA que no se hallan transcrito de forma inversa a cADN.

b. Selección y transformación de células huéspedes

Las células huéspedes se transforman o transfectan con los vectores de clonaje y expresión descritos aquí para PRO5775. Su producción y cultivo se realiza en medios convencionales con los nutrientes modificados según sea apropiado para inducir promotores, seleccionar transformantes, y amplificar los genes que codifican las secuencias deseadas. Las condiciones de cultivo, tipo de medio, temperatura, pH y demás condiciones, se seleccionan por personas expertas en la materia sin excesiva experimentación. En general, principios, protocolos, y técnicas prácticas para maximizar la productividad de los cultivos celulares se pueden encontrar en Mammalian Cell Biotechnology: A Practical Approach. Butler, ed. (IRL Press, 1991) y Sambrook y col., *supra*.

Los procedimientos para las transfecciones de células eucariotas y las transformaciones de células procariotas son conocidos por los entendidos en la materia, por ejemplo aquellos mediados por CaCl₂, CaPO₄, liposomas y electroporación. Según el huésped utilizado, se escogen las técnicas estándares apropiadas para las células. Generalmente para procariotas se utiliza el tratamiento con calcio que emplea cloruro de calcio, tal como se describe en Sambrook y col., *supra*, o la electroporación. La infección con *Agrobacterium tumefaciens* se utiliza para la transformación de ciertas células de plantas tal como se describe en Shaw y col., Gene, 23:315 (1983) y la patente internacional WO 89/05859 publicada el 29 de Junio de 1989. Para células de mamíferos sin paredes celulares, se puede utilizar el procedimiento de la precipitación con fosfato cálcico de Graham y van der Eb, Virology, 52:456-457 (1978). Los aspectos generales de los sistemas de transfección de células huéspedes de mamíferos se han descrito en la patente americana U.S. 4.399.216. Las transformaciones en levadura se llevan a cabo típicamente de acuerdo con el procedimiento de Van Solingen y col., J. Bact., 130:946 (1977) y Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). Sin embargo, se pueden utilizar otros procedimientos para introducir el ADN en las células, tal como la microinyección nuclear, la electroporación, la fusión de protoplastos bacterianos con células intactas, o policondones, por ejemplo polibreno y piornitina. Para varias técnicas de transformación de células de mamífero, véase Keown y col., Methods in Enzymology, 185:527-537 (1990) y Mansour y col., Nature, 336:348-352 (1988).

Las células huéspedes adecuadas incluyen las procariotas, las levaduras o las eucariotas, para el clonaje o la expresión del ADN en vectores. Las células procariotas adecuadas incluyen pero no se limitan a las eubacterias, los organismos Gram-negativos o Gram-positivos, por ejemplo, las Enterobacteriaceas como las cepas de *E. coli* de disponibilidad pública como la cepa de *E. coli* K12 MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* 31,537; la cepa *E. coli* W3110 (ATCC 27,325) y la cepa de *E. coli* K5 772 (ATCC 53,635). Otras células huéspedes procariotas adecuadas incluyen las enterobacteriaceas como *Escherichia*, por ejemplo, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, por ejemplo, *Salmonella typhimurium*, *Serratia*, por ejemplo, *Serratia marcesans* y *Shigella*, así como *Bacilli* como *B. subtilis* y *B. licheniformis* (*i.e.*, *B. licheniformis* 41P descrito en la patente DD 266.710 publicada el 12 de Abril de 1989), *Pseudomonas* tal como *P. aeruginosa*, y *Streptomyces*. Estos ejemplos son ilustrativos más que limitantes. La cepa W3110 es un huésped particularmente preferido o huésped parental porque es una cepa huésped común para fermentaciones de productos del ADN recombinante. Preferentemente, la célula huésped secreta cantidades mínimas de enzimas proteolíticas. La cepa W3110, puede ser modificada por efecto de mutaciones genéticas en los genes que codifican las proteínas endógenas del huésped, como ejemplos, de tales huéspedes se incluyen la cepa 1A2 de *E. coli* W3110, que tiene el genotipo completo *tonA*; la cepa 9E4 de *E. coli* W3110 que tiene el genotipo completo *tonA ptr3*; la cepa 27C7 de *E. coli* W3110 (ATCC 55,244) que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac)169 degP onepT kan'*; la cepa 37D6 de *E. coli* W3110 que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan'*; la cepa 40B4 de *E. coli* W3110 que es una cepa derivada de 37D6 con una mutación por delección *degP* que comporta la no resistencia a la kanamicina; y una cepa *E. coli* que tiene una proteasa periplasmática mutante descrita en la patente americana U.S. 4.946.783 publicada el 7 de Agosto de 1990. Alternativamente, también son adecuados los procedimientos de clonaje *in vitro*, por ejemplo, la PCR u otras reacciones de la polimerasa de ácidos nucleicos.

Además de los procariotas, algunos microbios eucariotas como los hongos filamentosos o las levaduras son adecuados como huéspedes para el clonaje o expresión de vectores que codifican PRO5775. *Saccharomyces cerevisiae* es un eucariota inferior utilizado comúnmente como microorganismo huésped. Entre otros se incluyen *Schizosaccharomyces pombe* (Beach y Nurse, Nature, 290:140 [1981]; patente europea EP 139,383 publicada el 2 de Mayo de 1985); el huésped *Kluyveromyces* (Patente U.S. Mo. 4.943.529; Fleer *et al.*, Bio/Technology, 9:968-975 (1991)) tal como, por ejemplo *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt y col., J. Bacteriol., 737 [1983]). *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilorum* (ATCC 36,906; Vanden Berg y col., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, y *K. marxianus*; yarrowia (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna y col., J. Basic. Microbiol., 28:265-278 [1988]; *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case y col., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]; *Schwanniomyces* tal como *Schwanniomyces occidentalis* (patente europea EP 394.538 publicada el 31 de Octubre de 1990); y hongos filamentosos tales como por ejemplo, *Neurospora*, *Penicilium*, *Tolypocladium* (patente internacional WO 91/00357 publicadaz el 10 de Enero de 1991), y *Aspergillus* como *A. nidulans* (Ballance y col., Biochem. Biophys. Res. Commun. 112:284-289 [1983]; Tiburn *et al.*, Gene, 26; 205-221 [1983]; Yelton y col., Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) y *A. niger* (Keelly y Hynes, EMBO J., 4:475-479 [1985]). Las levaduras metilotróficas son adecuadas aquí e incluyen, pero no se limitan a, las levaduras capaces de crecer en metanol seleccionadas a partir del género de *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis* y *Rhodotorula*. Una lista de especies específicos que son ejemplos de esta clase de levaduras se puede encontrar en C. Anthony, The Biochemistry of Methylotrrophs, 269 (1982).

Para la expresión de PRO5775 glicosilado las células huéspedes adecuada derivan de organismos multicelulares. Como ejemplos de células de invertebrados se incluyen las células de insectos como las de *Drosophila* S2 y *Spodoptera* Sf9, así como las células de plantas. Ejemplos de líneas celulares huéspedes de mamíferos de utilidad son las células de ovario de hámster chino (CHO) y las células COS. Más ejemplos específicos incluyen la línea de riñón de mono CV1 transformada con SV40 (COS-7, ATCC, CRL 1651); la línea de riñón embrionario humano (293 o células 293 subclonadas para el crecimiento de cultivos en suspensión, Graham y col., J. Gen. Virol. 36:59 (1977)); células de ovario de hámster chino (-DHFR (CHO), Urlaub y Chasin, Proc. Natl. Acad. Sci. USA, 77; 4216 (1980)); células de sertoli de ratón (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065); y un tumor mamario de ratón (MMT 060562, ATCC CCL51). La selección de una célula huésped apropiada es competencia del experto en la materia.

c. Selección y utilización de un vector replicable

El ácido nucleico (por ejemplo el cADN o el ADN genómico) que codifica PRO5775 puede insertarse en un vector replicable para su clonaje (amplificación del ADN) o para su expresión. Varios vectores son de disponibilidad pública. El vector puede ser un plásmido, cósmido, partícula viral o fago. La secuencia de ácidos nucleicos apropiada puede insertarse en el vector mediante diversos procedimientos. En general, el ADN se inserta en una(s) diana(s) de reconocimiento por nucleasas de restricción utilizando técnicas conocidas en la materia. Los componentes del vector generalmente incluyen, pero no se limitan a, uno o más secuencia señal, un origen de replicación, uno o más genes marcadores, un elemento intensificador de la transcripción, un promotor, y una secuencia terminadora de la transcripción. Para la construcción de vectores adecuados que contengan uno o más de esos componentes, se emplean técnicas estándares de ligación que son conocidas por los expertos en la materia.

El PRO5775 puede producirse de forma recombinante no sólo directamente, sino también como un polipéptido de fusión con un polipéptido heterólogo, que puede ser una secuencia señal u otro polipéptido con una diana de corte específica en el extremo N-terminal de la proteína o polipéptido maduro. En general, la secuencia señal puede ser un componente del vector, o puede ser una parte del ADN que codifica el PRO-5775 y que se inserta en el vector. La secuencia señal puede ser una secuencia señal procariota seleccionada, por ejemplo de entre el grupo de la fosfatasa alcalina, la penicilanasa, el lpp, o la enterotoxina II estable al calor. Para la secreción en levaduras, la secuencia señal puede ser, por ejemplo, la secuencia líder de la invertasa de levadura, la secuencia líder del factor alfa de *Saccharomyces* y del factor α de *Kluyveromyces*, la última descrita en la patente americana U.S. 5.010.182) o la secuencia de la fosfatasa ácida, la secuencia de la fluoamilasa *C. albicans* (patente europea EP 362.179 publicada el 4 de Abril de 1990), o la secuencia señal descrita en la patente internacional WO 90/13646 publicada el 15 de Noviembre de 1990. Para la expresión en células de mamífero pueden utilizarse secuencias señal de mamífero para dirigir la secreción de la proteína, como las secuencias señal de polipéptidos secretados de las mismas especies o relacionadas, por ejemplo los líderes secretores virales.

Tanto los vectores de expresión como de clonaje, contienen una secuencia de ácido nucleico que permite al vector replicarse en una o más células huéspedes seleccionadas. Estas secuencias son bien conocidas para una variedad de bacterias, levaduras y virus. El origen de replicación del plásmido pBR322 es adecuado para la mayoría de las bacterias Gram-negativas, el origen del plásmido 2 μ es adecuado para las levaduras y varios orígenes virales (SV40, poliomias, adenovirus, VSV, BPV) son adecuados para el clonaje de vectores en células de mamíferos.

Los vectores de expresión y clonaje contienen típicamente un gen de selección, también denominado marcador seleccionable. Los genes de selección más corrientes incluyen proteínas que (a) confieren resistencia a antibióticos u otras toxinas, por ejemplo ampicilina, nenomicina, metotrexate, o tetraciclina, (b) complementan deficiencias auxotróficas, o (c) aportan nutrientes críticos no disponibles en los medios complejos, por ejemplo, el gen que codifica la D-alanina racemasa de *Bacilli*.

Un ejemplo de marcadores seleccionables adecuados para células de mamífero son aquellos que permiten la identificación de las células competentes que han incorporado el ácido nucleico que codifica PRO5775, como el DHFR o la timidina quinasa. Una célula huésped adecuada cuando se emplea el DHFR de tipo salvaje es la línea celular CHO deficiente en la actividad DHFR, que se prepara y propaga tal como se describe en Urlaub y col., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). Un gen de selección adecuado para su uso en levaduras es el gen *trp1* presente en el plásmido de levadura YRp7 [Stinchcom y col., Nature, 282:39 (1979); Kingsman y col., Gene, 7:141 (1979); Tschemper y col., Gene, 10:157 (1980)]. El gen *trp1* proporciona un marcador de selección para la cepa mutante de levadura que carece de la capacidad para crecer en triptófano. ATCC No. 44076 o PEP4-1 [Jones, Genetics, 85:12 (1977)].

Los vectores de expresión y clonaje contienen normalmente un promotor unido operativamente a la secuencia de ácido nucleico que codifica PRO5775 para dirigir la síntesis del mRNA. Se pueden reconocer los promotores en una gran variedad de células huéspedes potenciales conocidas. Entre los promotores adecuados para su uso con los huéspedes procariotas se incluyen los promotores de: la β -lactamasa y el sistema promotor de la lactosa [Chan y col., Nature, 275:615 (1978); Goeddel y col., Nature, 281:544 (1979)], la fosfatasa alcalina, el sistema promotor del triptófano (*trp*) [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], y promotores híbridos como el promotor *tac* [deBoer Dalgarno (S.D.) secuencias operativamente unidas al ADN que codifica el PRO5775].

Entre los ejemplos de secuencias promotoras adecuadas para su uso en huéspedes de levadura se incluyen los promotores para la 3-fosfoglicerato quinasa [Hitzemanm y col., J. Biol. Chem., 255:2073 (1980)] u otras enzimas glicolíticas [Hess y col., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], tales como la enolasa, la gliceraldehído-3-fosfato dehidrogenasa, la hexoquinasa, la piruvato decarboxilasa, la fosfofructoquinasa, la glucosa-6-fosfato isomerasa, la 3-fosfatoglicerato mutasa, la piruvato quinasa, la triosefosfato isomerasa, la fosfoglucosa isomerasa, y la glucoquinasa.

Otro promotores de levaduras, son los promotores inducibles que tienen la ventaja adicional de controlar la transcripción según las condiciones de crecimiento, entre ellos están las regiones promotoras para la alcohol dehidrogenasa-3, el isocitocromo C, la fosfatasa ácida, enzimas degradadoras asociadas con el metabolismo del nitrógeno, la metalotioneína, la gliceraldehído-3-fosfato dehidrogenasa, y las enzimas responsables de la utilización de maltosa y galactosa. Los vectores y promotores adecuados para su uso en la expresión en levaduras se describen más extensamente en la patente europea EP 73.657.

La transcripción de vectores que poseen PRO5775 en células huéspedes de mamíferos se controla, por ejemplo, mediante promotores obtenidos de genomas víricos como el virus del polioma, el virus de la viruela aviar (patente inglesa UK 2.211.504 publicada el 5 de Julio de 1989), del adenovirus, (como el Adenovirus 2), del virus del papiloma bovino, del virus del sarcoma aviar, el citomegalovirus, el retrovirus, el virus de la hepatitis-B y del virus del simio 40 (SV40), o los promotores heterólogos de mamíferos que proporcionan promotores que son compatibles con los sistemas de la célula huésped, por ejemplo, el promotor de la actina o el de la inmunoglobulina y los promotores de choque térmico.

La transcripción del ADN que codifica PRO5775 en eucariotas superiores puede aumentarse por la inserción de una secuencia intensificadora en el vector. Los intensificadores son elementos que actúan en cis sobre el ADN, normalmente alrededor de 10 a 300 bp, y actúan en un promotor para aumentar su transcripción. Muchas secuencias intensificadoras son conocidas en los genes de mamíferos (globina, elastasa, albúmina, α -fetoproteína e insulina). Sin embargo, se usa frecuentemente una secuencia intensificadora de un virus de célula eucariota. Como ejemplos se incluyen la secuencia intensificadora del origen tardío de la replicación del virus SV40 (bp 100-270), la secuencia intensificadora del promotor temprano de citomegalovirus, la secuencia intensificadora del origen tardío de replicación del polioma y las secuencias intensificadoras de adenovirus. La secuencia intensificadora puede introducirse en el vector en posición 5' o 3' de la secuencia codificante de PRO5775, pero preferentemente en el extremo 5' del promotor.

Los vectores de expresión utilizados en células huéspedes eucariotas (levaduras, hongos, insectos, plantas, animales, humanos, o células nucleadas de otros organismos multicelulares) contienen también secuencias necesarias para la finalización de la transcripción y para la estabilización del mRNA. Estas secuencias están dispuestas comúnmente desde el extremo 5' y ocasionalmente desde el extremo 3' de las regiones no traducidas de los ADNs o cADNs eucariotas o virales. Estas regiones contienen secuencias de nucleótidos transcritas como fragmentos poliadenilados en la porción 3' no traducida del mRNA que codifica PRO5775.

Otros procedimientos, vectores y células huéspedes adecuados para su adaptación a la síntesis de PRO5775 en cultivos de células recombinantes de vertebrados se describen en Gething y col., Nature, 293:620-625 (1981); Mantei y col., Nature, 281:40-46 (1979); patente europea EP 117.060; y patente europea EP 117.058.

d. Detección de la Amplificación/Expresión génicas

La amplificación y/o expresión pueden cuantificarse directamente en la muestra, por ejemplo, mediante transferencia de Southern convencional, transferencia de Northern para cuantificar la transcripción de mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], transferencia de mancha (para análisis del ADN), o hibridación *in situ* mediante sondas marcadas apropiadas basadas en las secuencias proporcionadas aquí. De forma alternativa, se pueden utilizar anticuerpos para el reconocimiento de dúplex específicos, incluyendo el dúplex de ADN, dúplex de ARN, y dúplex híbridos ADN-ARN o ADN-proteína. Los anticuerpos a su vez pueden marcarse y el ensayo llevarse a cabo

con el dúplex unido a una superficie, de forma que después de la formación del dúplex en la superficie, se pueda detectar la presencia del anticuerpo unido al dúplex.

5 La expresión génica también puede cuantificarse por procedimientos inmunológicos, como la tinción inmunohistoquímica de células y de secciones de tejidos y por ensayo de cultivo celular o fluidos corporales para la cuantificación directa de la expresión del producto génico. Los anticuerpos adecuados para la tinción inmunohistoquímica y/o ensayo de muestras de fluido, pueden ser tanto monoclonales como policlonales y pueden prepararse de cualquier mamífero. Convenientemente, los anticuerpos se preparan contra una secuencia nativa del polipéptido PRO5775 o contra un péptido sintético basado en las secuencias de ADN proporcionadas en la presente invención o contra secuencias exógenas unidas al ADN de PRO5775 y que codifican un epitopo específico del anticuerpo.

IV. Purificación del Polipéptido

15 Las formas diversas del RPO5775 pueden recuperarse del medio de cultivo o de los lisados de las células huéspedes. Si está unido a membrana, debe liberarse de la membrana mediante una solución detergente adecuada (por ejemplo, Tritón-X-100) o por digestión enzimática. Las células empleadas en la expresión de PRO5775 se pueden romper por varios procedimientos físicos o químicos, como por ejemplo ciclos de congelación, sonicación, disrupción mecánica o agentes que provocan la lisis celular. En el caso de requerirse la purificación de PRO5775 a partir de proteína recombinantes celulares o polipéptidos, un ejemplo de los procedimientos a seguir se detallan a continuación: por fraccionamiento en columnas de intercambio iónico; precipitación en etanol; HPLC de fase inversa, cromatografía en sílice o resina de intercambio catiónico como DEAE; cromatografía de exclusión; SDS-PAGE; precipitación con sulfato amónico; filtración en gel, utilizando por ejemplo Sephadex G-75; columnas de proteína A Sepharose para extraer contaminantes como las IgG; y columnas de quelantes de metales que unen las formas etiquetadas de PRO5775. Se pueden emplear varios procedimientos de purificación de proteínas que son bien conocidos en la materia y se han descrito por ejemplo en Deutscher, *Methods in Enzymology*, 182 (1990); Scopes, *Purificación de Proteínas: Principios y Práctica*, Springer-Verlag, New York, (1982). Los pasos de purificación que se seleccionen, dependerán, por ejemplo, de la naturaleza del proceso de producción utilizado y del PRO5775 particular producido.

E. Amplificación de Genes que codifican los polipéptidos PRO5775 en Tejidos TumORAles y Líneas Celulares

30 La presente invención se basa en la identificación y caracterización de genes que se amplifican en ciertas células cancerosas.

35 El genoma de organismos procariotas y eucariotas está sometido aparentemente a dos requerimientos en conflicto. El primero es la preservación de la propagación del ADN como información genética en su forma original, para garantizar la herencia estable a través de múltiples generaciones. Y el segundo, que las células o los organismos han de ser capaces de adaptarse a cambios ambientales constantes. Los mecanismos adaptativos pueden incluir modificaciones cualitativas o cuantitativas de su material genético. Las modificaciones cualitativas incluyen las mutaciones de ADN, en las que las secuencias codificantes son alteradas resultando en una proteína estructural o funcionalmente diferente. 40 La amplificación génica es la modificación cuantitativa, por la cual el número actual de secuencia codificante completa, es decir un gen, aumenta, llevando a un aumento en el número de moldes disponibles para la transcripción, un número aumentado de transcritos traducibles y, finalmente, a un aumento en la abundancia de la proteína codificada por el gen amplificado.

45 El fenómeno de la amplificación y sus mecanismos subyacentes han sido investigados *in vitro* en algunos sistemas de cultivo de procariotas y eucariotas. El ejemplo mejor caracterizado de amplificación génica se refiere al cultivo de células eucariotas en medio que contienen concentraciones variables de un compuesto citotóxico, como el metotrexate (MTX). El MTX es un análogo del ácido fólico que interfiere en la síntesis de ADN al bloquear la enzima dihidrofolato reductasa (DHFR). Durante la exposición inicial a bajas concentraciones de MTX, la mayoría de las células (>99,9%) morirán. Un pequeño número de células sobrevive y son capaces de crecer en concentraciones en aumento de MTX al producir cantidades mayores de ARN y proteína DHFR. La base de esta sobreproducción es la amplificación de un gen único, el DHFR. Las copias adicionales del gen se encuentran en copias extracromosomales en la forma de pequeños, cromosomas supernumerarios (dobles minutos) o como copias integradas en los cromosomas.

55 La amplificación génica conlleva generalmente el desarrollo de la resistencia a compuestos citotóxicos (antibióticos para bacterias y agentes quimioterapéuticos para células eucariotas) y la transformación neoplásica. La transformación de células eucariotas como un acontecimiento espontáneo o debido a una exposición viral o químico/ambiental está asociada típicamente a cambios en el material genético de esas células. Uno de los cambios genéticos observados comúnmente en los procesos malignos en humanos son las mutaciones de la proteína p53. La proteína p53 controla 60 la transición de células de la fase estacionaria (G1) a la fase replicativa (S) y evita esta transición en presencia de daño en el ADN. En otras palabras, una de las consecuencias principales de la inutilización de p53 por mutación es la acumulación y propagación del daño en el ADN, por ejemplo, cambios genéticos. Los tipos de cambios genéticos más comunes en células neoplásicas son, además de las mutaciones puntuales, las amplificaciones y las alteraciones estructurales mayores, como son las translocaciones.

65 La amplificación de las secuencias de ADN puede indicar requerimientos específicos funcionales ilustrados en el sistema experimental de DHFR. Por tanto, la amplificación de ciertos oncogenes en los procesos malignos indicaría un papel causal de estos genes en el proceso de transformación maligna y mantenimiento del fenotipo transformado. Esta

hipótesis ha ganado soporte en estudios recientes. Por ejemplo, la proteína bcl-2 se ha hallado amplificada en ciertos tipos de linfoma de no Hodgkin. Esta proteína inhibe la apoptosis y lleva a la acumulación progresiva de células neoplásicas. Los miembros de esta familia génica de receptores de factores de crecimiento se han encontrado amplificados en varios tipos de cánceres y se ha sugerido que su sobreexpresión puede convertir las células neoplásicas en menos susceptibles a cantidades limitantes de factores de crecimiento disponible. Los ejemplos incluyen la amplificación del receptor de andrógenos en el cáncer de próstata recurrente durante la terapia de privación de andrógenos y la amplificación del receptor de factor de crecimiento homólogo, ERB2, en cáncer de mama. Últimamente, los genes implicados en la señalización intracelular y el control de la progresión del ciclo celular sufrir amplificación durante la transformación maligna. Esto queda ilustrado por la amplificación de bcl-I y ras en varias neoplasias epiteliales o linfoides.

Estos primeros estudios ilustran la posibilidad de identificar las secuencias de ADN amplificadas en neoplasias, este enfoque puede identificar genes importantes para la transformación maligna. El caso de ERB2 también demuestra la posibilidad de un punto de partida terapéutico, ya que las proteínas transformantes pueden representar dianas novedosas y específicas para la terapia tumoral.

Se utilizan diferentes técnicas para demostrar la amplificación de secuencias genómicas. El análisis clásico citogenético de preparaciones cromosómicas obtenidas a partir de células cancerosas es adecuado para la identificación de alteraciones estructurales mayores, como translocaciones, deleciones e inversiones. Las regiones genómicas amplificadas pueden visualizarse sólo si están implicadas regiones amplias con un número alto de copias o si está presente material extracromosómico. La citogenética ha sido la primera técnica en demostrar la asociación consistente en cambios cromosómicos específicos con neoplasias particulares pero no es adecuada para la identificación y el aislamiento de secuencias manejables de ADN. La técnica desarrollada más recientemente de la hibridación genómica comparativa (CGH) ha podido ilustrar el fenómeno extendido de la amplificación genómica en las neoplasias. En ésta, el ADN tumoral o normal se hibrida simultáneamente en las metafases de células normales y todo el genoma puede ser cribado mediante análisis de imagen para detectar las secuencias de ADN que están presentes en el tumor con una frecuencia aumentada. (Patente internacional WO93/18.186); Gray y col., Radiation Res., 137:275-289 [1994]. Como procedimiento de cribado, este tipo de análisis ha revelado un gran número de amplicones recurrentes (tramo de ADN amplificado) en una variedad de neoplasias humanas. Aunque la CGH es más sensible que el análisis citogenético clásico en la identificación de los tramos de ADN amplificados, no permite una identificación rápida y aislamiento de secuencias codificantes en el amplicón por las técnicas estándares de genética molecular.

Los procedimientos más sensibles para detectar la amplificación génica son los ensayos basados en la reacción en cadena de la polimerasa (PCR). Estos ensayos utilizan una cantidad muy pequeña de ADN tumoral como material de partida, son exquisitamente sensibles, proporcionan ADN que es susceptible de un análisis ulterior, por ejemplo por secuenciación, y son adecuados para el análisis de volúmenes con gran rendimiento.

Los ensayos arriba mencionados no son mutuamente excluyentes y frecuentemente se usan en combinación para identificar amplificaciones en neoplasias. Mientras el análisis citogenético y la CGH representan procedimientos de cribado para estudiar regiones amplificadas en el genoma entero, los ensayos basados en la PCR son los más adecuados para la identificación final de las secuencias codificantes, es decir los genes en las regiones amplificadas.

De acuerdo con la presente invención, tales genes han sido identificados por PCR cuantitativa (S. Gelmini y col., Clin. Chem., 43:752 [1997]), por comparación del ADN de una variedad de tumores primarios, incluyendo mama, pulmón, colon, próstata, cerebro, hígado, riñón, páncreas, bazo, timo, testículo, ovario, útero, etc., tumor o líneas celulares tumorales, con el ADN combinado de donantes sanos. La PCR cuantitativa se lleva a cabo en un instrumento TaqMan™ (ABI). Los cebadores específicos de los genes y las sondas fluorogénicas se diseñan en función de las secuencias codificantes de los ADNs.

Las líneas celulares de carcinoma de pulmón humano incluyen A549 (SRCC768), Calu-1 (SRCC769), Calu-6 (SRCC770), H157 (SRCC771), H441 (SRCC772), H460(SRCC773), SKMES-1(SRCC774), SW900(SRCC775), H522(SRCC832)y H810(SRCC833), están todas disponibles en el ATCC. Las células primarias de tumores humanos de pulmón normalmente derivan de adenocarcinomas, carcinomas de célula escamosa, carcinomas de célula grande, carcinomas de célula no pequeña, carcinomas de célula pequeña, y carcinomas bronco alveolares, e incluyen, por ejemplo, SRCC724 (adenocarcinoma, abreviado como "AdenoCa")(LT1), SECC725(carcinoma de célula escamosa, abreviado como "SqCCA")(LT1A), SRCC726 (adenocarcinoma)(LT2), SECC727 (adenocarcinoma)(LT3), SRCC728 (adenocarcinoma)(LT4), SRCC729 (carcinoma de célula escamosa)(LT6), SECC730 (carcinoma de célula adeno/escamosa)(LT7), SRCC731 (adenocarcinoma)(LT9), SRCC732 (carcinoma de célula escamosa)(LT10), SRCC733 /carcinoma de célula escamosa)(LT11), SRCC734 (adenocarcinoma)(LT12), SRCC735 (carcinoma de célula adeno/escamosa) (LT13), SRCC736 (carcinoma de célula escamosa)(LT15), SRCC737 (carcinoma de célula escamosa)(LT16), SRCC738 (carcinoma de célula escamosa)(LT17), SRCC739 (carcinoma de célula escamosa)(LT18), SRCC740 (carcinoma de célula escamosa)(LT19), SRCC741 (carcinoma de célula de pulmón, abreviado como "LC-Ca")(LT21), SRCC811 (adenocarcinoma)(LT22), SRCC825 (adenocarcinoma)(LT8), SRCC886 (adenocarcinoma) (LT25), SRCC887 (carcinoma de célula escamosa) (LT26), SRCC888 (carcinoma adeno-BAC) (LT27), SRCC889 (carcinoma de célula escamosa)(LT28), SRCC890 (carcinoma de célula escamosa)(LT29), SRCC891 (adenocarcinoma)(LT30), SRCC892 (carcinoma de célula escamosa)(LT31), SRCC894 (adenocarcinoma) (LT33). También se incluyen tumores de pulmón humano denominados como SRCC1125 [HF-000631], SRCC1127 [HF-000641], SRCC1129 [HF-000643], SRCC1133[HF-000840], SRCC1135[HF-000842], SRCC1227[HF-001291], SRCC1229[HF-001293],

ES 2 279 473 T3

SRCC11230[HF-001294], SRCC1231[HF-001295], SRCC1232[HF-001296], SRCC1233[HF-001297], SRCC1235 [HF-001299] y SRCC1236[HF-001300].

Las líneas celulares de cáncer de colon incluyen, por ejemplo, las líneas celulares de la ATCC SW480 (adenocarcinoma, SRCC776), SW620 (metástasis de nódulos linfáticos de adenocarcinoma de colon, SRCC777), Colo320 (carcinoma, SRCC778), HT29 (adenocarcinoma, SRCC779), HM7 (una variante de la línea celular de adenocarcinoma de colon del ATCC altamente productora de mucina, SRCC780, obtenida del Dr. Robert Warren, UCSF), CaWiDr (adenocarcinoma, SRCC781 HCT116 (carcinoma, SRCC782), SKCO1 (adenocarcinoma, SRCC783), SW403 (adenocarcinoma, SRCC784), LS174T (carcinoma, SRCC785), Colo205 (carcinoma, SRCC828), HCT15 (carcinoma, SRCC829), HCC2998 (carcinoma, SECC830), y KM12 (carcinoma, SRCC831). Los tumores de colon primarios incluyen los adenocarcinomas de colon denominados como CT2 (SRCC742), CT3 (SRCC743), CT8 (SRCC744), CT10 (SRCC754), CT12 (SRCC746), CT14 (SRCC747), CT15 (SRCC748), CT16 (SRCC749), CT17 (SRCC750), CT1 (SRCC751), CT4 (SRCC752), CT5 (SRCC753), CT6 (SRCC754), CT7 (SRCC755), CT9 (SRCC756), CT1 (SRCC757), CT18 (SRCC758), CT19 (adenocarcinoma, SRCC906), CT20 (adenocarcinoma, SRCC907), CT21 (adenocarcinoma, SRCC908), CT22 (adenocarcinoma, SRCC909), CT23 (adenocarcinoma, SRCC910), CT24 (adenocarcinoma, SRCC911), CT25 (adenocarcinoma, SRCC912), CT26 (adenocarcinoma, SRCC913), CT27 (adenocarcinoma, SRCC914), CT28 (adenocarcinoma, SRCC915), CT29 (adenocarcinoma, SRCC916), CT30 (adenocarcinoma, SRCC917), CT31 (adenocarcinoma, SRCC918), CT32 (adenocarcinoma, SRCC919), CT33 (adenocarcinoma, SRCC920), CT35 (adenocarcinoma, SRCC921), y CT36 (adenocarcinoma, SRCC922). También se incluyen los derivados de tumores de colon humanos denominados como SRCC 1051 [HF-000499], SRCC1052 [HF-000539], SRCC1053 [HF-000575], SRCC1054 [HF-000698], SRCC1059 [HF-000755], SRCC1060 [HF-000756], SRCC 1142 [HF-000762], SRCC1144 [HF-000789], SRCC1146 [HF-000795] y SRCC1148 [HF-000811].

Las líneas celulares de carcinoma de mama humano incluyen, por ejemplo, HBL100 (SRCC759), MB435s (SRCC760), T47D (SRCC761), MB468 (SRCC762), MB 175 (SRCC763), MB361 (SRCC764), BT20 (SRCC765), MCF7 (SRCC760) y SKBR3 (SRCC767), y de tumores de mama humano designados como SRCC1057 [HF-000545]. También se incluyen los tumores de mama humanos designados como SRCC1094, SRCC 1095, SRCC1096, SRCC1097, SRCC1098, SRCC1099, SRCC1100, SRCC1101, y tumor de mama-met-pulmón humano designado como SRCC893 [LT32].

Los tumores de recto humano incluyen SRCC989 [HF-000550] y SRCC982 [HF-000551].

Los tumores de riñón humano incluyen SRCC989 [HF-000611] y SRCC1014 [HF-000613].

Los tumores de testículo humano incluyen SRCC1001 [HF-000733] y el tumor marginal del testículo SRCC999 [HF-000716].

Los tumores de paratiroides humana incluyen SRCC1002 [HF-000831] y SRC1003 [000832].

Los tumores de nódulo linfático humano incluyen el SRCC 1004 [HF-000854], SRCC1005 [HF-000855] y SRCC1006 [000856].

F. Distribución tisular

Los resultados de los ensayos de amplificación génica de la presente invención pueden ser verificados posteriormente mediante otros análisis, tal como la determinación de la expresión de mRNA en varios tejidos humanos.

Como se ha indicado antes, la amplificación génica o/y la expresión génica en varios tejidos puede cuantificarse por técnicas convencionales como la transferencia de Southern, la transferencia de Northern para cuantificar la expresión del mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 [1980]), la transferencia de mancha (análisis del ADN), o la hibridación *in situ*, utilizando una sonda adecuada, basada en las secuencias que se proporcionan aquí. De forma alternativa, los anticuerpos pueden emplearse para el reconocimiento de dúplex específicos, incluidos los dúplex de ADN, de ARN, de híbridos ADN-ARN o de ADN-proteína.

La expresión génica en varios tejidos también puede cuantificarse mediante procedimiento inmunológicos, como la tinción inmunohistoquímica de secciones de tejido y el ensayo del cultivo celular o fluidos corporales, para cuantificar directamente la expresión del producto génico. Los anticuerpos útiles para la tinción inmunohistoquímica y/o el ensayo en muestras de fluidos pueden ser tanto monoclonales como policlonales, y pueden prepararse en cualquier mamífero. Convenientemente, los anticuerpos pueden prepararse contra la secuencia nativa del polipéptido PRO5775 o contra el péptido sintético basado en las secuencias proporcionadas aquí o contra la secuencia exógena unida a la secuencia de ADN PRO5775 y que codifica para el epitopo del anticuerpo específico. Las técnicas generales, para generar anticuerpos, y protocolos especiales para la transferencia de Northern e hibridación *in situ* se proporcionan aquí más abajo.

G. Mapeo cromosómico

Si la amplificación de un gen dado es relevante funcionalmente, el gen debería amplificarse más que las regiones genómicas vecinas que no son tan importantes para la supervivencia del tumor. Para su comprobación, el gen puede ser localizado en un cromosoma determinado, por ejemplo, mediante el ensayo de híbridos de radiación. El nivel

de amplificación se determina entonces en la localización identificada, y en las regiones vecinas. La amplificación selectiva o preferente en la región genómica en la que el gen se ha localizado es consistente con la posibilidad de que la amplificación génica observada promueva el crecimiento tumoral o su supervivencia. El mapeo cromosómico incluye tanto la región de trabajo determinada como el epicentro. Para más detalles véase, por ejemplo, Stewart y col., *Genome Research*, 7:422-433 (1997).

H. Estudios de unión del anticuerpo

Los resultados del estudio de amplificación génica pueden además verificarse mediante estudios de unión de anticuerpos, en los que se analiza la capacidad de los anticuerpos anti-PRO577 para inhibir la expresión de los polipéptidos PRO5775 en las células del tumor (cáncer). Ejemplos de anticuerpos son el policlonal, monoclonal, humanizado, biespecífico, y heteroconjugado, la preparación de los cuales se describe aquí más adelante.

Los resultados del estudio de amplificación génica pueden además llevarse a cabo con cualquier procedimiento de ensayo conocido, como los ensayos de unión competitiva, ensayos en sándwich directo o indirecto, y ensayos de inmunoprecipitación. Zola, *Monoclonal Antibodies: A Manual of Techniques*, pp. 147-158 (CRC Press, Inc., 1987).

Los ensayos de unión competitiva se basan en la capacidad de un estándar marcado para competir con el analito de la muestra a analizar por la unión con una cantidad limitada de anticuerpo. La cantidad de proteína diana (codificada por un gen amplificado en una célula tumoral) en la muestra a analizar es inversamente proporcional a la cantidad de estándar que se une al anticuerpo. Para facilitar la determinación de la cantidad de estándar que resulta unido, preferentemente, los anticuerpos se insolubilizan antes o después de la competición, así el estándar y el analito unidos al anticuerpo pueden separarse de forma adecuada del estándar y del analito que permanecen sin unir.

Los ensayos sándwich implican el uso de dos anticuerpos, cada uno capaz de unirse a diferentes regiones inmunogénicas o epítomos de la proteína a detectar. En un ensayo de sándwich, el analito de la muestra a analizar se une a un primer anticuerpo que está inmovilizado en un soporte sólido, y a continuación un segundo anticuerpo se une al analito, formando así un complejo de tres partes insoluble. Véase, por ejemplo, la patente americana U.S. 4.376.110. El segundo anticuerpo puede estar marcado con un grupo detectable (ensayos sándwich directos) o puede medirse usando un anticuerpo anti-inmunoglobulina que esté marcado con un grupo detectable (ensayos sándwich indirectos). Por ejemplo, un ensayo de sándwich es un ensayo de inmutación enzimática ELISA, en cuyo caso el grupo detectable es una enzima.

Para inmunohistoquímica, la muestra tumoral debe ser fresca o congelada o puede estar embebida en parafina y fijada con un conservante como formalina, por ejemplo.

I. Ensayos tumorales basados en células

Se pueden usar ensayos basados en células y modelos animales para tumores (por ejemplo, cánceres) para verificar los hallazgos del ensayo de amplificación génica, y mejorar el conocimiento de la relación entre los genes identificados aquí y el desarrollo de la patogénesis del crecimiento celular neoplásico. El papel de los productos génicos identificados aquí en el desarrollo y la patología de un tumor o cáncer puede analizarse usando células tumorales primarias o líneas celulares en las que se haya observado amplificación de dichos genes. Tales células incluyen, por ejemplo, células de cáncer de mama, colon y pulmón y las líneas celulares indicadas anteriormente.

En una aproximación diferente, se transfectan células de un tipo celular del que se conoce su implicación en un tumor en particular con los cADNs aquí descritos y se analiza la capacidad de estos cADNs de inducir un crecimiento desmedido. Entre las células adecuadas se incluyen, por ejemplo, líneas celulares tumorales estables tales como la línea celular B104-1-1 (línea celular NIH-3T3 transfectada de forma estable con el protooncogen *neu*) y células NIH-3T3 transfectadas con ras, que pueden ser transfectadas con el gen deseado y controladas respecto a su crecimiento tumorigénico. Dichas líneas celulares transfectadas pueden ser usadas para analizar la capacidad de anticuerpos poli o monoclonales o de mezclas de anticuerpos para inhibir el crecimiento celular tumorigénico al ejercer una acción citostática o citotóxica sobre el crecimiento de las células transformadas, o mediante la citotoxicidad celular dependiente del anticuerpo (ADCC). Las células transfectadas con las secuencias codificantes de los genes identificados aquí pueden usarse además para la identificación de fármacos candidatos para el tratamiento del cáncer.

Además, pueden usarse los cultivos primarios derivados de tumores en animales transgénicos (tal como se describe más adelante) en los ensayos basados en células aquí descritos, aunque se prefieren las líneas celulares estables. Son bien conocidas en el área las técnicas para obtener líneas celulares continuas a partir de animales transgénicos (véase, por ejemplo, Small y col., *Mol. Cell. Biol.* 5:642-618 [1985]).

J. Modelos animales

Es posible utilizar una variedad de modelos animales bien conocidos para entender en mayor profundidad el papel de los genes identificados aquí en el desarrollo y la patogénesis tumoral, y para analizar la eficacia de los agentes terapéuticos candidatos, incluyendo anticuerpos y otros antagonistas de los polipéptidos nativos, incluyendo moléculas pequeñas que actúan como antagonistas. La naturaleza *in vivo* de dichos modelos los hace particularmente útiles para predecir respuestas en pacientes humanos. Entre los modelos animales de tumores y cánceres (por ejemplo, cáncer

de mama, de colon, de próstata, de pulmón, etc.) se incluyen tanto animales no recombinantes como recombinantes (transgénicos). Entre los modelos animales no recombinantes se incluyen, por ejemplo, roedores, por ejemplo, modelos murinos. Tales modelos pueden generarse introduciendo células tumorales en ratones singénicos usando técnicas habituales, por ejemplo, inyección subcutánea, inyección en la vena de la cola, implantación en el bazo, implantación intraperitoneal, implantación bajo la cápsula renal, o implantación ortotópica, por ejemplo, células de cáncer de colon implantadas en un tejido colónico (véase, por ejemplo, la publicación PCT de la patente internacional WO97/33551, publicada el 18 de Septiembre de 1997).

Probablemente, la especie animal más frecuentemente utilizada en estudios oncológicos son ratones inmunodeficientes y, en particular, ratones *nude*. La observación de que los ratones *nude* con hipo/aplasia pueden actuar con éxito como un huésped para xenoinjertos de tumores humanos ha llevado a su uso generalizado para este propósito. El gen autosómico recesivo *nu* se ha introducido en un gran número de distintas cepas congénitas de ratones desnudos, incluyendo, por ejemplo, ASW, A/He, AKR, BALB/c, B10.LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, I/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII, y S JL. Además, se han criado y usado como receptores de xenoinjertos tumorales una amplia variedad de otros animales con defectos inmunológicos heredados distintos a parte del ratón *nude*. Para más detalles véase, por ejemplo, *The Nude Mouse in Oncology Research*, E. Boven and B. Winograd, eds., CRC Press, Inc., 1991.

Las células introducidas en dichos animales pueden derivarse de líneas celulares de un tumor/cáncer conocido, tales como, cualquiera de las líneas celulares tumorales enumeradas anteriormente, y, por ejemplo, la línea celular B104-1-1 (la línea celular NIH-3T3 transfectada de forma estable con el protooncogen *neu*); células NIH-3T3 transfectadas con *ras*; Caco-2 (ATCC HTB-37); una línea celular de adenocarcinoma de colon humano de grado II moderadamente bien diferenciado, las HT-29 (ATCC HTB-38), o a partir de tumores y cánceres. Las muestras de células tumorales o cancerosas pueden obtenerse a partir de pacientes que van a someterse a una cirugía, usando condiciones estándar, que implican congelación y almacenamiento en nitrógeno líquido (Karmali y col., Br. J. Cancer, 48:689-696 [1983]).

Las células tumorales pueden introducirse en animales, tales como ratones *nude*, mediante una variedad de procedimientos. El espacio subcutáneo (s.c.) en ratón es muy apropiado para la implantación de tumores. Los tumores pueden transplantarse s.c. como bloques sólidos, como biopsias con aguja usando un trócar o como suspensiones celulares. Para bloques sólidos o implantaciones por trócar, fragmentos de tejido tumoral de un tamaño adecuado se introducen en el espacio s.c. Las suspensiones celulares se preparan en fresco a partir de tumores primarios o de líneas celulares tumorales estables, y se inyectan subcutáneamente. Las células tumorales pueden inyectarse como implantes subdérmicos. En esta localización, el inóculo se deposita entre la parte inferior del tejido conectivo dérmico y el tejido s.c. Boven y Winograd (1991), *supra*.

Se pueden generar modelos animales de cáncer de mama, por ejemplo, mediante la implantación de células de neuroblastoma de rata (de las cuales se aisló inicialmente el oncogen *neu*) o de células NIH-3T3 transformadas con *neu* en animales *nude*, esencialmente tal como describe Drebin y col., PNAS USA, 83:9129-9133 (1986).

Análogamente, se pueden generar modelos animales de cáncer de colon transfiriendo células de cáncer de colon a animales, por ejemplo, ratones *nude*, lo que lleva a la aparición de tumores en estos animales. Se ha descrito un modelo de trasplante ortotópico de cáncer de colon humano en ratones *nude*, por ejemplo por Wang y col., Cancer Research, 54:4726-4728 (1994) y Too y col., Cancer Research, 55:681-684 (1995). Este modelo se basa en el llamado "METAMOUSE" vendido por AntiCancer, Inc., (San Diego, California).

Los tumores que se generan en los animales pueden extraerse y cultivarse *in vitro*. Las células procedentes de los cultivos *in vitro* pueden transferirse a animales. Dichos tumores pueden servir como dianas para análisis adicionales o para el análisis de fármacos. Alternativamente, los tumores resultantes de la transferencia pueden aislarse y se puede analizar el ARN procedente de las células antes de la transferencia y de las células aisladas después de una o más rondas de transferencia para determinar la expresión diferencial de los genes de interés. Dichas técnicas de transferencia pueden realizarse con cualquier línea celular tumoral o cancerosa conocida.

Por ejemplo, Meth A, CMS4, CMS5, CMS21, y WEHI-164 son fibrosarcomas inducidos químicamente en ratones hembra BALB/c (DeLeo y col., J. Exp. Med., 146:720[1977]), que proporcionan un sistema modelo altamente controlable para el estudio de las actividades antitumorales de diversos agentes (Palladino y col., J. Immuno., 138:4023-4032 [1987]). Brevemente, se propagan células tumorales *in vitro* en cultivo celular. Antes de la inyección en los animales, las líneas celulares se lavan y se resuspenden en tampón, a una densidad celular de alrededor de 10×10^6 a 10×10^7 cel/ml. Los animales se infectan entonces subcutáneamente con de 10 a 100 μ l de la suspensión celular, permitiendo la aparición de un tumor al cabo de una a tres semanas.

Además, el carcinoma de pulmón de Lewis (3LL) de ratón, que es uno de los tumores experimentales más ampliamente estudiados, puede utilizarse como un modelo tumoral investigacional. La eficacia de este modelo tumoral se ha correlacionado con los efectos beneficiosos en el tratamiento de pacientes humanos diagnosticados con carcinoma de células pequeñas de pulmón (SCCL). Este tumor puede introducirse en ratones normales mediante inyección de fragmentos de tumor de un ratón afectado o de células mantenidas en cultivo (Zupi y col., Br. J. Cancer, 41: suppl.4:309 [1980]), y existen evidencias que indican que los tumores pueden iniciarse a partir de la inyección de incluso una única célula y que una proporción muy alta de células tumorales infectadas sobreviven. Para más información sobre este modelo tumoral véase, Zacharski, Haemostasis, 16:300-320 [1986].

Una manera de evaluar la eficacia de un compuesto a analizar en un modelo animal sobre un tumor implantado es medir el tamaño del tumor antes y después del tratamiento. Tradicionalmente, se ha medido el tamaño de los tumores implantados con un portaobjetos calibrador en dos o tres dimensiones. La medida limitada a dos dimensiones no refleja de forma precisa el tamaño del tumor, por tanto, habitualmente se convierte en el volumen correspondiente usando una fórmula matemática. Sin embargo, la medida del tamaño del tumor es muy poco precisa. Los efectos terapéuticos de un fármaco candidato pueden describirse mejor como el retraso en el crecimiento inducido por el tratamiento y el retraso específico del crecimiento. Otra variable importante en la descripción del crecimiento del tumor es el tiempo de duplicación del volumen del tumor. Están disponibles programas informáticos para el cálculo y la descripción del crecimiento tumoral, tales como el programa descrito por Rygaard y Spang-Thomsen, Proc. 6th Int Workshop on Immune-Deficient Animals. Wu and Sheng eds., Basel, 1989,301. Cabe destacar, sin embargo, que al menos inicialmente, procesos de necrosis y respuestas inflamatorias pueden realmente dar lugar a un aumento del tamaño del tumor tras el tratamiento. Por tanto, estos cambios deben ser supervisados cuidadosamente, mediante una combinación de procedimientos morfométricos y análisis por citometría de flujo.

Se pueden generar modelos animales recombinantes (transgénicos) introduciendo la región codificante de los genes identificados aquí en el genoma de los animales de interés, usando técnicas estándar para la producción de animales transgénicos. Entre los animales que pueden ser objeto de manipulación transgénica se incluyen, sin limitaciones, ratones, ratas, conejos, cobayas, cordero, cabras, cerdos y primatos no humanos, e. j., babuinos, chimpancés y monos. Entre las técnicas conocidas en el área para introducir un transgen en dichos animales se incluye la microinyección pronuclear (Hoppe y Wanger, patente americana U.S. 4.873.191); la transferencia génica en líneas germinales mediante retrovirus (e. j., Van der Putten y col., Proc Natl. Aca. Sci. USA, 82:6148-615 [1985]; la recombinación homóloga en células madre embrionarias (Thompson y col., Cell, 56:313-321 [1989]); la electroporación de embriones (Lo, Mol. Cell Biol., 3:1803-1814 [1983]); la transferencia génica mediante esperma (Lavitrano y col., Cell, 57:717-73 [1989]). Véase, por ejemplo, la patente americana U.S. 4.736.866 para revisar el tema.

Para el objetivo de la presente invención, en los animales transgénicos se incluyen también aquellos que portan el transgen sólo en una parte de sus células (animales mosaico). El transgen puede integrarse como un único transgen, o en concatámeros, por ejemplo, tándem cabeza-cabeza, cabeza-cola. La introducción selectiva de un transgen en un tipo de célula en particular también es posible siguiendo, por ejemplo, la técnica de Lasko y col., Proc. Natl. Acad. Sci. USA, 89:6232-636 (1992).

La expresión del transgen en los animales transgénicos puede controlarse mediante técnicas estándar. Por ejemplo, puede utilizarse el análisis de transferencia de Southern o la amplificación por PCR para verificar la integración del transgen. El nivel de expresión de mRNA puede analizarse entonces usando técnicas como la hibridación *in situ*, el análisis por transferencia de Northern, la PCR, o la inmunocitoquímica. Los animales se examinan además para determinar signos de desarrollo canceroso o tumoral.

Alternativamente, se pueden generar animales genéticamente deficientes, los cuales tienen un gen anormal o alterado que codifica para un polipéptido PRO5775 identificado aquí, como resultado de una recombinación homóloga entre el gen endógeno que codifica por el polipéptido y un ADN genómico alterado que codifica por el mismo polipéptido introducido en una célula embrionaria del animal. Por ejemplo, puede usarse el cADN que codifica por el polipéptido PRO5775 para clonar el ADN genómico que codifica este polipéptido según técnicas establecidas. Una fracción del ADN genómico que codifica un polipéptido PRO5775 en particular puede ser eliminada o reemplazada por otro gen, tal como un gen que codifique un marcador seleccionable que puede utilizarse para hacer un seguimiento de la integración. Típicamente, se incluyen en el vector varias kilobases de ADN flanqueante no modificado (tanto en los extremos 5' como 3') [véase, por ejemplo, Thomas and Capecchi, Cell, 51:503 (1987) para la descripción de la recombinación homóloga de vectores]. El vector se introduce en una línea de células madre embrionarias (por ejemplo, por electroporación) y se seleccionan las células en las que el ADN introducido se ha recombinado homológamente con el ADN endógeno [véase, por ejemplo, Li y col., Cell, 69:915(1992)]. Las células seleccionadas se inyectan entonces en un blastocito de un animal (por ejemplo, un ratón o una rata) para formar quimeras de agregación [véase, por ejemplo, Bradley, en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E.J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A continuación un embrión quimérico puede implantarse en un animal adoptivo apropiado, una hembra seudopreñada, y el embrión puede llevarse a término para crear un animal genéticamente deficiente o "knock-out". La progenie que presenta el ADN recombinado homológamente en sus células germinales puede identificarse mediante técnicas estándares y usarse para criar animales en los que todas las células del animal contengan el ADN recombinado homológamente. Los animales knock-out pueden caracterizarse, por ejemplo, por su capacidad para defenderse contra ciertas condiciones patológicas y por el desarrollo de condiciones patológicas debido a la ausencia del polipéptido PRO5775.

La eficacia de anticuerpos que se unen específicamente a los polipéptidos aquí identificados y a otros fármacos candidatos, puede analizarse también para el tratamiento de tumores animales espontáneos. Una diana adecuada para dichos estudios es el carcinoma oral felino de células escamosas (SCC). El SCC oral felino es un tumor maligno altamente invasivo, es el tumor oral más común en gatos, ya que supone un 60% de los tumores orales descritos en esta especie. Raramente metastatiza en regiones distales, aunque esta baja incidencia de metástasis puede ser un mero reflejo del corto tiempo de supervivencia que muestran los gatos con este tumor. Habitualmente, estos tumores no son susceptibles de cirugía, principalmente debido a la anatomía de la cavidad oral del felino. En la actualidad, no existe un tratamiento efectivo para este tumor. Antes de entrar en el estudio, cada gato se somete a un examen clínico completo, a biopsia y es escaneado mediante tomografía computerizada (CT). Los gatos a los que se les diagnostica un tumor

oral sublingual de células escamosas se excluyen del estudio. La lengua puede llegar a paralizarse como resultado de dicho tumor, e incluso si el tratamiento destruye el tumor, los animales pueden no ser capaces de alimentarse por sí solos. Cada gato se trata de forma repetida, durante un largo período de tiempo. Se toman fotografías de los tumores diariamente durante el período de tratamiento, y en cada chequeo subsecuente. Tras el tratamiento, cada gato se somete a otro escáner CT. A partir de entonces, los escáneres por CT y las radiografías torácicas se evalúan cada 8 semanas. Se analizan los datos en busca de diferencias en la supervivencia, la respuesta y la toxicidad en comparación con los grupos control. Una respuesta positiva puede requerir evidencias de la regresión tumoral, preferentemente con una mejora en la calidad de vida/o un aumento en el tiempo de vida.

Además también pueden analizarse otros tumores animales espontáneos, tales como el fibrosarcoma, el adenocarcinoma, el linfoma, el condroma, el leiomiomasarcoma de perros, gatos y babuinos. De estos, el adenocarcinoma mamario en perros y gatos es el modelo preferido ya que su apariencia y comportamiento es muy similar al de los humanos. Sin embargo, el uso de este modelo está limitado por la baja incidencia de este tipo de tumores en animales.

15 K. Ensayos de cribado de fármacos candidatos

Los ensayos de cribado de fármacos candidatos están diseñados para identificar compuestos que se unan o formen complejos con los polipéptidos codificados por los genes identificados aquí, o de otro modo interfieran en la interacción de los polipéptidos con otras proteínas celulares. Tales ensayos de cribado incluirán ensayos susceptibles de estudios a gran escala o librerías químicas, haciéndolos particularmente adecuados para la identificación de pequeñas moléculas como fármacos candidatos. Entre las pequeñas moléculas contempladas se incluyen compuestos sintéticos orgánicos o inorgánicos, incluyendo péptidos, preferentemente péptidos solubles, fusiones de (poli)péptidos e inmunoglobulinas, y, en particular, anticuerpos, incluyendo, sin limitación, anticuerpos poli y monoclonales y fragmentos de anticuerpos, anticuerpos de una sola cadena, anticuerpos antiidiotípicos, y versiones quiméricas o humanizadas de dichos anticuerpos o fragmentos, así como anticuerpos humanos y fragmentos de los mismos. Los ensayos pueden realizarse en una variedad de formatos, incluyendo ensayos de unión de proteína-proteína, ensayos bioquímicos, inmunoensayos y ensayos basados en células, que están bien caracterizados en el área.

Todos los ensayos tienen en común que requieren el contacto del fármaco candidato con un polipéptido codificado por un ácido nucleico aquí identificado en condiciones adecuadas y durante un tiempo suficiente como para permitir que estos dos componentes interactúen.

En ensayos de unión, la interacción es la unión y el complejo formado puede así aislarse o detectarse en la mezcla de reacción. En una realización en particular, el polipéptido codificado por el gen identificado aquí o el fármaco candidato se inmoviliza sobre una fase sólida, por ejemplo, en una microplaca, mediante uniones covalentes o no covalentes. La unión no covalente generalmente se consigue mediante el recubrimiento de una superficie sólida con una solución del polipéptido y dejándolo secar. Alternativamente, un anticuerpo inmovilizado, por ejemplo, un anticuerpo monoclonal, específico para el polipéptido a ser inmovilizado puede usarse para fijarlo a una superficie sólida. El ensayo se realiza añadiendo el componente no inmovilizado, que puede ser marcado con un marcaje detectable, al componente inmovilizado, por ejemplo, la superficie recubierta que contiene el componente fijado. Cuando la reacción se ha completado, se eliminan los componentes que no han reaccionado, por ejemplo, lavando, y se detectan los complejos fijados en la superficie sólida. Cuando el componente no inmovilizado en un principio lleva un marcaje detectable, la detección del marcaje inmovilizado en la superficie indica que el complejo se ha formado. Si el componente no inmovilizado en un principio no lleva un marcaje, la formación del complejo debe detectarse, por ejemplo, usando un anticuerpo marcado que se una específicamente al complejo inmovilizado.

Si el compuesto candidato interactúa pero no se une a un polipéptido PRO5775 en particular codificado por un gen identificado aquí, su interacción con este polipéptido debe ser analizada mediante procedimientos bien conocidos para detectar interacciones proteína-proteína. Dichos ensayos incluyen aproximaciones tradicionales, tales como el entrecruzamiento, la coimmunoprecipitación, y la copurificación a través de gradientes o en columnas cromatográficas. Además, las interacciones proteína-proteína pueden controlarse mediante un sistema genético basado en levaduras descrito por Fields y colaboradores [Fields and Song, *Nature*, 340:245-246 (1989); Chien y col., *Proc. Natl. Acad. Sci. USA*, 88:9578-9582 (1991)] como se describe en Chevray and Nathans, *Proc. Natl. Acad. Sci. USA*, 89:5789-5793 (1991)]. Muchos activadores transcripcionales, tales como GAL4 de levadura, consisten en dos dominios modulares físicamente discretos, uno que actúa como el dominio de unión a ADN, mientras el otro funciona como el dominio de activación de la transcripción. El sistema de expresión en levadura descrito en las publicaciones antes mencionadas (generalmente denominados como "sistema del doble híbrido") toma ventaja de esta propiedad, y utiliza dos proteínas híbridas, una en la que la proteína diana está fusionada al dominio de unión del ADN de GAL4, y otra en la que las proteínas activadoras candidatas se fusionan al dominio de activación. La expresión de un gen reportero GAL1/*lacZ* bajo el control de un promotor activado por GAL4, depende de la reconstitución de la actividad GAL4 a través de la interacción proteína-proteína. Las colonias que contienen polipéptidos que interactúan se detectan con un sustrato cromogénico para la β -galactosidasa. Existe un equipo reactivo completo (MATCHMAKER™) disponible comercialmente por Clontech para la identificación de interacciones proteína-proteína entre dos proteínas específicas usando la técnica del doble híbrido. Este sistema también puede utilizarse para localizar dominios proteicos implicados en interacciones proteicas específicas así como para señalar los residuos aminoacídicos que son cruciales para estas interacciones.

ES 2 279 473 T3

Los compuestos que interfieren con la interacción de un gen identificado aquí que codifica PRO5775, PRO313, y otros componentes intra o extracelulares pueden ser analizados como sigue: habitualmente una mezcla de reacción se prepara conteniendo el producto del gen amplificado y el componente intra o extracelular en condiciones tales y durante un tiempo suficiente que se permita la interacción y la unión de los dos productos. Para determinar la capacidad de un compuesto a analizar de inhibir la unión, la reacción se realiza en ausencia y en presencia del compuesto a analizar. Además, puede añadirse un placebo a una tercera mezcla de reacción, para utilizarla como control positivo. La unión (formación del complejo) entre el compuesto a analizar y el componente intra o extracelular presente en la mezcla se controla tal como se describió anteriormente. La formación de un complejo en la(s) reacción(es) control pero no en la mezcla de reacción que contiene el compuesto a analizar indica que dicho compuesto interfiere en su interacción con su pareja de reacción.

Para analizar antagonistas, el polipéptido PRO5775 puede añadirse a una célula junto con el compuesto a analizar para una actividad en particular, y la capacidad del compuesto para inhibir la actividad de interés en presencia del polipéptido PRO5775 indica que el compuesto es un antagonista del polipéptido PRO5775. Alternativamente, se pueden detectar antagonistas mediante la combinación del polipéptido PRO5775 y un potencial antagonista con receptores del polipéptido PRO5775 unidos a membrana o receptores recombinantes bajo condiciones adecuadas para un ensayo de inhibición competitiva. El polipéptido PRO5775 puede marcarse, mediante por ejemplo radioactividad, de forma que el número de moléculas del polipéptido PRO5775 unidas al receptor pueden usarse para determinar la eficacia del potencial antagonista. El gen que codifica el receptor puede identificarse mediante numerosos procedimientos conocidos por los expertos en la materia, por ejemplo, utilización de un panel de ligandos y separación por FACS. Coligan y col., *Current Protocols in Immun.*, 1(2): capítulo 5(1991). Preferentemente, se emplea el clonaje de expresión en donde se prepara ARN poliadenilado a partir de una célula que responde al polipéptido PRO5775 y una librería de cADN creada a partir de este ARN se divide en grupos y se usa para transfectar células COS u otras células que no responden al polipéptido PRO5775. Las células transfectadas que han crecido en portaobjetos de cristal se exponen al polipéptido PRO5775 marcado. El polipéptido PRO5775 puede marcarse mediante distintos sistemas incluyendo la yodinación o la inclusión de un sitio de reconocimiento para una proteína quinasa de unión sitio-específica. Tras la fijación e incubación, los portaobjetos se someten a un análisis autorradiográfico. Los grupos positivos se identifican y se preparan subgrupos y se retransfectan usando un subagrupamiento interactivo y un nuevo proceso de análisis dando lugar, eventualmente, a un único clon que codifica para el receptor putativo.

Como una aproximación *alteARN*tiva para la identificación de un receptor, el polipéptido PRO5775 marcado puede unirse por fotoafinidad con preparaciones de membrana celular o extractos que expresan la molécula del receptor. El material entrecruzado se resuelve por PAGE y se expone a una película de rayos X. El complejo marcado que contiene el receptor puede cortarse, separarse en pequeños fragmentos y someterse a microsecuenciación de proteínas. La secuencia aminoacídica obtenida por microsecuenciación sería usada para diseñar un conjunto de sondas de oligonucleótidos degenerados para analizar una librería de cADN e identificar el gen que codifica el receptor putativo.

En otro ensayo para identificar antagonistas, se incubarían células de mamífero o una preparación de membrana que expresen el receptor con el polipéptido PRO5775 marcado en presencia del compuesto candidato. Puede así medirse la capacidad del compuesto de aumentar o bloquear esta interacción.

Ejemplos más específicos sobre potenciales antagonistas incluyen un oligonucleótido que se une a las fusiones de inmunoglobulina con el polipéptido PRO5775 y, en particular, anticuerpos incluyendo, sin limitación, anticuerpos poli y monoclonales y fragmentos de anticuerpo, anticuerpos de cadena sencilla, anticuerpos antiidiotípicos, y versiones quiméricas o humanizadas de dichos anticuerpos o fragmentos de los mismos, así como anticuerpos humanos y fragmentos de estos anticuerpos. Alternativamente, un potencial antagonista puede ser una proteína estrechamente relacionada, por ejemplo, una forma mutada del polipéptido PRO5775 que reconozca el receptor pero que no ejerza ningún efecto, por tanto, inhibiendo competitivamente la acción del polipéptido PRO5775.

Otro antagonista potencial del polipéptido PRO5775 es una construcción antisentido de ARN o ADN preparada usando la tecnología antisentido, en donde, por ejemplo, una molécula de ADN o ARN antisentido actúa bloqueando directamente la traducción de un mRNA mediante la hibridación con el mRNA diana y evitando la traducción de la proteína. La tecnología antisentido puede usarse para controlar la expresión génica a través de la formación de una triple hélice o del ADN o ARN antisentido, ambos procedimientos basados en la unión de un polinucleótido al ADN o al ARN. Por ejemplo, la región codificante 5' de la secuencia polinucleotídica, que codifica por el polipéptido maduro PRO5775 de esta invención se usa para diseñar un oligonucleótido antisentido de ARN de alrededor de 10 a 40 pares de bases de longitud. Se diseña un oligonucleótido de ADN que sea complementario a una región del gen implicado en la transcripción (triple hélice- véase, Lee y col., *Nucl. Acids Res.*, 6: 3073 (1979); Cooney y col., *Science*, 241:456 (1988); Dervan y col., *Science*, 251:1360 (1991)], evitando así la transcripción y la producción del polipéptido PRO5775. El oligonucleótido antisentido de ARN hibrida con el mRNA *in vivo* y bloquea la traducción de la molécula de mRNA en el polipéptido PRO5775 (antisentido-Okano, *Neurochem.*, 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). Los oligonucleótidos descritos anteriormente pueden también liberarse a células de forma que el ARN o ADN antisentido puede ser expresado *in vivo* para inhibir la producción del polipéptido PRO5775. Cuando se usa un ADN antisentido, se prefieren oligodeoxiribonucleótidos derivados a partir de la región de inicio de la traducción, por ejemplo, entre las posiciones de aproximadamente -10 a +10 de la secuencia nucleotídica del gen diana.

Las moléculas de ADN o ARN antisentido tienen generalmente una longitud de al menos aproximadamente 5 bases, de aproximadamente 10 bases, de aproximadamente 15 bases, de aproximadamente 20 bases, de aproximadamente 25 bases, de aproximadamente 30 bases, de aproximadamente 35 bases, de aproximadamente 40 bases, de aproximadamente 45 bases, de aproximadamente 50 bases, de aproximadamente 55 bases, de aproximadamente 60 bases, de aproximadamente 65 bases, de aproximadamente 70 bases, de aproximadamente 75 bases, de aproximadamente 80 bases, de aproximadamente 85 bases, de aproximadamente 90 bases, de aproximadamente 95 bases, de aproximadamente 100 bases, o más.

Entre los potenciales antagonistas se incluyen moléculas pequeñas que se unen al sitio activo, al sitio de unión al receptor, o del factor de crecimiento u otro sitio de unión relevante del polipéptido PRO5775, bloqueando así la actividad biológica normal del polipéptido PRO5775. Entre los ejemplos de pequeñas moléculas se incluyen, pero no está limitado a, pequeños péptidos o moléculas de tipo peptídico, preferentemente, péptidos solubles, y compuestos sintéticos orgánicos no peptídicos o inorgánicos.

Las ribozimas son moléculas de ARN enzimáticas capaces de catalizar la fragmentación específica del ARN. Las ribozimas actúan mediante hibridación a secuencias específicas del ARN diana complementario, seguido por el corte endonucleolítico. Los sitios de corte de la ribozima dentro de un potencial ARN diana se pueden identificar mediante técnicas conocidas. Para más detalles véase, por ejemplo, Rossi *Current Biology*, 4: 469-471 (1994), y la publicación PCT de la patente internacional 97/33.551 (publicada el 18 de Septiembre de 1997).

Las moléculas de ácido nucleico en formación de triple hélice usadas para inhibir la transcripción deberán ser de cadena sencilla y compuestas por desoxinucleótidos. La composición de bases de estos oligonucleótidos se diseña de forma que se promueva la formación de la triple hélice a través de las reglas de apareamiento de bases de Hoogsteen, las cuales generalmente requieren considerables tramos de purinas o pirimidinas en una cadena del dúplex. Para más detalles véase, por ejemplo, la publicación PCT de la patente internacional WO 97/33551, *supra*.

Estas pequeñas moléculas pueden identificarse mediante uno o más de los ensayos de análisis descritos aquí anteriormente y/o mediante cualquier otra técnica de análisis bien conocida por los expertos en la materia.

L. *Composiciones y procedimientos para el tratamiento de tumores.*

Las composiciones útiles en el tratamiento de tumores asociados con la amplificación de los genes identificados aquí, incluyen sin limitación, anticuerpos, pequeñas moléculas orgánicas e inorgánicas, péptidos, fosfopéptidos, moléculas de ribozima antisentido, moléculas de triple hélice, etc, que inhiban la expresión y/o la actividad del producto del gen diana.

Por ejemplo, moléculas de ARN y ARNs antisentido actúan para bloquear directamente la traducción del mRNA mediante la hibridación al mRNA diana y evitando así la traducción proteica. Cuando se usa un ADN antisentido, se prefieren oligodesoxiribonucleótidos derivados a partir del sitio de inicio de la traducción, por ejemplo, entre aproximadamente de la posición -10 a la +10 de la secuencia nucleotídica del gen diana.

Las ribozimas son moléculas de ARN enzimáticas capaces de catalizar la fragmentación específica del ARN. Las ribozimas actúan mediante hibridación a secuencias específicas del ARN diana complementario, seguido por el corte endonucleolítico. Los sitios de corte específicos de la ribozima dentro de un potencial ARN diana se pueden identificar mediante técnicas conocidas. Para más detalles véase, por ejemplo, Rossi, *Current Biology*, 4:469-471 (1994), y la publicación PCT de la patente internacional WO 97/33.551 (publicada el 18 de Septiembre de 1997).

Las moléculas de ácidos nucleicos en formación de triple hélice usadas para inhibir la transcripción suelen ser de cadena sencilla y compuestas por desoxinucleótidos. La composición de bases de estos oligonucleótidos se diseña de forma que promueva la formación de la triple hélice a través de las reglas de apareamiento de bases de Hoogsteen, las cuales generalmente requieren considerables tramos de purinas o pirimidinas en una cadena de un dúplex. Para más detalles véase, por ejemplo, la publicación PCT de patente internacional WO 97/33.551, *supra*.

Estas moléculas pueden identificarse mediante una o más de los ensayos de análisis descritos aquí anteriormente y/o mediante cualquier otra técnica de análisis bien conocida para aquellos expertos en el área.

M. *Anticuerpos*

Algunos de los fármacos candidatos más prometedores según la presente invención son anticuerpos y fragmentos de anticuerpos que pueden inhibir la producción o el producto del gen de los genes amplificados identificados aquí y/o reducir la actividad de los productos del gen.

1. *Anticuerpos policlonales*

Los procedimientos para la preparación de anticuerpos policlonales son conocidos por los expertos. Los anticuerpos policlonales pueden generarse en un mamífero, por ejemplo, mediante una o más inyecciones de un agente inmunizante y, si se desea, un adyuvante. Típicamente, el agente inmunizante y/o el adyuvante se inyectarán en el mamífero mediante múltiples inyecciones subcutáneas o intraperitoneales. El agente inmunizante puede incluir el

polipéptido PRO5775 o una fusión proteica del mismo. Puede ser útil la conjugación del agente inmunizante a una proteína de la que se conoce su poder inmunogénico en el mamífero que va a ser inmunizado. Entre los ejemplos de dichas proteínas inmunogénicas se incluyen, pero sin limitarse a, la hemocianina de lapa, la albúmina sérica, la tiroglobulina bovina, y el inhibidor de la tripsina de soja. Entre los ejemplos de los adyuvantes que pueden ser utilizados se incluyen el adyuvante completo de Freund's y el adyuvante MPL-TDM (monofosforil Lípido A, dicorinomicolato sintético de trehalosa). El protocolo de inmunización puede seleccionarse por los expertos en la materia sin excesiva experimentación.

2. Anticuerpos monoclonales

Los anticuerpos contra el PRO5775 pueden ser, alternativamente, anticuerpos monoclonales. Los anticuerpos monoclonales pueden prepararse usando procedimientos del hibridoma, como aquellos descritos por Kohler y Milstein, *Nature*, 256:495 (1975). En un procedimiento del hibridoma, un ratón, un hámster, u otro animal huésped apropiado, es típicamente inmunizado con un agente inmunizante que da lugar a linfocitos que producen o son capaces de producir anticuerpos que se unirán específicamente con el agente inmunizante. Alternativamente, los linfocitos pueden ser inmunizados *in vitro*.

El agente inmunizante típicamente incluirá el polipéptido PRO5775, incluyendo fragmentos, o una proteína de fusión de dicha proteína o un fragmento de los mismos. Generalmente, se usan tanto linfocitos de sangre periférica ("PBLs") si se desean células de origen humano, como células de bazo o de nódulos linfáticos si se desean fuentes mamíferas no humanas. Los linfocitos se fusionan entonces con una línea celular inmortalizada usando un agente de fusión adecuado, como polietilenglicol, para formar una célula de hibridoma [Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59-103]. Las líneas celulares inmortalizadas son habitualmente células de mamífero transformadas, particularmente células de mieloma de origen murino, bovino y humano. Habitualmente, se utilizan líneas celulares de mieloma de rata o ratón. Las células de hibridoma pueden cultivarse en un medio de cultivo adecuado que contenga preferentemente una o más sustancias que inhiban el crecimiento o la supervivencia de las células inmortalizadas no fusionadas. Por ejemplo, si las células parentales carecen de la enzima hipoxantina guanina fosforibosil transferasa (HGPRT o HPRT), el medio de cultivo de los hibridomas incluirá típicamente hipoxantina, aminopterina, y timidina (medio "HAT"), sustancias que evitan el crecimiento de las células deficientes en HGPRT.

Las líneas celulares inmortalizadas preferidas son aquellas que se fusionan de forma eficiente, soportan un nivel de expresión elevado y estable del anticuerpo por las células productoras de anticuerpo seleccionadas, y son sensibles a un medio como el medio HAT. Las líneas celulares inmortalizadas más preferibles son líneas de mieloma murino, que pueden obtenerse, por ejemplo, a partir del Salk Institute Cell Distribution Center, San Diego, California y a partir del American Type Culture Collection (ATCC), Manassas, Virginia. Las líneas celulares de mieloma humano de heteromieloma ratón-humano también se han descrito para la producción de anticuerpos monoclonales humanos [Kozbor, *J. Immunol.*, 133:3001 (1984); Brodeur y col., *Monoclonal Antibody Production Techniques and Applications*, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

El medio de cultivo en el que se cultivan las células de hibridoma puede analizarse para determinar la presencia de anticuerpos monoclonales dirigidos contra el PRO5775. Preferentemente, la especificidad de unión de los anticuerpos monoclonales producidos por las células de hibridoma se determina por inmunoprecipitación o mediante ensayos de unión *in vitro* como el radioinmunoensayo (RIA) o el ensayo de inmunoabsorción enzimática (ELISA). Tales técnicas y ensayos son conocidos en el área. La afinidad de unión del anticuerpo monoclonal puede, por ejemplo, determinarse a través del análisis de Scatchard de Munson y Pollard, *Anal. Biochem.*, 107:220 (1980).

Después de identificar las células de hibridoma deseadas, es posible subclonar los clones mediante procedimientos de selección por dilución limitante y crecerlos mediante procedimientos estándares [Goding, *supra*]. Entre los medios de cultivo adecuados para este propósito se incluyen, por ejemplo, el Dulbecco's Modified Eagle's Medium y el medio RPMI-1640. Alternativamente, las células de hibridoma pueden crecerse *in vivo* como ascitas en un mamífero.

Los anticuerpos monoclonales secretados por los subclones pueden aislarse o purificarse a partir del medio de cultivo o del fluido ascítico mediante procedimientos convencionales de purificación de inmunoglobulinas tales como, por ejemplo, proteína A-Sepharose, cromatografía de hidroxilapatita, electroforesis en gel, diálisis, o cromatografía de afinidad.

Los anticuerpos monoclonales pueden prepararse también mediante procedimientos de ADN recombinante, tales como los descritos en la patente americana U.S. 4.816.567. El ADN que codifica los anticuerpos monoclonales de la invención puede aislarse fácilmente y secuenciarse usando procedimientos convencionales (por ejemplo, usando sondas de oligonucleótidos que son capaces de unirse específicamente a genes que codifican la cadena pesada y la cadena ligera de anticuerpos murinos). Las células de hibridoma de la invención sirven como una fuente preferente de dicho ADN. Una vez aislado, el ADN puede introducirse en vectores de expresión, que son transfectados a continuación en células huésped, tales como las células COS de simio, las células de ovario de hámster chino (CHO), o las células de mieloma que de lo contrario no producirían proteína inmunoglobulina, para obtener la síntesis de anticuerpos monoclonales en las células huésped recombinantes. El ADN también puede modificarse, por ejemplo, sustituyendo la secuencia codificante de los dominios constantes de las cadenas pesada y ligera humanas en lugar de las secuencias murinas homólogas [Patente americana U.S. 4.816.567; Morrison y col., *supra*] o mediante unión covalente a la

secuencia codificante de la inmunoglobulina de toda o parte de la secuencia codificante de un polipéptido no inmunoglobulina. Dicho polipéptido no inmunoglobulínico puede sustituirse por los dominios constantes de un anticuerpo de la invención o puede sustituirse por los dominios variables de una región de un sitio de combinación con el antígeno de un anticuerpo de la invención para crear un anticuerpo bivalente quimérico.

5 Los anticuerpos pueden ser anticuerpos monovalentes. Los procedimientos para la preparación de anticuerpos monovalentes son bien conocidos en la materia. Por ejemplo, un procedimiento implica la expresión recombinante de la cadena ligera de una inmunoglobulina y la modificación de la cadena pesada. La cadena pesada generalmente se trunca en cualquier punto de la región Fc para prevenir el entrecruzamiento de la cadena pesada. Alternativamente, los residuos cisteína relevantes son sustituidos por otro residuo aminoacídico o se eliminan para prevenir el entrecruzamiento.

10 Para la preparación de anticuerpos monovalentes también son apropiados procedimientos *in vitro*. Puede conseguirse la digestión de anticuerpos para producir fragmentos de los mismos, particularmente, fragmentos Fab, usando técnicas rutinarias conocidas la materia.

15 3. Anticuerpos humanos y humanizados

Los anticuerpos contra el PRO5775 pueden además comprender anticuerpos humanizados o anticuerpos humanos. Formas humanizadas de anticuerpos no humanos (por ejemplo, murinos) son las inmunoglobulinas quiméricas, cadenas de inmunoglobulinas o fragmentos de los mismos (como Fv, Fab, Fab', F(ab')₂ u otras subsecuencias de unión a antígeno de los anticuerpos) que contienen una secuencia mínima derivada de una inmunoglobulina no humana. Entre los anticuerpos humanizados se incluyen inmunoglobulinas humanas (anticuerpo receptor) en las que se sustituyen residuos de una región determinante de la complementariedad (CDR) del receptor por los correspondientes residuos no humanos de un CDR de una especie no humana (anticuerpo donante) como ratón, rata o conejo que tengan la especificidad, afinidad y capacidad deseadas. En algunos casos, residuos Fv de la región de entramado de la inmunoglobulina humana se sustituyen por los correspondientes residuos no humanos. Los anticuerpos humanizados también pueden comprender residuos que no se encuentran ni en el anticuerpo receptor ni en el CDR importado ni en las secuencias de la región de entramado. En general, el anticuerpo humanizado comprenderá substancialmente todos o al menos uno, y típicamente dos, dominios variables, en los que todas o substancialmente todas las regiones CDR corresponden a las de una inmunoglobulina no humana y todas o esencialmente todas las regiones FR son las de una secuencia consenso de una inmunoglobulina humana. Óptimamente, el anticuerpo humanizado comprenderá al menos una porción de una región constante de una inmunoglobulina (Fc), típicamente aquella de una inmunoglobulina humana [Jones y col., Nature, 321:522-525 (1986); Riechmann y col., Nature, 332:323-329 (1988); y Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

35 Los procedimientos para la preparación de anticuerpos humanizados no humanos son bien conocidos en el área. Generalmente, un anticuerpo humanizado tiene uno o más de un residuo aminoacídico introducido en él a partir de una fuente que no es humana. Estos residuos aminoacídicos no humanos se denominan frecuentemente como residuos "importados", que típicamente se toman a partir de un dominio variable "importado". La humanización puede realizarse esencialmente, siguiendo el procedimiento de Winter y colaboradores [Jones y col., Nature, 321:522-525 (1986); Riechmann y col., Nature, 332:323-327 (1988); Verhoeyen y col., Science, 239:1534-1536 (1988)], y por sustitución de secuencias CDR o CDRs de roedores por las secuencias correspondientes de un anticuerpo humano. En concordancia, tales anticuerpos "humanizados" son anticuerpos quiméricos (patente americana U.S. 4.816.567), en donde substancialmente menos de un dominio variable humano intacto se han substituido por la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados son típicamente anticuerpos humanos en los que algunos residuos CDR y posiblemente algunos residuos FR se substituyen por residuos de regiones análogas en anticuerpos de roedores.

50 Los anticuerpos humanos pueden producirse también usando diversas técnicas conocidas en la materia, incluyendo librerías de expresión en fagos [Hoogenboom y Winter, J. Mol. Biol., 227:381 (1991); Marks y col., J. Mol. Biol., 222:581 (1991)]. Las técnicas de Cole y col., y Boerner y col., también están disponibles para la preparación de anticuerpos monoclonales humanos (Cole y col., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p.77 (1985) y Boerner y col., J. Immunol., 147(1): 86-95 (1991)]. Análogamente, los anticuerpos humanos pueden prepararse introduciendo los loci de inmunoglobulinas humanas en animales transgénicos, por ejemplo, ratones en los que los genes endógenos de las inmunoglobulinas se han inactivado parcial o totalmente. Tras el estímulo, se observa la producción del anticuerpo humano, lo que refleja estrechamente lo observado en humanos en todos los sentidos, incluyendo el reordenamiento génico, el ensamblaje, y el repertorio de anticuerpos. Esta aproximación se describe, por ejemplo, en las patentes americanas U.S. 5.545.807; 5.545.806; 5.569.825; 5.625.126; 5.633.425; 5.661.016, y en las siguientes publicaciones científicas: Marks y col., Bio/Technology, 10:779-783 (1992); Lonberg y col., Nature, 368:856-859 (1994); Morrison, Nature, 368:812-13 (1994); Fishwild y col., Nature Biotechnology, 14:845-51 (1996); Neuberger, Nature Biotechnology, 14:826 (1996); Lonberg y Huszar, Intern. Rev. Immunol., 13:65-93 (1995).

4. Terapia de profármacos mediados por enzimas dependientes de anticuerpos (ADEPT)

65 Los anticuerpos de la presente invención pueden usarse en ADEPT mediante la conjugación del anticuerpo a un enzima activadora de profármaco que convierte un profármaco (por ejemplo, un agente quimioterapéutico de tipo peptidilo, véase patente internacional WO 81/01.145) en un fármaco anticanceroso activo. Véase, por ejemplo, la patente internacional WO 88/07.378 y la patente americana U.S. 4.975.278.

El componente enzimático del inmunoconjugado útil para la ADEPT incluye cualquier enzima capaz de actuar sobre un profármaco de manera que lo convierta en su forma más activa y citotóxica.

Entre las enzimas que son útiles en el procedimiento de esta invención se incluyen, pero no está limitado a, la glicosidasa, la glucosa oxidasa, la lisozima humana, la glucuronidasa humana, la fosfatasa alcalina útil por convertir profármacos que contienen fosfato en los fármacos libres; la arilsulfatasa útil por convertir profármacos que contienen sulfato en los fármacos libres; citosina deaminasa útil por convertir la 5-fluorocitosina no tóxica en el fármaco anticanceroso 5-fluorouracilo; proteasas, tales como la proteasa de *Serratia*, la termolisina, la subtilisina, carboxipeptidasas (por ejemplo, carboxipeptidasa G2 y carboxipeptidasa A) y las catepsinas (tales como las catepsinas B y L), que son útiles por convertir profármacos que contienen péptidos en los fármacos libres; D-alanilcarboxipeptidasas, útiles por convertir profármacos que contienen sustituyentes de tipo D-aminoácido; las enzimas que hidrolizan carbohidratos como la β -galactosidasa y la neuraminidasa útiles por convertir profármacos glicosilados en fármacos libres; la β -lactamasa útil por convertir fármacos derivatizados con β -lactámicos en los fármacos libres; y las penicilin amidasas; tales como la penicilin varnidasa o la penicilin G amidasa, útiles por convertir fármacos derivatizados en sus aminas nitrogenadas con grupos fenoxiacetil o la fenilacetil, respectivamente, en fármacos libres. Alternativamente, pueden usarse anticuerpos con actividad enzimática también conocidos en el área como "abzimas" para convertir los profármacos de la invención en fármacos activos libres (véase, por ejemplo, Massey, Nature, 328:457-458 (1987)). Los conjugados anticuerpo-abzima pueden prepararse como se describe aquí para la liberación del abzima de una población de células tumorales.

Los enzimas de esta invención pueden unirse covalentemente a los anticuerpos contra el PRO5775 mediante técnicas bien conocidas en la materia como el uso de agentes de entrecruzamiento heterobifuncionales descritos anteriormente. Alternativamente, pueden generarse proteínas de fusión que comprendan al menos la región de unión al antígeno del anticuerpo de la invención unido a al menos una porción funcionalmente activa de una enzima de la invención usando técnicas de ADN recombinante bien conocidas en la materia (véase, por ejemplo, Neuberger y col., Nature, 312:604-608 (1984)).

5. Anticuerpos biespecíficos

Los anticuerpos biespecíficos son anticuerpos monoclonales, preferentemente humanos o humanizados, que tienen especificidades de unión para al menos dos antígenos diferentes. En el caso presente, una de las especificidades de unión es para el PRO5775 y la otra para cualquier otro antígeno, y preferentemente por una proteína de superficie celular o por un receptor o subunidad de un receptor.

Son conocidos en la materia los procedimientos para la preparación de anticuerpos biespecíficos. Tradicionalmente, la producción recombinante de anticuerpos biespecíficos se basa en la coexpresión de dos pares de cadena pesada/cadena ligera de inmunoglobulina, en donde las dos cadenas pesadas tienen distintas especificidades (Milstein y Cuello, Nature, 305:537-539 [1983]). Debido a la reordenación aleatoria de las cadenas pesadas y ligeras de inmunoglobulina, estos hibridomas (cuadromas) producen una mezcla potencial de diez moléculas de anticuerpo distintas, de los cuales sólo una presenta la estructura biespecífica correcta. La purificación de la molécula correcta se consigue habitualmente mediante procedimientos de cromatografía de afinidad. Se describen procedimientos similares en la patente internacional WO 93/08.829, publicada el 13 de Mayo de 1993, y en Traunecker y col., EMBO J., 10:3655-3659 (1991).

Los dominios variables de un anticuerpo con las especificidades de unión deseadas (sitios de combinación anticuerpo-antígeno) pueden fusionarse a secuencias de dominios constantes de una inmunoglobulina. La fusión preferente es con un dominio constante de una cadena pesada de inmunoglobulina, que comprenda al menos parte de las regiones bisagra, CH2 y CH3. Se prefiere tener presente la primera región constante de la cadena pesada (CH1) que contenga el sitio necesario para la unión de la cadena ligera en al menos una de las fusiones. Los ADNs que codifican las fusiones de la cadena pesada de la inmunoglobulina y, si se prefiere, la cadena ligera de la inmunoglobulina, se insertan en vectores de expresión separados, y se co-transfectan en un organismo huésped apropiado. Para más detalles sobre la generación de anticuerpos biespecíficos véase, por ejemplo, Suresh y col., Methods in Enzymology, 121:210 (1986).

Según otra aproximación descrita en la patente internacional WO 96/27.011, es posible diseñar la interfaz entre un par de moléculas de anticuerpo con el fin de maximizar el porcentaje de heterodímeros a recuperar a partir de un cultivo celular recombinante. La interfaz preferida comprende al menos una parte de la región CH3 de un dominio constante de un anticuerpo. En este procedimiento, una o más cadenas laterales pequeñas de aminoácidos de la interfaz de la primera molécula de anticuerpo se reemplaza por cadenas laterales más grandes (por ejemplo, tirosina o triptófano). Se crean "cavidades compensatorias de tamaño idéntico o similar al de las cadena(s) laterales más grandes en la interfaz de la segunda molécula de anticuerpo mediante la sustitución de las cadenas laterales aminoacídicas más grandes por otras más pequeñas (por ejemplo, alanina o treonina). Esto proporciona un mecanismo para incrementar el rendimiento del heterodímero frente a otros productos finales no deseados como los homodímeros.

Los anticuerpos biespecíficos pueden prepararse como anticuerpos de longitud completa o como fragmentos de anticuerpos (por ejemplo, anticuerpos biespecíficos F(ab')₂). Se han descrito en la literatura técnicas para la generación de anticuerpos biespecíficos a partir de fragmentos de anticuerpos. Por ejemplo, se pueden preparar anticuerpos biespecíficos usando ligamiento químico. Brennan y col., Science, 229:81 (1985) describen un procedimiento en don-

de anticuerpos intactos se rompen proteolíticamente para generar fragmentos $F(ab')_2$. Estos fragmentos se reducen en presencia del agente acomplejante ditiol y del arsenito sódico para estabilizar ditioles vecinos y prevenir la formación de disulfuros intermoleculares. Los fragmentos Fab' generados se convierten a continuación en derivados de tionitrobenzoato (TNB). Uno de los derivados Fab' -TNB se reconvierte a continuación en el tiol- Fab' a través de la reducción con mercaptoetilamina y se mezcla con una cantidad equimolar de otros derivados Fab' -TNB para formar el anticuerpo biespecífico. Los anticuerpos biespecíficos producidos pueden usarse como agentes para la inmovilización selectiva de enzimas.

Se pueden recuperar fragmentos Fab' de *E. coli* directamente y acoplarlos químicamente para formar anticuerpos biespecíficos. Shalaby y col., *J. Exp. Med.*, 175:217-225 (1992) describen la producción de una molécula de anticuerpo biespecífico $F(ab')_2$ totalmente humanizado. Cada fragmento Fab' se secreta separadamente a partir de *E. coli* y se somete a acoplamiento químico directo *in vitro* para formar el anticuerpo biespecífico. El anticuerpo biespecífico así formado es capaz de unirse a células que sobreexpresen el receptor ErbB2 y a células T humanas normales, así como desencadenar la actividad lítica de linfocitos citotóxicos humanos contra los tumores diana de mama humano.

Se han descrito también varias técnicas para preparar y aislar fragmentos de anticuerpos biespecíficos directamente a partir de cultivo de células recombinante. Por ejemplo, se han producido anticuerpos biespecíficos usando cremalleras de leucinas. Kostelny y col., *J. Immunol.*, 148(5):1547-1553 (1992). Los péptidos de la cremallera de leucina de las proteínas Fos y Jun se unen a las regiones Fab' de dos anticuerpos diferentes mediante fusión génica. Los homodímeros de anticuerpos se reducen en la región bisagra para formar monómeros y entonces oxidarse de nuevo para formar los heterodímeros del anticuerpo. Este procedimiento puede utilizarse también para la producción de homodímeros de anticuerpos. La tecnología de "diacuerpos" descrita por Hollinger y col., *Proc. Natl. Acad. Sci. USA*, 90:6444-6448 (1993) ha proporcionado un mecanismo alternativo para preparar fragmentos de anticuerpos biespecíficos. Los fragmentos comprenden un dominio variable de cadena pesada (V_H) conectado a un dominio variable de cadena ligera (V_L) mediante un engarce que sea lo suficientemente corto como para permitir el emparejamiento entre los dos dominios de la misma cadena. En concordancia, los dominios (V_H) y (V_L) de un fragmento se ven forzados a emparejarse con los dominios (V_H) y (V_L) complementarios de otro fragmento, formando así dos sitios de unión al antígeno. Se ha descrito otra estrategia para preparar fragmentos de anticuerpos biespecíficos mediante el uso de dímeros de Fv de cadena sencilla (sFv). Véase, Gruber y col., *J. Immunol.*, 152:5368 (1994).

Se contemplan anticuerpos con más de dos valencias. Por ejemplo se pueden preparar anticuerpos trispecíficos. Tutt y col., *J. Immunol.*, 147:60 (1991).

Los anticuerpos biespecíficos típicos pueden unirse a dos epítomos diferentes en un polipéptido determinado. Alternativamente, un brazo de un anti-polipéptido puede combinarse con un brazo que se une a una molécula activadora de un leucocito como una molécula del receptor de células T (por ejemplo, CD2, CD3, CD28, o B7), o receptores Fc para IgG (FcγR), tal como FcγRI (CD64), FcγRII (CD32) y FcγRIII (CD16) para así dirigir los mecanismos de defensa celular hacia la célula que expresa el polipéptido en particular. Los anticuerpos biespecíficos pueden usarse también para localizar agentes citotóxicos para las células que expresan un polipéptido en particular. Estos anticuerpos poseen un brazo de unión del polipéptido y un brazo que se une a un agente citotóxico o a un quelante radionucleico, tal como EOTUBE, DPTA, DOTA, o TETA. Otro anticuerpo biespecífico de interés se une al polipéptido y además se une al factor tisular (TF).

6. Anticuerpos heteroconjugados

Los anticuerpos heteroconjugados se componen de dos anticuerpos unidos covalentemente. Se ha propuesto que dichos anticuerpos, por ejemplo, dirigen las células del sistema inmune contra las células no deseadas [patente americana U.S. 4.676.980], y sirven para el tratamiento de la infección por VIH [patentes internacionales WO 91/00360; WO 92/200.373; patente europea EP 03.089]. Se considera que los anticuerpos se pueden preparar *in vitro* usando procedimientos conocidos en la química sintética de proteínas, incluyendo aquellos que implican agentes de entrecruzamiento. Por ejemplo, se pueden generar inmunotoxinas usando una reacción de intercambio de disulfuro o formando un enlace tioéter. Entre los ejemplos de agentes apropiados para este objetivo se incluye el iminotiolato y el metil-4-mercaptobutirimidato y aquellos descritos, por ejemplo, en la patente americana U.S. 4.676.980.

7. Diseño de funciones efectoras

Puede ser deseable modificar el anticuerpo de la invención con respecto a la función efectora, de manera que se incremente la efectividad del anticuerpo para el tratamiento del cáncer, por ejemplo. Por ejemplo, se pueden introducir residuo(s) de cisteína en la región Fc, permitiendo así la formación de enlaces disulfuro entre las cadenas en esta región. El anticuerpo homodimérico así generado puede presentar una mejora en la capacidad de internalización y/o un incremento de la capacidad de muerte celular mediada por el complemento y de la citotoxicidad celular dependiente del anticuerpo (ADCC). Véase, Caron y col., *J. Exp. Med.*, 176:1191-1195 (1992) y Shopes, *J. Immunol.*, 148:2918-2922 (1992). Los anticuerpos homodiméricos que muestran un incremento en la actividad antitumoral pueden prepararse también usando agentes de entrecruzamiento heterobifuncionales, tal como se describe en Wolf y col., *Cancer Research* 53:2560-2565 (1993). Alternativamente, se puede preparar un anticuerpo que tenga regiones Fc duales y que puedan por tanto mostrar un incremento en las capacidades de lisis mediada por el complemento y de ADCC. Véase, Stevenson y col., *Anti-Cancer Drug Design*, 3:219-230 (1989).

8. *Immunoconjugados*

La invención también se refiere a inmunoconjugados que comprenden un anticuerpo conjugado con un agente citotóxico tal como un agente quimioterapéutico, una toxina (por ejemplo, una toxina enzimáticamente activa de origen bacteriano, fúngico, de animales o plantas, o fragmentos de los mismos, o una pequeña molécula de toxina), o un isótopo radioactivo (por ejemplo, un radioconjugado).

Los agentes quimioterapéuticos útiles en la generación de dichos inmunoconjugados se han descrito anteriormente. Entre las toxinas proteicas enzimáticamente activas y los fragmentos de las mismas que pueden utilizarse se incluyen la cadena A de difteria, fragmentos activos no unidos de la toxina difterica, la toxina colérica, la toxina botulínica, la cadena A de la exotoxina (de *Pseudomonas aeruginosa*), la cadena A de la ricina, la cadena A de la abrina, la cadena A de la modicina, la alfa-sarcina, las proteínas de *Aleurites fordii*, las proteínas diantina, las proteínas de *Phytolaca americana* (PAPI, PAPII, y PAP-S), el inhibidor de la *Momordica charantia*, la curcina, la crotina, el inhibidor de *Sapaonaria officinalis*, la gelonina, la saponina, la mitogelina, la restrictocina, la fenomicina, la enomicina y los tricotesenos. Las pequeñas moléculas de toxinas incluyen, por ejemplo, caliqueamicinas, maitansinoides, palitoxina y CC1065. Para la producción de anticuerpos radioconjugados están disponibles una gran variedad de radionucleidos. Entre los ejemplos se incluyen Bi²¹², I¹³¹, In¹¹¹, Y⁹⁰, Re¹⁸⁶.

Se pueden preparar conjugados del anticuerpo y del agente citotóxico usando una variedad de agentes de acoplamiento de proteínas bifuncionales como el propionato de N-succinimidil-3-(2-piridiltiol) (SPDP), iminotiolano (IT), derivados bifuncionales de imidoésteres, (tal como el clorhidrato de dimetiladipimidato), ésteres activos (tal como el disuccinimidil suberato), aldehídos (tal como el glutaraldehído), compuestos bis-azido (como el bis(p-azidobenzoil) hexanediamina), derivados de bis-diazonio (como el bis-(p-diazo-niumbenzoil)-etilenediantino diisocianatos (tal como el tolueno 2,6-diisocianato), y compuestos de fluorina bi-activos (tal como el 1,5-difluoro-2,4-dinitrobenceno). Por ejemplo, una inmunotoxina de ricina se puede preparar tal como se describe en Viletta y col., Science, 238:1098 (1987). Un ejemplo de agente quelante para la conjugación de un radionucleido al anticuerpo es el ácido 1-isotiocianato-bencil-3-metildietil triaminopentaacético (MX-DTPA) marcado con carbono-14. Véase, la patente internacional WO94/11.026.

En otra realización, el anticuerpo puede conjugarse a un "receptor" (tal como la estreptavidina) para ser utilizado en el marcaje de tumores en donde el conjugado anticuerpo-receptor se administra al paciente, seguido de la eliminación del conjugado no unido de la circulación usando un agente aclarante y administrando a continuación un "ligando" (por ejemplo, avidina) que se conjuga a un agente tóxico (por ejemplo, un radionucleótido).

9. *Inmunoliposomas*

Los anticuerpos descritos aquí pueden formularse también en forma de inmunoliposomas. Los liposomas que contienen el anticuerpo se preparan mediante procedimientos conocidos en la materia, tales como los descritos por Epstein y col., Proc. Natl. Acad. Sci. USA, 82:3688 (1985); Hwang y col., Proc. Natl. Acad. Sci. USA, 77:4030 (1980); y las patentes americanas U.S. 4.485.045 y 4.544.545. Los liposomas que presentan un aumento en el tiempo de circulación se describen en la patente americana n° 5013556.

Se pueden generar liposomas particularmente útiles mediante el procedimiento de evaporación en fase reversa con una composición lipídica que comprende fosfatidilcolina, colesterol y fosfatidiletanolamina derivatizada con PEG (PEG-PE). Los liposomas se extrusionan a través de filtros de poro definido para dar lugar a liposomas con el diámetro deseado. Pueden conjugarse fragmentos Fab' del anticuerpo de la presente invención con los liposomas tal y como se describe en Martín y col., J. Biol. Chem., 257:286-288 (1982) a través de una reacción de intercambio de disulfuro. Opcionalmente, el liposoma puede contener un agente quimioterapéutico (tal como la doxorubicina). Véase, Gabizon y col., J. National Cancer Inst., 81(19):1484 (1989).

N. *Composiciones farmacéuticas*

Se pueden administrar los anticuerpos unidos específicamente al producto de un gen amplificado identificado aquí, así como otras moléculas identificadas en los estudios de análisis descritos aquí anteriormente, para el tratamiento de tumores, incluyendo cánceres, en forma de composiciones farmacéuticas.

Si la proteína codificada por el gen amplificado es intracelular y se usan los anticuerpos completos como inhibidores, se prefieren anticuerpos que se internalicen. Sin embargo, se pueden utilizar también lipofecciones o liposomas para vehicular el anticuerpo, o un fragmento del anticuerpo al interior de las células. En donde se usan los fragmentos de anticuerpo, se prefiere el fragmento inhibidor más pequeño que se una específicamente al dominio de unión de la proteína diana. Por ejemplo, se pueden diseñar moléculas peptídicas que retengan la capacidad de unirse a la secuencia proteica diana basándose en las secuencias de la región variable de un anticuerpo. Dichos péptidos pueden sintetizarse químicamente y/o producirse mediante la tecnología del ADN recombinante (véase, por ejemplo, Marasco y col., Proc. Natl. Acad. Sci. USA, 90:7889-7893 [(1993)]).

Se preparan formulaciones terapéuticas del anticuerpo para su almacenamiento mezclando el anticuerpo que presenta el grado deseado de pureza con vehículos opcionales farmacéuticamente aceptables, con excipientes o con estabilizantes (Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed. [1980]), en forma de formulaciones

liofilizadas o de soluciones acuosas. Los vehículos, excipientes o estabilizantes aceptables no han de ser tóxicos para el receptor a las dosis y concentraciones empleadas, e incluyen tampones como el fosfato, el citrato y otros ácidos orgánicos; antioxidantes incluyendo el ácido ascórbico y la metionina; conservantes (tales como el cloruro de octadecildimetilbenzil amonio, el cloruro de hexametonio; el cloruro de benzalconio, el cloruro de bencetonio; el alcohol fenólico, butílico o benzílico; los alquil parabenos tales como el metil o propil parabeno; el catecol; el resorcinol; el ciclohexanol; el 3-pentanol; y m-cresol); polipéptidos de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas, tales como la albúmina sérica, la gelatina, o las inmunoglobulinas; los polímeros hidrofílicos tales como al polivinilpirrolidona; los aminoácidos como la glicina, la glutamina, la asparragina, la histidina, la arginina, o la lisina; monosacáridos, disacáridos, y otros carbohidratos incluyendo la glucosa, la manosa, o las dextrinas; agentes quelantes tales como el EDTA; azúcares como la sacarosa, el manitol, la trehalosa o el sorbitol; contrarrestadores de iones formadores de sales tales como el sodio; complejos metálicos (por ejemplo, el complejos de proteína-Zn); y/o surfactantes no iónicos tales como TWEENTM, PLURONICSTM o el polietilenglicol (PEG).

De forma análoga, se pueden formular compuestos, que no son anticuerpos y que se han identificado mediante los ensayos de análisis de la presente invención, usando técnicas estándares bien conocidas en la materia.

La formulación aquí descrita también puede contener más de un compuesto activo si es necesario para la indicación en particular que vaya a ser tratada, preferentemente aquellos con actividades complementarias que no se afecten de forma adversa el uno con el otro. Alternativamente, o además, la composición puede comprender un agente citotóxico, una citoquina o un agente inhibidor del crecimiento. Tales moléculas están presentes de manera adecuadamente en combinación con cantidades que sean efectivas para el propósito que se pretende.

Los ingredientes activos pueden también encapsularse en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o por polimerización interfacial, por ejemplo, hidroximetilcelulosa o microcápsulas de gelatina y microcápsulas de poli-(metilmelacilato), respectivamente, en sistemas coloidales de liberación de fármacos (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas) o en macroemulsiones. Tales técnicas se describen en Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed. (1980).

Las formulaciones que van a usarse para la administración *in vivo* deben ser estériles. Esto se consigue fácilmente mediante filtración a través de membranas de filtración estériles.

Se pueden preparar preparaciones de liberación sostenida. Entre los ejemplos adecuados de preparaciones de liberación sostenida se incluyen matrices semipermeables de polímeros sólidos hidrofóbicos que contienen el anticuerpo, matrices que están en forma de artículos con formas determinadas, por ejemplo, películas o microcápsulas. Entre los ejemplos de matrices de liberación sostenida se incluyen poliésteres, hidrogeles (por ejemplo, poli(2-hidroxietil-metacrilato), o poli(vinilalcohol)), poliláctidos (patente americana U.S. 3.773.919), copolímeros de ácido L-glutámico y etil-L-glutamato, acetato de etilen-vinilo no degradable, copolímeros de ácido láctico-ácido glicólico degradables tales como el LUPRON DEPOTTM (microesferas inyectables compuestas de un copolímero de ácido láctico-ácido glicólico y acetato de leuprolide), y ácido poli-D-(-)3-hidroxibutírico. Mientras polímeros como el acetato de etilen-vinilo y el de ácido láctico-ácido glicólico permiten la liberación de moléculas durante 100 días, ciertos hidrogeles liberan proteínas durante periodos de tiempo más cortos. Cuando los anticuerpos encapsulados permanecen en el organismo durante un tiempo prolongado, se pueden desnaturalizar o agregar como resultado de la exposición a la humedad a 37°C, lo que se traduce en una pérdida de actividad biológica y en posibles cambios en la inmunogenicidad. Deben idearse estrategias racionales para la estabilización dependiendo del mecanismo implicado. Por ejemplo, si se descubre que el mecanismo de agregación se produce a través de la formación de un enlace S-S intermolecular a través de un intercambio tio-disulfuro, la estabilización puede lograrse modificando los residuos sulfidrilos, liofilizando a partir de soluciones ácidas, controlando el contenido de humedad, usando aditivos adecuados, y desarrollando composiciones específicas de matrices poliméricas.

50 O. Procedimientos de tratamiento

Se contempla que los anticuerpos y otros compuestos antitumorales de la presente invención pueden utilizarse para el tratamiento de varias condiciones, incluyendo aquellas caracterizadas por la sobreexpresión y/o activación de genes amplificados identificados aquí. Ejemplos de las condiciones o trastorno a ser tratados con dichos anticuerpos y otros compuestos, incluyendo, pero no limitándose a, pequeñas moléculas orgánicas e inorgánicas, péptidos, moléculas antisentido, etc., se incluyen los tumores benignos o malignos (por ejemplo, el carcinomas renal, el hepático, el de riñón, el de vejiga, el de mama, el gástrico, el ovárico, el colorectal, el de próstata, el pancreático, el pulmonar, el vulval, el tiroideo, el hepático; los sarcomas; los glioblastomas; y diversos tumores de cabeza y cuello); las leucemias y los tumores linfoides; otros trastornos de origen neuronal, glial, astrocital, hipotálamico y otros trastornos glandulares, macrofagales, epiteliales, estromales y blastocoélicos; y trastornos inflamatorios angiogénicos e inmunológicos.

Los agentes antitumorales de la presente invención, por ejemplo, anticuerpos, se administran en mamíferos, preferentemente en humanos, según procedimientos conocidos, tales como la administración intravenosa como un bolo o mediante infusión continua durante un periodo de tiempo determinado, mediante vía intramuscular, intraperitoneal, intracerebroespinal, subcutánea, intraarticular, intrasinovial, intratecal, oral, tópica, o inhalatoria. Se prefiere la administración intravenosa para la administración del anticuerpo.

Otros regímenes terapéuticos pueden combinarse con la administración de agentes anti-cancerosos, por ejemplo, los anticuerpos de la presente invención. Por ejemplo, el paciente a ser tratado con dichos agentes anti-cancerosos puede también recibir terapia de radiación. Alternativa o adicionalmente, puede administrarse al paciente un agente quimioterapéutico. Como guías para la preparación y administración de dosis de dichos agentes quimioterapéuticos pueden utilizarse las normas de acuerdo con las instrucciones del fabricante o según el criterio del experto en la materia. Las guías para la preparación y administración de dosis de dicho quimioterapéutico también se describen en Chemotherapy Service, Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992). El agente quimioterapéutico puede preceder o seguir a la administración del agente anti-tumoral, por ejemplo el anticuerpo, o puede administrarse simultáneamente. El anticuerpo puede combinarse con un compuesto anti-estrogénico como el tamoxifeno o uno anti-progeste-
 5 rona como la onapristona (véase la patente europea EP 616.812) en dosificaciones conocidas para dichas moléculas.
 10

También puede ser deseable administrar anticuerpos contra otros antígenos asociados, como los anticuerpos que se unen al ErbB2, EGFR, ErbB3, ErbB4, o al factor endotelial vascular (VEGF). Alternativa o adicionalmente, dos o más anticuerpos que se unen a los mismos o distintos antígenos descritos aquí pueden coadministrarse al paciente. A veces, puede ser beneficioso administrar una o más citocinas al paciente. En una realización preferida, los anticuerpos se coadministran con un agente inhibidor del crecimiento. Por ejemplo, el agente inhibidor del crecimiento puede administrarse en primer lugar, seguido de un anticuerpo de la presente invención. Sin embargo, la administración simultánea o la administración del anticuerpo de la presente invención en primer lugar también se contempla en la presente invención. Dosis adecuadas del agente inhibidor del crecimiento son las utilizadas en la presente y pueden disminuirse debido a la acción combinada (sinergia) del agente inhibidor del crecimiento y del anticuerpo de la presente invención.
 15
 20

Para la prevención o tratamiento de la enfermedad, la dosis adecuada de un agente anti-tumoral, por ejemplo, un anticuerpo dependerá del tipo de enfermedad a tratar, tal como se definió anteriormente, de la gravedad y del curso de la enfermedad tanto si el agente se administra con propósitos preventivos o terapéuticos, de terapias previas, de la historia clínica del paciente y de la respuesta al agente y de la discreción del médico responsable. El agente se administra de forma adecuada al paciente en una sola vez o durante una serie de tratamientos.
 25

Por ejemplo, dependiendo del tipo y gravedad de la enfermedad, una dosis candidata inicial para la administración al paciente será de aproximadamente 1 $\mu\text{g}/\text{Kg}$ a 15 mg/Kg (por ejemplo 0,1-20 mg/Kg) de anticuerpo, bien mediante, por ejemplo, una o más administraciones por separado, o mediante infusión continua. Una dosis diaria típica puede hallarse entre aproximadamente 1 $\mu\text{g}/\text{Kg}$ a 100 mg/Kg o más, dependiendo de los factores mencionados anteriormente. Para administraciones repetidas durante varios días o más, dependiendo de la condición, el tratamiento se prolonga hasta la supresión deseada de los síntomas de la enfermedad. Sin embargo, otros regímenes de dosificación pueden ser útiles. El progreso de esta terapia se monitoriza fácilmente mediante técnicas y ensayos convencionales.
 30
 35

P. Artículos de fabricación

En otra realización de la invención, se proporciona un artículo de fabricación que contiene materiales de utilidad para el diagnóstico de las enfermedades descritas anteriormente. El artículo de fabricación comprende un recipiente y una etiqueta. Recipientes adecuados incluyen, por ejemplo, botellas, viales, jeringas y tubos de ensayo. Los recipientes pueden estar formados de muy diversos materiales como por ejemplo el vidrio o el plástico. El recipiente contiene una composición que es efectiva para el diagnóstico o el tratamiento de la condición y puede tener un puerto de acceso estéril (por ejemplo, el recipiente puede ser una bolsa de solución intravenosa o un vial con un tapón agujereable por una aguja de inyección hipodérmica). El agente activo en la composición es generalmente un agente anti-tumoral capaz de interferir con la actividad de un producto génico identificado aquí, por ejemplo, un anticuerpo. La etiqueta sobre el recipiente, o en su interior, indica que la composición se utiliza para el diagnóstico o el tratamiento de la condición de elección. El artículo de fabricación puede comprender además un segundo recipiente que contiene un tampón aceptable farmacéuticamente, como el tampón fosfato salino, la solución de Ringer y la solución de dextrosa. Puede además incluir otros materiales deseables desde un punto de vista comercial y del usuario, incluyendo otros tampones, diluyentes, filtros, agujas, jeringas y prospectos con las instrucciones para su utilización.
 40
 45
 50

Q. Diagnóstico y pronóstico del tumor

Mientras que las proteínas de superficie celular, como los receptores del crecimiento sobreexpresados en ciertos tumores son dianas excelentes como fármacos candidatos o para el tratamiento de tumores (por ejemplo el cáncer), las mismas proteínas junto con proteínas secretadas codificadas por genes amplificados en células tumorales hallan una utilización adicional en el diagnóstico y pronóstico de tumores. Por ejemplo, los anticuerpos dirigidos contra productos proteicos de genes amplificados en células tumorales pueden utilizarse para el diagnóstico y pronóstico de tumores.
 55
 60

Por ejemplo, los anticuerpos, incluyendo los fragmentos de anticuerpos, pueden utilizarse para detectar cualitativa o cuantitativamente la expresión de proteínas codificadas por los genes amplificados ("productos de genes marcadores"). El anticuerpo está equipado preferentemente con un marcador detectable, por ejemplo un marcador fluorescente, y la unión puede monitorizarse mediante microscopía, citometría de flujo, fluorimetría u otras técnicas conocidas en la materia. Estas técnicas son particularmente adecuadas, si el gen amplificado codifica una proteína de superficie celular, por ejemplo, un factor de crecimiento. Dichos ensayos de unión se realizan esencialmente tal como se describe en la sección 5 de más arriba.
 65

La detección *in situ* de la unión del anticuerpo con los productos del gen marcador puede realizarse, por ejemplo, mediante microscopía de inmunofluorescencia o microelectrónica. Para dicho propósito, se obtiene un espécimen histológico del paciente, y se aplica el anticuerpo marcado a éste, preferentemente poniendo el anticuerpo sobre la muestra biológica. Este procedimiento también permite la determinación de la distribución del producto del gen marcador en el tejido examinado. Será evidente para aquellos expertos en la materia que una gran diversidad de procedimientos histológicos está fácilmente disponible para la detección *in situ*.

Los siguientes ejemplos se ofrecen únicamente con motivo ilustrativo, y no con intención de limitar de ninguna manera el ámbito de la presente invención.

Ejemplos

Los reactivos disponibles comercialmente referidos en los ejemplos se utilizaron de acuerdo con las instrucciones del fabricante salvo que se indique lo contrario. La fuente de las células se han identificado en los ejemplos siguientes, y a través de las especificaciones, por los números de acceso ATCC que corresponden al American Type Culture Collection, 1081 University Blvd., Manassas, VA 20110-2209. Todos los depósitos originales referidos en la presente solicitud se generaron según las indicaciones del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos con el Propósito de Procedimientos de Patentes y las Regulaciones que se desprenden (Tratado de Budapest). Ello asegura el mantenimiento de un cultivo viable del depósito durante 30 años desde la fecha del depósito. El depósito estará disponible por el ATCC bajo los términos del Tratado de Budapest, y sometido a un acuerdo entre Genentech Inc., y ATCC, que asegure la disponibilidad permanente y no restringida de la progenie del cultivo del depósito al público bajo expedición de la patente americana pertinente o tras poner a disposición pública de cualquier solicitud de patente U.S. o extranjera, cualquiera que sea la primera, y asegurar la disponibilidad de la progenie por la persona determinada por el Comisionado U.S. de Patentes y Marcas Registradas de acuerdo con la norma 35 USC § 122 y las normas del Comisionado según lo acordado (incluyendo 37 CFR § 1.14 con particular referencia a 886 OG 638).

Salvo que se especifique lo contrario, la presente invención utiliza procedimientos estándares de tecnología del ADN recombinante, tal como los descritos anteriormente y los descritos en los libros de texto siguientes: Sambrook y col., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press N.Y., 1989; Ausubel y col., *Current Protocols in Molecular Biology*, Green Publishing Associates and Wiley Intersciences, N.Y., 1989; Innis y col., *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., N.Y. 1990; Harlow y col., *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, 1988; Gait, *Oligonucleotide Synthesis*, IRL Press Oxford, 1984; R.I., Freshney, *Animal Cell Culture*, 1987; Colligan y col., *Current Protocols in Immunology*, 1991.

Ejemplo 1

Cribado por homología del dominio extracelular para identificar polipéptidos nuevos y el cADN que los codifica

Las secuencias del dominio extracelular (ECD) (incluyendo la secuencia de señal secretora, si la hubiera) desde aproximadamente 950 proteínas secretadas conocidas de la base de datos Swiss-Prot se utilizaron para buscar bases de datos EST. Las bases de datos de EST incluyeron las bases públicas (por ejemplo, Dayhoff, GenBank), y las bases de datos privadas (por ejemplo, LIFESEQ[®], Incyte Pharmaceuticalas, Palo Alto, CA). La búsqueda se realizó utilizando el programa informático BLAST o BLAST-2 (Altschul y col., *Methods in Enzymology*, 266:460-480 (1996)) como una comparación de las secuencias proteicas de ECD a 6 pautas de traducción de las proteínas EST. Aquellas comparaciones con una valoración de 70 (o en algunos casos 90) o superior que no codifican proteínas conocidas se agruparon y ensamblaron en secuencias de ADN consenso con el programa "phrap" (Phil Green, University of Washington, Seattle, Washington).

Mediante la utilización de este cribado por homología del dominio extracelular, se ensamblaron las secuencias consenso de ADN en relación con otras secuencias EST identificadas utilizando phrap. Además, las secuencias de ADN consenso obtenidas a menudo (pero no siempre) se ampliaron utilizando ciclos repetidos de BLAST o BLAST-2 y phrap para prolongar la secuencia consenso tan lejos como fuera posible mediante la utilización de fuentes de secuencias EST tal como se indicó más arriba.

A continuación, se sintetizaron los oligonucleótidos basados en las secuencias consenso obtenidas tal como se describió más arriba, y se utilizaron para identificar mediante PCR una librería de cADN que contenía las secuencia de interés y para utilizar como sondas para aislar un clon de la secuencia codificante de longitud completa para un polipéptido PRO. Los cebadores de PCR de sentido de transcripción 5' y 3' se hallaron generalmente en el intervalo de 20 a 30 nucleótidos y a menudo se diseñan para proporcionar un producto de PCR de aproximadamente 100-1000 pb de longitud. Las secuencias de la sonda son típicamente de una longitud de 40-55 pb. En algunos casos, se sintetizan oligonucleótidos adicionales cuando la secuencia consenso es superior a aproximadamente 1-1,5 Kb. Con el fin de cribar diversas librerías para un clon de longitud completa, el ADN de las librerías se rastreó mediante amplificación de ADN, tal como se especifica por Ausubel y col., *Current Protocols in Molecular Biology*, con la pareja de cebadores. Una librería positiva se utilizó a continuación para aislar los clones que codifican el gen de interés mediante la utilización del oligonucleótido de la sonda y uno de los cebadores de la pareja.

ES 2 279 473 T3

Las librerías de cADN se utilizaron para aislar los clones de cADN mediante procedimientos estándares mediante la utilización comercial de los reactivos disponibles tal como los de Invitrogen, San Diego, CA. Los cADNs se cebaron con el oligodT que contenía una diana NotI, unida por extremos romos con adaptadores semifosforilados SallI, cortados con NotI, de tamaño comprobado por electroforesis en gel y clonados en una orientación definida en un vector adecuado de clonaje (como pRKB o pRKD; pRK5B es un precursor de pRK5D que no contiene la diana SfiI; véase Holmes y col., Science, 253:1278-1280 (1991)) en las dianas únicas XhoI y NotI.

Ejemplo 2

10 *Aislamiento de clones de cADN mediante la utilización del análisis de algoritmos de señal*

Se identificaron diversas secuencias de ácido nucleico que codifican los diversos polipéptidos mediante la aplicación de un algoritmo de búsqueda de secuencia señal desarrollado por Genentech, Inc., (South San Francisco, CA) para ESTs así como para los fragmentos de EST agrupados y ensamblados a partir de bases de datos públicas (por ejemplo, GenBank) y/o privadas (LIFESEQ[®], Incyte Pharmaceuticals, Inc., Palo Alto, CA). El algoritmo de secuencia señal computa una valoración de secuencia señal basado en el carácter de los nucleótidos del ADN que rodean el primer(os) y opcionalmente el segundo(s) codón (codones) metionina (ATG) en el extremo 5' de la secuencia o el fragmento de secuencia a considerar. Los nucleótidos siguientes al primer(o) codón ATG deben codificar para al menos 35 aminoácidos no ambiguos sin ningún codón de parada. Si la primera pauta de lectura con el primer ATG contiene los aminoácidos requeridos, la segunda ya no se examina. Sin ninguna alcanza los requerimientos, la secuencia candidata no se valora. Con el fin de determinar si la secuencia EST contiene una secuencia señal auténtica, el ADN y las secuencias aminoacídicas correspondientes que rodean el codón ATG se valoran utilizando un grupo de siete sensores (parámetros de evaluación) conocidos por estar asociados con las señales de secreción. La utilización de este algoritmo dio como resultado la identificación de numerosas secuencias de ácido nucleico que codifican los polipéptidos.

25 Ejemplo 3

Aislamiento de los clones de cADN que codifican el PRO5775 humano

30 El ADN96869-2673 se identificó mediante aplicación del algoritmo de búsqueda de secuencia señal descrito en el Ejemplo 2 anterior. La utilización del algoritmo de secuencia señal descrito más arriba permitió la identificación de una secuencia de grupo EST a partir de la base de datos LIFESEQ[®], Incyte Pharmaceuticals, Palo Alto, CA, denominada aquí como CLU86443. Esta secuencia de grupo de EST se comparó a continuación con diversas bases de datos de secuencias expresadas dianas (EST) incluidas en bases de datos públicas (por ejemplo, GenBank) y una base de datos de EST ADN privada (LIFESEQ[®], Incyte Pharmaceuticals, Palo Alto, CA) para identificar homologías existentes. La búsqueda de homología se realizó utilizando el programa informático BLAST o BLAST2 (Altschul y col., Methods in Enzymology, 266:460-480 (1996)). Estas comparaciones que dieron como resultado un valor en BLAST de 70 (o en algunos casos 90) o superior que no codificaban proteínas conocidas se reunieron en una secuencia de ADN consenso con el programa "phrap" (Phil Green, University of Washington, Seattle, Washington). La secuencia consenso así obtenida se denominó ADN79860.

45 A la luz de una homología de secuencia observada entre las secuencia ADN79860 y una secuencia EST de Incyte abarcada por el clon no. 1614726H1 de la base de datos LIFESEQ[®], Incyte Pharmaceuticals, Palo Alto, CA, se obtuvo el clon n° 1614726H1 y se obtuvo y secuenció el inserto de cADN. La secuencia de este inserto de cADN se muestra en la Figura 31 (SEC ID NO:31) y se denominó aquí como ADN96869-2673.

La secuencia codificante completa del ADN96869-2673 se incluye en la Figura 1 (SEC ID N:1). El clon ADN96869-2673 contiene una única pauta abierta de lectura con un sitio aparente de inicio de transcripción en las posiciones nucleotídicas 193-195 y que finaliza en el codón de parada en las posiciones nucleotídicas 1660-1662 (Figura 1). El precursor polipeptídico predicho es de 489 aminoácidos (Figura 2; SEC ID NO:2). La proteína PRO5775 de tamaño completo que se muestra en la Figura 32 tiene un peso molecular estimado de aproximadamente 53.745 daltons y un pI de aproximadamente 8,36. El análisis de la secuencia PRO5775 de tamaño completo que se muestra en la Figura 2 (SEC ID NO:2) evidencia la presencia de una variedad importante de dominios polipeptídicos, en donde las localizaciones dadas para aquellos dominios polipeptídicos importantes son aproximadamente tal como se describió más arriba. El análisis de la secuencia PRO5775 de tamaño completo que se muestra en la Figura 2 pone en evidencia la presencia de: un péptido señal desde aproximadamente el aminoácido 1 al 29; un dominio transmembrana desde aproximadamente el aminoácido 381 al 399; sitios de N-glicosilación desde aproximadamente el aminoácido 133 al 137, desde aproximadamente el aminoácido 154 al 158, desde aproximadamente el aminoácido 232 al 236, desde aproximadamente el aminoácido 264 al 268, desde aproximadamente el aminoácido 386 al 390, desde aproximadamente el aminoácido 400 al 404, desde aproximadamente el aminoácido 410 al 414, desde aproximadamente el aminoácido 427 al 431; y sitios de N-miristoilación desde aproximadamente el aminoácido 58 al 64, desde aproximadamente el aminoácido 94 al 100, desde aproximadamente el aminoácido 131 al 137, desde aproximadamente el aminoácido 194 al 200, desde aproximadamente el aminoácido 251 al 257, desde aproximadamente el aminoácido 277 al 283, desde aproximadamente el aminoácido 281 al 287, desde aproximadamente el aminoácido 361 al 367, desde aproximadamente el aminoácido 399 al 405, desde aproximadamente el aminoácido 440 al 446, desde aproximadamente el aminoácido 448 al 454, desde aproximadamente el aminoácido 478 al 484. El clon ADN96869-2673 se ha depositado en el ATCC el 22 de Junio de 1999 y se le ha asignado el n° PTA-255 del depósito de ATCC.

ES 2 279 473 T3

Un análisis de la base de datos Dayhof (versión 35.45 SwissProt35), utilizando un análisis de alineamiento de secuencias WU-BLAST2 de la secuencia de tamaño completo que se muestra en la Figura 2 (SEC ID NO:2), evidenció la identidad de secuencia entre la secuencia aminoacídica PRO5775 y las secuencias Dayhoff siguientes: U94848_12, P_W57899, CV41KBPL_33, HSU60644_1, CVORF1L5L_3, VK04_VACCV, CVGRI90_41, VK04_VACCC y AF026124_1.

Ejemplo 4

Amplificación génica

Este ejemplo muestra que los genes que codifican PRO5775 se amplifican en el genoma de ciertos cánceres y/o líneas celulares humanas de pulmón, colon y/o mama. La amplificación se asocia con la sobreexpresión del producto génico lo que indica que los polipéptidos son dianas útiles para la intervención terapéutica en algunos cánceres como el colon, pulmón, mama y otros cánceres. Los agentes terapéuticos pueden tomar la forma de polipéptidos antagonistas de PRO5775, por ejemplo, anticuerpos quiméricos murinos-humanos, humanizados o humanos contra un polipéptido PRO5775.

El material de inicio para el cribado fue el ADN genómico aislado a partir de diversos cánceres. El ADN se cuantifica de forma precisa, por ejemplo, fluorimétricamente. Como un control negativo, el ADN se aisló de las células de 10 individuos sanos normales que se agruparon y utilizaron como controles del ensayo para la copia génica en individuos sanos (datos no mostrados). El ensayo de la nucleasa 5' (por ejemplo TaqMan™) y la PCR cuantitativa a tiempo real (por ejemplo, ABI Prism 7700 Sequence Detection System™ (Perkin Elmer, Applied Biosystems Division, Foster City, CA)) se utilizaron para hallar genes potencialmente amplificados en algunos cánceres. Los resultados se utilizaron para determinar si el ADN que codifica PRO5775 está sobreexpresado en cualquiera de los cánceres primarios de pulmón o colon, las líneas celulares de dichos cánceres o las líneas celulares de cáncer de mama que se cribaron. Los cánceres primarios de pulmón se obtuvieron de individuos con tumores del tipo y estadio tal como se indicó en la Tabla 6. Una explicación sobre las abreviaciones utilizadas para la denominación de los tumores primarios listados en la Tabla 6 y en los tumores primarios y las líneas celulares referidas en este ejemplo ya se ha proporcionado más arriba en la presente invención.

Los resultados de la TaqMan™ se expresan en unidades Ct delta (A). Una unidad corresponde a 1 ciclo de PCR o a aproximadamente una amplificación de 2 veces respecto a la normal, dos unidades corresponden a 4 veces, 3 unidades a 8 veces de amplificación y así sucesivamente. La cuantificación se obtuvo utilizando cebadores y una sonda fluorescente TaqMan™ derivada del gen codificante del PRO5775. Las regiones de PRO5775 que seguramente contienen secuencias de ácido nucleico único y que son las menos probables de tener intrones ajustados son las preferidas para generar el cebador y derivar la sonda, por ejemplo, las regiones 3' no traducidas. Las secuencias para los cebadores y las sondas (sentido de transcripción 5' y 3' y la sonda) utilizadas para el análisis de amplificación del gen de PRO5775 fueron las siguientes:

PRO5775 (ADN96869-2673):

96869.tm.f1:

5'-GGGGAACCATFCCAACATC-3'

(SEC ID NO:3)

96869.tm.p1

5'-CCATTCAGCAGGGTGAACCACAG-3'

(SEC ID NO:4)

96869.tmr1

5'-TCTCCGTGACCATGAACTTG-3'

(SEC ID NO:5)

La reacción del ensayo de la nucleasa 5' es una técnica basada en la PCR fluorescente que utiliza la actividad 5'-exonucleasa de la enzima Taq ADN polimerasa para monitorizar la amplificación a tiempo real. Se utilizan dos cebadores oligonucleotídicos para generar un amplicón típico de una reacción de PCR. Un tercer oligonucleótido, o sonda, se diseña para detectar una secuencia nucleotídica localizada entre los dos cebadores. La sonda no es extensible por la enzima Taq ADN polimerasa, y se marca con colorante fluorescente y un colorante fluorescente secuestrador. Cualquier emisión inducida por láser a partir del colorante reportero se secuestra por el colorante secuestrador cuando los dos colorantes se localizan estrechamente próximos como lo están en la sonda. Durante la reacción de amplificación, la enzima Taq ADN polimerasa corta la sonda de forma dependiente del molde. Los fragmentos de la sonda resultantes se disocian en solución, y la señal del colorante reportero se libera del efecto secuestrante del segundo fluoróforo. Una molécula del colorante reportero se libera de cada nueva molécula sintetizada, y la detección del colorante reportero no secuestrado proporciona la base para la interpretación cuantitativa de los resultados.

ES 2 279 473 T3

5 El procedimiento de la nucleasa 5' se realiza en un dispositivo de PCR cuantitativa a tiempo real como el de ABI Prism 7700TM Sequence Detection. El sistema consiste en un termociclador, un láser, una cámara de dispositivo acoplado a la carga (CCD) y un ordenador. El sistema amplifica las muestras en un formato de 96 pocillos en un termociclador. Durante la amplificación, la señal fluorescente inducida por el láser se recoge a tiempo real a través de cables de fibra óptica para los 96 pocillos, y se detecta en la CCD. El sistema incluye el programa para poner en marcha el instrumento y para analizar los resultados.

10 Los resultados del ensayo nucleasa 5' se expresan inicialmente como Ct, o el umbral del ciclo. Ello se define como el ciclo en donde se acumula la señal del reportero sobre el nivel de fondo de la fluorescencia. Los valores de ΔCt se utilizan como cuantificación cuantitativa del número relativo o del número de copias iniciales de una secuencia dada particular en una muestra de ácido nucleico cuando se comparan los resultados del ADN del cáncer con los resultados del ADN humano normal.

15 La Tabla 6 describe el estadio, estadio T y estadio N de diversos tumores primarios que se utilizaron para cribar el compuesto PRO5775 de la invención.

20 (Tabla pasa a página siguiente)

25

30

35

40

45

50

55

60

65

ES 2 279 473 T3

TABLA 6

Perfiles de tumores primarios de pulmón y colon

	Tumor primario	Estadio	Otro estadio	Estadio de Dukes	Estadio T	Estadio N
5	Tumor de pulmón humano AdenoCa (SRCC724) [LT1]	IIA			T1	N1
10	Tumor de pulmón humano SqCCa (SRCC725) [LT1a]	IIB			T3	N0
	Tumor de pulmón humano AdenoCa (SRCC726) [LT2]	IB			T2	N0
	Tumor de pulmón humano AdenoCa (SRCC727) [LT3]	IIIA			T1	N2
	Tumor de pulmón humano AdenoCa (SRCC728) [LT4]	IB			T2	N0
15	Tumor de pulmón humano SqCCa (SRCC729) [LT6]	IB			T2	N0
	Tumor de pulmón humano Aden/SqCCa (SRCC730) [LT7]	IA			T1	N0
	Tumor de pulmón humano AdenoCa (SRCC731) [LT9]	IB			T2	N0
20	Tumor de pulmón humano SqCCa (SRCC732) [LT10]	IIB			T2	N1
	Tumor de pulmón humano SqCCa (SRCC733) [LT11]	IIA			T1	N1
	Tumor de pulmón humano AdenoCa (SRCC734) [LT12]	IV			T2	N0
	Tumor de pulmón humano AdenoSqCCa (SRCC735) [LT13]	IB			T2	N0
25	Tumor de pulmón humano SqCCa (SRCC736) [LT15]	IB			T2	N0
	Tumor de pulmón humano SqCCa (SRCC737) [LT16]	IB			T2	N0
	Tumor de pulmón humano SqCCa (SRCC738) [LT17]	IIB			T2	N1
30	Tumor de pulmón humano SqCCa (SRCC739) [LT18]	IB			T2	N0
	Tumor de pulmón humano SqCCa (SRCC740) [LT19]	IB			T2	N0
	Tumor de pulmón humano LCCa (SRCC741) [LT21]	IIB			T3	N1
	Tumor de pulmón humano AdenoCa (SRCC811) [LT22]	1A			T1	N0
35	Tumor de pulmón humano AdenoCa (SRCC742) [CT2]		M1	D	pT4	N0
	Tumor de pulmón humano AdenoCa (SRCC743) [CT3]			B	pT3	N0
	Tumor de pulmón humano AdenoCa (SRCC744) [CT8]			B	T3	N0
40	Tumor de pulmón humano AdenoCa (SRCC745) [CT10]			A	pT2	N0
	Tumor de pulmón humano AdenoCa (SRCC746) [CT12]		MO, R1	B	T3	N0
	Tumor de pulmón humano AdenoCa (SRCC747) [CT14]		pMO, RO	B	T4	pN0
	Tumor de pulmón humano AdenoCa (SRCC748) [CT15]		M1, r2	D	T4	N2
45	Tumor de pulmón humano AdenoCa (SRCC749) [CT16]		pMO	B	pT3	pN0
	Tumor de pulmón humano AdenoCa (SRCC750) [CT17]			C1	PT3	pN1
	Tumor de pulmón humano AdenoCa (SRCC751) [CT1]		MO, R1	B	pT3	N0
50	Tumor de pulmón humano AdenoCa (SRCC752) [CT4]			B	pT3	N0
	Tumor de pulmón humano AdenoCa (SRCC753) [CT5]		G2	C1	pT3	pN0
	Tumor de pulmón humano AdenoCa (SRCC754) [CT6]		Pmo, RO	B	pT3	pN0
	Tumor de pulmón humano AdenoCa (SRCC755) [CT7]		G1	A	pT2	pNo
55	Tumor de pulmón humano AdenoCa (SRCC756) [CT9]		G3	D	pT4	pN2
	Tumor de pulmón humano AdenoCa (SRCC757) [CT11]			B	T3	N0
	Tumor de pulmón humano AdenoCa (SRCC758) [CT18]		MO, RO	B	pT3	pN0

60

Preparación del ADN

El ADN se preparó a partir de líneas celulares cultivadas, tumores primarios y sangre humana normal. El aislamiento se realizó utilizando el equipo de purificación, tampón y proteasas y todo el resto de Qiagen, de acuerdo con las instrucciones del fabricante y la descripción de más adelante.

65

ES 2 279 473 T3

Lisis del cultivo celular

- Las células se lavaron y tripsinizaron a una concentración de $7,5 \times 10^8$ por punto y se sedimentaron mediante centrifugación a 1000 rpm durante 5 minutos a 4°C, seguido de un nuevo lavado con ½ volumen de PBS y recentrifugación.
- 5 Los sedimentos se lavaron una tercera vez, las células resuspendidas se recogieron y lavaron 2x con PBS. Las células se resuspendieron a continuación en 10 ml de PBS. El tampón C1 se equilibró a 4°C. La proteasa de Qiagen #19155 se diluyó en 6,25 ml de ddH₂O fría a una concentración final de 20 mg/ml y se equilibró a 4°C. Se prepararon 10 ml de tampón G2 diluyendo la solución madre de ARNsa de Qiagen (10 mg/ml) a una concentración final de 200 µg/ml.
- 10 El tampón C1 (10 ml, 4°C) y ddH₂O (40 ml, 4°C) se añadieron a continuación a los 10 ml de suspensión celular, se mezclaron mediante inversión y se incubaron en hielo durante 10 minutos. Los núcleos celulares se sedimentaron mediante centrifugación en un rotor basculante Beckman a 2500 rpm a 4°C durante 15 minutos. El sobrenadante se descartó y los núcleos se resuspendieron con un vórtex en 2 ml de tampón C1 (a 4°C) y 6 ml de ddH₂O, seguido de una centrifugación a 4°C a 2500 rpm durante 15 minutos. Los núcleos se resuspendieron a continuación en el tampón residual con 200 µl por punto. Se añadió el tampón G2 (10 ml) a los núcleos resuspendidos mediante agitación suave con el vórtex. Después de finalizar la adición del tampón, se aplicó un vórtex vigoroso durante 30 minutos. Se añadió la proteasa de Qiagen (200 µl, preparada tal como se indicó más arriba) y se incubó a 50°C durante 60 minutos. La incubación y la centrifugación se repitieron hasta que los lisados estuvieron claros (por ejemplo, mediante una incubación adicional de 30-60 minutos, sedimentación a 3000xg durante 10 min, 4°C).
- 15
- 20

Preparación y lisis de la muestra de tumor sólido humano

- Las muestras de tumor se pesaron y colocaron en 50 ml de tubos cónicos y se mantuvieron en hielo. El procesamiento se limitó a no más de 250 mg de tejido por preparación (1 µl/preparación). La solución de proteasa se preparó extemporáneamente mediante dilución en 6,25 ml de ddH₂O fría a una concentración final de 20 mg/ml y se guardó a 4°C. El tampón G2 (20 ml) se preparó mediante dilución de la ADNsa A a una concentración final de 200 mg/ml (a partir de una solución madre de 100 mg/ml). El tejido tumoral se homogeneizó en 19 ml de tampón G2 durante 60 segundos mediante la utilización de la punta grande del politrón en una campana de flujo laminar TC con el fin de evitar inhalaciones de los aerosoles, y se mantuvo a temperatura ambiente. Entre las muestras, el politrón se lavó mediante centrifugación rápida 2x30 segundos cada en 2 l ddH₂O, seguido de tampón G2 (50ml). Si todavía había tejido en la punta del generador, el aparato se desmontó y limpió.
- 25
- 30

- Se añadió proteasa de Qiagen (preparada tal como se indicó anteriormente, 1,0 ml), seguido de un vórtex e incubación a 50°C durante 3 horas. La incubación y la centrifugación se repitieron hasta que los lisados estuvieron limpios (por ejemplo, incubando 30-60 minutos adicionales, sedimentando a 3000xg durante 10 min, a 4°C).
- 35

Preparación y lisis de sangre humana

- La sangre se obtuvo de voluntarios sanos mediante la utilización de protocolos estériles estándares y adición de citrato en el tubo para muestras de 10 ml. La proteasa de Qiagen se preparó extemporáneamente mediante dilución en 6,25 ml de ddH₂O fría a una concentración final de 20 mg/ml y se guardó a 4°C. El tampón G2 se preparó mediante dilución de la ARNsa a una concentración final de 200 µg/ml a partir de la solución madre de 100 mg/ml. La sangre (10 ml) se colocó en un tubo cónico de 50 ml y se añadieron 10 ml de tampón C1 y 30 ml de ddH₂O (previamente equilibrados a 4°C), se mezcló todo mediante inversión y se mantuvo en hielo durante 10 minutos. Los núcleos se sedimentaron con un rotor basculante Beckman a 2500 rpm, 4°C durante 15 minutos y se descartaron los sobrenadantes. Con ayuda del vórtex, se resuspendieron los núcleos en 2 ml de tampón C1 (4°C) y 6 ml de ddH₂O (4°C). Se volvió a mezclar con ayuda del vórtex tantas veces como fue necesario hasta lograr que el sedimento fuera blanco. A continuación, se resuspendieron los núcleos en el tampón residual con una punta de 200 µl. Se añadió tampón G2 (10 ml) a los núcleos resuspendidos mientras se agitaba suavemente con ayuda del vórtex, seguido de un vigoroso vórtex de 30 segundos. Se añadió la proteasa de Qiagen (200 µl) y se incubó a 50°C durante 60 minutos. La incubación y la centrifugación se repitieron hasta que los lisados estuvieron claros (por ejemplo, incubando 30-60 minutos adicionales, sedimentando a 3000xg durante 10 min, 4°C).
- 40
- 45
- 50

Purificación de lisados claros

- 55 (1) *Aislamiento del ADN genómico*

- El ADN genómico se equilibró (1 muestra por preparación de maxi) con 10 ml de tampón QBT. El tampón de elución QF se equilibró a 50°C. Las muestras se agitaron con ayuda del vórtex durante 30 s, luego se cargaron en columnas equilibradas y se dejaron caer por gravedad. Las columnas se lavaron con 2x15 ml de tampón QC. El ADN se eluyó en tubos Corex de 30 ml, silanizados y autoclavados, con 15 ml de tampón QF (50°C). Se añadió isopropanol (10,5 ml) a cada muestra, se taparon los tubos con película de parafina y se mezclaron repetidamente mediante inversión hasta que el ADN estuviera precipitado. Las muestras se sedimentaron mediante centrifugación en el rotor SS34 a 15.000 rpm durante 10 min a 4°C. Se marcó la localización del sedimento, se descartó el sobrenadante y se añadieron 10 ml de etanol al 70% (4°C). Las muestras se sedimentaron de nuevo por centrifugación en el rotor SS34 a 10.000 rpm durante 10 minutos a 4°C. La localización del sedimento se marcó y se eliminó el sobrenadante. Los tubos se colocaron en su rejilla de secado durante 10 minutos a 37°C, teniendo cuidado de no secar excesivamente las muestras.
- 60
- 65

ES 2 279 473 T3

Después del secado, los sedimentos se disolvieron en 1,0 ml de TE (pH 8,5) y se colocaron a 50°C durante 1-2 horas. Las muestras se mantuvieron a 4°C mientras continuó la disolución. La solución de ADN se transfirió a continuación a tubos de 1,5 ml con una aguja de diámetro 26 en una jeringa de tuberculina. La transferencia se repitió 5x con el fin de desmenuzar el ADN. Las muestras se colocaron a continuación a 50°C durante 1-2 horas.

5

(2) Cuantificación de ADN genómico y preparación del ensayo de amplificación génica

Los niveles de ADN en cada tubo se cuantificaron mediante espectrofotometría A260/A280 a una dilución 1:20 (5 μ l de ADN + 95 μ l de ddH₂O) mediante la utilización de cubetas de cuarzo de 0,1 ml en el espectrofotómetro Beckman DU640. Los cocientes A260/A280 se hallaron en el intervalo de 1,8-1,9. Cada muestra de ADN se diluyó posteriormente a aproximadamente 200 ng/ml en TE (pH 8,5). Si el material original estaba altamente concentrado (aproximadamente 700 ng/ μ l), el material se colocó a 50°C durante varias horas hasta su resuspensión.

La cuantificación del ADN fluorimétrica se realizó sobre el material diluido (20-600 ng/ml) según las instrucciones del fabricante tal como se modificaron más adelante. Ello se logró permitiendo que un fluorímetro Hoeffer DyNa Quant 200 se calentara durante aproximadamente 15 minutos. La solución de trabajo del colorante Hoechst (#H33258, 10 μ l, preparada en las 12 horas de su uso) se diluyó en 100 ml x tampón TNE. Se rellenó una cubeta de 2 ml con la solución del fluorímetro, se colocó en la máquina y se estableció el cero. Se añadió pGEM 3Zf(+) (2 μ l, lote #360851026) a 2 ml de la solución del fluorímetro y se calibró a 200 unidades. Se analizaron a continuación 2 μ l más del ADN de pGEM 3Zf(+) y la lectura se confirmó a 400 +/-10 unidades. Cada muestra se leyó a continuación por triplicado. Cuando las tres muestras se hallaron entre el 10% unas de otras, se calculó el promedio y dicho valor se utilizó como el valor de cuantificación.

La concentración determinada por fluorimetría se utilizó para diluir cada muestra a 10 ng/ μ l en ddH₂O. Ello se realizó simultáneamente en todas las muestras del molde para una sola placa de ensayo TaqMan™, y con material suficiente para realizar de 500-1000 ensayos. Estas muestras se analizaron por triplicado con cebadores TaqMan™ y se probaron con β -actina y GAPDH en una sola placa con ADN humano normal y controles sin molde. Las muestras diluidas se utilizaron siempre que el valor CT del ADN humano normal sustraído del ADN a analizar fuera +/- 1 Ct. El ADN genómico cualificado del lote y diluido se almacenó en alícuotas de 1,0 ml a -80°C. Las alícuotas se que a continuación iban a utilizarse en los ensayos de amplificación génica se guardaron a 4°C. Cada alícuota de 1 ml es suficiente para 8-9 placas o 64 tests.

Ensayo de amplificación génica

El compuesto PRO5775 de la invención se cribó en los tumores primarios siguientes y los valores de Δ Ct resultantes se plasmaron en la Tabla 7B.

TABLA 7B

40

Valores Δ Ct en tumores primarios de pulmón y colon y en modelos de líneas celulares	
Tumor primario	PRO5775
HF-000 631	1,97 1,70
HF-000 641	1,90 1,87
HF-000 643	1,13 1,21
HF-000 840	3,64 3,55
HF-000 842	2,56 2,42 2,12 2,88
HBL100	-
MB435s	-

65

ES 2 279 473 T3

	T47D	-
5	MB468	-
	MB175	-
	MB361	-
10	BT20	-
	MCF7	
	SKBR3	
15	SW480	
	SW620	
20	Colo320	
	HT29	
	HM7	
25	WiDr	
	HCT116	
	SKCO1	
30	SW403	
	LS174T	
35	Colo205	
	HCT15	
	HCC	
40	2998	
	KM12	
	A549	
45	Calu-1	
	Calu-6	
50	H157	
	H441	
	H460	
55	SKMES1	
	SW900	
	H522	
60	H810	
	SRCC	
65	1094	

ES 2 279 473 T3

5	SRCC 1095	
	SRCC 1096	
10	SRCC 1097	
15	SRCC 1098	
20	SRCC 1099	
	SRCC 1100	
25	SRCC 1101	
30	HF-000 545	
	HF-000 499	
35	HF-000 539	
40	HF-000 575	
45	HF-000 698	
	HF-000 756	
50	HF-000 762	2,01 1,04
55	HF-000 789	1,30 1,12
60	HF-000 795	
	HF-000 811	1,82 1,80
65	HF-000	

ES 2 279 473 T3

5	755	
	CT2	
	CT3	
	CT8	
10	CT10	
	CT12	
	CT14	
15	CT15	
	CT16	
20	CT17	
	CT1	
	CT4	
25	CT5	
	CT6	
	CT7	
30	CT9	
	CT11	
	CT18	
35	CT25	
	CT28	
	CT35	
40	HF-000 611	
45	HF-000 613	
	HF-00 1291	
50	HF-00 1293	2,12 2,09
55	HF-00 1294	2,15 1,99
	HF-00 1295	1,99 2,15
60	HF-00	4,62

65

ES 2 279 473 T3

	1296	4,78
5	HF-00 1297	
10	HF-00 1299	1,92 1,95
15	HF-00 1300	
	LT7	
	LT27	
20	LT13	
	LT1	
	LT2	
25	LT3	
	LT4	
	LT9	
30	LT12	
	LT22	
35	LT30	
	LT33	
	LT8	
40	LT21	
	LT1a	
	LT6	
45	LT10	
	LT11	
50	LT15	
	LT16	
	LT17	
55	LT18	
	LT19	
	LT26	
60	LT28	
	LT29	
65	LT31	

ES 2 279 473 T3

5	HF-000 854	
	HF-000 855	
10	HF-000 856	
15	HF-000 831	
	HF-000 832	
20	HF-000 550	
25	HF-000 551	
30	HF-000 733	
35	HF-000 716	

PRO5775 (ADN96869-2673)

Los valores ΔCt para el ADN96869-2673 en diversos tumores están indicados en la Tabla 7B. Un $\Delta Ct > 1$ se utilizó de forma típica como el valor umbral para la valoración de la amplificación, ya que representa doblar el número de copias génicas. La Tabla 7B indica que la amplificación significativa del ácido nucleico ADN96869-2673 que codifica PRO5775 tuvo lugar en: (1) en tumores primarios de pulmón: HF-000631, HF-000641, HF-000840, HF-001293, HF001294, HF-001295, HF-001296 y HF-001299; y (2) en tumores primarios de colon: HF-000762, HF-000789 y HF-000811.

Debido a que la amplificación del ADN96869-2673 tuvo lugar en diversos tumores, es altamente probable que desempeñe un papel importante en la formación o crecimiento del tumor. Como consecuencia, sería lógico esperar que los antagonistas (por ejemplo, los anticuerpos) dirigidos contra la proteína codificada por el ADN96869-2673 (PRO5775) que tengan una utilidad en la terapia contra el cáncer.

Ejemplo 5

Hibridación *in situ*

La hibridación *in situ* es una técnica potente y versátil para la detección y localización de secuencias de ácido nucleico dentro de la célula o en preparaciones de tejido. Puede ser útil, por ejemplo, identificar los sitios de la expresión génica, analizar la distribución tisular de la transcripción, la identidad y localizar la infección viral, seguir los cambios en la síntesis del mRNA específico y ayudar en el mapeo cromosómico.

La hibridación *in situ* se realizó según una versión optimizada del protocolo de Lu y Gillett, Cell Vision, 1:169-176 (1994), mediante ribosondas marcadas con P32 y generadas por PCR. En resumen, los tejidos humanos embebidos en parafina y fijados con formalina se seccionan, desparafinan, desproteinizan en proteinasa K (20 mg/ml) durante 15 minutos a 37°C, y además se procesan para la hibridación *in situ* tal como se describió por Lu y Gillett, *supra*. Una ribosonda antisentido marcada con UTP-P³² se generó a partir de un producto de PCR y se hibridó a 55°C durante toda la noche. Los portaobjetos se sumergieron en emulsión nuclear NTB2 de Kodak y se expusieron durante 4 semanas.

ES 2 279 473 T3

Síntesis de ribosondas-P³²

Se secaron al vacío 6,0 μ l de UTP-P32 (125 mCi) (Amersham BF 1002, SA<2000 Ci/mmol). A cada uno de los tubos que contenía el UTP-P32 seco se añadieron los ingredientes siguientes:

- 5 20 μ l de tampón de transcripción 5X
 1,0 μ l de DTT (100 mM)
10 2,0 μ l de NTP mix (2,5 mM:10 μ l de cada uno a 10 mM: GTP, CTP & ATP + 10 μ l H₂O)
 1,0 μ l de UTP (50 μ M)
 1,0 μ l de ARNsin
15 1,0 μ l de ADN molde (1 μ g)
 1,0 μ l de H₂O
20 1,0 μ l de ARN polimerasa (para productos de PCR T3 = AS, T7 = S, generalmente)

Los tubos se incubaron a 37°C durante una hora. Se añadió un total de 1,0 μ l de RQ1 ADNsa, seguido de incubación a 37°C durante 15 minutos. Se añadió un total de 90 μ l de TE (Tris 10 mM pH 7,6/EDTA 1 mM, pH 8,0), y la mezcla se pipeteó en papel DE81. La solución restante se cargó en una unidad de ultrafiltración MICROCON-50, y se centrifugó con el programa 10 (6 minutos). La unidad de filtración se invirtió en un segundo tubo y se centrifugó con el programa 2 (3 minutos). Después de la centrifugación rápida final, se añadieron un total de 100 μ l de TE, luego se pipeteó 1 μ l del producto final en papel DE81 y se contaron en 6 ml de BIOFLUORII™.

30 La sonda se migró en un gel TBE/urea. Un total de 1-3 μ l de la sonda o 5 μ l de ARN Mrk III se añadieron a 3 μ l de tampón de carga. Después de calentar en un bloque térmico a 95°C durante 3 minutos, el gel se colocó inmediatamente en hielo. Los pocillos del gel se nivelaron, se cargó la muestra y se migró a 180-250 voltios durante 45 minutos. El gel se envolvió con un plástico (marca SARAN™) y se expuso a una película XAR con una pantalla intensificadora en un congelador a -70°C durante una hora toda la noche.

35

Hibridación-P³²

A. Pretratamiento de secciones congeladas

40

Las preparaciones se sacaron del congelador, se colocaron en bandejas de aluminio, y se descongelaron a temperatura ambiente durante 5 minutos. Las bandejas se colocaron en un incubador de 55°C durante 5 minutos para reducir la condensación. Las preparaciones se fijaron durante 10 minutos en paraformaldehído al 4% en hielo en la campana de flujo, y se lavaron en 0,5XSSC durante 5 minutos, a temperatura ambiente (25 ml de 20XSSC + 975 ml SQ H₂O). Después de la desproteínización en 0,5 μ g/ml de proteinasa K durante 10 minutos a 37°C (12,5 μ l de una solución madre a 10 mg/ml en 250 ml de tampón libre de ARNsa precalentado), las secciones se lavaron en 0,5XSSC durante 10 minutos a temperatura ambiente. Las secciones se deshidrataron en series de etanol al 70%, 95% y 100% durante 2 minutos cada vez.

B. Pretratamiento de las secciones embebidas en parafina

Las preparaciones se desparafinaron, se colocaron en SQH₂O, y se lavaron dos veces en 2XSSC a temperatura ambiente, durante 5 minutos cada vez. Las secciones se desproteínizaron en 20 μ g/ml de proteinasa K (500 μ l de 10 mg/ml en 250 ml de tampón ARNsa libre de ARNsa a 37°C, 15 min) para el tejido embrionario humano, o 8xproteinasa K (100 μ l en 250 ml de tampón ARNsa, 37°C, 30 min) para tejidos en formalina. Los siguientes lavados en 0,5XSSC y la deshidratación se realizaron tal como se describió anteriormente.

55

C. Prehibridación

60 Los portaobjetos se colocaron en una caja de plástico con Tampón Box (4XSSC, formamida al 50%)-papel de filtro saturado. El tejido se cubrió con 50 μ l de tampón de hibridación (3,75 g de sulfato de dextrano + 6 ml de SQH₂O), se realizó un vórtex, y se calentó en el microondas durante 2 minutos con la tapa semiabierta. Después de enfriarse en hielo, se añadieron 18,75 ml de formamida, 3,75 ml de 20XSSC y 9 ml de SQH₂O, se realizó un vórtex y el tejido se incubó a 42°C durante 1-4 horas.

65

ES 2 279 473 T3

D. Hibridación

Se calentaron a 95°C durante 3 minutos, $1,0 \times 10^6$ cpm de la sonda y 1,0 μ l de tARN (50 mg/ml de solución madre). Las preparaciones se enfriaron en hielo y se añadió 48 μ l de tampón por preparación. Después del vórtex, se añadió 50 μ l de P³³ mix a 50 μ l de prehibridación en la preparación. Las preparaciones se incubaron durante toda la noche a 55°C.

E. Lavados

Los lavados se realizaron durante 2x10 min con 2XSSC, EDTA a temperatura ambiente (400 ml 20cSSC +16 ml de EDTA 0,25 M, Vf=4l), seguido por un tratamiento con ARNsa A a 37°C durante 30 minutos (500 μ l de 10 mg/ml en 250 ml de tampón ARNsa = 20 μ g/ml). Las preparaciones se lavaron 2x10 min con 2XSSC, EDTA a temperatura ambiente. Las condiciones de astringencia de los lavados fueron: 2 horas a 55°C, 0,1XSSC, EDTA (20 ml de 20XSSC + 16 ml EDTA, Vf=4l).

Ejemplo 6

Utilización de PRO5775 como una sonda de hibridación

El procedimiento siguiente describe la utilización de una secuencia nucleotídica que codifica un polipéptido PRO5775 como una sonda de hibridación.

El ADN que comprende la secuencia codificante de un polipéptido "PRO5775" de tamaño completo o maduro tal como se describe aquí y/o los fragmentos de lo mismo pueden utilizarse como una sonda para cribar los ADNs homólogos (como los que codifican de forma natural las variantes de PRO5775 en las librerías de cADN de tejido humano o las librerías genómicas de tejidos humanos).

La hibridación y lavado de las membranas que contienen los ADNs de las librerías se realiza en las siguientes condiciones de astringencia elevada. La hibridación de la sonda derivada de PRO5775 radiomarcada con las membranas se realiza en una solución de formamida al 50%, 5XSSC, SDS al 0,1%, pirofosfato sódico al 0,1%, fosfato sódico 50 mM, pH 6,8, solución de 2xDenhardt y sulfato de dextrano al 10% a 42°C durante 20 horas. El lavado de las membranas se realiza en una solución acuosa de 0,1XSSC y SDS al 0,1% a 42°C.

Los ADNs que tienen una identidad de secuencia deseada con el ADN que codifica la secuencia PRO5775 nativa de tamaño completo pueden identificarse mediante técnicas estándares conocidas en la materia.

Ejemplo 7

Expresión de los polipéptidos PRO5775 en *E. coli*

Este ejemplo ilustra la preparación de una forma no glicosilada de PRO5775 mediante expresión recombinante en *E. coli*.

La secuencia de ADN que codifica el polipéptido PRO de interés se amplificó inicialmente con cebadores de PCR seleccionados. Los cebadores deben contener dianas de enzimas de restricción que corresponden con las dianas de enzimas de restricción en el vector de expresión seleccionado. Pueden utilizarse diversos vectores de expresión. Un ejemplo de un vector es pBR322 (derivado de *E. coli*; véase Bolivar y col., Gene, 2:95 (1977)) que contiene genes para la resistencia a ampicilina y tetraciclina. El vector se digiere con la enzima de restricción y se desfosforila. Las secuencias de PCR amplificadas se ligan a continuación en el vector. El vector incluirá preferentemente secuencias que codifican un gen de resistencia a un antibiótico, un promotor trp, un líder poli-His (incluyendo los 6 primeros codones STII, la secuencia poli-His, y la diana de clonaje), la región codificante de PRO5775, el terminador transcripcional de lambda y un gen argU.

La mezcla de ligación se utiliza a continuación para transformar una cepa de *E. coli* mediante procedimientos descritos en Sambrook y col., *supra*. Los transformantes se identificaron por su capacidad para crecer en placas de LB y a continuación se seleccionaron las colonias resistentes al antibiótico. El ADN plasmídico puede aislarse, comprobarse por análisis de restricción y secuenciarse.

Los clones seleccionados pueden crecerse durante toda la noche en un medio de cultivo líquido como el LB suplementado con antibióticos. Los cultivos de toda la noche pueden consiguientemente utilizarse para inocular un cultivo a mayor escala. Las células se crecen a la densidad óptica deseada, tiempo durante el cual se activa el promotor de expresión.

Después de cultivar las células durante unas horas más, éstas pueden cosecharse mediante centrifugación. El sedimento celular obtenido por la centrifugación puede solubilizarse con diversos agentes conocidos en la materia, y la proteína PRO5775 puede a continuación purificarse con una columna quelante de metales en condiciones que permitan la unión fuerte de la proteína.

ES 2 279 473 T3

Las proteínas pueden expresarse en *E. coli* en una forma etiquetada con un poli-His según el procedimiento siguiente. El ADN que codifica la proteína se amplifica inicialmente con cebadores de PCR seleccionados. Los cebadores contienen dianas de enzimas de restricción que corresponden con las dianas de enzimas de restricción en el vector de expresión seleccionado, y también se proporcionan otras secuencias de utilidad para un inicio de traducción eficiente y fiable, una purificación rápida en una columna de quelación de metales y la eliminación proteolítica con enteroquinasa. Las secuencias etiquetadas con poli-His amplificadas por PCR se ligan a continuación en un vector de expresión, que se utiliza para transformar un huésped *E. coli* basado en la cepa 52 (W3110 fuhA(tonA)lon galE rpoHts(htpRts) clpP(lacIq)). Los transformantes se crecen en primer lugar en LB con carbenicilina 50 mg/ml a 30°C en agitación hasta alcanzar una densidad D.O. de 3-5 a 600 nm. Los cultivos se diluyen entonces 50-100 veces en medio CRAP (preparado mezclando 3,57 g de (NH₄)₂SO₄, 0,71 g de citrato sódico.2H₂O, 1,07 g de KCl, 5,36 g de extracto de levadura Difco, 5,36 g de Sheffield hycase SF en 500 ml de agua, así como PMOS 110 mM, pH 7,3, glucosa al 0,55% (p/v) y MgSO₄ 7 mM) y se crecieron por aproximadamente 20-30 horas a 30°C con agitación. Se tomaron muestras para verificar la expresión mediante análisis de SDS-PAGE, y se centrifugó todo el cultivo para sedimentar las células. Los sedimentos celulares se congelaron hasta la purificación y renaturalización.

La pasta de *E. coli* obtenida a partir de fermentaciones de 0,5-1 l (6-10 g de sedimentos) se resuspendieron en 10 volúmenes (p/v) en guanidina 7 M, Tris 20 mM, tampón pH 8. Se añadió sulfito sódico sólido y tetrionato sódico a concentraciones finales de 0,1 M y 0,02M, respectivamente, y la solución se agitó durante toda la noche a 4°C. Esta etapa dio como resultado una proteína desnaturalizada con todos los residuos de cisteína bloqueados por la sulfitolización. La solución se centrifugó a 40.000 rpm en una ultracentrífuga Beckman durante 30 min. El sobrenadante se diluyó con 3-5 volúmenes de tampón de columna quelante de metales (guanidina 6M, Tris 20 mM, pH 7,4) y se filtró a través de un filtro de 0,22 micras hasta su clarificación. El extracto clarificado se cargó en una columna de 5ml de Qiagen Ni²⁺-NTA quelante de metales equilibrada en el tampón de la columna quelante de metales. La columna se lavó con tampón adicional que contenía imidazol 50 mM (Calbiochem, grado Utrol), pH 7,4. Las proteínas se eluyeron con tampón que contenía imidazol. Las fracciones que contenían la proteína deseada se agruparon y guardaron a 4°C. La concentración de proteínas se estimó por su absorbancia a 280 nm utilizando el coeficiente de extinción calculado según su secuencia aminoacídica.

La proteína se renaturaliza diluyendo la muestra en tampón de renaturalización preparado al momento que consiste en: Tris 20 mM, pH 8,6; NaCl 0,3 M; urea 2,5 M; cisteína 5 mM, glicina 20 mM y EDTA 1 mM. Los volúmenes de renaturalización se eligen para que la concentración final se halle entre 50 a 100 microgramos/ml. La solución de renaturalización se agita suavemente a 4°C durante 12-36 horas. La reacción de renaturalización se secuestra mediante adición de TFA a una concentración final de 0,4% (pH de aproximadamente 3). Antes de la última purificación de la proteína, la solución se filtra a través de un filtro de 0,22 micras y se añade acetonitrilo a una concentración final de 2-10%. La proteína renaturalizada se cromatografía en una columna de fase inversa Poros R1/H con un tampón móvil de TFA al 0,1% por elución con un gradiente de acetonitrilo del 10 al 80%. Alícuotas de fracciones con una absorbancia A₂₈₀ se analizan en geles de poliacrilamida-SDS y las fracciones que contienen la proteína renaturalizada homogénea se agrupan. Por lo general, las especies renaturalizadas de manera adecuada de muchas proteínas se eluyen a concentraciones inferiores de acetonitrilo ya que estas especies son las más compactas con sus interiores hidrofóbicos resguardados de la interacción con la resina de fase inversa. Las especies agregadas se eluyen generalmente a concentraciones de acetonitrilo superiores. Además de resolver las formas mal renaturalizadas de las proteínas de la forma deseada, la etapa de fase inversa también elimina la endotoxina de las muestras.

Las fracciones que contenían la proteína renaturalizada deseada se agruparon y se eliminó el acetonitrilo con una corriente suave de nitrógeno dirigida a la solución. Las proteínas se formularon en Hepes 20 mM, pH 6,8 con cloruro sódico 0,14 M y manitol al 4% mediante diálisis o mediante filtración en gel con resina Superfine G25 (Pharmacia) equilibrada en el tampón de formulación y filtrada estérilmente.

Ejemplo 8

Expresión de PRO5775 en células de mamífero

Este ejemplo ilustra la preparación de una forma de glicosilación potencial de PRO5775 mediante la expresión recombinante de las células de mamífero.

El vector, pRK5 (véase la patente europea EP 307.247, publicada el 15 de Marzo de 1989), se utilizó como vector de expresión. Opcionalmente, el ADN de PRO5775 se ligó en pRK5 con enzimas de restricción seleccionadas para permitir la inserción del ADN de PRO5775 utilizando los procedimientos de ligación como los descritos por Sambrook y col., *supra*. El vector resultante se denominó pRK5-PRO5775.

En una realización, las células huésped seleccionadas pueden ser células 293. Las células 293 humanas (ATCC CCL 1573) se crecieron hasta la confluencia en placas de cultivo celular en medio como el DMEM suplementado con suero de ternera fetal y opcionalmente, componentes nutrientes y/o antibióticos. Aproximadamente 10 µg del ADN de pRK5-PRO5775 se mezclaron con aproximadamente 1 mg del ADN que codifica el gen VA ARN [Thimmappaya y col., Cell, 31:543 (1982)] y se disolvió en 500 µl de Tris-HCl 1 mM, EDTA 0,1 mM, CaCl₂ 0,277 M. A esta mezcla se añadió, gota a gota 500 µl de HEPES 50 mM (pH 7,35), NaCl 280 mM, NaPO₄ 1,5 mM, y se permitió la formación de un precipitado durante 10 min a 25°C. El precipitado se resuspendió y se añadió a las células 293 y se dejó reposar durante aproximadamente 4 horas a 37°C. El medio de cultivo se aspiró y se añadieron 2 ml de glicerol al 20% en PBS

ES 2 279 473 T3

durante 30 s. Las células 293 se lavaron a continuación con medio sin suero, se añadió medio nuevo y se incubaron las células durante aproximadamente 5 días.

Al cabo de aproximadamente 24 horas después de las transfecciones, se eliminó el medio de cultivo y se reemplazó con medio (sólo) o con medio y 200 $\mu\text{Ci/ml}$ de cisteína- S^{35} y 200 $\mu\text{Ci/ml}$ de metionina- S^{35} . Al cabo de 12 horas de incubación, el medio condicionado se recogió, se concentró en un filtro de centrifugación, se cargó en un gel al 15% de SDS. El gel procesado puede secarse y exponerse a una película durante un período de tiempo para poner en evidencia la presencia del polipéptido PRO5775. Los cultivos que contenían células transfectadas pueden incubarse durante más tiempo (en medio sin suero) y el medio se analiza en bioensayos seleccionados.

En una técnica alternativa, el ADN de PRO5775 puede introducirse en células 293 de forma transitoria gracias al procedimiento del dextrano de sulfato descrito por Somparvac y col., Proc. Natl. Acad. Sci., 12:7575 (1981). Las células 293 se crecieron a una densidad máxima en un frasco para centrifugado y se añadieron 700 μg del ADN PRK5-PRO5775. Las células se concentraron en primer lugar a partir del frasco centrifugador mediante centrifugación y se lavaron con PBS. El precipitado de ADN-dextrano se incubó con el precipitado celular durante 4 horas. Las células se trataron con glicerol al 20% durante 90 segundos, se lavaron con medio de cultivo de tejidos y se re-incubaron en el frasco de centrifugado que contenía medio de cultivo de tejidos, 5 $\mu\text{g/ml}$ de insulina bovina y 0,1 $\mu\text{g/ml}$ de transferrina bovina. Al cabo de 4 días, el medio condicionado se centrifugó y se filtró para eliminar las células y los restos celulares. La muestra que contenía el PRO5775 expresado puede a continuación expresarse y purificarse mediante un procedimiento seleccionado, como la diálisis y/o la cromatografía en columna.

En otra realización, el PRO5775 puede expresarse en células CHO. El vector pRK5-PRO5775 puede transfectarse en células CHO utilizando reactivos como CAPO₄ o DEAE-dextrano. Tal como se describió más arriba, los cultivos celulares pueden incubarse y el medio puede reemplazarse con medio de cultivo (sólo) o medio que contenga un marcaje radioactivo como la metionina S^{35} . Después de determinar la presencia del polipéptido PRO5775, el medio de cultivo puede reemplazarse con medio sin suero. Preferentemente, los cultivos se incuban durante aproximadamente 6 días, y luego el medio condicionado se cosecha. El medio que contenía el PRO5775 expresado puede concentrarse a continuación y purificarse mediante el procedimiento seleccionado.

El PRO5775 etiquetado epitópicamente puede también expresarse en células CHO. El PRO5775 puede subclonarse fuera del vector pRK5. A continuación, mediante PCR se puede fusionar el inserto del subclón en la misma pauta de lectura con una etiqueta epitópica como una etiqueta poli-His en un vector de expresión en Baculovirus. El inserto PRO5775 etiquetado epitópicamente con poli-His puede subclonarse en un vector dirigido por SV40 con un marcador de selección como el DHFR para la selección de clones estables. Por último, las células CHO pueden transfectarse (tal como se describió anteriormente) con el vector dirigido por SV40. El marcaje puede realizarse, tal como se describió anteriormente, para verificar la expresión. El medio de cultivo que contenía el PRO5775 etiquetado con poli-His puede concentrarse y purificarse mediante un procedimiento seleccionado, como por ejemplo cromatografía de afinidad con quelato de Ni^{2+} . La expresión de CHO y/o de células COS puede también lograrse mediante un procedimiento de expresión transitoria.

Las proteínas pueden expresarse en células CHO mediante un procedimiento de expresión estable o mediante un procedimiento transitorio. La expresión estable en células CHO se realiza mediante el procedimiento siguiente. Las proteínas se expresaron como una construcción de IgG (inmuno adhesina), en la cual las secuencias codificantes para las formas solubles (por ejemplo, dominios extracelulares) de las proteínas respectivas se fusionan a una secuencia de región constante de IgG1 que contiene la bisagra, los dominios CH2 y CH2 y/o una forma etiquetada con poli-His.

Después de la amplificación por PCR, los ADNs respectivos se subclonaron en un vector de expresión de CHO mediante técnicas estándares tal como se describe en Ausubel y col., Current Protocols of Molecular Biology, Unidad 3.16, John Wiley y Sons (1997). Los vectores de expresión en CHO se construyeron para tener dianas de restricción compatibles 5' y 3' del ADN de interés para permitir el traslado de los cADNs. El vector se utilizó para la expresión en células CHO tal como se describe en Lucas y col., Nucl. Acids. Res., 24:9 (1774-1779, (1996)), con el promotor temprano de SV40/intensificador para conducir la expresión del cADN de interés y la dihidrofolato reductasa (DHFR). La expresión de DHFR permite la selección para el mantenimiento estable del plásmido después de la transfección.

Se introdujeron 12 μg del ADN plasmídico en aproximadamente 10 millones de células CHO por medio de reactivos de transfección disponibles comercialmente, Superfect® (Qiagen), Dospert® o Fugene® (Boehringer Mannheim). Las células se crecieron tal como se describe en Lucas y col., *supra*. Aproximadamente, 3×10^7 células se congelan en un vial para su posterior crecimiento y producción tal como se describe más adelante.

Los viales que contenían el ADN plasmídico se descongelaron colocándolos en un baño maría y se mezclaron con el vórtex. Los contenidos se pipetearon en un tubo de centrifuga con 10 ml de medio y se centrifugaron a 1000 rpm durante 5 min. El sobrenadante se aspiró y las células se resuspendieron en 10 ml de medio selectivo (PS20 filtrado con filtro de 0,2 μm y suero bovino fetal filtrado con filtro de 0,2 μm). Las células se alicuotaron en frascos de centrifugado de 100 ml con 90 ml de medio selectivo. Al cabo de 1-2 días, las células se transfectaron en frascos de centrifugado de 250 ml rellenos con 150 ml de medio de crecimiento selectivo y se incubaron a 37°C. Al cabo de 2-3 días, se sembraron frascos de centrifugado de 250 ml, 500 ml y 2000 ml con 3×10^5 células/ml. El medio se cambió con uno nuevo por centrifugación y resuspensión en medio de producción. Aunque puede utilizarse cualquier medio para CHO adecuado, se utilizó un medio de producción descrito en la patente americana U.S. 5.122.469, publicado el

ES 2 279 473 T3

16 de Junio de 1992. Se sembró un frasco de centrifugado de 3 L con $1,2 \times 10^6$ células/ml. El día cero, se determinó el número de células y el pH. El día 1, se toma una muestra del frasco de centrifugado y se inicia la introducción de aire filtrado. El día 2, se toma una muestra del recipiente de cultivo, la temperatura se disminuye a 33°C y se añaden 30 ml de glucosa a 500 g/l y 0,6 ml de tensoactivo (por ejemplo, una emulsión de polidimetilsiloxano al 35%, Dow
5 Corning 365 Medical Grade Emulsion). Durante la producción, el pH se ajusta para mantenerlo a aproximadamente 7,2. Al cabo de 10 días, o hasta que la viabilidad disminuya por debajo del 70%, el cultivo celular se cosecha mediante centrifugación y se filtra a través de una membrana de $0,22 \mu\text{m}$. El filtrado se guarda a 4°C o se carga inmediatamente en columnas para su purificación.

10 Para las construcciones etiquetadas con poli-His, las proteínas se purificaron con una columna NTA-Ni²⁺ (Qia-gen). Antes de la purificación, se añade imidazol al medio condicionado a una concentración de 5 mM. El medio condicionado se bombea en una columna NTA-Ni²⁺ equilibrada con Hepes 20 mM, pH 7,4, tampón que contiene NaCl 0,3 M e imidazol 5 mM a una velocidad de flujo de 4-5 ml/min, a 4°C . Después de la carga, la columna se lavó con tampón de equilibrio adicional y se eluyó la proteína con tampón de equilibrio con tampón con imidazol 0,25 M.
15 La proteína altamente purificada se desaló a continuación en un tampón de almacenamiento con Hepes 10 mM, NaCl 0,14 M y manitol al 4%, pH 6,8, con una columna de 25 ml de G25 Superfine (Pharmacia) y se guardó a -80°C .

Las construcciones de inmunoadhesina (que contenían Fc) se purificaron a partir del medio condicionado de la manera siguiente. El medio condicionado se bombeó en una columna de Proteína A de 5 ml (Pharmacia) que se había
20 equilibrado con tampón fosfato Na 20 mM, pH 6,8. Después de cargar, la columna se lavó extensamente con tampón de equilibrio antes de la elución con ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutralizó inmediatamente recogiendo alícuotas de 1 ml en tubos con $275 \mu\text{l}$ de Tris 1 M, pH 9. La proteína altamente purificada se desaló a continuación en tampón de almacenamiento tal como se describió anteriormente para las proteínas etiquetadas con poli-His. Se valoró la homogeneidad mediante geles de poli-acrilamida-SDS y secuenciación aminoacídica N-terminal
25 por degradación de Edman.

Ejemplo 9

Expresión de PRO5775 en levaduras

30 El procedimiento siguiente describe la expresión recombinante de PRO5775 en levaduras.

En primer lugar, los vectores de expresión se construyeron para la producción intracelular o la secreción de PRO5775 a partir del promotor ADH2/GADPH. El ADN que codifica PRO5775 y el promotor se insertaron en las
35 dianas de restricción adecuadas en el plásmido seleccionado para dirigir la expresión intracelular de pPRO5775. Para la secreción, el ADN que codifica PRO5775 puede clonarse en el plásmido seleccionado, junto con el ADN que codifica el promotor ADH2/GADPH, un péptido señal de PRO5775 nativo u otro péptido señal de mamífero, o, por ejemplo, una secuencia líder/señal del factor alfa de levaduras, o de la invertasa secretora, y una secuencia engarce (si fuera necesario) para la expresión de PRO5775.

40 Las células de levaduras como la cepa de levadura AB110, puede transformarse con los plásmidos de expresión descritos anteriormente en el medio de fermentación seleccionado. Los sobrenadantes de levaduras transformadas pueden analizarse mediante precipitación con ácido tricloroacético al 10% y por separado mediante PAGE-SDS, seguido de una tinción de los geles con azul de Coomassie.

45 PRO5775 recombinante puede a continuación aislarse y purificarse extrayendo las células de levaduras del medio de fermentación mediante centrifugación y luego concentración del medio con filtros de cartucho seleccionados. El concentrado que contiene PRO5775 puede a continuación purificarse mediante resinas seleccionadas de cromatografía en columna.

Ejemplo 10

Expresión de PRO5775 en células de insecto infectadas por Baculovirus

55 El procedimiento siguiente describe la expresión recombinante de células de insecto infectadas con Baculovirus.

La secuencia codificante para PRO5775 se fusiona arriba de una etiqueta epitópica contenida en un vector de expresión de Baculovirus. Dichas etiquetas epitópicas incluyen las etiquetas de poli-His y las etiquetas de inmunoglobulina (como por ejemplo las regiones Fc de IgG). Puede utilizarse una variedad de plásmidos, incluyendo los
60 derivados de plásmidos disponibles comercialmente como pVL 1393 (Novagen). En resumen, la secuencia codificante de PRO5775 o la porción deseada de la secuencia codificante de PRO5775 [tal como la secuencia codificante del dominio extracelular de una proteína transmembrana o la secuencia codificante de la proteína madura si la proteína es extracelular] se amplifica mediante PCR con cebadores complementarios a las regiones 5' y 3'. El cebador 5' puede incorporar dianas de restricción flanqueantes (seleccionadas). El producto se digiere a continuación con las enzimas
65 de restricción seleccionadas y se subclona en el vector de expresión.

El baculovirus recombinante se generó mediante co-transfección del plásmido anterior y del ADN del virus BaculoGold™ (Pharmlingen) en células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711) mediante la utilización de

ES 2 279 473 T3

lipofectina (disponible comercialmente de GIBCO BRL). Al cabo de 4-5 días de incubación a 28°C, los virus liberados se recogen y utilizan para amplificaciones posteriores. La infección viral y la expresión de proteínas se realizaron tal como se describe en O'Reilly y col., *Baculovirus expression vectors: A Laboratory Manual*, Oxford: Oxford University Press (1994).

5 A continuación, es posible purificar el PRO5775 etiquetado con poli-His, por ejemplo mediante cromatografía de afinidad con quelato-Ni²⁺ de al manera siguiente. Se preparan los extractos a partir de las células Sf9 infectadas con el virus tal como se describe por Rupert y col., *Nature*, 362:175-179 (1993). En resumen, las células Sf9 se lavan, se resuspenden en tampón de sonicación (25 ml de Hepes, pH 7,9; MgCl₂ 12,5 mM; EDTA 0,1 mM; glicerol al 10%;
10 NP-40 al 0,1%; KCl 0,4 M), y se sonican dos veces durante 20 segundos en hielo. Los sonicados se clarifican por centrifugación, y el sobrenadante se diluye 50 veces en tampón de carga (fosfato 50 mM, NaCl 300 mM, glicerol al 10%, pH 7,8) y se filtran a través de un filtro de 0,45 µm. Se prepara una columna de agarosa NTA-Ni²⁺ (disponible comercialmente por Qiagen) con un volumen de lecho de 5 ml, se lava con 25 ml de agua y se equilibra con 25 ml de tampón de lavado. El extracto celular se carga en la columna a 0,5 ml por minuto. La columna se lava al nivel
15 basal A₂₈₀ con tampón de carga, en cuyo punto de fracción se inicia la recolección. A continuación, se lava la columna con un segundo tampón de lavado (fosfato 50 mM, NaCl 300 mM, glicerol al 10%, pH 6,0), que eluye la proteína unida inespecíficamente. Después de alcanzar el nivel basal A₂₈₀ de nuevo, la columna se revela con un gradiente de imidazol de 0 a 500 mM en el segundo tampón de lavado. Se recogen fracciones de 1 l y se analizan mediante SDS-PAGE y tinción de plata o transferencia de Western con NTA-Ni²⁺ conjugado con fosfatasa alcalina (Qiagen). Las
20 fracciones que contienen el eluido de PRO5775 etiquetado con poli-His, respectivamente, se agrupan y dializan contra el tampón de carga.

Alternativamente, la purificación del PRO5775 etiquetado con IgG (o Fc) puede realizarse mediante técnicas de cromatografía, incluyendo por ejemplo, columnas de cromatografía con Proteína A o proteína G.

25 Mientras que la expresión se realiza en realidad a una escala de 0,5-2 l, puede realizarse fácilmente a escalas mayores (por ejemplo, 8l). Las proteínas se expresan como una construcción IgG (inmunoadhesina) en la que la región extracelular de la proteína se fusiona a una secuencia de región constante de IgG1 que contiene la bisagra, los dominios CH₂ y CH₃ y/o las formas etiquetadas con poli-His.

30 Después de la amplificación por PCR, las secuencias codificantes respectivas se subclonaron en un vector de expresión de baculovirus (pb.PH.IgG pra fusiones IgG y pb.PH.His.c para proteínas etiquetadas con poli-His) y el vector y ADN de baculovirus Baculogold[®] (Pharmigen) se cotransfectaron en 10⁵ células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711), gracias a Lipofectin (Gibco BRL). Pb.PH.IgG y pb.PH.His son modificaciones de el vector de expresión de baculovirus disponible comercialmente pVL1393 (Pharmigen), con regiones del poliengarce modificadas para incluir las secuencias etiqueta His o Fc. Las células se crecieron en medio Hink TNM-FH suplementado con FBS al 10% 8Hycoone). Las células se incubaron durante 5 días a 28°C. El sobrenadante se cosechó y a continuación se utilizó para la primera amplificación viral mediante infección de células Sf9 en medio Hinks TNM-FH suplementado con FBS al 10% a una multiplicidad de infección (MOI) de aproximadamente 10. Las células se incubaron durante 3
40 días a 28°C. El sobrenadante se cosechó y la expresión de las construcciones en el vector de expresión de baculovirus se determinó mediante unión del lote de 1 ml de sobrenadante a 25 l de bolas de NCTA-Ni²⁺ (Qiagen) para las proteínas etiquetadas con histidina o bolas de Protein-A Sepharose CL-4B (Pharmacia) para proteínas etiquetadas con IgG seguido de un análisis por SDS-PAGE comparado con una concentración conocida de proteína estándar por tinción de azul de Coomassie.

45 El sobrenadante de la primera amplificación viral se utilizó para infectar el cultivo del fermentador (500 ml) de células Sf9 en medio ESF-921 (Sistemas de Expresión LLC) a una MOI de aproximadamente 0,1. Las células se incubaron durante 3 días a 28°C. El sobrenadante se cosechó y filtró. La unión del lote y el análisis de SDS-PAGE se repitieron cuantas veces fue necesario hasta confirmar la expresión del cultivo.

50 El sobrenadante de la primera amplificación viral se utilizó para infectar un cultivo en recipiente centrifugador (500 ml) de células Sf9 crecidas en medio ESF-921 (Expression Systems LLC) a aproximadamente una MOI De 0,1. Las células se incubaron durante 3 días a 28°C. El sobrenadante se cosechó y filtró. La unión del lote y del análisis por SDA-PAGE se repitió tantas veces como fue necesario, hasta confirmar la expresión del cultivo en el recipiente centrifugador.

55 El medio condicionado de las células transfectadas (0,5 a 3 l) se cosechó por centrifugación para eliminar las células y se filtró a través de filtros de 0,22 µm. Para las construcciones etiquetadas con poli-His, la construcción de la proteína se purificó con una columna NTA-Ni²⁺ (Qiagen). Antes de la purificación, se añadió imidazol al medio condicionado a una concentración de 5 mM. El medio condicionado se bombeó en una columna de NTA-Ni²⁺ de 6 ml equilibrada con Hepes 20 mM, pH 7,4, con NaCl 0,3 M e imidazol 5 mM a una velocidad de flujo de 4-5 ml/min, a 4°C. Después de la carga, la columna se lavó con tampón adicional de equilibrio y la proteína se eluyó con tampón de equilibrio que contenía imidazol 0,25 M. La proteína altamente purificada se desaló a continuación en un tampón de almacenamiento que contenía Hepes 10 mM, NaCl 0,14 M y manitol al 4%, pH 6,8, con una columna de 25 ml de G25 Superfine (Pharmacia) y se guardó a -80°C.

Las construcciones de proteínas de inmunoadhesina se purificaron del medio condicionado de la manera siguiente. El medio condicionado se bombeó en una columna de Proteína A de 5 ml (Pharmacia) que se había equilibrado con

ES 2 279 473 T3

tampón fosfato Na 20 mM, pH 6,8. Después de la carga, la columna se lavó extensamente con tampón de equilibrio antes de la elución con ácido cítrico, pH 3,5. La proteína eluida se neutralizó inmediatamente mediante cosecha de fracciones de 1 ml en tubos que contienen 275 ml de tampón Tris 1 M, pH 9. La proteína altamente purificada se desaló a continuación en tampón de almacenamiento tal como se describió más arriba para las proteínas etiquetadas con poli-His. La homogeneidad de las proteínas se verificó mediante electroforesis en gel de poli(acrilamida) (PAGE)-SDS y secuenciación aminoacídica N-terminal mediante degradación de Edman.

Alternativamente, se puede utilizar un procedimiento de baculovirus modificado que incorpora células High 5. En este procedimiento, el ADN que codifica la secuencia deseada se amplifica con sistemas adecuados, como la Pfu (Stratagene), o se fusiona cadena arriba (5'-de) de una etiqueta epitópica con un vector de expresión de baculovirus. Dichas etiquetas epitópicas incluyen etiquetas poli-His y etiquetas de inmunoglobulinas (como las regiones Fc o IgG). Es posible utilizar diversos plásmidos, incluyendo plásmidos derivados de plásmidos comerciales como el pIE1-1 (Novagen). Los vectores pIE1-1 y pIE1-2 se diseñaron para expresión constitutiva de proteínas recombinantes del promotor ie1 de baculovirus en células de insecto transformadas de forma estable. Los plásmidos difieren únicamente en la orientación de las dianas de clonaje múltiples y contiene todas las secuencias promotoras conocidas por ser importantes para la expresión de genes mediados por ie-1 en células de insecto no infectadas así como el elemento intensificador hr5. Los vectores pIE1-1 y pIE1-2 incluyen el sitio de inicio de traducción y pueden utilizarse para producir proteínas de fusión. En resumen, la secuencia deseada o la porción deseada de la secuencia (como la secuencia codificante del dominio extracelular de una proteína transmembrana) se amplificó mediante PCR con cebadores complementarios a las regiones 5' y 3'. El cebador 5' puede incorporar dianas de restricción flanqueantes (seleccionadas). El producto se digiere a continuación con las enzimas de restricción seleccionadas y se subclona en el vector de expresión. Por ejemplo, los derivados de pIE1-1 pueden incluir la región Fc de la IgG humana (pb.PH.IgG) o una etiqueta de 8 histidinas cadena abajo (3'-de) la secuencia deseada. Preferentemente, la construcción del vector se secuencia para su confirmación.

Las células High 5 se crecen a una confluencia del 50% en las siguientes condiciones: 27°C, sin CO₂, sin pen/strep. Para cada placa de 150 mm, se mezclan 30 µg del vector basado en pIE que contenía la secuencia con 1 ml de medio Ex-Cell (Medio: Ex-Cell 401 + 1/100 L-Glu JRH Biosciences #14401-78P (nota: este medio es sensible a la luz), y en un tubo separado, se mezclan 100 µl de CellFectin (CellFECTIN (GIBcoBRL#10362-010) (se agita con el vórtex para mezclar) y se añade 1 ml de medio Ex-Cell y se mezcla. Las dos soluciones se combinan y se permite la incubación a temperatura ambiente durante 15 minutos. Se añaden 8 ml de medio Ex-Cell a los 2 ml de ADN/CellFECTIN y todo ello se dispone sobre las células high 5 que habían sido lavadas una vez con medio Ex-Cell. La placa se incuba a continuación en la oscuridad durante 1 hora a temperatura ambiente. La mezcla de ADN/CellFECTIN se aspira y las células se lavan de nuevo con Ex-Cell para eliminar el exceso de CellFECTIN, se añaden 30 ml de medio Ex-CELL fresco y las células se incuban durante 3 días a 28°C. El sobrenadante se cosecha y se determina la expresión de la secuencia en el vector de expresión de baculovirus mediante unión del lote de 1 ml de sobrenadante a 25 ml de bolas de NTA-Ni²⁺ (Qiagen) para las proteínas etiquetadas con histidina o bolas de Protein-A Sepharose CL-4 (Pharmacia) para proteínas etiquetadas con IgG seguido por análisis de PAGE-SDS comparando con una concentración conocida de proteínas estándares por tinción de azul de Coomassie.

El medio condicionado de las células transfectadas (0,5 a 3 l) se cosechó mediante centrifugación para eliminar las células y se filtró a través de un filtro de 0,22 µm. Para las construcciones etiquetadas con poli-His, la proteína que comprende la secuencia se purificó con una columna NTA-Ni²⁺ (Qiagen). Antes de la purificación, se añadió imidazol al medio condicionado a una concentración de 5 mM. El medio condicionado se bombeó en una columna de NTA-Ni²⁺ de 6 ml equilibrada con HEPES 20 mM, pH 7,4, y tampón que contenía NaCl 0,3 M e imidazol 5 mM a una velocidad de flujo de 4-5 min a 48°C. Después de la carga, la columna se lavó con tampón de equilibrio y la proteína se eluyó con tampón de equilibrio que contenía imidazol. La proteína altamente purificada se desaló a continuación en un tampón de almacenamiento que contenía HEPES 10 mM, NaCl 0,14 M y manitol al 14%, pH 6,8 con una columna de 25 ml G25 Superfine (Pharmacia) y se guardó a -80°C.

Las construcciones de inmunoadhesina (que contienen Fc) de las proteínas se purificaron del medio condicionado de la manera siguiente. El medio condicionado se bombeó en una columna de 5 ml de Protein A (Pharmacia) que había sido equilibrada con tampón fosfato N 20 mM, pH 6,8. Después de la carga, la columna se lavó extensamente con tampón de equilibrio antes de la elución con ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutralizó inmediatamente recogiendo fracciones de 1 ml en tubos que contenían ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutralizó inmediatamente recogiendo fracciones de 1 ml en tubos que contenían 275 ml de tampón Tris 1 M, pH 9. La proteína altamente purificada se desaló en el tampón de almacenamiento tal como se describió anteriormente para las proteínas etiquetadas con poli-His. La homogeneidad de la secuencia se confirmó mediante geles de poli(acrilamida)-SDS y secuenciación aminoacídica N-terminal por degradación de Edman y otros procedimientos analíticos según sea necesario o a requerimiento.

Ejemplo 11

Preparación de anticuerpos que se unen a PRO5775

Este ejemplo ilustra la preparación de anticuerpos monoclonales que pueden unirse específicamente a PRO5775.

Las técnicas para la producción de los anticuerpos monoclonales son conocidas en la materia y se describen, por ejemplo, en Goding, *supra*. Los inmunógenos que pueden utilizarse incluyen las proteínas de fusión PRO5775

ES 2 279 473 T3

purificadas que contienen PRO5775 y las células que expresan el PRO5775 recombinante en la superficie celular. La selección del inmunogen puede realizarse por el experto en la materia sin demasiada experimentación.

5 Se inmunizaron ratones, como los Balb/c, con el inmunógeno PRO5775 emulsionado en adyuvante completo de Freund y se inyectaron subcutáneamente o intraperitonealmente en una cantidad de 1-100 microgramos. Alternativamente, el inmunógeno se emulsionó en adyuvante MPL-TDM (Ribi Immunochemical Research, Hamilton, MT) y se inyectaron en los cojinetes de las patas traseras. Los ratones inmunizados se volvieron a inyectar 10 a 12 días después con inmunógeno adicional emulsionado en el adyuvante seleccionado. Los ratones también pueden volverse a inyectar con inyecciones de inmunización adicionales durante varias semanas. De este modo, es posible obtener muestras de suero periódicamente del ratón mediante sangrado retro-orbital para los ensayos ELISA con el fin de detectar anti-PRO5775.

15 Después de detectar un título de anticuerpos adecuado, los animales "positivos" para los anticuerpos pueden volverse a inyectar con una inyección intravenosa final de PRO5775. De tres a cuatro días más tarde, los ratones se sacrifican y se obtienen las células del bazo. Éstas se fusionan (por medio de polietilén glicol al 35%) con una línea celular de mieloma murino como P3X63AgU.1, disponible de ATCC No.CRL 1597. Las fusiones generan células de hibridoma que pueden sembrarse en placas de 96 pocillos con medio HAT (hipoxantina, aminopterina y timidina) para inhibir la proliferación de células no-fusionadas, híbridos de mieloma e híbridos de células del bazo.

20 Las células de hibridoma se analizarán en un ensayo ELISA para su reactividad contra PRO5775. La determinación de células de hibridoma "positivas" que secreten los anticuerpos monoclonales contra PRO5775 deseados se halla dentro del ámbito de la materia.

25 Las células de hibridoma positivas pueden inyectarse intraperitonealmente en ratones Balb/c para producir ascitas que contengan anticuerpos monoclonales anti-PRO5775. La purificación de anticuerpos monoclonales en las ascitas puede lograrse mediante la utilización la precipitación con persulfato amónico, seguido de cromatografía de exclusión en gel. Alternativamente, puede utilizarse la cromatografía de afinidad basada en la unión del anticuerpo a la proteína A o proteína G.

30 *Depósito del material*

Los materiales siguientes se han depositado en el American Type Culture Collection, 10801 University Blvd. Manassas, VA 20110-2209, USA (ATCC):

35	Material	ATCC Depósito No:	Fecha del depósito
	ADN96869-2673	PTA-255	22 de Junio de 1999

40 Estos depósitos se realizaron según las consideraciones del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos con el Propósito de Procedimiento de Patentes y sus Regulaciones (Tratado de Budapest). Ello asegura el mantenimiento de un cultivo viable del depósito de 30 años desde la fecha del depósito. El depósito estará disponible por el ATCC bajo los términos del Tratado de Budapest, y sometido a un acuerdo entre Genentech, Inc. y ATCC, que asegure la disponibilidad permanente y sin restricciones de la progenie del cultivo depositado para el público según la expedición de la patente americana pertinente o tras haber permanecido abierta al público de cualquier solicitud americana o extranjera, quien quiera que sea el primero y asegure una disponibilidad de la progenie al determinado por el Comisionado U.S. de Patentes y Marcas según el acuerdo 35 USC § 122 y las normas del Comisionado según lo acordado (incluyendo la 37§CFR1.14 con la referencia especial a 886 OG 638).

50 El cesionario de la presente solicitud está de acuerdo en que si un cultivo de los materiales en depósito pereciera o se perdiera o se destruyera cuando se cultiva en las condiciones adecuadas, los materiales se reemplazarán prontamente tras notificación con otro cultivo del mismo. La disponibilidad del material depositado no debe considerarse como una licencia para la práctica de la invención en contravención con los derechos garantizados bajo la autoría de cualquier gobierno de acuerdo con sus leyes de patentes.

55 La especificación escrita precedente se considera suficiente para permitir a un experto en la materia la práctica de la invención. La presente invención no está limitada en su ámbito por la construcción depositada, ya que la realización depositada pretende ser una ilustración sencilla de ciertos aspectos de la invención y cualquier construcción que sea equivalente funcionalmente se hallará dentro del ámbito de esta invención. El depósito del material no constituye una admisión de que la descripción escrita contenida aquí sea inadecuada para permitir la práctica de cualquier aspecto de la invención, incluyendo el mejor modo, ni debe considerarse como limitante del ámbito de las reivindicaciones las ilustraciones específicas que se presentan.

65

REIVINDICACIONES

1. Procedimiento de diagnóstico en un mamífero, que comprende la detección del nivel de expresión de un gen que codifica un polipéptido PRO5775 (a) en una muestra a analizar de células del tejido obtenido del mamífero, y (b) en una muestra control de células del tejido normal conocido del mismo tipo celular, en donde el nivel de expresión superior en la muestra a analizar comparado con la muestra control, es indicativo de la presencia del tumor en el mamífero del cual se obtienen las células del tejido a analizar, y en donde dicho polipéptido PRO5775 es un polipéptido que tiene al menos una identidad de secuencia aminoacídica del 80% con:

- (i) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2),
- (ii) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal,
- (iii) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), con el péptido señal, o
- (iv) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal, o
- (v) la secuencia aminoacídica codificada por la secuencia codificante de PRO5775 de longitud completa del ADN96869-2673, depositada en el ATCC con el número de acceso PTA-255,

en donde, dicha identidad de secuencia aminoacídica se determina con la utilización del programa informático ALIGN-2.

2. Procedimiento según la reivindicación 1, en donde el nivel de identidad es de al menos el 90%.

3. Procedimiento según la reivindicación 2, en donde el nivel de identidad es de al menos el 95%.

4. Procedimiento según la reivindicación 3, en donde dicho polipéptido PRO5775 es un polipéptido que tiene:

- (i) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2),
- (ii) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal,
- (iii) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), con el péptido señal,
- (iv) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal, o
- (v) la secuencia aminoacídica codificada por la secuencia codificante de PRO5775 de longitud completa del ADN96869-2673, depositada en el ATCC con el número de acceso PTA-255.

5. Procedimiento según cualquiera de las reivindicaciones anteriores, en donde el péptido señal va desde el aminoácido 1 al 29±5 de la Figura 2 (SEC ID NO:2).

6. Procedimiento según cualquiera de las reivindicaciones anteriores, en donde el dominio extracelular con el péptido señal va desde el aminoácido 1 al 380±5 de la Figura 2 (SEC ID NO:2), o el dominio extracelular sin el péptido señal va desde el aminoácido 30±5 de la Figura 2 (SEC ID NO:2).

7. Procedimiento de diagnóstico de un tumor en un mamífero, que comprende (a) el contacto de un anticuerpo con una muestra a analizar de células de tejido obtenidas del mamífero, y (b) la detección de la formación de un complejo entre dicho anticuerpo y el polipéptido PRO5775 tal como se definió en cualquiera de las reivindicaciones 1 a 6 en la muestra a analizar, en donde la formación de un complejo es indicativa de la presencia de un tumor en dicho mamífero, y en donde el anticuerpo es capaz de unirse específicamente a un polipéptido que tiene:

- (i) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2),
- (ii) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal,
- (iii) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), con el péptido señal,
- (iv) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal, o

ES 2 279 473 T3

(v) la secuencia aminoacídica codificada por la secuencia codificante de PRO5775 de longitud completa del ADN96869-2673, depositada en el ATCC con el número de acceso PTA-255.

5 8. Procedimiento según la reivindicación 7, en donde el péptido señal va desde el aminoácido 1 al 29±5 de la Figura 2 (SEC ID NO:2).

9. Procedimiento según la reivindicación 7 o la reivindicación 8, en donde el dominio extracelular con el péptido señal va desde el aminoácido 1 al 380±5 de la Figura 2 (SEC ID NO:2), o el dominio extracelular sin el péptido señal va desde el aminoácido 30±5 al 380±5 de la Figura 2 (SEC ID NO:2).

10. Procedimiento según la reivindicación 7, en donde el anticuerpo mencionado está marcado de modo detectable.

11. Procedimiento según la reivindicación 7, en donde la muestra de células del tejido a analizar se obtiene de un individuo sospechoso de tener un crecimiento o proliferación de células neoplásicas.

12. Equipo de diagnóstico del cáncer que comprende un anticuerpo y un vehículo en un paquete adecuado, en donde el anticuerpo es capaz de unirse específicamente a un polipéptido que tiene:

20 (i) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2),

(ii) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal,

25 (iii) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), con el péptido señal,

(iv) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal, o

30 (v) la secuencia aminoacídica codificada por la secuencia codificante de PRO5775 de longitud completa del ADN96869-2673, depositada en el ATCC con el número de acceso PTA-255.

13. Equipo según la reivindicación 12, en donde el péptido señal va desde el aminoácido 1 al 29±5 de la Figura 2 (SEC ID NO:2).

35 14. Equipo según la reivindicación 12 o la reivindicación 13, en donde en donde el dominio extracelular con el péptido señal va desde el aminoácido 1 al 380±5 de la Figura 2 (SEC ID NO:2), o el dominio extracelular sin el péptido señal va desde el aminoácido 30±5 al 380±5 de la Figura 2 (SEC ID NO:2).

40 15. Equipo según cualquiera de las reivindicaciones 12 a 14, que además comprende las instrucciones para la utilización de dicho anticuerpo para detectar la presencia de un polipéptido, tal como se definió en cualquiera de las reivindicaciones 1 a 6, en una muestra sospechosa de contener el mismo.

16. Anticuerpo para utilizar en el diagnóstico *in vivo*, en donde el anticuerpo es capaz de unirse específicamente a un polipéptido que tiene:

45 (i) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2),

(ii) la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal,

50 (iii) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), con el péptido señal,

(iv) el dominio extracelular de la secuencia aminoacídica que se muestra en la Figura 2 (SEC ID NO:2), sin el péptido señal, o

55 (v) la secuencia aminoacídica codificada por la secuencia codificante de PRO5775 de longitud completa del ADN96869-2673, depositada en el ATCC con el número de acceso PTA-255.

60 17. Anticuerpo según la reivindicación 16, en donde el péptido señal va desde el aminoácido 1 al 29±5 de la Figura 2 (SEC ID NO:2).

18. Anticuerpo según la reivindicación 16 o la reivindicación 17, en donde en donde el dominio extracelular con el péptido señal va desde el aminoácido 1 al 380±5 de la Figura 2 (SEC ID NO:2), o el dominio extracelular sin el péptido señal va desde el aminoácido 30±5 al 380±5 de la Figura 2 (SEC ID NO:2).

65 19. Anticuerpo según cualquiera de las reivindicaciones 16 a 18, en donde se trata del diagnóstico de un tumor.

20. Anticuerpo según la reivindicación 19, en donde se trata del diagnóstico de un tumor de colon o de pulmón.

ES 2 279 473 T3

21. Anticuerpo según cualquiera de las reivindicaciones 16 a 18, el cual es un anticuerpo monoclonal.

22. Anticuerpo según la reivindicación 21 que comprende una región de determinación de la complementariedad (CDR) o una región de entramado (FR).

5

23. Anticuerpo según la reivindicación 21, en donde el anticuerpo mencionado es un anticuerpo humanizado.

24. Anticuerpo según cualquiera de las reivindicaciones 16 a 18, el cual es un fragmento de anticuerpo o un anticuerpo de una sola cadena.

10

25. Composición que comprende un anticuerpo tal como se definió en cualquiera de las reivindicaciones 16 a 24 en una mezcla junto con un vehículo farmacéuticamente aceptable.

15

20

25

30

35

40

45

50

55

60

65

FIGURA 1

CCAGCTGCAGAGAGGAGGAGGTGAGCTGCAGAGAAAGAGGAGGTGGTGTGGAGCAACAGGCAGCACCGAGCCTGCC
CCGTGAGCTGAGGGCTGACATCTGCGGCTGGAATCAGGATAGACACCAGGCAGGACCCCCAGAGATGCTGAAAG
CCTCTTTGAAAGCAAGCATGCCCCACATGGCCATGCTCCAGCCGCCCCCGCCCCGAGCAGAGAGGCT
GGCAGTTGCAAGTCCCTGGGAGGGCTGGCTGTGCTGTGGCTGGGCTCCGTGGCTCTTATCTGGCTCCTGTGGCAA
GTGCCCGCTCCTCCACTGGGGCCAGGTGCAAGCCCAAGGACGTGCCAGGTCTGTGGAGCATGGCTCCAGCCCCA
GCTTGGAGGCCCTGGAAGCAGAGGCCAGGCCAGCAGAGGACTCCTGCCAGCTTGTCTTGTGGAAAGCATGCC
CAGGACTTGCATCTGCAGCCGGCAGGCCCTCTGCCAGCCTCTGGCCAGGCTGTGGCTGCGCTGCTGGACT
GCCCCAGAGAGGCTCCAGGTGGCTTCACTACTGCTTCCCTCAGAGGCTGACATCGGGTCAAGGACTCGTCT
TCCAGGCTGGAGAGGCTCTTCTGCAGAGCTGCAGCAGCTGCTGGCAAGAACATTTCCCTGGCTGTGGCCACC
AGCAGCCGACATGGCCAGGACATCCACCGACTGCAAGTTCTGGCTGCCAGGTTGCCATGACGACAGGTTG
CCCATGGGGGCTCACCAAGGGTGTFTTGCAGCTCCAAATCTGGGTTGTGGATGGAGCGCACATAACAGGGG
MTGCCAACATGACCTGGCGTCTCTGACGCAAGGTGAAGGAGCTTGGCGCTGTCTATAACTGCAGCCACTG
GCCCCAGACCTGGAGAGAGCCTTCCAGACTACTGGCTACTGGGGTGCACAGGCTGTCTCCCCAAACTGG
CCTCAGAACCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT
TACTTCTCAGGCTGGCCACCAGCCTCTGTGCCAGGGCCGACCCGGGACTTGGAGGGCTGTCTGGCGGAGT
GGAGCGCCAGGAGTTGATCTATGCTCCGTGATGGAGTATTTCCCAACAGCGCTTCAGCCACCCCGGAGG
TACTGGCCGGTGTGGACAAAGCCCTGCGGGCCGAGCTTGGCAAGGGGTTGGCGTGGGCTGCTGGTCCG
TGCCGACTCAACACGGAGCCCCACCATGTTCCCTTCCCTGCGGTCCCTGCAGGGCTCAGCAACCCGGGCCAAC
GTCTCTGTGGAGGTGAAGTCTTCACTGGTGGGTAAGGAAACCAATTCACATCCCATTCAGGAGGTGAACCA
AGCAATTCATGCTCAAGGAAAGGCAAGCTTACATAGGCACTTCCAACTGCTGGAGGATTAATTCAGCAGCAG
CCGGGGTGGGCTTGGTGTCCACCCAGGCTTGGCCGACAGCCGGGGGCAAGGTGCAGAGCAGCTGGGG
CAGCTCTTTGAGCCGGACTGGGTTCCCGCTAAGCCCTTGGCTTGGACGGACAGGCTCCGGCCAGGACTGCGTT
TGGCAGGCTGAGGGGGCTCTTTTCTCTGGGGAGCCCGCCCGGCAAGGCTTCCCTCTGACCCGGGCT
GGCTTCCAGCCGCTTCCCTCCCAAGCAGCTCCGGTCCGCACTGGCCCAAGAGCCGCTGGACCCGGGGCT
CGCAACCCGGCCGCTCTCTGATTTCCGAGTCCAGCCCGCCCTGAGCCCACTCTCTCAGGGAGCCCTCA
GGAAGCCCTTCCCTGACTCCTGGCCACAGCCAGGCTTAAAAAACTGCTGCTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURA 2

NPPRRPWDRRAGTLQVIGALAVLWLGSAVALYICLWQVPRPPTMQVQPKDVPKSNWEGSSPAWEPLKAEARQQRD
 SCQLVLYVESIPQDLPSAAGSPBAQPLGQANLQLLDTAQESVHVASYWNSLTGPDIGVMDSSQIQRALIQKIQQL
 LGRNLSLAVATSSPTLARTSTDLQVLAARGAHVROVPMGRRTAGVLSKFWVVDGRHIYNGSAMMDWRSLTQVKE
 LGAVITNCSNLAQQLKTPQITWYLVNFKAVLFCWNPQNFSSHPNLFQPFSLFDGVVITATYFSASPPALCFQGR
 TROLEALLAVMGSAQEPYASVMYFPFTRFSHFPRYKPVLDNLRRAAFEGQVYRVLLVGCCLNTEPTMFFYLE
 ELQALSNFAAVSVVDVKVFIYVQGNHNIFFSRVHSEKFNVTKAAVIGTSNWSEDTFSSITAGVGLVVTQSPCAQ
 PGRATVQEQRLQKFRDNBSRYAVGLOGQAPGQDCVMQG

Péptido señal:	Aminoácidos 1-29
Dominio transmembrana:	Aminoácidos 381-399
Sitios de N-glicosilación:	Aminoácidos 133-137;154-158;232-236; 264-268;386-390;400-404;410-414;427-431
Sitios de N-miristilación:	Aminoácidos 58-64;94-100;131-137; 194-200;251-257;277-283;281-287;361-367; 399-405;440-446;448-454;478-484

ES 2 279 473 T3

LISTA DE SECUENCIAS

- <110> Genentech Inc.
- 5 <120> COMPOSICIONES Y PROCEDIMIENTOS PARA EL TRATAMIENTO DEL TUMOR
- <130> CMD/FP6297329
- 10 <140> EP 05018356.5
<141> 2000-02-11
- <150> PCT/US99/05028
- 15 <151> 1999-03-11
- <150> US 60/123.972
<151> 1999-03-11
- 20 <150> US 60/133.459
<151> 1999-05-11
- 25 <150> PCT/US99/12252
<151> 1999-06-02
- <150> US 60/140.650
- 30 <151> 1999-06-22
- <150> US 60/140.653
<151> 1999-06-22
- 35 <150> US 60/144.758
<151> 1999-07-20
- 40 <150> US 60/145.698
<151> 1999-07-26
- <150> US 60/146.222
<151> 1999-08-17
- 45 <150> US 60/151.689
<151> 1999-08-31
- 50 <150> PCT/US99/20111
<151> 1999-09-01
- 55 <150> PCT/US99/21090
<151> 1999-09-15
- 60 <150> PCT/US99/28313
<151> 1999-11-30
- <150> PCT/US99/28301
- 65 <151> 1999-12-01

ES 2 279 473 T3

<150> PCT/US99/28634

<151> 1999-12-01

5 <150> PCT/US99/00219

<151> 2000-01-05

<160> 5

10

<210> 1

<211> 1964

<212> ADN

15 <213> *Homo sapiens*

<400> 1

20

```
ccagctgcag agaggaggag gtgagctgca gagaagagga ggttgggtgtg 50
gagcaacaggc agcaaccgagc ctgccccgtg agctgagggc ctgcagctctg 100
cggctggaat caggtatagac accaaggcag gacccccaga gatgctgaag 150
cctcttttga aagcagcagt ggccccaca tggccatgct ceatgcccgc 200
ccgcccggcg tgggacagag aggtctggac gttgcaggtc ctgggagcgc 250
tggtctgtct gtggtctggc tccgtgctc ttatctgctt cctgtggcaa 300
gtgcccctgc ctcccactg gggccaggtg cagccccagg acgtgcccag 350
gtcctgggag catggctcca gccagcttg ggagccccctg gaagcagagg 400
ccaggcagca gagggactcc tgcagcttg tcttgttga aagcatcccc 450
caggacctgc catctgcagc cggcagcccc tctgcccagc ctctgggcca 500
ggcctggttg cagctgcttg acactgccc ggagagcgtc cacgtggctt 550
catactactg gtccctcaca gggcctgaca tgggggtcaa cgactcgtct 600
tcccagctgg gagagctctt tctgcagaag ctgcagcagc tcttgggagc 650
gaacatttcc ctggctgttg ccaccagcag cccgacactg gccaggacat 700
ccaccgacct gcaggttctg gctgcccag gtgcccctgt acgacaggty 750
cccattgggc ggctcaccag ggggtgtttg cactccaaat tctgggttgt 800
ggatggacgg cacatataca tgggcagtgc caacatggac tggcggcttc 850
tgacgcaggt gaaggagctt ggcgctgtca tctataactg cagccacctg 900
gcccagacc tggagaagac ctccagacc tactgggtac tgggggtgcc 950
caaggctgtc ctccccaaa cctggcctca gaacttctca tctcacttca 1000
accgtttcca gcccttccac ggcctctttg atgggggtgc caccactgcc 1050
tacttctcag cgtcggcacc agcactctgt ccccagggcc gcaccggga 1100
cctggaggcg ctgctggcgg tcatggggag cggcccaggag ttcattctatg 1150
cctccgtgat ggggtatttc ccaccacgc gcttcagcca cccccgaggy 1200
tactggccgg tgctggaca cgcgctcggg gcggcagcct tcggcaaggg 1250
```

65

ES 2 279 473 T3

5 cgctgcgctg cgcctgctgg teggctgcgg actcaacacg gaccccacca 1300
 tgttccccta cctggggtcc ctgcaggcgc tcagcaacc cggggccaac 1350
 gtctctgtgg acgtgaaagt cttcatcgtg ccgggtggga accattccaa 1400
 catcccattc agcagggtga accacagcaa gtctcatggtc acggagaagg 1450
 cagcctacat aggcacctcc aactggctgg aggattactt cagcagcacg 1500
 10 ggggggtgg gcttgggtgt caccagagc cctggcgcgc agcccgcggg 1550
 ggccacggtg caggagcagc tgcggcagct ctttgagcgg gactggagtt 1600
 cgcgtaccgc cgtcggcctg gacggacagg ctccgggcca ggactgcgtt 1650
 15 tggcagggct gggggggccc tctttttctc tegggacce cgcgccgac 1700
 ggcctctccc ctctgacccc ggcctgggct tcagccgctt cctcccgca 1750
 gacgcccggg tccgactgc gccaggagcc gcctgcgacc gcccgggcgt 1800
 20 cgcacaccgc ccgctgctc tctgatttcc gactccagcc cccctgagc 1850
 cccacctcct ccaggagcc ctccaggaag ccccttcct gactcctggc 1900
 ccacaggcca ggcctaaaaa aaactcgtgg cttcaaaaaa aaaaaaaaaa 1950
 25 aaaaaaaaaa aaaa 1964

<210> 2
 30 <211> 489
 <212> PRT
 <213> *Homo sapiens*
 35 <400> 2

Met Pro Pro Arg Arg Pro Trp Asp Arg Glu Ala Gly Thr Leu Gln
 1 5 10 15
 40 Val Leu Gly Ala Leu Ala Val Leu Trp Leu Gly Ser Val Ala Leu
 20 25 30
 Ile Cys Leu Leu Trp Gln Val Pro Arg Pro Pro Thr Trp Gly Gln
 35 40 45
 Val Gln Pro Lys Asp Val Pro Arg Ser Trp Glu His Gly Ser Ser
 50 55 60
 Pro Ala Trp Glu Pro Leu Glu Ala Glu Ala Arg Gln Gln Arg Asp
 65 70 75
 50 Ser Cys Gln Leu Val Leu Val Glu Ser Ile Pro Gln Asp Leu Pro
 80 85 90
 Ser Ala Ala Gly Ser Pro Ser Ala Gln Pro Leu Gly Gln Ala Trp
 95 100 105
 55 Leu Gln Leu Leu Asp Thr Ala Gln Glu Ser Val His Val Ala Ser
 110 115 120

60

65

ES 2 279 473 T3

Tyr Tyr Trp Ser Leu Thr Gly Pro Asp Ile Gly Val Asn Asp Ser
 125 130 135
 Ser Ser Gln Leu Gly Glu Ala Leu Leu Gln Lys Leu Gln Gln Leu
 140 145 150
 5 Leu Gly Arg Asn Ile Ser Leu Ala Val Ala Thr Ser Ser Pro Thr
 155 160 165
 Leu Ala Arg Thr Ser Thr Asp Leu Gln Val Leu Ala Ala Arg Gly
 170 175 180
 10 Ala His Val Arg Gln Val Pro Met Gly Arg Leu Thr Arg Gly Val
 185 190 195
 Leu His Ser Lys Phe Trp Val Val Asp Gly Arg His Ile Tyr Met
 200 205 210
 15 Gly Ser Ala Asn Met Asp Trp Arg Ser Leu Thr Gln Val Lys Glu
 215 220 225
 Leu Gly Ala Val Ile Tyr Asn Cys Ser His Leu Ala Gln Asp Leu
 230 235 240
 20 Glu Lys Thr Phe Gln Thr Tyr Trp Val Leu Gly Val Pro Lys Ala
 245 250 255
 Val Leu Pro Lys Thr Trp Pro Gln Asn Phe Ser Ser His Phe Asn
 260 265 270
 25 Arg Phe Gln Pro Phe His Gly Leu Phe Asp Gly Val Pro Thr Thr
 275 280 285
 Ala Tyr Phe Ser Ala Ser Pro Pro Ala Leu Cys Pro Gln Gly Arg
 290 295 300
 Thr Arg Asp Leu Glu Ala Leu Leu Ala Val Met Gly Ser Ala Gln
 305 310 315
 30 Glu Phe Ile Tyr Ala Ser Val Met Glu Tyr Phe Pro Thr Thr Arg
 320 325 330
 Phe Ser His Pro Pro Arg Tyr Trp Pro Val Leu Asp Asn Ala Leu
 335 340 345
 35 Arg Ala Ala Ala Phe Gly Lys Gly Val Arg Val Arg Leu Leu Val
 350 355 360
 Gly Cys Gly Leu Asn Thr Asp Pro Thr Met Phe Pro Tyr Leu Arg
 365 370 375
 40 Ser Leu Gln Ala Leu Ser Asn Pro Ala Ala Asn Val Ser Val Asp
 380 385 390
 Val Lys Val Phe Ile Val Pro Val Gly Asn His Ser Asn Ile Pro
 395 400 405
 45 Phe Ser Arg Val Asn His Ser Lys Phe Met Val Thr Glu Lys Ala
 410 415 420
 Ala Tyr Ile Gly Thr Ser Asn Trp Ser Glu Asp Tyr Phe Ser Ser
 425 430 435
 Thr Ala Gly Val Gly Leu Val Val Thr Gln Ser Pro Gly Ala Gln
 440 445 450
 50 Pro Ala Gly Ala Thr Val Gln Glu Gln Leu Arg Gln Leu Phe Glu
 455 460 465
 Arg Asp Trp Ser Ser Arg Tyr Ala Val Gly Leu Asp Gly Gln Ala
 470 475 480
 55 Pro Gly Gln Asp Cys Val Trp Gln Gly
 485

<210> 3

<211> 19

60 <212> ADN

<213> Secuencia artificial

<400> 3

65

ggggaacct tccaacatc

ES 2 279 473 T3

	<210> 4	
	<211> 23	
	<212> ADN	
5	<213> Secuencia artificial	
	<400> 4	
10	ccattcagca ggggaacca cag	23
	<210> 5	
	<211> 20	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Sonda oligonucleotídica sintética	
	<400> 5	
25	tctccgtgac catgaactg	20
30		
35		
40		
45		
50		
55		
60		
65		