



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 361 044**

51 Int. Cl.:
C12N 15/12 (2006.01)
C12N 9/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **98903409 .5**
96 Fecha de presentación : **08.01.1998**
97 Número de publicación de la solicitud: **0972024**
97 Fecha de publicación de la solicitud: **19.01.2000**

54 Título: **Un supresor de tumor designado TS10q23.3.**

30 Prioridad: **30.01.1997 US 791115**

45 Fecha de publicación de la mención BOPI:
13.06.2011

45 Fecha de la publicación del folleto de la patente:
13.06.2011

73 Titular/es: **The Board of Regents - The University
of Texas System
201 West 7th Street
Austin, Texas 78701, US**

72 Inventor/es: **Steck, Peter;
Pershouse, Mark, A.;
Jasser, Samar, A.;
Yung, W., K., Alfred y
Tavtigian, Sean, V.**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 361 044 T3

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

DESCRIPCIÓN

Un supresor de tumor designado TS10q23.3.

5 **Antecedentes de la invención****I. Campo de la invención**

La presente invención se refiere a los campos de la oncología, genética y biología molecular. Más concreto la invención se refiere a la identificación, en el cromosoma 10 humano, de un gen supresor de tumor. Los defectos en este gen están asociados con el desarrollo de cánceres, tales como gliomas.

II. Técnica relacionada

La oncogénesis fue descrita por Foulds (1958) como un proceso biológico multietapas, el cual actualmente se sabe que ocurre mediante la acumulación de daño genético. A un nivel molecular, el proceso multietapas de la tumorigénesis implica la interrupción de los efectores reguladores tanto positivos como negativos (Weinberg, 1989). Se ha presupuesto que la base molecular para los carcinomas de colon humano, por Vogelstein y colaboradores (1990), implica un número de oncogenes, genes supresores de tumor y genes de reparación. Igualmente, los defectos que llevan al desarrollo del retinoblastoma se han relacionado con otro gen supresor de tumor (Lee *et al.*, 1987). Aún se han identificado otros oncogenes y supresores de tumor en una diversidad de otras enfermedades. Desafortunadamente, queda un número insuficiente de cánceres tratables, y los efectos del cáncer son catastróficos (más de medio millón de muertes por año solamente en los Estados Unidos).

Un ejemplo de la devastadora naturaleza del cáncer implica que tumores que surgen de células del linaje astrocítico son los tumores primarios más comunes del sistema nervioso central (Russell & Rubinstein, 1989). La mayoría de estos tumores ocurren en la población adulta. Los tumores cerebrales primarios también explican el cáncer sólido más común en la población paciente pediátrica y la segunda causa importante de muertes por cáncer en niños menores de 15 años de edad. En 1994 se diagnosticaron unos estimados 18.500 nuevos casos de tumores cerebrales primarios (Boring *et al.*, 1994). Los estudios epidemiológicos muestran que la incidencia de tumores cerebrales está aumentando y representa la tercera causa más común de muerte por cáncer entre pacientes de 18 a 35 años de edad. Debido a su localización dentro del cerebro y a la infiltración típica de las células tumorales dentro de los tejidos de alrededor, con frecuencia la intervención terapéutica con éxito para tumores cerebrales primarios está limitada. Desafortunadamente, aproximadamente dos terceras partes de estos individuos afectados sucumbirán a la enfermedad en dos años. Los tumores intracraneales más comunes en adultos surgen de células del linaje glial y ocurre aproximadamente en una frecuencia de 48% de glioblastoma multiforme (GBM), 21% de astrocitomas (A) (anaplástica (AA) y de bajo grado) y 9% de ependimomas y oligodendrogliomas (Levin *et al.*, 1993).

Los estudios genéticos han implicado diversos genes, y sus correspondientes productos proteicos, en la oncogénesis de los tumores cerebrales primarios. Entre las diversas alteraciones informadas están: amplificación del receptor del factor de crecimiento epidérmico y uno de sus ligandos, transformación del factor alfa de crecimiento, *N-myc*; gli, maduración por corte y empalme y expresión de los receptores del factor de crecimiento de fibroblasto alterado, y pérdida de función de p53, p16, *Rb*, genes 1 y 2 de neurofibromatosis, *DCC* y genes supresores de tumor putativos en los cromosomas 4, 10, 17 (no p53), 19, 22 y X (Wong *et al.*, 1987; El-Azouzi *et al.*, 1989; Nishi *et al.*, 1991; James *et al.*, 1988; Kamb *et al.*, 1994; Henson *et al.*, 1994; Yamaguchi *et al.*, 1994; Bianchi *et al.*, 1994; Ransom *et al.*, 1992; Rasheed *et al.*, 1992; Scheck y Coons, 1993; Von Demling *et al.*, 1994; Rubio *et al.*, 1994; Ritland *et al.*, 1995).

Las alteraciones más frecuentes incluyen la amplificación del receptor de EGF ($\approx 40\%$), pérdida de función de p53 ($\approx 50\%$), p16 ($\approx 50\%$), *Rb* ($\approx 30\%$) y deleciones en el cromosoma 10 ($>90\%$). Además, el grado y nivel de malignidad histológica de los tumores astrocíticos guarda correlación con la acumulación aumentada del daño genético similar al carcinoma de colon. Además, algunos cambios parecen ser relativamente específicos a linaje o grado. Por ejemplo, pérdidas en el cromosoma 19q ocurren predominantemente en oligodendrogliomas, mientras las deleciones en el cromosoma 10 y la amplificación y mutación del receptor de EGF ocurren principalmente en GBMs. La deleción de una copia entera o segmentos del cromosoma 10 está fuertemente indicada como el suceso genético más común asociado con la forma más común de tumores cerebrales primarios, GBMs.

La citogenética y estudios de deleción alélica en GBMs posteriores claramente han demostrado alteraciones genéticas moleculares frecuentes y extensas asociadas con el cromosoma 10 (Binger *et al.*, 1988; Ransom *et al.*, 1992; Rasheed *et al.*, 1992; James *et al.*, 1988; Fujimoto *et al.*, 1989; Fults *et al.*, 1990, 1993; Karlbom *et al.*, 1993; Rasheed *et al.*, 1995; Sonoda *et al.*, 1996; Albarosa *et al.*, 1996). Los análisis citogenéticos han mostrado claramente la alteración del cromosoma 10 como una ocurrencia común en GBMs, con el 60% de tumores que presentan alteración. Los estudios de deleción alélica de GBMs también han revelado desequilibrios alélicos muy frecuentes asociados con el cromosoma 10 (90%). Sin embargo, las pérdidas son tan extensas y frecuentes que mediante estos análisis se podría definir inequívocamente la no sublocalización cromosómica de una pérdida constante.

Varios estudios recientes han implicado la región 10q25-26, específicamente una región de 17 cM desde D10S190 a D10S216. Una región telomérica desde D10S587 a D10S216 estaba implicada mediante análisis de deleción alélica usando una serie de gliomas de grado bajo y alto que presentan solamente una pérdida parcial de cromosoma 10

(Rasheed *et al.*, 1995). Esta región (≈ 1 cM) estaba perdida o era no informativa en 11 GBMs, 4 AAs, 1 A y 1 oligodendroglioma, sugiriendo la localización de una región candidata. Este estudio también ilustró que las deleciones en el cromosoma 10 ocurren en gliomas grado inferior. Albarosa *et al.* (1996) sugieren una región candidata centromérica basada en una pequeña deleción alélica en un tumor cerebral pediátrico de los marcadores D10S221 a D10S209. Saya *et al.*, usando una serie de GBMs, han sugerido dos regiones comunes de deleciones, 10q26 y 10q24 (D10S192).

El brazo corto del cromosoma 10 también ha estado implicado en contener otro gen supresor de tumor. Los estudios primero proporcionaron evidencia funcional de un gen supresor de tumor en 10p en glioma (Steck *et al.*, 1995) lo cual más tarde se mostró para próstata (Sánchez *et al.*, 1995; Murakami *et al.*, 1996). El último estudio ha implicado una región de 11 cM entre D10S1172 y D10S527. Los estudios de deleción alélica de gliomas han mostrado deleción extensa en 10p, pero de nuevo, no se ha alcanzado una localización firme (Karlsson *et al.*, 1993; Kimmelman *et al.*, 1996; estas regiones del cromosoma 10 se muestran en la Fig. 1, más adelante). Además, también se ha mostrado que la amplificación del receptor de EGF ocurre casi exclusivamente en tumores que tenían deleciones en el cromosoma 10, sugiriendo una posible relación entre estas alteraciones genéticas (Von Deimling *et al.*, 1992).

También se han informado de deleciones en el brazo largo, de particularmente 10q24, para carcinomas de próstata, renales, uterinos, de pulmón de célula pequeña, endometriales, meningioma y leucemias graves de célula T (Carter *et al.*, 1990; Morita *et al.*, 1991; Herbst *et al.*, 1984; Jones *et al.*, 1994; Rempel *et al.*, 1993; Peiffen *et al.*, 1995; Petersen *et al.*, 1997). Recientemente, estudios detallados sobre el carcinoma de próstata han demostrado que (1) el brazo corto y largo del cromosoma 10 parece totalmente contener genes supresores de tumor y (2) la localización del gen supresor del brazo largo se mapea en el límite 10q23-24 (Gray *et al.*, 1995; Ittmann, 1996; Trybus *et al.*, 1996). La región de la deleción común identificada por estos tres grupos se centra alrededor de D10S215 y se extiende aproximadamente 10 cM (Fig. 1). La región se solapa con nuestra región candidata, sin embargo, se informó de no localización adicional dentro de la región para el carcinoma de próstata. Las pérdidas alélicas asociadas con el carcinoma de próstata también parecen ocurrir en solamente el aproximadamente 30-40% de los tumores examinados. Además, las deleciones se observaron en tumores de fases más avanzadas, similares a GBMs, y se pueden relacionar con la capacidad de producir metástasis (Nihei *et al.*, 1995; Komiya *et al.*, 1996). La combinación de estos resultados sugiere que los cánceres humanos múltiples implican la región 10q23-24.

Las diferencias en las localizaciones de las regiones candidatas sugieren diversas posibilidades. Primero, es posible la presencia de dos o más genes supresores de tumor en 10q. Segundo, no todas las deleciones pueden efectuarse en el locus del gen supresor de tumor. Estas alternativas no son mutuamente exclusivas. A favor de la última posibilidad, se sugirió que un centrómero latente potencial ocurre en 10q25 el cual puede dar lugar a las alteraciones genéticas, particularmente rotura (Vouillaire *et al.*, 1993).

A pesar de toda esta información, se mantuvo difícil de localizar la identidad del gen (o genes) implicado con la supresión del tumor relacionado con 10q23-24. Sin identificación de un gen específico y deducción de la proteína para la cual codifica, es imposible comenzar el desarrollo de una terapia eficaz de direccionamiento ("targeting") de este producto. Por tanto, es un objetivo importante aislar el(los) supresor(es) de tumor localizado(s) en esta región y determinar su estructura y función.

Compendio de la invención

Por lo tanto, es un objetivo de la presente invención proporcionar un supresor de tumor, designado TS10q23.3. También es un objetivo proporcionar ADNs que representen todo o parte de un gen que codifica TS10q23.3. También es un objetivo proporcionar métodos para usar estas composiciones.

De acuerdo con los anteriores objetivos, se ha proporcionado, en una realización, un supresor de tumor designado TS10q23.3. El polipéptido tiene, en un ejemplo, la secuencia de aminoácidos como la expuesta en la Fig. 7 o Fig. 9 o en la SEC. ID. N°: 1. También se ha proporcionado un péptido aislado que tiene 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400 o más residuos consecutivos de un supresor de tumor designado TS10q23.3. El péptido también puede estar conjugado con una molécula vehículo, por ejemplo, KLH o BSA.

En otra realización, se ha proporcionado un anticuerpo monoclonal que se une inmunológicamente a un supresor de tumor designado TS10q23.3. El anticuerpo puede ser reactivo no cruzado con otros polipéptidos humanos, o puede unirse a TS10q23.3 no humano, pero no a TS10q23.3 humano. Además el anticuerpo puede comprender un marcaje detectable, tal como un marcaje fluorescente, un marcaje quimioluminiscente, un radiomarcaje o una enzima. También abarca las células de hibridoma y las líneas celulares que producen tales anticuerpos.

En otra realización, se incluyen antisueros policlonales, los anticuerpos de los cuales se unen inmunológicamente a un supresor de tumor designado TS10q23.3. Los antisueros pueden derivarse de cualquier animal, pero preferentemente es de además de humano, de ratón o perro.

En aún otra realización, se ha proporcionado un ácido nucleico aislado que comprende una región, o el complemento de la misma, que codifica un supresor de tumor designado TS10q23.3 o una variante alélica, de la misma. La región que codifica el supresor de tumor puede derivarse de cualquier mamífero pero, en realizaciones particulares, está seleccionada entre secuencias de murine, caninas y humanas. Las mutaciones del ácido nucleico pueden ser mu-

ES 2 361 044 T3

tantes de delección, mutantes de inserción, mutantes de desplazamiento de la pauta de lectura, mutantes sin sentido, mutantes de sentido erróneo (“missense”) o mutantes de maduración por corte y empalme. En una realización particular, el supresor de tumor tiene la secuencia de aminoácidos de la Fig. 9. El ácido nucleico puede ser ADN genómico, ADN complementario o ARN.

5 En realizaciones adicionales, el ácido nucleico comprende un ADN complementario y comprende además un promotor enlazado de manera operable a la región o al complemento del mismo, que codifica el supresor de tumor. Elementos adicionales incluyen una señal de poliadenilación y un origen de replicación.

10 También se pueden emplear vectores vitales tales como retrovirus, adenovirus, herpesvirus, virus vaccinia y virus adenoasociados. El vector puede ser “desnudo” o estar empaquetado en una partícula de virus. Alternativamente, el ácido nucleico puede comprender un vector de expresión empaquetado en un liposoma.

15 Se describen diversos tamaños de ácidos nucleicos, pero no son limitantes: aproximadamente 1.212 bases, aproximadamente 1.500 bases, aproximadamente 2.000 bases, aproximadamente 3.500 bases, aproximadamente 5.000 bases, aproximadamente 10.000 bases, aproximadamente 15.000 bases, aproximadamente 20.000 bases, aproximadamente 25.000 bases, aproximadamente 30.000 bases, aproximadamente 35.000 bases, aproximadamente 40.000 bases, aproximadamente 45.000 bases, aproximadamente 50.000 bases, aproximadamente 75.000 bases y aproximadamente 100.000 bases.

20 En aún otra realización más, se ha proporcionado un oligonucleótido aislado de entre 10 y 50 bases consecutivas de un ácido nucleico, o complementario al mismo, que codifica un supresor de tumor designado TS10q23.3. El oligonucleótido puede ser de 15 bases de longitud, 17 bases de longitud, 20 bases de longitud, 25 bases de longitud o 50 bases de longitud.

25 En aún otra realización más, se ha proporcionado un método de diagnóstico de un cáncer o ascenso del desarrollo de cáncer o metástasis en un sujeto humano que comprende la determinación en una muestra de dicho sujeto si hay una mutación, delección o inserción puntual en una secuencia de nucleótidos de TS10q23.3 de dicho sujeto en comparación con la región codificadora de la secuencia de nucleótidos mostrada en la SEC. ID. N°: 2, siendo dicha mutación, delección o inserción puntual indicativa de cáncer o ascenso del desarrollo de cáncer o metástasis. El cáncer puede ser de cerebro, pulmón, hígado, bazo, riñón, ganglio linfático, intestino delgado, páncreas, células sanguíneas, colon, estómago, pecho, endometrio, próstata, testículo, ovario, piel, cabeza y cuello, esófago, médula ósea, glioblastoma, meduloblastoma, astrocitoma, oligodendroglioma, ependimoma. En realizaciones preferidas, el cáncer es un cáncer de próstata o cáncer de pecho. En otra realización preferida, el cáncer es cáncer de cerebro, por ejemplo, un glioma. La muestra es una muestra de tejido o fluido.

30 En un formato, el método implica ensayar un ácido nucleico de la muestra. El método puede comprender además someter la muestra a condiciones adecuadas para amplificar el ácido nucleico. Alternativamente, el método puede comprender poner en contacto la muestra con un anticuerpo que se une inmunológicamente a un TS10q23.3, por ejemplo, en un método ELISA. La comparación, a pesar del formato, puede incluir comparar la expresión de TS10q23.3 con la expresión de TS10q23.3 en muestras de no cáncer. La comparación puede considerar niveles de expresión de TS10q23.3. Alternativamente, la comparación puede implicar la evaluación de la estructura del gen de TS10q23.3, proteína o transcrito. Tales formatos pueden incluir secuenciación, hibridación de oligonucleótido de tipo natural, hibridación de oligonucleótido mutante, SSCP, PCR y protección de ARNasa. Las realizaciones particulares incluyen la evaluación de la hibridación del oligonucleótido mutante o de tipo natural donde el oligonucleótido está configurado en una matriz, chip o lámina.

35 También se ha descrito un método para alterar el fenotipo de una célula tumoral que comprende la etapa de poner en contacto la célula con un supresor de tumor designado TS10q23.3 bajo condiciones que permitan la absorción del supresor de tumor por la célula tumoral. La célula tumoral puede derivarse de un tejido tal como tejido de cerebro, pulmón, hígado, bazo, riñón, ganglio linfático, intestino delgado, células sanguíneas, páncreas, colon, estómago, pecho, endometrio, próstata, testículo, ovario, piel, cabeza y cuello, esófago, médula ósea y tejido sanguíneo. El fenotipo se puede seleccionar entre proliferación, migración, inhibición por contacto, crecimiento en agar blando o ciclo celular. El supresor de tumor puede estar encapsulado en un liposoma o estar libre.

40 Además se ha descrito un método para alterar el fenotipo de una célula tumoral que comprende la etapa de poner en contacto la célula con un ácido nucleico (i) que codifica un supresor de tumor designado TS10q23.3 y (ii) un promotor activo en la célula tumoral, en la que el promotor se enlaza de manera operable a la región que codifica el supresor de tumor, bajo condiciones que permitan la absorción del ácido nucleico por la célula tumoral. El fenotipo puede ser proliferación, migración, inhibición por contacto, crecimiento en agar blando o ciclo celular. El ácido nucleico puede estar encapsulado en un liposoma. Si el ácido nucleico es un vector vírico tal como retrovirus, adenovirus, virus adenoasociado, virus vaccinia y herpesvirus, puede estar encapsulado en una partícula vírica.

45 En una realización adicional, se ha proporcionado un polipéptido supresor de tumor tal como se ha definido en las reivindicaciones para usarse en el tratamiento de cáncer.

En aún una realización adicional, se ha proporcionado un ácido nucleico tal como se ha definido en las reivindicaciones para usarse en el tratamiento de cáncer, en el que dicho ácido nucleico es capaz de expresar el polipéptido supresor de tumor tal como se ha definido en las reivindicaciones en las células tumorales asociadas con dicho cáncer.

5

En más aún una realización adicional, se ha proporcionado un mamífero no humano transgénico en el cual ambas copias de gen codificador de TS10q23.3 están interrumpidas o reemplazadas con otro gen.

Se ha descrito un método de determinación de la fase de cáncer que comprende las etapas de (i) obtener una muestra de un sujeto; y (ii) determinar la expresión de un supresor de tumor TS10q23.3 funcional en células de la muestra. El cáncer puede ser un cáncer cerebral y la fase se distingue entre grado bajo y glioma. La determinación puede comprender ensayar un ácido nucleico o polipéptido TS10q23.3 en la muestra.

También se ha descrito un método que predice la metástasis tumoral que comprende las etapas de (i) obtener una muestra de un sujeto; y (ii) determinar la expresión de un supresor de tumor TS10q23.3 en células de la muestra. El cáncer se puede distinguir como metástasis y no metástasis. La determinación puede comprender ensayar un ácido nucleico o polipéptido TS10q23.3 en la muestra.

Se ha descrito un método de investigación de una sustancia candidata para la actividad antitumoral que comprende las etapas de (i) proporcionar una célula que carezca de polipéptido TS10q23.3 funcional; (ii) poner en contacto la célula con la sustancia candidata; y (iii) determinar el efecto de la sustancia candidata sobre la célula. La célula puede ser una célula tumoral, por ejemplo, una célula tumoral que tenga una mutación en la región codificadora de TS10q23.3.7. La mutación puede ser un mutante de delección, un mutante de inserción, un mutante de desplazamiento de la pauta de lectura, un mutante sin sentido, un mutante de sentido erróneo o un mutante de maduración por corte y empalme. La determinación puede comprender la comparación de una o más características de la célula en presencia de la sustancia candidata con características de una célula en ausencia de la sustancia candidata. La característica puede ser: expresión de TS10q23.3, actividad de fosfatasa, proliferación, metástasis, inhibición por contacto, crecimiento en agar blando, regulación del ciclo celular, formación tumoral, progresión tumoral e invasión de tejido. La sustancia candidata puede ser un agente quimioterapéutico o radioterapéutico o ser seleccionada entre una pequeña biblioteca molecular. La molécula puede estar en contacto *in vitro* o *in vivo*.

Además se ha descrito un método de investigación de una sustancia candidata para la actividad antiquinasa que comprende las etapas de (i) proporcionar un polipéptido TS10q23.3 que comprenda al menos un sitio de tirosin quinasa; (ii) poner en contacto la célula con la sustancia candidata; y (iii) determinar el efecto de la sustancia candidata sobre la fosforilación del sitio. La determinación puede comprender la comparación de una o más características de la célula en presencia de la sustancia candidata con características de una célula en ausencia de la sustancia candidata. La característica puede ser el estatus de fosforilación de TS10q23.3, expresión de TS10q23.3, actividad fosfatasa, proliferación, metástasis, inhibición por contacto, crecimiento en agar blando, regulación del ciclo celular, formación tumoral, progresión tumoral e invasión de tejido. La sustancia candidata puede ser un agente quimioterapéutico o radioterapéutico o ser seleccionada entre una pequeña biblioteca molecular. La célula puede estar en contacto *in vitro* o *in vivo*.

Breve descripción de los dibujos

Los siguientes dibujos forman parte de la presente memoria y se incluyen para demostrar más ciertos aspectos de la presente invención. La invención se puede entender mejor en referencia a uno o más de estos dibujos en combinación con la descripción detallada de las realizaciones específicas presentadas en la presente memoria:

Fig. 1- *Localización del Loci del Supresor de Tumor Candidato en el Cromosoma 10 Humano*. Diversos loci en el cromosoma 10 humano han estado implicados como posibles sitios para la actividad supresora de tumor. Se representan estas localizaciones, y el grupo informante.

Fig. 2- *Ilustración de Delecciones Homocigotas en Líneas Celulares de Glioma*. Se investigaron en diversas líneas celulares de glioma la presencia de delecciones en ambas copias de loci en el cromosoma 10. Los loci están indicados en el eje vertical y las líneas celulares están enumeradas a través del eje horizontal. La pérdida homocigota está indicada por un óvalo oscurecido.

Fig. 3- *Ilustración de Regiones del Cromosoma 10: Presencia o Ausencia de Marcadores Microsatélites de ADN en el Clon Híbrido*. Se ilustran las regiones de presencia (círculo cerrado) o ausencia (círculo abierto) en el cromosoma 10 del ADN correspondiente a marcadores microsatélites específicos al cromosoma 10 de once subclones del clon híbrido U251.N10.7 de la célula somática que se transfirieron en células A9 de ratón. Los híbridos de células somáticas U251.N10.6 y U251.N10.8 son clones completamente suprimidos, que presentan no crecimiento o pequeño en agarosa blanda, y los subclones U251.10.5A y C están parcialmente suprimidos (Steck *et al.*, 1995). La diferencia entre los clones completamente suprimidos y los clones parcialmente suprimidos proporciona una localización funcional del gen del supresor de tumor. Las posibles regiones que contienen el gen supresor de tumor están indicadas por los recuadros tramados. El recuadro tramado en 10q23.3 se solapa con las delecciones homocigotas y la región implicada por el análisis de delección alélica (véase Fig. 2 y Fig. 4).

Fig. 4- *Mapa de Delección del Cromosoma 10 en Gliomas Humanos*. La región unida por los marcadores D10S551 a D10S583 está localizada en una región de 10 cM. Se muestra que los microsatélites son los más probablemente enlazados y se mapean en su localización cromosómica aproximada en base al mapa de híbrido de tal radiación como lo descrito por Gyapay *et al.*, 1994. La región del cromosoma 10 que no está implicada en astrocitomas anaplásicos y un glioma se muestra en las regiones recuadradas del tumor. La región crítica definida a partir del análisis de delección homocigota y no excluida se muestra mediante este análisis por la barra sólida en el lado derecho.

Fig. 5- *Mapeado de BAC 106d16*. Se ilustra el mapeado de BAC designado 106d16, y la demostración de la delección homocigota mediante transferencia Southern. Se representa el mapa parcial de la restricción de 106d16. La ilustración de la transferencia muestra la delección homocigótica de la banda nº 14 de *Eco* (Mr aprox. 11 kb) en células EFC-2.

Fig. 6- *Secuencia codificadora y regiones flanqueadoras 5' y 3' de TS10q23.3*. La región codificadora está en negrita como la primera en el codón de terminación de marco.

Fig. 7- *Secuencia de Aminoácidos Pronosticada del Producto TS10q23.3*. Las abreviaciones son: A, alanina; C, cisteína; D, ácido aspártico; E, ácido glutámico; F, fenilalanina; G, glicina; H, histidina; I, isoleucina; K, lisina; L, leucina; M, metionina; N, asparagina; P, prolina; Q, glutamina; R, arginina; S, serina; T, treonina; V, valina; W, triptófano; Y, tirosina. El sitio consenso de fosfatasa está en negrita; los sitios de fosforilación de tirosina están en cursiva y subrayados.

Fig. 8- *Análisis de Delección de 10q23.3*. Se volvió a analizar la presencia del gen de TS10q23.3 en la línea de glioma inicialmente indicada como que tiene delecciones homocigóticas en 10q23.3. El óvulo en oscuro indica que la región del gen está presente; el óvulo abierto indica una delección homocigótica en la región del gen. *- indica exones atrapados.

Fig. 9- *Comparación de Homología de TS10q23.3 Humano con Homólogos de Ratón y Perro*. El codón ATG de iniciación y el aminoácido de metionina están designados en la posición de comienzo (1). El codón de terminación es TGA (1.210). Las alteraciones entre las secuencias humanas y de ratón o perro en el nivel genómico o de aminoácidos están designadas por una estrella en la secuencia comparada. Sin embargo, no se observaron cambios en la secuencia de aminoácidos.

Fig. 10- *Secuencia de Exones y Regiones Intrónicas de Alrededor de TS10q23.3*. Los exones están indicados con letras mayúsculas comenzando en la posición uno y los intrones están designados con letras minúsculas; excepto para el primer exón donde el codón de iniciación comienza en la posición uno y el límite exón/intrón 3' está en la posición 79 y 80, respectivamente. La designación con letra minúscula (Tabla 4) corresponde a la numeración de la secuencia presentada en esta figura, excepto para el primer exón. Las mutaciones para U87 y U138 están en el primer residuo de intrón G [G+I>T] después del exón (G y H, respectivamente). Para T98G y KE, las mutaciones puntuales están en las posiciones 46 y 28 del exón B, respectivamente. Para las células LnCap, la mutación es una delección de las bases 16 y 17 en el primer intrón.

Figs. 11A-G- *Análisis de las Estructuras Secundarias en TS10q23.3*. Fig. 11A: Gráfico de hidrofiliidad; Fig. 11B: Gráfico de probabilidad superficial; Fig. 11C: Gráfico de flexibilidad; Fig. 11D: Gráfico de índice antigénico; Fig. 11E: Gráfico de hélice anfifílica; Fig. 11F: Gráfico de lámina anfifílica; Fig. 11G: Gráfico de estructura secundaria.

Figs. 12A-I- *Comparación de las Características Pronosticadas en TS10q23.3 y Mutantes Puntuales T98G y KE*. Fig. 12A: Gráfico de hidrofiliidad de los residuos 1-60 del polipéptido de tipo natural; Fig. 12B: Gráfico de probabilidad superficial de los residuos 1-60 del polipéptido de tipo natural; Fig. 12C: Gráfico de estructura secundaria de los residuos 1-60 del polipéptido de tipo natural; Fig. 12D: Gráfico de hidrofiliidad de los residuos 1-60 del mutante KE; Fig. 12E: Gráfico de probabilidad superficial de los residuos 1-60 del mutante KE; Fig. 12F: Gráfico de estructura secundaria de los residuos 1-60 del mutante KE; Fig. 12G: Gráfico de hidrofiliidad de los residuos 1-60 del mutante T98G; Fig. 12H: Gráfico de probabilidad superficial de los residuos 1-60 del mutante T98G; Fig. 12I: Gráfico de estructura secundaria de los residuos 1-60 del mutante T98G. La mutación T98G (Leu→Arg) en el residuo 42 da como resultado la pérdida de la estructura secundaria en hélice propuesta de TS10q23.3. La mutación en KE en el residuo 36 (Gly→Glu) aumentaría de manera significativa la longitud de la estructura helicoidal propuesta en la región. Ambas mutaciones afectarían a la misma estructura helicoidal. También, surgen cambios menores en la hidrofiliidad y en la probabilidad superficial.

Compendio de las secuencias

SEC. ID. Nº: 1=secuencia pronosticada para TS10q23.3; SEC. ID. Nº: 2=secuencia de gen humano; SEC. ID. Nº: 3=secuencia de gen de ratón; SEC. ID. Nº: 4=secuencia de gen de perro; SEC. ID. Nº: 5=secuencia peptídica humana; SEC. ID. Nº: 6=secuencia peptídica de ratón; SEC. ID. Nº: 7=secuencia peptídica de perro; SEC. ID. Nº: 8=exón a; SEC. ID. Nº: 9=exón b; SEC. ID. Nº: 10=exón c; SEC. ID. Nº: 11=exón d; SEC. ID. Nº: 12=exón e; SEC. ID. Nº: 13=exón f; SEC. ID. Nº: 14=exón g; SEC. ID. Nº: 15=exón h; SEC. ID. Nº: 16=exón i; SEC. ID. Nº: 17=un motivo de los residuos 88 a 98; SEC. ID. Nº: 18=dominio catalítico conservado de una protein tirosin fosfatasa (Denu *et al.*, 1996); SEC. ID. Nº: 19=residuos 1-60 del polipéptido de tipo natural TS10q23.3; SEC. ID. Nº: 20=residuos 1-60 del polipéptido TS10q23.3 del mutante T98G; SEC. ID. Nº: 21=residuos 1-60 del polipéptido TS10q23.3 del mutante KE.

Descripción detallada de las realizaciones preferidas

I. La presente Invención

5 Tal como se ha planteado anteriormente, un número de diferentes grupos han mostrado evidencia de una actividad supresora de tumor asociada con la región 10q del cromosoma humano 10. A pesar de esta considerable cantidad de trabajo, no se ha determinado la identidad del gen o genes responsables para esta actividad. Los anteriores usaron un enfoque funcional que implica la transferencia de los cromosomas o fragmentos cromosómicos que se sospecha que albergan el(los) gen(es) supresor(es) de tumor dentro de las células de glioma tumorígeno. Estos esfuerzos permitieron la definición de la actividad biológica de el(los) gen(es) supresor(es) de tumor putativo(s) y ayudan en la localización de tales. Los cromosomas 2 y 10 se transfirieron en las células de glioma U251 y los cromosomas 2 y 10 dentro de las células LG-11. Se mostró que dentro de las células LG-11 tenían copias no intactas del cromosoma 10 y posteriormente se encontró que el punto de rotura ocurría en 10q24. La transferencia del cromosoma 10 dio como resultado células híbridas que exponían un fenotipo suprimido, presentando una pérdida de tumorigenicidad (no formación de tumor) y pérdida de la capacidad de crecer en agarosa blanda (disminuido 50x a 1000x; Pershouse *et al.*, 1993). El índice de crecimiento exponencial del híbrido era similar a la célula parental, aunque la densidad de saturación de la célula híbrida era de manera significativa (10x a 20x) menor que las células parentales. La transferencia del cromosoma 2 dio como resultado células híbridas que actuaban similarmente a las células parentales.

20 Un objetivo de estos estudios era localizar el gen supresor del cromosoma 10 mediante fragmentación del cromosoma 10 etiquetado con neomicina y, posteriormente, transferir el cromosoma fragmentado dentro de las células de glioma. Sin embargo, los inventores observaron que algunas de las células híbridas tenían reorganizaciones cromosómicas experimentadas espontáneamente para producir células híbridas que retenían solamente diversas regiones del cromosoma 10 insertado (Pershouse *et al.*, 1993). A continuación, los inventores subclonaron los híbridos y los analizaron, además de iniciar estudios de fragmentación (Steck *et al.*, 1995). Se siguió la pista de la retención del cromosoma 10 insertado o sus fragmentos mediante marcadores RFLP informativos y análisis FISH. Interesantemente, se sometió a reorganización solamente el cromosoma insertado. La inserción de una copia entera del cromosoma 10 dio como resultado la propiedad transformada de la célula híbrida para proliferar en agarosa blanda y formar tumores en ratones desnudos.

30 Estos dos fenotipos parecen ahora ser parcialmente separables mediante el presente análisis. Algunos subclones (U251.N10.5a-j), los cuales revelaron una pérdida de una porción muy importante del brazo largo del cromosoma 10, crecieron en agarosa blanda pero fallaron en formar tumores en ratones nude, indicando así que un locus supresor de tumor reside en la porción restante del cromosoma (10pter a 10q11). Por lo contrario, los clones que retenían una región distal del brazo largo, 10q24 a 10q26, fallaron en proliferar tanto en agarosa blanda como en ratones nude (véase Fig. 4). Esto sugiere otra región supresora fenotípica que reside en la región distal del cromosoma. La carencia de material asociado a 10 adicional sugeriría además que el material del cromosoma 10 restante es responsable del fenotipo biológico alterado. Estos resultados implican la presencia de dos regiones supresoras fenotípicamente independientes en el cromosoma 10 implicado en la progresión del glioma (Steck *et al.*, 1995).

40 De acuerdo con la presente invención, los inventores han usado en la actualidad diversas estrategias independientes para localizar un gen supresor de tumor, designado TS10q23.3, que está implicado en gliomas, cáncer de pecho, cáncer de próstata y otros cánceres. Estos enfoques, descritos en mayor detalle en los siguientes Ejemplos, incluían (i) la identificación de deleciones homocigóticas en una serie de líneas celulares de glioma humanos; (ii) determinación de una(s) región(es) consistente(s) de retención en clones suprimidos para la tumorigenicidad; y (iii) estudios de deleción alélica sobre diversos grados de glioma y muestras normales correspondientes. Con el gen en mano, ahora llega a ser posible explotar la información codificada por el gen para desarrollar enfoques novedosos de diagnóstico y terapéuticos relacionados con el cáncer humano.

50 II. El Supresor de Tumor 10q23.3

De acuerdo con la presente invención, se ha identificado un supresor de tumor, codificado por un gen en el locus 10q23.3, y designado en la presente memoria como TS10q23.3. Esta molécula es capaz de suprimir fenotipos tumorales en diversas células. El término supresor de tumor es bien conocido por los expertos en la técnica. Ejemplos de otros supresores de tumor son p53, Rb y p16, por nombrar unos pocos. Mientras que estas moléculas son estructuralmente distintas, forman grupo de moléculas funcionalmente relacionadas, del cual TS10q23.3 es un miembro. Los usos para los cuales estos otros supresores de tumor están siendo explotados en la actualidad son igualmente aplicables en la presente memoria.

60 Además de la molécula TS10q23.3 entera, la presente invención también se refiere a fragmentos del polipéptido que tiene función fosfatasa o fragmentos adecuados como antígeno tal como el definido en las reivindicaciones.

Los fragmentos que incluyen el terminal N de la molécula se puede generar por ingeniería genética de sitios de terminación de traducción dentro de la región codificadora (discutido a continuación). Alternativamente, el tratamiento de la molécula TS10q23.3 con enzimas proteolíticas, conocidas como proteasas, pueden producir una diversidad de fragmentos N-terminal, C-terminal e internos. Los ejemplos de los fragmentos pueden incluir residuos contiguos de la secuencia de TS10q23.3 dados en la Fig. 7 y Fig. 9, de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,

ES 2 361 044 T3

23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, 100, 200, 300, 400 o más aminoácidos de longitud. Estos fragmentos se pueden purificar de acuerdo con métodos conocidos, tales como precipitación (por ejemplo, sulfato de amonio), HPLC, cromatografía de intercambio iónico, cromatografía de afinidad (incluyendo cromatografía de inmunoafinidad) o diversas separaciones por tamaño (sedimentación, electroforesis en gel, filtración en gel).

5

A. Características estructurales del Polipéptido

El gen para TS10q23.3 codifica un polipéptido de 403 aminoácidos. El peso molecular pronosticado de esta molécula es de 47.122 con un pI resultante de 5,86. Por tanto, como mínimo, esta molécula se puede usar como un patrón en el ensayo donde se están examinando el peso molecular y el pI.

Un sitio consenso de fosfatasa localizado en los residuos 122-131, que se corresponde perfectamente a la secuencia consenso de tirosin fosfatasa (PTP): [I/V]HCxAGxxR[S/T]G. Fuera de los dominios activos, las secuencias difieren enormemente. Las PTPs continúan a través de intermediarios de fosfoenzima. La reacción enzimática implica la formación del intermediario fosforil-cisteína después del ataque nucleófilo del átomo de fósforo del sustrato mediante el anión tiolato de cisteína. La reacción se puede representar como un proceso químico en dos etapas: la transferencia del fosforil a la enzima acompañado por liberación rápida del producto desfosforilado; y la hidrólisis del intermediario tiol-fosfato concomitante con liberación rápida del fosfato. Para formar el complejo componente catalítico, la enzima se une y reacciona con el dianoión del sustrato que contiene fosfato. Sobre la enzima, un ácido aspártico debe estar protonado y la cisteína nucleófila debe no estar protonada (anión de tiolato) para la fosforil transferencia a la enzima. También hay que fijarse en los sitios de fosforilación de tirosina potenciales localizados en los residuos 233-240 y 308-315 y en los sitios de fosforilación de cAMP localizados en los residuos 128, 164, 223 y 335. Se sabe que las fosfatasas tienen sitios de quinasa, y la actividad fosfatasa de estas enzimas se pueden modular mediante fosforilación en estos sitios. Generalmente las fosfatasas de proteína están divididas en dos categorías: fosfatasas de serina/treonina y fosfatasas de tirosina. Ciertas fosfatasas de tirosina también tienen actividad frente a la fosfoserina y fosfotreonina.

La interacción entre quinastas y fosfatasa, y los diversos estados de fosforilación de los polipéptidos se han demostrado como rasgos importantes en la fisiología celular. A través de una diversidad de diferentes mecanismos, quinastas y fosfatasas actúan en diferentes rutas dentro de las células que están implicadas en la señalización, almacenamiento de energía y regulación celular. Puesto que se ha demostrado que la identificación de una función tirosin quinasa intrínseca en la proteína transformada de src (Collett & Erickson, 1978), el papel de la fosforilación, particularmente sobre los residuos de tirosina, juega un papel crítico en el control de la proliferación celular y la inducción del cáncer (Hunter, 1991; Bishop, 1991). Los papeles que juegan las fosfatasas de proteína en la regulación del crecimiento, así como en muchas otras actividades biológicas y bioquímicas, se han correlacionado con el estado de fosforilación de las moléculas biológicamente importantes (Cohen, 1994).

También se debería mencionar que los más o menos 60 aminoácidos del terminal N de la molécula muestra alguna homología con la tensina, una proteína citoesquelética implicada en placas de adhesión. Esto sugiere que TS10q23.3 puede estar implicado en fenómenos de superficie celular tales como inhibición por contacto, invasión, migración o señalización célula-célula. Las mutaciones puntuales de TS10q23.3 identificadas en ciertas líneas celulares tumorales afectan a las estructuras secundarias propuestas en esta región.

45

B. Aspectos Funcionales

Cuando la presente solicitud se refiere a la función de TS10q23.3 o actividad de "tipo natural", significa que la molécula en cuestión tiene la capacidad de inhibir la transformación de la célula desde un estado normalmente regulado de proliferación a estado maligno, es decir, se asocia con cualquier tipo de regulación anormal del crecimiento, o de inhibir la transformación de una célula desde un estado anormal a un estado altamente maligno, por ejemplo, para prevenir metástasis o crecimiento tumoral invasivo. Otros fenotipos que se pueden considerar para ser regulados por el producto de gen de TS10q23.3 normal son angiogénesis, adhesión, migración, señalización célula-célula, crecimiento celular, proliferación celular, crecimiento dependiente de densidad, crecimiento dependiente de anclaje y otros. La determinación de qué moléculas poseen esta actividad se puede determinar usando ensayos familiares para los expertos en la técnica. Por ejemplo, la transferencia de genes que codifican TS10q23.3, o variantes de los mismos, dentro de las células que no tienen un producto TS10q23.3 funcional, y por lo tanto presentan control de crecimiento dañado, identificará, en virtud de la supresión del crecimiento, aquellas moléculas que tienen función de TS10q23.3.

Tal como se ha planteado anteriormente, hay una indicación de que TS10q23.3 es una fosfatasa. La porción de la proteína localizada en los residuos 88 a 98 es una correspondencia perfecta para el dominio catalítico conservado de la proteína tirosin fosfatasa. También, las diana de quinasa putativa están localizadas en la molécula, lo cual es otra característica de las fosfatasas. Debido a que otros supresores de tumor se han identificado con este tipo de actividad, será deseable determinar la función fosfatasa en el papel de supresión de tumor de TS10q23.3. Esto también puede ser un enfoque fructífero para desarrollar ensayos de investigación de la ausencia de función de TS10q23.3 o en el desarrollo de terapias de cáncer, por ejemplo, en direccionamiento de la función de la fosfatasa de TS10q23.3, direccionamiento del sustrato sobre el cual actúa TS10q23.3 y/o direccionamiento de la quinasa o quinastas que actúan sobre TS10q23.3.

C. Variantes de TS10q23.3

Las variantes de secuencia de aminoácidos del polipéptido pueden ser variantes sustitucionales, insercionales o de delección. Las variantes de delección carecen de uno o más residuos de la proteína nativa los cuales no son esenciales para la función o actividad inmunógena y se ilustran por las variantes que carecen de una secuencia transmembrana descrita anteriormente. Otro tipo común de variante de delección es una que carece de secuencias señal secretora o secuencias señal que dirigen una proteína a unirse a una parte particular de una célula. Los mutantes insercionales típicamente implican la adición de material en un punto no terminal en el polipéptido. Esto puede incluir la inserción de un epítipo inmunoreactivo o simplemente un residuo único. Las adiciones terminales, denominadas proteínas de fusión, se discuten a continuación.

Las variantes sustitucionales típicamente contienen el intercambio de un aminoácido por otro en uno o más sitios dentro de la proteína, y se pueden diseñar para modular una o más propiedades del polipéptido, tal como la estabilidad frente a la escisión proteolítica, sin la pérdida de otras funciones o propiedades. Las sustituciones de esta clase preferentemente son conservadoras, es decir, un aminoácido está reemplazado con uno de forma y carga similar. Las sustituciones conservadoras son bien conocidas en la técnica e incluyen, por ejemplo, los cambios de alanina por serina; arginina por lisina; asparagina por glutamina o histidina; aspartato por glutamato; cisteína por serina; glutamina por asparagina; glutamato por aspartato; glicina por prolina; histidina por asparagina o glutamina; isoleucina por leucina o valina; leucina por valina o isoleucina; lisina por arginina; metionina por leucina o isoleucina; fenilalanina por tirosina, leucina o metionina; serina por treonina; treonina por serina; triptófano por tirosina; tirosina por triptófano o fenilalanina; y valina por isoleucina o leucina.

Lo siguiente es una discusión basada en el cambio de los aminoácidos de una proteína para crear un equivalente, o incluso una molécula de segunda generación mejorada. Por ejemplo, ciertos aminoácidos se pueden sustituir por otros aminoácidos en una estructura proteica sin pérdida apreciable de la capacidad de unión interactiva con estructuras tales como, por ejemplo, regiones de unión a antígeno de anticuerpos o sitios de unión sobre moléculas de sustrato. Puesto que es la capacidad interactiva y la naturaleza de una proteína lo que define esa actividad funcional biológica de la proteína, se pueden hacer ciertas sustituciones de aminoácidos en una secuencia de proteína, y su secuencia codificadora de ADN subyacente, y no obstante obtener una proteína con propiedades similares. Se pueden hacer diversos cambios en las secuencias de ADN de genes sin pérdida apreciable de su utilidad biológica o actividad, tal como se discute a continuación. La Tabla 1 muestra los codones que codifican aminoácidos particulares.

En la realización de tales cambios, se puede considerar el índice hidropático de los aminoácidos. La importancia del índice hidropático de los aminoácidos en conferir función biológica interactiva en una proteína generalmente se entiende en la técnica (Kyte & Doolittle, 1982). Se acepta que el carácter hidropático relativo del aminoácido contribuye a la estructura secundaria de la proteína resultante, la cual por turnos define la interacción de la proteína con otras moléculas, por ejemplo, enzimas, sustratos, receptores, anticuerpos de ADN, antígenos y similares.

Se ha asignado a cada aminoácido un índice hidropático en base a sus características de hidrofobicidad y carga (Kyte & Doolittle, 1982), estas son: isoleucina (+4,5); valina (+4,2); leucina (+3,8); fenilalanina (+2,8); cisteína/cistina (+2,5); metionina (+1,9); alanina (+1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparagina (-3,5); lisina (-3,9) y arginina (4,5).

En la técnica se sabe que ciertos aminoácidos se pueden sustituir por otros aminoácidos que tienen un índice o resultado hidropático similar y todavía dan como resultado una proteína con actividad biológica similar, es decir, todavía obtienen una proteína funcionalmente equivalente biológica. En la realización de tales cambios, es preferida la sustitución de aminoácidos cuyos índices hidropáticos están dentro de ± 2 , particularmente se prefieren aquellas que están dentro de ± 1 , e incluso son más particularmente preferidas aquellas dentro de $\pm 0,5$.

También se entiende en la técnica que la sustitución de aminoácidos similares se puede hacer eficazmente en base a la hidrofiliidad. El documento de Patente U.S. 4.554.101 plantea que la mayor hidrofiliidad promedio local de una proteína, como la regida por la hidrofiliidad de sus aminoácidos adyacentes, se correlaciona con una propiedad biológica de la proteína. Tal como se detalla en el documento de Patente U.S. 4.554.101, se han asignado los siguientes valores de hidrofiliidad a los residuos de aminoácidos: arginina (+3,0); lisina (+3,0); aspartato (+3,0 \pm 1); glutamato (+3,0 \pm 1); serina (+3,0); asparagina (+0,2); glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,5 \pm 1); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5); triptófano (-3,4).

Se entiende que un aminoácido puede ser sustituido por otro que tenga un valor de hidrofiliidad similar y todavía obtienen una proteína biológicamente equivalente e inmunológicamente equivalente. En tales cambios, se prefiere la sustitución de aminoácidos cuyos índices hidropáticos están dentro de ± 2 , particularmente se prefieren aquellos que están dentro de ± 1 , e incluso son más particularmente preferidos aquellos dentro de $\pm 0,5$.

Tal como se ha esbozado anteriormente, las sustituciones de aminoácidos generalmente se basan en la similitud relativa de los sustituyentes de la cadena lateral de aminoácidos, por ejemplo, su hidrofobicidad, hidrofiliidad, carga, tamaño y similares. Sustituciones ilustrativas que toman diversas de las características anteriormente mencionadas en consideración son bien conocidas por los expertos en la técnica e incluyen: arginina y lisina; glutamato y aspartato; serina y treonina; glutamina y asparagina; y valina, leucina e isoleucina.

Los polipéptidos también se pueden preparar mediante el uso de miméticos de péptidos. Los miméticos son moléculas que contienen péptidos que mimetizan elementos de la estructura secundaria de la proteína. Véase, por ejemplo, Johnson *et al.*, "Peptide Turn Mimetics" in BIOTECHNOLOGY AND PHARMACY, Pezzuto *et al.*, Eds., Chapman and Hall, Nueva York (1993). Las razones subyacentes detrás del uso de los miméticos del péptido es que la cadena principal del péptido de las proteínas existe sobre todo para orientar las cadenas laterales de aminoácidos de tal modo que facilite las interacciones moleculares, tal como aquellas del anticuerpo y antígeno. Se espera que un mimético de péptido permita las interacciones moleculares similares a la molécula natural. Estos principios se pueden usar, junto con los principios anteriormente esbozados, para producir por ingeniería moléculas de segunda generación que tengan muchas de las propiedades naturales de TS10q23.3, pero con características alteradas e incluso mejoradas.

D. Conmutación de Dominio

Tal como se describe en los ejemplos, los presentes inventores han identificado homólogos putativos de murine y caninos del gen humano de TS10q23.3. Además, se han identificado mutaciones en TS10q23.3 las cuales se cree que alteran su función. Estos estudios son importantes por al menos dos razones. En primer lugar, proporcionan una expectativa razonable de que puedan existir todavía otros homólogos, variantes alélicas y mutantes de este gen en especies relacionadas, tales como rata, conejo, mono, gibón, chimpancé, simio, babuino, vaca, cerdo, caballo, oveja y gato. Sobre el aislamiento de estos homólogos, variantes y mutantes, y junto con otros análisis, se pueden identificar ciertos dominios activos y funcionales. En segundo lugar, esto proporcionará un punto de partida para análisis mutacionales adicionales de la molécula. Una forma en la cual se puede explotar esta información es en "conmutación de dominio".

La conmutación de dominio implica la generación de moléculas quiméricas usando diferentes pero, en este caso, polipéptidos relacionados. Al comparar las secuencias de ratón, perro y humanas para TS10q23.3 con el TS10q23.3 de otras especies, y con mutantes y variantes alélicas de estos polipéptidos, se pueden hacer predicciones para las regiones funcionalmente significativas de estas moléculas. A continuación, es posible conmutar dominios relacionados de estas moléculas en un esfuerzo para determinar la capacidad crítica de estas regiones para la función de TS10q23.3. Estas moléculas pueden tener valor adicional en el sentido de que estas "quimeras" se pueden distinguir de las moléculas naturales, mientras que posiblemente proporcionan la misma función.

En base a la identidad de secuencia, al nivel de aminoácidos, de las secuencias de ratón, perro y humano, se puede deducir que incluso pequeños cambios en la secuencia primaria de la molécula afectarán a la función. Análisis adicionales de mutaciones y su efecto pronosticado sobre la estructura secundaria ampliará este entendimiento.

Otro aspecto estructural de TS10q23.3 que proporciona sustrato fértil para los experimentos de conmutación de dominio es el dominio similar a fosfatasa de tirosina y los sitios putativos de fosforilación de tirosina. Este dominio puede estar sustituido por otros dominios de fosfatasa para alterar la especificidad de esta función. Mediante esta observación se garantiza una investigación adicional de la homología entre TS10q23.3 y otras fosfatasas.

E. Proteínas de Fusión

Una clase especializada de variante insercional es la proteína de fusión. Esta molécula generalmente tiene toda o una porción sustancial de la molécula nativa, enlazada al terminal N o C, a todo o una porción de un segundo polipéptido. Por ejemplo, las fusiones típicamente emplean secuencias líder de otras especies para permitir la expresión recombinante de una proteína en un hospedante heterólogo. Otra fusión útil incluye la adición de un dominio inmunológicamente activo, tal como un epítipo de anticuerpo, para facilitar la purificación de la proteína de fusión. La inclusión de un sitio de escisión en o cerca de la junta de fusión facilitará la separación del polipéptido extraño después de la purificación. Otras fusiones útiles incluyen el enlace de los dominios funcionales, tal como los sitios activos de las enzimas, dominios de glicosilación, señales diana celulares o regiones de transmembrana. Una fusión particular de interés incluiría una construcción de delección que carezca del sitio de fosfatasa de TS10q23.3 pero conteniendo otras regiones que podrían unirse a la molécula sustrato. La fusión a un polipéptido que se puede usar para la purificación del complejo sustrato-TS10q23.3 serviría para aislar el sustrato para la identificación y el análisis.

F. Purificación de Proteínas

Será deseable purificar TS10q23.3 o variantes del mismo. Las técnicas de purificación de proteína son bien conocidas por los expertos en la técnica. Estas técnicas implican, en un nivel, la fraccionación en bruto del entorno celular a fracciones de polipéptidos y no polipéptidos. Al haber separado el polipéptido a partir de otras proteínas, el polipéptido de interés se puede purificar más usando técnicas cromatográficas y electroforéticas para alcanzar la purificación parcial o completa (o purificación a homogeneidad). Los métodos analíticos particularmente adecuados para la preparación de un péptido puro son: cromatografía de intercambio iónico, cromatografía de exclusión; electroforesis en gel de poliácridamida; enfoque isoeléctrico. Un método particularmente eficiente de purificación de péptidos es la cromatografía líquida de proteína rápida o incluso la HPLC.

Ciertos aspectos de la presente invención conciernen a la purificación, y a la purificación sustancial, de una proteína o péptido codificado. El término “proteína o péptido purificado” tal como se usa en la presente memoria, pretende referirse a una composición, capaz de aislarse de otros componentes, en la que la proteína o el péptido están purificados a cualquier nivel relativo a su estado obtenible de manera natural. Por lo tanto, una proteína o péptido purificado también se refiere a una proteína o péptido, libre del ambiente en el cual puede ocurrir de manera natural.

Generalmente, “purificado” se referirá a una composición de proteína o péptido que se ha sometido a fraccionación para separar otros diversos componentes, y dicha composición sustancialmente retiene su actividad biológica expresada. Cuando se usa el término “sustancialmente purificado”, esta designación se referirá a una composición en la cual la proteína o péptido forman el componente principal de la composición, de modo que constituya aproximadamente 50%, aproximadamente 60%, aproximadamente 70%, aproximadamente 80%, aproximadamente 90%, aproximadamente 95% o más de las proteínas en la composición.

Diversos métodos para cuantificar el nivel de purificación de la proteína o péptido serán conocidos por los expertos en la técnica teniendo en cuenta la presente descripción. Estos incluyen, por ejemplo, determinar la actividad específica de una fracción activa, o valorar la cantidad de polipéptidos dentro de una fracción mediante análisis SDS/PAGE. Un método preferido para valorar la pureza de una fracción es calcular la actividad específica de la fracción, para compararlo con la actividad específica del extracto inicial, y así calcular el nivel de pureza, valorado en la presente memoria mediante un “número de veces la purificación”. La unidades actuales usadas para representar la cantidad de actividad dependerán, por supuesto, de la técnica de ensayo particular elegida para seguir la purificación y si el péptido o la proteína expresados presentan o no una actividad detectable.

Diversas técnicas adecuadas para usar en la purificación de proteína serán bien conocidas por los expertos en la técnica. Estas incluyen, por ejemplo, precipitación con sulfato de amonio, PEG, anticuerpos y similares o mediante desnaturalización por calor, seguido de centrifugación; etapas de cromatografía tales como intercambio iónico, filtración en gel, fase inversa, cromatografía de hidroxilapatita y de afinidad; enfoque isoeléctrico; electroforesis en gel; y combinaciones de dichas y otras técnicas. Tal como generalmente se conoce en la técnica, se cree que se puede cambiar el orden de conducción de las diversas etapas de purificación, o que se pueden omitir ciertas etapas, y todavía dar como resultado un método adecuado para la precipitación de una proteína o péptido sustancialmente purificado.

No hay un requerimiento general de que la proteína o péptido siempre sean proporcionados en su estado más purificado. En efecto, se contempla que productos menos sustancialmente purificados tendrán utilidad en ciertas realizaciones. La purificación parcial se puede conseguir usando menos etapas de purificación en combinación, o utilizando diferentes formas del mismo esquema de purificación general. Por ejemplo, se aprecia que una cromatografía en columna de intercambio catiónico realizada utilizando un aparato de HPLC generalmente dará como resultado una purificación “unas veces” mayor que la misma técnica utilizando un sistema de cromatografía de baja presión. Los métodos que presentan un grado inferior de purificación relativa pueden tener ventajas en la recuperación total del producto proteico, o en el mantenimiento de la actividad de una proteína expresada.

Se sabe que la migración de un polipéptido puede variar, algunas veces de manera significativa, con diferentes condiciones de SDS/PAGE (Capaldi *et al.*, 1977). Por lo tanto, se apreciará que bajo diferentes condiciones de electroforesis, pueden variar los pesos moleculares aparentes de productos de expresión purificados o parcialmente purificados.

La Cromatografía Líquida de Alta Resolución (HPLC, del Inglés “High Performance Liquid Chromatography”) se caracteriza por una separación muy rápida con extraordinaria resolución de los picos. Esto se alcanza mediante el uso de partículas muy finas y alta presión para mantener un caudal adecuado. La separación se puede conseguir en cuestión de minutos, o como mucho en una hora. Además, se necesita solamente un volumen muy pequeño de la muestra porque las partículas son tan pequeñas y cerradamente empaquetadas que el volumen en vacío es una fracción muy pequeña del volumen del lecho. Además, la concentración de la muestra no necesita ser muy grande porque las bandas son tan estrechas que hay muy poca dilución de la muestra.

La cromatografía en gel, o cromatografía de criba molecular, es un tipo especial de cromatografía de partición que se basa en el peso molecular. La teoría detrás de la cromatografía en gel es que la columna, la cual se prepara con diminutas partículas de una sustancia inerte que contiene poros pequeños, separa las moléculas más grandes de las moléculas más pequeñas a medida que pasan a través de o alrededor de los poros, dependiendo de su tamaño. Mientras el material del cual están hechas las partículas no absorban las moléculas, el único factor que determina el caudal es el tamaño. Por lo tanto, las moléculas se eluyen de la columna en tamaño descendente, siempre que la forma sea relativamente constante. La cromatografía en gel es insuperable para separar moléculas de diferente tamaño porque la separación es independiente de todos los otros factores tales como pH, fuerza iónica, temperatura, etc. Prácticamente tampoco hay absorción, hay menos difusión de zona y el volumen de elución está relacionado con una simple cuestión con el peso molecular.

La cromatografía de afinidad es un procedimiento cromatográfico que depende de la afinidad específica entre una sustancia a aislar y una molécula a la que se puede unir específicamente. Esto es una interacción de tipo ligando-receptor. El material de columna está sintetizado mediante acoplamiento covalente de uno de las parejas de la unión a una matriz insoluble. Entonces el material de la columna es capaz de absorber específicamente la sustancia de la disolución. La elución ocurre cambiando las condiciones por aquellas en las que no ocurrirá la unión (modificar pH, fuerza iónica, temperatura, etc).

Un tipo particular de cromatografía de afinidad útil en la purificación del carbohidrato que contiene compuestos es la cromatografía de afinidad de lectina. Las lectinas son una clase de sustancias que se unen a una diversidad de polisacáridos y glicoproteínas. Las lectinas normalmente se acoplan a agarosa mediante bromuro de cianógeno. La conconavalina A acoplada a Sefarosa fue el primer material de este tipo en usarse y se ha usado mucho en el aislamiento de polisacáridos y glicoproteínas. Otras lectinas que se han incluido lectina de lenteja, aglutinina de germen de trigo las cuales han sido útiles en la purificación de residuos de N-acetil glucosaminil y lectina de *Helix pomatia*. Las propias lectinas son purificadas usando cromatografía de afinidad con ligandos de carbohidrato. La lactosa se ha usado para purificar lectinas a partir de semilla de ricino y cacahuetes; la maltosa ha sido útil en la extracción de lectinas de lentejas y judía blanca (“jack bean”); la N-acetil-D galactosamina se usa para purificar lectinas de soja; la N-acetil glucosaminil se une a lectinas de germen de trigo; la D-galactosamina se ha usado en la obtención de lectinas de almejas y la L-fucosa se unirá a lectinas de loto.

La matriz debería ser una sustancia que por sí misma no absorba moléculas a cualquier extensión significativa y que tenga un amplio intervalo de estabilidad química, física y térmica. El ligando se debería acoplarse de tal forma que no afecte a sus propiedades de unión. El ligando también debería proporcionar unión relativamente ajustada. Y debería ser posible eluir la sustancia sin destruir la muestra o el ligando. Una de las formas más comunes de cromatografía de afinidad es la cromatografía de inmovilización. A continuación se discute la generación de anticuerpos que serían adecuados para usarse de acuerdo con la presente invención.

G. Péptidos Sintéticos

La presente invención también describe péptidos relacionados con TS10q23.3 más pequeños para usar en diversas realizaciones de la presente invención. A causa de su tamaño relativamente pequeño, los péptidos de la invención también se pueden sintetizar en disolución o sobre un soporte sólido de acuerdo con las técnicas convencionales. Diversos sintetizadores automáticos están comercialmente disponibles y se pueden usar de acuerdo con los protocolos conocidos. Véase, por ejemplo, Stewart y Young, (1984); Tam *et al.*, (1983); Marrisfield, (1986); y Barany y Merufield (1979).

Secuencias cortas de péptido, o bibliotecas de péptidos que se solapan, normalmente desde aproximadamente 6 hasta aproximadamente 35 a 50 aminoácidos, los cuales corresponden a las regiones seleccionadas descritas en la presente memoria, pueden ser sintetizadas y a continuación investigadas en ensayos de investigación diseñados para identificar péptidos reactivos. Alternativamente, se puede emplear tecnología de ADN recombinante en la que una secuencia de nucleótidos que codifica un péptido de la invención se inserta dentro de un vector de expresión, se transforma o se somete a transfección dentro de una célula hospedante apropiada y cultivada bajo condiciones adecuadas para la expresión.

H. Composiciones de antígeno

La presente invención también proporciona el uso de proteínas de TS10q23.3 o fragmentos de polipéptido como antígenos para la inmunización de animales en relación con la producción de los anticuerpos. Se prevé que o bien TS10q23.3 o porciones del mismo se acoplarán, adherirán, unirán, conjugarán o se enlazarán químicamente a uno o más agentes vía enlaces (“linkers”), polienlaces (“polylinkers”) o aminoácidos derivatizados. Esto se puede realizar de modo que se produce una composición biespecífica o multivalente o de vacuna. Se prevé además que los métodos usados en la preparación de estas composiciones serán familiares para los expertos en la técnica y deberían ser adecuados para la administración a animales, es decir, farmacéuticamente aceptables. Agentes preferidos son los vehículos que son Hemocianina de Lapa (KLH, del Inglés “Keyhole Limpet Hemonyannin”) o albúmina de suero bovina (BSA, del Inglés “Bovine Serum Albumin”).

III. Ácidos Nucleicos

La presente invención también proporciona, en otra realización, genes que codifican TS10q23.3. Se han identificado los genes para la molécula TS10q23.3 humana, canina y de murine. Uno de los expertos normales en la técnica podría, usando estos dos ácidos nucleicos, fácilmente identificar homólogos relacionados en diversas otras especies (por ejemplo, rata, conejo, mono, gibón, chimpancé, simio, babuino, vaca, cerdo, caballo, oveja, gato y otras especies). El encuentro de homólogos de ratón y perro para este gen hace probable que otras especies más estrechamente relacionadas con los humanos tendrán, de hecho, también un homólogo.

Además, debería estar claro que la presente invención no está limitada a los ácidos nucleicos específicos descritos en la presente memoria. Tal como se discute a continuación, “un gen de TS10q23.3” puede contener una diversidad de bases diferentes y más aún producir un correspondiente polipéptido que es funcionalmente indistinguible, y en algunos casos estructuralmente, de los genes humanos y de ratón descritos en la presente memoria.

Igualmente, cualquier referencia a un ácido nucleico debería leerse como que abarca una célula hospedante que contiene ese ácido nucleico y, en algunos casos, capaz de expresar el producto de ese ácido nucleico. Además de las

ES 2 361 044 T3

consideraciones terapéuticas, las células que expresan los ácidos nucleicos de la presente invención pueden probar ser útiles en el contexto de la investigación para agentes que inducen, reprimen, inhiben, aumentan, interfieren con, bloquean, derogan, estimulan o mejoran la función de TS10q23.3.

5

A. Ácidos Nucleicos que Codifican 10q23.3

El gen humano descrito en las Figs. 6 y 9, y el gen de murine descrito en la Fig. 9 son genes de TS10q23.3 de la presente invención. Los ácidos nucleicos de acuerdo con la presente invención pueden codificar un gen de TS10q23.3 entero, o cualquier fragmento de la secuencia de TS10q23.3 que tenga función fosfatasa, presentada en la presente memoria. El ácido nucleico puede ser derivado del ADN genómico, es decir, clonado directamente del genoma de un organismo particular. El ácido nucleico puede comprender complementariedad.

ADN (ADNc). También se ha descrito un ADNc más un intrón natural o un intrón derivado de otro gen; tales moléculas producidas por ingeniería algunas veces se refieren como "mini-genes". Como mínimo, estos y otros ácidos nucleicos de la presente invención se pueden usar como patrones de peso molecular en, por ejemplo, electroforesis en gel.

El término "ADNc" intenta referirse a ADN preparado usando ARN mensajero (ARNm) como plantilla. La ventaja de usar un ADNc, a diferencia de ADN genómico o ADN polimerizado a partir de una plantilla de ARN genómico, no o parcialmente procesada, es que el ADNc principalmente contiene secuencias codificadoras de la proteína correspondiente. Puede haber veces en las que se prefiera la secuencia genómica completa o parcial, tal como donde se requieren las regiones no codificadoras para la expresión óptima o donde las regiones no codificadoras tales como intrones son para ser sometidos a direccionamiento en una estrategia antisentido.

También se describe que un TS10q23.3 dado de unas especies dadas se puede representar mediante variantes naturales que tienen secuencias de ácido nucleico algo diferentes pero, no obstante, codifican la misma proteína (véase la Tabla 1 de a continuación).

Tal como se usa en esta solicitud, el término "un ácido nucleico que codifica un TS10q23.3" se refiere a una molécula de ácido nucleico que se ha aislado libre del ácido nucleico celular total. En realizaciones preferidas, la invención concierne a una secuencia de ácido nucleico esencialmente como la expuesta en las Figs. 6 y 9. El término "como la expuesta en la Fig. 6 ó 9" quiere decir que la secuencia de ácidos nucleicos sustancialmente corresponde a una porción de la Fig. 6 ó 9. El término "codón funcionalmente equivalente" se usa en la presente memoria para referirse a los codones que codifican el mismo aminoácido, tal como los seis codones para arginina o serina (Tabla 1, de a continuación), y también se refiere a los codones que codifican aminoácidos biológicamente equivalentes, tal como se discute en las siguientes páginas.

40

TABLA 1

Aminoácidos			Codones
Alanina	Ala	A	GCA GCC GCG GCU
Cisteina	Cys	C	UGC UGU
Ácido Aspártico	Asp	D	GAC GAU
Ácido Glutámico	Glu	E	GAA GAG
Fenilalanina	Phe	F	UUC UUU
Glicina	Gly	G	GGA GGC GGG GGU
Histidina	Hys	H	CAC CAU
Isoleucina	Ile	I	AUA AUC AUU
Lisina	Lys	K	AAA AAG
Leucina	Leu	L	UUA UUG CUA CUC CUG CUU
Metionina	Met	M	AUG
Asparagina	Asn	N	AAC AAU
Prolina	Pro	P	CCA CCC CCG CCU
Glutamina	Gln	Q	CAA CAG
Arginina	Arg	R	AGA AGG CGA CGC CGG CGU
Serina	Ser	S	AGC AGU UCA UCC UCG UCU
Treonina	Thr	T	ACA ACC ACG ACU
Valina	Val	V	GUA GUC GUG GUU
Triptófano	Trp	W	UGG
Tirosina	Tyr	Y	UAC UAU

65

Teniendo en cuenta la degeneración del código genético, las secuencias que tienen al menos aproximadamente 50%, normalmente al menos aproximadamente 60%, más normalmente aproximadamente 70%, lo más normalmente

aproximadamente 80%, preferentemente al menos aproximadamente 90% y los más preferentemente aproximadamente 95% de nucleótidos que son idénticos a los nucleótidos de la Fig. 9 serán secuencias que son “como la expuesta en la Fig. 9”. Secuencias que son esencialmente la misma que las expuestas en la Fig. 9 también se pueden definir funcionalmente como secuencias que son capaces de hibridar con un segmento de ácido nucleico que contiene el complemento de la Fig. 9 bajo condiciones estándar.

Los segmentos de ADN pueden ser aquellos que codifican péptidos y proteínas de TS10q23.3 equivalentes biológicamente funcionales, tal como se ha descrito anteriormente. Tales secuencias pueden surgir como una consecuencia de la redundancia de codón y la equivalencia funcional de los aminoácidos que se sabe que ocurren de manera natural dentro de las secuencias de ácido nucleicos y las proteínas así codificadas. Alternativamente, se puede crear proteínas o péptidos funcionalmente equivalentes vía la aplicación de la tecnología de ADN recombinante, en la cual los cambios en la estructura de la proteína se pueden realizar por ingeniería, basada en consideraciones de las propiedades de los aminoácidos que se intercambian. Los cambios diseñados por el hombre se pueden introducir a través de la aplicación de las técnicas de mutagénesis dirigidas a sitio o se pueden introducir al azar e investigar más tarde para la función deseada, tal como se describe a continuación.

B. Sondas y Cebadores de Oligonucleótidos

Naturalmente, la presente invención también abarca segmentos de ADN que son complementarios, o esencialmente complementarios, a la secuencia expuesta en las Figs. 6 y 9. Las secuencias de ácido nucleico que son “complementarias” a aquellas que son capaces de emparejarse a las bases de acuerdo con las reglas de complementariedad estándar de Watson-Crick. Tal como se usa en la presente memoria, el término “secuencias complementarias” quiere decir secuencias de ácido nucleico que son sustancialmente complementarias, ya que se pueden valorar mediante la misma comparación de nucleótido anteriormente expuesto, o tan definido como ser capaz de hibridar con el segmento de ácido nucleico de las Figs. 6 y 9 bajo condiciones relativamente estrictas tales como las descritas en la presente memoria.

Alternativamente, los segmentos de hibridación pueden ser oligonucleótidos de 10 a 50 bases consecutivas del ácido nucleico de la invención. Las secuencias de 17 bases de longitud deberían ocurrir solamente una vez en el genoma humano y, por lo tanto, ser suficiente para especificar una secuencia diana única. Aunque oligómeros más cortos son más fáciles de realizar e incrementar *in vivo* la accesibilidad, numerosos otros factores están implicados en la determinación de la especificidad de la hibridación. Tanto la afinidad de unión como la especificidad de secuencia de un oligonucleótido a su diana complementaria aumentan con el aumento de la longitud. Se contempla que se usarán los oligonucleótidos ilustrativos de 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45 ó 50. También se describen polinucleótidos más largos que codifican 250, 500, 1.000, 1.212, 1.500, 2.000, 2.500, 3.000 ó 3.431 bases y más largos. Tales oligonucleótidos encontrarán uso, por ejemplo, como sondas en las transferencias Southern y Northern y como cebadores en reacciones de amplificación.

Las condiciones de hibridación adecuadas serán bien conocidas por los expertos en la técnica. En ciertas aplicaciones, por ejemplo, sustitución de aminoácidos mediante mutagénesis dirigida a sitio, se aprecia que se requieren condiciones de inferior rigor. Bajo estas condiciones, la hibridación puede ocurrir aún cuando las secuencias de sonda y la cadena diana no sean perfectamente complementarias, pero estén mal emparejadas en una o más posiciones. Las condiciones pueden resultar menos rigurosas mediante el incremento de la concentración de sal y la disminución de la temperatura. Por ejemplo, una condición de rigor medio se podría proporcionar mediante NaCl aproximadamente 0,1 a 0,25 M a temperaturas de aproximadamente 37°C a aproximadamente 55°C, mientras que una condición de rigor bajo se podría proporcionar mediante sal aproximadamente 0,15 M a aproximadamente 0,9 M, a temperaturas que oscilan entre aproximadamente 20°C a aproximadamente 55°C. Por tanto, las condiciones de hibridación se pueden manipular fácilmente, y así generalmente será un método de elección que depende de los resultados deseados.

En otras realizaciones, se puede alcanzar la hibridación bajo condiciones de, por ejemplo, Tris-HCl 50 mM (pH 8,3), KCl 75 mM, MgCl₂ 3 mM, ditiotreitól 10 mM, a temperaturas entre aproximadamente 20°C a aproximadamente 37°C. Otras condiciones de hibridación utilizadas podrían incluir aproximadamente Tris HCl 10 mM (pH 8,3), KCl 50 mM, MgCl₂ 15 μM, a temperaturas que oscilan desde aproximadamente 40°C a aproximadamente 72°C. También se pueden usar formamida y SDS para alterar las condiciones de hibridación.

Un método de uso de las sondas y cebadores de la presente invención está en la investigación de genes relacionados con TS10q23.3 o, más particularmente, homólogos de TS10q23.3 de otras especies. La existencia de un homólogo de murine fuertemente sugiere que otros homólogos de TS10q23.3 humano se descubrirán en especies más estrechamente relacionadas, y quizás más remotas, que ratón. Normalmente, el ADN diana será una biblioteca genómica o de ADNc, aunque la investigación puede implicar el análisis de moléculas de ARN. Al variar el rigor de la hibridación, y la región de la sonda, se pueden descubrir diferentes niveles de homología.

Otro modo de explotar las sondas y cebadores de la presente invención es en mutagénesis dirigida a sitio o específica de sitio. La mutagénesis específica de sitio es una técnica útil en la preparación de péptidos individuales, o péptidos o proteínas equivalentes biológicamente funcionales, a través de mutagénesis específica del ADN subyacente. La técnica proporciona además una capacidad lista para preparar y ensayar variantes de la secuencia, incorporando una o más de las consideraciones anteriores, mediante la introducción de uno o más cambios de secuencia de nucleótidos

dentro del ADN. La mutagénesis específica a sitio permite la producción de mutantes a través del uso de las secuencias específicas de oligonucleótidos que codifican la secuencia de ADN de la mutación deseada, así como un número suficiente de nucleótidos adyacentes, para proporcionar una secuencia cebador de tamaño suficiente y complejidad de secuencia para formar un dúplex estable en ambos lados de la junta de delección a atravesar. Típicamente, se prefiere un cebador de aproximadamente 17 a 25 nucleótidos de longitud, con aproximadamente 5 a 10 residuos en ambos lados de la junta de la secuencia a alterar.

La técnica típicamente emplea un vector bacteriófago que existe tanto en forma de una cadena sencilla como doble cadena. Típicamente los vectores útiles en la mutagénesis dirigida a sitio incluyen vectores tales como el fago M13. Estos vectores de fago están comercialmente disponibles y su uso generalmente es bien conocido por los expertos en la técnica. También se emplean de forma rutinaria en la mutagénesis dirigida a sitio los plásmidos de doble cadena, lo cual elimina la etapa de transferencia del gen de interés de un fago a un plásmido.

En general, la mutagénesis dirigida a sitio se realiza mediante la primera obtención de un vector de cadena sencilla, o la fusión de dos cadenas de un vector de doble cadena que incluye dentro de su secuencia una secuencia de ADN que codifica la proteína deseada. Sintéticamente se prepara un cebador de oligonucleótido que soporta la secuencia mutada deseada. A continuación, este cebador se alineó con la preparación de ADN de cadena sencilla, tomando en cuenta el nivel de error de emparejamiento cuando se seleccionan las condiciones de hibridación, y se somete a enzimas de polimerización de ADN tal como el fragmento Klenow de polimerasa I de *E. coli*, para completar la síntesis de la cadena que soporta la mutación. Por tanto, se forma un heterodúplex en el que una cadena codifica la secuencia no mutada original y la segunda cadena soporta la mutación deseada. Este vector heterodúplex a continuación se usa para transformar las células apropiadas, tales como las células de *E. coli*, y los clones se seleccionan de modo que incluyan vectores recombinantes que soportan la disposición de la secuencia mutada.

La preparación de las variantes de secuencia del gen seleccionado que usa la mutagénesis dirigida a sitio se proporciona como un medio de producción de especies potencialmente útiles y no se supone que es limitante, ya que hay otras formas en las cuales se puede obtener las variantes de secuencia de los genes. Por ejemplo, se puede tratar los vectores recombinantes que codifican el gen deseado con agentes mutágenos, tal como hidroxilamina, para obtener variantes de secuencia.

C. Construcciones Antisentido

En algunos casos, los supresores de tumor mutantes no pueden ser no-funcionales. Más bien, pueden tener funciones aberrantes que no se pueden superar mediante terapia genética de reemplazamiento, incluso si la molécula “de tipo natural” se expresa en cantidades por encima del polipéptido mutante. El tratamiento antisentido es una forma de dirigir esta situación. La tecnología antisentido también se puede usar para la función “knockout” de TS10q23.3 en el desarrollo de líneas celulares o ratones transgénicos para los propósitos de estudio, diagnóstico e investigación.

La metodología antisentido toma ventaja del hecho de que los ácidos nucleicos tienden a emparejarse con secuencias “complementarias”. Por complementario, se quiere decir polinucleótidos que son capaces de emparejamiento de base de acuerdo con las reglas de complementariedad estándar de Watson-Crick. Es decir, que las purinas más grandes se emparejarán por base con las pirimidinas más pequeñas para formar combinaciones de guanina emparejada con citosina (G:C) y adenina emparejada con o bien timina (A:T) en el caso de ADN, o adenina emparejada con uracilo (A:U) en el caso de ARN. La inclusión de bases menos comunes tales como inosina, 5-metilcitosina, 6-metiladenina, hipoxantina y otras en las secuencias de hibridación no interfiere con el emparejamiento.

El direccionamiento del ADN de doble cadena (dc) con polinucleótidos conduce a la formación de triple hélice; el direccionamiento de ARN conducirá a la formación de doble hélice. Los polinucleótidos antisentido, cuando se introducen en una célula diana, específicamente se unen a su polinucleótido diana e interfiere con la transcripción, procesamiento de ARN, transporte, traducción y/o estabilidad. Las construcciones de ARN antisentido, o ADN que codifica tales ARNs antisentido, se pueden emplear para inhibir la transcripción o traducción del gen o ambas dentro de una célula hospedante, o bien *in vitro* o *in vivo*, tal como dentro de un animal hospedante, incluyendo un sujeto humano.

Las construcciones antisentido pueden estar diseñadas para unirse al promotor y a otras regiones control, exones, intrones o incluso límites exón-intrón de un gen. Se contempla que las construcciones antisentido más eficaces incluirán regiones complementarias a juntas de corte y empalme intrón/exón. Por tanto, se propone que una realización preferida incluye una construcción antisentido con complementariedad a regiones dentro de 50-200 bases de una junta de corte y empalme intrón-exón. Se ha observado que algunas secuencias exón pueden estar incluidas en la construcción sin afectar seriamente la propia selectividad de la diana. La cantidad de material exónico incluido variará dependiendo de las secuencias exón e intrón particulares usadas. Se puede fácilmente ensayar si demasiado ADN exón está incluido simplemente ensayando las construcciones *in vitro* para determinar si la función celular normal está afectada o si la expresión de genes relacionados que tienen secuencias complementarias está afectada.

Tal como se ha planteado anteriormente, “complementario” o “antisentido” quiere decir secuencias de polinucleótidos que son sustancialmente complementarias sobre su longitud entera y tienen muy pocos errores de emparejamiento de base. Por ejemplo, secuencias de quince bases de longitud se pueden calificar complementarias cuando tienen

nucleótidos complementarios en trece o catorce posiciones. Naturalmente, las secuencias que son completamente complementarias serán secuencias que son enteramente complementarias por toda su longitud entera y no tienen malos emparejamientos de base. También se contemplan otras secuencias con niveles inferiores de homología. Por ejemplo, se podría diseñar una construcción antisentido que ha limitado regiones de alta homología, pero también contiene una región no homóloga (por ejemplo, ribozima, véase a continuación). Estas moléculas, aunque teniendo menos del 50% de homología, se unirían a secuencias diana bajo condiciones apropiadas.

Puede ser ventajoso combinar porciones de ADN genómico con ADNc o secuencias sintéticas para generar construcciones específicas. Por ejemplo, si se desea un intrón en la construcción final, se necesitará usar un clon genómico. El ADNc o un polinucleótido sintetizado pueden proporcionar sitios de restricción más conveniente para la porción restante de la construcción, por lo tanto, se usaría para el resto de la secuencia.

D. Ribozimas

Otro enfoque para dirigir el supresor de tumor mutante “dominante negativo” es a través del uso de ribozimas. Aunque tradicionalmente las proteínas se han usado para catálisis de ácidos nucleicos, han surgido otra clase de macromoléculas útiles en este intento. Los ribozimas son complejos ARN-proteína que escinden ácidos nucleicos de un modo específico a sitio. Los ribozimas tienen dominios catalíticos específicos que poseen actividad endonucleasa (Kim y Cook, 1987; Gerlach *et al.*, 1987; Forster y Simons, 1987). Por ejemplo, un gran número de ribozimas aceleran las reacciones de transferencia de fosfoéster con un alto nivel de especificidad, sometiendo a escisión con frecuencia solamente a uno de los varios fosfoésteres en un sustrato de oligonucleótido (Cook *et al.*, 1981; Michael y Westhof, 1990; Reinhold-Hurek y Shub, 1992). Esta especificidad se ha atribuido al requerimiento de que el sustrato se une vía interacciones de emparejamiento de base específicas a la secuencia guía interna (“IGS”, del Inglés “Internal Guide Sequence”) de la ribozima antes de la reacción química.

La catálisis de ribozima principalmente se ha observado como parte de reacciones de escisión/ligación específicas a secuencia que implican los ácidos nucleicos (Joyce, 1989; Cook *et al.*, 1981). Por ejemplo, el documento de Patente U.S. N° 5.354.855 informa que ciertas ribozimas pueden actuar como endonucleasas con una especificidad de secuencia mayor que la de las ribonucleasas conocidas y el enfoque de las enzimas de restricción de ADN. Por tanto, la inhibición mediada por ribozima específica a secuencia de la expresión de gen particularmente puede ser adecuada para las aplicaciones terapéuticas (Scalon *et al.*, 1991; Sarver *et al.*, 1990). Recientemente, se informó que las ribozimas provocaban cambios genéticos en algunas líneas celulares a las cuales se aplicaron; los genes alterados incluían los oncogenes H-ras, c-fos y genes de VIH. La mayoría de este trabajo implicaba la modificación de un ARNm diana, basado en un codón mutante específico que se escinde mediante una ribozima específica.

E. Vectores para la Clonación, Transferencia y Expresión del Gen

Dentro de ciertas realizaciones se emplearon vectores de expresión para expresar el producto de polipéptido TS10q23.3, el cual a continuación se puede purificar y, por ejemplo, usarse para vacunar animales para generar antisueros o anticuerpo monoclonal con el cual se pueden conducir estudios adicionales. En otras realizaciones, los vectores de expresión se pueden usar en terapia de gen. La expresión requiere que se proporcionen señales adecuadas en los vectores, y que incluyan diversos elementos regulatorios, tales como potenciadores/promotores a partir tanto de fuentes víricas como mamíferas que conducen la expresión de los genes de interés en las células hospedantes. También se definen elementos diseñados para optimizar la estabilidad y la capacidad de traducción del ARN mensajero en las células hospedantes. También se proporcionan las condiciones para el uso de un número de marcadores de selección de fármaco dominantes para establecer clones celulares permanentes y estables que expresan los productos, ya que es un elemento que relaciona la expresión de los marcadores de selección de fármaco con la expresión del polipéptido.

(i) Elementos Reguladores

A lo largo de toda esta solicitud, el término “construcción de expresión” se supone que incluye cualquier tipo de construcción genética que contenga un ácido nucleico que codifique un producto de gen en el cual parte o toda la secuencia codificadora de ácido nucleico es capaz de ser transcrita. El transcrito se puede someter a traducción en una proteína, pero no es necesario. En ciertas realizaciones, la expresión incluye tanto transcripción de un gen como traducción de ARNm en un producto de gen. En otras realizaciones, la expresión solamente incluye transcripción del ácido nucleico que codifica un gen de interés.

En realizaciones preferidas, el ácido nucleico que codifica un producto de gen está bajo control transcripcional de un promotor. Un “promotor” se refiere a una secuencia de ADN reconocido por la maquinaria sintética de la célula, o la maquinaria sintética introducida, requerida para iniciar la transcripción específica de un gen. La frase “bajo control transcripcional” quiere decir que el promotor está en la localización y orientación correcta en relación con el ácido nucleico para controlar la iniciación de la ARN polimerasa y la expresión del gen.

ES 2 361 044 T3

El término promotor se usará en la presente memoria para referirse a un grupo de módulos de control transcripcional que se agrupan alrededor del sitio de iniciación para la ARN polimerasa II. La mayoría de las creencias sobre cómo los promotores se organizan deriva de los análisis de diversos promotores víricos, que incluyen aquellos para la timidin quinasa (tk) de HSV y las unidades de transcripción temprana de SV40. Estos estudios, aumentados por trabajos más recientes, han mostrado que los promotores están compuestos de módulos funcionales discretos, consistiendo cada uno de aproximadamente 7-20 pb de ADN, y conteniendo uno o más sitios de reconocimiento para el activador transcripcional o las proteínas represoras.

Al menos un módulo en cada promotor funciona para localizar el sitio de comienzo para la síntesis de ARN. El mejor ejemplo conocido de esto es la caja TATA, pero en algunos promotores que carecen de una caja TATA, tal como el promotor para el gen de deoxinucleotidil transferasa terminal de mamífero y el promotor para los genes tardíos de SV40, un elemento discreto que recubre el sitio de comienzo propio ayuda a fijar el lugar de iniciación.

Elementos promotores adicionales regulan la frecuencia de la iniciación transcripcional. Típicamente, estos están localizados en la región en dirección 5' ("upstream") de 30-110 pb del sitio de comienzo, aunque se ha mostrado también recientemente que un número de promotores contienen elementos funcionales en dirección 3' ("downstream") del sitio de comienzo. El espacio entre los elementos promotor frecuentemente es flexible, de manera que la función del promotor se preserva cuando los elementos se invierten o mueven relativos a otro. En el promotor tk, el espacio entre elementos promotor se puede aumentar en 50 pb antes de que la actividad empiece a declinar. Dependiendo del promotor, parece que los elementos individuales pueden funcionar o bien de manera cooperativa o independiente para activar la transcripción.

No se cree que el promotor particular empleado para controlar la expresión de una secuencia de ácido nucleico de interés sea importante, siempre que sea capaz de dirigir la expresión del ácido nucleico en la célula diana. Por tanto, donde una célula humana es seleccionada como diana, es preferible colocar la región codificadora de ácido nucleico adyacente a y debajo del control de un promotor que sea capaz de ser expresado en una célula humana. Generalmente hablando, tal promotor debería incluir o bien un promotor humano o vírico.

En diversas realizaciones, se pueden usar el promotor de gen temprano inmediato de citomegalovirus (CMV) humano, el promotor temprano de SV40, la repetición terminal larga del virus del sarcoma de Rous, el promotor de insulina de rata y gliceraldehído-3-fosfato deshidrogenasa para obtener expresión de alto nivel de la secuencia codificadora de interés. También se contempla el uso de otros promotores celulares de mamífero o celulares víricos o de fago bacteriano los cuales son bien conocidos en la técnica para alcanzar la expresión de una secuencia codificadora de interés, proporcionó que los niveles de expresión sean suficientes para un propósito dado.

Al emplear un promotor con propiedades bien conocidas, se puede optimizar el nivel y modelo de expresión de la proteína de interés seguido de la transfección o transformación. Además, la selección de un promotor que se regula en respuesta a señales fisiológicas específicas puede permitir la expresión inducible del producto de gen. Las Tablas 2 y 3 enumeran varios elementos/promotores que se pueden emplear, en el contexto de la presente invención, para regular la expresión del gen de interés. Esta lista no intenta ser exhaustiva de todos los posibles elementos implicados en el ascenso de la expresión del gen, pero, simplemente, ser ilustrativa de los mismos.

Los potenciadores son elementos genéticos que incrementan la transcripción a partir de un promotor localizado en una posición distante en la misma molécula de ADN. Los potenciadores se organizan muy parecidos a los promotores. Es decir, están compuestos de muchos elementos individuales, cada uno de los cuales se unen a una o más proteínas transcripcionales.

La distinción básica entre potenciadores y promotores es operacional. Una región potenciadora como un todo debe ser capaz de estimular la transcripción en una distancia, esto no necesita ser cierto para una región promotor o sus elementos componentes. Por otro lado, un promotor debe tener uno o más elementos que dirijan la iniciación de la síntesis de ARN en un sitio particular y en una orientación particular, mientras que los potenciadores carecen de estos sitios específicos. Los promotores y los potenciadores con frecuencia están solapados y son contiguos, pareciendo con frecuencia tener una organización modular muy similar.

A continuación hay una lista de promotores víricos, promotores/potenciadores celulares y promotores/potenciadores inducibles que podrían usarse en combinación con el ácido nucleico que codifica un gen de interés en una construcción de expresión (Tabla 2 y Tabla 3). Adicionalmente, también se podría usar cualquier combinación promotor/potenciador (tal como por la Base de Datos de Promotor Eucariota EPDB, del Inglés "Eukaryotic Promoter Data Base") para conducir la expresión del gen. Las células eucariotas pueden soportar transcripción citoplasmática a partir de ciertos promotores bacterianos si se proporciona la polimerasas bacteriana apropiada, o bien como parte del complejo de liberación o como una construcción de expresión genética adicional.

ES 2 361 044 T3

TABLA 2

POTENCIADOR/PROMOTOR	
5	Cadena Pesada de Inmunoglobulina
	Cadena Ligera de Inmunoglobulina
	Receptor de Célula T
	HLA DQ α y DQ β
	β -Interferón
10	Interleukina-2
	Receptor de Interleukina-2
	MHC Clase II 5
	MHC Clase II HLA-DR α
15	β -Actina
	Quinasa de Creatina de Músculo
	Prealbumina (Transtiretin)
	Elastasa I
	Metalotioneina
20	Colagenasa
	Gen de Albumina
	α -Fetoproteína
	τ -Globina
	β -Globina
25	e-fos
	c-HAS-ras
	Insulina
	Molécula de Adhesión Celular Neural (NCAM, del Inglés "Neural Cell Adhesion Molecule")
30	α 1-Antitripsina
	H2B(THZB) Histona
	Colágeno de Ratón o Tipo I
	Proteínas Reguladas por Glucosa (GRP94 y GRP78)
35	Hormona de Crecimiento de Rata
	Amiloide A de Suero Humano (SAA, del Inglés "Serum Amyloid A")
	Troponina I (TNI)
	Factor de Crecimiento Derivado de Plaqueta
	Distrofia Muscular Duchenne
40	SV40
	Polioma
	Retrovirus
	Virus del Papiloma
	Virus de Hepatitis B
45	Virus de la Inmunodeficiencia Humana
	Citomegalovirus
	Virus de Leucemia de Simio Gibón

TABLA 3

Elemento	Inductor
MT II	Metales pesados de Éster de Forbol
MMTV ("House Mammary Tumor Virus")(virus de tumor mamario de ratón)	Glucocorticoides
13-Interferón	Poli(ri)X Poli(rc)
Adenovirus 5 E2	Ela

ES 2 361 044 T3

TABLA 3 (continuación)

	Elemento	Inductor
5	c-jun	Éster de Forbol (TPA), H ₂ O ₂
	Colagenasa	Éster de Forbol (TPA)
	Estromelisin	Éster de Forbol (TPA), IL-1
	SV40	Éster de Forbol (TPA)
10	Gen MX Murine	Interferón, Virus de Enfermedad de Newcastle
	Gen GRP78	A23187
	α -2-Macroglobulina	IL-6
	Vimeatina	Suero
	Gen H2kB de Clase I MHC	Interferón
15	HSP70	Ela, Antígeno T largo de SV40
	Proliferina	Éster de Forbol-TPA
	Factor de Necrosis Tumoral	FMA
	Gen de la Hormona α Estimuladora de Tiroides	Hormona de Tiroides
20	Caja de Insulina E	Glucosa

Si se emplea un inserto de ADNc, típicamente se deseará incluir una señal de poliadenilación para efectuar la apropiada poliadenilación del transcrito del gen. Se cree que la naturaleza de la señal de poliadenilación es crucial para la exitosa práctica de la invención, y se puede emplear cualquiera de tales secuencias tales como las señales de la hormona de crecimiento humana y de poliadenilación de SV40. También se contempla un terminador como un elemento del casete de expresión. Estos elementos pueden servir para aumentar los niveles de mensaje y para minimizar la lectura a través del casete dentro de otras secuencias.

(ii) Marcadores Seleccionables

En ciertas realizaciones de la invención, las células contienen las construcciones de ácido nucleico de la presente invención, una célula se puede identificar *in vitro* o *in vivo* al incluir un marcador en la construcción de expresión. Tales marcadores conferirían un cambio identificable a la célula que permite la fácil identificación de células que contienen la construcción de expresión. Normalmente la inclusión de un marcador de selección de fármaco ayuda en la clonación y en la selección de transformantes, por ejemplo, los genes que confieren resistencia a neomicina, puomicina, higromicina, DHFR, GPT, zeocina e histidinol son marcadores seleccionables útiles. Alternativamente, se pueden emplear enzimas tales como la timidin quinasa (tk) del virus del herpes simple o la cloramfenicol acetil-transferasa (CAT). También se pueden emplear marcadores inmunológicos. No se cree que el marcador seleccionable empleado sea importante, siempre que sea capaz de ser expresado simultáneamente con el ácido nucleico que codifica un producto de gen. Ejemplos adicionales de marcadores seleccionables son bien conocidos por los expertos en la técnica.

(iii) Construcciones Multigen e IRES

En ciertas realizaciones de la invención, el uso de elementos de sitios de unión a ribosoma internos (IRES, del Inglés "Internal Ribosome Entry Sites") se usa para crear mensajes multigen, o policistrónicos. Los elementos IRES son capaces de eludir el modelo de escaneo de ribosoma de la traducción dependiente de Cap metilado 5' y la traducción del empiece en sitios internos (Pelletier y Sonenberg, 1988). Se han descrito (Pelletier y Sonenberg, 1988) elementos IRES a partir de dos miembros de la familia de los picanovirus (polio y encefalomiocarditis), así como IRES a partir de un mensaje de mamífero (Macejak y Sarnow, 1991). Los elementos IRES se pueden enlazar a marcos de lectura abierta heterólogos. Los marcos de lectura abierta múltiples se pueden transcribir juntos, cada uno separado por un IRES, creando mensajes policistrónicos. En virtud del elemento IRES, cada marco de lectura abierta es accesible a ribosomas para la traducción eficiente. Los genes múltiples se puede expresar eficientemente usando un promotor/potenciador sencillo para transcribir un mensaje sencillo.

Cualquier marco de lectura abierta heterólogo se puede enlazar a los elementos IRES. Esto incluye genes para proteínas secretadas, proteínas multi-subunidad, codificadas por genes independientes, proteínas unidas a membrana o intracelulares y marcadores seleccionables. De esta forma, se puede producir simultáneamente por ingeniería diversas proteínas dentro de una célula con una construcción sencilla y un marcador seleccionable sencillo.

(iv) Liberación de Vectores de Expresión

Hay un número de formas por las cuales se pueden introducir vectores de expresión dentro de las células. En ciertas realizaciones de la invención, la construcción de expresión comprende un virus o una construcción producida por ingeniería derivada de un genoma vírico. La capacidad de ciertos virus para entrar en las células vía endocitosis

mediada por receptor, para integrarse dentro del genoma celular hospedante y expresar genes víricos establemente y eficientemente los han hecho candidatos atractivos para la transferencia de genes extraños dentro de las células mamíferas (Ridgeway, 1988; Nicolas y Rubenstein, 1988; Baichwal y Sugden, 1986; Temin, 1986). Los primeros virus usados como vectores de gen fueron los virus de ADN que incluyen los papovavirus (virus de simio 40, virus del papiloma bovino y poliovirus)(Ridgeway, 1988; Baichwal y Sugden, 1986) y adenovirus (Ridgeway, 1988; Baichwal y Sugden, 1986). Estos tienen una capacidad relativamente baja para las secuencias extrañas de ADN y tienen un espectro hospedante restringido. Además, sus efectos citopáticos y potenciales oncogénicos en las células permisivas elevan los asuntos de seguridad. Pueden acomodar solamente hasta 8 kb de material genético extraño pero se pueden introducir fácilmente en una diversidad de líneas celulares y animales de laboratorio (Nicolas y Rubenstein, 1988; Temin, 1986).

Uno de los métodos preferidos para la liberación *in vivo* implica el uso de un vector de expresión de adenovirus. “Vector de expresión de adenovirus” se supone que incluye aquellas construcciones que contienen secuencias de adenovirus suficientes para (a) soportar el empaquetamiento de la construcción y (b) para expresar un polinucleótido antisentido que se ha clonado en el mismo. En este contexto, la expresión no requiere que se sintetice el producto de gen.

El vector de expresión comprende una forma de adenovirus producida por ingeniería genética. El conocimiento de la organización genética del adenovirus, un virus de ADN de doble cadena, lineal y de 36 kb, permite la sustitución de piezas largas del ADN adenovírico con secuencias extrañas hasta 7 kb (Grunhaus y Horwitz, 1992). A diferencia del retrovirus, la infección adenovírica de las células hospedantes no da como resultado integración cromosómica porque el ADN adenovírico puede replicarse de una manera episomal sin genotoxicidad potencial. Además, los adenovirus son estructuralmente estables, y se ha detectado no reorganización del genoma después de la extensa amplificación. El adenovirus puede infectar prácticamente todas las células epiteliales a pesar de su fase del ciclo celular. Hasta ahora, la infección adenovírica parece estar relacionada solamente con enfermedades leves tales como enfermedad respiratoria aguda en humanos.

El adenovirus es particularmente adecuado para usarse como vector de transferencia de gen debido a su genoma de medio tamaño, facilidad de manipulación, alto título (“titer”), amplio intervalo de célula diana y alta capacidad de infección. Ambos extremos del genoma vírico contienen repeticiones invertidas (ITRs, del Inglés “Inverted Repeats”) de 100-200 pares de bases, las cuales son elementos *cis* necesarios para la replicación y empaquetamiento del ADN. Las regiones tempranas (E, del Inglés “Early”) y tardías (L, del Inglés “Late”) del genoma contienen diferentes unidades de transcripción que se dividen mediante el comienzo de la replicación de ADN vírico. La región E1 (E1A y E1B) codifica proteínas responsables para la regulación de la transcripción del genoma vírico y pocos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la replicación del ADN vírico. Estas proteínas están implicadas en la replicación de ADN, la expresión del gen tardío y cierre de la célula hospedante (Renan, 1990). Los productos de los genes tardíos, que incluyen la mayoría de las proteínas de la cápsida vírica, se expresan solamente después del procesamiento significativo de un transcrito primario único emitido por el promotor tardío principal (MLP, del Inglés “Major Late Promoter”). El MLP, (localizado en 16.8 m.u.) es particularmente eficaz durante la fase tardía de la infección, y todo el ARNm’s emitido a partir de este promotor posee una secuencia líder tripartita 5’ (TPL, del Inglés “Tripartite Leader”) que los convierte en ARNm’s preferidos para la traducción.

En un sistema actual, se genera adenovirus recombinantes a partir de recombinación homóloga entre el vector lanzadera y el vector provirus. Debido a la posible recombinación entre dos vectores províricos, se puede generar el adenovirus de tipo natural a partir de este proceso. Por lo tanto, es crítico aislar un clon sencillo del virus a partir de una placa individual y examinar su estructura genómica.

La generación y propagación de los actuales vectores de adenovirus, los cuales son deficientes en la replicación, depende de una única línea celular ayudante, designada 293, que se transforma a partir de células de riñón embrionarias humanas mediante fragmentos de ADN de Ad5 y expresa constitutivamente proteínas E1 (Graham *et al.*, 1977). Puesto que la región E3 es prescindible a partir del genoma de adenovirus (Jones y Shenk, 1978), los vectores de adenovirus actuales, con la ayuda de las células 293, llevan ADN extraño en cualquiera de las regiones E1, D3 o ambas (Graham y Prevec, 1991). En la naturaleza, el adenovirus puede empaquetarse aproximadamente al 105% del genoma del tipo natural (Ghosh-Choudhury *et al.*, 1987), proporcionando la capacidad para aproximadamente 2 extra kb de ADN combinado con los aproximadamente 5,5 kb de ADN que es reemplazable en las regiones E1 y E3, la capacidad máxima del actual vector de adenovirus está por debajo de 7,5 kb, o aproximadamente 15% de la longitud total del vector. Más del 80% del genoma vírico del adenovirus permanece en la cadena principal del vector y es la fuente de la citotoxicidad aguantada por el vector. Además, la deficiencia de replicación del virus delecionado por E1 es incompleta. Por ejemplo, se ha observado la unión de la expresión del gen vírico con los vectores actualmente disponibles en altas multiplicidades de la infección (MOI, del Inglés “Multiplicities of Infection”)(Mulligan, 1993).

Las líneas celulares ayudantes pueden derivarse de las células humanas tales como células de riñón embrionarias humanas, células musculares, células hematopoyéticas u otras células mesenquimales o epiteliales embrionarias humanas. Alternativamente, las células ayudantes pueden derivarse de las células de otras especies mamíferas que son permisivas para los adenovirus humanos. Tales células incluyen, por ejemplo, células Vero u otras células mesenquimales o epiteliales embrionarias de mono. Tal como se ha planteado anteriormente, la línea de célula ayudante preferida es 293.

Recientemente, Racher *et al.*, (1995) describió métodos mejorados para el cultivo de células 293 y la propagación de adenovirus. En un formato, los agregados celulares naturales se hacen crecer mediante inoculación de células individuales en matraces de centrifugadora siliconizados de 1 litro (Techne, Cambridge, UK) que contienen 100-200 ml de medio. Después de agitación a 40 rpm, la viabilidad de la célula se estima con azul de trypan. En otro formato, se emplea microvehículos de Fibra-Cel (Bibby Sterlin, Stone, UK) (5g/l) tal como sigue. Se añade un inóculo de célula, en resuspensión en 5 ml de medio, al vehículo (50 ml) en un matraz Erlenmeyer de 250 ml y se deja reposar, con agitación ocasional, durante 1 a 4 h. A continuación, se reemplaza el medio con 50 ml de medio fresco y se inicia la agitación. Para la producción de virus, las células se dejan crecer al 80% aproximadamente de confluencia, después de dicho tiempo el medio se reemplaza (al 25% del volumen final) y se añade el adenovirus a un MOI de 0,05. Los cultivos se dejan reposar durante toda la noche, después de lo cual se aumenta el volumen al 100% y se comienza la agitación durante otras 72 h.

Además del requerimiento de que el vector de adenovirus sea defectuoso en replicación, o al menos condicionalmente defectuoso, se cree que la naturaleza del vector de adenovirus es crucial para la práctica exitosa de la invención. El adenovirus puede ser de cualquiera de los 42 serotipos diferentes conocidos o subgrupos A-F. El adenovirus tipo 5 del subgrupo C es el material de comience preferido para obtener el vector de adenovirus defectuoso en replicación condicional para usar en la presente invención. Esto es debido a que el Adenovirus tipo 5 es un adenovirus humano sobre el cual se conoce mucha información bioquímica y genética, e históricamente se ha usado para la mayoría de construcciones que emplean adenovirus como vector.

Tal como se ha planteado anteriormente, el vector típico de acuerdo con la presente invención es defectuoso en replicación y no tendrá una región E1 de adenovirus. Por tanto, lo más conveniente será introducir el polinucleótido que codifica el gen de interés en la posición a partir de la cual se han separado las secuencias que codifican E1. Sin embargo, la posición de la inserción de la construcción dentro de las secuencias de adenovirus no es crítica para la invención. También se puede insertar el polinucleótido que codifica el gen de interés en vez de la región E3 delecionada en vectores de reemplazamiento de E3 como los descritos por Karlsson *et al.*, (1986) o en la región E4 donde una línea celular ayudante o virus ayudante complementa el defecto de E4.

El adenovirus es fácil de hacer crecer y manipular y presenta amplio intervalo hospedante *in vitro* e *in vivo*. Este grupo de virus se puede obtener en altos títulos, por ejemplo, 10^9 - 10^{11} unidades formadoras de placa por ml, y son altamente infecciosos. El ciclo de vida de los adenovirus no requiere integración dentro del genoma de la célula hospedante. Los genes extraños liberados por vectores de adenovirus son episomales y, por lo tanto, tienen baja genotoxicidad para las células hospedantes. Se ha informado de no efectos secundarios en estudios de vacunación con adenovirus tipo natural (Couch *et al.*, 1963; Top *et al.*, 1971), demostrando su seguridad y potencial terapéutico como vectores de transferencia de gen *in vivo*.

Los vectores de adenovirus se han usado en expresión de gen eucariota (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) y desarrollo de vacuna (Grunhaus y Horwitz, 1992; Graham y Prevec, 1992). Recientemente, estudios animales sugieren que los adenovirus recombinantes se podrían usar para terapia de gen (Stratford-Perricaudet y Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Los estudios en la administración de adenovirus recombinante a diferentes tejidos incluyen instilación de tráquea (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), inyección muscular (Ragot *et al.*, 1993), inyecciones intravenosas periféricas (Herz y Gerard, 1993) e inoculación estereotáctica dentro del cerebro (Le Gal La Salle *et al.*, 1993).

Los retrovirus son un grupo de virus de ARN de cadena sencilla caracterizados por una capacidad para convertir su ARN a ADN de doble cadena en células infectadas por un proceso de transcripción inversa (Coffin, 1990). Entonces, el ADN resultante se integra establemente dentro de los cromosomas celulares como un provirus y dirige la síntesis de proteínas víricas. La integración da como resultado la retención de las secuencias del gen vírico en la célula recipiente y sus descendientes. El genoma retrovídico contiene tres genes, gag, pol y env que codifican las proteínas de la cápsida, la enzima polimerasa y los componentes de la envoltura, respectivamente. Una secuencia encontrada en dirección 5' desde el gen gag contiene una señal para empaquetar el genoma dentro de viriones. Dos secuencias de repetición terminal larga (LTR, del Inglés "Long Terminal Repeat") están presentes en los extremos 5' y 3' del genoma vírico. Estas contienen fuertes secuencias promotor y potenciador y también se requieren para la integración en el genoma celular hospedante (Coffin, 1990).

Para construir un vector retrovídico, se inserta un ácido nucleico que codifica un gen de interés dentro del genoma vírico en el lugar de ciertas secuencias víricas para producir un virus que sea defectuoso en replicación. Para producir viriones, se construye una línea celular empaquetadora que contiene los genes gag, pol y env pero sin los componentes de empaquetamiento y LTR (Mann *et al.*, 1983). Cuando se introduce un plásmido recombinante que contiene un ADNc, junto con las secuencias LTR y de empaquetamiento retrovirales en esta línea celular (por ejemplo mediante precipitación de fosfato de calcio), la secuencia de empaquetamiento permite la transcripción de ARN del plásmido recombinante a empaquetar dentro de las partículas víricas, las cuales a continuación se secretan dentro de los medios de cultivo (Nicolas y Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). A continuación, se recoge los medios que contienen los retrovirus recombinantes, opcionalmente se concentran, y se usan para transferir gen. Los vectores retrovídicos son capaces de infectar una amplia diversidad de tipos celulares. Sin embargo, la integración y la expresión estable requieren la división de las células hospedantes (Paskind *et al.*, 1975).

Recientemente se desarrolló un enfoque novedoso diseñado para permitir el direccionamiento específico de los vectores de retrovirus basado en la modificación química de un retrovirus mediante la adición química de residuos de lactosa a la envoltura vírica. Esta modificación podría permitir la infección específica de hepatocitos vía receptores de sialoglicoproteína.

Se diseñó un enfoque diferente para el direccionamiento de retrovirus recombinantes en el cual se usaron anticuerpos biotinilados frente a una proteína de la envoltura retroviral y frente a un receptor celular específico. Los anticuerpos se acoplaron vía los componentes de biotina mediante el uso de estreptavidina (Roux *et al.*, 1989). Usando anticuerpos frente a los muy importantes antígenos de clase I y clase II del complejo de histocompatibilidad, demostraron la infección de una diversidad de células humanas que soportaban aquellos antígenos de superficie con un virus ecotrópico *in vitro* (Roux *et al.*, 1989).

Hay ciertas limitaciones para el uso de vectores de retrovirus en todos los aspectos de la presente invención. Por ejemplo, los vectores de retrovirus normalmente se integran dentro de los sitios al azar en el genoma celular. Esto puede conducir a mutagénesis insercional a través de la interrupción de los genes hospedantes o a través de la inserción de secuencias reguladoras víricas que pueden interferir con la función de los genes franqueantes (Varmus *et al.*, 1981). Otro asunto con el uso de los vectores de retrovirus defectuosos es la apariencia potencial del virus componente de replicación de tipo natural en las células de empaquetamiento. Esto puede dar como resultado los sucesos de recombinación en los cuales la secuencia intacta de los virus recombinantes se inserta en dirección 5' a partir de la secuencia gag, pol, env integrada en el genoma de la célula hospedante. Sin embargo, las nuevas líneas celulares de empaquetamiento en realidad están disponibles de modo que deberían disminuir mayormente la similitud de recombinación (Markowitz *et al.*, 1988; Hersdorffer *et al.*, 1990).

Se pueden emplear otros vectores víricos como construcciones de expresión en la presente invención. Se pueden emplear vectores derivados de virus tales como el virus vaccinia (Ridgeway, 1988; Baichwal y Sugden, 1986; Coupar *et al.*, 1988) virus adenoasociados (AAV, del Inglés "Adeno-associated Virus") (Ridgeway, 1988; Baichwal y Sugden, 1986; Hermonat y Muzycska, 1984) y herpesvirus. Ofrecen varios rasgos atractivos para diversas células mamíferas (Friedmann, 1989; Ridgeway, 1988; Baichwal y Sugden, 1986; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

Con el reciente reconocimiento de los virus defectuosos de la hepatitis B, el nuevo entendimiento se ganó dentro de la relación función-estructura de diferentes secuencias víricas. Estudios *in vitro* mostraron que los virus podrían retener la capacidad para el empaquetamiento dependiente de ayudante y la transcripción inversa a pesar de la delección de hasta 80% de su genoma (Horwich *et al.*, 1990). Esto sugiere que grandes porciones del genoma podrían ser reemplazadas con material genético extraño. El hepatotropismo y la persistencia (integración) eran propiedades particularmente atractivas para la transferencia de gen dirigida a riñón. Chang *et al.* recientemente introdujeron el gen de cloramfenicol acetiltransferasa (CAT) dentro del genoma del virus de hepatitis B de pato en el lugar de las secuencias codificadoras superficiales o presuperficiales de la polimerasa. Se sometieron a cotransfección con virus tipo natural dentro de una línea celular de hepatoma de aves. Se usaron los medios de cultivo que contenían altos títulos de los virus recombinantes para infectar hepatocitos de pato primarios. Se detectó la expresión del gen CAT estable durante al menos 24 días después de la transfección (Chang *et al.*, 1991).

Para efectuar la expresión de las construcciones de gen sentido y antisentido, la construcción de expresión se debe liberar dentro de la célula. Esta liberación se puede conseguir *in vitro*, como en procedimientos de laboratorio para la transformación de líneas celulares, o *in vivo* o *ex vivo*, como en el tratamiento de ciertos estados de enfermedad. Un mecanismo para la liberación es vía infección vírica donde la construcción de expresión es encapsulada en una partícula vírica infecciosa.

Varios métodos no víricos para la transferencia de las construcciones de expresión dentro de las células de mamífero cultivadas también están contemplados por la presente invención. Estos incluyen precipitación con fosfato de calcio (Graham y Van Der Eb, 1973; Chen y Okayama, 1987; Rippe *et al.*, 1990) DEAE-dextrano (Gopal, 1985), electroporación (Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984), microinyección directa (Harland y Weintraub, 1985), liposomas cargados de ADN (Nicolau y Sene, 1982; Fraley *et al.*, 1979) y complejos lipofectamina-ADN, sonicación celular (Fechheimer *et al.*, 1987), bombardeo de gen usando microproyectiles de alta velocidad (Yang *et al.*, 1990) y transfección mediada por receptor (Wu y Wu, 1987; Wu y Wu, 1988). Algunas de las técnicas se pueden adaptar con éxito para usar *in vivo* o *ex vivo*.

Una vez se haya liberado la construcción de expresión dentro de la célula se puede colocar el ácido nucleico que codifica el gen de interés y expresarse en diferentes sitios. En ciertas realizaciones, el ácido nucleico que codifica el gen se puede integrar de manera estable dentro del genoma de la célula. Esta integración puede ser en la localización y orientación cognadas vía recombinación homóloga (reemplazamiento de gen) o se puede integrar en un localización no específica al azar (aumento de gen). En aún más realizaciones el ácido nucleico se puede mantener de manera estable en la célula como un segmento de ADN episomal separado. Tales segmentos de ácido nucleico o "episomas" codifican secuencias suficientes para permitir la permanencia y replicación independiente de o en sincronización con el ciclo celular hospedante. Cómo se libere la construcción de expresión a una célula y dónde se mantenga el ácido nucleico en la célula es dependiente del tipo de construcción de expresión empleada.

En otra realización más de la invención, la construcción de expresión puede simplemente consistir en ADN recombinante desnudo o plásmidos. Se puede realizar la transferencia de la construcción mediante cualquiera de los

métodos anteriormente mencionados los cuales permeabilizan físicamente o químicamente la membrana celular. Esto es particularmente aplicable para transferir *in vitro* pero también se puede aplicar para usar *in vivo*. Dubensky *et al.* (1984) inyectaron con éxito ADN de poliomavirus en la forma de precipitados de fosfato de calcio dentro de hígado y bazo de ratones adultos o recién nacidos demostrando la replicación vírica activa e infección grave. Benvenisty y Neshif (1986) también demostraron que la inyección intraperitoneal directa de los plásmidos precipitados con fosfato de calcio da como resultado la expresión de los genes transfectados. Se prevé que el ADN codificador de un gen de interés también puede ser transferido de una manera similar *in vivo* y expresar el producto de gen.

En aún otra realización de la invención para la transferencia de una construcción de expresión de ADN desnudo dentro de las células puede implicar bombardeo de partícula. Este método depende de la capacidad de acelerar microproyectiles revestidos de ADN a una alta velocidad permitiéndolos perforar las membranas celulares y entrar en las células sin matarlas (Klein *et al.*, 1987). Se han desarrollado varios dispositivos para acelerar pequeñas partículas. Uno de tales dispositivos depende de una descarga a alto voltaje para generar una corriente eléctrica, la cual por turnos proporciona la fuerza motriz (Yang *et al.*, 1990). Los microproyectiles usados han consistido en sustancias biológicamente inertes tales como gotas de tungsteno u oro.

Se han bombardeado *in vivo* órganos seleccionados que incluyen el riñón, la piel y el tejido muscular de ratas y ratones (Yang *et al.*, 1990; Zelenin *et al.*, 1991). Esto puede requerir exposición quirúrgica del tejido o células, para eliminar cualquier tejido intermedio entre el arma y el órgano diana, es decir, tratamiento *ex vivo*. De nuevo, se puede liberar ADN que codifica un gen particular vía este método.

En una realización adicional de la invención, la construcción de expresión puede ser atrapada en un liposoma. Los liposomas son estructuras vesiculares caracterizadas por una membrana bicapa de fosfolípido y un medio acuoso interno. Los liposomas multilamelar tienen capas lipídicas múltiples separadas por medio acuoso. Se forman de manera espontánea cuando los lípidos están suspendidos por encima de la solución acuosa. Los componentes lipídicos experimentan auto reorganización antes de la formación de estructuras cerradas y atrapan agua y solutos disueltos entre las bicapas lipídicas (Ghosh y Bachhawat, 1991). También se contemplan los complejos lipofectamina-ADN.

La liberación de ácido nucleico mediada por liposoma y la expresión de ADN extraño *in vitro* ha sido muy exitosa. Wong *et al.* (1980) demostraron la viabilidad de la liberación mediada por liposoma y la Expresión de ADN extraño en células cultivadas de embrión de pollito, HeLa y hepatoma. Nicolau *et al.* (1987) consiguieron exitosa transferencia de gen mediada por liposoma en ratas después de la inyección intravenosa.

En ciertas realizaciones de la invención, el liposoma puede estar formado en complejo con un virus de hemaglutinación (HVJ). Esto ha demostrado que facilita la fusión con la membrana celular y promueve la entrada celular del ADN encapsulado en liposoma (Kaneda *et al.*, 1989). En otras realizaciones, el liposoma puede estar formado en complejo o emplearse junto con proteínas cromosómicas de no-histidina nuclear (HMG-1)(Kato *et al.*, 1991). En aún más realizaciones, el liposoma se puede formar en complejo o emplear junto con tanto HVJ como HMG-1. En el sentido de que las construcciones de expresión se han empleado con éxito en la transferencia y expresión del ácido nucleico *in vitro* e *in vivo*, entonces son aplicables para la presente invención. Si se emplea un promotor bacteriano en la construcción de ADN, también será deseable incluir dentro del liposoma una polimerasa bacteriana apropiada.

Otros construcciones de expresión que se pueden emplear para liberar un ácido nucleico que codifica un gen particular dentro de las células son los vehículos de liberación mediados por receptor. Estos toman la ventaja de la absorción selectiva de macromoléculas mediante endocitosis mediada por receptor en casi todas las células eucariotas. Debido a la distribución específica a tipo celular de diversos receptores, la liberación puede ser altamente específica (Wu y Wu, 1993).

Los vehículos de direccionamiento de gen mediado por receptor generalmente consisten en dos componentes: un ligando específico a receptor celular y un agente de unión a ADN. Se han usado varios ligandos para la transferencia de gen mediada por receptor. Los ligando más extensamente caracterizados son los asialoorosomucoideos (ASOR)(Wu y Wu, 1987) y la transferrina (Wagner *et al.*, 1990). Recientemente, se ha usado una neoglicoproteína sintética, la cual reconoce el mismo receptor que ASOR, como vehículo de liberación de gen (Ferkol *et al.*, 1993; Perales *et al.*, 1994) y también se ha usado factor de crecimiento epidérmico (EGF, del Inglés "Epidermal Growth Factor") para liberar genes a células de carcinoma escamoso (Myers, EPO 0273085).

En otras realizaciones, el vehículo de liberación puede comprender un ligando y un liposoma. Por ejemplo, Nicolau *et al.* (1987) emplearon lactosil ceramida, un asialgangliosido terminal de galactosa, incorporado dentro de los liposomas y observaron un incremento en la absorción del gen de insulina por los hepatocitos. Por tanto, es viable que un ácido nucleico que codifica un gen particular también pueda ser liberado específicamente dentro de un tipo celular tal como las células pulmonares, epiteliales o tumorales, mediante cualquier número de sistemas receptor-ligando con o sin liposomas. Por ejemplo, se puede usar el factor de crecimiento epidérmico (EGF) como receptor para la liberación mediada de un ácido nucleico que codifica un gen en muchas células tumorales que presentan regulación al alza del receptor de EGF. Se puede usar manosa para identificar el receptor de manosa sobre células de riñón. También, los anticuerpos a CD5 (CLL), CD22 (linfoma), CD25 (leucemia de célula T) y MAA (melanoma) se pueden usar igualmente como restos de direccionamiento.

La transferencia del gen se puede realizar más fácilmente bajo condiciones *ex vivo*. La terapia de gen *ex vivo* se refiere al aislamiento de las células a partir de un animal, la liberación de un ácido nucleico dentro de las células *in vitro*, y a continuación el retorno de las células modificadas de vuelta en un animal. Esto puede implicar la separación quirúrgica de tejidos/órganos de un animal o el cultivo primario de células y tejidos.

Se pueden preparar cultivos celulares de mamífero primarios de diversas formas. Para que las células se mantengan viables mientras *in vitro* y en contacto con la construcción de expresión, es necesario asegurar que las células mantienen contacto con la relación correcta de oxígeno y dióxido de carbono y nutrientes pero están protegidas de la combinación microbiana. Las técnicas de cultivo celular están bien documentadas (Freshner, 1992).

Una realización de lo anterior implica el uso de transferencia de gen para inmortalizar células para la producción de proteínas. Se puede transferir el gen para la proteína de interés tal como se ha descrito anteriormente dentro de las células hospedantes apropiadas seguido de cultivo de células bajo las condiciones apropiadas. De esta manera se puede emplear el gen para prácticamente cualquier polipéptido. La generación de vectores de expresión recombinantes, y los elementos incluidos en la misma, se han discutido anteriormente. Alternativamente, la proteína a producir puede ser una proteína endógena normalmente sintetizada por la célula en cuestión.

Ejemplos de líneas celulares hospedantes de mamífero útiles son las células Vero y HeLa y las líneas celulares de ovario de hámster chino, células W138, BHK, COS-7, 293, HepG2, NIH3T3, RIN y MDCK. Además, se puede elegir una cepa celular hospedante que module la expresión de las secuencias insertadas, o modifique y procese el producto de gen de la manera deseada. Tales modificaciones (por ejemplo, glicosilación) y procesamiento (por ejemplo, escisión) de productos proteicos pueden ser importantes para la función de la proteína. Diferentes células hospedantes tienen mecanismos característicos y específicos para el procesamiento post-traducción y la modificación de las proteínas. Se pueden elegir líneas celulares apropiadas o sistemas hospedantes para asegurar la correcta modificación y procesamiento de la proteína extraña expresada.

Se pueden usar un número de sistemas de selección que incluyen, pero no se limitan a, genes de timidin quinasa de HSV, fosforibosiltransferasa de hipoxantina-guanina y fosforibosiltransferasa, en células *tk-*, *hgprt-* o *apri-*, respectivamente. Además, se puede usar resistencia antimetabolito como la base de selección para *dhfr*, que confiere resistencia a; *gpt*, que confiere resistencia a ácido micofenólico; *neo*, que confiere resistencia al aminoglicosido G418; y *hygro*, que confiere resistencia a higromicina.

Se pueden propagar células animales *in vitro* de dos modos: como células dependientes de no-anclaje que crecen en suspensión por todo el volumen del cultivo o como células dependientes de anclaje que requieren la unión a un sustrato sólido para su propagación (es decir, un tipo monocapa de crecimiento celular).

Los cultivos en suspensión o dependientes de no-anclaje a partir de líneas celulares establecidas continuas son los medios más ampliamente usados de producción a gran escala de células y productos celulares. Sin embargo, las células cultivadas en suspensión tienen limitaciones, tales como el potencial tumorigeno y menor producción de proteína que las células T adherentes.

El cultivo en suspensión a gran escala de células de mamífero en matraces agitados es un método común para la producción de proteínas recombinantes. Dos diseños de reactor de cultivo en suspensión están en amplio uso (el reactor agitado y el reactor de suspensión de aire). El diseño agitado ha sido usado con éxito en una capacidad de 8.000 litros para la producción de interferón. Las células crecen en un tanque de acero inoxidable con una relación altura-diámetro de 1:1 a 3:1. El cultivo normalmente se mezcla con uno o más agitadores, basado en discos con filos o patrones de hélice marina. Se han descrito sistemas de agitador que ofrecen menos fuerzas de rotura que las hojas. La agitación se puede conducir o bien directamente o indirectamente mediante controladores magnéticamente acoplados. Los controladores indirectos reducen el riesgo de contaminación microbiana a través de sellos sobre los mangos del agitador.

El reactor de suspensión de aire, también inicialmente descrito para la fermentación microbiana y adaptado más tarde para el cultivo de mamíferos, depende de una corriente de gas para tanto mezclar como oxigenar el cultivo. La corriente de gas entra en una sección más elevada del reactor y conduce la circulación. El gas se suelta en la superficie de cultivo, causando que el líquido más denso libre de las pompas de gas viaje hacia abajo en la sección de la esquina inferior del reactor. La ventaja principal de este diseño es la simplicidad y la carencia de la necesidad de la mezcla mecánica. Típicamente, la relación altura-diámetro es de 10:1. El reactor de suspensión de aire es relativamente fácil agrandar a escala, tiene buena transferencia de masa de gases y genera relativamente bajas fuerzas de rotura.

Los anticuerpos de la presente invención son particularmente útiles para el aislamiento de antígenos mediante inmunoprecipitación. La inmunoprecipitación implica la separación del componente antígeno diana a partir de una mezcla de complejo, y se usa para discriminar o aislar cantidades mínimas de proteína. Para el aislamiento de las proteínas de membrana las células deben estar solubilizadas dentro de las micelas de detergente. Se prefieren las sales no iónicas, ya que otros agentes tales como las sales biliares, precipitan en pH ácido o en presencia de cationes bivalentes. A continuación, los anticuerpos y sus usos se discuten más.

ES 2 361 044 T3

III. Generación de Anticuerpos Reactivos con TS10q23.3

En otro aspecto, la presente invención contempla un anticuerpo que es inmunoreactivo con una molécula de TS10q23.3 de la presente invención, o cualquier porción de la misma. Un anticuerpo puede ser un anticuerpo policlonal o monoclonal. En una realización preferida, un anticuerpo es un anticuerpo monoclonal. Los medios para preparar y caracterizar los anticuerpos son bien conocidos en la técnica (véase, por ejemplo, Howell y Lane, 1988).

En resumen, se prepara un anticuerpo policlonal mediante inmunización de un animal con un inmunógeno que comprende un polipéptido de la presente invención y recogida de los antisueros de ese animal inmunizado. Se puede usar un amplio intervalo de especies animales para la producción de antisueros. Típicamente un animal usado para la producción de anti-antisueros es un animal no humano que incluye conejos, ratones, ratas, hámsters, cerdos y caballos. Debido al relativamente gran volumen de sangre de los conejos, un conejo es una elección preferida para la producción de anticuerpos policlonales.

Se pueden preparar anticuerpos, tanto policlonales como monoclonales, específicos para las isoformas de antígeno usando técnicas de inmunización convencionales, ya que serán generalmente conocidas por los expertos en la técnica. Se puede usar una composición que contiene epítopos antigénicos de los compuestos de la presente invención para inmunizar uno o más animales experimentales, tales como conejo o ratón, los cuales entonces procederán a producir anticuerpos específicos frente a los compuestos de la presente invención. Se pueden obtener antisueros policlonales, después de dejar el tiempo para la generación de anticuerpo, simplemente haciendo sangrar al animal y preparando muestras de suero a partir de toda la sangre.

Se propone que los anticuerpos monoclonales de la presente invención encontrarán aplicación útil en los procedimientos inmunológicos estándares, tales como los métodos ELISA y transferencia Western y en procedimientos inmunohistoquímicos tales como tinción de tejido, así como en otros procedimientos que pueden utilizar anticuerpos específicos a epítopos de antígeno relacionados con TS10q23.3. Adicionalmente, se propone que se pueden utilizar anticuerpos monoclonales específicos al TS10q23.3 particular de diferentes especies en otras aplicaciones útiles.

En general, se pueden usar anticuerpos tanto policlonales como monoclonales frente a TS10q23.3 en una diversidad de realizaciones. Por ejemplo, se pueden emplear en protocolos de clonación de anticuerpo para obtener ADNc o genes que codifican otro TS10q23.3. También se pueden usar en estudios de inhibición para analizar los efectos de péptidos relacionados con TS10q23.3 en células o animales. Los anticuerpos anti-TS10q23.3 también serán útiles en estudios de inmunolocalización para analizar la distribución de TS10q23.3 durante diversos sucesos celulares, por ejemplo, determinar la distribución celular o específica a tejido de los polipéptidos TS10q23.3 bajo diferentes puntos en el ciclo celular. Una aplicación particularmente útil de tales anticuerpos está en la purificación de TS10q23.3 nativo o recombinante, por ejemplo, usando una columna de afinidad a anticuerpo. La operación de todas estas técnicas inmunológicas será bien conocida por los expertos en la técnica en vista de la presente descripción.

Los medios para preparar y caracterizar anticuerpos son bien conocidos en la técnica (véase, por ejemplo, Harlow y Lane, 1998). En los ejemplos de a continuación, se dan ejemplos más específicos de la preparación de anticuerpo monoclonal.

Tal como se conoce bien en la técnica, una composición dada puede variar en su inmunogenicidad. Por lo tanto, con frecuencia es necesario estimular el sistema inmune hospedante, ya que se puede alcanzar acoplando un inmunógeno de péptido o polipéptido a un vehículo. Vehículos ilustrativos y preferidos son la hemocianina de lapa (KLH) y la albúmina de suero bovina (BSA). También se pueden usar como vehículos otras albúminas tales como ovalbumina, albúmina de suero de ratón o albúmina de suero de conejo. Los medios para conjugar un polipéptido a una proteína vehículo son bien conocidos en la técnica e incluyen glutaraldehído, éster de m-maleimidobencil-N-hidroxisucinimida, carbodiimida y bencidina bis-biazotizada.

Tal como también es conocido en la técnica, se puede aumentar la inmunogenicidad de una composición particular de inmunógeno mediante el uso de estimuladores no-específicos de la respuesta inmune, conocidos como adyuvantes. Adyuvantes ilustrativos y preferidos incluyen adyuvante de Freund completo (un estimulador no específico de la respuesta inmune que contiene *Mycobacterium tuberculosis* muerto), adyuvantes de Freund incompletos y adyuvante de hidróxido de aluminio.

La cantidad de composición de inmunógeno usada en la producción de anticuerpos policlonales varía sobre la naturaleza del inmunógeno así como el animal usado para la inmunización. Se pueden usar una diversidad de rutas para administrar el inmunógeno (subcutánea, intramuscular, intradérmica, intravenosa e intraperitoneal). Se puede hacer un seguimiento de la producción de los anticuerpos policlonales muestreando sangre del animal inmunizado en diversos puntos después de la inmunización. También se puede administrar una segunda inyección, refuerzo. El proceso de estimulación y producción de título ("titering") se repite hasta que se alcanza un título adecuado. Cuando se obtiene un nivel deseado de inmunogenicidad, se puede sangrar el animal inmunizado y aislar el suero y almacenar, y/o el animal se puede usar para generar mAbs.

Los MAb se pueden preparar fácilmente a través del uso de técnicas bien conocidas, tales como las ilustradas en el documento de Patente U.S. 4.196.265. Típicamente, esta técnica implica la inmunización de un animal adecuado con una composición de inmunógeno seleccionada, por ejemplo, una proteína TS10q23.3 purificada o parcialmente

ES 2 361 044 T3

purificada, polipéptido o péptido o célula que expresa alto niveles de TS10q23.3. La composición de inmunización se administra de una manera eficaz para estimular las células productoras de anticuerpo. Roedores tales como ratones y ratas son los animales preferidos, sin embargo, también es posible el uso de células de conejo, oveja, rana. El uso de ratas puede proporcionar ciertas ventajas (Goding, 1986), pero se prefieren los ratones, siendo el ratón BALB/c el más preferido ya que este es el más usado de manera rutinaria y generalmente da un porcentaje mayor de fusiones estables.

Después de la inmunización, se seleccionan células somáticas con el potencial para producir anticuerpos, específicamente B-linfocitos (células B), para usarse en el protocolo de generación de mAb. Estas células se pueden obtener a partir de hígados, amígdalas o ganglios linfáticos sometidos a biopsia, o a partir de una muestra de sangre periférica. Se prefieren las células de bazo y las células sanguíneas periféricas, las primeras porque son una fuente rica de células productoras de anticuerpo que están en la fase plasmoblasto de división, y las últimas porque la sangre periférica es fácilmente accesible. Con frecuencia, se habrá inmunizado un panel de animales y se separará el bazo del animal con los más altos título de anticuerpo y se obtendrá los linfocitos de bazo mediante homogenización del bazo con una jeringuilla. Típicamente, un bazo de un ratón inmunizado contiene aproximadamente de 5×10^7 a 2×10^3 linfocitos.

A continuación, los linfocitos B productores de anticuerpo del animal inmunizado se fusionan con las células de un mieloma celular inmortal, generalmente una de la misma especie que el animal que se inmunizó. Las líneas celulares de mieloma adecuadas para usar en los procedimientos de fusión productora de hibridoma preferentemente son no productoras de anticuerpo, tienen alta eficacia de fusión y deficiencias de enzima que entonces resultan incapaces de crecer en ciertos medios selectivos que soportan el crecimiento de solamente la células fusionadas deseadas (hibridomas).

Se puede usar cualquiera de un número de células de mieloma, ya que son conocidas por los expertos en la técnica (Goding, 1986; Campbell, 1984). Por ejemplo, si el animal inmunizado es un ratón, se pueden usar P3-X63/Ag8, P3-X63-Ag8.653, NSI/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 y S194/5XX0 Bul; para ratas, se pueden usar R210.RCY3, Y3-Ag 1.2.3, IR983F y 4B210; y U-266, GM1500-GRG2, LICR-LON-HMy2 y UC729-6 son todas útiles en la conexión con las fusiones celulares.

Los métodos para generar híbridos de células de ganglio linfático o bazo productoras de anticuerpo y células de mieloma normalmente comprenden mezcla de células somáticas con células de mieloma en una relación 2:1, aunque la relación puede variar desde aproximadamente 20:1 hasta aproximadamente 1:1, respectivamente, en presencia de un agente o agentes (químico o eléctrico) que promueven la fusión de membranas celulares. Se han descrito los métodos de fusión que usan virus Sendai (Kohler y Milstein, 1975; 1976), y aquellos que usan polietilenglicol (PEG), tal como PEG al 37% (v/v), por Geffer *et al.* (1977). El uso de métodos de fusión eléctricamente inducidos también es apropiado (Goding, 1986).

Los procedimientos de fusión normalmente producen híbridos viables a bajas frecuencias, alrededor de 1×10^{-6} a 1×10^{-8} . Sin embargo, esto no supone un problema, ya que los híbridos fusionados viables se diferencian de las células no fusionadas parentales (particularmente las células de mieloma no fusionadas que normalmente continuarían dividiéndose indefinidamente) mediante el cultivo en un medio selectivo. El medio selectivo es generalmente uno que contiene un agente que bloquea la síntesis *de novo* de nucleótidos en los medios de cultivo de tejido. Agentes ilustrativos y preferidos son la aminopterina, el metotrexato y la azaserina. La aminopterina y el metotrexato bloquean la síntesis *de novo* de tanto purinas como pirimidinas, mientras que la azaserina bloquea solamente la síntesis de purina. Si se usa la aminopterina o el metotrexato, los medios se complementan con hipoxantina y timidina como fuente de nucleótidos (medio HAT). Si se usa la azaserina, los medios se complementan con hipoxantina.

El medio de selección preferido es HAT. Solamente las células capaces de manejar rutas de rescate de nucleótido son capaces de sobrevivir en el medio HAT. Las células de mieloma son defectuosas en enzimas clave de la ruta de rescate, por ejemplo, la hipoxantina fosforibosil transferasa (HPRT, del Inglés "Hypoxanthine Phosphoribosyl Transferase"), y no pueden sobrevivir. Las células B pueden manejar esta ruta, pero tienen una duración limitada en el cultivo y generalmente mueren dentro de aproximadamente dos semanas. Por lo tanto, las únicas células que pueden sobrevivir en los medios selectivos son aquellos híbridos formados a partir de mieloma y células B.

Este cultivo proporciona una población de hibridomas a partir de los cuales se seleccionan los hibridomas específicos. Típicamente, la selección de hibridomas se realiza mediante el cultivo de células por dilución de clon único en placas de microtiter, seguido de ensayo de los sobrenadantes clonales individuales (después de aproximadamente dos a tres semanas) para la reactividad deseada. El ensayo debería ser sensible, simple y rápido, tal como los radioinmunoensayos, los inmunoensayos enzimáticos, los ensayos de citotoxicidad, los ensayos de placa, los ensayos de inmunounión en punto y similares.

A continuación, se diluirían en serie los hibridomas seleccionados y se clonarían dentro de las líneas celulares individuales productoras de anticuerpo, tales clones a continuación se pueden propagar indefinidamente para proporcionar mAbs. Las líneas celulares se pueden explotar para la producción de mAb de dos formas básicas. Se puede inyectar una muestra del hibridoma (con frecuencia dentro de la cavidad peritoneal) en un animal histocompatible del tipo que se usa para proporcionar las células somáticas y de mieloma para la fusión original. El animal inyectado desarrolla tumores que secretan el anticuerpo monoclonal específico producido por el híbrido celular fusionado. Los fluidos corporales del animal, tales como suero o fluido de ascitis, a continuación se pueden sacar para proporcionar mAbs en

alta concentración. Las líneas celulares individuales también podrían cultivarse *in vitro*, donde los mAbs se secretan de manera natural en el medio de cultivo a partir del cual se pueden obtener fácilmente en altas concentraciones. Los mAbs producidos por otros modos se pueden purificar más, si se desea, usando filtración, centrifugación y diversos métodos cromatográficos tales como HPLC o cromatografía de afinidad.

5

IV. *Diagnosis de Cánceres que Implican TS10q23.3*

Los presentes inventores han determinado que las alteraciones en TS10q23.3 están asociadas con la malignidad. Por lo tanto, se puede emplear TS10q23.3 y el gen correspondiente como indicador de diagnóstico o de pronóstico de cáncer. Más específicamente, mutaciones puntuales, deleciones, inserciones o perturbaciones reguladoras que se relacionan con TS10q23.3 pueden causar cáncer o promover el desarrollo de cáncer, causar o promover progresión tumoral en un sitio principal, y/o causar o promover metástasis. Otros fenómenos asociados con la malignidad que pueden estar afectados por la expresión de TS10q23.3 incluyen angiogénesis e invasión de tejido.

15

A. *Diagnosis Genética*

Una realización de la presente invención comprende un uso de la región codón de la secuencia de nucleótidos mostrada en la SEC. ID. N°: 2, una variante alélica, un anticuerpo o un oligonucleótido como el definido en las reivindicaciones para determinar la expresión de TS10q23.3. Esto puede comprender la determinación de ese nivel de TS10q23.3. Una realización adicional de la invención comprende un método de diagnóstico de cáncer o ascenso del desarrollo del cáncer o metástasis que comprende la determinación de alteraciones específicas en el producto expresado. Obviamente, este tipo de ensayo tiene importancia en la diagnosis de cánceres relacionados. Dicho cáncer puede implicar cánceres del cerebro (glioblastomas, meduloblastoma, astrocitoma, oligodendroglioma, ependimomas), pulmón, hígado, bazo, riñón, páncreas, intestino delgado, células sanguíneas, ganglio linfático, colon, pecho, endometrio, estómago, próstata, testículo, ovario, piel, cabeza y cuello, esófago, médula ósea, sangre u otro tejido. En particular, la presente invención se refiere a la diagnosis de gliomas.

La muestra biológica puede ser cualquier tejido o fluido. Diversas realizaciones incluyen células de la piel, músculo, fascia, cerebro, próstata, pecho, endometrio, pulmón, cabeza y cuello, páncreas, intestino delgado, células sanguíneas, hígado, testículos, ovarios, colon, piel, estómago, esófago, bazo, ganglio linfático, médula ósea o riñón. Otras realizaciones incluyen muestras de fluido tales como sangre periférica, fluido linfático, ascitis, fluido seroso, efusión pleural, esputo, fluido cerebroespinal, fluido lacrimal, heces u orina.

35

El ácido nucleico usado se aísla de las células contenidas en la muestra biológica, de acuerdo con las metodologías estándar (Sambrook *et al.*, 1989). El ácido nucleico puede ser ADN genómico o ARN celular completo o fraccionado. Si se usa ARN, se puede desear convertir el ARN a un ADN complementario. En una realización, el ARN es ARN celular completo; en otro, es ARN de poli A. Normalmente, se amplifica el ácido nucleico.

40

Dependiendo del formato, se identifica el ácido nucleico de interés en la muestra directamente usando amplificación o con un segundo ácido nucleico conocido después de la amplificación. Luego, se detecta el producto identificado. En ciertas aplicaciones, la detección se puede realizar por medio visuales (por ejemplo, tinción con bromuro de etidio de un gel). Alternativamente, la detección puede implicar identificación indirecta del producto vía quimioluminiscencia, escintigrafía radioactiva o radiomarcaje o marcaje fluorescente o incluso vía un sistema que usa señales de impulsos eléctricos o térmicos (Affymax Technology; Bellus, 1994).

45

Después de la detección, se puede comparar los resultados vistos en un paciente dado con un grupo de referencia estadísticamente significativo de pacientes normales y pacientes que tienen patologías relacionadas con TS10q23.3. De esta forma, es posible correlacionar la cantidad o clase de TS10q23.3 detectado con diversos estados clínicos.

50

Diversos tipos de defectos son para ser identificados. Por tanto, "alteraciones" debería ser leído como que incluyen deleciones, inserciones, mutaciones puntuales y duplicaciones. Las mutaciones puntuales dan como resultado codones de terminación, mutaciones de desplazamiento de la pauta de lectura o sustituciones de aminoácidos. Mutaciones somáticas son aquellas que ocurren en tejidos de no línea germinal. El tejido de línea germinal puede ocurrir en cualquier tejido y se hereda. Las mutaciones dentro y fuera de la región codificadora también pueden afectar a la cantidad de TS10q23.3 producido, ambas por alteración de la transcripción del gen o en la desestabilización o de otra manera alterar el procesamiento de o bien el transcrito (ARNm) o la proteína.

55

Una diversidad de ensayos diferentes están contemplados en este aspecto, que incluyen pero no se limitan a, hibridación fluorescente *in situ* (FISH, del Inglés "Fluorescent *In Situ* Hybridization"), secuenciación directa de ADN, análisis PFGE, transferencia Southern o Northern, análisis de conformación de cadena sencilla (SSCA, del Inglés "Single-Stranded Conformation Analysis"), ensayo de protección de ARNasa, oligonucleótido específico a alelo (ASO, del Inglés "Allele-specific Oligonucleotide"), análisis de transferencia puntual, electroforesis en gel de gradiente de desnaturalización, RFLP y PCR-SSCP.

65

(i) *Cebadores y Sondas*

El término cebador, tal como se define en la presente memoria, supone que abarca cualquier ácido nucleico que sea capaz de cebar la síntesis de un ácido nucleico naciente en un proceso dependiente de plantilla. Típicamente, los cebadores son oligonucleótidos desde diez a veinte pares de bases de longitud, pero se pueden emplear secuencias más largas. Los cebadores pueden ser proporcionados en forma de cadena doble o de cadena sencilla, aunque se prefiere la forma de cadena sencilla. Las sondas se definen de manera diferente, aunque pueden actuar como cebadores. Las sondas, aunque quizás capaces de cebar, están diseñadas para unirse al ADN diana o ARN y no necesitan ser usadas en un proceso de amplificación.

En las realizaciones preferidas, las sondas y los cebadores están marcados con especies radioactivas (^{32}P , ^{14}C , ^{35}S , ^3H u otro marcaje), con un fluoróforo (rodamina, fluoresceína) o un quimioluminiscente (luciferasa).

(ii) *Métodos de Amplificación Dependientes de Plantilla*

Un número de procesos dependientes de plantilla están disponibles para amplificar las secuencias marcadoras presentes en una muestra de plantilla dada. Uno de los mejores métodos de amplificación conocidos es la reacción en cadena de polimerasa (referida como PCRTM, del Inglés “Polymerase Chain Reaction”) la cual se describe en detalle en el documento de Patente U.S. N°s 4.683.195, 4.683.202 y 4.800.159 y en Innis *et al.*, 1990.

En resumen, en la PCR, se preparan dos secuencias cebadoras que son complementarias a las regiones en cadenas complementarias opuestas de la secuencia marcadora. Se añade un exceso de deoxinucleósido trifosfatos a una mezcla de reacción junto con una ADN polimerasa, por ejemplo, *Taq* polimerasa. Si la secuencia marcadora está presente en una muestra, los cebadores se unirán al marcador y la polimerasa causará que los cebadores se extiendan a lo largo de la secuencia marcadora añadiendo nucleótidos. Al elevar o bajar la temperatura de la mezcla de reacción, los cebadores extendidos se disociarán del marcador para formar productos de reacción, los cebadores en exceso se unirán al marcador y a los productos de reacción y el proceso se repetirá.

Se puede realizar un procedimiento de amplificación PCR con transcriptasa inversa para cuantificar la cantidad de ARNm amplificado. Los métodos de ARN de transcripción inversa dentro del ADNc son bien conocidos y descritos en Sambrook *et al.*, 1989. Los métodos alternativos para la transcripción inversa utilizan ADN polimerasas dependientes de ARN termoestables. Estos métodos están descritos en el documento WO 90/07641 presentado el 21 de Diciembre de 1990. Las metodologías de reacción en cadena con polimerasa son bien conocidas en la técnica.

Otro método para la amplificación es la reacción en cadena con ligasa (“LCR”, del Inglés “Ligase Chain Reaction”), descrita en el documento EPO N° 320 308. En la LCR, se preparan dos pares sondas complementarios, y en presencia de la secuencia diana, cada par se unirá a cadenas complementarias opuestas a la diana de modo que estén contiguos. En presencia de una ligasa, los dos pares sondas se enlazarán para formar una unidad única. Al hacer oscilar la temperatura, como en la unión de PCR de las unidades ligadas se disocian de la diana y a continuación sirven como “secuencias diana” para la unión de los pares sondas en exceso. El documento de Patente U.S. 4.883.750 describe un método similar a la LCR para unir pares sondas a una secuencia diana.

La Qbeta Replicasa, descrita en la Solicitud de PCT N° PCT/US87/00880, también se puede usar como otro método más de amplificación en la presente invención. En este método, se añade una secuencia replicativa de ARN que tiene una región complementaria a la de una diana a una muestra en presencia de una ARN polimerasa. La polimerasa copiará la secuencia replicativa que entonces se puede detectar.

También puede ser útil un método de amplificación isotérmica, en el cual se usan las endonucleasas de restricción y las ligasas para alcanzar la amplificación de las moléculas diana que contienen nucleótido 5'-[alfa-tio]-trifosfatos en una cadena de un sitio de restricción en la amplificación de los ácidos nucleicos en la presente invención, Walker *et al.* (1992).

La Amplificación con Desplazamiento de Cadena (SDA, del Inglés “Strand Displacement Amplification”) es otro método para llevar a cabo la amplificación isotérmica de los ácidos nucleicos que implica rondas múltiples de desplazamiento y síntesis de cadena, es decir, traducción del “nick”. Un método similar, denominado Reacción en Cadena de Reparación (RCR, del Inglés “Repair Chain Reaction”), implica alineación de varias sondas por toda una región diana para la amplificación, seguido de una reacción de reparación en la cual solamente están presentes dos de las cuatro bases. Se pueden añadir las otras dos bases como derivados biotinilados para la detección fácil. Un enfoque similar se usa en SDA. Las secuencias específicas a diana también se pueden detectar usando una reacción de sonda cíclica (CPR, del Inglés “Cyclic Probe Reaction”). En la CPR, se hibrida una sonda que tiene secuencias 3' y 5' de ADN no específico y una media secuencia del ARN específico con ADN que está presente en una muestra. Tras la hibridación, la reacción se trata con ARNasa H y los productos de la sonda identificados como productos característicos que se liberan después de la digestión. La plantilla original se alineó a otra sonda cíclica y se repite la reacción.

Se pueden usar otros métodos de amplificación más descritos en la Solicitud GB N° 2.202.328, y en la solicitud PCT N° PCT/US89/01025. En la primera solicitud, se usan cebadores “modificados” en una síntesis dependiente de enzima y plantilla, similar a PCR. Los cebadores se pueden modificar marcando con un resto de captura (por ejemplo,

biotina) y/o un resto detector (por ejemplo, enzima). En la última solicitud, se añaden un exceso de sondas marcadas a una muestra. En presencia de la secuencia diana, la sonda se une y se escinde catalíticamente. Después de la escisión, se libera la secuencia diana intacta para unirse por sonda en exceso. La escisión de la sonda marcada señala la presencia de la secuencia diana.

Otros procedimientos de amplificación de ácido nucleico incluyen sistemas de amplificación basados en transcripción (TAS, del Inglés "Transcription-based Amplification Systems"), que incluyen amplificación basada en secuencia de ácido nucleico (NASBA, del Inglés "Nucleic Acid Sequence Based Amplification") y 3SR (Kwoh *et al.*, 1989; Gingeras *et al.*, Solicitud PCT WO 88/10315).

En NASBA, los ácidos nucleicos se pueden preparar para la amplificación mediante extracción estándar por fenol/cloroformo, desnaturalización por calor de una muestra clínica, tratamiento con tampón lisis y columnas de minipin para el aislamiento de ADN y ARN o extracción con cloruro de guanidino de ARN. Estas técnicas de amplificación implican la alineación de un cebador que tiene secuencias específicas a diana. Después de la polimerización, se digieren los híbridos de ADN/ARN con ARNasa H mientras que las moléculas de ADN de doble cadena se desnaturalizan de nuevo por calor. En cualquier caso el ADN de cadena sencilla está hecho completamente de doble cadena mediante la adición de segundo cebador específico a diana, seguido de polimerización. A continuación, las moléculas de ADN de doble cadena se transcriben de manera múltiple mediante una ARN polimerasa tal como T7 o SP6. En una reacción cíclica isotérmica, los ARN son transcritos a la inversa en ADN de cadena sencilla, el cual a continuación se convierte en ADN de doble cadena, y a continuación transcritos una vez de nuevo con una ARN polimerasa tal como T7 o SP6. Los productos resultantes, truncados o completos, indican las secuencias específicas a diana.

Davey *et al.*, el documento EPO N° 239.822 describe un proceso de amplificación de ácido nucleico que implica sintetizar cíclicamente ARN de cadena sencilla ("ARNcs"), ADNcs y ADN de doble cadena (ADNdc), los cuales se pueden usar de acuerdo con la presente invención. El ARNcs es una plantilla para un primer oligonucleótido cebador, el cual se alarga mediante transcriptasa inversa (ADN polimerasa dependiente de ARN). A continuación, se separa el ARN del dúplex ADN:ARN resultante mediante la acción de la ribonucleasa H (ARNasa H, una ARNasa específica para el ARN en el dúplex con o bien ADN o ARN). El ADNcs resultante es una plantilla para un segundo cebador, el cual también incluye las secuencias de un promotor de ARN polimerasa (ilustrado por la ARN polimerasa T7) 5' para su homología a la plantilla. A continuación, se extiende este cebador mediante la ADN polimerasa (ilustrado por el fragmento largo "Klenow" de la ADN polimerasa I de *E. coli*), dando como resultado una molécula de ADN de doble cadena ("ADNdc"), que tiene una secuencia idéntica a la del ARN original entre los cebadores y que tiene adicionalmente, en un extremo, una secuencia promotor. Esta secuencia promotor se puede usar mediante la ARN polimerasa apropiada para realizar muchas copias ARN del ADN. Estas copias entonces pueden entrar de nuevo en el ciclo que conduce a una amplificación muy rápida. Con la elección apropiada de las enzimas, esta amplificación se puede hacer isotérmicamente sin adición de las enzimas en cada ciclo. Debido a la naturaleza cíclica de este proceso, la secuencia de comienzo puede ser elegida para estar en forma de o bien ADN o ARN.

Miller *et al.*, Solicitud PCT WO 89/06700 describen un esquema de amplificación de secuencia de ácido nucleico basado en la hibridación de una secuencia promotor/cebador con un ADN de cadena sencilla diana ("ADNcs") seguido de transcripción de muchas copias de ARN de la secuencia. Este esquema no es cíclico, es decir, no se producen nuevas plantillas a partir de los transcritos de ARN resultantes. Otros métodos de amplificación incluyen "RACE" y "PCR de un lado" (Frohman, M.A., En: PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Academic Press, N.Y., 1990; Ohara *et al.*, 1989).

También se pueden usar en la etapa de amplificación métodos basados en la unión de dos (o más) oligonucleótidos en presencia del ácido nucleico que tiene la secuencia del "di-oligonucleótido" resultante, amplificando de ese modo el di-oligonucleótido. (Wu *et al.*, (1989)).

(iii) *Transferencia Southern/Northern*

Las técnicas de transferencia son bien conocidas por los expertos en la técnica. La transferencia Southern implica el uso de ADN como diana, mientras que la transferencia Northern implica el uso de ARN como diana. Cada uno proporciona diferentes tipos de información, aunque la transferencia de ADN es análoga, en muchos aspectos, a transferencia de especies de ARN.

En resumen, una sonda se usa para identificar una especie de ADN o ARN que se ha inmovilizado en una matriz adecuada, con frecuencia un filtro de nitrocelulosa. Las diferentes especies deberían estar espacialmente separadas para facilitar el análisis. Esto con frecuencia se consigue mediante electroforesis en gel de especies de ácido nucleico seguido de "transferencia" sobre el filtro.

Posteriormente, la diana sometida a transferencia se incuba con una sonda (normalmente marcada) bajo condiciones que promueven la desnaturalización y la rehibridación. Debido a que la sonda está diseñada para emparejar de base con la diana, la sonda unirá una porción de la secuencia diana bajo condiciones de renaturalización. A continuación, se separa la sonda no unida, y la detección se consigue tal como se ha descrito anteriormente.

(iv) Métodos de Separación

Normalmente es deseable, en una fase u otra, separar el producto de amplificación de la plantilla y el cebador en exceso con el propósito de determinar si ha ocurrido la amplificación específica. En una realización, los productos de amplificación son separados mediante electroforesis en gel de agarosa, agarosa-acrilamida o poliacrilamida usando métodos estándar. Véase Sambrook *et al.*, 1989.

Alternativamente, se pueden emplear las técnicas cromatográficas para efectuar la separación. Hay muchas clases de cromatografía que se pueden usar en la presente invención: absorción, partición, intercambio iónico y criba molecular, y muchas técnicas especializadas para usarlas que incluyen cromatografía en columna, en papel, de capa fina y de gas (Freifelder, 1982).

(v) Métodos de Detección

Los productos se pueden visualizar para confirmar la amplificación de las secuencias marcadoras. Un método de visualización típico implica la tinción de un gel con bromuro de etidio y la visualización bajo luz UV. Alternativamente, si los productos de amplificación están íntegramente marcados con nucleótidos radio o fluorométricamente marcados, entonces los productos de amplificación se pueden exponer a película de rayos X o visualizar bajo el apropiado espectro estimulador, después de la separación.

En una realización, la visualización se alcanza indirectamente. Después de la separación de los productos de amplificación, se pone en contacto una sonda de ácido nucleico marcada con la secuencia marcadora amplificada. La sonda preferentemente se conjuga con el cromóforo pero se puede radiomarcarse. En otra realización, la sonda se conjuga a una pareja de unión, tal como un anticuerpo o biotina, y el otro miembro del par de unión lleva un resto detectable.

En una realización, la detección es mediante una sonda marcada. Las técnicas implicadas son bien conocidas por los expertos en la técnica y se pueden encontrar en muchos libros estándares o protocolos moleculares. Véase Sambrook *et al.*, 1989. Por ejemplo, el cromóforo o sondas radiomarcadas o cebadores identifican la diana durante o después de la amplificación.

Un ejemplo de lo anterior está descrito en el documento de Patente U.S. Nº 5.219.721, el cual describe un aparato y método para la electroforesis automatizada y la transferencia de los ácidos nucleicos. El aparato permite la electroforesis y la transferencia sin manipulación externa del gel y es idealmente adecuado para llevar a cabo los métodos de acuerdo con la presente invención.

Además, los productos de amplificación anteriormente descritos pueden someterse a análisis de secuencia para identificar clases específicas de variaciones usando técnicas de análisis de secuencia estándar. Dentro de ciertos métodos, se lleva a cabo análisis exhaustivos de genes mediante análisis de secuencia usando series cebadores diseñados para la óptima secuenciación (Pignon *et al.*, 1994). La presente invención proporciona métodos por los cuales se pueden usar cualquier o todos estos tipos de análisis. Usando las secuencias descritas en la presente memoria, los cebadores de oligonucleótido pueden estar diseñados para permitir la amplificación de las secuencias por todo el gen de TS10q23.3 que a continuación se pueden analizar mediante secuenciación directa.

(vi) Componentes del kit

Todos los materiales esenciales y reactivos requeridos para detectar y secuenciar el TS10q23.3 y variantes del mismo se pueden reunir en un kit. Esto generalmente comprenderá los cebadores y las sondas preseleccionadas. También incluidas pueden estar las enzimas adecuadas para la amplificación de ácidos nucleicos que incluyen diversas polimerasas (RT, Taq, Sequenase™ etc.), deoxinucleótidos y tampones para proporcionar la mezcla de reacción necesaria para la amplificación. Dichos kits también generalmente comprenderán, en medios adecuados, distintos recipientes para cada reactivo y enzima individual así como para cada cebador o sonda.

(vii) Diseño y Consideraciones Teóricas para RT-PCR Cuantitativa Relativa

La transcripción inversa (RT) de ARN o ADNc seguida de PCR cuantitativa relativa (RT-PCR) se puede usar para determinar las concentraciones relativas de las especies de ARNm específicas aisladas de los pacientes. Al determinar que la concentración de una especie de ARNm específica varía, se demuestra que el gen codificador de las especies de ARNm específico se expresa de manera diferente.

En la PCR, el número de moléculas del ADN diana amplificado aumenta mediante un factor que aproxima dos con todo el ciclo de la reacción hasta que algún reactivo llega a ser limitante. A partir de entonces, el índice de amplificación llega a estar cada vez más disminuido hasta que hay no incremento en la diana amplificada entre ciclos. Si se traza una gráfica en la cual el número de ciclo está sobre el eje X y el logaritmo de la concentración del ADN diana amplificado está sobre el eje Y, se forma una línea curvada de forma característica al conectar los puntos trazados. Empezando

con el primer ciclo, la pendiente de la línea es positiva y constante. Esto se dice que es la porción lineal de la curva. Después de que un reactivo llega a ser limitante, la pendiente de la línea empieza a disminuir y finalmente llega a ser cero. En este punto la concentración del ADN diana amplificado llega a ser asintótico a algún valor fijado. Esto se dice que es la porción meseta de la curva.

5

La concentración del ADN diana en la porción lineal de la amplificación por PCR es directamente proporcional a la concentración inicial de la diana antes de empezar la reacción. Al determinar la concentración de los productos amplificados del ADN diana en las reacciones PCR que han completado el mismo número de ciclos y están en sus intervalos lineales, es posible determinar las concentraciones relativas de la secuencia diana específica en la mezcla de ADN original. Si las mezclas de ADN son ADNc sintetizados a partir de ARN aislados de diferentes tejidos o células, se pueden determinar las abundancias relativas del ARNm específico a partir del cual se deriva la secuencia diana para los respectivos tejidos o células. Esto proporcionalmente directo entre la concentración de los productos de PCR y las abundancias relativas de ARNm es únicamente cierto en el intervalo lineal de la reacción de PCR.

10

La concentración final del ADN diana en la porción meseta de la curva se determina mediante la disponibilidad de los reactivos en la mezcla de reacción y es independiente de la concentración original del ADN diana. Por lo tanto, la primera condición que debe encontrarse antes de que las abundancias relativas de una especie de ARNm se puedan determinar mediante RT-PCR para una colección de poblaciones de ARN es que las concentraciones de los productos de PCR amplificados se deben muestrear cuando las reacciones de PCR están en la porción lineal de sus curvas.

15

La segunda condición que debe encontrarse para un experimento de RT-PCR para determinar con éxito las abundancias relativas de una especie de ARNm particular es que las concentraciones relativas de los ADNc amplificados deben ser normalizados a algún patrón independiente. El objetivo de un experimento de RT-PCR es determinar la abundancia de una especie de ARNm particular relativa a la abundancia promedio de todas las especies de ARNm en la muestra. En los experimentos descritos a continuación, se usaron los ARNm para β -actina, asparagina sintetasa y lipocortina II como patrones externos e internos a los cuales se compara la abundancia relativa de otros ARNm.

20

La mayoría de los protocolos para la PCR competitiva utilizan los patrones de PCR internos que son aproximadamente tan abundantes como la diana. Estas estrategias son eficaces si los productos de las amplificaciones de PCR se muestrean durante sus fases lineales. Si los productos se muestrean cuando las reacciones están aproximándose a la fase meseta, entonces el producto menos abundante llega a ser relativamente sobre representado. Las comparaciones de las abundancias relativas se dirigen a muchas muestras de ARN diferentes, tales como es el caso cuando al examinar las muestras de ARN para la expresión diferencial, se llegan a deformar de tal forma que hace que las diferencias en las abundancias relativas de los ARN parezcan menos de lo que son en realidad. Esto no es un problema significativo si el patrón interno es mucho más abundante que la diana. Si el patrón interno es mucho más abundante que la diana, entonces se puede realizar comparaciones lineales entre las muestras de ARN.

25

La anterior discusión describe consideraciones teóricas para un ensayo de RT-PCR para materiales clínicamente derivados. Los problemas inherentes en las muestras clínicas son que son de cantidad variable (haciendo la normalización problemática), y que son de calidad variable (necesitando la amplificación conjunta de un control interno fiable, preferentemente de mayor tamaño que la diana). Ambos problemas se superan si la RT-PCR se hace como una RT-PCR cuantitativa relativa con un patrón interno en el cual el patrón interno es un fragmento de ADNc amplificable que es mayor que el fragmento de ADNc diana y en el cual la abundancia del ARNm codificador del patrón interno es aproximadamente 5-100 veces mayor que el ARNm codificador de la diana. Este ensayo mide la abundancia relativa, no la abundancia absoluta de la respectiva especie de ARNm.

30

Otros estudios se pueden hacer usando un ensayo de RT-PCR cuantitativa relativa convencional con un protocolo estándar externo. Estos ensayos muestrean los productos PCR en la porción lineal de sus curvas de amplificación. El número de ciclos PCR que son óptimos para muestrear debe ser empíricamente determinado para cada fragmento de ADNc diana. Además, los productos de transcriptasa inversa de cada población de ARN aislada de las diversas muestras de tejido deben ser cuidadosamente normalizadas para concentraciones iguales de ADNc amplificables. Esta consideración es muy importante puesto que el ensayo mide la abundancia absoluta de ARNm. La abundancia absoluta de ARNm se puede usar como una medida de expresión de gen diferencial solamente en muestras normalizadas. Si la determinación empírica del intervalo lineal de la curva de amplificación y la normalización de las preparaciones de ADNc es tediosa y son procesos que consumen tiempo, los ensayos de RT-PCR resultantes pueden ser superiores a los derivados del ensayo de RT-PCR cuantitativa relativa con un patrón interno.

35

Una razón para esta ventaja es que sin el patrón/competidor interno, todos los reactivos se pueden convertir en un producto de PCR sencillo en el intervalo lineal de la curva de amplificación, incrementando así la sensibilidad del ensayo. Otra razón es que con solamente un producto de PCR, la exposición del producto sobre un gel electroforético u otro método de exposición llega a ser menos complejo, tiene menos antecedentes y es más fácil de interpretar.

40

(viii) *Tecnologías de Chip*

45

Específicamente contemplada por los presentes inventores son las tecnologías de ADN basadas en chip tales como las descritas por Hacia *et al.* (1996) y Shoemaker *et al.* (1996). En resumen, estas tecnologías implican métodos cuantitativos para analizar grandes números de genes rápidamente y exactamente. Al etiquetar genes con oligonucleó-

tidos o usar las formaciones de sonda fijada, se puede emplear tecnología de chip para segregar moléculas diana como formaciones de alta densidad e investigar estas moléculas sobre la base de la hibridación. Véase también Pease *et al.* (1994); Fodor *et al.* (1991).

5

B. *Inmunodiagnosis*

Los anticuerpos de la presente invención se pueden usar en la caracterización del contenido de TS10q23.3 de tejidos sanos y enfermos, a través de técnicas tales como ELISA y transferencia Western. Esto puede proporcionar una investigación de presencia o ausencia de malignidad o como un indicador de futuro cáncer.

El uso de los anticuerpos de la presente invención, se contempla en un ensayo ELISA. Por ejemplo, se inmovilizan anticuerpos anti-TS10q23.3 sobre una superficie seleccionada, preferentemente una superficie que presenta una afinidad de proteína tal como los pocillos de una placa microtiter de poliestireno. Después de lavar para separar el material incompletamente adsorbido, es deseable unir o revestir los pocillos de la placa de ensayo con una proteína no específica que se sabe que es antigénicamente neutral en relación con los antisueros de ensayo tal como albúmina de suero bovina (BSA), caseína o disoluciones de leche en polvo. Esto permite el bloqueo de sitios de adsorción no específicos sobre la superficie de inmovilización y así reduce los antecedentes causados por la unión no específica de antígeno sobre la superficie.

Después de la unión del anticuerpo al pocillo, el revestimiento con un material no reactivo para reducir los antecedentes, y el lavado para separar el material no unido, se conecta la superficie de inmovilización con la muestra a ensayar de una manera propicia a la formación de complejo inmune (antígeno/anticuerpo).

Después de la formación de inmunocomplejos específicos entre la muestra de ensayo y el anticuerpo unido, y el posterior lavado, se puede determinar la incidencia e incluso la cantidad de formación de inmunocomplejo sometiendo la misma a un segundo anticuerpo que tiene especificidad para TS10q23.3 que difiere del primer anticuerpo. Las condiciones apropiadas preferentemente incluyen dilución de la muestra con diluyentes tales como BSA, gama globulina bovina (BGG, del Inglés "Bovine Gamma Globulin") y disolución salina tamponada con fosfato (PBS)/Tween[®]. Estos agentes añadidos también tienden a asistir en la reducción de antecedentes no específicos. A continuación, se deja incubar los antisueros capeados durante aproximadamente 2 a aproximadamente 4 h., a temperaturas preferentemente alrededor de aproximadamente 25° a aproximadamente 27°C. Después de la incubación, se lava la superficie conectada a antisuero para separar el material no formado en inmunocomplejo. Un procedimiento de lavado preferido incluye lavado con una disolución tal como PBS/Tween[®] o tampón de borato.

Para proporcionar un medio de detección, el segundo anticuerpo preferentemente tendrá una enzima asociada que generará un desarrollo de color tras la incubación con un sustrato cromogénico apropiado. Por tanto, por ejemplo, se deseará poner en contacto e incubar la superficie unida al segundo anticuerpo con un IgG antihumano conjugado con peroxidasa o ureasa durante un periodo de tiempo y bajo condiciones que favorecen el desarrollo de la formación del inmunocomplejo (por ejemplo, incubación durante 2 h a temperatura ambiente en una disolución que contiene PBS tal como PBS/Tween[®]).

Después de la incubación con el segundo anticuerpo etiquetado con enzima, y posterior al lavado para separar el material no unido, se cuantifica la cantidad de marcaje mediante incubación con un sustrato cromogénico tal como urea y púrpura de bromocresol o el ácido 2,2'-azino-di-(3-etil-benzotiazolin)-6-sulfónico (ABTS) y H₂O₂, en el caso de peroxidasa como marcaje de enzima. A continuación, se alcanza la cuantificación mediante la medición del nivel de generación de color, por ejemplo, usando un espectrofotómetro de espectro visible.

El formato precedente se puede alterar al unir primero la muestra a la placa de ensayo. A continuación, el anticuerpo primario se incuba con la placa de ensayo, seguido de la detección del anticuerpo primario unido usando un segundo anticuerpo marcado con especificidad para el anticuerpo primario.

Las composiciones de anticuerpo de la presente invención encontrará gran uso en los análisis de inmunotransferencia o transferencia Western. Los anticuerpos se pueden usar como reactivos primarios de alta afinidad para la identificación de proteínas inmovilizadas sobre una matriz soporte sólida, tal como nitrocelulosa, nylon o combinaciones de los mismos. Junto con la inmunoprecipitación, seguido de electroforesis en gel, estos se pueden usar como un reactivo de etapa única para usar en la detección de antígenos frente a los cuales los reactivos secundarios usados en la detección del antígeno causan un antecedente adverso. Los métodos de detección inmunológicamente basados para usar junto con transferencia Western incluyen anticuerpos secundarios enzimáticamente-, radiomarcado o fluorescentemente etiquetados frente al resto de toxina se consideraron que eran de uso particular en este aspecto.

V. *Métodos para la Investigación de Compuestos Activos*

También se describe el uso de TS10q23.3 y fragmentos activos, y ácidos nucleicos que codifican los mismos, en la investigación de compuestos para la actividad en o bien la estimulación de la actividad de TS10q23.3, la superación de la carencia de TS10q23.3 o el bloqueo del efecto de una molécula TS10q23.3 mutante. Estos ensayos pueden producir el uso de una diversidad de diferentes formatos y pueden depender de la clase de "actividad" para la cual se está

ES 2 361 044 T3

conduciendo la investigación. Las “lecturas de salida” funcionales contempladas incluyen la unión a un compuesto, la inhibición de la unión a un sustrato, ligando, receptor u otra pareja de unión mediante un compuesto, actividad fosfatasa, actividad anti-fosfatasa, fosforilación de TS10q23.3, desfosforilación de TS10q23.3, inhibición o estimulación de una señalización célula-célula, crecimiento, metástasis, división celular, migración celular, formación de colonia en agar blando, inhibición por contacto, capacidad de invasión, angiogénesis, apoptosis, progresión tumoral u otro fenotipo maligno.

A. *Ensayos In Vitro*

Se describe la investigación de los compuestos que se unen a la molécula de TS10q23.3 o fragmento de los mismos. El polipéptido o fragmento puede ser cualquiera libre en disolución, fijado a un soporte, expresado en o sobre la superficie de una célula. O bien el polipéptido o el compuesto pueden estar marcados, permitiendo de ese modo determinar la unión.

El ensayo puede medir la inhibición de la unión de TS10q23.3 a un sustrato natural o artificial o pareja de unión. Se pueden realizar ensayos de unión competitiva en los cuales se marque uno de los agentes (TS10q23.3, pareja de unión o compuesto). Normalmente, el polipéptido será la especie marcada. Se puede medir la cantidad de marcaje libre frente al marcaje unido para determinar la unión o inhibición de la unión.

En el documento WO 84/03564 se describe otra técnica para la investigación de alto rendimiento de los documentos. Se sintetizan grandes números de compuestos de ensayo de pequeño péptido sobre un sustrato sólido, tal como alfileres de plástico u alguna otra superficie. Se hacen reaccionar los compuestos de ensayo de péptido con TS10q23.3 y se lavan. El polipéptido unido es detectado mediante diversos métodos.

El TS10q23.3 purificado puede ser revestido directamente sobre placas para usarse en las anteriormente mencionadas técnicas de investigación de fármaco. Sin embargo, se pueden usar los anticuerpos no neutralizantes al polipéptido para inmovilizar el polipéptido a una fase sólida. Además, se pueden usar las proteínas de fusión que contienen una región reactiva (preferentemente una región terminal) para enlazar la región activa de TS10q23.3 a una fase sólida.

Diversas líneas celulares que contienen mutaciones de tipo natural o naturales o producidas por ingeniería en TS10q23.3 se pueden usar para estudiar diversos atributos funcionales de TS10q23.3 y cómo un compuesto candidato afecta a estos atributos. Métodos para las mutaciones producidas por ingeniería están descritos en otra parte de este documento, ya que son mutaciones que ocurren de manera natural en TS10q23.3 que llevan a contribuir y/o por otra parte causar malignidad. En tales ensayos, el compuesto se formularía apropiadamente, se daría su naturaleza bioquímica, y se pondría en contacto con una célula diana. Dependiendo del ensayo, se puede requerir cultivo. A continuación, se puede examinar la célula en virtud de un número de diferentes ensayos fisiológicos. Alternativamente, se puede realizar el análisis molecular en el cual se puede explorar la función de TS10q23.3, o rutas relacionadas. Esto puede implicar ensayos tales como aquellos para la expresión de proteína, función enzimática, utilización de sustrato, estados de fosforilación de diversas moléculas incluyendo TS10q23.3, niveles de cAMP, expresión de ARNm (incluyendo exposición diferencial de toda la célula o ARN poliA) y otros.

B. *Ensayos In Vivo*

También se describe el uso de diversos modelos animales. En la presente memoria, la igualdad vista entre TS10q23.3 humano y de ratón proporciona una excelente oportunidad para examinar la función de TS10q23.3 en un sistema animal completo donde normalmente se expresa. Al desarrollar o aislar líneas celulares mutantes que fallan para expresar TS10q23.3 normal, se puede generar modelos de cáncer en ratones que serán altamente indicadores de cánceres en humanos y otros mamíferos. Estos modelos pueden emplear la administración ortotópica o sistemática de células tumorales para cánceres mímicos primarios y/o metastático. Alternativamente, se puede inducir cánceres en animales al proporcionar agentes conocidos por ser responsables de ciertos sucesos asociados con la transformación maligna y/o progresión tumoral. Finalmente, los animales transgénicos (discutidos a continuación) que carecen de un TS10q23.3 tipo natural se pueden utilizar como modelos para el desarrollo y tratamiento del cáncer.

El tratamiento de animales con compuestos de ensayo implicará la administración del compuesto, en una forma apropiada, al animal. La administración será mediante cualquier ruta que se pudiera utilizar para propósitos clínicos o no-clínicos, incluyendo pero no limitado a oral, nasal, bucal, rectal, vaginal o local. Alternativamente, la administración puede ser por instilación intratraqueal, instilación bronquial, inyección intradérmica, subcutánea, intramuscular, intraperitoneal o intravenosa. Específicamente contemplados están la inyección intravenosa sistemática, la administración regional vía suministro sanguíneo o linfático y la inyección intratumoral.

La determinación de la eficacia de un compuesto *in vivo* puede implicar una diversidad de diferentes criterios. Tales criterios incluyen, pero no se limitan a, supervivencia, reducción de la carga o masa tumoral, la detención o la desaceleración de la progresión tumoral, la eliminación de tumores, la inhibición de la prevención de metástasis, el nivel de actividad incrementado, el mejoramiento en la función efectora inmune y la mejorada toma de comida.

C. Diseño Racional de Fármaco

El objetivo del diseño racional del fármaco es producir análogos estructurales de polipéptidos biológicamente activos o compuestos con los cuales interactúan (agonistas, antagonistas, inhibidores, parejas de unión, etc.). Al crear tales análogos, es posible crear fármacos que son más activos o estables que las moléculas naturales, las cuales tienen diferente susceptibilidad a la alteración o las cuales puede afectar a la función de otras diversas moléculas. En un enfoque, se generaría una estructura tridimensional para TS10q23.3 o un fragmento del mismo. Esto se podría conseguir mediante cristalografía por rayos X, modelización por ordenador o mediante una combinación de ambos enfoques. Un enfoque alternativo, “escaneado con alanina”, implica el reemplazamiento al azar de residuos por toda la molécula con alanina, y el resultante afecta a la función determinada.

También es posible aislar un anticuerpo específico a TS10q23.3, seleccionado mediante un ensayo funcional, y a continuación solucionar su estructura de cristal. En principio, este enfoque produce un farmacóforo sobre el cual se puede basar el posterior diseño de fármaco. Es posible evitar la cristalografía proteica toda junta mediante la generación de anticuerpos anti-idiotípicos a un anticuerpo farmacológicamente activo y funcional. Como una imagen espejo de una imagen espejo, el sitio de unión del anti-idiotipo se esperaría que fuera un análogo del antígeno original. A continuación, el anti-idiotipo se podría usar para identificar y aislar los péptidos de los bancos de péptidos producidos químicamente o biológicamente. A continuación, los péptidos seleccionados servirían como el farmacóforo. Los anti-idiotipos se pueden generar usando los métodos descritos en la presente memoria para la producción de anticuerpos, usando un anticuerpo como antígeno.

Por tanto, se puede diseñar fármacos que tengan actividad TS10q23.3 mejorada o que actúen como estimuladores, inhibidores, agonistas, antagonistas o TS10q23.3 o moléculas afectadas por la función de TS10q23.3. En virtud de la disponibilidad de las secuencias de TS10q23.3 clonadas, se pueden producir cantidades suficientes de TS10q23.3 para realizar estudios cristalográficos. Además, el conocimiento de las secuencias de polipéptidos permite las predicciones empleadas por ordenador de las relaciones estructura-función.

VI. Enfermedades Relacionadas con Tratamiento de 10q23.3

La presente invención se puede usar en el tratamiento de cáncer. Los tipos de cáncer que se pueden tratar, de acuerdo con la presente invención, están limitados únicamente por la participación de TS10q23.3. Por la participación, incluso no es un requerimiento que TS10q23.3 sea mutado o anormal (la sobreexpresión de este supresor de tumor en realidad puede superar otras lesiones dentro de la célula). Por tanto, se contempla que una amplia diversidad de tumores puedan ser tratados usando terapia de TS10q23.3, incluyendo cánceres del cerebro (glioblastoma, astrocitoma, oligodendroglioma, ependimomas), pulmón, hígado, bazo, riñón, ganglio linfático, páncreas, intestino delgado, células sanguíneas, colon, estómago, pecho, endometrio, próstata, testículo, ovario, piel, cabeza y cuello, esófago, médula ósea, sangre u otro tejido.

En muchos contextos, no es necesario que se mate la célula tumoral o que sea inducida a experimentar muerte celular normal o “apoptosis”. Además, para conseguir un tratamiento significativo, todo eso requiere que el crecimiento tumoral se ralentice a algún nivel. Puede ser que se bloquee completamente el crecimiento tumoral, sin embargo, o que se alcance alguna regresión tumoral. La terminología clínica tal como carga de “remisión” y “reducción de tumor” también se contempla dado su uso normal.

A. Terapias Basadas en la Genética

Una de las realizaciones terapéuticas contempladas por los presentes inventores es la intervención, al nivel molecular, en los sucesos implicados en la tumorigénesis de algunos cánceres. Específicamente, los presentes inventores intentan proporcionar, a una célula cancerosa, una construcción de expresión capaz de proporcionar TS10q23.3 a esa célula. Debido a que los genes humanos, de ratón y perro todos codifican el mismo polipéptido, cualquiera de estos ácidos nucleicos se podría usar en terapia humana, como podría cualquiera de las variantes de secuencia de gen anteriormente discutidas las cuales codificarían el mismo, o un polipéptido biológicamente equivalente. La prolongada discusión de los vectores de expresión y los elementos genéticos empleados en la presente invención está incorporada dentro de esta sección por referencia. Los vectores de expresión particularmente preferidos son los vectores víricos tales como los adenovirus, virus adeno asociados, herpesvirus, virus vaccinia y retrovirus. También preferido es el vector de expresión liposomalmente encapsulado.

Aquellos expertos en la técnica son conscientes de cómo aplicar la liberación de gen a situaciones *in vivo* y *ex vivo*. Para los vectores víricos, generalmente se preparará un banco de vector vírico. Dependiendo de la clase de virus y el título alcanzable, se liberaría 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} partículas infecciosas al paciente. Figuras similares pueden ser extrapoladas para formulaciones liposomales u otras no víricas mediante la comparación de las eficiencias de absorción relativa. A continuación, se discute la formulación como una composición farmacéuticamente aceptable.

Se contemplan diversas rutas para diversos tipos tumorales. La sección de a continuación contiene una extensa lista de posibles rutas. Para prácticamente cualquier tumor, se contempla la liberación sistemática. Esto se probará

especialmente importante para el ataque del cáncer microscópico o metastático. Si se puede identificar la masa tumoral discreta, se pueden tomar una diversidad de enfoques directos, locales y regionales. Por ejemplo, el tumor puede ser directamente inyectado con el vector de expresión. Se puede tratar un lecho tumoral antes de, durante o después de la resección. Después de la resección, generalmente se liberaría el vector mediante un catéter izquierdo en lugar de la siguiente cirugía. Se puede utilizar la basculación del tumor para introducir el vector dentro del tumor al inyectar una vena o arteria de soporte. También se puede utilizar una ruta de suministro sanguíneo distal.

También se describe la terapia de gen *ex vivo*. Este enfoque es particularmente adecuado, aunque no limitado, al tratamiento de cánceres asociados a médula ósea. En una realización *ex vivo*, las células del paciente se separan y se mantienen fuera del cuerpo durante al menos algún periodo de tiempo. Durante este periodo, se libera una terapia, después de la cual se reintroducen las células dentro del paciente; con optimismo, se han matado cualquiera de las células tumorales en la muestra.

El trasplante de médula ósea autóloga (ABMT, del Inglés “Autologous Bone Marrow Transplant”) es un ejemplo de terapia de gen *ex vivo*. Básicamente, la noción bajo el ABMT es que el paciente servirá como su propio donador de médula ósea. Por tanto, se puede liberar una dosis normalmente letal de irradiación o quimioterapia al paciente para matar las células tumorales, y repoblar la médula ósea con las propias células del paciente que se han mantenido (y quizás expandido) *ex vivo*. Debido a que, la médula ósea con frecuencia se contamina con células tumorales, es deseable purgar la médula ósea de estas células. El uso de la terapia de gen para conseguir este objetivo es incluso otra forma más en que TS10q23.3 se puede utilizar de acuerdo con la presente invención.

B. Inmunoterapias

Las inmunoterapéuticas, generalmente, depende del uso de las células efectoras inmunes y las moléculas para identificar y destruir las células cancerosas. El efector inmune puede ser, por ejemplo, un anticuerpo específico para algún marcador sobre la superficie de una célula tumoral. El anticuerpo solo puede servir como efector de la terapia o puede reclutar otras células para efectuar en realidad la matanza celular. El anticuerpo también se puede conjugar a un fármaco o toxina (quimioterapéutica, radionúclido, cadena de ricina A, toxina del cólera, toxina de pertussis, etc.) y sirven simplemente como un agente de direccionamiento. Alternativamente, el efector puede ser un linfocito que lleva una molécula de superficie que interactúa, o bien directamente o indirectamente, con una diana celular de tumor. Diversas células efectoras incluyen células T citotóxicas y células NK.

Es improbable que el TS10q23.3 de la presente invención pudiera servir como una diana para un efector inmune dado que (i) es improbable que se exprese sobre la superficie de la célula y (ii) que la presencia, no ausencia, de TS10q23.3 está asociada con el estado normal. Sin embargo, es posible que formas mutantes particulares de TS10q23.3 se puedan identificar mediante inmunoterapia, o bien usando anticuerpos, conjugados de anticuerpo o células efectoras inmunes.

Un escenario más probable es que la inmunoterapia se pudiera usar como parte de una terapia combinada, junto con terapia de gen diana a TS10q23.3. A continuación, se discute el enfoque general para la terapia combinada. Generalmente, la célula tumoral debe soportar algún marcador que es susceptible a direccionamiento, es decir, no está presente en la mayoría de las otras células. Muchos marcadores tumorales existen y algunos de estos pueden ser adecuados para direccionamiento en el contexto de la presente invención. Los marcadores tumorales comunes incluyen antígeno carcinoembrionario, antígeno específico de próstata, antígeno asociado a tumor urinario, antígeno fetal, tirosinasa (p97), gp68, TAG-72, HMFG, Antígeno Sialyl Lewis, MucA, MucB, PLAP, receptor de estrógeno, receptor de laminina, *erb B* y p155.

C. Terapia de Proteína

Otro enfoque de terapia es la provisión, a un sujeto, de polipéptido TS10q23.3, fragmentos activos, péptidos sintéticos, miméticos u otros análogos del mismo. La proteína se puede producir por medios de expresión recombinante o, si es demasiado pequeña, generar mediante un sintetizador de péptido automatizado. Las formulaciones se seleccionarían en base a la ruta de administración y al propósito que incluyen, pero no se limitan a, formulaciones liposomales y preparaciones farmacéuticas clásicas.

D. Terapia Combinada con Inmunoterapia, Quimio o Radioterapia Tradicional

La resistencia celular tumoral a agentes que dañan el ADN representa un problema muy importante en la oncología clínica. Un objetivo del actual estudio del cáncer es encontrar las formas para mejorar la eficacia de la quimio y radioterapia. Una forma es mediante la combinación de dichas terapias tradicionales con la terapia de gen. Por ejemplo, el gen de la timidin quinasa del herpes simple (*HS-tk*), cuando se liberó a tumores cerebrales mediante un sistema de vector retroviral, con éxito indujeron susceptiblemente al agente ganciclovir antiviral (Culver *et al.*, 1992). En el contexto de la presente invención, se contempla que la terapia de reemplazamiento de TS10q23.3 se podría usar igualmente junto con la intervención quimio o radioterapéutica. También puede demostrar eficacia para combinar la terapia de gen de TS10q23.3 con inmunoterapia, tal como se ha descrito anteriormente.

ES 2 361 044 T3

Para matar células, inhibir el crecimiento celular, inhibir la metástasis, inhibir la angiogénesis o de otra manera invertir o reducir el fenotipo maligno de las células tumorales, usando los métodos y composiciones de la presente invención, generalmente se pondría en contacto una célula “diana”(“target”) con una construcción de expresión de TS10q23.3 y al menos otro agente. Estas composiciones serían proporcionadas en una cantidad combinada eficaz para matar o inhibir la proliferación de la célula. Este proceso puede implicar poner en contacto las células con la construcción de expresión y el(los) agente(s) o factor(es) al mismo tiempo. Esto se puede alcanzar poniendo en contacto la célula con una composición sencilla o formulación farmacológica que incluye ambos agentes, o poniendo en contacto la célula con dos composiciones o formulaciones distintas, al mismo tiempo, en el que una composición incluye la construcción de expresión y la otra incluye el agente.

Alternativamente, el tratamiento con terapia de gen puede preceder o seguir al otro tratamiento con agente mediante intervalos que oscilan desde minutos a semanas. Si el otro agente y construcción de expresión se aplican separadamente a la célula, generalmente se aseguraría que un significativo periodo de tiempo no expiraría entre el tiempo de cada liberación, de modo que el agente y la construcción de expresión serían incluso capaces de ejercer un efecto ventajosamente combinado sobre la célula. En tales ejemplos, se contempla que se pondría en contacto la célula con ambas modalidades dentro de aproximadamente 12-24 horas de cada uno y, más preferentemente, dentro de aproximadamente 6-12 horas de cada uno, con un tiempo de demora de solamente aproximadamente 12 horas que es lo más preferido. En algunas situaciones, puede ser deseable extender el periodo de tiempo para el tratamiento de manera significativa, sin embargo, donde varios días (2, 3, 4, 5, 6 ó 7) a varias semanas (1, 2, 3, 4, 5, 6, 7 ó 8) transcurren entre las respectivas administraciones.

También es concebible que será deseable más de una administración de o bien TS10q23.3 o el otro agente. Se pueden emplear diversas combinaciones, donde TS10q23.3 es “A” y el otro agente es “B”, tal como se ilustra a continuación:

A/B/A B/A/B B/B/A A/A/B B/A/A A/B/B B/B/B/A B/B/A/B

A/A/B/B A/B/A/B A/B/B/A B/B/A/A B/A/B/A B/A/A/B B/B/B/A

A/A/AB B/A/A/A A/B/A/A A/AB/A A/B/B/B B/A/B/B B/B/A/B

Se contempla otras combinaciones. De nuevo, para alcanzar la muerte celular, se liberan ambos agentes a una célula en una cantidad combinada eficaz para matar la célula.

Los agentes o factores adecuados para usar en una terapia combinada son cualquier compuesto químico o método de tratamiento que induce daño de ADN donde se suministra a una célula. Tales agentes y factores incluyen radiación y ondas que inducen daño de ADN tales como irradiación γ , rayos X, irradiación UV, microondas, emisiones electrónicas y similares. Una diversidad de compuestos químicos, también descritos como “agentes quimioterapéuticos”, funcionan para inducir daño de ADN, todos los cuales se intenta que sean de uso en los métodos de tratamiento combinado descritos en la presente memoria. Los agentes quimioterapéuticos contemplados para ser de uso, incluyen, por ejemplo, adriamicina, 5-fluorouracilo (5FU), etopósido (VP-16), camptotecina, actinomicina-D, mitomicina C, asplatina (CDDP) e incluso peróxido de hidrógeno. La invención también abarca el uso de una combinación de uno o más agentes que dañan el ADN, compuestos actuales o basados en radiación, tales como el uso de rayos X con cisplatina o el uso de cisplatina con etopósido. En ciertas realizaciones, particularmente se prefiere el uso de cisplatina en combinación con una construcción de expresión de TS10q23.3 como este compuesto.

En el tratamiento de cáncer de acuerdo con la invención, se pondrían en contacto las células tumorales con un agente además de la construcción de expresión. Esto se puede alcanzar irradiando el sitio de tumor localizado con radiación tal como rayos X, luz UV, rayos γ o incluso microondas. Alternativamente, las células tumorales se pueden poner en contacto con el agente mediante la administración al sujeto de una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un compuesto tal como, adriamicina, 5-fluorouracilo, etopósido, camptotecina, actinomicina D, mitomicina C o, más preferentemente, cisplatina. El agente se puede preparar y usar como una composición terapéutica combinada, o kit, combinándolo con una construcción de expresión de TS10q23.3, tal como se ha descrito anteriormente.

Agentes que directamente entrecruzan ácidos nucleicos, específicamente ADN, se conciben para facilitar el daño de ADN que conduce a una combinación sinérgica, antineoplásica con TS10q23.3. Se pueden usar agentes tales como cisplatina, y otros agentes alquilantes de ADN. La cisplatina ha sido ampliamente usada para tratar cáncer, con dosis eficaces usadas en aplicaciones clínicas de 20 mg/m² durante 5 días cada tres semanas durante un total de tres cursos. La cisplatina no se absorbe oralmente por lo tanto se debe liberar vía inyección intravenosamente, subcutáneamente, intratumoralmente o intraperitonealmente.

Los agentes que dañan ADN también incluyen compuestos que interfieren con la replicación de ADN, mitosis y segregación cromosómica. Tales compuestos quimioterapéuticos incluyen adriamicina, también conocida como doxo-

rubicina, etopósido, verapamila, podofilotoxina y similares. Ampliamente usado en un entorno clínico para el tratamiento de neoplasmos, estos compuestos se administran a través de inyecciones de bolos intravenosamente a dosis que oscilan entre 25-75 mg/m² en intervalos de 21 días para adriamicina, a 35-50 mg/m² para etopósido intravenosamente o el doble de la dosis intravenosa oralmente.

5 Los agentes que interrumpen la síntesis y fidelidad de los precursores de ácidos nucleicos y subunidades también conducen al daño de ADN. Se han desarrollado como tales un número de precursores de ácido nucleico. Particularmente útil son los agentes que han experimentado ensayo exhaustivo y están fácilmente disponibles. Como tales, agentes tales como 5-fluorouracilo (5-FU), preferentemente se usan mediante tejido neoplásico, haciendo a este agente particularmente útil para el direccionamiento a células neoplásicas. Aunque bastante tóxico, 5-FU, es aplicable en un amplio rango de vehículos, incluyendo local, sin embargo siendo comúnmente usada la administración intravenosa con dosis que oscilan entre 3 y 15 mg/kg/día.

15 Otros factores que causan daño de ADN y se han usado extensivamente incluyen los comúnmente conocidos como rayos γ , rayos X y/o la liberación dirigida de radioisótopos a células tumorales. Otras formas de factores que dañan ADN también se contemplan tal como microondas e irradiación UV. Lo más probable es que todos estos factores efectúan un amplio intervalo de daño de ADN, sobre los precursores de ADN, la replicación y reparación de ADN, y el ensamblaje y mantenimiento de cromosomas. Los intervalos de dosificación para el intervalo de rayos X desde las dosis diarias de 50 a 200 roentgens durante periodos de tiempo prolongados (3 a 4 semanas), a dosis únicas de 2.000 a 20 6.000 roentgens. Los intervalos de dosificación para radioisótopos varían ampliamente, y dependen de la vida útil del isótopo, la fuerza y tipo de radiación emitida y la absorción mediante las células neoplásicas.

25 El experto en la técnica se dirige a "Remington's Pharmaceutical Sciences" 15ª Edición, capítulo 33, en particular páginas 624-652. Alguna variación en la dosificación ocurrirá necesariamente dependiendo de la condición del sujeto a tratar. La persona responsable de la administración, en un suceso, determinará la dosis apropiada para el sujeto individual. Además, para la administración humana, las preparaciones deberían encontrar patrones de esterilidad, pirogenicidad, seguridad general y pureza tal como lo requerido por la Oficina de FDA de Patrones Biológicos.

30 Los inventores proponen que la liberación regional de las construcciones de expresión de TS10q23.3 a pacientes con cánceres relacionados con 10q23.3 será un método muy eficiente para la liberación de un gen terapéuticamente eficaz para contrarrestar la enfermedad clínica. Igualmente, la quimio o radioterapia puede estar dirigida a una región afectada particular del cuerpo de los sujetos. Alternativamente, la liberación sistemática de la construcción de expresión y/o el agente puede ser apropiada en ciertas circunstancias, por ejemplo, donde ha ocurrido la metástasis extensiva.

35 Además de la combinación de terapias de direccionamiento a TS10q23.3 con quimio y radioterapias, también se contempla que la combinación con otras terapias de gen será ventajosa. Por ejemplo, el direccionamiento de TS10q23.3 y mutaciones p53 o p16 al mismo tiempo pueden producir un tratamiento anticáncer mejorado. Cualquier otro gen relacionado con tumor concebiblemente puede estar sometido a direccionamiento de esta manera, por ejemplo, p21, 40 Rb, APC, DCC, NF-1, NF-2, BCRA2, p16, FHIT, WT-1, MEN-I, MEN-II, BRCA1, VHL, FCC, MCC, *ras*, *myc*, *neu*, *raf*, *erb*, *src*, *fms*, *jun*, *trk*, *ret*, *gsp*, *hst*, *bcl* y *abl*.

45 También se debería indicar que cualquiera de las terapias anteriores puede probar ser útil por ella misma en el tratamiento de un TS10q23.3. En este aspecto, la referencia a compuestos quimioterapéuticos y terapia de gen de no TS10q23.3 en combinación también debería leerse como una reflexión de que estos enfoques se pueden emplear separadamente.

50 E. Formulaciones y Rutas para la Administración a Pacientes

Si se contemplan aplicaciones clínicas, será necesario preparar composiciones farmacéuticas (vectores de expresión, reservas de virus, proteínas, anticuerpos y fármacos) en una forma apropiada para la aplicación deseada. Generalmente, esto implicará la preparación de composiciones que están esencialmente libres de pirógenos, así como otras impurezas que podrían ser perjudiciales para humanos o animales.

55 Generalmente se deseará emplear sales y tampones apropiados para hacer los vectores de liberación estables y permitir la absorción mediante células diana. Los tampones también se emplearán cuando se introduzcan las células recombinantes dentro de un paciente. Las composiciones acuosas de la presente invención comprenden una cantidad eficaz del vector a células, disueltas o dispersas en un vehículo farmacéuticamente aceptable o medio acuoso. Tales composiciones también se refieren como inóculos. La frase "farmacéuticamente o farmacológicamente aceptable" se refiere a entidades moleculares y composiciones que no producen reacciones adversas, alérgicas u otras perjudiciales cuando se administra a un animal o un humano. Tal como se ha usado en la presente memoria, "vehículo farmacéuticamente aceptable" incluye alguno y todos los disolventes, medios de dispersión, revestimientos, agentes antibacterianos y antifúngicos, agentes que retrasan la absorción y similares. El uso de tales medios y agentes para las sustancias farmacéuticamente activas es bien conocido en la técnica. Excepto en el grado en el que cualquier medio convencional o agente es incompatible con los vectores o células de la presente invención, se contempla su uso en las composiciones terapéuticas. Los ingredientes activos complementarios también se pueden incorporar en las composiciones.

ES 2 361 044 T3

Las composiciones activas de la presente invención pueden incluir preparaciones farmacéuticas clásicas. La administración de estas composiciones de acuerdo con la presente invención será vía cualquier ruta común siempre que el tejido diana esté disponible vía esa ruta. Esto incluye oral, nasal, bucal, rectal, vaginal o local. Alternativamente, la administración puede ser mediante inyección ortotópica, intradérmica, subcutánea, intramuscular, intraperitoneal o intravenosa. Tales composiciones normalmente se administrarían como composiciones farmacéuticamente aceptables, descritas anteriormente.

Los compuestos activos también se pueden administrar parenteralmente o intraperitonealmente. Las disoluciones de los compuestos activos como base libre o sales farmacológicamente aceptables se pueden preparar en agua adecuadamente mezclada con un tensioactivo, tal como hidroxipropilcelulosa. Las dispersiones también se pueden preparar en glicerol, polietilenglicoles líquidos y mezclas de los mismos y en aceites. Bajo condiciones normales de almacenamiento y uso, estas preparaciones contienen un conservante para prevenir el crecimiento de los microorganismos.

Las formas farmacéuticas adecuadas para el uso inyectable incluyen disoluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de las disoluciones o dispersiones inyectables estériles. En todos los casos la forma debe ser estéril y debe ser fluida hasta el punto de que existe fácil para usar con jeringuilla. Debe ser estable bajo las condiciones de producción y almacenamiento y debe preservarse frente la acción contaminante de microorganismos, tales como bacterias y hongos. El vehículo puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido y similares), mezclas adecuadas de los mismos y aceites vegetales. La fluidez apropiada se puede mantener, por ejemplo, mediante el uso de un revestimiento, tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de dispersión y mediante el uso de tensioactivos. La prevención de la acción de los microorganismos se puede traer aproximadamente mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro de sodio. La absorción prolongada de las composiciones inyectables se puede traer aproximadamente mediante el uso en las composiciones de agentes retardantes de la absorción, por ejemplo, monoesterato de aluminio y gelatina.

Las disoluciones inyectables estériles se preparan mediante la incorporación de los compuestos activos en la cantidad requerida en el disolvente apropiado con diversos de los otros ingredientes anteriormente enumerados, tal como se requiere, seguido de la esterilización filtrada. Generalmente, las dispersiones se preparan mediante la incorporación de los diversos ingredientes activos esterilizados dentro de un vehículo estéril el cual contiene el medio de dispersión básico y los otros ingredientes requeridos a partir de los anteriormente enumerados. En el caso de los polvos estériles para la preparación de disoluciones inyectables estériles, los métodos preferidos de preparación son las técnicas de secado al vacío y liofilización que producen un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una disolución filtrada previamente estéril del mismo.

Tal como se ha usado en la presente memoria, el “vehículo farmacéuticamente aceptable” incluye alguno y todos los disolventes, medios de dispersión, revestimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardante de la absorción y similares. Es bien conocido en la técnica el uso de tales medios y agentes para las sustancias farmacéuticas activas. Excepto en la medida de que cualquier medio convencional o agente es incompatible con el ingrediente activo, se contempla su uso en las composiciones terapéuticas. También se pueden incorporar ingredientes activos complementarios dentro de las composiciones.

Para la administración oral los polipéptidos de la presente invención se pueden incorporar con excipientes y usar en forma de enjagües bucales no ingeribles y dentífricos. Un enjagüe bucal se puede preparar incorporando el ingrediente activo en la cantidad requerida en un disolvente apropiado, tal como una disolución de borato de sodio (Disolución de Dobell). Alternativamente, el ingrediente activo se puede incorporar dentro de un lavado antiséptico que contiene borato de sodio, glicerina y bicarbonato de potasio. El ingrediente activo también se puede dispersar en dentífricos, incluyendo: geles, pastas, polvos y suspensiones. El ingrediente activo se puede añadir en una cantidad terapéuticamente eficaz a un dentífrico en pasta que puede incluir agua, ligantes, abrasivos, agentes aromatizantes, agentes espumantes y humectantes.

Las composiciones de la presente invención pueden estar formuladas en una forma neutral o de sal. Las sales farmacéuticamente aceptables incluyen las sales de adición de ácido (formadas con los grupos de amino libre de la proteína) y las cuales se forman con ácidos inorgánicos tales como, por ejemplo, ácidos hidroc্লórico o fosfórico o tales ácidos orgánicos como acético, oxálico, tartárico, mandélico y similares. Las sales formadas con los grupos de carboxilo libres también pueden derivarse de bases inorgánicas tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio o férrico, y tales bases orgánicas como isopropilamina, trimetilamina, histidina, procaina y similares.

Tras la formulación, las disoluciones se administrarán de una manera compatible con la formulación de dosificación y en tal cantidad como sea terapéuticamente eficaz. Las formulaciones fácilmente se administran en una diversidad de formas de dosificación tales como disoluciones inyectables, cápsulas de liberación de fármaco y similares. Para la administración parenteral en una disolución acuosa, por ejemplo, la disolución debería ser adecuadamente tamponada si fuera necesario y el diluyente líquido primero se vuelve isotónico con suficiente disolución salina o glucosa. Estas disoluciones acuosas particulares son especialmente adecuadas para la administración intravenosa, intramuscular, subcutánea e intraperitoneal. En esta conexión, los medios acuosos estériles que se pueden emplear serán conocidos por

los expertos en la técnica teniendo en cuenta la presente descripción. Por ejemplo, se podría disolver una dosificación en 1 ml de disolución de NaCl isotónica y o bien se añade a 1.000 ml de fluido de hipodermoclisis o se inyecta en el sitio propuesto de la infusión, (véase, por ejemplo, "Remington's Pharmaceutical Sciences" 15ª Edición, páginas 1.035-1.038 y 1.570-1.580). Alguna variación en la dosificación ocurrirá necesariamente dependiendo de la condición del sujeto a tratar. La persona responsable para la administración, en cualquier suceso, determinará la dosis apropiada para el sujeto individual. Además, para la administración humana, las preparaciones deberían encontrar patrones de esterilidad, pirogenicidad, seguridad general y de pureza tal como se requiere por la Oficina FDA de Patrones Biológicos.

VII. Animales Transgénicos/Animales "Knockout"

Se pueden producir animales transgénicos los cuales contienen un transgen funcional que codifica un polipéptido TS10q23.3 funcional o variantes del mismo. Los animales transgénicos que expresan transgenes TS10q23.3, líneas celulares recombinantes derivadas de tales animales y embriones transgénicos pueden ser útiles en métodos para la investigación y la identificación de agentes que inducen o reprimen la función de TS10q23.3. Los animales transgénicos de la presente invención también se pueden usar como modelos para el estudio de indicaciones tales como cánceres.

Se introduce un transgen de TS10q23.3 dentro de un hospedante no humano para producir un animal transgénico que expresa un gen de TS10q23.3 humano o de murine. El animal transgénico se produce mediante la integración del transgen dentro del genoma de una manera que permita la expresión del transgen. Los métodos para producir animales transgénicos están descritos de manera general por Wagner y Hoppe (documento de Patente U.S. N° 4.873.191), Brinster *et al.* 1985 y en "Manipulating the Mouse Embryo; A Laboratory Manual" 2º edición (eds., Hogan, Beddington, Costantini y Long, Cold Spring Harbor Laboratory Press, 1994).

Puede ser deseable reemplazar el TS10q23.3 endógeno mediante recombinación homóloga entre el transgen y el gen endógeno; o se puede eliminar el gen endógeno mediante la delección como en la preparación de animales "knockout". Típicamente, se transfiere un gen de TS10q23.3 flanqueado por secuencias genómicas mediante microinyección dentro de un huevo fertilizado. Los huevos microinyectados se implantan dentro de un hospedante femenino, y se investiga la progenie para la expresión del transgen. Los animales transgénicos se pueden producir a partir de los huevos fertilizados de un número de animales que incluyen, pero no se limitan a reptiles, anfibios, pájaros, mamíferos y peces. Los ratones transgénicos que se generan sobreexpresan TS10q23.3 o expresan una forma mutante del polipéptido. Alternativamente, en una realización preferida la ausencia de un TS10q23.3 en ratones "knockout" permite el estudio de los efectos que la pérdida de la proteína de TS10q23.3 tiene sobre una célula *in vivo*. Los ratones "Knockout" también proporcionan un modelo para el desarrollo de cánceres relacionados con TS10q23.3.

Tal como se ha indicado anteriormente, los animales transgénicos y las líneas celulares derivadas de tales animales pueden encontrar uso en ciertos experimentos de ensayo. En este aspecto, los animales transgénicos y las líneas celulares capaces de expresar TS10q23.3 de tipo natural o mutante pueden ser expuestos a sustancias de ensayo. En estas sustancias de ensayo se puede investigar la capacidad de aumentar la expresión de TS10q23.3 de tipo natural y o funcionan o afectan a la expresión o función del TS10q23.3 mutante.

VIII. Ejemplos

Los siguientes ejemplos se incluyen para demostrar las realizaciones preferidas de la invención. Se debería apreciar por los expertos en la técnica que las técnicas descritas en los ejemplos que siguen representan técnicas descubiertas por el inventor para funcionar bien en la práctica de la invención, y por tanto se puede considerar que constituyen modos preferidos para su práctica

Ejemplo 1

Delecciones Homocigotas en las Líneas Celulares de Glioma

Los inventores han examinado el ADN de una serie de 21 líneas celulares de glioma y cultivos primarios, junto con células normales, para identificar delecciones homocigotas de material genómico en el cromosoma 10. Se eligieron marcadores para su localización aproximada en o cerca de regiones previamente implicadas (Fig. 1). Las células analizadas fueron generadas en el Departamento de Neuro-Oncología UTMDACC (LG11, EFC-2, PL-1, PC-1, JW, FG-2, FG-0, NG-1, PH-2, KE, PC-3 y D77), estaban comercialmente disponibles (U138, A172, U373, U87, U251, U118 y T98G), o se obtuvieron de colaboradores (13 wk astro, D54-MG). Los marcadores se obtuvieron de Research Genetics, Huntsville, AL, o se sintetizaron de la secuencia informada. Una vez la línea celular, EFC-2, reveló una gran delección homocigota asociada con cuatro marcadores de alrededor D10S215 (Fig. 2). Esta delección también se observó por FISH usando YAC 746h6, el cual mapea la región. Otras tres líneas celulares (D-54, A172 y LG11) también demostraron delecciones homocigotas en AFMA086WG9 (AFM086), de ese modo implicando fuertemente la región que contiene un gen supresor de tumor putativo (Fig. 2). Las delecciones en las reacciones PCRTM se realizaron en presencia de dos pares de cebadores (multiplexados) para asegurar condiciones de amplificación apropiadas. Todas las delecciones se confirmaron mediante (al menos) reacciones triplicadas. Esta misma región también se ha implicado

en el carcinoma de próstata (Gray *et al.*, 1995). Las deleciones homocigotas en las líneas celulares también se han usado para definir un locus del gen supresor de tumor en 3p21.3 en carcinoma de pulmón de célula pequeña (Daly *et al.*, 1993; Kok *et al.*, 1994; Wei *et al.*, 1996).

5 Ejemplo 2

Retención del Loci 10q en las Células Híbridas Suprimidas

10 La segunda estrategia de los inventores fue examinar las regiones del cromosoma 10 que se retenían en los clones híbridos suprimidos, pero ausentes en los clones revertidos ("revertant"). Este análisis extendió el estudio previo de los inventores, mostrando la presencia de dos loci supresores de tumor en el cromosoma 10 y analizando las regiones que se retenían. Los híbridos que retenían todo o porciones de 10q fallaron en crecer en la agarosa blanda y en ratones nude (clones "completamente" suprimidos), mientras que las células híbridas que perdieron la mayoría del
15 cromosoma 10q insertado crecieron en agarosa blanda, pero eran no tumorigénicas (clones "parcialmente" suprimidos; Steck *et al.*, 1995; Fig. 3, lado derecho). Se mostró que clones originales U251N10.6, N10.7 y N10.8 previamente retenían solamente fragmentos de 10q (Pershouse *et al.*, 1993; Steck *et al.*, 1995). Usando marcadores microsatélites informativos adicionales, se identificaron tres regiones retenidas en todos los tres clones suprimidos; una región de 22 cM desde D10S219 a D10S110, una región de 14 cM desde D10S192 a D10S187 y una región de 18 cM desde
20 D10S169 a través de D10S1134 (Fig. 3).

Para evitar esta limitación, el cromosoma 10 etiquetado para resistencia a neomicina originalmente transferido desde el híbrido U251.N10.7 fue "rescatado" mediante la transferencia del cromosoma mediado por microcélula dentro de las células A9 de ratón. Esto permite que todos los marcadores microsatélites humanos sean informativos para la
25 presencia del cromosoma 10. La bases para este análisis es que todos los subclones "completamente" suprimidos deberían retener una región común y esta región estar delecionada en los subclones "parcialmente" suprimidos. Un impulso adicional fue que N10.7 manifestó considerable heterogeneidad en el tamaño del cromosoma 10 retenido, tal como se determinó mediante FISH usando sondas específicas al cromosoma 10. Además, en las células híbridas usadas para este rescate se ensayó en primer lugar el crecimiento en agarosa blanda y mostraron no formación de
30 colonia. Todos los híbridos de ratón que contenían el cromosoma 10 humano transferido contenían el brazo corto del cromosoma 10. La misma región era retenida en los clones "parcialmente" suprimidos (N10.5a-j) que crecían en agarosa blanda (Steck *et al.*, 1995), excluyendo así esta región (10pter-10q11) ya que contiene el gen supresor de tumor 10q. El examen de las regiones de 10q retenidas ilustró la considerable heterogeneidad (Fig. 3). La mayoría de los clones mostraron deleciones o bien parciales o extensas de 10q23-26. En todos los subclones examinados solamente
35 dos regiones fueron retenidas. La región más centromérica retenida implicaba los marcadores D10S210 y D10S219. Sin embargo, estos marcadores estaban ausentes en los clones N10.6 y/o N10.8 originales, excluyendo esta región (Fig.3). La otra región era centromérica de D4S536 pero telomérica de D10S215 (≈ 4 cM). Los marcadores AFM086 y D10S536 estaban retenidos en todos los clones examinados (región recuadrada en la Fig. 3). Estos marcadores estaban ausentes en los clones parcialmente suprimidos (N10.5a-j). Estos resultados demostraron que una región común, de
40 alrededor de AFM086, es retenida en todas las células híbridas que son suprimidas de manera fenotípica. Esta misma región está delecionada en varias líneas celulares de glioma.

Este análisis tiene varias limitaciones. En primer lugar, no se pueden analizar la actividad biológica en los clones rescatados, por lo tanto algunos cambios en el cromosoma 10 que pueden haber ocurrido durante o después de la
45 transferencia no se podrían detectar. Para dirigir parcialmente este suceso, se realizó el análisis de los inventores tan pronto como se pudieron recoger los clones. Además, la retención de esta porción del cromosoma solamente puede "corregir" una deleción artefacto *in vitro*. Por consiguiente, se realizaron estudios de deleción alélica para determinar si esta región estaba implicada en gliomas. Además, se sugirió una región alternativa mediante este análisis en D10S1158, donde todos los clones excepto uno (C7) retuvieron esta región. Sin embargo, la región retenida en AFM086 también
50 presentó deleciones homocigóticas, estando implicadas de ese modo por dos métodos alternativos en comparación con D10S1158. También es interesante indicar que la región del gen supresor de tumor parece que preferentemente se retiene, mientras se fragmenta el resto de 10q.

55 Ejemplo 3

Análisis de Deleción Alélica de 10q

Se realizó un estudio de deleción alélica sobre ADN de una serie de 53 especímenes de glioma y linfocitos de pa-
60 ciente correspondientes usando marcadores microsatélites específicos para el cromosoma 10. Este estudio se emprendió para determinar si nuestra región crítica también estaba implicada en los especímenes de glioma. Se observaron extensas deleciones en la mayoría de los especímenes derivados de GBM, con 30 de 38 GBMs que presentan deleción de la mayoría o de todos los marcadores del cromosoma 10. Se observaron deleciones menos extensas en la mayoría si los especímenes derivaban de astrocitomas anaplásticos, mientras que se observaron deleciones infrecuentes en los
65 astrocitomas y en la mayoría de los oligodendrogliomas (Fig. 4 y datos no mostrados). La mayoría de los marcadores usados en este análisis se mapearon en 10q23-26 (Gyapay *et al.*, 1994). Similar a otros estudios, no se podría demostrar convincentemente una región común de deleción, debido a las grandes deleciones en la mayoría de las muestras de GBM (Fults *et al.*, 1993; Rasheed *et al.*, 1995).

Sin embargo, para los especímenes GBM examinados, todos excepto una muestra de tumor (N° 9; Fig. 4) revelaron deleciones que implicaban la región desde D10S579 a D10S541. Además, solamente un AA mostró una deleción en la región crítica de los inventores, y no astrocitomas. Dos oligodendrogliomas presentaron deleciones dentro de la región crítica, pero ambos fueron diagnosticados como malignos. Este estudio presenta diversas posibilidades. En primer lugar, las deleciones que implican la región crítica de los inventores ocurren predominantemente en GBMs y no en tumores de grado inferior. Esto implicaría que la pérdida del gen supresor de tumor en el cromosoma 10q en la región crítica de los inventores representaría una alteración genética asociada con la progresión a GBM. En apoyo a esta hipótesis, aún cuando las deleciones ocurren en 10q en tumores de grado inferior, no se identificó región común de deleción en 10q para estos especímenes. Estas observaciones apoyarían, de nuevo, la sugerencia anterior de los inventores de que la deleción del gen supresor de tumor 10q predominantemente está asociada con GBMs y no todas las deleciones en 10q afectan al gen supresor de tumor. La región D10S216 a D10S587, sugerida por Rasheed, mostró deleciones extensas, pero varios GBMs presentaron retención de heterocigosidad en esta región (tumores N° 2, N° 9, N° 13, N° 26; Fig.4). Además, si se excluyen los tumores de grado bajo de su estudio, la región de los inventores está implicada en todos GBMs. Esta combinación de enfoques independientes fuertemente sugiere un gen supresor de tumor 10q que mapea en la región D10S215 a D10S541, específicamente en AFM086.

Ejemplo 4

20 Mapeado de la Región del Gen del Supresor de Tumor Candidato

La región crítica que los inventores han identificado está centrada en AFM086 y está bordeada por D10S215 y S10S541 (Figs. 2 y 8). Esta región es relativamente pequeña, estando contenida dentro de varios YACs individuales (787d7; 746h8; 934d3). FISH que aplica con YAC 746h8 en las propagaciones de la metafase EFC-2 muestra que la deleción homocigótica está contenida dentro del YAC ya que se observó parcialmente el YAC y YACs adjuntos sobre ambos lados estaban presentes. Se han aislado cromosomas artificiales bacterianos (BACs, del Inglés “Bacterial Artificial Chromosome”) o PACs para todos los marcadores en la región (Fig. 8). El BAC que sigue a la región estaba construido a partir de las secuencias finales de BACs que mapeaban en la región. Se han identificado varios rasgos notables. En primer lugar, se identificaron dos BACs solapados (46b12 y 2f20) y verifican la integridad genómica de 106d16. En segundo lugar, se identificó un sitio *NotI* en un extremo de los BACs. La presencia del sitio *NotI* y la digestión de restricción coincidente con *SacII*, *EagI* y *BssHIII* sugiere la presencia de una isla CpG dentro de 106d16.

Se usaron los fragmentos *EcoRI* de BAC 106d16 para examinar la extensión de las deleciones homocigóticas, mediante transferencia Southern, en las células de glioma que previamente se muestra que habían delecionado homocigóticamente AFM086 (Figs. 2 y 5). El lado derecho (fragmento 14 de *EcoRI*) contiene la probable isla CpG y está presente en tres de las cuatro líneas celulares. Se usó un fragmento *NotI/EcoRI* (N° 3) como sonda en una transferencia Southern que contiene varios BACs y la línea celular de glioma (Fig. 2). No se han detectado deleciones al lado telomérico (lado derecho) usando sondas del 46b12, excepto para las células EFC-2. Sin embargo, se han observado deleciones homocigóticas adicionales en las células dentro de la región definida por 106d16 (65 kb). Se observa una deleción homocigótica para la banda 3 para las células LG11 y EFC-2, pero no para las células de glioma adicionales o controles normales. Se ha observado que 106d16 está presente en todas las células (EFC-2 presenta una banda de migración alterada), que sugiere que la deleción homocigótica está contenida enteramente dentro de 106d16.

45 Ejemplo 5

Identificación de Genes Expresados dentro de la Región Crítica

Se generaron fragmentos *EcoRI* a partir de BAC 106d16 y se separaron por tamaño mediante electroforesis en gel de agarosa. Se ligaron bandas individuales o conjuntos de bandas de tamaño similar dentro de pSPL3 (GIBCO, Gaithersburg, MD). Se identificaron exones putativos tal como lo describe el fabricante. Se sometieron a maduración por corte y empalme apropiadamente dos exones dentro del vector atrapante. Los exones se derivan del conjunto de banda 2, 3, 4, 5 y banda 7. Se determinó la secuencia de los exones atrapados y se definieron por la secuencia vector atrapante conocida. Usando búsquedas BLAST de la base de datos de etiqueta de secuencia expresada (dbEST), se identificaron cinco etiquetas potenciales de secuencia expresada (ESTs, del Inglés “Expressed Sequence Tag”). Se observaron dos ESTs (gb/H92038, AA009519) que contenían o bien uno o ambos de los exones (si bien es cierto que una EST estaba en la orientación equivocada).

Se generaron cebadores de secuenciación a partir de las ESTs y se usaron para definir los límites de exón-intrón putativos usando BAC46b12 como plantilla. Se identificaron nueve exones. Las diferencias de secuencia entre las ESTs y la plantilla genómica fueron corregidas. Todos los exones estaban contenidos dentro de BAC 46b12. Se generaron cebadores a partir de las secuencias de intrón adyacentes a los exones para formar unidades amplicón para cada exón. Dos de los exones se correspondían a los exones atrapados de las secuencias BAC 106d16 *EcoRI*. La secuencia del gen se muestra en la Fig. 6. La lectura de aminoácidos pronosticada se definió por la presencia de un sitio de comienzo ATG, codones de terminación TGA y TAA en el marco, la presencia de codones de terminación múltiples en todos los tres marcos de lectura en otro lugar en la secuencia, nueve sitios de maduración por corte y empalme, y la presencia de señales Kozak cerca del sitio de iniciación. La secuencia de aminoácidos 403 se muestra en la Fig. 7 y Fig. 9. El peso molecular pronosticado es de 47.122 con un pl de 5.86.

ES 2 361 044 T3

Se sugiere un posible papel funcional para el producto proteico mediante su homología de secuencia a varios motivos de proteína. Un motivo crítico a partir de los residuos 88 a 98[IHCKAGKGRG](SEC. ID. N°: 17) tiene una correspondencia exacta para el dominio catalítico conservado de una protein tirosin fosfatasa [(I/V)HCxAGxxR (S/T)G](SEC. ID. N°: 18)(Denu *et al.*, 1996). Se identificaron varios otros motivos que concordarían con la función fosfatasa del gen supresor de tumor.

Se generaron amplicones (productos de PCRTM generados a partir de diversas regiones del gen) a partir de ADN cebado al azar. La secuencia de los amplicones correspondían a la secuencia de ADN. Se usaron amplicones no solapados para sondar transferencias Northern de tejido normal derivadas de diversos órganos (Clontech, Palo Alto, CA; transferencias multitejido). Todos los amplicones identificaron una banda principal en 5,5 a 6 kb sobre las transferencias Northern y varias bandas menores. El mensaje se expresó en todos los tejidos examinados (corazón, cerebro, placenta, pulmón, hígado, esqueleto, músculo, riñón, páncreas, bazo, timo, próstata, testículos, ovario, intestino delgado, colon y linfocitos de sangre periférica).

Ejemplo 6

Análisis Mutacional

Los análisis mutacionales inicialmente han avanzado sobre dos frentes. En primer lugar, se analizó la presencia del gen candidato en las líneas celulares de glioma que inicialmente mostraron tener deleciones homocigóticas. Tal como se muestra en la Fig. 8, todas la líneas celulares que presentaron deleción de AFM086 tenían deleciones homocigóticas de exones múltiples del gen candidato. Además, las deleciones ocurrieron en medio del gen, definiendo así los límites de deleción (deleciones similares en todas las líneas celulares) entre exones B y G. Las deleciones que afectan a la mitad del gen además indican que el gen identificado representa el gen diana para la mutación.

También se realizaron análisis preliminares para las mutaciones de secuencia en una serie de líneas celulares de glioma. Se observaron mutaciones y/o deleciones en todas excepto tres líneas celulares de glioma examinadas (Tabla 4). La referencia al número base en la tabla se refiere al exón, no a las secuencias enteras, es decir, a la 98ª base del exón G para U251.

TABLA 4

MUTACIONES IDENTIFICADAS EN EL GEN CANDIDATO				
	Células	Tipo de Célula	Mutación	Efecto Pronosticado
1	U87	Glioma	Exón c de junta de corte y empalme:G+1>T	Variante de corte y empalme
2	U138	Glioma	Exón h del sitio de corte y empalme: G+1>T	Variante de corte y empalme
3	U251	Glioma	Exón G de adición de 2 pb; 98 ins TT	
4	U373	Glioma	Exón G de desplazamiento de pauta de lectura	
5	EFC-2	Glioma	-todos los exones	No producto
6	D54	Glioma	-exones C-I	No producto
7	A172	Glioma	-exones C-I	No producto
8	LG11	Glioma	-exones B-1	No producto
9	T98G	Glioma	Exón B de sentido erróneo; T46->G	leu>arg
10	KE	Glioma	Exón B de sentido erróneo; G28->A	gly>glu
11	F60	Glioma	Exón H de mutación terminal; C202->T	Parada terminal
12	D77	Glioma	No mutación (heterogéneo para 10q)	
13	PC-3	Bajo grado	No mutación	
14	PH-2	Bajo grado	No mutación	
15	nLnCap	Próstata	Exón A de deleción, 16-17 del AA; mutación B, C53->T	silencio

Además, se encontraron deleciones de exones en LnCap, una línea celular de próstata. Las células de glioma que fallaron en mostrar una mutación/deleción eran derivadas de tumores de bajo grado (PC-3 y PH-2) donde se espera no deleción alélica del cromosoma 10 y se ha observado para estas células. Las otras células (D77) eran un cultivo celular primario, y se mostró que el cromosoma 10 era heterocigótico a partir de un polimorfismo de 1 pb dentro del género. Una línea celular de cáncer de pecho también mostró una mutación. Este análisis inicial apoya la conclusión de los inventores de que la pérdida de un gen supresor de tumor 10q representa un marcador molecular crítico para la enfermedad y la progresión del glioblastoma.

Referencias

- 5 **Albarosa et al.**, Deletion mapping of gliomas suggest the presence of two small regions for candidate tumor-suppressor genes in a 17-cM interval on chromosome 10q. *Am. J. Genet.*, 58:1.260-1.267, 1996.
- 10 **Baichwal y Sudgen**, Vectors for gene transfer derived from animal DNA viruses: Transient and stable expression of transferred genes. En: Gene Transfer, Kucherlapati R. ed, New York, *Plenum Press*, pp. 117-148, 1986.
- 15 **Barany y Merrifield**, The Peptides, Gross and Meienhofer, eds., *Academic Press*, New York, pp. 1-284, 1979.
- 20 **Benvenisty y Neshif**, Direction introduction of genes into rats and expression of the genes. *Proc. Nat. Acad. Sci. USA*, 83:9.551-9.555, 1986.
- 25 **Bianchi et al.**, Mutations in transcript isoforms of the neurofibromatosis 2 gene in multiple human tumor types. *Nature Genetics*, 6:185-192, 1994.
- 30 **Bigner et al.**, Specific chromosomal abnormalities in malignant human gliomas. *Cancer Res.*, 48:405-411, 1988.
- 35 **Bigner et al.**, Specific chromosomal abnormalities in malignant human gliomas. *Cancer Res.* 48:405-411, 1988.
- 40 **Bishop, J. M.** Molecular themes in oncogenesis. *Cell*, 64:2.351-248, 1991.
- 45 **Boring et al.**, *Cancer Statistics*, 1994 CA, 43:7-26, 1994.
- 50 **Brinster et al.**, *Proc. Nat'l Acad. Sci. USA*, 82:4.438-4.442, 1985.
- 55 **Capaldi et al.**, *Biochem. Biophys. Res. Comm.*, 76:425, 1977.
- 60 **Carter et al.**, Allelic loss of chromosome 16q and 10q in human prostate carcinoma. *Proc. Natl. Acad. Sci. USA*, 87:8.751-8.755, 1990.
- 65 **Chang et al.**, Foreign gene delivery and expression in hepatocytes using a hepatitis B virus vector. *Hepatology*, 14:124A, 1991.
- 70 **Chen y Okayama**, High-efficiency transfection of mammalian cells by plasmid DNA. *Mol. Cell. Biol.*, 7:2.745-2.752, 1987.
- 75 **Coffin**, Retroviridae and Their Replication. En: Virology, Fields *et al.*, eds., *Raven Press*, New York, pp. 1.437-1.500, 1990.
- 80 **Cohen, P.**, The discovery of protein phosphatases: From chaos and confusion to an understanding of their role in cell regulation and human disease. *Bioessays*, 61-583-588, 1994.
- 85 **Collet et al.**, Protein Kinase activity associated with the avian sarcoma virus src gene product. *Proc. Natl. Acad. Sci. USA*, 75:2.021-2.024, 1978.
- 90 **Coupar et al.**, A general method for the construction of recombinant vaccinia virus expressing multiple foreign genes, *Gene*, 68:1-10, 1988.
- 95 **Culver et al.**, *In vivo* gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors. *Science*, 256:1.550-1.552, 1992.
- 100 **Daly et al.**, A homozygous deletion on chromosome 3 in small cell lung cancer cell line correlates with a region of tumor suppressive activity. *Oncogene* 8:1.721-1.729, 1993.
- 105 **Denu et al.**, Form and function in protein dephosphorylation. *Cell* 87:361-364, 1996.
- 110 **Dubensky et al.**, Direct transfection of viral and plasmid DNA into the liver or spleen of mice. *Proc. Nat. Acad. Sci. USA*, 81:7.529-7.533, 1984.
- 115 **El-Azouzi et al.**, Loss of distinct regions on the short arm of chromosome 17 associated with tumorigenesis of human astrocytomas, *Proc. Natl. Acad. Sci. USA*, 86: 7.186-7.190, 1989.
- 120 EP 329 822, **Davey et al.** Fearron y Vogelstein, A genetic model for colorectal tumorigenesis. *Cell*, 61:759-767, 1990.
- 125 **Fechheimer et al.**, Transfection of mammalian cells with plasmid DNA by scrape loading and sonication loading. *Proc. Natl. Acad. Sci. USA*, 84:8.463-8.467, 1987.

ES 2 361 044 T3

- Ferkol et al.**, Regulation of the phosphoenolpyruvate carboxykinase/human factor IX gene introduced into the livers of adult rats by receptor-mediated gene transfer. *FASEB J.*, 7:1.081-1.091, 1993.
- Fodor et al.**, Light-directed, spatially addressable parallel chemical synthesis. *Science*, 251:767-773, 1991.
- 5 **Foulds**, The natural history of cancer. *J. Chronic Dis.*, 8:2-37, 1958.
- Fraley et al.**, Entrapment of a bacterial plasmid in phospholipid vesicles: Potential for gene transfer. *Proc. Natl. Acad. Sci. USA*, 76:3.348-3.352, 1979.
- 10 **Freshner**, Animal Cell Culture: A Practical Approach, 2^a ed., Oxford/New York, *IRL Press, Oxford University Press*, 1992.
- Friedmann**, Progress toward human gene therapy. *Science*, 244:1.275-1.281, 1989.
- 15 **Frohmann**, En: PCR Protocols: A Guide To Methods And Applications, *Academic Press*, N.Y., 1990.
- Fujimoto et al.**, Loss of heterozygosity on chromosome 10 in human glioblastoma multiforme. *Genomics*, 4:210-214, 1989.
- 20 **Fults & Pedone**, Deletion mapping of the long arm of chromosome 10 in glioblastoma and multiforme. *Genes Chromosom, Cancer* 7:173-177, 1993.
- Fults et al.**, Allelotype of human malignant astrocytoma, *Cancer Res.*, 50:5.784-5.789, 1990. GB 2 202 328.
- 25 **Gefter et al.**, *Somatic Cell Genet.*, 3:231-236, 1977.
- Ghosh-Choudhury et al**, *EMBO J.*, 6:1.733-1.739, 1987.
- 30 **Ghosh y Bachhawat**, Targeting of Liposomes to Hepatocytes. En: Liver Diseases, Targeted Diagnosis and Therapy using Specific Receptors and Ligands. Wu *et al.*, eds., *Marcel Dekker*, New York, pp. 87-104, 1991.
- Goding**, 1986, En: Monoclonal Antibodies: Principles and Practice, 2^a ed, *Academic Press*, Orlando, Fla., pp. 60-61 y 71-74, 1986.
- 35 **Gomez-Foix et al.**, *J. Biol. Chem.*, 267:25.129-25.134, 1992.
- Gopal**, Gene transfer method for transient gene expression, stable transfection, and cotransfection of suspension cell cultures. *Mol. Cell. Biol.*, 5:1.188-1.190, 1985.
- 40 **Graham y Prevec**, En: Methods in Molecular Biology: Gene Transfer and Expression Protocol, E.J. Murray, ed., *Humana Press*, Clifton, NJ, 7:109-128, 1991.
- Graham y van der Eb**, A new technique for the assay of infectivity of human adenovirus 5 DNA *Virology*, 52:456-467, 1973.
- 45 **Graham et al.**, Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.*, 36:59-72, 1977.
- 50 **Gray et al.**, Loss of chromosomal region 10q23-25 in prostate cancer. *Cancer Res.*, 55:4.800-4.803, 1995.
- Grunhaus y Horwitz**, Adenovirus as cloning vector. *Seminar in Virology*, 3:237-252, 1992.
- Gyapay et al.**, The 1993-94 Genethon human genetic linkage map. *Nat. Genet.*, 7:246-339, 1994.
- 55 **Hacia et al.**, Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nature Genetics*, 14:441-447, 1996.
- Harland y Weintraub**, Translation of mammalian mRNA injected into *Xenopus* oocytes is specifically inhibited by antisense RNA. *J. Cell Biol.*, 101:1.094-1.099, 1985.
- 60 **Henson et al.**, The retinoblastoma gene is involved in malignant progression of astrocytomas. *Ann. Neurol.*, 36:714-721, 1994.
- 65 **Herbst et al.**, Loss of heterozygosity for 10q22-qter in malignant melanoma progression. *Cancer Res.*, 54:3.111-3.114, 1994.

ES 2 361 044 T3

Hermonat y Muzycska, Use of adenoassociated virus as a mammalian DNA cloning vector: Transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Nat. Acad. Sci. USA*, 81:6.466-6.470, 1984.

Hersdorffer et al., *DNA Cell Biol.*, 9:713-723, 1990.

Herz y Gerard, *Proc. Nat'l Acad Sci. USA*, 90:2.812-2.816, 1993.

Horwich, et al., Synthesis of hepadnavirus particles that contain replication-defective duck hepatitis B virus genomes in cultured HuH7 cells. *J. Virol.*, 64:642-650, 1990.

Hunter, T., Cooperation between oncogenes. *Cell*, 64-249-270, 1991.

Ittmann, Allelic loss on chromosome 10 in prostate adenocarcinoma. *Cancer Res.*, 56:2.143-2.147, 1996.

James et al., Clonal genomic alterations in glioma malignancy stages. *Cancer Res.*, 48:5.546-5.551, 1988.

Johnson et al., Peptide Turn Mimetics. En: *Biotechnology And Pharmacy*, Pezzuto et al., eds., Chapman and Hall, New York, 1993.

Jones y Shenk, *Cell*, 13:181-188, 1978.

Kamb et al., A cell cycle regulator potentially involved in genesis of many tumor types. *Science*, 264:436-440, 1984.

Kaneda et al., Increased expression of DNA cointroduced with nuclear protein in adult rat liver. *Science*, 243:375-378, 1989.

Karlbom et al., Loss of heterozygosity in malignant gliomas involves at least three distinct regions on chromosome 10. *Hum. Genet.*, 92:169-174, 1993.

Karisson et al., *EMBO J.*, 5:2.377-2.385, 1986.

Kato et al., Expression of hepatitis B virus surface antigen in adult rat liver. *J. Biol. Chem.*, 266:3.361-3.364, 1991.

Kimmelman et al., Loss of heterozygosity of chromosome 10p in human gliomas. *Genomics* 34:250-254, 1996.

Klein et al., High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature*, 327:70-73, 1987.

Kok et al., A homozygous deletion in a small cell lung cancer cell line involving a 3p21 region with a marked instability in yeast artificial chromosomes. *Cancer Res.* 54:4.183-4.187, 1994.

Komiya et al., Allelic losses at loci on chromosome 10 are associated with metastasis and progression of human prostate cancer. *Genes Chromo. Cancer* 17:245-253, 1996.

Kwoh et al., *Proc. Nat. Acad Sci. USA*, 86:1.173, 1989.

Kyte y Doolittle, *J. Mol. Biol.*, 157(1):105-132, 1982.

Le **Gal La Salle et al.**, *Science*, 259:988-990, 1993.

Lee et al., Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science*, 235:1.394-1.399, 1987.

Levin et al., Neoplasms of the central nervous system. En: *Cancer: Principles & Practice of Oncology*, 4^a ed., DeVita et al., eds., J.B. Lippincott Co., Philadelphia, 1993.

Levrero et al., *Gene*, 101:165-202, 1991.

Macejak y Samow, *Nature*, 353:90-94, 1991.

Manipulating the Mouse Embryo: A Laboratory Manual, 2^o ed., Hogan et al., eds., *Cold Spring Harbor Laboratory Press*, 1994.

Mann et al., Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, 33:153-159, 1983.

Markowitz et al., *J. Virol.*, 62:1.120-1.124, 1988.

ES 2 361 044 T3

Merrifield, *Science*, 232:341-347, 1986.

Morita et al., Common regions of deletion on chromosomes 5q, 6q and 10q in renal cell carcinoma. *Cancer Res.*, 51:5.817-5.820, 1991.

Mulligan, *Science*, 260:926-932, 1993.

Murakami et al., Suppression of the malignant phenotype of human prostate cancer cell line PPC-1 by introduction of normal fragments of human chromosome 10. *Cancer Res.*, 56:2.157-2.160, 1996.

Myers, EP 0273085.

Nicolas y Rubinstein, En: Vectors: A survey of molecular cloning vectors and their uses, Rodriguez and Denhardt, eds., Stoneham:Butterworth, pp. 494-513, 1988.

Nicolau y Sene, Liposome-mediated DNA transfer in eukaryotic cells. *Biochim. Biophys. Acta*, 721:185-190, 1982.

Nihei et al., Localization of a metastasis suppressor gene(s) for rat prostatic cancer to the long arm of human chromosome 10. *Genes Chromosom. Cancer*, 14:112-119, 1995.

Nishi et al., Differential expression of two types of the neurofibromatosis type 1 gene transcripts related to neural differentiation. *Oncogene*, 6:1.555-1.559, 1991.

Paskind et al., Dependence of moloney murine leukemia virus production on cell growth. *Virology*, 67:242-248, 1975.

PCT/US87/00880.

PCT/US89/01025.

Pease et al., Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. USA*, 91:5.022-5.026, 1994.

Peiffer et al., Allelic loss of sequences from the long arm of chromosome 10 and replication errors in endometrial cancers. *Cancer Res.*, 55:1.922-1.926, 1995.

Pelletier y Sonenberg, *Nature*, 334:320-325, 1988.

Perales et al., Gene transfer in vivo: Sustained expression and regulation of genes introduced into the liver by receptor-targeted uptake. *Proc. Natl. Acad. Sci.* 91:4.086-4.090, 1994.

Pershouse et al., Analysis of the functional role of chromosome 10 loss in human glioblastomas. *Cancer Res.* 53: 5.043-5.050, 1993.

Petersen et al., Small-cell lung cancer is characterized by a high incidence of deletions on chromosomes 3p, 4q, 5q, 13q y 17p. *Brit. J. Cancer* 75:79-86, 1997.

Pignon et al., *Hum. Mutat.*, 3:126-132, 1994.

Racher et al., *Biotechnology Techniques*, 9:169-174, 1995.

Ragot et al., *Nature*, 361:647-650, 1993.

Ransom et al., Correlation of cytogenetic analysis and loss of heterozygosity studies in human diffuse astrocytomas and mixed oligo-astrocytomas. *Genes Chromosom. Cancer* 5:357-374, 1992.

Rasheed et al., Loss of heterozygosity for 10q loci in human gliomas. *Genes Chromo. Cancer*, 5:75-82, 1992.

Rasheed et al., *Oncogene*, 11:2.243-2.246, 1995. *Remington's Pharmaceutical Sciences*, 15^a ed., pp. 1.035-.038 y 1.570-1.580.

Rempel et al., Loss of heterozygosity for loci on chromosome 10 is associated with morphologically malignant meningioma progression. *Cancer Res.*, 53:2.386-2.392, 1993.

Rich et al., *Hum. Gene Ther.*, 4:461-476, 1993.

ES 2 361 044 T3

Ridgeway, Mammalian Expression Vectors, En: Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Rodriguez *et al.*, eds., Stoneham:Butterworth, pp. 467-492, 1988.

5 **Ritland *et al.***, Region-specific loss of heterozygosity on chromosome 19 is related to the morphologic type of human glioma. *Genes Chromo. Cancer*, 12:277-282, 1995.

Rosenfeld *et al.*, In vivo transfer of the human cystic fibrosis transmembrane conductance regulator gene to the airway epithelium. *Cell*, 68:143-155, 1992.

10 **Rosenfeld *et al.***, *Science*, 252:431-434, 1991.

Roux *et al.*, A versatile and potentially general approach to the targeting of specific cell types by retroviruses: Application to the infection of human cells by means of major histocompatibility complex class I and class II antigens by mouse ecotropic murine leukemia virus-derived viruses. *Proc. Nat'l Acad. Sci. USA*, 86:9.079-9.083, 1989.

15 **Rubio *et al.***, Analysis of the neurofibromatosis 2 gene in human ependymonas and astrocytomas, *Cancer Res.*, 54:45-47, 1994.

20 **Russell y Rubinstein**, Tumors of the neuroepithelial origin, En: Pathology of Tumors of the Nervous System, 5^a ed., Williams and Wilkins, eds., pp. 82-219, 1989.

Sambrook *et al.*, En: Molecular Cloning: A Laboratory Manual 2^a Ed., *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, NY, 1989.

25 **Sanchez *et al.***, Tumor suppression and apoptosis of human prostate carcinoma mediated by a genetic locus within human chromosome 10pter-q11. *Proc. Natl. Acad. Sci. USA*, pp. 2.551-2.556, 1996.

Sarver *et al.*, *Science*, 247:1.222-1.225, 1990.

30 **Scheck y Coons**, Expression of the tumor suppressor gene DCC in human gliomas. *Cancer Res.*, 53:5.605-5.609, 1993.

Shoemaker *et al.*, Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular barcoding strategy. *Nature Genetics* 14:450-456, 1996.

35 **Sonoda *et al.***, *Cancer Res.*, 55:2.166-2.168, 1995.

Steck *et al.*, Evidence for two tumor suppressive loci on chromosome 10 involved in glioblastomas. *Genes Chromosom. Cancer* 712:255-261, 1995.

40 **Stewart y Yound**, Solid Phase Peptide Synthesis, 2^a ed., Pierce Chemical Co., 1984.

Stratford-Perricaudet y Perricaudet, Gene transfer into animals: the promise of adenovirus. En: Human Gene Transfer, O. Cohen-Haguenuer *et al.*, eds., John Libbey Eurotext, Francia, pp. 51-61, 1991.

45 **Stratford-Perricaudet *et al.***, Evaluation of the transfer and expression in mice of an enzyme-encoding gene using a human adenovirus vector. *Hum. Gene. Ther.*, 1:241-256, 1990.

Tam *et al.*, *J. Am. Chem. Soc.*, 105:6.442, 1983.

50 **Temin**, Retrovirus vectors for gene transfer: Efficient integration into and expression of exogenous DNA in vertebrate cell genome. En: Gene Transfer, Kucherlapati R., ed. New York, *Plenum Press*, pp. 149-188, 1986.

55 **Trybus *et al.***, Distinct areas of allelic loss on chromosomal regions 10p and 10q in human prostate cancer. *Cancer Res.* 56:2.263-2.267, 1996.

Tur-Kaspa *et al.*, Use of electroporation to introduce biologically active foreign genes into primary rat hepatocytes. *Mol. Cell Biol.*, 6:716-718, 1986.

60 U.S. 4.873.191, **Wagner y Hoppe** U.S. 5.279.721

Varmus *et al.*, *Cell*, 25:23-36, 1981.

Von **Deimling *et al.***, Deletion mapping of chromosome 19 in human gliomas. *Int. J. Cancer*, 57:676-680, 1994.

65 **Voullaire *et al.***, A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am. J. Hum. Genet.*, 52:1.153-1.163.

ES 2 361 044 T3

Wagner et al., *Science*, 260: 1.510-1.513, 1993.

Walker et al., *Proc. Nat'l Acad. Sci. USA*, 89:392-396, 1992.

5 **Wei et al.**, Construction of a 600-kilobase cosmid clone containing and generation of a transcriptional map surrounding the lung cancer tumor suppressor gene (TSG) locus on human chromosome 3p21.3: progress toward the isolation of a lung cancer TSG. *Cancer Res.* 56:1.487-1.494, 1996.

Weinberg, Positive and negative controls on cell growth. *Biochemistry*, 28:8.263-8.269, 1989.

10 WO 88/10351, Gingeras *et al.*, WO 89/06700, Miller *et al.*, WO 90/07641, presentado 21 de Diciembre, 1990.

Wong et al., Appearance of b-lactanase activity in animal cells upon liposome mediated gene transfer, *Gene*, 10:87-94, 1980.

15 **Wong et al.**, Increased expression of the EGF receptor gene in malignant gliomas is invariably associated with gene amplification. *Proc. Natl. Acad. Sci. USA*, 84:6.899-6.903, 1987.

Wu y Wu, *Adv. Drug Delivery Rev.*, 12:159-167, 1993.

20 **Wu y Wu**, Evidence for targeted gene delivery to HepG2 hepatoma cells *in vitro*. *Biochemistry*, 27:887-892, 1988.

Wu y Wu, Receptor-mediated *in vitro* gene transfections by a soluble DNA carrier system. *J. Biol. Chem.*, 262:4.429-4.432, 1987.

25 **Wu et al.**, *Genomics*, 4:560, 1989.

Yamaguchi et al., Differential expression of two fibroblast growth factor-receptor genes is associated with malignant progression in human astrocytomas. *Proc. Natl. Acad. Sci. USA*, 91:484-488, 1994.

30 **Yang et al.**, *In vivo* and *in vitro* gene transfer to mammalian somatic cells by particle bombardment. *Proc. Natl. Acad. Sci. USA*, 87:9.568-9.572, 1990.

35 **Zelenin et al.**, High-velocity mechanical DNA transfer of the chloramphenicol acetyltransferase gene into rodent liver, kidney and mammary gland cells in organ explants and *in vivo*. *FEBS Lett.*, 280:94-96, 1991.

40

45

50

55

60

65

ES 2 361 044 T3

REIVINDICACIONES

1. Un polipéptido supresor de tumor designado TS10q23.3 seleccionado entre el grupo que consiste en:

- (a) el polipéptido mostrado en la SEC. ID. N°: 1;
- (b) un fragmento del polipéptido (a) que tiene función fosfatasa; y
- (c) una variante alélica del polipéptido (a).

2. Un fragmento del polipéptido mostrado en la SEC. ID. N°: 1 adecuado como antígeno para anticuerpos provocadores específicos para uno o más polipéptidos como el de la reivindicación 1, en el que dicho fragmento incluye residuos contiguos de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400 ó más aminoácidos de longitud.

3. El polipéptido de la reivindicación 1 o fragmento de la reivindicación 2 el cual se conjuga a una molécula vehículo.

4. Una proteína de fusión que comprende el polipéptido de la reivindicación 1 o el fragmento de la reivindicación 2.

5. Un anticuerpo que se une específicamente a un polipéptido de la reivindicación 1 o a un fragmento de la reivindicación 2.

6. El anticuerpo de la reivindicación 5 el cual es un anticuerpo monoclonal.

7. El anticuerpo de la reivindicación 6, el cual está marcado con una marca detectable.

8. El anticuerpo de la reivindicación 6, el cual es útil en procedimientos inmunoquímicos o inmunohistoquímicos.

9. El anticuerpo de la reivindicación 8, en el que dicho procedimiento inmunoquímico es un método ELISA o transferencia Western.

10. El anticuerpo de la reivindicación 8, en el que dicho procedimiento inmunohistoquímico es tinción de tejido.

11. El anticuerpo de cualquiera de las reivindicaciones 6 a 10, en el que dicho anticuerpo es un reactivo no cruzado con otros polipéptidos humanos.

12. Una célula de hibridoma que procede de un anticuerpo monoclonal de la reivindicación 6.

13. Un método para producir anticuerpos que se unen específicamente a un polipéptido de la reivindicación 1 o un fragmento de la reivindicación 2 que comprende el uso de dicho polipéptido o dicho fragmento como inmunógeno.

14. Un ácido nucleico seleccionado del grupo que consiste en:

- (a) una secuencia de nucleótido que codifica el polipéptido mostrado en la SEC. ID. N°: 1;
- (b) la secuencia codificadora para un polipéptido supresor de tumor mostrado en SEC. ID. N°: 2 que comienza en el nucleótido 1.035 (designado posición +1); SEC. ID. N°: 3 que comienza en el nucleótido 751 (designado posición +1); o SEC. ID. N°: 4 que comienza en el nucleótido 109 (designado posición +1);
- (c) la secuencia de nucleótidos del exón B, C, D, E, F, G, H o I mostrado en la Figura 10;
- (d) una secuencia de nucleótidos que es un fragmento de la secuencia de nucleótido (a) o (b) y la cual codifica un polipéptido que tiene una función fosfatasa; y
- (e) una secuencia de nucleótidos que codifica una variante alélica de un polipéptido codificado por la secuencia de nucleótido (a) o (b);

o la cadena complementaria de dicha secuencia de nucleótido.

15. El ácido nucleico de la reivindicación 14, en el que el polipéptido codificado es un polipéptido de murine, canino o humano.

ES 2 361 044 T3

16. Un ácido nucleico que comprende

- (i) una secuencia de nucleótido tal como se define en cualquiera de (a) a (b) de la reivindicación 14, en la que dicha secuencia de nucleótidos es un ADN complementario; y
- (ii) además comprende un promotor enlazado de manera operable a dicha secuencia de nucleótidos.

5

17. El ácido nucleico de la reivindicación 16, que comprende además una señal de poliadenilación enlazada de manera operable a dicha secuencia de codificación.

10

18. El ácido nucleico de la reivindicación 17, que comprende además un origen de replicación.

19. Un vector de expresión o un vector vírico que comprende el ácido nucleico de la reivindicación 18.

15

20. Un oligonucleótido de entre 10 y 50 bases consecutivas del ácido nucleico de la reivindicación 14 ó 15, o complementario al mismo.

21. El oligonucleótido de la reivindicación 20 útil para la cuantificación de un ARNm que codifica el polipéptido TS10q23.3 de la reivindicación 1 o para la secuenciación de ADN.

20

22. Un método de diagnóstico del cáncer o ascenso del desarrollo de cáncer o metástasis en un sujeto humano que comprende la determinación en una muestra de dicho sujeto si hay una mutación puntual, delección o inserción en una secuencia de nucleótidos de TS10q23.3 de dicho sujeto en comparación con la región codificadora de la secuencia de nucleótido mostrada en la SEC. ID. N°: 2, siendo dicha mutación puntual, delección o inserción indicativa de cáncer o ascenso del desarrollo de cáncer o metástasis.

25

23. El método de la reivindicación 22, en el que dicho cáncer está seleccionado entre el grupo que consiste en cerebro, glioblastoma, meduloblastoma, astrocitoma, oligodendroglioma, ependimoma, pulmón, hígado, bazo, riñón, páncreas, intestino delgado, células sanguíneas, ganglio linfático, colon, pecho, endometrio, estómago, próstata, testículo, ovario, piel, cabeza y cuello, esófago y médula ósea.

30

24. El método de la reivindicación 22, en el que dicha mutación, delección o inserción puntual está seleccionada entre el grupo que consiste en:

35

(a) un ácido nucleico que tiene una mutación de maduración por corte y empalme ("splicing") que da como resultado un cambio de G a T en la posición +1 en la junta de maduración por corte y empalme del exón C de la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

40

(b) un ácido nucleico que tiene una mutación de maduración por corte y empalme que da como resultado un cambio de G a T en la posición +1 en la junta de maduración por corte y empalme del exón H de la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

45

(c) un ácido nucleico el cual tiene una inserción de 2 pares de bases de TT en la posición 98 del exón G de la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

50

(d) un ácido nucleico el cual tiene una mutación de desplazamiento de la pauta de lectura en el exón G de la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

55

(e) un ácido nucleico el cual tiene una mutación de sentido erróneo que da como resultado un cambio de T a G en la posición 46 del exón B en la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

60

(f) un ácido nucleico el cual tiene una mutación de sentido erróneo que da como resultado un cambio de G a A en la posición 28 del exón B en la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

65

(g) un ácido nucleico el cual tiene una mutación sin sentido que da como resultado un cambio de C a T en la posición 202 del exón H en la secuencia codificadora mostrada en la Figura 10 para dicho polipéptido supresor de tumor;

(h) un ácido nucleico en el que dicha mutación es una delección de AA en posiciones 16 y 17 del exón A mostrado en la Figura 10; y

70

(i) un ácido nucleico en el que dicha mutación es un cambio de C a T en la posición 53 en el exón B mostrado en la Figura 10.

ES 2 361 044 T3

25. El uso de

(a) la región codificadora de la secuencia de nucleótidos mostrada en SEC. ID. Nº: 2;

5 (b) una variante alélica de (a);

(c) un anticuerpo de cualquiera de las reivindicaciones 5 a 11; o

10 (d) un oligonucleótido de la reivindicación 20 ó 21

para determinar la expresión de una secuencia de nucleótidos TS10q23.3 en las células de una muestra.

15 26. El uso de la reivindicación 25, que comprende el ensayo de un ácido nucleicos TS10q23.3 o polipéptido en la muestra.

27. El uso de la reivindicación 26 que comprende la cuantificación de la cantidad de ARNm de TS10q23.3.

20 28. Una composición farmacéutica que comprende una construcción de expresión que comprende un ácido nucleico de la reivindicación 14, un polipéptido de la reivindicación 1 o un anticuerpo de la reivindicación 5.

29. El uso de un polipéptido de la reivindicación 1 para la producción de un medicamento para el tratamiento del cáncer.

25 30. Un polipéptido de la reivindicación 1 para el uso en el tratamiento de cáncer.

31. El uso de un ácido nucleico de cualquiera de las reivindicaciones 14 a 19 para la producción de un medicamento para el tratamiento de cáncer, en el que dicho ácido nucleico es capaz de expresar el polipéptido de la reivindicación 1 en las células tumorales asociadas con dicho cáncer.

30 32. Un ácido nucleico de cualquiera de las reivindicaciones 14 a 19 para usar en el tratamiento de cáncer, en el que dicho ácido nucleico es capaz de expresar el polipéptido de la reivindicación 1 en las células tumorales asociadas con dicho cáncer.

35 33. Un mamífero no humano transgénico en el cual ambas copias del gen codificador del supresor de tumor TS10q23.3 de la reivindicación 1 están interrumpidas o reemplazadas con otro gen.

40

45

50

55

60

65

Localización del Loci del Supresor de Tumor Candidato en el Cromosoma 10

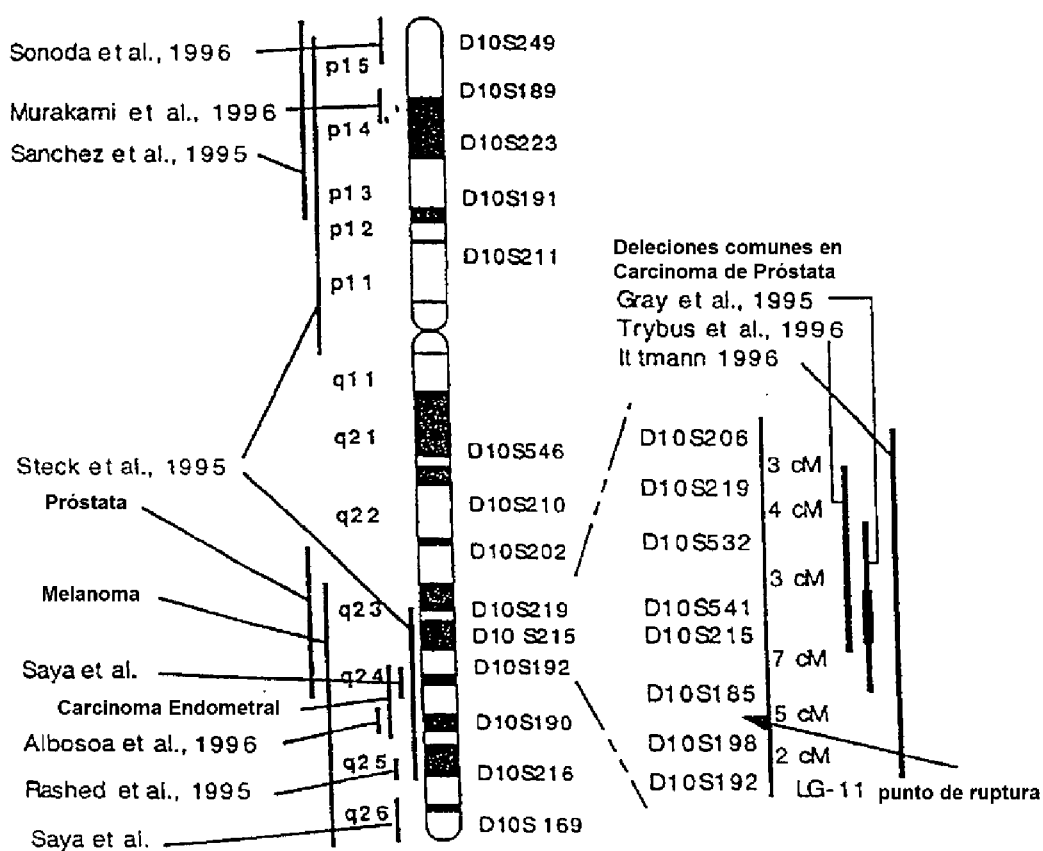


FIG. 1

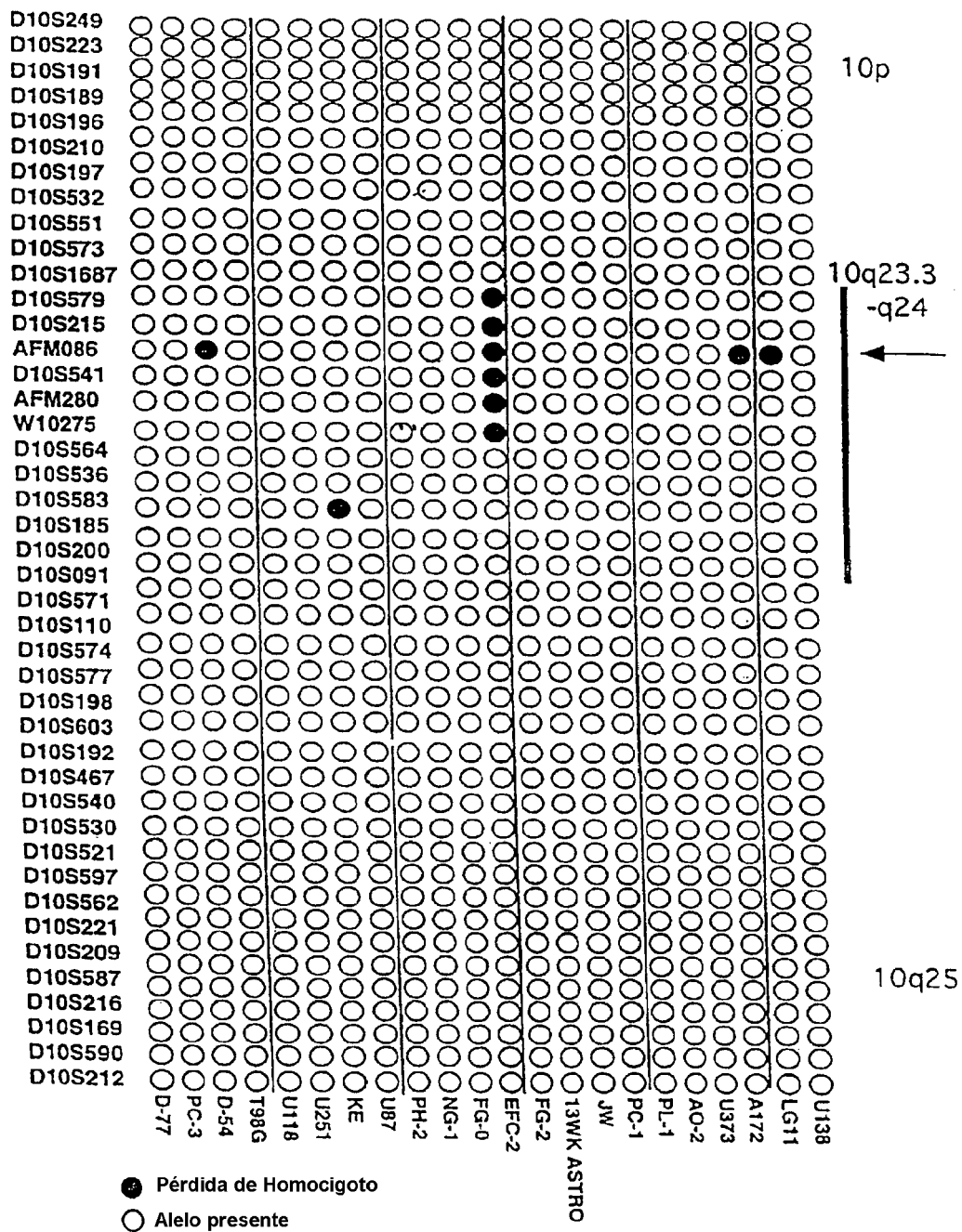


FIG. 2

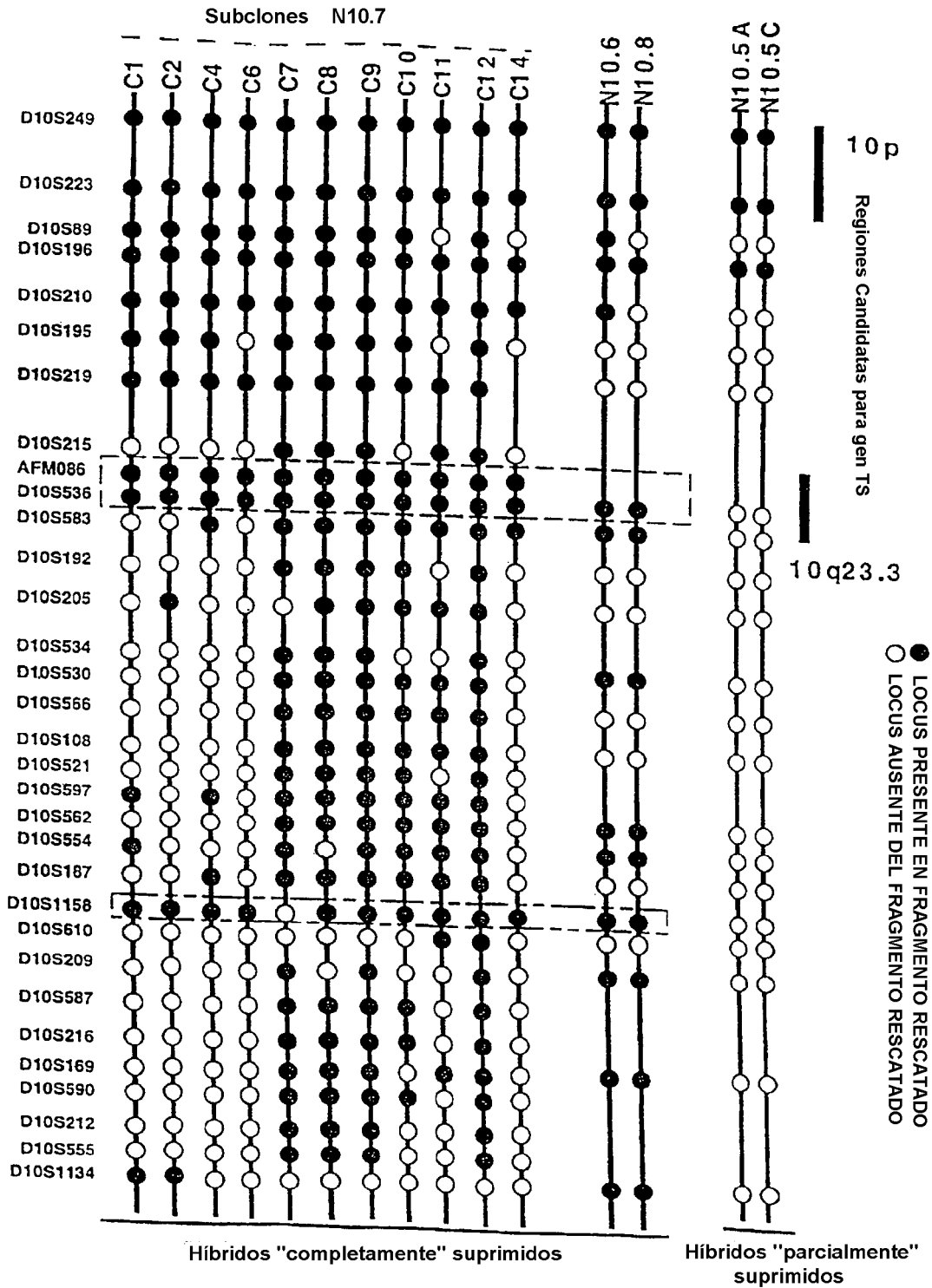


FIG. 3

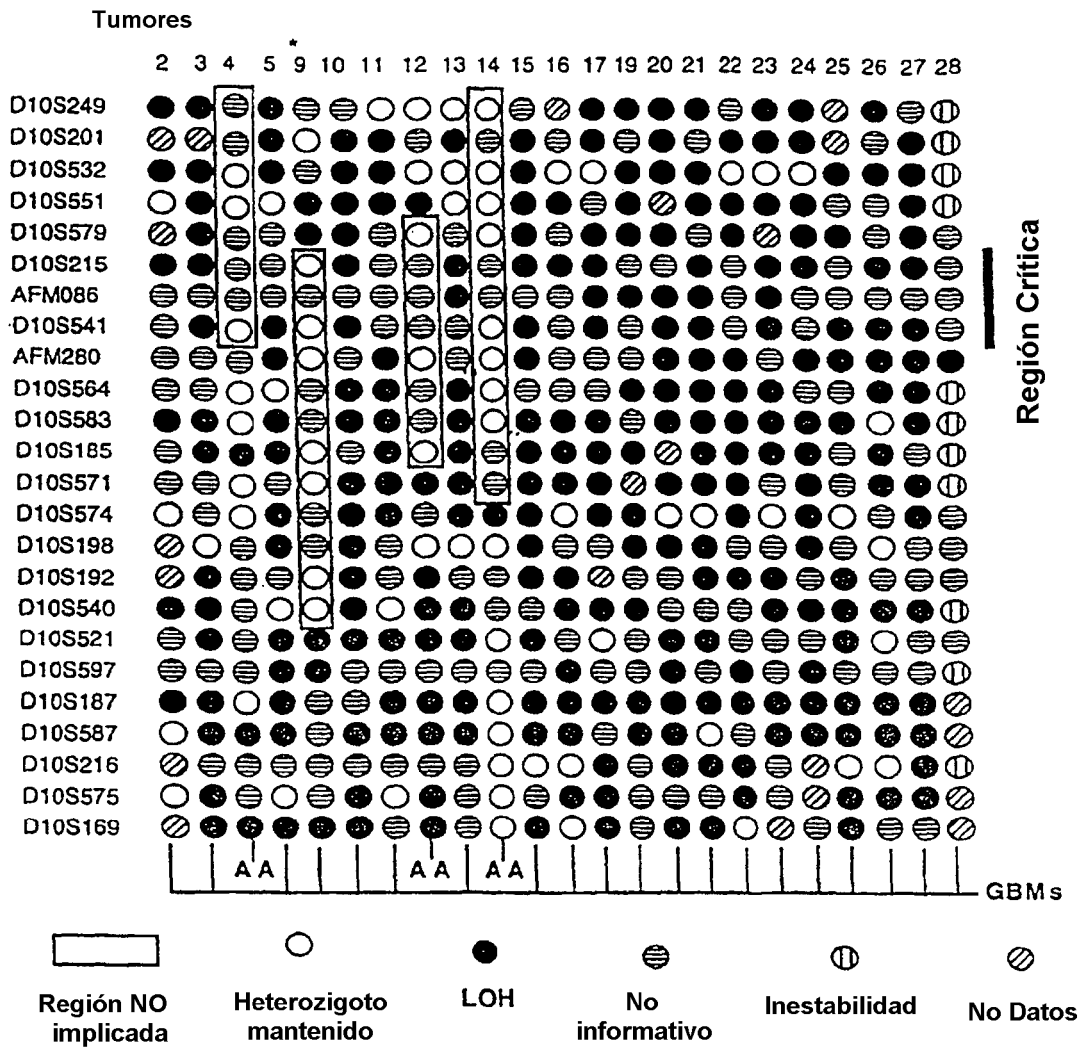


FIG. 4

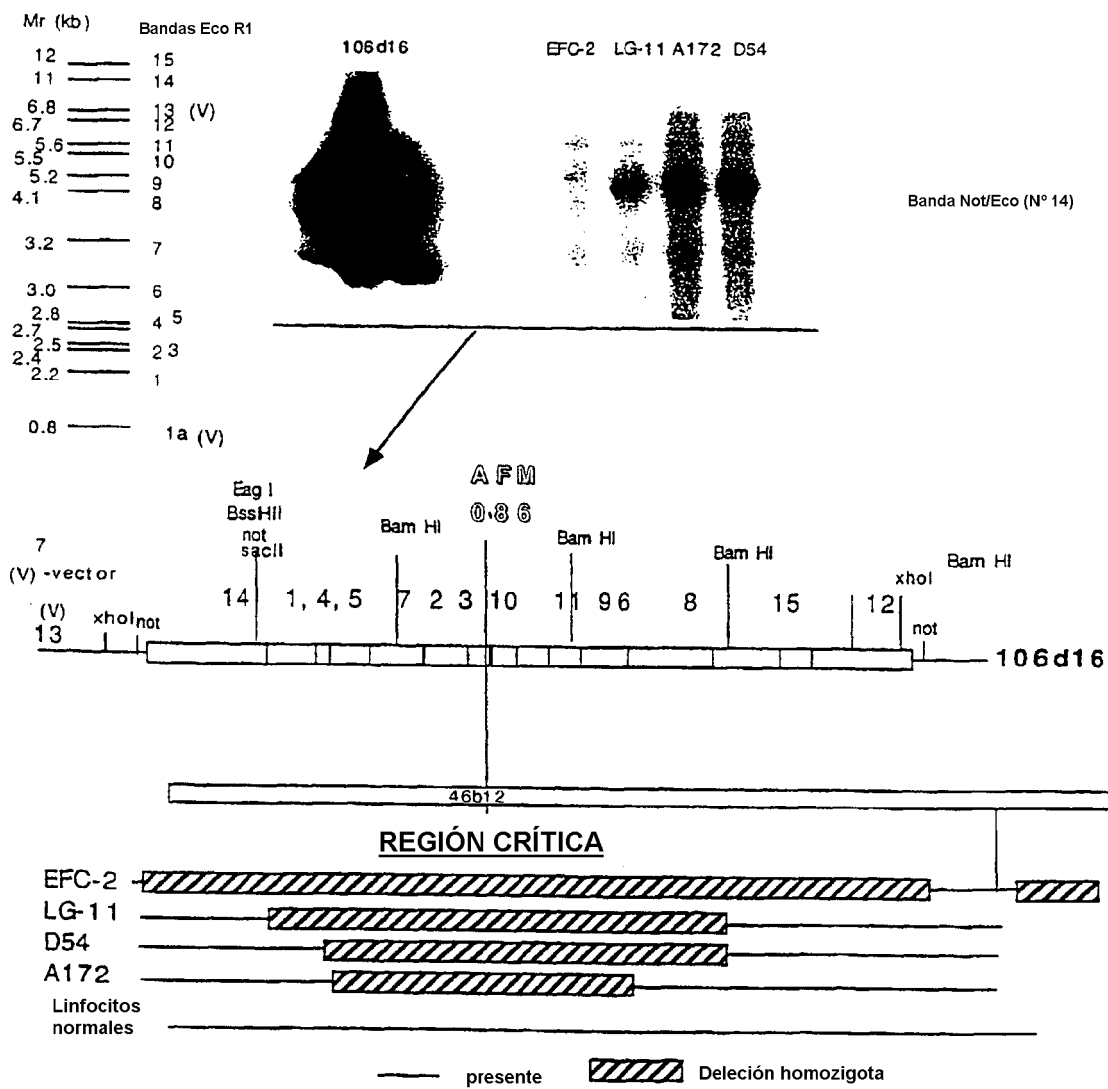


FIG. 5

1 M T A I I K E I V S R N K R R Y Q E D G F D L D L T Y I Y P
 31 N I I A M G F P A E R L E C V Y R N N A I D D V V R F L D S K
 61 H K N H Y K I Y N L C A E R H Y D T A K F D N C S R V D F A Q N H V F A
 91 E D H N P P Q L E L I K P F C E D L L H R G R Y V Y Y Y S G G T Y
 121 A I H C K A G K G R T R D K K G V T I P S Q R R Y V Y Y S G G T Y
 151 A L D F Y G E V R T R V A L L F H K M S G P T R R E D K F M Y
 181 L L K N H L D Y R P V A L L F H K M S G P T R R E D K F M Y
 211 C N P Q F V V C Q L K V K I Y S S F H K Q N K M L K K D K M Y
 241 F E F P P Q P L P V C G D I K V E F F H K Q N K M L K K D K M Y
 271 F H F W V N T F F I F G P E E T S E K V E N G S L C D D Q E I
 301 D S I C S I E R A D N D K E Y L V L T L T V K E N D L D K A N K S
 331 D K A N R Y F S P N F K V K L Y F T K Y R Y S D T P S S D P E A S
 351 S S T S V T P D V S D N E P D H Y R Y S D T P S S D P E A S
 391 P F D E D Q H T Q I T K V

FIG. 7

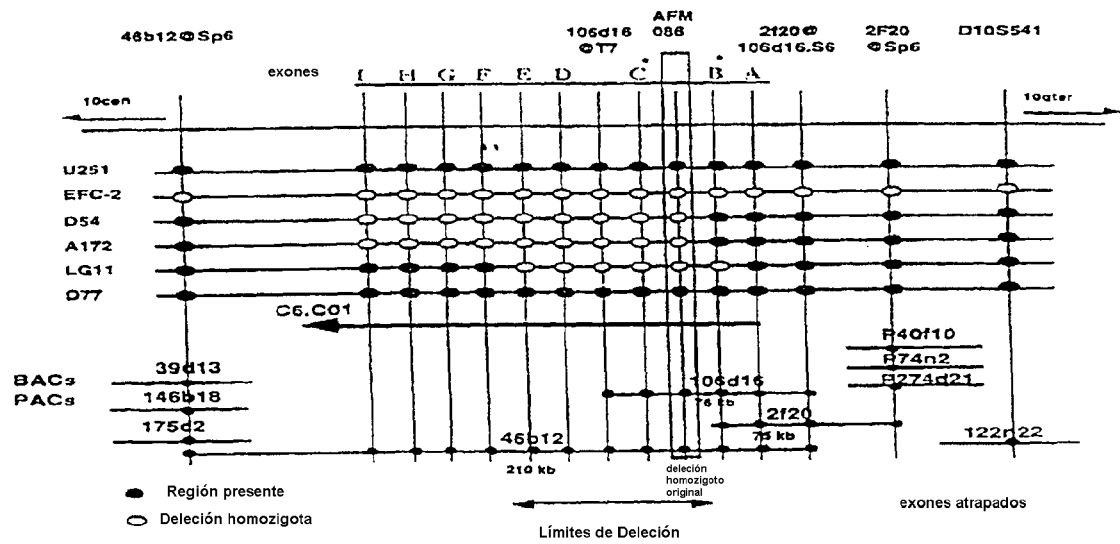


FIG. 8

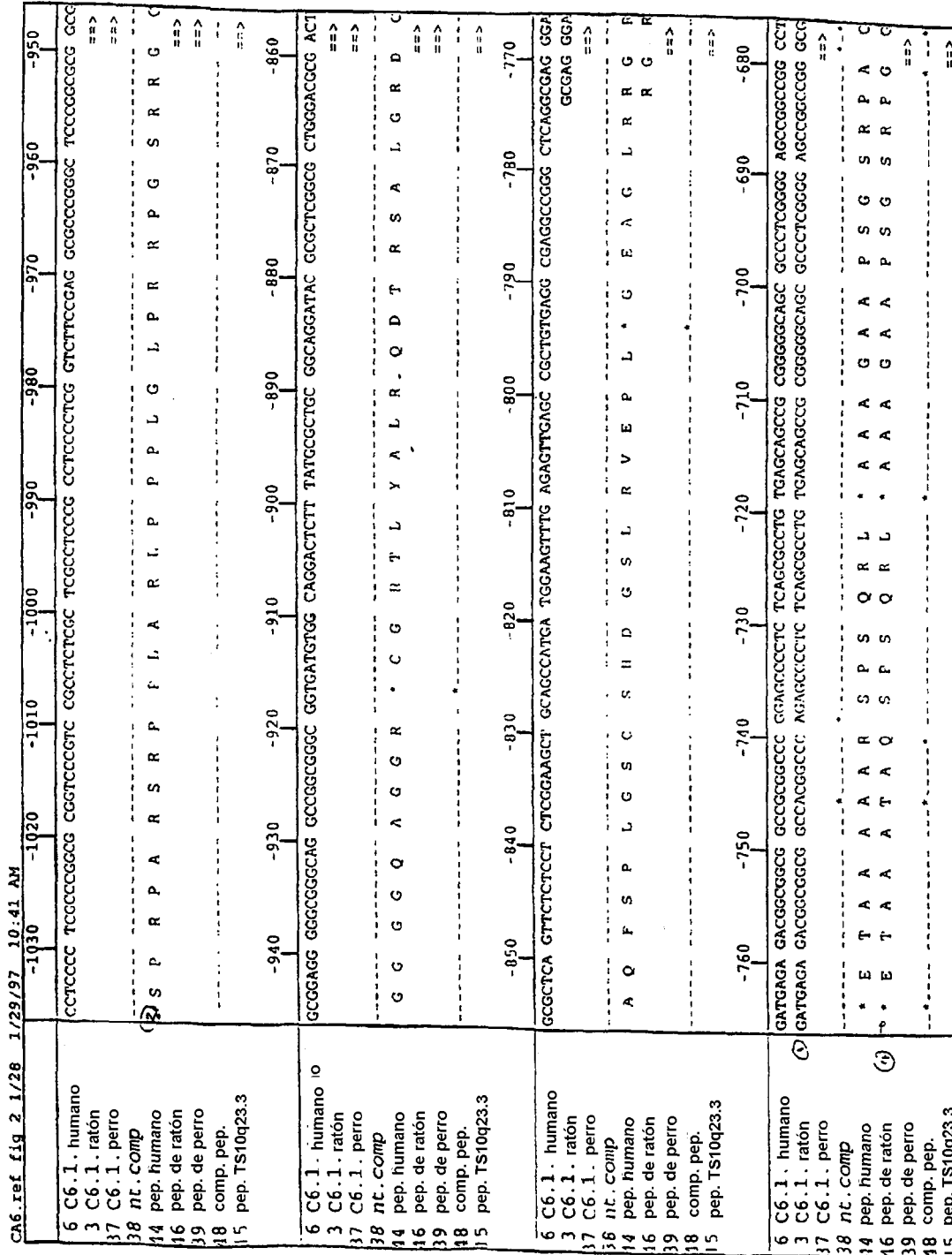


FIG. 9

CA6.ref fig 2 1/28 1/29/97 10:41 AM	-670	-660	-650	-640	-630	-620	-610	-600	-590
6 C6.1. humano	GGGGGG	CGGCAGCGGC	GGCGTTTCTC	GCCTCCTCTT	CGTCITTTCT	AACCGTGGAG	CCCTCTCTC	GGCTTCTCT	GAAGGGAAG
3 C6.1. ratón	GGGGGG	CGGCAGCGGC	GGCGGGCCTC	GCCTCCTCTT	CGTCITTTCT	AACCGGGCAG	CTTCTGACCA	G-CTT---CG	GAGAGAGACG
37 C6.1. perro									==>
38 nt.comp									==>
44 pep.humano	G G G S G	G V S R L	L L F V F	S N R A A	S S S A S	S S A S P	E R E G ?		
16 pep.de ratón	G G G S G	G G P R L	L L V V C	S N R A A	S S E Q ?	? ? ? E R	D G G ?		
39 pep.de perro									==>
48 comp.pep.									==>
15 pep. TS10q23.3									==>
6 C6.1. humano	-580	-570	-560	-550	-540	-530	-520	-510	-500
3 C6.1. ratón	---GAAG	CGTGGGCTC	GGCGGGAGC	CGGTGAGGC	CGCGGGGG	ACCTCCCGCT	CCTGGAGCGG	GGGGGAGAG	CGG
37 C6.1. perro	GAAGAAG	CGTGGGCTC	GAGCGGGAGC	CGGCCAGGC	T-----	CGGGGCTGC	ACCTCCCGCT	CCTGGAGCGG	GGGGGAGAG
38 nt.comp									==>
44 pep.humano	? S R G L	G R E P A	E A R R R	R R R R R	R R R R R	L P L L P	L L E R G	E E A A	
16 pep.de ratón	R S R G L	E R E P A	Q A ? ? ?	R R L H L	L P L L E	R R G G E	E A A		
39 pep.de perro									==>
48 comp.pep.									==>
15 pep. TS10q23.3									==>
6 C6.1. humano	-490	-480	-470	-460	-450	-440	-430	-420	-410
3 C6.1. ratón	CGGGGC	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG
37 C6.1. perro	CGGGGC	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG
38 nt.comp									==>
44 pep.humano	A A A A A	A A A P G	R G S E S	P V T I S	R A G N A	G E L V			
16 pep.de ratón	A A ? ? A	? ? ? P G	R G S E S	P V T I A	R A G N A	G E L L			
39 pep.de perro									==>
48 comp.pep.									==>
15 pep. TS10q23.3									==>
6 C6.1. humano	-400	-390	-380	-370	-360	-350	-340	-330	-320
3 C6.1. ratón	TCTCTCC	CCTTCTACTG	CCTCCACAC	GGCGGGGG	GGCGGGGC	ATCCAGGGAC	CGGGCCCGT	TTTAAACCTC	CGGTCCCGC
37 C6.1. perro	TCTCTCC	CCTTCTCTG	CCTCCACAC	GGCGGGGG	GGCGGGGC	GTCCAGGGAC	CGGGCCCGT	GTTAAGCCTC	CGGTCCCGC
38 nt.comp									==>
44 pep.humano	S P L L L	P P T R R	R R R R R	R R R R R	R R R R R	I Q G P G	P V L N L	P S A A A	
16 pep.de ratón	S P L L L	P P T R R	R R R R R	R R R R R	R R R R R	V O G P G	P V L S L	P S A A A	
39 pep.de perro									==>
48 comp.pep.									==>
15 pep. TS10q23.3									==>

FIG. 9 (cont.)

CA6.ref fig 2 1/28 1/29/97 10:41 AM	-310	-300	-290	-280	-270	-260	-250	-240	-230
6 C6.1. humano	CCGCACC	CCCGTGGCC	CGGCTCCGG	AGGCCCGGG	CGGAGGCAG	CGTTCGGAG	ATTATTCTGTC	TTCTCCCCAT	TCGGTGGCG
3 C6.1. ratón	CCGCACC	CCCGTGGCC	CGGCTCCGG	AGGCCCGGG	CGGAGGCAG	CGTTCGGAG	ATTATTCTGTC	TTCTCCCCAT	TCGGTGGCG
37 C6.1. perro	CCGCACC	CCCGTGGCC	CGGCTCCGG	AGGCCCGGG	CGGAGGCAG	CGTTCGGAG	ATTATTCTGTC	TTCTCCCCAT	TCGGTGGCG
38 nt.comp	==>	==>	==>	==>	==>	==>	==>	==>	==>
44 pep.humano	A P V A R A P E A A G G S R S E D Y S S S P H S A A A	A P P L A R A P E A A G G S R C E D Y P S S P H S A A S							
46 pep.ratón									
39 pep.perro									
48 comp.pep.									
15 pep.TS10q23.3									
6 C6.1. humano	-220	-210	-200	-190	-180	-170	-160	-150	-140
3 C6.1. ratón	CTGCCAG	GCCTCTGGCT	GCTGAGGAGA	AGCAGGCCA	GTCGCTGCAA	CCATCCAGCA	GCCGCCGAG	CAGCCATTAC	CCGGCTGGG
37 C6.1. perro	CTGCCAG	GCCTCTGGCT	GCTGAGGAGA	AGCAGGCCA	GTCGCTGCAA	CCATCCAGCA	GCCGCCGAG	CAGCCATTAC	CCGGCTGGG
38 nt.comp	==>	==>	==>	==>	==>	==>	==>	==>	==>
44 pep.humano	A R P L A A E E K Q A Q S L Q P S S S R R S S H Y P A A V Q	A R P L A A E E K Q A Q S L Q P S S S R R S S H Y P A A V Q							
46 pep.ratón									
39 pep.perro									
48 comp.pep.									
15 pep.TS10q23.3									
6 C6.1. humano	-130	-120	-110	-100	-90	-80	-70	-60	-50
3 C6.1. ratón	AGAGCCA	AGCGGGCA	GAGCAGGG	CATCAGTAC	CGCCAATCC	AGAGCCATTT	CCATCTGCA	GAAGAAGCC	CGCCACCAGC
37 C6.1. perro	AGAGCCA	AGCGGGCA	GAGCAGGG	CATCAGTAC	CGCCAATCC	AGAGCCATTT	CCATCTGCA	GAAGAAGCC	CGCCACCAGC
38 nt.comp	==>	==>	==>	==>	==>	==>	==>	==>	==>
44 pep.humano	S Q A A A E R G A S A T A K S R A I S I L Q K K P R H Q Q I	G Q A A A E R G A S A T A K S R A I S I L Q K K P R H Q Q L							
46 pep.ratón									
39 pep.perro									
48 comp.pep.									
15 pep.TS10q23.3									
6 C6.1. humano	-40	-30	-20	-10	0	10	20	30	40
3 C6.1. ratón	TTCTGCC	ATCTCTCTCC	TCCTTTTCT	TCAGCCACAG	GCTCCACAG	ATGACAGCA	TCATCAAGA	GATCGTTAGC	AGAAACAAA
37 C6.1. perro	TTCTGCC	ATCTCTCTCC	TCCTTTTCT	TCAGCCACAG	GCTCCACAG	ATGACAGCA	TCATCAAGA	GATCGTTAGC	AGAAACAAA
38 nt.comp	==>	==>	==>	==>	==>	==>	==>	==>	==>
44 pep.humano	L P S L S S F F F S H R L P D M T A I I K E I V S R N K R R	L P S L S S F F F S H R L P D M T A I I K E I V S R N K R R							
46 pep.ratón									
39 pep.perro									
48 comp.pep.									
15 pep.TS10q23.3									

FIG. 9 (cont.)

	50	60	70	80	90	100	110	120	130
6 C6.1. humano	GATATCA	AGAGGATGCA	TTCGACTTAG	ACTTGACCTA	TATTTATCCA	AACATTTATG	CTATGGGATT	TCCTGCAGAA	AGACTTGAAG
3 C6.1. ratón	GATATCA	AGAGGATGCA	TTCGACTTAG	ACTTGACCTA	TATTTATCCA	AATATTTATG	CTATGGGATT	TCCTGCAGAA	AGACTTGAAG
37 C6.1. perro	GCTACCA	GGAGGATGG	TTCGACTTGG	ACTTGACCTA	TATTTATCCC	ACATTTATG	CTATGGGGTT	TCCTGCAGAA	AGACTTGAAG
38 nt.comp	Y Q	E D G	F D L D	L T Y	I Y P	N I I A	M G F	P A E	R L E G
44 pep.humano	Y Q	E D G	F D L D	L T Y	I Y P	N I I A	M G F	P A E	R L E G
46 pep.de ratón	Y Q	E D G	F D L D	L T Y	I Y P	N I I A	M G F	P A E	R L E G
39 pep.de perro	Y Q	E D G	F D L D	L T Y	I Y P	N I I A	M G F	P A E	R L E G
48 comp.pep.	Y Q	E D G	F D L D	L T Y	I Y P	N I I A	M G F	P A E	R L E G
15 pep.TS10q23.3	140	150	160	170	180	190	200	210	220
6 C6.1. humano	TATACAG	GAACAATATT	GATGATGAG	TAAGTTTTT	GGATTCARAG	CATAAAAACC	ATTACAAGAT	ATACAATCTT	TGTGCTGAAA
3 C6.1. ratón	TATACAG	GAACAATATT	GATGATGAG	TAAGTTTTT	GGATTCARAG	CATAAAAACC	ATTACAAGAT	ATACAATCTA	TGTGCTGAGA
37 C6.1. perro	TATACAG	GAACAATATT	GATGATGAG	TAAGTTTTT	GGATTCARAG	CATAAAAACC	ATTACAAGAT	ATACAATCTG	TGTGCTGAAA
38 nt.comp	Y R	N N I	D D V	V R F	L D S	K H K	N H Y	K I Y	N L C
44 pep.humano	Y R	N N I	D D V	V R F	L D S	K H K	N H Y	K I Y	N L C
46 pep.de ratón	Y R	N N I	D D V	V R F	L D S	K H K	N H Y	K I Y	N L C
39 pep.de perro	Y R	N N I	D D V	V R F	L D S	K H K	N H Y	K I Y	N L C
48 comp.pep.	Y R	N N I	D D V	V R F	L D S	K H K	N H Y	K I Y	N L C
15 pep.TS10q23.3	230	240	250	260	270	280	290	300	310
6 C6.1. humano	ATTATGA	CACCGCCAAA	TTTAATTGCA	GAGTTGCACA	ATATCCITTT	GAAGACCATA	ACCCACCACA	GCTAGAACCT	ATCAAAACCT
3 C6.1. ratón	ATTATGA	CACCGCCAAA	TTTAATTGCA	GAGTTGCACA	GTATCCITTT	GARGACCATA	ACCCACCACA	GCTAGAACCT	ATCAAAACCT
37 C6.1. perro	ATTATGA	TACCGCCAAA	TTTAACTGCA	GAGTTGCACA	GTATCCITTT	GAAGACCATA	ATCCACCACA	GCTAGAACCT	ATCAAAACCT
38 nt.comp	Y D	T A K	F N C	R V A	Q Y P	F E D	H N P	P Q L	E L I
44 pep.humano	Y D	T A K	F N C	R V A	Q Y P	F E D	H N P	P Q L	E L I
46 pep.de ratón	Y D	T A K	F N C	R V A	Q Y P	F E D	H N P	P Q L	E L I
39 pep.de perro	Y D	T A K	F N C	R V A	Q Y P	F E D	H N P	P Q L	E L I
48 comp.pep.	Y D	T A K	F N C	R V A	Q Y P	F E D	H N P	P Q L	E L I
15 pep.TS10q23.3	320	330	340	350	360	370	380	390	400
6 C6.1. humano	GTAAGA	TCTTGACCAA	TGGCTAAGTG	AGATGACAA	TCATGTTGCA	GCAATTCACCT	GTAAGCTGG	AAAGGACCA	ACTGGTGTAA
3 C6.1. ratón	GTAAGA	TCTTGACCAA	TGGCTAAGTG	AGATGACAA	TCATGTTGCA	GCAATTCACCT	GTAAGCTGG	AAAGGACCG	ACTGGTGTAA
37 C6.1. perro	GTAAGA	TCTTGACCAA	TGGCTAAGTG	AGATGACAA	TCATGTTGCA	GCAATTCACCT	GTAAGCTGG	AAAGGACCA	ACTGGTGTAA
38 nt.comp	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T
44 pep.humano	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T
46 pep.de ratón	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T
39 pep.de perro	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T
48 comp.pep.	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T
15 pep.TS10q23.3	E D	L D Q	W L S	E D D	N H V	A A I	H C K	A G K	G R T

FIG. 9 (cont.)

CA6.ref	Fig 2	1/28	1/29/97	10:41	AM	410	420	430	440	450	460	470	480	490
6	C6.1.	humano	TATGTCC	ATATTTATTA	CATCGGGCA	AATTTTTRAA	GGCACAAGAG	GCCTAGATT	TCATATGGGA	AGTAAGGACC	AGAGACAANA	AGC		
3	C6.1.	ratón	TTTGTGC	ATATTTATTTG	CATCGGGCA	AATTTTTRAA	GGCACAAGAG	GCCTAGATT	TTTATGGGGA	AGTAAGGACC	AGAGACAANA	AGC		
37	C6.1.	perro	TTTGTGC	ATATTTATTA	CATCGGGCA	AATTTTCTAAA	GGCACAAGAG	GCCTAGATT	TCATATGGGA	AGTAAGGACC	AGAGACAANA	AGC		
38	nt.	comp												
44	pep.	humano	C	A	Y	L	L	H	R	G	K	F	L	K
46	pep.	de ratón	C	A	Y	L	L	H	R	G	K	F	L	K
39	pep.	de perro	C	A	Y	L	L	H	R	G	K	F	L	K
48	comp.	pep.												
15	pep.	TS10q23.3	C	A	Y	L	L	H	R	G	K	F	L	K
			500	510	520	530	540	550	560	570	580			
6	C6.1.	humano	GAGTAAC	TATTTCCAGT	CAGAGGGCT	ATGTATATA	TTATAGCTAC	CTGTTAAGA	ATCATCTGGA	TTATAGACCA	GTGGCACTGT	TGT		
3	C6.1.	ratón	GAGTCAC	AATTTCCAGT	CAGAGGGCT	ATGTATATA	TTATAGCTAC	CTGTTAAGA	ATCATCTGGA	TTATAGACCA	GTGGCACTGT	TGT		
37	C6.1.	perro	GAGTAAC	TATTTCCAGT	CAGAGGGCT	ATGTATATA	TTATAGCTAC	CTGTTAAGA	ATCATCTGGA	TTATAGACCA	GTGGCACTGT	TGT		
38	nt.	comp												
44	pep.	humano	V	T	I	P	S	Q	R	R	Y	V	Y	Y
46	pep.	de ratón	V	T	I	P	S	Q	R	R	Y	V	Y	Y
39	pep.	de perro	V	T	I	P	S	Q	R	R	Y	V	Y	Y
48	comp.	pep.												
15	pep.	TS10q23.3	V	T	I	P	S	Q	R	R	Y	V	Y	Y
			590	600	610	620	630	640	650	660	670			
6	C6.1.	de humano	TTTCACAA	GATGATGTTT	GAACATATTC	CAATGTTTCAG	TGGGGAACT	TGCATCTC	AGTTTGTGT	CTGCCAGCTA	AAGGTGAAGA	TAT		
3	C6.1.	de ratón	TTTCACAA	GATGATGTTT	GAACATATTC	CAATGTTTCAG	TGGGGAACT	TGCATCTC	AGTTTGTGT	CTGCCAGCTA	AAGGTGAAGA	TAT		
37	C6.1.	de perro	TTTCACAA	GATGATGTTT	GAACATATTC	CAATGTTTCAG	TGGGGAACT	TGCATCTC	AGTTTGTGT	CTGCCAGCTA	AAGGTGAAGA	TCT		
38	nt.	comp												
44	pep.	humano	H	K	M	M	F	E	T	I	P	M	F	S
46	pep.	de ratón	H	K	M	M	F	E	T	I	P	M	F	S
39	pep.	de perro	H	K	M	M	F	E	T	I	P	M	F	S
48	comp.	pep.												
15	pep.	TS10q23.3	H	K	M	M	F	E	T	I	P	M	F	S
			680	690	700	710	720	730	740	750	760			
6	C6.1.	humano	ATTCCTC	CAATTCAGGA	CCCACAGGAC	GGGAGACAA	GTTTCATGTAC	TTTGAGTTCC	CTCAGCCGTT	ACCTGTGTGT	GGTGATATCA	AAG		
3	C6.1.	ratón	ATTCCTC	CAATTCAGGA	CCCACAGGAC	GGGAGACAA	GTTTCATGTAC	TTTGAGTTCC	CTCAGCCGTT	ACCTGTGTGT	GGTGATATCA	AAG		
37	C6.1.	perro	ATTCCTC	CAATTCAGGA	CCCACAGGAC	GGGAGACAA	GTTTCATGTAC	TTTGAGTTCC	CTCAGCCGTT	ACCTGTGTGT	GGTGATATCA	AAG		
38	nt.	comp												
44	pep.	humano	S	S	N	S	G	P	T	R	R	E	D	K
46	pep.	de ratón	S	S	N	S	G	P	T	R	R	E	D	K
39	pep.	de perro	S	S	N	S	G	P	T	R	R	E	D	K
48	comp.	pep.												
15	pep.	TS10q23.3	S	S	N	S	G	P	T	R	R	E	D	K

FIG. 9 (cont.)

	770	780	790	800	810	820	830	840	850
6 C6.1. humano	TAGAGTT	CTTCACAAA	CAGAACAGA	TGCTAAAAA	GGACAAAATG	TTTCACTTTT	GGTAAATAC	ATTCTTATA	CCAGGACCAG AGG
3 C6.1. ratón	TAGAGTT	CTTCACAAA	CAGAACAGA	TGCTAAAAA	GGACAAAATG	TTTCACTTTT	GGTAAATAC	ATTCTTATA	CCAGGACCAG AGG
37 C6.1. perro	TAGAGTT	CTTCACAAA	CAGAACAGA	TGCTAAAAA	GGACAAAATG	TTTCACTTTT	GGTAAATAC	ATTCTTATA	CCAGGACCAG AGG
38 nt.comp									
44 pep.humano	E F F H K	Q N K M	L K K	D K M	F H F	W V N T	F F I	P G P	E E
46 pep.de ratón	E F F H K	Q N K M	L K K	D K M	F H F	W V N T	F F I	P G P	E E
39 pep.de perro	E F F H K	Q N K M	L K K	D K M	F H F	W V N T	F F I	P G P	E E
48 comp.pep.									
15 pep.TS10q23.3	E F F H K	Q N K M	L K K	D K M	F H F	W V N T	F F I	P G P	E E
	860	870	880	890	900	910	920	930	940
6 C6.1. humano	AAACCTC	AGAAAAGTA	GAAATGGAA	GTCCTATGTA	TCAGAAATC	GATAGCATTT	GCAGTATAGA	GGGTGCAGAT	AATGACAAAGG AAT
3 C6.1. ratón	AAACCTC	AGAAAAGTG	GAAATGGAA	GTCCTATGTA	TCAGAAATC	GATAGCATTT	GCAGTATAGA	GGGTGCAGAT	AATGACAAAGG AGT
37 C6.1. perro	AAACCTC	AGAAAAGTA	GAAATGGAA	GTCCTATGTA	TCAGAAATC	GATAGCATTT	GCAGTATAGA	GGGTGCAGAT	AATGACAAAGG AAT
38 nt.comp									
44 pep.humano	T S E K V	E N G S	L C D	Q E I	D S I	C S I	E R A	D N D	K E Y
46 pep.de ratón	T S E K V	E N G S	L C D	Q E I	D S I	C S I	E R A	D N D	K E Y
39 pep.de perro	T S E K V	E N G S	L C D	Q E I	D S I	C S I	E R A	D N D	K E Y
48 comp.pep.									
15 pep.TS10q23.3	T S E K V	E N G S	L C D	Q E I	D S I	C S I	E R A	D N D	K E Y
	950	960	970	980	990	1000	1010	1020	1030
6 C6.1. humano	ATCTAGT	ACTCACTTTA	ACAAAATG	ATCTGACAA	AGCAATATA	GACAAAGCCA	ACCGATACTT	TTCTCCAAAT	TTTAAAGTGA AGC
3 C6.1. ratón	ATCTTGT	ACTCACCTTA	ACAAAATG	ATCTGACAA	AGCAATATA	GACAAAGCCA	ACCGATACTT	TTCTCCAAAT	TTTAAAGTGA AAC
37 C6.1. perro	ATCTAGT	ACTCACTTTA	ACAAAATG	ATCTGACAA	AGCAATATA	GACAAAGCCA	ACCGATACTT	TTCTCCAAAT	TTTAAAGTGA AGC
38 nt.comp									
44 pep.humano	L V L T L	T K N D	L D K	A N K	D K A	N R Y	F S P	N F K	V K L
46 pep.de ratón	L V L T L	T K N D	L D K	A N K	D K A	N R Y	F S P	N F K	V K L
39 pep.de perro	L V L T L	T K N D	L D K	A N K	D K A	N R Y	F S P	N F K	V K L
48 comp.pep.									
15 pep.TS10q23.3	L V L T L	T K N D	L D K	A N K	D K A	N R Y	F S P	N F K	V K L
	1040	1050	1060	1070	1080	1090	1100	1110	1120
6 C6.1. humano	TGTACTT	CACAAAACA	GTAGAGGAGC	CGTCAATCC	AGAGGCTAGC	AGTTCAACTT	CTGTACACC	AGATGTTAGT	GACAAATGAC CTG
3 C6.1. ratón	TATACTT	TACAAAACA	GTAGAGGAGC	CATCAATCC	AGAGGCTAGC	AGTTCAACTT	CTGTACTCC	AGATGTTAGT	GACAAATGAC CTG
37 C6.1. perro	TGTACTT	CACAAAACCT	GTAGAGGAGC	CATCAAAACC	GGAGGCTAGC	AGTTCAACTT	CTGTGAGCC	AGATGTTAGT	GACAAATGAC CTG
38 nt.comp									
44 pep.humano	Y F T K T	V E E P	S N P	E A S	S S T	S V T	P D V	S D N	E P D
46 pep.de ratón	Y F T K T	V E E P	S N P	E A S	S S T	S V T	P D V	S D N	E P D
39 pep.de perro	Y F T K T	V E E P	S N P	E A S	S S T	S V T	P D V	S D N	E P D
48 comp.pep.									
15 pep.TS10q23.3	Y F T K T	V E E P	S N P	E A S	S S T	S V T	P D V	S D N	E P D

FIG. 9 (cont.)

CA6.ref	fig	2	1/28	1/29/97	10:41	AM	1130	1140	1150	1160	1170	1180	1190	1200	1210
6	C6.1.	humano													
3	C6.1.	ratón													
37	C6.1.	perro													
38	nt.	comp													
44	pep.	humano													
46	pep.	de ratón													
39	pep.	de perro													
48	comp.	pep.													
15	pep.	TS10q23.3													
<p>ATCATTA TAGATAATCT GACACCCTG ACTCTGATCC AGAAGATGAA CCTTTTGATG AGATTCAGCA TACACAAATT ACAAAAGTCT GAA ATCATTA TAGATAATCT GACACCCTG ACTCTGATCC AGAAGATGAA CCTTTTGATG AGATTCAGCA TACACAAATT ACAAAAGTCT GA ATCATTA TAGATAATCT GACACCCTG ACTCTGATCC AGAAGATGAA CCTTTTGATG AGA</p>															
<p>H Y R Y S D T T D S D P E N E P F D E D Q H T Q I T K V * H Y R Y S D T T D S D P E N E P F D E D Q H T Q I T K V * H Y R Y S D T T D S D P E N E P F D E D Q H T Q I T K V * H Y R Y S D T T D S D P E N E P F D E D Q H T Q I T K V *</p>															
1220	1230	1240	1250	1260	1270	1280	1290	1300							
<p>TTTTTT TTATCAAGAG GGATPAAAACA CCATGAAAT AACCTTGAAT AACCTGAAAA TGGACCTTTT TTTTTTAAAT GGCAATAGGA CNT <== <== <== <== <== <==</p>															
1310	1320	1330	1340	1350	1360	1370	1380	1390							
<p>TGTTGCA GATTACCAGT TATAGNACA ATTCTCTTTT CCTGACCAAT CTGTGTTTAC CCTATACATC CACAGGGTTT TGACACTTGT TGT <== <== <== <== <== <==</p>															
1400	1410	1420	1430	1440	1450	1460	1470	1480							
<p>CCAGTTG AAAAAAGTT GTGTAGCTGT GTCATGTATA TACCCTTTTG TGTCAAAGG ACAATTTAAA TTCAAATTAGG ATTAATAAAG ATG <== <== <== <== <== <==</p>															
6	C6.1.	humano													
3	C6.1.	ratón													
37	C6.1.	perro													
38	nt.	comp													
44	pep.	humano													
46	pep.	de ratón													
39	pep.	de perro													
48	comp.	pep.													
15	pep.	TS10q23.3													

FIG. 9 (cont.)

CAG.ref fig 2 1/28 1/29/97 10:41 AM		1850	1860	1870	1880	1890	1900	1910	1920	1930
6 C6.1. humano		GCCGCTG	TCACGTGCTG	TTGTTTGGC	ATTTTTTTTT	AAAGCATATTT	GGTGCTAGAA	AGGCAGCTA	ANGGAAGTGA	ATCTGTATTG
3 C6.1. ratón		<==								
37 C6.1. perro		<==								
38 nt. comp		<==								
44 pep. humano		<==								
46 pep. de ratón		<==								
39 pep. de perro		<==								
48 comp. pep.		<==								
15 pep. TS10q23.3		<==								
6 C6.1. humano		1940	1950	1960	1970	1980	1990	2000	2010	2020
3 C6.1. ratón		TACAGGA	ATGMACTTC	TGCACATCT	TAAATCCAC	AAATGAAGG	ATATATAAAT	ATGTGATAG	GTAAGAACA	CAGCAACAAT
37 C6.1. perro		<==								
38 nt. comp		<==								
44 pep. humano		<==								
46 pep. de ratón		<==								
39 pep. de perro		<==								
48 comp. pep.		<==								
15 pep. TS10q23.3		<==								
6 C6.1. humano		2030	2040	2050	2060	2070	2080	2090	2100	2110
3 C6.1. ratón		TTAACCA	TATAAATGTG	GAGGCTATCA	ACAAGAATG	GCCITTCAAAC	ATTATATAAA	TTGACANTGA	TTTTATTAAAT	ATGTTTTTCTC
37 C6.1. perro		<==								
38 nt. comp		<==								
44 pep. humano		<==								
46 pep. de ratón		<==								
39 pep. de perro		<==								
48 comp. pep.		<==								
15 pep. TS10q23.3		<==								
6 C6.1. humano		2120	2130	2140	2150	2160	2170	2180	2190	2200
3 C6.1. ratón		TGTAAAA	AAAAAA							
37 C6.1. perro		<==								
38 nt. comp		<==								
44 pep. humano		<==								
46 pep. de ratón		<==								
39 pep. de perro		<==								
48 comp. pep.		<==								
15 pep. TS10q23.3		<==								

FIG. 9 (cont.)

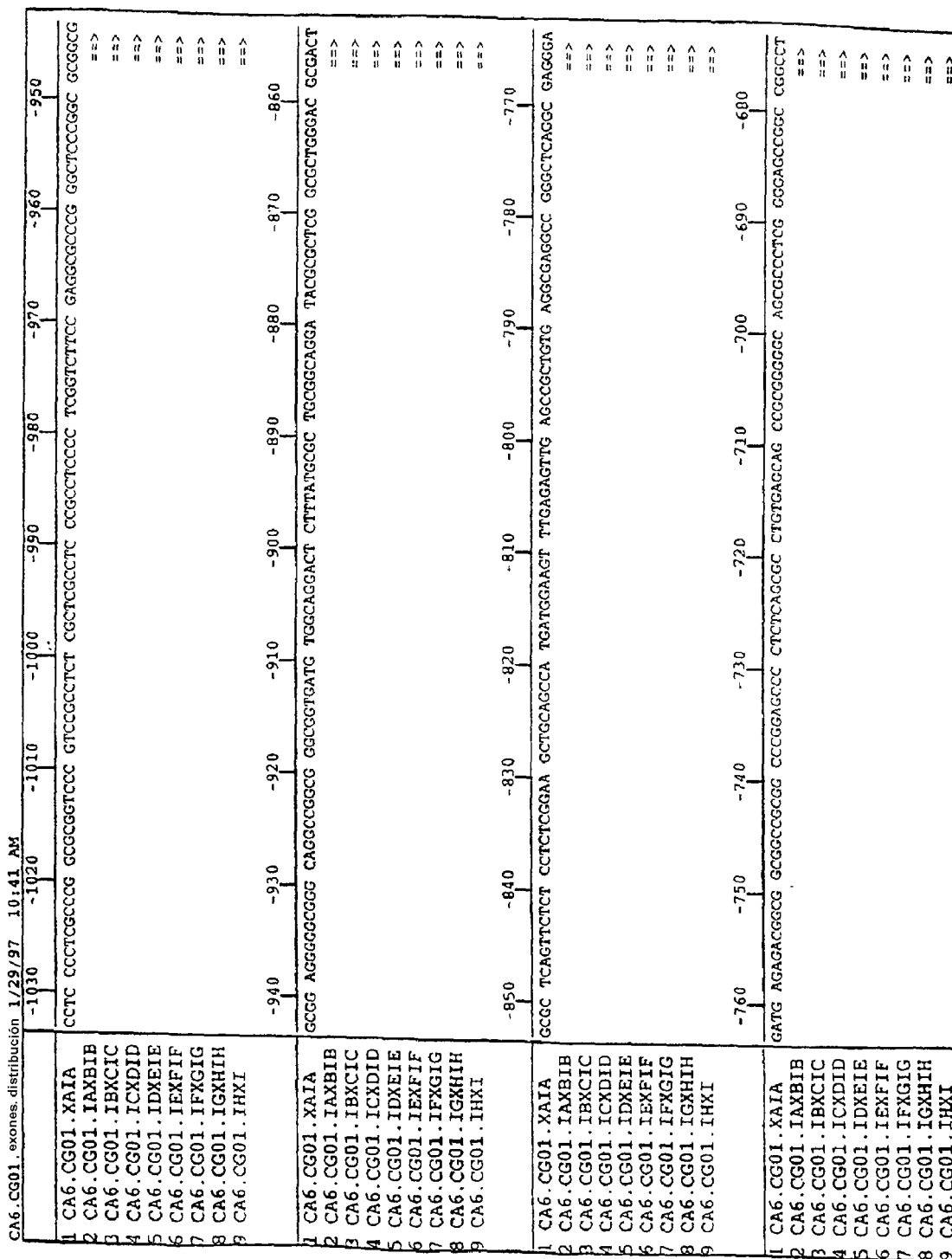


Fig.10

CA6..CG01..exones, distribución 1/29/97 10:41 AM		-670	-660	-650	-640	-630	-620	-610	-600	-590	
1	CA6..CG01..XAIA	GGCG	GGCGGGCAGC	GGCGGGTTT	CTCGCTCCT	CTTCGGTCTTT	TCTAACCGTG	CAGCCTCTTC	CTCGGCTTCT	CCTGAAAAGGG	AAGGTG
2	CA6..CG01..IAXBIB										
3	CA6..CG01..IBXCIC										
4	CA6..CG01..ICXDID										
5	CA6..CG01..IDXEIE										
6	CA6..CG01..IEXFIF										
7	CA6..CG01..IFXGIG										
8	CA6..CG01..IGXHIH										
9	CA6..CG01..IHXI										
1	CA6..CG01..XAIA										
2	CA6..CG01..IAXBIB										
3	CA6..CG01..IBXCIC										
4	CA6..CG01..ICXDID										
5	CA6..CG01..IDXEIE										
6	CA6..CG01..IEXFIF										
7	CA6..CG01..IFXGIG										
8	CA6..CG01..IGXHIH										
9	CA6..CG01..IHXI										
1	CA6..CG01..XAIA										
2	CA6..CG01..IAXBIB										
3	CA6..CG01..IBXCIC										
4	CA6..CG01..ICXDID										
5	CA6..CG01..IDXEIE										
6	CA6..CG01..IEXFIF										
7	CA6..CG01..IFXGIG										
8	CA6..CG01..IGXHIH										
9	CA6..CG01..IHXI										
1	CA6..CG01..XAIA										
2	CA6..CG01..IAXBIB										
3	CA6..CG01..IBXCIC										
4	CA6..CG01..ICXDID										
5	CA6..CG01..IDXEIE										
6	CA6..CG01..IEXFIF										
7	CA6..CG01..IFXGIG										
8	CA6..CG01..IGXHIH										
9	CA6..CG01..IHXI										

FIG. 10 (cont.)

CA6.CG01	exones, distribución	1/29/97	10:41 AM
1	CA6.CG01.XA1A	-310	-230
2	CA6.CG01.IAXB1B	-300	-240
3	CA6.CG01.IBXC1C	-290	-250
4	CA6.CG01.ICXD1D	-280	-260
5	CA6.CG01.IDXE1E	-270	-270
6	CA6.CG01.IEXF1F	-260	-280
7	CA6.CG01.IFXG1G	-250	-290
8	CA6.CG01.IGXH1H	-240	-300
9	CA6.CG01.IHXI	-230	-310
1	CA6.CG01.XA1A	-220	-140
2	CA6.CG01.IAXB1B	-210	-150
3	CA6.CG01.IBXC1C	-200	-160
4	CA6.CG01.ICXD1D	-190	-170
5	CA6.CG01.IDXE1E	-180	-180
6	CA6.CG01.IEXF1F	-170	-190
7	CA6.CG01.IFXG1G	-160	-200
8	CA6.CG01.IGXH1H	-150	-210
9	CA6.CG01.IHXI	-140	-220
1	CA6.CG01.XA1A	-130	-50
2	CA6.CG01.IAXB1B	-120	-60
3	CA6.CG01.IBXC1C	-110	-70
4	CA6.CG01.ICXD1D	-100	-80
5	CA6.CG01.IDXE1E	-90	-90
6	CA6.CG01.IEXF1F	-80	-100
7	CA6.CG01.IFXG1G	-70	-110
8	CA6.CG01.IGXH1H	-60	-120
9	CA6.CG01.IHXI	-50	-130
1	CA6.CG01.XA1A	-40	40
2	CA6.CG01.IAXB1B	-30	30
3	CA6.CG01.IBXC1C	-20	20
4	CA6.CG01.ICXD1D	-10	10
5	CA6.CG01.IDXE1E	0	0
6	CA6.CG01.IEXF1F	10	10
7	CA6.CG01.IFXG1G	20	20
8	CA6.CG01.IGXH1H	30	30
9	CA6.CG01.IHXI	40	40

FIG. 10 (cont.)

	50	60	70	80	90	100	110	120	130	
CA5.CG01.exones.distribucion 1/29/97 10:41 AM	50	60	70	80	90	100	110	120	130	
1 CA6.CG01.XA1A	ATCA	AGAGGATGGA	TTGCACTTAG	ACTTGACCTG	TATCCATTTC	TGCGGCTGCT	CCCTTTTACC	TTTCTGTAC	TCCTTTAGAA	CGTGGG
2 CA6.CG01.IAXB1B	TGAA	GGCGTATACA	GGAAACAATAT	TGATGATGTA	GTAAGTAAAG	aatgctttga	ttttctcattt	caaatattga	tgtttataatt	catggt
3 CA6.CG01.IBXC1C	taag	tatgtttttct	tatttgtatg	cttgcaaaata	lcttcaaaaa	caactatttaa	gtgaaagtta	tctgctttgt	agagtggagt	agagtt
4 CA6.CG01.ICXD1D	aggt	atgaatgtac	tgtaactagt	tgataaatt	aaaccgcata	gactgtatct	tactgtcata	acaataatga	gtcatccaga	ttatcg
5 CA6.CG01.IDXE1E	TATC	AAACCCCTTT	GTGAGATCT	GTGCAAAATG	CTAATGTAG	ATCAACAATCA	TGTTGACGCA	ATTCATGTA	AGCTGGAA	GGGACG
6 CA6.CG01.IEXF1F	ACCT	GTTAANGAAT	CATCTGGAT	ATAGACCTG	GGCATTG	TGATCAAGA	TGATGTTGA	ATGATTTCCA	ATGTTTCAAG	GGCGAA
7 CA6.CG01.IFXG1G	CAAT	TCAGGACCCA	CACGACGGGA	AGACAAGTTC	ATGTACTTTC	AGTTCCCTCA	CGCTTACCT	TGCTGTGGTG	ATATCAAAAT	AGAGTT
8 CA6.CG01.IGXH1H	AGAG	GAAACCTCAG	AAAAATGAGA	AAATGGAAT	CTATGTGTC	AGAANTCCA	TAGCAATTCG	ASPTATAGAGC	GTCAGATAA	TGACAA
9 CA6.CG01.IHX1	GGCT	AGCAGTTCAA	CTTCTGTAA	ACCAGATGTT	AGTGCATG	AACTGTATCA	TTATAGATAT	TCTGACACCA	CTGACTCTGA	TCCAGA
140	150	160	170	180	190	200	210	220		
1 CA6.CG01.XA1A	AGTA	GACGGATGCG	AAATGTCCG	TAGTTTGGGT	GACTATAACA	TTTTAACCCCTG	GTCAGGTTGC	TAGGTCATAT	ATTTTGTGTT	TTC
2 CA6.CG01.IAXB1B	gtgt	tttccatttag	aaaagatttc	taagccacag	aaaagataac	tttggatgt	aaactattat	tytaggtctc	tataatcatt	tttttg
3 CA6.CG01.IBXC1C	aaag	atacatttta	acagaattgt	atctctaaac	cgattaagtc	agaagctcca	agagactctt	tagatccatt	agaagtgta	gtgatg
4 CA6.CG01.ICXD1D	agtg	agatacatat	tatctttaa	aattatctt	aaaatttca	aaaattttaa	tttcaactgt	gtgttttag	aaaaagtatt	gcataa
5 CA6.CG01.IDXE1E	AACT	GGTAAATGA	TATGTGATA	TTATTATAT	CGGGCAAT	TTTTAAAGGC	ACAAGAGGCC	CTAGATTTCT	ATGGGAAAT	AAGGAC
6 CA6.CG01.IEXF1F	CTTG	CAGtaagtgc	tggaaattc	tcatctctc	atgatctga	acagtttct	taccatattc	tagaagtta	cataaaaatt	tagaaa
7 CA6.CG01.IFXG1G	CTTC	CACAAACAGA	ACAAAGTCT	AAAAAGTt	tqactttac	tttcaattgg	agaataatcc	aaaataaaga	cagattaaa	gctata
8 CA6.CG01.IGXH1H	GGAA	TATCTAGTAC	TACTTTTAC	AAAAATGAT	CTTGCACAAG	CAATAAAGA	CAAGCCCAAC	CGATACTTTT	TCCCAARTTT	TAAAGT
9 CA6.CG01.IHX1	GAAT	GAACCTTTTG	ATGACAGTCA	GCATACACAA	ATTACAAAAG	Tctgaatttt	tttttatcaa	gagggataaa	acacccatgaa	aatuaa
230	240	250	260	270	280	290	300	310		
<==										
1 CA6.CG01.XA1A	ctta	cgtacctaa	tggacttcag	ggggatacac	ttcattttag	agaactgac	cttatacatt	acataatcag	gtacttatgt	gataac
2 CA6.CG01.IAXB1B	aggt	aaacatttg	tggcacagat	tcaatgttact	tgatctgctt	tasaatgactt	ggcatctagc	ccataattga	gcccataacc	gtgtgg
3 CA6.CG01.IBXC1C	agct	attaatattg	tcaggaagac	taagtycag	catagactaa	gcaatccagga	aaatctctag	actaataata	gataaagag	agggtt
4 CA6.CG01.ICXD1D	CAGA	GACAAAAGG	taagtattct	tttgatgttt	tctctctct	ctctctggt	ctgagaattt	attggaatac	agatcttggg	ttctct
5 CA6.CG01.IDXE1E	agaa	atttaccaca	tttgaattt	atgcccagga	ctatatctt	gagycatttg	adcaaatba	ttagctttgt	tgttcaactc	attggg
6 CA6.CG01.IEXF1F	tttt	attttatgac	atgtaagga	ctataattg	ttttctatta	tttctctct	actctggcat	tgggtgagaca	ttataa	
7 CA6.CG01.IFXG1G	cagt	taacttaaac	atttctggg	ggttggtag	ttgtatgat	gtgatgtgac	tttaattcta	ggagcacagc	tgatgagaa	cttgcct
8 CA6.CG01.IGXH1H	cttg	aataaactga	aaatggacct	ttttttttt	aatggcaata	ggacattgtg	tcagattctc	agttataagga	acaattctct	tttctc
320	330	340	350	360	370	380	390	400		
<==										
1 CA6.CG01.XA1A	atct	cttggactcc	ataaaatgt	ggtccaccag	ttataacct	ggattccatt	acagttgat	ttttgtctta	tttccatggt	ggggat
2 CA6.CG01.IAXB1B	taat	tgaagtga	attcacagta	gagctctgt	taaacacta	atageatctt	ccatggaggt	atacttcaga	gtgaataata	ttttgt
3 CA6.CG01.IBXC1C	tacc	tactatttga	ggcagtttgt	ctaattagtaa	gcaatccacag	ggagaaagc	agaaactact	taactctct	gtgttgagga	atgaca
4 CA6.CG01.ICXD1D	tttt	tctttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttt
5 CA6.CG01.IDXE1E	ctaa	agaagcaba	agcaatgggt	tttaatgtag	tcgaagccaa	attatattta	tgaagaaat	attctgtgt	ataaccacc	aaatcc
6 CA6.CG01.IEXF1F	gggt	aaataatctt	ggttgaagga	aaagcccta	tggcaattga	acattagag	aaattttct	abcaaggtg	gttaactgag	agaaa
7 CA6.CG01.IFXG1G	tgac	agtttttaa	cttatgtatt	attccgaagc	agtgtttacg	taccagtaac	atgaaagttt	ctataaaaat	acccaatgta	caacag
8 CA6.CG01.IGXH1H	gacc	aactttgtt	tacctataac	atccacaggg	tttgacct	tggttccag	ttgaaaaag	glttgttagc	tgtgtcatgt	atatac
9 CA6.CG01.IHX1										

FIG. 10 (cont.)

CA6.CG01. exones. distribución 1/29/97 10:41 AM		410	420	430	440	450	460	470	480	490
1	CA6.CG01.XAIA	<==								
2	CA6.CG01.IAXBIB	tagg	cttaaaatcc	tagagtggat	ttattcágtt	aaatttacc	acactaagat	gtnatgact	gtnatgact	atltttatgt
3	CA6.CG01.IBXCIC	ttat	ctctgtcttc	tagagctcatt	gactgaaaaa	gctgttaggg	cattctctaa	ctgtacatca	ccraagttat	ttaaaaattgc
4	CA6.CG01.ICXDID	taaa	aggtacgaaa	ggatataac						
5	CA6.CG01.IDXEIE	NNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
6	CA6.CG01.IEXFIF	agcc	caattctga	ctapagtat	ggaagaacct	gtcccc				
7	CA6.CG01.IFXGIG	ttag	catggacca	atatttfaaa	aatlttttgt	ctataggtag	aaatgagatc	tgttctgtgg	tcctatgag	tgacacaaaac
8	CA6.CG01.IGXHHH	gtca	aaaaagctgc	atttttccct	ttcttaattc	tttgttgttt	gctgaaatct	ggggcaaaagg	tgcgggaggg	ggctaaatga
9	CA6.CG01.IHXI	cttt	tttgttcaaa	aggacaltta	aaattcaatt	aggatttaata	aggatggcac	ttccccgttt	tattccagtt	ttataaaaaag
		500	510	520	530	540	550	560	570	580
1	CA6.CG01.XAIA	<==								
2	CA6.CG01.IAXBIB	aatt	ttaaagttacc	actgtgcata	tgntaccaa	ctacctgaag	aantatttgg	ttggtacnag	anatataaa	aggaatcgcct
3	CA6.CG01.IBXCIC	aagt	ggcttgtctt	gtctangaca	ganttttraag	gactgcccac	ctgattgata	gaactagtgg	accittatett	taactttttg
4	CA6.CG01.ICXDID	<==								
5	CA6.CG01.IDXEIE	NNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
6	CA6.CG01.IEXFIF	<==								
7	CA6.CG01.IFXGIG	<==								
8	CA6.CG01.IGXHHH	tatg	aagtaggaat	gggagaggaa	agaaatagat	gggaactcag	tcatttggga	atgattcata	tggaaatgttt	ttactgtctt
9	CA6.CG01.IHXI	caga	ctgatgtgta	tacgtaggaa	ttttttccct	ttgtttctg	tcaccaactg	aagtggctaa	agagctttgt	gatatactgg
		590	600	610	620	630	640	650	660	670
1	CA6.CG01.XAIA	<==								
2	CA6.CG01.IAXBIB	tcca	aggctaattcc	agttttataa	titttgcataa	tttcttaact	gcgaatataca	tttattana	acaatttatt	ctcccag
3	CA6.CG01.IBXCIC	tttg	acttgggant	aaaangttga	aaaggtaaaa	ggaagga				
4	CA6.CG01.ICXDID	<==								
5	CA6.CG01.IDXEIE	NNNN	Ntctctcttc	cccacctca	gtcncgtgaa	aacaggtttt	aaagatagtt	gctaactctt	atctctctca	aatctctta
6	CA6.CG01.IEXFIF	<==								
7	CA6.CG01.IFXGIG	<==								
8	CA6.CG01.IGXHHH	<==								
9	CA6.CG01.IHXI	ttct	acctctttgc	acttgtggca	acagataagt	tggcagttgg	ctaagagagg	tttccgaaaag	gtttt	

FIG. 10 (cont.)

FIG. 11A

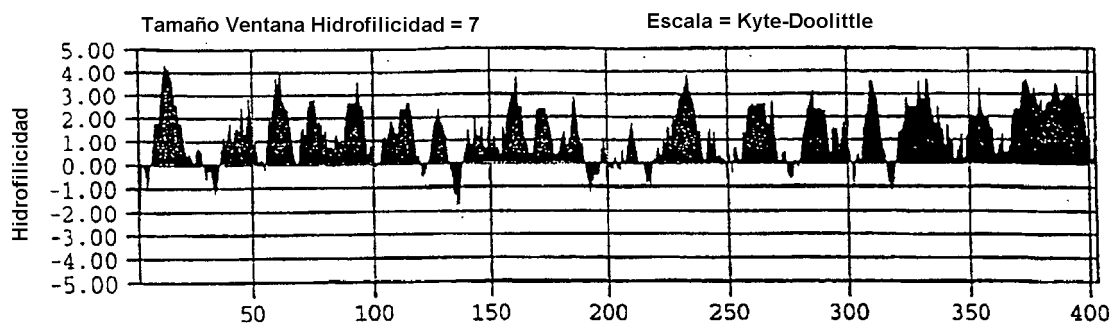


FIG. 11B

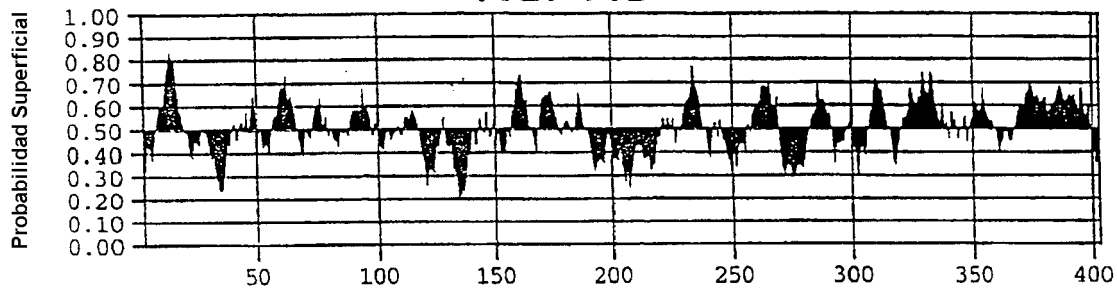


FIG. 11C

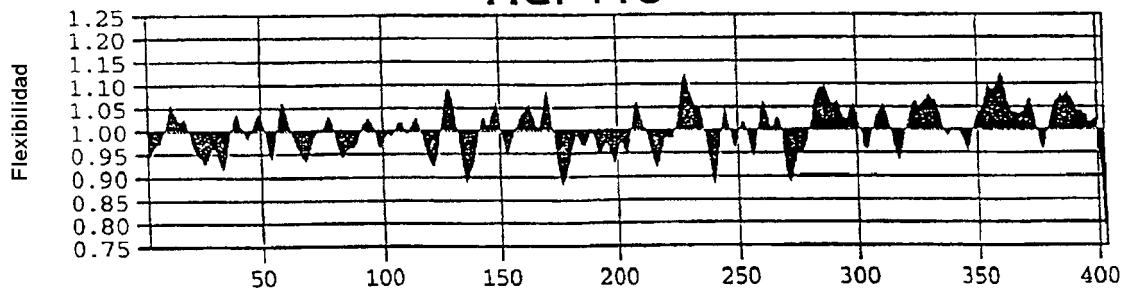


FIG. 11D

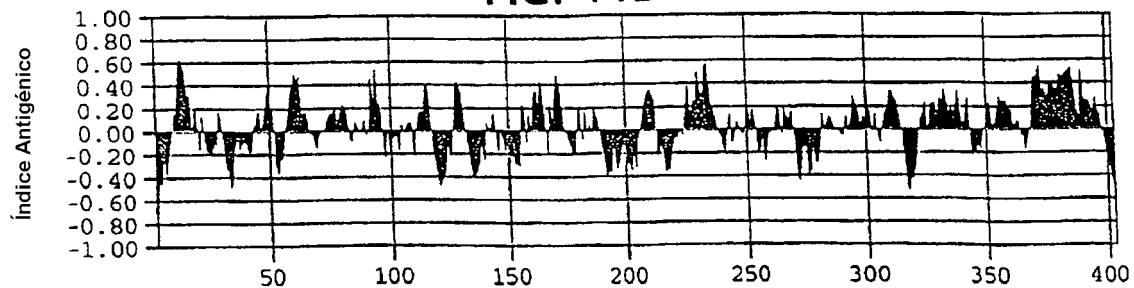


FIG. 11E

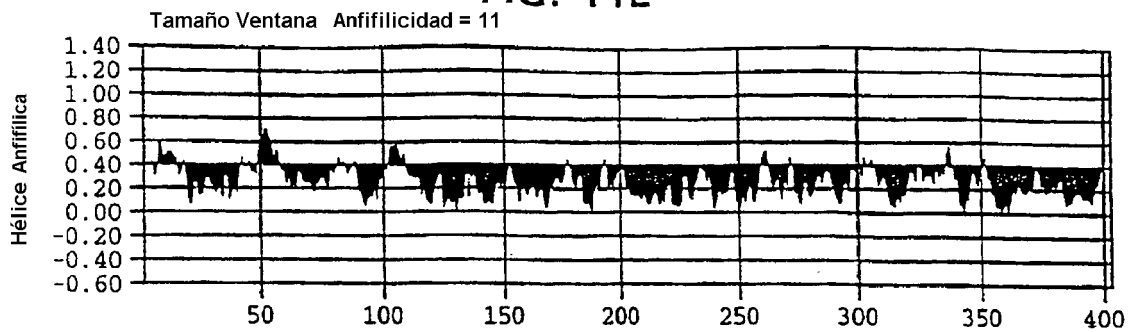


FIG. 11F

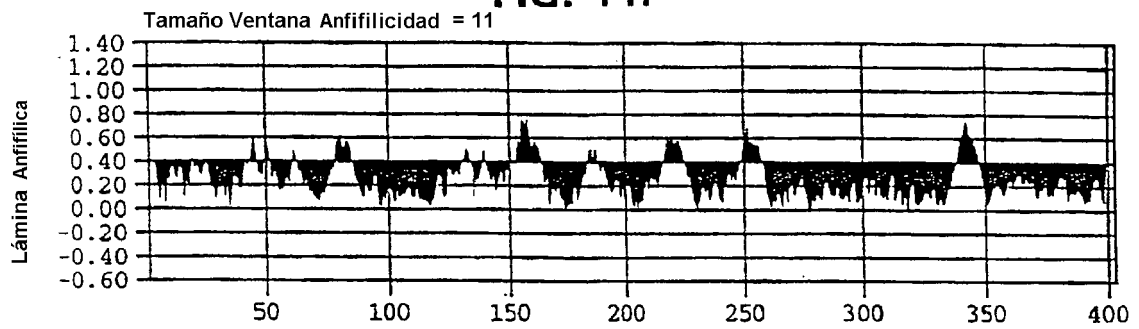


FIG. 11G

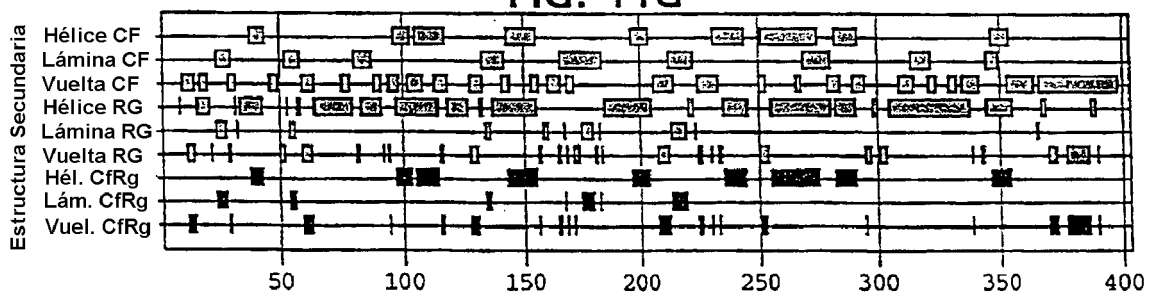


FIG. 12A

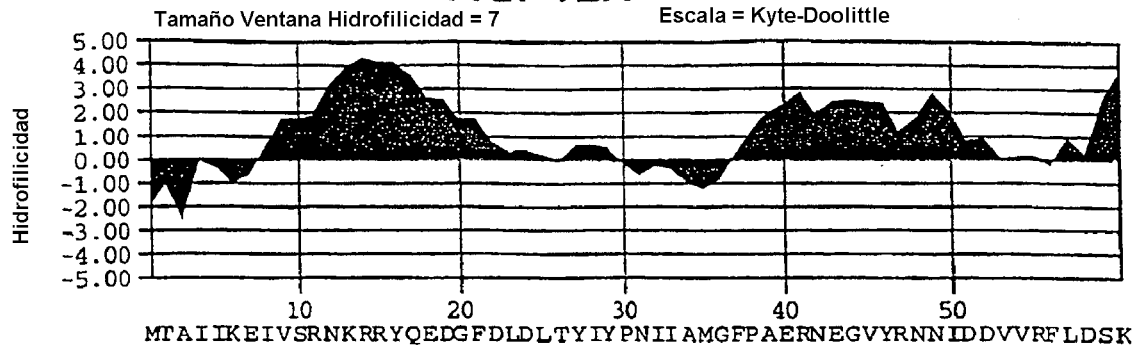


FIG. 12B

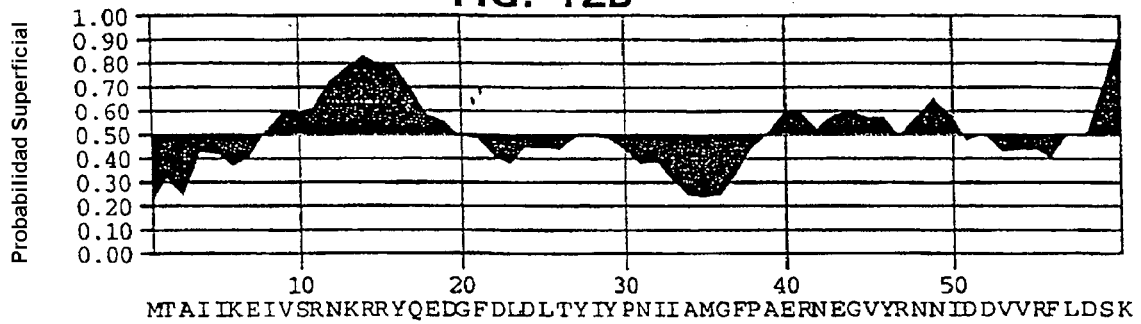


FIG. 12C

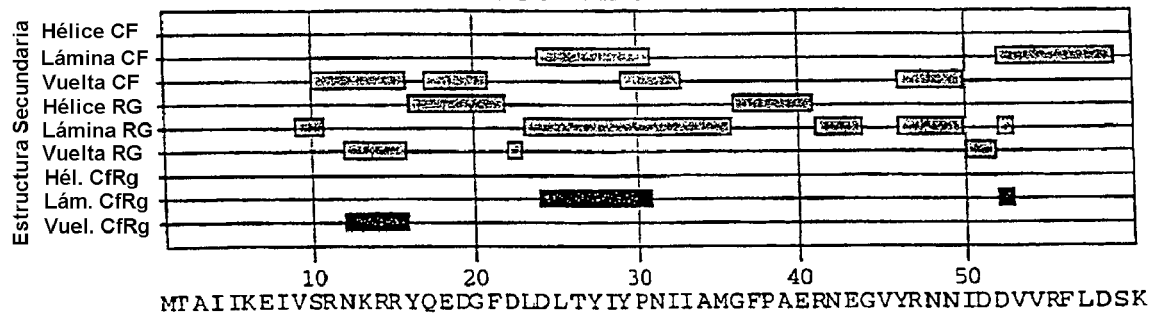


FIG. 12D

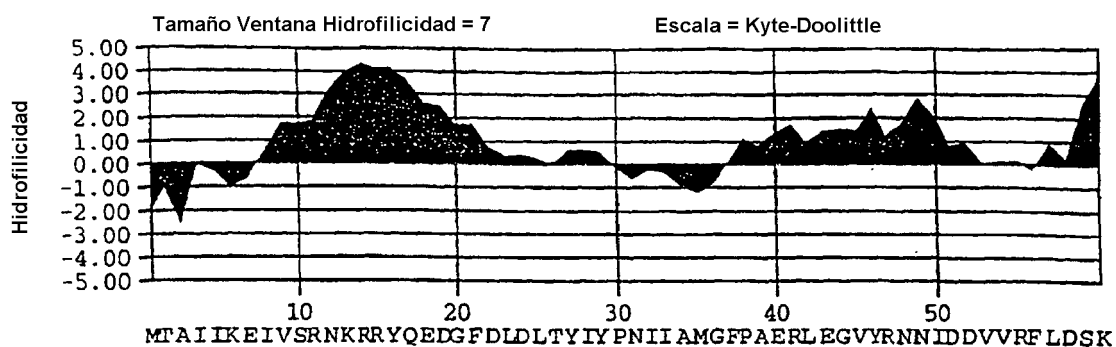


FIG. 12E

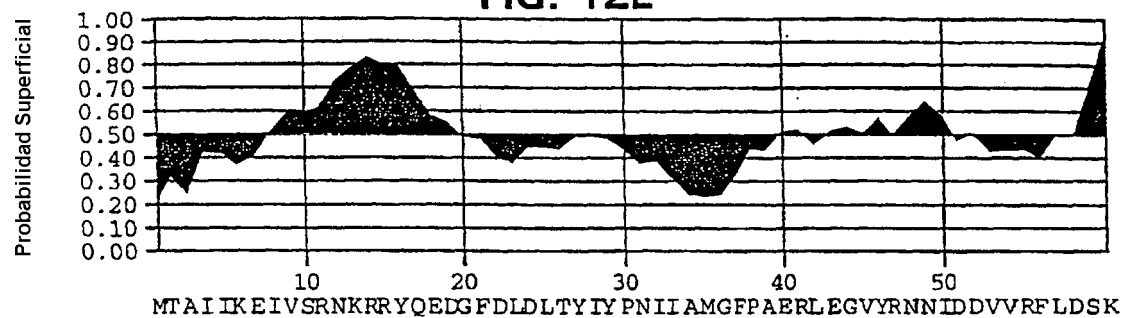


FIG. 12F

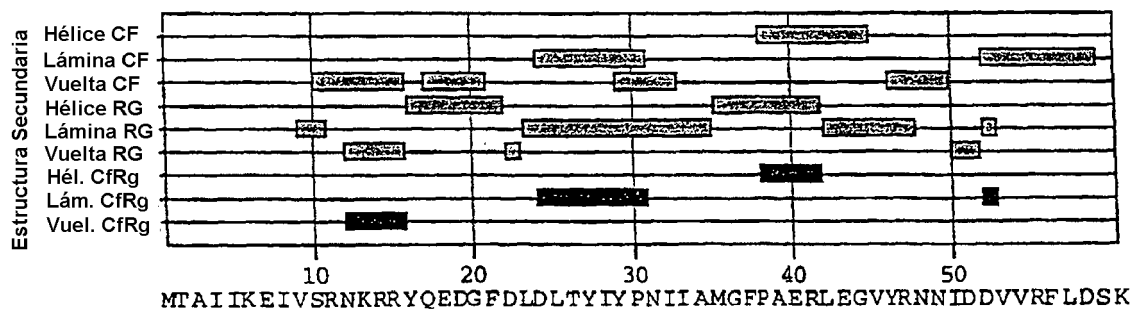


FIG. 12G

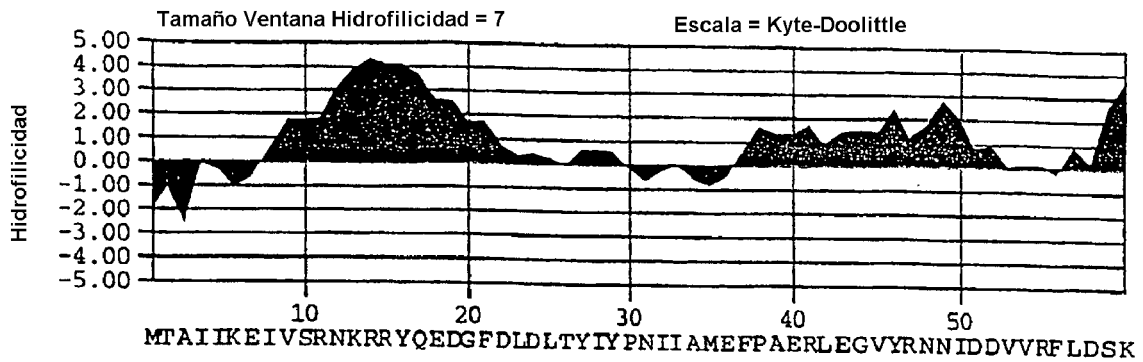


FIG. 12H

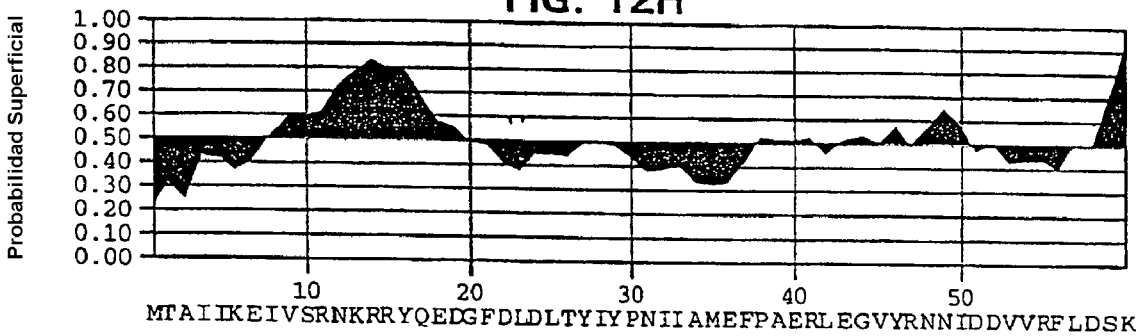
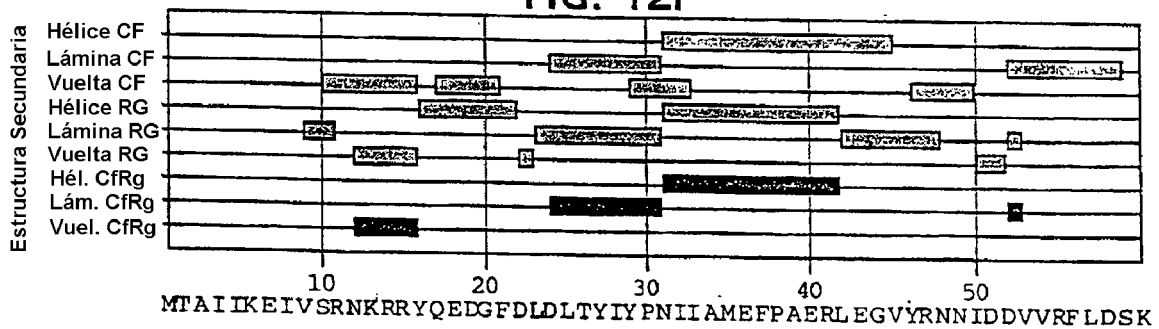


FIG. 12I



ES 2 361 044 T3

LISTA DE SECUENCIAS

(i) SOLICITANTE:

- 5 (A) NOMBRE: Board of Regents, The University of Texas System
(B) CALLE: 201 West 7th Street
(C) CIUDAD: Austin
(D) ESTADO: Texas
10 (E) PAIS: USA
(F) CÓDICO POSTAL (ZIP): 78701
- 15 (A) NOMBRE: Myriad Genetics, Inc.
(B) CALLE: 320 Wakara Way
(C) CIUDAD: Sal Lake City
(D) ESTADO: Utah
20 (E) PAÍS: USA
(F) CÓDIO POSTAL(ZIP): 84108
- 25 (A) NOMBRE: Peter Steck
(B) CALLE: 5421 Holly
(C) CIUDAD: Bellaire
(D) ESTADO: Texas
30 (E) PAÍS: USA
(F) CÓDIO POSTAL(ZIP): 77401
- 35 (A) NOMBRE: Mark A. Pershouse
(B) CALLE: 4065 McDermed
(C) CIUDAD: Houston
(D) ESTADO: Texas
40 (E) PAÍS: USA
(F) CÓDIO POSTAL(ZIP): 77025
- 45 (A) NOMBRE: Samar A. Jasser
(B) CALLE: 9411 Kristin
(C) CIUDAD: Houston
(D) ESTADO: Texas
50 (E) PAÍS: USA
(F) CÓDIO POSTAL(ZIP): 77031
- 55 (A) NOMBRE: W.K. Alfred Yung
(B) CALLE: 4141 Byron
(C) CIUDAD: Houston
(D) ESTADO: Texas
60 (E) PAÍS: USA
(F) CÓDIO POSTAL(ZIP): 77005
- 65 (A) NOMBRE: Sean V. Tantigian
(B) CALLE: 557 East 1st Avenue N° 3
(C) CIUDAD: Salt Lake City

ES 2 361 044 T3

(D) ESTADO: Utah

(E) PAÍS: USA

(F) CÓDIGO POSTAL(ZIP): 84103

5 (ii) TÍTULO DE LA INVENCION: UN SUPRESOR DE TUMOR DESIGNADO TS10q23.3

(iii) NÚMERO DE SECUENCIAS: 21

10 (iv) FORMA LEGIBLE POR ORDENADOR:

(A) TIPO DE MEDIO: Floppy disk

(B) ORDENADOR: IBM PC compatible

(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS

15 (D) PROGRAMA INFORMÁTICO: PatentIn Release N° 1.0, Versión N° 1.30 (EPO)

(v) DATOS DE LA SOLICITUD ACTUAL:

NÚMERO DE SOLICITUD: Desconocido

20 (vi) DATOS DE LA SOLICITUD ANTERIOR:

(A) NÚMERO DE LA SOLICITUD: US 08/791.115

(B) DATOS DE PRESENTACIÓN: 30 de Enero de 1997

25 (2) INFORMACIÓN PARA LA SEC. ID. N°: 1:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

30 (A) LONGITUD: 403 aminoácidos

(B) TIPO: aminoácidos

(C) ESTADO DE LA CADENA:

35 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 1:

40

45

50

55

60

65

ES 2 361 044 T3

Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
 1 5 10 15
 5 Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
 20 25 30
 10 Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
 35 40 45
 Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys His Lys Asn His
 50 55 60
 15 Tyr Lys Ile Tyr Asn Leu Cys Ala Glu Arg His Tyr Asp Thr Ala Lys
 65 70 75 80
 20 Phe Asn Cys Arg Val Ala Gln Tyr Pro Phe Glu Asp His Asn Pro Pro
 85 90 95
 Gln Leu Glu Leu Ile Lys Pro Phe Cys Glu Asp Leu Asp Gln Trp Leu
 100 105 110
 25 Ser Glu Asp Asp Asn His Val Ala Ala Ile His Cys Lys Ala Gly Lys
 115 120 125
 30 Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg Gly Lys
 130 135 140
 Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val Arg Thr
 145 150 155 160
 35 Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr Val Tyr
 165 170 175
 40 Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala
 180 185 190
 Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly
 195 200 205
 45 Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile
 210 215 220
 50 Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr
 225 230 235 240
 Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys Val Glu

55

60

65

ES 2 361 044 T3

				245					250					255		
	Phe	Phe	His	Lys	Gln	Asn	Lys	Met	Leu	Lys	Lys	Asp	Lys	Met	Phe	His
				260					265					270		
5	Phe	Trp	Val	Asn	Thr	Phe	Phe	Ile	Pro	Gly	Pro	Glu	Glu	Thr	Ser	Glu
			275					280					285			
10	Lys	Val	Glu	Asn	Gly	Ser	Leu	Cys	Asp	Gln	Glu	Ile	Asp	Ser	Ile	Cys
		290					295					300				
15	Ser	Ile	Glu	Arg	Ala	Asp	Asn	Asp	Lys	Glu	Tyr	Leu	Val	Leu	Thr	Leu
	305					310					315					320
20	Thr	Lys	Asn	Asp	Leu	Asp	Lys	Ala	Asn	Lys	Asp	Lys	Ala	Asn	Arg	Tyr
					325					330					335	
25	Phe	Ser	Pro	Asn	Phe	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Lys	Thr	Val	Glu
				340					345					350		
30	Glu	Pro	Ser	Asn	Pro	Glu	Ala	Ser	Ser	Ser	Thr	Ser	Val	Thr	Pro	Asp
			355					360					365			
35	Val	Ser	Asp	Asn	Glu	Pro	Asp	His	Tyr	Arg	Tyr	Ser	Asp	Thr	Thr	Asp
		370					375					380				
40	Ser	Asp	Pro	Glu	Asn	Glu	Pro	Phe	Asp	Glu	Asp	Gln	His	Thr	Gln	Ile
	385					390					395					400
45	Thr	Lys	Val													

(2) INFORMACIÓN PARA LA SEC. ID. N°: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA

- (A) LONGITUD: 3.160 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 2:

45	CCTCCCCTCG	CCCGGC	CGCG	TCCCGTCCGC	CTCTCGCTCG	CCTCCCGCCT	CCCCTCGGTC	60
	TTCCGAGGCG	CCCGGGCTCC	CGGCGCGGGC	GCGGAGGGGG	CGGGCAGGCC	GGCGGGCGGT	120	
50	GATGTGGCAG	GACTCTTTAT	GCGCTGCGGC	AGGATACGCG	CTCGGCCTG	GGACGCGACT	180	
	GCGCTCAGTT	CTCTCCTCTC	GGAAGCTGCA	GCCATGATGG	AAGTTTGAGA	GTTGAGCCGC	240	
55	TGTGAGGCGA	GGCCGGGCTC	AGGCGAGGGA	GATGAGAGAC	GGCGGC	GCGGCCCGGA	300	
	GCCCCCTCTCA	GCGCCTGTGA	GCAGCCGCGG	GGGCAGCGCC	CTCGGGGAGC	CGGCCGGCCT	360	
60	GCGGCGGCGG	CAGCGGCGGC	GTTTCTCGCC	TCCTCTTCGT	CTTTTCTAAC	CGTGCAGCCT	420	
	CTTCCTCGGC	TTCTCCTGAA	AGGGAAGGTG	GAAGCCGTGG	GCTCGGGCGG	GAGCCGGCTG	480	
65	AGGCGCGGGC	GCGGCGGCGG	CGGCACCTCC	CGCTCCTGGA	GCGGGGGGGA	GAAGCGGCGG	540	
	CGGCGGCGGC	CGCGGCGGCT	GCAGCTCCAG	GGAGGGGGTC	TGAGTCGCCT	GTCACCATTT	600	
	CCAGGGCTGG	GAACGCCGGA	GAGTTGGTCT	CTCCCCTTCT	ACTGCCTCCA	ACACGGCGGC	660	

ES 2 361 044 T3

	GGCGGCGGCG	GCACATCCAG	GGACCCGGGC	CGGTTTTAAA	CCTCCCGTCC	GCCGCCGCCG	720
	CACCCCCCGT	GGCCCCGGGCT	CCGGAGGCCG	CCGGCGGAGG	CAGCCGTTCC	GAGGATTATT	780
5	CGTCTTCTCC	CCATTCCGCT	GCCGCCGCTG	CCAGGCCTCT	GGCTGCTGAG	GAGAAGCAGG	840
	CCCAGTCGCT	GCAACCATCC	AGCAGCCGCC	GCAGCAGCCA	TTACCCGGCT	GCGGTCCAGA	900
10	GCCAAGCGGC	GGCAGAGCGA	GGGGCATCAG	CTACCCGCAA	GTCCAGAGCC	ATTTCCATCC	960
	TGCAGAAGAA	GCCCCGCCAC	CAGCAGCTTC	TGCCATCTCT	CTCCTCCTTT	TTCTTCAGCC	1020
	ACAGGCTCCC	AGACATGACA	GCCATCATCA	AAGAGATCGT	TAGCAGAAAC	AAAAGGAGAT	1080
15	ATCAAGAGGA	TGGATTGAC	TTAGACTTGA	CCTATATTTA	TCCAAACATT	ATTGCTATGG	1140
	GATTTCTGCT	AGAAAGACTT	GAAGGCGTAT	ACAGGAACAA	TATTGATGAT	GTAGTAAGGT	1200
20	TTTTGGATTG	AAAGCATAAA	AACCATTACA	AGATATACAA	TCTTTGTGCT	GAAAGACATT	1260
	ATGACACCGC	CAAATTTAAT	TGCAGAGTTG	CACAATATCC	TTTTGAAGAC	CATAACCCAC	1320
	CACAGCTAGA	ACTTATCAAA	CCCTTTTGTG	AAGATCTTGA	CCAATGGCTA	AGTGAAGATG	1380
25	ACAATCATGT	TGCAGCAATT	CACTGTAAAG	CTGGAAAGGG	ACGAACTGGT	GTAATGATAT	1440
	GTGCATATTT	ATTACATCGG	GGCAAATTTT	TAAAGGCACA	AGAGGCCCTA	GATTTCTATG	1500
30	GGGAAGTAAG	GACCAGAGAC	AAAAAGGGAG	TAACTATTCC	CAGTCAGAGG	CGCTATGTGT	1560
	ATTATTATAG	CTACCTGTTA	AAGAATCATC	TGGATTATAG	ACCAGTGGCA	CTGTTGTTTC	1620
	ACAAGATGAT	GTTTGAAACT	ATTCCAATGT	TCAGTGGCGG	AACTTGCAAT	CCTCAGTTTG	1680
35	TGGTCTGCCA	GCTAAAGGTG	AAGATATATT	CCTCCAATTC	AGGACCCACA	CGACGGGAAG	1740
	ACAAGTTCAT	GTACTTTGAG	TTCCCTCAGC	CGTTACCTGT	GTGTGGTGAT	ATCAAAGTAG	1800
40	AGTTCTTCCA	CAAACAGAAC	AAGATGCTAA	AAAAGGACAA	AATGTTTCAC	TTTTGGGTAA	1860
	ATACATTCTT	CATACCAGGA	CCAGAGGAAA	CCTCAGAAAA	AGTAGAAAAT	GGAAGTCTAT	1920
	GTGATCAAGA	AATCGATAGC	ATTTGCAGTA	TAGAGCGTGC	AGATAATGAC	AAGGAATATC	1980
45	TAGTACTTAC	TTTAAACAAA	AATGATCTTG	ACAAAGCAAA	TAAAGACAAA	GCCAACCGAT	2040
	ACTTTTCTCC	AAATTTTAAG	GTGAAGCTGT	ACTTCACAAA	AACAGTAGAG	GAGCCGTCAA	2100
50	ATCCAGAGGC	TAGCAGTTCA	ACTTCTGTAA	CACCAGATGT	TAGTGACAAT	GAACCTGATC	2160
	ATTATAGATA	TTCTGACACC	ACTGACTCTG	ATCCAGAGAA	TGAACCTTTT	GATGAAGATC	2220
55	AGCATACACA	AATTACAAAA	GTCTGAATTT	TTTTTTTATCA	AGAGGGATAA	AACACCATGA	2280
	AAATAAACTT	GAATAAACTG	AAAATGGACC	TTTTTTTTTTT	TAATGGCAAT	AGGACATTGT	2340
	GTCAGATTAC	CAGTTATAGG	AACAATTCCTC	TTTTCTGAC	CAATCTTGTT	TTACCCTATA	2400
60	CATCCACAGG	GTTTTGACAC	TTGTTGTCCA	GTTGAAAAAA	GGTTGTGTAG	CTGTGTCATG	2460
	TATATACCTT	TTTGTGTCAA	AAGGACATTT	AAAATTCAAT	TAGGATTAAT	AAAGATGGCA	2520
65	CTTTCCCGTT	TTATTCCAGT	TTTATAAAAA	GTGGAGACAG	ACTGATGTGT	ATACGTAGGA	2580

ES 2 361 044 T3

	ATTTTTTCCT	TTTGTGTTCT	GTCACCAACT	GAAGTGGCTA	AAGAGCTTTG	TGATATACTG	2640
	G TTCACATCC	T ACCCCTTTG	C ACTTGTGGC	A ACAGATAAG	T TTGCAGTTG	G CTAAGAGAG	2700
5	G TTTCCGAAA	G GTTTTGCTA	C CATTCTAAT	G GCATGTATTC	G GGTTAGGGC	A AATGGAGGGG	2760
	A ATGCTCAGA	A AAGGAAATAA	T TTTTATGCTG	G GACTCTGGAC	C CATATACCAT	C CTCCAGCTAT	2820
	T TACACACAC	C TTTCTTTAG	C ATGCTACAG	T TATTAATCT	G GACATTCGA	G GGAATTGGCC	2880
10	G CTGTCACTG	C TTGTTGTTT	G GCGCATTTTT	T TTTAAAGCA	T TATTGGTGCT	A AGAAAAGGCA	2940
	G GCTAAAGGAA	G TGAATCTGT	A TTGGGGTAC	A AGGAATGAAC	C TTTCTGCAAC	A ATCTTAAGAT	3000
	C CCACAAATGA	A AGGGATATAA	A AAATAATGTC	A ATAGSTAAGA	A AACACAGCAA	C CAATGACTTA	3060
	A ACCATATAAA	T TGTGGAGGCT	A TCAACAAAG	A AATGGGCTTG	A AAACATTATA	A AAAATTGACA	3120
15	A ATGATTTATT	A AAATATGTTT	T TCTCAATTGT	A AAAAAAAAAA			3160

20 (2) INFORMACIÓN PARA LA SEC. ID. N°: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1.962 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencillo
- (D) TOPOLOGÍA: lineal

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 3:

	GCGAGGGAGA	TGAGAGACGG	CGGCGGCCAC	GGCCCAGAGC	CCCTCTCAGC	GCCTGTGAGC	60
35	A AGCCGCGGGG	G GCAGCGCCCT	C GGGGAGCCG	G GCCGGGCGGC	G GCGGCGGCA	G GCGGCGGCGG	120
	G GCCTCGCCTC	C CTCGTCTCT	G GTTCTAACCG	G GGCAGCTTCT	G GAGCAGCTTC	G GGAGAGAGAC	180
40	G GGTGGAAGAA	G GCCGTGGGCT	C GAGCGGGAG	C CCGGCGCAGG	C CTCGGCGGCT	G GCACCTCCCG	240
	C CTCCTGGAGC	G GGGGGGAGA	A AGCGGCGGCG	G GCGGCCGCGG	C TCCGGGGAG	G GGGGTCGGAG	300
	T TCGCCTGTCA	C CCATTGCCAG	G GGCTGGGAAC	G GCCGGAGAGT	T TGCTCTCTCC	C CCTTCTCCTG	360
45	C CCTCCAACAC	G GCGGCGGGG	G GCGGCGGCAC	G GTCCAGGGAC	C CCGGGCCGGT	G GTTAAGCCTC	420
	C CCGTCCGCCG	C CCGCCGCACC	C CCCCCTGGCC	C CGGGCTCCGG	A AGGCCGCCGG	A AGGAGGCAGC	480
50	C CGCTGCGAGG	A ATTATCCGTC	T TTCTCCCCAT	T TCCGCTGCCT	C CGGCTGCCAG	G GCCTCTGGCT	540
	G GCTGAGGAGA	A AGCAGGCCCA	G GTCTCTGCAA	C CCATCCAGCA	G GCCGCCGCAG	C CAGCCATTAC	600
55	C CCGGCTGCGG	T TCCAGGGCCA	A AGCGGCAGCA	G GAGCGAGGGG	C ATCAGCGAC	C GCCAAGTCC	660
	A AGAGCCATTT	C CCATCCTGCA	G GAAGAAGCCT	C GCCACCAGC	A AGCTTCTGCC	A ATCTCTCTCC	720
	T TCCTTTTTTCT	T TCAGCCACAG	G GCTCCAGAC	A ATGACAGCCA	T TCATCAAAGA	G GATCGTTAGC	780
60	A AGAAACAAAA	G GGAGATATCA	A AGAGGATGGA	T TCGACTTAG	A ACTTGACCTA	T TATTTATCCA	840
	A AATATTATTG	C CTATGGGATT	T TCCTGCAGAA	A GACTTGAAG	G GTGTATACAG	A GAACAAATAT	900
65	G GATGATGTAG	T TAAGGTTTTT	G GGATTCAAAG	C CATAAAAACC	A ATTACAAGAT	A ATACAATCTA	960

ES 2 361 044 T3

	TGTGCTGAGA GACATTATGA CACCGCCAAA TTAACTGCA GAGTTGCACA GTATCCTTTT	1020
	GAAGACCATA ACCCACCACA GCTAGAACTT ATCAAACCCT TCTGTGAAGA TCTTGACCAA	1080
5	TGGCTAAGTG AAGATGACAA TCATGTTGCA GCAATTCCTT GTAAAGCTGG AAAGGGACGG	1140
	ACTGGTGTAA TGATTTGTGC ATATTTATTG CATCGGGGCA AATTTTAAA GGCACAAGAG	1200
10	GCCCTAGATT TTTATGGGGA AGTAAGGACC AGAGACAAAA AGGGAGTCAC AATTCCTCAGT	1260
	CAGAGGCGCT ATGTATATTA TTATAGCTAC CTGCTAAAAA ATCACCTGGA TTACAGACCC	1320
	GTGGCACTGC TGTTTCACAA GATGATGTTT GAACTATTC CAATGTTTCAG TGGCGGAACT	1380
15	TGCAATCCTC AGTTTGTGGT CTGCCAGCTA AAGGTGAAGA TATATTCCTC CAATTCAGGA	1440
	CCCACGCGGC GGGAGGACAA GTTCATGTAC TTTGAGTTCC CTCAGCCATT GCCTGTGTGT	1500
20	GGTGATATCA AAGTAGAGTT CTTCCACAAA CAGAACAAGA TGCTCAAAAA GGACAAAATG	1560
	TTTCACTTTT GGGTAAATAC GTTCTTCATA CCAGGACCAG AGGAAACCTC AGAAAAAGTG	1620
25	GAAAATGGAA GTCTTTGTGA TCAGGAAATC GATAGCATTT GCAGTATAGA GCGTGCAGAT	1680
	AATGACAAGG AGTATCTTGT ACTCACCTA ACAAAAAACG ATCTTGACAA AGCAAACAAA	1740
	GACAAGGCCA ACCGATACTT CTCTCCAAAT TTTAAGGTGA AACTATACTT TACAAAACA	1800
30	GTAGAGGAGC CATCAAATCC AGAGGCTAGC AGTTCAACTT CTGTGACTCC AGATGTTAGT	1860
	GACAATGAAC CTGATCATTG TAGATATTCT GACACCACTG ACTCTGATCC AGAGAATGAA	1920
35	CCTTTTGATG AAGATCAGCA TACACAAAT ACAAAGTCT GA	1962

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 4:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 1.291 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 4:

	CCGCCCCCGG CCAGGCCCGG GGCCGCTGC AGCCTGCGGA GGAGGCCCGG CCGCCCCCGG	60
50	CTCCTGCCGT CTCTCTCCTC CTTCTCTCC AGCCACCGGC TCCAGACAT GACAGCCATC	120
	ATCAAGGAGA TCGTCAGCAG AAACAAAAGG CGCTACCAGG AGGATGGGTT CGACTTGGAC	180
55	TTGACCTATA TTTATCCCAA CATTATTGCT ATGGGGTTTC CTGCAGAAAG ACTTGAAGGC	240
	GTATACAGGA ACAATATTGA TGATGTAGTA AGGTTTTTGG ATTCAAAGCA TAAAAACCAT	300
60	TACAAGATAT ACAATCTGTG TGCTGAAAGA CATTATGATA CCGCCAAATT TAACTGCAGA	360
	GTTGCACAGT ATCCTTTTGA AGACCATAAT CCACCACAGC TAGAACTTAT CAAACCCTTT	420
	TGTGAAGATC TTGACCAATG GCTAAGTGAA GATGACAATC ATGTTGCAGC AATTCCTGT	480
65	AAAGCTGGAA AGGGACGAAC TGGTGTAAATG ATTTGTGCAT ATTTATTACA TCGGGGCAAA	540
	TTTCTAAAGG CACAAGAGGC CCTAGATTTT TATGGGGAAG TAAGGACCAG AGACAAAAAG	600

ES 2 361 044 T3

GGAGTAACTA TTCCCAGTCA GAGGCGCTAT GTGTATTATT ATAGCTACCT GTTAAAGAAT 660
 CATCTGGATT ATAGACCAGT GGCCTGTGTTG TTTCACAAGA TGATGTTTGA AACTATTCCA 720
 5 ATGTTTCAGTG GCGGAACTTG CAATCCTCAG TTTGTGGTCT GCCAGCTAAA GGTGAAGATC 780
 TATTCCTCCA ATTCAGGACC CACACGACGG GAAGACAAGT TCATGTACTT TGAGTTCCTT 840
 10 CAGCCATTGC CTGTGTGCGG TGACATCAAA GTAGAGTTCT TCCACAAACA GAACAAGATG 900
 CTA AAAAAGG ACAAATGTT TCACTTTTGG GTAAACACAT TCTTCATACC AGGACCAGAG 960
 GAAACCTCAG AAAAAGTAGA AAATGGAAGT CTATGTGATC AAGAAATTGA TAGTATTTGC 1020
 15 AGTATAGAAC GTGCAGATAA TGACAAGGAA TATCTAGTAC TCACTTTAAC AAAAAATGAT 1080
 CTCGACAAAG CAAATAAAGA CAAGGCCAAC CGATATTTTT CTCCAAATTT TAAGGTGAAG 1140
 20 CTGTACTTCA CAAAACTGT AGAGGAGCCA TCAAACCCGG AGGCTAGCAG TTCAACTTCT 1200
 GTGACGCCAG ATGTTAGTGA CAATGAACCT GATCATTATA GATATTCTGA CACCACTGAC 1260
 TCTGACCCAG AGAATGAACC CTTTGATGAA G 1291
 25

(2) INFORMACIÓN PARA LA SEC. ID. N°: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 742 aminoácidos

(B) TIPO: aminoácido

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 5:

40 Ser Pro Arg Pro Ala Arg Ser Arg Pro Pro Leu Ala Arg Leu Pro Pro
 1 5 10 15
 45 Pro Leu Gly Leu Pro Arg Arg Pro Gly Ser Arg Arg Gly Gly Gly Gly
 20 25 30
 Gly Gly Gln Ala Gly Gly Arg Cys Gly Arg Thr Leu Tyr Ala Leu Arg
 35 40 45
 50 Gln Asp Thr Arg Ser Ala Leu Gly Arg Asp Cys Ala Gln Phe Ser Pro
 50 55 60
 Leu Gly Ser Cys Ser His Asp Gly Ser Leu Arg Val Glu Pro Leu Gly
 65 70 75 80
 55 Glu Ala Gly Leu Arg Arg Gly Arg Glu Thr Ala Ala Ala Ala Ala Arg
 85 90 95
 60 Ser Pro Ser Gln Arg Leu Ala Ala Ala Gly Ala Ala Pro Ser Gly Ser
 100 105 110
 Arg Pro Ala Cys Gly Gly Gly Ser Gly Gly Val Ser Arg Leu Leu Phe
 115 120 125
 65 Val Phe Ser Asn Arg Ala Ala Ser Ser Ser Ala Ser Pro Glu Arg Glu
 130 135 140
 Gly Ser Arg Gly Leu Gly Arg Glu Pro Ala Glu Ala Arg Arg Arg Arg

ES 2 361 044 T3

	145					150										155				160
5	Arg	Arg	His	Leu	Pro	Leu	Leu	Glu	Arg	Gly	Gly	Glu	Ala	Ala	Ala	Ala				
					165					170					175					
	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Pro	Gly	Arg	Gly	Ser	Glu	Ser	Pro	Val				
				180					185					190						
10	Thr	Ile	Ser	Arg	Ala	Gly	Asn	Ala	Gly	Glu	Leu	Val	Ser	Pro	Leu	Leu				
			195				200						205							
	Leu	Pro	Pro	Thr	Arg	Arg	Arg	Arg	Arg	Arg	His	Ile	Gln	Gly	Pro	Gly				
		210					215					220								
15	Pro	Val	Leu	Asn	Leu	Pro	Ser	Ala	Ala	Ala	Ala	Pro	Pro	Val	Ala	Arg				
		225				230					235					240				
	Ala	Pro	Glu	Ala	Ala	Gly	Gly	Gly	Ser	Arg	Ser	Glu	Asp	Tyr	Ser	Ser				
20				245						250					255					
	Ser	Pro	His	Ser	Ala	Ala	Ala	Ala	Ala	Arg	Pro	Leu	Ala	Ala	Glu	Glu				
				260					265					270						
25	Lys	Gln	Ala	Gln	Ser	Leu	Gln	Pro	Ser	Ser	Ser	Arg	Arg	Ser	Ser	His				
			275					280					285							
	Tyr	Pro	Ala	Ala	Val	Gln	Ser	Gln	Ala	Ala	Ala	Glu	Arg	Gly	Ala	Ser				
		290					295					300								
30	Ala	Thr	Ala	Lys	Ser	Arg	Ala	Ile	Ser	Ile	Leu	Gln	Lys	Lys	Pro	Arg				
		305				310					315					320				
	His	Gln	Gln	Leu	Leu	Pro	Ser	Leu	Ser	Ser	Phe	Phe	Phe	Ser	His	Arg				
35				325						330					335					
	Leu	Pro	Asp	Met	Thr	Ala	Ile	Ile	Lys	Glu	Ile	Val	Ser	Arg	Asn	Lys				
				340					345					350						
40	Arg	Arg	Tyr	Gln	Glu	Asp	Gly	Phe	Asp	Leu	Asp	Leu	Thr	Tyr	Ile	Tyr				
			355					360					365							
	Pro	Asn	Ile	Ile	Ala	Met	Gly	Phe	Pro	Ala	Glu	Arg	Leu	Glu	Gly	Val				
		370					375					380								
45	Tyr	Arg	Asn	Asn	Ile	Asp	Asp	Val	Val	Arg	Phe	Leu	Asp	Ser	Lys	His				
		385				390					395				400					
	Lys	Asn	His	Tyr	Lys	Ile	Tyr	Asn	Leu	Cys	Ala	Glu	Arg	His	Tyr	Asp				
50				405						410					415					
	Thr	Ala	Lys	Phe	Asn	Cys	Arg	Val	Ala	Gln	Tyr	Pro	Phe	Glu	Asp	His				
				420					425					430						
55	Asn	Pro	Pro	Gln	Leu	Glu	Leu	Ile	Lys	Pro	Phe	Cys	Glu	Asp	Leu	Asp				
			435					440					445							
	Gln	Trp	Leu	Ser	Glu	Asp	Asp	Asn	His	Val	Ala	Ala	Ile	His	Cys	Lys				
60		450					455					460								
	Ala	Gly	Lys	Gly	Arg	Thr	Gly	Val	Met	Ile	Cys	Ala	Tyr	Leu	Leu	His				
		465				470					475					480				
65	Arg	Gly	Lys	Phe	Leu	Lys	Ala	Gln	Glu	Ala	Leu	Asp	Phe	Tyr	Gly	Glu				
				485						490					495					

ES 2 361 044 T3

Val Arg Thr Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg
 500 505 510
 5 Tyr Val Tyr Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg
 515 520 525
 Pro Val Ala Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met
 530 535 540
 10 Phe Ser Gly Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys
 545 550 555 560
 Val Lys Ile Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys
 565 570 575
 Phe Met Tyr Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile
 580 585 590
 20 Lys Val Glu Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys
 595 600 605
 Met Phe His Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu
 610 615 620
 25 Thr Ser Glu Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp
 625 630 635 640
 Ser Ile Cys Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val
 645 650 655
 30 Leu Thr Leu Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala
 660 665 670
 Asn Arg Tyr Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys
 675 680 685
 35 Thr Val Glu Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val
 690 695 700
 Thr Pro Asp Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp
 705 710 715 720
 Thr Thr Asp Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His
 725 730 735
 45 Thr Gln Ile Thr Lys Val
 740

(2) INFORMACIÓN PARA LA SEC. ID. N°: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA

(A) LONGITUD: 645 aminoácidos

(B) TIPO: aminoácido

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 6:

Arg Glu Thr Ala Ala Ala Thr Ala Gln Ser Pro Ser Gln Arg Leu Ala
 1 5 10 15

Ala Ala Gly Ala Ala Pro Ser Gly Ser Arg Pro Gly Gly Gly Gly Gly

ES 2 361 044 T3

Gly Lys Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg
 370 375 380
 5 Gly Lys Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val
 385 390 395 400
 Arg Thr Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr
 405 410 415
 10 Val Tyr Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro
 420 425 430
 Val Ala Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe
 15 435 440 445
 Ser Gly Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val
 450 455 460
 20 Lys Ile Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe
 465 470 475 480
 Met Tyr Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys
 25 485 490 495
 Val Glu Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys Met
 500 505 510
 30 Phe His Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu Thr
 515 520 525
 Ser Glu Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp Ser
 530 535 540
 35 Ile Cys Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val Leu
 545 550 555 560
 Thr Leu Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala Asn
 40 565 570 575
 Arg Tyr Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys Thr
 580 585 590
 45 Val Glu Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val Thr
 595 600 605
 Pro Asp Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp Thr
 610 615 620
 50 Thr Asp Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His Ser
 625 630 635 640
 55 Gln Ile Thr Lys Val
 645

(2) INFORMACIÓN PARA LA SEC. ID. N°: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 430 aminoácidos

(B) TIPO: aminoácido

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

ES 2 361 044 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 7:

5	Pro	Pro	Ala	Ala	Arg	Pro	Gly	Ala	Ala	Cys	Ser	Leu	Arg	Arg	Arg	Pro
	1			5						10					15	
	Arg	Arg	Pro	Pro	Leu	Leu	Pro	Ser	Leu	Ser	Ser	Phe	Leu	Ser	Ser	His
			20						25					30		
10	Arg	Leu	Pro	Asp	Met	Thr	Ala	Ile	Ile	Lys	Glu	Ile	Val	Ser	Arg	Asn
			35					40					45			
	Lys	Arg	Arg	Tyr	Gln	Glu	Asp	Gly	Phe	Asp	Leu	Asp	Leu	Thr	Tyr	Ile
15		50					55					60				
	Tyr	Pro	Asn	Ile	Ile	Ala	Met	Gly	Phe	Pro	Ala	Glu	Arg	Leu	Glu	Gly
	65					70					75					80
20	Val	Tyr	Arg	Asn	Asn	Ile	Asp	Asp	Val	Val	Arg	Phe	Leu	Asp	Ser	Lys
				85						90					95	
	His	Lys	Asn	His	Tyr	Lys	Ile	Tyr	Asn	Leu	Cys	Ala	Glu	Arg	His	Tyr
25			100						105					110		
	Asp	Thr	Ala	Lys	Phe	Asn	Cys	Arg	Val	Ala	Gln	Tyr	Pro	Phe	Glu	Asp
			115					120					125			
	His	Asn	Pro	Pro	Gln	Leu	Glu	Leu	Ile	Lys	Pro	Phe	Cys	Glu	Asp	Leu
30		130					135					140				
	Asp	Gln	Trp	Leu	Ser	Glu	Asp	Asp	Asn	His	Val	Ala	Ala	Ile	His	Cys
	145					150					155					160
	Lys	Ala	Gly	Lys	Gly	Arg	Thr	Gly	Val	Met	Ile	Cys	Ala	Tyr	Leu	Leu
35					165					170					175	
	His	Arg	Gly	Lys	Phe	Leu	Lys	Ala	Gln	Glu	Ala	Leu	Asp	Phe	Tyr	Gly
				180					185					190		
40	Glu	Val	Arg	Thr	Arg	Asp	Lys	Lys	Gly	Val	Thr	Ile	Pro	Ser	Gln	Arg
			195					200					205			
	Arg	Tyr	Val	Tyr	Tyr	Tyr	Ser	Tyr	Leu	Leu	Lys	Asn	His	Leu	Asp	Tyr
		210					215					220				
45	Arg	Pro	Val	Ala	Leu	Leu	Phe	His	Lys	Met	Met	Phe	Glu	Thr	Ile	Pro
		225				230					235					240
	Met	Phe	Ser	Gly	Gly	Thr	Cys	Asn	Pro	Gln	Phe	Val	Val	Cys	Gln	Leu
					245					250					255	
50	Lys	Val	Lys	Ile	Tyr	Ser	Ser	Asn	Ser	Gly	Pro	Thr	Arg	Arg	Glu	Asp
				260					265					270		
	Lys	Phe	Met	Tyr	Phe	Glu	Phe	Pro	Gln	Pro	Leu	Pro	Val	Cys	Gly	Asp
55			275					280					285			
	Ile	Lys	Val	Glu	Phe	Phe	His	Lys	Gln	Asn	Lys	Met	Leu	Lys	Lys	Asp
		290					295					300				
60	Lys	Met	Phe	His	Phe	Trp	Val	Asn	Thr	Phe	Phe	Ile	Pro	Gly	Pro	Glu
	305					310					315					320
	Glu	Thr	Ser	Glu	Lys	Val	Glu	Asn	Gly	Ser	Leu	Cys	Asp	Gln	Glu	Ile
					325					330					335	

65

ES 2 361 044 T3

Asp Ser Ile Cys Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu
 340 345 350
 Val Leu Thr Leu Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys
 5 355 360 365
 Ala Asn Arg Tyr Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr
 370 375 380
 10 Lys Thr Val Glu Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser
 385 390 395 400
 Val Thr Pro Asp Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser
 15 405 410 415
 Asp Thr Thr Asp Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu
 420 425 430

(2) INFORMACIÓN PARA LA SEC. ID. N°: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1.257 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 8:

CCTCCCCTCG CCCGGCGCGG TCCCGTCCGC CTCTCGCTCG CCTCCCGCCT CCCCTCGGTC 60
 TTCCGAGGCG CCCGGGCTCC CGGCGCGGCG GCGGAGGGGG CGGGCAGGCC GGCGGGCGGT 120
 35 GATGTGGCAG GACTCTTTAT GCGCTGCGGC AGGATACGCG CTCGGCGCTG GGACCGGACT 180
 GCGCTCAGTT CTCTCCTCTC GGAAGCTGCA GCCATGATGG AAGTTTGAGA GTTGAGCCGC 240
 40 TGTGAGGCGA GGCCGGGCTC AGGCGAGGGA GATGAGAGAC GCGGCGGGCC GCGGCCCGGA 300
 GCCCCTCTCA GCGCCTGTGA GCAGCCGCGG GGGCAGCGCC CTCGGGGAGC CGGCCGGCCT 360
 45 GCGGCGGGCG CAGCGGCGGC GTTCTCTGCC TCCTCTTCGT CTTTCTAAC CGTGACGCCT 420
 CTTCTCTGGC TTCTCCTGAA AGGGAAGGTG GAAGCCGTGG GCTCGGGCGG GAGCCGGCTG 480
 AGGCGCGGGC GCGGCGGGCG CGGCACCTCC CGCTCCTGGA GCGGGGGGA GAAGCGGGCG 540
 CGGCGGGCGC CGCGGCGGCT GCAGCTCCAG GGAGGGGGTC TGAGTCGCCT GTCACCATTT 600
 CCAGGGCTGG GAACGCCGA GAGTTGGTCT CTCCCCTTCT ACTGCCTCCA ACACGGCGGC 660
 55 GCGGGCGGGC GCACATCCAG GGACCCGGGC CGGTTTTAAA CCTCCCGTCC GCCGCCGCCG 720
 CACCCCCCGT GGCCCGGGCT CCGGAGGCCG CCGGCGGAGG CAGCCGTTCC GAGGATTATT 780
 60 CGTCTTCTCC CCATTCCGCT GCCGCCGCTG CCAGGCCTCT GGCTGCTGAG GAGAAGCAGG 840
 CCCAGTCGCT GCAACCATCC AGCAGCCGCC GCAGCAGCCA TTACCCGGCT GCGGTCCAGA 900
 GCCAAGCGGC GGCAGAGCGA GGGGCATCAG CTACCGCCAA GTCCAGAGCC ATTTCCATCC 960
 65 TGCAGAAGAA GCCCCGCCAC CAGCAGCTTC TGCCATCTCT CTCCTCCTTT TTCTTCAGCC 1020

ES 2 361 044 T3

	ACAGGCTCCC AGACATGACA GCCATCATCA AAGAGATCGT TAGCAGAAAC AAAAGGAGAT	1080
	ATCAAGAGGA TGGATTGAC TTAGACTTGA CCTGTATCCA TTTCTGCGGC TGCTCCTCTT	1140
5	TACCTTTCTG TCACTCTCTT AGAACGTGGG AGTAGACGGA TGCGAAAATG TCCGTAGTTT	1200
	GGGTGACTAT AACATTTAAC CCTGGTCAGG TTGCTAGGTC ATATATTTTG TGTTTCC	1257

10 (2) INFORMACIÓN PARA LA SEC. ID. Nº: 9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1.084 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 9:

	GAGACATAGC CAGCTCTTAA ATCTGACTTC CAGATTTTCA CTGTGTCTTC TTTTTTCTGT	60
	AACGTGTTGC CTTTTTTAGC CATGAAAAAT TAGAAGTTGA ACTCTTGTCT TTTCAGGCAG	120
25	GTGTCAATTT TGGGGTTTTG TTTTGATTTT TGGTTTTTGA CATAAAGTAC TTTAGTTCTG	180
	TGATGTATAA ACCGTGAGTT TCTGTTTTTC TCATATACCT GAATACTGTC CATGTGGAAG	240
30	TTACCTTTTA TCTTTACCAG TATTAACACA TAAATGGTTA TACATAAATA CATTGACCAC	300
	CTTTTATTAC TCCAGCTATA GTGGGGAAAG CTTTCTTTTC ATAAGTAGCT AATGTTTTAA	360
35	AAAGTATTCT TTTAGTTTGA TTGCTGCATA TTTCAGATAT TTCTTTCCTT AACTAAAGTA	420
	CTCAGATATT TATCCAAACA TTATTGCTAT GGGATTTCCCT GCAGAAAGAC TTGAAGGCGT	480
40	ATACAGGAAC AATATTGATG ATGTAGTAAG GTAAGAATGC TTTGATTTTC TATTTCAAAT	540
	ATTGATGTTT ATATTCATGT TGTGTTTTCA TTTAGAAAAG ATTTCTAAGC CACAGAAAAA	600
45	GATACTTTGT GATGTAAACT ATTATTGTAG TGCTCTATAA TCATTTTTTG GCTTACCGTA	660
	CCTAATGGAC TTCAGGGGGA TACAGTTCAT TTGATAAGAA CTGACCTTAT ACATTACATA	720
50	ATCAGGTACT TATGTGATAT CATTTCCCTGG ACTCCATAAA ATGCTGGTCA CCAGGTTTAA	780
	TACCTGGATT CCATTACAGT GTGATTTTTG TCTTATTTCA TAGTTGGGGA TTAGGCTTAA	840
55	AATCCTAGAG TGGATTTATT CAGTTAAAT TATTCACACT AAGATGTGAT GACTAATACT	900
	GTATATTTTT ATGTAGACCA AATTTTAAGG TACCACTGTG CATATGTTAC CAACTACCTG	960
60	AAGAATATTT GGTGGTACA GAATATATAA AGGAATCGCT GGTGTTCCAA GGCTAATCCA	1020
	GTTTTATAAT TTTGCATAAT TTCCTAACTG CGAATATCAT TTATTTAAAC AATTTATTCT	1080
	CCAG	1084

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 10:

65 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1.104 pares de bases
- (B) TIPO: ácido nucleico

ES 2 361 044 T3

(C) ESTADO DE LA CADENA: sencilla

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC.ID. Nº: 10:

5	GAATTAATAG TTAGTACGTG GATCTTTCAA ATATCAAAAG TTTTCAGTTT GATGGGAAAA	60
	TGATGTCTGA ATTTTCAGGG TTATTTTAA GAGTACTTGA TTATGACTGT CTTGTAAATC	120
10	TCTATGAGCT AGGTATACTT GCACTAAATG CTAATGCTTT TTAAAGAAGT TATGTCTTAA	180
	TATTCAGTCT CATTATGTTA GGTGGAAGAT AGAAGATTAT GAAAATATTC TCTGAAAAGC	240
15	TCTGGTTTTA CTTAGATTG TATAAATCTG TGTAATGTAA TAATTATTTA AGAATGACAT	300
	GATTACTACT CTAACCCAT AGAAGGGGTA TTTGTTGGAT TATTTATTTT CACTTAAATG	360
20	GTATTTGAGA TTAGGAAAA GAAAATCTGT CTTTGGTTT TTCTTGATAG TATTAATGTA	420
	ATTTCAAATG TTAGCTCATT TTTGTTAATG GTGGCTTTTT GTTTGTTTGT TTTGTTTTAA	480
25	GTTTTTGGGA TTCAAAGCAT AAAAACCATT ACAAGATATA CAATCTGTAA GTATGTTTTC	540
	TTATTTGTAT GCTTGCAAAT ATCTTCTAAA ACAACTATTA AGTGAAAGTT ATCTGCTTGT	600
30	TAGAGTGAGG TAGAGTTAAA GATACATTTT AACAGAATTG TATTCCTAAA CCGATTAAGT	660
	CAAGAAGTCC AAGAGCATTG TTAGATCATT TAGAAAGTGT AGTGATGAGG TAAAACATTG	720
35	TTGGCACAGA TTCATGTTAC TTGATCTGCT TTAATGACT TGGCATCTAG CCCATATTTG	780
	AGCCCATAAC CGTGTGGTAA TTTGAAGTGT AATTCACAGT AGAGCTTCTG TTAAAGCACT	840
40	AATAGCATCT TCCATGGAGG TATACTTCAG AGTGAATATA ATTTGTTTA TCCTGTGICT	900
	CTAGAGCTAT TGACTGAAAA AGCTGTTAGG GCATTCCTTA ACTGTACATC ACCTAAGTTA	960
45	TTTAAAATTG CTGAATTAAG TGGCTTGTCT TGTCTAGACA GATTTAAGG ACTGCCACC	1020
	TGATTGATAG AACTAGTTGA CCTTATCTTT AACTTTTTGT TTTCTTTGA CTTGGGATAA	1080
50	AAGTTGAAAA GGTAAAAGGA AGGA	1104

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 656 pares de bases

(B) TIPO: ácido nucleico

(C) ESTADO DE LA CADENA: sencilla

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 11:

55	TTGCATACAC TTAATCTTTT AAGCTTTGGT TTTATTATTA TAATATGGGG GTGATAACAG	60
	TATCTACTTA ATAGAATTCT TGTTATTAAC ATGAAATAAT TAATGTTAAA CACAGCATAA	120
60	TATGTGTCAC ATTATAAAGA TTCAGGCAAT GTTTGTTAGT ATTAGTACTT TTTTTCTTC	180
	CTAAGTGCAA AAGATAACTT TATATCACTT TAAACTTTT CTTTGTAGTTG TGCTGAAAGA	240
65	CATTATGACA CCGCCAAATT TAATTGCAGA GGTAGGTATG AATGTACTGT ACTATGTTGT	300
	ATAACTTAAA CCCGATAGAC TGTATCTTAC TGTCATAACA ATAATGAGTC ATCCAGATTA	360

ES 2 361 044 T3

	TCGAGTGAGA TACATATTTA TCTTAAGAAT TATCTTTAAA AATTTCAAAA ATTTTAATTT	420
	TACTGTTGTG TTTTAGGAAA AAGTATTGCA TAAAGCTATT AATATTGTCA GGAAGACTAA	480
5	AGTGCAGCAT AGACTAAGCA ATCAGGAAAA TTCCTAGACT AAAAATAGTA TAAGGAGAGG	540
	GTTTACCTAC TATTTGAGGC AGTTGGTCTA ATAGTAAGCA ATCACAGGGA GGAAAGCAGA	600
10	AACTACTTAA CTCTTCTGTG TTGAGGAATG ACATAAAAGG TATGAAAGGA TATAAC	656

(2) INFORMACIÓN PARA LA SEC. ID. N°: 12:

- 15 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 808 pares de bases
 - (B) TIPO: ácido nucleico
 - 20 (C) ESTADO DE LA CADENA: sencilla
 - (D) TOPOLOGÍA: lineal
- (ix) RASGO:
- 25 (A) NOMBRE/CLAVE: base modificada
 - (B) LOCALIZACIÓN: 463..754
 - (C) OTRA INFORMACIÓN: /nota="N=C, G, A ó T"

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 12:

30	ATACATTATT TTTCTCTGGA ATCCAGTGTT TCTTTTAAAT ACCTGTTAAG TTTGTATGCA	60
	ACATTTCTAA AGTTACCTAC TTGTTAATTA AAAATTCAAG GGTTTTTTTTT TCTTATTCTG	120
35	AGGTTATCTT TTTACCACAG TTGCACAATA TCCTTTTGAA GACCATAACC CACCACAGCT	180
	AGAACTTATC AAACCCTTTT GTGAAGATCT TGACCAATGG CTAAGTGAAG ATGACAATCA	240
40	TGTTGCAGCA ATTCACTGTA AAGCTGGAAA GGGACGAACT GGTGTAATGA TATGTGCATA	300
	TTTATTACAT CGGGGCAAAT TTTTAAAGGC ACAAGAGGCC CTAGATTTCT ATGGGGAAGT	360
45	AAGGACCAGA GACAAAAGG TAAGTPATTT TTTGATGTTT TTCCTTTCCT CTCCTGGAT	420
	CTGAGAATTT AFTGGAAAAC AGATTTTGGG TTTCTTTTTT TCNNNNNNNNN NNNNNNNNNN	480
	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	540
50	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	600
	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	660
55	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	720
	NNNNNNNNNN NTCCTCCCTC CCCACCCTCA GTCNCTGGAA AACAGGTTTT AAAGATAGTT	780
60	GCTAATCCTT ATTTCTTCTA AATTTTTA	808

(2) INFORMACIÓN PARA LA SEC. ID. N°: 13:

- 65 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 670 pares de bases
 - (B) TIPO: ácido nucleico

ES 2 361 044 T3

(C) ESTADO DE LA CADENA: sencilla

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 13:

5	ATATGATAAT TGTTTTAAGG GAGGAGAGTT ATTCTGATAT CCTTGTATTG ATATTGCTCT	60
	TATTTATTAT TGAGCTGGAT TTAAGTATTA ATCATTTAAG GTCAAATTC TAATGTATAA	120
10	TATGTTCTTA AATGGCTACG ACCCAGTTAC CATAGCAATT TAGTGAAATA ACTATAATGG	180
	AACATTTTTT TTCAATTTGG CTTCTCTTTT TTTTCTGTCC ACCAGGGAGT AACTATTCCC	240
15	AGTCAGAGGC GCTATGTGTA TTATTATAGC TACCTGTAA AGAATCATCT GGATTATAGA	300
	CCAGTGGCAC TGTTGTTTCA CAAGATGATG TTGAAACTA TTCCAATGTT CAGTGGCGGA	360
20	ACTTGCAGTA AGTGCTTGG AATTCTCATC CTTCCATGTA TTGGAACAGT TTTCTTAACC	420
	ATATCTAGAA GTTTACATAA AAATTTAGAA AAGAAATTA CCACATTTGA AATTTATGCA	480
25	GGAGACTATA TTTCTGAAGC ATTTGAACAA ATTAATTAGC TTTGTTGTTT AACTCATTGG	540
	GCTAAAGAAG CCAAAGCAA TGGGTTTTAA TGTAGTCGAA GCCAAATTAT ATTTATGAAA	600
30	GAAATATTCT GTGTTATAAC CCACCAAATA CAGCCCAATT TCTGACTAGA TGTATGGAAG	660
	AACCTGTCCC	670

(2) INFORMACIÓN PARA LA SEC. ID. N°: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 35 (A) LONGITUD: 661 pares de bases
 (B) TIPO: ácido nucleico
 (C) ESTADO DE LA CADENA: sencilla
 40 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 14:

45	ATATTTTTAT TTCATTTATT TCAGTTGATT TGCTTGAGAT CAAGATTGCA GATACAGAAT	60
	CCATATTTTCG TGTATATTGC TGATATTAAT CATTAAAATC GTTTTTGACA GTTTGACAGT	120
50	TAAAGGCATT TCCTGTGAAA TAATACTGGT ATGTATTTAA CCATGCAGAT CCTCAGTTTG	180
	TGGTCTGCCA GCTAAAGGTG AAGATATATT CCTCCAATTC AGGACCCACA CGACGGGAAG	240
55	ACAAGTTCAT GTACTTTGAG TTCCCTCAGC CGTTACCTGT GTGTGGTGAT ATCAAAGTAG	300
	AGTTCTTCCA CAAACAGAAC AAGATGCTAA AAAAGGTTTG TACTTTACTT TCATTGGGAG	360
60	AAATATCCAA AATAAGGACA GATTAAAAGC TATATTTTAT TTTATGACAT GTAAGGAACT	420
	ATAATTTGTT TTCTATTAGA TCTGCAGGTG TTTTGCTTAC TCTGGCATTG GTGAGACATT	480
65	ATAAGGGTAA ATAATCCTGT TTGAAGGAAA AGGCCTTATG GCATTGTAAC ATTAGAGGAA	540
	TTTTTCTTAA CAAGGATGGT TAACTGAGAA GAAATTAGCA TGGGACCAAT ATTTTAAAAA	600
	TTTTTGGTCT ATAGGTAGAA ATGAGATCTG TTCTGTGGTC TTATGTAGTG ACACAAACCA	660
	C	661

ES 2 361 044 T3

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 739 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(ix) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 15:

5	GTGTTACCT TTATTCAGAA TATCAAATGA TAGTTTATTT TGTTGACTTT TTGCAAATGT	60
15	TTAACATAGG TGACAGATTT TCTTTTTTAA AAAAATAAAA CATCATTAAT TAAATATGTC	120
20	ATTTCAATTC TTTTCTTTT CTTTTTTTTT TTTTTTTAGG ACAAATGTT TCACTTTTGG	180
25	GTAATACAT TCTTCATACC AGGACCAGAG GAAACCTCAG AAAAAGTAGA AAATGGAAGT	240
30	CTATGTGATC AAGAAATCGA TAGCATTTGC AGTATAGAGC GTGCAGATAA TGACAAGGAA	300
35	TATCTAGTAC TTAATTTAAC AAAAAATGAT CTTGACAAAG CAAATAAAGA CAAAGCCAAC	360
40	CGATACTTTT CTCCAAATTT TAAGGTCAGT TAAATTAAC ATTTTGTGGG GGTGGTGAC	420
45	TTGTATGTAT GTGATGTGTG TTTAATTCTA GGAGTACAGC TGATGAAGAA CTTGCTTGAC	480
50	AAGTTTTTAA CTTATGTATT ATTTGGAAGC AGTGTTCAG TAGCAGTAAC ATGAAAGTTT	540
55	CTAATAAAAT ACCCAATGTA CACAGCGTCA AAAAAGCTGC ATTTTTCCTT TTCCTAATTC	600
60	TTTGTGTTT GCTGAAATCT GGGGCAAAGG TGC GGAGGG GGCTAAATGA CTGGGATATG	660
65	AAGTAGGAAT GGGAGAGGAA AGAAATAGAT GGGAACTCAG TCATTTGGGA ATGATTCATA	720
70	TGGAATGTTT TTAATGCTT	739

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 970 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTADO DE LA CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 16:

ES 2 361 044 T3

(B) LOCALIZACIÓN: 1

(C) OTRA INFORMACIÓN:/nota="puede ser o bien I=Isoleucina o V=Valina"

5 (ix) RASGO:

(A) NOMBRE/CLAVE: sitio modificado

(B) LOCALIZACIÓN: 4..8

(C) OTRA INFORMACIÓN:/nota="X=Cualquier aminoácido"

10

(ix) RASGO:

(A) NOMBRE/CLAVE: sitio modificado

(B) LOCALIZACIÓN: 10

15

(C) OTRA INFORMACIÓN:/nota="puede ser o bien S=Serina o T=Treonina"

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 18:

20

Ile His Cys Xaa Ala Gly Xaa Xaa Arg Ser Gly
1 5 10

25 (2) INFORMACIÓN PARA LA SEC. ID. Nº: 19:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 60 aminoácidos

30

(B) TIPO: aminoácido

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 19:

40

Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
1 5 10 15

Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
20 25 30

45

Ile Ala Met Gly Phe Pro Ala Glu Arg Asn Glu Gly Val Tyr Arg Asn
35 40 45

50

Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys
50 55 60

(2) INFORMACIÓN PARA LA SEC. ID. Nº: 20:

55

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 60 aminoácidos

(B) TIPO: aminoácido

60

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. Nº: 20:

65

ES 2 361 044 T3

1 Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
 1 5 10 15
 5 Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
 20 25 30
 10 Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
 35 40 45
 15 Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys
 50 55 60

(2) INFORMACIÓN PARA LA SEC. ID. N°: 21:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 60 aminoácidos

(B) TIPO: aminoácido

(C) ESTADO DE LA CADENA:

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 21:

1 Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
 1 5 10 15
 35 Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
 20 25 30
 40 Ile Ala Met Glu Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
 35 40 45
 50 Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys
 50 55 60