



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 328 448**

51 Int. Cl.:

C12N 15/55 (2006.01) **C12N 9/16** (2006.01)
C12N 5/10 (2006.01) **C07K 16/40** (2006.01)
C12Q 1/68 (2006.01) **G01N 33/50** (2006.01)
G01N 33/577 (2006.01) **G01N 33/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **01922897 .2**

96 Fecha de presentación : **30.03.2001**

97 Número de publicación de la solicitud: **1294899**

97 Fecha de publicación de la solicitud: **26.03.2003**

54

Título: **16836, Un miembro de la familia de la fosfolipasa C humana y sus usos.**

30

Prioridad: **31.03.2000 US 193921 P**

45

Fecha de publicación de la mención BOPI:
13.11.2009

45

Fecha de la publicación del folleto de la patente:
13.11.2009

73

Titular/es: **MILLENNIUM PHARMACEUTICALS, Inc.**
75 Sidney Street
Cambridge, Massachusetts 02139, US

72

Inventor/es: **Meyers, Rachel, A. y**
Hunter, John, J.

74

Agente: **Elzaburu Márquez, Alberto**

ES 2 328 448 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

16836, Un miembro de la familia de la fosfolipasa C humana y sus usos.

5 Solicitudes relacionadas

Esta solicitud reivindica la prioridad de la solicitud de Patente Provisional de Estados Unidos con el número 60/193.921 presentada el 31 de marzo de 2000.

10 Antecedentes de la invención

La fosfolipasa específica de fosfoinosítidos (PI-PLC) media las acciones celulares de una diversidad de hormonas, neurotransmisores y factores de crecimiento. La activación de PI-PLC es una de las primeras respuestas a diversas señales extracelulares. La activación dependiente de agonista de PI-PLC ocasiona la hidrólisis de fosfatidilinositol 4,5-bisfosfato (PIP₂) de la membrana, generando los segundos mensajeros inositol 1,4,5-trisfosfato (IP₃) y diacilglicerol (DAG). El IP₃ se une a receptores intracelulares específicos para inducir la movilización de Ca²⁺, mientras que el DAG media la activación de una familia de isozimas de la proteína quinasa C. Este proceso catalítico se regula fuertemente por fosforilación reversible y unión de proteínas reguladoras (Rhee *et al.* (1997) *J. Biol. Chem.* 272:15045-15048).

En mamíferos hay al menos seis isoformas diferentes de PI-PLC, que difieren en su estructura de dominios, regulación y distribución tisular. Basándose en el tamaño molecular, la inmuno-reactividad y la secuencia de aminoácidos, se han clasificado varios subtipos. En general, la identidad de secuencia entre los subtipos es baja, aunque todas las isoformas comparten dos dominios conservados que constituyen el dominio catalítico de PLC, denominado X e Y: la región X abarca aproximadamente 170 restos y la región Y aproximadamente 260. El orden de estas dos regiones es siempre igual (NH₂-X-Y-COOH), pero la separación es variable. En los subtipos de PLC-beta, los dominios X e Y están separados por un tramo de 70-120 aminoácidos ricos en Ser, Thr y restos ácidos, mientras que sus 450 restos C-terminales son ricos en restos básicos. En los subtipos de PLC-gamma, hay un inserto de más de 400 restos que contienen un dominio SH3 y dos dominios SH2. Las PLC muestran poca similitud en la región N-terminal de 300 restos que precede al dominio X. Las PI-PLC tienen un dominio C2 C-terminal del dominio catalítico. Se cree que el dominio C2 está implicado en la unión de fosfolípidos dependiente de calcio (Rhee *et al.* (1997) *J. Biol. Chem.* 272:15045-15048).

35 Compendio de la invención

La presente invención se basa, en parte, en el descubrimiento de un nuevo miembro de la familia de fosfolipasa C, denominado en la presente memoria "16836". La secuencia de nucleótidos del ADNc que codifica 16836 se muestra en la SEC ID N°: 1, y la secuencia de aminoácidos de un polipéptido 16836 se muestra en la SEC ID N°: 2. Además, las secuencias de nucleótidos de la región codificante se representan en la SEC ID N°: 3.

Por consiguiente, en un aspecto, la invención se refiere a una molécula de ácido nucleico que codifica una proteína o polipéptido 16836, por ejemplo, una parte biológicamente activa de la proteína 16836. En una realización preferida, la molécula de ácido nucleico aislada codifica un polipéptido que tiene la secuencia de aminoácidos de la SEC ID N°: 2. En otras realizaciones, la invención proporciona moléculas de ácido nucleico de 16836 aisladas que tienen la secuencia de nucleótidos mostrada en la SEC ID N°: 1, SEC ID N°: 3, o la secuencia del inserto de ADN del plásmido depositado con el Número de Acceso de la ATCC 1774. En otras realizaciones, la invención proporciona moléculas de ácido nucleico que son sustancialmente idénticas a la secuencia de nucleótidos mostrada en la SEC ID N°: 1, SEC ID N°: 3, o la secuencia del inserto de ADN del plásmido depositado con el Número de Acceso de la ATCC 1774. En otras realizaciones, la invención proporciona una molécula de ácido nucleico que hibrida en las condiciones rigurosas descritas en la presente memoria con una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEC ID N°: 1, SEC ID N°: 3, o la secuencia del inserto de ADN del plásmido depositado con el Número de Acceso de la ATCC 1774, donde el ácido nucleico codifica una proteína 16836 de longitud completa o un fragmento activo de la misma.

En un aspecto relacionado, la invención proporciona además construcciones de ácido nucleico que incluyen una molécula de ácido nucleico 16836 descrita en la presente memoria. En ciertas realizaciones, las moléculas de ácido nucleico de la invención se unen operativamente a secuencias reguladoras nativas o heterólogas. También se incluyen vectores y células hospedadoras que contienen las moléculas de ácido nucleico 16836 de la invención, por ejemplo, vectores y células hospedadoras adecuadas para producir moléculas de ácido nucleico y polipéptidos 16836.

En otro aspecto relacionado, la invención, proporciona fragmentos de ácido nucleico adecuados como cebadores o sondas de hibridación para la detección de ácidos nucleicos que codifican 16836.

En otro aspecto relacionado, se proporcionan moléculas de ácido nucleico aisladas que son antisentido con respecto a una molécula de ácido nucleico que codifica 16836.

ES 2 328 448 T3

En otro aspecto, la invención se refiere a polipéptidos 16836, y fragmentos biológicamente activos o antigénicos de los mismos que son útiles, por ejemplo, como reactivos o dianas en ensayos aplicables al tratamiento y diagnóstico de trastornos relacionados o mediados por 16836. En otra realización, la invención proporciona polipéptidos 16836 que tienen una actividad 16836. Son polipéptidos preferidos proteínas 16836 que incluyen al menos uno de: un dominio de factor intercambiador de nucleótidos de guanina de Ras (RasGEF), un dominio "X" de fosfolipasa C específico de fosfatidilinositol, un dominio "Y" de fosfolipasa C específico de fosfatidilinositol, un dominio C2 o un dominio de asociación de Ras (RalGDS/AF-6) y, preferiblemente, que tienen actividad 16836, por ejemplo, una actividad 16836 como se describe en la presente memoria.

En otras realizaciones, la invención proporciona polipéptidos 16836, por ejemplo, un polipéptido 16836 que tiene la secuencia de aminoácidos mostrada en la SEC ID N°: 2 o la secuencia de aminoácidos codificada por el inserto de ADNc del plásmido depositado con el Número de Acceso de ATCC 1774; una secuencia de aminoácidos que es sustancialmente idéntica a la secuencia de aminoácidos mostrada en la SEC ID N°: 2 o la secuencia de aminoácidos codificada por el inserto de ADNc del plásmido depositado con el Número de Acceso de ATCC 1774; o una secuencia de aminoácidos codificada por una molécula de ácido nucleico que tiene una secuencia de nucleótidos que hibrida en las condiciones rigurosas descritas en la presente memoria con una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEC ID N°: 1, SEC ID N°: 3, o la secuencia del inserto de ADN del plásmido depositado por el Número de Acceso de la ATCC 1774, donde el ácido nucleico codifica una proteína 16836 de longitud completa o un fragmento activo de la misma.

En un aspecto relacionado, la invención proporciona además construcciones de ácido nucleico que incluyen una molécula de ácido nucleico de 16836 descrita en la presente memoria.

En un aspecto relacionado, la invención proporciona polipéptidos 16836 o fragmentos unidos operativamente a polipéptidos que no son 16836 para formar proteínas de fusión.

En otro aspecto, la invención se refiere a anticuerpos y fragmentos de unión a antígeno de los mismos, que reaccionan con, o más preferiblemente se unen específicamente a polipéptidos 16836.

En otro aspecto, la invención proporciona métodos para seleccionar agentes, por ejemplo, compuestos, que modulan la expresión o actividad de los polipéptidos o ácidos nucleicos de 16836.

En otro aspecto, la invención proporciona un proceso para modular la expresión o actividad de polipéptidos o ácidos nucleicos de 16836, por ejemplo, usando los compuestos seleccionados. En la descripción también se incluyen los métodos que implican el tratamiento de afecciones relacionadas con una actividad o expresión aberrante de los polipéptidos o ácidos nucleicos de 16836, tales como afecciones que implican una proliferación o diferenciación celular aberrante o deficiente.

En otro aspecto, la invención proporciona métodos para modular la actividad de una célula que expresa 16836, por ejemplo, una célula hiperproliferativa o aberrante que expresa 16836. El método incluye poner en contacto la célula con un compuesto (por ejemplo, un compuesto identificado usando los métodos descritos en la presente memoria) que modula la actividad, o expresión, del polipéptido o ácido nucleico de 16836, modulando de esta manera la actividad celular. Preferiblemente, la actividad de la célula se modula inhibiendo la proliferación o induciendo la diferenciación o destrucción de la célula que expresa 16836.

En una realización preferida, la etapa de contacto es eficaz *in vitro* o *ex vivo*. En otras realizaciones, la etapa de contacto se realiza *in vivo*, por ejemplo, en un sujeto (por ejemplo, un mamífero, por ejemplo, un ser humano), como parte de un protocolo terapéutico o profiláctico.

En una realización preferida, la célula que expresa 16836 se encuentra en un tejido o célula cancerosa, por ejemplo un tumor sólido, o un tumor de tejido blando, o una lesión metastásica. En un ejemplo, la célula que expresa 16836 se encuentra en un tejido canceroso del pulmón, colon o mama.

En otras realizaciones, la célula que expresa 16836 es una célula ósea (por ejemplo un osteoclasto), una célula cerebral, una célula testicular, una célula cardíaca, una célula de músculo esquelético o una célula inmune, por ejemplo, una célula de un linaje mielóide, linfóide o eritroide, o una célula precursora de la misma.

En una realización preferida, el agente, por ejemplo, el compuesto, es un inhibidor de un polipéptido 16836. Preferiblemente, el inhibidor se elige entre un péptido, un fosfopéptido, un peptidomimético, por ejemplo un análogo de fosfonato de un sustrato peptídico, una molécula orgánica pequeña, una molécula inorgánica pequeña y un anticuerpo (por ejemplo, un anticuerpo conjugado con un resto terapéutico seleccionado entre una citotoxina, un agente citotóxico y un ion metálico radiactivo).

En una realización preferida, el agente, por ejemplo, el compuesto, es un inhibidor de un ácido nucleico de 16836, por ejemplo, una molécula antisentido, una ribozima o una molécula en triple hélice.

En una realización preferida, el agente, por ejemplo, el compuesto, se administra en combinación con un agente citotóxico. Los ejemplos de agentes citotóxicos incluyen agentes anti-microtúbulos, un inhibidor de la topoisomerasa

ES 2 328 448 T3

I, un inhibidor de la topoisomerasa II, un antimetabolito, un inhibidor mitótico, un agente alquilante, un agente intercalante, un agente capaz de interferir con una ruta de transducción de señales, un agente que promueve la apoptosis o necrosis, y radiación.

5 En otro aspecto, la invención se refiere a un método para tratar o prevenir, en un sujeto, un trastorno caracterizado por la actividad o expresión aberrante de un ácido nucleico o polipéptido 16836. En una realización, el método incluye administrar al sujeto una cantidad eficaz de un agente, por ejemplo, un compuesto que modula la actividad o expresión de un polipéptido o ácido nucleico de 16836 de tal forma que se mejore o prevenga el trastorno.

10 En una realización preferida, el trastorno es un trastorno de proliferación o diferenciación celular, por ejemplo, cáncer de pulmón, cáncer de colon o cáncer de mama.

En una realización preferida, el sujeto es un ser humano. Por ejemplo, el sujeto es un ser humano que padece o tiene riesgo de padecer un trastorno de proliferación y/o diferenciación celular como se describe en la presente memoria.

15 En otras realizaciones, el sujeto es un ser humano que padece o con riesgo de padecer un trastorno óseo, un trastorno cerebral, un trastorno cardiovascular, un trastorno reproductivo, un trastorno del músculo esquelético o un trastorno inmune como se describe en la presente memoria.

20 En una realización preferida, el agente, por ejemplo, el compuesto, es un péptido, un fosfopéptido, una molécula pequeña, un anticuerpo o cualquier combinación de los mismos. En otra realización, el agente es una molécula antisentido, una ribozima, una molécula en triple hélice, un ácido nucleico de 16836 o cualquier combinación de los mismos.

25 En una realización preferida, el agente, por ejemplo, el compuesto, se administra en combinación con un agente citotóxico. Los ejemplos de agentes citotóxicos incluyen agentes anti-microtúbulos, un inhibidor de la topoisomerasa I, un inhibidor de la topoisomerasa II, un antimetabolito, un inhibidor mitótico, un agente alquilante, un agente intercalante, un agente capaz de interferir con una ruta de transducción de señales, un agente que promueve la apoptosis o necrosis, y radiación.

30 La invención también proporciona ensayos para determinar la actividad o la presencia o ausencia de polipéptidos o moléculas de ácido nucleico de 16836 en una muestra biológica, incluyendo para el diagnóstico de enfermedades. Preferiblemente, la muestra biológica incluye una célula o tejido canceroso o precanceroso. Por ejemplo, el tejido canceroso puede ser un tumor sólido, un tumor de tejido blando, o una lesión metastásica. En un ejemplo, el tejido canceroso procede del pulmón, colon, o mama. En otras realizaciones, la muestra biológica contiene una célula ósea (por ejemplo un osteoclasto), una célula cerebral, una célula testicular, una célula cardíaca, una célula de músculo esquelético o una célula inmune, por ejemplo, una célula de un linaje mieloide, linfoide, o eritroide, o una célula precursora de la misma.

40 En otro aspecto, la invención proporciona ensayos para determinar la presencia o ausencia de una alteración genética en una molécula de polipéptido o de ácido nucleico de 16836, incluyendo para el diagnóstico de enfermedades. Preferiblemente, la muestra biológica incluye una célula o tejido canceroso o precanceroso. Por ejemplo, el tejido canceroso puede ser un tumor sólido, un tumor de tejido blando, o una lesión metastásica. En un ejemplo, el tejido canceroso procede del pulmón, colon, o mama. En otras realizaciones, la muestra biológica contiene una célula ósea (por ejemplo, un osteoclasto), una célula cerebral, una célula testicular, una célula del músculo esquelético o una célula inmune, por ejemplo, una célula de un linaje mieloide linfoide o eritroide, o una célula precursora de la misma.

50 En otro aspecto, la invención, se refiere a un método de diagnóstico o de determinación del estadio de un trastorno mediado por 16836, por ejemplo, un trastorno de proliferación y/o diferenciación celular, en un sujeto. El método incluye evaluar la expresión o actividad de un ácido nucleico o polipéptido 16836, diagnosticando o determinando el estadio del trastorno de esta manera. En una realización preferida, la expresión o actividad se compara con un valor de referencia, por ejemplo, una diferencia en el nivel de expresión o actividad del ácido nucleico o polipéptido 16836 con respecto a un sujeto normal o a un grupo de sujetos normales indica el trastorno o el estadio del trastorno.

55 En una realización preferida, el sujeto es un ser humano. Por ejemplo, el sujeto es un ser humano que padece o tiene riesgo de padecer un trastorno de proliferación y/o diferenciación celular como se describe en la presente memoria.

60 En una realización, la etapa de evaluación se realiza *in vitro* o *ex vivo*. Por ejemplo, se obtiene una muestra, por ejemplo, una muestra de sangre o tejido, una biopsia, del sujeto. Preferiblemente, la muestra contiene una célula cancerosa.

65 También se incluye en la descripción que la etapa de evaluación se realiza *in vivo*. Por ejemplo, por medio de la administración al sujeto de un agente marcado de forma detectable que interacciona con el ácido nucleico o polipéptido asociado con 16836, de tal forma que se genere una señal relativa al nivel de actividad o expresión del ácido nucleico o polipéptido 16836.

ES 2 328 448 T3

En realizaciones preferidas, el método se realiza: sobre una muestra de un sujeto, una muestra de un ser humano; por ejemplo, una muestra de un paciente que padece o tiene riesgo de padecer un trastorno de proliferación y/o diferenciación celular como se describe en la presente memoria; para determinar si el individuo del que se recogió el ácido nucleico o la proteína diana debe recibir un fármaco u otro tratamiento; para diagnosticar en un individuo un trastorno o la predisposición a la resistencia al tratamiento o para determinar el estadio de una enfermedad o trastorno.

En otro aspecto adicional, la invención proporciona métodos para evaluar la eficacia de un tratamiento de un trastorno, por ejemplo, un trastorno de proliferación y/o diferenciación celular, por ejemplo, un cáncer. El método incluye: tratar a un sujeto, por ejemplo, a un paciente o un animal, con un protocolo bajo evaluación (por ejemplo, tratando a un sujeto con uno o más de: quimioterapia, radiación y/o un compuesto identificado usando los métodos descritos en la presente memoria); y evaluar la expresión de un ácido nucleico o polipéptido 16836 antes y después del tratamiento. Un cambio, por ejemplo, una reducción o aumento en el nivel de un ácido nucleico (por ejemplo un ARNm) o polipéptido 16836 después de tratamiento, con respecto al nivel de expresión previo al tratamiento, indica la eficacia del tratamiento del trastorno.

En una realización preferida, la etapa de evaluación incluye obtener una muestra (por ejemplo una muestra de tejido, por ejemplo una biopsia o una muestra de fluido) a partir del sujeto, antes y después del tratamiento y comparar el nivel de expresión de un ácido nucleico (por un ejemplo un ARNm) o polipéptido 16836 antes y después del tratamiento.

En otro aspecto, la invención proporciona métodos para evaluar la eficacia de un agente terapéutico o profiláctico (por ejemplo un agente anti-neoplásico). El método incluye: poner en contacto una muestra con un agente (por ejemplo, un compuesto identificado usando los métodos descritos en la presente memoria, un agente citotóxico) y evaluando la expresión o actividad de un ácido nucleico o polipéptido 16836 en la muestra antes y después de la etapa de contacto. Un cambio, por ejemplo, una reducción o aumento del nivel de ácido nucleico (por ejemplo ARNm) o polipéptido 16836 en la muestra obtenida después de la etapa de contacto, con respecto al nivel de expresión en la muestra antes de la etapa de contacto, indica la eficacia del agente. El nivel de expresión de ácido nucleico o polipéptido 16836 puede detectarse por cualquier método descrito en la presente memoria.

En una realización preferida, la muestra incluye células obtenidas a partir de un tejido canceroso donde se obtiene un polipéptido o ácido nucleico de 16836.

En una realización preferida, la muestra es una muestra de tejido (por ejemplo una biopsia), un fluido corporal, una célula cultivada (por ejemplo, una línea de células tumorales).

En otras realizaciones, la muestra incluye una célula ósea (por ejemplo un osteoclasto), una célula testicular, una célula cardíaca, una célula de músculo esquelético o una célula inmune, por ejemplo, una célula de un linaje mielóide, linfóide o eritroide, o una célula precursora de la misma.

En otro aspecto, la invención se refiere a un método para identificar un agente que modula actividad o expresión de un polipéptido o ácido nucleico de 16836. El método incluye las etapas de: poner en contacto el polipéptido o ácido nucleico de 16836 con un agente; y determinar el efecto del agente sobre la actividad o expresión del polipéptido o ácido nucleico. En una realización, el método incluye poner en contacto un polipéptido 16836 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido 16836 para catalizar la hidrólisis de fosfatidilinositol. En otra realización, el método incluye poner en contacto un polipéptido 16836 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido 16836 de asociarse con una proteína Ras. En otra realización, el método incluye poner contacto un polipéptido 16836 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido 16836 de mediar la actividad de intercambio de nucleótidos de guanina. El agente puede ser un péptido, un fosfopéptido, una molécula pequeña, un anticuerpo o cualquier combinación de los mismos. Además, el agente puede ser una molécula antisentido, una ribozima, una molécula de triple hélice, un ácido nucleico de 16836 o cualquier combinación de los mismos.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada y a partir de las reivindicaciones.

Breve descripción de los dibujos

Las Figuras 1A-1J representan una secuencia de ADNc (SEC ID N°: 1) y la secuencia de aminoácidos prevista (SEC ID N°: 2) de 16836 humano. La fase de lectura abierta iniciada con metionina de 16836 humano (sin las regiones 5' y 3' no traducidas) se muestra como secuencia codificante SEC ID N°: 3.

La Figura 2 representa un gráfico de hidropatía de 16836 humano. Los restos hidrófobos relativos se muestran encima de la línea horizontal de trazos, y los restos hidrófilos relativos están debajo de la línea horizontal de trazos. Los restos de cisteína (cys) se indican por líneas verticales cortas justo debajo del gráfico de hidropatía. Se indican los números correspondientes a la secuencia de aminoácidos de polipéptido 16836 humano. Los polipéptidos de la inven-

ES 2 328 448 T3

ción incluyen fragmentos que incluyen: todo o parte de una secuencia hidrófoba, es decir, una secuencia por encima de la línea de trazos, por ejemplo, la secuencia desde aproximadamente el aminoácido 200-230, desde aproximadamente 430-450, y desde aproximadamente 830-840 de la SEC ID N°: 2; todo o parte de una secuencia hidrófila, es decir, una secuencia por debajo de la línea de trazos, por ejemplo, la secuencia desde aproximadamente el aminoácido 330-350, desde aproximadamente 610-630, y desde aproximadamente 1120-1140 de la SEC ID N°: 2; una secuencia que incluye una Cys, o un sitio de glicosilación.

La Figura 3A representa un alineamiento del dominio del factor intercambiador de nucleótidos de guanina de Ras (RasGEF) de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de SMART. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 4), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 35-338 de la SEC ID N°: 2.

La Figura 3B representa un alineamiento del dominio PI-PLC-X de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de PFAM. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 5), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 900-1048 de la SEC ID N°: 2.

La Figura 3C representa un alineamiento del dominio PI-PLC-Y de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de PFAM. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 6), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 1171-1184 de la SEC ID N°: 2.

La Figura 3D representa un alineamiento del dominio PI-PLC-Y de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de PFAM. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 7), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 1261-1353 de la SEC ID N°: 2.

La Figura 3E representa un alineamiento del dominio C2 de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de PFAM. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 8), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 1378-1460 de la SEC ID N°: 2.

La Figura 3F representa un alineamiento del dominio de asociación de Ras (RalGDS/AF-6) de 16836 humano con una secuencia de aminoácidos consenso procedente de un modelo oculto de Markov (HMM) de PFAM. La secuencia superior es la secuencia de aminoácidos consenso (SEC ID N°: 9), mientras que la secuencia de aminoácidos inferior corresponde a los aminoácidos 1640-1745 de la SEC ID N°: 2.

La Figura 4 es un gráfico de barras que representa la expresión del ARN de 16836 en un panel de tejidos humanos normales, incluyendo células óseas, hígado fetal, médula ósea, tráquea, piel, músculo esquelético, testículos, detectado usando el análisis TaqMan.

La Figura 5 es un gráfico de barras que representa la expresión del ARN de 16836 en un panel de células pulmonares normales y células de tumor de pulmón, con respecto a células normales. La expresión se detectó usando el análisis TaqMan.

La Figura 6A es un gráfico de barras que representa la asociación de los niveles de expresión de 16836 con mutaciones de *kras* en muestras de tumor (T) y normales (N). Las columnas indicadas como "GTT" corresponden a muestras de *kras* mutante, mientras que las columnas indicadas como "GGT" corresponden a muestras de tipo silvestre. La expresión se detectó usando el análisis TaqMan.

La Figura 6B es un gráfico de barras que representa la correlación de la expresión de 16836 y mutaciones de *kras* en líneas de células tumorales. Las columnas indicadas como "kras" corresponden a muestras con *kras* mutante; las columnas indicadas como "wt" corresponden a muestras de tipo silvestre. La expresión se detectó usando el análisis TaqMan.

La Figura 7 es un gráfico de barras que representa la expresión de 16836 en tejidos de mama, pulmón, colon e hígado normales y malignos.

La Figura 8 es un gráfico de barras que representa la expresión de 16836 en muestras de tejido de mama normal y maligno, incluyendo carcinoma ductal infiltrante (IDC), carcinoma ductal (DC), carcinoma lobular invasivo (ILC) y carcinoma ductal *in situ* (DCIS).

La Figura 9 es un gráfico de barras que representa la expresión de 16836 en líneas de células cancerosas, incluyendo líneas de células de cáncer de mama humano, líneas de células de cáncer de colon humano y líneas de células de cáncer de pulmón humano.

Descripción detallada

La secuencia de 16836 humano (Figs. 1A-1J; SEC ID N°: 1), que tiene una longitud de aproximadamente 10.172 nucleótidos incluyendo regiones no traducidas, contiene una secuencia codificante prevista de aproximadamente 5.430 nucleótidos, incluyendo el codon de terminación (nucleótidos indicados como codificantes de la SEC ID N°: 1 en las Figs. 1A-1J; SEC ID N°: 3). La secuencia codificante incluye una proteína de 1809 aminoácidos (SEC ID N°: 2).

16836 humano contiene las siguientes regiones o características estructurales: un dominio de factor intercambiable de nucleótidos de guanina de Ras (RasGEF) (Fig. 3A; identificador SMART RasGEF) localizado aproximadamente en los restos aminoacídicos 35-338 de la SEC ID N°: 2; un dominio "X" de fosfolipasa C específica de fosfatidilinositol (PI-PLC-X) (Fig. 3B; Acceso PFAM PF00388) localizado aproximadamente en los restos aminoacídicos 900-1048 de la SEC ID N°: 2; un dominio "Y" de fosfolipasa C específica de fosfatidilinositol (PI-PLC-Y) (Figs. 3C y 3D; Acceso PFAM PF00387) localizado aproximadamente en los restos aminoacídicos 1171-1184 y 1261-1353 de la SEC ID N°: 2; un dominio C2 (Fig. 3E; Acceso PFAM PF00168) localizado aproximadamente en los restos aminoacídicos 1378-1460 de la SEC ID N°: 2; y un dominio de asociación a Ras (RalGDS/AF-6) (RA) (Fig. 3F; Acceso PFAM PF00788) localizado aproximadamente en los restos aminoacídicos 1640-1745 de la SEC ID N°: 2;

La proteína 16836 también incluye los siguientes dominios: 16 sitios de N-glicosilación previstos (PS00001) localizados aproximadamente en los aminoácidos 285-288, 300-303, 395-398, 419-422, 583-586, 719-722, 752-755, 764-767, 770-773, 784-787, 817-820, 1115-1118, 1191-1194, 1224-1227, 1450-1453 y 1498-1501 de la SEC ID N°: 2; seis sitios de fosforilación de proteína quinasa dependiente de AMPc y GMPc (PS00004) localizados aproximadamente en los aminoácidos 303-306, 559-562, 1162-1165, 1277-1280, 1707-1710 y 1764-1767 de la SEC ID N°: 2; 30 sitios de fosforilación de proteína quinasa C previstos (PS00005) localizados aproximadamente en los aminoácidos 160-162, 177-179, 180-182, 268-270, 301-303, 306-308, 342-344, 384-386, 410-412, 501-503, 562-564, 571-573, 590-592, 613-615, 653-655, 686-688, 733-735, 736-738, 964-966, 1195-1197, 1203-1205, 1227-1229, 1258-1260, 1280-1282, 1492-1494, 1520-1522, 1705-1707, 1711-1713, 1715-1717 y 1769-1771 de la SEC ID N°: 2; 37 sitios de fosforilación de caseína quinasa II previstos (PS00006) localizados aproximadamente en los aminoácidos 124-127, 173-176, 180-183, 190-193, 239-242, 251-254, 268-271, 290-293, 295-298, 306-309, 323-326, 437-440, 607-610, 626-629, 721-724, 736-739, 754-757, 766-769, 819-822, 849-852, 889-892, 927-930, 1054-1057, 1102-1105, 1117-1120, 1216-1219, 1230-1233, 1238-1241, 1266-1269, 1299-1302, 1364-1367, 1618-1621, 1638-1641, 1667-1670, 1749-1752, 1790-1793 y 1800-1803 de la SEC ID N°: 2; 16 sitios de N-miristoilación previstos (PS00008) localizados aproximadamente en los aminoácidos 71-76, 146-151, 264-269, 338-343, 416-421, 484-489, 587-592, 705-710, 820-825, 1194-1199, 1209-1214, 1234-1239, 1501-1506, 1541-1546, 1756-1761 y 1797-1802 de la SEC ID N°: 2; un sitio de amidación previsto (PS00009) localizado aproximadamente en los aminoácidos 557-560 de la SEC ID N°: 2; y un dominio de superenrollamiento previsto aproximadamente en los aminoácidos 874-901 de la SEC ID N°: 2.

Si se desea información general con respecto a los identificadores PFAM, números de identificación de dominio de prefijo PS y prefijo PF, véase Sonnhammer *et al.* (1997) Protein 28: 405-420 y <http://www.psc.edu/general/software/packages/pfam/pfam.html>.

Un plásmido que contenía la secuencia de nucleótidos que codificaba 16836 humano (clon "Fbh16836FL") se depositó en la Colección Americana de Cultivos Tipo (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, el 19 de abril de 2000 y se le asignó el Número de Acceso 1774. Este depósito se mantendrá bajo los términos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos para los Propósitos de Procedimientos de Patente. Este depósito se realizó simplemente por conveniencia de los especialistas en la técnica y no es una admisión de que se necesita un depósito según 35 U.S.C. §112.

TABLA 1

Resumen de Información de Secuencia para 16836

ADNc	ORF	Polipéptido	Figura	Número de Acceso de la ATCC
SEC ID N°: 1.	SEC ID N°: 3.	SEC ID N°: 2.	Figs. 1A-1J	1774

TABLA 2

Resumen de dominios de 16836

Dominio	Localización en la SEC ID N°: 2
RasGEF	aproximadamente los restos aminoacídicos 35-338 de la SEC ID N°: 2
PI-PLC-X	aproximadamente los restos aminoacídicos 900-1048 de la SEC ID N°: 2
PI-PLC-Y	aproximadamente los restos aminoacídicos 1171-1184 y 1261-1353 de la SEC ID N°: 2
C2	aproximadamente los restos aminoacídicos 1378-1460 de la SEC ID N°: 2
Asociación de Ras (RA)	aproximadamente los restos aminoacídicos 1640-1745 de la SEC ID N°: 2

La proteína 16836 contiene un número significativo de características estructurales en común con miembros de la familia de la fosfolipasa C. El término “familia”, cuando hace referencia a las moléculas de proteína y de ácido nucleico de la invención, se refiere a dos o más proteínas o moléculas de ácido nucleico que tienen un dominio o motivo estructural común y que tienen una homología de secuencias de aminoácidos o de nucleótidos suficiente como se define en la presente memoria. Estos miembros de la familia pueden ser naturales o no naturales y pueden proceder de la misma especie o de especies diferentes. Por ejemplo, una familia puede contener una primera proteína de origen humano así como otras proteínas distintas de origen humano o, como alternativa, puede contener homólogos de origen no humano, por ejemplo, proteínas de rata o de ratón. Los miembros de una familia también pueden tener características funcionales comunes.

Los miembros de la familia de proteínas conocidas como fosfolipasa C específica de fosfoinosítido (PI-PLC) se caracterizan por una secuencia de aminoácidos que cataliza la hidrólisis de fosfatidil-inositol-4,5-bisfosfato (PIP₂) para generar los segundos mensajeros inositol 1,4,5-trisfosfato (IP₃) y diacilglicerol (DAG). El IP₃ puede difundirse al interior de la superficie del retículo endoplásmico donde se puede unir a un receptor de IP₃, induciendo la liberación de Ca²⁺ desde las reservas intracelulares al citoplasma. DAG permanece en la membrana celular donde puede servir para activar la enzima proteína quinasa C. Tanto la liberación de Ca²⁺ como la activación de la proteína quinasa C están implicadas en acontecimientos celulares tales como proliferación, diferenciación, secreción y migración.

Los miembros de la familia de PI-PLC generalmente comparten uno o más de los siguientes dominios: un dominio PI-PLC-X; un dominio PI-PLC-Y; y un dominio C2. Un dominio PI-PLC-X es un subdominio de PI-PLC que, junto con un subdominio de PI-PLC-Y, constituye el sitio catalítico de una fosfolipasa, por ejemplo, un dominio que cataliza la hidrólisis de fosfatidilinositol. Un dominio C2 es un dominio que puede mediar la unión a fosfolípidos dependiente de calcio.

La estructura del dominio de 16836 es similar a la estructura del dominio de una proteína codificada por el gen de *C. elegans* PLC210 (Shibatohge *et al.* (1998) *J. Biol. Chem.* 273:6218-6222). PLC210 se aisló en una selección doble híbrido de levadura para efectores de la función de ras. PLC210 se une preferiblemente a ras unido a GTP (activo), lo que sugiere que funciona en la propagación o amplificación de señales para la proliferación celular. Además de un papel para 16836 y PLC210 como efectores de ras, su pertenencia a la familia PI-PLC también sugiere un papel funcional para la proteína en la señalización de proliferación.

16836 y PLC210 parecen contener dominios no encontrados en miembros de las tres clases conocidas de PI-PLC. En primer lugar, la región N terminal lleva un dominio, un dominio del factor intercambiador de nucleótidos de guanina de Ras (RasGEF), homólogo a una familia de factores intercambiadores de nucleótidos de guanina para ras. En segundo lugar, la región C terminal contiene una estructura para la unión de ras, un dominio de asociación a ras (RA). Un dominio RA es un dominio que puede mediar la unión a una proteína ras. El dominio RA se asocia

ES 2 328 448 T3

preferentemente con un ras activado, por ejemplo el dominio RA puede asociarse específicamente con ras unido a GTP. Además, el dominio RA interacciona con la región efectora de ras. De esta manera, 16836 y PLC210 comprenden una nueva clase de PI-PLC.

5 Ras puede regular la actividad de 16836 modulando uno o mas de: (1) una activación de la actividad de 16836; o (2) una translocación inducida por ras de 16836 a un compartimento de membrana específico que contiene sustratos sobre los que puede actuar 16836. La mayor velocidad de renovación de fosfoinosítido observada en las células transformadas con ras sugiere una estimulación persistente de PI-PLC por ras activado. Además, los anticuerpos anti-PI-PLC inhiben la mitogénesis inducida por ras. De esta manera, la regulación de PI-PLC por ras puede participar en el control de la proliferación y/o diferenciación celular.

Una molécula 16836 puede incluir un dominio “del factor intercambiador de nucleótidos de guanina de Ras (Ras-GEF)” o regiones homólogas con un dominio “RasGEF”.

15 Como se usa en la presente memoria, la expresión “dominio de RasGEF” incluye una secuencia de aminoácidos de aproximadamente 50-400 restos aminoacídicos de longitud y que tiene una puntuación bit score para el alineamiento de la secuencia en el perfil del dominio RasGEF (SMART HMM) de al menos 5. Preferiblemente, un dominio RasGEF incluye al menos aproximadamente 80-350 aminoácidos, más preferiblemente aproximadamente 150-325 restos aminoacídicos, o aproximadamente 250-320 aminoácidos y tiene una puntuación bit para el alineamiento de la secuencia con el dominio RasGEF (HMM) de al menos 15 o mayor. Al dominio RasGEF (HMM) se le ha asignado el identificador SMART RasGEF (<http://smart.embl-heidelberg.de/>). Al dominio RasGEF (HMM) se le ha asignado el Número de Acceso de PFAM PF00617 (<http://genome.wustl.edu/Pfam/html>). En la Figura 3A se representa un alineamiento del dominio RasGEF (aminoácidos 35-338 de la SEC ID N°: 2) de 16836 humano con una secuencia de aminoácidos consenso (SEC ID N°: 4) procedente de un modelo oculto de Markov.

25 En una realización preferida, el polipéptido o proteína 16836 de la descripción tiene un “dominio RasGEF” o una región que incluye al menos aproximadamente 80-350, más preferiblemente aproximadamente 150-325 o 250-320 restos aminoacídicos y tiene al menos una homología de aproximadamente 50%, 60%, 70% 80% 90% 95%, 99% o 100% con un “dominio RasGEF” por ejemplo un dominio RasGEF de 16836 humano (por ejemplo los restos 35-338 de la SEC ID N°: 2).

Un polipéptido 16836 de la descripción puede incluir además un “dominio PI-PLC-X” o regiones homólogas a un “dominio PI-PLC-X”.

35 Como se usa en la presente memoria, la expresión “dominio PI-PLC-X” incluye una secuencia de aminoácidos de aproximadamente 30-250 restos aminoacídicos de longitud y que tiene una puntuación bit score para el alineamiento de la secuencia con el perfil del dominio PI-PLC-X (PFAM HMM) de al menos 50. Preferiblemente, un dominio PI-PLC-X incluye al menos aproximadamente 100-220 aminoácidos, más preferiblemente aproximadamente 120-200 restos aminoacídicos, o aproximadamente 130-170 aminoácidos y tiene valor de bit score para el alineamiento de la secuencia con el dominio PI-PLC-X (HMM) de al menos 240 o mayor. Al dominio PI-PLC-X (HMM) se le ha asignado el Número de Acceso de PFAM PF00388 (<http://genome.wustl.edu/Pfam/html>). En la Figura 3B se representa un alineamiento del dominio PI-PLC-X (aminoácidos 900-1048 de la SEC ID N°: 2) de 16836 humano con una secuencia de aminoácidos consenso (SEC ID N°: 5) procedente de un modelo oculto de Markov.

45 En una realización preferida, el polipéptido o proteína 16836 de la descripción tiene un “dominio PI-PLC-X” o una región que incluye al menos aproximadamente 100-220, más preferiblemente aproximadamente 120-200 o 130-170 restos aminoacídicos y tiene al menos una homología de aproximadamente 50%, 60%, 70% 80% 90% 95%, 99% o 100% con un “dominio PI-PLC-X”, por ejemplo el dominio PI-PLC-X de 16836 humano (por ejemplo los restos 900-1048 de la SEC ID N°: 2).

50 Una molécula 16836 de la descripción puede incluir además un “dominio PI-PLC-Y” o regiones homólogas con un “dominio PI-PLC-Y”.

55 Como se usa en la presente memoria, la expresión “dominio PI-PLC-Y” incluye una secuencia de aminoácidos de aproximadamente 40-300 restos aminoacídicos de longitud y que tiene una puntuación bit score para el alineamiento de la secuencia con el perfil del dominio PI-PLC-Y (PFAM HMM) de al menos 10. Preferiblemente, un dominio PI-PLC-Y incluye al menos aproximadamente 60-260 aminoácidos, más preferiblemente aproximadamente 80-250 restos aminoacídicos, o aproximadamente 90-200 aminoácidos y tiene valor de bit score para el alineamiento de la secuencia con el dominio PI-PLC-X (HMM) de al menos 140 o mayor. Al dominio PI-PLC-Y (HMM) se le ha asignado el Número de Acceso de PFAM PF00387 (<http://genome.wustl.edu/Pfam/html>). En las Figuras 3C y 3D se representan los alineamientos de los dominios PI-PLC-Y (aminoácidos 1171-1184 y 1261-1353 de la SEC ID N°: 2) de 16836 con secuencias de aminoácidos consenso (SEC ID N°: 6 y SEC ID N°: 7) procedentes de un modelo oculto de Markov.

65 En una realización preferida, el polipéptido o proteína 16836 de la descripción tiene un “dominio PI-PLC-Y” o una región que incluye al menos aproximadamente 60-260, más preferiblemente aproximadamente 80-250 o 90-200 restos aminoacídicos y tiene al menos una homología de aproximadamente 50%, 60%, 70% 80% 90% 95%, 99% o 100% con un “dominio PI-PLC-Y”, por ejemplo un PI-PLC-Y de 16836 humano (por ejemplo los restos 1171-1184 y 1261-1353 de la SEC ID N°: 2).

ES 2 328 448 T3

Una molécula de 16836 de la descripción puede incluir además un “dominio C2” o regiones homólogas con un “dominio C2”.

Como se usa en este documento, la expresión “dominio C2” incluye una secuencia de aminoácidos de aproximadamente 20-200 restos aminoacídicos de longitud y que tiene una puntuación bit score para el alineamiento de la secuencia con el perfil del dominio C2 (PFAM HMM) de al menos 15. Preferiblemente, un dominio C2 incluye al menos aproximadamente 50-120 aminoácidos, más preferiblemente aproximadamente 70-100 restos aminoacídicos, o aproximadamente 80-90 aminoácidos y tiene una puntuación bit score para el alineamiento de la secuencia con el dominio C2 (HMM) de al menos 35 o mayor. Al dominio C2 (HMM) se le ha asignado el Número de Acceso de PFAM PF00618 (<http://genome.wustl.edu/Pfam/>.html). En la Figura 3E se representa un alineamiento del dominio C2 (aminoácidos 1378-1460 de la SEC ID N°: 2) de 16836 humano con una secuencia de aminoácidos consenso (SEC ID N°: 8) procedente de un modelo oculto de Markov.

En una realización preferida, el polipéptido o proteína 16836 de la descripción tiene un “dominio C2” o una región que incluye al menos aproximadamente 50-120, más preferiblemente aproximadamente 70-100 u 80-90 restos aminoacídicos y tiene al menos una homología de aproximadamente 50%, 60%, 70% 80% 90% 95%, 99% o 100% con un “dominio C2” por ejemplo un dominio C2 de 16836 humano (por ejemplo los restos 1378-1460 de la SEC ID N°: 2).

Una molécula de 16836 de la descripción puede incluir además un “dominio de asociación a Ras (RA)” o regiones homólogas a un “dominio RA”.

Como se usa en la presente memoria, la expresión “dominio RA” incluye una secuencia de aminoácidos de aproximadamente 20-200 restos aminoacídicos de longitud y que tiene una puntuación bit score para el alineamiento de la secuencia con el perfil del dominio RA (PFAM HMM) de al menos 2. Preferiblemente, un dominio RA incluye al menos aproximadamente 50-140 aminoácidos, más preferiblemente aproximadamente 80-120 restos aminoacídicos, o aproximadamente 90-110 aminoácidos y tiene una puntuación bit score para el alineamiento de la secuencia con el dominio RA (HMM) de al menos 3 o mayor. Al dominio RA (HMM) se le ha asignado el Número de Acceso de PFAM PF00788 (<http://genome.wustl.edu/Pfam/>.html). En la Figura 3F se representa un alineamiento del dominio RA (aminoácidos 1640-1745 de la SEC ID N°: 2) de 16836 humano con una secuencia de aminoácidos consenso (SEC ID N°: 9) procedente de un modelo oculto de Markov.

En una realización preferida, un polipéptido o proteína 16836 de la descripción tiene un “dominio RA” o una región que incluye al menos aproximadamente 50-140, más preferiblemente aproximadamente 80-120 o 90-110 restos aminoacídicos y tiene al menos una homología de aproximadamente 50%, 60%, 70%, 80%, 90%, 95%, 99% o 100% con un “dominio RA” por ejemplo, un dominio RA de 16836 humano (por ejemplo los restos 1640-1745 de la SEC ID N°: 2).

Para identificar la presencia de un dominio “PI-PLC-X”, un dominio “PI-PLC-Y”, un dominio “C2” o un dominio “RA” en una secuencia de proteína 16836 y hacer la determinación de que un polipéptido o proteína de interés tiene un perfil particular, la secuencia de aminoácidos de la proteína puede buscarse frente a la base de datos PFAM de HMM (por ejemplo, la base de datos Pfam, edición 2.1) usando los parámetros por defecto (http://www.sanger.ac.uk/Software/PFAM/HMM_search). Por ejemplo, el programa hmmsf, que está disponible como parte del paquete de programas de búsqueda HMMER, es una familia de programas por defecto específicos para MILPAT0063 y una puntuación de 15 es la puntuación umbral por defecto para determinar un acierto (hit). Como alternativa, la puntuación umbral para determinar un acierto puede estar reducida, (por ejemplo, a 8 bits). Puede encontrarse una descripción de la base de datos PFAM en Sonhammer *et al.* (1997) *Proteins* 28(3): 405-420 y puede encontrarse una descripción detallada de HMM, por ejemplo, en Gribskov *et al.* (1990) *Meth. Enzymol.* 183: 146-159; Gribskov *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 4355-4358; Krogh *et al.* (1994) *J. Mol. Biol.* 235: 1501-1531; y Stultz *et al.* (1993) *Protein Sci.* 2: 305-314, cuyos contenidos se incorporan en la presente memoria como referencia.

Para identificar la presencia de un dominio “RasGEF” en una secuencia de proteína 16836 y hacer la determinación de que un polipéptido o proteína de interés tiene un perfil particular, la secuencia de aminoácidos de la proteína puede buscarse frente a una base de datos SMART (Simple Modular Architecture Research Tool, <http://smart.embl-heidelberg.de/>) de HMM como se describe en Schultz *et al.* (1998), *Proc. Natl. Acad. Sci. USA* 95: 5857 y Schultz *et al.* (200) *Nucl. Acids Res* 28: 231. La base de datos contiene dominios identificados por perfilado con los modelos ocultos de Markov del programa de búsqueda HMMer2 (R. Durbin *et al.* (1998) *Biological sequence analysis: probabilistic models of proteins and nucleic acids.* Cambridge University Press.; <http://hmmer.wustl.edu/>). La base de datos también se anota y controla extensivamente por expertos para mejorar la exactitud. Se realizó una búsqueda frente a la base de datos HMM que dio como resultado la identificación de un dominio “RasGEF” en la secuencia de aminoácidos del polipéptido 16836 humano aproximadamente en los restos 35 a 338 de la SEC ID N° 2 (véase la Figura 3A).

Un miembro de la familia de 16836 de la descripción puede incluir un dominio RasGEF, un dominio PI-PLC-X, al menos uno y preferiblemente dos dominios PI-PLC-Y, un dominio C2 y/o un dominio RA.

Además, un miembro de la familia de 16836 de la descripción puede incluir al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, 10, 11, 12, 13, 14, 15 y preferiblemente 16 sitios de N-glicosilación (PS00001); al menos uno, dos, tres, cuatro, cinco y preferiblemente seis sitios de fosforilación de proteína quinasa dependiente de AMPc y

ES 2 328 448 T3

GMPc (PS00004); al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 y preferiblemente 30 sitios de fosforilación de proteína quinasa C (PS00005); al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 y preferiblemente 37 sitios de fosforilación de caseína quinasa II (PS00006);
5 al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, 10, 11, 12, 13, 14, 15 y preferiblemente 16 sitios de N-miristoilación (PS00008); al menos un sitio de amidación (PS00009); y al menos un dominio de superenrollamiento.

Tanto en tumores de pulmón como en tumores de colon son comunes mutaciones de activación del oncogen *kras*. Las mutaciones en *kras* producen un aumento de la proliferación celular, supuestamente mediado por efectores de la función de *ras*. 16836 parece ser un nuevo efector de *ras*. Una asociación de 16836 con la ruta de *ras* se confirma adicionalmente por su mayor expresión en tumores (por ejemplo, tumores de pulmón, mama y colon) y particularmente en tumores con mutaciones *kras* activadoras. La expresión de 16836 puede ser un componente necesario para la ruta de señalización de *kras* en las células tumorales y los aumentos en los niveles de 16836 pueden contribuir el proceso de tumorigénesis.
15

Como los polipéptidos 16836 de la invención pueden modular las actividades mediadas por 16836, pueden ser útiles para el desarrollo de nuevos agentes de diagnóstico y terapéuticos para trastornos relacionados o mediados por 16836, como se describe más adelante.

Como se usa en este documento, una “actividad de 16836”, “actividad biológica de 16836” o “actividad funcional de 16836” se refiere a una actividad ejercida por una proteína, polipéptido o molécula de ácido nucleico de 16836. Por ejemplo, una actividad de 16836 puede ser una actividad ejercida por 16836 en un entorno fisiológico, por ejemplo una célula que responde a 16836 o en un sustrato de 16836, por ejemplo, un sustrato proteico. Una actividad de 16836 puede determinarse *in vivo* o *in vitro*. En una realización, una actividad de 16836 es una actividad directa, tal como una asociación con una molécula diana de 16836. Una “molécula diana” o “compañero de unión” es una molécula con la que se une o interacciona la proteína 16836 en la naturaleza, por ejemplo, un fosfatidilinositol o una proteína *ras*.
25

La actividad de 16836 también puede ser una actividad indirecta, por ejemplo, una actividad de señalización celular medida por interacción de la proteína 16836 con un ligando de 16836. Basándose en la similitudes de secuencias descritas anteriormente además de los patrones de expresión de 16836, es previsible que las moléculas de 16836 de la presente invención tengan actividades biológicas similares a los miembros de la familia de la fosfolipasa C y proteínas de asociación a *ras*. Los miembros de la familia de la fosfolipasa C juegan papeles importantes en la transducción de señales. Una PLC activada puede catalizar la hidrólisis de PIP₂, un componente minoritario de la membrana plasmática, para producir DAG e IP₃. IP₃ provoca la liberación de calcio desde las reservas intracelulares y aumenta la entrada de calcio desde el fluido extracelular. Los iones de calcio regulan directamente enzimas diana y afectan indirectamente a otras enzimas por medio del funcionamiento como un segundo mensajero e interacción con proteínas de unión a calcio, tales como la troponina C y la calmodulina. Por ejemplo, los iones de calcio regulan la contracción muscular, la degradación del glucógeno y la exocitosis. DAG, un producto de la hidrólisis por PI-PLC, actúa como segundo mensajero por activación de la proteína quinasa C. La proteína quinasa C activada fosforila un gran número de proteínas intracelulares en los restos de serina y treonina y modula diferentes rutas de señalización. Por ejemplo, la fosforilación de la glucógeno sintetasa por la proteína quinasa C detiene la síntesis de glucógeno. Además, la proteína quinasa C controla la división y proliferación celular. Las dos rutas forman parte de mecanismos de transducción de señales transmembrana, que regulan procesos celulares que incluyen la secreción, actividad neural, metabolismo, diferenciación y proliferación. La presencia de un dominio RA en 16836 sugiere que es un efector de *ras*, que participa en propagación o amplificación de señales de proliferación celular transducidas por *ras*.
30
35
40
45

Las proteínas 16836 de la presente invención pueden tener una o más de las siguientes actividades: (1) actividad de metabolización de fosfolípidos, por ejemplo, la capacidad de catalizar la hidrólisis de PIP₂ para producir DAG e IP₃; (2) la capacidad de asociarse con *ras*, por ejemplo, *ras* activado; (3) la capacidad de propagar la transducción de señales mediada por *ras*; (4) la capacidad de mediar la actividad de intercambio de nucleótidos de guanina; (5) la capacidad de transducir señales de membrana; (6) la capacidad de modular la proliferación; (7) la capacidad de modular la diferenciación; (8) la capacidad de modular la secreción; (9) la capacidad de modular la migración celular; (10) la capacidad de modular el metabolismo; (11) la capacidad de modular la percepción sensorial; o (12) la capacidad de modular la fertilización.
50
55

Las moléculas de 16836 pueden actuar como nuevas dianas de diagnóstico y agentes terapéuticos para controlar uno o más trastornos de la proliferación y/o diferenciación celular o trastornos asociados con el metabolismo óseo.

Los ejemplos de trastornos de la proliferación y/o diferenciación celular incluyen cáncer, por ejemplo, carcinoma, sarcoma, trastornos metastásicos o trastornos neoplásicos hematopoyéticos, por ejemplo leucemias. Un tumor metastásico puede producirse a partir de una multitud de tipos de tumores primarios, incluyendo pero sin limitación los de próstata, colon, pulmón, mama e hígado.
60

Como se usa en la presente memoria, los términos “cáncer”, “hiperproliferativo” y “neoplásico” se refieren a células que tienen la capacidad de crecimiento autónomo. Los ejemplos de estas células incluyen células que tienen un estado o condición anómala caracterizada por un crecimiento celular que prolifera rápidamente. Las patologías hiperproliferativas y neoplásicas pueden clasificarse como patológicas, es decir, que caracterizan o constituyen una patología, o pueden clasificarse como no patológicas, es decir, una desviación de lo normal pero no asociado con una
65

patología. Se entiende que el término incluye todos los tipos de crecimientos cancerosos o procesos oncogénicos, tejidos metastásicos o células, tejidos u órganos transformados por una malignidad, independientemente del tipo o etapa histopatológica de invasión. Las células “patológicas hiperproliferativas” se producen en patologías caracterizadas por un crecimiento tumoral maligno. Los ejemplos de células hiperproliferativas no patológicas incluyen la proliferación de células asociadas con la reparación de heridas.

Los términos “cáncer” o “neoplasmas” incluyen malignidades de los diversos sistemas de órganos, tales como las que afectan al pulmón, mama, tiroides, sistema linfático, tracto gastrointestinal y tracto genitourinario, así como adenocarcinomas que incluyen malignidades tales como la mayoría de los cánceres de colon, carcinoma de células renales, cáncer de próstata y/o tumores testiculares, carcinoma no microcítico de pulmón, cáncer del intestino delgado y cáncer del esófago.

El término “carcinoma” se reconoce en la técnica y se refiere a malignidades de tejidos epiteliales o endocrinos que incluyen carcinomas del sistema respiratorio, carcinoma del sistema gastrointestinal, carcinomas del sistema genitourinario, carcinomas testiculares, carcinomas de mama, carcinomas prostáticos, carcinomas del sistema endocrino y melanomas. Los ejemplos de carcinomas incluyen los que se forman a partir de tejidos del cuello del útero, pulmón, próstata, mama, cabeza y cuello, colon y ovario. El término también incluye carcinosarcomas, por ejemplo, que incluyen tumores malignos compuestos de tejidos carcinomatosos y sarcomatosos. Un “adenocarcinoma” se refiere a un carcinoma procedente de un tejido glandular o en el cual las células tumorales forman estructuras glandulares reconocibles.

El término “sarcoma” se reconoce en la técnica y se refiere a tumores malignos de origen mesenquimático.

Otros ejemplos de trastornos proliferativos incluyen trastornos neoplásicos hematopoyéticos. Como se usa en la presente memoria, la expresión “trastornos neoplásicos hematopoyéticos” incluye enfermedades que implican células hiperplásicas/neoplásicas de origen hematopoyético. Un trastorno neoplásico hematopoyético puede proceder de linajes mielóide, linfóide o eritroide, o células precursoras de los mismos. Preferiblemente, la enfermedad se produce a partir de leucemias agudas poco diferenciadas, por ejemplo, leucemia eritroblástica y leucemia megacarioblástica aguda. Otros trastornos mieloides ilustrativos incluyen, pero sin limitación, leucemia promielóide aguda (APML), leucemia mielógena aguda (AML) y leucemia mielógena crónica (CML) (revisado en Vaickus, L. (1991) Crit Rev. in Oncol./Hematol. 11: 267-97); las malignidades linfoides incluyen, pero sin limitación, leucemia linfoblástica aguda (ALL) que incluye ALL de linaje B y ALL de linaje T, leucemia linfocítica crónica (CLL), leucemia promielocítica (PLL), leucemia de células pilosas (BLL) y macroglobulinemia de Waldenstrom (WM). Otras formas de linfomas malignos incluyen, pero sin limitación, linfoma no Hodgkin y variantes del mismo, linfomasa de células T periféricas, leucemia/linfoma de células T adultas (ATL), linfoma de células T cutáneas (CTCL), leucemia linfocítica de células granulares grandes (LGF), enfermedad de Hodgkin y enfermedad de Reed-Sternberg.

Se detecta expresión de ARNm de 16836 en células óseas (por ejemplo, osteoclastos), testículos, cerebro, músculo esquelético, mama, corazón e hígado fetal (Figura 4). Por consiguiente, las moléculas de la invención pueden mediar trastornos que implican actividades aberrantes de esas células, por ejemplo, trastornos óseos, trastornos reproductivos, trastornos cardiovasculares, trastornos del músculo esquelético o trastornos inmunes como se describe con más detalle más adelante.

La presencia del ARN o la proteína 16836 puede usarse para identificar una célula o tejido, u otra muestra biológica, que procede de osteoclastos, testículos, cerebro, músculo esquelético, mama, corazón o hígado fetal o que es de origen humano. La expresión puede determinarse evaluando el ARN, por ejemplo, por hibridación de una sonda específica de 16836, o con un anticuerpo específico para 16836.

La proteína 16836, fragmentos de la misma y derivados y otras variantes de la secuencia en la SEC ID N°: 2 de la misma se denominan colectivamente “polipéptidos o proteínas de la invención” o “polipéptidos o proteínas 16836”. Las moléculas de ácido nucleico que codifican dichos polipéptidos o proteínas se denominan colectivamente “ácido nucleicos de la invención” o “ácidos nucleicos de 16836.” Las moléculas de 16836 se refieren a ácidos nucleicos, polipéptidos y anticuerpos de 16836.

Como se usa en la presente memoria, la expresión “molécula de ácido nucleico” incluye moléculas de ADN (por ejemplo, un ADNc o un ADN genómico), moléculas de ARN (por ejemplo, un ARNm) y análogos del ADN o ARN. Un análogo de ADN o ARN puede sintetizarse a partir de análogos de nucleótidos. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferiblemente es un ADN bicatenario.

La expresión “molécula de ácido nucleico aislada” o “molécula de ácido nucleico purificada” incluye moléculas de ácido nucleico que están separadas de otras moléculas de ácido nucleico presentes en la fuente natural del ácido nucleico. Por ejemplo, con respecto al ADN genómico, el término “aislado” incluye moléculas de ácido nucleico que están separadas del cromosoma con el cual está asociado naturalmente el ADN genómico. Preferiblemente, un ácido nucleico “aislado” carece de secuencias que flanquean de forma natural el ácido nucleico (es decir, secuencias localizadas en los extremos 5' y/o 3' del ácido nucleico) en el ADN genómico del organismo del que procede el ácido nucleico. Por ejemplo, en diversas realizaciones, la molécula de ácido nucleico aislada puede contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb o 0,1 kb de secuencias de nucleótidos 5' y/o 3' que flanquean de forma natural la molécula de ácido nucleico en el ADN genómico de la célula de la que procede el ácido nucleico.

Además, una molécula de ácido nucleico “aislada”, tal como una molécula de ADNc, puede carecer sustancialmente de otro material celular, o medio de cultivo cuando se produce por técnicas recombinantes, o carece sustancialmente de precursores químicos u otros agentes químicos cuando se sintetiza químicamente.

5 Como se usa en la presente memoria, la expresión “hibrida en condiciones de baja rigurosidad, rigurosidad media, alta rigurosidad o rigurosidad muy alta” describe condiciones de hibridación y lavado. Pueden encontrarse pautas para realizar reacciones de hibridación en Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6, que se incorpora como referencia. Esa referencia describe métodos acuosos y no acuosos y pueden usarse ambos. Las condiciones de hibridación específicas a las que se hace referencia en la presente memoria son las siguientes:
 10 1) condiciones de hibridación de baja rigurosidad en cloruro sódico/citrato sódico (SSC) 6X a aproximadamente 45°C, seguido de dos lavados en SSC 0,2X, SDS al 0,1% al menos a 50°C (la temperatura de los lavados puede aumentarse a 55°C en las condiciones de baja rigurosidad); 2) condiciones de hibridación de rigurosidad media en SSC 6X a aproximadamente 45°C, seguido de uno o más lavados en SSC 0,2X, SDS al 0,1% a 60°C; 3) condiciones de hibridación de rigurosidad elevada en SSC 6X a aproximadamente 45°C, seguido de uno o más lavados en SSC
 15 0,2X, SDS al 0,1% a 65°C; y preferiblemente 4) son condiciones de hibridación de rigurosidad muy alta fosfato sódico 0,5 M, SDS al 7% a 65°C, seguido de uno o más lavados en SSC 0,2X, SDS al 1% a 65°C. Las condiciones de rigurosidad muy alta (4) son las condiciones preferidas y las que deben usarse a menos que se especifique otra cosa.

Preferiblemente, una molécula de ácido nucleico aislada de la invención que hibrida en las condiciones de rigurosidad descritas en la presente memoria con la secuencia de la SEC ID N°: 1 o la SEC ID N°: 3, corresponde a una molécula de ácido nucleico natural.

Como se usa en la presente memoria, una molécula de ácido nucleico “natural” se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que se produce de forma natural. Por ejemplo, una molécula de ácido nucleico natural puede codificar una proteína natural.

Como se usa en la presente memoria, las expresiones “gen” y “gen recombinante” se refiere a moléculas de ácido nucleico que incluyen al menos una fase de lectura abierta que codifica una proteína 16836. El gen puede incluir opcionalmente secuencias no codificantes, por ejemplo, secuencias reguladoras e intrones. Preferiblemente, un gen codifica una proteína 16836 de mamífero o un derivado de la misma.

Un polipéptido o proteína “aislada” o “purificada” carece sustancialmente de material celular u otra proteína contaminante de la fuente celular o tisular de la que procede la proteína, o carece sustancialmente de precursores químicos u otros agentes químicos cuando se sintetiza químicamente. “Carece sustancialmente” significa que la preparación de una proteína 16836 tiene una pureza de al menos 10%. En una realización preferida, la preparación de la proteína 16836 tiene menos de aproximadamente 30%, 20%, 10% y más preferiblemente 5% (en peso seco) de material distinto de la proteína 16836 (también denominado en la presente memoria “proteína contaminante”) o de precursores químicos o agentes químicos que no son 16836. Cuando la proteína 16836 o la parte biológicamente activa de la misma se produce de forma recombinante, preferiblemente también carece sustancialmente de medio de cultivo, es decir, el medio de cultivo representa menos de aproximadamente 20%, más preferiblemente menos de aproximadamente 10% y aún más preferiblemente menos de aproximadamente 5% del volumen de la preparación de la proteína. La invención incluye preparaciones aisladas o purificadas de al menos 0,01, 0,1, 1,0 y 10 miligramos en peso seco.

Un resto aminoacídico “no esencial” es un resto que puede estar alterado con respecto a la secuencia de tipo silvestre de 16836 sin anular o alterar sustancialmente la actividad de 16836. Preferiblemente, la alteración no altera sustancialmente la actividad de 16836, por ejemplo, la actividad es al menos 20%, 40%, 60%, 70% o 80% del tipo silvestre. Un resto aminoacídico “esencial” es un resto que, cuando se altera con respecto a la secuencia de tipo silvestre de 16836, tiene como resultado la anulación de la actividad de 16836 de tal forma que está presente menos de 20% de la actividad de tipo silvestre. Por ejemplo, se prevé que los restos aminoacídicos conservados de 16836 sean particularmente susceptibles de alteración.

Una “sustitución conservativa de aminoácido” es una en la que el resto aminoacídico se reemplaza por un resto aminoacídico que tiene una cadena lateral similar. En la técnica se han definido familias de restos aminoacídicos que tienen cadenas laterales similares. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas β (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). De esta manera, un resto aminoacídico no esencial previsto en una proteína 16836 se reemplaza preferiblemente por otro resto aminoacídico de la misma familia de cadenas laterales. Como alternativa, en otra realización, pueden introducirse mutaciones aleatoriamente a lo largo de toda o parte de una secuencia codificante de 16836, tal como por mutagénesis de saturación, y en los mutantes resultantes puede explorarse la actividad biológica de 16836 para identificar mutantes que retienen la actividad. Después de la mutagénesis de la SEC ID N°: 1 o la SEC ID N°: 3, la proteína codificada puede expresarse de forma recombinante y puede determinarse la actividad de la proteína.

Como se usa en la presente memoria, una “parte biológicamente activa” de una proteína 16836 incluye un fragmento de una proteína 16836 que participa en una interacción, por ejemplo, una interacción intramolecular o inter-

molecular. Una interacción inter-molecular puede ser una interacción de unión específica o una interacción enzimática (por ejemplo, la interacción puede ser transitoria y se forma o se rompe un enlace covalente). Una interacción intermolecular puede ser entre una molécula de 16836 y una molécula que no es de 16836 o entre una primera molécula 16836 y una segunda molécula de 16836 (por ejemplo, una interacción de dimerización). Las partes biológicamente activas de una proteína 16836 incluyen péptidos que comprende secuencias de aminoácidos suficientemente homólogas o derivadas de la secuencia de aminoácidos de la proteína 16836, por ejemplo, la secuencia de aminoácidos mostrada en la SEC ID N°: 2, que incluye menos aminoácidos que las proteínas 16836 de longitud completa y presentan al menos una actividad de una proteína 16836. Típicamente, las partes biológicamente activas comprenden un dominio o motivo con al menos una actividad de la proteína 16836, por ejemplo, la capacidad de catalizar la hidrólisis de fosfatidilinositol o asociarse con Ras. Una parte biológicamente activa de una proteína 16836 puede ser un polipéptido que tiene, por ejemplo, una longitud de 10, 25, 50, 100, 200 o más aminoácidos. Las partes biológicamente activas de una proteína 16836 pueden usarse como dianas para desarrollar agentes que modulan una actividad mediada por 16836, por ejemplo, la capacidad de catalizar la hidrólisis de fosfatidilinositol o asociarse con Ras.

Se realizan cálculos de homología o identidad de secuencia entre secuencias (los términos se usan indistintamente en la presente memoria) como se indica a continuación.

Para determinar el porcentaje de identidad de dos secuencias de aminoácidos, o de dos secuencias de ácido nucleico, se alinean las secuencias para conseguir una comparación óptima (por ejemplo, pueden introducirse huecos en una o tanto en la primera como en la segunda secuencias de aminoácidos o de ácido nucleico para un alineamiento óptimo y pueden descartarse secuencias no homólogas con fines comparativos). En una realización preferida, la longitud de una secuencia de referencia alineada con fines comparativos es al menos 30%, preferiblemente al menos 40%, más preferiblemente al menos 50%, 60% e incluso más preferiblemente al menos 70%, 80%, 90% o 100% de la longitud de la secuencia de referencia. Después se comparan los restos aminoácidos o nucleótidos en las posiciones de aminoácidos o en las posiciones de nucleótidos correspondientes. Cuando una posición de la primera secuencia está ocupada por el mismo resto aminoácido o nucleótido que la correspondiente posición en la segunda secuencia, entonces las moléculas son idénticas en esa posición (como se usa en la presente memoria, "identidad" de aminoácido o ácido nucleico es equivalente a "homología" de aminoácido o ácido nucleico).

El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos, y la longitud de cada hueco, que necesitan introducirse para un alineamiento óptimo de las dos secuencias.

La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias pueden realizarse usando un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman y Wunsch ((1970) *J. Mol. Biol.* 48: 444-453) que se ha incorporado en el programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), usando una matriz Blosum 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 ó 4 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina usando el programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. Una serie de parámetros particularmente preferidos (y el que debe usarse a menos que se especifique otra cosa) son una matriz de puntuación Blossum 62 con una penalización de hueco de 12, una penalización de extensión de hueco de 4, y una penalización de hueco de desplazamiento de fase de 5.

El porcentaje de identidad entre dos secuencias de aminoácidos o de nucleótidos puede determinarse usando el algoritmo de E. Meyers y W. Miller ((1989) *CABIOS*, 4: 11-17) que se ha incorporado en el programa ALIGN (versión 2.0), usando una tabla de restos de peso de PAM120, una penalización de longitud de hueco de 12 y una penalización de hueco de 4.

Las secuencias de ácido nucleico y de proteína descritas en la presente memoria pueden usarse como "secuencia problema" para realizar una búsqueda frente a bases de datos públicas para identificar, por ejemplo, otros miembros de la familia o secuencias relacionadas. Estas búsquedas pueden realizarse usando los programas NBLAST y XBLAST (versión 2.0) de Altschul, *et al.* (1990) *J. Mol. Biol.* 215: 403-10. Las búsquedas de nucleótidos BLAST pueden realizarse con el programa NBLAST, puntuación = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas a moléculas de ácido nucleico de 16836 de la invención. Las búsquedas de proteínas BLAST pueden realizarse con el programa XBLAST (puntuación = 50, longitud de palabra = 3) para obtener secuencias de aminoácidos homólogas a moléculas de proteína 16836 de la invención. Para obtener alineamientos con huecos con fines comparativos, puede utilizarse Gapped BLAST como se describe en Altschul *et al.*, (1997) *Nucleic Acids Res.* 25: 3389-3402. Cuando se utilizan programas BLAST y Gapped BLAST, pueden usarse los parámetros por defecto de los respectivos programas (por ejemplo, XBLAST y NBLAST). Véase <http://www.ncbi.nlm.nih.gov>.

Los polipéptidos 16836 particulares de la presente invención tienen una secuencia de aminoácidos sustancialmente idéntica a la secuencia de aminoácidos de la SEC ID N°: 2. En el contexto de una secuencia de aminoácidos, la expresión "sustancialmente idéntica" se usa para hacer referencia a un primer aminoácido que contiene un número suficiente o mínimo de restos aminoácidos que son i) idénticos a, o ii) sustituciones conservativas de restos aminoácidos alineados en una segunda secuencia de aminoácidos de tal forma que la primera y la segunda secuencias de aminoácidos pueden tener un dominio estructural común y/o una actividad funcional común. Por ejemplo, las secuencias de ami-

noácidos que contienen un dominio estructural común que tiene una identidad de al menos aproximadamente 60% o 65%, probablemente una identidad de 75%, más probablemente una identidad de 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% con la SEC ID N°: 2 se denominan sustancialmente idénticas.

5 En el contexto de una secuencia de nucleótidos, la expresión “sustancialmente idéntica” se usa en la presente memoria para hacer referencia a una primera secuencia de ácido nucleico que contiene un número suficiente o mínimo de nucleótidos que son idénticos a nucleótidos alineados en una segunda secuencia de ácido nucleico de tal forma que la primera y segunda secuencias de ácido nucleico codifican un polipéptido que tiene una actividad funcional común, o codifican un dominio de polipéptido estructural común o una actividad de polipéptido funcional común. Por ejemplo,
10 las secuencias e nucleótidos que tienen una identidad de al menos aproximadamente 60% o 65%, probablemente una identidad de 75%, más probablemente una identidad de 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% con la SEC ID N°: 1 ó 3 se denominan sustancialmente idénticas.

15 “Expresión defectuosa o expresión aberrante”, como se usa en la memoria, se refiere a un patrón de expresión génica que no es de tipo silvestre a nivel del ARN o de la proteína. Incluye: expresión a niveles que no son de tipo silvestre, es decir, una expresión excesiva o insuficiente; un patrón de expresión que difiere del tipo silvestre en términos del momento o fase en la que se expresa el gen, por ejemplo, aumento o disminución de la expresión (en comparación con el tipo silvestre) en un periodo o fase del desarrollo predeterminado; un patrón de expresión que difiere del tipo silvestre en términos de expresión alterada, por ejemplo aumentado o disminuido (en comparación con el tipo silvestre) en un tipo celular o tipo tisular predeterminado; un patrón de expresión que difiere del tipo silvestre en términos del tamaño de corte y empalme, secuencia de aminoácidos traducida, modificación postraduccional o actividad biológica del polipéptido expresado; un patrón de expresión que difiere del tipo silvestre en términos del efecto de un estímulo ambiental o estímulo extracelular sobre la expresión del gen, por ejemplo, un patrón de aumento o disminución de la expresión (en comparación con el tipo silvestre) en presencia de un aumento o disminución en la
25 intensidad del estímulo.

“Sujeto”, como se usa en la presente memoria, se refiere a animales humanos y no humanos. La expresión “animales no humanos” de la invención incluye todos los vertebrados, por ejemplo mamíferos tales como primates no humanos (particularmente primates superiores), ovejas, perros, roedores (por ejemplo rata o ratón), cobaya, cabra,
30 cerdo, gato, conejos, vacas y no mamíferos, tales como pollos, anfibios, reptiles, etc. En una realización preferida, el sujeto es un ser humano. En otra realización, el sujeto es un animal experimental o un animal adecuado como modelo de enfermedad.

Una “preparación purificada de células”, como se usa en la presente memoria, se refiere a una preparación *in vitro*
35 de células. En el caso de células de organismos multicelulares (por ejemplo, plantas y animales), una preparación purificada de células es una subserie de células obtenidas a partir del organismo, no el organismo intacto entero. En el caso de un microorganismo unicelular (por ejemplo, células cultivadas y células microbianas), consiste en una preparación de al menos 10% y más preferiblemente 50% de las células objeto.

40 A continuación se describen con detalle diversos aspecto de la invención.

Moléculas de ácido nucleico aisladas

En un aspecto, la invención proporciona una molécula de ácido nucleico aislada o purificada que codifica un poli-
45 péptido 16836 descrito en la presente memoria, por ejemplo, una proteína 16836 de longitud completa o un fragmento de la misma, por ejemplo una parte biológicamente activa de la proteína 16836. También se incluye un fragmento de ácido nucleico adecuado para uso como sonda de hibridación que puede usarse, por ejemplo, para identificar una molécula de ácido nucleico que codifica un polipéptido de la invención, ARNm de 16836 y fragmentos adecuados para uso como cebadores, por ejemplo cebadores de PCR para la amplificación o mutación de moléculas de ácido
50 nucleico.

En una realización, una molécula de ácido nucleico aislada de la invención incluye la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o una parte de cualquiera de estas secuencias de nucleótidos. En una realización, la molécula de ácido nucleico incluye secuencias que codifican la proteína 16836 humana (es decir, “la región codificante”
55 de la SEC ID N°: 1, como se muestra en la SEC ID N°: 3), así como secuencias 5’ no traducidas. Como alternativa, la molécula de ácido nucleico puede incluir sólo la región codificante de la SEC ID N°: 1 (por ejemplo, la SEC ID N°: 3) y, por ejemplo, y ninguna secuencia no flanqueante de las que normalmente acompañan a la secuencia objeto. La descripción también incluye cuando la molécula de ácido nucleico codifica una secuencia correspondiente a un fragmento de la proteína desde aproximadamente el aminoácido 35-338, 900-1048, 1171-1184, 1261-1353, 1378-1460 ó
60 1640-1745 de la SEC ID N°: 2.

En otra realización, una molécula de ácido nucleico aislada de la invención incluye una molécula de ácido nucleico que es un complemento de la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o la SEC ID N°: 3, o una parte de cualquiera de estas secuencias de nucleótidos. En otras realizaciones, la molécula de ácido nucleico de la invención
65 es suficientemente complementaria a la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o SEC ID N°: 3, como para que pueda hibridar (por ejemplo en condiciones rigurosas descritas en la presente memoria) con la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o 3, formando de esta manera un dúplex estable.

ES 2 328 448 T3

En una realización, una molécula de ácido nucleico aislada de la presente invención incluye una secuencia de nucleótidos que tiene una homología de al menos aproximadamente: 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o mayor con la longitud entera de la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o SEC ID N°: 3, o una parte, preferiblemente de la misma longitud de cualquiera de estas secuencias de nucleótidos.

Fragmentos de Ácido Nucleico de 16836

Una molécula de ácido nucleico de la invención puede incluir sólo una parte de la secuencia de ácido nucleico de la SEC ID N°: 1 ó 3. Por ejemplo, dicha molécula de ácido nucleico puede incluir un fragmento que puede usarse como sonda o cebador o un fragmento que codifica una parte de una proteína 16836, por ejemplo, una parte inmunogénica o biológicamente activa de una proteína 16836. Un fragmento puede comprender los nucleótidos de la SEC ID N°: 1 que codifica un dominio de fosfolipasa C, RA o RasGEF de la proteína 16836 humano. La secuencia de nucleótidos determinada a partir de la clonación del gen 16836 permite la generación de sondas y cebadores diseñados para uso en la identificación y/o clonación de otros miembros de la familia de 16836, o fragmentos de los mismos, así como homólogos de 16836, o fragmentos de la misma, procedentes de otras especies.

En otra realización, un ácido nucleico incluye una secuencia de nucleótidos que incluye parte o toda la región codificante y se extiende a la región no codificante 5' o 3' (o ambas). Otras realizaciones incluyen un fragmento que incluye una secuencia de nucleótidos que codifica un fragmento de aminoácidos descrito en la presente memoria. Los fragmentos de ácido nucleico pueden codificar un dominio o sitio específico descrito en la presente memoria o fragmentos del mismo, particularmente fragmentos del mismo que tienen una longitud de al menos 100 aminoácidos, por ejemplo al menos 1200, 1400, 1600 ó 1800 aminoácidos de longitud. Los fragmentos también incluyen secuencias de ácido nucleico que corresponden a secuencias de aminoácidos específicas descritas anteriormente o fragmentos de las mismas. No debe considerarse que los fragmentos de ácido nucleico incluyen los fragmentos que pueden haberse descrito antes de la invención.

Un fragmento de ácido nucleico puede incluir una secuencia correspondiente a un dominio, región o sitio funcional descrito en la presente memoria. Un fragmento de ácido nucleico también puede incluir uno o más dominios, regiones o sitios funcionales descritos en la presente memoria. De esta manera, por ejemplo, un fragmento de ácido nucleico de 16836 puede incluir una secuencia correspondiente a un dominio de fosfolipasa C, RA o RasGEF.

Se proporcionan sonda y cebadores de 16836. Típicamente, una sonda/cebador es un oligonucleótido aislado o purificado. El oligonucleótido típicamente incluye una región de secuencia de nucleótidos que hibrida en las condiciones rigurosas descritas en la presente memoria con al menos aproximadamente 7, 12 ó 15, preferiblemente 20 ó 25, más preferiblemente aproximadamente 30, 35, 40, 45, 50, 55, 60, 65 ó 75 nucleótidos consecutivos de una secuencia con sentido o antisentido de la SEC ID N°: 1 ó la SEC ID N°: 3, o de una variante alélica natural o mutante de la SEC ID N°: 1 o SEC ID N°: 3.

En una realización preferida, el ácido nucleico es una sonda que tiene una longitud de al menos 5 ó 10, y menos de 200, más preferiblemente menos de 100 o menos de 50 pares de bases. Debe ser idéntica o diferir en uno o menos de 5 ó 10 bases de una secuencia descrita en la presente memoria. Si se necesita un alineamiento para esta comparación, las secuencias deben alinearse para una homología máxima. Se consideran diferencias las secuencias de "bucles" de deleciones o inserciones, o desacoplamiento.

Una sonda o cebador puede proceder de la cadena con sentido o antisentido de un ácido nucleico que codifica: un dominio de factor intercambiador de nucleótidos de guanina de Ras (RasGEF) (aproximadamente los restos aminoácidos 35-338 de la SEC ID N°: 2); un dominio "X" de fosfolipasa C específico de fosfatidilinositol (PI-PLC-X) (aproximadamente los restos aminoácidos 900-1048 de la SEC ID N°: 2); un dominio "Y" de fosfolipasa C específico de fosfatidilinositol (PI-PLC-Y) (aproximadamente los restos aminoácidos 1171-1184 o 1261-1353 de la SEC ID N°: 2); un dominio C2 (aproximadamente los restos aminoácidos 1378-1460 de la SEC ID N°: 2); o un dominio de asociación con Ras (RalGDS/AF-6) (RA) (aproximadamente los restos aminoácidos 1640-1745 de la SEC ID N°: 2).

En otra realización, se proporciona una serie de cebadores, por ejemplo, cebadores adecuados para uso en una PCR, que pueden usarse para amplificar una región seleccionada de una secuencia de 16836, por ejemplo, un dominio, región, sitio u otra secuencia descrita en la presente memoria. Los cebadores deben tener una longitud de al menos 5, 10 ó 50 pares de bases y una longitud menor de 100 o menor de 200 pares de bases. Los cebadores deben ser idénticos o difieren en una base de una secuencia descrita en la presente memoria o de una variante natural. Por ejemplo, se proporcionan cebadores adecuados para amplificar todo o parte de cualquiera de las siguientes regiones: un dominio de factor intercambiador de nucleótidos de guanina de Ras (RasGEF) (aproximadamente los restos aminoácidos 35-338 de la SEC ID N°: 2); un dominio "X" de fosfolipasa C específico de fosfatidilinositol (PI-PLC-X) (aproximadamente los restos aminoácidos 900-1048 de la SEC ID N°: 2); un dominio "Y" de fosfolipasa C específico de fosfatidilinositol (PI-PLC-Y) (aproximadamente los restos aminoácidos 1171-1184 o 1261-1353 de la SEC ID N°: 2); un dominio C2 (aproximadamente los restos aminoácidos 1378-1460 de la SEC ID N°: 2); o un dominio de asociación con Ras (RalGDS/AF-6) (RA) (aproximadamente los restos aminoácidos 1640-1745 de la SEC ID N°: 2).

ES 2 328 448 T3

Un fragmento de ácido nucleico puede codificar un epítipo que lleva una región de un polipéptido descrito en la presente memoria.

5 Un fragmento de ácido nucleico que codifica una “parte biológicamente activa de un polipéptido 16836” puede prepararse por aislamiento de una parte de la secuencia de nucleótidos de la SEC ID N° 1 ó 3, que codifica un polipéptido que tiene una actividad biológica de 16836 (por ejemplo las actividades biológicas de las proteínas 16836 se describen en la presente memoria) que expresa la parte codificada de la proteína 16836 (por ejemplo, por expresión recombinante *in vitro*) y evaluación de la actividad de la parte codificada de la proteína 16836. Por ejemplo, un fragmento de ácido nucleico que codifica una parte biológicamente activa de 16836 incluye un dominio de fosfolipasa
10 C, RA o RasGEF, por ejemplo, los restos aminoácidos aproximadamente 35-338, 900-1048, 1171-1184, 1261-1353, 1378-1460 ó 1640-1745 de la SEC ID N°: 2. Un fragmento de ácido nucleico que codifica una parte biológicamente activa de un polipéptido 16836 puede comprender una secuencia de nucleótidos que tiene una longitud mayor de 300 o más nucleótidos.

15 En realizaciones preferidas, el fragmento de ácido nucleico incluye una secuencia de nucleótidos que es distinta de la secuencia de AF117948, AL050031, AW243053, AW272589, AW272590, A02325, V89699, AC03885 o WO/58473.

20 En realizaciones preferidas, el fragmento comprende la región codificante de 16836, por ejemplo, la secuencia de nucleótidos de la SEC ID N°: 3.

En realizaciones preferidas, un ácido nucleico incluye una secuencia de nucleótidos que tiene un longitud de aproximadamente 4000, 4500, 5000, 6000, 7000, 8000, 9000, 10000 o más nucleótidos e hibrida en condiciones de rigurosidad descritas en la presente memoria con una molécula de ácido nucleico de la SEC ID N°: 1 o la SEC ID N°:
25 3.

Variantes de Ácido Nucleico de 16836

30 La descripción incluye además moléculas de ácido nucleico que difieren de la secuencia de nucleótidos mostrada en la SEC ID N°: 1 o la SEC ID N°: 3. Estas diferencias pueden deberse a la degeneración del código genético (y producir un ácido nucleico que codifica las mismas proteínas 16836 que se han codificado por la secuencia de nucleótidos descrita en la presente memoria). En otra realización, una molécula de ácido nucleico aislada de la invención tiene una secuencia de nucleótidos que codifica una proteína que tiene una secuencia de aminoácidos que difiere al menos en uno, pero en menos de 5, 10, 20, 50, ó 100 restos aminoácidos mostrados en la SEC ID N°: 2. Si se necesita
35 un alineamiento para esta comparación, las secuencias deben alinearse para una homología máxima. Se consideran diferencias las secuencias de “bucles” de deleciones o inserciones, o desacoplamiento.

Pueden seleccionarse ácidos nucleicos de la descripción para que tengan codones que son preferidos o no preferidos para un sistema de expresión particular. Por ejemplo, el ácido nucleico puede ser uno en el que al menos un codón,
40 preferiblemente al menos 10% o 20% de los codones se han alterado de forma que la secuencia se optimiza con respecto a la expresión en *E. coli*, células de levadura, células humanas, células de insecto o células CHO.

Las variantes del ácido nucleico puede ser naturales, tales como variantes alélicas (mismo locus), homólogas (diferente locus) y ortólogas (diferente organismo) o pueden ser no naturales. Las variantes no naturales pueden obtenerse
45 por técnicas de mutagénesis, incluyendo las aplicadas a polinucleótidos, células u organismos. Las variantes pueden contener sustituciones, deleciones, inversiones e inserciones de nucleótidos. Puede producirse una variación en cualquiera o tanto en la región codificante como en la región no codificante. Las variaciones pueden producir sustituciones de aminoácidos conservativas y no conservativas (en comparación con el producto codificado).

50 Pueden identificarse ortólogos, homólogos y variantes alélicas usando métodos conocidos en la técnica. Estas variantes comprenden una secuencia de nucleótidos que codifica un polipéptido que tiene una identidad de 50%, al menos aproximadamente 55%, típicamente al menos aproximadamente 70-75%, más típicamente al menos aproximadamente 80-85% y aún más típicamente al menos aproximadamente 90-95% o mayor con la secuencia de nucleótidos mostrada en la SEC ID N°: 2 o un fragmento de esta secuencia. Estas moléculas de ácido nucleico pueden identificarse
55 fácilmente como secuencias capaces de hibridar en condiciones de rigurosidad descritas en la presente memoria, con la secuencia de nucleótidos mostrada en la SEC ID N° 2 o un fragmento de la secuencia. Las moléculas de ácido nucleico correspondientes a ortólogos, homólogos y variantes alélicas de los ADNc de 16836 de la invención pueden aislarse adicionalmente por mapeo en el mismo cromosoma o locus que el gen de 16836.

60 Las variantes preferidas incluyen las que están correlacionadas con la capacidad de catalizar la hidrólisis de fosfatidilinositol o asociarse con Ras.

Las variantes alélicas de 16836, por ejemplo 16836 humana, incluyen tanto proteínas funcionales como proteínas no funcionales. Las variantes alélicas funcionales son variantes de secuencias de aminoácidos naturales de la proteína
65 16836 dentro de una población que mantiene la capacidad de catalizar la hidrólisis de fosfatidilinositol o asociarse con Ras. Las variantes alélicas funcionales típicamente contendrán sólo la sustitución conservativa de uno o más aminoácidos de la SEC ID N°: 2, o la sustitución, deleción o inserción de restos no críticos en regiones no críticas

de la proteína. Las variantes alélicas no funcionales son variantes de la secuencia de aminoácidos naturales de la proteína 16836, por ejemplo humana, dentro de una población que no tiene la capacidad de catalizar la hidrólisis de fosfatidilinositol o asociarse con Ras. Las variantes alélicas no funcionales contendrán típicamente una sustitución, delección, o inserción no conservativas, o el truncamiento prematuro de la secuencia de aminoácidos de la SEC ID N°: 2, o una sustitución, inserción, o delección en restos críticos o regiones críticas de la proteína.

Moléculas de Ácido Nucleico Antisentido, Ribozimas y Moléculas de Ácido Nucleico de 16836 Modificadas

En otro aspecto, la invención se refiere a una molécula de ácido nucleico aislada que es antisentido con respecto a 16836. Un ácido nucleico “antisentido” puede incluir una secuencia de nucleótidos que es complementaria a un ácido nucleico “con sentido” que codifica una proteína, por ejemplo, complementaria a la cadena codificante de una molécula de ADNc bicatenaria o complementaria a una secuencia de ARNm. El ácido nucleico antisentido puede ser complementario a una cadena codificante de 16836 entera, o sólo a una parte de la misma (por ejemplo, la región codificante de 16836 humana correspondiente a la SEC ID N°: 3). En otra realización, la molécula de ácido nucleico antisentido es antisentido con respecto a una “región no codificante” de la cadena codificante de una secuencia de nucleótidos que codifica 16836 (por ejemplo, las regiones 5' y 3' no traducidas).

Un ácido nucleico antisentido puede diseñarse de tal forma que sea complementario a la región codificante entera del ARNm de 16836, pero más preferiblemente es un oligonucleótido que es antisentido con respecto a solamente una parte de la región codificante o no codificante del ARNm de 16836. Por ejemplo, el oligonucleótido antisentido puede ser complementario a la región que rodea el sitio de inicio de la traducción del ARNm de 16836, por ejemplo, entre las regiones -10 y +10 de la secuencia de nucleótidos del gen diana de interés. Un oligonucleótido antisentido puede tener, por ejemplo, aproximadamente 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 o más nucleótidos de longitud.

Se puede construir un ácido nucleico antisentido de la invención usando síntesis química y reacciones de ligamiento enzimático usando procedimientos conocidos en la técnica. Por ejemplo, un ácido nucleico antisentido (por ejemplo, un oligonucleótido antisentido) puede sintetizarse químicamente usando nucleótidos naturales o nucleótidos modificados de diversa forma diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre los ácidos nucleicos con sentido y antisentido, por ejemplo, pueden usarse derivados fosforotioato y nucleótidos sustituidos con acridina. El ácido nucleico antisentido también puede producirse biológicamente usando un vector de expresión en el que se ha subclonado un ácido nucleico en una orientación antisentido (es decir, el ARN transcrito a partir del ácido nucleico insertado tendrá una orientación antisentido con respecto a un ácido nucleico diana de interés, descrito adicionalmente en la siguiente subsección).

Las moléculas de ácido nucleico antisentido de la invención típicamente se administran a un sujeto (por ejemplo, por inyección directa en un sitio de un tejido), o se generan *in situ* de forma que hibriden o se unan al ARNm y/o ADN genómico celular que codifica una proteína 16836 para inhibir de esta manera la expresión de la proteína, por ejemplo, por medio de la inhibición de la transcripción y/o la traducción. Como alternativa, pueden modificarse moléculas de ácido nucleico antisentido para dirigirse a células seleccionadas y después administrarse sistémicamente. Para la administración sistémica, las moléculas antisentido pueden modificarse de tal forma que se unan específicamente a receptores o antígenos expresados en la superficie de una célula seleccionada, por ejemplo, por unión de las moléculas de ácido nucleico antisentido a péptidos o anticuerpos que se unen a receptores o antígenos de la superficie celular. Las moléculas de ácido nucleico antisentido también pueden administrarse a células usando los vectores descritos en la presente memoria. Para conseguir suficientes concentraciones intracelulares de las moléculas antisentido, se prefieren construcciones de vectores en las que la molécula de ácido nucleico antisentido se pone bajo el control de un promotor pol II o pol III fuerte.

En otra realización, la molécula de ácido nucleico antisentido de la invención es una molécula de ácido nucleico α -anomérica. Una molécula de ácido nucleico α -anomérica forma híbridos bicatenarios específicos con ARN complementario en el que, al contrario de las unidades β habituales, las cadenas van paralelas entre sí (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15: 6625-6641). La molécula de ácido nucleico antisentido también puede comprender un 2'-O-metilribonucleótido (Inoue *et al.* (1987) *Nucleic Acids Res.* 15: 6131-6148) o un análogo de ARN-ADN quimérico (Inoue *et al.* (1987) *FEBS Lett.* 215: 327-330).

En otra realización, un ácido nucleico antisentido de la invención es una ribozima. Una ribozima que tiene especificidad por un ácido nucleico codificante de 16836 puede incluir una o más secuencias complementarias a la secuencia de nucleótidos de un ADNc de 16836 descrito en la presente memoria (es decir, la SEC ID N°: 1 o SEC ID N°: 3), y una secuencia que tiene una secuencia catalítica conocida responsable de la escisión del ARNm (véase la Patente de Estados Unidos N° 5.093.246 o Haselhoff y Gerlach (1988) *Nature* 334: 585-591). Por ejemplo, puede construirse un derivado del ARN L-19 IVS de *Tetrahymena* en el que la secuencia de nucleótidos del sitio activo es complementaria a la secuencia de nucleótidos a escindir en un ARNm codificante de 16836. Véase, por ejemplo, Cech *et al.* Patente de Estados Unidos N° 4.987.071; y Cech *et al.* Patente de Estados Unidos N° 5.116.742. Como alternativa, el ARNm de 16836 puede usarse para seleccionar un ARN catalítico que tiene una actividad ribonucleasa específica de un grupo de moléculas de ARN. Véase, por ejemplo, Bartel, D. y Szostak, J.W. (1993) *Science* 261: 1411-1418.

ES 2 328 448 T3

La expresión del gen de 16836 puede inhibirse por dirección a secuencias de nucleótidos complementarias a la región reguladora de la proteína 16836 (por ejemplo, el promotor y/o potenciadores de 16836) para formar estructuras en triple hélice que impiden la transcripción del gen de 16836 en células diana. Véase, en general, Helene, C. (1991) *Anticancer Drug Des.* 6: 569-84; Helene, C. i (1992) *Ann. N.Y. Acad. Sci.* 660: 27-36; y Maher, L. J. (1992) *Bioassays* 14: 807-15. Las secuencias potenciales que pueden fijarse como dianas para la formación de triples hélices pueden aumentarse creando una denominada molécula de ácido nucleico "switchback". Las moléculas switchback se sintetizan de una forma 5'-3', 3'-5' alterna, de tal manera que forman pares de bases primero con una cadena de un dúplex y después con la otra, eliminando la necesidad de que esté presente un tramo considerable de purinas o pirimidinas en una cadena de dúplex.

La invención también proporciona moléculas de sonda y cebador oligonucleotídicas marcadas de forma detectable. Típicamente, estos marcadores son quimioluminiscentes, fluorescentes, radiactivos o colorimétricos.

Polipéptidos 16836 Aislados

En otro aspecto, la invención se refiere a una proteína 16836 aislada, o un fragmento, por ejemplo, una parte biológicamente activa de la misma, para uso como inmunógenos o antígenos para inducir o ensayar (o más generalmente para unirse a) anticuerpos anti-16836. La proteína 16836 puede aislarse a partir de fuente celulares o tisulares usando técnicas de purificación de proteínas convencionales. La proteína 16836 o sus fragmentos pueden producirse por técnicas de ADN recombinantes o pueden sintetizarse químicamente.

Los polipéptidos de la invención incluyen los que se producen como resultado de la existencia de múltiples genes, acontecimientos de transcripción alternativos, acontecimientos de corte y empalme de ARN alternativos y acontecimientos traduccionales y post-traduccionales alternativos. El polipéptido puede expresarse en sistemas, por ejemplo, células cultivadas, que producen sustancialmente las mismas modificaciones post-traduccionales presentes cuando se expresa el polipéptido en una célula nativa, o en sistemas que tienen como resultado la alteración u omisión de modificaciones post-traduccionales, por ejemplo glicosilación o escisión, presentes cuando se expresa en una célula nativa.

En una realización preferida, el polipéptido 16836 tiene una o más de las siguientes características:

(i) tiene la capacidad de catalizar la hidrólisis de fosfatidilinositol;

(ii) tiene la capacidad de asociarse con ras, preferiblemente ras activado (unido a GTP);

(iii) tiene la capacidad de mediar la actividad de intercambio de nucleótidos de guanina;

(iv) tiene un peso molecular, por ejemplo, un peso molecular deducido, preferiblemente ignorando cualquier contribución de modificaciones post-traduccionales, composición de aminoácidos u otras características físicas de un polipéptido 16836, por ejemplo un polipéptido de la SEC ID N°: 2;

(v) tiene una similitud de secuencia global de al menos 80, 90 ó 95%, con un polipéptido de la SEC ID N°: 2;

(vi) puede encontrarse en células óseas, células de hígado fetal, células cerebrales, músculo esquelético, testículos, mama o corazón;

(vii) tiene un dominio RasGEF con una similitud de secuencia global de aproximadamente 70%, 80%, 90% o 95% con los restos aminoacídicos 35-338 de la SEC ID N°: 2;

(viii) tiene un dominio PI-PLC-X con una similitud de secuencia global de aproximadamente 70%, 80%, 90% o 95% con los restos aminoacídicos 900-1048 de la SEC ID N°: 2;

(ix) tiene un dominio PI-PLC-Y con una similitud de secuencia global de aproximadamente 70%, 80%, 90% o 95% con los restos aminoacídicos 1171-1184 o 1261-1353 de la SEC ID N°: 2;

(x) tiene un dominio C2 con una similitud de secuencia global de aproximadamente 70%, 80%, 90% o 95% con los restos aminoacídicos 1378-1460 de la SEC ID N°: 2;

(xi) tiene un dominio RA con una similitud de secuencia global de aproximadamente 70%, 80%, 90% o 95% con los restos aminoacídicos 1640-1745 de la SEC ID N°: 2;

(xii) puede colocalizarse con la proteína ras; o

(xiii) tiene al menos 70%, preferiblemente 80% y aún más preferiblemente 90% de las cisteínas encontradas en la secuencia de aminoácidos de la proteína nativa.

En una realización preferida, la proteína 16836, o sus fragmentos, difiere de la secuencia correspondiente de la SEC ID N°: 2. En una realización, difiere en al menos uno pero menos de 15, 10 ó 5 restos aminoacídicos. En otra difiere de

ES 2 328 448 T3

la secuencia correspondiente de la SEC ID N°: 2 en al menos un resto pero menos de 20%, 15%, 10% o 5% de los restos en ella difieren de la secuencia correspondiente de la SEC ID N°: 2. (Si esta comparación requiere un alineamiento de las secuencias, deben alinearse para una homología máxima. Las secuencias de “bucles” de deleciones o inserciones, o desacoplamiento se consideran diferencias). Las diferencias son, preferiblemente, diferencias o cambios en un resto no esencial o una sustitución conservativa. En una realización preferida, las diferencias no están en los dominios de fosfolipasa C, RA o RasGEF. En otra realización preferida una o más diferencias están en los dominios de fosfolipasa C, RA o RasGEF.

Otras realizaciones incluyen una proteína que contiene uno o más cambios en la secuencia de aminoácidos, por ejemplo, un cambio en un resto aminoacídico que no es esencial para la actividad. Estas proteínas 16836 difieren en la secuencia de aminoácidos de la SEC ID N°: 2, aunque retienen la actividad biológica.

En una realización, la proteína incluye una secuencia de aminoácidos con una homología de al menos aproximadamente 80%, 85%, 90%, 95%, 98% o mayor con la SEC ID N°: 2.

Se proporciona una proteína 16836 o un fragmento que varía de la secuencia de la SEC ID N°: 2 en regiones definidas por los aminoácidos aproximadamente 1-57, 233-899, 1049-1170, 1185-1260, 1354-1377, 1461-1639 ó 1746-1809 de la SEC ID N°: 2 en al menos uno pero menos de 15, 10 ó 5 restos aminoácidos de la proteína o fragmento, pero que no difiere de la SEC ID N°: 2 en regiones definidas por los aminoácidos aproximadamente 35-338, 900-1048, 1171-1184, 1261-1353, 1378-1460 ó 1640-1745 de la SEC ID N°: 2. (Si esta comparación requiere un alineamiento de las secuencias, deben alinearse para una homología máxima. Las secuencias de “bucles” de deleciones o inserciones, o desacoplamiento se consideran diferencias). En algunas realizaciones, la diferencia está en un resto no esencial o es una sustitución conservativa, mientras que en otras la diferencia está en un resto esencial o es una sustitución no conservativa.

En una realización, una parte biológicamente activa de una proteína 16836 incluye un dominio de fosfolipasa C, RA o RasGEF. Además, pueden prepararse otras partes biológicamente activas en las que están delecionadas otras regiones de la proteína por técnicas recombinantes y evaluarse con respecto a una o más de las actividades funcionales de una proteína 16836 nativa.

En una realización preferida, la proteína 16836 tiene una secuencia de aminoácidos mostrada en la SEC ID N°: 2. En otras realizaciones, la proteína 16836 es sustancialmente idéntica a la SEC ID N°: 2. En otra realización, la proteína 16836 es sustancialmente idéntica a la SEC ID N°: 2 y retiene la actividad funcional de la proteína de la SEC ID N°: 2, como se describe con detalle en las subsecciones anteriores.

Proteínas 16836 Quiméricas o de Fusión

En otro aspecto, la invención proporciona proteínas 16836 quiméricas o de fusión. Como se usa en la presente memoria, una “proteína quimérica” o “proteína de fusión” 16836 incluye un polipéptido 16836 unido a un polipéptido que no es 16836. Un “polipéptido que no es 16836” se refiere a un polipéptido que tiene una secuencia de aminoácidos correspondiente a una proteína que no es sustancialmente homóloga a la proteína 16836, por ejemplo, una proteína que es diferente de la proteína 16836 y que procede del mismo organismo o de un organismo diferente. El polipéptido 16836 de la proteína de fusión puede corresponder a todo o a una parte, por ejemplo, a un fragmento descrito en la presente memoria de una secuencia de aminoácidos de 16836. En una realización preferida, una proteína de fusión de 16836 incluye al menos una (o dos) partes biológicamente activas de una proteína 16836. El polipéptido que no es 16836 puede fusionarse al extremo N o al extremo C del polipéptido 16836.

La proteína de fusión puede incluir un resto que tiene alta afinidad por un ligando. Por ejemplo, la proteína de fusión puede ser una proteína de fusión GST-16836 en la que las secuencias de 16836 están fusionadas al extremo C de las secuencias de GST. Estas proteínas de fusión pueden facilitar la purificación de 16836 recombinante. Como alternativa, la proteína de fusión puede ser una proteína 16836 que contiene una secuencia señal heteróloga en su extremo N. En ciertas células hospedadoras (por ejemplo, células hospedadoras de mamífero), la expresión y/o secreción de 16836 puede aumentarse por medio del uso de una secuencia señal heteróloga.

Las proteínas de fusión pueden incluir todo o parte de una proteína sérica, por ejemplo, una región constante de IgG o albúmina de suero humano.

Las proteínas de fusión de 16836 de la invención pueden incorporarse en composiciones farmacéuticas y administrarse a un sujeto *in vivo*. Las proteínas de fusión de 16836 pueden usarse para afectar a la biodisponibilidad de un sustrato de 16836. Las proteínas de fusión de 16836 pueden ser terapéuticamente útiles para el tratamiento trastornos producidos, por ejemplo, por i) la mutación o modificación aberrante de un gen que codifica una proteína 16836; ii) la regulación alterada del gen de 16836; y (iii) una modificación post-traducciona aberrante de una proteína 16836.

Además, las proteínas de fusión de 16836 de la invención pueden usarse como inmunógenos para producir anticuerpos anti-16836 en un sujeto, para purificar ligandos de 16836 y en ensayos de exploración para identificar moléculas que inhiben la interacción de 16836 con sustrato de 16836.

Se dispone en el mercado de vectores de expresión que ya codifican una resto de fusión (por ejemplo, un polipéptido de GST). Un ácido nucleico que codifica 16836 puede clonarse en dicho vector de expresión de tal forma que el resto de fusión se una en fase a la proteína 16836.

5 *Anticuerpos Anti-16836*

En otro aspecto, la invención proporciona un anticuerpo anti-16836, o un fragmento del mismo (por ejemplo, un fragmento de unión a antígenos del mismo). El término “anticuerpo” como se usa en la presente memoria, se refiere a una molécula inmunoglobulina o parte inmunológicamente activa de la misma, es decir, una parte de unión a antígeno. Como se usa en la presente memoria, el término “anticuerpo” se refiere a una proteína que comprende al menos una, y preferiblemente dos, regiones variables de cadena pesada (H) (abreviadas en la presente memoria como VH), y al menos una y preferiblemente dos regiones variables de cadena ligera (L) (abreviadas en la presente memoria como VL). Las regiones VH y VL pueden subdividirse adicionalmente en regiones de hipervariabilidad, denominadas “regiones determinantes de complementariedad” (“CDR”), intercaladas con regiones que están más conservadas denominadas “regiones flanqueantes” (FR). El grado de la región flanqueante y las CDR se ha definido de forma precisa (véase, Kabat, E. A., *et al.* (1991) *Sequences of Proteins of Immunological Interest*, Quinta Edición, U. S. Department of Health and Human Services, NIH Publication No. 91-3242, y Chothia, C. *et al.* (1987) *J. Mol. Biol.* 196: 901-917, que se incorpora en la presente memoria como referencia). Cada VH y VL se compone de tres CDR y cuatro FR, dispuestas desde el extremo amino al extremo carboxi en el siguiente orden: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4.

El anticuerpo anti-16836 puede incluir además una región constante de cadena pesada y ligera, para formar de esta manera una cadena de inmunoglobulina pesada y ligera, respectivamente. En una realización, el anticuerpo es un tetrámero de dos cadenas pesadas de inmunoglobulina y dos cadenas ligeras de inmunoglobulina, donde las cadenas pesada y ligeras de inmunoglobulina están interconectadas, por ejemplo, por enlaces disulfuro. La región constante de cadena pesada comprende tres dominios, CH1, CH2 y CH3. La región constante de cadena ligera comprende un dominio, CL. La región variable de las cadenas pesadas y ligeras contiene un dominio de unión que interacciona con un antígeno. Las regiones constantes de los anticuerpos típicamente median la unión del anticuerpo a tejidos o factores del hospedador, incluyendo diversas células del sistema inmune (por ejemplo células efectoras) y el primer componente (Clq) del sistema clásico de complemento.

Como se usa en la presente memoria, el término “inmunoglobulina” se refiere a una proteína que consiste en uno o más polipéptidos sustancialmente codificados por genes de inmunoglobulina. Los genes de inmunoglobulina humanos reconocidos incluyen los genes de región constante kappa, lambda, alfa (IgA1 e IgA2), gamma (IgG1, IgG2, IgG3, IgG4), delta, épsilon y mu, así como la miríada de genes de región variable de inmunoglobulina. Las “cadenas ligeras” de inmunoglobulina de longitud completa (de aproximadamente 25 KDa o 214 aminoácidos) se codifican por un gen de región variable en el extremo NH₂ (aproximadamente 110 aminoácidos) y un gen de región constante kappa o lambda en el extremo COOH. Las “cadenas pesadas” de inmunoglobulina de longitud completa (aproximadamente 50 KDa o 446 aminoácidos), se codifican de forma similar por un gen de región variable (de aproximadamente 116 aminoácidos) y uno de los otros genes de región constante mencionados anteriormente, por ejemplo gamma (que codifica aproximadamente 330 aminoácidos).

La expresión “fragmento de unión a antígeno” de un anticuerpo (o simplemente “parte de anticuerpo” o “fragmento”), como se usa en la presente memoria, se refiere a uno o más fragmentos de un anticuerpo de longitud completa que retienen la capacidad de unirse específicamente al antígeno, por ejemplo, el polipéptido 16836 o un fragmento del mismo. Los ejemplos de fragmentos de unión a antígeno del anticuerpo anti-16836 incluyen, pero sin limitación: (i) un fragmento Fab, un fragmento monovalente que consiste en los dominios VL, VH, CL y CH1; (ii) un fragmento F(ab')₂, un fragmento bivalente que comprende dos fragmentos Fab unidos por un enlace disulfuro a la región de bisagra; (iii) un fragmento Fd que consiste en los dominios VH y CH1; (iv) un fragmento Fv que consiste en los dominios VL y VH de un solo brazo de un anticuerpo, (v) un fragmento dAb (Ward *et al.*, (1989) *Nature* 341: 544-546), que consiste en un dominio VH; y (vi) una región determinante de complementariedad (CDR) aislada. Además, aunque los dos dominios del fragmento Fv, VL y VH se codifican por dos genes distintos, pueden unirse, usando métodos recombinantes, por un enlazador sintético que permite que se obtengan como una sola cadena proteica en la que las regiones VL y VH están emparejadas para formar moléculas monovalentes (conocidas como Fv monocatenario (scFv); véase, por ejemplo, Bird *et al.* (1988) *Science* 242: 423-426; y Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85: 5879-5883). Dentro de la expresión “fragmento de unión a antígeno” de un anticuerpo también se incluyen estos anticuerpos monocatenarios. Estos fragmentos de anticuerpo se obtienen usando técnicas convencionales conocidas por los especialistas en la técnica y los fragmentos se exploran con respecto a la utilidad de la misma manera que los anticuerpos intactos.

El anticuerpo anti-16836 puede ser un anticuerpo policlonal o monoclonal. En otras realizaciones, el anticuerpo puede producirse de forma recombinante, por ejemplo, puede producirse por presentación en fagos o por métodos combinatorios.

La presentación en fagos y los métodos combinatorios para generar anticuerpos anti-16836 se conocen en la técnica (como se describe, por ejemplo, en Ladner *et al.* Patente de los Estados Unidos N° 5.223.409; Kang *et al.* Publicación Internacional N° WO 92/18619; Dower *et al.* Publicación Internacional N° WO 91/17271; Winter *et al.* Publicación Internacional N° WO 92/20791; Markland *et al.* Publicación Internacional N° WO 92/15679; Breitling *et al.* Publicación

Internacional N° WO 93/01288; McCafferty *et al.* Publicación Internacional N° WO 92/01047; Garrard *et al.* Publicación Internacional N° WO 92/09690; Ladner *et al.* Publicación Internacional N° WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9: 1370-1372; Hay *et al.* (1992) *Hum Antibod Hybridomas* 3: 81-85; Huse *et al.* (1989) *Science* 246: 1275-1281. Griffiths *et al.* (1993) *EMBO J* 12: 725-734. Hawkins *et al.* (1992) *J Mol Biol* 226: 889-896; Clackson
 5 *et al.* (1991) *Nature* 352: 624-628. Gram *et al.* (1992) *PNAS* 89: 3576-3580; Garrard *et al.* (1991) *Bio/Technology* 9: 1373-1377; Hoogenboom *et al.* (1991) *Nuc AcidRes* 19: 4133-4137; y Barbas *et al.* (1991) *PNAS* 88: 7978-7982, cuyos contenidos se incorporan en la presente memoria como referencia).

En una realización, el anticuerpo anti-16836 es un anticuerpo completamente humano (por ejemplo un anticuerpo
 10 obtenido en un ratón que se ha modificado por ingeniería genética para producir un anticuerpo a partir de una secuencia de inmunoglobulina humana) o un anticuerpo no humano, por ejemplo un anticuerpo de roedor (ratón o rata), cabra, primate (por ejemplo mono), o camello. Preferiblemente, el anticuerpo no humano es de un roedor (un anticuerpo de ratón o de rata). En la técnica se conocen métodos para producir anticuerpos de roedor.

Pueden generarse anticuerpos monoclonales humanos usando ratones transgénicos que llevan genes de inmunoglobulina humana en lugar del sistema de ratón. Se usan esplenocitos de estos ratones transgénicos inmunizados con el antígeno de interés para producir hibridomas que secretan mAb humanos con afinidades específicas por epítomos de una proteína humana (véase, por ejemplo, Wood *et al.* Solicitud Internacional WO 91/00906, Kuchelapati *et al.* Publicación PCT WO 91/10741; Lonberg *et al.* Solicitud Internacional N° WO 92/03918; Kay *et al.* Solicitud Internacional N° WO 92/03917; Lonberg, N. *et al.* 1994 *Nature* 368: 856-859; Green, L. L. *et al.* 1994 *Nature Medicine* 7: 13-21; Morrison, S. L. *et al.* 1994 *Proc. Natl. Acad. Sci. USA* 81: 6851-6855; Bruggeman *et al.* 1993 *Year Immunol.* 7: 33-40; Tuailon *et al.* 1993 *PNAS* 90: 3720-3724; Bruggeman *et al.* 1991) *Eur J. Immunol.* 21: 1323-1326).

Un anticuerpo anti-16836 puede ser uno en el que la región variable, o una parte de la misma, por ejemplo, las
 25 CDR, se genera en un organismo no humano, por ejemplo, una rata o ratón. Están dentro de la invención anticuerpos quiméricos, con CDR injertadas y humanizadas. Están dentro de la invención anticuerpos generados en un organismo no humano, por ejemplo, una rata o ratón, y después modificados, por ejemplo, en la región flanqueante variable o constante, para reducir la antigenicidad en un ser humano.

Los anticuerpos quiméricos se pueden producir por técnicas de ADN recombinante conocidas en este campo. Por
 ejemplo, un gen que codifica la región constante Fc de una molécula de anticuerpo monoclonal murino (o de otra especie) se digiere con enzimas de restricción para retirar la región que codifica la Fc murina y se sustituye por la parte equivalente de un gen que codifica una región constante Fc humana (véase Robinson *et al.*, Publicación de Patente Internacional PCT/US86/02269; Akira, *et al.*, Solicitud de Patente Europea 184,187; Taniguchi, M., Solicitud de Patente Europea 171.696; Morrison *et al.*, Solicitud de Patente Europea 173.494; Neuberger *et al.*, Solicitud Internacional WO 86/01533; Cabilly *et al.* Patente de Estados Unidos N° 4.816.567; Cabilly *et al.*, Solicitud de Patente Europea 125.023; Better *et al.* (1988 *Science* 240: 1041-1043); Liu *et al.* (1987) *PNAS* 84: 3439-3443; Liu *et al.*, 1987, *J. Immunol.* 139: 3521-3526; Sun *et al.* (1987) *PNAS* 84: 214-218; Nishimura *et al.*, 1987, *Canc. Res.* 47: 999-1005; Wood *et al.* (1985) *Nature* 314: 446-449; y Shaw *et al.*, 1988, *J. Natl Cancer Inst.* 80: 1553-1559).

Un anticuerpo humanizado o con CDR injertadas tendrá al menos una o dos, pero generalmente las tres CDR del
 receptor (de cadenas de inmunoglobulina pesadas o ligeras) reemplazadas por una CDR del donante. Un anticuerpo puede reemplazarse por al menos una parte de una CDR no humana o sólo algunas de las CDR pueden reemplazarse por CDR no humanas. Sólo es necesario reemplazar el número de CDR necesarias para unir el anticuerpo humanizado a una 16836 o un fragmento de la misma. Preferiblemente, el donante será un anticuerpo de roedor, por ejemplo, un anticuerpo de rata o ratón, y el receptor será una región flanqueante humana o una región flanqueante humana consenso. Típicamente, la inmunoglobulina que proporciona las CDR se denomina “donante” y la inmunoglobulina que proporciona la región flanqueante se denomina “aceptor.” En una realización, la inmunoglobulina donante es no humana (por ejemplo de roedor). La región flanqueante del aceptador es una región flanqueante natural (por ejemplo, humana) o una región flanqueante consenso, o una secuencia con una identidad de aproximadamente 85% o mayor, preferiblemente 90%, 95%, 99% o mayor.

Como se usa en la presente memoria, la expresión “secuencia consenso” se refiere a la secuencia formada a partir de los aminoácidos (o nucleótidos) que aparecen con más frecuencia en una familia de secuencias relacionadas (Véase, por ejemplo, Winnaker, *From Genes to Clones* (Verlagsgesellschaft, Weinheim, Alemania 1987). En una familia de proteínas, cada posición en la secuencia consenso está ocupada por el aminoácido que aparece con más frecuencia en esa posición de la familia. Si dos aminoácidos aparecen con la misma frecuencia, cualquiera puede incluirse en la secuencia consenso. Una “región flanqueante consenso” se refiere a la región flanqueante en la secuencia de inmunoglobulina consenso.

Un anticuerpo puede humanizarse por métodos conocidos en la técnica. Los anticuerpos humanizados pueden generarse reemplazando secuencias de la región variable de Fv que no están implicadas directamente en la unión al antígeno con secuencias equivalentes de regiones variables de Fv humanas. Se proporcionan métodos generales para generar anticuerpos humanizados por Morrison, S. L., 1985, *Science* 229: 1202-1207, por Oi *et al.*, 1986, *BioTechniques* 4: 214, y por Queen *et al.* documentos US 5.585.089, US 5.693.761 y US 5.693.762, cuyos contenidos se incorporan en la presente memoria como referencia. Estos métodos incluyen el aislamiento, manipulación y expresión de las secuencias de ácido nucleico que codifican todo o parte de las regiones variables de Fv de inmunoglobulina de al menos una de una cadena pesada o ligera. Las fuentes de dicho ácido nucleico son bien conocidas por los especialistas

en la técnica y, por ejemplo, pueden obtenerse a partir de un hibridoma que produce un anticuerpo contra un polipéptido 16836 o un fragmento del mismo. El ADN recombinante que codifica el anticuerpo humanizado, o fragmento del mismo, después puede clonarse en un vector de expresión apropiado.

5 Los anticuerpos humanizados o con injertos de CDR pueden producirse por injerto de CDR o sustitución de CDR, donde pueden reemplazarse una, dos o todas las CDR de una cadena de inmunoglobulina. Véase, por ejemplo, la Patente de Estados Unidos 5.225.539; Jones *et al.* 1986 Nature 321: 552-525; Verhoeyan *et al.* 1988 Science 239: 1534; Beidler *et al.* 1988 J. Immunol. 141: 4053-4060; Winter, documento US 5.225.539, cuyo contenido se incorpora expresamente en la presente memoria como referencia. Winter describe un método de injerto de CDR que puede
10 usarse para preparar los anticuerpos humanizados de la presente invención (Solicitud de Patente del Reino Unido GB 2188638A, presentada el 26 de marzo de 1987; Winter documento US 5.225.539), cuyo contenido se incorpora expresamente como referencia.

También se incluyen dentro del alcance de la invención anticuerpos humanizados en los que se han sustituido, 15 delecionado o añadido aminoácidos específicos. Los anticuerpos humanizados preferidos tienen sustituciones de aminoácidos en la región flanqueante, tal como para mejorar la unión al antígeno. Por ejemplo, un anticuerpo humanizado tendrá restos flanqueante idénticos al resto flanqueante del donante o a otro aminoácido distinto del resto flanqueante del receptor. Para generar estos anticuerpos, un número pequeño seleccionado de restos flanqueantes del aceptor de la cadena de inmunoglobulina humanizada pueden reemplazarse por los aminoácidos correspondientes del donante. Las 20 localizaciones preferidas de las sustituciones incluyen restos aminoácidos adyacentes a la CDR, o que son capaces de interactuar con una CDR (véase, por ejemplo, el documento US 5.585.089). Se describen criterios para seleccionar aminoácidos del donante en el documento US 5.585.089, por ejemplo, en las columnas 12-16 del documento US 5.585.089, por ejemplo las columnas 12-16 del documento US 5.585.089, cuyos contenidos se incorporan en la presente memoria como referencia. Se describen otras técnicas para humanizar anticuerpos en Padlan *et al.* documento 25 EP 519596 A1, publicado el 23 de diciembre de 1992.

En realizaciones preferidas, un anticuerpo puede obtenerse inmunizando con antígeno 16836 purificado, o un fragmento del mismo, por ejemplo, un fragmento descrito en la presente memoria, tejido, por ejemplo, preparaciones de tejido en bruto, células enteras, preferiblemente células vivas, células lisadas o fracciones celulares.

30 Una proteína 16836 de longitud completa o un fragmento peptídico antigénico de 16836 puede usarse como inmunógeno o puede usarse para identificar anticuerpos anti-16836 obtenidos con otros inmunógenos, por ejemplo, células, preparaciones de membrana y similares. El péptido antigénico de 16836 debe incluir al menos 8 restos aminoácidos de la secuencia de aminoácidos mostrada en la SEC ID N°: 2 e incluye un epítipo de 16836. Preferiblemente, el péptido antigénico incluye al menos 10 restos aminoácidos, más preferiblemente al menos 15 restos aminoácidos, incluso más preferiblemente al menos 20 restos aminoácidos y aún más preferiblemente al menos 30 restos aminoácidos.

Pueden usarse fragmentos de 16836 que incluyen restos aproximadamente 330-350, aproximadamente 610-630 o 40 aproximadamente 1120-1140 para obtener anticuerpos contra regiones hidrófilas de la proteína 16836, por ejemplo, usados como inmunógenos o usados para caracterizar la especificidad de un anticuerpo. De forma similar, pueden usarse fragmentos de 16836 que incluyen los restos aproximadamente 200-230, aproximadamente 430-450, o aproximadamente 830-840 para obtener un anticuerpo contra una región hidrófoba de la proteína 16836. Un fragmento de 16836 que incluye los restos aproximadamente 35-338 puede usarse para obtener un anticuerpo contra la región 45 RasGEF de la proteína 16836. Un fragmento de 16836 que incluye los restos aproximadamente 35-338 puede usarse para fabricar un anticuerpo contra la región RasGEF de la proteína 16836. Un fragmento de 16836 que incluye los restos aproximadamente 900-1048 puede usarse para obtener un anticuerpo contra la región PI-PLC-X de la proteína 16836. Un fragmento de 16836 que incluye los restos aproximadamente 1171-1184 o 1261-1353 puede usarse para obtener un anticuerpo contra la región PI-PLC-Y de la proteína 16836. Un fragmento de 16836 que incluye los restos 50 aproximadamente 1378-1460 puede usarse para obtener un anticuerpo contra la región C2 de la proteína 16836. Un fragmento de 16836 que incluye los restos aproximadamente 1640-1745 puede usarse para obtener un anticuerpo contra la región RA de la proteína 16836.

Se proporcionan anticuerpos reactivos con o específicos para cualquiera de estas regiones, u otras regiones o 55 dominios descritos en la presente memoria.

Los anticuerpos que se unen sólo a la proteína 16836 nativa, sólo a la proteína 16836 desnaturalizada o no nativa de otra manera o que se unen a ambas proteínas están dentro de la invención. También están dentro de la invención anticuerpos con epítopos lineales o conformacionales. Los epítopos conformacionales algunas veces pueden identificarse 60 identificando anticuerpos que se une a la proteína nativa pero no a la proteína 16836 desnaturalizada.

Los epítopos preferidos incluidos por el péptido antigénico son regiones de 16836 que están localizadas en la superficie de la proteína, por ejemplo regiones hidrófilas, así como regiones con alta antigenicidad. Por ejemplo, puede usarse un análisis de probabilidad de localización en superficie de Emini de la secuencia de la proteína 16836 65 humana para indicar las regiones que tienen una probabilidad particularmente elevada de localizarse en la superficie de la proteína 16836 y, por lo tanto, es probable que constituyan restos en la superficie útiles para dirigirse a la producción de anticuerpos.

En realizaciones preferidas los anticuerpos pueden unirse a uno o más antígenos purificados, tejidos, por ejemplo secciones tisulares, células enteras, preferiblemente células vivas, células lisadas o fracciones celulares.

5 El anticuerpo anti-16836 puede ser un anticuerpo monocatenario. Un anticuerpo monocatenario (scFV) puede obtenerse por ingeniería genética (véase, por ejemplo, Colcher, D. *et al.* (1999) *Ann N Y Acad Sci* 880: 263-80; y Reiter, Y. (1996) *Clin Cancer Res* 2: 245-52). El anticuerpo monocatenario puede dimerizarse o multimerizarse para generar anticuerpos multivalentes que tienen especificidades por diferentes epítomos de la misma proteína 16836 diana.

10 En una realización preferida, el anticuerpo tiene: función efectora; y puede fijar el complemento. En otra realización, el anticuerpo no recluta células efectoras; ni se fija al complemento.

15 En una realización preferida, el anticuerpo tiene una capacidad reducida o nula de unirse a un receptor Fc. Por ejemplo, es un isotipo o subtipo, fragmento u otro mutante, que no soporta la unión a un receptor Fc, por ejemplo, tiene una región de unión al receptor Fc mutagenizada o delecionada.

En una realización preferida, un anticuerpo anti-16836 altera (por ejemplo aumenta o reduce) la capacidad de un polipéptido 16836 de catalizar la hidrólisis del fosfatidilinositol o asociarse con Ras.

20 El anticuerpo puede acoplarse a una toxina, por ejemplo, una toxina polipeptídica, por ejemplo ricina o la toxina diftérica o un fragmento activo de la misma, o un núcleo radiactivo, o un agente de formación de imágenes, por ejemplo, un agente radiactivo, enzimático y otro agente, por ejemplo, un agente de formación de imágenes, por ejemplo, un agente de contraste de RMN. Se prefieren marcadores que producen emisiones radiactivas detectables o fluorescencia.

25 Un anticuerpo anti-16836 (por ejemplo, anticuerpo monoclonal) puede usarse para aislar 16836 por técnicas convencionales, tales como cromatografía de afinidad o inmunoprecipitación. Además, un anticuerpo anti-16836 puede usarse para detectar la proteína 16836 (por ejemplo, en un lisado celular o sobrenadante celular) para evaluar la abundancia y patrón de expresión de la proteína. Los anticuerpos anti-16836 pueden usarse diagnósticamente para controlar niveles de proteínas en un tejido como parte de un procedimiento de ensayo clínico, por ejemplo, para determinar la eficacia de un régimen de tratamiento dado. La detección puede facilitarse por acoplamiento (es decir unión física) del anticuerpo o una sustancia detectable (es decir marcaje con anticuerpos). Los ejemplos de sustancias detectables incluyen diversas enzimas, grupos prostéticos, materiales fluorescentes, materiales luminescentes, materiales bioluminescentes y materiales radiactivos. Los ejemplos de las enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina, β -galactosidasa, o acetilcolinesterasa; los ejemplos de los complejos de grupos prostéticos adecuados incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de los materiales fluorescentes adecuados incluyen umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, diclorotriazinilamina fluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de un material bioluminiscente incluye luminol; los ejemplos de materiales bioluminescentes incluyen luciferasa, luciferina y acuorina y los ejemplos de materiales radiactivos adecuados incluyen ^{125}I , ^{131}I , ^{35}S o ^3H . La invención también incluye un ácido nucleico que codifica un anticuerpo anti-16836, por ejemplo, un anticuerpo anti-16836 descrito en la presente memoria. También se incluyen vectores que incluyen el ácido nucleico y células transformadas con el ácido nucleico, particularmente células que son útiles para producir un anticuerpo, por ejemplo, células de mamífero, por ejemplo células CHO o células linfáticas.

45 La invención también incluye líneas celulares, por ejemplo, hibridomas, que fabrican un anticuerpo anti-16836, por ejemplo, y un anticuerpo descrito en la presente memoria, y un método para usar dichas células para obtener un anticuerpo 16836.

Células Hospedadoras de Vectores de Expresión Recombinantes y Células Obtenidas por Ingeniería Genética

50 En otro aspecto, la invención incluye vectores, preferiblemente vectores de expresión, que contienen un ácido nucleico que codifica un polipéptido descrito en la presente memoria. Como se usa en la presente memoria, el término “vector” se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido y puede incluir un plásmido, cósmido o vector viral. El vector puede replicarse de forma autónoma o puede integrarse en un ADN del hospedador. Los vectores virales incluyen, por ejemplo, retrovirus con defectos en la replicación, adenovirus y virus adenoasociados.

55 Un vector puede incluir un ácido nucleico de 16836 en una forma adecuada para la expresión del ácido nucleico en una célula hospedadora. Preferiblemente, el vector de expresión recombinante incluye una o más secuencias reguladoras unidas operativamente a la secuencia de ácido nucleico a expresar. El término “secuencia reguladora” incluye promotores, potenciadores y otros elementos de control de la expresión (por ejemplo señales de poliadenilación). Las secuencias reguladoras incluyen las que dirigen la expresión constitutiva de una secuencia de nucleótidos, así como secuencias inducibles y/o reguladoras con especificidad de tejido. El diseño del vector de expresión puede depender de factores tales como la elección de la célula hospedadora a transformar, el nivel de expresión de proteína deseado y similares. Los vectores de expresión de la invención pueden introducirse en las células hospedadoras para producir de esta manera proteínas o polipéptidos, incluyendo proteínas o polipéptidos de fusión, codificados por ácidos nucleicos como se describe en la presente memoria (por ejemplo, proteínas 16836, formas mutantes de proteínas 16836, proteínas de fusión y similares).

ES 2 328 448 T3

Los vectores de expresión recombinantes de la invención pueden diseñarse para la expresión de proteínas 16836 en células procarióticas o eucarióticas. Por ejemplo, pueden expresarse polipéptidos de la invención en *E. coli*, células de insecto (por ejemplo usando vectores de expresión de baculovirus), células de levadura o células de mamífero. Se describen células hospedadora adecuadas adicionalmente en Goeddel, (1990) Gene Expression Technology: Methods
5 in Enzymology 185, Academic Press, San Diego, CA. Como alternativa, el vector de expresión recombinante puede transcribirse y traducirse *in vitro*, por ejemplo, usando secuencias reguladoras del promotor de T7 y polimerasa de T7.

La expresión de proteínas en procariotas se realiza la mayoría de las veces en *E. coli* con vectores que contienen
10 promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o proteínas que no son proteínas de fusión. Los vectores de fusión añaden varios aminoácidos a una proteína codificada en los mismos, normalmente en el extremo amino de la proteína recombinante. Estos vectores de fusión típicamente tiene tres fines: 1) aumentar la expresión de la proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante; y 3) ayudar en la purificación de la proteína recombinante actuando como ligando en la purificación por afinidad. A menudo, se introduce un sitio de escisión proteolítico en la unión del resto de fusión y la proteína recombinante para permitir la
15 separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Estas enzimas y sus secuencias de reconocimiento afines incluyen el Factor Xa, trombina y enteroquinasa. Los vectores de expresión de fusión típicos incluyen pGEX (Pharmacia Biotech Inc; Smith, D.B. y Johnson, K. S. (1988) Gene 67: 31-40), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ) que fusionan glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, a la proteína recombinante
20 diana.

En ensayos de actividad de 16836 pueden usarse proteínas de fusión purificadas (por ejemplo, ensayos directos o ensayos competitivos descritos con detalle más adelante) o para generar anticuerpos específicos para proteínas 16836.
25 En una realización preferida, puede usarse una proteína de fusión expresada en un vector de expresión retroviral de la presente invención para infectar células de médula ósea, que posteriormente se transplantan en receptores irradiados. Posteriormente, se examina la patología del receptor después de que haya transcurrido un tiempo suficiente (por ejemplo, seis semanas).

Para maximizar la expresión de proteínas recombinantes en *E. coli*, se expresa la proteína en una bacteria hospedadora con una capacidad deteriorada de escindir proteolíticamente la proteína recombinante (Gottesman, S. (1990) *Methods Enzymol.* 1990:Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California 119-128). Otra estrategia es alterar la secuencia de ácido nucleico del ácido nucleico a insertar en un vector de expresión de forma que los codones individuales para cada aminoácido sean los utilizados preferentemente en *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20: 2111-2118). Esta alteración de secuencias de ácido nucleico de la invención puede realizarse por técnicas de síntesis de ADN convencionales.
30
35

El vector de expresión de 16836 puede ser un vector de expresión de levadura, un vector para la expresión en células de insecto, por ejemplo, un vector de expresión de baculovirus o un vector adecuado para la expresión en células de mamífero.
40

Cuando se usa en células de mamíferos, las funciones de control del vector de expresión pueden proporcionarse por elementos reguladores virales. Por ejemplo, los promotores usados comúnmente proceden de poliovirus, Adenovirus 2, citomegalovirus y virus de simio 40.
45

En otra realización, el promotor es un promotor inducible, por ejemplo, un promotor regulado por una hormona esteroidea, por una hormona polipeptídica (por ejemplo, por medio de una ruta de transducción de señales), o por un polipéptido heterólogo (por ejemplo, los sistemas inducibles por tetraciclina “Tet-On” y “Tet-Off”; véase, por ejemplo, Clontech Inc., CA, Gossen y Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89: 5547, y Paillard (1989) *Human Gene Therapy* 9: 983).
50

En otra realización, el vector de expresión de mamífero recombinante es capaz de dirigir la expresión del ácido nucleico preferencialmente en un tipo celular particular (por ejemplo, se usan elementos reguladores con especificidad de tejido para expresar el ácido nucleico). Los ejemplos no limitantes de promotores con especificidad de tejido adecuado incluyen el promotor de albúmina (con especificidad de hígado; Pinkert *et al.* (1987) *Genes Dev.* 1: 268-277), promotores con especificidad linfoide (Calame y Eaton (1988) *Adv. Immunol.* 43: 235-275), en particular promotores de receptores de células T (Winoto y Baltimore (1989) *EMBO J.* 8: 729-733) e inmunoglobulinas (Banerji *et al.* (1983) *Cell* 33: 729-740; Queen y Baltimore (1983) *Cell* 33: 741-748), promotores con especificidad de neurona (por ejemplo, el promotor del neurofilamento; Byrne y Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), promotores con especificidad de páncreas (Edlund *et al.* (1985) *Science* 230: 912-916), y promotores con especificidad de glándula mamaria (por ejemplo, promotor de suero de leche; Patente de Estados Unidos N° 4.873.316 y Publicación de Solicitud Europea N° 264.166). También se incluyen promotores regulados en el desarrollo, por ejemplo, los promotores de *hox* murinos (Kessel y Gruss (1990) *Science* 249: 374-379) y el promotor de α -fetoproteína (Campes y Tilghman (1989) *Genes Dev.* 3: 537-546).
55
60
65

La invención también proporciona un vector de expresión recombinante que comprende una molécula de ADN de la invención clonada en el vector de expresión en una orientación antisentido. Pueden elegirse secuencias reguladoras (por ejemplo, promotores y/o potenciadores virales) unidos operativamente a un ácido nucleico clonado en la

orientación antisentido que dirijan la expresión constitutiva, específica de tejido o específica de tipo celular del ARN antisentido en una diversidad de tipos celulares. El vector de expresión antisentido puede estar forma de un plásmido recombinante, fagémico o virus atenuado.

5 Otro aspecto de la invención proporciona una célula hospedadora que incluye una molécula de ácido nucleico descrita en la presente memoria, por ejemplo, una molécula de ácido nucleico de 16836 dentro de un vector de expresión recombinante o una molécula de ácido nucleico de 16836 que contiene secuencias que permiten que se recombine homológamente en un sitio específico del genoma de la célula hospedadora. Las expresiones “célula hospedadora” y “célula hospedadora recombinante” se usan indistintamente en la presente memoria. Estos términos se refieren no sólo a la célula objeto particular, sino a la descendencia o posible descendencia de dicha célula. Como pueden realizarse ciertas modificaciones en generaciones sucesivas debido a una mutación o influencias ambientales, dicha descendencia efectivamente puede no ser idéntica a la célula parental, pero aún así se incluye dentro del alcance de la expresión como se usa en la presente memoria.

15 Una célula hospedadora puede ser cualquier célula procariótica o eucariótica. Por ejemplo, una proteína 16836 puede expresarse en células bacterianas (tales como *E. coli*), células de insecto, células de levadura o mamífero (tales como células de ovario de hámster chino (CHO) o células COS). Los especialistas en la técnica conocen otras células hospedadoras adecuadas.

20 El ADN del vector puede introducirse en células hospedadoras por técnicas de transformación o transfección convencionales. Como se usa en la presente memoria, los términos “transformación” y “transfección” pretenden hacer referencia a una diversidad de técnicas reconocidas en la técnica para introducir un ácido nucleico extraño (por ejemplo, un ADN) en una célula hospedadora, incluyendo coprecipitación con fosfato cálcico o cloruro cálcico, transfección mediada por DEAE-dextrano, lipofección o electroporación.

25 Una célula hospedadora de la invención puede usarse para producir (es decir expresar) una proteína 16836. Por consiguiente, la invención proporciona además métodos para producir una proteína 16836 usando las células hospedadoras de la invención. En una realización, el método incluye cultivar la célula hospedadora de la invención (en la que se ha introducido un vector de expresión recombinante que codifica una proteína 16836) en un medio adecuado de tal forma que se produzca la proteína 16836. En otra realización, el método incluye además aislar una proteína 16836 a partir del medio o de la célula hospedadora.

30 En otro aspecto, la invención se refiere a una célula o preparación purificada de células que incluye un transgen de 16836 o que expresa erróneamente de otra manera 16836. La preparación de células puede consistir en células humanas o no humanas, por ejemplo, células de roedor, por ejemplo, células de ratón o rata, células de conejo o células de cerdo. En realizaciones preferidas, la célula o las células incluyen un transgen de 16836 por ejemplo, una forma heteróloga de una proteína 16836, por ejemplo, un gen derivado de humanos (en el caso de una célula no humana). El transgen de 16836 puede expresarse de forma defectuosa, por ejemplo, tener una expresión excesiva o insuficiente, por ejemplo sobre-expresarse o infra-expresarse. En otras realizaciones preferidas, la célula o las células incluyen un gen que expresa de forma errónea una proteína 16836 endógena, por ejemplo, un gen cuya expresión se altera, por ejemplo, un knockout. Estas células pueden servir como modelo para estudiar trastornos que están relacionados con alelos de 16836 mutados o expresados de forma errónea o para uso en la exploración de fármacos.

45 En otro aspecto, la invención se refiere a una célula humana, por ejemplo, una célula madre hematopoyética, transformada con un ácido nucleico que codifica un polipéptido 16836 objeto.

50 También se proporcionan células, preferiblemente células humanas, por ejemplo células hematopoyéticas o fibroblastos humanos, en los que una proteína 16836 endógena está bajo el control de una secuencia reguladora que normalmente no controla la expresión del gen 16836 endógeno. Las características de expresión de un gen endógeno dentro de una célula, por ejemplo, una línea celular o microorganismo, pueden modificarse insertando un elemento regulador de ADN heterólogo en el genoma de la célula de tal forma que el elemento regulador insertado esté unido operativamente al gen de 16836 endógeno. Por ejemplo, un gen de 16836 endógeno que es “silencioso transcripcionalmente”, por ejemplo, no se expresa normalmente o se expresa sólo a niveles muy bajos, puede activarse por inserción de un elemento regulador que es capaz de promover la expresión de un producto génico que normalmente se expresa en esa célula. Pueden usarse técnicas tales como recombinaciones homólogas dirigidas para insertar el ADN heterólogo como se describe, por ejemplo, en Chappel, documento US 5.272.071; documento WO 91/06667, publicado el 16 de mayo de 1991.

60 En una realización preferida, pueden usarse células recombinantes descritas en la presente memoria para la terapia de reemplazo en un sujeto. Por ejemplo, un ácido nucleico que codifica un polipéptido 16836 unido operativamente a un promotor inducible (por ejemplo, un promotor regulado por el receptor de hormonas esteroideas) se introduce en una célula recombinante humana o no humana, por ejemplo, de mamífero, por ejemplo porcina. La célula se cultiva y se encapsula en un material biocompatible, tal como arginato de polilisisina, y posteriormente se implanta en el sujeto. Véase, por ejemplo, Lanza (1996) Nat. Biotechnol. 14: 1107; Joki *et al.* (2001) Nat. Biotechnol. 19: 35; y Patente de Estados Unidos N° 5.876.742. La producción del polipéptido 16836 puede regularse en el sujeto por medio de la administración de un agente (por ejemplo, una hormona esteroidea) al sujeto. En otra realización preferida, las células recombinantes implantadas expresan y secretan un anticuerpo específico para un polipéptido 16836. El anticuerpo puede ser cualquier anticuerpo o cualquier derivado de anticuerpo descrito en la presente memoria.

Usos

Las moléculas de ácido nucleico, proteínas, homólogos de proteína y anticuerpos descritos en la presente memoria pueden usarse en uno o más de los siguientes métodos: a) ensayos de exploración; b) medicina predictiva (por ejemplo, ensayos de diagnóstico, ensayos de pronóstico, ensayos clínicos de monitorización y farmacogenética); y c) métodos de tratamiento (por ejemplo, terapéuticos y profilácticos).

Las moléculas de ácido nucleico aisladas de la invención pueden usarse, por ejemplo, para expresar una proteína 16836 (por ejemplo, por medio de un vector de expresión recombinante en una célula hospedadora en aplicaciones de terapia génica), para detectar un ARNm de 16836 (por ejemplo, en una muestra biológica) o una alteración genética en un gen de 16836, y para modular la actividad de 16836, como se describe adicionalmente más adelante. Las proteínas 16836 pueden usarse para tratar trastornos caracterizados por una producción insuficiente o excesiva de un sustrato de 16836 o la producción de inhibidores de 16836. Además, las proteínas 16836 pueden usarse para seleccionar sustratos de 16836 naturales, para buscar fármacos o compuestos que modulan la actividad de 16836, así como para tratar trastornos caracterizados por una producción insuficiente o excesiva de proteína 16836 o la producción de formas de proteína 16836 que tienen una actividad reducida, aberrante o indeseada en comparación con la proteína 16836 de tipo silvestre (por ejemplo, trastornos de la proliferación y/o diferenciación celular). Además, los anticuerpos anti-16836 de la invención pueden usarse para detectar y aislar proteínas 16836, regular la biodisponibilidad de las proteínas 16836 y modular la actividad de 16836.

Se proporciona un método para evaluar en un compuesto la capacidad de interactuar con, por ejemplo, unirse a un polipéptido 16836 objeto. El método incluye: poner en contacto el compuesto con el polipéptido 16836 objeto; y evaluar la capacidad del compuesto de interactuar con, por ejemplo, de unirse o formar un complejo con el polipéptido 16836 objeto. Este método puede realizarse *in vitro*, por ejemplo en un sistema sin células, o *in vivo*, por ejemplo, en un ensayo de trampa de interacción doble híbrido. Este método puede usarse para identificar moléculas naturales que interactúan con el polipéptido 16836 objeto. También puede usarse para encontrar inhibidores naturales o sintéticos del polipéptido 16836 objeto. Más adelante se describen con más detalle métodos de exploración.

Ensayos de Exploración

La invención proporciona métodos (también denominados en la presente memoria “ensayos de exploración”) para identificar moduladores, es decir, compuestos o agentes candidatos o de ensayo (por ejemplo proteínas, péptidos, peptidomiméticos, peptoides, moléculas pequeñas u otros fármacos) que se unen a proteínas 16836, tienen un efecto estimulador o inhibidor, por ejemplo, sobre la expresión de 16836 o la actividad de 16836, o tienen un efecto estimulador o inhibidor, por ejemplo, sobre la expresión o la actividad de un sustrato de 16836. Los compuestos identificados de esta manera pueden usarse para modular la actividad de productos génicos diana (por ejemplo genes de 16836) en un protocolo terapéutico, para elaborar la función biológica del producto génico diana, o para identificar compuestos que alteran las interacciones normales con el gen diana.

En una realización, la invención proporciona ensayos para explorar compuestos candidatos o de ensayo que son sustratos de una proteína o polipéptido 16836 o una parte biológicamente activa del mismo. En otra realización, la invención proporciona ensayos para explorar compuestos candidatos o de ensayo que se unen o modulan una actividad de una proteína o polipéptido 16836 o una parte biológicamente activa del mismo.

Las actividades de una proteína 16836 que pueden evaluarse en un ensayo de exploración incluyen la capacidad de catalizar la hidrólisis de fosfatidilinositol, la capacidad de asociarse con ras, preferiblemente ras activado (unido a GTP) y la capacidad de mediar la actividad de intercambio de nucleótidos de guanina.

En una realización, puede ensayarse la actividad de asociación con ras de una proteína 16836 como se indica a continuación. En primer lugar, se prepara un polipéptido que contiene un fragmento de unión a ras de 16836, por ejemplo, aproximadamente los aminoácidos 1640-1745 de la SEC ID N°: 2. Por ejemplo, el fragmento de unión a ras de 16836 puede fusionarse a una secuencia de aminoácidos que permite la purificación del polipéptido y/o la detección de la unión a ras. En segundo lugar, se evalúan los compuestos de ensayo o candidatos con respecto a su capacidad de modular la unión a ras por el polipéptido. En un ejemplo, se usa una exploración doble híbrido en levadura para detectar la interacción entre una proteína 16836 y ras, así como una modulación de esta interacción en presencia de compuestos candidatos o de ensayo. Véase, por ejemplo, Shibatohe *et al.* (1998) *J. Biol. Chem.* 273: 6218 como ejemplo de dicho ensayo.

Los compuestos de ensayo de la presente invención pueden obtenerse usando cualquiera de las numerosas estrategias en los métodos de bibliotecas combinatorias conocidos en la técnica, incluyendo: bibliotecas biológicas; bibliotecas de peptoides (bibliotecas de moléculas que tienen las funcionalidades de péptidos, pero con un nuevo esqueleto no peptídico que son resistentes a la degradación enzimática pero que siguen siendo bioactivas; véase, por ejemplo, Zuckermann, R. N. *et al.* (1994) *J. Med Chem.* 37: 2678-85); bibliotecas en fase de solución o en fase sólida paralelas dirigibles espacialmente; métodos de bibliotecas sintéticas que requieren desconvolución; el método de biblioteca “una perla-un compuesto”; y los métodos de bibliotecas sintéticas que usan la selección de cromatografía de afinidad. Las estrategias de biblioteca biológica y biblioteca de peptoides se limitan a bibliotecas de péptidos, mientras que las otras cuatro estrategias son aplicables a bibliotecas de péptidos, oligómeros no peptídicos o compuestos de molécula pequeña (Lam (1997) *Anticancer Drug Des.* 12:145).

ES 2 328 448 T3

Pueden encontrarse ejemplos de métodos para la síntesis de bibliotecas moleculares en la técnica, por ejemplo, en: DeWitt *et al.* (1993) Proc. Natl. Acad. Sci. USA 90: 6909; Erb *et al.* (1994) Proc. Natl. Acad. Sci. USA 91: 11422; Zuckermann *et al.* 1994. J. Med. Chem. 37: 2678; Cho *et al.* (1993) Science 261: 1303; Carrell *et al.* (1994) Angew. Chem. Int. Ed. Engl. 33: 2059; Carell *et al.* (1994) Angew. Chem. Int. Ed. Engl. 33: 2061; y Gallop *et al.* 1994 J. Med. Chem. 37: 1233.

Pueden presentarse bibliotecas de compuestos en solución (por ejemplo Houghten (1992) Biotechniques 13: 412-421) o en perlas (Lam (1991) Nature 354: 82-84), chips (Fodor (1993) Nature 364: 555-556), bacteria (Ladner, Patente de Estados Unidos N° 5.223.409), esporas (Ladner Patente de Estados Unidos N° 5.223.409), plásmidos (Cull *et al.* (1992) Proc Natl Acad Sci USA 89: 1865-1869) o en fagos (Scott y Smith (1990) Science 249: 386-390; Devlin (1990) Science 249: 404-406; Cwirla *et al.* (1990) Proc. Natl. Acad. Sci. 87: 6378-6382; Felici (1991) J. Mol. Biol. 222: 301-310; Ladner *supra.*).

En una realización, el ensayo es un ensayo basado en células en el que una célula que expresa una proteína 16836 o una parte biológicamente activa de la misma se pone en contacto con un compuesto de ensayo y se determina la capacidad del compuesto de modular la actividad de 16836. La determinación de la capacidad del compuesto de ensayo de modular la actividad de 16836 puede realizarse controlando, por ejemplo, la hidrólisis de fosfatidilinositol o la asociación de Ras. La célula, por ejemplo, puede ser de mamífero, por ejemplo, humana.

También puede evaluarse la capacidad del compuesto de ensayo de modular la unión de 16836 a un compuesto, por ejemplo, un sustrato de 16836, o de unirse a 16836. Esto puede conseguirse, por ejemplo, por acoplamiento del compuesto, por ejemplo el sustrato, con un marcador radioisotópico o enzimático de tal forma que la unión del compuesto, por ejemplo, el sustrato a 16836 pueda determinarse por medio de la detección del compuesto marcado, por ejemplo, el sustrato en un complejo. Como alternativa, 16836 podría acoplarse con un marcador radioisotópico o enzimático para controlar la capacidad de un compuesto de ensayo de modular la unión de 16836 a un sustrato de 16836 en un complejo. Por ejemplo, los compuestos (por ejemplo, sustratos de 16836) pueden marcarse con ¹²⁵I, ³⁵S, ¹⁴C, o ³H, directamente o indirectamente, y el radioisótopo puede detectarse por contacto directo de radioemisión por recuento de centelleo. Como alternativa, los compuestos pueden marcarse enzimáticamente, por ejemplo, con peroxidasa de rábano picante, fosfatasa alcalina, o luciferasa, y el marcador enzimático puede detectarse determinando la conversión de un sustrato apropiado en producto.

Puede evaluarse la capacidad de un compuesto (por ejemplo sustrato de 16836) de interactuar con 16836 con o sin el marcaje de cualquiera de los agentes que interactúan. Por ejemplo, puede usarse un microfisiómetro para detectar la interacción de un compuesto con 16836 sin el marcaje del compuesto o el 16836. McConnell, H. M. *et al.* (1992) Science 257: 1906-1912. Como se usa en la presente memoria, un "microfisiómetro" (por ejemplo, Cytosensor) es un instrumento analítico que mide la velocidad a la que una célula acidifica su entorno usando un sensor potenciométrico dirigible por luz (LAPS). Los cambios en esta velocidad de acidificación pueden usarse como indicador de la interacción entre un compuesto y 16836.

En otra realización, se proporciona un ensayo sin células en el que una proteína 16836 o una parte biológicamente activa de la misma se pone en contacto con un compuesto de ensayo y se evalúa la capacidad del compuesto de ensayo de unirse a la proteína 16836 o a la parte biológicamente activa de la misma. Las partes biológicamente activas preferidas de las proteínas 16836 a usar en los ensayos de la presente invención incluyen fragmentos que participan en interacciones con moléculas que no son 16836, por ejemplo, fragmentos con altas puntuaciones de probabilidad de localización en superficie.

En los ensayos sin células de la invención pueden usarse formas solubles y/o unidas a la membrana de proteínas aisladas (por ejemplo, proteínas 16836 o partes biológicamente activas de las mismas). Cuando se usan formas unidas en la membrana de la proteína, puede ser deseable utilizar un agente de solubilización. Los ejemplos de estos agentes de solubilización incluyen detergentes no iónicos tales como n-octilglucósido, n-dodecilglucósido, n-dodecilmaltósido, octanoil-N-metilglucamida, decanoil-N-metilglucamida, Triton[®] X-100, Triton[®] X-114, Thesit[®], isotridecilo (etilenglicoléter)_n, 3-[(3-colamidopropil)dimetilamonio]-1-propano sulfonato (CHAPS), 3-[(3-colamidopropil)dimetilamonio]-2-hidroxi-1-propano sulfonato (CHAPSO) o N-dodecil-N,N-dimetil-3-amonio-1-propano sulfonato.

Los ensayos sin células implican preparar una mezcla de reacción de la proteína del gen diana y el compuesto de ensayo en condiciones y durante un periodo de tiempo suficiente para permitir que los dos componentes interactúen y se unan, formando de esta manera un complejo que puede retirarse y/o detectarse.

También puede detectarse la interacción entre dos moléculas, por ejemplo, usando transferencia de energía de fluorescencia (FET) (véase, por ejemplo, Lakowicz *et al.*, Patente de Estados Unidos N° 5.631.169; Stavrianopoulos, *et al.*, Patente de Estados Unidos N° 4.868.103). Se selecciona un marcador fluoróforo en la primera molécula "donadora" de tal forma que su energía fluorescente emitida se absorba por un marcador fluorescente en una segunda molécula "aceptora", que a su vez puede emitir fluorescencia debido a la energía absorbida. Como alternativa, la molécula de proteína "donadora" simplemente puede utilizar la energía fluorescente natural de los restos de triptófano. Se eligen marcadores que emiten diferentes longitudes de onda de luz, de tal forma que el marcador de la molécula "aceptora" puede diferenciarse de la de la "donadora". Como la eficacia de transferencia de energía entre los marcadores está relacionada con la distancia que separa las moléculas, puede evaluarse la relación espacial entre las moléculas. En una situación en la que se produce la unión entre las moléculas, la emisión fluorescente del marcador de la

molécula “aceptora” en el ensayo debe ser máxima. Convenientemente puede medirse un acontecimiento de unión FET por medios de detección fluorométrica convencionales bien conocidos en la técnica (por ejemplo, usando un fluorímetro).

5 En otra realización, la determinación de la capacidad de la proteína 16836 de unirse a una molécula diana puede realizarse usando el Análisis de Interacción Biomolecular (BIA) a tiempo real (véase, por ejemplo, Sjolander, S. y Urbaniczky, C. (1991) *Anal. Chem.* 63: 2338-2345 y Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5: 699-705). La “resonancia de plasmón superficial” o “BIA” detecta interacciones bioespecíficas a tiempo real, sin marcar ninguna de las moléculas que interactúan (por ejemplo, BIAcore). Los cambios en la masa en la superficie de unión (que
10 indica un acontecimiento de unión) producen alteraciones del índice de refracción de luz cerca de la superficie (el fenómeno óptico de resonancia de plasmón superficial (SPR)), dando como resultado una señal detectable que puede usarse como indicación de reacciones a tiempo real entre moléculas biológicas.

15 En una realización, el producto del gen diana o la sustancia de ensayo se ancla en una fase sólida. Al final de la reacción pueden detectarse los complejos de producto génico diana/compuesto ensayo anclados en la fase sólida. Preferiblemente, el producto génico diana puede anclarse en una superficie sólida y el compuesto de ensayo (que no se ancla) puede marcarse, directamente o indirectamente, con marcadores detectables descritos en la presente memoria.

20 Puede ser deseable inmovilizar 16836, o un anticuerpo anti-16836 o su molécula diana para facilitar la separación entre las formas complejadas y las formas no complejadas de una o las dos proteínas, así como para acomodar la automatización del ensayo. La unión de un compuesto de ensayo a una proteína 16836, o la interacción de una proteína 16836 con una molécula diana en presencia y ausencia de un compuesto candidato, pueden realizarse en cualquier recipiente adecuado para contener los reactivos. Los ejemplos de estos recipientes incluyen placas de microtitulación,
25 tubos de ensayo y tubos de microcentrífuga. En una realización, puede proporcionarse una proteína de fusión que añade un dominio que permite a una o a las dos proteínas unirse a una matriz. Por ejemplo, pueden adsorberse proteínas de fusión glutatión-S-transferasa/16836 o proteínas de fusión de glutatión-S-transferasa/diana en perlas de glutatión sepharose (Sigma Chemical, St. Louis, MO) o placas de microtitulación modificadas con glutatión, que después se combinan con el compuesto de ensayo o el compuesto de ensayo y la proteína diana no adsorbida o proteína 16836,
30 y la mezcla puede incubarse en condiciones que conducen a la formación de complejos (por ejemplo, en condiciones fisiológicas para sales y pH). Después de la incubación, se lavan las perlas o los pocillos de la placa de microtitulación para retirar cualquier componente que no se haya unido, se inmoviliza la matriz en el caso de las perlas y se determinan los complejos directa o indirectamente, por ejemplo, como se descrito anteriormente. Como alternativa, los complejos pueden disociarse de la matriz y puede determinarse el nivel de unión o actividad de 16836 usando técnicas
35 convencionales.

Otras técnicas para inmovilizar una proteína 16836 o una molécula diana en matrices incluyen el uso de conjugación de biotina y estreptavidina. Puede prepararse moléculas diana o proteínas 16836 biotiniladas a partir de biotina-NHS (N-hidroxi-succinimida) usando técnicas conocidas en este campo (por ejemplo, un kit de biotinilación, Pierce
40 Chemicals, Rockford, IL), e inmovilizarse en los pocillos de placas de 96 pocillos recubiertas con estreptavidina (Pierce Chemical).

Para realizar el ensayo, el componente no inmovilizado se añade a la superficie recubierta que contiene el componente anclado. Después de completarse la reacción, se retiran los componentes que no han reaccionado (por ejemplo
45 por lavado) en condiciones tales que los complejos formados permanezcan inmovilizados en la superficie sólida. La detección de complejos anclados en la superficie sólida puede realizarse de varias maneras. Cuando el componente no inmovilizado previamente está premarcado, la detección del marcador inmovilizado en la superficie indica que se formaron complejos. Cuando el componente no inmovilizado previamente no está premarcado, puede usarse un marcador indirecto para detectar complejos anclados en la superficie; por ejemplo usando un anticuerpo marcado específico para el componente inmovilizado (el anticuerpo, a su vez, puede marcarse directamente o marcarse indirectamente, por
50 ejemplo, con un anticuerpo anti-Ig marcado).

En una realización, este ensayo se realiza utilizando anticuerpos reactivos con la proteína 16836 o moléculas diana que no interfieren con la unión de la proteína 16836 a su molécula diana. Estos anticuerpos pueden modificarse en los
55 pocillos de la placa y la diana no unida o la proteína 16836 puede quedar atrapada en los pocillos por conjugación con el anticuerpo. Los métodos para detectar estos complejos, además de los descritos anteriormente para los complejos inmovilizados con GST, incluyen inmunodetección de complejos usando anticuerpos reactivos con la proteína 16836 o molécula diana, así como ensayos ligados a enzimas que se basan en la detección de una actividad enzimática asociada con la proteína 16836 o molécula diana.

60 Como alternativa, los ensayos sin células se pueden realizar en una fase líquida. En dicho ensayo, los productos de reacción se separan de los componentes que no han reaccionado, por cualquiera de varias técnicas convencionales, incluyendo pero sin limitación: centrifugación diferencial (véase por ejemplo, Rivas, G., y Minton, A. P., (1993) *Trends Biochem Sci* 18: 284-7); cromatografía (cromatografía de exclusión molecular, cromatografía de intercambio iónico); electroforesis (véase por ejemplo, Ausubel, F. *et al.*, eds. *Current Protocols in Molecular Biology* 1999, J. Wiley: Nueva York); e inmunoprecipitación (véase por ejemplo, Ausubel, F. *et al.*, eds. (1999) *Current Protocols in Molecular Biology*, J. Wiley: Nueva York). Estas resinas y técnicas cromatográficas se conocen por los especialistas en la técnica (véase por ejemplo, Heegaard, N. H., (1998) *J Mol Recognit* 11: 141-8; Hage, D. S., y Tweed, S. A. (1997)

J Chromatogr B Biomed Sci Appl. 699: 499-525). Además, también puede utilizarse convenientemente transferencia de energía de fluorescencia, como se describe en este documento, para detectar la unión sin purificación adicional del complejo de la solución.

5 En una realización preferida, el ensayo incluye poner en contacto la proteína 16836 o una parte biológicamente activa de la misma con un compuesto conocido que se une a 16836 para formar una mezcla de ensayo, poner en contacto la mezcla de ensayo con un compuesto de ensayo y determinar la capacidad del compuesto de ensayo de interactuar con una proteína 16836, donde la determinación de la capacidad del compuesto de ensayo de interactuar con una proteína 16836 incluye la determinación de la capacidad del compuesto de ensayo de unirse preferentemente a 16836
10 o a una parte biológicamente activa de la misma o modular la actividad de una molécula diana en comparación con el compuesto conocido.

Los productos génicos diana de la invención pueden interactuar *in vivo* con una o más macromoléculas celulares o extracelulares tales como proteínas. Para los fines de este análisis, estas macromoléculas celulares y extracelulares se denominan en este documento “compañeros de unión.” Los compuestos que rompen dichas interacciones pueden ser útiles para regular la actividad del producto génico diana. Estos compuestos pueden incluir, pero sin limitación, moléculas tales como anticuerpos, péptidos y moléculas pequeñas. Los genes/productos diana preferidos para uso en esta realización son los genes de 16836 identificados en la presente memoria. En una realización alternativa, la invención proporciona métodos para determinar la capacidad del compuesto de ensayo de modular la actividad de una proteína 16836 por medio de la modulación de la actividad de un efecto aguas abajo de una molécula diana de 16836.
20 Por ejemplo, puede determinarse la actividad de la molécula efectora en una diana apropiada o puede determinarse la unión del efector a una diana apropiada, como se ha descrito previamente.

Para identificar compuestos que interfieren con la interacción entre el producto génico diana y su compañero de unión celular o extracelular, se prepara una mezcla de reacción que contiene el producto génico diana y el compañero de unión, en condiciones y durante un periodo de tiempo suficiente para permitir que los dos productos formen un complejo. Para ensayar un agente inhibidor, la mezcla de reacción se proporciona en presencia y ausencia del compuesto de ensayo. El compuesto de ensayo puede incluirse inicialmente en la mezcla de reacción o puede añadirse en un periodo de tiempo posterior a la adición del gen diana y su compañero de unión celular o extracelular. Las mezclas de reacción de control se incuban sin el compuesto de ensayo o con un placebo. Después se detecta la formación de cualquier complejo entre el producto génico diana y el compañero de unión celular o extracelular. La formación de un complejo en la reacción de control, pero no en la mezcla de reacción que contiene el compuesto de ensayo, indica que el compuesto interfiere con la interacción del producto génico diana y el compañero de unión interactivo. Además, la formación de complejos dentro de mezclas de reacción que contienen el compuesto de ensayo y el producto génico diana normal también puede compararse con la formación de complejos dentro de mezclas de reacción que contienen el compuesto de ensayo y el producto génico diana mutante. Esta comparación puede ser importante en los casos en los que deseable identificar compuestos que alteran interacciones de mutantes pero no productos génicos diana normales.
35

Estos ensayos pueden realizarse en un formato heterogéneo u homogéneo. Los ensayos heterogéneos implican el anclaje del producto génico diana o el compañero de unión en una fase sólida y la detección de complejos anclados en la fase sólida al final de la reacción. En los ensayos homogéneos, la reacción entera se realiza en una fase líquida. En cualquier estrategia, el orden de adición de los reactivos puede variarse para obtener diferente información sobre los compuestos que se están ensayando. Por ejemplo, pueden identificarse compuestos de ensayo que interfieren con la interacción entre los productos génicos diana y los compañeros de unión, por ejemplo, por competición realizando la reacción en presencia de la sustancia de ensayo. Como alternativa, pueden ensayarse compuestos de ensayo que rompen los complejos preformados, por ejemplo, compuestos con mayores constantes de unión que desplazan uno de los componentes del complejo, añadiendo el compuesto de ensayo a la mezcla de reacción después de que se hayan formado los complejos. Los diversos formatos se describen brevemente más adelante.
50

En un sistema de ensayo heterogéneo, el producto génico diana o el compañero de unión celular o extracelular interactivo, se ancla en una superficie sólida (por ejemplo, una placa de microtitulación), mientras que la especie no anclada se marca directa o indirectamente. Las especies ancladas pueden inmovilizarse por uniones covalentes o no covalentes. Como alternativa, puede usarse un anticuerpo inmovilizado específico para la especie a anclar para anclar la especie a la superficie sólida.
55

Para realizar el ensayo, el compañero de la especie inmovilizada se expone a la superficie recubierta con o sin el compuesto de ensayo. Después de completarse la reacción, se retiran los componentes que no han reaccionado (por ejemplo por lavado) y cualquier complejo formado permanecerá inmovilizado en la superficie sólida. Cuando la especie no inmovilizada está pre-mercada, la detección del marcador inmovilizado en la superficie indica que se formaron complejos. Cuando la especie no inmovilizada no está pre-mercada, puede usarse un marcador indirecto para detectar complejos anclados en la superficie; por ejemplo usando un anticuerpo marcado específico para la especie inicialmente no inmovilizada (el anticuerpo, a su vez, puede marcarse directamente o marcarse indirectamente con, por ejemplo, un anticuerpo anti-Ig marcado). Dependiendo del orden de adición de los componentes de reacción, pueden detectarse compuestos de ensayo que inhiben la formación de complejos o que rompen los complejos preformados.
65

Como alternativa, la reacción puede realizarse en una fase líquida en presencia o ausencia del compuesto de ensayo, los productos de reacción pueden separarse de los componentes que no han reaccionado y los complejos pueden detectarse; por ejemplo, usando un anticuerpo inmovilizado específico para uno de los componentes de unión para anclar cualquier complejo formado en solución, y un anticuerpo marcado específico para el otro compañero para detectar complejos anclados. De nuevo, dependiendo del orden de adición de los reactivos a la fase líquida, pueden identificarse compuestos de ensayo que inhiben complejos o que rompen complejos preformados.

En una realización alternativa de la invención, puede usarse un ensayo homogéneo. Por ejemplo, se prepara un complejo preformado del producto génico diana y el producto del compañero de unión celular o extracelular interactivo en el que los productos génicos diana o sus compañeros de unión están marcados, pero la señal generada por el marcador se inactiva debido a la formación de complejos (véase, por ejemplo, la Patente de Estados Unidos N° 4.109.496 que utiliza esta estrategia para inmunoensayos). La adición de una sustancia de ensayo que compite y desplaza una de las especies del complejo preformado dará como resultado la generación de una señal por encima del nivel de fondo. De esta manera, pueden identificarse sustancias de ensayo que alteran la interacción de producto génico diana-compañero de unión.

En otro aspecto, las proteínas 16836 pueden usarse como “proteínas cebo” en un ensayo doble híbrido o en un ensayo triple híbrido (véase, por ejemplo, la Patente de los Estados Unidos Núm. 5.283.317; Zervos *et al.* (1993) *Cell* 72: 223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268: 12046-12054. Bartel *et al.* (1993) *BioTechniques* 14: 920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1 693-1696; y Brent documento WO 94/10300), para identificar otras proteínas que se unen o interaccionan con 16836 (“proteínas de unión a 16836” o “16836-bp”) y están implicadas en la actividad de 16836. Estas 16836-bp pueden ser activadores o inhibidores de señales por las proteínas 16836 o dianas de 16836 tales como, por ejemplo, elementos aguas abajo de una ruta de señalización mediada por 16836.

El sistema doble híbrido se basa en la naturaleza modular de la mayoría de los factores de transcripción, que consisten en dominios de unión y activación de ADN separables. En resumen, el ensayo utiliza dos construcciones de ADN diferentes. En una construcción, el gen que codifica una proteína 16836 se fusiona a un gen que codifica el dominio de unión al ADN de un factor de transcripción conocido (por ejemplo, GAL-4). En la otra construcción, una secuencia de ADN de una biblioteca de secuencias de ADN, que codifica una proteína no identificada (“presa” o “muestra”) se fusiona a un gen que codifica el dominio de activación del factor de transcripción conocido. (Como alternativa: la proteína 16836 puede fusionarse al dominio activador.) Si las proteínas “cebo” y “presa” pueden interactuar, *in vivo*, formando un complejo dependiente de 16836, los dominios de unión al ADN y de activación del factor de transcripción se acercan mucho. Esta proximidad permite la transcripción de un gen indicador (por ejemplo, LacZ) que se une operablemente a un sitio regulador de la transcripción que responde al factor de transcripción. La expresión del gen indicador puede detectarse y pueden aislarse colonias de células que contienen el factor de transcripción funcional y usarse para obtener el gen clonado que codifica la proteína que interacciona con la proteína 16836.

En otra realización, se identifican moduladores de la expresión de 16836. Por ejemplo, una célula o una mezcla sin células se pone en contacto con un compuesto candidato y se evalúa la expresión del ARNm o la proteína 16836 con respecto a nivel de expresión del ARNm o proteína 16836 en ausencia del compuesto candidato. Cuando la expresión del ARNm o proteína 16836 es mayor en presencia del compuesto candidato que en su ausencia, el compuesto candidato se identifica como un estimulador de la expresión de la proteína o ARNm de 16836. Como alternativa, cuando la expresión del ARNm o la proteína 16836 es menor (menor de una forma estadísticamente significativa) en presencia del compuesto candidato que en su ausencia, se identifica el compuesto candidato como un inhibidor de la expresión del ARNm o la proteína 16836. El nivel de expresión del ARNm o la proteína 16836 puede determinarse por métodos descritos en la presente memoria para detectar el ARNm o la proteína 16836.

En otro aspecto, la invención se refiere a una combinación de dos o más de los ensayos descritos en la presente memoria. Por ejemplo, puede identificarse un agente modulador usando un ensayo basado en células o sin células, y puede confirmarse la capacidad del agente de modular la actividad de una proteína 16836 *in vivo*, por ejemplo en un animal.

Esta invención se refiere además a nuevos agentes identificados por los ensayos de exploración descritos anteriormente. Por consiguiente, está dentro del alcance de esta invención usar adicionalmente un agente identificado como se describe en la presente memoria (por ejemplo, un agente modulador de 16836, una molécula de ácido nucleico de 16836 antisentido, un anticuerpo específico para 16836 o un compañero de unión de 16836) en un modelo animal apropiado para determinar la eficacia, toxicidad, efectos secundarios o mecanismo de acción de tratamiento con dicho agente. Además, pueden usarse nuevos agentes identificados por los ensayos de exploración descritos anteriormente para los tratamientos como se describe en la presente memoria.

Ensayos de detección

Pueden usarse partes o fragmentos de las secuencias de ácido nucleico identificadas en la presente memoria como reactivos polinucleotídicos. Por ejemplo, estas secuencias pueden usarse para: (i) mapear sus genes respectivos en un cromosoma, por ejemplo, para localizar regiones génicas asociadas con enfermedades genéticas o para asociar 16836 con una enfermedad; (ii) identificar a un individuo a partir de una muestra biológica muy pequeña (tipificación de tejidos); y (iii) ayudar en la identificación forense de una muestra biológica.

ES 2 328 448 T3

Medicina predictiva

La presente invención también se refiere al campo de la medicina predictiva en la que se usan ensayos de diagnóstico, ensayos de pronóstico y ensayos clínicos de monitorización con fines de pronóstico (predictivos) para tratar de esta manera a un individuo.

En general, la invención proporciona un método para determinar si un sujeto tiene riesgo de padecer un trastorno relacionado con una lesión o la expresión defectuosa de un gen que codifica 16836.

Estos trastornos incluyen, por ejemplo, un trastorno asociado con la expresión defectuosa de un gen de 16836; un trastorno del sistema de señalización de ras; o un trastorno asociado con la hidrólisis anómala de fosfatidilinositol.

El método incluye uno o más de los siguientes:

detectar en un tejido del sujeto, la presencia o ausencia de una mutación que afecta la expresión del gen de 16836, o detectar la presencia o ausencia de una mutación en una región que controla la expresión del gen, por ejemplo, una mutación en la región de control 5';

detectar, en un tejido del sujeto, la presencia o ausencia de una mutación que altera la estructura del gen de 16836;

detectar, en un tejido del sujeto, la expresión defectuosa del gen de 16836, a nivel de ARNm, por ejemplo, detectar un nivel de tipo no silvestre de un ARNm;

detectar, en un tejido del sujeto, la expresión defectuosa del gen, a nivel de la proteína, por ejemplo, detectando el nivel de tipo no silvestre de un polipéptido 16836;

En realizaciones preferidas, el método incluye: averiguar la existencia de al menos uno de: una delección de uno o más nucleótidos del gen 16836; una inserción de uno o más nucleótidos en el gen, una mutación puntual, por ejemplo, una sustitución de uno o más nucleótidos del gen, una reordenación cromosómica general del gen, por ejemplo, una translocación, inversión o delección.

Por ejemplo, la detección de la lesión genética puede incluir: (i) proporcionar una sonda/cebador que incluye un oligonucleótido que contiene una región de secuencia de nucleótidos que hibrida con una secuencia con sentido o antisentido de la SEC ID N°: 1, o mutantes naturales de la misma o secuencias flanqueante 5' o 3' asociadas de forma natural con el gen de 16836; (ii) exponer la sonda/cebador al ácido nucleico del tejido; y detectar, por hibridación, por ejemplo, hibridación *in situ*, de la sonda/cebador al ácido nucleico la presencia o ausencia de la lesión genética.

En realizaciones preferidas que detectan la expresión defectuosa se incluye la determinación de la existencia de al menos uno de: una alteración en el nivel de transcrito de ARN mensajero del gen de 16836; la presencia de un patrón de corte y empalme de tipo no silvestre de un transcrito de ARN mensajero del gen; o un nivel de tipo no silvestre de 16836.

Los métodos de la invención pueden usarse antes del nacimiento o para determinar si la descendencia de un sujeto tiene riesgo de un trastorno.

En realizaciones preferidas el método incluye determinar la estructura de un gen de 16836, siendo indicativa una estructura anómala del riesgo de padecer el trastorno.

En realizaciones preferidas el método incluye poner en contacto una muestra del sujeto con un anticuerpo contra la proteína 16836 o un ácido nucleico que hibrida específicamente con el gen. Estas y otras realizaciones se describen más adelante.

Ensayos de Diagnóstico y Pronóstico

Los ensayos de diagnóstico y pronóstico de la invención incluyen un método para evaluar el nivel de expresión de moléculas de 16836 y para identificar variaciones y mutaciones en la secuencia de las moléculas de 16836.

Control y Perfil de Expresión

La presencia, nivel o ausencia de proteína o ácido nucleico de 16836 en una muestra biológica puede evaluarse obteniendo una muestra biológica de un sujeto de ensayo y poniendo en contacto la muestra biológica con un compuesto o un agente capaz de detectar la proteína o el ácido nucleico de 16836 (por ejemplo, ARNm, ADN genómico) que codifica la proteína 16836 de tal manera que se detecte la presencia la de proteína o ácido nucleico de 16836 en la muestra biológica. La expresión "muestra biológica" incluye tejidos, células y fluidos biológicos aislados a partir de un sujeto, así como tejidos, células y fluidos presentes dentro de un sujeto. Una muestra biológica preferida es suero. El nivel de expresión del gen de 16836 puede medirse de varias formas, incluyendo pero sin limitación: medición del ARNm codificado por los genes de 16836; medición de la cantidad de proteína codificada por los genes de 16836; o medición de la actividad de la proteína codificada por los genes de 16836.

ES 2 328 448 T3

El nivel de ARNm correspondiente al gen de 16836 en una célula puede determinarse tanto *in situ* como por medio de formatos *in vitro*.

5 El ARNm aislado puede usarse en ensayos de hibridación o amplificación que incluyen, pero sin limitación, análisis de Southern o Northern, análisis de reacción en cadena de la polimerasa y matrices de sondas. Un método de diagnóstico preferido para la detección de los niveles de ARNm implica poner en contacto el ARNm aislado con una molécula de ácido nucleico (sonda) que puede hibridar con el ARNm codificado por el gen que se está detectando. La sonda de ácido nucleico puede ser, por ejemplo, un ácido nucleico de 16836 de longitud completa, tal como el ácido nucleico de la SEC ID N° 1, o una parte del mismo, tal como un oligonucleótido de al menos 7, 15, 30, 50, 100, 250 ó
10 500 nucleótidos de longitud y suficiente para hibridar específicamente en condiciones rigurosas con el ARNm o ADN genómico de 16836. La sonda puede disponerse en una dirección de una matriz, por ejemplo una matriz descrita más adelante. En la presente memoria se describen otras sondas adecuadas para uso en los ensayos de diagnóstico.

15 En un formato, se inmoviliza ARNm (o ADNc) en una superficie y se pone en contacto con las sondas, por ejemplo procesando el ARNm aislado en un gel de agarosa y transfiriendo el ARNm del gel a una membrana, tal como una membrana de nitrocelulosa. En un formato alternativo, las sondas se inmovilizan en una superficie y el ARNm (o ADNc) se pone en contacto con las sondas, por ejemplo, en una matriz de chips de genes bidimensional descrita más adelante. Un especialista en la técnica puede adaptar métodos de detección de ARNm conocidos para uso en la detección del nivel de ARNm codificado por los genes de 16836.
20

El nivel de ARNm en una muestra que se codifica por uno de 16836 puede evaluarse con amplificación de ácidos nucleicos, por ejemplo, por rtPCR (Mullis (1987) Patente de Estados Unidos N° 4.683.202), reacción en cadena de la ligasa (Barany (1991) Proc. Natl. Acad. Sci. USA 88: 189-193), replicación de secuencia autosostenida (Guatelli *et al.*, (1990) Proc. Natl. Acad. Sci. USA 87: 1874-1878), sistema de amplificación transcripcional (Kwoh *et al.*, (1989), Proc. Natl. Acad. Sci. USA 86: 1173-1177), Q-Beta Replicasa (Lizardi *et al.*, (1988) Bio/Technology 6: 1197), replicación por círculo rodante (Lizardi *et al.*, Patente de Estados Unidos N° 5.854.033) o cualquier otro método de amplificación de ácidos nucleicos, seguido de la detección de las moléculas amplificadas usando técnicas conocidas en este campo. Como se usa en la presente memoria, los cebadores de amplificación se definen como un par de moléculas de ácido nucleico que pueden hibridar con regiones 5' o 3' de un gen (cadenas más y menos, respectivamente o vice-versa) y contener una región corta entre ellos. En general, los cebadores de amplificación tienen una longitud de aproximadamente 10 a 30 nucleótidos y flanquean una región con una longitud de aproximadamente 50 a 200 nucleótidos. En condiciones apropiadas y con reactivos apropiados, estos cebadores permiten la amplificación de una molécula de ácido nucleico que comprende la secuencia de nucleótidos flanqueada por los cebadores.
30

35 Para los métodos *in situ*, puede prepararse/procesarse una célula o muestra de tejido e inmovilizarse en un soporte, típicamente un portaobjetos de vidrio, y después ponerse en contacto con una sonda que puede hibridar con el ARNm que codifica el gen de 16836 que se está analizando.

40 En otra realización, los métodos además ponen en contacto una muestra de control con un compuesto o agente capaz de detectar ARNm o ADN genómico de 16836 y comparan la presencia del ARNm o ADN genómico de 16836 en la muestra de control con la presencia de ARNm o ADN genómico de 16836 en la muestra de ensayo. En otra realización, se usa análisis en serie de expresión génica, como se describe en la Patente de Estados Unidos N° 5.695.937, para detectar niveles de transcrito de 16836.

45 Pueden usarse diversos métodos para determinar el nivel de proteína codificada por 16836. En general, estos métodos incluyen poner en contacto un agente que se une selectivamente a la proteína, tal como un anticuerpo con una muestra, para evaluar el nivel de proteína en la muestra. En una realización preferida, el anticuerpo lleva un marcador detectable. Los anticuerpos pueden ser policlonales, o más preferiblemente, monoclonales. Puede usarse un anticuerpo intacto, o un fragmento del mismo (por ejemplo Fab o F(ab')₂). El término "marcado", con respecto a la sonda o anticuerpo, pretende incluir el marcaje directo de la sonda o anticuerpo por acoplamiento (es decir, unión física) de una sustancia detectable a la sonda o anticuerpo, así como el marcaje indirecto de la sonda o anticuerpo por reactividad con una sustancia detectable. En la presente memoria se proporcionan ejemplos de sustancias detectables.
50

55 Los métodos de detección pueden usarse para detectar la proteína 16836 en una muestra biológica *in vitro* así como *in vivo*. Las técnicas *in vitro* para la detección de la proteína 16836 incluyen ensayos de inmunoabsorbente ligado a enzima (ELISA), inmunoprecipitaciones, inmunofluorescencia, inmunoensayo enzimático (EIA), radioinmunoensayo (RIA) y análisis de transferencia de Western. Las técnicas *in vivo* para la detección de la proteína 16836 incluyen la introducción en un sujeto de un anticuerpo anti-16836 marcado. Por ejemplo, el anticuerpo puede marcarse con un marcador radiactivo cuya presencia y localización en un sujeto puede detectarse por técnicas convencionales de formación de imágenes. En otra realización, la muestra se marca, por ejemplo, se biotinila y después se pone en contacto con el anticuerpo, por ejemplo, un anticuerpo anti-16836 colocado en una matriz de anticuerpo (como se describe más adelante). La muestra puede detectarse, por ejemplo, con avidina acoplada a un marcador fluorescente.
60

65 En otra realización, los métodos incluyen además poner en contacto la muestra de control con un compuesto o agente capaz de detectar la proteína 16836, y comparar la presencia de la proteína 16836 en la muestra de control con la presencia de la proteína 16836 en la muestra de ensayo.

La invención también incluyen kits para detectar la presencia de 16836 en una muestra biológica. Por ejemplo, el kit puede incluir un compuesto o agente capaz de detectar una proteína o ARNm de 16836 en una muestra biológica; y un patrón. El compuesto o agente puede estar empaquetado en un recipiente adecuado. El kit puede comprender además instrucciones para usar el kit para detectar la proteína o ácido nucleico de 16836.

Para kits basados en anticuerpo, el kit puede incluir: (1) un primer anticuerpo (por ejemplo unido a un soporte sólido) que se une a un polipéptido correspondiente a un marcador de la invención; y, opcionalmente (2) un segundo anticuerpo diferente que se une al polipéptido o al primer anticuerpo y se conjuga con un agente detectable.

Para kits basados en oligonucleótidos, el kit puede incluir: (1) un oligonucleótido, por ejemplo, un oligonucleótido marcado de forma detectable que hibrida con una secuencia de ácido nucleico que codifica un polipéptido correspondiente a un marcador de la invención o (2) un par de cebadores útiles para amplificar una molécula de ácido nucleico que corresponde a un marcador de la invención. El kit también puede incluir un agente tamponante, un conservante o un agente estabilizador de proteínas. El kit también puede incluir componentes necesarios para detectar el agente detectable (por ejemplo una enzima o un sustrato). El kit también puede contener una muestra de control o una serie de muestras de control que pueden ensayarse y compararse con la muestra de ensayo contenida. Cada componente del kit puede encerrarse dentro de un recipiente individual y todos los diversos recipientes pueden estar dentro de un solo envase, junto con instrucciones para interpretar los resultados de los ensayos realizados usando el kit.

Los métodos de diagnóstico descritos en la presente memoria pueden identificar sujetos que tienen o con riesgo de desarrollar una enfermedad o trastorno asociado con la expresión o actividad defectuosa, aberrante o indeseada de 16836. Como se usa en la presente memoria, el término “indeseada” incluye un fenómeno indeseado implicado en una respuesta biológica tal como cáncer o una proliferación celular no regulada.

En una realización, se identifica una enfermedad o trastorno asociado con la expresión o actividad de 16836 aberrante o indeseada. Se obtiene una muestra de ensayo de un sujeto y se evalúa la proteína o ácido nucleico de 16836 (por ejemplo, ARNm o ADN genómico), donde el nivel, por ejemplo, la presencia o ausencia de proteína o ácido nucleico de 16836 es diagnóstico de un sujeto que tiene riesgo de desarrollar una enfermedad o trastorno asociado con la expresión o actividad aberrante o indeseada de 16836. Como se usa en la presente memoria una “muestra de ensayo” se refiere a una muestra biológica obtenida a partir de un sujeto de interés, incluyendo un fluido biológico (por ejemplo, suero), una muestra celular o un tejido.

Los ensayos de pronósticos descritos en la presente memoria pueden utilizarse para determinar si a un sujeto se le puede administrar un agente (por ejemplo un agonista, antagonista, peptidomimético, proteína, péptido, ácido nucleico, molécula pequeña, u otro candidato de fármaco) para tratar una enfermedad o trastorno asociado con la expresión o actividad aberrante o indeseada de 16836. Por ejemplo, estos métodos pueden usarse para determinar si un sujeto se puede tratar eficazmente con un agente para un trastorno relacionado con la proliferación o diferenciación celular.

En otro aspecto, la invención se refiere a un método informático que tiene una pluralidad de registros de datos codificados digitalmente. Cada registro de datos incluye un valor que representa el nivel de expresión de 16836 en una muestra, y un descriptor de la muestra. El descriptor de la muestra puede ser un identificador de la muestra, un sujeto a partir del cual se obtuvo la muestra (por ejemplo un paciente), un diagnóstico o un tratamiento (por ejemplo un tratamiento preferido). En una realización preferida, el registro de datos incluye además valores que representan el nivel de expresión de genes distintos de 16836 (por ejemplo, otros genes asociados con un trastorno relacionado con 16836, u otros genes en una matriz). El registro de datos puede estructurarse como una tabla, por ejemplo una tabla que forma parte de una base de datos tal como una base de datos relacional (por ejemplo, una base de datos SQL de los entornos de bases de datos Oracle o Sybase).

También se refiere a un método para evaluar una muestra. El método incluye proporcionar una muestra, por ejemplo, a partir del sujeto y determinar un perfil de expresión génica de la muestra, donde el perfil incluye un valor que representa el nivel de expresión de 16836. El método puede incluir además comparar el valor o el perfil (es decir múltiples valores) con un valor de referencia o perfil de referencia. El perfil de expresión génica de la muestra puede obtenerse por cualquiera de los métodos descritos en la presente memoria (por ejemplo, proporcionando un ácido nucleico a partir de la muestra y poniendo en contacto el ácido nucleico con una matriz). El método puede usarse para diagnosticar un trastorno de proliferación y/o diferenciación celular en un sujeto donde un aumento en la expresión de 16836 es una indicación de que el sujeto tiene o es propenso a tener un trastorno de la proliferación y/o diferenciación celular. El método puede usarse para monitorizar un tratamiento para el trastorno de la proliferación y/o diferenciación celular en un sujeto. Por ejemplo, el perfil de expresión génica puede determinarse para muestra de un sujeto sometido a tratamiento. El perfil puede compararse con un perfil de referencia o con un perfil obtenido a partir del sujeto antes del tratamiento o antes de la aparición del trastorno (véase, por ejemplo, Golub *et al.* (1999) Science 286: 531).

En otro aspecto, la invención se refiere a un método para evaluar un compuesto de ensayo (véase también la sección de “Ensayos de Exploración”, anterior). El método incluye proporcionar una célula y un compuesto de ensayo; poner en contacto el compuesto de ensayo con la célula; obtener un perfil de expresión del sujeto para la célula con la que ha entrado en contacto; y comparar el perfil de expresión del sujeto con uno o más perfiles de referencia. Los perfiles

incluyen un valor que representa el nivel de expresión de 16836. En una realización preferida, el perfil de expresión del sujeto se compara con un perfil diana, por ejemplo, un perfil para una célula normal o para el estado deseado de una célula. El compuesto de ensayo se evalúa favorablemente si el perfil de expresión del sujeto es más similar al perfil diana que un perfil de expresión obtenido a partir de una célula con la que no ha contactado.

5 En otro aspecto, la invención se refiere a un método para evaluar a un sujeto. El método incluye: a) obtener una muestra a partir de un sujeto, por ejemplo, a partir de un cuidador sanitario, por ejemplo, un cuidador sanitario que obtiene la muestra del sujeto; b) determinar el perfil de expresión de un sujeto para la muestra. Opcionalmente, el método incluye además una o las dos etapas: c) comparar el perfil de expresión del sujeto con uno o más perfiles de expresión de referencia; y d) seleccionar el perfil de referencia más similar al perfil de referencia del sujeto. El perfil de expresión del sujeto y los perfiles de referencia incluyen un valor que representa el nivel de expresión de 16836. Puede usarse una diversidad de medidas estadísticas rutinarias para comparar dos perfiles de referencia. Una medida posible es la longitud del vector de distancia que es la diferencia entre los dos perfiles. Cada uno de los perfiles del sujeto y de referencia se representa como un vector multidimensional, donde cada dimensión es un valor en el perfil.

15 El método puede incluir además transmitir un resultado a un cuidador sanitario. El resultado puede ser el perfil de expresión del sujeto, un resultado de una comparación del perfil de expresión del sujeto con otro perfil, un perfil de referencia más similar o un descriptor de cualquiera de los mencionados anteriormente. El resultado puede transmitirse a través de una red informática, por ejemplo, el resultado puede estar en forma de una transmisión informática, por ejemplo una señal de datos informáticos incluidos en una onda portadora.

También se refiere a un medio informático que tiene un código ejecutable para realizar las siguientes etapas: recibir el perfil de expresión de un sujeto; acceder a una base de datos de perfiles de expresión de referencia; y i) seleccionar un perfil de referencia correspondiente que sea el más similar al perfil de expresión del sujeto o ii) determinar al menos una puntuación de comparación para la similitud del perfil de expresión del sujeto con al menos un perfil de referencia. El perfil de expresión del sujeto, y los perfiles de expresión de referencia incluyen un valor que representa el nivel de expresión de 16836.

30 *Composiciones farmacéuticas*

Los ácidos nucleicos y polipéptidos, fragmentos de los mismos, así como los anticuerpos anti-16836 (también denominados en este documento “compuestos activos”) de la invención pueden incorporarse en composiciones farmacéuticas. Estas composiciones típicamente incluyen la molécula de ácido nucleico, proteína o anticuerpo y un vehículo farmacéuticamente aceptable. Como se usa en la presente memoria, la expresión “vehículo farmacéuticamente aceptable” incluye disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes para retrasar la absorción e isotónicos y similares, compatibles con la administración farmacéutica. También pueden incorporarse en las composiciones compuestos activos complementarios.

40 Una composición farmacéutica se formula de manera que sea compatible con la vía de administración para la que está destinada. Los ejemplos de vías de administración incluyen la administración parenteral, por ejemplo intravenosa, intradérmica, subcutánea, oral (por ejemplo, inhalación), transdérmica (tópica), transmucosa y rectal. Las soluciones o suspensiones usadas para aplicación parenteral, intradérmica o subcutánea pueden incluir los componentes siguientes: un diluyente estéril tal como agua para la inyección, solución salina, aceites fijos, polietilenglicoles, glicerina, propilenglicol u otros disolventes sintéticos; agentes antibacterianos tales como alcohol bencílico o metilparabenos; antioxidantes tales como ácido ascórbico o bisulfito sódico; agentes quelantes tales como ácido etilendiaminotetraacético; tampones tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro sódico dextrosa. El pH puede ajustarse con ácidos o bases, tales como ácido clorhídrico o hidróxido sódico. La preparación parenteral puede encerrarse en ampollas, jeringas desechables o viales de múltiples dosis fabricados de vidrio o plástico.

55 Las composiciones farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas estériles (solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables. Para la administración intravenosa, los vehículos adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor EL™ (BASF, Parsippany, NJ) o solución salina tamponada con fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida en tal medida que pueda administrarse por una jeringuilla. Debe ser estable en las condiciones de fabricación y almacenamiento y debe conservarse contra 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) y mezclas adecuadas de los mismos. La fluidez apropiada puede mantenerse, por ejemplo, por medio del uso de un recubrimiento tal como lecitina, por el mantenimiento del tamaño de partículas requerido en el caso de dispersión y por el uso de tensioactivos. La prevención de la acción de microorganismos puede conseguirse por diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol y cloruro sódico en la composición. La absorción prolongada de las composiciones inyectables puede producirse incluyendo en la composición un agente que retrasa la absorción, por ejemplo, monoestearato de aluminio y gelatina.

ES 2 328 448 T3

Pueden prepararse soluciones inyectables estériles incorporando el compuesto activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de ingredientes enumerados anteriormente, cuando se requiera, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los demás ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos de preparación preferidos son secado al vacío y liofilización que produce un polvo del ingrediente activo más cualquier ingrediente deseado adicional a partir de una solución filtrada de forma estéril previamente del mismo.

Las composiciones orales generalmente incluyen un diluyente inerte o un vehículo comestible. Para la administración terapéutica oral, el compuesto activo puede incorporarse con excipientes y usarse en forma de comprimidos, trociscos o cápsulas, por ejemplo cápsulas de gelatina. También pueden prepararse composiciones orales usando un vehículo líquido para uso como un colutorio. Pueden incluirse agentes aglutinantes farmacéuticamente compatibles y/o materiales adyuvantes como parte de la composición. Los comprimidos, píldoras, cápsulas, trociscos y similares pueden contener cualquiera de los siguientes ingredientes o compuestos de una naturaleza similar: un aglutinante tal como una celulosa microcristalina, goma de tragacanto o gelatina; un excipiente tal como almidón o lactosa, un agente desintegrante tal como ácido algínico, Primogel o almidón de maíz; un lubricante tal como estearato de magnesio o Sterotes; un deslizante tal como dióxido de silicio coloidal; un agente edulcorante tal como sacarosa o sacarina; o un agente aromatizante tal como menta, salicilato de metilo o aroma de naranja.

Para la administración por inhalación, se administran los compuestos en forma de una pulverización de aerosol desde un recipiente o dispensador a presión que contiene un propulsor adecuado, por ejemplo, un gas tal como dióxido de carbono, o un nebulizador.

La administración sistémica también puede realizarse por medios transmucosos o transdérmicos. Para la administración transmucosa o transdérmica, en la formulación se usan penetrantes apropiados para la barrera a atravesar. Estos penetrantes generalmente se conocen en la técnica e incluyen, por ejemplo, para la administración transmucosa, detergentes, sales biliares y derivados del ácido fusídico. La administración transmucosa puede realizarse por medio del uso de pulverizaciones nasales o supositorios. Para la administración transdérmica, se formulan los compuestos activos en pomadas, ungüentos, geles o cremas como se conoce en general en la técnica.

Los compuestos también pueden prepararse en forma de supositorios (por ejemplo, con bases de supositorio convencionales tales como manteca de cacao y otros glicéridos) o enemas de retención para la administración rectal.

En una realización, se preparan los compuestos activos con vehículos que protegerán al compuesto frente a la eliminación rápida del cuerpo, tal como una formulación de liberación controlada, incluyendo implantes y sistemas de liberación microencapsulados. Pueden utilizarse polímeros biodegradables y biocompatibles, tales como vinil acetato de etileno, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Para los especialistas en la técnica serán evidentes métodos para la preparación de estas formulaciones. Los materiales también pueden obtenerse en el mercado en Alza Corporation y Nova Pharmaceuticals, Inc. También pueden usarse suspensiones de liposomas (incluidos los liposomas dirigidos a células infectadas con anticuerpos monoclonales contra los antígenos virales) como vehículos farmacéuticamente aceptables. Éstas pueden prepararse de acuerdo con métodos conocidos por los especialistas en la técnica, por ejemplo, como se describe en la Patente de Estados Unidos N° 4.522.811.

Es ventajoso formular composiciones orales o parenterales en forma de dosificación unitaria para facilitar la administración y uniformidad de la dosificación. La forma de dosificación unitaria como se usa en la presente memoria se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para el sujeto a tratar; conteniendo cada unidad una cantidad predeterminada del compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido.

La toxicidad y la eficacia terapéutica de estos compuestos puede determinarse por procedimientos farmacéuticos convencionales en cultivos celulares o animales experimentales, por ejemplo para determinar la DL50 (la dosis letal a 50% de la población) y la DE50 (la dosis terapéuticamente eficaz en 50% de la población). La relación de dosis entre los efectos tóxicos y terapéuticos es el índice terapéutico y puede expresarse como la relación DL50/DE50. Se prefieren compuestos que presentan altos índices terapéuticos. Aunque pueden usarse compuestos que presentan efectos secundarios tóxicos, debe tenerse cuidado de diseñar un sistema de liberación que dirija dichos compuestos al sitio del tejido afectado para minimizar las lesiones potenciales en las células no infectadas y, de esta manera, reducir los efectos secundarios.

Los datos obtenidos a partir de los ensayos de cultivos celulares y los estudios animales pueden usarse para formular un intervalo de dosificación para uso en seres humanos. La dosificación de estos compuestos está preferiblemente dentro de un intervalo de concentraciones circulantes que incluyen la DE50 con poca toxicidad o sin toxicidad. La dosificación puede variar dentro de este intervalo dependiendo de la forma de dosificación empleada y de la vía de administración utilizada. Para cualquier compuesto usado en el método de la invención, la dosis terapéuticamente eficaz puede estimarse inicialmente a partir de ensayos de cultivos celulares. Una dosis puede formularse en modelos de animales para conseguir un intervalo de concentraciones plasmáticas circulantes que incluye la CI50 (es decir, la concentración del compuesto de ensayo que consigue una inhibición semimáxima de los sínto-

mas) como se determina en el cultivo celular. Esta información puede usarse para determinar de forma más precisa las dosis útiles en humanos. Los niveles en plasma pueden medirse, por ejemplo, por cromatografía líquida de alta resolución.

5 Como se define en la presente memoria, una cantidad terapéuticamente eficaz de proteína o polipéptido (es decir, una dosificación eficaz) varía de aproximadamente 0,001 a 30 mg/kg de peso corporal, preferiblemente de aproximadamente 0,01 a 25 mg/kg de peso corporal, más preferiblemente de aproximadamente 0,1 a 20 mg/kg de peso corporal e incluso más preferiblemente de aproximadamente 1 a 10 mg/kg, de 2 a 9 mg/kg, de 3 a 8 mg/kg, de 4 a 7 mg/kg o de 5 a 6 mg/kg de peso corporal. La proteína o polipéptido puede administrarse una vez por semana durante un periodo
10 comprendido entre aproximadamente 1 y 10 semanas, preferiblemente entre 2 y 8 semanas, más preferiblemente entre aproximadamente 3 y 7 semanas e incluso más preferiblemente durante aproximadamente 4, 5 ó 6 semanas. El especialista en la técnica apreciará que ciertos factores pueden influir en la dosificación y programa de tiempo necesario para tratar eficazmente a un sujeto, incluyendo, pero sin limitación, la gravedad de la enfermedad o el trastorno, los tratamientos previos, la salud general y/o la edad del sujeto, y otras enfermedades presentes. Además, el tratamiento
15 de un sujeto con una cantidad terapéuticamente eficaz de una proteína, polipéptido o anticuerpo puede incluir un solo tratamiento o, preferiblemente, puede incluir una serie de tratamientos.

En el caso de los anticuerpos, la dosificación preferida es de 0,1 mg/kg de peso corporal (generalmente de 10 mg/kg a 20 mg/kg). Si el anticuerpo va actuar en el cerebro, normalmente es apropiada una dosificación de 50 mg/kg a 100 mg/kg. Generalmente, los anticuerpos parcialmente humanos y los anticuerpos completamente humanos tienen una mayor vida media dentro del cuerpo humano que otros anticuerpos. Por consiguiente, con frecuencia es posible el uso de menos dosificaciones y una administración menos frecuente. Pueden usarse modificaciones tales como lipidación para estabilizar anticuerpos y para aumentar la captación y penetración en el tejido (por ejemplo en el cerebro). Cruikshank *et al.* ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14: 193) describen
25 un método para lipidación de anticuerpos.

La presente invención incluye agentes que modulan la expresión o actividad. Un agente puede ser, por ejemplo, una molécula pequeña. Por ejemplo, estas moléculas pequeñas incluyen, pero sin limitación, péptidos, peptidomiméticos (por ejemplo peptoides), aminoácidos, análogos de aminoácido, polinucleótidos, análogos de polinucleótido, nucleótidos, análogos de nucleótido, compuestos orgánicos e inorgánicos (es decir incluyendo compuesto heteroorgánicos y organometálicos) que tienen un peso molecular menor de aproximadamente 10.000 gramos por mol, compuestos orgánicos o inorgánicos que tienen un peso molecular menor de aproximadamente 5.000 gramos por mol, compuestos orgánicos o inorgánicos que tienen un peso molecular menor de aproximadamente 1.000 gramos por mol, compuestos orgánicos o inorgánicos que tienen un peso molecular menor de aproximadamente 500 gramos por mol y sales, ésteres y otras formas farmacéuticamente aceptables de estos compuestos.
35

Las dosis ejemplares incluyen cantidades en miligramos o microgramos de molécula pequeña por kilogramo de peso del sujeto o de muestra (por ejemplo, de aproximadamente 1 microgramo por kilogramo a aproximadamente 500 miligramos por kilogramo, de aproximadamente 100 microgramos por kilogramo a aproximadamente 5 miligramos por kilogramo o de aproximadamente 1 microgramo por kilogramo a aproximadamente 50 microgramos por kilogramo). Además, debe entenderse que las dosis apropiadas de una molécula pequeña dependen de la potencia de la molécula pequeña con respecto a la expresión o actividad a modular. Cuando una o más de estas moléculas pequeñas se van administrar a un animal (por ejemplo, un ser humano) para modular la expresión o actividad de un polipéptido o ácido nucleico de la invención, un médico, veterinario o investigador puede prescribir, por ejemplo, una dosis relativamente baja al principio, aumentando posteriormente la dosis hasta que se obtenga la respuesta apropiada. Además, se entiende que el nivel dosificación específico para cualquier sujeto animal particular dependerá del peso corporal, el estado de salud general, el sexo, y la dieta del sujeto, el momento de administración, la vía de administración, la velocidad de excreción, cualquier combinación con fármacos y el grado de expresión o actividad a modular.
45

50 Un anticuerpo (o un fragmento del mismo) puede conjugarse con un resto terapéutico tal como una citotoxina, un agente terapéutico o un ion metálico radiactivo. Una citotoxina o agente citotóxico incluye cualquier agente que es perjudicial para las células. Los ejemplos incluyen taxol, citocalasina B, gramicidina D, bromuro de etidio, emetina, mitomicina, etopósido, tenopósido, vincristina, vinblastina, colquicina, doxorubicina, daunorrubicina, dihidroxiantra-cindiona, mitoxantrona, mitramicina, actinomicina D, 1-deshidrotestosterona, glucocorticoides, procaína, tetracaína, lidocaína, propranolol y puomicina y análogos u homólogos de los mismos. Los agentes terapéuticos incluyen, pero sin limitación, antimetabolitos (por ejemplo, metotrexato, 6-mercaptopurina, 6-tioguanina, citarabina, 5-fluorouracilo decarbazina), agentes alquilantes (por ejemplo, mecloretamina, tioepa clorambucilo, melfalán, carmustina (BSNU) y lomustina (CCNU), ciclofosfamida, busulfán, dibromomanitol, estreptozocina, mitomicina C y cis-diclorodiamina platino (II) (DDP) cisplatino), antraciclinas (por ejemplo, daunorrubicina (anteriormente daunomicina) y doxorubicina), antibióticos (por ejemplo, dactinomicina (anteriormente actinomicina)), bleomicina, mitramicina y antramicina (AMC) y agentes antimitóticos (por ejemplo, vincristina y vinblastina).
60

Los conjugados de la invención pueden usarse para modificar una respuesta biológica dada y no debe considerarse que el resto de fármaco limita los agentes terapéuticos químicos clásicos. Por ejemplo, el resto de fármaco puede ser una proteína o un polipéptido que posee una actividad biológica deseada. Estas proteínas pueden incluir, por ejemplo, una toxina tal como abrina, ricina A, exotoxina de *Pseudomonas* o toxina diftérica; una proteína tal como factor de necrosis tumoral, α -interferón, β -interferón, factor de crecimiento nervioso, factor de crecimiento derivado de plaquetas, activador de plasminógeno tisular; o modificadores de la respuesta biológica tales como, por ejemplo, linfoquinas,
65

interleuquina-1 (“IL-1”), interleuquina-2 (“IL-2”), interleuquina-6 (“IL-6”), factor estimulador de colonias de granulocitos y macrófagos (“GM-CSF”), factor estimulador de colonias de granulocitos (“G-CSF”) u otros factores de crecimiento.

5 Como alternativa, un anticuerpo puede conjugarse con un segundo anticuerpo para formar heteroconjugados de anticuerpo como describe Segal en la Patente de Estados Unidos N° 4.676.980.

10 Las moléculas de ácido nucleico de la invención pueden insertarse en vectores y usarse como vectores de terapia génica. Los vectores de terapia génica pueden administrarse a un sujeto, por ejemplo, por inyección intravenosa, administración local (véase la Patente de Estados Unidos N° 5.328.470) o por inyección estereotáctica (véase, por ejemplo, Chen *et al.*) (1994) Proc. Natl. Acad. Sci. USA 91: 3054-3057). La preparación farmacéutica del vector de terapia génica puede incluir el vector de terapia génica en un diluyente aceptable o puede comprender una matriz de liberación lenta en la que se incluye el vehículo de liberación del gen. Como alternativa, cuando el vector de liberación génica completo puede producirse intacto a partir de células recombinantes, por ejemplo, vectores retrovirales, la preparación farmacéutica puede incluir una o más células que producen el sistema de liberación génica.

15 Las composiciones farmacéuticas pueden incluirse en un recipiente, envase o dispensador junto con instrucciones para la administración.

20

Métodos de Tratamiento

25 La presente invención proporciona métodos profilácticos y terapéuticos para tratar a un sujeto con riesgo (o susceptible de padecer) un trastorno o que tiene un trastorno asociado con la actividad o expresión aberrante o indeseada de 16836. Como se usa en la presente memoria, el término “tratamiento” se define como la aplicación o administración de un agente terapéutico a un paciente, o la aplicación o administración de un agente terapéutico a un tejido o línea celular aislada de un paciente, que tiene una enfermedad, un síntoma de enfermedad o una predisposición hacia una enfermedad, con el objetivo de curar, sanar, aliviar, calmar, alterar, remediar, mejorar, paliar o afectar a la enfermedad, los síntomas de la enfermedad o la predisposición a una enfermedad. Un agente terapéutico incluye, pero sin limitación, moléculas pequeñas, péptidos, anticuerpos, ribozimas y oligonucleótidos antisentido.

30 Con respecto a los métodos de tratamiento tanto profilácticos como terapéuticos, estos tratamientos pueden adaptarse o modificarse específicamente, basándose en el conocimiento obtenido del campo de la farmacogenómica. La “farmacogenómica”, como se usa en la presente memoria, se refiere a la aplicación de tecnologías genómicas tales como secuenciación de genes, genética estadística y el análisis de expresión génica a fármacos en desarrollo clínico y en el mercado. Más específicamente, el término se refiere al estudio de cómo los genes de un paciente determinan su respuesta a un fármaco (por ejemplo, un “fenotipo de respuesta al fármaco” de un paciente o un “genotipo de respuesta al fármaco”). De esta manera, otro aspecto de la invención proporciona métodos para adaptar un tratamiento profiláctico o terapéutico de un individuo con las moléculas de 16836 de la presente invención o moduladores de 16836 de acuerdo con ese genotipo de respuesta al fármaco del individuo. La farmacogenómica permite a un médico dirigir tratamientos profilácticos o terapéuticos a pacientes que se beneficiarán más del tratamiento y para evitar el tratamiento de pacientes que experimentarán efectos secundarios tóxicos relacionados con el fármaco.

35

40 En un aspecto, la invención proporciona un método para prevenir en un sujeto, una enfermedad o afección asociada con una actividad o expresión aberrante o indeseada de 16836, por medio de la administración al sujeto de 16836 o un agente que modula la expresión de 16836 o al menos una actividad de 16836. Los sujetos con riesgo de padecer una enfermedad que se produce o a la que contribuye una actividad o expresión aberrante o indeseada de 16836 pueden identificarse, por ejemplo, por cualquiera de o una combinación de ensayos de diagnóstico o pronóstico como se describen en la presente memoria. La administración de un agente profiláctico puede realizarse antes de la manifestación de los síntomas característicos de la aberración de 16836, de tal forma que se prevenga una enfermedad o trastorno o, como alternativa, retrasase su progresión. Dependiendo del tipo de aberración de 16836, por ejemplo, puede usarse para tratar al sujeto una proteína 16836, un agonista de 16836 o un antagonista de 16836. El agente apropiado puede determinarse basándose en ensayos de exploración descritos en la presente memoria.

45

50 Es posible que algunos trastornos de 16836 puedan producirse, al menos en parte por un nivel anómalo de producto génico, o por la presencia de un producto génico que presenta una actividad anormal. Como tal, la reducción en el nivel y/o actividad de dichos productos génicos producirá la mejoría de los síntomas del trastorno.

55

60 Como se muestra en la Figura 4, la expresión del ARNm de 16836 se detecta en células óseas (por ejemplo, osteoclastos), testículos, cerebro, músculo esquelético, mama, corazón e hígado fetal. Por consiguiente, las moléculas de la invención pueden mediar trastornos que implican actividades aberrantes de estas células, por ejemplo, trastornos óseos, trastornos cerebrales, trastornos de mama, trastornos reproductivos (por ejemplo, trastornos testiculares), trastornos cardiovasculares, trastornos del músculo esquelético o trastornos inmunes.

65

Como se describe, el tratamiento satisfactorio de trastornos de 16836 pueden producirse por técnicas que sirven para inhibir la expresión o actividad de productos génicos diana. Por ejemplo, los compuestos, por ejemplo, un agen-

te identificado usando un ensayo descrito anteriormente que demuestra presentar una actividad moduladora negativa, pueden usarse de acuerdo con la invención para prevenir y/o mejorar síntomas de trastornos de 16836. Estas moléculas pueden incluir, pero sin limitación, péptidos, fosfopéptidos, moléculas orgánicas o inorgánicas pequeñas o anticuerpos, (incluyendo, por ejemplo, anticuerpos policlonales, monoclonales, humanizados, anti-idiotípicos, quiméricos o monocatenarios, Fab, F(ab')₂ y fragmentos de bibliotecas de expresión de Fab, moléculas scFV y fragmentos de unión a epítipo de las mismas).

Además, de acuerdo con la invención también pueden usarse moléculas antisentido y ribozimas que inhiben la expresión del gen diana para reducir el nivel de expresión génica diana, reduciendo de esta manera eficazmente el nivel de actividad génica diana. Además, pueden utilizarse moléculas de triple hélice para reducir el nivel de actividad génica diana. Las moléculas antisentido, ribozimas y de triple hélice se han descrito anteriormente.

Es posible que el uso de moléculas antisentido, ribozimas y/o de triple hélice para reducir o inhibir la expresión del gen mutante también reduzca o inhiba la transcripción (triple hélice) y/o la traducción (antisentido, ribozima) de un ARNm producido por alelos de genes diana normales, de tal forma que la concentración de producto génico diana normal presente puede ser menor de lo necesario para un fenotipo normal. En estos casos, pueden introducirse moléculas de ácido nucleico que codifican y expresan polipéptidos de genes diana que presentan una actividad de gen diana normal en las células por medio de un método de terapia génica. Como alternativa, en casos en los que el gen diana codifica una proteína extracelular, puede ser preferible co-administrar una proteína de un gen diana normal en la célula o tejido o mantener el nivel requerido de actividad del gen diana celular o tisular.

Otro método por medio del cual pueden utilizarse moléculas de ácido nucleico en el tratamiento o prevención de una enfermedad caracterizada por la expresión de 16836 es por medio del uso de moléculas de aptámero específicas para la proteína 16836. Los aptámeros son moléculas de ácido nucleico que tienen una estructura terciaria que les permite unirse específicamente a ligandos proteicos (véase, por ejemplo, Osborne, *et al.* (1997) *Curr. Opin. Chem Biol.* 1: 5-9; y Patel, D. J. (1997) *Curr Opin Chem Biol* 1: 32-46). Como las moléculas de ácido nucleico en muchos casos pueden introducirse más convenientemente en células diana que las moléculas de proteína terapéuticas, los aptámeros ofrecen un método por el medio del cual puede reducirse específicamente la actividad de la proteína 16836 sin la introducción de fármacos u otras moléculas que pueden tener efectos pluripotent.

Pueden generarse anticuerpos que sean específicos para una producto génico diana y que reduzcan la actividad del producto génico diana. Estos anticuerpos, por lo tanto, pueden administrarse en casos en los que sean apropiadas para el tratamiento de trastornos de 16836 técnicas moduladoras negativas. Para una descripción de anticuerpos, véase la sección de Anticuerpos anterior.

En circunstancias en las que la inyección en un animal o un ser humano de una proteína o epítipo de 16836 para estimular la producción de anticuerpos es perjudicial para el sujeto, es posible generar una respuesta inmune contra 16836 por medio del uso de anticuerpos anti-idiotípicos (véase, por ejemplo, Herlyn, D. (1999) *Ann Med* 31: 66-78; y Bhattacharya-Chatterjee, M., y Foon, K.A. (1998) *Cancer Treat Res.* 94: 51-68). Si se introduce un anticuerpo anti-idiotípico en un mamífero o en un ser humano, deberían estimular la producción de anticuerpos anti-anti-idiotípicos, que debería ser específicos para la proteína 16836. También pueden generarse de esta manera vacunas dirigidas a una enfermedad caracterizada por la expresión de 16836.

En casos en los que el antígeno diana es intracelular y se usan anticuerpos enteros, pueden preferirse anticuerpos de internalización. Puede usarse lipofectina o liposomas para administrar el anticuerpo o un fragmento de la región Fab que se une al antígeno diana en las células. Cuando se usan fragmentos del anticuerpo, se prefiere el menor fragmento inhibitor que se une al antígeno diana. Por ejemplo, pueden usarse péptidos que tienen una secuencia de aminoácidos correspondiente a la región Fv del anticuerpo. Como alternativa, también pueden administrarse anticuerpos neutralizadores monocatenarios que se unen a antígenos diana intracelulares. Estos anticuerpos monocatenarios pueden administrarse, por ejemplo, por expresión de secuencias de nucleótidos que codifican anticuerpos monocatenarios dentro de la población de células diana (véase, por ejemplo, Marasco *et al.* (1993) *Proc. Natl. Acad Sci. USA* 90: 7889-7893).

Los compuestos identificados que inhiben la expresión, síntesis y/o actividad del gen diana pueden administrarse a un paciente a dosis terapéuticamente eficaces para prevenir, tratar o mejorar trastornos de 16836. Una dosis terapéuticamente eficaz se refiere a la cantidad del compuesto suficiente para producir una mejoría de los síntomas de los trastornos. La toxicidad y la eficacia terapéutica de estos compuestos pueden determinarse por procedimientos farmacéuticos convencionales como se ha descrito anteriormente.

Los datos obtenidos a partir de los ensayos de cultivo de células y los estudios animales pueden utilizarse para formular un intervalo de dosificación para uso en seres humanos. La dosificación de estos compuestos está preferiblemente dentro de un intervalo de concentraciones circulantes que incluyen la DE₅₀ con toxicidad baja o nula. La dosificación puede variar dentro de este intervalo dependiendo de la forma de dosificación empleada y la vía de administración utilizada. Para cualquier compuesto usado en el método de la invención, la dosis terapéuticamente eficaz puede estimarse inicialmente a partir de ensayos de cultivo de células. Una dosis puede formularse en modelos animales para conseguir un intervalo de concentraciones en el plasma circulantes que incluye la CI₅₀ (es

decir, la concentración del compuesto de ensayo que consigue una inhibición semimáxima de los síntomas) como se determina en cultivo de celular. Esta información puede usarse para determinar de forma más precisa las dosis útiles en seres humanos. Los niveles en plasma pueden medirse, por ejemplo, por cromatografía líquida de alta resolución.

Otro ejemplos de determinación de dosis eficaz para un individuo es la capacidad de ensayar directamente niveles de compuesto “libre” y “unido” en el suero del sujeto de ensayo. Estos ensayos pueden utilizar miméticos de anticuerpo y/o “biosensores” que se han creado por medio de técnicas de moldeado molecular. El compuesto que puede modular la actividad de 16836 se usa como plantilla o “molécula de molde”, para organizar espacialmente monómeros polimerizables antes de su polimerización con reactivos catalíticos. La posterior eliminación de la molécula moldeada deja una matriz de polímero que contiene una “imagen negativa” repetida del compuesto y puede volverse a unir selectivamente a la molécula en condiciones de ensayo biológico. Puede verse una revisión detallada de esta técnica en Ansell, R. J. *et al* (1996) *Current Opinion in Biotechnology* 7: 89-94 y en Shea, K. J. (1994) *Trends in Polymer Science* 2: 166-173. Estas matrices de afinidad “moldeadas” son susceptibles de ensayos de unión a ligando, con lo que el componente de anticuerpo monoclonal inmovilizado se reemplaza por una matriz moldeada de forma apropiada. Un ejemplo del uso de estas matrices de esta manera puede verse en Vlatakis, G. *et al* (1993) *Nature* 361: 645-647. Por medio del uso del marcaje de isótopos, la concentración “libre” del compuesto que modula la expresión o actividad de 16836 puede controlarse fácilmente y usarse en cálculos de CI_{50} .

Estas matrices de afinidad “moldeadas” también pueden diseñarse para incluir grupos fluorescentes cuyas propiedades de emisión de fotones cambian de forma medible tras la unión local y selectiva del compuesto diana. Estos cambios pueden ensayarse fácilmente a tiempo real usando dispositivos de fibra óptica apropiados, permitiendo a su vez que la dosis en un sujeto de ensayo se optimice rápidamente basándose en CI_{50} individual. En Kriz, D. *et al* (1995) *Analytical Chemistry* 67: 2142-2144 se describe un ejemplo rudimentario de dicho “biosensor”.

Otro aspecto de la invención se refiere a métodos para modular la expresión o actividad de 16836 con fines terapéuticos. Por consiguiente, en una realización ilustrativa, el método modulador de la invención implica poner en contacto una célula con una proteína 16836 o un agente que modula una o más de las actividades de la proteína 16836 asociada con la célula. Un agente que modula la actividad de la proteína 16836 puede ser un agente como se describe en la presente memoria, tal como un ácido nucleico o una proteína, una molécula diana natural de una proteína 16836 (por ejemplo un sustrato o receptor de 16836), un anticuerpo de 16836, un agonista o antagonista de 16836, un peptidomimético de un agonista o antagonista de 16836, u otra molécula pequeña.

En una realización, el agente estimula una o más de las actividades de 16836. Los ejemplos de estos agentes estimuladores incluyen la proteína 16836 activa y una molécula de ácido nucleico que codifica 16836. En otra realización, el agente inhibe una o más actividades de 16836. Los ejemplos de estos agentes inhibidores incluyen moléculas de ácido nucleico de 16836 antisentido, anticuerpos anti-16836 e inhibidores de 16836. Estos métodos moduladores pueden realizarse *in vitro* (por ejemplo, por cultivo de la célula con el agente) o, como alternativa, *in vivo* (por ejemplo, administrando el agente a un sujeto). Como tal, la presente invención proporciona métodos de tratamiento de un individuo que padece una enfermedad o trastorno caracterizado por una expresión o actividad aberrante o indeseada de una molécula de proteína o ácido nucleico de 16836. En una realización, el método implica administrar un agente (por ejemplo, un agente identificado por un ensayo de exploración descrito en la presente memoria) o una combinación de agentes que modulan (por ejemplo, regulan positivamente o regula negativamente) la expresión o actividad de 16836. En otra realización, el método implica administrar una molécula de proteína o ácido nucleico de 16836 como terapia para compensar la expresión o actividad de 16836 reducida, aberrante o indeseada.

La estimulación de la actividad de 16836 es deseable en situaciones en las que 16836 está regulada negativamente de forma anómala y/o en las que es probable que un aumento de la actividad de 16836 tenga un efecto beneficioso. Por ejemplo, la estimulación de la actividad de 16836 es deseable en situaciones en las que 16836 está regulada negativamente y/o en las que es probable que un aumento de la actividad de 16836 tenga un efecto beneficioso. De forma similar, la inhibición de la actividad de 16836 es deseable en situaciones en las que 16836 está regulada positivamente de forma anómala y/o en las que es probable que una reducción de la actividad de 16836 tenga un efecto beneficioso.

Farmacogenómica

Las moléculas de 16836 de la presente invención, así como agentes o moduladores que tienen un efecto estimulador o inhibidor sobre la actividad de 16836 (por ejemplo, la expresión del gen de 16836) identificada por un ensayo de exploración descrito en la presente memoria pueden administrarse a individuos para tratar (profiláctica o terapéuticamente) trastornos asociados con 16836 (por ejemplo, trastornos de la proliferación y/o diferenciación celular) asociados con una actividad aberrante o indeseada de 16836. Junto con dicho tratamiento, puede considerarse la farmacogenómica (es decir, el estudio de la relación entre el genotipo de un individuo y la respuesta del individuo a un compuesto o fármaco extraño). Las diferencias en el metabolismo de los agentes terapéuticos pueden conducir a una toxicidad severa o fallo terapéutico por alteración de la relación entre la dosis y la concentración sanguínea del fármaco farmacológicamente activo. De esta manera, un médico puede considerar la aplicación del conocimiento

obtenido en estudios farmacogenómicos relevantes para determinar si hay que administrar una molécula de 16836 o un modulador de 16836 así como para adaptar la dosis y/o régimen terapéutico de tratamiento con una molécula de 16836 o modulador de 16836.

5 La farmacogenómica trata de variaciones hereditarias clínicamente significativas en la respuesta a fármacos debidas a una alteración de la disposición del fármaco y a una acción anómala en personas afectadas. Véase, por ejemplo, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23: 983-985 y Linder, M. W. *et al.* (1997) *Clin. Chem.* 43: 254-266. En general, pueden diferenciarse dos tipos de condiciones farmacogenéticas. Las condiciones genéticas transmitidas como un solo factor que altera la forma en la que actúan los fármacos en el cuerpo (acción del fármaco alterada) o condiciones genéticas transmitidas como factores únicos que alteran la forma en la que el cuerpo actúa sobre los fármacos (metabolismo del fármaco alterado). Estas condiciones farmacogenéticas pueden producirse como defectos genéticos raros o como polimorfismos naturales. Por ejemplo, la deficiencia en glucosa-6-fosfato deshidrogenasa (G6PD) es una enzimopatía hereditaria común en la que la complicación clínica principal es la hemólisis después de la ingestión de fármacos oxidantes (antimaláricos, sulfonamidas, analgésicos, nitrofuranos) y el consumo de habas.

Una estrategia farmacogenómica para identificar genes que predicen la respuesta a un fármaco, conocida como “asociación a nivel del genoma” se basa principalmente en el mapa de alta resolución del genoma humano consistente en marcadores relacionados con genes ya conocidos (por ejemplo, un mapa de marcadores genéticos “bialélicos” que consiste en 60.000-100.000 sitios polimórficos o variables en el genoma humano, de los que cada uno tiene dos variantes). Dicho mapa genético de alta resolución puede compararse con un mapa del genoma de cada uno de un número estadísticamente significativo de pacientes que participan en un ensayo de fármacos de fase II/III para identificar marcadores relacionados con una respuesta al fármaco o efecto secundario observado particular. Como alternativa, este mapa de alta resolución puede generarse a partir de una combinación de unos diez millones de polimorfismos de un solo mononucleótido (SNP) conocidos en el genoma humano. Como se usa en la presente memoria, un “SNP” es una alteración común que se produce en una base en una sola base de nucleótido en un tramo de ADN. Por ejemplo, un SNP puede aparecer una vez por cada 1000 bases de ADN. Un SNP puede implicarse en un proceso de enfermedad; sin embargo, la inmensa mayoría puede no estar asociados con la enfermedad. Dado un mapa genético basado en la aparición de dicho SNP, los individuos pueden agruparse en categorías genéticas dependiendo de un patrón particular de SNP en su genoma individual. De esta manera, pueden adaptarse regímenes de tratamiento a grupos de individuos genéticamente similares, teniendo en cuenta rasgos que pueden ser comunes entre dichos individuos genéticamente similares.

Como alternativa, puede utilizarse un método denominado “estrategia de gen candidato” para identificar genes que predicen una respuesta a un fármaco. De acuerdo con este método, si se conoce un gen que codifica la diana de un fármaco (por ejemplo, una proteína 16836 de la presente invención), se pueden identificar bastante fácilmente todas las variantes comunes de ese gen en la población y puede determinarse si la posesión de una versión del gen frente a otra está asociada con una respuesta particular al fármaco.

Como alternativa, puede utilizarse un método denominado “perfilado de expresión génica” para identificar genes que predicen la respuesta al fármaco. Por ejemplo, la expresión génica de un animal al que se le ha dosificado un fármaco (por ejemplo, una molécula de 16836 o un modulador de 16836 de la presente invención) puede dar una indicación de si se han activado rutas génicas relacionadas con la toxicidad.

La información generada a partir de más de una de las estrategias farmacogenómicas anteriores puede utilizarse para determinar los regímenes de dosificación y tratamiento apropiados para el tratamiento profiláctico o terapéutico de un individuo. Este conocimiento, cuando se aplica a la dosificación o selección de fármacos, puede evitar reacciones adversas o el fallo terapéutico y por lo tanto mejorar la eficacia terapéutica o profiláctica cuando se trata un sujeto con una molécula de 16836 o un modulador de 16836, tal como un modulador identificado por uno de los ensayos de exploración ilustrativos descritos en la presente memoria.

La presente invención también proporciona métodos para identificar nuevos agentes, o combinaciones, que se basan en la identificación de agentes que modulan la actividad de uno o más de los productos génicos codificados por uno o más de los genes de 16836 de la presente invención, donde estos productos pueden asociarse con resistencia de las células a un agente terapéutico. Específicamente, la actividad de las proteínas codificadas por los genes de 16836 de la presente invención puede usarse como base para identificar agentes para solucionar la resistencia al agente. Por medio del bloqueo de la actividad de una o más de las proteínas de resistencia, las células diana, por ejemplo, células humanas, se volverán sensibles al tratamiento con un agente al que eran resistentes las células diana no modificadas.

La monitorización de la influencia de agentes (por ejemplo, fármacos) sobre la expresión o actividad de una proteína 16836 puede aplicarse en ensayos clínicos. Por ejemplo, la eficacia de un agente determinado por un ensayo de exploración como se describe en la presente memoria para aumentar la expresión del gen de 16836, los niveles de proteína 16836 o regular positivamente la actividad de 16836 pueden monitorizarse en ensayos clínicos de sujetos que presenta un aumento de la expresión génica de 16836, niveles de proteína o una actividad de 16836 regulada negativamente. Como alternativa, la eficacia de un agente determinada por un ensayo de exploración para reducir la expresión del gen de 16836, niveles de proteína o la actividad de 16836 regulada negativamente puede monitorizarse en ensayos clínicos de sujetos que presentan un aumento de la expresión del gen de 16836, niveles de proteína o una

actividad de 16836 regulada positivamente. En estos ensayos clínicos, la expresión o actividad de un gen de 16836 y, preferiblemente, otros genes que se han implicado en, por ejemplo, un trastorno asociado con 16836, puede usarse como “indicadores” o marcadores del fenotipo de una célula particular.

5

Ejemplos

Ejemplo 1

10 *Identificación y Caracterización de ADNc de 16836 Humano*

La secuencia de 16836 humana (Figs. 1A-1J; SEC ID N°: 1), que tiene una longitud de aproximadamente 10.172 nucleótidos incluyendo regiones no traducidas, contiene una secuencia codificante iniciada por metionina prevista de aproximadamente 5.430 nucleótidos, incluyendo el codon de terminación (nucleótidos indicados como “codificantes” de la SEC ID N°: 1 en las Figs. 1; SEC ID N°: 3). La secuencia codificante codifica una proteína de 1809 aminoácidos (SEC ID N°: 2).

15

Ejemplo 2

20

Distribución Tisular de ARNm de 16836 por Análisis TaqMan

Se determinó la expresión endógena del gen de 16836 humano usando el Perkin-Elmer/ABI 7700 Sequence Detection System que emplea la tecnología TaqMan. En resumen, la tecnología TaqMan se basa en una RT-PCR convencional con la adición de un tercer oligonucleótido con especificidad de gen (denominado sonda) que tiene un colorante fluorescente acoplado a su extremo 5' (típicamente 6-FAM) y un colorante de inactivación en el extremo 3' (típicamente TAMRA). Cuando el oligonucleótido marcado con fluorescencia está intacto, se inactiva la señal fluorescente del colorante 5'. Según avanza la PCR, la actividad nucleolítica de 5' a 3' de la taq polimerasa digiere el cebador marcado, produciendo un nucleótido libre marcado con 6-FAM, que ahora se detecta como señal fluorescente. El ciclo de PCR donde primero se libera y se detecta la fluorescencia es directamente proporcional a la cantidad de partida del gen de interés en la muestra de ensayo, proporcionando de esta manera una forma de cuantificar la concentración de plantilla inicial. Las muestras pueden controlarse internamente por la adición de una segunda serie de cebadores/sondas específicas para un gen de mantenimiento (housekeeping) tal como GAPDH que se ha marcado con un fluoróforo diferente en el extremo 5' (típicamente VIC).

35

Para determinar el nivel de 16836 en diversos tejidos humanos se diseñó un conjunto de cebador/sonda usando el software Primer Express (Perkin-Elmer) y la información de la secuencia de ADNc primaria. Se preparó ARN total a partir de una serie de tejidos humanos usando el kit RNeasy de Qiagen. Se preparó ADNc de primera cadena a partir de un μg de ARN total usando un cebador de oligo dT y la transcriptasa inversa Superscript II (Gibco/BRL). El ADNc obtenido a partir de aproximadamente 50 ng de ARN total se usó para la reacción TaqMan.

40

Los tejidos normales ensayados incluían los tejidos humanos proporcionados en la Figura 4, incluyendo células óseas (por ejemplo osteoclastos y osteoblastos), hígado, hígado, fetal, cerebro, tráquea, músculo esquelético, corazón, tiroides, piel, testículos, mama y placenta, entre otros. La expresión se encontró principalmente en osteoclastos, testículos, cerebro, músculo esquelético, mama, corazón e hígado fetal (Figura 4).

45

Se observó un aumento de expresión de 16836 en tumores de pulmón (“pulmón T”; carcinoma microcítico (SCC) y adenocarcinoma (AC)) en comparación con el tejido pulmonar normal (N) (Figura 5). Se encontró una asociación de una expresión elevada de 16836 con mutaciones activadoras de *kras* tanto en muestras de tumor de pulmón (Figura 6A) como en líneas de células de tumor derivadas de mama, colon y pulmón (Figura 6B).

50

En un panel que comprende tejidos normales (N) y tumorales (T) de mama, pulmón, colon e hígado, se observó una mayor expresión de 16836 en los tejidos tumorales, especialmente en tumores de mama, pulmón y colon (Figura 7). En la Figura 8 se representa un análisis de la expresión de 16836 en tumor en comparación con tejidos normales de mama. En la Figura 9 se representa un análisis de la expresión de 16836 en varias líneas celulares transformadas. Estas líneas celulares incluían las líneas de células cancerosas de mama humanas (MCF-7, ZR75, T47D y MDA), líneas de células de cáncer de colon humanas (DLD-1, SW480, SW620, HCT116, HT 29 y Colo205), y líneas de células de cáncer de pulmón humanas (NCI-H125 y A549).

55

La incidencia de la expresión asociada a tumores de 16836 en tumores del colon, mama y pulmón se evaluó por hibridación *in situ*. Se detectó una expresión moderada de 16836 en células de tumor de colon (expresión en 0/2 muestras normales; expresión en 2/3 muestras de tumor; y expresión en 0/2 muestras de metástasis). Se observó una expresión ligeramente positiva focalmente en el epitelio normal en una muestra de tumor de mama (expresión en 0/2 muestras normales; y expresión en epitelio normal en 1/4 muestras de tumor). Se observó una expresión moderada en células inflamatorias de muestras de tumor de pulmón (expresión en 0/2 muestras normales; y expresión en 2/4 muestras de tumor).

65

ES 2 328 448 T3

Ejemplo 3

Distribución Tisular de ARNm de 16836 por Análisis de Northern

5 Pueden realizarse hibridaciones de transferencia de Northern con las diversas muestras de ARN en condiciones convencionales y lavarse en condiciones rigurosas, es decir, SSC 0,2x a 65°C. Puede usarse una sonda de ADN correspondiente a todo o a parte del ADNc de 16836 (SEC ID N°: 1). El ADN se marcó radiactivamente con ³²P-dCTP usando el kit Prime-It (Stratagene, La Jolla, CA) de acuerdo con las instrucciones del proveedor. Pueden soldarse filtros que contienen ARNm de tejidos hematopoyéticos y endocrinos de ratón y líneas de células cancerosas (Clontech, Palo Alto, CA) en solución de hibridación ExpressHyb (Clontech) y lavarse con alta rigurosidad de acuerdo con las recomendaciones del fabricante.

15 Se realizaron hibridaciones de transferencia de Northern usando la sonda de 16836 marcada y muestras de ARN procedentes de diversos tumores de pulmón y líneas de células tumorales. Se detectó un transcrito de 16836 con un tamaño mayor de 7,5 kb en algunas de estas muestras. Los resultados fueron los siguientes: NHBE (negativo); A549 (positivo muy fuerte); NCI-H69 (negativo); NCI-H125 (positivo fuerte); NCI-H322 (positivo); y NCI-H460 (positivo).

Ejemplo 4

Expresión Recombinante de 16836 en Células Bacterianas

20 En este ejemplo, se expresa 16836 como un polipéptido de fusión de glutatión-S-transferasa (GST) recombinante en *E. coli* y el polipéptido de fusión se aísla y caracteriza. Específicamente, se fusiona 16836 a GST y se expresa este polipéptido de fusión en *E. coli*, por ejemplo, la cepa PEB199. La expresión de la proteína de fusión GST-16836 en PEB199 se induce con IPTG. El polipéptido de fusión recombinante se purifica a partir de lisados bacterianos brutos de la cepa de PEB 199 inducida por cromatografía de afinidad en perlas de glutatión. Usando análisis electroforético en gel de poliacrilamida del polipéptido purificado a partir de los lisados bacterianos, se determina el peso molecular del polipéptido de fusión resultante.

Ejemplo 5

Expresión de Proteína 16836 Recombinante en Células COS

35 Para expresar el gen de 16836 en células COS, se usa el vector pcDNA/Amp por Invitrogen Corporation (San Diego, CA). Este vector contiene un origen de replicación de SV40, un gen de resistencia a ampicilina, un origen de replicación de *E. coli*, un promotor de CMV seguido de una región polienlazadora y un intrón y un sitio de poliadenilación de SV40. Un fragmento de ADN que codifica la proteína 16836 entera y un marcador HA (Wilson *et al.* (1984) *Cell* 37: 767) o un marcador FLAG fusionado en fase a su extremo 3' del fragmento, se clona en la región polienlazadora del vector, poniendo de esta manera la expresión de la proteína recombinante bajo el control del promotor de CMV.

45 Para construir el plásmido, la secuencia de ADN de 16836 se amplifica por PCR usando dos cebadores. El cebador 5' contiene el sitio de restricción de interés seguido de aproximadamente 20 nucleótidos de la secuencia codificante de 16836 que empieza desde el codón de iniciación; la secuencia del extremo 3' contiene secuencias complementarias al otro sitio de restricción de interés, un codón de terminación de la traducción, el marcador HA o el marcador FLAG y los últimos 20 nucleótidos de la secuencia codificante de 16836. El fragmento amplificado por PCR y el vector pcDNA/Amp se digieren con las enzimas de restricción apropiadas y el vector se desfosforila usando la enzima CIAP (New England Biolabs, Beverly, MA). Preferiblemente, los dos sitios de restricción elegidos son diferentes de forma que el gen de 16836 se inserta en la orientación correcta. La mezcla de ligamiento se utiliza para transformar células *E. coli* (pueden usarse cepas HB101, DH5α y SURE, disponibles en Stratagene Cloning Systems, La Jolla, CA), el cultivo transformado se pone en placas con medio de ampicilina y se seleccionan las colonias resistentes. El ADN plasmídico se aísla a partir de los transformantes y se examina por análisis de restricción con respecto a la presencia del fragmento correcto.

55 Posteriormente se transfectan células COS con el ADN del plásmido pcDNA/Amp-16836 usando métodos de coprecipitación con fosfato cálcico o cloruro de cálcico, transfección mediada por DEAE-dextrano, lipofección, o electroporación. Pueden encontrarse otros métodos adecuados para transfectar células hospedadoras en Sambrook, J., Fritsh, E. F., y Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. La expresión del polipéptido 16836 se detecta por radiomarcaje (puede usarse ³⁵S-metionina o ³⁵S-cisteína disponible en NEN, Boston, MA) e inmunoprecipitación (Harlow, E. y Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) usando el anticuerpo monoclonal específico de HA. En resumen, las células se marcan durante 8 horas con ³⁵S-metionina (o ³⁵S-cisteína). Después se recogen los medios de cultivo y las células se lisan usando detergentes (tampón RIPA, NaCl 150 mM, NP-40 al 1%, SDS al 0,1%, DOC al 0,5%, Tris 50 mM, pH 7,5). Tanto el lisado celular como el medio de cultivo se precipitan con un anticuerpo monoclonal específico de HA. Los polipéptidos precipitados después se analizan por SDS-PAGE.

ES 2 328 448 T3

Como alternativa, el ADN que contiene la secuencia codificante de 16836 se clona directamente en el polienlazador del vector pcDNA/Amp usando los sitios de restricción apropiados. El plásmido resultante se introduce por transfección en células COS de la manera descrita anteriormente y la expresión del polipéptido 16836 se detecta por radiomarcaje e inmunoprecipitación usando un anticuerpo monoclonal específico de 16836.

5

Equivalentes

Los especialistas en la técnica reconocerán o podrán averiguar sin usar nada más que una experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención descrita en la presente memoria. Estos equivalentes pretenden incluirse en las siguientes reivindicaciones.

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada seleccionada entre el grupo consistente en:

- a) una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que tienen una identidad de al menos 80% con la longitud entera de la secuencia de nucleótidos de la SEC ID N°: 1 o la SEC ID N°: 3;
- b) una molécula de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2; y
- c) una molécula de ácido nucleico que codifica un fragmento de un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2, donde el fragmento comprende al menos 1200 aminoácidos contiguos de la SEC ID N°: 2.

2. La molécula de ácido nucleico aislada de la reivindicación 1, que se selecciona entre el grupo consistente en:

- a) un ácido nucleico que comprende la secuencia de nucleótidos de la SEC ID N°: 1, SEC ID N°: 3; y
- b) una molécula de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2.

3. La molécula de ácido nucleico de la reivindicación 1, que comprende además secuencias de ácido nucleico de vectores.

4. La molécula de ácido nucleico de la reivindicación 1 que comprende además secuencias de ácido nucleico que codifican un polipéptido heterólogo.

5. Una célula hospedadora que contiene la molécula de ácido nucleico de la reivindicación 1.

6. La célula hospedadora de la reivindicación 5 que es una célula hospedadora de mamífero.

7. Una célula hospedadora de mamífero no humano que contiene la molécula de ácido nucleico de la reivindicación 1.

8. Un polipéptido aislado seleccionado entre el grupo consistente en:

- a) un polipéptido que se codifica por una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que tiene una identidad de al menos 80% con la longitud entera de la secuencia de nucleótidos de la SEC ID N°: 1 o la SEC ID N°: 3; y
- b) un fragmento de un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2, donde el fragmento comprende al menos 1200 aminoácidos contiguos de la SEC ID N°: 2.

9. El polipéptido aislado de la reivindicación 8 que comprende la secuencia de aminoácidos de la SEC ID N°: 2.

10. El polipéptido de la reivindicación 8 que comprende además secuencias de aminoácidos heterólogas.

11. Un anticuerpo o parte inmunológicamente activa del mismo, que se une selectivamente a un polipéptido seleccionado entre el grupo consistente en:

- a) un polipéptido que se codifica por una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que tienen una identidad de al menos 80% con la longitud entera de la secuencia de nucleótidos de la SEC ID N°: 1 o la SEC ID N°: 3; y
- b) un fragmento de un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2, donde el fragmento comprende al menos 1200 aminoácidos contiguos de la SEC ID N°: 2.

12. El anticuerpo o parte inmunológicamente activa del mismo de la reivindicación 11, donde el anticuerpo está marcado de forma detectable.

13. El anticuerpo o parte inmunológicamente activa del mismo de la reivindicación 12, donde el marcador detectable se selecciona entre el grupo consistente en:

- a) enzimas;
- b) grupos prostéticos;

ES 2 328 448 T3

- c) materiales fluorescentes;
- d) materiales luminiscentes;
- 5 e) materiales bioluminiscentes; y
- f) materiales radiactivos.

10 14. El anticuerpo o parte inmunológicamente activa del mismo de la reivindicación 11, donde el anticuerpo se selecciona entre el grupo consistente en:

- a) anticuerpo monoclonal;
- b) anticuerpo policlonal;
- 15 c) fragmento Fab;
- d) fragmento F(ab')₂;
- 20 e) anticuerpo quimérico;
- f) anticuerpo humanizado; y
- g) anticuerpo humano

25 15. Un método para producir un polipéptido seleccionado entre el grupo consistente en:

- a) un polipéptido que comprende la secuencia de aminoácidos de la SEC ID N°: 2; y
 - 30 b) un polipéptido que comprende un fragmento de la secuencia de aminoácidos de la SEC ID N°: 2, donde el fragmento comprende al menos 1200 aminoácidos contiguos de la SEC ID N°: 2;
- comprendiendo el método cultivar la célula hospedadora de la reivindicación 5 en condiciones en las que se exprese la molécula de ácido nucleico.

35 16. Un método para detectar la presencia de un polipéptido de la reivindicación 8 en una muestra, que comprende:

- a) poner en contacto la muestra con un compuesto que se une selectivamente a un polipéptido de la reivindicación 8; y
- 40 b) determinar si el compuesto se une al polipéptido en la muestra.

17. El método de la reivindicación 16, donde el compuesto que se une al polipéptido es un anticuerpo.

45 18. Un kit que comprende un anticuerpo que se une selectivamente a un polipéptido de la reivindicación 8 e instrucciones de uso.

19. Un método para detectar la presencia de una molécula de ácido nucleico de la reivindicación 1 en una muestra, que comprende las etapas de:

- 50 a) poner en contacto la muestra con una sonda de ácido nucleico o cebador que hibrida selectivamente con la molécula de ácido nucleico; y
- b) determinar si la sonda de ácido nucleico o cebador se une a una molécula de ácido nucleico en la muestra.

55 20. El método de la reivindicación 19, donde la muestra comprende moléculas de ARNm y se pone en contacto con una sonda de ácido nucleico.

60 21. Un kit que comprende una sonda de ácido nucleico o cebador que hibrida selectivamente con una molécula de ácido nucleico de la reivindicación 1 e instrucciones de uso.

22. Un método para identificar un compuesto que se une a un polipéptido de la reivindicación 8, que comprende las etapas de:

- 65 a) poner en contacto un polipéptido, o una célula que expresa un polipéptido de la reivindicación con un compuesto de ensayo; y
- b) determinar si el polipéptido se une al compuesto de ensayo.

ES 2 328 448 T3

23. El método de la reivindicación 22, donde la unión del compuesto de ensayo del polipéptido se detecta por un método seleccionado entre el grupo consistente en:

- a) detección de la unión por detección directa de la unión del compuesto de ensayo/polipéptido; y
- b) detección de unión usando un ensayo de unión competitiva.

24. Un método para modular la actividad de un polipéptido de la reivindicación 8, que comprende poner en contacto un polipéptido o una célula que expresa un polipéptido de la reivindicación 8 con un anticuerpo que se une al polipéptido en una concentración suficiente para modular la actividad del polipéptido.

25. Un método para identificar un agente que modula la actividad de un polipéptido de la reivindicación 8 que comprende:

- a) poner en contacto un polipéptido de la reivindicación 8 con un agente de ensayo; y
- b) determinar el efecto del compuesto de ensayo sobre la actividad del polipéptido para identificar de esta manera un agente que modula la actividad del polipéptido.

26. Uso de un agente que modula la actividad o expresión de un ácido nucleico de la reivindicación 1 o un polipéptido de la reivindicación 8 en la fabricación de un medicamento para el tratamiento o prevención de un trastorno de la proliferación o diferenciación celular **caracterizado** por una actividad o expresión aberrante de un ácido nucleico de la reivindicación 1 o un polipéptido de la reivindicación 8 en un sujeto, comprendiendo el tratamiento la administración al sujeto de una cantidad eficaz del agente que modula la actividad o expresión de un ácido nucleico de la reivindicación 1 o un polipéptido de la reivindicación 8 de tal forma que se mejore o prevenga el trastorno de proliferación o diferenciación celular, donde el agente se selecciona entre el grupo consistente en un polipéptido de la reivindicación 8, una molécula de ácido nucleico de la reivindicación 1, un anticuerpo, una ribozima y una molécula antisentido.

27. El uso de la reivindicación 26, donde el trastorno de la proliferación o diferenciación celular es cáncer de pulmón, cáncer de colon o cáncer de mama.

28. Un método *in vitro* para identificar un agente que modula la actividad o expresión de un polipéptido de la reivindicación 8 o una molécula de ácido nucleico de la reivindicación 1, comprendiendo el método:

- a) poner en contacto el polipéptido o ácido nucleico con un agente; y
- b) determinar el efecto del agente sobre la actividad o expresión del polipéptido o ácido nucleico; donde el agente se selecciona entre el grupo consistente en un polipéptido de la reivindicación 8, una molécula de ácido nucleico de la reivindicación 1, un anticuerpo, una ribozima y una molécula antisentido.

29. El método de la reivindicación 28, que comprende poner en contacto un polipéptido de la reivindicación 8 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido de catalizar la hidrólisis de fosfatidilinositol, o que comprende poner en contacto un polipéptido de la reivindicación 8 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido de asociarse con una proteína Ras, o que comprende poner en contacto un polipéptido de la reivindicación 8 con el agente y determinar el efecto del agente sobre la capacidad del polipéptido de mediar la actividad de intercambio de nucleótidos de guanina.

30. Un método *in vitro* para modular la actividad de una célula que expresa un polipéptido de la reivindicación 8 o una molécula de ácido nucleico de la reivindicación 1, que comprende poner en contacto la célula con una cantidad de un agente que modula la actividad o expresión de un ácido nucleico de la reivindicación 1 o un polipéptido de la reivindicación 8, de tal forma que se modula la actividad de la célula; donde el agente se selecciona entre el grupo consistente en un polipéptido de la reivindicación 8, una molécula de ácido nucleico de la reivindicación 1, un anticuerpo, una ribozima y una molécula antisentido.

ES 2 328 448 T3

inicio de la SEC ID N°:1

gtatatttatt	aagacagggg	caaaaacaca	tatactaatt	gtgaccatga	tattggccag	60
gtgtgggtcac	taaagaatgg	tgggcactgg	catccagtg	agaactcatg	ccccatctaa	120
cagtagaaga	tctagccatt	tgataccatg	taggaatgct	cagtgttagc	agattgtttc	180
attatttcagg	agatgctgga	acttcagatg	tatatgtaaa	gctttcctga	atttaaacag	240
ggatgtcag	atgaaacagg	tctgttgact	gaatctggcc	catggccaat	tgattttgtga	300
cctccctaca	ttaggcactt	ggtggcactt	ttctattgca	ttetaactgt	cctgaaggcc	360
cgagggtgaa	ggggagatgc	agaantaaggc	ggagatggaa	caatgagatg	ccaagagatg	420
aaagacttca	agaaggatgg	atgattgcct	atltcagaga	ttctgcagag	ggcmamagaa	480
atgggaatac	gmaaaaggcta	ttagattaaa	gtaaatatt	tattttggcac	agcactacgt	540
ctacaggctg	tgttctttta	agaaatgac	aaataacatg	ggaaatagtt	aagcaaatta	600
aaatactttg	tttactctga	ctctngtaat	tgtctttaat	attctattat	gcaatgtaaa	660
tctccaaaag	gagaagagat	ttactgagat	tttttaaact	ttcatgatca	aggagtccct	720
aggtcataca	atcccttcta	gatagaatgt	tttgaggaaa	acaccttggg	aaagctggat	780
aagagactga	atgagtcact	gggagtctat	ggaggcctga	tgtgggagact	ggaatgcaca	840
tggtagttag	caaatgcaca	atggccagc	gcagcagctt	ggcaagttaa	gaaagtggc	900
caacctatta	ggaactagga	atggccagct	tagtaatctt	tttaagacaa	aacctttta	960
gcattattcc	tttgctgaaa	acaaaccagc	gaaaacaaaa	acaaaaacaa	aacaaacaaa	1020
aaacaaaaaa	aagcatcaag	gggtgccctc	ctctgatggg	tgaaatctcc	ttgctttaac	1080
attccaggtc	tcctttatca	ggccagctga	cctttcttat	tcgccagttg	ttccctggcc	1140
tttctcattc	cccagttatt	tectatgctt	actccctgct	ctaaactcca	ctggctgttc	1200
ttattgcttg	catggtgagt	tgtctcccca	gtgctccag	ttagagaaat	gctgtctttc	1260
caaggcacag	gtcaaacctt	gctttgccta	ttaagccttc	cctagccate	aagctaccag	1320
tcagttcttt	tccttacaca	actggatcag	ttacttctga	aggacttgca	tgggacttag	1380
tgtgttgag	atggcctgta	tgatttggag	ccaagagaat	aaaactatgt	cttgagatct	1440
gcagctctct	acccccattt	gccctcacag	tacacccaaa	tagggttagg	ataatattgc	1500
ccagaaaac	agaagagcaa	gggcataact	ccaateacat	tgcatttcat	agattgtggt	1560
ctatagttat	acaaccacct	attattcatt	gagggtaaa	gagtgtattc	tcttttatag	1620
acttgcctaa	cagtaaaag	catcatttct	agacactagg	ctgagtatgt	tacatgttta	1680
tcattggtgaa	tcccagaatc	tcctaagaa	tagtgttatt	attccattt	tacagaggag	1740
ggagctgtct	tgaaacattc	aagtcagctt	gttcacactt	aatgaccagc	tctctatctg	1800
gtgaaggctg	agtcaggtgt	acctgacctc	atgtacctga	ccccaaacct	gggctctttc	1860
tgtcccagca	caccatctcc	atggccatgc	tggccaactag	aaatttaatg	acagagaatt	1920
aaaggtaaga	ctcacactgt	gccttcaagt	ggcaaatata	taccaccaga	cagaggaa	1980
cagataaaa	ctagttaggg	tctacacagt	ttgaatatca	taggactcgg	tgcattcttg	2040
ggatcataba	ggacotagca	cagctcaaga	gaggccctct	tgagtgagat	ttggaggatg	2100
caacatcatg	ggggcacatg	cctggcccgc	agggcagtg	gagtcttcta	ctccctggcc	2160
attcgtcagt	gtcaccgaag	gaggtgacct	tcatttccac	cctaactctaa	gctcttggct	2220
tttctgtctc	ctttactaga	ctgtgaactc	tgaggagaga	aactatttta	tttatccagc	2280
cctgtcatag	tttctgggat	ataaacagta	actatttgtt	ggtttgatga	atgaatgaca	2340
agtcagccaa	ggcaagggtc	tcaagggtgac	atgatagaga	cagcatgata	gagacagcac	2400
ctttattccc	ccatcagaga	gctggattac	ctcactgcat	gttatgtgtc	tcctccgttt	2460
gtttatacac	tatccagaag	tctctggctt	ctcagaatgc	tyacaaatt	ccaagtcctc	2520
ccttcaataa	tgtcaaacct	tataattttc	agccctcacc	gaagattgtg	aaatcctgca	2580
gcagcacaag	aggtgtctgc	tgccctcac	actttcccca	catcttccaa	cacactgcct	2640
tgcacagagc	aggtctotata	tacatgcaca	ttgaatgact	gcattgaaat	agattggaga	2700
catggttagg	aacctgtcct	aacagtgcaa	tcttcttgag	acaatgcttg	gtgaattcgg	2760
gtgataaaa	gagaagaaat	caggatagtg	acagtgacat	agatggggca	agaagagtga	2820
gcaggatctg	tcaacttatt	ggatattgga	ggtgagggg	gaggaagaga	aaaatgatt	2880
caaatacttt	gatctcagct	gacagtttg	aggtattgac	ccaaagggga	aataggggag	2940
ttctctttcc	ccatgacctt	ctcatctca	tttcagttta	agcatggaga	aaagatttcc	3000
cataaaaagtc	aggctgaaac	actccatgga	gggagacaga	cctgggacag	tatttcagta	3060
atgtccagga	aaaccagaag	gaaagggttg	ccttcatctc	tctgcagaag	agctttgtct	3120
tgctgaggag	gagaaatttg	agctgaatgt	atctctgggt	ggatgactt	gtaagttctc	3180
tgtttgatag	cagagttggt	tgaatgacct	atcttataac	tacacagagc	ttttccaaat	3240
agatggtgct	tgtagaatgt	tcaggttaat	gcattgacct	agtgtcaga	gtgtttgcac	3300
ttggagcctc	tgagccttct	ttcttaataa	gtcatcttgg	cattttgtca	acaagttaat	3360
ttgctttct	aagaaactag	attaataact	cctgtacttt	gtgctattca	cagatctaca	3420

Fig. 1A

ES 2 328 448 T3

ctg cac atc cct ggc tgt aag gtg gtt caa ggg tgt ggg gtg ttt ctg	4337
Leu His Ile Pro Gly Cys Lys Val Val Gln Gly Cys Gly Val Phe Leu	
205 210 215	
aag gag ctc tgt gaa gtg ctt gac ggc gcc tcc ggt ctc atg aag ctt	4385
Lys Glu Leu Cys Glu Val Leu Asp Gly Ala Ser Gly Leu Met Lys Leu	
220 225 230	
tgc ccg cgg tcc aat tcc caa gaa gaa act tta gag ttt gta gca gat	4433
Cys Pro Arg Ser Asn Ser Gln Glu Glu Thr Leu Glu Phe Val Ala Asp	
235 240 245	
tac agt gga caa gat aat ttc tta caa cga gtg gga caa aat ggc tta	4481
Tyr Ser Gly Gln Asp Asn Phe Leu Gln Arg Val Gly Gln Asn Gly Leu	
250 255 260 265	
aag aat tcg gag aag gag tcc act gtc aac agc atc ttt cag gtc atc	4529
Lys Asn Ser Glu Lys Ser Thr Val Asn Ser Ile Phe Gln Val Ile	
270 275 280	
cgg agc tgc aat cga agt ctg gag aca gac gag gag gac agc ccc agt	4577
Arg Ser Cys Asn Arg Ser Leu Glu Thr Asp Glu Glu Asp Ser Pro Ser	
285 290 295	
gaa gga aac agc tcc agg aaa agc tcc ttg aag gat aaa agc cga tgg	4625
Glu Gly Asn Ser Ser Arg Lys Ser Ser Leu Lys Asp Lys Ser Arg Trp	
300 305 310	
cag ttt ata att gga gat ttg ttg gat tca gac aat gac atc ttt gag	4673
Gln Phe Ile Ile Gly Asp Leu Leu Asp Ser Asp Asn Asp Ile Phe Glu	
315 320 325	
caa tcc aaa gaa tac gac tct cat ggt tca gag gac tca cag aag gcc	4721
Gln Ser Lys Glu Tyr Asp Ser His Gly Ser Glu Asp Ser Gln Lys Ala	
330 335 340 345	
ttc gac cat ggg acg gag ctc atc cct tgg tac gtg ctg tcc atc caa	4769
Phe Asp His Gly Thr Glu Leu Ile Pro Trp Tyr Val Leu Ser Ile Gln	
350 355 360	
gcc gat gtg cac cag ttc ctg ctg cag ggg gcc acg gtc atc cac tac	4817
Ala Asp Val His Gln Phe Leu Leu Gln Gly Ala Thr Val Ile His Tyr	
365 370 375	
gac cag gac aca cac ctc tct gcc cgc tgc ttc ctc cag ctt cag ccc	4865
Asp Gln Asp Thr His Leu Ser Ala Arg Cys Phe Leu Gln Leu Gln Pro	
380 385 390	
gac aat agc acc ttg acc tgg gta aag ccc aca act gcc tcc cca gcc	4913
Asp Asn Ser Thr Leu Thr Trp Val Lys Pro Thr Thr Ala Ser Pro Ala	
395 400 405	
agc agt aaa gca aaa ctt ggt gta ctt aat aac aca gct gag cct gga	4961
Ser Ser Lys Ala Lys Leu Gly Val Leu Asn Asn Thr Ala Glu Pro Gly	
410 415 420 425	

Fig. 1C

ES 2 328 448 T3

aaa ttc cca cta ctg ggt aat gct gga tta agt agc ctg acg gaa ggg	5009
Lys Phe Pro Leu Leu Gly Asn Ala Gly Leu Ser Ser Leu Thr Glu Gly	
430 435 440	
gtc ttg gat ctt ttt gca gtg aag gct gta tac atg ggc cac cct ggc	5057
Val Leu Asp Leu Phe Ala Val Lys Ala Val Tyr Met Gly His Pro Gly	
445 450 455	
att gat ata cac act gtg tgt gtt cag aac aaa ctg ggt agc atg ttc	5105
Ile Asp Ile His Thr Val Cys Val Gln Asn Lys Leu Gly Ser Met Phe	
460 465 470	
ctg tca gag act ggt gtg aca ttg ctc tat ggg ctt cag acc aca gac	5153
Leu Ser Glu Thr Gly Val Thr Leu Leu Tyr Gly Leu Gln Thr Thr Asp	
475 480 485	
aac aga tta ttg cac ttc gtg gca cca aag cac aca gct aaa atg ctc	5201
Asn Arg Leu Leu His Phe Val Ala Pro Lys His Thr Ala Lys Met Leu	
490 495 500 505	
ttc agc gga tta ttg gaa ctc act aga gct gtg aga aag atg agg aaa	5249
Phe Ser Gly Leu Leu Glu Leu Thr Arg Ala Val Arg Lys Met Arg Lys	
510 515 520	
ttc cct gac caa aga cag cag tgg ctg cgg aaa cag tac gtc agc ctt	5297
Phe Pro Asp Gln Arg Gln Gln Trp Leu Arg Lys Gln Tyr Val Ser Leu	
525 530 535	
tat cag gag gat gga cgg tat gaa ggc cca act ttg gct cac gct gtg	5345
Tyr Gln Glu Asp Gly Arg Tyr Glu Gly Pro Thr Leu Ala His Ala Val	
540 545 550	
gag ttg ttt ggt ggc aga cgg tgg agt gct cga aac ccc agc ccc gga	5393
Glu Leu Phe Gly Gly Arg Arg Trp Ser Ala Arg Asn Pro Ser Pro Gly	
555 560 565	
aca tca gca aag aat gct gag aag ccc aat atg cag aga aac aat acc	5441
Thr Ser Ala Lys Asn Ala Glu Lys Pro Asn Met Gln Arg Asn Asn Thr	
570 575 580 585	
ctg ggc ata agc act acc aag aaa aag aag aaa atc ctc atg agg ggt	5489
Leu Gly Ile Ser Thr Thr Lys Lys Lys Lys Lys Ile Leu Met Arg Gly	
590 595 600	
gag agt gga gag gta act gac gat gag atg gca acc cga aag gcc aag	5537
Glu Ser Gly Glu Val Thr Asp Asp Glu Met Ala Thr Arg Lys Ala Lys	
605 610 615	
atg cac aaa gag tgt cga agc cgg agt ggt tct gat cct caa gac att	5595
Met His Lys Glu Cys Arg Ser Arg Ser Gly Ser Asp Pro Gln Asp Ile	
620 625 630	
aat gaa caa gaa gaa tca gag gtg aat gcc atc gct aac cct cca aac	5633
Asn Glu Gln Glu Glu Ser Glu Val Asn Ala Ile Ala Asn Pro Pro Asn	
635 640 645	

Fig. 1D

ES 2 328 448 T3

ccc ctc cct tcc aga aga gcc cac tct ttg acc aca gct ggg tcc ccc Pro Leu Pro Ser Arg Arg Ala His Ser Leu Thr Thr Ala Gly Ser Pro 650 655 660 665	5681
aac ttg gct gcc ggg acg tca tct ccc atc agg cca gtg tcc tcc cct Asn Leu Ala Ala Gly Thr Ser Ser Pro Ile Arg Pro Val Ser Ser Pro 670 675 680	5729
gtg ctg tct tct tca aac aag agc cca tcc agt gct tgg agc agt agt Val Leu Ser Ser Ser Asn Lys Ser Pro Ser Ser Ala Trp Ser Ser 685 690 695	5777
agc tgg cac ggg cgg atc aaa ggc ggc atg aag gga ttt cag agc ttc Ser Trp His Gly Arg Ile Lys Gly Gly Met Lys Gly Phe Gln Ser Phe 700 705 710	5825
atg gtt tca gat agc aac atg agt ttt gtt gaa ttt gtt gag ctg ttc Met Val Ser Asp Ser Asn Met Ser Phe Val Glu Phe Val Glu Leu Phe 715 720 725	5873
aaa tca ttc agt gtc agg agc cgc aag gac ctg aag gat ctg ttt gat Lys Ser Phe Ser Val Arg Ser Arg Lys Asp Leu Lys Asp Leu Phe Asp 730 735 740 745	5921
gtc tat gca gtg ccc tgc aac cga tct ggc tcc gag tca gcc cca ctc Val Tyr Ala Val Pro Cys Asn Arg Ser Gly Ser Glu Ser Ala Pro Leu 750 755 760	5969
tac acc aac ctg aca att gat gaa aac acc agc gat ctt cag cct gac Tyr Thr Asn Leu Thr Ile Asp Glu Asn Thr Ser Asp Leu Gln Pro Asp 765 770 775	6017
cta gat ctg ttg acc aga aat gtc tcg gat ttg ggg ttg ttc att aag Leu Asp Leu Leu Thr Arg Asn Val Ser Asp Leu Gly Leu Phe Ile Lys 780 785 790	6065
agt aaa cag cag cta tgg gac aac cag agg cag ata tct gat gcc att Ser Lys Gln Gln Leu Ser Asp Asn Gln Arg Gln Ile Ser Asp Ala Ile 795 800 805	6113
gct gct gca agc att gtg aca aat ggc act ggg att gag agc aca tct Ala Ala Ala Ser Ile Val Thr Asn Gly Thr Gly Ile Glu Ser Thr Ser 810 815 820 825	6161
ctg ggc att ttt ggg gtg ggc ata ctt cag ctc aac gat ttc ctc gtg Leu Gly Ile Phe Gly Val Gly Ile Leu Gln Leu Asn Asp Phe Leu Val 830 835 840	6209
aat tgc caa gga gaa cac tgc act tat gat gaa atc ctc agc atc atc Asn Cys Gln Gly Glu His Cys Thr Tyr Asp Glu Ile Leu Ser Ile Ile 845 850 855	6257
cag aag ttc gag cct agc atc agt atg tgt cat cag gga cta atg tca Gln Lys Phe Glu Pro Ser Ile Ser Met Cys His Gln Gly Leu Met Ser 860 865 870	6305

Fig. 1E

ES 2 328 448 T3

ttt gaa ggg ttt gcc agg ttt ctg atg gat aaa gaa aat ttt gcc tca Phe Glu Gly Phe Ala Arg Phe Leu Met Asp Lys Glu Asn Phe Ala Ser 875 880 885	6353
aaa aat gat gag tca cag gag aac att aaa gaa ctg cag cta ccc ctc Lys Asn Asp Glu Ser Gln Glu Asn Ile Lys Glu Leu Gln Leu Pro Leu 890 895 900 905	6401
tca tac tat tac atc gaa tct tcg cac aat acc tac ctc acg ggc cat Ser Tyr Tyr Tyr Ile Glu Ser Ser His Asn Thr Tyr Leu Thr Gly His 910 915 920	6449
cag ctc aaa gga gaa tcc tcg gta gaa ctc tac agc cag gtc ctt ttg Gln Leu Lys Gly Glu Ser Ser Val Glu Leu Tyr Ser Gln Val Leu Leu 925 930 935	6497
caa ggc tgt cga agt gta gaa ttg gac tgc tgg gac gga gac gat ggg Gln Gly Cys Arg Ser Val Glu Leu Asp Cys Trp Asp Gly Asp Asp Gly 940 945 950	6545
atg ccc atc att tat cat gga cat acg ctg aca acc aag atc ccc ttc Met Pro Ile Ile Tyr His Gly His Thr Leu Thr Thr Lys Ile Pro Phe 955 960 965	6593
aag gaa gtg gtt gaa gcc att gat cgc agt gcc ttc atc aac tct gac Lys Glu Val Val Glu Ala Ile Asp Arg Ser Ala Phe Ile Asn Ser Asp 970 975 980 985	6641
ctg cca atc atc ata tcg att gag aac cac tgt tca ttg oct cag caa Leu Pro Ile Ile Ile Ser Ile Glu Asn His Cys Ser Leu Pro Gln Gln 990 995 1000	6689
cga aaa atg gca gaa att ttc aag act gtg ttt gga gaa aag ctg gtg Arg Lys Met Ala Glu Ile Phe Lys Thr Val Phe Gly Glu Lys Leu Val 1005 1010 1015	6737
act aaa ttc tta ttt gag act gat ttc tca gat gat cca atg ctt cct Thr Lys Phe Leu Phe Glu Thr Asp Phe Ser Asp Asp Pro Met Leu Pro 1020 1025 1030	6785
tca cct gac caa ctc aga aag aaa gtt ctt ctt aaa aac aag aag cta Ser Pro Asp Gln Leu Arg Lys Lys Val Leu Leu Lys Asn Lys Lys Leu 1035 1040 1045	6833
aaa gcc cat cag acg cca gtg gat atc tta aag caa aag gct cat cag Lys Ala His Gln Thr Pro Val Asp Ile Leu Lys Gln Lys Ala His Gln 1050 1055 1060 1065	6881
tta gca tct atg caa gtg cag gct tat aat ggt ggg aat gcc aac ccc Leu Ala Ser Met Gln Val Gln Ala Tyr Asn Gly Gly Asn Ala Asn Pro 1070 1075 1080	6929
cga cct gcc aat aat gag gaa gag gaa gat gag gag gac gaa tat gat Arg Pro Ala Asn Asn Glu Glu Glu Glu Asp Glu Glu Asp Glu Tyr Asp 1085 1090 1095	6977

Fig. 1F

ES 2 328 448 T3

tat gac tat gaa tcc ctt tct gat gac aac att ctg gaa gac aga cct Tyr Asp Tyr Glu Ser Leu Ser Asp Asp Asn Ile Leu Glu Asp Arg Pro 1100 1105 1110	7025
gaa aat aaa tca tgt aat gac aag ctt cag ttt gaa tat aat gaa gaa Glu Asn Lys Ser Cys Asn Asp Lys Leu Gln Phe Glu Tyr Asn Glu Glu 1115 1120 1125	7073
atc cca aag agg ata aag aaa gca gat aac tct gct tgc aac aaa gga Ile Pro Lys Arg Ile Lys Lys Ala Asp Asn Ser Ala Cys Asn Lys Gly 1130 1135 1140 1145	7121
aag gtt tat gat atg gaa ctg gga gaa gaa ttt tat ctt gat cag aat Lys Val Tyr Asp Met Glu Leu Gly Glu Glu Phe Tyr Leu Asp Gln Asn 1150 1155 1160	7169
aaa aag gaa agc aga cag att gca cca gag ctt tct gac ctt gta atc Lys Lys Glu Ser Arg Gln Ile Ala Pro Glu Leu Ser Asp Leu Val Ile 1165 1170 1175	7217
tat tgt caa gca gta aaa ttt cca gga ctg tca act cta aat gca tct Tyr Cys Gln Ala Val Lys Phe Pro Gly Leu Ser Thr Leu Asn Ala Ser 1180 1185 1190	7265
ggc tct agc aga gga aaa gaa agg aaa agc agg aag tcc att ttt ggc Gly Ser Ser Arg Gly Lys Glu Arg Lys Ser Arg Lys Ser Ile Phe Gly 1195 1200 1205	7313
aac aat ccg ggc aga atg agc cca ggg gag aca gca tca ttt aac aaa Asn Asn Pro Gly Arg Met Ser Pro Gly Glu Thr Ala Ser Phe Asn Lys 1210 1215 1220 1225	7361
aca tct gga aaa agt tcc tgt gaa ggc att cga cag acc tgg gag gaa Thr Ser Gly Lys Ser Ser Cys Glu Gly Ile Arg Gln Thr Trp Glu Glu 1230 1235 1240	7409
tct tct tcc cct ctc aac cca acc acg tcc ctc agt gct atc att aga Ser Ser Ser Pro Leu Asn Pro Thr Thr Ser Leu Ser Ala Ile Ile Arg 1245 1250 1255	7457
act ccc aaa tgt tat cat atc tcg tcg ctg aat gaa aat gcc gcc aaa Thr Pro Lys Cys Tyr His Ile Ser Ser Leu Asn Glu Asn Ala Ala Lys 1260 1265 1270	7505
cgt ctg tgt cgc agg tat tct cag aaa ctg acc cag cac acc gcc tgt Arg Leu Cys Arg Arg Tyr Ser Gln Lys Leu Thr Gln His Thr Ala Cys 1275 1280 1285	7553
cag ctg ctg aga act tac cct gct gcc acc cgc atc gac tct tcc aac Gln Leu Leu Arg Thr Tyr Pro Ala Ala Thr Arg Ile Asp Ser Ser Asn 1290 1295 1300 1305	7601
ccg aac ccc ctc atg ttc tgg ctc cat ggg ata cag ctt gtg gca ctc Pro Asn Pro Leu Met Phe Trp Leu His Gly Ile Gln Leu Val Ala Leu 1310 1315 1320	7649

Fig. 1G

ES 2 328 448 T3

aag cag ctt ctg cag caa att ctg aca aat gaa caa gac atc aaa cct Lys Gln Leu Leu Gln Gln Ile Leu Thr Asn Glu Gln Asp Ile Lys Pro 1550 1555 1560	8369
ggt acc aca gac tat ttt ttg atg gaa gaa aaa tat ttt ata tct aaa Val Thr Thr Asp Tyr Phe Leu Met Glu Glu Lys Tyr Phe Ile Ser Lys 1565 1570 1575	8417
gaa aag aat gaa tgt agg aaa caa cca ttc cag aga gcc att ggt cca Glu Lys Asn Glu Cys Arg Lys Gln Pro Phe Gln Arg Ala Ile Gly Pro 1580 1585 1590	8465
gaa gag gag atc atg caa att tta agc agc tgg ttt cca gaa gag gga Glu Glu Glu Ile Met Gln Ile Leu Ser Ser Trp Phe Pro Glu Glu Gly 1595 1600 1605	8513
tac atg ggc agg att gtc tta aaa acc cag cag gaa aac cta gaa gag Tyr Met Gly Arg Ile Val Leu Lys Thr Gln Gln Glu Asn Leu Glu Glu 1610 1615 1620 1625	8561
aaa aac att gtt caa gat gac aaa gag gtg atc ttg agc tca gag gag Lys Asn Ile Val Gln Asp Asp Lys Glu Val Ile Leu Ser Ser Glu Glu 1630 1635 1640	8609
gag agt ttc ttt gtc caa gtg cat gat gtt tct cca gag caa cct cga Glu Ser Phe Phe Val Gln Val His Asp Val Ser Pro Glu Gln Pro Arg 1645 1650 1655	8657
aca gtc atc aaa gca ccc cgc gtc agc act gca cag gat gtc att cag Thr Val Ile Lys Ala Pro Arg Val Ser Thr Ala Gln Asp Val Ile Gln 1660 1665 1670	8705
cag acc tta tgc aaa gcc aaa tat tcc tac agc atc ctg agc aac ccc Gln Thr Leu Cys Lys Ala Lys Tyr Ser Tyr Ser Ile Leu Ser Asn Pro 1675 1680 1685	8753
aat cca agc gac tat gtg ctt ttg gaa gag gtg gtg aaa gac act acc Asn Pro Ser Asp Tyr Val Leu Leu Glu Glu Val Val Lys Asp Thr Thr 1690 1695 1700 1705	8801
aac aag aag act acc aca cca aag tcc tct cag cgg gtc ctt ctg gat Asn Lys Lys Thr Thr Thr Pro Lys Ser Ser Gln Arg Val Leu Leu Asp 1710 1715 1720	8849
cag gag tgt gtg ttt caa gcc caa agc aag tgg aaa ggt gca gga aaa Gln Glu Cys Val Phe Gln Ala Gln Ser Lys Trp Lys Gly Ala Gly Lys 1725 1730 1735	8897
ttc atc ctt aag cta aag gag cag gtg cag gca tct cga gaa gat aaa Phe Ile Leu Lys Leu Lys Glu Gln Val Gln Ala Ser Arg Glu Asp Lys 1740 1745 1750	8945
aag aaa ggc att tct ttc gca agt gaa ctc aag aag ctc acc aag tca Lys Lys Gly Ile Ser Phe Ala Ser Glu Leu Lys Lys Leu Thr Lys Ser 1755 1760 1765	8993

Fig. 11

ES 2 328 448 T3

```

act aaa cag ccc cga gga ctt aca tca cct tet cag etc ttg acc tca      9041
Thr Lys Gln Pro Arg Gly Leu Thr Ser Pro Ser Gln Leu Leu Thr Ser
1770                               1775                               1780                               1785

gaa agt atc caa acc aag gag gag aaa cct gtg ggt ggc ttg tcc tcc      9089
Glu Ser Ile Gln Thr Lys Glu Glu Lys Pro Val Gly Gly Leu Ser Ser
                               1790                               1795                               1800

agt gac aca atg gat tac cga cag tgactaaggg cagcatgttt aaccagggtg      9143
Ser Asp Thr Met Asp Tyr Arg Gln
                               1805
                               ↖ final de la SEC ID N°: 3
                               ↖ final de la SEC ID N°: 2
aagatcttta agcaagaagt taaagagtga acatggtgga aaaaatataa ttattttcat      9203
cagactraaa ctggaattg atgatttctg aactgaagcc ttcacacatg tgagatccat      9263
gctgaggaga agcaaaatgg cacagggcta gttgccacca accaatttac tgatgaatga      9323
agccagggg actgccattt tataaatgtc agcagttgga aaaatcgtca cgaattgact      9383
tagagcaagg gtcagcaagc ttgtctgtaa agggccaaac agtaaatatt ttagggctgg      9443
gggccatasa atatgttgca accacccaat tetgccattg tagtgcaaaa gcagccatag      9503
acaacacara catgaacgaa cgtggctgta ttccaataaa actttattta tggacactga      9563
aaaaaaaaaa aaaaagggcg gccgctcgag totagagggc mcrtwaaac ccgctsathc      9623
cccyccacgw gcygcmwca wadaatgcag ccatygtgtg tttkcccytc yyccgtgect      9683
tccttgaycc tgrmaartga agcactcccw gwwaarttc acwaataaaa tgwkaawty      9743
gcatcgttct rgtttgrgta ggtchcawty tattottgkg kettatcatg tctggartcc      9803
ccgggtaccg agctcgaatt aattcctctt ccgcttctc ygcctactga ctgctgccc      9863
tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggtawtc      9923
cacagaatca ggggartaac gcaggaaga acatgtgagc aaaaggccag caaaaggcca      9983
ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag gctcggccc cctgacgagc      10043
atcacaaaaa tcgacgctca agtcagaggt ggcgaaacc gacaggacta taaagatacc      10103
caggcgtttc cccctggaag ctccctcctg cgctctcctg ttcgaccct gccgcttacc      10163
ggatacctg
                               ↖ final de la SEC ID N°: 1

```

Fig. 1J

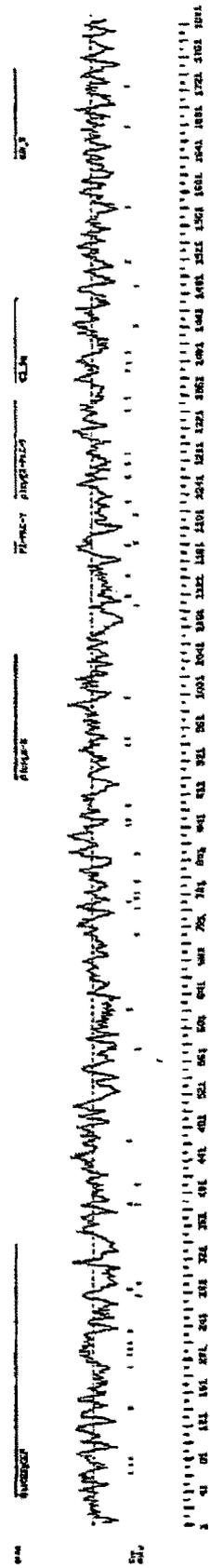


Fig. 2

ES 2 328 448 T3

RasGEF_3: dominio 1 de 1, de 35 a 338: puntuación 17,0, E = 1,7e-07
 SEC ID Nº: 4 *->l1llldpeelAeQITl1ldfelFrkIepsEllgsvwgkrskkapsplap
 l+++ peetA L+ + ++ ++ p ++l ++ + ++++ + p
 16836 35 LIQF-PEEVASILMEQEQTIIYLRVLPVDYLCFLTRDLGTPECQSSLP 80
qnleafierfNevsnwVateIlkqttlkpkR
 + + + + ++++++ le+++ rfNevs+wV+ Il++ + ++
 16836 81 cikasisasilttqngehNALEDLVMRFNEVSSVWTWLILTAGS--MEEK 128
 aeVlskFIkVakhcreLnNFnSlmAivsAlssspIsRLkkTWeklsskyk
 ev+s+ ++Vak c++++N+n +m+ ++L s + + ++ +
 16836 129 REVFSYLVHVAKCCWNMGNYNAVMEFLAGLRSRKVLKMWQFMDQSD---I 175
 klfeeLeelldpserNfknYReaLkscnksp..nvqppcvPflGvyLkDL
 ++ +L++ + ++ ++YR+++ + + p+ +v ++c Gv+Lk L
 16836 176 ETMRSLKDAMAQH-ESSCEYRKVVTRALHIPgcKVVQGC----GVFLKEL 220
 tfidegnpdfleht.....kglvNFaK
 + +g + + ++++++++ + + +++++ ++ +tgL N eK
 16836 221 CEVLDGASGLMKLCPrsnsqeetlefvadysgqdnflqrvqgNGLKNSEK 270
 rriakiilreirqlqsacqpynlkpnr.....ndiqellra
 + i + ir + + ++ ++ ++++++++ +++++ ++
 16836 271 ESTVNSIFQVIRSCNR--SLETDEEDSpsegnsrkslkdKSRWQFIIG 318
 srplevlpeeedelyelSlriEPrv<-*
 + l++++ ++e+S + +
 16836 319 ----DLLSDN-DIFEQSKYDShG 338

Fig. 3A

PI-PLC-X: dominio 1 de 1, de 900 a 1048: puntuación 243,2, E = 3,5e-69
 SEC ID Nº: 5 *->dmsiPLsHYfisSshntYltgkQlwGkssvesYrqqLdaGcRcvELD
 ++ PLs+Y+i SshntYltg+Ql+G+ssve Y q+L++GcR vELD
 16836 900 ELQLPLSYYYIESSHNTYLTGHQLKGESSVELYSQVLLQGCRSVELD 946
 cwdGkpddepiIyHGhtltleikikdVleaIkdfafkPtSpyPvIlSlen
 cwdG d++piIyHGhtlt++i++k+V+eaI af +S P+I+S+en
 16836 947 CWDG-DDGMPITTYHGHTLTKIPFKEVVEAIDRSAFI-NSDLPITIIISIEN 994
 HcnsddqQrkmakyfkeiFgdmLtkPtldslttepqlpLpSlkdrlrgKI
 Hc+ + qQrkma++fk++Fg++L+tk l+++ + Lps+++lr K
 16836 995 HCSLP-QQRKMAEIFKTVEGKLVTKF-LFETDFSDDPMLPSPDQLRKKV 1042
 LLknkk<-*
 LLknkk
 16836 1043 LLKNKK 1048

Fig. 3B

ES 2 328 448 T3

PI-PLC-Y: dominio 1 de 2, de 1171 a 1184: puntuación 13,9, E = 0,012
 SEC ID N° 6 *->ElsnLvnYiqsikF<-*

```

    Els+Lv+Y+q +kF
16836 1171  ELSDLVIYCQAVKF 1184
  
```

Fig. 3C

PI-PLC-Y: dominio 2 de 2, de 1261 a 1353: puntuación 141,5, E = 1,4e-38

```

    SEC ID N° 7 *->syeisSFsErkvkakkllkespvefVkyNkrqLsRvYPkGtRvDSSN
    +y+isS++E+ +ak+l + + ++ qL+R YP+ tR DSSN
16836 1261  CYHISSLNEN---AAKRLCRRYSQKLTQHTACQLLRTYPAATRIDSSN 1305
  
```

```

    fmPqvfwNaGCCQmVALNfQtsDlpmqiNdGmFeyNggqPdGsfsksGYLLK
    +P +fW G Q+VALN+QT Dlp+ +N +mFe+Ngg +GY+LK
16836 1306  PNPMLFWLHGIQLVALNYQTDDLPLHLNAAMFEANGG-----CGYVLK 1348
  
```

```

    PeflR<-*
    P+ l
16836 1349  PPVLW 1353
  
```

Fig. 3D

C2: dominio 1 de 1, de 1378 a 1460: puntuación 39,8, E = 6,1e-08

```

    SEC ID N° 8 *->LtVtvieArnLpkmDkvngrlsDPYVksllgdkkdlkkfkTk.vvk
    + t+ ++n + +n+ P ++v++lg++ d f+Tk++++
16836 1378  YSLTIVSQNVCP---SNS-MGSPCIEVDVLGMPLDSCHEFTKpIHR 1420
  
```

```

    ktNGLNpWneEtFvFekvplpelasktLrfaVyDedfrsrdDfiGqvt<
    +t LNP Wn E+F F v+++t+ l + LrfaV + ++ + +q++
16836 1421  NT--LNPWN-EQFLFH-VHFEDLVF--LRFVAVENNSA---VTAQRI 1460
  
```

```

    -*
16836 - -
  
```

Fig. 3E

RA: dominio 1 de 1, de 1460 a 1745: puntuación 3,0, E = 0,085

```

    SEC ID N° 9 *->dqgvlrvyfqdllsvtpgvayKtIrvssedtapdVigeaLeKfr...
    +++ v++ d++ p ++ i++ +ta+dViq L K + +
16836 1640  EEESFFVQVHDVS---PEQPRTVIKAPRVSTAQDVIQQLCKAKysy 1683
  
```

```

    .lddRMedpeeYaLvevlltr.ega..lesggkerkLpddenPlqlrln
    + l + ++p +Y+L e+ + ++ ++ +r+L dte+++q++
16836 1684  siLSN--PNPSDYVLLLEEVDKdtTNKktTTPKSSQRVLLDQECVFOAQS 1731
  
```

```

    lprddrrsvrqsslrFlLkrdd<-*
    ++ F+lK +++
16836 1732  WKGA-----GKFILKLEQ 1745
  
```

Fig. 3F

expresión de 16836 en tejidos humanos normales

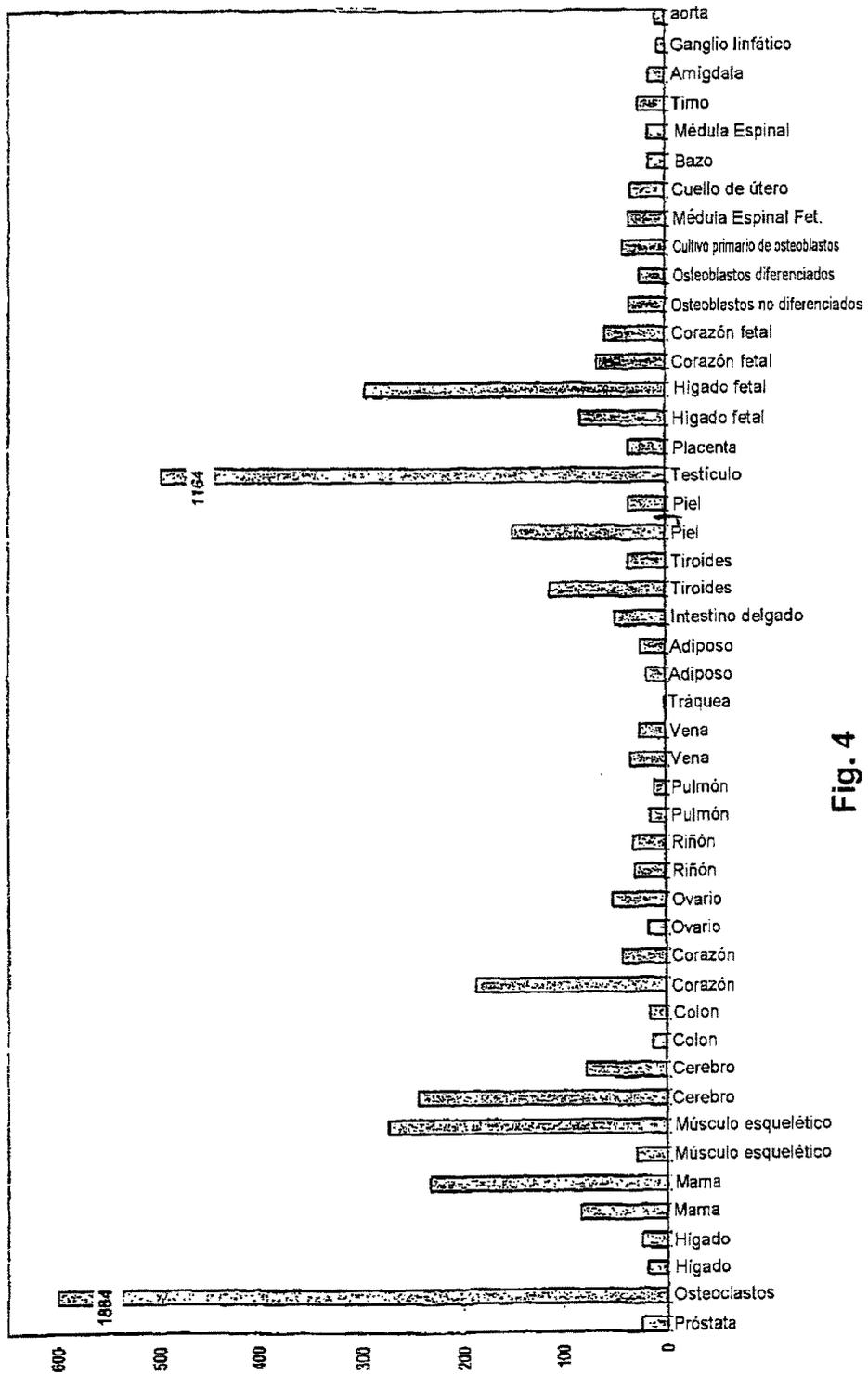


Fig. 4

Expresión de 16836 en Pulmón

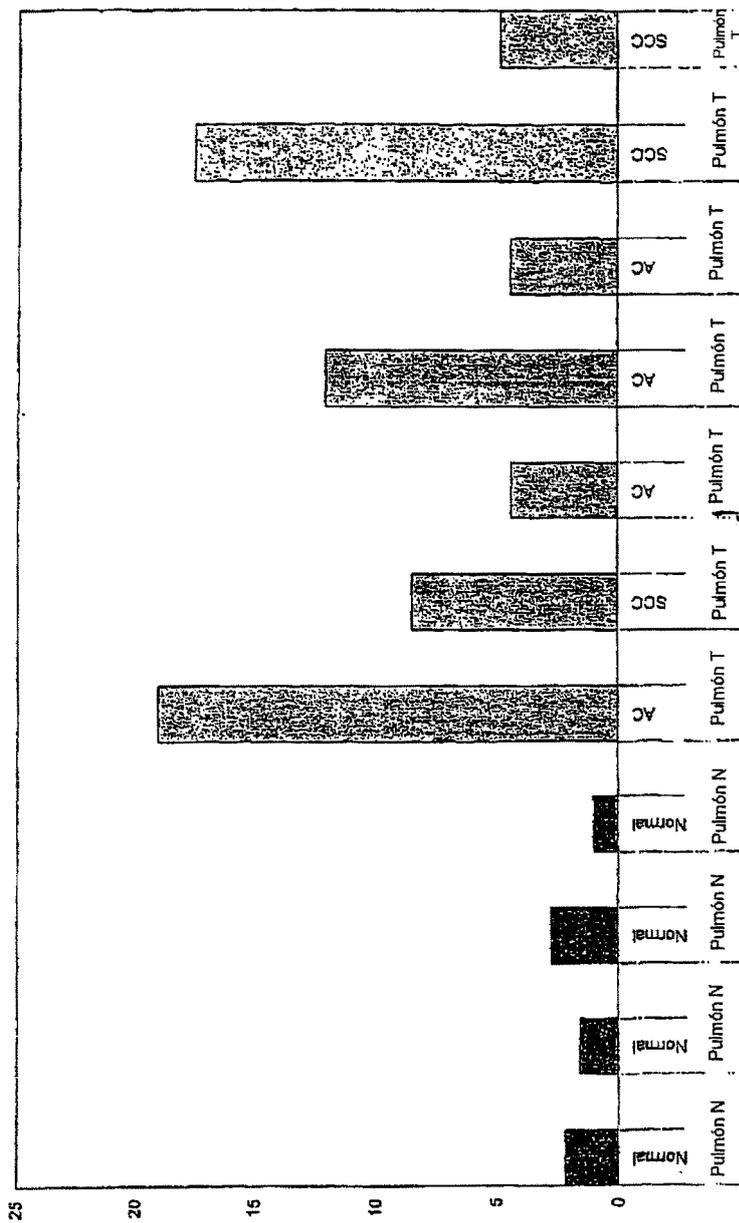


Fig. 5

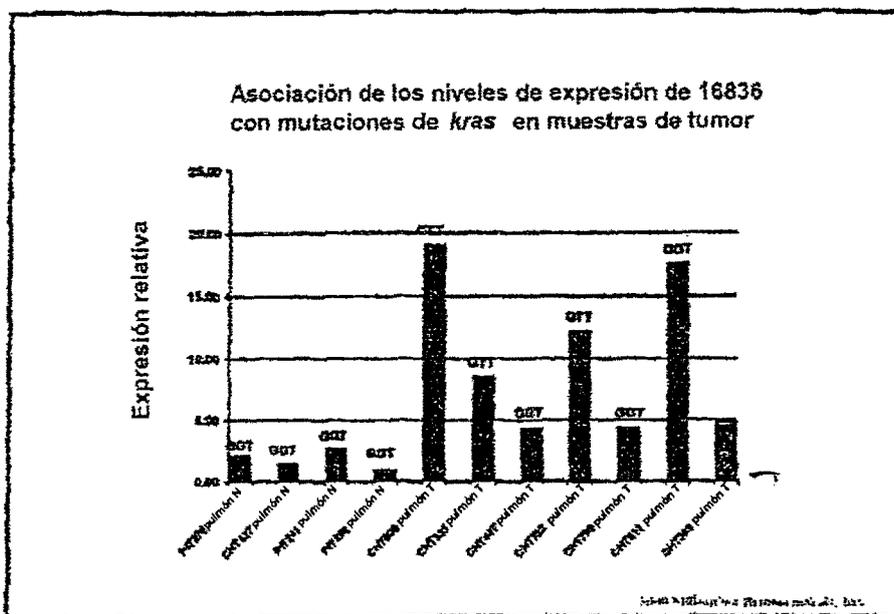


Fig. 6A

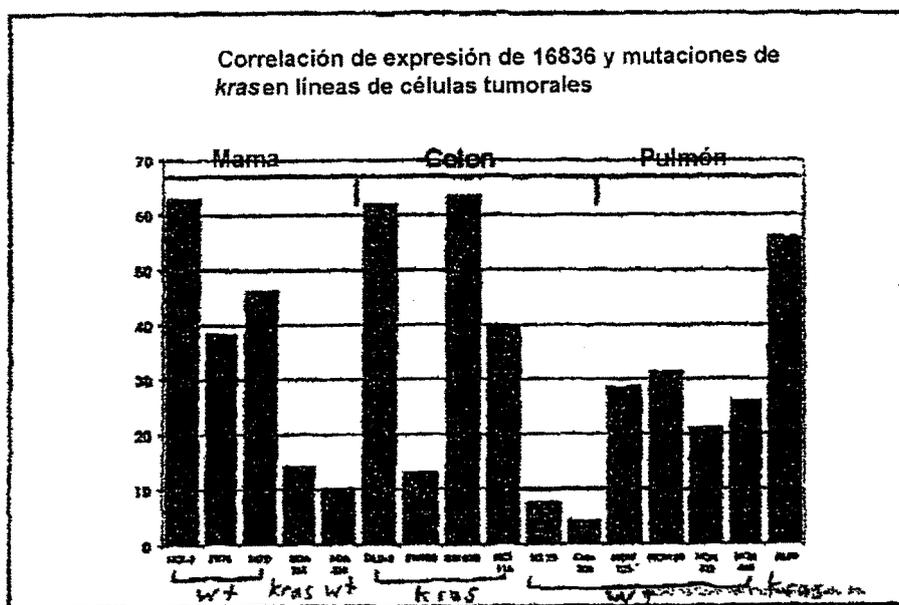


Fig. 6B

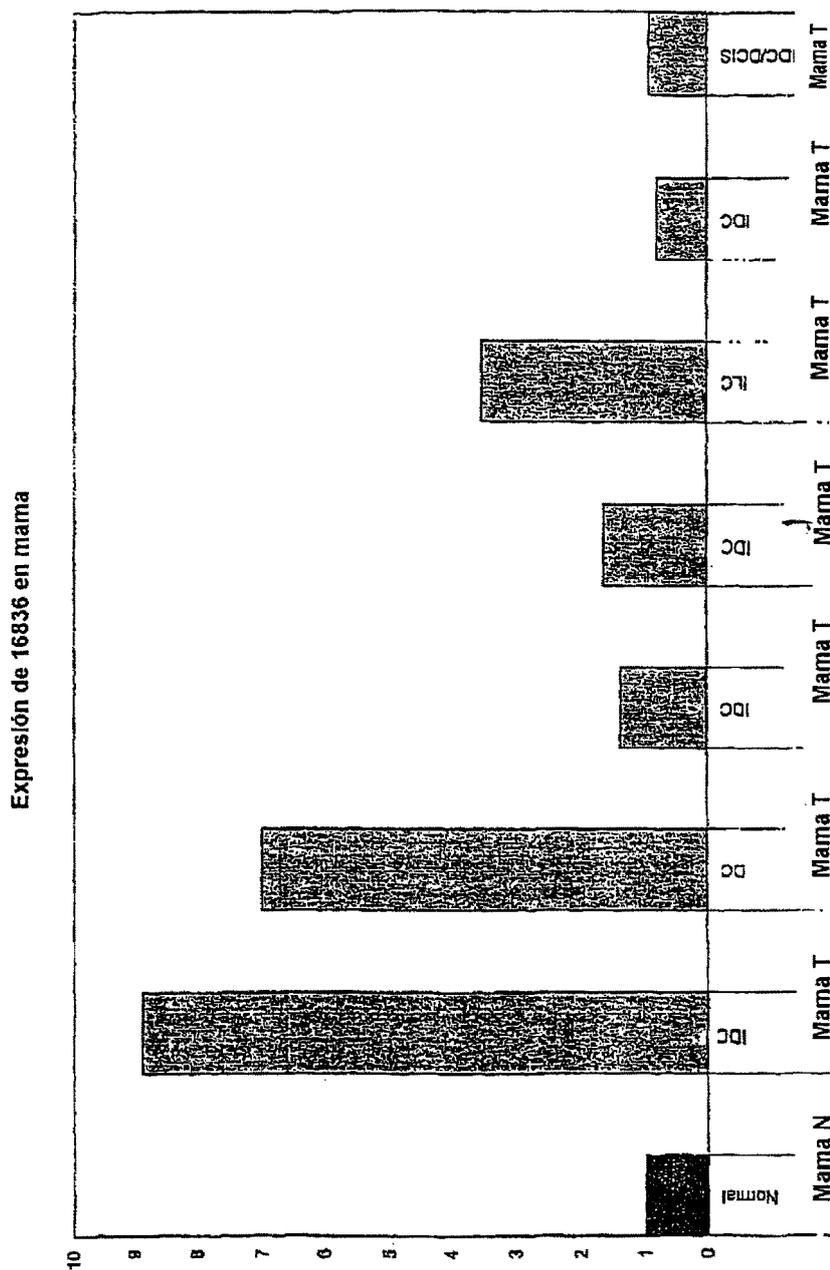


Fig. 8

Expresión de 16836

