



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 327 903**

51 Int. Cl.:

**C12N 15/12** (2006.01)

**C12N 15/54** (2006.01)

**C12N 9/12** (2006.01)

**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **01968005 .7**

96 Fecha de presentación : **17.08.2001**

97 Número de publicación de la solicitud: **1356040**

97 Fecha de publicación de la solicitud: **29.10.2003**

54 Título: **14189, una cinasa humana y usos de la misma.**

30 Prioridad: **18.08.2000 US 641690**

45 Fecha de publicación de la mención BOPI:  
**05.11.2009**

45 Fecha de la publicación del folleto de la patente:  
**05.11.2009**

73 Titular/es:  
**Bayer Schering Pharma Aktiengesellschaft**  
**Müllerstrasse 170-178**  
**13353 Berlin, DE**

72 Inventor/es: **Hodge, Martin, R.**

74 Agente: **Carpintero López, Mario**

**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

14189, una cinasa humana y usos de la misma.

5 **Campo de la invención**

La invención se refiere a secuencias de ácido nucleico de cinasa humana novedosas y a las moléculas de proteína codificadas. Se proporcionan también vectores, células hospedadoras y procedimientos recombinantes para preparar y usar las moléculas novedosas.

10 **Antecedentes de la invención**

Se conocía desde finales del siglo XIX que el fosfato se asociaba estrechamente a una molécula, por ejemplo una proteína. Desde entonces, se han encontrado una variedad de ligamientos covalentes de fosfato con proteínas. Los más comunes implican la esterificación de fosfato con serina, treonina y tirosina, ligándose cantidades menores con lisina, arginina, histidina, ácido aspártico, ácido glutámico y cisteína. La aparición de moléculas fosforiladas, por ejemplo proteínas, implica la existencia de una o más cinasas, por ejemplo proteína cinasas, capaces de fosforilar diversas moléculas, por ejemplo residuos aminoácidos en proteínas, y también de fosfatasa, por ejemplo proteína fosfatasa, capaces de hidrolizar diversas moléculas fosforiladas, por ejemplo, residuos aminoácidos fosforilados en proteínas.

Las proteína cinasas desempeñan papeles críticos en la regulación de los cambios bioquímicos y morfológicos asociados al crecimiento y división celulares (D'Urso y col. (1990) *Science* 250: 786-791; Birchmeier y col. (1993) *Bioessays* 15: 185-189). Por ejemplo, se ha mostrado que estas cinasas participan en la transmisión de señales desde receptores de factor de crecimiento (Sturgill y col. (1988) *Nature* 344: 715-718; Gómez y col. (1991) *Nature* 353: 170-173), el control de la entrada de células en mitosis (Nurse (1990) *Nature* 344: 503-508; Maller (1991) *Curr. Opin. Cell Biol.* 3: 269-275) y en la regulación de la formación de haces de actina (Husain-Chishti y col. (1988) *Nature* 334: 718-721). Las proteína cinasas sirven como receptores de factor de crecimiento y transductores de señal y se han implicado en la transformación y malignidad celulares (Hunter y col. (1992) *Cell* 70: 375-387; Posada y col. (1992) *Mol. Biol. Cell* 3: 583-592; Hunter y col. (1994) *Cell* 79: 573-582). Las alteraciones en los genes de cinasa y sus productos pueden conducir a una proliferación celular desregulada, un sello del cáncer. La modulación de estos genes y sus actividades reguladoras puede permitir el control de la proliferación e invasión de células tumorales.

Las proteína cinasas pueden dividirse en cuatro grupos diferentes basados en su similitud aminoácida o especificidad por residuos de serina/treonina o tirosina. Se ha descrito también un pequeño número de cinasas de especificidad dual. Dentro de la amplia clasificación, las cinasas pueden subdividirse adicionalmente en familias cuyos miembros comparten un alto grado de identidad de secuencia aminoácida del dominio catalítico y tienen también propiedades bioquímicas similares. La mayoría de los miembros de la familia de proteína cinasa comparten también rasgos estructurales fuera del dominio cinasa que reflejan sus papeles celulares particulares. Estos incluyen dominios reguladores que controlan la actividad cinasa o la interacción con otras proteínas (Hanks y col. (1988) *Science* 241: 42-52).

Las cinasas reguladas por señal extracelular/proteína cinasas activadas por mitógeno (ERK/MAPK) y cinasas dirigidas por ciclina (Cdk) representan dos grandes familias de serina-treonina cinasas (véase Songyang y col. (1996) *Mol. Cell. Biol.* 16: 6486-6493). Ambos tipos de cinasas funcionan en el crecimiento celular, la división celular y la diferenciación celular en respuesta a estímulos extracelulares. Los miembros de la familia ERK/MAPK son participantes críticos en las rutas de señalización intracelular. Los activadores cadena arriba, así como los componentes de ERK/MAPK, se fosforilan después del contacto de las células con factores de crecimiento u hormonas o en respuesta a estresantes celulares, por ejemplo, calor, luz ultravioleta y citocinas inflamatorias. Estas cinasas transportan mensajes que se han transmitido desde la membrana plasmática al citoplasma mediante cinasas cadena arriba hasta el núcleo, donde fosforilan factores de transcripción y efectúan la modulación de la transcripción génica (Karin y col. (1995) *Curr. Biol.* 5: 747-757). Los sustratos de la familia ERK/MAPK incluyen c-fos, c-jun, APF2 y los miembros de la familia ETS Elk1, Sap1a y c-Ets-1 (citados en Brott y col. (1998) *Proc. Natl. Acad. Sci. USA* 95: 963-968).

Las rutas de transducción de señal que emplean miembros de la familia ERK/MAPK de serina/treonina cinasas están ampliamente conservadas entre los eucariotas. La multiplicidad de estas rutas permite a la célula responder a estímulos extracelulares divergentes iniciando un amplio conjunto de respuestas en el intervalo de crecimiento celular a apoptosis. Las rutas de ERK/MAPK comprenden un módulo de señalización núcleo de tres niveles en el que las ERK/MAPK se regulan por MAPK/ERK cinasas (MEK), y las MEK a su vez se regulan por MAPK cinasas (MAPKKK). Las rutas de ERK/MAPK activadas por estrés en mamífero se han implicado en numerosas funciones fisiológicas importantes, incluyendo crecimiento y proliferación celulares, respuestas inflamatorias y apoptosis. Por ejemplo, la activación de la ruta de señalización de ERK1,2 por un factor de crecimiento mitogénico, un promotor tumoral o mediante transformación, suprime la expresión del gen de decorina en fibroblastos, que a su vez puede promover la proliferación y migración de células normales y malignas (Laine y col. (2000) *Biochem. J.* 349: 19-25).

Las Cdk regulan las transiciones entre etapas sucesivas del ciclo celular. Se controla la actividad de estas moléculas mediante eventos de fosforilación y mediante asociación con ciclina. La actividad Cdk está regulada negativamente mediante la asociación de moléculas inhibidoras pequeñas (Dylnacht (1997) *Nature* 389: 148-152). Las dianas de Cdk incluyen diversos activadores transcripcionales tales como p110Rb, p107, y factores de transcripción tales como p53,

## ES 2 327 903 T3

E2F y ARN polimerasa II, así como diversas proteínas citoesqueléticas y proteínas de señalización citoplasmáticas (citadas en Brott y col. (1998) *Proc. Natl. Acad. Sci. USA* 95: 963-968).

5 Las proteína cinasas desempeñan papeles críticos en el crecimiento celular, particularmente en la transducción de señales para proliferación, diferenciación y apoptosis celulares. Por lo tanto, los polinucleótidos y proteínas de proteína cinasa novedosos son útiles para modular el crecimiento, diferenciación y/o desarrollo celulares.

10 La entrada a la base de datos DATABASE EMBL, HEIDELBERG, n° de acceso BE464560 a la base de datos de 30 de julio de 2007, da a conocer un EST humano que es similar a la cinasa 7 regulada por señal extracelular y que es un 99,8% idéntico, de la posición 513-3, al polinucleótido mostrado por la SEQ ID NO:1 de la posición 1377-1887.

15 ABE, MK y col, *Molecular and Cellular Biology*, vol. 19 n° 2, febrero de 1999, páginas 1301-1302 dan a conocer la cinasa 7 regulada por señal de rata (ERK7) que es un 69,4% idéntica al polipéptido designado como SEQ ID NO: 2. Además, se da a conocer la expresión de dicha ERK7 fusionada con un marcador HA en células COS, el análisis de transferencia Northern de dicha ERK7, así como un procedimiento para identificar un anticuerpo anti-MAPK activa que se une a dicha ERK7.

### Sumario de la invención

20 Se proporcionan moléculas de ácido nucleico aisladas correspondientes a secuencias de ácido nucleico de cinasa. Adicionalmente, están comprendidas secuencias aminoácidas correspondientes a los polinucleótidos. En particular, la presente invención proporciona moléculas de ácido nucleico aisladas que comprenden secuencias nucleotídicas que codifican la secuencia aminoácida mostrada en la SEQ ID NO: 2 o la secuencia nucleotídica que codifica la secuencia de ADN depositada en el hospedador bacteriano como número de acceso ATCC PTA-2333. Se proporcionan adicionalmente polipéptidos cinasa que tienen secuencias aminoácidas codificadas por las moléculas de ácido nucleico de la invención.

30 La presente invención proporciona también vectores y células hospedadoras para la expresión recombinante de las moléculas de ácido nucleico de la invención. Se dan a conocer también procedimientos de preparación de dichos vectores y células hospedadoras y para usarlos para la producción de los polipéptidos o péptidos de la invención mediante técnicas recombinantes.

35 Las moléculas de cinasa de la presente invención son útiles para modular el crecimiento celular, la proliferación celular y/o rutas metabólicas celulares, particularmente para regular una o más proteínas implicadas en el crecimiento, proliferación y metabolismo. En consecuencia, en un aspecto esta invención proporciona moléculas de ácido nucleico aisladas que codifican proteínas cinasa.

40 Otro aspecto de esta invención se refiere a proteínas y polipéptidos de cinasa aislados o recombinantes. Las proteínas y polipéptidos de cinasa preferidos poseen al menos una actividad biológica poseída por las proteínas cinasa de origen natural de la invención.

45 Se dan a conocer moléculas de ácido nucleico y polipéptidos variantes sustancialmente homólogos con las secuencias nucleotídicas y aminoácidas de la presente invención. Adicionalmente, se dan a conocer fragmentos y fragmentos sustancialmente homólogos de las secuencias nucleotídicas y aminoácidas de la presente invención.

Se dan a conocer anticuerpos y fragmentos de anticuerpo que se unen selectivamente a polipéptidos de cinasa y fragmentos de los mismos. Dichos anticuerpos son útiles para detectar los polipéptidos de cinasa así como para modular el crecimiento, proliferación y metabolismo celulares.

50 En otro aspecto, la presente invención proporciona un procedimiento para detectar la presencia de una molécula de ácido nucleico de la invención en una muestra biológica poniendo en contacto la muestra biológica con una sonda de ácido nucleico o cebador que hibrida selectivamente con la misma, de tal modo que se detecte la presencia de la molécula de ácido nucleico en la muestra biológica.

55 En aún otro aspecto, se da a conocer un procedimiento para modular la actividad cinasa que comprende poner en contacto una célula con un agente que modula (inhibe o estimula) la actividad o expresión de cinasa de tal modo que se module la actividad o expresión de cinasa en la célula.

60 En una realización, se usan los procedimientos dados a conocer para tratar un sujeto que tiene un trastorno caracterizado por una actividad o expresión de ácido nucleico aberrante de las proteínas de cinasa de la invención mediante la administración al sujeto de un agente que es un modulador de cinasa.

65 Se da a conocer un ensayo diagnóstico para identificar la presencia o ausencia de una lesión o mutación genética caracterizada por al menos una de las siguientes: (1) modificación o mutación aberrante del gen que codifica una proteína de cinasa de la invención; (2) desregulación del gen que codifica una proteína de cinasa de la invención; y (3) modificación postraduccional aberrante de una proteína de cinasa de la invención, en el que la forma de tipo silvestre del gen codifica una proteína con actividad cinasa.

En otro aspecto, la invención proporciona un procedimiento para identificar un compuesto que se une a una proteína de cinasa de la invención. En general, dichos procedimientos suponen poner en contacto la proteína de cinasa con un compuesto de ensayo e identificar aquellos compuestos que se unen a la proteína de cinasa.

Se dan a conocer también procedimientos para identificar un compuesto que modula la expresión de un gen de cinasa de la invención midiendo la expresión de la secuencia de cinasa en presencia y ausencia del compuesto.

Resultarán evidentes otros rasgos y ventajas de la invención a partir de la descripción detallada y las reivindicaciones.

## Breve descripción de los dibujos

La Figura 1 proporciona la secuencia nucleotídica (SEQ ID NO:1) y la secuencia aminoacídica (SEQ ID NO: 2) de la proteína cinasa humana novedosa de la invención, h14189.

La Figura 2 muestra un alineamiento de secuencias nucleotídicas del marco de lectura abierto de h14189 (SEQ ID NO:1) con los nucleótidos 151-1804 de la cinasa 7 regulada por señal extracelular de rata, ERK7 (número de acceso AF078798; SEQ ID NO: 3). Estas secuencias exhiben aproximadamente 77% de identidad.

La Figura 3 muestra un alineamiento de la secuencia aminoacídica h14189 (SEQ ID NO: 2) con cinasa 7 regulada por señal extracelular de rata, ERK7 (número de acceso 4220888; SEQ ID NO: 4). Estas secuencias exhiben un 66,6% de identidad y un 69,2% de similitud.

Las Figuras 4A y 4B muestran la expresión del gen h14189 determinada mediante PCR cuantitativa TaqMan® en diversos tejidos y tipos celulares. Se observaron los mayores niveles de expresión de h14189 en tejido de riñón y pulmón normales, células HepG2 transfectadas con HBV (HepG2.2.15), muestras de hígado fibrótico (NF/NDR), una línea celular de eritroleucemia humana transformada (K562) y células de sangre periférica movilizadas (MPB CD34+). Se observaron también niveles significativos de expresión de h14189 en células hepáticas estrelladas humanas, bazo, amígdala, sangre de cordón umbilical, médula ósea y fibroblastos dérmicos.

## Descripción detallada de la invención

La presente invención está basada, al menos en parte, en el descubrimiento de moléculas novedosas, designadas en la presente memoria como moléculas de ácido nucleico y polipéptido de "cinasa", que desempeñan un papel, o actúan, en rutas de señalización asociadas a rutas de crecimiento celular, proliferación celular, diferenciación y/o a rutas metabólicas celulares. Se describen estas rutas de crecimiento, proliferación y metabólicas en Lodish y col. (1995) "Molecular Cell Biology" (Scientific American Books Inc., Nueva York, NY) y "Stryer Biochemistry" (W. H. Freeman, NY). En una realización, las moléculas de cinasa de la invención modulan la actividad de una o más proteínas implicadas en el crecimiento o diferenciación celulares, por ejemplo, el crecimiento y diferenciación de células cardíacas, hepáticas estrelladas o fibroblastos dérmicos. En otra realización, las moléculas de cinasa de la presente invención son capaces de modular el estado de fosforilación de una molécula de cinasa o el estado de fosforilación de una o más proteínas implicadas en el crecimiento o diferenciación celulares, por ejemplo, el crecimiento o diferenciación de células cardíacas, hepáticas estrelladas o fibroblastos dérmicos como se describe, por ejemplo, en Lodish y col. y Stryer, *supra*. Además, las cinasas de la presente invención son diana de medicamentos descritos en Goodman y Gilman (1996) "The Pharmacological Basis of Therapeutics" (9ª ed.), ed. Hartman y Limbard, cuyos contenidos se incorporan a la presente memoria como referencia. Particularmente, las cinasas de la presente invención pueden modular la fosforilación de proteínas celulares y moléculas proteicas en tejidos y células que incluyen, pero sin limitación: pulmón, riñón, bazo, hígado fetal, hígado normal, hígado fibrótico, nódulos linfáticos, amígdala, células HepG2, células Hep3, granulocitos, fibroblastos dérmicos y pulmonares, células hepáticas estrelladas, linfocitos T CD8+, linfocitos T, linfocitos B CD19+, células CD34+ de sangre periférica movilizadas, médula ósea en reposo de adulto, células GPA+, células de médula ósea, sangre de cordón umbilical, hígado infectado por virus de la hepatitis C (HCV) y líneas celulares de eritroleucemia transformadas como se muestra en las Figuras 4A y 4B.

Como se usa en la presente memoria, el término "cinasa" incluye una proteína, polipéptido u otra molécula no proteica que sea capaz de modular su propio estado de fosforilación o el estado de fosforilación de una proteína, polipéptido u otra molécula no proteica diferente. Las cinasas pueden tener especificidad por (concretamente especificidad para fosforilar) residuos de serina/treonina, residuos de tirosina o tanto residuos de serina/treonina como tirosina, por ejemplo las cinasas de especificidad dual. Como se designa en la presente memoria, las cinasas tales como proteína cinasas incluyen preferiblemente un dominio catalítico de aproximadamente 200-400 residuos aminoacídicos de longitud, preferiblemente de aproximadamente 200-300 residuos aminoacídicos de longitud, o más preferiblemente de aproximadamente 250-300 residuos aminoacídicos de longitud, que incluyen preferiblemente 5-20, más preferiblemente 5-15, o lo más preferiblemente 11 aminoácidos altamente conservados separados por secuencias de aminoácidos con conservación reducida o mínima designados en la presente memoria como subdominios o motivos. La especificidad de una cinasa para fosforilación de tirosina o serina/treonina puede predecirse mediante la secuencia de dos de los subdominios (VIb y VIII) en los que están conservados diferentes residuos en cada clase (como se describe, por ejemplo, en Hanks y col. (1988) *Science* 241: 42-52). Estos subdominios se describen también con más detalle en la presente memoria.

Las moléculas de proteína y ácido nucleico de cinasa novedosas de la presente invención pertenecen a una familia de moléculas que tienen ciertos rasgos estructurales y funcionales conservados. El término “familia”, cuando se designa en la presente memoria, se entiende que indica dos o más moléculas de proteína o ácido nucleico que tienen un dominio estructural o motivo común y que tienen suficiente identidad de secuencia aminoacídica o nucleotídica como se define en la presente memoria. Dichos miembros de familia pueden ser de origen natural o no natural y pueden ser de la misma o diferentes especies. 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. Los miembros de una familia pueden tener también características funcionales comunes.

Una realización de la invención caracteriza moléculas de ácido nucleico de cinasa, preferiblemente moléculas de ácido nucleico de cinasa humana, que se identificaron basándose en un motivo de consenso o dominio proteico característico de una familia de proteínas de cinasa. Específicamente, se proporciona el gen humano novedoso, denominado clon h14189 (SEQ ID NO: 1; Figura 1). Dichas secuencias se designan en la presente memoria como secuencias de “cinasa”, indicando que los marcos abiertos de lectura comparten similitud de secuencia con los de los genes de cinasa.

Las moléculas de ácido nucleico aisladas de la presente invención codifican un polipéptido de proteína cinasa eucariótica (SEQ ID NO: 2; Figura 1) que reside en la familia de proteína cinasas eucarióticas. Las proteína cinasas eucarióticas (descritas, por ejemplo, en Hanks y col. (1995) *FASEB J.* 9: 576-596) son enzimas que pertenecen a una extensa familia de proteínas que comparten un núcleo o dominio catalítico conservado común tanto a serina/treonina como a tirosina proteína cinasas. Hay una serie de regiones conservadas en el núcleo o dominio catalítico de las proteína cinasas. Una de estas regiones, localizada en el extremo N-terminal del dominio catalítico, es un tramo rico en glicina de residuos en la cercanía de un residuo de lisina que se ha mostrado que está implicado en la unión a ATP. Es una secuencia de consenso (SEQ ID NO: 11) para esta región:

**[LIV]-G-{P}-G-{P}-[FYWMGSTNH]-[SGA]-{PW}-[LIVCAT]-{PD}-x-  
[GSTACLIVMFY]-x(5,18)-[LIVMFYWCSTAR]-[AIVP]-[LIVMFAGCKR]-K [K  
se une a ATP]**

En este y los siguientes patrones de secuencia de consenso, cada elemento en el patrón está separado por un guión (-), los corchetes rectos [ ] indica los residuos particulares que se aceptan en esa posición, los corchetes curvados { } indican los residuos que no se aceptan en esa posición; x indica que cualquier residuo es aceptado en esa posición; la repetición de un elemento particular se indica siguiendo el elemento con un valor numérico o un intervalo numérico incluido entre paréntesis (concretamente, anteriormente x(5,18) indica que está presente cualesquiera de 5-18 residuos en el elemento, y puede aceptarse cualquier residuo en cada una de estas posiciones de residuo 5-18); y se usa el código de una letra estándar de la IUPAC para los aminoácidos. En el polipéptido h14189 expuesto en la SEQ ID NO: 2, esta región de unión a ATP reside en los aminoácidos 19-42 (véase la Figura 1). El residuo K en la posición 29 y/o 42 puede estar implicado en la unión a ATP.

Otra región, localizada en la parte central del núcleo o dominio catalítico, contiene un residuo de ácido aspártico conservado que es importante para la actividad catalítica de la enzima (Knighton y col. (1991) *Science* 253: 407-414). Se han descrito dos patrones característicos para esta región: uno específico de serina/treonina cinasas y otro de tirosina cinasas. Es una secuencia de consenso para las serina/treonina cinasas (SEQ ID NO: 12):

**[LIVMFYC]-x-[HY]-x-D-[LIVMFY]-K-x(2)-N-[LIVMFYCT](3)  
[D es un residuo de sitio activo]**

Es una secuencia de consenso para las tirosina cinasas (SEQ ID NO: 13):

**[LIVMFYC]-x-[HY]-x-D-[LIVMFY]-[RSTAC]-x(2)-N-[LIVMFYC](3)  
[D es un residuo de sitio activo]**

El patrón característico del polipéptido h14189 de la presente invención se parece más al de una serina/treonina cinasa y reside en los aminoácidos 133-145 (véase la Figura 1).

Se descubrieron las moléculas de cinasa de la presente invención basándose en una secuencia de ADNc novedosa identificada en una colección de tejido pulmonar. Este clon, h14189, codifica un transcrito de ARNm de aproximada-

mente 1,9 kb que tiene la correspondiente secuencia de ADNc expuesta en la SEQ ID NO: 1 (véase la Figura 1). Este transcrito tiene un marco abierto de lectura de 1635 nucleótidos incluyendo el codón de detención (nucleótidos 48-1682 de la SEQ ID NO: 1), que codifica una proteína de 544 aminoácidos (SEQ ID NO: 2) que tiene un peso molecular de aproximadamente 59,83 kDa. La molécula no parece contener un segmento transmembrana como se predice por MEMSAT.

Se usó el programa de análisis Prosite para predecir diversos sitios en la proteína h14189. Se predijo un sitio de *N*-glucosilación en los aminoácidos (aa) 148-151. Se predijo un sitio de unión a glucosaminoglucano en los aa 447-450. Se predijo un sitio de fosforilación de proteína cinasa dependiente de AMPc y GMPc en los aa 291-294. Se predijeron sitios de fosforilación de proteína cinasa C en los aa 57-59, 150-152, 192-194, 352-354, 403-405 y 447-449. Se predijeron sitios de fosforilación de caseína cinasa II en los aa 3-6, 57-60, 75-78, 161-164, 238-241, 273-276, 331-334 y 352-355. Se predijo un sitio de fosforilación de tirosina en los aa 81-89. Se predijeron sitios de *N*-miristoilación en los aa 157-162, 327-332, 347-352, 360-365, 450-455, 478-483 y 509-514. Se predijo un sitio de amidación en los aa 497-500, y se predijo una secuencia de unión a células RGD en los aa 470-472. Se predijo una región de unión a ATP característica de proteína cinasa en los aa 19-42.

La proteína h14189 posee dos dominios de proteína cinasa eucariótica, abarcando el primero los aa 13-201 y el segundo los aa 269-304, como se predijo mediante HMMer, versión 2. Se predijeron también un dominio de serina cinasa y un dominio de tirosina cinasa para la proteína h14189 mediante análisis con HMMer, abarcando los aa 13-304 y aa 13-307, respectivamente. Sin embargo, hay un grado de similitud mucho mayor entre h14189 y el dominio de serina cinasa de consenso (puntuación de 278,4 y valor de  $E$  de  $9,5 \times 10^{-80}$ ) que entre esta proteína y el dominio de tirosina cinasa de consenso (puntuación de 19,7 y valor de  $E$  de  $4 \times 10^{-13}$ ). Además, el patrón característico de h14189, VVHRDQKPSNVLL (residuos aa 133-145 de la Figura 1 y SEQ ID NO: 2), coincide más estrechamente con la secuencia de consenso de serina/treonina cinasas (mostrada anteriormente) que con la de tirosina cinasas (mostrada anteriormente)).

El análisis BLAST de las bases de datos de patentes y públicas de ácido nucleico y proteína indica que la proteína cinasa h14189 comparte la mayor similitud con varias proteína cinasas reguladas por señal extracelular (ERK) y proteína cinasas activadas por mitógeno (MAPK). La proteína cinasa h14189 exhibe la homología más estrecha con una proteína de rata conocida como cinasa ERK7 (número de acceso AF078798; SEQ ID NO: 4). El análisis GAP de los marcos abiertos de lectura de los ADNc que codifican la proteína de cinasa h14189 de la invención y la cinasa ERK7 de rata (número de acceso AF078798) muestra que estas secuencias nucleotídicas comparten aproximadamente un 77,2% de identidad (véase la Figura 2). El análisis GAP de la proteína de cinasa h14189 de la invención y la proteína de cinasa ERK7 de rata (número de acceso 4220888) demuestra que estas proteínas comparten aproximadamente un 66,6% de identidad (véase la Figura 3).

Se depositó un plásmido que contiene el inserto de ADNc de cinasa humana 14189 en la American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, Virginia, el 3 de agosto de 2000, y se le asignó el número de acceso PTA-2333. Este depósito se mantendrá según los términos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes. Este depósito se hizo simplemente por conveniencia para los expertos en la técnica, y no es una admisión de que se requiera un depósito según el título 35 del U.S.C. sección 112.

Los polipéptidos de cinasa dados a conocer tienen una secuencia aminoacídica suficientemente idéntica con la secuencia aminoacídica de SEQ ID NO: 2, o un dominio de la misma. El término "suficientemente idéntica" se usa en la presente memoria para designar una primera secuencia aminoacídica o nucleotídica que contiene un número suficiente o mínimo de residuos aminoacídicos o nucleotídicos idénticos o equivalentes (por ejemplo, con una cadena lateral similar) a una segunda secuencia aminoacídica o nucleotídica, de tal modo que las primera y segunda secuencias aminoacídica o nucleotídica tengan un dominio estructural común (por ejemplo, un dominio de cinasa eucariótica, dominio de serina cinasa y/o dominio de tirosina cinasa) y/o actividad funcional común. Por ejemplo, las secuencias aminoacídicas o nucleotídicas que contienen un dominio estructural común que tienen al menos aproximadamente un 80% de identidad, preferiblemente un 85% de identidad, más preferiblemente un 90% de identidad y lo más preferiblemente un 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad, se definen en la presente memoria como suficientemente idénticas.

Para determinar la identidad porcentual de dos secuencias aminoacídicas o dos ácidos nucleicos, se alinean las secuencias con fines de comparación óptima. La identidad porcentual entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (concretamente, identidad porcentual = número de posiciones idénticas/número total de posiciones (por ejemplo, posiciones superpuestas)  $\times$  100). En una realización, las dos secuencias son de la misma longitud. La identidad porcentual entre dos secuencias puede determinarse usando técnicas similares a las descritas a continuación, permitiendo o no huecos. Al calcular la identidad porcentual, se cuentan típicamente las coincidencias exactas.

La determinación de la identidad porcentual entre dos secuencias puede conseguirse usando un algoritmo matemático. Es un ejemplo preferido no limitante de un algoritmo matemático utilizado para la comparación de dos secuencias el algoritmo de Karlin y Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87: 2264, modificado en Karlin y Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90: 5873-5877. Dicho algoritmo se incorpora a los programas NBLAST y XBLAST de Altschul y col. (1990) *J. Mol. Biol.* 215: 403. Las búsquedas de nucleótidos BLAST pueden efectuarse con el pro-

grama NBLAST, puntuación = 100, longitud de palabra = 12, para obtener secuencias nucleotídicas homólogas de la molécula de ácido nucleico de cinasa de la invención. Las búsquedas de proteína BLAST pueden efectuarse con el programa XBLAST, puntuación = 50, longitud de palabra = 3, para obtener secuencias aminoácidas homólogas a la molécula de proteína de cinasa de la invención. Para obtener alineamientos con hueco con fines comparativos, puede utilizarse BLAST con huecos como se describe en Altschul y col. (1997) *Nucleic Acids Res.* 25: 3389. Como alternativa, puede usarse PSI-BLAST para efectuar una búsqueda iterativa que detecte relaciones distantes entre moléculas. Cuando se utilizan los programas BLAST, BLAST con huecos y PSI-BLAST, pueden usarse los parámetros por defecto de los programas respectivos (por ejemplo, XBLAST y NBLAST). Véase <http://www.ncbi.nlm.nih.gov>. Otro ejemplo preferido de algoritmo utilizado para la comparación de secuencias es el algoritmo de Myers y Miller (1988) *CABIOS* 4: 11-17. Dicho algoritmo se incorpora al programa ALIGN (versión 2.0), que es parte del paquete de software de alineamiento de secuencias GCG. Cuando se utiliza el programa ALIGN para comparar secuencias aminoácidas, pueden usarse una tabla de residuos ponderados PAM120, una penalización por longitud de hueco de 1, y una penalización de hueco de 4. Es un programa preferido el Pairwise Alignment Program (Sequence Explorer), usando parámetros por defecto.

En consecuencia, otra realización de la invención se caracteriza por proteínas y polipéptidos de cinasa aislados que tienen una actividad proteína cinasa. Como se usa intercambiamente en la presente memoria, una "actividad proteína cinasa", "actividad biológica de una proteína cinasa" o "actividad funcional de una proteína cinasa" designa una actividad ejercida por una proteína, polipéptido o molécula de ácido nucleico de cinasa sobre una célula sensible a cinasa, determinada *in vivo* o *in vitro* según técnicas de ensayo estándar. Una actividad cinasa puede ser una actividad directa, tal como una asociación con o una actividad enzimática en una segunda proteína, o una actividad indirecta, tal como una actividad de señalización celular mediada por la interacción de la proteína de cinasa con una segunda proteína. En una realización, una actividad cinasa incluye al menos una o más de las siguientes actividades: (1) modular (estimular y/o potenciar o inhibir) la proliferación, crecimiento y/o metabolismo celular (por ejemplo, en aquellas células en que se expresa la secuencia, incluyendo: pulmón, riñón, bazo, hígado fetal, hígado normal e hígado fibrótico, nódulos linfáticos y amígdala; células HepG2; células HepG2 transfectadas con HBV; células Hep3; granulocitos; fibroblastos dérmicos y pulmonares; células hepáticas estrelladas; linfocitos T CD8+; linfocitos T; linfocitos B CD19+; células CD34+ de sangre periférica movilizadas; médula ósea de adulto en reposo; células GPA+ de médula ósea, cordón umbilical; hígado infectado con el virus de la hepatitis C (HCV) y líneas celulares de eritroleucemia transformadas); (2) regular la transmisión de señales desde los receptores celulares, incluyendo receptores de factor de crecimiento; (3) modular al entrada de células en mitosis; (4) modular la diferenciación celular; (5) modular la muerte celular y (6) regular la función citoesquelética, incluyendo la formación de haces de actina.

Son bien conocidos en la técnica ensayos para medir la actividad cinasa, dependiendo de la cinasa particular. Están disponibles protocolos de ensayo específicos en fuentes estándar conocidas por el experto en la técnica. Por ejemplo, véanse "Kinases" en Ausubel y col., eds. (1994-1998) "Current Protocols in Molecular Biology" (3) y las referencias citadas en los mismos, los siguientes dos sitios de recursos de internet: 1) [http://www.sdsc.edu/Kinases/pkr/pk\\_protocols.html](http://www.sdsc.edu/Kinases/pkr/pk_protocols.html), y 2) [http://www.sdsc.edu/Kinases/pkr/pk\\_protocols/tyr\\_synpep\\_assay.html](http://www.sdsc.edu/Kinases/pkr/pk_protocols/tyr_synpep_assay.html).

Una molécula de ácido nucleico o proteína de cinasa "aislada" o "purificada", o una porción biológicamente activa de la misma, está sustancialmente exenta de otro material celular o medio de cultivo cuando se produce mediante técnicas recombinantes, o sustancialmente exenta de precursores químicos u otros productos químicos cuando se sintetiza químicamente. Preferiblemente, un ácido nucleico "aislado" está exento de secuencias (preferiblemente secuencias que codifican proteína) que flanquean naturalmente al ácido nucleico (concretamente, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del organismo del que deriva el ácido nucleico. Con fines de la invención, "aislado" cuando se usa para designar moléculas de ácido nucleico excluye cromosomas aislados. Por ejemplo, en diversas realizaciones, la molécula de ácido nucleico de cinasa 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 nucleotídicas que flanquean naturalmente a la molécula de ácido nucleico en ADN genómico de la célula de la que deriva el ácido nucleico. Una proteína de cinasa que está sustancialmente exenta de material celular incluye preparaciones de proteína de cinasa que tienen menos de aproximadamente 30%, 20%, 10% ó 5% (en peso seco) de proteína no cinasa (también designada en la presente memoria como "proteína contaminante"). Cuando la proteína de cinasa o porción biológicamente activa de la misma se produce recombinantemente, preferiblemente el medio de cultivo representa menos de aproximadamente 30%, 20%, 10% ó 5% del volumen de la preparación de proteína. Cuando la proteína de cinasa se produce mediante síntesis química, preferiblemente las preparaciones de proteína tienen menos de aproximadamente 30%, 20%, 10% ó 5% (en peso seco) de precursores químicos o productos químicos no cinasas.

Las cinasas desempeñan un papel en rutas de señalización asociadas al crecimiento celular. Por ejemplo, las proteínas cinasas están implicadas en la regulación de la transmisión de señal de receptores celulares tales como receptores de factor de crecimiento, la entrada de células en mitosis y la regulación de la función citoesquelética. Las rutas de transducción de señal que emplean miembros de la familia ERK/MAPK de proteína serina/treonina cinasas están ampliamente conservadas entre los eucariotas. Como se observa anteriormente, las rutas de ERK/MAPK activadas por estrés en mamíferos se han implicado en numerosas funciones fisiológicas importantes, incluyendo el crecimiento y proliferación celulares, respuestas inflamatorias y apoptosis.

La inhibición o sobreestimulación de la actividad cinasa implicadas en rutas de señalización asociadas al crecimiento celular puede conducir a un crecimiento celular alterado, que a su vez puede conducir a trastornos relacionados con el crecimiento celular. Como se usa en la presente memoria, un "trastorno relacionado con el crecimiento celular"

incluye un trastorno, enfermedad o afección caracterizado por una desregulación, por ejemplo una regulación positiva o una regulación negativa, del crecimiento celular. La desregulación del crecimiento celular puede ser debida a una desregulación de la proliferación celular, de la progresión del ciclo celular, de la diferenciación celular y/o a hipertrofia celular. Los ejemplos de dichos trastornos relacionados con el crecimiento celular incluyen, pero sin limitación, 5 trastornos cardiovasculares tales como insuficiencia cardíaca, hipertensión, fibrilación auricular, cardiomiopatía dilatada, cardiomiopatía idiopática o angina; trastornos proliferativos o trastornos diferenciativos tales como cáncer, por ejemplo, melanoma, cáncer de próstata, cáncer de cuello del útero, cáncer de mama, cáncer de colon o sarcoma. Las composiciones de la invención son útiles en el diagnóstico y/o tratamiento de dichos trastornos, así como de fibrosis hepática y otros trastornos relacionados con el hígado. Están comprendidos también los trastornos asociados a las 10 siguientes células o tejidos: pulmón, riñón, bazo, hígado fetal, hígado normal e hígado fibrótico, nódulos linfáticos y amígdala, células HepG2, granulocitos, fibroblastos dérmicos y pulmonares, células hepáticas estrelladas, linfocitos T CD8+, linfocitos T, linfocitos B CD 19+, células CD34+ de sangre periférica movilizada, médula ósea y sangre de cordón umbilical.

En vista de la función biológica de las moléculas de cinasa de la invención, estas moléculas y moduladores de las 15 mismas pueden usarse en procedimientos dirigidos a la modulación, diagnóstico y tratamiento de trastornos, incluyendo pero sin limitación, trastornos hepáticos, trastornos fibróticos y trastornos inmunitarios, inflamatorios, respiratorios y hematológicos. Los trastornos que implican el hígado incluyen, pero sin limitación, lesión hepática, ictericia y colestasis tales como formación de bilirrubina y bilis, insuficiencia hepática y cirrosis, hipertensión portal incluyendo 20 ascitis, derivaciones portosistémicas y esplenomegalia; trastornos infecciosos tales como hepatitis vírica, incluyendo infección por hepatitis A-E e infección por otros virus de hepatitis; síndromes clinicopatológicos tales como estado de portador, infección asintomática, hepatitis vírica aguda, hepatitis vírica crónica y hepatitis fulminante; hepatitis autoinmunitaria; enfermedad hepática inducida por medicamentos y toxinas tales como enfermedad hepática alcohólica; errores innatos del metabolismo y enfermedad hepática pediátrica tal como hemocromatosis, enfermedad de 25 Wilson, deficiencia de antitripsina  $\alpha_1$  y hepatitis neonatal; enfermedad del tracto biliar intrahepático tal como cirrosis biliar secundaria, cirrosis biliar primaria, colangitis esclerosante primaria y anomalías del árbol biliar; trastornos circulatorios tales como flujo sanguíneo insuficiente al hígado, incluyendo compromiso arterial hepático y obstrucción y trombosis de la vena portal, flujo sanguíneo insuficiente a través del hígado, incluyendo congestión pasiva y necrosis centrilobular y peliosis hepática, obstrucción del flujo de salida de vena hepática, incluyendo trombosis de 30 vena hepática (síndrome de Budd-Chiari) y enfermedad venooclusiva; enfermedad hepática asociada al embarazo tal como preeclampsia y eclampsia, hígado graso agudo del embarazo y colestasis intrahepática del embarazo; complicaciones hepáticas de trasplante de órgano o médula ósea tales como toxicidad medicamentosa después de trasplante de médula ósea, enfermedad del hospedador frente al injerto y rechazo hepático y daños no inmunitarios por aloinjertos hepáticos; tumores y afecciones tumorales tales como hiperplasias nodulares, adenomas y tumores malignos, 35 incluyendo carcinoma primario del hígado y tumores metastásicos.

Los trastornos o enfermedades fibróticos incluyen fibrosis en general, por ejemplo, enfermedad pulmonar obstructiva crónica, fibrosis pulmonar idiopática, glomerulofibrosis progresiva, sarcoidosis, fibrosis quística, fibrosis/cirrosis, incluyendo cirrosis secundaria de alcoholismo crónico, cirrosis secundaria de hepatitis de tipo B o hepatitis de tipo C 40 y cirrosis biliar primaria; trastornos hepáticos dados a conocer anteriormente, particularmente fibrosis hepática y otras enfermedades fibróticas, así como fibrosis asociada al tratamiento de quemaduras y cicatrices.

Los trastornos inmunitarios incluyen, pero sin limitación, enfermedades y trastornos inflamatorios crónicos, enfermedades intestinales inflamatorias tales como enfermedad de Crohn y colitis ulcerosa, artritis reumatoide, incluyendo 45 enfermedad de Lyme, diabetes insulino dependiente, autoinmunidad específica de órgano, incluyendo esclerosis múltiple, tiroiditis de Hashimoto y enfermedad de Grave, dermatitis de contacto, psoriasis, rechazo de injerto, enfermedad de hospedador frente al injerto, sarcoidosis, afecciones atópicas tales como asma y alergia, incluyendo rinitis alérgica, alergias gastrointestinales, incluyendo alergias alimentarias, eosinofilia, conjuntivitis, glomerulonefritis, ciertas sensibilidades patogénicas tales como helmínticas (por ejemplo, leishmaniasis), ciertas infecciones víricas, incluyendo 50 VIH, e infecciones bacterianas, incluyendo tuberculosis y lepra lepromatosa.

Los trastornos respiratorios incluyen, pero sin limitación, apnea, asma, particularmente asma bronquial, beriliosis, bronquiectasis, bronquitis, bronconeumonía, fibrosis quística, difteria, disnea, enfisema, enfermedad pulmonar obstructiva crónica, aspergilosis broncopulmonar alérgica, neumonía, edema pulmonar agudo, tos ferina, faringitis, 55 atelectasia, granulomatosis de Wegener, legionelosis, pleuresía, fiebre reumática y sinusitis.

Los trastornos hematológicos incluyen, pero sin limitación, anemias, incluyendo anemia de células falciformes y hemolítica, hemofilias, incluyendo de tipos A y B, leucemias, talasemias, esferocitosis, enfermedad de Von Willebrand, enfermedad granulomatosa crónica, deficiencia de glucosa-6-fosfato deshidrogenasa, trombosis, anomalías 60 y deficiencias del factor de coagulación, incluyendo deficiencias de factor VII y IX, hemartrosis, hematemesis, hematomas, hematuria, hemocromatosis, hemoglobinuria, síndrome hemolítico-urémico, trombocitopenias, incluyendo trombocitopenia asociada a VIH, telangiectasia hemorrágica, púrpura trombocitopénica idiopática, microangiopatía trombótica, hemosiderosis.

Se describen diversos aspectos de la invención con más detalle en las siguientes subsecciones.

## I. Moléculas de ácido nucleico aisladas

Un aspecto de la invención está relacionado con moléculas de ácido nucleico aisladas que comprenden secuencias nucleotídicas que codifican las proteínas de cinasa de la presente invención o porciones biológicamente activas de las mismas, así como moléculas de ácido nucleico suficientes para usar como sondas de hibridación para identificar ácidos nucleicos que codifican cinasa (por ejemplo, ARNm de cinasa) y fragmentos para uso como cebadores de PCR para la amplificación o mutación de moléculas de ácido nucleico de cinasa. Como se usa en la presente memoria, se pretende que el término “molécula de ácido nucleico” incluya moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo, ARNm) y análogos de ADN o ARN generados usando análogos nucleotídicos. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferiblemente es ADN bicatenario.

Las secuencias nucleotídicas que codifican las proteínas de cinasa de la presente invención incluyen la secuencia expuesta en la SEQ ID NO: 1, la secuencia nucleotídica del inserto de ADNc del plásmido depositado en la ATCC con número de acceso PTA-2333 (designado como “el ADNc de ATCC PTA-2333”), y complementos de las mismas. La secuencia aminoacídica de la proteína de cinasa codificada por estas secuencias nucleotídicas se expone en la SEQ ID NO: 2.

Se dan a conocer también moléculas de ácido nucleico que son fragmentos de las secuencias nucleotídicas de cinasa de la invención. Se pretende por “fragmento” una porción de la secuencia nucleotídica que codifica una proteína de cinasa de la invención. Un fragmento de secuencia nucleotídica de cinasa puede codificar una porción biológicamente activa de la proteína de cinasa, o puede ser un fragmento que puede usarse como sonda de hibridación o cebador de PCR usando procedimientos dados a conocer a continuación. Puede prepararse una porción biológicamente activa de una proteína de cinasa de la invención aislando una porción de una de las secuencias nucleotídicas de cinasa de la invención que expresa la porción codificada de la proteína de cinasa (por ejemplo, mediante expresión recombinante *in vitro*) y evaluando la actividad de la porción codificada de la proteína de cinasa. Generalmente, las moléculas de ácido nucleico que son fragmentos de una secuencia nucleotídica de cinasa comprenden al menos 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 325, 350, 375, 400, 425, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1450, 1500, 1550, 1600 ó 1620 nucleótidos o hasta el número de nucleótidos presentes en una secuencia nucleotídica de cinasa entera dada a conocer en la presente memoria (por ejemplo, 1635 nucleótidos para la SEQ ID NO: 1), dependiendo del uso pretendido.

Un fragmento de una secuencia nucleotídica de cinasa que codifica una porción biológicamente activa de la proteína de cinasa de la invención codificará al menos 20, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 475, 500, 525, 530 ó 540 aminoácidos contiguos, o hasta el número total de aminoácidos presentes en una proteína de cinasa completa de la invención (por ejemplo, 544 aminoácidos para la SEQ ID NO: 2). Los fragmentos de una secuencia nucleotídica de cinasa que son útiles como sondas de hibridación para cebadores de PCR generalmente no tienen que codificar una porción biológicamente activa de una proteína de cinasa.

Se dan a conocer moléculas de ácido nucleico que son variantes de las secuencias nucleotídicas de cinasa. “Variantes” de las secuencias nucleotídicas de cinasa incluyen aquellas secuencias que codifican las proteínas de cinasa dadas a conocer en la presente memoria, pero que difieren de forma conservativa debido a la degeneración del código genético. Estas variantes alélicas de origen natural pueden identificarse con el uso de técnicas de biología molecular bien conocidas, tales como reacción en cadena de la polimerasa (PCR) y técnicas de hibridación como se indican a continuación. Las variantes de secuencias nucleotídicas incluyen también secuencias nucleotídicas derivadas sintéticamente que se han generado, por ejemplo, usando mutagénesis dirigida a sitio, pero que codifican todavía las proteínas de cinasa dadas a conocer en la presente invención como se debate a continuación. Generalmente, las variantes de secuencia nucleotídica de la invención tendrán al menos aproximadamente un 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% de identidad con las secuencias nucleotídicas dadas a conocer en la presente memoria. Una variante de secuencia nucleotídica de cinasa codificará una proteína de cinasa que tiene una secuencia aminoacídica que tiene al menos aproximadamente un 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% de identidad con una secuencia aminoacídica de una proteína de cinasa dada a conocer en la presente memoria.

Además de la secuencia nucleotídica de cinasa mostrada en la SEQ ID NO: 1 y de la secuencia nucleotídica de ADNc de ATCC PTA-2333, se apreciará por los expertos en la técnica que pueden existir polimorfismos de secuencia de ADN que conducen a cambios en las secuencias aminoacídicas de las proteínas de cinasa dentro de una población (por ejemplo, la población humana). Dicho polimorfismo genético en un gen de cinasa puede existir entre individuos dentro de una población debido a una variación alélica natural. Un alelo es uno de un grupo de genes que aparece alternativamente en un locus genético dado. Como se usa en la presente memoria, los términos “gen” y “gen recombinante” designan moléculas de ácido nucleico que comprenden un marco abierto de lectura que codifica una proteína de cinasa, preferiblemente una proteína de cinasa de mamífero. Como se usa en la presente memoria, la frase “variante alélica” designa una secuencia nucleotídica que aparece en un locus de cinasa o un polipéptido codificado por la secuencia nucleotídica. Dichas variaciones alélicas naturales pueden dar como resultado típicamente una variación de 1-5% de la secuencia nucleotídica del gen de cinasa. Se entiende que están dentro del alcance de la invención todas y cada una de dichas variaciones nucleotídicas y polimorfismos o variaciones aminoacídicas resultantes en una secuencia de cinasa que son el resultado de una variación alélica natural y que no alteran la actividad funcional de proteínas de cinasa.

Además, se dan a conocer también moléculas de ácido nucleico que codifican proteínas de cinasa de otras especies (homólogos de cinasa), que tienen una secuencia nucleotídica que difiere de la de las secuencias de cinasa dadas a conocer en la presente memoria. Las moléculas de ácido nucleico correspondientes a variantes alélicas naturales y homólogos de ADNc de cinasa de la invención pueden aislarse basándose en su identidad con los ácidos nucleicos de cinasa humana dados a conocer en la presente memoria usando los ADNc humanos, o una porción de los mismos, como sonda de hibridación según técnicas de hibridación estándar en condiciones de hibridación rigurosas como se da a conocer a continuación.

Además de las variantes alélicas de origen natural de la secuencia de cinasa que pueden existir en la población, el experto en la técnica apreciará adicionalmente que pueden introducirse cambios mediante mutación en las secuencias nucleotídicas de la invención, conduciendo así a cambios en la secuencia aminoacídica de la proteína de cinasa codificada, sin alterar la actividad biológica de la proteína de cinasa. Por tanto, puede crearse una molécula de ácido nucleico aislada que codifica una proteína de cinasa que tiene una secuencia que difiere de la de SEQ ID NO:2 introduciendo una o más sustituciones, adiciones o deleciones nucleotídicas en la secuencia nucleotídica dada a conocer en la presente memoria, de tal modo que se introduzcan una o más sustituciones, adiciones o deleciones aminoacídicas en la proteína codificada. Las mutaciones pueden introducirse mediante técnicas estándar tales como mutagénesis dirigida a sitio y mutagénesis mediada por PCR. Dichas variantes de secuencias nucleotídicas están también comprendidas por la presente invención.

Por ejemplo, pueden hacerse preferiblemente sustituciones aminoacídicas en uno o más residuos aminoacídicos predichos, preferiblemente no esenciales. Un residuo aminoacídico “no esencial” es un residuo que puede alterarse de la secuencia de tipo silvestre de una proteína de cinasa (por ejemplo, la secuencia SEQ ID NO: 2) sin alterar la actividad biológica, mientras que un residuo aminoacídico “esencial” es necesario para la actividad biológica. Una “sustitución aminoacídica conservativa” es una en que el residuo aminoacídico se reemplaza por un residuo aminoacídico que tiene una cadena lateral similar. Se han definido en la técnica familias de residuos 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 no cargadas (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 de ramificación beta (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). Dichas sustituciones no se harían para residuos aminoacídicos conservados o para residuos aminoacídicos que residan en un dominio de proteína conservado, tal como el dominio de proteína cinasa eucariótica crítico del clon dado a conocer, donde dichos residuos son esenciales para la actividad proteica.

Como alternativa, pueden hacerse variantes de secuencias nucleotídicas de cinasa introduciendo mutaciones aleatoriamente en toda o parte de la secuencia de codificación de cinasa, tal como mediante mutagénesis por saturación, y en los mutantes resultantes puede cribarse la actividad biológica de cinasa para identificar los mutantes que retienen actividad. Después de la mutagénesis, la proteína codificada puede expresarse recombinantemente, y puede determinarse la actividad de la proteína usando técnicas de ensayo estándar.

Por tanto, las secuencias nucleotídicas de la invención incluyen la SEQ ID NO: 1 y la secuencia nucleotídica del ADNc de ATCC. Las secuencias nucleotídicas de cinasa de la invención, y los fragmentos y variantes de las mismas, pueden usarse como sondas y/o cebadores para identificar y/o clonar homólogos de cinasa en otros tipos celulares, por ejemplo, de otros tejidos, así como homólogos de cinasa de otros mamíferos. Dichas sondas pueden usarse para detectar transcritos o secuencias genómicas que codifican la misma o idénticas proteínas. Estas sondas pueden usarse como parte de un kit de ensayo de diagnóstico para identificar células o tejidos que exhiben una expresión aberrante de una proteína de cinasa de la invención, tal como midiendo los niveles de un ácido nucleico que codifica cinasa en una muestra de células de un sujeto, por ejemplo, detectando los niveles de ARNm de cinasa o determinando si un gen de cinasa genómico se ha mutado o eliminado.

De esta manera, pueden usarse procedimientos tales como PCR, hibridación y similares para identificar dichas secuencias que tienen una identidad sustancial con las secuencias de la invención. Véanse, por ejemplo, Sambrook y col. (1989) “Molecular Cloning: Laboratory Manual” (2ª ed., Cold Spring Harbor Laboratory Press, Plainview, NY) e Innis, y col. (1990) “PCR Protocols: A Guide to Methods and Applications” (Academic Press, NY). Se dan a conocer también secuencias nucleotídicas de cinasa aisladas basándose en su identidad de secuencia con las secuencias nucleotídicas de cinasa expuestas en la presente memoria o fragmentos y variantes de las mismas.

En un procedimiento de hibridación, puede usarse toda o parte de una secuencia nucleotídica de cinasa conocida para cribar colecciones de ADNc o genómicas. Los procedimientos para la construcción de dichas colecciones de ADNc y genómicas son generalmente conocidos en la técnica, y se dan a conocer en Sambrook y col. (1989) “Molecular Cloning: A Laboratory Manual” (2ª ed., Cold Spring Harbor Laboratory Press, Plainview, NY). Las denominadas sondas de hibridación pueden ser fragmentos de ADN genómico, fragmentos de ADNc, fragmentos de ARN u otros oligonucleótidos y pueden estar marcadas con un grupo detectable tal como <sup>32</sup>P o cualquier otro marcador detectable tal como otros radioisótopos, un compuesto fluorescente, una enzima o un cofactor enzimático. Las sondas para hibridación pueden prepararse marcando oligonucleótidos sintéticos basados en secuencias nucleotídicas de cinasa conocidas dadas a conocer en la presente memoria. Pueden usarse adicionalmente cebadores degenerados diseñados basándose en residuos nucleotídicos o aminoacídicos conservados de una secuencia nucleotídica de cinasa conocida

o una secuencia aminoacídica codificada. La sonda comprende típicamente una región de secuencia nucleotídica que hibrida en condiciones rigurosas con al menos aproximadamente 12, preferiblemente al menos aproximadamente 25, más preferiblemente aproximadamente 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 ó 400 nucleótidos consecutivos de la secuencia nucleotídica de cinasa de la invención o un fragmento o variante de la misma. La preparación de sondas para hibridación es generalmente conocida en la técnica y se da a conocer en Sambrook y col. (1989) "Molecular Cloning: A Laboratory Manual" (2ª ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York).

Por ejemplo, una molécula de ácido nucleico de cinasa anteriormente no identificada hibrida en condiciones rigurosas con una sonda que es una molécula de ácido nucleico que comprende la secuencia nucleotídica de cinasa de la invención o un fragmento de la misma.

Como se usa en la presente memoria, se pretende que el término "hibrida en condiciones rigurosas" describa condiciones para hibridación y lavado en las que secuencias nucleotídicas que tienen al menos 60%, 65%, 70%, preferiblemente 75% o más identidad entre sí permanezcan típicamente hibridadas entre sí. Dichas condiciones rigurosas son conocidas por los expertos en la técnica y pueden encontrarse en "Current Protocols in Molecular Biology" (John Wiley & Sons, Nueva York (1989)), 6.3.1-6.3.6. Es un ejemplo preferido no limitante de condiciones de hibridación rigurosas la hibridación en 6X cloruro de sodio/citrato de sodio (SSC) aproximadamente a 45°C, seguido de uno o más lavados con 0,2 X SSC, 0,1% de SDS a 50-65°C. En otra realización preferida, las condiciones rigurosas comprenden hibridación en 6 X SSC a 42°C, seguido de lavado con 1 X SSC a 55°C. Preferiblemente, una molécula de ácido nucleico aislada que hibrida en condiciones rigurosas con una secuencia de cinasa de la invención, corresponde a una molécula de ácido nucleico de origen natural. Como se usa en la presente memoria, una molécula de ácido nucleico "de origen natural" designa una molécula de ARN o ADN que tiene una secuencia nucleotídica que aparece en la naturaleza (por ejemplo, codifica una proteína natural).

La presente invención comprende también moléculas de ácido nucleico no codificantes, concretamente, moléculas que son complementarias de un ácido nucleico codificante que codifica una proteína, por ejemplo, complementarias de la cadena de codificación de una molécula de ADN bicatenaria, o complementarias de una secuencia de ARNm. En consecuencia, un ácido nucleico no codificante puede estar unido por enlace de hidrógeno con un ácido nucleico codificante. El ácido nucleico no codificante puede ser complementario de una cadena codificante de cinasa entera o de la región codificante de proteína (o marco abierto de lectura). Las regiones no codificantes son las secuencias 5' y 3' que flanquean la región de codificación y no se traducen en aminoácidos.

Las moléculas de ácido nucleico no codificantes dadas a conocer se administran típicamente a un sujeto o se generan *in situ* de tal modo que hibridan con o se unen a ARNm celular y/o ADN genómico que codifica una proteína cinasa para inhibir así la expresión de la proteína, por ejemplo, inhibiendo la transcripción y/o traducción. Un ejemplo de una vía de administración de moléculas de ácido nucleico no codificantes de la invención incluye la inyección directa en un sitio de tejido. Como alternativa, las moléculas de ácido nucleico no codificantes pueden modificarse para orientarse a células seleccionadas y administrarse después sistémicamente. Por ejemplo, las moléculas no codificantes pueden ligarse a péptidos o anticuerpos formando un complejo que se une específicamente a receptores o antígenos expresados en una superficie celular seleccionada. Las moléculas de ácido nucleico no codificantes pueden suministrarse también a células usando los vectores descritos en la presente memoria. Para conseguir concentraciones intracelulares suficientes de las moléculas no codificantes, se prefieren constructos de vector en los que la molécula de ácido nucleico no codificante está dispuesta bajo el control de un promotor fuerte pol II o pol III.

Se dan a conocer ribozimas, que son moléculas de ARN catalíticas con actividad ribonucleasa que son capaces de escindir un ácido nucleico monocatenario, tal como un ARNm, con el que tienen una región complementaria. Las ribozimas (por ejemplo, ribozimas de cabeza de martillo (descritas en Haselhoff y Gerlach (1988) *Nature* 334: 585-591)) pueden usarse para escindir catalíticamente transcritos de ARNm de cinasa para inhibir así la traducción de ARNm de cinasa. Puede diseñarse una ribozima que tenga especificidad por un ácido nucleico que codifica cinasa basándose en la secuencia nucleotídica del ADN de cinasa dada a conocer en la presente memoria (por ejemplo, SEQ ID NO: 1). Véanse, por ejemplo, Cech y col., patente de EE.UU. n° 4.987.071 y Cech y col., patente de EE.UU. n° 5.116.742. Como alternativa, el ARNm de cinasa puede usarse para seleccionar un ARN catalítico que tenga una actividad ribonucleasa específica de un grupo de moléculas de ARN. Véase, por ejemplo, Bartel y Szostak (1993) *Science* 261: 1411-1418.

Se dan a conocer moléculas de ácido nucleico que forman estructuras de triple hélice. Por ejemplo, la expresión del gen de cinasa puede inhibirse orientando secuencias nucleotídicas complementarias de la región reguladora de la proteína de cinasa (por ejemplo, el promotor y/o potenciadores de cinasa) para formar estructuras de triple hélice que evitan la transcripción del gen de cinasa en células diana. Véanse en general Helene (1991) *Anticancer Drug Des.* 6 (6): 569; Helene (1992) *Ann. N.Y. Acad. Sci.* 660: 27; y Maher (1992) *Bioassays* 14 (12): 807.

En otras realizaciones, las moléculas de ácido nucleico de la invención pueden modificarse en el resto básico, el resto azúcar o la cadena principal de fosfato para mejorar, por ejemplo, la estabilidad, hibridación o solubilidad de la molécula.

II. *Proteínas de cinasa y anticuerpos anti-cinasa aislados*

Las proteínas de cinasa están también comprendidas dentro de la presente invención. Se pretenden por “proteína de cinasa” proteínas que tienen la secuencia aminoacídica expuesta en la SEQ ID NO: 2.

“Fragmentos” o “porciones biológicamente activas” incluyen fragmentos peptídicos adecuados para uso como inmunógenos para crear anticuerpos anti-cinasa. Los fragmentos incluyen péptidos que comprenden secuencias aminoacídicas suficientemente idénticas a o derivadas de las secuencias aminoacídicas de una proteína de cinasa de la invención y que exhiben al menos una actividad de una proteína de cinasa, pero que incluyen menos aminoácidos que las proteínas de cinasa enteras dadas a conocer en la presente memoria. Típicamente, las porciones biológicamente activas comprenden un dominio o motivo con al menos una actividad de proteína de cinasa. Una porción biológicamente activa de una proteína de cinasa puede ser un polipéptido que es, por ejemplo, de 20, 25, 50, 100 ó más aminoácidos de longitud. Dichas porciones biológicamente activas pueden prepararse mediante técnicas recombinantes y evaluarse una o más de las actividades funcionales de una proteína cinasa nativa.

Se pretenden por “variantes” proteínas o polipéptidos que tienen una secuencia aminoacídica que es al menos de aproximadamente 70%, preferiblemente de aproximadamente un 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% idéntica a la secuencia aminoacídica de SEQ ID NO: 2. Las variantes incluyen también polipéptidos codificados por el inserto de ADNc del plásmido depositado en la ATCC con número de acceso PTA-2333, o polipéptidos codificados por una molécula de ácido nucleico que hibrida con una molécula de ácido nucleico de SEQ ID NO: 1 o un complemento de la misma en condiciones rigurosas. Dichas variantes retendrán generalmente la actividad funcional de las proteínas de cinasa de la invención. Las variantes incluyen polipéptidos que difieren en la secuencia aminoacídica debido a la variación alélica natural o a mutagénesis.

La invención proporciona también proteínas quiméricas o de fusión de cinasa. Como se usa en la presente memoria, una “proteína quimérica” o “proteína de fusión” de cinasa comprende un polipéptido de cinasa ligado operativamente a un polipéptido no de cinasa. Un “polipéptido de cinasa” designa un polipéptido que tiene una secuencia aminoacídica correspondiente a una proteína de cinasa, mientras que un “polipéptido no de cinasa” designa un polipéptido que tiene una secuencia aminoacídica correspondiente a una proteína que no es sustancialmente idéntica a la proteína de cinasa, por ejemplo, una proteína que es diferente de la proteína de cinasa y que deriva del mismo organismo o uno diferente. En la proteína de fusión, se pretende que el término “ligado operativamente” que indique que el polipéptido de cinasa y el polipéptido no de cinasa están fusionados en fase entre sí. El polipéptido no de cinasa puede fusionarse con el extremo N o el extremo C del polipéptido de cinasa.

Es una proteína de fusión útil la proteína de fusión GST-cinasa, en la que las secuencias de cinasa están fusionadas con el extremo C de las secuencias GST. Dichas proteínas de fusión pueden facilitar la purificación de proteínas de cinasa recombinantes.

En aún otra realización, la proteína de fusión es una proteína de fusión cinasa-inmunoglobulina en la que toda o parte de una proteína de cinasa está fusionada con secuencias derivadas de un miembro de la familia de proteínas inmunoglobulinas. Las proteínas de fusión de cinasa-inmunoglobulina de la invención pueden incorporarse a composiciones farmacéuticas y administrarse a un sujeto para inhibir una interacción entre un ligando de cinasa y una proteína de cinasa en la superficie de una célula, suprimiendo así la transducción de señal mediada por cinasa *in vivo*. Las proteínas de fusión de cinasa-inmunoglobulina pueden usarse para afectar a la biodisponibilidad de un ligando afín por cinasa. La inhibición de la interacción ligando de cinasa/cinasa puede ser terapéuticamente útil, tanto para tratar trastornos proliferativos y diferenciativos como para modular (por ejemplo, promover o inhibir) la supervivencia celular. Además, las proteínas de fusión de cinasa-inmunoglobulina de la invención pueden usarse como inmunógenos para producir anticuerpos anti-cinasa en un sujeto, para purificar ligandos de cinasa y en ensayos de cribado para identificar moléculas que inhiben la interacción de una proteína de cinasa con un ligando de cinasa.

Preferiblemente, se produce una proteína quimérica o de fusión de cinasa de la invención mediante técnicas de ADN recombinante estándar. Por ejemplo, los fragmentos de ADN que codifican las diferentes secuencias polipeptídicas pueden ligarse en fase entre sí, o puede sintetizarse el gen de fusión, tal como con sintetizadores de ADN automatizados. Como alternativa, puede llevarse a cabo la amplificación por PCR de fragmentos génicos usando cebadores de anclaje que dan lugar a salientes complementarios entre dos fragmentos génicos consecutivos, que pueden asociarse posteriormente y reamplificarse generando una secuencia génica quimérica (véase, por ejemplo, Ausubel y col., eds. (1995) “Current Protocols in Molecular Biology”) (Greene Publishing y Wiley-Interscience, NY). Además, puede clonarse un ácido nucleico que codifica cinasa en un vector de expresión comercialmente disponible de tal modo que esté ligado en fase con un resto de fusión existente. Las variantes de esta proteína de cinasa pueden funcionar como agonistas de cinasa (miméticos) o como antagonistas de cinasa. Las variantes de la proteína de cinasa pueden generarse mediante mutagénesis, por ejemplo, mutación puntual discreta o truncamiento de la proteína de cinasa. Un agonista de la proteína de cinasa puede retener sustancialmente las mismas o un subconjunto de las actividades biológicas de la forma de origen natural de la proteína de cinasa. Un antagonista de la proteína de cinasa puede inhibir una o más de las actividades de la forma de origen natural de la proteína de cinasa, por ejemplo, mediante unión competitiva a un miembro cadena abajo o cadena arriba de una cascada de señalización celular que incluye la proteína de cinasa. Por tanto, pueden desencadenarse efectos biológicos específicos mediante el tratamiento con una variante de función limitada. El tratamiento de un sujeto con una variante que tiene un subconjunto de las actividades biológicas de la

forma de origen natural de la proteína puede tener menos efectos secundarios en un sujeto respecto al tratamiento con la forma de origen natural de las proteínas de cinasa.

Las variantes de la proteína de cinasa que funcionan como agonistas de cinasa o antagonistas de cinasa pueden identificarse cribando en colecciones combinatorias de mutantes, por ejemplo mutantes de truncamiento, de la proteína de cinasa la actividad agonista o antagonista de proteína de cinasa. Se genera una colección variada de variantes de cinasa mediante mutagénesis combinatoria a nivel de ácido nucleico y se codifica por una colección génica variada. Puede producirse una colección variada de variantes de cinasa, por ejemplo, ligando enzimáticamente una mezcla de oligonucleótidos sintéticos con secuencias génicas de tal modo que sea expresable un conjunto degenerado de secuencias de cinasa potenciales como polipéptidos individuales o, como alternativa, en forma de un conjunto de proteínas de fusión más grandes (por ejemplo, para exhibición en fago) que contienen el conjunto de secuencias de cinasa en las mismas. Existe una variedad de procedimientos que pueden usarse para producir colecciones de variantes de cinasa potenciales a partir de una secuencia oligonucleotídica degenerada. Puede efectuarse la síntesis química de una secuencia génica degenerada en un sintetizador de ADN automático, y ligarse después el gen sintético con un vector de expresión apropiado. El uso de un conjunto degenerado de genes permite la provisión, en una mezcla, de todas las secuencias que codifican el conjunto deseado de secuencias de cinasa potenciales. Los procedimientos para sintetizar oligonucleótidos degenerados son conocidos en la técnica (véanse, por ejemplo, Narang (1983) *Tetrahedron* 39: 3; Itakura y col. (1984) *Annu. Rev. Biochem.* 53: 323; Itakura y col. (1984) *Science* 198: 1056; Ike y col. (1983) *Nucleic Acid Res.* 11:477).

Además, pueden usarse colecciones de fragmentos de la secuencia de codificación de proteína de cinasa para generar una población variada de fragmentos de cinasa para el cribado y posterior selección de variantes de una proteína de cinasa. Puede generarse una colección de fragmentos de secuencia de codificación tratando un fragmento de PCR bicatenario de una secuencia de codificación de cinasa con una nucleasa en condiciones en que aparece mellado sólo aproximadamente una vez por molécula, desnaturalizando el ADN bicatenario, renaturalizando el ADN para formar ADN bicatenario que puede incluir pares codificantes/no codificantes de diferentes productos mellados, retirando las porciones monocatenarias de los dúplex reformados mediante tratamiento con nucleasa S1, y ligando la colección de fragmentos resultante en un vector de expresión. Mediante este procedimiento, puede derivarse una colección de expresión que codifica fragmentos N-terminales e internos de diversos tamaños de la proteína de cinasa.

Son conocidos en la técnica varias técnicas para cribar los productos génicos de colecciones combinatorias realizadas mediante mutaciones puntuales o truncamiento y para cribar en colecciones de ADNc los productos génicos que tengan una propiedad seleccionada. Dichas técnicas son adaptables para un cribado rápido de las colecciones génicas generadas mediante mutagénesis combinatoria de proteínas de cinasa. Las técnicas más ampliamente usadas, que son susceptibles de análisis de alto rendimiento para cribado de colecciones génicas grandes, incluyen típicamente clonar la colección génica en vectores de expresión replicables, transformar las células apropiadas con la colección de vectores resultante y expresar los genes combinatorios en condiciones en que la detección de una actividad deseada facilite el aislamiento del vector que codifica el gen cuyo producto se detectó. La mutagénesis de conjunto recursiva (REM), una técnica que potencia la frecuencia de mutantes funcionales en las colecciones, puede usarse en combinación con los ensayos de cribado para identificar variantes de cinasa (Arkin y Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave y col. (1993) *Protein Engineering* 6 (3): 327-331).

Puede usarse un polipéptido de cinasa aislado de la invención como inmunógeno para generar anticuerpos que se unen a proteínas de cinasa usando técnicas estándar para la preparación de anticuerpos policlonales y monoclonales. Puede usarse la proteína de cinasa completa o, como alternativa, la invención proporciona fragmentos peptídicos antígenicos de proteínas de cinasa para uso como inmunógenos. El péptido antígeno de una proteína de cinasa comprende al menos 8, preferiblemente 10, 15, 20 ó 30 residuos aminoácidos de la secuencia aminoácida mostrada en la SEQ ID NO: 2 y comprende un epítipo de una proteína de cinasa de tal modo que un anticuerpo creado contra el péptido forme un complejo inmunitario específico con la proteína de cinasa. Los epítipos preferidos comprendidos por el péptido antígeno son regiones de una proteína de cinasa que están localizadas en la superficie de la proteína, por ejemplo, regiones hidrófilas. El análisis de la región de codificación del polipéptido 14189 predice regiones hidrófilas desde el aminoácido aproximadamente 10 al aminoácido aproximadamente 22, desde el aminoácido aproximadamente 30 hasta el aminoácido aproximadamente 40, desde el aminoácido aproximadamente 50 hasta el aminoácido aproximadamente 60, desde el aminoácido aproximadamente 68 hasta el aminoácido aproximadamente 72, desde el aminoácido aproximadamente 80 hasta el aminoácido aproximadamente 90, desde el aminoácido aproximadamente 132 hasta el aminoácido aproximadamente 142, desde el aminoácido aproximadamente 168 hasta el aminoácido aproximadamente 177, desde el aminoácido aproximadamente 180 hasta el aminoácido aproximadamente 188, desde el aminoácido aproximadamente 213 hasta el aminoácido aproximadamente 250, desde el aminoácido aproximadamente 258 hasta el aminoácido aproximadamente 265, desde el aminoácido aproximadamente 290 hasta el aminoácido aproximadamente 330, desde el aminoácido aproximadamente 335 hasta el aminoácido aproximadamente 340, desde el aminoácido aproximadamente 350 hasta el aminoácido aproximadamente 420, desde el aminoácido aproximadamente 428 hasta el aminoácido aproximadamente 440, desde el aminoácido aproximadamente 472 hasta el aminoácido aproximadamente 480 y desde el aminoácido aproximadamente 485 hasta el aminoácido aproximadamente 508.

En consecuencia, se dan a conocer anticuerpos policlonales y monoclonales anti-cinasa que se unen a una proteína de cinasa. Los anticuerpos policlonales anti-cinasa pueden prepararse inmunizando un sujeto adecuado (por ejemplo, conejo, cabra, ratón u otro mamífero) con un inmunógeno de cinasa. El título de anticuerpo anti-cinasa en el sujeto inmunizado puede controlarse con el tiempo mediante técnicas estándar, tales como con un ensayo de inmunosorción

ligado a enzima (ELISA) usando proteína de cinasa inmovilizada. En un momento apropiado después de la inmunización, por ejemplo cuando los títulos de anticuerpo anti-cinasa son máximos, pueden obtenerse células productoras de anticuerpos a partir del sujeto y usarse para preparar anticuerpos monoclonales mediante técnicas estándar, tales como la técnica de hibridoma descrita originalmente por Kohler y Milstein (1975) *Nature* 256: 495-497, la técnica de hibridoma de linfocitos B humanos (Kozbor y col. (1983) *Immunol. Today* 4: 72), la técnica de hibridoma de EBV (Cole y col. (1985) en "Monoclonal Antibodies and Cancer Therapy", ed. Reisfeld y Sell (Alan R. Liss, Inc., Nueva York, NY), pág. 77-96) o técnicas de trioma. La tecnología para producir hibridomas es bien conocida (véanse, en general, Coligan y col., eds. (1994) "Current Protocols in Immunology" (John Wiley & Sons, Inc., Nueva York, NY); Galfre y col. (1977) *Nature* 266: 55052; Kenneth (1980) en "Monoclonal Antibodies: A New Dimension In Biological Analices" (Plenum Publishing Corp., NY; y Lerner (1981) *Yale J. Biol. Med.*, 54: 387-402).

Como alternativa a preparar hibridomas secretores de anticuerpo monoclonal, puede identificarse y aislarse un anticuerpo monoclonal anti-cinasa cribando una colección combinatoria de inmunoglobulinas recombinantes (por ejemplo, una colección de exhibición en fago de anticuerpo) con una proteína de cinasa, para aislar así miembros de la colección de inmunoglobulinas que se unan a la proteína de cinasa. Están comercialmente disponibles kits para generar y cribar colecciones de exhibición en fago (por ejemplo, el Pharmacia Recombinant Phage Antibody System, n° de catálogo 27-9400-01; y el Stratagene SurfZAP™ Phage Display Kit, n° de catálogo 240612). Adicionalmente, pueden encontrarse ejemplos de procedimientos y reactivos particularmente susceptibles de uso para generar y cribar una colección de exhibición de anticuerpos, por ejemplo, en la patente de EE.UU. n° 5.223.409; las publicaciones PCT n° WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; 93/01288; WO 92/01047; 92/09690 y 90/02809; Fuchs y col. (1991) *Bio/Technology* 9: 1370-1372; Hay y col. (1992) *Hum. Antibod. Hybridomas* 3: 81-85; Huse y col. (1989) *Science* 246: 1275-1281; Griffiths y col. (1993) *EMBO J.* 12: 725-734.

Adicionalmente, se dan a conocer anticuerpos recombinantes anti-cinasa, tales como anticuerpos monoclonales quiméricos y humanizados, que comprenden tanto porciones humanas como no humanas, que pueden prepararse usando técnicas de ADN recombinante estándar. Dichos anticuerpos monoclonales quiméricos y humanizados pueden producirse mediante técnicas de ADN recombinante conocidas en la técnica, por ejemplo, usando procedimientos descritos en las publicaciones PCT n° WO 86/01533 y WO 87/02671; solicitudes de patente europea n° 184.187, 171.496, 125.023 y 173.494; patentes de EE.UU. n° 4.816.567 y 5.225.539; solicitud de patente europea 125.023; Better y col. (1988) *Science* 240: 1041-1043; Liu y col. (1987) *Proc. Natl. Acad. Sci. USA* 84: 3439-3443; Liu y col. (1987) *J. Immunol.* 139: 3521-3526; Sun y col. (1987) *Proc. Natl. Acad. Sci. USA* 84: 214-218; Nishimura y col. (1987) *Canc. Res.* 47: 999-1005; Wood y col. (1985) *Nature* 314: 446-449; Shaw y col. (1988) *J. Natl. Cancer Inst.* 80: 1553-1559; Morrison (1985) *Science* 229: 1202-1207; Oi y col. (1986) *Bio/Techniques* 4: 214; Jones y col. (1986) *Nature* 321: 552-525; Verhoeyan y col. (1988) *Science* 239: 1534 y Beidler y col. (1988) *J. Immunol.* 141: 4053-4060.

Los anticuerpos completamente humanos son particularmente deseables para el tratamiento terapéutico de pacientes humanos. Dichos anticuerpos pueden producirse usando ratones transgénicos que son incapaces de expresar los genes de cadena pesada y ligera de inmunoglobulina endógenos pero que pueden expresar los genes de cadena pesada y ligera humanos. Véanse, por ejemplo, Lonberg y Huszar (1995) *Int. Rev. Immunol.* 13: 65-93; y las patentes de EE.UU. n° 5.625.126; 5.633.425; 5.569.825; 5.661.016 y 5.545.806. Además, puede encargarse a compañías tales como Abgenix, Inc. (Fremont, CA) proporcionar anticuerpos humanos dirigidos contra un antígeno seleccionado usando una tecnología similar a la descrita anteriormente.

Pueden generarse anticuerpos completamente humanos que reconocen un epítipo seleccionado usando una técnica designada como "selección guiada". En este enfoque, se usa un anticuerpo monoclonal no humano seleccionado, por ejemplo, un anticuerpo de murino, para guiar la selección de un anticuerpo completamente humano que reconoce al mismo epítipo. Esta tecnología se describe en Jespers y col. (1994) *Bio/Technology* 12: 899-903).

Puede usarse un anticuerpo anti-cinasa (por ejemplo, anticuerpo monoclonal) para aislar proteínas de cinasa mediante técnicas estándar, tales como cromatografía por afinidad o inmunoprecipitación. Un anticuerpo anti-cinasa puede facilitar la purificación de proteína de cinasa natural a partir de células y de proteína de cinasa producida recombinantemente expresada en células hospedadoras. Además, puede usarse un anticuerpo anti-cinasa para detectar proteína de cinasa (por ejemplo, en un lisado celular o sobrenadante celular) para evaluar la abundancia y el patrón de expresión de la proteína de cinasa. Pueden usarse anticuerpos anti-cinasa en diagnóstico para controlar los niveles de proteína en 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 acoplando el anticuerpo con una sustancia detectable. Los ejemplos de sustancias detectables incluyen diversas enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radiactivos. Los ejemplos de enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina,  $\beta$ -galactosidasa o acetilcolinesterasa; los ejemplos de complejos de grupo prostético incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de materiales fluorescentes adecuados incluyen umbeliferona, fluoresceína, fluoresceinisotiocianato, rodamina, diclorotriazinilaminofluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de material luminiscente incluye luminol; los ejemplos de materiales bioluminiscentes incluyen luciferasa, luciferina y ecuorina; y los ejemplos de material radiactivo adecuado incluyen  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  o  $^3\text{H}$ .

Además, puede conjugarse un anticuerpo (o fragmento del mismo) con un resto terapéutico tal como una citotoxina, un agente terapéutico o un ión metálico radiactivo. Una citotoxina o agente citotóxico incluye cualquier agente que se perjudicial para células. Los ejemplos incluyen taxol, citocalasina B, gramicidina D, bromuro de etidio, emetina,

mitomicina, etopósido, tenopósido, vincristina, vinblastina, colquicina, doxorubicina, daunorubicina, dihidroxiantra-  
cindiona, mitoxantrona, mitramicina, actinomicina D, 1-deshidrotestosterona, glucocorticoides, procaína, tetracaína,  
lidocaína, propranolol y puromicina y análogos u homólogos de los mismos. Los agentes terapéuticos incluyen, pero  
sin limitación, antimetabolitos (por ejemplo, metotrexato, 6-mercaptopurina, 6-tioguanina, citarabina, 5-fluorouracil-  
decarbazona), agentes alquilantes (por ejemplo, mecloretamina, tioepa-clorambucilo, melfalán, carmustina (BSNU) y  
lomustina (CCNU), ciclofosfamida, busulfano, dibromomanitol, estreptoizotocina, mitomicina C y cis-diclorodiami-  
naplatino (II) (DDP) cisplatino), antraciclinas (por ejemplo, daunorubicina (anteriormente daunomicina) y doxorubi-  
cina), antibióticos (por ejemplo, dactinomicina (anteriormente actinomicina), bleomicina, mitramicina y antramicina  
(AMC)), y agentes antimitóticos (por ejemplo, vincristina y vinblastina). Los conjugados pueden usarse para modificar  
una respuesta biológica dada, el resto medicamentoso no ha de considerarse limitado a agentes terapéuticos químicos  
clásicos. Por ejemplo, el resto medicamentoso puede ser una proteína o polipéptido que posea una actividad biológica  
deseada. Dichas proteínas pueden incluir, por ejemplo, una toxina tal como abrina, ricina A, exotoxina de *Pseudomo-  
nas* o toxina de difteria, una proteína tal como factor de necrosis tumoral, interferón alfa, interferón beta, factor de  
crecimiento nervioso, factor de crecimiento derivado de plaquetas, activador de plasminógeno de tejido o modifica-  
dores de la respuesta biológica tales como, por ejemplo, linfocinas, interleucina-1 ("IL-1"), interleucina-2 ("IL-2"),  
interleucina-6 ("IL-6"), factor estimulante de colonia en macrofase de granulocitos ("GM-CSF"), factor estimulante  
de colonia de granulocitos ("G-CSF") u otros factores de crecimiento.

Las técnicas para conjugar dicho resto terapéutico con anticuerpos son bien conocidas, véanse, por ejemplo, Amon  
y col., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", en "Monoclonal Antibodies And  
Cancer Therapy", Reisfeld y col. (eds.), pág. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom y col., "Antibodies For Drug  
Delivery", en "Controlled Drug Delivery" (2ª Ed.), Robinson y col. (eds.), pág. 623-53 (Marcel Dekker, Inc. 1987);  
Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", en "Monoclonal Antibodies '84:  
Biological And Clinical Applications", Pinchera y col. (eds.), pág. 475-506 (1985); "Analysis, Results, And Future  
Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", en "Monoclonal Antibodies For  
Cancer Detection And Therapy", Baldwin y col. (eds.), pág. 303-16 (Academic Press 1985) y Thorpe y col., "The  
Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62: 119-58 (1982). Como  
alternativa, puede conjugarse un anticuerpo con un segundo anticuerpo para formar un heteroconjugado de anticuerpos  
como se describe por Segal en la patente de EE.UU. nº 4.676.980.

### III. Vectores de expresión recombinantes y células hospedadoras

Otro aspecto de la invención está relacionado con vectores, preferiblemente vectores de expresión, que contienen un  
ácido nucleico que codifica una proteína de cinasa de la invención. "Vector" designa una molécula de ácido nucleico  
capaz de transportar otro ácido nucleico al que está ligado, tal como un "plásmido", un bucle de ADN bicatenario  
circular en el que pueden ligarse segmentos de ADN adicionales, o un vector vírico en el que pueden ligarse segmentos  
de ADN adicionales al genoma vírico. Los vectores son útiles para replicación autónoma en una célula hospedadora o  
pueden integrarse en el genoma de una célula hospedadora tras la introducción en la célula hospedadora, y replicarse  
así junto con el genoma hospedador (por ejemplo, vectores de mamífero no episómicos). Los vectores de expresión son  
capaces de dirigir la expresión de genes a los que están ligados operativamente. En general, los vectores de expresión de  
utilidad en técnicas de ADN recombinante están a menudo en forma de plásmidos (vectores). Sin embargo, se pretende  
que la invención incluya dichas otras formas de vectores de expresión tales como vectores víricos (por ejemplo,  
retrovirus defectivos de replicación, adenovirus y virus adenoasociados) que sirven para funciones equivalentes.

Los vectores de expresión recombinantes de la invención comprenden un ácido nucleico de la invención en una for-  
ma adecuada para expresión del ácido nucleico en una célula hospedadora. Esto significa que los vectores de expresión  
recombinantes incluyen una o más secuencias reguladoras, seleccionadas basándose en las células hospedadoras que  
se van a usar para expresión, ligadas operativamente a la secuencia de ácido nucleico a expresar. Se pretende que "liga-  
do operativamente" signifique que la secuencia nucleotídica de interés está ligada a la(s) secuencia(s) reguladora(s) de  
manera que permita la expresión de la secuencia nucleotídica (por ejemplo, en un sistema de transcripción/traducción  
*in vitro* o en una célula hospedadora cuando el vector se introduce en la célula hospedadora). Se pretende que el  
término "secuencia reguladora" incluya promotores, potenciadores y otros elementos de control de la expresión (por  
ejemplo, señales de poliadenilación). Véase, por ejemplo, Goeddel (1990) en "Gene Expression Technology: Methods  
in Enzymology" 185 (Academic Press, San Diego, CA). Las secuencias reguladoras incluyen aquellas que dirigen la  
expresión constitutiva de una secuencia nucleotídica en muchos tipos de células hospedadoras y aquellas que dirigen  
la expresión de la secuencia nucleotídica sólo en ciertas células hospedadoras (por ejemplo, secuencias reguladoras  
específicas de tejido). Se apreciará por los expertos en la técnica que 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,  
etc. Los vectores de expresión de la invención pueden introducirse en células hospedadoras para producir así proteí-  
nas o péptidos, incluyendo proteínas o péptidos de fusión, codificados por ácidos nucleicos como se describen en la  
presente memoria (por ejemplo, proteínas de cinasa, formas mutantes de proteínas de cinasa, proteínas de fusión, etc.).

Los vectores de expresión recombinantes de la invención pueden diseñarse para la expresión de proteína de cinasa  
en células hospedadoras procarióticas o eucarióticas. La expresión de proteínas en procariotas se lleva a cabo lo más  
a menudo en *E. coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de  
proteínas de fusión o no de fusión. Los vectores de fusión añaden una serie de aminoácidos a la proteína codificada en  
los mismos, habitualmente al extremo amino de la proteína recombinante. Los vectores de expresión de fusión típicos

incluyen pGEX (Pharmacia Biotech Inc; Smith y Johnson (1988) *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ) que fusionan glutation-S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, con la proteína recombinante diana. Los ejemplos de vectores de expresión de *E. coli* no de fusión incluyen pTrc (Amann y col. (1988) *Gene* 69: 301-315) y pET 11d (Studier y col. (1990) en “Gene Expression Technology: Methods in Enzymology” 185 (Academic Press, San Diego, CA), pág. 60-89). Las estrategias para maximizar la expresión de proteína recombinante en *E. coli* pueden encontrarse en Gottesman (1990) en “Gene Expression Technology: Methods in Enzymology” 185 (Academic Press, CA), pág. 119-128 y Wada y col. (1992) *Nucleic Acids Res.* 20: 2111-2118. La expresión del gen diana a partir del vector pTrc se basa en la transcripción por ARN polimerasa de hospedador a partir de un promotor de fusión trp-lac híbrido.

Las células hospedadoras eucarióticas adecuadas incluyen células de insecto (los ejemplos de vectores de baculovirus disponibles para expresión de proteínas en células de insecto cultivadas (por ejemplo, células Sf9) incluyen la serie pAc (Smith y col. (1983) *Mol. Cell Biol.* 3: 2156-2165) y la serie pVL (Lucklow y Summers (1989) *Virology* 170: 31-39)); células de levadura (los ejemplos de vectores para expresión en levadura *S. cerevisiae* incluyen pYepSecl (Baldari y col. (1987) *EMBO J.* 6: 229-234), pMFa (Kurjan y Herskowitz (1982) *Cell* 30: 933-943), pJRY88 (Schultz y col. (1987) *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, CA) y pPicZ (Invitrogen Corporation, San Diego, CA)); o células de mamífero (los vectores de expresión de mamífero incluyen pCDM8 (Seed (1987) *Nature* 329: 840) y pMT2PC (Kaufman y col. (1987) *EMBO J.* 6: 187: 195)). Las células de mamífero adecuadas incluyen células de ovario de hámster chino (CHO) o células de riñón de simio transformadas SV40 (COS). En células de mamífero, las funciones de control del vector de expresión se proporcionan a menudo por elementos reguladores víricos. Por ejemplo, los promotores usados normalmente derivan de polioma, adenovirus 2, citomegalovirus y virus de simio 40. Para otros sistemas de expresión adecuados tanto para células procarióticas como eucarióticas, véanse los capítulos 16 y 17 de Sambrook y col. (1989) “Molecular Cloning: A Laboratory Manual” (2ª ed., Cold Spring Harbor Laboratory Press, Plainview, NY). Véase, Goeddel (1990) en “Gene Expression Technology: Methods 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 T7 y polimerasa T7.

Los términos “célula hospedadora” y “célula hospedadora recombinante” se usan intercambiabilmente en la presente memoria. Se entiende que dichos términos designan no sólo la célula objeto particular, sino la progenie o progenie potencial de dicha célula. Debido a que pueden aparecer ciertas modificaciones en generaciones sucesivas debido mutaciones o influencias ambientales, dicha progenie puede no ser idéntica de hecho a la célula original, pero está incluida todavía dentro del alcance del término como se usa en la presente memoria.

En una realización, el vector de expresión es un vector de expresión recombinante de mamífero que comprende elementos reguladores específicos de tejido que dirigen la expresión del ácido nucleico preferiblemente a un tipo celular particular. Los promotores específicos de tejido adecuados incluyen el promotor de albúmina (específico de hígado; Pinkert y col. (1987) *Genes Dev.* 1: 268-277), promotores específicos de linfoides (Calame y Eaton (1988) *Adv. Immunol.* 43: 235-275), promotores particulares de receptores de linfocitos T (Winoto y Baltimore (1989) *EMBO J.* 8: 729-733) e inmunoglobulinas (Banerji y col. (1983) *Cell* 33: 729-740; Queen y Baltimore (1983) *Cell* 33: 741-748), promotores específicos de neuronas (por ejemplo, el promotor de neurofilamento; Byrne y Ruddell (1989) *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), promotores específicos de páncreas (Edlund y col. (1985) *Science* 230: 912-916) y promotores específicos de glándula mamaria (por ejemplo, promotor de suero de leche; patente de EE.UU. n° 4.873.316 y publicación de solicitud de patente europea n° 264.166). Los promotores regulados por el desarrollo están también comprendidos, por ejemplo, el promotor de homeosecuencia de murino (hox) (Kessel y Gruss (1990) *Science* 249: 374-379), el promotor de  $\alpha$ -fetoproteína (Campes y Tilghman (1989) *Genes Dev.* 3: 537-546) y similares.

La invención proporciona adicionalmente un vector de expresión recombinante que comprende una molécula de ADN de la invención clonada en el vector de expresión en orientación no codificante. Es decir, la molécula de ADN está ligada operativamente a una secuencia reguladora de manera que permita la expresión (mediante transcripción de la molécula de ADN) de una molécula de ARN que está no codificante al ARNm de cinasa. Las secuencias reguladoras ligadas operativamente al ácido nucleico clonado en la orientación no codificante pueden elegirse para dirigir la expresión continua de la molécula de ARN no codificante en una variedad de tipos celulares, por ejemplo promotores y/o potenciadores víricos, o pueden elegirse secuencias reguladoras para dirigir la expresión constitutiva, específica de tejido o específica de tipo celular del ARN no codificante. El vector de expresión no codificante puede estar en forma de un plásmido, fagémido o virus atenuado recombinante en el que los ácidos nucleicos no codificantes se producen bajo el control de una región reguladora de alta eficacia, cuya actividad puede determinarse mediante el tipo celular en el que se introduce el vector. Para un debate de la regulación de expresión del gen usando genes no codificantes, véase Weintraub y col. (1986) “Reviews - Trends in Genetics”, vol. 1 (1).

El ADN de vector puede introducirse en células procarióticas o eucarióticas mediante técnicas de transformación o transfección convencionales. Como se usa en la presente memoria, se pretende que los términos “transformación” y “transfección” designen una variedad de técnicas reconocidas en la técnica para introducir ácido nucleico extraño (por ejemplo, ADN) en una célula hospedadora, incluyendo coprecipitación con fosfato de calcio o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección o electroporación. Pueden encontrarse procedimientos adecuados para transformar o transfectar células hospedadoras en Sambrook y col. (1989) “Molecular Cloning: A Laboratory Manual” (2ª ed., Cold Spring Harbor Laboratory Press, Plainview, NY) y otros manuales de laboratorio.

Para la transfección estable de células de mamífero, es sabido que, dependiendo del vector de expresión y de la técnica de transfección usados, sólo una pequeña fracción de las células puede integrar el ADN extraño en su genoma. Para identificar y seleccionar estos integrantes, se introduce generalmente un gen que codifica un marcador seleccionable (por ejemplo, de resistencia a antibióticos) en las células hospedadoras junto con el gen de interés.

- 5 Los marcadores seleccionables preferidos incluyen aquellos que confieren resistencia a medicamentos tales como G418, higromicina y metotrexato. Puede introducirse ácido nucleico que codifica un marcador seleccionable en una célula hospedadora en el mismo vector que codifica la proteína de cinasa o puede introducirse en un vector separado. Las células transfectadas establemente con el ácido nucleico introducido pueden identificarse mediante selección por medicamentos (por ejemplo, las células que han incorporado el gen marcador seleccionable sobrevivirán, mientras que las demás células mueren).

- 15 Puede usarse una célula hospedadora de la invención, tal como una célula hospedadora procariótica o eucariótica en cultivo, para producir (concretamente expresar) proteína de cinasa. En consecuencia, se dan a conocer adicionalmente procedimientos para producir proteína de cinasa usando las células hospedadoras de la invención. En un ejemplo, el procedimiento comprende cultivar la célula hospedadora de la invención, en la que se ha introducido un vector de expresión recombinante que codifica una proteína de cinasa, en un medio adecuado de tal modo que se produzca la proteína de cinasa. En otro ejemplo, el procedimiento comprende adicionalmente aislar la proteína de cinasa del medio o de la célula hospedadora.

- 20 Las células hospedadoras de la invención pueden usarse también para producir animales transgénicos no humanos. Por ejemplo, es una célula hospedadora de la invención un ovocito fertilizado no humano o célula madre embrionaria en el que se han introducido secuencias de codificación de cinasa. Dichas células hospedadoras pueden usarse entonces para crear animales transgénicos no humanos en los que se han introducido secuencias de cinasa exógenas en su genoma o animales recombinantes homólogos en los que se han alterado secuencias de cinasa endógenas. Dichos animales son útiles para estudiar la función y/o actividad de los genes y proteínas de cinasa y para identificar y/o evaluar los moduladores de la actividad cinasa. Como se usa en la presente memoria, un "animal transgénico" es un animal no humano, preferiblemente un mamífero, más preferiblemente un roedor tal como una rata o ratón, en el que una o más de las células del animal incluyen un transgén. Otros ejemplos de animales transgénicos incluyen primates no humanos, ovejas, perros, vacas, cabras, pollos, anfibios, etc. Un transgén es un ADN exógeno que está integrado en el genoma de una célula a partir de la cual se desarrolla un animal transgénico y que permanece en el genoma del animal maduro, dirigiendo así la expresión de un producto génico codificado en uno o más tipos o tejidos celulares del animal transgénico. Como se usa en la presente memoria, un "animal recombinante homólogo" es un animal no humano, preferiblemente un mamífero, más preferiblemente un ratón, en el que se ha alterado un gen de cinasa endógeno mediante recombinación homóloga entre el gen endógeno y una molécula de ADN exógena introducida en una célula del animal, por ejemplo, una célula embrionaria del animal, antes del desarrollo del animal.

- Puede crearse un animal transgénico no humano introduciendo ácido nucleico que codifica cinasa en los pronúcleos masculinos de un ovocito fertilizado, por ejemplo, mediante microinyección o infección retroviral, y permitiendo que el ovocito se desarrolle en un animal hembra pseudopreñado de adopción. La secuencia de ADN de cinasa puede introducirse en forma de un transgén en el genoma de un animal no humano. Como alternativa, puede aislarse un homólogo del gen de cinasa de ratón basándose en la hibridación y usarse como transgén. Pueden incluirse también secuencias intrónicas y señales de poliadenilación en el transgén para aumentar la eficacia de expresión del transgén. Puede(n) ligarse operativamente una(s) secuencia(s) reguladora(s) específica(s) de tejido con el transgén de cinasa para dirigir la expresión de la proteína de cinasa a células particulares. Los procedimientos para generar animales transgénicos mediante manipulación embrionaria y microinyección, particularmente animales tales como ratones, se han convertido en convencionales en la técnica y se describen, por ejemplo, en las patentes de EE.UU. n° 4.736.866, 4.870.009 y 4.873.191 y en Hogan (1986) "Manipulating the Mouse Embryo" (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1986). Se usan procedimientos similares para la producción de otros animales transgénicos. Puede identificarse un animal transgénico fundador basándose en la presencia del gen de cinasa en su genoma y/o la expresión de ARNm de cinasa en tejidos o células de los animales. Puede usarse después un animal transgénico fundador para criar animales adicionales que portan el transgén. Además, los animales transgénicos que portan un transgén que codifica el gen de cinasa pueden criarse adicionalmente hasta otros animales transgénicos que portan otros transgenes.

- 55 Para crear un animal recombinante homólogo, se prepara un vector que contiene al menos una porción de un gen de cinasa u homólogo del gen en el que se ha introducido una delección, adición o sustitución para alterar así, por ejemplo, desestabilizar funcionalmente, el gen de cinasa. En una realización preferida, el vector se diseña de modo que, tras la recombinación homóloga, el gen de cinasa endógeno está funcionalmente desestabilizado (concretamente, ya no codifica una proteína funcional; dichos vectores se designan también como vectores "con inactivación génica"). Como alternativa, el vector puede diseñarse de tal modo que, tras la recombinación homóloga, el gen de cinasa endógeno esté mutado o alterado de otro modo pero siga codificando una proteína funcional (por ejemplo, la región reguladora cadena arriba puede alterarse para alterar así la expresión de la proteína de cinasa endógena). En el vector de recombinación homólogo, la porción alterada del gen de cinasa está flanqueada en sus extremos 5' y 3' por ácido nucleico adicional del gen de cinasa para permitir que aparezca recombinación homóloga entre el gen de cinasa exógeno portado por el vector y un gen de cinasa endógeno en una célula madre embrionaria. El ácido nucleico de cinasa flanqueante adicional es de longitud suficiente para una recombinación homóloga exitosa con el gen endógeno. Típicamente, se incluyen en el vector varios kilobases de ADN flanqueante (en ambos extremo 5' y 3') (véase, por ejemplo, Thomas y Capecchi (1987) *Cell* 51: 503 para una descripción de vectores de recombinación homólogos). Se introduce el vector

en una línea de células madre embrionarias no humanas (por ejemplo, mediante electroporación), y se seleccionan las células en las que se ha recombinado homológamente el gen de cinasa introducido con el gen de cinasa endógeno (véase, por ejemplo, Li y col. (1992) *Cell* 69: 915). Se inyectan después las células seleccionadas en un blastocito de un animal no humano (por ejemplo, un ratón) para formar quimeras de agregación (véase, por ejemplo, Bradley (1987) en "Teratocarcinomas and Embryonic Stem Cells: A Practical Approach", ed. Robertson (IRL, Oxford), pág. 113-152). Puede implantarse después un embrión quimérico en un animal hembra pseudopreñado de adopción y llevarse a término el embrión. Puede usarse la progenie que alberga el ADN recombinado homológamente en sus células germinales para criar animales en los que todas las células del animal contienen el ADN recombinado homológamente mediante transmisión de línea germinal del transgén. Se describen adicionalmente procedimientos para construir vectores de recombinación homólogos y animales recombinantes homólogos en Bradley (1991) *Current Opinion in BioTechnology* 2:823-829 y las publicaciones PCT n° WO 90/11354, WO 91/01140, WO 92/0968 y WO 93/04169.

Pueden producirse animales transgénicos no humanos que contienen sistemas seleccionados que permiten la expresión regulada del transgén. Es un ejemplo de dicho sistema el sistema de recombinasa cre/l oxP del bacteriófago P1. Para una descripción del sistema de recombinasa cre/l oxP, véase, por ejemplo, Lakso y col. (1992) *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Es otro ejemplo de sistema de recombinasa el sistema de recombinasa FLP de *Saccharomyces cerevisiae* (O'Gorman y col. (1991) *Science* 251: 1351-1355). Si se usa un sistema de recombinasa cre/l oxP para regular la expresión del transgén, son necesarios animales que contengan transgenes que codifiquen tanto la recombinasa Cre como una proteína seleccionada. Dichos animales pueden proporcionarse mediante la construcción de animales transgénicos "dobles", por ejemplo, apareando dos animales transgénicos, uno que contiene un transgén que codifica una proteína seleccionada y el otro que contiene un transgén que codifica una recombinasa.

Pueden producirse también clones de los animales transgénicos no humanos descritos en la presente memoria según los procedimientos descritos en Wilmut y col. (1997) *Nature* 385: 810-813 y las publicaciones PCT n° WO 97/07668 y WO 97/07669.

#### IV. Composiciones farmacéuticas

Pueden incorporarse moléculas de ácido nucleico de cinasa y proteínas de cinasa de la invención, y moduladores de la misma (por ejemplo, anticuerpos anti-cinasa) (también designados en la presente memoria como "compuestos activos") a composiciones farmacéuticas adecuadas para administración. Dichas composiciones comprenden típicamente la molécula de ácido nucleico, proteína o moduladores (por ejemplo, anticuerpo o moléculas pequeñas y un vehículo farmacéuticamente aceptable). Como se usa en la presente memoria, se pretende que la frase "vehículo farmacéuticamente aceptable" incluya todos y cada uno de los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción y similares, compatibles con la administración farmacéutica. El uso de dichos medios y agentes para sustancias farmacéuticamente activas es bien conocido en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con el compuesto activo, se contempla el uso del mismo en las composiciones. Pueden incorporarse también compuestos activos suplementarios a las composiciones.

Las composiciones son útiles para tratar cualquiera de los trastornos debatidos en la presente memoria. Las composiciones se proporcionan en cantidades terapéuticamente eficaces. Se pretende por "cantidades terapéuticamente eficaces" una cantidad suficiente para modular la respuesta deseada. Como se define en la presente memoria, una cantidad terapéuticamente eficaz de proteína o polipéptido (concretamente, una dosificación eficaz) está en el intervalo 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 y aún más preferiblemente de aproximadamente 1 a 10 mg/kg, 2 a 9 mg/kg, 3 a 8 mg/kg, 4 a 7 mg/kg ó 5 a 6 mg/kg de peso corporal.

El experto en la técnica apreciará que ciertos factores pueden influir en la dosificación necesaria para tratar eficazmente un sujeto, incluyendo pero sin limitación, la gravedad de la enfermedad o trastorno, tratamientos anteriores, la salud general y/o edad del sujeto y otras enfermedades presentes. Además, el tratamiento 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 un ejemplo preferido, se trata un sujeto con anticuerpo, proteína o polipéptido en el intervalo de entre aproximadamente 0,1 a 20 mg/kg de peso corporal una vez por semana durante entre aproximadamente 1 a 10 semanas, preferiblemente entre 2 a 8 semanas, más preferiblemente entre aproximadamente 3 a 7 semanas, y aún más preferiblemente durante aproximadamente 4, 5 ó 6 semanas. Se apreciará también que la dosificación eficaz de anticuerpo, proteína o polipéptido usada para el tratamiento puede aumentarse o reducirse durante el transcurso de un tratamiento particular. Pueden resultar cambios de dosificación y resultan evidentes a partir de los resultados de los ensayos de diagnóstico como se describen en la presente memoria.

Se dan a conocer agentes que modulan la expresión o la actividad. Un agente puede ser, por ejemplo, una molécula pequeña. Por ejemplo, dichas moléculas pequeñas incluyen, pero sin limitación, péptidos, peptidomiméticos, aminoácidos, análogos aminoácidos, polinucleótidos, análogos polinucleótidos, nucleótidos, análogos nucleotídicos, compuestos orgánicos o inorgánicos (concretamente, incluyendo compuestos heteroorgánicos y organometálicos) que tengan un peso molecular menor de aproximadamente 10.000 g por mol, compuestos orgánicos o inorgánicos que tengan un peso molecular menor de aproximadamente 5.000 g por mol, compuestos orgánicos o inorgánicos que tengan un peso molecular menor de aproximadamente 1.000 g por mol, compuestos orgánicos o inorgánicos que ten-

gan un peso molecular menor de aproximadamente 500 g por mol, y sales, ésteres y otras formas farmacéuticamente aceptables de dichos compuestos.

Se entiende que las dosis apropiadas de agentes de molécula pequeña dependen de una serie de factores dentro del conocimiento del médico, veterinario o investigador experto en la técnica. La(s) dosis de molécula pequeña variará(n), por ejemplo, dependiendo de la identidad, tamaño y condición del sujeto o muestra que se esté tratando, dependiendo adicionalmente de la vía mediante la que se va a administrar la composición, si es aplicable, y del efecto que desea el facultativo que tenga la molécula pequeña sobre el ácido nucleico o polipéptido de la invención. Las dosis ejemplares incluyen cantidades de miligramos o microgramos de la molécula pequeña por kilogramo de peso de sujeto o muestra (por ejemplo, aproximadamente 1  $\mu\text{g/kg}$  a aproximadamente 500 mg/kg, aproximadamente 100  $\mu\text{g/kg}$  a aproximadamente 5 mg/kg, o aproximadamente 1  $\mu\text{g/kg}$  a aproximadamente 50  $\mu\text{g/kg}$ ). Se entiende además 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. Dichas dosis apropiadas pueden determinarse usando los ensayos descritos en la presente memoria. Cuando se van a administrar una o más de estas moléculas pequeñas 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 inicio, aumentando posteriormente la dosis hasta obtener una respuesta apropiada. Además, se entiende que el nivel de dosis específico para cualquier sujeto animal particular dependerá de una variedad de factores, incluyendo la actividad del compuesto específico empleado, la edad, peso corporal, salud general, género y dieta del sujeto, el momento de administración, la vía de administración, la tasa de excreción, cualquier combinación de medicamentos y el grado de expresión o actividad a modular.

Se formula una composición farmacéutica para ser compatible con su vía de administración pretendida. Los ejemplos de vías de administración incluyen administración parenteral, por ejemplo, intravenosa, intradérmica, subcutánea, oral (por ejemplo, por inhalación), transdérmica (tópica), transmucosa y rectal. Las disoluciones o suspensiones usadas para administración parenteral, intradérmica o subcutánea pueden incluir los siguientes componentes: un diluyente estéril tal como agua para inyecciones, disolución salina, aceites no volátiles, polietilenglicoles, glicerina, propilenglicol u otros disolventes sintéticos; agentes antibacterianos tales como alcohol bencílico o metilparabenos; antioxidantes tales como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como ácido etilendiaminotetraacético; tampones tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa. El pH puede ajustarse con ácidos o bases tales como ácido clorhídrico o hidróxido de sodio. La preparación parenteral puede incluirse en ampollas, jeringuillas desechables o viales de dosis múltiples hechos de vidrio o plástico.

Las composiciones farmacéuticas adecuadas para uso inyectable incluyen disoluciones acuosas estériles (cuando son hidrosolubles) o dispersiones y polvos estériles para la preparación extemporánea de disoluciones o dispersiones inyectables estériles. Para administración intravenosa, los vehículos adecuados incluyen disolución salina fisiológica, agua bacteriostática, Cremophor EL<sup>TM</sup> (BASF; Parsippany, NJ) o disolución salina tamponada con fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida en la medida en que exista una fácil inyectabilidad. Debe ser estable en las condiciones de fabricación y almacenamiento y debe conservarse frente a 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 o polietilenglicol líquido, y similares) y mezclas adecuadas de los mismos. Puede mantenerse la fluidez apropiada, por ejemplo, mediante el uso de recubrimientos tales como lecitina, mediante el mantenimiento del tamaño de partícula necesario en el caso de dispersión y mediante el uso de tensioactivos. Puede conseguirse la prevención de la acción de microorganismos mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutano, fenol, ácido ascórbico, timerosal y similares. En muchos casos, será preferible incluir en la composición agentes isotónicos, por ejemplo, azúcares y polialcoholes tales como manitol, sorbitol, cloruro de sodio. La absorción prolongada de las composiciones inyectables puede producirse mediante la inclusión en composición de un agente que retarde la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las disoluciones inyectables estériles pueden prepararse incorporando el compuesto activo (por ejemplo, una proteína de cinasa o anticuerpo anti-cinasa) en la cantidad necesaria a un disolvente apropiado con uno o una combinación de ingredientes enumerados anteriormente, según sea necesario, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el compuesto activo a un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes necesarios de los enumerados anteriormente. En el caso de polvos estériles para la preparación de disoluciones inyectables estériles, son procedimientos de preparación preferidos secado a vacío y liofilización, que proporcionan un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una disolución del mismo esterilizada por filtración anteriormente.

Las composiciones orales incluyen generalmente un diluyente inerte o un vehículo comestible. Pueden introducirse en cápsulas de gelatina o comprimirse en comprimidos. Con fines de administración terapéutica oral, el compuesto activo puede incorporarse a excipientes y usarse en forma de comprimidos, trociscos o cápsulas. Las composiciones orales pueden prepararse también usando un vehículo líquido para uso como colutorio, en el que el compuesto en el vehículo fluido se administra por vía oral y se enjuaga, escupe o traga. Pueden incluirse agentes aglutinantes y/o materiales coadyuvantes farmacéuticamente compatibles como parte de la composición. Los comprimidos, cápsulas, trociscos y similares pueden contener cualquiera de los siguientes ingredientes o compuestos de naturaleza similar: un aglutinante tal como celulosa microcristalina, goma de tragacanto o gelatina; un excipiente tal como almidón o lactosa; un agente disgregante 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 piperita, salicilato de metilo o aroma de naranja. Para administración mediante inhalación, los compuestos se suministran en forma de un pulverizador en aerosol a partir de un envase a presión o dispensador que contiene un propelente adecuado, por ejemplo, un gas tal como dióxido de carbono, o un nebulizador.

5 La administración sistémica puede ser también por medios transmucosos o transdérmicos. Para administración transmucosa o transdérmica, se usan en la formulación penetrantes apropiados para la barrera a permear. Dichos penetrantes son generalmente conocidos en la técnica e incluyen, por ejemplo para administración transmucosa, detergentes, sales biliares y derivados de ácido fusídico. La administración transmucosa puede realizarse mediante el uso de pulverizadores nasales o supositorios. Para administración transdérmica, se formulan los compuestos activos  
10 en ungüentos, pomadas, geles o cremas como se conoce en general en la técnica. Los compuestos pueden prepararse también en forma de supositorios (por ejemplo, con bases de supositorio convencionales tales como manteca de cacao y otros glicéridos) o enemas de retención para suministro rectal.

Los compuestos activos pueden prepararse con vehículos que protegerán al compuesto frente a una eliminación  
15 rápida del cuerpo, tales como una formulación de liberación controlada, incluyendo implantes y sistemas de suministro microencapsulados. Pueden usarse polímeros biodegradables biocompatibles tales como acetato de etilvinilo, polianhídridos, poli(ácido glicólico), colágeno, poliortoésteres y poli(ácido láctico). Los procedimientos para la preparación de dichas formulaciones serán evidentes para los expertos en la técnica. Los materiales pueden obtenerse también comercialmente a partir de Alza Corporation y Nova Pharmaceuticals, Inc. Pueden usarse también suspensiones liposómicas (incluyendo liposomas orientados a células infectadas con anticuerpos monoclonales de antígenos  
20 víricos) como vehículos farmacéuticamente aceptables. Estos pueden prepararse según procedimientos conocidos en la técnica, por ejemplo, como se describen en la patente de EE.UU. n° 4.522.811.

Es especialmente ventajoso formular composiciones orales o parenterales en forma de dosificación unitaria para facilitar la administración y la uniformidad de la dosificación. Forma de dosificación unitaria como se usa en la presente memoria designa unidades físicamente discretas adecuadas como dosificaciones unitarias para el sujeto que se va a tratar, conteniendo cada unidad una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico necesario. Dependiendo del tipo y la gravedad de la enfermedad, de aproximadamente 1  $\mu\text{g/kg}$  a aproximadamente 15 mg/kg (por ejemplo, 0,1 a 20 mg/kg) de anticuerpo es una dosificación inicial candidata para administración al paciente tanto, por ejemplo, mediante una o más  
30 administraciones separadas o mediante infusión continua. Una dosificación diaria típica podría estar en el intervalo de aproximadamente 1  $\mu\text{g/kg}$  a aproximadamente 100 mg/kg o más, dependiendo de los factores mencionados anteriormente. Para administraciones repetidas durante varios días o más tiempo, dependiendo de la afección, el tratamiento se mantiene hasta que aparece la supresión deseada de los síntomas de enfermedad. Sin embargo, pueden ser útiles otros regímenes de dosificación. Se controla fácilmente la progresión de esta terapia mediante técnicas y ensayos convencionales. Se da a conocer un régimen de dosificación ejemplar en el documento 94/04188. Las especificaciones de las formas de dosificación unitarias están dictadas por, y dependen directamente de, las características únicas del compuesto activo y el efecto terapéutico particular a conseguir, y de las limitaciones inherentes a la técnica de combinación de dicho compuesto activo para el tratamiento de individuos.

Las moléculas de ácido nucleico de la invención pueden insertarse en vectores y usarse como vectores de terapia  
40 génica. Los vectores de terapia génica pueden suministrarse a un sujeto, por ejemplo, mediante inyección intravenosa, administración local (patente de EE.UU. n° 5.328.470) o mediante inyección estereotáctica (véase, por ejemplo, Chen y col. (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 embebe el vehículo de suministro génico. Como alternativa, cuando el vector de suministro génico puede producirse intacto a partir de células recombinantes, por ejemplo vectores retrovíricos, la preparación farmacéutica puede incluir una o más células que producen el sistema de suministro génico.

50 Las composiciones farmacéuticas pueden incluirse en un envase, paquete o dispensador junto con instrucciones de administración.

## V. Usos y procedimientos

55 Las moléculas de ácido nucleico, proteínas homólogos proteicos y anticuerpos descritos en la presente memoria pueden usarse en uno o más de los siguientes procedimientos: (a) ensayos de cribado; (b) ensayos de detección (por ejemplo, cartografía cromosómica, tipado de tejido, biología forense); (c) medicina predictiva (por ejemplo, ensayos de diagnóstico, ensayos de pronóstico, control de ensayos clínicos y farmacogenómica) y (d) procedimientos de tratamiento (por ejemplo, terapéutico y profiláctico). La molécula de ácido nucleico aislada de la invención puede usarse para expresar proteína de cinasa (por ejemplo, mediante un vector de expresión recombinante en una célula hospedadora en aplicaciones de terapia génica), para detectar ARNm de cinasa (por ejemplo, en una muestra biológica) o una  
60 lesión genética en un gen de cinasa, y para modular la actividad de cinasa. Además, la proteína de cinasa puede usarse para cribar medicamentos o compuestos que modulan el crecimiento y/o el metabolismo celular así como para tratar trastornos caracterizados por una producción insuficiente o excesiva de proteína de cinasa o la producción de formas de proteína de cinasa que tienen actividad reducida o aberrante en comparación con proteína de tipo silvestre de cinasa. Además, los anticuerpos anti-cinasa de la invención pueden usarse para detectar y aislar proteínas de cinasa y modular la actividad de cinasa.

## A. Ensayos de cribado

La invención proporciona un procedimiento (también designado en la presente memoria como un “ensayo de cribado”) para identificar moduladores, concretamente candidatos o compuestos de ensayo o agentes (por ejemplo, péptidos, peptidomiméticos, moléculas pequeñas u otros medicamentos) que se unen a las proteínas de cinasa de la invención o tienen un efecto estimulante o inhibidor, por ejemplo, sobre la expresión de cinasa o la actividad cinasa.

Los compuestos de ensayo pueden obtenerse usando cualquiera de los numerosos enfoques de procedimientos de colecciones combinatorias conocidos en la técnica, incluyendo colecciones biológicas, colecciones en fase sólida o fase de disolución paralelas espacialmente referenciables, procedimientos de colección sintética que requieren desconvolución, el procedimiento de colección “una perla-un compuesto” y procedimientos de colección sintética que usan selección por cromatografía de afinidad. El enfoque de colección biológica está limitado a colecciones peptídicas, mientras que los otros cuatro enfoques son aplicables a colecciones de compuestos de péptidos, oligómeros no peptídicos o moléculas pequeñas (Lam (1997) *Anticancer Drug Des.* 12: 145).

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

Las colecciones de compuestos pueden presentarse en disolución (por ejemplo, Houghten (1992) *Bio/Techniques* 13: 412-421), o en perlas (Lam (1991) *Nature* 354: 82-84), chips (Fodor (1993) *Nature* 364: 555-556), bacterias (patente de EE.UU. n° 5.223.409), esporas (patentes de EE.UU. n° 5.571.698; 5.403.484. y 5.223.409), plásmidos (Cull y col. (1992) *Proc. Natl. Acad. Sci. USA* 89: 865-1869), o fago (Scott y Smith (1990) *Science* 249: 386-390; Devlin (1990) *Science* 249: 404-406; Cwirla y col. (1990) *Proc. Natl. Acad. Sci. USA* 87: 6378-6382 y Felici (1991) *J. Mol. Biol.* 222: 301-310).

Puede conseguirse determinar la capacidad del compuesto de ensayo de unirse a la proteína de cinasa, por ejemplo, acoplando el compuesto de ensayo con un radioisótopo o marcaje enzimático de tal modo que la unión del compuesto de ensayo a la proteína de cinasa o porción biológicamente activa de la misma pueda determinarse detectando el compuesto marcado en un complejo. Por ejemplo, los compuestos de ensayo pueden marcarse con  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$  o  $^3\text{H}$ , directa o indirectamente, y detectarse el radioisótopo mediante recuento directo de la radioemisión o mediante recuento de centelleo. Como alternativa, los compuestos de ensayo pueden marcarse enzimáticamente, por ejemplo, con peroxidasa de rábano picante, fosfatasa alcalina o luciferasa, y detectarse el marcador enzimático mediante la determinación de la conversión de un sustrato apropiado en producto.

De manera similar, puede determinarse la capacidad de la proteína de cinasa de unirse a o interaccionar con una molécula diana de cinasa. Se pretende por “molécula diana” una molécula con la que una proteína de cinasa se une o interacciona en la naturaleza. En una realización preferida, la capacidad de la proteína de cinasa de unirse a o interaccionar con una molécula diana de cinasa puede determinarse controlando la actividad de la molécula diana. Por ejemplo, puede controlarse la actividad de la molécula diana detectando la inducción de un segundo mensajero celular de la diana (por ejemplo,  $\text{Ca}^{2+}$  intracelular, diacilglicerol, IP3, etc.), detectando la actividad catalítica/enzimática de la diana sobre un sustrato apropiado, detectando la inducción de un gen informador (por ejemplo, un elemento regulador sensible a cinasa ligado operativamente a un ácido nucleico que codifica un marcador detectable, por ejemplo luciferasa) o detectando una respuesta celular, por ejemplo, crecimiento, diferenciación o proliferación celulares.

En aún otra realización, es un ensayo de la presente invención un ensayo exento de células que comprende poner en contacto una proteína de cinasa o porción biológicamente activa de la misma con un compuesto de ensayo y determinar la capacidad del compuesto de ensayo de unirse a la proteína de cinasa o porción biológicamente activa de la misma. La unión del compuesto de ensayo a la proteína de cinasa puede determinarse directa o indirectamente como se describe anteriormente. En una realización preferida, el ensayo incluye poner en contacto la proteína de cinasa o porción biológicamente activa de la misma con un compuesto conocido que se une a proteína de cinasa formando 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 unirse preferiblemente a proteína de cinasa o una porción biológicamente activa de la misma en comparación con el compuesto conocido.

En otra realización, es un ensayo un ensayo exento de células que comprende poner en contacto la proteína de cinasa o porción biológicamente activa de la misma con un compuesto de ensayo y determinar la capacidad del compuesto de ensayo de modular (por ejemplo, estimular o inhibir) la actividad de la proteína de cinasa o porción biológicamente activa de la misma. Puede conseguirse determinar la capacidad del compuesto de ensayo de modular la actividad de la proteína de cinasa, por ejemplo, determinando la capacidad de la proteína de cinasa de unirse a una molécula diana de cinasa como se describe anteriormente para determinar la unión directa. En una realización alternativa, puede conseguirse determinar la capacidad del compuesto de ensayo de modular la actividad de una proteína de cinasa determinando la capacidad de la proteína de cinasa de modular adicionalmente una molécula diana de cinasa. Por ejemplo, puede determinarse la actividad catalítica/enzimática de la molécula diana sobre un sustrato apropiado como se describe anteriormente.

En aún otra realización, el ensayo exento de células comprende poner en contacto la proteína de cinasa o porción biológicamente aceptable de la misma con un compuesto conocido que se une a una proteína de cinasa formando 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 unirse preferiblemente a o modular la actividad de una molécula diana de cinasa.

En los ensayos anteriormente mencionados, puede ser deseable inmovilizar una proteína de cinasa o su molécula diana para facilitar la separación de las formas complejadas de las no complejadas de una o ambas de las proteínas, así como para servir para la automatización del ensayo. En una realización, puede proporcionarse una proteína de fusión que añade un dominio que permite a una o ambas proteínas unirse a una matriz. Por ejemplo, las proteínas de fusión glutation-S-transferasa/cinasa o proteínas de fusión glutation-S-transferasa/diana pueden adsorberse sobre perlas de Sepharose de glutation (Sigma Chemical, St. Louis, MO) o placas de microvaloración derivatizadas con glutation, que se combinan después con el compuesto de ensayo o el compuesto de ensayo y la proteína diana o proteína de cinasa no adsorbida, y se incuba la mezcla en condiciones tendentes a la formación de complejo (por ejemplo, en condiciones fisiológicas de sal y pH). Después de la incubación, se lavan las perlas o pocillos de la placa de microvaloración para retirar cualquier componente no unido y se mide la formación de complejo directa o indirectamente, por ejemplo como se describe anteriormente. Como alternativa, pueden disociarse los complejos de la matriz, y determinarse el nivel de unión o la actividad cinasa usando técnicas estándar.

Pueden usarse también otras técnicas para inmovilizar proteínas sobre matrices en los ensayos de cribado de la invención. Por ejemplo, pueden inmovilizarse proteína cinasa o su molécula diana utilizando la conjugación de biotina y estreptavidina. Pueden prepararse moléculas de cinasa o moléculas diana biotiniladas a partir de biotina-NHS (*N*-hidroxisuccinimida) usando técnicas bien conocidas en la técnica (por ejemplo, kit de biotinilación, Pierce Chemicals, Rockford, IL), e inmovilizarse en los pocillos de placas de 96 pocillos recubiertas con estreptavidina (Pierce Chemicals). Como alternativa, los anticuerpos reactivos con una proteína de cinasa o moléculas diana, pero que no interfieren con la unión de la proteína de cinasa a su molécula diana, pueden derivatizarse en los pocillos de la placa, y atraparse en los pocillos la diana o proteína de cinasa no unida mediante conjugación de anticuerpos. Los procedimientos para detectar dichos complejos, además de los descritos anteriormente para complejos inmovilizados en GST, incluyen la inmunodetección de complejos usando anticuerpos reactivos con la proteína de cinasa o molécula diana, así como ensayos ligados a enzima que se basan en la detección de una actividad enzimática asociada a la proteína de cinasa o la molécula diana.

En otro ejemplo, se identifican los moduladores de expresión de cinasa en un procedimiento en el que se pone en contacto una célula con un compuesto candidato y se determina la expresión de ARNm o proteína de cinasa en la célula respecto a la expresión de ARNm o proteína de cinasa en una célula en ausencia del compuesto candidato. Cuando la expresión es mayor (mayor de forma estadísticamente significativa) en presencia del compuesto candidato que en su ausencia, el compuesto candidato se identifica como estimulante de la expresión de ARNm o proteína de cinasa. Como alternativa, cuando la expresión es menor (menor de forma estadísticamente significativa) que en su ausencia, el compuesto candidato se identifica como un inhibidor de la expresión de ARNm o proteína de cinasa. El nivel de expresión de ARNm o proteína de cinasa en las células puede determinarse mediante procedimientos descritos en la presente memoria para detectar ARNm o proteína de cinasa.

En aún otro aspecto de la invención, las proteínas de cinasa pueden usarse como “proteínas cebo” en un ensayo de doble híbrido o ensayo de triple híbrido (véanse, por ejemplo, la patente de EE.UU. n° 5.283.317; Zervos y col. (1993) *Cell* 72: 223-232; Madura y col. (1993) *J. Biol. Chem.* 268: 12046-12054; Bartel y col. (1993) *Bio/Techniques* 14: 920-924; Iwabuchi y col. (1993) *Oncogene* 8: 1693-1696 y la publicación PCT n° WO 94/10300), para identificar otras proteínas que se unen o interaccionan con proteína de cinasa (“proteínas de unión a cinasa” o “cinasa-bp”) y modulan la actividad cinasa. Dichas proteínas de unión a cinasa es también probable que estén implicadas en la propagación de señales por las proteínas de cinasa como, por ejemplo, elementos cadena arriba o cadena abajo de una ruta de señalización.

## B. Ensayos de detección

Pueden usarse porciones o fragmentos de una secuencia de ADNc identificada en la presente memoria (y la correspondiente secuencia génica completa) de numerosas maneras como reactivos polinucleotídicos. Por ejemplo, la secuencia puede usarse para: (1) cartografiar el gen respectivo en un cromosoma; (2) identificar un individuo de una muestra biológica mínima (tipado de tejido) y (3) ayudar a la identificación forense de una muestra biológica. Estas aplicaciones se describen en las subsecciones siguientes.

### 1. Cartografía cromosómica

Puede usarse la secuencia de ADNc de cinasa aislada de la invención para cartografiar el gen de cinasa respectivo en un cromosoma, facilitando así la localización de regiones génicas asociadas a una enfermedad genética. Puede usarse el análisis informático de la secuencia de cinasa para seleccionar rápidamente cebadores de PCR (preferiblemente de 15-25 pb de longitud) que no cubren más de un exón en el ADN genómico, simplificando así el proceso de amplificación.

Otras estrategias de cartografía que pueden usarse de forma similar para cartografiar una secuencia de cinasa en su cromosoma incluyen hibridación *in situ* (descrita en Fan y col. (1990) *Proc. Natl. Acad. Sci. USA* 87: 6223-27), precribado con cromosomas clasificados por flujo marcados y preselección mediante hibridación con colecciones de

ADNc específico de cromosoma. Además, puede usarse hibridación *in situ* con fluorescencia (FISH) de una secuencia de ADN con una dispersión cromosómica en metafase para proporcionar una localización cromosómica precisa en una etapa. Para una revisión de esta técnica, véase Verma y col. (1988) "Human Chromosomes: A Manual of Basic Techniques" (Pergamon Press, NY). La técnica FISH puede usarse con una secuencia de ADN tan corta como de 500 ó 600 bases. Sin embargo, los clones mayores de 1.000 bases tienen una mayor probabilidad de unirse a una localización cromosómica única con una intensidad de señal suficiente para una detección sencilla. Preferiblemente, 1.000 bases y más, preferiblemente 2.000 bases, bastarán para conseguir buenos resultados en un intervalo de tiempo razonable.

Pueden usarse individualmente reactivos para cartografía cromosómica para marcar un solo cromosoma o un solo sitio en ese cromosoma, o pueden usarse paneles o reactivos para marcar múltiples sitios y/o múltiples cromosomas. En realidad, se prefieren los reactivos correspondientes a las regiones de no codificación de los genes con fines de cartografía. Las secuencias de codificación es más probable que estén conservadas dentro de las familias génicas, aumentando así la posibilidad de hibridaciones cruzadas durante la cartografía cromosómica.

Otra estrategia para cartografiar la localización cromosómica de genes de cinasa usa polipéptidos y fragmentos y secuencias de la presente invención y anticuerpos específicos de los mismos. Esta cartografía puede llevarse a cabo detectando específicamente la presencia de un polipéptido de cinasa en miembros de un panel de híbridos de células somáticas entre células de una primera especie de animal de la que se genera la proteína y células de una segunda especie de animal, y determinando después cuál(es) híbrido(s) de células somáticas expresa(n) el polipéptido y anotando el(los) cromosoma(s) de la primera especie animal que lo contiene. Para ejemplos de esta técnica, véanse, Pajunen y col. (1988) *Cytogenet. Cell. Genet.* 47: 37-41 y Van Keuren y col. (1986) *Hum. Genet.* 74: 34-40. Como alternativa, puede determinarse la presencia de un polipéptido de cinasa en los híbridos de células somáticas ensayando una actividad o propiedad del polipéptido, por ejemplo, actividad enzimática, como se describe en Bordelon-Riser y col. (1979) *Somatic Cell Genetics* 5: 597-613 y Owerbach y col. (1978) *Proc. Natl. Acad. Sci. USA* 75: 5640-5644.

Una vez se ha cartografiado una secuencia en una localización cromosómica precisa, la posición física de la secuencia en el cromosoma puede correlacionarse con los datos del mapa genético. (Dichos datos se encuentran, por ejemplo, en V. McKusick, "Mendelian Inheritance in Man", disponible en línea en la Johns Hopkins University Welch Medical Library). La relación entre genes y enfermedad cartografiados en la misma región cromosómica puede identificarse después mediante análisis de ligamiento (co-herencia de genes físicamente adyacentes), descrito, por ejemplo, en Egeland y col. (1987) *Nature* 325: 783-787.

Además, pueden determinarse las diferencias en las secuencias de ADN entre individuos afectados y no afectados por una enfermedad asociada al gen de cinasa. Si se observa una mutación en algunos o todos los individuos afectados, pero no en los individuos no afectados, entonces es probable que la mutación sea el agente causante de la enfermedad particular. La comparación de individuos afectados y no afectados implica generalmente en primer lugar buscar alteraciones estructurales en los cromosomas tales como deleciones o translocaciones que sean visibles en dispersiones cromosómicas o detectables usando PCR basada en esa secuencia de ADN. Por último, puede efectuarse la secuenciación completa de los genes de varios individuos para confirmar la presencia de una mutación y distinguir las mutaciones de los polimorfismos.

## 2. Tipado de tejido

Las secuencias de cinasa de la presente invención pueden usarse también para identificar individuos a partir de muestras biológicas mínimas. El ejército de los Estados Unidos, por ejemplo, está considerando el uso de polimorfismos de longitud del fragmento de restricción (RFLP) para la identificación de su personal. En esta técnica, se digiere un ADN genómico del individuo con una o más enzimas de restricción y se sondea en una transferencia Southern para proporcionar bandas únicas de identificación. La secuencia nucleotídica de la presente invención es útil como un marcador de ADN adicional para RFLP (descrito en la patente de EE.UU. n° 5.272.057).

Además, las secuencias de la presente invención pueden usarse para proporcionar una técnica alternativa para determinar la secuencia de ADN real base a base de porciones seleccionadas del genoma de un individuo. Por tanto, la secuencia de cinasa de la invención puede usarse para preparar dos cebadores de PCR de los extremos 5' y 3' de la secuencia. Estos cebadores pueden usarse entonces para amplificar el ADN de un individuo y posteriormente secuenciarlo.

Los paneles de secuencias de ADN correspondientes de individuos, preparados de esta manera, pueden proporcionar identificaciones individuales únicas, ya que cada individuo tendrá un conjunto único de dichas secuencias de ADN debido a diferencias alélicas. Las secuencias de cinasa de la invención representan únicamente porciones del genoma humano. Aparece variación alélica en cierto grado en las regiones de codificación de estas secuencias, y en mayor grado en las regiones de no codificación. Se estima que la variación alélica entre humanos individuales aparece con una frecuencia de aproximadamente una vez cada 500 bases. La secuencia descrita en la presente memoria puede usarse, en cierta medida, como patrón frente al que puede compararse el ADN de un individuo con fines de identificación. Las secuencias de no codificación de SEQ ID NO: 1 pueden proporcionar cómodamente una identificación individual positiva con un panel de quizás 10 a 1.000 cebadores, que proporcionan cada uno una secuencia amplificada no de codificación de 100 bases. Si se usan secuencias de codificación predichas, tales como las de SEQ ID NO: 1, sería un número más apropiado de cebadores para identificación individual positiva de 500 a 2.000.

### 3. Uso de secuencias de cinasa parciales en biología forense

Las técnicas de identificación basadas en ADN pueden usarse también en biología forense. De esta manera, puede usarse la tecnología de PCR para amplificar secuencias de ADN tomadas de muestras biológicas muy pequeñas tales como tejidos, por ejemplo, pelo, piel o fluidos corporales, por ejemplo, sangre, saliva o semen encontrado en una escena del delito. Puede compararse entonces la secuencia amplificada con un patrón, permitiendo así la identificación del origen de la muestra biológica.

Las secuencias de la presente invención pueden usarse para proporcionar reactivos polinucleotídicos, por ejemplo cebadores de PCR, orientados a loci específicos en el genoma humano, que pueden potenciar la fiabilidad de identificaciones forenses basadas en ADN proporcionando, por ejemplo, otro “marcador de identificación” que es único para un individuo particular. Como se menciona anteriormente, la información de secuencia básica real puede usarse para identificación como una alternativa exacta a patrones formados mediante fragmentos generados por enzimas de restricción. Las secuencias orientadas a regiones no de codificación de SEQ ID NO: 1 son particularmente apropiadas para este uso, ya que aparecen grandes números de polimorfismos en las regiones de no codificación, haciendo más fácil diferenciar a individuos usando esta técnica. Un ejemplo de dichos reactivos polinucleotídicos útiles incluye la secuencia de cinasa o porciones o fragmentos de la misma; derivados de las regiones de no codificación de la SEQ ID NO:1 que tienen una longitud de al menos 20 ó 30 bases.

Las secuencias de cinasa descritas en la presente memoria pueden usarse adicionalmente para proporcionar reactivos polinucleotídicos, por ejemplo, sondas marcadas o marcables que pueden usarse, por ejemplo, en una técnica de hibridación *in situ* para identificar un tejido específico. Esto puede ser muy útil en casos en que se presenta a un patólogo forense un tejido de origen desconocido. Pueden usarse paneles de dichas sondas de cinasa para identificar tejidos por tipo de especie y/u órgano.

De forma similar, estos reactivos, por ejemplo cebadores o sondas de cinasa, pueden usarse para cribar la contaminación en cultivos de tejido (concretamente, cribar la presencia de una mezcla de tipos diferentes de células en un cultivo).

### C. Medicina predictiva

La presente descripción está también relacionada con el campo de la medicina predictiva en la que se usan ensayos de diagnóstico, ensayos de pronóstico, farmacogenómica y control de ensayos clínicos con fines de pronóstico (predictivos) para tratar así profilácticamente a un individuo. Estas aplicaciones se describen en las subsecciones siguientes.

#### 1. Ensayos de diagnóstico

Se dan a conocer ensayos de diagnóstico para detectar la expresión de proteína y/o ácido nucleico de cinasa así como la actividad cinasa en el contexto de una muestra biológica. Un procedimiento ejemplar para detectar la presencia o ausencia de proteína de cinasa de la invención en una muestra biológica implica obtener una muestra biológica de un sujeto de ensayo y poner en contacto la muestra biológica con un compuesto o un agente capaz de detectar la proteína o ácido nucleico de cinasa (por ejemplo, ARNm, ADN genómico) que codifica la proteína de cinasa de la invención, de tal modo que se detecta la presencia de proteína de cinasa en la muestra biológica. Los resultados obtenidos con una muestra biológica del sujeto de ensayo pueden compararse con los resultados obtenidos con una muestra biológica de un sujeto de control.

Es un agente para detectar ARNm o ADN genómico de cinasa una sonda de ácido nucleico marcada capaz de hibridar con ARNm o ADN genómico de cinasa de la invención. La sonda de ácido nucleico puede ser, por ejemplo, un ácido nucleico de cinasa total o parcial, tal como el ácido nucleico de SEQ ID NO: 1 o una porción del mismo tal como una molécula de ácido nucleico de al menos 15, 30, 50, 100, 250 ó 500 nucleótidos de longitud y suficiente para hibridar específicamente en condiciones rigurosas con ARNm o ADN genómico de cinasa de la invención. Se describen en la presente memoria otras sondas adecuadas para uso en los ensayos de diagnóstico.

Es un agente preferido para detectar la proteína de cinasa un anticuerpo capaz de unirse a la proteína de cinasa de la invención, preferiblemente un anticuerpo con un marcaje 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)N<sub>2</sub>). Se pretende que el término “marcado”, con respecto a la sonda o el anticuerpo, comprenda marcaje directo de la sonda o anticuerpo mediante acoplamiento (concretamente, ligamiento físico) de una sustancia detectable a la sonda o anticuerpo, así como marcaje indirecto de la sonda o anticuerpo mediante reactividad con otro reactivo que está marcado directamente. Los ejemplos de marcaje indirecto incluyen la detección de un anticuerpo primario usando un anticuerpo secundario marcado fluorescentemente y el marcaje terminal de una sonda de ADN con biotina de modo que pueda detectarse con estreptavidina marcada fluorescentemente.

Se pretende que el término “muestra biológica” incluya tejidos, células y fluidos biológicos aislados de un sujeto. Es decir, el procedimiento de detección de la invención puede usarse para detectar ARNm, proteína o ADN genómico de cinasa en una muestra biológica *in vitro*. Por ejemplo, las técnicas *in vitro* para la detección de ARNm de

cinasa incluyen hibridaciones Northern e hibridaciones *in situ*. Las técnicas *in vitro* para la detección de proteína de cinasa incluyen ensayos de inmunosorción ligados a enzima (ELISA), transferencias Western, inmunoprecipitaciones e inmunofluorescencia. Las técnicas *in vitro* para la detección de ADN genómico de cinasa incluyen hibridaciones Southern. Además, las técnicas *in vivo* para la detección de proteína de cinasa incluyen introducir en un sujeto un anticuerpo anti-cinasa marcado. Por ejemplo, el anticuerpo puede marcarse con un marcador radiactivo cuya presencia y localización en un sujeto pueden detectarse mediante técnicas de formación de imagen estándar.

En un ejemplo, la muestra biológica contiene moléculas de proteína del sujeto de ensayo. En una realización, la muestra biológica puede contener moléculas de ARNm del sujeto de ensayo o moléculas de ADN genómico del sujeto de ensayo. Las muestras biológicas pueden obtenerse de sangre, suero, células o tejido de un sujeto.

Se dan a conocer kits para detectar la presencia de proteínas de cinasa en una muestra biológica (una muestra de ensayo). Dichos kits pueden usarse para determinar si un sujeto padece o tiene un riesgo elevado de desarrollar un trastorno asociado a la expresión aberrante de proteína de cinasa. Por ejemplo, el kit puede comprender un compuesto o agente marcado capaz de detectar proteína o ARNm de cinasa en una muestra biológica y medios para determinar la cantidad de una proteína de cinasa en la muestra (por ejemplo, un anticuerpo anti-cinasa tal como un anticuerpo anti-cinasa h14189 o una sonda oligonucleotídica que se une a ADN que codifica una proteína de cinasa de la invención tal como la de SEQ ID NO: 1). Los kits puede incluir también instrucciones para observar que el sujeto ensayado padece o tiene riesgo de desarrollar un trastorno asociado a la expresión aberrante de la secuencia de cinasa de la invención si la cantidad de proteína o ARNm de cinasa es superior o inferior a un nivel normal.

Para kits basados en anticuerpos, el kit puede comprender, por ejemplo: (1) un primer anticuerpo (por ejemplo, unido a un soporte sólido) que se une a la proteína de cinasa de la invención y, opcionalmente, (2) un segundo anticuerpo diferente que se une a la proteína de cinasa de la invención o al primer anticuerpo y se conjuga con un agente detectable. Para kits basados en oligonucleótidos, el kit puede comprender, por ejemplo: (1) un oligonucleótido, por ejemplo, un oligonucleótido marcado detectablemente que hibrida con una secuencia de ácido nucleico de cinasa de la invención o (2) un par de cebadores útiles para amplificar una molécula de ácido nucleico de cinasa de la invención.

El kit puede comprender también, por ejemplo, un agente de tamponación, un conservante o un agente estabilizante de proteína. El kit puede comprender también los componentes necesarios para detectar el agente detectable (por ejemplo, una enzima o un sustrato). El kit puede contener también 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 está habitualmente introducido en un envase individual, y todos los diversos envases están dentro de un solo paquete junto con instrucciones para observar si el sujeto ensayado padece o tiene riesgo de desarrollar un trastorno asociado a la expresión aberrante de proteínas de cinasa de la invención.

## 2. Otros ensayos de diagnóstico

Se da a conocer un procedimiento de análisis de una pluralidad de sondas de captura. El procedimiento puede usarse, por ejemplo, para analizar la expresión génica. El procedimiento incluye: proporcionar una matriz bidimensional que tiene una pluralidad de direcciones, siendo cada dirección de la pluralidad posicionalmente distinguible de las otras direcciones de la pluralidad, y teniendo cada dirección de la pluralidad una sonda de captura única, por ejemplo, una secuencia de ácido nucleico o peptídica; poner en contacto la matriz con un ácido nucleico preferiblemente purificado, polipéptido preferiblemente purificado o anticuerpo de cinasa, y evaluar así la pluralidad de sondas de captura. Se detecta la unión, por ejemplo en el caso de un ácido nucleico la hibridación, con una sonda de captura en una dirección de la pluralidad, por ejemplo, mediante la señal generada por un marcador unido al ácido nucleico, polipéptido o anticuerpo de cinasa. Las sondas de captura pueden ser un conjunto de ácidos nucleicos de una muestra seleccionada, por ejemplo, una muestra de ácidos nucleicos derivada de un tejido o célula de control o no estimulado.

El procedimiento puede incluir poner en contacto el ácido nucleico, polipéptido o anticuerpo de cinasa con una primera matriz que tiene una pluralidad de sondas de captura y una segunda matriz que tiene una diferente pluralidad de sondas de captura. Pueden compararse los resultados de cada hibridación, por ejemplo, para analizar las diferencias en expresión entre una primera y segunda muestras. La primera pluralidad de sondas de captura puede ser de una muestra de control, por ejemplo, una muestra de tipo silvestre, normal o no enferma y no estimulada, por ejemplo, una muestra de fluido biológico, tejido o célula. La segunda pluralidad de sondas de captura puede ser de una muestra experimental, por ejemplo, una muestra de tipo mutante, con riesgo de enfermedad, en estado enfermo o estimulada, por ejemplo, una muestra de fluido corporal, tejido o célula.

La pluralidad de sondas de captura puede ser una pluralidad de sondas de ácido nucleico cada una de las cuales hibrida específicamente con un alelo de una secuencia de cinasa de la invención. Dichos procedimientos pueden usarse para diagnosticar a un sujeto, por ejemplo, para evaluar el riesgo de enfermedad o trastorno, para evaluar la idoneidad de un tratamiento seleccionado para un sujeto o para evaluar si un sujeto tiene una enfermedad o trastorno.

El procedimiento puede usarse para detectar polimorfismos de un solo nucleótido (SNP) como se describe a continuación.

Se da a conocer un procedimiento de análisis de una pluralidad de sondas. El procedimiento es útil, por ejemplo, para analizar la expresión génica. El procedimiento incluye: proporcionar una matriz bidimensional que tiene una pluralidad de direcciones, siendo cada dirección de la pluralidad posicionalmente distinguible de otras direcciones de la pluralidad, y teniendo cada dirección de la pluralidad una sonda de captura única, por ejemplo, en el que las sondas de captura son de una célula o sujeto que expresa un polipéptido de cinasa de la invención o de una célula o sujeto en que se ha desencadenado una respuesta mediada por cinasa, por ejemplo, mediante la puesta en contacto de la célula con un ácido nucleico o proteína de cinasa de la invención o la administración a la célula o el sujeto de un ácido nucleico o proteína de cinasa de la invención; poner en contacto la matriz con una o más sondas de prueba, en las que la sonda de prueba puede ser un ácido nucleico, polipéptido o anticuerpo (que es preferiblemente distinto del ácido nucleico o polipéptido de cinasa de la invención); proporcionar una matriz bidimensional que tiene una pluralidad de direcciones, siendo cada dirección posicionalmente distinguible de las demás direcciones de la pluralidad, y teniendo cada dirección de la pluralidad una sonda de captura única, por ejemplo, en la que las sondas de captura son de una célula o sujeto que no expresa una secuencia de cinasa de la invención (o no la expresa tanto como en el caso de la pluralidad de sondas de captura positiva de cinasa) o de una célula o sujeto en que la respuesta mediada por cinasa no se ha desencadenado (o se ha desencadenado en menor extensión que en la primera muestra); poner en contacto la matriz con una o más sondas de prueba (que es preferiblemente distinta del ácido nucleico o polipéptido de cinasa de la invención) y evaluar así la pluralidad de sondas de captura. Se detecta la unión (por ejemplo, en el caso de un ácido nucleico, la hibridación) con una sonda de captura en una dirección de la pluralidad, por ejemplo, mediante una señal generada por un marcador unido al ácido nucleico, polipéptido o anticuerpo.

Se da a conocer un procedimiento de análisis de una secuencia de cinasa de la invención, por ejemplo, análisis de la estructura, función o similitud con otras secuencias de ácido nucleico o aminoácidas. El procedimiento incluye: proporcionar una secuencia de ácido nucleico o aminoácida de cinasa, por ejemplo, la secuencia 14189 expuesta en la SEQ ID NO:1 o la SEQ ID NO: 2 o una porción de las mismas; comparar la secuencia de cinasa con una o más, preferiblemente una pluralidad de secuencias, de una colección de secuencias, por ejemplo, una base de datos de secuencias de ácido nucleico o proteína; para analizar así la secuencia de cinasa de la invención.

El procedimiento puede incluir evaluar la identidad de secuencia entre una secuencia de cinasa de la invención, por ejemplo, la secuencia 14189, y una secuencia de base de datos. El procedimiento puede efectuarse accediendo a la base de datos en un segundo sitio, por ejemplo, en internet.

Se da a conocer un conjunto de oligonucleótidos útil, por ejemplo, para identificar SNP o para identificar alelos específicos de una secuencia de cinasa de la invención, por ejemplo, la secuencia 14189. El conjunto incluye una pluralidad de oligonucleótidos, cada uno de los cuales tiene un nucleótido diferente en una posición de interrogación, por ejemplo, un SNP o el sitio de una mutación. En un ejemplo preferido, los oligonucleótidos de la pluralidad son idénticos en secuencia entre sí (excepto por diferencias en longitud). Los oligonucleótidos pueden dotarse de marcadores diferenciales, de tal modo que un oligonucleótido que hibride con un alelo proporcione una señal que sea distinguible de un oligonucleótido que hibride con un segundo alelo.

### 3. Ensayos de pronóstico

Los procedimientos descritos en la presente memoria pueden utilizarse además como ensayos de diagnóstico o pronóstico para identificar sujetos que tienen o tienen riesgo de desarrollar una enfermedad o trastorno asociado a la proteína de cinasa, la expresión de ácido nucleico de cinasa o la actividad cinasa de la invención. Los ensayos de pronóstico pueden usarse con fines de pronóstico o predictivos para tratar así profilácticamente un individuo antes del inicio de un trastorno caracterizado por o asociado a la proteína de cinasa, expresión de ácido nucleico de cinasa o actividad cinasa de la invención.

Por tanto, se da a conocer un procedimiento en el que se obtiene una muestra de ensayo de un sujeto, y se detecta la proteína o ácido nucleico de cinasa de la invención (por ejemplo, ARNm, ADN genómico), en el que la presencia de proteína o ácido nucleico de cinasa es un diagnóstico de que un sujeto tiene o tiene riesgo de desarrollar una enfermedad o trastorno asociado a la expresión o actividad aberrante de cinasa. Como se usa en la presente memoria, una “muestra de ensayo” designa una muestra biológica obtenida de un sujeto de interés. Por ejemplo, una muestra de ensayo puede ser una muestra celular de fluido biológico o tejido.

Además, en el uso de los ensayos de pronóstico descritos en la presente memoria, se dan a conocer procedimientos para determinar si puede administrarse a un sujeto un agente específico (por ejemplo, un agonista, antagonista, peptidomimético, proteína, péptido, ácido nucleico, molécula pequeña u otro candidato a medicamento) o clase de agentes (por ejemplo, agentes del tipo que reduce la actividad cinasa) para tratar eficazmente una enfermedad o trastorno asociado a una expresión o actividad aberrante de cinasa. De esta manera, se obtiene una muestra de ensayo y se detecta la proteína o ácido nucleico de cinasa. La presencia de proteína o ácido nucleico de cinasa de la invención es un diagnóstico de que puede administrarse a un sujeto el agente para tratar un trastorno asociado a la expresión o actividad aberrante de cinasa.

Los procedimientos pueden usarse también para detectar lesiones genéticas o mutaciones en un gen de cinasa de la invención, determinando así si un sujeto con el gen lesionado tiene riesgo de trastorno caracterizado por una proliferación y/o diferenciación celular aberrante. Los procedimientos incluyen detectar, en una muestra de células del

sujeto, la presencia o ausencia de una lesión genética o mutación caracterizada por al menos una de una alteración que afecta a la integridad de un gen que codifica una proteína de cinasa, o la expresión incorrecta del gen de cinasa de la invención. Por ejemplo, dichas lesiones genéticas o mutaciones pueden detectarse valorando la existencia de al menos una de: (1) una delección de uno o más nucleótidos de un gen de cinasa; (2) una adición de uno o más nucleótidos a un gen de cinasa; (3) una sustitución de uno o más nucleótidos de un gen de cinasa; (4) una redistribución cromosómica de un gen de cinasa; (5) una alteración en el nivel de un transcrito de ARN mensajero de un gen de cinasa; (6) una modificación aberrante de un gen de cinasa, tal como del patrón de metilación del ADN genómico; (7) la presencia de un patrón de corte y empalme de tipo no silvestre de un transcrito de ARN mensajero de un gen de cinasa; (8) un nivel de tipo no silvestre de una proteína de cinasa; (9) una pérdida alélica de un gen de cinasa y (10) una modificación postraducciona inapropiada de una proteína de cinasa. Como se describe en la presente memoria, hay un gran número de técnicas de ensayo conocidas en la técnica que pueden usarse para detectar lesiones en un gen de cinasa. Puede utilizarse en los ensayos de pronóstico descritos en la presente memoria cualquier tipo de célula o tejido en el que se expresen las proteínas de cinasa.

La detección de la lesión puede implicar el uso de una sonda/cebador en una reacción en cadena de la polimerasa (PCR) (véanse, por ejemplo, las patentes de EE.UU. n° 4.683.195 y 4.683.202) tal como una PCR de anclaje o PCR RACE o, como alternativa, en una reacción en cadena de ligamiento (LCR) (véanse, por ejemplo, Landegran y col. (1988) *Science* 241: 1077-1080 y Nakazawa y col. (1994) *Proc. Natl. Acad. Sci. USA* 91: 360-364), pudiendo ser la última particularmente útil para detectar mutaciones puntuales en el gen de cinasa (véanse, por ejemplo, Abravaya y col. (1995) *Nucleic Acids Res.* 23: 675-682). Se prevé que PCR y/o LCR pueden ser deseables para usar como etapa de amplificación preliminar junto con cualquiera de las técnicas usadas para detectar mutaciones descritas en la presente memoria.

Los procedimientos de amplificación alternativos incluyen replicación de secuencia automantenida (Guatelli y col. (1990) *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), sistema de amplificación transcripcional (Kwoh y col. (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173-1177), replicasa Q-Beta (Lizardi y col. (1988) *Bio/Technology* 6: 1197) o cualquier otro procedimiento de amplificación de ácido nucleico, seguido de la detección de las moléculas amplificadas usando técnicas bien conocidas por los expertos en la técnica. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si dichas moléculas están presentes en muy bajos números.

Las mutaciones en un gen de cinasa de una célula de muestra pueden identificarse mediante alteraciones en los patrones de escisión de enzimas de restricción de la muestra de ensayo aislada y ADN de control digeridos con una o más endonucleasas de restricción. Además, el uso de ribozimas específicas de secuencia (véase, por ejemplo, la patente de EE.UU. n° 5.498.531) puede usarse para puntuar la presencia de mutaciones específicas mediante el desarrollo o pérdida de un sitio de escisión de ribozima.

Las mutaciones genéticas en una molécula de cinasa pueden identificarse hibridando una muestra y ácidos nucleicos de control, por ejemplo ADN o ARN, con matrices de alta densidad que contienen cientos o miles de sondas oligonucleotídicas (Cronin y col. (1996) *Human Mutation* 7: 244-255; Kozal y col. (1996) *Nature Medicine* 2: 753-759). Puede usarse cualquiera de una variedad de reacciones de secuenciación conocidas en la técnica para secuenciar directamente el gen de cinasa y detectar mutaciones comparando la secuencia del gen de cinasa de la muestra con la correspondiente secuencia de tipo silvestre (control). Los ejemplos de reacciones de secuenciación incluyen aquellas basadas en técnicas desarrolladas por Maxim y Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74: 560) o Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74: 5463). Se contempla también que puede utilizarse cualquiera de una variedad de procedimientos de secuenciación automatizados cuando se efectúan los ensayos de diagnóstico ((1995) *Bio/Techniques* 19: 448), incluyendo la secuenciación por espectrometría de masas (véanse, por ejemplo, la publicación PCT n° WO 94/16101; Cohen y col. (1996) *Adv. Chromatogr.* 36: 127-162 y Griffin y col. (1993) *Appl. Biochem. Biotechnol.* 38: 147-159).

Otros procedimientos para detectar mutaciones en el gen de cinasa incluyen procedimientos en los que se usa la protección de agentes de escisión para detectar bases emparejadas erróneamente en heterodúplex de ARN/ARN o ARN/ADN (Myers y col. (1985) *Science* 230: 1242). Véanse también Cotton y col. (1988) *Proc. Natl. Acad. Sci. USA* 85: 4397; Saleeba y col. (1992) *Methods Enzymol.* 217: 286-295. En un ejemplo preferido, el ADN o ARN de control pueden estar marcados para detección.

La reacción de escisión de emparejamiento erróneo emplea una o más enzimas de "reparación de emparejamiento erróneo de ADN" que reconocen pares de bases emparejados erróneamente en ADN bicatenario en sistemas definidos para detectar y cartografiar mutaciones puntuales en ADN de cinasa obtenidos a partir de muestras de células. Véase, por ejemplo, Hsu y col. (1994) *Carcinogenesis* 15: 1657-1662. Según un ejemplo, se hibrida una sonda basada en una secuencia de cinasa, por ejemplo, una secuencia de cinasa de tipo silvestre, con un ADNc u otro producto de ADN de una(s) célula(s) diana. Se trata el dúplex con una enzima de reparación de emparejamiento erróneo de ADN y los productos de escisión, si los hubiera, pueden detectarse mediante protocolos de electroforesis o similares. Véase, por ejemplo, la patente de EE.UU. n° 5.459.039.

Las alteraciones en la movilidad electroforética se usarán para identificar mutaciones en genes de cinasa. Por ejemplo, puede usarse el polimorfismo de conformación monocatenaria (SSCP) para detectar diferencias en la movilidad electroforética entre ácidos nucleicos de tipo mutante y silvestre (Orita y col. (1989) *Proc. Natl. Acad. Sci. USA* 86: 2766; véanse también Cotton (1993) *Mutat. Res.* 285: 125-144; Hayashi (1992) *Genet. Anal. Tech. Appl.* 9: 73-79).

Puede potenciarse la sensibilidad del ensayo usando ARN (en lugar de ADN), en el que la estructura secundaria es más sensible a un cambio de secuencia. En otra realización, el procedimiento objeto utiliza análisis de heterodúplex para separar moléculas de heterodúplex bicatenarias basándose en los cambios de movilidad electroforética (Keen y col. (1991) *Trends Genet.* 7: 5).

Se ensaya el movimiento de fragmentos de tipo mutante o silvestre en geles de poliacrilamida que contienen un gradiente desnaturalizante usando electroforesis en gel de gradiente desnaturalizante (DGGE) (Myers y col. (1985) *Nature* 313: 495). Cuando se usa el DGGE como procedimiento de análisis, se modificará el ADN para asegurar que no se desnaturaliza completamente, por ejemplo, añadiendo una abrazadera de GC de aproximadamente 40 pb de ADN rico en GC de alto punto de fusión mediante PCR. Puede usarse un gradiente de temperatura en lugar de un gradiente desnaturalizante para identificar diferencias en la movilidad de ADN de control y muestra (Rosenbaum y Reissner (1987) *Biophys. Chem.* 265: 12753).

Los ejemplos de otras técnicas para detectar mutaciones puntuales incluyen, pero sin limitación, hibridación selectiva de oligonucleótidos, amplificación selectiva o extensión selectiva de cebador. Por ejemplo, pueden prepararse cebadores oligonucleotídicos en los que la mutación conocida se dispone centralmente y se hibrida entonces con ADN diana en condiciones que permitan la hibridación sólo si se encuentra un emparejamiento perfecto (Saiki y col. (1986) *Nature* 324: 163); Saiki y col. (1989) *Proc. Natl. Acad. Sci. USA* 86: 6230). Se hibridan dichos oligonucleótidos específicos de alelo con ADN diana amplificado por PCR o una serie de diferentes mutaciones cuando los oligonucleótidos están unidos a la membrana de hibridación e hibridados con ADN diana marcado.

Como alternativa, puede usarse la tecnología de amplificación específica de alelo, que depende de la amplificación selectiva por PCR, junto con la presente invención. Los oligonucleótidos usados como cebadores para amplificación específica pueden portar la mutación de interés en el centro de la molécula, de modo que la amplificación depende de la hibridación diferencial (Gibbs y col. (1989) *Nucleic Acids Res.* 17: 2437-2448), o en el extremo 3' de un cebador en que, en condiciones apropiadas, el emparejamiento erróneo puede prevenir o reducir la extensión por polimerasa (Prossner (1993) *Tibtech* 11: 238). Además, puede ser deseable introducir un sitio de restricción novedoso en la región de la mutación para crear una detección basada en escisión (Gasparini y col. (1992) *Mol. Cell Probes* 6: 1). Se prevé que en ciertas realizaciones la amplificación pueda efectuarse también usando ligasa Taq para amplificación (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88: 189). En dichos casos, aparecerá ligamiento sólo si hay un emparejamiento perfecto en el extremo 3' de la secuencia 5', haciendo posible detectar la presencia de una mutación conocida en un sitio específico observando la presencia o ausencia de amplificación.

Los procedimientos descritos en la presente memoria pueden efectuarse, por ejemplo, utilizando kits de diagnóstico preenvasados que comprenden al menos un reactivo de ácido nucleico o anticuerpo de sonda descrito en la presente memoria, que puede usarse convenientemente, por ejemplo, en instalaciones clínicas para diagnosticar a pacientes que exhiben síntomas o un historial familiar de enfermedad o dolencia que implica un gen de cinasa.

#### 4. Farmacogenómica

Los agentes o moduladores que tienen un efecto estimulante o inhibidor sobre la actividad cinasa (por ejemplo, expresión del gen de cinasa) como se identifica mediante un ensayo de cribado descrito en la presente memoria, pueden administrarse a individuos para tratar (profiláctica o terapéuticamente) trastornos asociados a una actividad cinasa aberrante, así como para modular el crecimiento, diferenciación y/o metabolismo celulares. Junto con dicho tratamiento, puede considerarse la farmacogenómica (concretamente, el estudio de la relación entre el genotipo de un individuo y la respuesta del individuo ante un compuesto extraño o medicamento) del individuo. Las diferencias en el metabolismo del producto terapéutico pueden conducir a toxicidad grave o insuficiencia terapéutica, alterando la relación entre la dosis y la concentración sanguínea del medicamento farmacológicamente activo. Por tanto, la farmacogenómica del individuo permite la selección de agentes eficaces (por ejemplo medicamentos) para tratamientos profilácticos o terapéuticos basados en la consideración del genotipo del individuo. Dicha farmacogenómica puede usarse adicionalmente para determinar dosificaciones y regímenes terapéuticos apropiados. En consecuencia, puede determinarse la actividad de la proteína de cinasa, la expresión del ácido nucleico de cinasa o el contenido de mutación de los genes de cinasa en un individuo, seleccionando así el(los) agente(s) apropiado(s) para el tratamiento terapéutico o profiláctico del individuo.

La farmacogenómica trata con variaciones hereditarias significativas de la respuesta a medicamentos debidas a una disposición alterada a los medicamentos y una acción anormal en las personas afectadas. Véase, por ejemplo, Linder (1997) *Clin. Chem.* 43 (2): 254-266. En general, pueden diferenciarse dos tipos de condiciones farmacogenéticas. Las afecciones genéticas transmitidas como factores individuales que alteran el modo en que actúan los medicamentos sobre el cuerpo se designan como "acción alterada del medicamento". Las afecciones genéticas transmitidas como factores individuales que alteran el modo en que actúa el cuerpo sobre los medicamentos se designan como "metabolismo alterado del medicamento". Estas afecciones farmacogenéticas pueden aparecer como defectos o polimorfismos raros. Por ejemplo, la deficiencia de glucosa-6-fosfato deshidrogenasa (G6PD) es una enzimopatía heredada común en la que la complicación clínica principal es hemólisis después de la ingestión de medicamentos oxidantes (antimaláricos, sulfonamidas, analgésicos, nitrofuranos) y el consumo de judías verdes.

Un enfoque farmacogenómico para identificar genes que predicen la respuesta a medicamentos, conocido como “asociación del genoma completo”, se basa principalmente en un mapa de alta resolución del genoma humano que consiste en los marcadores relacionados con genes ya conocidos (por ejemplo, un mapa de marcadores génicos “bia-  
 5 lélicos” que consiste en 60.000-100.000 sitios polimórficos o variables en el genoma humano, cada uno de los cuales 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 toman parte en un ensayo de medicamento de fase II/III para identificar los marcadores asociados a una respuesta a medicamento o efecto secundario observado particular. Como alternativa, dicho mapa de alta resolución puede generarse a partir de una combinación de unos diez millones de polimorfismos nucleotídicos individuales (SNP) conocidos en el genoma humano. Como se usa en la presente memoria,  
 10 un “SNP” es una alteración común que aparece en una sola base nucleotídica en un tramo de ADN. Por ejemplo, un SNP puede aparecer una vez cada 1.000 bases de ADN. Un SNP puede estar implicado en un proceso patológico, sin embargo, la amplia mayoría pueden no estar asociados a enfermedad. Dado un mapa genético basado en la aparición de dichos SNP, los individuos pueden agruparse en categorías genéticas dependiendo del patrón particular de SNP en su genoma individual. De tal manera, los regímenes de tratamiento pueden ajustarse a grupos de individuos genéticamente  
 15 similares, teniendo en cuenta rasgos que pueden ser comunes entre dichos individuos genéticamente similares.

Como alternativa, puede utilizarse un procedimiento denominado el “enfoque del gen candidato” para identificar genes que predicen la respuesta a medicamento. Según este procedimiento, si un gen que codifica una diana de medicamento es conocido (por ejemplo, una proteína de cinasa de la presente invención), todas las variantes comunes de ese gen pueden identificarse bastante fácilmente en la población y puede determinarse si tener una versión del gen frente a otra está asociado a una respuesta a medicamento particular.

Como alternativa, puede utilizarse un procedimiento denominado “perfil de expresión génica” para identificar genes que predicen la respuesta a medicamento. Por ejemplo, la expresión génica de un animal dosificado con un medicamento (por ejemplo, una molécula de cinasa o modulador de cinasa 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 uno o más de los enfoques farmacogenómicos anteriores puede usarse para determinar la dosificación y regímenes de tratamiento adecuados para el tratamiento profiláctico o terapéutico de un individuo. Este conocimiento, cuando se aplica a la dosificación o selección de medicamento, puede evitar reacciones  
 30 adversas o insuficiencia terapéutica y potenciar por tanto la eficacia terapéutica o profiláctica cuando se trata un sujeto con una molécula de cinasa o modulador de cinasa de la invención, tal como un modulador identificado mediante uno de los ensayos de cribado ejemplares descritos en la presente memoria.

Se dan a conocer adicionalmente procedimientos para identificar nuevos agentes, o combinaciones, que están basados en identificar agentes que modulan la actividad de uno o más de los productos génicos codificados por uno o más de los genes de cinasa de la presente invención, en los que estos productos pueden estar asociados a resistencia de las células a un agente terapéutico. Específicamente, la actividad de las proteínas codificadas por los genes de cinasa de la presente invención puede usarse como base para identificar agentes para superar la resistencia a agente. Al bloquear  
 40 la actividad de una o más de las proteínas de resistencia, las células diana se volverán sensibles al tratamiento con un agente al que las células diana no modificadas eran resistentes.

Controlar la influencia de agentes (por ejemplo medicamentos) sobre la expresión o actividad de una proteína de cinasa puede aplicarse en ensayos clínicos. Por ejemplo, puede controlarse la eficacia de un agente determinada por un ensayo de cribado como se describe en la presente memoria para aumentar la expresión del gen de cinasa, los niveles de proteína o regular positivamente la actividad cinasa, en ensayos clínicos de sujetos que exhiben una expresión de gen de cinasa y niveles de proteína reducidos o que regulan negativamente la actividad cinasa. Como alternativa, puede controlarse la eficacia de un agente determinada por un ensayo de cribado como se describe en la presente memoria para reducir la expresión del gen de cinasa, los niveles de proteína o regular negativamente la actividad cinasa, en ensayos clínicos de sujetos que exhiben una expresión del gen de cinasa y niveles de proteínas aumentados, o que regulan positivamente la actividad cinasa. En dichos ensayos clínicos, la expresión o actividad de un gen de cinasa, y preferiblemente otros genes que se han implicado, por ejemplo, en un trastorno asociado a cinasa, pueden usarse como “lectura” de los marcadores del fenotipo de una célula particular.

Como ejemplo ilustrativo, la actividad de enzimas metabolizantes de medicamento es un determinante importante tanto de la intensidad como de la duración de la acción del medicamento. El descubrimiento de polimorfismos genéticos de enzimas metabolizantes de medicamentos (por ejemplo, *N*-acetiltransferasa 2 (NAT 2) y enzimas CYP2D6 y CYP2C19 del citocromo P450) ha proporcionado una explicación de porqué algunos pacientes no obtienen los efectos de medicamento esperados o muestran una respuesta a medicamento exagerada y una grave toxicidad después de tomar la dosis estándar y segura de un medicamento. Estos polimorfismos se expresan en dos fenotipos en la población, el metabolizante extenso (EM) y el metabolizante malo (PM). La prevalencia de PM es diferente entre diferentes poblaciones. Por ejemplo, el gen que codifica CYP2D6 es altamente polimórfico y se han identificado varias mutaciones en PM que conducen todas a la ausencia de CYP2D6 funcional. Los metabolizantes malos de CYP2D6 y CYP2C19 experimentan con bastante frecuencia una respuesta a medicamento exagerada y efectos secundarios cuando reciben dosis estándar. Si un metabolito es el resto terapéutico activo, un PM no mostrará respuesta terapéutica, como se demuestra por el efecto analgésico de la codeína mediado por su metabolito formado por CYP2D6 morfina. En el otro extremo están los denominados metabolizadores ultrarrápidos, que no responden a dosis estándar. Recientemente, se ha identificado que la base molecular del metabolismo ultrarrápido es debida a la amplificación del gen CYP2D6.

Por tanto, pueden determinarse la actividad de la proteína de cinasa, la expresión del ácido nucleico de cinasa o el contenido de mutación de los genes de cinasa en un individuo para así seleccionar agente(s) apropiado(s) para el tratamiento terapéutico o profiláctico del individuo. Además, pueden usarse estudios de farmacogenética para aplicar el genotipado de alelos polimórficos que codifican enzimas metabolizantes de medicamento para la identificación del fenotipo de sensibilidad a medicamento de un individuo. Este conocimiento, cuando se aplica a la selección de dosificación o medicamento, puede evitar reacciones adversas o insuficiencia terapéutica y potencia por tanto la eficacia terapéutica o profiláctica cuando se trata un sujeto con un modulador de cinasa, tal como un modulador identificado mediante uno de los ensayos de cribado descritos en la presente memoria.

## 5. Control de los efectos durante ensayos clínicos

El control de la influencia de los agentes (por ejemplo, medicamentos o compuestos) sobre la expresión o actividad de los genes de cinasa (por ejemplo, la capacidad de modular la proliferación y/o diferenciación celular aberrante) puede aplicarse no sólo al cribado básico de medicamentos, sino también en ensayos clínicos. Por ejemplo, puede controlarse la eficacia de un agente, determinada mediante un ensayo de cribado como se describe en la presente memoria, de aumentar o reducir la expresión del gen de cinasa, los niveles de proteína o la actividad de proteína, en ensayos clínicos de sujetos que exhiben una expresión del gen de cinasa, niveles de proteína o actividad de proteína reducidas o aumentadas. En dichos ensayos clínicos, la expresión o actividad de cinasa y preferiblemente la de otros genes que se han implicado, por ejemplo, en un trastorno de proliferación celular, pueden usarse como marcador del crecimiento y diferenciación celulares.

Por ejemplo, y no a modo de limitación, pueden identificarse genes que se modulan en células mediante el tratamiento con un agente (por ejemplo, compuesto, medicamento o molécula pequeña) que modula la actividad cinasa (por ejemplo, como se identifica en un ensayo de cribado descrito en la presente memoria). Por tanto, para estudiar el efecto de agentes sobre trastornos de proliferación celular, por ejemplo en un ensayo clínico, pueden aislarse células y prepararse el ARN y analizarse los niveles de expresión de genes de cinasa y otros genes implicados en el trastorno. Los niveles de expresión génica (concretamente, el patrón de expresión génica) pueden cuantificarse mediante análisis de transferencia Northern o PCR-FI, como se describe en la presente memoria o, como alternativa, midiendo la cantidad de proteína producida mediante uno de los procedimientos descritos en la presente memoria, o midiendo los niveles de actividad de genes de cinasa u otros genes. De este modo, el patrón de expresión génica puede servir como marcador, indicativo de la respuesta fisiológica de las células al agente. En consecuencia, este estado de respuesta puede determinarse antes, y en varios puntos durante el tratamiento del individuo con el agente.

Se da a conocer un procedimiento para controlar la eficacia del tratamiento de un sujeto con un agente (por ejemplo, un agonista, antagonista, peptidomimético, proteína, péptido, ácido nucleico, molécula pequeña u otro candidato a medicamento identificado mediante los ensayos de cribado descritos en la presente memoria) que comprende las etapas de (1) obtener una muestra de preadministración de un sujeto antes de la administración del agente; (2) detectar el nivel de expresión de una proteína, ARNm o ADN genómico de cinasa en la muestra de preadministración; (3) obtener una o más muestras de postadministración del sujeto; (4) detectar el nivel de expresión o actividad de la proteína, ARNm o ADN genómico de cinasa en las muestras de postadministración; (5) comparar el nivel de expresión o actividad de la proteína, ARNm o ADN genómico de cinasa en la muestra de preadministración con la proteína, ARNm o ADN genómico de cinasa en la muestra o muestras de postadministración y (6) alterar la administración del agente al sujeto en consecuencia para producir el efecto deseado, concretamente por ejemplo, un aumento o reducción de la expresión o actividad de una proteína de cinasa.

## C. Procedimientos de tratamiento

Se dan a conocer procedimientos tanto profilácticos como terapéuticos de tratamiento de un sujeto con riesgo de (o sensible a) un trastorno o que tienen un trastorno asociado a una expresión o actividad de cinasa aberrante. Adicionalmente, las composiciones descritas en la presente memoria encuentran uso en el tratamiento de los trastornos descritos en la presente memoria.

### 1. Procedimientos profilácticos

Se da a conocer un procedimiento para prevenir en un sujeto una enfermedad o afección asociada a una expresión o actividad aberrante de cinasa mediante la administración al sujeto de un agente que modula la expresión de cinasa o al menos una actividad del gen de cinasa. Los sujetos con riesgo de una enfermedad que está causada, o contribuida, por una expresión o actividad de cinasa aberrante pueden identificarse, por ejemplo, mediante una cualquiera o una combinación de ensayos de diagnóstico o pronóstico como se describen en la presente memoria. La administración de un agente profiláctico puede darse antes de la manifestación de síntomas característicos de la aberración de cinasa, de tal modo que se previene una enfermedad o trastorno o, como alternativa, se retarda su progresión. Dependiendo del tipo de aberración de cinasa, por ejemplo, puede usarse un agonista de cinasa o antagonista de cinasa para tratar el sujeto. El agente apropiado puede determinarse basándose en los ensayos de cribado descritos en la presente memoria.

## 2. Procedimientos terapéuticos

Se dan a conocer procedimientos de modulación de la expresión o actividad de cinasa con fines terapéuticos. El procedimiento modulador implica poner en contacto una célula con un agente que modula una o más de las actividades de la actividad de la proteína de cinasa asociada a la célula. Un agente que modula la actividad de proteína de cinasa puede ser un agente como se describe en la presente memoria, tal como un ácido nucleico o proteína, un ligando afín de origen natural de una proteína de cinasa, un péptido, un peptidomimético de cinasa u otra molécula pequeña. En una realización, el agente estimula una o más de las actividades biológicas de la proteína de cinasa. Los ejemplos de dichos agentes estimulantes incluyen proteína de cinasa activa y una molécula de ácido nucleico que codifica una proteína de cinasa que se ha introducido en la célula. El agente puede inhibir una o más de las actividades biológicas de la proteína de cinasa. Los ejemplos de dichos agentes inhibidores incluyen moléculas de ácido nucleico de cinasa no codificantes y anticuerpos anti-cinasa.

Estos procedimientos moduladores pueden efectuarse *in vitro* (por ejemplo, cultivando la célula con el agente) o, como alternativa, *in vivo* (por ejemplo, administrando el agente a un sujeto). Como tales, se dan a conocer procedimientos de tratamiento de un individuo aquejado de una enfermedad o trastorno caracterizado por una expresión o actividad aberrante de una proteína o molécula de ácido nucleico de cinasa. Se da a conocer la administración de un agente (por ejemplo, un agente identificado mediante un ensayo de cribado descrito en la presente memoria) o una combinación de agentes que modulan (por ejemplo, regulan positivamente o regulan negativamente) la expresión o actividad de cinasa. Se da a conocer la administración de una proteína o molécula de ácido nucleico de cinasa como terapia para compensar una expresión o actividad reducida o aberrante de cinasa.

La estimulación de la actividad cinasa es deseable en situaciones en que una proteína de cinasa está anormalmente regulada negativamente y/o en la que una actividad aumentada es probable que tenga un efecto beneficioso. A la inversa, la inhibición de la actividad cinasa es deseable en situaciones en que la actividad cinasa está anormalmente regulada positivamente y/o en que la actividad cinasa reducida es probable que tenga un efecto beneficioso.

Esta invención se ilustra adicionalmente por los siguientes ejemplos, que no deben considerarse como limitantes.

## Parte experimental

### Análisis de expresión génica

Se preparó el ARN total de diversos tejidos humanos mediante un procedimiento de extracción en una etapa usando RNA STAT-60 según las instrucciones del fabricante (TelTest, Inc). Se trató cada preparación de ARN con ADNasa I (Ambion) a 37°C durante 1 hora. Se determinó que el tratamiento con ADNasa I estaba completado si la muestra requería al menos 38 ciclos de amplificación por PCR para alcanzar un nivel umbral de fluorescencia usando  $\beta$ -2 microglobulina como referencia de amplificación interna. Se confirmó la integridad de las muestras de ARN después de tratamiento con ADNasa I mediante electroforesis en gel de agarosa y tinción con bromuro de etidio. Después de la extracción con fenol, se preparó ADNc a partir de la muestra usando SUPERScript™ Choice System siguiendo las instrucciones del fabricante (GibcoBRL). Se realizó la transcripción inversa ficticia sin transcriptasa inversa de un control negativo de ARN para cada muestra de ARN.

Se midió la expresión de cinasa novedosa mediante PCR cuantitativa TaqMan® (Perkin Elmer Applied Biosystems) en ADNc preparado a partir de los siguiente tejidos y líneas celulares humanos como se muestra en las Figuras 4A y 4B: pulmón; riñón; corazón; bazo; hígado fetal, hígado normal (NDR 200, PIT 260) e hígado fibrótico (LF/NDR 191, LF/NDR 194, LF/NDR 079, NDR 141, NDR 190, NDR 192, NDR 225, NDR 249, NDR 250, NDR 251, PIT 324, PIT 345), nódulos linfáticos y amígdala; hepatocitos no cultivados; células HepG2; células HepG2 cultivadas en presencia de TGF-beta; células HepG2 transfectadas con HBV; células Hep3 cultivadas con oxígeno normal y bajo; granulocitos normales; fibroblastos dérmicos y pulmonares cultivos con y sin TGF-beta; células estrelladas quiescentes, en reposo y sometidas a pasada; células mononucleares de sangre periférica activadas por fitohemaglutinina (PBMC); linfocitos T CD3+; CD4+ y CD8+; la línea de linfocitos T (Molt 4), células Th1 y Th2 estimuladas durante 24 horas con anticuerpo anti-CD3 y células JY; linfocitos B CD19+ en reposo y activados con lipopolisacárido; células CD14+; células CD14-/CD15+; células CD34+ de sangre periférica inmovilizada (MPB CD34+; médula ósea en reposo adulta (ABM CD34+; médula ósea adulta movilizada (MBM CD11b-); células mononucleares de médula ósea (BM MNC); cordón umbilical neonatal; células eritroides; megacariocitos; neutrófilos; células hepáticas infectadas con el virus de la hepatitis C (HCV); líneas celulares humanas transformadas, incluyendo de eritroleucemia (K562) y leucemia promielocítica aguda (HL60). Se usó un control sin molde (NTC) como muestra de referencia.

## ES 2 327 903 T3

Se diseñaron las sondas mediante el software PrimerExpress (PE Biosystems) basado en la secuencia del gen de cinasa h14189. Los cebadores y sondas para el análisis de expresión de h14189 y  $\beta$ -2-microglobulina son los siguientes:

|    |                                                      |                             |                 |
|----|------------------------------------------------------|-----------------------------|-----------------|
| 5  | Cebador de codificación de h14189                    | ACCGCAGCCGCGTCT             | (SEQ ID NO: 5)  |
| 10 | Cebador inverso de h14189                            | TCTCTCTCGAGGTGCCGCT         | (SEQ ID NO: 6)  |
|    | Sonda TaqMan de h14189                               | TCAGATGATCCTGGAGTGTGGAGGCA  | (SEQ ID NO: 7)  |
| 15 | Cebador de codificación de $\beta$ -2 microglobulina | CACCCCCACTGAAAAAGATGA       | (SEQ ID NO: 8)  |
| 20 | Cebador inverso de $\beta$ -2 microglobulina         | CTTA ACTATCTTGGGCTGTGACAAAG | (SEQ ID NO: 9)  |
|    | Sonda TaqMan de $\beta$ -2 microglobulina            | TATGCCTGCCGTGTGAACACGTG     | (SEQ ID NO: 10) |

25 Se marcó la sonda del gen de cinasa h14189 usando FAM (6-carboxifluoresceína) y se marcó la sonda de referencia de  $\beta$ -2 microglobulina con un tinte fluorescente diferente, VIC. El marcaje diferencial del gen diana y del gen de referencia interna permitió así la medida en el mismo pocillo. Se añadieron cebadores de codificación e inverso y las sondas para tanto el gen de  $\beta$ -2 microglobulina como diana a TaqMan® Universal PCR Master Mix (PE Applied Biosystems). Aunque la concentración final de cebador y sonda podían variar, cada una era internamente consistente dentro de un experimento dado. Un experimento típico contenía 200 nM de cebadores de codificación e inversos más 100 nM de sonda para  $\beta$ -2 microglobulina y 600 nM de cebadores de codificación e inversos más 200 nM de sonda para el gen diana. Se llevaron a cabo los experimentos en matriz TaqMan en un sistema de detección de secuencia ABI PRISM 7700 (PE Applied Biosystems). Las condiciones del ciclador térmico fueron las siguientes: mantenido durante 30 2 min a 50°C y 10 min a 95°C, seguido de PCR de dos etapas durante 40 ciclos de 95°C durante 15 s, seguido de 60°C durante 1 min.

Se usó el siguiente procedimiento para calcular cuantitativamente la expresión del gen de cinasa h14189 en diversos tejidos respecto a la expresión de  $\beta$ -2 microglobulina en el mismo tejido. Se define el valor de ciclo umbral (Ct) como el ciclo en que se detecta un aumento estadísticamente significativo de fluorescencia. Un valor de Ct es indicativo de una mayor concentración de ARNm. El valor de Ct del gen de cinasa se normaliza restando el valor de Ct del gen de  $\beta$ -2 microglobulina, obteniéndose el valor de  $\Delta$ Ct usando la siguiente fórmula:  $\Delta$ Ct = Ct<sub>cinasa</sub> - Ct <sub>$\beta$ -2 microglobulina</sub>. Se calibra entonces la expresión frente a una muestra de ADNc, mostrando un nivel comparativamente bajo de expresión del gen de cinasa. Se resta entonces el valor de  $\Delta$ Ct de la muestra calibradora de la  $\Delta$ Ct de cada muestra de tejido según la siguiente fórmula:  $\Delta\Delta$ Ct =  $\Delta$ Ct<sub>muestra</sub> -  $\Delta$ Ct<sub>calibrador</sub>. Se calcula entonces la expresión relativa usando la fórmula aritmética dada por  $2^{-\Delta\Delta$ Ct}. Se representa entonces gráficamente la expresión del gen de cinasa diana en cada uno de los tejidos como se debate como más detalle a continuación.

Las Figuras 4A y 4B muestran la expresión del gen h14189 determinada mediante PCR cuantitativa TaqMan® en 50 diversos tejidos y tipos celulares. Como se muestra en las Figuras 4A y 4B, se observaron los niveles más altos de expresión de h14189 en tejido de riñón y pulmón normales, células HepG2 transfectadas con HBV (HepG2.2.15), muestras de hígado fibróticas, una línea celular de eritroleucemia humana transformada (K562) y células de sangre periférica movilizada (MPB CD34+). Se observaron también niveles significativos de expresión de h14189 en células hepáticas estrelladas humanas, bazo, amígdala, sangre de cordón umbilical, médula ósea y fibroblastos dérmicos.

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada seleccionada del grupo constituido por:

a) una molécula de ácido nucleico que comprende una secuencia nucleotídica que es idéntica a la secuencia nucleotídica de SEQ ID NO: 1, el inserto de ADNc del plásmido depositado en la ATCC con número de acceso PTA-2333, o un complemento del mismo;

b) una molécula de ácido nucleico que codifica un polipéptido que comprende la secuencia aminoacídica de SEQ ID NO: 2, o una secuencia aminoacídica codificada por el inserto de ADNc del plásmido depositado en la ATCC con número de acceso PTA-2333.

2. La molécula de ácido nucleico de la reivindicación 1, que comprende adicionalmente secuencias de ácido nucleico de vector.

3. La molécula de ácido nucleico de la reivindicación 1 ó 2, que comprende adicionalmente secuencias de ácido nucleico que codifican un polipéptido heterólogo.

4. Una célula hospedadora que contiene la molécula de ácido nucleico de cualquiera de las reivindicaciones 2 ó 3.

5. Un polipéptido aislado codificado por un ácido nucleico según la reivindicación 1.

6. El polipéptido de la reivindicación 5, que comprende adicionalmente secuencias aminoacídicas heterólogas.

7. Un procedimiento para detectar la presencia de una molécula de ácido nucleico de cualquiera de las reivindicaciones 1 a 3 en una muestra, que comprende las etapas de:

a) poner en contacto la muestra con una sonda o cebador de ácido nucleico que hibrida selectivamente con dicha molécula de ácido nucleico; y

b) determinar si la sonda o cebador de ácido nucleico se une a una molécula de ácido nucleico en la muestra.

8. Un procedimiento para identificar un compuesto que se une a un polipéptido de la reivindicación 5 ó 6, que comprende las etapas de:

a) poner en contacto un polipéptido, o una célula que expresa un polipéptido, de la reivindicación 5 ó 6 según pueda ser el caso, con un compuesto de ensayo; y

b) determinar si el polipéptido se une al compuesto de ensayo.

Fichero de entrada Fbh14189a.seq; Fichero de salida Fbh14189a.pat  
Longitud de secuencia 1914

|                                                                                 |                 |      |
|---------------------------------------------------------------------------------|-----------------|------|
| GTTCCGACCMCGCGTCCGCGGACGCGTGGGCGGCGTCCTGGCCGCC                                  | M C T V V D P R | 8    |
| ATT GTC CGG AGA TAC CTA CTC AGG CGG CAG CTC GGG CAG GGG GCC TAT GGC ATT GTG TGG |                 | 71   |
| K A V D R R T G E V V A I K K I F D A F                                         |                 | 28   |
| AAG GCA GTG CAC CGG AGG ACT GGT GAG GTC GTG GCC ATC AAG AAA ATC TTT GAT GCT TTT |                 | 131  |
| R D K T D A O R T F R E I T L L O E F G                                         |                 | 48   |
| AGG GAT AAG ACA GAT GCC CAG AGA ACA TTC CGG GAA ATC ACG CTC CTC CAG GAG TTT GGG |                 | 191  |
| D H P N I I S L L D V I R A E N D R D I                                         |                 | 68   |
| GAC CAT CCC AAC ATC ATC AGC CTC CTT GAC GTG ATC CGG GCA GAG AAC GAC AGG GAC ATT |                 | 251  |
| Y L V F E F M D T D L N A V I R K G G L                                         |                 | 88   |
| TAC CTG GTG TTT GAG TTT ATG GAC ACT GAC CTG AAC GCA GTC ATC CGG AAG GGC GGC CTG |                 | 311  |
| L O D V H V R S I F Y O L L R A T R F L                                         |                 | 108  |
| CTG CAG GAC GTC CAC GTG CGC TCC ATC TTC TAC CAG CTC CTG CGG GCC ACC CGG TTC CTC |                 | 371  |
| H S G H V V H R D O K P S N V L L D A N                                         |                 | 128  |
| CAC TCG GGG CAC GTT GTG CAC CGG GAC CAG AAG CCG TCC AAT GTG CTC CTG GAT GCC AAC |                 | 431  |
| C T V K L C D F G L A R S L G D L P E G                                         |                 | 148  |
| TGC ACA GTG AAG CTG TGT GAC TTT GGC CTG GCC CGC TCC CTG GGC GAC CTC CCT GAG GGG |                 | 491  |
| P E D O A V T E Y V A T R W Y R A P E V                                         |                 | 168  |
| CCT GAG GAC CAG GCC GTG ACA GAG TAC GTG GCC ACA CGC TGG TAC CGA GCA CCG GAG GTG |                 | 551  |
| L L S S H R Y T L G V D M W S L G C I L                                         |                 | 188  |
| CTG CTC TCT TCG CAC CGA TAC ACC CTT GGG GTG GAC ATG TGG AGT CTG GGC TGT ATC CTG |                 | 611  |
| G E M L R G R P L F P G T S T L H O L E                                         |                 | 208  |
| GGG GAG ATG CTG CGG GGG AGA CCC CTG TTC CCC GGC ACG TCC ACC CTN CAC CAG CTG GAG |                 | 671  |
| L I L E T I P P P S E E D L L A L G S G                                         |                 | 228  |
| CTG ATC CTG GAG ACC ATC CCA CCG CCA TCT GAG GAG GAC CTC CTG GCT CTC GGC TCA GGC |                 | 731  |
| C R A S V L H O L G S R P R O T L D A L                                         |                 | 248  |
| TGC CGT GCC TCT GTG CTG CAC CAG CTG GGG TCC CGG CCA CGA CAG ACG CTG GAT GCC CTC |                 | 791  |
| L P P D T S P E A L D L L R R L L V F A                                         |                 | 268  |
| CTA CCG CCA GAC ACC TCC CCA GAG GCC TTG GAC CTC CTT AGG CGA CTC CTG GTG TTC GCC |                 | 851  |
| P D K R L S A T O A L O H P Y V O R F H                                         |                 | 288  |
| CCG GAC AAG CGG TTA AGC GCG ACC CAG GCA CTG CAG CAC CCC TAC GTG CAG AGG TTC CAC |                 | 911  |
| C P S D E W A R E A D V R P R A H E G V                                         |                 | 308  |
| TGC CCC AGC GAC GAG TGG GCA CGA GAG GCA GAT GTG CGG CCC CGG GCA CAC GAA GGG GTC |                 | 971  |
| O L S V P E Y R S R V Y O M I L E C G G                                         |                 | 328  |
| CAG CTC TCT GTG CCT GAG TAC CGC AGC CGC GTC TAT CAG ATG ATC CTG GAG TGT GGA GGC |                 | 1031 |
| S S G T S R E K G P E G V S P S O A H L                                         |                 | 348  |
| AGC AGC GCC ACC TCG AGA GAG AAG GGC CCG GAG GGT GTC TCC CCA AGC CAG GCA CAC CTG |                 | 1091 |
| H K P R A D P O L P S R T P V O G P R P                                         |                 | 368  |
| CAC AAA CCC AGA GCC GAC CCT CAG CTG CCT TCT AGG ACA CCT GTG CAG GGT CCC AGA CCC |                 | 1151 |
|                                                                                 |                 | 388  |
|                                                                                 |                 | 1211 |

FIG. 1A.

|                                                                                 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|---------------------------------------------------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| R                                                                               | P   | O   | S   | S   | P   | G   | H   | D   | P   | A   | E   | H   | E   | S   | P   | R   | A   | A   | K   | 408  |
| AGG                                                                             | CCC | CAG | AGC | AGC | CCA | GGC | CAT | GAC | CCT | GCC | GAG | CAC | GAG | TCC | CCC | CGT | GCA | GCC | AAG | 1271 |
| N                                                                               | V   | P   | R   | Q   | N   | S   | A   | P   | L   | L   | Q   | T   | A   | L   | L   | G   | N   | G   | E   | 428  |
| AAC                                                                             | GTT | CCC | AGG | CAG | AAC | TCC | GCT | CCC | CTG | CTC | CAA | ACT | GCT | CTC | CTA | GGG | AAT | GGG | GAA | 1331 |
| R                                                                               | P   | P   | G   | A   | K   | E   | A   | P   | P   | L   | T   | L   | S   | L   | V   | K   | P   | S   | G   | 448  |
| AGG                                                                             | CCC | CCT | GGG | GCG | AAG | GAA | GCG | CCC | CCC | TTC | ACA | CTC | TGC | CTG | GTG | AAG | CCA | AGC | GGG | 1391 |
| R                                                                               | G   | A   | A   | P   | S   | L   | T   | S   | Q   | A   | A   | A   | Q   | V   | A   | N   | O   | A   | L   | 468  |
| AGG                                                                             | GGA | GCT | GCG | CCC | TCC | CTG | ACC | TCC | CAG | GCT | GCG | GCT | CAG | GTG | GCC | AAC | CAG | GCC | CTG | 1451 |
| I                                                                               | R   | G   | D   | W   | N   | R   | G   | G   | G   | V   | R   | V   | A   | S   | V   | O   | O   | V   | P   | 488  |
| ATC                                                                             | CGG | GGT | GAC | TGG | AAC | CGG | GGC | GGT | GGG | GTG | AGG | GTG | GCC | AGC | GTA | CAA | CAG | GTC | CCT | 1511 |
| P                                                                               | R   | L   | P   | P   | E   | A   | R   | P   | G   | R   | R   | M   | F   | S   | T   | S   | A   | L   | O   | 508  |
| CCC                                                                             | CGG | CTT | CCT | CCG | GAG | GCC | CGG | CCC | GGC | CGG | AGG | ATG | TTC | AGC | ACC | TCT | GCC | TTC | CAG | 1571 |
| G                                                                               | A   | O   | G   | G   | A   | R   | A   | L   | L   | G   | G   | Y   | S   | Q   | A   | Y   | G   | T   | V   | 528  |
| GGT                                                                             | GCC | CAG | GGG | GGT | GCC | AGG | GCT | TTC | CTT | GGA | GGC | TAC | TCC | CAA | GCC | TAC | GGG | ACT | GTC | 1631 |
| C                                                                               | H   | S   | A   | L   | G   | H   | L   | P   | L   | L   | E   | G   | H   | H   | V   | *   |     |     |     | 545  |
| TGC                                                                             | CAC | TCG | GCA | CTG | GGC | CAC | CTG | CCC | CTG | CTG | GAG | GGG | CAC | CAT | GTG | TGA |     |     |     | 1682 |
| GCGGCCCTACTCCCTTCACCTGGCCCTCTGTTCCCTGCCCCAGCCCTTCCCCAGACCCCTCTCCAGTCTCCTGCACCCC |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 1761 |
| TTAGCCCTCCCTGCTTTGCCCTGGCCCGTTGAAGTTCCAGGGAGCTTGCCCGGGTCTCCTCGGGGAGCAGATGAGGGCC |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 1840 |
| CTGCCCCCGCCCACTGACTTTCTCCAATAAAGNCATGTCTGCCCCCCCCNAAAAAAAAAAAAAAAAAAAAA         |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 1914 |

FIG. 1B.

GAP de: FrGcgManager\_453\_GID0yhnfT control: 7821 dc: 1 a: 1998

AF078798 en Genbank

a: FrGcgManager\_453\_HID058iyl control: 2142 de 1 a 1635

Fbh14189a ORF de - Importación completo

Tabla de comparación de símbolos /ddm\_local/gcg/gcg\_9.1/gcgcore/data/rundata/nwsgapdna.cmp  
Control comparativo: 8760

Ponderación de hueco: 12 Emparejamiento medio: 10,000  
Ponderación de longitud: 4 Emparejamiento erróneo medio: 0,000

Calidad: 12076 Longitud: 2022  
Relación: 7,386 Huecos: 13  
Porcentaje de similitud: 77,219 Porcentaje de identidad: 77,157

Umbrales de exhibición de emparejamiento para el/(los) emparejamientos:

| = IDENTIDAD  
: = 5  
. = 1

FrGcgManager\_453\_GID0yhnfT x FrGcgManager\_453\_HID058iyl June 20, 19100 14:05 ..

```

      .      .      .      .      .
151 ACCATGTGTGCTGCCGAGGTGGACCGTCATGTATCCCAGAGATACCTGAT 200
      ||||| ||| ||||| || | || ||||| |||
1  ....ATGTGC.ACCGTAGTGGACCCTCGCATTGTCCGAGATACCTACT 44
      .      .      .      .      .
201 CAAGCGGAGGCTTGGGAAGGGGGCCTACGGCATTGTATGGAAGGCCATGG 250
      || |||| ||| ||| ||||| ||| ||||| ||||| |||
45 CAGGCGGCAGCTCGGGCAGGGGGCCTATGGCATTGTGTGGAAGGCAGTGG 94
      .      .      .      .      .
251 ACCGGAGGACTGGTGAGGTAGTGGCCATCAAGAAAATCTTTGATGCCTTT 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
95 ACCGGAGGACTGGTGAGGTCTGGCCATCAAGAAAATCTTTGATGCCTTT 144
      .      .      .      .      .
301 AGGGACCAGACAGATGCTCAGAGGACCTCCGTGAAATCATGCTTCTCCG 350
      ||||| ||||| ||||| ||| ||||| ||||| ||| |||||
145 AGGGATAAGACAGATGCCCAGAGAACATTCCGGGAAATCACGCTCCTCCA 194
      .      .      .      .      .
351 GGAGTTTGGGGGCCATCCCAACATCATCCGCCTGCTTGATGTAATCCCAG 400
      ||||| ||||| ||||| ||||| ||||| ||| ||||| |
195 GGAGTTTGGGGACCATCCCAACATCATCAGCCTCCTTGACGTGATCCGGG 244
      .      .      .      .      .
401 CAAAGAATGACAGGGATATTTACCTGGTGTGAGTCCATGGACACCGAC 450

```

**FIG. 2A.**

245 CAGAGAACGACAGGGACATTTACCTGGTGTGTTGAGTTTATGGACACTGAC 294  
 451 CTGAACGCGGTCATCCAGAAGGGCAGACTGTTGGAGGACATCCACAAACG 500  
 295 CTGAACGCAGTCATCCGGAAGGGCGGCCTGCTGCAGGACGTCCACGTGCG 344  
 501 TTGCATCTTTTACCAGCTCCTGAGAGCCACCAAGTTTATCCATTCAGGGC 550  
 345 CTCCATCTTCTACCAGCTCCTGCGGGCCACCCGGTTCCTCCACTCGGGGC 394  
 551 GCGTCATTACCGGGACCAGAAGCCAGCCAACGTTCTATTGGATGCTGCT 600  
 395 ACGTTGTGCACCGGGACCAGAAGCCGTCCAATGTGCTCCTGGATGCCAAC 444  
 601 TGCCGGGTGAACTATGTGACTTTGGCCTGGCACGCTCCCTCAGTGACTT 650  
 445 TGCACAGTGAAGCTGTGTGACTTTGGCCTGGCCCGCTCCCTGGGCGACCT 494  
 651 CCCTGA.AGGCCT..GGGCCAGGCCCTGACAGAATATGTGGCCACACGCT 697  
 495 CCCTGAGGGGCCTGAGGACCAGGCCGTGACAGAGTACGTGGCCACACGCT 544  
 698 GGTACCGAGCTCCAGAGGTGCTTCTGTCTTCCCGATGGTATACCCCTGGG 747  
 545 GGTACCGAGCACCGGAGGTGCTGCTCTCTTCGCACCGATACACCCTTGGG 594  
 748 GTGGACATGTGGAGCCTGGGCTGCATACTGGGAGAGATGCTTCGAGGGCA 797  
 595 GTGGACATGTGGAGTCTGGGCTGTATCCTGGGGGAGATGCTGCGGGGGAG 644  
 798 GCCACTGTTTCCGGGTACATCTACTTTCCACCAGCTGGAGCTGATCCTGG 847  
 645 ACCCCTGTTCCCCGGCACGTCCACCCTNCACCAGCTGGAGCTGATCCTGG 694  
 848 AGACCATTCCATTGCCCTCCATGGAGGAGCTCCAGGGCCTTGATCAGAC 897  
 695 AGACCATCCCACCGCCATCTGAGGAGGACCTCCTGGCTCTCGGCTCAGGC 744  
 898 TACAGTGCTTTGATTCTGCAGAATCTTGGATCCAGGCCACGGCAGACGCT 947  
 745 TGCCGTGCCTCTGTGCTGCACCAGCTGGGGTCCCAGGCCACGACAGACGCT 794  
 948 GGACGCCCTCCTGCCGCCAGACACCCCCCAGAAGCCCTGGACCTCCTCA 997  
 795 GGATGCCCTCCTACCGCCAGACACCTCCCAGAGGCCTTGACCTCCTTA 884  
 998 AGCGACTCTTGGCATTGCTCCGGACAAACGGCTTAGTGCAGAGCAGGCT 1047  
 845 GGCGACTCCTGGTGTTCGCCCCGGACAAGCGGTTAAGCGCGACCCAGGCA 894

**FIG. 2B.**

1048 CTTCAACACCCTTACGTGCAGAGATTCCATTCGCCCCGACCGCAGTGCGAC 1097  
|| || ||||| ||||||||| ||||| ||||| | ||||| |  
895 CTGCAGCACCCCTACGTGCAGAGGTTCCACTGCCCCAGCGACGAGTGGGC 944  
1098 ACGGGGGTCCGATGTGCGGCTCCCGTACACGAAGGAGACCAGCTCTCTG 1147  
||| | | ||||||||| || || ||||||||| | |||||||||  
945 ACGAGAGGCAGATGTGCGGCCCCGGGCACACGAAGGGGTCCAGCTCTCTG 994  
1148 CACCAGAGTATCGCAACCGCCTGTACCAGATGATCCTGGAGCGGAGACGG 1197  
|| ||||| ||||| ||||| | || ||||||||| || ||| |  
995 TGCCTGAGTACCGCAGCCGCGTCTATCAGATGATCCTGGAGTGTGGAGGC 1044  
1198 AACAGCCGCAGCCCTCGAGAGGAAGACTTGGGGGTGTGGCCTCGCGGGC 1247  
| |||| ||| | | ||||| | | || ||| ||| | ||| |  
1045 AGCAGCGGCACCTCGAGAGAGAAGGGCCCGAGGGTGT...CTCCC...C 1088  
1248 AGAGCTCAGGGCTTCCCAGAGGCAATCGCTCAAGCCCGAGTCTCTCCCC 1297  
| ||| |||| | | || || || || || || || || || ||  
1089 A.AGC.CAGG.....CA....CACCTGCACAAACCCAGAGCCGACCCCTC 1126  
1298 AGGTCTTGGCGGAGACGCCTGCGCGGAAACGCGGACCCAAACCTCAGAAT 1347  
|| | | || |||| || || | | |||||| || ||||  
1127 AGCTGCCTTCTAGGACACCTGTGCAGGGTCCAGACCCAGGCCCCAGAGC 1176  
1348 GGCCATGGTCACGA...TCCCGAGCATGTG.....GAA 1377  
||| || || || | ||||||| | | |||  
1177 AGCCCAGGCCATGACCCTGCCGAGCACGAGTCCCCCGTGCAGCCAAGAA 1226  
1378 .GTTTCGAGGCAGAGTTTACAGCCCCCTGTACCAACTTCGCGCCGAGGCA 1426  
|||| |||||||| || | |||||| |||| | | | |||||  
1227 CGTTCCAGGCAGAACTCCGCTCCCTGCTCAAAGTGTCTCTTAGGGA 1276  
1427 GCGGGGAAAGGCCCCCAGGGGCCACAGGGGAGCCACCCCTCCGCACCCTCA 1476  
||||||| ||||| ||||| | | ||| |||| ||| |||  
1277 ATGGGGAAAGGCCCCCTGGGGCGAAGGAAGCGCCCCCTTGACACTCTCG 1326  
1477 GGGGTGAAGACTCACGTTAGGGCGGTGGCGCCCTCCCTGACTTCACAGGC 1526  
||||||| | || |||| | ||||||||| || |||||  
1327 CTGGTGAAGCCAAGCGGGAGGGGAGCTGCGCCCTCCCTGACCTCCCAGGC 1376  
1527 CGCGGCTCAGGCGGCCAATCAGCCTCTGATCCGCAGTGATCCGGCCCCGGG 1576  
||||||| ||||| |||| | |||||| |||| | |||||  
1377 TGCGGCTCAGGTGGCCAACAGGCCCTGATCCGGGTGACTGGAACCGGG 1426  
1577 GCGGTGGGGCAAGGGCTGTGCGGCGCGGACGGGTCCCTTCCCGCCTGCCC 1626  
||||||| |||| | | ||| | || |||||| |||| || ||  
1427 GCGGTGGGGTGAGGGTGGCCAGCGTACAACAGGTCCCTCCCCGGCTTCCT 1476  
1627 CGGGAGGCCCCGGAACCCCGACCCGGCCGAAGGATGTTGGCATCTCGGT 1676

**FIG. 2C.**

1477 CCGGAGG.....CCCGGCCCGGCCGAGGATGTTTCAGCACCTCTGC 1517  
1677 CTCGCAGGGGGCCAGGGTGCAGCCAGAGCTGCTCTTGCGGCTACTCCC 1726  
1518 CTTGCAGGGTGCCAGGGGGGTGCCAGGGCTTTGCTTGAGGCTACTCCC 1567  
1727 AGGCCTACGGGACCGTGTGCCGCTCCGCGCTGGGCGCCTGCCTCTGCTC 1776  
1568 AAGCCTACGGGACTGTCTGCCACTCGGCACTGGGCCACCTGCCCCTGCTG 1617  
1777 CCCGGACCGCGTGCCTGAGCCGCCACCAACCTCCTTGCGGCAAACTGGC 1826  
1618 GAGGGGCACCATGTGTGA..... 1635

•  
•  
•

FIG. 2D.

GAP de: FrGcgManager\_454\_LIDsNLRN\_ control:9383 de: 1 a: 546

4220888 en Genbank

a: FrGcgManager\_454\_MID0hcGgX control:6114 de 1 a 545

Fbh14189 aa - Importación - completo

Tabla de comparación de símbolos /prod/ddm/seqanal/BLAST/matrix/aa/BLOSUM62

Control comparativo: 1102

Matriz preparada con matblas a partir de blosum62.ijj

Ponderación de hueco: 12 Emparejamiento medio: 2,778

Ponderación de longitud: 4 Emparejamiento erróneo medio: -2,248

Calidad: 1623

Longitud: 555

Relación: 2,978

Huecos: 5

Porcentaje de similitud: 69,216 Porcentaje de identidad: 66,604

Umbrales de exhibición de emparejamiento para el/(los) emparejamientos:

| = IDENTIDAD

: = 2

. = 1

FrGcgManager\_454\_LIDsNLRN\_ x FrGcgManager\_454\_MID0hcGgX June 20, 19100 14:28 ..

```

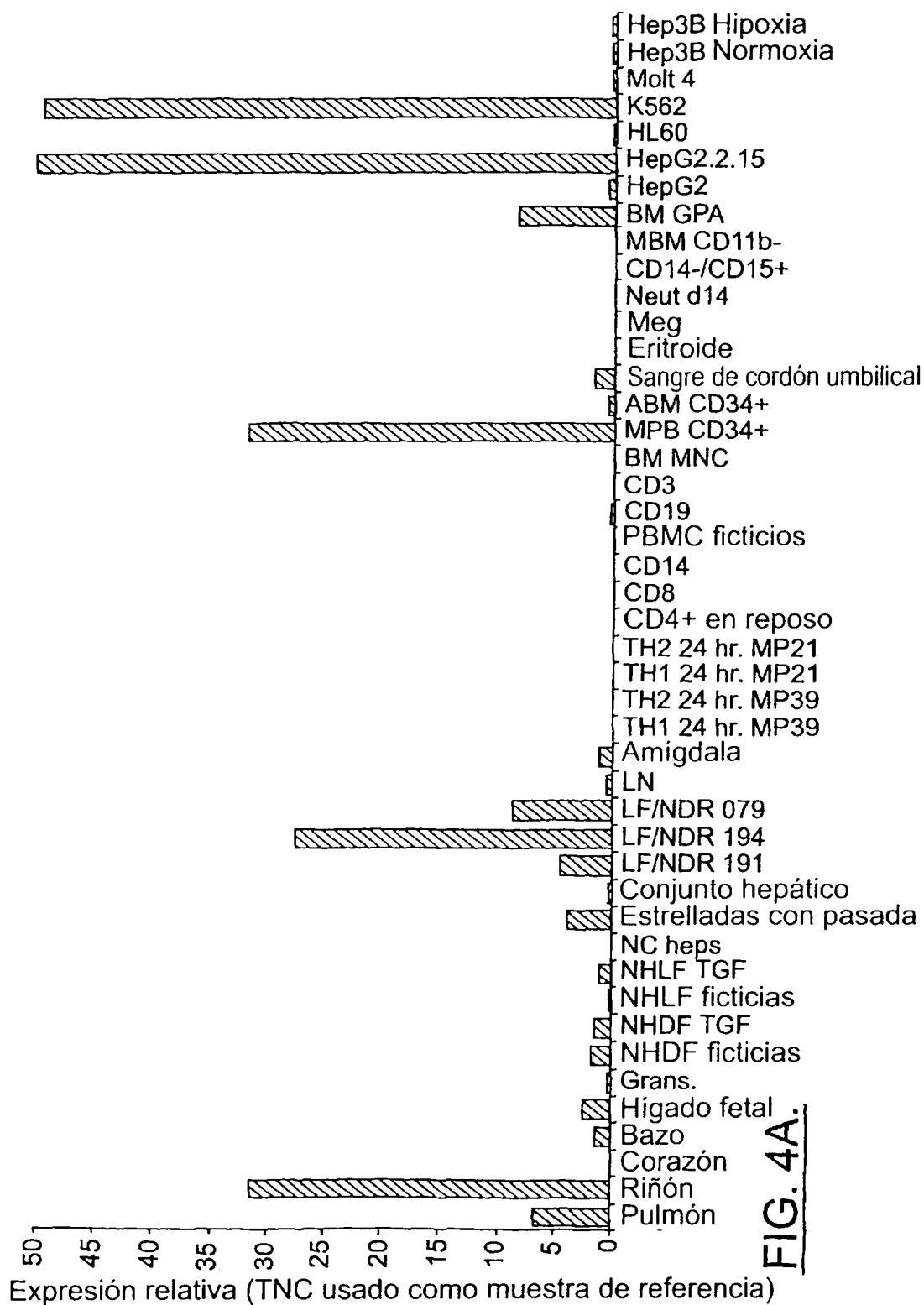
1  MCAA.EVDRHVSQRYLIKRR.LKGAYGIVWKAMDRRTGEVVAIKKIFDAFR 50
   ||  ||  : .|||:|.||.|||||||.|||||||.|||||||.||||||
1  MCTV.VDPRIVRRYLLRRQLGQAYGIVWKA.VDRRTGEVVAIKKIFDAFR 49
   ||  ||  : .|||:|.||.|||||||.|||||||.|||||||.||||||
51 DQ.TDAQRTFREIMLLREFGGHPNIIRLLDVIPAKNDRDIYLVFESMDTDL 100
   |.|||||||.|||.|||.|||||.|||||.|||||||.|||||||.||||
50 DKTDAQRTFREITLLQ.EFGDHPNIISLLDVIRAENDRDIYLVFEFMDTDL 99
   |.|||||||.|||.|||.|||||.|||||.|||||||.|||||||.||||
101 NAVIQKGR.LLEDIHKRCIFYQLLRATKFIHSGRVIHRDQKPANVLLDAAC 150
   |||.|| |::|:| | |||||::|:| | |::|:| | |||||.|||||
100 NAVIRKGG.LLQDVHVR.SIFYQLLRATRFLHSGHVVRDQKPSNVLLDANC 149
   |||.|| |::|:| | |||||::|:| | |::|:| | |||||.|||||
151 RVKLCDFGLARSL.SDFPEG.LGQALTEYVATRWYRAPEVLLSSRWYTPGV 199
   |||.|| |::|:| | |||||::|:| | |::|:| | |||||.|||||
150 TVKLCDFGLARSL.GDLPEGPEDQAVTEYVATRWYRAPEVLLSSHRYTLGV 199
   |||.|| |::|:| | |||||::|:| | |::|:| | |||||.|||||
200 DMWSLGCILGEMLRGQPLFPGTSTFHQLELILETIPLPSMEELQGLGSDY 249
   || | | |::| | | |||| |::|:| | | | |::|:| | |||||.
200 DMRGLATIQGQILPGGQAQPGPPRQHQLEHTIXRTPXPSEEEHLGRGPGE 249
   || | | |::| | | |||| |::|:| | | | |::|:| | |||||.
250 SALILQNLGSRPRQTL.DALLPPDTPPEALDLLKRL.LAFAPDKRLSAEQAL 299
   || | | |::| | | |||| |::|:| | | | |::|:| | |||||.
250 RPCILHTPAHRPRQTL.DALLPPDTSPEALDLLRRL.LVFAPDKRLSATQAL 299
   || | | |::| | | |||| |::|:| | | | |::|:| | |||||.

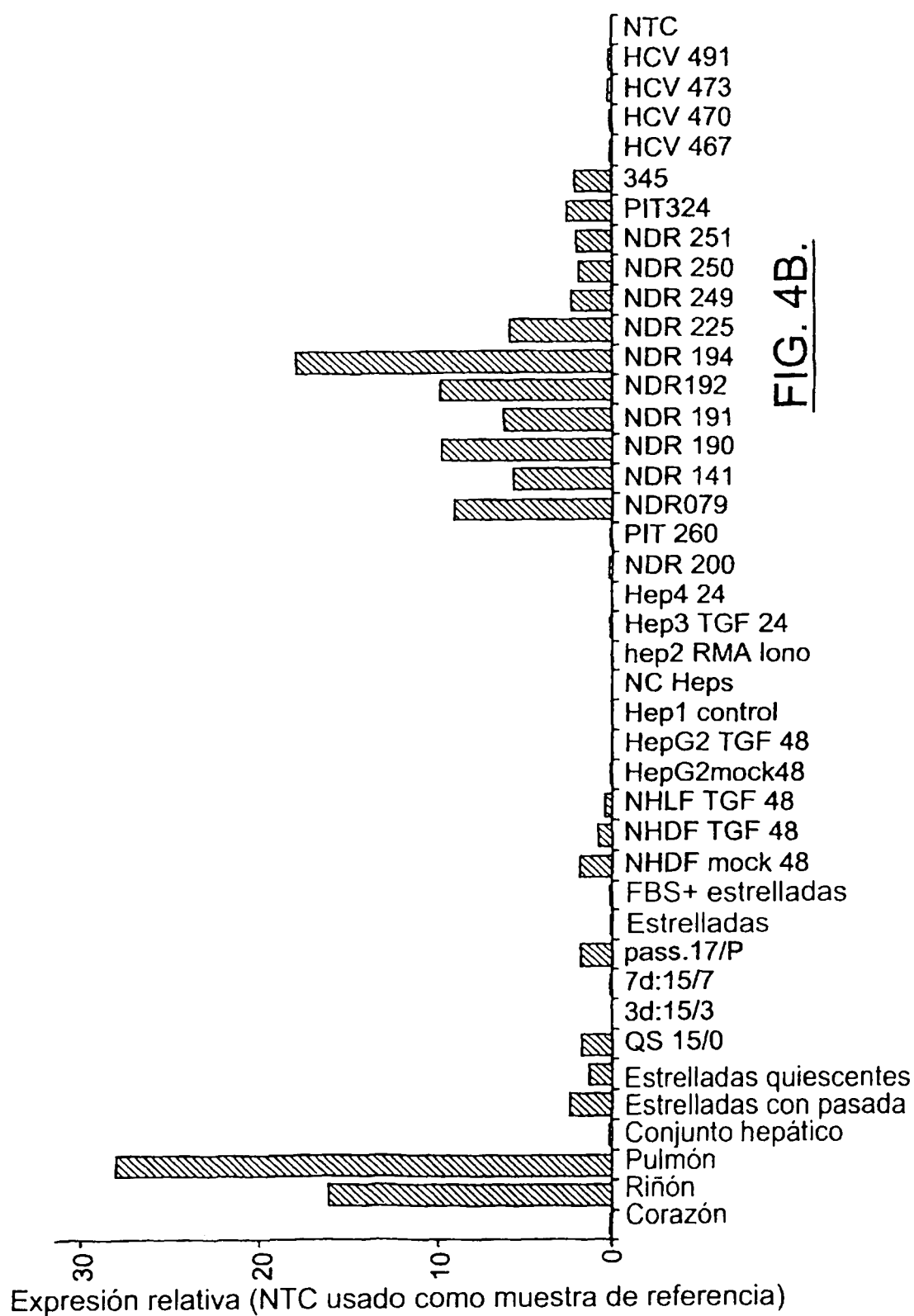
```

**FIG. 3A.**

300 QHPYVQRFHCPDREWTRGSDVRLPVHEGDQLSAPEYRNRLYQMILERRRN 349  
 ||||| ||| .||| ||| ||| |||||.|||  
 300 QHPYVQRFHCPSDEWAREADVPRRAHEGVQLSVPEYRSRVYQMILECGGS 349  
 350 SRSPREEDLGVVASRAELRASQRQSLKPGVLPQVLAETPARKRGPKPONG 399  
 |.||. |. || || ||. ||. |:||.  
 350 SGTSREKGPEGVS.....PSQAH LHKPRADPQLPSRTPVQGP RP RPQSS 393  
 400 HGHDP EHV E.....VRRQSSDPLYQLPPPGSGERPPGATGEPPSAPSG 442  
 |||| | |||.||| |.||||| || |  
 394 PGHDP AEHESPRAAKNVPRONSAPLLQTALLGNGERPPGAKEAPPLTSL 443  
 443 VKTHVRAVAPSLTSQAAAQAANQPLIRSDPARGGGPRAVGARRVPSRLPR 492  
 || | ||||| ||| ||| | ||| | :|| |||  
 444 VKPSGRGAAPSLTSQAAAQVANQALIRGDWNRGGGV RVASXNKVPPRLP. 492  
 493 EAP EPRPGRRMFGISVSQGAQGAARAALGGYSQAYGTVCRSALGRLPLLP 542  
 || ||||| | |||| || ||||| ||||| |||||  
 493 ..PEARPGRRMFSTSALQGAQGGARALLGGYSQAYGTVCHSALGHLPLLE 540  
 543 GPRA. 546  
 |  
 541 Ghhv\* 545

**FIG. 3B.**





# ES 2 327 903 T3

## LISTA DE SECUENCIAS

|    |                                                                                                                                                           |  |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | <110> Hodge, Martin R.                                                                                                                                    |  |
| 5  | <120> 14189, una cinasa humana novedosa y usos de la misma                                                                                                |  |
|    | <130> 035800/238036 (5800-153-1)                                                                                                                          |  |
|    | <140> US09/641.690                                                                                                                                        |  |
|    | <141> 18-08-2000                                                                                                                                          |  |
| 10 | <160> 13                                                                                                                                                  |  |
|    | <170> FastSEQ for Windows versión 4.0                                                                                                                     |  |
|    | <210> 1                                                                                                                                                   |  |
| 15 | <211> 1914                                                                                                                                                |  |
|    | <212> ADN                                                                                                                                                 |  |
|    | <213> <i>Homo sapiens</i>                                                                                                                                 |  |
|    | <220>                                                                                                                                                     |  |
| 20 | <221> CDS                                                                                                                                                 |  |
|    | <222> (48)...(1682)                                                                                                                                       |  |
|    | <221> misc_feature                                                                                                                                        |  |
|    | <222> 719, 1874, 1893                                                                                                                                     |  |
| 25 | <223> n = A, T, C o G                                                                                                                                     |  |
|    | <400> 1                                                                                                                                                   |  |
| 30 | gttccgaccm cgcgtccgcg gacgcgtggg cgggcgtcct ggccgcc atg tgc acc 56<br>Met Cys Thr<br>1                                                                    |  |
| 35 | gta gtg gac cct cgc att gtc cgg aga tac cta ctc agg cgg cag ctc 104<br>Val Val Asp Pro Arg Ile Val Arg Arg Tyr Leu Leu Arg Arg Gln Leu<br>5 10 15         |  |
| 40 | ggg cag ggg gcc tat ggc att gtg tgg aag gca gtg gac cgg agg act 152<br>Gly Gln Gly Ala Tyr Gly Ile Val Trp Lys Ala Val Asp Arg Arg Thr<br>20 25 30 35     |  |
| 45 | ggt gag gtc gtg gcc atc aag aaa atc ttt gat gct ttt agg gat aag 200<br>Gly Glu Val Val Ala Ile Lys Lys Ile Phe Asp Ala Phe Arg Asp Lys<br>40 45 50        |  |
| 50 | aca gat gcc cag aga aca ttc cgg gaa atc acg ctc ctc cag gag ttt 248<br>Thr Asp Ala Gln Arg Thr Phe Arg Glu Ile Thr Leu Leu Gln Glu Phe<br>55 60 65        |  |
| 55 | ggg gac cat ccc aac atc atc agc ctc ctt gac gtg atc cgg gca gag 296<br>Gly Asp His Pro Asn Ile Ile Ser Leu Leu Asp Val Ile Arg Ala Glu<br>70 75 80        |  |
| 60 | aac gac agg gac att tac ctg gtg ttt gag ttt atg gac act gac ctg 344<br>Asn Asp Arg Asp Ile Tyr Leu Val Phe Glu Phe Met Asp Thr Asp Leu<br>85 90 95        |  |
| 65 | aac gca gtc atc cgg aag ggc ggc ctg ctg cag gac gtc cac gtg cgc 392<br>Asn Ala Val Ile Arg Lys Gly Gly Leu Leu Gln Asp Val His Val Arg<br>100 105 110 115 |  |
| 70 | tcc atc ttc tac cag ctc ctg cgg gcc acc cgg ttc ctc cac tcg ggg 440<br>Ser Ile Phe Tyr Gln Leu Leu Arg Ala Thr Arg Phe Leu His Ser Gly<br>120 125 130     |  |
| 75 | cac gtt gtg cac cgg gac cag aag ccg tcc aat gtg ctc ctg gat gcc 488                                                                                       |  |

# ES 2 327 903 T3

|    |                                                                                                                                    |      |
|----|------------------------------------------------------------------------------------------------------------------------------------|------|
|    | His Val Val His Arg Asp Gln Lys Pro Ser Asn Val Leu Leu Asp Ala                                                                    |      |
|    | 135 140 145                                                                                                                        |      |
| 5  | aac tgc aca gtg aag ctg tgt gac ttt ggc ctg gcc cgc tcc ctg ggc<br>Asn Cys Thr Val Lys Leu Cys Asp Phe Gly Leu Ala Arg Ser Leu Gly | 536  |
|    | 150 155 160                                                                                                                        |      |
| 10 | gac ctc cct gag ggg cct gag gac cag gcc gtg aca gag tac gtg gcc<br>Asp Leu Pro Glu Gly Pro Glu Asp Gln Ala Val Thr Glu Tyr Val Ala | 584  |
|    | 165 170 175                                                                                                                        |      |
| 15 | aca cgc tgg tac cga gca ccg gag gtg ctg ctc tct tcg cac cga tac<br>Thr Arg Trp Tyr Arg Ala Pro Glu Val Leu Ser Ser His Arg Tyr     | 632  |
|    | 180 185 190 195                                                                                                                    |      |
| 20 | acc ctt ggg gtg gac atg tgg agt ctg ggc tgt atc ctg ggg gag atg<br>Thr Leu Gly Val Asp Met Trp Ser Leu Gly Cys Ile Leu Gly Glu Met | 680  |
|    | 200 205 210                                                                                                                        |      |
| 25 | ctg cgg ggg aga ccc ctg ttc ccc ggc acg tcc acc ctn cac cag ctg<br>Leu Arg Gly Arg Pro Leu Phe Pro Gly Thr Ser Thr Leu His Gln Leu | 728  |
|    | 215 220 225                                                                                                                        |      |
| 30 | gag ctg atc ctg gag acc atc cca ccg cca tct gag gag gac ctc ctg<br>Glu Leu Ile Leu Glu Thr Ile Pro Pro Ser Glu Glu Asp Leu Leu     | 776  |
|    | 230 235 240                                                                                                                        |      |
| 35 | gct ctc ggc tca ggc tgc cgt gcc tct gtg ctg cac cag ctg ggg tcc<br>Ala Leu Gly Ser Gly Cys Arg Ala Ser Val Leu His Gln Leu Gly Ser | 824  |
|    | 245 250 255                                                                                                                        |      |
| 40 | cgg cca cga cag acg ctg gat gcc ctc cta ccg cca gac acc tcc cca<br>Arg Pro Arg Gln Thr Leu Asp Ala Leu Leu Pro Pro Asp Thr Ser Pro | 872  |
|    | 260 265 270 275                                                                                                                    |      |
| 45 | gag gcc ttg gac ctc ctt agg cga ctc ctg gtg ttc gcc ccg gac aag<br>Glu Ala Leu Asp Leu Arg Arg Leu Val Phe Ala Pro Asp Lys         | 920  |
|    | 280 285 290                                                                                                                        |      |
| 50 | cgg tta agc gcg acc cag gca ctg cag cac ccc tac gtg cag agg ttc<br>Arg Leu Ser Ala Thr Gln Ala Leu Gln His Pro Tyr Val Gln Arg Phe | 968  |
|    | 295 300 305                                                                                                                        |      |
| 55 | cac tgc ccc agc gac gag tgg gca cga gag gca gat gtg cgg ccc cgg<br>His Cys Pro Ser Asp Glu Trp Arg Glu Ala Asp Val Arg Pro Arg     | 1016 |
|    | 310 315 320                                                                                                                        |      |
| 60 | gca cac gaa ggg gtc cag ctc tct gtg cct gag tac cgc agc cgc gtc<br>Ala His Glu Gly Val Gln Leu Ser Val Pro Glu Tyr Arg Ser Arg Val | 1064 |
|    | 325 330 335                                                                                                                        |      |
| 65 | tat cag atg atc ctg gag tgt gga ggc agc agc ggc acc tcg aga gag<br>Tyr Gln Met Ile Leu Glu Cys Gly Gly Ser Ser Gly Thr Ser Arg Glu | 1112 |
|    | 340 345 350 355                                                                                                                    |      |
| 70 | aag ggc ccg gag ggt gtc tcc cca agc cag gca cac ctg cac aaa ccc<br>Lys Gly Pro Glu Gly Val Ser Pro Ser Gln Ala His Leu His Lys Pro | 1160 |
|    | 360 365 370                                                                                                                        |      |
| 75 | aga gcc gac cct cag ctg cct tct agg aca cct gtg cag ggt ccc aga<br>Arg Ala Asp Pro Gln Leu Pro Ser Arg Thr Pro Val Gln Gly Pro Arg | 1208 |
|    | 375 380 385                                                                                                                        |      |
| 80 | ccc agg ccc cag agc agc cca ggc cat gac cct gcc gag cac gag tcc<br>Pro Arg Pro Gln Ser Ser Pro Gly His Asp Pro Ala Glu His Glu Ser | 1256 |
|    | 390 395 400                                                                                                                        |      |
| 85 | ccc cgt gca gcc aag aac gtt ccc agg cag aac tcc gct ccc ctg ctc                                                                    | 1304 |

# ES 2 327 903 T3

```

Pro Arg Ala Ala Lys Asn Val Pro Arg Gln Asn Ser Ala Pro Leu Leu
405                      410                      415

caa act gct ctc cta ggg aat ggg gaa agg ccc cct ggg gcg aag gaa 1352
Gln Thr Ala Leu Leu Gly Asn Gly Glu Arg Pro Pro Gly Ala Lys Glu
420                      425                      430                      435

gcg ccc ccc ttg aca ctc tcg ctg gtg aag cca agc ggg agg gga gct 1400
Ala Pro Pro Leu Thr Leu Ser Leu Val Lys Pro Ser Gly Arg Gly Ala
440                      445                      450

gcg ccc tcc ctg acc tcc cag gct gcg gct cag gtg gcc aac cag gcc 1448
Ala Pro Ser Leu Thr Ser Gln Ala Ala Ala Gln Val Ala Asn Gln Ala
455                      460                      465

ctg atc cgg ggt gac tgg aac cgg ggc ggt ggg gtg agg gtg gcc agc 1496
Leu Ile Arg Gly Asp Trp Asn Arg Gly Gly Gly Val Arg Val Ala Ser
470                      475                      480

gta caa cag gtc cct ccc cgg ctt cct ccg gag gcc cgg ccc ggc cgg 1544
Val Gln Gln Val Pro Pro Arg Leu Pro Pro Glu Ala Arg Pro Gly Arg
485                      490                      495

agg atg ttc agc acc tct gcc ttg cag ggt gcc cag ggg ggt gcc agg 1592
Arg Met Phe Ser Thr Ser Ala Leu Gln Gly Ala Gln Gly Gly Ala Arg
500                      505                      510                      515

gct ttg ctt gga ggc tac tcc caa gcc tac ggg act gtc tgc cac tcg 1640
Ala Leu Leu Gly Gly Tyr Ser Gln Ala Tyr Gly Thr Val Cys His Ser
520                      525                      530

gca ctg ggc cac ctg ccc ctg ctg gag ggg cac cat gtg tga 1682
Ala Leu Gly His Leu Pro Leu Leu Glu Gly His His Val *
535                      540

gccgccctac tcccttcacc tggccctctg ttcttgcccc agcccttccc ccagaccctt 1742
ctccagtcctc ctgcaccctc tagccctccc tgctttgcct ggcccgttga agttccaggg 1802
agcttgccccg ggtctcctcg ggggagcaga tgagggccct gcccccgccc cactgacttt 1862
ctccaataaaa gncatgtctg cccccccccc naaaaaaaaaa aaaaaaaaaa aa 1914

```

<210> 2

40 <211> 544

<212> PRT

<213> *Homo sapiens*

45 <400> 2

```

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

```

# ES 2 327 903 T3

Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Val Leu Leu Ser Ser  
 180 185 190  
 His Arg Tyr Thr Leu Gly Val Asp Met Trp Ser Leu Gly Cys Ile Leu  
 195 200 205  
 Gly Glu Met Leu Arg Gly Arg Pro Leu Phe Pro Gly Thr Ser Thr Leu  
 210 215 220  
 His Gln Leu Glu Leu Ile Leu Glu Thr Ile Pro Pro Pro Ser Glu Glu  
 225 230 235 240  
 Asp Leu Leu Ala Leu Gly Ser Gly Cys Arg Ala Ser Val Leu His Gln  
 245 250 255  
 Leu Gly Ser Arg Pro Arg Gln Thr Leu Asp Ala Leu Leu Pro Pro Asp  
 260 265 270  
 Thr Ser Pro Glu Ala Leu Asp Leu Arg Arg Leu Leu Val Phe Ala  
 275 280 285  
 Pro Asp Lys Arg Leu Ser Ala Thr Gln Ala Leu Gln His Pro Tyr Val  
 290 295 300  
 Gln Arg Phe His Cys Pro Ser Asp Glu Trp Ala Arg Glu Ala Asp Val  
 305 310 315 320  
 Arg Pro Arg Ala His Glu Gly Val Gln Leu Ser Val Pro Glu Tyr Arg  
 325 330 335  
 Ser Arg Val Tyr Gln Met Ile Leu Glu Cys Gly Gly Ser Ser Gly Thr  
 340 345 350  
 Ser Arg Glu Lys Gly Pro Glu Gly Val Ser Pro Ser Gln Ala His Leu  
 355 360 365  
 His Lys Pro Arg Ala Asp Pro Gln Leu Pro Ser Arg Thr Pro Val Gln  
 370 375 380  
 Gly Pro Arg Pro Arg Pro Gln Ser Ser Pro Gly His Asp Pro Ala Glu  
 385 390 395 400  
 His Glu Ser Pro Arg Ala Ala Lys Asn Val Pro Arg Gln Asn Ser Ala  
 405 410 415  
 Pro Leu Leu Gln Thr Ala Leu Leu Gly Asn Gly Glu Arg Pro Pro Gly  
 420 425 430  
 Ala Lys Glu Ala Pro Pro Leu Thr Leu Ser Leu Val Lys Pro Ser Gly  
 435 440 445  
 Arg Gly Ala Ala Pro Ser Leu Thr Ser Gln Ala Ala Ala Gln Val Ala  
 450 455 460  
 Asn Gln Ala Leu Ile Arg Gly Asp Trp Asn Arg Gly Gly Gly Val Arg  
 465 470 475 480  
 Val Ala Ser Val Gln Gln Val Pro Pro Arg Leu Pro Pro Glu Ala Arg  
 485 490 495  
 Pro Gly Arg Arg Met Phe Ser Thr Ser Ala Leu Gln Gly Ala Gln Gly  
 500 505 510  
 Gly Ala Arg Ala Leu Leu Gly Gly Tyr Ser Gln Ala Tyr Gly Thr Val  
 515 520 525  
 Cys His Ser Ala Leu Gly His Leu Pro Leu Leu Glu Gly His His Val  
 530 535 540

<210> 3

<211> 1998

<212> ADN

<213> *Rattus norvegicus*

<220>

<221> CDS

<222> (154)...(1794)

<400> 3

cttagtagag agagagagag agagagagag agagagacag acagacagac aagaaaagaa 60  
 aagaaaagga aaggaaaagga aagaaaagaa aaaacaaagg ggcaggacta agaaaagaat 120  
 cggtaggacaa agagccagcg tctgcctctc acc atg tgt gct gcc gag gtg gac 174  
 Met Cys Ala Ala Glu Val Asp  
 1 5

cgt cat gta tcc cag aga tac ctg atc aag cgg agg ctt ggg aag ggg 222  
 Arg His Val Ser Gln Arg Tyr Leu Ile Lys Arg Arg Leu Gly Lys Gly  
 10 15 20

# ES 2 327 903 T3

|    |                                                                 |      |
|----|-----------------------------------------------------------------|------|
|    | gcc tac ggc att gta tgg aag gcc atg gac cgg agg act ggt gag gta | 270  |
|    | Ala Tyr Gly Ile Val Trp Lys Ala Met Asp Arg Arg Thr Gly Glu Val |      |
|    | 25 30 35                                                        |      |
| 5  | gtg gcc atc aag aaa atc ttt gat gcc ttt agg gac cag aca gat gct | 318  |
|    | Val Ala Ile Lys Lys Ile Phe Asp Ala Phe Arg Asp Gln Thr Asp Ala |      |
|    | 40 45 50 55                                                     |      |
| 10 | cag agg acc ttc cgt gaa atc atg ctt ctc cgg gag ttt ggg ggc cat | 366  |
|    | Gln Arg Thr Phe Arg Glu Ile Met Leu Leu Arg Glu Phe Gly Gly His |      |
|    | 60 65 70                                                        |      |
|    | ccc aac atc atc cgc ctg ctt gat gta atc cca gca aag aat gac agg | 414  |
|    | Pro Asn Ile Ile Arg Leu Leu Asp Val Ile Pro Ala Lys Asn Asp Arg |      |
|    | 75 80 85                                                        |      |
| 15 | gat att tac ctg gtg ttt gag tcc atg gac acc gac ctg aac gcg gtc | 462  |
|    | Asp Ile Tyr Leu Val Phe Glu Ser Met Asp Thr Asp Leu Asn Ala Val |      |
|    | 90 95 100                                                       |      |
| 20 | atc cag aag ggc aga ctg ttg gag gac atc cac aaa cgt tgc atc ttt | 510  |
|    | Ile Gln Lys Gly Arg Leu Leu Glu Asp Ile His Lys Arg Cys Ile Phe |      |
|    | 105 110 115                                                     |      |
|    | tac cag ctc ctg aga gcc acc aag ttt atc cat tca ggg cgc gtc att | 558  |
|    | Tyr Gln Leu Leu Arg Ala Thr Lys Phe Ile His Ser Gly Arg Val Ile |      |
| 25 | 120 125 130 135                                                 |      |
|    | cac cgg gac cag aag cca gcc aac gtt cta ttg gat gct gct tgc cgg | 606  |
|    | His Arg Asp Gln Lys Pro Ala Asn Val Leu Leu Asp Ala Ala Cys Arg |      |
|    | 140 145 150                                                     |      |
| 30 | gtg aaa cta tgt gac ttt ggc ctg gca cgc tcc ctc agt gac ttc cct | 654  |
|    | Val Lys Leu Cys Asp Phe Gly Leu Ala Arg Ser Leu Ser Asp Phe Pro |      |
|    | 155 160 165                                                     |      |
|    | gaa ggc ctg ggc cag gcc ctg aca gaa tat gtg gcc aca cgc tgg tac | 702  |
|    | Glu Gly Leu Gly Gln Ala Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr |      |
| 35 | 170 175 180                                                     |      |
|    | cga gct cca gag gtg ctt ctg tct tcc cga tgg tat acc cct ggg gtg | 750  |
|    | Arg Ala Pro Glu Val Leu Leu Ser Ser Arg Trp Tyr Thr Pro Gly Val |      |
|    | 185 190 195                                                     |      |
| 40 | gac atg tgg agc ctg ggc tgc ata ctg gga gag atg ctt cga ggg cag | 798  |
|    | Asp Met Trp Ser Leu Gly Cys Ile Leu Gly Glu Met Leu Arg Gly Gln |      |
|    | 200 205 210 215                                                 |      |
|    | cca ctg ttt ccg ggt aca tct act ttc cac cag ctg gag ctg atc ctg | 846  |
|    | Pro Leu Phe Pro Gly Thr Ser Thr Phe His Gln Leu Glu Leu Ile Leu |      |
| 45 | 220 225 230                                                     |      |
|    | gag acc att cca ttg ccc tcc atg gag gag ctc cag ggc ctt gga tca | 894  |
|    | Glu Thr Ile Pro Leu Pro Ser Met Glu Glu Leu Gln Gly Leu Gly Ser |      |
|    | 235 240 245                                                     |      |
| 50 | gac tac agt gct ttg att ctg cag aat ctt gga tcc agg cca cgg cag | 942  |
|    | Asp Tyr Ser Ala Leu Ile Leu Gln Asn Leu Gly Ser Arg Pro Arg Gln |      |
|    | 250 255 260                                                     |      |
| 55 | acg ctg gac gcc ctc ctg ccg cca gac acc ccc cca gaa gcc ctg gac | 990  |
|    | Thr Leu Asp Ala Leu Leu Pro Pro Asp Thr Pro Pro Glu Ala Leu Asp |      |
|    | 265 270 275                                                     |      |
|    | ctc ctc aag cga ctc ttg gca ttt gct ccg gac aaa cgg ctt agt gca | 1038 |
|    | Leu Leu Lys Arg Leu Leu Ala Phe Ala Pro Asp Lys Arg Leu Ser Ala |      |
| 60 | 280 285 290 295                                                 |      |

# ES 2 327 903 T3

5 gag cag gct ctt caa cac cct tac gtg cag aga ttc cat tgc ccc gac 1086  
 Glu Gln Ala Leu Gln His Pro Tyr Val Gln Arg Phe His Cys Pro Asp  
 300 305 310

10 cgc gag tgg aca cgg ggg tcc gat gtg cgg ctc ccg gta cac gaa gga 1134  
 Arg Glu Trp Thr Arg Gly Ser Asp Val Arg Leu Pro Val His Glu Gly  
 315 320 325

15 gac cag ctc tct gca cca gag tat cgc aac cgc ctg tac cag atg atc 1182  
 Asp Gln Leu Ser Ala Pro Glu Tyr Arg Asn Arg Leu Tyr Gln Met Ile  
 330 335 340

20 ctg gag cgg aga cgg aac agc cgc agc cct cga gag gaa gac ttg ggg 1230  
 Leu Glu Arg Arg Arg Asn Ser Arg Ser Pro Arg Glu Glu Asp Leu Gly  
 345 350 355

25 gtt gtg gcc tcg cgg gca gag ctc agg gct tcc cag agg caa tcg ctc 1278  
 Val Val Ala Ser Arg Ala Glu Leu Arg Ala Ser Gln Arg Gln Ser Leu  
 360 365 370 375

30 aag ccc gga gtc ctc ccc cag gtc ttg gcg gag acg cct gcg cgg aaa 1326  
 Lys Pro Gly Val Leu Pro Gln Val Leu Ala Glu Thr Pro Ala Arg Lys  
 380 385 390

35 cgc gga ccc aaa cct cag aat ggc cat ggt cac gat ccc gag cat gtg 1374  
 Arg Gly Pro Lys Pro Gln Asn Gly His Gly His Asp Pro Glu His Val  
 395 400 405

40 gaa gtt cgc agg cag agt tca gac ccc ctg tac caa ctt ccg ccg cca 1422  
 Glu Val Arg Arg Gln Ser Ser Asp Pro Leu Tyr Gln Leu Pro Pro Pro  
 410 415 420

45 ggc agc ggg gaa agg ccc cca ggg gcc aca ggg gag cca ccc tcc gca 1470  
 Gly Ser Gly Glu Arg Pro Pro Gly Ala Thr Gly Glu Pro Pro Ser Ala  
 425 430 435

50 ccc tca ggg gtg aag act cac gtt agg gcg gtg gcg ccc tcc ctg act 1518  
 Pro Ser Gly Val Lys Thr His Val Arg Ala Val Ala Pro Ser Leu Thr  
 440 445 450 455

55 tca cag gcc gcg gct cag gcg gcc aat cag cct ctg atc cgc agt gat 1566  
 Ser Gln Ala Ala Ala Gln Ala Ala Asn Gln Pro Leu Ile Arg Ser Asp  
 460 465 470

60 ccg gcc cgg ggc ggt ggg cca agg gct gtc gcc gcg cga cgg gtc cct 1614  
 Pro Ala Arg Gly Gly Gly Pro Arg Ala Val Gly Ala Arg Arg Val Pro  
 475 480 485

65 tcc cgc ctg ccc cgg gag gcc ccg gaa ccc cga ccc ggc cga agg atg 1662  
 Ser Arg Leu Pro Arg Glu Ala Pro Glu Pro Arg Pro Gly Arg Arg Met  
 490 495 500

70 ttt ggc atc tcg gtc tcg cag ggg gcc cag ggt gca gcc aga gct gct 1710  
 Phe Gly Ile Ser Val Ser Gln Gly Ala Gln Gly Ala Ala Arg Ala Ala  
 505 510 515

75 ctt ggc gcc tac tcc cag gcc tac ggg acc gtg tgc cgc tcc gcg ctg 1758  
 Leu Gly Gly Tyr Ser Gln Ala Tyr Gly Thr Val Cys Arg Ser Ala Leu  
 520 525 530 535

80 ggc cgc ctg cct ctg ctc ccc gga ccg cgt gcg tga gccgccacc 1804  
 Gly Arg Leu Pro Leu Leu Pro Gly Pro Arg Ala \*  
 540 545

85 aacctccttg cggcaaaactg gcgcgggccc ccagacccca gcgcacatc tcccttctct 1864  
 gcctccccgc ctctgagttc caggagggtt ctctgagcct ctctcccccta agggggagta 1924  
 gatggagggtt tccactccac ccccaacaat ttattccaat aaagttttat ctttcttaaa 1984  
 aaaaaaaaaa aaaa 1998

<210> 4

65 <211> 546

<212> PRT

<213> *Rattus norvegicus*

# ES 2 327 903 T3

<400> 4

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|    | Met | Cys | Ala | Ala | Glu | Val | Asp | Arg | His | Val | Ser | Gln | Arg | Tyr | Leu | Ile |
|    | 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| 5  | Lys | Arg | Arg | Leu | Gly | Lys | Gly | Ala | Tyr | Gly | Ile | Val | Trp | Lys | Ala | Met |
|    |     |     |     | 20  |     |     |     | 25  |     |     |     |     |     | 30  |     |     |
|    | Asp | Arg | Arg | Thr | Gly | Glu | Val | Val | Ala | Ile | Lys | Lys | Ile | Phe | Asp | Ala |
|    |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| 10 | Phe | Arg | Asp | Gln | Thr | Asp | Ala | Gln | Arg | Thr | Phe | Arg | Glu | Ile | Met | Leu |
|    | 50  |     |     |     |     |     | 55  |     |     |     | 60  |     |     |     |     |     |
|    | Leu | Arg | Glu | Phe | Gly | Gly | His | Pro | Asn | Ile | Ile | Arg | Leu | Leu | Asp | Val |
|    | 65  |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |     |
|    | Ile | Pro | Ala | Lys | Asn | Asp | Arg | Asp | Ile | Tyr | Leu | Val | Phe | Glu | Ser | Met |
|    |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |     |
| 15 | Asp | Thr | Asp | Leu | Asn | Ala | Val | Ile | Gln | Lys | Gly | Arg | Leu | Leu | Glu | Asp |
|    |     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
|    | Ile | His | Lys | Arg | Cys | Ile | Phe | Tyr | Gln | Leu | Leu | Arg | Ala | Thr | Lys | Phe |
|    |     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
|    | Ile | His | Ser | Gly | Arg | Val | Ile | His | Arg | Asp | Gln | Lys | Pro | Ala | Asn | Val |
| 20 |     | 130 |     |     |     |     | 135 |     |     |     | 140 |     |     |     |     |     |
|    | Leu | Leu | Asp | Ala | Ala | Cys | Arg | Val | Lys | Leu | Cys | Asp | Phe | Gly | Leu | Ala |
|    | 145 |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |     |
|    | Arg | Ser | Leu | Ser | Asp | Phe | Pro | Glu | Gly | Leu | Gly | Gln | Ala | Leu | Thr | Glu |
|    |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |
| 25 | Tyr | Val | Ala | Thr | Arg | Trp | Tyr | Arg | Ala | Pro | Glu | Val | Leu | Leu | Ser | Ser |
|    |     |     |     | 180 |     |     |     | 185 |     |     |     |     | 190 |     |     |     |
|    | Arg | Trp | Tyr | Thr | Pro | Gly | Val | Asp | Met | Trp | Ser | Leu | Gly | Cys | Ile | Leu |
|    |     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
|    | Gly | Glu | Met | Leu | Arg | Gly | Gln | Pro | Leu | Phe | Pro | Gly | Thr | Ser | Thr | Phe |
| 30 |     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
|    | His | Gln | Leu | Glu | Leu | Ile | Leu | Glu | Thr | Ile | Pro | Leu | Pro | Ser | Met | Glu |
|    | 225 |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |     |
|    | Glu | Leu | Gln | Gly | Leu | Gly | Ser | Asp | Tyr | Ser | Ala | Leu | Ile | Leu | Gln | Asn |
|    |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |     |
| 35 | Leu | Gly | Ser | Arg | Pro | Arg | Gln | Thr | Leu | Asp | Ala | Leu | Leu | Pro | Pro | Asp |
|    |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |     |
|    | Thr | Pro | Pro | Glu | Ala | Leu | Asp | Leu | Leu | Lys | Arg | Leu | Leu | Ala | Phe | Ala |
|    |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
|    | Pro | Asp | Lys | Arg | Leu | Ser | Ala | Glu | Gln | Ala | Leu | Gln | His | Pro | Tyr | Val |
| 40 |     | 290 |     |     |     |     | 295 |     |     |     | 300 |     |     |     |     |     |
|    | Gln | Arg | Phe | His | Cys | Pro | Asp | Arg | Glu | Trp | Thr | Arg | Gly | Ser | Asp | Val |
|    | 305 |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |     |
|    | Arg | Leu | Pro | Val | His | Glu | Gly | Asp | Gln | Leu | Ser | Ala | Pro | Glu | Tyr | Arg |
|    |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |     |
| 45 | Asn | Arg | Leu | Tyr | Gln | Met | Ile | Leu | Glu | Arg | Arg | Arg | Asn | Ser | Arg | Ser |
|    |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |     |
|    | Pro | Arg | Glu | Glu | Asp | Leu | Gly | Val | Val | Ala | Ser | Arg | Ala | Glu | Leu | Arg |
|    |     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
|    | Ala | Ser | Gln | Arg | Gln | Ser | Leu | Lys | Pro | Gly | Val | Leu | Pro | Gln | Val | Leu |
| 50 |     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
|    | Ala | Glu | Thr | Pro | Ala | Arg | Lys | Arg | Gly | Pro | Lys | Pro | Gln | Asn | Gly | His |
|    | 385 |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |     |
|    | Gly | His | Asp | Pro | Glu | His | Val | Glu | Val | Arg | Arg | Gln | Ser | Ser | Asp | Pro |
|    |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |     |
| 55 | Leu | Tyr | Gln | Leu | Pro | Pro | Pro | Gly | Ser | Gly | Glu | Arg | Pro | Pro | Gly | Ala |
|    |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |     |
|    | Thr | Gly | Glu | Pro | Pro | Ser | Ala | Pro | Ser | Gly | Val | Lys | Thr | His | Val | Arg |
|    |     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
|    | Ala | Val | Ala | Pro | Ser | Leu | Thr | Ser | Gln | Ala | Ala | Ala | Gln | Ala | Ala | Asn |
| 60 |     |     | 450 |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
|    | Gln | Pro | Leu | Ile | Arg | Ser | Asp | Pro | Ala | Arg | Gly | Gly | Gly | Pro | Arg | Ala |
|    | 465 |     |     |     | 470 |     |     |     |     | 475 |     |     |     | 480 |     |     |
|    | Val | Gly | Ala | Arg | Arg | Val | Pro | Ser | Arg | Leu | Pro | Arg | Glu | Ala | Pro | Glu |

# ES 2 327 903 T3

485                      490                      495  
 Pro Arg Pro Gly Arg Arg Met Phe Gly Ile Ser Val Ser Gln Gly Ala  
                                  500                      505                      510  
 5 Gln Gly Ala Ala Arg Ala Ala Leu Gly Gly Tyr Ser Gln Ala Tyr Gly  
                                  515                      520                      525  
 Thr Val Cys Arg Ser Ala Leu Gly Arg Leu Pro Leu Leu Pro Gly Pro  
                                  530                      535                      540  
 10 Arg Ala  
                                  545

<210> 5

<211> 15

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador oligonucleotídico

<400> 5

accgcagccg cgtct 15

<210> 6

<211> 19

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador oligonucleotídico

<400> 6

tctctctcga ggtgccgct 19

<210> 7

<211> 26

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador oligonucleotídico

<400> 7

tcagatgac ctggagtgtg gaggca 26

<210> 8

<211> 21

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador oligonucleotídico

<400> 8

cacccccact gaaaaagatg a 21

<210> 9

<211> 26

<212> ADN

## ES 2 327 903 T3

<213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleotídico  
 5 <400> 9  
           cttaactatc ttgggctgtg acaaag           26  
 10 <210> 10  
       <211> 24  
       <212> ADN  
 15 <213> Secuencia artificial  
       <220>  
       <223> Cebador oligonucleotídico  
 20 <400> 10  
           tatgcctgcc gtgtgaacca cgtg           24  
       <210> 11  
 25 <211> 17  
       <212> PRT  
       <213> Secuencia artificial  
       <220>  
 30 <223> Secuencia de consenso implicada en la unión a ATP  
       <221> VARIANTE  
       <222> (1)...(1)  
 35 <223> L puede ser también I o V  
       <221> VARIANTE  
       <222> (3)...(3)  
       <223> Xaa = cualquier aminoácido distinto de P  
 40 <221> VARIANTE  
       <222> (5)...(5)  
       <223> Xaa = cualquier aminoácido distinto de P  
       <221> VARIANTE  
 45 <222> (6)...(6)  
       <223> F puede ser también Y, W, M, G, S, T, N o H  
       <221> VARIANTE  
 50 <222> (7)...(7)  
       <223> S puede ser también G o A  
       <221> VARIANTE  
       <222> (8)...(8)  
 55 <223> Xaa = cualquier aminoácido distinto de P y W  
       <221> VARIANTE  
       <222> (9)...(9)  
 60 <223> L puede ser también I, V, C, A o T  
       <221> VARIANTE  
       <222> (10)...(10)  
       <223> Xaa = cualquier aminoácido distinto de P y D  
 65 <221> VARIANTE  
       <222> (11)...(11)

## ES 2 327 903 T3

- <223> Xaa = cualquier aminoácido  
 <221> VARIANTE  
 <222> (12)...(12)
- 5 <223> G puede ser también S, T, A, C, L, I, V, M, F o Y  
 <221> VARIANTE  
 <222> (13)...(13)
- 10 <223> Xaa = los siguientes 5-18 aminoácidos pueden ser cualquier aminoácido  
 <221> VARIANTE  
 <222> (14)...(14)
- <223> L puede ser también I, V, M, F, Y, W, C, S, T, A o R
- 15 <221> VARIANTE  
 <222> (15)...(15)
- <223> A puede ser también I, V o P
- 20 <221> VARIANTE  
 <222> (16)...(16)
- <223> L puede ser también I, V, M, F, A, G, C, K o R
- 25 <400> 11
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gly | Xaa | Gly | Xaa | Phe | Ser | Xaa | Leu | Xaa | Xaa | Gly | Xaa | Leu | Ala | Leu |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
- Lys**
- 30 <210> 12  
 <211> 13  
 <212> PRT
- 35 <213> Secuencia artificial  
 <220>  
 <223> Secuencia de consenso para serina/treonina cinasas  
 <221> VARIANTE
- 40 <222> (1)...(1)  
 <223> La L en posición 1 puede ser I, V, M, F, Y o C.  
 <221> VARIANTE
- 45 <222> (2)...(2)  
 <223> La X en posición 2 puede ser cualquier aminoácido  
 <221> VARIANTE  
 <222> (3)...(3)
- 50 <223> La H en posición 3 puede ser Y.  
 <221> VARIANTE  
 <222> (6)...(6)
- 55 <223> La L en posición 6 puede ser I, V, M, F o Y.  
 <221> VARIANTE  
 <222> (8)...(8)
- <223> La X en posición 8 puede ser cualquier aminoácido.
- 60 <221> VARIANTE  
 <222> (9)...(9)
- <223> La X en posición 9 puede ser cualquier aminoácido  
 <221> VARIANTE
- 65 <222> (11)...(11)  
 <223> La L en posición 11 puede ser I, V, M, F, Y, C o T.

## ES 2 327 903 T3

- <221> VARIANTE  
 <222> (12)...(12)  
 <223> La L en posición 12 puede ser I, V, M, F, Y, C o T.
- 5 <221> VARIANTE  
 <222> (13)...(13)  
 <223> La L en posición 13 puede ser I, V, M, F, Y, C o T.
- 10 <221> VARIANTE  
 <222> (4)...(4)  
 <223> La X en posición 4 puede ser cualquier aminoácido
- 15 <400> 12  

|     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Xaa | His | Xaa | Asp | Leu | Lys | Xaa | Xaa | Asn | Leu | Leu | Leu |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |
- <210> 13  
 20 <211> 13  
 <212> PRT  
 <213> Secuencia artificial
- 25 <220>  
 <223> Secuencia de consenso para tirosina cinasas  
 <221> VARIANTE  
 <222> (1)...(1)
- 30 <223> La L en posición 1 puede ser también I, V, M, F, Y o C.  
 <221> VARIANTE  
 <222> (2)...(2)
- 35 <223> La X en posición 2 puede ser cualquier aminoácido  
 <221> VARIANTE  
 <222> (3)...(3)  
 <223> La H en posición 3 puede ser Y.
- 40 <221> VARIANTE  
 <222> (6)...(6)  
 <223> La L en posición 6 puede ser I, V, M, F o Y.
- 45 <221> VARIANTE  
 <222> (7)...(7)  
 <223> La R en posición 7 puede ser S, T, A o C.
- 50 <221> VARIANTE  
 <222> (8)...(8)  
 <223> La X en posición 8 puede ser cualquier aminoácido
- 55 <221> VARIANTE  
 <222> (9)...(9)  
 <223> La X en posición 9 puede ser cualquier aminoácido
- 60 <221> VARIANTE  
 <222> (11)...(11)  
 <223> La L en posición 11 puede ser I, V, M, F, Y o C.
- 65 <221> VARIANTE  
 <222> (12)...(12)  
 <223> La L en posición 12 puede ser I, V, M, F, Y o C.
- <221> VARIANTE  
 <222> (13)...(13)

## ES 2 327 903 T3

<223> La L en posición 13 puede ser I, V, M, F, Y o C.

<221> VARIANTE

<222> (4)...(4)

5 <223> La X en posición 4 puede ser cualquier aminoácido

<400> 13

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 10 | Leu | Xaa | His | Xaa | Asp | Leu | Arg | Xaa | Xaa | Asn | Leu | Leu | Leu |
|    | 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |

15

20

25

30

35

40

45

50

55

60

65