

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 637 592**

51 Int. Cl.:

C12N 9/12 (2006.01)
C12N 15/62 (2006.01)
C07K 14/47 (2006.01)
A61K 38/00 (2006.01)
A61K 31/713 (2006.01)
C12Q 1/68 (2006.01)
G01N 33/574 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.04.2007** **E 12150524 (2)**

97 Fecha y número de publicación de la concesión europea: **17.05.2017** **EP 2450437**

54 Título: **Defectos de genes y quinasa ALK mutante en tumores sólidos humanos**

30 Prioridad:

14.04.2006 US 792364 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
13.10.2017

73 Titular/es:

CELL SIGNALING TECHNOLOGY, INC. (100.0%)
3 Trask Lane
Danvers, MA 01923, US

72 Inventor/es:

RIKOVA, KLARISA;
HAACK, HERBERT;
GU, TING-LEI;
POSSEMATO, ANTHONY;
GUO, AILAN;
YU, JIAN;
SULLIVAN, LAURA y
MACNEIL, JOAN

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 637 592 T3

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

DESCRIPCIÓN

Defectos de genes y quinasa ALK mutante en tumores sólidos humanos

Campo de la invención

La invención se refiere en general a proteínas y genes implicados en el cáncer, y al tratamiento del cáncer.

5 Antecedentes de la invención

Muchos cánceres se caracterizan por alteraciones en las vías de señalización celular que conducen a un control aberrante de los procesos celulares, o al crecimiento y proliferación incontrolados de las células. Estas alteraciones son causadas a menudo por los cambios en la actividad de proteínas de señalización concretas, tales como las quinastas. Entre estos cánceres se encuentran tumores sólidos, como carcinoma de pulmón de células no pequeñas (CPCNP). El CPCNP es la principal causa de muerte por cáncer en los Estados Unidos, y representa aproximadamente 87% de todos los cánceres de pulmón. Existen aproximadamente 151.000 nuevos casos de CPCNP en los Estados Unidos cada año, y se estima que más de 120.000 pacientes morirán anualmente a causa de esta enfermedad solo en los Estados Unidos. Véase "Cancer Facts and Figures 2005", American Cancer Society. El CPCNP, que comprende tres subtipos distintos, a menudo solo se detecta después de que haya metastatizado, y por lo tanto la tasa de mortalidad es de 75% en los dos años siguientes al diagnóstico.

Se sabe que las deleciones y/o translocaciones génicas que dan como resultado proteínas de fusión de quinastas con actividad de señalización aberrante pueden conducir directamente a ciertos cánceres. Por ejemplo, se ha demostrado directamente que la oncoproteína BCR-ABL, una proteína de fusión de tirosina quinasa, es el agente causal en la leucemia mielógena crónica humana (LMC). La oncoproteína BCR-ABL, que se encuentra en al menos 90-95% de los casos de LMC, es generada por la translocación de secuencias génicas de la proteína tirosina quinasa c-Abl en el cromosoma 9 a las secuencias de BCR en el cromosoma 22, produciendo el denominado cromosoma Filadelfia. Véase, p. ej., Kurzrock et al., *N. Engl. J. Med.* 319: 990-998 (1988). La translocación se observa también en casos de leucemia linfocítica aguda y CPCNP.

Se han descrito translocaciones y deleciones génicas que conducen a proteínas mutantes o de fusión implicadas en una variedad de otros cánceres. Por ejemplo, Falini et al., *Blood* 99(2): 409-426 (2002), revisan las translocaciones que se sabe que se producen en los cánceres hematológicos, incluyendo la fusión NPM-ALK encontrada en LACG. Hasta la fecha, se han descrito solo un número limitado de translocaciones, deleciones de genes, y proteínas mutantes que aparecen en los cánceres de pulmón, incluyendo la translocación t(15;19) que implica Notch3. Véase Dang et al., *J. Natl. Can. Instit.* 92(16): 1355-1357 (2000). Se han encontrado defectos en la expresión y/o actividad de la Proteína 6 de Unión a ARN (EML-4) en carcinomas de pulmón de células pequeñas y de células no pequeñas. Véase. Drabkin et al., *Oncogene* 8(16): 2589-97 (1999). Sin embargo, hasta la fecha, no se han descrito translocaciones o deleciones en el cáncer CPCNP humano que impliquen proteínas quinastas.

Se han descrito defectos en la expresión de la quinasa ALK resultante de la fusión de NPM a ALK en el linfoma anaplásico de células. Véanse Morris et al., 1994; Shiota et al., 1994. Se han descrito la fusión de ALK a moesina, a la cadena pesada 9 de la miosina no muscular (Tort et al. 2001), a la cadena pesada de clartrina (Touriol et al., 2000; Bridge et al., 2001), a tropomiosina 3 (TPM3) (Lamant et al., 1999), al gen fusionado a TRK (TGF) (Hernández et al., *Am. J. Path.* 160(4): 1487-1493 (2002)) y a otros genes. En particular, se informó sobre la fusión de TGF-ALK en el linfoma no sólido, pero hasta la fecha esta fusión no se ha descrito en los tumores sólidos. Se ha descrito el papel general de ALK en el cáncer. Véase Pulford et al., *J. Cell Physiol.* 199(3): 330-358 (2004). Sin embargo, hasta la fecha, no se han descrito defectos en la expresión y/o activación de EML-4.

La identificación de mutaciones en cánceres humanos es altamente deseable, ya que puede conducir al desarrollo de nuevos agentes terapéuticos que se dirigen a tales proteínas de fusión o mutantes, y a nuevos diagnósticos para la identificación de pacientes que tienen tales mutaciones genéticas. Por ejemplo, BCR-ABL se ha convertido en una diana para el desarrollo de agentes terapéuticos para el tratamiento de la leucemia. Más recientemente, Gleevec® (mesilato de imatinib, STI-571), un inhibidor de molécula pequeña de la quinasa ABL, ha sido aprobado para el tratamiento de la LMC. Este fármaco es el primero de una nueva clase de agentes anti-proliferativos diseñados para interferir en las rutas de señalización que dirigen el crecimiento de células tumorales. El desarrollo de este fármaco representa un avance significativo sobre las terapias convencionales para la LMC y la LLA, la quimioterapia y la radiación, que se ven afectadas por los efectos secundarios bien conocidos y tienen a menudo un efecto limitado, ya que no pueden dirigirse específicamente a las causas subyacentes de las enfermedades malignas. Del mismo modo, se han descrito los reactivos y métodos para detectar específicamente la proteína de fusión BCR-ABL en pacientes, con el fin de identificar a los pacientes con más probabilidades de responder a inhibidores específicos, como Gleevec®.

Por consiguiente, sigue existiendo la necesidad de la identificación de nuevas mutaciones de genes, tales como translocaciones o deleciones, que dan como resultado proteínas de fusión o mutantes implicadas en la progresión de cánceres humanos, concretamente tumores sólidos, incluyendo cánceres de pulmón como el CPCNP, y el desarrollo de nuevos reactivos y métodos para el estudio y la detección de tales proteínas de fusión. La identificación de tales proteínas de fusión permitirá deseablemente, entre otras cosas, nuevos métodos de selección

de pacientes para terapias dirigidas, así como para el escrutinio de nuevos fármacos que inhiban tales proteínas mutantes/de fusión.

Compendio de la invención

La invención se reivindica en las reivindicaciones 1-6. En esta memoria se describen mutaciones novedosas por delección génica que se producen en el cromosoma humano 2 que dan como resultado proteínas de fusión que combinan parte de la Quinasa del Linfoma Anaplásico (ALK) con una proteína secundaria que se han identificado ahora en el tumor sólido de carcinoma de pulmón de células no pequeñas (CPCNP) humano. Las proteínas secundarias que participan en las fusiones de ALK incluyen Proteína de Tipo 4 Asociada a Microtúbulos de Equinodermo (EML-4) y Gen Fusionado a TRK (TFG). Las quinasas ALK mutantes/de fusión han sido observadas en la actualidad en muestras de pacientes con carcinoma de pulmón de células no pequeñas.

Por consiguiente, la descripción proporciona, en parte, polinucleótidos aislados y vectores que codifican los polipéptidos de ALK mutantes/de fusión descritos, sondas y análisis para detectarlos, polipéptidos de ALK mutante/de fusión aislados, polipéptidos mutantes recombinantes, y reactivos para la detección de los polinucleótidos y polipéptidos de ALK mutantes. La identificación descrita de estas nuevas quinasas ALK mutantes y de las translocaciones/delecciones permite nuevos métodos para determinar la presencia de polinucleótidos o polipéptidos de ALK mutantes en una muestra biológica, y métodos para el escrutinio de compuestos que inhiben la proteína quinasa mutante. También describe métodos para inhibir la progresión de un cáncer caracterizado por la expresión de polipéptidos de ALK mutantes, que son proporcionados por la invención. Los aspectos y realizaciones de la invención se describen con más detalle a continuación.

Breve descripción de los dibujos

Fig. 1A - muestra las localizaciones del gen de EML-4 y del gen de ALK sobre el cromosoma 2 (panel A), y las localizaciones de los dominios de las proteínas EML-4 y ALK completas, así como las de la proteína de fusión EML4-ALK (variante corta) (panel B); la unión por fusión se produce en los aminoácidos 233-234, y la proteína de fusión incluye el dominio quinasa (pero no los dominios transmembrana y extracelulares) de ALK. También se muestra (en el panel B) la secuencia de ADN (y de la proteína) de la región de unión por fusión del exón 6/intron 6 de EML4/exón 20 de ALK (SEQ ID NO: 7 y SEQ ID NO: 8, respectivamente).

Fig. 1B - muestra las localizaciones del gen de EML-4 y el gen de ALK sobre el cromosoma 2 (panel A), y las localizaciones de los dominios de las proteínas EML-4 y ALK completas, así como las de la proteína de fusión EML4-ALK (variante larga) (panel B); la unión por fusión se produce en los aminoácidos 495-496, y la proteína de fusión incluye el dominio quinasa (pero no los dominios transmembrana y extracelulares) de ALK. También se muestra (en el panel B) la secuencia de ADN (y de la proteína) de la región de unión por fusión del exón 13 de EML4/exón 20 de ALK (SEQ ID NO: 24 y SEQ ID NO: 25, respectivamente).

Fig. 1C - muestra las localizaciones del gen de TFG sobre el cromosoma 6 y del gen ALK sobre el cromosoma 2 (panel A), y las localizaciones de los dominios de las proteínas TFG y ALK completas, así como las de la proteína de fusión de TFG-ALK (panel B); la unión por fusión se produce en los aminoácidos 138-139, y la proteína de fusión incluye el dominio quinasa (pero no los dominios transmembrana y extracelulares) de ALK. También se muestra (en el panel B) la secuencia el ADN (y de la proteína) de la región de unión por fusión del exón 3 de TFG/exón 20 de ALK (SEQ ID NO: 26 y SEQ ID NO: 27, respectivamente).

Fig. 2A - es la secuencia de aminoácidos (código de 1 letra) de la proteína de fusión EML4-ALK humana (variante corta) (SEQ ID NO: 1) (panel superior) indicándose también la secuencia de ADN codificante (SEQ ID NO: 2) (panel inferior); los residuos del radical EML-4 están en cursiva, mientras que los residuos del dominio quinasa ALK están en negrita.

Fig. 2B - es la secuencia de aminoácidos (código de 1 letra) de la proteína de fusión EML4-ALK humana (variante larga) (SEQ ID NO: 18) (panel superior) indicándose también la secuencia de ADN codificante (SEQ ID NO: 19) (panel inferior); los residuos del radical EML-4 están en cursiva, mientras que los residuos del dominio quinasa ALK están en negrita.

Fig. 2C - es la secuencia de aminoácidos (código de 1 letra) de la proteína de fusión de TFG-ALK humana (SEQ ID NO: 20) (panel superior) indicándose también la secuencia de ADN codificante (SEQ ID NO: 21) (panel inferior); los residuos del radical TFG están en cursiva, mientras que los residuos del dominio quinasa ALK están en negrita.

Fig. 3A-3B - es la secuencia de aminoácidos (código de 1 letra) de la proteína EML-4 humana (SEQ ID NO: 3) (Núm. de Acceso SwissProt 061936) indicándose también la secuencia de ADN codificante (SEQ ID NO: 4) (Núm. de Acceso GenBank NM019063); los residuos conservados en el mutante por delección de la variante corta están subrayados, mientras que los residuos conservados en la variante larga están en cursiva.

Fig. 4A-4B - es la secuencia de aminoácidos (código de 1 letra) de la quinasa ALK humana (SEQ ID NO: 5) (Núm. de Acceso SwissProt Q9UM73) indicándose también la secuencia de ADN codificante (SEQ ID NO: 6) (Núm. de Acceso GenBank HSU66559); los residuos conservados en los mutantes por delección están subrayados, mientras

que los residuos del dominio quinasa están en negrita.

Fig. 4C-4D - es la secuencia de aminoácidos (código de 1 letra) de la proteína TFG humana (SEQ ID NO: 22) (Núm. de Acceso SwissProt Q92734) indicándose también la secuencia de ADN codificante (SEQ ID NO: 23) (Núm. de Acceso GenBank NM006070); los residuos conservados en el mutante por delección están subrayados.

Fig. 5 - son geles que representan (A) la detección de ALK a través del producto 5' RACE con cebadores ALK después de 2 rondas de PCR; UAP significa Cebador de Amplificación Universal, GSP significa Cebador Específico del Gen, (B) la detección del gen de fusión formado por el mutante por delección de EML-4 y ALK mediante RT-PCR, (C) la detección del gen de fusión EML4-ALK (variantes corta y larga) en muestras de tumor CPCNP humano por medio de 5' RACE, y (D) la detección del gen de fusión de TFG-ALK en muestras de tumor CPCNP humano por medio de 5' RACE.

Fig. 6 - es una imagen que representa la detección del gen de fusión formado por la translocación de EML-4 y ALK en células H2228 mediante el análisis FISH empleando una sonda de rotura de doble color (naranja/verde) que comprende sondas para los lados opuestos del punto de rotura 2p23 del gen ALK; los tamaños y localizaciones de las sondas se muestran en el panel superior.

Descripción detallada de la invención

De acuerdo con la invención, se han identificado ahora delecciones y translocaciones de genes previamente desconocidas que dan como resultado proteínas de fusión de quinasa mutantes, que combinan parte de la Quinasa del Linfoma Anaplásico (ALK) con una porción de una proteína secundaria, en el tumor sólido de carcinoma de pulmón de células no pequeñas (CPCNP) humano. Las proteínas secundarias que participan en las fusiones de ALK descubierta incluyen la Proteína de Tipo 4 Asociada con Microtúbulos de Equinodermo (EML-4) y el Gen Fusionado a TRK (TFG).

Los dos delecciones descritas, que se produce entre los genes EML4 y ALK en el cromosoma 2, producen proteínas de fusión que combinan el extremo N de EML-4, una proteína de unión a microtúbulos de 401 aminoácidos, con el dominio quinasa y el extremo C de ALK, una tirosina quinasa de membrana de 1.620 aminoácidos. Se espera que las proteínas de fusión EML4-ALK resultantes, que tienen 796 aminoácidos (variante corta) y 1059 aminoácidos (variante larga), respectivamente, y conservan la actividad de la quinasa ALK, dirijan la proliferación y la supervivencia de un subgrupo de tumores sólidos humanos, incluyendo CPCNP.

La translocación descrita, que se produce entre el gen de TFG sobre el cromosoma 6 y el gen ALK sobre el cromosoma 2, produce una proteína de fusión que combina el extremo N de TFG, una proteína de 400 aminoácidos, con el dominio quinasa y el extremo C de ALK, una tirosina quinasa de membrana de 1620 aminoácidos. La proteína de fusión de TFG-ALK resultante, que tiene 701 aminoácidos, ha sido observada previamente en el linfoma humano no sólido (Hernández *et al.* (2002), más arriba), pero no ha sido previamente descrita en tumores sólidos. La proteína de fusión de TFG-ALK conserva la actividad de la quinasa ALK, y se espera que dirija la proliferación y la supervivencia de un subgrupo de tumores sólidos humanos, incluyendo CPCNP.

Aunque se han descrito unas pocas translocaciones o delecciones de genes que dan como resultado proteínas de fusión aberrantes en CPCNP, incluyendo la translocación t(15;19) que implica Notch3 (véase Dang *et al.*, más arriba), los mutantes por delección y las proteínas de fusión de EML4-ALK descritas en la actualidad son novedosos. Del un modo similar, el mutante por translocación y la proteína de fusión de TFG-ALK, aunque no conocidos en tumores no sólidos como el linfoma, son novedosos en el tumor sólido CPCNP. EML-4 es una proteína asociada a los microtúbulos que se expresa en la mayoría de los tejidos humanos. Hasta la fecha, no se ha informado de defectos en la expresión y/o actividad de EML-4. ALK es una tirosina quinasa de membrana, y es expresada, en seres humanos, en cerebro y tejidos del SNC, también en intestino delgado y testículos, pero no en células linfoides normales. Desempeña un papel importante en el desarrollo y la función normales del sistema nervioso (Iwahara *et al.*, 1997).

Se han encontrado defectos en la expresión y/o activación de ALK en linfoma anaplásico de células grandes y neuroblastoma (véanse Morris *et al.*, 1994, Osajima-Hakomori *et al.*, 2005). Se ha descrito la fusión de ALK a moesina, a la cadena pesada 9 de miosina no muscular, a la cadena pesada de clatrina, a tropomiosina 3 (TPM3), al gen fusionado a TRK (TFG), y a otros genes. Véanse Tort *et al.*; Touriol *et al.*, Hernández *et al.*, más arriba). Curiosamente, la fusión descrita de EML-4 a ALK (variante corta) se produce precisamente en el mismo punto en la ALK de tipo salvaje (aminoácido 1058) que se ha descrito anteriormente para otros mutantes de fusión de ALK.

Como se describe adicionalmente más adelante, los mutantes por delección de EML4-ALK y las proteínas de fusión expresadas se han aislado y secuenciado en la actualidad, y se han producido los ADNc para la expresión de las proteínas de fusión. Por consiguiente, la descripción proporciona, en parte, polinucleótidos aislados que codifican polipéptidos de fusión de EML4-ALK, sondas de ácidos nucleicos que hibridan con tales polinucleótidos, y métodos, vectores y células anfitrionas para la utilización de tales polinucleótidos para producir polipéptidos de ALK mutantes recombinantes. La descripción también proporciona, en parte, polipéptidos aislados que comprenden secuencias de aminoácidos que codifican polipéptidos de fusión de EML4-ALK, polipéptidos mutantes recombinantes, y reactivos aislados que se unen específicamente a y/o detectan polipéptidos de fusión de EML4-ALK, pero no se unen a ni

detectan EML-4 de tipo salvaje o ALK de tipo salvaje. Estos aspectos de la descripción, que se describen con más detalle a continuación, serán útiles, entre otras cosas, en el estudio adicional de los mecanismos de los cánceres dirigidos por la expresión/actividad de la quinasa ALK mutante, para la identificación de tumores sólidos (por ejemplo, carcinomas incluyendo carcinomas de pulmón y sarcomas) y otros cánceres caracterizados por las mutaciones por delección y translocación y/o proteínas de fusión de ALK descritas, o por la expresión/actividad de la quinasa ALK mutante, y en la práctica de los métodos de la descripción descritos adicionalmente a continuación.

La identificación de los mutantes de la quinasa ALK novedosos y las mutaciones por delección y translocación del gen tiene implicaciones importantes para el diagnóstico y el tratamiento potenciales de tumores sólidos, tales como CPCNP, que se caracterizan por una o más de estas proteínas de fusión. El CPCNP, por ejemplo, a menudo solo se detecta después de que haya metastatizado, y de este modo la tasa de mortalidad es de 75% en los dos años siguientes al diagnóstico. En consecuencia, sería muy deseable la capacidad para identificar, tan temprano como sea posible, los pacientes que tienen mutaciones génicas que pueden conducir a CPCNP.

Por lo tanto, el descubrimiento de las proteínas de fusión de EML4-ALK (variantes cortas y largas) que resultan de la delección génica y la proteína de fusión de TGF-ALK que resulta de la translocación génica, que se espera que dirijan la proliferación y la supervivencia de un tumor sólido, CPCNP, permite importantes métodos nuevos para identificar con precisión tumores sólidos de mamíferos, incluyendo cánceres de pulmón (tales como CPCNP), así como otros tipos de cáncer, en los que se expresa una proteína de fusión de ALK (tal como EML4-ALK o TFG-ALK). Es muy probable que estos tumores respondan a inhibidores de la actividad quinasa de la proteína ALK mutante, tales como WHI-131 o WHI-154. La capacidad para identificar, tan pronto como sea posible, cánceres que están dirigidos por una quinasa ALK mutante será de gran ayuda en la determinación clínica de qué agente terapéutico o combinación de agentes terapéuticos, será la más adecuada para un paciente concreto, contribuyendo de ese modo a evitar la prescripción de inhibidores dirigidos otras quinasas que no son, de hecho, la molécula de señalización principal que dirige el cáncer.

Por consiguiente, la descripción proporciona, en parte, métodos para detectar la presencia de un polinucleótido mutante de ALK y/o polipéptido de fusión en un cáncer utilizando reactivos específicos para la fusión y específicos para el mutante de la descripción. Tales métodos pueden ponerse en práctica, por ejemplo, para identificar un tumor sólido, tal como CPCNP, que es probable que responda a un inhibidor de la actividad de la quinasa ALK de la proteína mutante. La descripción también proporciona, en parte, métodos para determinar si un compuesto inhibe la progresión de un cáncer caracterizado por un polipéptido de fusión de EML4-ALK. La invención proporciona adicionalmente un compuesto para usar en la inhibición de la progresión de un tumor sólido que expresa un polipéptido de fusión de EML4-ALK mediante la inhibición de la expresión y/o actividad del polipéptido mutante. Tales métodos y compuestos se describen con más detalle a continuación.

Los aspectos, ventajas y realizaciones adicionales de la invención se describen con más detalle a continuación.

Definiciones.

Según se utiliza en la presente memoria, los siguientes términos tienen los significados indicados.

"Anticuerpo" o "anticuerpos" se refiere a todos los tipos de inmunoglobulinas, incluyendo IgG, IgM, IgA, IgD, e IgE, incluyendo F_{ab} o fragmentos de reconocimiento del antígeno de los mismos, incluyendo anticuerpos quiméricos, policlonales, y monoclonales. El término "anticuerpo humanizado", según se utiliza en la presente memoria, se refiere a moléculas de anticuerpo en las que los aminoácidos han sido sustituidos en las regiones que no se unen al antígeno con el fin de parecerse más estrechamente a un anticuerpo humano, a la vez que todavía conservan la capacidad de unión original.

El término "biológicamente activo" se refiere a una proteína que tiene las funciones estructurales, reguladoras o bioquímicas de una molécula de origen natural. Del mismo modo, "inmunológicamente activo" se refiere a la capacidad del polipéptido de fusión de EML4-ALK o TFG-ALK natural, recombinante, o sintético, o cualquier oligopéptido del mismo, para inducir una respuesta inmunitaria específica en animales o células apropiados y para unirse con anticuerpos específicos.

El término "muestra biológica" se utiliza en su sentido más amplio, y representa cualquier muestra biológica que se sospecha que contiene polinucleótidos o polipéptidos de fusión de ALK o fragmentos de los mismos (incluyendo polinucleótidos y polipéptidos de fusión de EML4-ALK y TFG-ALK), y puede comprender una célula, los cromosomas aislados de una célula (p. ej., una dispersión de cromosomas en metafase), ADN genómico (en disolución o unido a un soporte sólido, por ejemplo para el análisis Southern), ARN (en disolución o unido a un soporte sólido, por ejemplo para análisis northern), ADNc (en disolución o unido a un soporte sólido), un extracto de células, sangre, orina, médula ósea, o un tejido, y similares.

"Caracterizado por" con respecto a un cáncer y un polinucleótido y polipéptido de ALK mutante significa un cáncer en el que están presentes una delección o translocación génica y/o un polipéptido de fusión expresado que implica ALK en comparación con un cáncer en el que no están presentes semejante delección génica y/o polipéptido de fusión. La presencia de un polipéptido mutante puede dirigir, en su totalidad o en parte, el crecimiento y la supervivencia de tal cáncer.

- "Consenso" se refiere a una secuencia de ácido nucleico que ha sido re-secuenciada para resolver bases no exigidas, o que ha sido ampliada utilizando XL-PCR™ (Perkin Elmer, Norwalk, Conn.) en la dirección 5' y/o 3' y re-secuenciada, o que ha sido ensamblada a partir de secuencias solapantes de más de un clon Incyte utilizando el sistema de ensamblaje de fragmentos GELVIEW™ (GCG, Madison, Wis.), o que ha sido tanto ampliada como ensamblada.
- 5 "Agente terapéutico inhibidor de quinasa ALK" significa cualquier composición que comprende uno o más compuestos, químicos o biológicos, que inhiben, ya sea directa o indirectamente, la expresión y/o actividad de quinasa ALK tipo salvaje o truncada, ya sea sola y/o como parte de una proteína de fusión (tal como, proteínas de fusión de EML4-ALK y proteínas de fusión de TGF-ALK).
- 10 "Derivado" se refiere a la modificación química de una secuencia de ácido nucleico que codifica un polinucleótido de fusión descrito o el propio polipéptido codificado. Sería ilustrativa de tales modificaciones la sustitución de hidrógeno por un grupo alquilo, acilo, o amino. Un derivado de ácido nucleico codificaría un polipéptido que conservaría las características biológicas esenciales de la molécula natural.
- 15 "Marca detectable" con respecto a un polipéptido, polinucleótido, o reactivo descrito en la presente memoria significa una modificación química, biológica, u otra, incluyendo pero no limitada a modificaciones de fluorescencia, masa, residuo, colorante, radioisótopo, marca, o etiqueta, etc., por medio de la cual se puede detectar la presencia de la molécula de interés.
- "Expresión" o "expresado" con respecto a un polipéptido de fusión de ALK en una muestra biológica significa expresado significativamente en comparación con una muestra de control en la que este polipéptido de fusión no es expresado de manera significativa.
- 20 "Péptido marcado con isótopo pesado" (utilizado indistintamente con péptido AQUA) significa un péptido que comprende al menos una marca de isótopo pesado, que es adecuada para la cuantificación o detección absolutas de una proteína como se describe en el documento WO/03016861, "Absolute Quantification of Proteins and Modified Forms Thereof by Multistage Mass Spectrometry" (Gygi *et al.*), comentado adicionalmente más adelante. El término
- 25 "detecta específicamente" con respecto a un péptido AQUA significa que el péptido solamente detectará y cuantificará polipéptidos y proteínas que contienen la secuencia del péptido AQUA y no detectará sustancialmente polipéptidos y proteínas que no contienen la secuencia del péptido AQUA.
- "Aisladas" (o "sustancialmente purificadas") se refiere a secuencias de ácidos nucleicos o de aminoácidos que son retiradas de su medio natural, aisladas o separadas. Se encuentran preferiblemente libres en al menos 60%, más preferiblemente libres en 75%, y lo más preferiblemente libres en 90% o más de otros componentes con los que están asociadas de forma natural.
- 30 "Mimético" se refiere a una molécula, cuya estructura es desarrollada a partir del conocimiento de la estructura de un polipéptido de fusión de ALK o porciones del mismo y, como tal, es capaz de efectuar algunas o todas las acciones de las moléculas de tipo proteína asociadas con la translocación.
- 35 "ALK mutante" o polinucleótido o polipéptido de "fusión" se refiere a un polinucleótido o polipéptido de fusión que implica ALK y una proteína secundaria (p. ej., EML-4 o TFG), como se describe en la presente memoria.
- "Polinucleótido" (o "secuencia de nucleótidos") se refiere a un oligonucleótido, nucleótido, o polinucleótido, y fragmentos o porciones de los mismos, y a ADN o ARN de origen genómico o sintético, que puede ser de hebra sencilla o de doble hebra, y representan la hebra efectora o antisentido.
- 40 "Polipéptido" (o "secuencia de aminoácidos") se refiere a una secuencia de oligopéptido, péptido, polipéptido, o proteína, y fragmentos o porciones de la misma, y a moléculas de origen natural o sintético. Cuando en la presente memoria se cita "secuencia de aminoácidos" para referirse a una secuencia de aminoácidos de una molécula de proteína de origen natural, no se pretende que "secuencia de aminoácidos" y términos similares, tales como
- 45 "polipéptido" o "proteína", limiten la secuencia de aminoácidos a la secuencia de aminoácidos nativa, completa asociada con la molécula de proteína citada.
- "Polinucleótido de fusión de EML4-ALK" se refiere a la secuencia de ácido nucleico de un producto génico mutante por delección o un polinucleótido de fusión (variante corta o larga) de EML4-ALK sustancialmente purificado como se describe en la presente memoria, obtenido a partir de cualquier especie, particularmente de mamífero, incluyendo bovino, ovino, porcino, equino, y preferentemente humano, de cualquier origen ya sea natural, sintético, semisintético, o recombinante.
- 50 "Polipéptido de fusión de EML4-ALK" se refiere a la secuencia de aminoácidos de un polipéptido de fusión de EML4-ALK (variante corta o larga) sustancialmente purificado descrito en la presente memoria, obtenido a partir de cualquier especie, particularmente de mamífero, incluyendo bovino, ovino, porcino, equino, y, preferiblemente, humano, de cualquier origen ya sea natural, sintético, semisintético, o recombinante.
- 55 "Polinucleótido de fusión de TFG-ALK" se refiere a la secuencia de ácido nucleico de un producto génico mutante

por translocación o polinucleótido de fusión de TFG-ALK sustancialmente purificado como se describe en la presente memoria, obtenido a partir de cualquier especie, particularmente de mamífero, incluyendo bovino, ovino, porcino, murino, equino, y, preferiblemente, humano, de cualquier origen ya sea natural, sintético, semisintético, o recombinante.

- 5 "Polipéptido de fusión de TFG-ALK" se refiere a la secuencia de aminoácidos de un polipéptido de fusión de TFG-ALK sustancialmente purificado descrito en la presente memoria, obtenido a partir de cualquier especie, particularmente de mamífero, incluyendo bovino, ovino, porcino, murino, equino, y preferentemente humano, de cualquier origen ya sea natural, sintético, semisintético, o recombinante.

- 10 Los términos "se une específicamente a" (o "que se une específicamente" o "unión específica") en referencia a la interacción de un anticuerpo y una proteína o péptido, significan que la interacción depende de la presencia de una estructura concreta (es decir, el determinante antigénico o epítipo) en la proteína; en otras palabras, el anticuerpo está reconociendo y se está uniendo a una estructura de proteína específica en lugar de a las proteínas en general. El término "no se une" con respecto a la unión de un anticuerpo a secuencias o determinantes antigénicos distintos de aquel para el cual es específico significa que no reacciona sustancialmente en comparación con la unión del anticuerpo al determinante antigénico o secuencia para el cual es específico el anticuerpo.

- 15 El término "condiciones rigurosas" con respecto a las condiciones de hibridación de secuencias o sondas es la "rigurosidad" que existe en un intervalo de aproximadamente T_m menos 5°C (5°C por debajo de la temperatura de fusión (T_m) de la sonda o secuencia) a aproximadamente 20°C a 25°C por debajo de la T_m . Las condiciones rigurosas típicas son: incubación durante la noche a 42°C en una disolución que comprende: formamida al 50%, 5 X SSC (NaCl 750 mM, citrato trisódico 75 mM), fosfato de sodio 50 mM (pH 7,6), 5X solución de Denhardt, sulfato de dextrano al 10%, y 20 microgramos /ml de ADN de esperma de salmón sometido a cizalla, desnaturalizado, seguido de lavado de los filtros en 0,1 X SSC a aproximadamente 65°C. Como entenderán los expertos en la técnica, la rigurosidad de la hibridación se puede alterar con el fin de identificar o detectar secuencias de polinucleótidos idénticas o relacionadas.

- 25 Una "variante" de un polipéptido de ALK mutante, se refiere a una secuencia de aminoácidos que es alterada por uno o más aminoácidos. La variante puede tener cambios "conservativos", en donde el aminoácido sustituido tiene propiedades estructurales o químicas similares, p. ej., sustitución de leucina por isoleucina. Más raramente, una variante puede tener cambios "no conservativos", p. ej., sustitución de una glicina por triptófano. Las variaciones menores similares también pueden incluir deleciones o inserciones de aminoácidos, o ambas. Las pautas para determinar qué residuos de aminoácidos pueden ser sustituidos, insertados, o suprimidos sin anular la actividad biológica o inmunológica se pueden encontrar usando programas de ordenador bien conocidos en la técnica, por ejemplo, el soporte lógico DNASTAR.

A. Identificación de quinasas ALK mutantes en tumores sólidos humanos.

- 35 Las nuevas deleciones de genes humanos descritas en la presente memoria, que se producen en el cromosoma 2 y dan como resultado la expresión de dos variantes de la proteína de fusión que combinan el extremo N de EML-4 con el dominio quinasa y el extremo C de ALK, se identificaron sorprendentemente durante el examen de perfiles de péptidos fosforilados globales en extractos de líneas celulares de carcinoma de pulmón de células no pequeñas (CPCNP) (incluyendo H2228) y tumores sólidos de pacientes. El CPCNP, un tumor sólido, es un subtipo de cáncer de pulmón. Las proteínas implicadas en estas fusiones por delección se muestran en las Figuras 1A-1B, panel A.

- 40 El perfil de fosforilación de la línea celular H2228 se dilucidó primero usando una técnica recientemente descrita para el aislamiento y caracterización mediante espectrometría de masas de péptidos modificados a partir de mezclas complejas (véase la Publicación de Patente de los Estados Unidos Núm. 20030044848, Rush et al., "Immunoaffinity Isolation of Modified Peptides from Complex Mixtures" (técnica "IAP"), como se describe adicionalmente en el Ejemplo 1 de más abajo. La aplicación de la técnica IAP utilizando un anticuerpo específico de fosfotirosina (CELL SIGNALING TECHNOLOGY, INC., Beverly, MA, 2003/04 Núm. de Cat. 9411), identificó que la línea celular H2228 expresa la quinasa ALK, pero que la proteína estaba aparentemente truncada. El escrutinio identificó muchas otras quinasas activadas en la línea celular, incluyendo algunas que se sabe que están activadas en el cáncer de pulmón. El análisis de la secuencia 5' a ALK mediante 5'RACE identificó a continuación que la quinasa se fusionaba al extremo N de EML-4 (véase la Fig. 6).

- 50 El examen posterior de 154 muestras de tumor de pacientes de CPCNP utilizando el mismo enfoque de perfilado de la fosforilación global no solo confirmó la presencia de la mutación EML4-ALK (variante corta) en una población de esos pacientes, sino que también reveló la presencia de una segunda EML4-ALK (variante larga) y la presencia de la mutación TFG-ALK en otras poblaciones de pacientes (véase el Ejemplo 1B y 1C).

- 55 La confirmación de que las proteínas de ALK mutantes están dirigiendo la proliferación y la supervivencia celular en estos tumores de CPCNP se puede establecer mediante la inhibición de las células utilizando el silenciamiento con ARNip (véase el Ejemplo 3).

Los genes de fusión de EML4-ALK (variantes cortas y largas) y el gen de fusión de TFG-ALK se amplificaron por PCR, se aislaron, y se secuenciaron (véase el Ejemplo 3). Como se muestra en el panel B de las Figuras 1A-1B, la

deleción de EML4-ALK combina el extremo N de EML-4 de tipo salvaje (aminoácidos 1-233 en la variante corta, o aminoácidos 1-495 en la variante larga) con el dominio quinasa y el extremo C de ALK de tipo salvaje (aminoácidos 1057-1620) (véanse también los SEQ ID NO: 3 y 5). La unión por fusión se produce justo en el extremo C con respecto al dominio transmembrana de ALK de tipo salvaje (véanse las Figuras 1A-1B). Los polipéptidos de fusión de EML4-ALK conservan los 233 o 495 aminoácidos N-terminales de EML-4, respectivamente, lo que incluye el dominio de bobina en espiral de esta proteína. Las proteínas de fusión de EML4-ALK resultantes, que comprenden 796 aminoácidos (variante corta) o 1059 aminoácidos (variante larga), respectivamente (véase el panel B de las Figuras 1A-1B y las Figuras 2A-2B (SEQ ID NO: 1 y 18)), conservan la actividad de la quinasa ALK. Los exones implicados y la unión de fusión se muestran en las Figuras 1A-1B (panel B). La unión por fusión incluye el intrón 6 de EML-4, que sigue al exón 6 (variante corta) o al exón 13 de EML-4 (variante larga).

Como se muestra en el panel B de la Figura 1C, la translocación de TFG-ALK combina el extremo N de TFG de tipo salvaje (aminoácidos 1-138) con el dominio quinasa y el extremo C de ALK de tipo salvaje (aminoácidos 1057-1620) (véanse también los SEQ ID NO: 22 y 5, y el panel B de la Figura 1C y la Figura 4C (SEQ ID NO: 20 y 1)). La unión por fusión se produce justo en el extremo C con respecto al dominio transmembrana de ALK de tipo salvaje (véase la Figura 1C) y conserva la actividad de la quinasa ALK. Los exones implicados y la unión por fusión se muestran en la Figura 1C (panel B). La unión de fusión incluye el exón 3 del TFG y el exón 20 de ALK.

Se utilizaron sondas FISH para detectar la presencia de la proteína de fusión de EML4-ALK (variante corta) en un grupo de 400 muestras de tumores de CPCNP humanos embebidas en parafina (véanse los Ejemplos 6 y 7; Figura 6). La incidencia de esta mutación de variante corta en este tamaño de muestra fue muy baja. Sin embargo, la expresión de las proteínas de fusión de EML4-ALK (variantes cortas y largas), así como la proteína de fusión de TFG-ALK, se detectó con una mayor incidencia utilizando la técnica IAP para examinar los perfiles de fosforilación global de otro grupo de 154 muestras congeladas de tumores de CPCNP humanos de pacientes (véase el Ejemplo 1B).

B. Polinucleótidos aislados.

La presente descripción proporciona, en parte, polinucleótidos aislados que codifican polipéptidos de fusión de EML4-ALK, sondas de nucleótidos que hibridan con tales polinucleótidos, y métodos, vectores y células anfitrionas para la utilización de tales polinucleótidos para producir polipéptidos de fusión recombinantes.

A menos que se indique lo contrario, todas las secuencias de nucleótidos determinadas por secuenciación de una molécula de ADN en la presente memoria se determinaron utilizando un secuenciador de ADN automatizado (tal como el Modelo 373 de Applied Biosystems, Inc.), y se determinaron todas las secuencias de aminoácidos de polipéptidos codificados por moléculas de ADN determinadas en la presente memoria utilizando un secuenciador automatizado de péptidos (véase el Ejemplo 2). Como es conocido en la técnica para cualquier secuencia de ADN determinada mediante este enfoque automatizado, cualquier secuencia de nucleótidos determinada en la presente memoria puede contener algunos errores. Las secuencias de nucleótidos determinadas mediante automatización son típicamente idénticas al menos en aproximadamente 90%, más típicamente idénticas en al menos aproximadamente 95% a al menos aproximadamente 99,9% a la secuencia de nucleótidos real de la molécula de ADN secuenciada. La secuencia real puede determinarse más precisamente por otros enfoques, incluyendo métodos manuales de secuenciación de ADN bien conocidos en la técnica. Como también es conocido en la técnica, una sola inserción o deleción en una secuencia de nucleótidos determinada en comparación con la secuencia real causará un desplazamiento de marco en la traducción de la secuencia de nucleótidos de manera que la secuencia de aminoácidos predicha codificada por una secuencia de nucleótidos determinada será completamente diferente de la secuencia de aminoácidos realmente codificada por la molécula de ADN secuenciada, comenzando en el punto de semejante inserción o deleción.

A menos que se indique lo contrario, cada secuencia de nucleótidos expuesta en la presente memoria se presenta como una secuencia de desoxirribonucleótidos (abreviados A, G, C y T). Sin embargo, por "secuencia de nucleótidos" de una molécula de ácido nucleico o polinucleótido se entiende, para una molécula de ADN o polinucleótido, una secuencia de desoxirribonucleótidos, y para una molécula de ARN o polinucleótido, la secuencia correspondiente de ribonucleótidos (A, G, C y U), donde cada desoxirribonucleótido timidina (T) en la secuencia de desoxirribonucleótidos especificada es remplazado por el ribonucleótido uridina (U). Por ejemplo, se pretende que la referencia a una molécula de ARN que tiene la secuencia del SEQ ID NO: 2 mostrada utilizando abreviaturas de desoxirribonucleótidos indique una molécula de ARN que tiene una secuencia en la que cada desoxirribonucleótido A, G o C del SEQ ID NO: 2 ha sido reemplazado por el correspondiente ribonucleótido A, G o C, y cada desoxirribonucleótido T ha sido reemplazado por un ribonucleótido U.

En una realización, la descripción proporciona un polinucleótido aislado que comprende una secuencia de nucleótidos idéntica en al menos 95% a una secuencia seleccionada del grupo que consiste en:

(a) una secuencia de nucleótidos que codifica un polipéptido de fusión Proteína de Tipo 4 Asociada a Microtúbulos de Equinodermo/Quinasa de Linfoma Anaplásico (EML4-ALK) que comprende la secuencia de aminoácidos del SEQ ID NO: 1 o del SEQ ID NO: 18;

(b) una secuencia de nucleótidos que codifica un polipéptido de fusión de EML4-ALK, comprendiendo dicha secuencia de nucleótidos la secuencia de nucleótidos del SEQ ID NO: 2 o SEQ ID NO: 19;

(c) una secuencia de nucleótidos que codifica un polipéptido de fusión de EML4-ALK que comprende la secuencia de aminoácidos N-terminal de EML-4 (residuos 1-233 del SEQ ID NO: 3 o residuos 1 a 495 del SEQ ID NO: 3) y el dominio quinasa ALK (residuos 1116 a 1383 del SEQ ID NO: 5);

(d) una secuencia de nucleótidos que comprende la secuencia de nucleótidos N-terminal de EML-4 (nucleótidos 1-700 del SEQ ID NO: 4 o los nucleótidos 1-1486 del SEQ ID NO: 4) y la secuencia de nucleótidos de dominio quinasa ALK (nucleótidos 3348-4149 del SEQ ID NO: 6);

(e) una secuencia de nucleótidos que comprende al menos seis nucleótidos contiguos que abarcan la unión por fusión (nucleótidos 700-701 del SEQ ID NO: 2 o los nucleótidos 1486-1487 del SEQ ID NO: 19) de un polinucleótido de fusión de EML4-ALK;

(f) una secuencia de nucleótidos que codifica un polipéptido que comprende al menos seis aminoácidos contiguos que abarcan la unión por fusión (residuos 233-234 del SEQ ID NO: 1 o los residuos 495-496 del SEQ ID NO: 18) de un polipéptido de fusión de EML4-ALK; y

(g) una secuencia de nucleótidos complementaria a cualquiera de las secuencias de nucleótidos de (a) - (f).

Utilizando la información proporcionada en la presente memoria, tal como la secuencia de nucleótidos de la Figura 2 (SEQ ID NO: 2), se puede obtener una molécula de ácido nucleico de la presente descripción que codifica un polipéptido de ALK mutante de la descripción utilizando procedimientos de clonación y escrutinio convencionales, tales como aquellos para clonar ADNc utilizando ARNm como material de partida. El polinucleótido de fusión de EML4-ALK (variante corta), ilustrativo de la descripción, descrito en la Figura 2 (SEQ ID NO: 2) se aisló a partir de ADN genómico de una línea celular de CPCNP humano (como se describe adicionalmente en el Ejemplo 2 a continuación). El gen de fusión también se puede identificar genotecas de ADN genómico o ADNc en otros tipos de cáncer, incluyendo tumores sólidos, en los que se produce una delección descrita del gen EML4-ALK (cromosoma 2).

Las secuencias de nucleótidos determinadas de los genes de fusión de EML4-ALK (SEQ ID NO: 2 y 19) codifican proteínas de fusión de quinasa de 796 aminoácidos (variante corta) y 1059 aminoácidos (variante larga), respectivamente (véanse las Figuras 2A-B (SEQ ID NO: 1 y 18) y las Figuras 1A-B). Los polinucleótidos de fusión de EML4-ALK comprenden la porción de la secuencia de nucleótidos de EML-4 de tipo salvaje (véase la Figura 3 (SEQ ID NO: 4)) que codifica el extremo N (aminoácidos 1-233 (variante corta) o aminoácidos 1-495 (variante larga)) de esa proteína con la porción de la secuencia de nucleótidos de ALK de tipo salvaje (véase la Figura 4 (SEQ ID NO: 6)) que codifica el dominio quinasa y el extremo C de esa proteína. Véanse las Figuras 1A-B. El dominio quinasa comprende los residuos 292-568 de la proteína de fusión de la variante corta (codificada por los nucleótidos 874-1704 del polinucleótido de fusión de la variante corta) y los residuos 555-831 de la proteína de fusión de la variante larga (codificada por los nucleótidos 1663-2494 del polinucleótido de fusión de la variante larga). Véanse las Figuras 2A-2B.

Como se ha indicado, la presente descripción proporciona, en parte, la forma madura de las proteínas de fusión de EML4-ALK. De acuerdo con la hipótesis de la señal, las proteínas secretadas por células de mamíferos tienen una secuencia señal o líder secretora que se escinde de la proteína madura una vez que se ha iniciado la exportación de la cadena de proteína en crecimiento a través del retículo endoplásmico rugoso. La mayoría de las células de mamífero e incluso las células de insecto escinden las proteínas secretadas con la misma especificidad. Sin embargo, en algunos casos, la escisión de una proteína secretada no es enteramente uniforme, lo que da como resultado dos o más especies maduras sobre la proteína. Adicionalmente, hace tiempo que se sabe que la especificidad de escisión de una proteína secretada está determinada en última instancia por la estructura primaria de la proteína completa, es decir, ésta es inherente en la secuencia de aminoácidos del polipéptido.

Por polipéptido de EML4-ALK maduro que tiene la secuencia de aminoácidos codificada, p. ej., por el clon de ADNc depositado, se quiere significar la forma madura de esta proteína de fusión producida por la expresión en una célula de mamífero (p. ej., células 3T3, como se describe a continuación) del marco de lectura abierto completo codificado por la secuencia de ADN humana del clon depositado o de otro clon que codifica el polipéptido de fusión maduro.

Como se ha indicado, los polinucleótidos de la presente descripción pueden estar en forma de ARN, tal como ARNm, o en forma de ADN, incluyendo, por ejemplo, ADNc y ADN genómico obtenidos por clonación o producidos sintéticamente. El ADN puede ser de doble hebra o de hebra sencilla. El ADN o ARN de hebra sencilla puede ser la hebra codificante, también conocida como hebra efectora, o puede ser la hebra no codificante, también conocida como hebra antisentido.

Los polinucleótidos aislados de la descripción son moléculas de ácido nucleico, ADN o ARN, que han sido separadas de su ambiente nativo. Por ejemplo, las moléculas de ADN recombinante contenidas en un vector se consideran aisladas para los fines de la presente descripción. Otros ejemplos de moléculas de ADN aisladas incluyen moléculas de ADN recombinante mantenidas en células anfitrionas heterólogas o moléculas de ADN purificadas (parcial o sustancialmente) en disolución. Las moléculas de ARN aisladas incluyen transcritos de ARN *in*

vivo o *in vitro* de las moléculas de ADN de la presente descripción. Las moléculas de ácido nucleico aisladas de acuerdo con la presente descripción incluyen adicionalmente tales moléculas producidas sintéticamente.

Los polinucleótidos aislados de la descripción incluyen la molécula de ADN mostrada en las Figuras 2A-B (SEQ ID NO: 2 y 19), moléculas de ADN que comprenden la secuencia codificante para las proteínas de fusión de EML4-ALK maduras mostradas en las Figuras 1A-B (SEQ ID NO: 1 y 18), y moléculas de ADN que comprenden una secuencia sustancialmente diferente de las descritas anteriormente pero que, debido a la degeneración del código genético, todavía codifican un polipéptido de ALK mutante de la descripción. El código genético es bien conocido en la técnica, por lo tanto, sería rutinario para un experto en la técnica generar tales variantes degeneradas.

En otra realización, la descripción proporciona un polinucleótido aislado que codifica el polipéptido de fusión de EML4-ALK que comprende la secuencia de nucleótidos de fusión de EML4-ALK contenida en el clon de ADNc depositado anteriormente descrito. Preferiblemente, tal molécula de ácido nucleico codificará el polipéptido de fusión maduro codificado por el clon de ADNc depositado u otro clon que expresa una proteína de fusión de EML4-ALK completa descrita en la presente memoria. En otra realización, la descripción proporciona una secuencia de nucleótidos aislada que codifica un polipéptido de fusión de EML4-ALK que comprende la secuencia de aminoácidos N-terminal de EML-4 (residuos 1-233 del SEQ ID NO: 3 o residuos 1-495 del SEQ ID NO: 3) y el dominio quinasa ALK (residuos 1116-1383 del SEQ ID NO: 5). En una realización, el polipéptido que comprende el dominio quinasa ALK comprende los residuos 1057-1620 del SEQ ID NO: 5 (véase la Figura 1, panel B). En otra realización, la secuencia de aminoácidos N-terminal de EML-4 y el dominio quinasa ALK antes mencionados están codificados por secuencias de nucleótidos que comprende los nucleótidos 1-700 del SEQ ID NO: 4 o los nucleótidos 1-1486 del SEQ ID NO: 4 y los nucleótidos 3171-4860 del SEQ ID NO: 6, respectivamente.

La descripción proporciona adicionalmente polinucleótidos aislados que comprenden secuencias de nucleótidos que tienen una secuencia complementaria a uno de los polinucleótidos de ALK mutante de la descripción. Tales moléculas aisladas, particularmente moléculas de ADN, son útiles como sondas para el mapeo de genes, mediante hibridación *in situ* con cromosomas, y para detectar la expresión de una proteína de fusión de EML4-ALK en el tejido humano, por ejemplo, mediante análisis de transferencia Northern, como se describe adicionalmente en la Sección F a continuación.

La presente descripción se dirige adicionalmente a fragmentos de las moléculas de ácido nucleico aisladas descritas en la presente memoria. Por fragmento de un polinucleótido de EML4-ALK aislado de la descripción se quieren significar fragmentos de al menos aproximadamente 15 nucleótidos, y más preferiblemente al menos aproximadamente 20 nucleótidos, aún más preferiblemente al menos aproximadamente 30 nucleótidos, e incluso más preferiblemente, al menos aproximadamente 40 nucleótidos de longitud, que son útiles como sondas y cebadores de diagnóstico como se comenta en la presente memoria. Por supuesto, también son útiles fragmentos más grandes de aproximadamente 50-1500 nucleótidos de longitud de acuerdo con la presente descripción, como son los fragmentos correspondientes a la mayoría, si no todas, de las secuencias de nucleótidos de ALK mutantes de los ADNc depositados o mostrados en la Figura 2 (SEQ ID NO: 2) o de otro clon que expresa el polinucleótido de fusión mostrado en las Figuras 2A-B (SEQ ID NO: 2 o 19). Por fragmento de al menos 20 nucleótidos de longitud, por ejemplo, se quieren significar fragmentos que incluyen 20 o más bases contiguas de las secuencias de nucleótidos respectivas de los que derivan los fragmentos. La generación de tales fragmentos de ADN es rutinaria para el experto en la técnica, y se puede llevar a cabo, a modo de ejemplo, mediante escisión con endonucleasas de restricción o cizallamiento por sonicación de ADN obtenible a partir del clon de ADNc depositado o sintetizado de acuerdo con la secuencia descrita en la presente memoria. Alternativamente, tales fragmentos se pueden generar directamente sintéticamente.

Los fragmentos de ácido nucleico o sondas preferidos de la presente descripción incluyen moléculas de ácido nucleico que codifican la unión por fusión de los productos génicos de fusión de EML4-ALK (véanse las Figuras 1A-B, panel B). Por ejemplo, en ciertas realizaciones preferidas, un polinucleótido aislado de la descripción comprende una secuencia/fragmento de nucleótidos que comprende al menos seis nucleótidos contiguos que incluyen la unión por fusión (nucleótidos 700-701 del SEQ ID NO: 2 o los nucleótidos 1486-1487 del SEQ ID NO: 19) de un polinucleótido de fusión de EML4-ALK (véanse las Figuras 1A-B, panel B (SEQ ID NO: 8 y 25)). En otra realización preferida, un polinucleótido aislado de la descripción comprende una secuencia/fragmento de nucleótidos que codifica un polipéptido que comprende al menos seis aminoácidos contiguos que abarcan la unión por fusión (residuos 233-234 del SEQ ID NO: 1 o residuos 495-496 del SEQ ID NO: 18) de un polipéptido de fusión de EML4-ALK (véanse las Figuras 1A-B, panel inferior (SEQ ID NO: 7 y 24)).

En otro aspecto, la descripción proporciona un polinucleótido aislado que hibrida en condiciones de hibridación rigurosas con una porción de un polinucleótido de ALK mutante de la descripción como se describe en la presente memoria. Por "condiciones de hibridación rigurosas" se entiende una incubación durante la noche a 42°C en una disolución que comprende: formamida al 50%, 5 X SSC (NaCl 750 mM, citrato trisódico 75 mM), fosfato de sodio 50 mM (pH 7,6), 5X solución de Denhardt, sulfato de dextrano al 10%, y 20 microgramos/ml de ADN de esperma de salmón sometido a cizalla, desnaturalizado, seguido de lavado de los filtros en 0,1 X SSC a aproximadamente 65°C.

Por polinucleótido que se hibrida con una "porción" de un polinucleótido se quiere significar un polinucleótido (ADN o ARN) que hibrida con al menos aproximadamente 15 nucleótidos (nt), y más preferiblemente al menos

aproximadamente 20 nt, aún más preferiblemente al menos aproximadamente 30 nt, e incluso más preferiblemente aproximadamente 30-70 nt del polinucleótido de referencia. Estos son útiles como sondas y cebadores de diagnóstico como se comentó anteriormente y con más detalle a continuación.

Por supuesto, los polinucleótidos que hibridan a una porción mayor del polinucleótido de referencia (p. ej., el polinucleótido de fusión de EML4-ALK maduro descrito en la Figura 2 (SEQ ID NO: 2)), por ejemplo, una porción de 50-750 nt de longitud, o incluso a toda la longitud del polinucleótido de referencia, también son útiles como sondas, según la presente descripción, como lo son los polinucleótidos correspondientes a la mayoría, si no todas, las secuencias de nucleótidos de los ADNc depositados o las secuencias de nucleótidos mostradas en las Figuras 2A-B (SEQ ID NO: 2 o 19), o las Figuras 1A-B (panel B)) (SEQ ID NO: 7 y 24).

Por porción de un polinucleótido de "al menos 20 nucleótidos de longitud", por ejemplo, se quieren significar 20 o más nucleótidos contiguos de la secuencia de nucleótidos del polinucleótido de referencia. Como se indica, tales porciones son útiles diagnósticamente o bien como una sonda según las técnicas de hibridación de ADN convencionales o bien como cebadores para la amplificación de una secuencia diana por reacción en cadena de la polimerasa (PCR), como se describe, por ejemplo, en MOLECULAR CLONING, A Laboratory Manual, 2ª ed., Sambrook, J., Fritsch, E.F. y Maniatis, T., eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York (1989), cuya descripción completa se incorpora a esta memoria por referencia. Por supuesto, un polinucleótido que hibrida solo con una secuencia poli A (tal como la extensión poli(A) 3' terminal de la secuencia de EML4-ALK mostrada en la Figura 2 (SEQ ID NO: 2)) o con un tramo complementario de residuos de T (o U), no se incluiría en un polinucleótido de la descripción utilizado para hibridar con una porción de un ácido nucleico de la descripción, puesto que tal polinucleótido hibridaría con cualquier molécula de ácido nucleico que contuviera un tramo de poli (A) o el complemento del mismo (p. ej., prácticamente cualquier clon de ADNc de doble hebra).

Como se ha indicado, las moléculas de ácido nucleico de la presente descripción, que codifican un polipéptido de ALK mutante de la descripción, pueden incluir, pero no se limitan a las que codifican la secuencia de aminoácidos del polipéptido maduro, por sí mismas; la secuencia codificante para el polipéptido maduro y secuencias adicionales, tales como las que codifican la secuencia líder o secretora, tal como una secuencia de pre-, o pro- o pre-pro-proteína; la secuencia codificante del polipéptido maduro, con o sin las secuencias codificantes adicionales anteriormente mencionadas, junto con, secuencias no codificantes adicionales, incluyendo por ejemplo, pero no limitadas a intrones y secuencias 5' y 3' no codificantes, tales como las secuencias transcritas, no traducidas que juegan un papel en la transcripción, el procesamiento del ARNm, incluyendo las señales de empalme y poliadenilación, por ejemplo -- unión al ribosoma y estabilidad de ARNm; y una secuencia codificante adicional que codifica aminoácidos adicionales, tales como las que proporcionan funcionalidades adicionales.

De este modo, la secuencia que codifica el polipéptido puede fusionarse a una secuencia marcadora, tal como una secuencia que codifica un péptido que facilita la purificación del polipéptido fusionado. En ciertas realizaciones preferidas de este aspecto de la descripción, la secuencia de aminoácidos marcadora es un péptido de hexa-histidina, tal como la etiqueta proporcionada en un vector pQE (Qiagen, Inc.), entre otras, muchas de las cuales están disponibles en el mercado. Como describen Gentz et al., Proc. Natl. Acad. Sci. USA 86: 821-824 (1989), por ejemplo, la hexa-histidina proporciona una purificación conveniente de la proteína de fusión. La etiqueta "HA" es otro péptido útil para la purificación que corresponde a un epítipo derivado de la proteína hemaglutinina de la gripe, que ha sido descrita por Wilson et al., Cell 37: 767 (1984). Como se comenta más adelante, otras de tales proteínas de fusión incluyen el propio polipéptido de fusión de EML4-ALK fusionado a Fc en el extremo N o C.

La presente descripción se refiere adicionalmente a variantes de las moléculas de ácido nucleico de la presente descripción, que codifican porciones, análogos o derivados de un polipéptido de fusión de EML4-ALK descrito en la presente memoria. Las variantes pueden aparecer de forma natural, tal como una variante alélica natural. Por "variante alélica" se quiere significar una de las varias formas alternativas de un gen que ocupa un locus dado en un cromosoma de un organismo. Véase, p. ej. Genes II, Lewin, B., ed., John Wiley & Sons, Nueva York (1985). Las variantes de origen no natural se pueden producir utilizando mecanismos de mutagénesis conocidos en la técnica.

Tales variantes incluyen las producidas por sustituciones, deleciones o adiciones de nucleótidos. Las sustituciones, deleciones o adiciones pueden implicar uno o más nucleótidos. Las variantes pueden estar alteradas en las regiones codificantes, las regiones no codificantes, o ambas. Las alteraciones en las regiones codificantes pueden producir sustituciones, deleciones o adiciones de aminoácidos conservativas o no conservativas. Especialmente preferidas entre éstas son las sustituciones, adiciones y deleciones silenciosas, que no alteran las propiedades y actividades (p. ej., actividad quinasa) de los polipéptidos de ALK mutantes descritos en la presente memoria. También se prefieren especialmente a este respecto las sustituciones conservativas.

Realizaciones adicionales de la descripción incluyen polinucleótidos aislados que comprenden una secuencia de nucleótidos idéntica al menos en 90%, y más preferiblemente idéntica al menos en 95%, 96%, 97%, 98% o 99%, a un polinucleótido de ALK mutante de la descripción (por ejemplo, una secuencia de nucleótidos que codifica el polipéptido de fusión de EML4-ALK que tiene la secuencia de aminoácidos completa mostrada en la Figura 2 (SEQ ID NO: 1, o una secuencia de nucleótidos que codifica el extremo N de EML-4 y el dominio quinasa ALK (véase la Figura 1, panel B; y las Figuras 3 y 4); o un nucleótido complementario a tales secuencias ilustrativas).

Por polinucleótido que tiene una secuencia de nucleótidos al menos, por ejemplo, "idéntica" en 95% a una secuencia de nucleótidos de referencia que codifica un polipéptido de ALK mutante se quiere significar que la secuencia de nucleótidos del polinucleótido es idéntica a la secuencia de referencia excepto que la secuencia de polinucleótidos puede incluir hasta cinco mutaciones puntuales por cada 100 nucleótidos de la secuencia de nucleótidos de referencia que codifica el polipéptido de ALK mutante. En otras palabras, para obtener un polinucleótido que tiene una secuencia de nucleótidos idéntica al menos en 95% a una secuencia de nucleótidos de referencia, hasta 5% de los nucleótidos en la secuencia de referencia puede ser suprimido o sustituido por otro nucleótido, o se pueden insertar en la secuencia de referencia varios nucleótidos hasta 5% de los nucleótidos totales de la secuencia de referencia. Estas mutaciones de la secuencia de referencia pueden producirse en las posiciones 5' o 3' terminales de la secuencia de nucleótidos de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas individualmente entre los nucleótidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia.

Como cuestión práctica, que cualquier molécula de ácido nucleico concreta sea idéntica en al menos 90%, 95%, 96%, 97%, 98% o 99% a, por ejemplo, las secuencias de nucleótidos mostradas en las Figuras 2A-B (SEQ ID NOS: 2 y 19) o a la secuencia de nucleótidos de los clones de ADNc depositados descritos anteriormente se puede determinar convencionalmente usando programas de ordenador conocidos tales como el programa Bestfit (Wisconsin Sequence Analysis Package, Versión 8 para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711. Bestfit utiliza el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981), para encontrar el mejor segmento de homología entre dos secuencias. Cuando se usa Bestfit o cualquier otro programa de alineamiento de secuencias para determinar si una secuencia concreta es, por ejemplo, idéntica en 95% a una secuencia de referencia de polinucleótido de fusión de EML4-ALK o una secuencia de polinucleótidos de ALK truncado de acuerdo con la presente descripción, los parámetros se ajustan, por supuesto, de manera que el porcentaje de identidad se calcule sobre la longitud completa de la secuencia de nucleótidos de referencia y que se permitan huecos en la homología de hasta 5% del número total de nucleótidos en la secuencia de referencia.

La presente descripción incluye en su alcance moléculas de ácido nucleico idéntica en al menos 90%, 95%, 96%, 97%, 98% o 99% a la secuencia de ácido nucleico mostrada en la Figura 2 (SEQ ID NO: 2), o a las secuencias de ácido nucleico de los ADNc depositados, con independencia de si codifican un polipéptido que tiene actividad de la quinasa ALK. Esto es porque incluso cuando una molécula de ácido nucleico concreta no codifica un polipéptido de fusión que tiene actividad de la quinasa ALK, un experto en la técnica todavía sabría cómo utilizar la molécula de ácido nucleico, por ejemplo, como sonda de hibridación o cebador para la reacción en cadena de la polimerasa (PCR). Los usos de las moléculas de ácido nucleico de la presente descripción que no codifican un polipéptido que tiene quinasa incluyen, entre otros, (1) el aislamiento del gen de delección de EML4-ALK, o el gen de ALK truncado, o variantes alélicas del mismo en una genoteca de ADNc; (2) la hibridación *in situ* (p. ej., "FISH") a dispersiones cromosómicas en metafase para proporcionar una localización cromosómica precisa del gen de supresión de EML4-ALK o el gen de ALK truncado, como describen Verma et al., *HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES*, Pergamon Press, Nueva York (1988); y el análisis de transferencia Northern para detectar la proteína de fusión de EML4-ALK o la expresión de ARNm de la quinasa ALK truncada en tejidos específicos.

Se prefieren, sin embargo, moléculas de ácido nucleico que tienen secuencias idénticas en al menos 95% a un polipéptido de ALK mutante de la descripción o a la secuencia de ácido nucleico de los ADNc depositados que de hecho, codifican un polipéptido de fusión que tiene actividad de la quinasa ALK. Tal actividad puede ser similar, pero no necesariamente idéntica, a la actividad de una proteína de fusión de EML4-ALK descrita en la presente memoria (ya sea la proteína completa, la proteína madura, o un fragmento de proteína que conserva la actividad de quinasa), como se mide en un análisis biológico concreto. Por ejemplo, la actividad de la quinasa ALK se puede examinar determinando su capacidad para fosforilar uno o más sustratos de péptidos que contienen tirosina, por ejemplo, Sustrato 1 o 2 del Receptor de Insulina (IRS1, IRS2), que son sustratos para la quinasa ALK.

Debido a la degeneración del código genético, un experto normal en la técnica reconocerá inmediatamente que un gran número de las moléculas de ácido nucleico que tienen una secuencia idéntica en al menos 90%, 95%, 96%, 97%, 98%, o 99 % a la secuencia de ácido nucleico de los ADNc depositados o las secuencias de ácidos nucleicos mostradas en las Figuras 2A-B (SEQ ID NO: 2 y 19) codificarán un polipéptido mutante que tenga actividad de ALK. De hecho, puesto que las variantes degeneradas de estas secuencias de nucleótidos codifican todas el mismo polipéptido, esto estará claro para el experto en la técnica incluso sin realizar el análisis comparativo anteriormente descrito. Adicionalmente se reconocerá en la técnica que, para tales moléculas de ácido nucleico que no son variantes degeneradas, un número razonable también codificará un polipéptido que conserve la actividad de la quinasa ALK. Esto es porque el experto en la técnica es plenamente consciente de las sustituciones de aminoácidos que es menos probable o no es probable que afecten significativamente a la función de la proteína (p. ej., remplazando un aminoácido alifático por un segundo aminoácido alifático).

Por ejemplo, las pautas concernientes a cómo realizar sustituciones de aminoácidos fenotípicamente silenciosas son proporcionadas por Bowie et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions", *Science* 247: 1306-1310 (1990), que describen dos enfoques principales para estudiar la tolerancia de una secuencia de aminoácidos al cambio. El primer método se basa en el proceso de evolución, en el que las mutaciones son aceptadas o rechazadas por la selección natural. El segundo enfoque utiliza ingeniería genética

para introducir cambios de aminoácidos en posiciones específicas de un gen clonado y selecciones o escrutinio para identificar secuencias que mantienen la funcionalidad. Estos estudios han revelado que las proteínas son sorprendentemente tolerantes a las sustituciones de aminoácidos. Los expertos en la técnica familiarizados con tales técnicas también aprecian qué cambios de aminoácidos es probable que sean permisivos en una cierta posición de la proteína. Por ejemplo, los residuos de aminoácidos más escondidos requieren cadenas laterales no polares, mientras que se conservan generalmente pocas características de las cadenas laterales de la superficie. Otras de tales sustituciones fenotípicamente silenciosas son descritas por Bowie *et al.*, más arriba, y en las referencias allí citadas.

Los métodos para la secuenciación de ADN que son bien conocidos y están generalmente disponibles en la técnica se pueden usar para poner en práctica cualquiera de las realizaciones de polinucleótidos de la descripción. Los métodos pueden emplear enzimas tales como el fragmento Klenow de la ADN polimerasa I, SEQUENASE® (US Biochemical Corp., Cleveland, Ohio), polimerasa Taq (Perkin Elmer), polimerasa T7 termoestable (Amersham, Chicago, IL), o combinaciones de polimerasas recombinantes y exonucleasas de corrección de pruebas, tales como el sistema de amplificación elongasa comercializado por Gibco BRL (Gaithersburg, MD). Preferiblemente, el procedimiento se automatiza con máquinas tales como Hamilton Micro Lab 2200 (Hamilton, Reno, Nev.), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) y los secuenciadores de ADN ABI 377 (Perkin Elmer).

Las secuencias de polinucleótidos que codifican un polipéptido de ALK mutante de la descripción se pueden ampliar utilizando una secuencia parcial de nucleótidos y empleando diversos métodos conocidos en la técnica para detectar secuencias aguas arriba tales como promotores y elementos reguladores. Por ejemplo, un método que se puede emplear, la PCR para "sitio de restricción", utiliza cebadores universales para recuperar la secuencia desconocida adyacente a un locus conocido (Sarkar, G., *PCR Methods Applic.* 2: 318-322 (1993)). En particular, el ADN genómico se amplifica primero en presencia del cebador para la secuencia conectora y un cebador específico para la región conocida. Los cebadores ilustrativos son los descritos en el Ejemplo 2 en la presente memoria. Las secuencias amplificadas se someten después a una segunda ronda de PCR con el mismo cebador conector y otro cebador específico interno con respecto al primero. Los productos de cada ronda de PCR se transcriben con una ARN polimerasa apropiada y se secuencian utilizando transcriptasa inversa.

También se puede utilizar la PCR inversa para amplificar o ampliar secuencias utilizando cebadores divergentes basados en una región conocida (Triglia *et al.*, *Nucleic Acids Res.* 16:8186 (1988)). Los cebadores se pueden diseñar utilizando el soporte lógico OLIGO 4.06 Primer Analysis (National Biosciences Inc., Plymouth, Minn.), u otro programa apropiado, para que tengan 22-30 nucleótidos de longitud, tengan un contenido de GC de 50% o más, e hibriden con la secuencia diana a temperaturas de aproximadamente 68-72°C. El método utiliza varias enzimas de restricción para generar un fragmento adecuado en la región conocida de un gen. El fragmento es circularizado a continuación por medio de ligación intramolecular y utilizado como un molde para la PCR.

Otro método que se puede utilizar es la PCR de captura que implica la amplificación por PCR de fragmentos de ADN adyacentes a una secuencia conocida en ADN de cromosomas artificiales humanos y de levadura (Lagerstrom *et al.*, *PCR Methods Applic.* 1: 111-119 (1991)). En este método, también se pueden utilizar múltiples digestiones y ligaciones con enzimas de restricción para colocar una secuencia de doble hebra manipulada genéticamente en una porción desconocida de la molécula de ADN antes de realizar la PCR. Otro método que se puede usar para recuperar secuencias desconocidas es el descrito por Parker *et al.*, *Nucleic Acids Res.* 19: 3055-3060 (1991)). Adicionalmente, se pueden utilizar PCR, cebadores anidados y bibliotecas PROMOTERFINDER® para pasear por el ADN genómico (Clontech, Palo Alto, Calif.). Este procedimiento evita la necesidad de escrutar bibliotecas y es útil para encontrar uniones intrón/exón.

Cuando se realiza el escrutinio para determinar ADNc completos, es preferible usar bibliotecas que hayan sido seleccionadas por tamaño para incluir ADNc más grandes. Además, se prefieren las bibliotecas de cebado al azar, ya que contendrán más secuencias que contengan regiones 5' de genes. El uso de una biblioteca cebada aleatoriamente puede ser especialmente preferible para situaciones en las que una biblioteca de oligo d(T) no produce un ADNc completo. Las bibliotecas genómicas pueden ser útiles para la ampliación de secuencias en las regiones reguladoras no transcritas 5' y 3'.

Los sistemas de electroforesis capilar, que están disponibles comercialmente, se pueden utilizar para analizar el tamaño o confirmar la secuencia de nucleótidos de los productos de secuenciación o PCR. En particular, la secuenciación capilar puede emplear polímeros verticales para la separación electroforética, cuatro colorantes fluorescentes diferentes (uno para cada nucleótido) que son activados por láser, y detección de las longitudes de onda emitidas por medio de una cámara con un dispositivo de carga acoplada. La intensidad de luz de salida se puede convertir en señal eléctrica utilizando el soporte lógico apropiado (p. ej., GENOTYPER™ y SEQUENCE NAVIGATOR™, Perkin Elmer) y todo el proceso desde la carga de las muestras al análisis por ordenador y la presentación de datos electrónicos se pueden controlar por ordenador. La electroforesis capilar es especialmente preferible para la secuenciación de pequeñas porciones de ADN que podrían estar presentes en cantidades limitadas en una muestra concreta.

C. Vectores y células anfitrionas.

La presente descripción también proporciona vectores recombinantes que comprenden un polinucleótido aislado de la presente descripción, células anfitrionas que están manipuladas genéticamente con los vectores recombinantes, y la producción de polipéptidos de EML4-ALK recombinantes o fragmentos de los mismos mediante técnicas recombinantes.

- 5 Los constructos recombinantes se pueden introducir en células anfitrionas utilizando técnicas bien conocidas tales como infección, transducción, transfección, transvección, electroporación y transformación. El vector puede ser, por ejemplo, un vector fágico, plasmídico, viral o retroviral. Los vectores retrovirales pueden ser de replicación competente o de replicación defectuosa. En este último caso, la propagación viral generalmente se producirá solo en células anfitrionas de complementación.
- 10 Los polinucleótidos se pueden unir a un vector que contiene un marcador seleccionable para la propagación en un anfitrión. Generalmente, se introduce un vector plasmídico en un producto precipitado, tal como un producto precipitado con fosfato de calcio, o en un complejo con un lípido cargado. Si el vector es un virus, éste se puede empaquetar in vitro utilizando una línea celular de empaquetamiento apropiada y transducirlo a continuación a células anfitrionas.
- 15 Se prefieren los vectores que comprenden regiones de control que actúan en cis con respecto al polinucleótido de interés. Los factores que actúan en trans apropiados pueden ser suministrados por el anfitrión, suministrados por un vector de complementación o suministrados por el propio vector tras la introducción en el anfitrión. En ciertas realizaciones preferidas a este respecto, los vectores proporcionan una expresión específica, que puede ser inducible y/o específica del tipo celular. Son particularmente preferidos entre tales vectores aquellos inducibles por factores ambientales que son fáciles de manipular, tales como la temperatura y aditivos nutrientes.
- 20

Los vectores de expresión útiles en la presente descripción incluyen vectores cromosómicos, episómicos y derivados de virus, p. ej., vectores derivados de plásmidos bacterianos, bacteriófagos, episomas de levadura, elementos cromosómicos de levadura, virus tales como baculovirus, papovavirus, virus vaccinia, adenovirus, virus de la viruela aviar, virus de pseudorrabia y retrovirus, y vectores derivados de combinaciones de los mismos, tales como cósmidos y fagémidos.

25

- El inserto de ADN que comprende un polinucleótido de EML4-ALK o de la descripción debe estar conectado operativamente a un promotor apropiado, tal como el promotor PL del fago lambda, los promotores lac, trp y tac de *E. coli*, los promotores temprano y tardío de SV40 y los promotores de LTR retrovirales, por nombrar unos pocos. Otros promotores adecuados son conocidos por el experto en la técnica. Los constructos de expresión contendrán adicionalmente sitios para el inicio de la transcripción, la terminación y, en la región transcrita, un sitio de unión al ribosoma para la traducción. La porción codificante de los transcritos maduros expresados por los constructos incluirá preferiblemente un inicio de la traducción al principio y un codón de terminación (UAA, UGA o UAG) situado apropiadamente al final del polipéptido que se va a traducir.
- 30

- Como se ha indicado, los vectores de expresión incluirán preferiblemente al menos un marcador seleccionable. Tales marcadores incluyen los genes de la dihidrofolato reductasa o de resistencia a neomicina para el cultivo de células eucariotas y de resistencia a tetraciclina o ampicilina para el cultivo en *E. coli* y otras bacterias. Los ejemplos representativos de los anfitriones apropiados incluyen, pero no se limitan a, células bacterianas, tales como células de *E. coli*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura; células de insecto tales como células de *Drosophila* S2 y de *Spodoptera* Sf9; células animales tales como células CHO, COS y de melanoma de Bowes; y células vegetales. Los medios de cultivo apropiados y las condiciones para las células anfitrionas descritas anteriormente son conocidos en la técnica.
- 35 40

- Entre los vectores preferidos para su uso en bacterias se incluyen pQE70, pQE60 y pQE-9, asequibles de Qiagen; vectores pBS, vectores Phagescript, vectores Bluescript, pNH8A, pNH16a, pNH18A, pNH46A, asequibles de Stratagene; y ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 asequibles de Pharmacia. Entre los vectores eucarióticos preferidos se encuentran pWLNEO, pSV2CAT, pOG44, pXT1 y pSG asequibles de Stratagene; y pSVK3, pBPV, pMSG y pSVL asequibles de Pharmacia. Otros vectores adecuados serán fácilmente evidentes para el experto en la técnica.
- 45

- Entre los promotores bacterianos conocidos adecuados para su uso en la presente descripción se incluyen los promotores lacI y lacZ de *E. coli*, los promotores T3 y T7, el promotor gpt, los promotores PR y PL de lambda y el promotor trp. Los promotores eucarióticos adecuados incluyen el promotor inmediato temprano de CMV, el promotor de la timidina quinasa de HSV, los promotores tempranos y tardíos de SV40, los promotores de LTR retrovirales, tales como los del virus del sarcoma de Rous (RSV), y los promotores de metalotioneína, tales como el promotor de metalotioneína 1 de ratón.
- 50

- En la levadura, *Saccharomyces cerevisiae*, se pueden utilizar varios vectores que contienen promotores constitutivos o inducibles tales como el factor alfa, la alcohol oxidasa, y PGH. Para las revisiones, véanse Ausubel et al. (1989) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, NY, y Grant et al., *Methods Enzymol.* 153: 516-544 (1997).
- 55

La introducción del constructo en la célula anfitriona se puede efectuar mediante transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, transfección mediada por lípidos catiónicos, electroporación, transducción, infección u otros métodos. Tales métodos se describen en muchos manuales de laboratorio convencionales, tales como Davis et al., BASIC METHODS IN MOLECULAR BIOLOGY (1986).

- 5 La transcripción de ADN que codifica un polipéptido de fusión de EML4-ALK de la presente descripción por eucariotas superiores se puede incrementar insertando una secuencia potenciadora en el vector. Los potenciadores son elementos de ADN que actúan en cis, generalmente de aproximadamente 10 a 300 pb que actúan para aumentar la actividad transcripcional de un promotor en un tipo dado de célula anfitriona. Los ejemplos de los potenciadores incluyen el potenciador de SV40, que está ubicado en el lado tardío del origen de replicación en los pares de bases 100 a 270, el potenciador de promotor temprano de citomegalovirus, el potenciador de polioma en el lado tardío del origen de replicación, y los intensificadores de adenovirus.

Para la secreción de la proteína traducida al lumen del retículo endoplásmico, al espacio periplásmico o al entorno extracelular, se pueden incorporar señales de secreción apropiadas al polipéptido expresado. Las señales pueden ser endógenas al polipéptido o pueden ser señales heterólogas.

- 15 El polipéptido se puede expresar en una forma modificada, tal como una proteína de fusión (p. ej., una fusión con GST), y puede incluir no sólo señales de secreción, sino también regiones funcionales heterólogas adicionales. Por ejemplo, se puede añadir una región de aminoácidos adicionales, particularmente aminoácidos cargados, al extremo N del polipéptido para mejorar la estabilidad y persistencia en la célula anfitriona, durante la purificación, o durante la posterior manipulación y almacenamiento. Asimismo, se pueden añadir radicales peptídicos al polipéptido para facilitar la purificación. Tales regiones se pueden retirar antes de la preparación final del polipéptido. La adición de radicales peptídicos a los polipéptidos para suscitar la secreción o excreción, para mejorar la estabilidad y para facilitar la purificación, entre otros, son mecanismos familiares y rutinarios en la técnica. Una proteína de fusión preferida comprende una región heteróloga de inmunoglobulina que es útil para solubilizar proteínas.

- 25 Los polipéptidos de fusión de EML4-ALK se pueden recuperar y purificar a partir de cultivos de células recombinantes por medio de métodos bien conocidos incluyendo la precipitación con sulfato de amonio o etanol, la extracción con ácido, la cromatografía de intercambio aniónico o catiónico, la cromatografía en fosfocelulosa, la cromatografía de interacción hidrófoba, la cromatografía de afinidad, la cromatografía en hidroxilapatita y la cromatografía con lectina. Lo más preferiblemente, se emplea cromatografía líquida de alta resolución ("HPLC") para la purificación. Los polipéptidos de la presente descripción incluyen productos purificados naturalmente, productos de procedimientos sintéticos químicos, y productos producidos por técnicas recombinantes a partir de un anfitrión procariontico o eucariótico, incluyendo, por ejemplo, células bacterianas, de levadura, de plantas superiores, de insectos y de mamíferos. Dependiendo del anfitrión empleado en el procedimiento de producción recombinante, los polipéptidos de la presente descripción pueden estar glicosilados o pueden no estar glicosilados. Además, los polipéptidos de la descripción también pueden incluir un residuo de metionina modificado inicial, en algunos casos como resultado de procesos mediados por el anfitrión.

- Por consiguiente, en una realización, la descripción proporciona un método para producir un polipéptido de fusión de EML4-ALK recombinante mediante el cultivo de una célula anfitriona recombinante (como se ha descrito anteriormente) en condiciones adecuadas para la expresión del polipéptido de fusión y recuperar el polipéptido. Las condiciones de cultivo adecuadas para el crecimiento de las células anfitrionas y la expresión de los polipéptidos recombinantes a partir de tales células son bien conocidas por los expertos en la técnica. Véase, p. ej., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel FM et al., Eds., Volumen 2, Capítulo 16, Wiley Interscience.

D. Polipéptidos aislados.

- La descripción también proporciona, en parte, polipéptidos de quinasa ALK mutantes aislados y fragmentos de los mismos. En una realización, la descripción proporciona un polipéptido aislado que comprende una secuencia de aminoácidos idéntica en al menos 95% a una secuencia seleccionada del grupo que consiste en:

(a) una secuencia de aminoácidos que codifica un polipéptido de fusión de EML4-ALK que comprende la secuencia de aminoácidos del SEQ ID NO: 1 o el SEQ ID NO: 18;

- (b) una secuencia de aminoácidos que codifica un polipéptido de fusión de EML4-ALK que comprende la secuencia de aminoácidos N-terminal de EML-4 (residuos 1-233 del SEQ ID NO: 3 o residuos 1-495 del SEQ ID NO: 3) y el dominio quinasa ALK (residuos 1116-1383 del SEQ ID NO: 5); y

(c) una secuencia de aminoácidos que codifica un polipéptido que comprende al menos seis aminoácidos contiguos que abarcan la unión por fusión (residuos 233-234 del SEQ ID NO: 1 o residuos 495-496 del SEQ ID NO: 18) de un polipéptido de fusión de EML4-ALK.

- En una realización preferida, se proporcionan polipéptidos de ALK mutantes recombinantes de la descripción, que pueden ser producidos utilizando un vector recombinante o célula anfitriona recombinante como se ha descrito anteriormente.

Se reconocerá en la técnica que algunas secuencias de aminoácidos del polipéptido de fusión de EML4-ALK o del polipéptido de la quinasa ALK activo truncado se pueden variar sin efecto significativo de la estructura o la función de la proteína mutante. Si se contemplan tales diferencias en la secuencia, debe recordarse que habrá zonas críticas en la proteína que determinan la actividad (p. ej., el dominio quinasa ALK). En general, es posible reemplazar los residuos que forman la estructura terciaria, siempre que se utilicen residuos que realicen una función similar. En otros casos, el tipo de residuo puede ser completamente insignificante si la alteración se produce en una región no crítica de la proteína.

Por lo tanto, la descripción incluye adicionalmente variaciones de un polipéptido de fusión de EML4-ALK que conservan la actividad de la quinasa ALK sustancial o que incluyen otras regiones de las proteínas EML-4 o ALK, tales como las porciones de proteínas comentadas a continuación. Tales mutantes incluyen delecciones, inserciones, inversiones, repeticiones y sustituciones tipo (p. ej., la sustitución de un residuo hidrófilo por otro, pero no de fuertemente hidrófilos por fuertemente hidrófobos como norma). Pequeños cambios o tales sustituciones de aminoácidos "neutros" tendrán generalmente poco efecto sobre la actividad.

Por lo general se consideran sustituciones conservativas los reemplazos, de uno por otro, entre los aminoácidos alifáticos Ala, Val, Leu e Ile; el intercambio de residuos hidroxilados Ser y Thr, el cambio de los residuos ácidos Asp y Glu, la sustitución entre los residuos amida Asn y Gln, el cambio de residuos alcalinos Lys y Arg y los reemplazos entre los residuos aromáticos Phe, Tyr. Los ejemplos de las sustituciones de aminoácidos conservativas conocidas por los expertos en la técnica son: Aromáticos: fenilalanina triptófano tirosina; Hidrófilos: leucina isoleucina valina; Polares: glutamina asparagina; Alcalinos: arginina lisina histidina; Ácidos: ácido aspártico ácido glutámico; Pequeñas: alanina serina treonina metionina glicina. Como se indica con detalle más arriba, más directrices se pueden encontrar pautas adicionales concernientes a qué cambios de aminoácidos sean fenotípicamente silenciosos (esto es, no es probable que tengan un efecto perjudicial significativo sobre una función) en Bowie *et al.*, *Science* 247, más arriba.

Los polipéptidos de la presente descripción se proporcionan preferiblemente en una forma aislada, y preferiblemente están sustancialmente purificados. Una versión producida de forma recombinante de un polipéptido de fusión de EML4-ALK de la descripción se puede purificar sustancialmente mediante el método de una etapa descrito por Smith y Johnson, *Gene* 67: 31-40 (1988).

Los polipéptidos de la presente descripción incluyen los polipéptidos de fusión de EML4-ALK de las Figuras 2A-B (SEQ ID NO: 1 y 18) (incluyan o no una secuencia líder), una secuencia de aminoácidos que codifica un polipéptido de fusión de EML4-ALK que comprende la secuencia de aminoácidos N-terminal de EML-4 (residuos 1-233 del SEQ ID NO: 3 o residuos 1 a 495 del SEQ ID NO: 3) y el dominio quinasa ALK (residuos 1116 a 1383 del SEQ ID NO: 5), y una secuencia de aminoácidos que codifica un polipéptido que comprende al menos seis aminoácidos contiguos que abarcan la unión por fusión (residuos 233-234 del SEQ ID NO: 1 o residuos 495-496 del SEQ ID NO: 18) de un polipéptido de fusión de EML4-ALK (véanse también las Figuras 1A-B, panel inferior), así como polipéptidos que tienen una similitud de al menos 90%, preferiblemente una similitud de al menos 95%, y aún más preferiblemente una similitud de al menos 96%, 97%, 98% o 99% con los descritos anteriormente.

Por "% de similitud" para dos polipéptidos se quiere significar una puntuación de similitud producida por la comparación de las secuencias de aminoácidos de los dos polipéptidos utilizando el programa Bestfit (Wisconsin Sequence Analysis Package, Versión 8 para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711) y los ajustes por defecto para determinar la similitud. Bestfit utiliza el algoritmo de homología local de Smith y Waterman (*Advances in Applied Mathematics* 2: 482-489 (1981)) para encontrar el mejor segmento de similitud entre dos secuencias.

Por polipéptido que tiene una secuencia de aminoácidos al menos, por ejemplo, "idéntica" en 95% a una secuencia de aminoácidos de referencia de un polipéptido de fusión de EML4-ALK de la descripción se quiere significar que la secuencia de aminoácidos del polipéptido es idéntica a la secuencia de referencia excepto que la secuencia del polipéptido puede incluir hasta cinco alteraciones de aminoácidos por cada 100 aminoácidos de la secuencia de aminoácidos de referencia del polipéptido de ALK mutante. En otras palabras, para obtener un polipéptido que tiene una secuencia de aminoácidos idéntica en al menos 95% a una secuencia de aminoácidos de referencia, se pueden suprimir o sustituir hasta 5% de los residuos de aminoácidos en la secuencia de referencia por otro aminoácido, o pueden insertarse en la secuencia de referencia varios aminoácidos hasta 5% de los residuos de aminoácidos totales en la secuencia de referencia. Estas alteraciones de la secuencia de referencia pueden producirse en las posiciones amino o carboxi terminales de la secuencia de aminoácidos de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas o bien individualmente entre residuos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia.

Cuando se utiliza Bestfit o cualquier otro programa de alineamiento de secuencias para determinar si una secuencia particular es, por ejemplo, idéntica en 95% a una secuencia de referencia de acuerdo con la presente descripción, los parámetros se ajustan, por supuesto, de tal manera que el porcentaje de identidad se calcula sobre la longitud completa de la secuencia de aminoácidos de referencia y que se permiten huecos en la homología de hasta 5% del número total de residuos de aminoácidos en la secuencia de referencia.

Un polipéptido de fusión de EML4-ALK de la presente descripción se puede utilizar como un marcador de peso molecular en geles de SDS-PAGE o en columnas de filtración en gel de tamiz molecular, por ejemplo, utilizando métodos bien conocidos por los expertos en la técnica.

Como se describe adicionalmente en detalle a continuación, los polipéptidos de la presente descripción también se pueden utilizar para generar reactivos específicos de polipéptidos de fusión, tales como anticuerpos policlonales y monoclonales, o reactivos específicos de polipéptidos truncados, que son útiles en análisis para detectar la expresión del polipéptido de ALK mutante como se describe a continuación, o como agonistas y antagonistas capaces de potenciar o inhibir la función/actividad de la proteína ALK mutante. Adicionalmente, tales polipéptidos se pueden utilizar en el sistema de levadura de dos híbridos para "capturar" el polipéptido de fusión de EML4-ALK o proteínas de unión al polipéptido de la quinasa ALK truncada, que son también agonistas y antagonistas candidato de acuerdo con la presente descripción. El sistema de dos híbridos de levadura es descrito por Fields y Song, *Nature* 340: 245-246 (1989).

En otro aspecto, la descripción proporciona un péptido o polipéptido que comprende una porción portadora de epítipo de un polipéptido de la descripción, por ejemplo, un epítipo que comprende la unión por fusión de un polipéptido de fusión de EML4-ALK (véanse las Figuras 1A-B, panel inferior). El epítipo de esta porción del polipéptido es un epítipo inmunogénico o antigénico de un polipéptido de la descripción. Un "epítipo inmunogénico" se define como una porción de una proteína que provoca una respuesta de anticuerpos cuando la proteína completa es el inmunógeno. Se cree que estos epítipos inmunogénicos están confinados a unos pocos loci en la molécula. Por otro lado, una región de una molécula de proteína a la que puede unirse un anticuerpo se define como un "epítipo antigénico". El número de epítipos inmunogénicos de una proteína generalmente es menor que el número de epítipos antigénicos. Véase, por ejemplo, Geysen et al., *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1983). La producción de anticuerpos específicos para el polipéptido de fusión de la descripción se describe con más detalle a continuación.

Los anticuerpos generados por péptidos o polipéptidos que portan epítipos antigénicos son útiles para detectar una proteína imitada, y se pueden utilizar anticuerpos para diferentes péptidos para el seguimiento del destino de diversas regiones de un precursor de proteína que experimenta el procesamiento post-traducciona. Los péptidos y anticuerpos anti-péptido se pueden utilizar en una variedad de análisis cualitativos o cuantitativos para determinar la proteína imitada, por ejemplo, en análisis de competición, ya que se ha demostrado que incluso los péptidos cortos (p. ej., aproximadamente 9 aminoácidos) se pueden unir y desplazar los péptidos más grandes en análisis de inmunoprecipitación. Véase, por ejemplo, Wilson et al., *Cell* 37:767-778 (1984) en 777. Los anticuerpos anti-péptido de la descripción también son útiles para la purificación de la proteína imitada, por ejemplo, mediante cromatografía de adsorción utilizando métodos bien conocidos en la técnica. Los formatos de análisis inmunológicos se describen con más detalle a continuación.

Los polipéptidos de ALK mutantes recombinantes también están dentro del alcance de la presente descripción, y se pueden producir utilizando polinucleótidos de la descripción, como se describe en la Sección B anterior. Por ejemplo, la descripción proporciona, en parte, un método para producir un polipéptido de fusión de EML4-ALK recombinante mediante el cultivo de una célula anfitriona recombinante (como se describe anteriormente) en condiciones adecuadas para la expresión del polipéptido de fusión y recuperar el polipéptido. Las condiciones de cultivo adecuadas para el crecimiento de las células anfitrionas y la expresión de los polipéptidos recombinantes a partir de tales células son bien conocidas por los expertos en la técnica.

E. Reactivos específicos de mutantes

Los reactivos específicos del polipéptido de ALK mutante útiles en la práctica de los métodos descritos incluyen, entre otros, anticuerpos específicos de polipéptidos de fusión y péptidos AQUA (péptidos marcados con isótopos pesados) que corresponden a, y adecuados para la detección y cuantificación de, la expresión del polipéptido de fusión de EML4-ALK en una muestra biológica de un cáncer, tal como un tumor sarcoma o carcinoma sólido de mamífero. También son útiles los reactivos específicos de truncamiento, tales como anticuerpos, péptidos AQUA, o sondas de ácidos nucleicos, adecuados para la detección de la presencia o ausencia de un polinucleótido o polipéptido de quinasa ALK truncado de la descripción. Un reactivo específico de polipéptido de fusión es cualquier reactivo, biológico o químico, capaz de unirse específicamente a, detectar y/o cuantificar la presencia/nivel de polipéptido de fusión de EML4-ALK expresado en una muestra biológica. El término incluye, pero no se limita a, los reactivos de anticuerpos y péptidos AQUA preferidos mencionados a continuación, y reactivos equivalentes están dentro del alcance de la presente descripción.

Anticuerpos.

Los reactivos adecuados para su uso en la práctica de los métodos de la descripción incluyen un anticuerpo específico del polipéptido de fusión de EML4-ALK y un anticuerpo específico del polipéptido de fusión de TFG-ALK. Un anticuerpo específico de una fusión de la descripción es uno o varios anticuerpos aislado que se unen específicamente a un polipéptido de fusión de EML4-ALK de la descripción (p. ej., SEQ ID NO: 1) pero no se unen sustancialmente a EML-4 de tipo salvaje o a ALK de tipo salvaje, o se unen específicamente a un polipéptido de fusión de TFG-ALK descrito en la presente memoria (p. ej., SEQ ID NO: 20), pero no se unen sustancialmente a

TFG de tipo salvaje o ALK de tipo salvaje. Otros reactivos adecuados incluyen anticuerpos específicos de epítomos que se unen específicamente a un epítomo en el dominio extracelular de la secuencia de la proteína ALK de tipo salvaje (cuyo dominio no está presente en la quinasa ALK activa, troncada descrita en la presente memoria), y por lo tanto son capaces de detectar la presencia (o ausencia) de ALK de tipo salvaje en una muestra.

5 Los anticuerpos específicos del polipéptido de fusión de EML4-ALK o TFG-ALK humanos también pueden unirse a secuencias peptídicas epitópicas altamente homólogas y equivalentes en otras especies de mamíferos, por ejemplo murinos o de conejo, y viceversa. Los anticuerpos útiles en la práctica de los métodos de la descripción incluyen (a) anticuerpos monoclonales, (b) anticuerpos policlonales purificados que se unen específicamente al polipéptido diana (p. ej., la unión por fusión del polipéptido de fusión de EML4-ALK (véanse las Figuras 1A-B, panel inferior) o el
10 polipéptido de fusión de TFG-ALK (véase la Figura 1C, panel inferior), (c) anticuerpos como los descritos en los apartados (a) - (b) anteriores que se unen epítomos o sitios de fosforilación equivalentes y altamente homólogas en otras especies no humanas (p. ej., ratón, rata), y (d) fragmentos de los apartados (a) - (c) anteriores que se unen al antígeno (o más preferiblemente al epítomo) unido mediante los anticuerpos ilustrativos descritos en la presente memoria.

15 El término "anticuerpo" o "anticuerpos", según se utiliza en la presente memoria se refiere a todos los tipos de inmunoglobulinas, incluyendo IgG, IgM, IgA, IgD, e IgE. Los anticuerpos pueden ser monoclonales o policlonales y pueden ser de cualquier especie de origen, incluyendo (por ejemplo) ratón, conejo, caballo, o ser humano, o pueden ser anticuerpos quiméricos. Véanse, p. ej., M. Walker et al., *Molec. Immunol.* 26: 403-11 (1989); Morrison et al., *Proc. Natl. Acad. Sci.* 81: 6851 (1984); Neuberger et al., *Nature* 312:604 (1984)). Los anticuerpos pueden ser
20 anticuerpos monoclonales recombinantes producidos de acuerdo con los métodos descritos en la Patente de los Estados Unidos Núm. 4.474.893 (Reading) o la Patente de los Estados Unidos Núm. 4.816.567 (Cabilly et al.). Los anticuerpos también pueden ser anticuerpos específicos construidos químicamente preparados de acuerdo con el método descrito en la Patente de los Estados Unidos Núm. 4.676.980 (Segel et al.).

25 El sitio epitópico preferido de un anticuerpo específico de un polipéptido de fusión de EML4-ALK de la descripción es un fragmento peptídico que consiste esencialmente en aproximadamente 11 a 17 aminoácidos de una secuencia del polipéptido de fusión de EML4-ALK humano (SEQ ID NO: 1 y 18) cuyo fragmento abarca la unión por fusión (que se produce en los residuos 233-234 en la proteína de fusión de variante corta y los residuos 495-496 en la proteína de fusión de variante larga (véanse las Figuras 1A-B (panel inferior)). Se apreciará que los anticuerpos que se unen específicamente péptidos/epítomos más cortos o más largos que abarcan la unión de fusión de un polipéptido de
30 fusión de EML4-ALK están dentro del alcance de la presente descripción.

Del mismo modo, el sitio epitópico preferido de un anticuerpo específico del polipéptido de fusión de TFG-ALK útil en la práctica de los métodos descritos es un fragmento peptídico que consiste esencialmente en aproximadamente 11 a 17 aminoácidos de la secuencia del polipéptido de fusión de TFG-ALK humano (SEQ ID NO: 20), cuyo fragmento abarca la unión por fusión (que se produce en los residuos 137-138 (véase la Figura 1C (panel inferior)).

35 La descripción no se limita al uso de anticuerpos, sino que incluye moléculas equivalentes, tales como dominios de unión a proteínas o aptámeros de ácidos nucleicos, que se unen, de una manera específica de la proteína de fusión o la proteína troncada, esencialmente al mismo epítomo al que se une el anticuerpo específico del polipéptido de fusión de EML4-ALK o TFG-ALK útil en los métodos de la descripción. Véase, p. ej., Neuberger et al., *Nature* 312: 604 (1984). Tales reactivos distintos de anticuerpos equivalentes se pueden emplear adecuadamente en los
40 métodos de la descripción descritos adicionalmente más adelante.

Los anticuerpos policlonales útiles en la práctica de los métodos de la descripción se pueden producir de acuerdo con técnicas convencionales inmunizando un animal adecuado (p. ej., conejo, cabra, etc.) con un antígeno que incluye un epítomo específico de la proteína de fusión deseada (p. ej., la unión por fusión de una proteína de fusión de ALK descrita en la presente memoria), recogiendo el suero inmune del animal, y separando los anticuerpos
45 policlonales del suero inmune, y purificando los anticuerpos policlonales que tienen la especificidad deseada, de acuerdo con procedimientos conocidos. El antígeno puede ser un antígeno peptídico sintético que comprende la secuencia epitópica deseada, seleccionado y construido de acuerdo con técnicas bien conocidas. Véanse, p. ej., ANTIBODIES: A LABORATORY MANUAL, capítulo 5, págs. 75-76, Harlow y Lane Eds., Cold Spring Harbor Laboratory (1988); Czernik, *Methods in Enzymology*, 201: 264-283 (1991); Merrifield, *J. Am. Chem. Soc.* 85: 21-49 (1962)). Los anticuerpos policlonales producidos como se describe en la presente memoria se pueden escutar y
50 aislar como se describe adicionalmente más adelante.

Los anticuerpos monoclonales también se pueden emplear de manera beneficiosa en los métodos de la descripción, y se pueden producir en líneas celulares de hibridoma de acuerdo con la técnica bien conocida de Kohler y Milstein. *Nature* 265: 495-97 (1975); Kohler y Milstein, *Eur. J. Immunol.* 6: 511 (1976); véase también, CURRENT
55 PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel et al. Eds. (1989). Los anticuerpos monoclonales producidos de este modo son altamente específicos, y mejoran la selectividad y especificidad de los métodos de análisis proporcionados por la descripción. Por ejemplo, se puede inyectar una disolución que contiene el antígeno apropiado (p. ej., un péptido sintético que comprende la unión por fusión de polipéptido de fusión de EML4-ALK) en un ratón y, después de un tiempo suficiente (de acuerdo con técnicas convencionales), el ratón se sacrifica y se
60 obtienen las células del bazo obtenidas. A continuación las células del bazo se immortalizan fusionándolas con

células de mieloma, típicamente en presencia de polietilenglicol, para producir células de hibridoma. Los hibridomas de fusión de conejo, por ejemplo, se pueden producir como se describe en la Patente de los Estados Unidos Núm. 5.675.063, K. Knight, Expedida el 7 de Octubre de 1997. Las células de hibridoma se cultivan a continuación en un medio de selección adecuado, tal como hipoxantina-aminopterina-timidina (HAT), y el sobrenadante se escruta para determinar los anticuerpos monoclonales que tienen la especificidad deseada, como se describe a continuación. El anticuerpo secretado se puede recuperar del sobrenadante de cultivo de tejidos mediante métodos convencionales tales como precipitación, cromatografía de intercambio iónico o de afinidad, o similares.

Los fragmentos Fab monoclonales también se pueden producir en *Escherichia coli* mediante mecanismos recombinantes conocidos por los expertos en la técnica. Véanse, p. ej., W. Huse, *Science* 246: 1275-81 (1989); Mullinax et al., *Proc. Natl. Acad. Sci.* 87: 8095 (1990). Si se prefieren anticuerpos monoclonales de un isotipo para una aplicación concreta, se pueden preparar los isotipos concretos directamente, mediante selección a partir de la fusión inicial, o se pueden preparar secundariamente, a partir de un hibridoma parental que secreta un anticuerpo monoclonal de isotipo diferente mediante el uso de la técnica de selección sib para aislar variantes de cambio de clase (Steplewski, et al., *Proc. Natl. Acad. Sci.*, 82: 8653 (1985); Spira et al., *J. Immunol. Methods*, 74: 307 (1984)). El sitio de combinación con el antígeno del anticuerpo monoclonal se puede clonar mediante PCR y producir anticuerpos de cadena sencilla como anticuerpos recombinantes presentados en fagos o anticuerpos solubles en *E. coli* (véase, p. ej., ANTIBODY ENGINEERING PROTOCOLOS, 1995, Humana Press, Sudhir Paul editor.)

Más aún, la Patente de los Estados Unidos Núm. 5.194.392, Geysen (1990) describe un método general para detectar o determinar la secuencia de monómeros (aminoácidos u otros compuestos) que es un equivalente topológico del epítipo (es decir, un "mimótipo") que es complementaria a un parátipo concreto (sitio de unión al antígeno) de un anticuerpo de interés. Más en general, este método implica la detección o determinación de una secuencia de monómeros que es un equivalente topográfico de un ligando que es complementario al sitio de unión de ligando de un receptor concreto de interés. Del mismo modo, la Patente de los Estados Unidos Núm. 5.480.971, de Houghten et al. (1996) describe oligopéptidos peralquilados con alquilo C₁-C lineal y conjuntos y bibliotecas de tales péptidos, así como métodos para utilizar tales conjuntos y bibliotecas de oligopéptidos para determinar la secuencia de un oligopéptido peralquilado que se une preferentemente a una molécula aceptora de interés. De este modo, los análogos no peptídicos de los péptidos portadores de epítopos de la descripción también se pueden elaborar de forma rutinaria por medio de estos métodos.

Los anticuerpos útiles en los métodos de la descripción, ya sean policlonales o monoclonales, se pueden escrutar para determinar la especificidad del epítipo y la proteína de fusión de acuerdo con mecanismos convencionales. Véase, p. ej., Czernik et al, *Methods in Enzymology*, 201: 264-283 (1991). Por ejemplo, los anticuerpos se pueden escrutar frente a una biblioteca de péptidos mediante ELISA para asegurar la especificidad tanto para el antígeno deseado como, si se desea, para la reactividad solamente con, p. ej., un polipéptido de fusión de EML4-ALK de la descripción y no con EML-4 de tipo salvaje o ALK de tipo salvaje. Los anticuerpos también se pueden someter a ensayo mediante transferencia Western frente a preparaciones celulares que contienen la proteína diana para confirmar la reactividad con la única diana deseada y para asegurar que no hay unión apreciable a otras proteínas de fusión que implican ALK. La producción, escrutinio, y uso de anticuerpos específicos para la proteína de fusión son conocidos por los expertos en la técnica, y han sido descritos. Véase, p. ej., la Publicación de Patente de los Estados Unidos Núm. 20050214301, Wetzal et al., 29 de Septiembre de 2005.

Los anticuerpos específicos para el polipéptido de fusión útiles en los métodos de la descripción pueden exhibir alguna reactividad cruzada limitada con epítopos de fusión similares en otras proteínas de fusión o con los epítopos de EML-4 de tipo salvaje, TFG de tipo salvaje, y ALK de tipo salvaje que forman la unión por fusión. Esto no es inesperado ya que la mayoría de los anticuerpos muestran algún grado de reactividad cruzada, y los anticuerpos anti-péptido a menudo reaccionarán de forma cruzada con epítopos que tienen alta homología o identidad con el péptido inmunizante. Véase, p. ej., Czernik, más arriba. La reactividad cruzada con otras proteínas de fusión se caracteriza fácilmente por transferencia Western junto con marcadores de peso molecular conocido. Las secuencias de aminoácidos de las proteínas de reacción cruzada se pueden examinar para identificar los sitios altamente homólogos o idénticos a la secuencia de polipéptido de fusión de EML4-ALK o TFG-ALK a la que se une el anticuerpo. La reactividad cruzada no deseable se puede eliminar mediante selección negativa utilizando purificación de anticuerpos en columnas peptídicas (p. ej., seleccionando anticuerpos que se unen, a EML-4 de tipo salvaje y/o a ALK de tipo salvaje).

Los anticuerpos específicos del polipéptido de fusión de EML4-ALK de la descripción (y los anticuerpos específicos del polipéptido de fusión de TFG-ALK) que son útiles en la práctica de los métodos descritos en la presente memoria son idealmente específicos para el polipéptido de fusión humano, pero no se limitan solo a la unión a la especie humana, *per se*. La descripción incluye la producción y uso de anticuerpos que también se unen a epítopos conservados y altamente homólogos o idénticos en otras especies de mamíferos (p. ej., ratón, rata, mono). Las secuencias altamente homólogas o idénticas en otras especies pueden identificarse fácilmente mediante comparaciones de secuencias patrón, por ejemplo utilizando BLAST, con una secuencia de polipéptido de fusión de EML4-ALK humano descrita en la presente memoria (SEQ ID NO: 1 y 18) o una secuencia del polipéptido de fusión de TFG-ALK humano descrita en la presente memoria (SEQ ID NO: 20).

Los anticuerpos empleados en los métodos de la descripción se pueden caracterizar adicionalmente, y validar, para

su uso en un formato de análisis concreto, por ejemplo citometría de flujo (FC), inmunohistoquímica (IHC), y/o inmunocitoquímica (ICC). El uso de anticuerpos específicos para el polipéptido de fusión de ALK en tales métodos se describe adicionalmente en la sección F a continuación. Los anticuerpos también pueden conjugarse ventajosamente con colorantes fluorescentes (p. ej., Alexa 488, PE), o marcas tales como puntos cuánticos, para su uso en análisis multiparamétricos junto con otros anticuerpos de transducción de señales (fosfo-AKT, fosfo-Erk 1/2) y/o marcadores celulares (citoqueratina), como se describe adicionalmente en la sección F a continuación.

A poner en práctica los métodos de la descripción, también se pueden examinar ventajosamente la expresión y/o actividad de EML-4 de tipo salvaje, TFG de tipo salvaje, y/o ALK de tipo salvaje en una muestra biológica dada utilizando anticuerpos (ya sea fosfoespecíficos o totales) para estas proteínas de tipo salvaje. Por ejemplo, los anticuerpos específicos totales y del sitio de fosforilación de ALK están disponibles en el mercado (véase CELL SIGNALING TECHNOLOGY, INC., Beverly MA, 2005/06 Núm. de Catálogo 3341, 3342). Tales anticuerpos también pueden producirse de acuerdo con métodos convencionales, como se ha descrito anteriormente. Las secuencias de aminoácidos de EML-4, TFG, y ALK humanos están publicadas (véanse las Figuras 3A y 4A-4C, y los Núm. de acceso SwissProt referenciados), como lo están las secuencias de estas proteínas de otras especies.

La detección de la expresión y/o activación de EML-4 de tipo salvaje, TFG, ALK de tipo salvaje, junto con la expresión del polipéptido de fusión de EML4-ALK y/o TFG-ALK, en una muestra biológica (p. ej., una muestra tumoral) puede proporcionar información acerca de si la proteína de fusión por sí sola está dirigiendo el tumor, o si ALK de tipo salvaje también está activada y dirige el tumor. Tal información tiene utilidad clínica para evaluar si es probable que la elección como diana de la proteína de fusión o la proteína o proteínas de tipo natural, o ambas, sea lo más beneficioso para inhibir la progresión del tumor, y para seleccionar un agente terapéutico apropiado o una de sus combinaciones. Los anticuerpos específicos para el dominio extracelular de la quinasa ALK de tipo salvaje, que no está presente en la quinasa ALK activa truncada descrita en la presente memoria, pueden ser particularmente útiles para determinar la presencia/ausencia de la quinasa ALK mutante.

Se entenderá que se puede utilizar más de un anticuerpo en la práctica de los métodos descritos anteriormente. Por ejemplo, se pueden emplear simultáneamente uno o más anticuerpos específicos para el polipéptido de fusión de EML4-ALK junto con uno o más anticuerpos específicos para otra quinasa, receptor, o sustrato de quinasa que se sospecha que está, o potencialmente está, activado en un cáncer en el es expresado el polipéptido de fusión de EML4-ALK para detectar la actividad de tales otras moléculas de señalización en una muestra biológica que comprende células de tal cáncer.

Los expertos en la técnica apreciarán que los polipéptidos de fusión de EML4-ALK de la presente descripción y los fragmentos portadores de epítomos de unión por fusión de los mismos descritos anteriormente pueden combinarse con porciones del dominio constante de inmunoglobulinas (IgG), dando como resultado polipéptidos quiméricos. Estas proteínas de fusión facilitan la purificación y muestran un incremento de la vida media *in vivo*. Esto se ha demostrado, p. ej., para proteínas quiméricas que consisten en los dos primeros dominios del polipéptido CD4 humano y diversos dominios de las regiones constantes de las cadenas pesadas o ligeras de inmunoglobulinas de mamíferos (EPA 394.827; Traunecker et al., *Nature* 331: 84-86 (1988)). Las proteínas de fusión que tienen una estructura dimérica unida por puentes disulfuro debido a la porción IgG también pueden ser más eficaces en la unión y neutralización de otras moléculas que el polipéptido de fusión de EML4-ALK monomérico solo (Fountoulakis et al., *J. Biochem* 270: 3958-3964 (1995)).

Péptidos marcados con isótopos pesados (Péptidos AQUA)

Los reactivos específicos para el polipéptido fusión de EML4-ALK o TFG-ALK útiles en la práctica de los métodos descritos también pueden comprender péptidos marcados con isótopos pesados adecuados para la cuantificación absoluta del polipéptido de fusión de ALK expresado o el polipéptido de quinasa ALK truncado en una muestra biológica. Se han descrito la producción y el uso de péptidos AQUA para la cuantificación absoluta de proteínas (AQUA) en mezclas complejas. Véanse WO/03016861, "Absolute Quantification of Proteins and Modified Forms Thereof by Multistage Mass Spectrometry", Gygi et al. y también Gerber et al. *Proc. Natl. Acad. Sci. USA* 100: 6940-5 (2003) (cuyas enseñanzas se incorporan en la presente memoria por referencia, en su totalidad).

La metodología AQUA emplea la introducción de una cantidad conocida de al menos un patrón peptídico marcado con isótopo pesado (que tiene una firma única detectable por cromatografía LC-SRM) en una muestra biológica digerida con el fin de determinar, por comparación con el patrón de péptido, la cantidad absoluta de un péptido con la misma secuencia y modificación de proteína en la muestra biológica. En pocas palabras, la metodología AQUA tiene dos fases: selección del patrón interno peptídico y la validación y desarrollo de métodos; y la aplicación utilizando los patrones internos peptídicos validados para detectar y cuantificar una proteína diana en la muestra. El método es una técnica poderosa para detectar y cuantificar un péptido/proteína dados dentro de una mezcla biológica compleja, tal como un producto lisado celular, y se puede emplear, p. ej., para cuantificar el cambio en la fosforilación de proteínas como resultado del tratamiento con fármacos, o para cuantificar diferencias en el nivel de una proteína en diferentes estados biológicos.

Generalmente, para desarrollar un patrón interno adecuado, se selecciona un péptido concreto (o péptido modificado) dentro de una secuencia de proteína diana basándose en su secuencia de aminoácidos y en la proteasa

concreta que se va a utilizar para la digestión. A continuación el péptido se genera mediante síntesis de péptidos en fase sólida de manera que un residuo es remplazado por ese mismo residuo que contiene isótopos estables C^{13} , N^{15}). El resultado es un péptido que es químicamente idéntico a su contraparte nativa formada mediante proteólisis, pero es fácilmente distinguible mediante MS a través de un desplazamiento de masas de 7 Da. El péptido patrón interno AQUA recién sintetizado se evalúa a continuación mediante LC-MS/MS. Este procedimiento proporciona información cualitativa sobre la retención de péptido mediante cromatografía de fase inversa, eficacia de ionización, y fragmentación por medio de disociación inducida por colisión. Los iones de los fragmentos informativos y abundantes para los conjuntos de péptidos nativos y patrones internos se seleccionan y a continuación se verifican específicamente en una sucesión rápida como una función de la retención cromatográfica para formar un método de verificación de la reacción seleccionada (LC-SRM) basándose en el perfil único del patrón peptídico.

La segunda fase de la estrategia AQUA es su implementación para medir la cantidad de una proteína o proteína modificada a partir de mezclas complejas. Los productos lisados celulares completos son fraccionados típicamente por medio de electroforesis en gel de SDS-PAGE, y las regiones del gel coincidentes con la migración de proteína son escindidas. Este procedimiento está seguido por la proteólisis en gel en presencia de los péptidos AQUA y análisis LC-SRM. (véase Gerber *et al.* más arriba). Los péptidos AQUA se agregan a la mezcla de péptidos compleja obtenida por digestión del producto lisado celular completo con una enzima proteolítica y se someten a purificación por inmovinoafinidad como se ha descrito anteriormente. El tiempo de retención y patrón de fragmentación del péptido nativo formado por digestión (p. ej., tripsinización) son idénticos a los del péptido patrón interno AQUA determinado previamente; de este modo, el análisis LC-MS/MS utilizando un experimento SRM da como resultado una medición altamente específica y sensible tanto del patrón interno como del analito directamente a partir de las mezclas de péptidos extremadamente complejas.

Puesto que se añade una cantidad absoluta del péptido AQUA (p. ej., 250 fmol), la razón de las áreas bajo la curva se puede utilizar para determinar los niveles de expresión precisos de una proteína o forma fosforilada de una proteína en el producto lisado celular original. Además, el patrón interno está presente durante la digestión en el gel a medida que se forman los péptidos nativos, de manera que la eficacia de extracción del péptido de los pedazos de gel, las pérdidas absolutas durante la manipulación de la muestra (incluyendo la centrifugación a vacío), y la variabilidad durante la introducción en el sistema LC-MS no afectan a la razón determinada de abundancias de péptidos nativos y AQUA.

Un patrón de péptido AQUA se desarrolla para una secuencia conocida previamente identificada por medio del método IAP-LC-MS/MS en una proteína diana. Si el sitio es modificado, se puede desarrollar un péptido AQUA que incorpora la forma modificada del residuo concreto dentro del sitio, y se desarrolla un segundo péptido AQUA que incorpora la forma no modificada del residuo. De esta manera, se pueden utilizar los dos patrones para detectar y cuantificar las formas del sitio tanto modificadas como no modificadas en una muestra biológica.

Patrones internos peptídicos también se pueden generar examinando la secuencia de aminoácidos primaria de una proteína y determinando los límites de los péptidos producidos por la escisión con proteasa. Alternativamente, se puede digerir realmente una proteína con una proteasa y a continuación se puede secuenciar un fragmento peptídico concreto producido. Las proteasas adecuadas incluyen, pero no se limitan a, serina proteasas (p. ej., tripsina, hepsina), metaloproteasas (p. ej., PUMP1), quimotripsina, catepsina, pepsina, termolisina, carboxipeptidasas, etc.

Una secuencia peptídica dentro de una proteína diana se selecciona de acuerdo con uno o más criterios para optimizar el uso del péptido como patrón interno. Preferiblemente, se selecciona el tamaño del péptido para reducir al mínimo las posibilidades de que la secuencia peptídica se repita en otros lugares de otras proteínas distintas de la diana. De este modo, un péptido tiene preferiblemente al menos aproximadamente 6 aminoácidos. El tamaño del péptido también se optimiza para maximizar la frecuencia de ionización. Por lo tanto, no se prefieren péptidos más largos de aproximadamente 20 aminoácidos. El intervalo preferido es de aproximadamente 7 a 15 aminoácidos. También se selecciona una secuencia peptídica que no es probable que sea químicamente reactiva durante la espectrometría de masas, por lo tanto se evitan secuencias que comprenden cisteína, triptófano, o metionina.

Se puede seleccionar una secuencia peptídica que no incluye una región modificada de la región diana de manera que el patrón interno peptídico se pueda utilizar para determinar la cantidad de todas las formas de la proteína. Alternativamente, puede ser deseable detectar un patrón interno peptídico que abarca un aminoácido modificado y cuantificar solamente la forma modificada de la proteína diana. Se pueden utilizar juntos patrones peptídicos para regiones tanto modificadas como no modificadas, para determinar el alcance de una modificación en una muestra particular (es decir, para determinar qué fracción de la cantidad total de proteína está representada por la forma modificada). Por ejemplo, se pueden usar patrones peptídicos para la forma tanto fosforilada como no fosforilada de una proteína que se sabe que está fosforilada en un sitio concreto para cuantificar la cantidad de la forma fosforilada en una muestra.

El péptido se marca utilizando uno o más aminoácidos marcados (es decir, la marca es una parte real del péptido) o menos preferiblemente, las marcas se pueden anclar después de la síntesis de acuerdo con métodos convencionales. Preferiblemente, la marca es una marca que altera la masa seleccionada basándose en las siguientes consideraciones: La masa debe ser única con respecto a las masas de los fragmentos desplazados

producidos mediante análisis de MS a las regiones del espectro con fondo bajo; el componente de firma de la masa iónica es la porción del radical de marcaje que muestra preferiblemente una firma de masa iónica única en el análisis MS; la suma de las masas de los átomos constituyentes de la marca es preferiblemente única y diferente de los fragmentos de todos los posibles aminoácidos. Como resultado, los aminoácidos y péptidos marcados se distinguen fácilmente de los no marcados por el patrón de masa iónica en el espectro de masas resultante. Preferiblemente, el componente de firma de masa iónica confiere una masa a un fragmento de proteína que no coincide con la masa del residuo para ninguno de los 20 aminoácidos naturales.

La marca debe ser robusta en las condiciones de fragmentación del MS y no experimentar una fragmentación desfavorable. La química de marcaje debe ser eficaz en un intervalo de condiciones, particularmente condiciones de desnaturalización, y la etiqueta marcada permanece preferiblemente soluble en el sistema de tampón de MS de elección. La marca no suprime preferiblemente la eficacia de ionización de la proteína y no es químicamente reactiva. La marca puede contener una mezcla de dos o más especies isotópicamente distintas para generar un patrón de espectrometría de masas único en cada posición de fragmento marcado. Los isótopos estables, tales como ^2H , ^{13}C , ^{15}N , ^{17}O , ^{18}O , o ^{34}S , se encuentran entre las marcas preferidas. También se pueden preparar pares de patrones internos peptídicos que incorporan una marca isotópica diferente. Los residuos de aminoácidos preferidos en los que se puede incorporar una marca isotópica pesada incluyen leucina, prolina, valina, y fenilalanina.

Los patrones internos peptídicos se caracterizan de acuerdo con su razón de masa a carga (m/z), y preferiblemente, también de acuerdo con su tiempo de retención en una columna cromatográfica (p. ej., una columna de HPLC). Los patrones internos que eluyen simultáneamente con péptidos no marcados de secuencia idéntica se seleccionan como patrones internos óptimos. El patrón interno se analiza a continuación mediante la fragmentación del péptido por cualquier medio adecuado, por ejemplo, mediante disociación inducida por colisión (CID) utilizando, p. ej., argón o helio como gas de colisión. Los fragmentos se analizan a continuación, por ejemplo, mediante espectrometría de masas de múltiples etapas (MS^n) para obtener un espectro de iones fragmento, para obtener una firma de fragmentación peptídica. Preferiblemente, los fragmentos peptídicos tienen diferencias significativas en las razones de m/z para permitir que los picos correspondientes a cada fragmento se separen bien, y se obtenga una firma que sea única para el péptido diana. Si no se obtiene una firma de fragmento adecuada en la primera etapa, se realizan etapas adicionales de MS hasta que se obtiene una firma única.

Los iones fragmento en los espectros MS/MS y MS^3 son típicamente altamente específicos para el péptido de interés, y, junto con los métodos LC, permiten un medio altamente selectivo para detectar y cuantificar un péptido/proteína diana en una mezcla compleja de proteínas, tal como un producto lisado celular, que contiene muchos miles o decenas de miles de proteínas. Se puede analizar cualquier muestra biológica que contenga potencialmente una proteína/péptido diana de interés. Se emplean preferiblemente extractos celulares brutos o parcialmente purificados. Generalmente, la muestra tiene al menos 0,01 mg de proteína, típicamente una concentración de 0,1-10 mg/ml, y se puede ajustar a una concentración de tampón y pH deseados.

A continuación se añade una cantidad conocida de un patrón interno del péptido marcado, preferiblemente aproximadamente 10 femtomoles, correspondiente a una proteína diana que se va a detectar/cuantificar a una muestra biológica, tal como un producto lisado celular. La muestra enriquecida se digiere a continuación con una o más proteasas durante un período de tiempo adecuado para permitir la digestión. Después se realiza una separación (p. ej., mediante HPLC, HPLC de fase inversa, electroforesis capilar, cromatografía de intercambio iónico, etc.) para aislar el patrón interno marcado y su péptido diana correspondiente de otros péptidos en la muestra. La LC por microcapilaridad es un método preferido.

Cada péptido aislado se examina a continuación mediante el seguimiento de una reacción seleccionada en el MS. Esto implica el uso del conocimiento previo obtenido mediante la caracterización del patrón interno peptídico y requiere después el MS para controlar continuamente un ión específico en el espectro MS/MS o MS^n tanto para el péptido de interés como para el patrón interno. Después de la elución, se calculan el área bajo la curva (AUC) para los picos tanto del patrón peptídico como del péptido diana. La razón de las dos áreas proporciona la cuantificación absoluta que se puede normalizar para el número de células utilizado en el análisis y el peso molecular de la proteína, para proporcionar el número exacto de copias de la proteína por célula. Los detalles adicionales de la metodología AQUA son descritos por Gygi et al., y Gerber et al. más arriba.

Los patrones peptídicos internos AQUA (péptidos marcados con isótopos pesados) se pueden producir de manera deseable, como se ha descrito anteriormente, para detectar y cuantificar cualquier sitio único (p. ej., la unión por fusión dentro de un polipéptido de fusión de EML4-ALK descrito) en de un polipéptido de ALK mutante de la descripción. Por ejemplo, se puede preparar un fosfopéptido AQUA que corresponde a la secuencia de unión por fusión de un polipéptido de fusión de EML4-ALK (véanse las Figuras 1A-B (panel inferior)) o que corresponde al punto de truncamiento de cualquiera de EML4, TFG, o ALK. Se pueden producir patrones peptídicos para la unión por fusión de EML4-ALK o TFG-ALK y se pueden emplear tales patrones en la metodología AQUA para detectar y cuantificar la unión por fusión (es decir, la presencia de polipéptido de fusión de EML4-ALK) en una muestra biológica.

Por ejemplo, un péptido AQUA ilustrativo de la descripción comprende la secuencia de aminoácidos INQVYR (véase la Figura 1, panel inferior), que corresponde a los tres aminoácidos que flanquean inmediatamente cada lado de la

unión por fusión en el polipéptido de fusión de EML4-ALK (véase el SEQ ID NO: 7). Se apreciará que también se pueden construir péptidos AQUA más grandes que comprenden la secuencia de unión por fusión (y residuos adicionales aguas abajo o aguas arriba de la misma). Del mismo modo, se puede construir alternativamente un péptido AQUA más pequeño que comprende menos de la totalidad de los residuos de tal secuencia (pero que todavía comprende el punto de la propia unión por fusión). Tales péptidos AQUA más grandes o más pequeños están dentro del alcance de la presente descripción, y la selección y producción de péptidos AQUA preferidos se pueden llevar a cabo como se describió anteriormente (véase Gygi *et al.*, Gerber *et al.*, más arriba).

Sondas de ácido nucleico

Los reactivos de fusión específica proporcionados por la descripción también incluyen sondas y cebadores de ácidos nucleicos adecuados para la detección de un polinucleótido de EML4-ALK o un polinucleótido de quinasa ALK truncado, como se describe con detalle en la Sección B anterior. Tales sondas deseables incluyen, entre otras, sondas de punto de rotura que corresponden a ambos lados de los puntos de rotura de los genes de EML4 de tipo salvaje y/o ALK de tipo salvaje que producen la fusión. El uso específico de tales sondas en análisis tales como la hibridación fluorescente *in situ* (FISH) o la amplificación mediante reacción en cadena de la polimerasa (PCR) se describe en la sección F de más abajo. Se pueden preparar sondas de punto de ruptura similares para detectar la presencia del polinucleótido de fusión de TFG-ALK (véase la Figura 1C (SEQ ID NO: 21)).

F. Aplicaciones de diagnóstico y formatos de análisis.

Los métodos de la descripción se pueden llevar a cabo en una variedad de formatos de análisis diferentes conocidos por los expertos en la técnica.

Inmunoanálisis

Los inmunoanálisis útiles en la práctica de los métodos de la descripción pueden ser inmunoanálisis homogéneos o inmunoanálisis heterogéneos. En un análisis homogéneo la reacción inmunológica normalmente implica un reactivo específico del polipéptido quinasa ALK mutante (p. ej., un anticuerpo específico del polipéptido de fusión de EML4-ALK), un analito marcado, y la muestra biológica de interés. La señal que surge de la marca se modifica, directa o indirectamente, tras la unión del anticuerpo al analito marcado. Tanto la reacción inmunológica como la detección del grado de la misma se llevan a cabo en una solución homogénea. Las marcas inmunoquímicas que se pueden emplear incluyen radicales libres, radioisótopos, colorantes fluorescentes, enzimas, bacteriófagos, coenzimas, etcétera. También se pueden emplear ventajosamente marcas de nanocristales semiconductores, o "puntos cuánticos", y su preparación y uso han sido bien descritos. Véase, en general, K. Barovsky, *Nanotech. Law y Bus.* 1(2): Artículo 14 (2004) y patentes allí citadas.

En un enfoque de análisis heterogéneo, los reactivos son por lo general la muestra biológica, un reactivo específico del polipéptido de quinasa ALK mutante (p. ej., un anticuerpo específico de la fusión de EML4-ALK), y medios adecuados para producir una señal detectable. Se pueden utilizar muestras biológicas descritas adicionalmente a continuación. El anticuerpo generalmente se inmoviliza sobre un soporte, tal como una cuenta, placa o portaobjetos, y se pone en contacto con la muestra que se sospecha que contiene el antígeno en una fase líquida. El soporte se separa después de la fase líquida y se examina ya sea la fase de soporte o la fase líquida para determinar una señal detectable empleando los métodos para producir tal señal. La señal está relacionada con la presencia del analito en la muestra biológica. Los métodos para producir una señal detectable incluyen el uso de marcas radiactivas, marcas fluorescentes, marcas enzimáticas, puntos cuánticos, etcétera. Por ejemplo, si el antígeno que se va a detectar contiene un segundo sitio de unión, se puede conjugar un anticuerpo que se une a ese sitio con un grupo detectable y se puede añadir a la solución de reacción en fase líquida antes de la etapa de separación. La presencia del grupo detectable sobre el soporte sólido indica la presencia del antígeno en la muestra de ensayo. Los ejemplos de inmunoanálisis adecuados son radioinmunoanálisis, métodos de inmunofluorescencia, inmunoanálisis ligados a enzimas, y similares.

Los formatos de inmunoanálisis y variaciones de los mismos, que pueden ser útiles para llevar a cabo los métodos descritos en la presente memoria, son bien conocidos en la técnica. Véase, en general E. Maggio, *Enzyme-Immunoassay*, (1980) (CRC Press, Inc., Boca Raton, Fla.); véanse también, p. ej., la Patente de los Estados Unidos Núm. 4.727.022 (Skold *et al.*, "Methods for Modulating Ligand-Receptor Interactions and their Application"); la Patente de los Estados Unidos Núm. 4.659.678 (Forrest *et al.*, "Immunoassay of Antigens"); la Patente de los Estados Unidos Núm. 4.376.110 (David *et al.*, "Immunometric Assays Using Monoclonal Antibodies"). Las condiciones adecuadas para la formación de complejos de reactivo-anticuerpo son bien conocidas por los expertos en la técnica. Véase *id.* Se pueden utilizar anticuerpos monoclonales específicos del polipéptido de fusión de EML4-ALK en un análisis "de dos sitios" o "sándwich", con una sola línea celular de hibridoma que sirve como fuente tanto para el anticuerpo monoclonal marcado como para el anticuerpo monoclonal unido. Tales análisis se describen en la Patente de los Estados Unidos Núm. 4.376.110. La concentración de reactivo detectable debe ser suficiente de manera que la unión del polipéptido de fusión de EML4-ALK o TFG-ALK sea detectable en comparación con el fondo.

Los anticuerpos útiles en la práctica de los métodos descritos en la presente memoria se pueden conjugar con un soporte sólido adecuado para un análisis de diagnóstico (p. ej., cuentas, placas, portaobjetos o pocillos formados a partir de materiales tales como látex o poliestireno) de acuerdo con técnicas conocidas, tales como la precipitación. Los anticuerpos u otros reactivos de unión al polipéptido de fusión de ALK o al polipéptido de quinasa ALK truncado se pueden conjugar del mismo modo con grupos detectables tales como radiomarcas (p. ej., ³⁵S, ¹²⁵I, ¹³¹I), marcas enzimáticas (p. ej., peroxidasa de rábano picante, fosfatasa alcalina), y marcas fluorescentes (p. ej., fluoresceína) de acuerdo con técnicas conocidas.

Los análisis basados en células, tales como la citometría de flujo (FC), la inmunohistoquímica (IHC), o inmunofluorescencia (IF) son particularmente deseables en la práctica de los métodos de la descripción, ya que dichos formatos de ensayo son clínicamente adecuados, permiten la detección de la expresión del polipéptido de quinasa ALK mutante *in vivo*, y evitan el riesgo de cambios por artefactos en la actividad resultante de la manipulación de las células obtenidas a partir de, p. ej., una muestra tumoral con el fin de obtener extractos. De acuerdo con ello, en alguna realización preferida, los métodos de la descripción se implementan en un formato de análisis de citometría de flujo (FC), inmunohistoquímica (IHC), o inmunofluorescencia (IF).

La citometría de flujo (FC) puede emplearse para determinar la expresión de polipéptido de quinasa ALK mutante en un tumor de mamífero antes, durante, y después del tratamiento con un fármaco dirigido a inhibir la actividad de la quinasa ALK. Por ejemplo, las células tumorales de una muestra de médula ósea se pueden analizar mediante citometría de flujo para determinar la expresión y/o la activación del polipéptido de fusión de EML4-ALK o TFG-ALK, así como para determinar marcadores que identifican tipos de células cancerosas, etc., si así se desea. La citometría de flujo se puede llevar a cabo de acuerdo con métodos convencionales. Véase, p. ej., Chow et al, *Citometry (Communications in Clinical Cytometry)* 46: 72-78 (2001). Brevemente y a modo de ejemplo, se puede emplear el siguiente protocolo para el análisis citométrico: fijación de las células con paraformaldehído al 2% durante 10 minutos a 37°C seguido de permeabilización en metanol del 90% durante 30 minutos sobre hielo. A continuación las células se pueden teñir con el anticuerpo específico del polipéptido de fusión de EML4-ALK o TFG-ALK primario, lavar y marcar con un anticuerpo secundario marcado con fluorescencia. Las células se poseían analizar después en un citómetro de flujo (p. ej., Beckman Coulter FC500) de acuerdo con los protocolos específicos del aparato utilizado. Tal análisis identificaría el nivel de polipéptido de fusión de EML4-ALK o TFG-ALK expresado en el tumor. Un análisis similar después del tratamiento del tumor con un agente terapéutico inhibidor de ALK revelaría la sensibilidad de un tumor que expresa el polipéptido de fusión de ALK al inhibidor selectivo de la quinasa ALK.

La tinción inmunohistoquímica (IHC) también se puede emplear para determinar la expresión y/o estado de activación del polipéptido de quinasa ALK mutante en un cáncer de mamífero (p. ej., un tumor sólido tal como CPCNP) antes, durante, y después del tratamiento con un fármaco dirigido a inhibir la actividad de la quinasa ALK. La IHC se puede llevar a cabo de acuerdo con técnicas bien conocidas. Véase, p. ej., ANTIBODIES: A LABORATORY MANUAL, Capítulo 10, Harlow y Lane Eds., Cold Spring Harbor Laboratory (1988). En resumen, y a modo de ejemplo, se prepara tejido incluido en parafina (p. ej., tejido tumoral de una biopsia) para la tinción inmunohistoquímica mediante desparafinado de las secciones de tejido con xileno seguido de etanol; hidratación en agua, después PBS; desmascaramiento del antígeno mediante calentamiento del portaobjetos en tampón de citrato de sodio; incubación de las secciones en peróxido de hidrógeno; bloqueo en disolución de bloqueo; incubación del portaobjetos en anticuerpo para el polipéptido de fusión anti-EML4-ALK o anti-TFG-ALK primario y anticuerpo secundario; y, finalmente, detección utilizando método de avidina/biotina ABC de acuerdo con las instrucciones del fabricante.

Los análisis de inmunofluorescencia (IF) también se pueden emplear para determinar la expresión y/o el estado de activación del polipéptido de quinasa ALK mutante en un cáncer de mamífero antes, durante, y después del tratamiento con un fármaco dirigido a inhibir la actividad de la quinasa ALK. La IF se puede llevar a cabo de acuerdo con técnicas bien conocidas. Véase, p. ej., J.M. Polak y S. Van Noorden (1997) INTRODUCTION TO IMMUNOCYTOCHEMISTRY, 2ª ed.; ROYAL MICROSCOPY SOCIETY MICROSCOPY HANDBOOK 37, Bioscientific/Springer-Verlag. En resumen, y a modo de ejemplo, se pueden fijar muestras de pacientes en paraformaldehído seguido de metanol, bloquear con una solución de bloqueo tal como suero de caballo, incubar con el anticuerpo primario contra el polipéptido de fusión de EML4-ALK o TFG-ALK seguido de un anticuerpo secundario marcado con un colorante fluorescente tal como Alexa 488 y analizar con un microscopio de epifluorescencia.

Los anticuerpos empleados en los análisis descritos anteriormente se pueden conjugar ventajosamente con colorantes fluorescentes (p. ej., Alexa 488, PE), u otras marcas, tales como puntos cuánticos, para su uso en análisis multiparamétricos junto con otros anticuerpos de transducción de la señal (p. ej., EGFR, fosfo-AKT, fosfo-Erk 1/2) y/o marcadores celulares (p. ej., citoqueratina).

Una variedad de otros protocolos, incluyendo el análisis de inmunoabsorción ligado a enzimas (ELISA), radioinmunoanálisis (RIA), y clasificación celular activada por fluorescencia (FACS), para medir el polipéptido de quinasa ALK mutante es conocida en la técnica y proporciona una base para el diagnóstico de niveles alterados o anormales de la expresión polipéptido de fusión de EML4-ALK o TFG-ALK. Los valores normales o patrón para la expresión de estos polipéptido de fusión se establecen mediante la combinación de fluidos corporales o extractos de células tomados de sujetos mamíferos normales, preferiblemente seres humanos, con un anticuerpo para el polipéptido de fusión de EML4-ALK o TFG-ALK en condiciones adecuadas para la formación de complejos. La

cantidad de formación de complejo patrón se puede cuantificar por medio de diversos métodos, pero preferiblemente mediante métodos fotométricos. Las cantidades de polipéptido de fusión de EML4-ALK o TFG-ALK expresadas en sujetos, de control, y en muestras de enfermedad de tejidos de biopsia se comparan con los valores patrón. La desviación entre los valores patrón y del sujeto establece los parámetros para diagnosticar la enfermedad.

5 Análisis de péptidos y ácidos nucleicos

Del mismo modo, se pueden preparar péptidos AQUA para la detección/cuantificación del polipéptido de quinasa ALK mutante expresado en una muestra biológica que comprende células de un tumor y utilizarlos en análisis AQUA convencionales, como se ha descrito en detalle en la sección E anterior. Por consiguiente, en algunas realizaciones preferidas de los métodos de la descripción, el reactivo específico del polipéptido de fusión de ALK comprende un fosfopéptido marcado con un isótopo pesado (péptido AQUA) correspondiente a una secuencia peptídica que comprende la unión por fusión de un polipéptido de fusión de EML4-ALK o de un polipéptido de fusión TFG-ALK, como se ha descrito anteriormente en la Sección E.

Los reactivos de específicos del polipéptido de ALK mutantes útiles en la práctica de los métodos de la descripción también pueden ser ARNm, oligonucleótidos o sondas de ADN que pueden hibridarse directamente con, y detectar, los transcritos de expresión de los polipéptidos de fusión o truncados en una muestra biológica. Tales sondas se comentan con detalle en la Sección B anterior. En resumen, y a modo de ejemplo, se pueden sondear muestras de pacientes, fijadas con formalina, incluidas en parafina con una sonda de ARN marcada con fluoresceína seguido de lavados con formamida, SSC y PBS y análisis con un microscopio de fluorescencia. También se prefieren las sondas FISH, incluyendo las sondas de punto de rotura, que permiten la detección fluorescente de reordenamientos de genes, tales como las mutaciones por delección de EML4-ALK en el cromosoma 2 (véase el Ejemplo 6).

Los polinucleótidos que codifican el polipéptido de quinasa ALK mutante también se pueden utilizar para propósitos de diagnóstico. Los polinucleótidos que se pueden utilizar incluyen secuencias de oligonucleótidos, moléculas de ARN y ADN antisentido, y PNA. Los polinucleótidos se pueden utilizar para detectar y cuantificar la expresión génica en tejidos de tumores sólidos sometidos a biopsia en los que la expresión del polipéptido de fusión de EML4-ALK o TFG-ALK o del polipéptido de quinasa ALK activo truncado puede ser correlacionada con la enfermedad. Por ejemplo, se puede utilizar un análisis de diagnóstico para distinguir entre la ausencia, presencia, y exceso de expresión del polipéptido de fusión de EML4-ALK o TFG-ALK, y para verificar la regulación de los niveles del polipéptido de fusión de ALK durante la intervención terapéutica.

En una realización preferida, se puede utilizar la hibridación con sondas de PCR que son capaces de detectar secuencias de polinucleótidos, incluyendo secuencias genómicas, que codifican un polipéptido de fusión de ALK o un polipéptido de quinasa ALK truncado, o moléculas estrechamente relacionadas, para identificar secuencias de ácido nucleico que codifican el polipéptido de ALK mutante. La construcción y el uso de tales sondas se describen en la Sección B anterior. La especificidad de la sonda, ya sea elaborada a partir de una región altamente específica, p. ej., 10 nucleótidos únicos en la unión por fusión, ya sea una región menos específica, p. ej., la región codificante 3', y la rigurosidad de la hibridación o amplificación (máxima, alta, intermedia o baja) determinará si la sonda identifica solo secuencias de origen natural que codifican el polipéptido de ALK mutante, alelos, o secuencias relacionadas.

Las sondas también se pueden utilizar para la detección de secuencias relacionadas, y deben contener preferiblemente al menos 50% de los nucleótidos de cualquiera de las secuencias codificantes del polipéptido de ALK mutante. Las sondas de hibridación objeto de la descripción pueden ser ADN o ARN y estar derivadas de las secuencias de nucleótidos de los SEQ ID NO: 2, 19, y 21, abarcando lo más preferiblemente la unión por fusión (véanse las Figuras 1A-C, el panel inferior y los SEQ ID NO: 7, 24, y 26), o de secuencias genómicas incluyendo el promotor, elementos potenciadores, intrones y de los polipéptidos EML-4, TFG, y ALK de origen natural, como se ha descrito adicionalmente en la Sección B anterior.

Por ejemplo, se puede utilizar el polinucleótido de fusión de EML4-ALK de la descripción en el análisis Southern o northern, transferencia puntual, u otras tecnologías basadas en membranas; en tecnologías de PCR; o en análisis con tiras reactivas, punción, ELISA o chips que utilizan fluidos o tejidos de biopsias de pacientes para detectar la expresión alterada del polipéptido de ALK. Tales métodos cualitativos o cuantitativos son bien conocidos en la técnica. En un aspecto concreto, las secuencias de nucleótidos que codifican un polipéptido de ALK mutante de la descripción pueden ser útiles en análisis que detectan la activación o la inducción de diversos tipos de cáncer, incluyendo carcinomas de pulmón. Los polinucleótidos de ALK mutante se pueden marcar por medio de métodos convencionales, y añadir a una muestra de fluido o tejido de un paciente en condiciones adecuadas para la formación de complejos de hibridación. Después de un período de incubación adecuado, la muestra se lava y la señal se cuantifica y se compara con un valor patrón. Si la cantidad de señal en la muestra obtenida mediante biopsia o extraída está significativamente alterada con respecto a la de una muestra de control comparable, las secuencias de nucleótidos se han hibridado con secuencias de nucleótidos en la muestra, y la presencia de niveles alterados de secuencias de nucleótidos que codifican el polipéptido de fusión de EML4-ALK o el polipéptido de quinasa ALK truncado en la muestra indica la presencia de la enfermedad asociada. Tales análisis también se pueden utilizar para evaluar la eficacia de un régimen de tratamiento terapéutico concreto en estudios con animales, en pruebas clínicas, o en el seguimiento del tratamiento de un paciente individual.

Con el fin de proporcionar una base para el diagnóstico de enfermedades caracterizadas por la expresión del polipéptido de quinasa ALK mutante, se establece un perfil normal o patrón para expresión. Esto se puede lograr mediante la combinación de fluidos corporales o extractos celulares tomados de sujetos normales, ya sean animales o seres humanos, con una secuencia, o un fragmento de la misma, que codifica un polipéptido de fusión de EML4-ALK o TFG-ALK, en condiciones adecuadas para la hibridación o la amplificación. La hibridación patrón se puede cuantificar comparando los valores obtenidos de sujetos normales con los de un experimento en el que se utiliza una cantidad conocida de un polinucleótido sustancialmente purificado. Los valores patrón obtenidos a partir de muestras normales se pueden comparar con los valores obtenidos a partir de muestras de pacientes que son sintomáticos para la enfermedad. La desviación entre los valores patrón y del sujeto se utiliza para establecer la presencia de la enfermedad.

Una vez que se establece la enfermedad y se inicia un protocolo de tratamiento, se pueden repetir los análisis de hibridación sobre una base regular para evaluar si el nivel de expresión en el paciente comienza a aproximarse al que se observa en el paciente normal. Los resultados obtenidos a partir de análisis sucesivos se pueden utilizar para demostrar la eficacia del tratamiento a lo largo de un período que oscila de varios días a meses.

Los usos diagnósticos adicionales para los polinucleótidos de ALK mutante de la descripción pueden implicar la utilización de la reacción en cadena de la polimerasa (PCR), un formato de análisis preferido que es convencional para los expertos en la técnica. Véase, p. ej., MOLECULAR CLONING, A LABORATORY MANUAL, segunda edición, Sambrook, J., Fritsch, E.F. y Maniatis, T., eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York (1989). Los oligómeros para la PCR se pueden sintetizar químicamente, generarse enzimáticamente o producirse a partir de una fuente recombinante. Los oligómeros consistirán preferiblemente en dos secuencias de nucleótidos, una con orientación efectora (5' a 3') y otra antisentido (3' a 5'), empleadas en condiciones optimizadas para la identificación de un gen o afección específicos. Los mismos dos oligómeros, conjuntos anidados de oligómeros, o incluso una reserva degenerada de oligómeros pueden emplearse en condiciones menos rigurosas para la detección y/o cuantificación de secuencias de ADN o ARN estrechamente relacionadas.

Los métodos que también se pueden utilizar para cuantificar la expresión de un polipéptido de fusión de ALK o polipéptido de quinasa ALK truncado incluyen el radiomarcaje o la biotilación de nucleótidos, la amplificación simultánea de un ácido nucleico de control, y curvas patrón sobre la que se interpolan los resultados experimentales (Melby et al., *J. Immunol. Methods*, 159: 235-244 (1993); Duplaa et al. *Anal. Biochem.* 229-236 (1993)). La velocidad de cuantificación de múltiples muestras se puede acelerar mediante la ejecución del análisis en un formato ELISA en el que el oligómero de interés se presenta en varias diluciones y una respuesta espectrofotométrica o colorimétrica proporciona una cuantificación rápida.

En otra realización de la descripción, los polinucleótidos de ALK mutante de la descripción, así como la región genómica adyacente próxima y distante de ellos, se pueden utilizar para generar sondas de hibridación que son útiles para el mapeo de la secuencia genómica de origen natural. Las secuencias se pueden mapear con respecto a un cromosoma concreto o a una región específica del cromosoma utilizando técnicas bien conocidas. Tales técnicas incluyen la hibridación fluorescente *in situ* (FISH), FACS, o construcciones de cromosomas artificiales, tales como cromosomas artificiales de levadura, cromosomas artificiales bacterianos, construcciones P1 bacterianas o bibliotecas de ADNc de cromosomas individuales, como fue revisado por Price, C.M., *Blood Rev.* 7: 127-134 (1993), y Trask, B.J., *Trends Genet.* 7: 149-154 (1991).

En una realización preferida, se emplea FISH (como describen Verma et al. HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES, Pergamon Press, Nueva York, NY (1988)) y se puede correlacionar con otras técnicas física de mapeo de cromosomas y datos de mapas genéticos. Se pueden encontrar ejemplos de datos genéticos en 1994 Genome Issue of *Science* (265: 1981f). La correlación entre la ubicación del gen que codifica el polipéptido de fusión de EML4-ALK o TFG-ALK o el polipéptido de la quinasa ALK activo truncado sobre un mapa cromosómico físico y una enfermedad específica, o la predisposición a una enfermedad específica, pueden ayudar a delimitar la región de ADN asociada con esa enfermedad genética. Las secuencias de nucleótidos de la descripción sujeto se pueden utilizar para detectar diferencias en las secuencias génicas entre individuos normales, portadores, o afectados. Se pueden emplear sondas FISH punto de rotura de dos colores, por ejemplo, para detectar la presencia o ausencia de genes de EML-4, TFG, y/o ALK mutantes en una muestra.

La hibridación *in situ* de preparaciones cromosómicas y las técnicas físicas de mapeo tales como el análisis de ligamiento utilizando marcadores cromosómicos establecidos se puede utilizar para ampliar los mapas genéticos. A menudo la ubicación de un gen sobre el cromosoma de otra especie de mamífero, tal como ratón, puede revelar marcadores asociados incluso si el número o el brazo de un cromosoma humano concreto no es conocido. Se pueden asignar nuevas secuencias a los brazos cromosómicos, o porciones de los mismos, mediante mapeo físico. Esto proporciona una información valiosa a los investigadores en busca de genes de enfermedades utilizando clonación posicional u otras técnicas de descubrimiento de genes. Una vez que la enfermedad o el síndrome han sido localizados groseramente mediante ligamiento genético a una región genómica concreta, por ejemplo, AT a 11q22-23 (Gatti et al., *Nature* 336: 577-580 (1988)), cualquiera de las secuencias asignada a esa zona puede representar genes asociados o reguladores para una investigación adicional. La secuencia de nucleótidos de la descripción sujeto también se puede utilizar para detectar diferencias en la localización cromosómica debido a translocación, inversión, etc., entre individuos normales, portadores, o afectados.

Otros métodos adecuados para la detección de ácidos nucleicos, tales como sondas de oligonucleótidos conjugadas de unión al surco menor (véase, p. ej., la Patente de los Estados Unidos Núm. 6.951.930, "Hybridization-Triggered Fluorescent Detection of Nucleic Acids ") son conocidos por los expertos en la técnica.

Muestras biológicas

- 5 Las muestras biológicas útiles en la práctica de los métodos de la descripción se pueden obtener de cualquier mamífero en el que está presente o en desarrollo un cáncer caracterizado por la expresión de un polipéptido de fusión de EML4-ALK o TFG-ALK. En una realización, el mamífero es un ser humano, y el ser humano puede ser un candidato para un agente terapéutico inhibidor de ALK para el tratamiento de un cáncer, p. ej., CPCNP. El candidato humano puede ser un paciente que está siendo tratado actualmente con, o está siendo considerado para el
- 10 tratamiento con, un inhibidor de la quinasa ALK, tal como WHI-131 y/o WHI-154. En otra realización, el mamífero es un animal grande, tal como un caballo o una vaca, mientras que en otras realizaciones, el mamífero es un animal pequeño, tal como un perro o un gato, todos los cuales son conocidos por desarrollar cánceres, incluyendo carcinomas de pulmón.

- Cualquier muestra biológica que comprende células (o extractos de células) de un cáncer de mamífero es adecuada para su uso en los métodos de la descripción. En el caso del polipéptido de fusión de EML-ALK, cualquier tipo de cáncer, ya sea sólido o no sólido, será adecuado. En el caso de TFG-ALK, los tumores sólidos están dentro del alcance de los métodos de la descripción. Por ejemplo, la muestra biológica puede comprender células obtenidas de una efusión, tal como una efusión pleural. Se sabe que las efusiones pleurales (líquido que se forma fuera del pulmón en la cavidad torácica y que contiene células cancerosas) se forman en muchos pacientes con cáncer de
- 20 pulmón avanzado (incluyendo CPCNP), y la presencia de tal efusión es predictiva de un resultado desfavorable y un corto tiempo de supervivencia. Los mecanismos convencionales para la obtención de muestras de efusión pleural se han descrito y son bien conocidos en la técnica (véase Sahn, *Clin Chest Med.* 3(2): 443-52 (1982)). También se pueden obtener células tumorales circulantes a partir de suero utilizando marcadores tumorales, marcadores de la proteína citoqueratina u otros métodos de selección negativa como se ha descrito (véase Ma et al., *Anticancer Res.* 23(1A): 49-62 (2003)). Las muestras de suero y de médula ósea pueden ser particularmente preferidas para los
- 25 pacientes con leucemia. Para los cánceres que implican tumores sólidos, tales como sarcomas y carcinomas, la muestra biológica puede comprender células obtenidas de una biopsia de tumor, que se puede obtener de acuerdo con técnicas clínicas convencionales. Por ejemplo, se ha observado la expresión aberrante de ALK en un espectro de cánceres, incluidos neuroblastomas y cáncer neuroectodérmico. Véase, p. ej., Pulford *et al.*, más arriba. El mutante por translocación de TFG-ALK, sin embargo, solo se ha descrito en linfoma y no observado previamente en tumores sólidos.

- Una muestra biológica puede comprender células (o extractos de células) de un cáncer en el que se expresa y/o activa que un polipéptido de fusión de ALK, pero no la quinasa ALK de tipo salvaje. Alternativamente, la muestra puede comprender células de un cáncer en el que se expresan y/o activan tanto el polipéptido de ALK mutante como la
- 35 quinasa ALK de tipo salvaje, o en el que se expresan y/o activan la quinasa ALK y/o EML-4 y/o TFG de tipo salvaje pero no el polipéptido de ALK mutante.

- Los extractos celulares de las muestras biológicas anteriores se pueden preparar, en bruto o parcialmente (o totalmente) purificados, de acuerdo con técnicas convencionales, y se pueden utilizar en los métodos de la descripción. Alternativamente, se pueden utilizar muestras biológicas que comprenden células completas en
- 40 formatos de análisis preferidos tales como inmunohistoquímica (IHC), citometría de flujo (FC), inmunofluorescencia (IF), e hibridación fluorescente *in situ* (FISH) como se ha descrito adicionalmente más arriba. Tales análisis de células completas son ventajosos ya que minimizan la manipulación de la muestra de células tumorales y por lo tanto reducen los riesgos de alteración del estado de señalización/activación *in vivo* de las células y/o introducir señales artefacto. Los análisis de células completas también son ventajosos ya que caracterizan la expresión y la
- 45 señalización únicamente en células tumorales, en lugar de una mezcla de células tumorales y normales.

- Al poner en práctica el método descrito para determinar si un compuesto inhibe la progresión de un tumor caracterizado por un gen de fusión y/o un polipéptido de fusión de EML4-ALK o TFG-ALK, también se pueden emplear ventajosamente muestras biológicas que comprenden células de modelos de trasplante o xenoinjertos de médula ósea de mamíferos. Los xenoinjertos preferidos (o receptores de trasplante) son mamíferos pequeños, tales
- 50 como ratones, que albergan tumores humanos que expresan un polipéptido de quinasa ALK mutante. Los xenoinjertos que albergan tumores humanos son bien conocidos en la técnica (véase Kal, *Cancer Treat Res.* 72: 155-69 (1995)) y la producción de xenoinjertos de mamíferos que albergan tumores humanos está bien descrita (véase Winograd et al., *In vivo.* 1(1): 1-13 (1987)). De un modo similar están bien descritos la generación y el uso de modelos de trasplante de médula ósea (véanse, p. ej., Schwaller, et al, *EMBO J.* 17: 5321-333 (1998); Kelly et al,
- 55 *Blood* 99: 310-318 (2002)). Por "cáncer caracterizado por" un polinucleótido de fusión y/o un polipéptido de fusión de EML4-ALK o TFG-ALK se quiere significar un cáncer en el que están presentes tal gen y/o polipéptido expresado de ALK mutante, en comparación con un cáncer en el que no están presentes tal gen de fusión y/o polipéptido de fusión.

- En la evaluación de la presencia del polinucleótido o la expresión del polipéptido de ALK mutante en una muestra biológica que comprende células de un tumor de cáncer de mamífero, se puede emplear de manera deseable una
- 60

muestra de control que representa una célula en la que no se producen tales translocación y/o proteína de fusión para fines comparativos. Idealmente, la muestra de control comprende células de un subgrupo de cáncer concreto (p. ej., CPCNP) que es representativo del subgrupo en el que no se produce la mutación (p. ej., mutación por delección de EML4-ALK) y/o no se expresa el polipéptido de fusión. La comparación del nivel en la muestra de control frente a la muestra biológica de ensayo identifica de ese modo si está/están presentes el polinucleótido y/o el polipéptido de ALK mutante. Alternativamente, puesto que polinucleótido y/o polipéptido de fusión de EML4-ALK y/o TFG-ALK pueden no estar presentes en la mayoría de los cánceres, cualquier tejido que no exprese de manera similar tal polipéptido de ALK mutante (o albergue el polinucleótido mutante) se puede emplear como control.

Los métodos descritos a continuación tendrán utilidad de diagnóstico valiosa para los cánceres caracterizados por el polinucleótido y/o polipéptido de ALK mutante, y las decisiones de tratamiento correspondientes a la misma. Por ejemplo, las muestras biológicas se pueden obtener de un sujeto al que no se ha diagnosticado previamente que tiene un cáncer caracterizado por una mutación por delección y/o un polipéptido de fusión de EML4-ALK, ni tampoco se ha sometido a un tratamiento para tal cáncer, y el método se emplea para identificar diagnósticamente un tumor en tal sujeto como perteneciente a un subgrupo de tumores (p. ej., CPNM) en el que está presente y/o se expresa un polinucleótido y/o polipéptido de fusión de EML4-ALK. Los métodos de la descripción también se pueden emplear para controlar la progresión o la inhibición de un cáncer que expresa el polipéptido de quinasa ALK mutante después del tratamiento de un sujeto con una composición que comprende un agente terapéutico o una combinación de agentes terapéuticos inhibidores de quinasa ALK.

Semejante análisis de diagnóstico se puede llevar a cabo después de o antes de los procedimientos de evaluación preliminar o vigilancia quirúrgica. El método de identificación de la descripción se puede emplear ventajosamente como diagnóstico para identificar pacientes que tienen cáncer, tales como CPCNP, dirigido por las proteínas de fusión de EML4-ALK y/o TFG-ALK o por la quinasa ALK truncada, cuyos pacientes sería muy probable que respondieran a agentes terapéuticos dirigidos a inhibir la actividad de la quinasa ALK, tales como WHI-131 y/o WHI-154 o sus análogos. La capacidad para seleccionar tales pacientes sería también útil en la evaluación clínica de la eficacia de futuros agentes terapéuticos dirigidos ALK así como en la prescripción futura de tales fármacos a los pacientes.

Diagnóstico

La capacidad de identificar selectivamente cánceres en los que está/están presentes un polinucleótido de fusión y/o polipéptido de fusión de EML4-ALK y/o TFG-ALK permite importantes nuevos métodos para identificar con precisión tales tumores con fines diagnósticos, así como obtener información útil en la determinación de si es probable que un tumor de este tipo responda a una composición terapéutica inhibidora de ALK, o es probable que sea parcialmente o totalmente insensible a un inhibidor dirigido a una quinasa diferente cuando se administra como un agente único para el tratamiento del cáncer.

Por consiguiente, en una realización, la descripción proporciona un método para detectar la presencia de un polinucleótido de ALK mutante y/o su polipéptido de ALK mutante codificado en una muestra biológica de un cáncer de mamífero, comprendiendo dicho método las etapas de:

(a) obtener una muestra biológica de un cáncer de mamífero; y

(b) utilizar al menos un reactivo que detecta un polinucleótido de fusión, o su polipéptido de fusión codificado, que comprende parte de ALK con parte de una proteína secundaria para determinar si un polinucleótido mutante de ALK y/o su polipéptido de ALK mutante codificado están presentes en dicha biológica muestra.

En algunas realizaciones preferidas, el cáncer es un tumor sólido sarcoma o carcinoma, mientras que en otra realización el carcinoma es un carcinoma de pulmón, tal como CPCNP. En otra realización preferida, el polipéptido de ALK mutante es un polipéptido de fusión que comprende los residuos 1116-1383 de ALK (SEQ ID NO: 5) con una porción de dicha proteína secundaria. En otra realización preferida, la proteína secundaria se selecciona del grupo que consiste en EML-4 (SEQ ID NO: 3) y la proteína del Gen Fusionado a TRK (TFG) (SEQ ID NO: 22). En otra realización preferida adicional, el polipéptido de fusión comprende los residuos 1-233 o los residuos 1 a 495 de EML-4 (SEQ ID NO: 3) o los residuos 1-138 de TFG (SEQ ID NO: 22).

En otras realizaciones preferidas, el polinucleótido de fusión comprende un polinucleótido de fusión de EML4-ALK (SEQ ID NO: 2 o 19) o un polinucleótido de fusión de TFG-ALK (SEQ ID NO: 21) mientras que en otra realización preferida adicional el polipéptido de fusión comprende un polipéptido de fusión de EML4-ALK (SEQ ID NO: 1 o 18) o un polipéptido de fusión de TFG-ALK (SEQ ID NO: 20). En otra realización preferida más, el polinucleótido de fusión es un polinucleótido o polipéptido de fusión de EML4-ALK descritos anteriormente.

En más realizaciones preferidas, el método emplea un reactivo que comprende un polinucleótido de fusión de EML4-ALK y/o al menos un reactivo específico del polipéptido de fusión de EML4-ALK (anticuerpo o péptido AQUA), como se ha descrito anteriormente. En algunas realizaciones preferidas, el reactivo comprende un reactivo aislado que se une específicamente a, o detecta, un polipéptido de fusión de TFG-ALK (SEQ ID NO: 20) o un polinucleótido de fusión TFG-ALK (SEQ ID NO: 21), pero no se une a, ni detecta, TFG de tipo salvaje ni ALK de tipo salvaje. En otras realizaciones preferidas, el reactivo es una sonda para la reacción en cadena de la polimerasa (PCR) de una sonda

de hibridación fluorescente *in situ* (FISH). Ciertas realizaciones preferidas emplean un péptido marcado con isótopo pesado (AQUA) que comprende la secuencia de aminoácidos de la unión por fusión de polipéptido de fusión de TFG-ALK o un punto de truncamiento dentro de ALK de tipo salvaje.

5 En algunas realizaciones preferidas, los métodos de diagnóstico de la descripción se implementan en un formato de análisis de citometría de flujo (FC), inmunohistoquímica (IHC), o inmuno-fluorescencia (IF), como se ha descrito anteriormente. En otra realización preferida, se detecta la actividad del polipéptido de fusión de EML4-ALK o TFG-ALK. En otras realizaciones preferidas, los métodos de diagnóstico de la descripción se implementan en un formato de análisis de hibridación fluorescente *in situ* (FISH) o de reacción en cadena de la polimerasa (PCR), como se ha descrito anteriormente.

10 La descripción proporciona adicionalmente un método para determinar si un compuesto inhibe la progresión de un cáncer caracterizado por un polinucleótido de fusión y/o polipéptido de fusión de EML4-ALK o TFG-ALK, comprendiendo dicho método la etapa de determinar si dicho compuesto inhibe la expresión y/o actividad de dicho polipéptido de fusión de EML4-ALK o TFG-ALK en dicho cáncer. En una realización preferida la inhibición de la expresión y/o actividad del polipéptido de fusión de ALK se determina utilizando al menos un reactivo que detecta un polinucleótido o polipéptido de fusión de EML4-ALK de la descripción y/o un polinucleótido o polipéptido de fusión de TFG-ALK descrito en la presente memoria. Los compuestos adecuados para la inhibición de la actividad de la quinasas ALK se comentan con más detalle en la Sección G a continuación.

20 Sondas polinucleotídicas de ALK mutante y los reactivos específicos del polipéptido útiles en la práctica de los métodos de la descripción se describen con más detalle en las secciones B y D anteriores. En una realización preferida, el reactivo específico del polipéptido de fusión de ALK comprende un anticuerpo específico del polipéptido de fusión. En otra realización preferida, el reactivo específico del polipéptido de fusión comprende un fosfopéptido marcado con un isótopo pesado (péptido AQUA) que corresponde a la unión por fusión de un polipéptido de fusión de ALK (véanse las Figuras 1A-C (panel inferior)). En otra realización preferida, el reactivo específico del polinucleótido de fusión comprende una sonda FISH que corresponde a la unión por fusión de un gen de fusión de ALK y/o puntos de rotura de los genes de EML4, TFG, o ALK de tipo salvaje.

25 Los métodos de la descripción descritos anteriormente también pueden comprender opcionalmente la etapa de determinar el nivel de expresión o activación de otras quinasas, tales como ALK y EGFR de tipo salvaje, o de otras moléculas de señalización aguas abajo en dicha muestra biológica. El perfilado tanto de la expresión/activación del polipéptido de fusión de ALK como de la expresión/activación de otras quinasas y rutas en una muestra biológica dada puede proporcionar información valiosa sobre qué quinasas y rutas están dirigiendo la enfermedad, y por lo tanto qué régimen terapéutico es probable que sea el de mayor beneficio.

Escrutinio de compuestos

35 El descubrimiento de los polipéptidos de fusión de EML4-ALK novedosos descritos en la presente memoria también permite el desarrollo de nuevos compuestos que inhiben la actividad de esta proteína de ALK mutante, en particular su actividad de la quinasas ALK. Por consiguiente, la descripción también proporciona, en parte, un método para determinar si un compuesto inhibe la progresión de un cáncer caracterizado por un polinucleótido de fusión y/o polipéptido de fusión de EML4-ALK, comprendiendo dicho método la etapa de determinar si dicho compuesto inhibe la expresión y/o actividad de dicho polipéptido de fusión de EML4-ALK en dicho cáncer. En una realización preferida, la inhibición de la expresión y/o actividad del polipéptido de fusión de EML4-ALK se determina utilizando al menos un reactivo que detecta un polinucleótido de fusión y/o polipéptido de fusión de la descripción. Los reactivos preferidos de la descripción se han descrito anteriormente. Los compuestos adecuados para la inhibición de la actividad de la quinasas ALK se describen en más detalle en la Sección G a continuación.

45 El compuesto puede, por ejemplo, ser un inhibidor de quinasas, tal como un inhibidor de molécula pequeña o anticuerpo inhibidor. Éste puede ser un pan-inhibidor de quinasas con actividad contra varias quinasas diferentes, o un inhibidor específico de una quinasas. Los compuestos inhibidores de quinasas ALK se comentan con más detalle en la Sección G de más abajo. Las muestras biológicas de los pacientes se pueden tomar antes y después del tratamiento con el inhibidor y después analizarse, utilizando los métodos descritos anteriormente, para determinar el efecto biológico del inhibidor sobre la actividad de la quinasas ALK, incluyendo la fosforilación de la proteína sustrato aguas abajo. Tal análisis farmacodinámico puede ser útil en la determinación de la dosis biológicamente activa del fármaco que puede ser preferible a una dosis máxima tolerable. Tal información también sería útil en las presentaciones para la aprobación de fármacos mediante la demostración del mecanismo de acción del fármaco. La identificación de los compuestos con tales características inhibitoras deseadas se describen adicionalmente en la Sección G de más abajo.

G. Inhibición terapéutica de cánceres

55 De acuerdo con la presente invención, se ha demostrado ahora que la progresión de un cáncer de tumor sólido de mamífero (CPCNP) en el que se expresa la proteína de fusión de EML4-ALK puede ser inhibida, *in vivo*, mediante la inhibición de la actividad de la quinasas ALK de tal cáncer. De un modo similar al descrito en la presente memoria, la

actividad de un cáncer de tumor sólido de mamífero en el que se expresa la proteína de fusión de TFG-ALK puede ser inhibida de manera similar, *in vivo*, mediante la inhibición de la actividad de la quinasa ALK de tal cáncer. La actividad ALK en los cánceres caracterizados por la expresión de un polipéptido de ALK mutante puede ser inhibida poniendo en contacto el cáncer (p. ej., un tumor) con un agente terapéutico inhibidor de la quinasa ALK, tal como un inhibidor de quinasa de molécula pequeña como WHI-131 o WHI-154. Como se describe adicionalmente en el Ejemplo 2 a continuación, la inhibición del crecimiento de tumores que expresan la proteína de fusión de ALK, por ejemplo, puede llevarse a cabo mediante la inhibición de la quinasa de fusión utilizando un tipo ilustrativo de agente terapéutico inhibidor de ALK, ARNip. Por consiguiente, la descripción proporciona, en parte, un inhibidor de ALK para usar en la inhibición de la progresión de un cáncer que expresa el polipéptido de fusión de EML4-ALK mediante la inhibición de la expresión y/o actividad de la quinasa ALK mutante en el cáncer.

Un agente terapéutico inhibidor de la quinasa ALK puede ser cualquier composición que comprende al menos un compuesto, biológico o químico, que inhibe, directa o indirectamente, la expresión y/o actividad de la quinasa ALK *in vivo*, incluyendo las clases ilustrativas de compuestos descritas a continuación. Tales compuestos incluyen agentes terapéuticos que actúan directamente sobre la propia quinasa ALK, o sobre proteínas o moléculas que modifican la actividad de ALK, o que actúan indirectamente mediante la inhibición de la expresión de ALK. Tales composiciones también pueden incluir composiciones que comprenden solamente un único compuesto inhibidor de la quinasa ALK, así como composiciones que comprenden múltiples agentes terapéuticos (incluyendo aquellos contra otras RTK), que también pueden incluir un agente terapéutico no específico como agente quimioterapéutico o inhibidor general de la transcripción.

Inhibidores de molécula pequeña

Un agente terapéutico inhibidor de ALK útil en la invención es un inhibidor de pequeña molécula dirigido, tal como WHI-131 y WHI-154, o sus análogos. WHI-131 y WHI-154 son inhibidores de ALK específicos de molécula pequeña de tipo quinazolina, y sus propiedades han sido descritas. Véase Marzec et al., *Lab. Invest.* 85(12): 1544-54 (2005). Se ha demostrado que estos compuestos inducen la apoptosis y suprimen la proliferación en células de linfoma. Otros inhibidores de quinasas específicos de molécula pequeñas son bien conocidos en la técnica. Por ejemplo, Gleevec® (STI-571, imatinib), que se une específicamente a, y bloquea el sitio de unión a ATP de la quinasa de fusión BCR-ABL (así como otras quinasas) impidiendo de este modo la fosforilación y la activación de esta enzima, se encuentra disponible en el mercado y sus propiedades son bien conocidas. Véase, p. ej., Dewar et al., *Blood* 105(8): 3127-32 (2005). Otros inhibidores de ALK de molécula pequeña están actualmente en desarrollo por Novartis, Inc., y Cephalon, Inc.

Los inhibidores específicos de molécula pequeña son una clase de moléculas que inhiben típicamente la actividad de su enzima diana uniéndose específicamente, y a menudo de forma irreversible, al sitio catalítico de la enzima, y/o uniéndose a una hendidura de unión a ATP u otro sitio de unión dentro de la enzima que impide que la enzima adopte una conformación necesaria para su actividad. Los inhibidores de molécula pequeña pueden ser diseñados racionalmente utilizando el modelado cristalográfico de rayos X o por ordenador de la estructura tridimensional de la quinasa ALK, o se pueden encontrar mediante escrutinio de alto rendimiento de bibliotecas de compuestos para la inhibición de ALK. Tales métodos son bien conocidos en la técnica, y han sido descritos. La especificidad de la inhibición de ALK puede confirmarse, p. ej., examinando la capacidad de dicho compuesto para inhibir la actividad de ALK, pero no otra actividad quinasa, en un panel de quinasas, y/o examinando la inhibición de la actividad de ALK en una muestra biológica que comprende células de carcinoma de pulmón, como se ha descrito anteriormente. Tales métodos de escrutinio se describen adicionalmente más adelante.

Anticuerpo inhibidores

Los agentes terapéuticos inhibidores de quinasas ALK útiles en la invención también pueden ser anticuerpos dirigidos que se unen específicamente a sitios o dominios catalíticos o de unión críticos necesarios para la actividad de ALK, e inhiben la quinasa bloqueando el acceso de ligandos, sustratos o moléculas secundarias para ALK, y/o evitando que la enzima adopte una conformación necesaria para su actividad. La producción, escrutinio y/o uso terapéutico de anticuerpos humanizados específicos de dianas han sido bien descritos. Véase Merluzzi et al., *Adv Clin Path.* 4(2): 77-85 (2000). Se encuentran disponibles tecnologías y sistemas comerciales, tales como Morphosys, Inc.'s Human Combinatorial Antibody Library (HuCAL®), para la generación y el escrutinio de alto rendimiento de anticuerpos humanizados inhibidores específicos de dianas.

Se ha descrito la producción de diversos anticuerpos dirigidos anti-quinasas receptoras y su uso para inhibir la actividad del receptor elegido como diana. Véanse, p. ej., la Publicación de Patente de los Estados Unidos Núm. 20040202655, "Antibodies to IgF-I Receptor for the Treatment of Cancers", 14 de Octubre de 2004, Morton et al.; Publicación de Patente de los Estados Unidos Núm. 20040086503, "Human anti-Epidermal Growth Factor Receptor Single-Chain Antibodies", 15 de Abril de 2004, Raisch et al.; Publicación de Patente de los Estados Unidos Núm. 20040033543, "Treatment of Renal Carcinoma Using Antibodies Using Antibodies Against EGFr," 19 de Febrero de 2004, Schwab et al. Los métodos normalizados para la producción y el uso de anticuerpos inhibidores de la actividad tirosina quinasa receptora son conocidos en la técnica. Véase, p. ej., la Patente Europea Núm. EP1423428, "Antibodies that Block Receptor Tyrosine Kinase Activation, Methods of Screening for and Uses Thereof", 2 de Junio de 2004 Borges et al.

También se pueden emplear los enfoques de presentación de fagos para generar anticuerpos inhibidores específicos de ALK, y protocolos para la construcción de la biblioteca de bacteriófagos y la selección de anticuerpos recombinantes se proporcionan en el texto de referencia bien conocido CURRENT PROTOCOLS IN IMMUNOLOGY, Colligan et al. (Eds.), John Wiley & Sons, Inc. (1992-2000), Capítulo 17, Sección 17.1. Véanse también la Patente de los Estados Unidos Núm. 6.319.690, 20 de Noviembre de 2001, Little et al.; la Patente de los Estados Unidos Núm. 6.300.064, 9 de Octubre de 2001, Knappik et al.; la Patente de los Estados Unidos Núm. 5.840.479, 24 de Noviembre de 1998, Little et al.; la Publicación de Patente de los Estados Unidos Núm. 20030219839, 27 de Noviembre de 2003, Bowdish et al.

Se puede producir una biblioteca de fragmentos de anticuerpos presentados en la superficie de bacteriófagos (véase, p. ej., la Patente de los Estados Unidos Núm. 6.300.064, 9 de Octubre de 2001, Knappik et al.) y se puede escrutar para determinar la unión a una forma dimérica soluble de una proteína tirosina quinasa receptoras (como ALK). Un fragmento de anticuerpo que se une a la forma dimérica soluble de la RTK utilizado para el escrutinio se identifica como una molécula candidato para bloquear la activación constitutiva de la RTK diana en una célula. Véase la Patente Europea Núm. EP1423428, Borges et al., más arriba.

Los anticuerpos dirigidos a la unión de ALK identificados en el escrutinio de bibliotecas de anticuerpos como se ha descrito anteriormente se puede escrutar adicionalmente a continuación para determinar su capacidad para bloquear la actividad de ALK, tanto en un análisis de quinasa *in vitro* como en líneas celulares y/o tumores *in vivo*. La inhibición de ALK puede confirmarse, por ejemplo, examinando de la capacidad de tal agente terapéutico de anticuerpo para inhibir la actividad de la quinasa ALK, pero no otra actividad quinasa, en un panel de quinasas, y/o examinando la inhibición de la actividad de ALK en una muestra biológica que comprende células de cáncer, como se ha descrito anteriormente. Los métodos para el escrutinio de tales compuestos para la inhibición de la quinasa ALK se describen más arriba.

Inhibidores indirectos

Los compuestos inhibidores de ALK útiles en la práctica de los métodos descritos también pueden ser compuestos que inhiben indirectamente la actividad de ALK mediante la inhibición de la actividad de proteínas o moléculas distintas de la propia quinasa ALK. Tales agentes terapéuticos inhibidores pueden ser inhibidores dirigidos que modulan la actividad de quinasas reguladoras clave que fosforilan o desfosforilan (y por lo tanto activan o desactivan) la propia ALK, o interfieren en la unión de ligandos. Al igual que con otras tirosina quinasas receptoras, ALK regula la señalización aguas abajo a través de una red de proteínas adaptadoras y quinasas aguas abajo, incluyendo STAT5 y AKT. Como resultado, la inducción del crecimiento y la supervivencia celulares por la actividad de ALK puede inhibirse eligiendo como diana estas proteínas que interactúan o aguas abajo. Los fármacos actualmente en desarrollo que podrían usarse de esta manera incluyen la Wartmanina.

La actividad de la quinasa ALK también se puede inhibir indirectamente mediante el uso de un compuesto que inhibe la unión de una molécula activadora necesaria para que ALK adopte su conformación activa. Del mismo modo, por ejemplo, se han descrito la producción y el uso de anticuerpos anti-PDGF para regular a la baja la tirosina quinasa receptora de PDGF. Véase la Publicación de Patente de los Estados Unidos Núm. 20030219839, "Anti-PDGF Antibodies and Methods for Producing Engineered Antibodies," Bowdish *et al.*

Los inhibidores indirectos de la actividad de ALK se pueden diseñar racionalmente utilizando modelado cristalográfico de rayos X o por ordenador de la estructura tridimensional de ALK, o se pueden encontrar mediante escrutinio de alto rendimiento de bibliotecas de compuestos para la inhibición de enzimas reguladoras clave aguas arriba y/o moléculas de unión necesarias, que da como resultado la inhibición de la actividad de la quinasa ALK. Tales enfoques son bien conocidos en la técnica, y se han descrito. La inhibición de ALK por tales agentes terapéuticos se puede confirmar, por ejemplo, examinando la capacidad del compuesto para inhibir la actividad de ALK, pero no otra actividad quinasa, en un panel de quinasas, y/o examinando la inhibición de la actividad de ALK en una muestra biológica que comprende células cancerosas, p. ej., células de CPCNP, como se ha descrito anteriormente. Los métodos para la identificación de compuestos que inhiben un cáncer caracterizado por un polinucleótido de fusión y/o un polipéptido de fusión de EML4-ALK o TFG-ALK se describen adicionalmente más adelante.

Inhibidores anti-sentido y/o de la transcripción

Los agentes terapéuticos inhibidores de ALK también pueden comprender compuestos inhibidores anti-sentido y/o de la transcripción que inhiben la actividad de la quinasa ALK mediante el bloqueo de la transcripción del gen que codifica ALK y/o los genes de fusión de EML4-ALK o TFG-ALK o genes de ALK truncados. Por ejemplo, se ha descrito la inhibición de diversas quinasas receptoras, incluyendo VEGFR, EGFR, e IGFR, y FGFR, por agentes terapéuticos antisentido para el tratamiento del cáncer. Véanse, p. ej., las Patentes de los Estados Unidos Núms. 6.734.017; 6.710.174; 6.617.162; 6.340.674; 5.783.683; 5.610.288.

Los oligonucleótidos antisentido pueden ser diseñados, construidos, y empleados como agentes terapéuticos contra genes diana de acuerdo con técnicas conocidas. Véanse, p. ej. Cohen, J., Trends in Pharmacol. Sci. 10(11): 435-437 (1989); Marcus-Sekura, *Anal. Biochem.* 172: 289-295 (1988); Weintraub, H., *Sci. AM.* págs. 40-46 (1990); Van

Der Krol et al., *BioTechniques* 6(10): 958-976 (1988); Skorski et al., *Proc. Natl. Acad. Sci. USA* (1994) 91:4504-4508. Recientemente se ha descrito la inhibición del crecimiento de carcinoma humano *in vivo* utilizando un inhibidor de ARN antisentido de EGFR. Véase la Publicación de Patente de los Estados Unidos Núm. 20040047847, "Inhibition of Human Squamous Cell Carcinoma Growth in vivo by Epidermal Growth Factor Receptor Antisense RNA Transcribed from a Pol III Promoter," 11 de Marzo de 2004, He *et al.* Del mismo modo, se puede preparar un agente terapéutico inhibidor de ALK que comprende al menos un oligonucleótido antisentido contra un gen de ALK de mamífero (véase la Figura 4 (SEQ ID NO: 6)) o un polinucleótido de fusión de EML4-ALK o TFG-ALK o un polinucleótido de ALK truncado (véanse las Figuras 2A-C (SEQ ID NO: 2, 19, y 21)) de acuerdo con los métodos descritos anteriormente. Las composiciones farmacéuticas que comprenden compuestos antisentido inhibidores de ALK se pueden preparar y administrar como se describe adicionalmente a continuación.

ARN de interferencia pequeño

Las composiciones de moléculas de ARN de interferencia pequeño (ARNip), que inhiben la traducción, y por lo tanto la actividad de ALK a través del proceso de interferencia de ARN, también se pueden emplear deseablemente en la invención. La interferencia de ARN, y el silenciamiento selectivo de la expresión de la proteína diana mediante la introducción de moléculas pequeñas exógenas de ARN de doble hebra que comprenden la secuencia complementaria al ARNm que codifica la proteína diana, ha sido bien descrita. Véanse, p. ej. la Publicación de Patente de los Estados Unidos Núm. 20040038921, "Composition and Method for Inhibiting Expression of a Target Gene", 26 de febrero de 2004, Kreutzer *et al.*; Publicación de Patente de los Estados Unidos Núm. 20020086356, "RNA Sequence-Specific Mediators of RNA Interference", 12 de Junio, 2003, Tuschl *et al.*; la Publicación de Patente de los Estados Unidos Núm. 20040229266, "RNA Interference Mediating Small RNA Molecules," 18 de Noviembre de 2004, Tuschl *et al.*

Por ejemplo, como se muestra en la presente (véase el Ejemplo 2), silenciamiento mediado por ARNip de la expresión de la proteína de fusión de EML4-ALK en una línea celular de CPCNP humano que expresa la proteína de fusión inhibió selectivamente la progresión de la enfermedad en esas células, pero no en células de control que no expresan la proteína de ALK mutante.

Se ha demostrado que las moléculas de ARN de doble hebra (ARNdh) bloquean la expresión génica en un mecanismo regulador muy conservado conocido como interferencia de ARN (ARNi). Brevemente, la RNasa III Dicer procesa el ARNdh a ARN de interferencia pequeños (ARNip) de aproximadamente 22 nucleótidos, que sirven como secuencias guía para inducir la escisión del ARNm específica de la diana por un complejo de silenciamiento inducido por ARN RISC (véase Hammond y otros, *Nature* (2000) 404: 293-296). El ARNi implica una reacción de tipo catalítico por medio de la cual se generan nuevos ARNip a través de la escisión sucesiva de ARNdh más largo. De este modo, a diferencia del antisentido, el ARNi degrada el ARN diana de una manera no estequiométrica. Cuando se administra a una célula u organismo, se ha demostrado que el ARNdh exógeno dirige la degradación específica de la secuencia del ARN mensajero endógeno (ARNm) a través del ARNi.

En la actualidad se encuentra disponible en el mercado una amplia variedad de productos de ARNip específicos de dianas, que incluyen vectores y sistemas para su expresión y su uso en células de mamíferos. Véase, p. ej., Promega, Inc. (promega.com); Dharmacon, Inc. (dharmacon.com). Se encuentran disponibles manuales técnicos detallados sobre el diseño, construcción y uso de ARNdh para ARNi. Véase, p. ej., Dharmacon's "RNAi Technical Reference & Application Guide"; Promega's "RNAi: A Guide to Gene Silencing." Los productos de ARNip inhibidores de ALK también se encuentran disponibles en el mercado, y pueden emplearse adecuadamente en la invención. Véase, p. ej. Dharmacon, Inc., Lafayette, CO (Núms. de Cat. M-003162-03, MU-003162-03, D-003162-07 a -10 (ARNip siGENOME™ SMARTselection y SMARTpool®).

Se ha establecido recientemente que los ARNdh pequeños de menos de 49 nucleótidos de longitud, y preferiblemente de 19-25 nucleótidos, que comprenden al menos una secuencia que es sustancialmente idéntica a parte de una secuencia de ARNm diana, y cuyo ARNdh tiene de manera óptima al menos un saliente de 1-4 nucleótidos en un extremo, son los más eficaces en la mediación de ARNi en mamíferos. Véase la Publicación de Patente de los Estados Unidos Núm. 20040038921, Kreutzer *et al.*, más arriba; la Publicación de Patente de los Estados Unidos Núm. 20040229266, Tuschl *et al.*, más arriba. La construcción de tales ARNdh, y su uso en preparaciones farmacéuticas para silenciar la expresión de una proteína diana, *in vivo*, se describen en detalle en tales publicaciones.

Si se conoce la secuencia del gen que se va a elegir como diana en un mamífero, se puede producir ARN de 21-23 nt, por ejemplo, y se puede someter a ensayo para determinar su capacidad para mediar en ARNi en una célula de mamífero, tal como una célula humana o de otro primate. Esas moléculas de ARN de 21-23 nt que se ha demostrado que median en ARNi pueden someterse a ensayo, si se desea, en un modelo animal apropiado para evaluar adicionalmente su eficacia *in vivo*. Los sitios diana que se conocen, por ejemplo sitios diana que se ha determinado que son sitios diana eficaces basándose en estudios con otras moléculas de ácido nucleico, por ejemplo ribozimas o antisentido, o aquellas dianas que se sabe que están asociadas con una enfermedad o afección, tales como aquellos sitios que contienen mutaciones o deleciones, se pueden utilizar para diseñar moléculas de ARNip que se dirigen también a estos sitios.

Alternativamente, las secuencias de ARNd_h eficaz se pueden diseñar/predecir racionalmente escrutando el ARNm diana de interés para determinar los sitios diana, por ejemplo mediante el uso de un algoritmo de plegamiento por ordenador. La secuencia diana se puede analizar *in silico* en una lista de todos los fragmentos o subsecuencias de una longitud concreta, por ejemplo, fragmentos de 23 nucleótidos, utilizando un comando Perl adaptado o programas de análisis de secuencias comerciales tales como Oligo, MacVector o GCG Wisconsin Package.

Se pueden utilizar varios parámetros para determinar qué sitios son los sitios diana más adecuados dentro de la secuencia de ARN diana. Estos parámetros incluyen, pero no se limitan a la estructura secundaria o terciaria del ARN, la composición de base de nucleótidos de la secuencia diana, el grado de homología entre diversas regiones de la secuencia diana, o la posición relativa de la secuencia diana dentro del transcrito de ARN. Sobre la base de estas determinaciones, se puede seleccionar cualquier número de sitios diana dentro del transcrito de ARN para escrutar las moléculas de ARN_i para determinar la eficacia, por ejemplo, mediante el uso de análisis de escisión del ARN *in vitro*, cultivo celular, o modelos animales. Véase, p. ej., la Publicación de Patente de los Estados Unidos Núm. 20030170891, 11 de Septiembre 2003, McSwiggen J. También se ha descrito recientemente un algoritmo para la identificación y la selección de sitios diana de ARN_i. Véase la Publicación de Patente de los Estados Unidos Núm. 20040236517 "Selection of Target Sites for Antisense Attack of RNA," 25 de Noviembre de 2004, Drica *et al.*

Las técnicas de transferencia génica utilizadas comúnmente incluyen métodos con fosfato de calcio, DEAE-dextrano, electroporación y microinyección y virales (Graham *et al.* (1973) *Viol.* 52: 456; McCutchan *et al.*, (1968)), *J. Natl. Cáncer Inst.* 41: 351; Chu *et al.* (1987), *Nucl. Acids Res.* 15: 1311; Fraley *et al.* (1980), *J. Biol. Chem.* 255: 10431; Capecchi (1980), *Cell* 22: 479). El ADN también se puede introducir en las células usando liposomas catiónicos (Felgner *et al.* (1987), *Proc. Natl. Acad. Sci. USA* 84: 7413). Las formulaciones de lípidos catiónicos disponibles en el mercado incluyen Tfx 50 (Promega) o Lipofectamine 200 (Life Technologies). Alternativamente, los vectores virales se pueden emplear para liberar ARNd_h en una célula y mediar en ARN_i. Véase la Publicación de Patente de los Estados Unidos Núm. 20040023390, "ARN_i-mediated Gene Silencing with Viral Vectors," 4 de Febrero de 2004, Davidson *et al.*

Los sistemas de transfección y de expresión con vectores para ARN_i en células de mamífero están disponibles en el mercado y han sido bien descritos. Véase, p. ej., el sistema Dharmacon, Inc., DharmaFECT™; Promega, Inc., el sistema siSTRIKE™ U6 Hairpin; véanse también Gou *et al.* (2003) *FEBS*. 548, 113-118; Sui, G. *et al.* A DNA vector-based RNAi technology to suppress gene expression in mammalian cells (2002) *Proc. Natl. Acad. Sci.* 99, 5515-5520; Yu *et al.* (2002) *Proc. Natl. Acad. Sci.* 99, 6047-6052; Paul, C. *et al.* (2002) *Nature Biotechnology* 19, 505-508; McManus *et al.* (2002) *RNA* 8, 842-850.

La interferencia con ARN_i en un mamífero utilizando moléculas de ARNd_h preparadas se puede llevar a cabo a continuación mediante la administración de una preparación farmacéutica que comprende el ARNd_h al mamífero. La composición farmacéutica se administra a una dosis suficiente para inhibir la expresión del gen diana. El ARNd_h puede administrarse típicamente a una dosis de menos de 5 mg de ARNd_h por kilogramo de peso corporal por día, y es suficiente para inhibir o suprimir la expresión del gen diana completamente. En general, una dosis adecuada de ARNd_h estará en el intervalo de 0,01 a 2,5 miligramos por kilogramo de peso corporal del receptor por día, preferiblemente en el intervalo de 0,1 a 200 microgramos por kilogramo de peso corporal por día, más preferiblemente en el intervalo de 0,1 a 100 microgramos por kilogramo de peso corporal por día, incluso más preferiblemente en el intervalo de 1,0 a 50 microgramos por kilogramo de peso corporal por día, y lo más preferiblemente en el intervalo de 1,0 a 25 microgramos por kilogramo de peso corporal por día. Una composición farmacéutica que comprende el ARNd_h se administra una vez al día, o en múltiples subdosis, por ejemplo, utilizando formulaciones de liberación sostenida bien conocidas en la técnica. La preparación y administración de tales composiciones farmacéuticas pueden llevarse a cabo de acuerdo con técnicas convencionales, como se describe adicionalmente a continuación.

Tal ARNd_h se puede utilizar después para inhibir la expresión y la actividad de ALK en un cáncer, preparando una preparación farmacéutica que comprende una cantidad terapéuticamente eficaz de tal ARNd_h, como se describe más arriba, y administrando la preparación a un sujeto humano que tiene un cáncer que expresa la proteína de fusión de EML4-ALK o TFG-ALK o la quinasa ALK activa truncada, por ejemplo, a través de inyección directa en el tumor. Recientemente se ha descrito la inhibición similar de otras tirosina quinasas receptoras, tales como VEGFR y EGFR utilizando inhibidores ARN_i. Véanse la Publicación de Patente de los Estados Unidos Núm. 20040209832, 21 de Octubre de 2004, McSwiggen *et al.*; Publicación de Patente de los Estados Unidos Núm. 20030170891, 11 de Septiembre de 2003, McSwiggen; Publicación de Patente de los Estados Unidos Núm. 20040175703, 9 de Septiembre de 2004, Kreutzer *et al.*

Composiciones terapéuticas; Administración.

Las composiciones terapéuticas inhibidoras de la quinasa ALK útiles en la invención se pueden administrar a un mamífero mediante cualquier medio conocido en la técnica incluyendo, pero no limitado a las rutas oral o peritoneal, incluyendo la administración intravenosa, intramuscular, intraperitoneal, subcutánea, transdérmica, vías respiratorias (aerosol), rectal, vaginal y tópica (incluyendo bucal y sublingual).

Para la administración oral, se proporcionarán generalmente un agente terapéutico inhibidor de ALK en forma de comprimidos o cápsulas, como polvo o gránulos, o como una solución o suspensión acuosa. Los comprimidos para uso oral pueden incluir los ingredientes activos mezclados con excipientes farmacéuticamente aceptables tales como diluyentes inertes, agentes disgregantes, agentes aglutinantes, agentes lubricantes, agentes edulcorantes, agentes aromatizantes, agentes colorantes y conservantes. Los diluyentes inertes adecuados incluyen carbonato de sodio y de calcio, fosfato de sodio y de calcio, y lactosa, mientras que el almidón de maíz y el ácido algínico son agentes disgregantes adecuados. Los agentes aglutinantes pueden incluir almidón y gelatina, mientras que el agente lubricante, si está presente, será generalmente estearato de magnesio, ácido esteárico o talco. Si se desea, los comprimidos pueden recubrirse con un material tal como monoestearato de glicerilo o diestearato de glicerilo, para retrasar la absorción en el tracto gastrointestinal.

Las cápsulas para uso oral incluyen cápsulas de gelatina dura en las que el ingrediente activo se mezcla con un diluyente sólido, y cápsulas de gelatina blanda en las que los ingredientes activos se mezclan con agua o un aceite tal como aceite de cacahuete, parafina líquida o aceite de oliva. Para uso intramuscular, intraperitoneal, subcutáneo e intravenoso, las composiciones farmacéuticas descritas en la presente memoria se proporcionarán generalmente en soluciones o suspensiones acuosas estériles, tamponadas a un pH y una isotonicidad apropiados. Los vehículos acuosos adecuados incluyen disolución de Ringer y cloruro sódico isotónico. El portador puede consistir exclusivamente en un tampón acuoso ("exclusivamente" significa que no se encuentran presentes agentes auxiliares o sustancias encapsulantes que pudieran afectar o mediar la absorción del agente terapéutico inhibidor de ALK). Tales sustancias incluyen, p. ej., estructuras micelares, tales como liposomas o cápsidas, como se describe a continuación. Las suspensiones acuosas pueden incluir agentes suspensores tales como derivados de celulosa, alginato de sodio, polivinilpirrolidona y goma de tragacanto, y un agente humectante tal como lecitina. Los conservantes adecuados para suspensiones acuosas incluyen p-hidroxibenzoato de etilo y n-propilo.

Las composiciones terapéuticas inhibidoras de quinasa ALK también pueden incluir formulaciones encapsuladas para proteger el agente terapéutico (p. ej. un compuesto de ARNdh) frente a la rápida eliminación del organismo, tal como una formulación de liberación controlada, incluyendo implantes y sistemas de liberación microencapsulados. Se pueden utilizar polímeros biocompatibles, biodegradables, tales como etileno-acetato de vinilo, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y poli(ácido láctico). Los métodos para la preparación de tales formulaciones serán evidentes para los expertos en la técnica. Los materiales también se pueden obtener comercialmente de Alza Corporation y Nova Pharmaceuticals, Inc. Las suspensiones liposomales (incluyendo liposomas dirigidos a células infectadas con anticuerpos monoclonales para antígenos virales) también se pueden utilizar como portadores farmacéuticamente aceptables. Estas se pueden preparar de acuerdo con métodos conocidos por los expertos en la técnica, p. ej., como se describe en la Patente de los Estados Unidos Núm. 4.522.811; La publicación PCT WO 91/06309; y la Publicación de Patente Europea EP-A-43075. Una formulación encapsulada puede comprender una proteína de la cubierta viral. La proteína de la cubierta viral puede derivar de o asociarse con un virus, tal como un virus de polio, o puede ser parcial o totalmente artificial. Por ejemplo, la proteína de la cubierta puede ser una Proteína Viral 1 y/o la Proteína Viral 2 del virus del polio, o un derivado de las mismas.

Las composiciones inhibidoras de ALK también pueden comprender un vehículo de liberación, incluyendo liposomas, para la administración a un sujeto, portadores y diluyentes y sus sales, y/o pueden estar presentes en formulaciones farmacéuticamente aceptables. Por ejemplo, los métodos para la liberación de moléculas de ácido nucleico son descritos por Akhtar et al., 1992, *Trends Cell Bio.*, 2, 139; DELIVERY STRATEGIES FOR ANTISENSE OLIGONUCLEOTIDE THERAPEUTICS, ed. Akhtar, 1995, Maurer et al., 1999, *Mol. Membr. Biol.*, 16, 129-140; Hofland and Huang, 1999, *Handb. Exp. Pharmacol.*, 137, 165-192; y Lee et al., 2000, *ACS Symp. Ser.*, 752, 184-192. Beigelman et al., Patente de los Estados Unidos Núm. 6.395.713 y Sullivan et al., PCT WO 94/02595 describen adicionalmente los métodos generales para la liberación de moléculas de ácido nucleico. Estos protocolos se pueden utilizar para la liberación de prácticamente cualquier molécula de ácido nucleico.

Se pueden administrar agentes terapéuticos inhibidores de ALK a un tumor de mamífero mediante una variedad de métodos conocidos por los expertos en la técnica, incluyendo, pero no limitados a, encapsulación en liposomas, mediante iontoforesis, o mediante incorporación en otros vehículos, tales como hidrogeles, ciclodextrinas, nanocápsulas biodegradables, y microesferas bioadhesivas, o mediante vectores proteínicos (O'Hare y Normand, Publicación Internacional PCT Núm. WO 00/53722). Alternativamente, la combinación de agente terapéutico/vehículo se libera localmente mediante inyección directa o mediante el uso de una bomba de infusión. La inyección directa de la composición, ya sea subcutánea, intramuscular, o intradérmica, puede tener lugar utilizando metodologías de agujas y jeringas convencionales, o tecnologías sin agujas tales como las descritas por Conry et al., 1999, en *Clin. Cancer Res.*, 5, 2330-2337 y Barry et al., Publicación PCT Internacional Núm. WO 99/31262.

Las formulaciones farmacéuticamente aceptables de agentes terapéuticos inhibidores de la quinasa ALK incluyen sales de los compuestos anteriormente descritos, p. ej., sales de adición de ácido, por ejemplo, sales de ácido clorhídrico, bromhídrico, acético, y ácido bencenosulfónico. Una composición o formulación farmacológica se refiere a una composición o formulación en una forma adecuada para la administración, p. ej., administración sistémica, en una célula o paciente, incluyendo por ejemplo un ser humano. Las formas adecuadas dependen, en parte, del uso o de la ruta de entrada, por ejemplo oral, transdérmica, o mediante inyección. Tales formas no deberían impedir que la composición o formulación alcance una célula diana. Por ejemplo, las composiciones farmacológicas inyectadas en

el torrente sanguíneo deben ser solubles. Se conocen otros factores en la técnica, e incluyen consideraciones tales como toxicidad y formas que impiden que la composición o formulación ejerzan su efecto.

Las rutas de administración que conducen a la absorción sistémica (es decir absorción sistémica o acumulación de fármacos en el torrente sanguíneo seguido de distribución en todo el organismo), son deseables e incluyen, sin limitación: intravenosa, subcutánea, intraperitoneal, inhalación, oral, intrapulmonar e intramuscular. Cada una de estas rutas de administración expone el agente terapéutico inhibidor de ALK a un tejido enfermo o tumor accesibles. Se ha demostrado que la tasa de entrada de un fármaco en la circulación es una función del peso molecular o el tamaño. El uso de un liposoma u otro portador de fármaco que comprende los compuestos para usar según la presente invención puede localizar potencialmente el fármaco, por ejemplo, en ciertos tipos de tejidos, tales como los tejidos del sistema reticuloendotelial (RES). Una formulación de liposomas que puede facilitar la asociación del fármaco con la superficie de las células, tales como, linfocitos y macrófagos también es útil. Este enfoque puede proporcionar una liberación mejorada del fármaco en las células diana mediante el aprovechamiento de la especificidad del reconocimiento inmunitario de los macrófagos y linfocitos de células anormales, tales como las células cancerosas.

Por "formulación farmacéuticamente aceptable" se quiere significar, una composición o formulación que permite la distribución eficaz de las moléculas de ácido nucleico para usar según la presente invención en la localización física más adecuada para su actividad deseada. Los ejemplos no limitantes de agentes los adecuados para la formulación con las moléculas de ácido nucleico para usar según la presente invención incluyen: inhibidores de P-glicoproteína (tales como Pluronic P85), que pueden mejorar la entrada de fármacos en el SNC (Jolliet-Riant y Tillement, 1999, *Fundam. Clin. Pharmacol.*, 13, 16-26); polímeros biodegradables, tales como microesferas poli(DL-lactida-coglicolida) para la administración de liberación sostenida después de la implantación intracerebral (Emerich et al., 1999, *Cell Transplant*, 8, 47-58) (Alkermes, Inc. Cambridge, Mass.); y nanopartículas cargadas, tales como las elaboradas de poli(cianoacrilato de butilo), que pueden liberar fármacos a través de la barrera hematoencefálica y pueden modificar mecanismos de absorción neuronales (*Prog. Neuro-Psychopharmacol Biol Psychiatry*, 23, 941-949, 1999). Otros ejemplos no limitantes de estrategias de liberación de los compuestos inhibidores de ALK para usar según la invención incluyen el material descrito por Boado et al., 1998, en *J. Pharm. Sci.*, 87, 1308-1315; Tyler et al., 1999, *FEBS Lett.*, 421, 280-284; Pardridge et al., 1995, *PNAS USA*, 92, 5592-5596; Boado, 1995, *Adv. Drug Delivery Rev.*, 15, 73-107; Aldrian-Herrada et al., 1998, *Nucleic Acids Res.*, 26, 4910-4916; y Tyler et al., 1999, *PNAS USA*, 96, 7053-7058.

Las composiciones terapéuticas que comprenden liposomas modificados en la superficie que contienen lípidos de poli(etilenglicol) (modificados con PEG o liposomas de circulación prolongada o liposomas encubiertos) también se pueden emplear adecuadamente para usar según la invención. Estas formulaciones ofrecen un método para aumentar la acumulación de fármacos en tejidos diana. Esta clase de portadores de fármacos resiste opsonización y la eliminación por el sistema mononuclear fagocítico (MPS o RES), permitiendo de este modo tiempos de circulación de la sangre más largos y un aumento de la exposición en el tejido para el fármaco encapsulado (Lasic et al. *Chem. Rev.* 1995, 95, 2601-2627; Ishiwata et al., *Chem. Pharm. Bull.* 1995, 43, 1005-1011). Se ha demostrado que tales liposomas se acumulan selectivamente en los tumores, presumiblemente por la extravasación y captura en los tejidos diana neovascularizados (Lasic et al., *Science* 1995, 267, 1275-1276; Oku et al., 1995, *Biochim. Biophys. Acta*, 1238, 86-90). Los liposomas de circulación prolongada mejoran la farmacocinética y la farmacodinámica del ADN y el ARN, particularmente en comparación con los liposomas catiónicos convencionales que se sabe que se acumulan en tejidos del MPS (Liu et al., *J. Biol. Chem.* 1995, 270, 24864-24870; Choi et al., *Publicación Internacional PCT Núm. WO 96/10391*; Ansell et al., *Publicación Internacional PCT Núm. WO 96/10390*; Holland et al., *Publicación Internacional PCT Núm. WO 96/10392*). También es probable que los liposomas de circulación prolongada protejan los fármacos de la degradación por nucleasas en un mayor grado en comparación con los liposomas catiónicos, basándose en su capacidad para evitar la acumulación en los tejidos del MPS metabólicamente agresivos tales como el hígado y el bazo.

Las composiciones terapéuticas pueden incluir una cantidad farmacéuticamente eficaz de los compuestos deseados en un portador o diluyente farmacéuticamente aceptable. Los portadores o diluyentes aceptables para uso terapéutico son bien conocidos en la técnica farmacéutica, y se describen, por ejemplo, en REMINGTON'S PHARMACEUTICAL SCIENCES, Mack Publishing Co. (A.R. Gennaro, ed. 1985). Por ejemplo, se pueden proporcionar conservantes, estabilizadores, colorantes y agentes aromatizantes. Estos incluyen benzoato de sodio, ácido sórbico y ésteres de ácido p-hidroxibenzoico. Además, se pueden utilizar antioxidantes y agentes suspensorios.

Una dosis farmacéuticamente eficaz es la dosis requerida para prevenir, inhibir la aparición o tratar (aliviar un síntoma en cierta medida, preferentemente todos los síntomas) de un estado de enfermedad. La dosis farmacéuticamente eficaz depende del tipo de enfermedad, la composición usada, la vía de administración, el tipo de mamífero que está siendo tratado, las características físicas del mamífero específico bajo consideración, la medicación concurrente y otros factores que reconocerán los expertos en las técnicas médicas. En general, se administra una cantidad entre 0,1 mg/kg y 100 mg/kg de peso corporal/día de los ingredientes activos dependiendo de la potencia del polímero cargado negativamente.

Los niveles de dosificación del orden de aproximadamente 0,1 mg a aproximadamente 140 mg por kilogramo de

- peso corporal por día son útiles en el tratamiento de las afecciones indicadas anteriormente (aproximadamente 0,5 mg a aproximadamente 7 g por paciente por día). La cantidad de ingrediente activo que se puede combinar con los materiales portadores para producir una sola forma de dosificación varía dependiendo del anfitrión tratado y del modo concreto de administración. Las formas de dosificación unitarias generalmente contienen entre
- 5 aproximadamente 1 mg y aproximadamente 500 mg de un ingrediente activo. Se entiende que el nivel de dosificación específico para cualquier paciente concreto depende de una variedad de factores incluyendo la actividad del compuesto específico empleado, la edad, el peso corporal, la salud general, el sexo, la dieta, el tiempo de administración, la ruta de administración y la tasa de excreción, la combinación de fármacos y la gravedad de la enfermedad concreta que experimenta la terapia.
- 10 Para la administración a animales no humanos, la composición también se puede añadir al pienso animal o al agua de bebida. Puede ser conveniente formular las composiciones de pienso animal y de agua de bebida de modo que el animal tome una cantidad terapéuticamente apropiada de la composición junto con su dieta. También puede ser conveniente presentar la composición como una premezcla para la adición al pienso o al agua de bebida.
- 15 Un agente terapéutico inhibidor de ALK para usar según la invención puede comprender un único compuesto como se ha descrito anteriormente, o una combinación de múltiples compuestos, ya sea en la misma clase de inhibidor (es decir, anticuerpo inhibidor), o en diferentes clases (es decir, anticuerpos inhibidores e inhibidores de molécula pequeña). Tal combinación de compuestos puede aumentar el efecto terapéutico global en la inhibición de la progresión de un cáncer que expresa la proteína de fusión. Por ejemplo, la composición terapéutica puede ser un inhibidor de molécula pequeña, tal como WHI-131 y/o WHI-154 solo, o combinado con otros inhibidores dirigidos a la
- 20 actividad de ALK y/u otros inhibidores de molécula pequeña. La composición terapéutica también puede comprender uno o más agente quimioterapéuticos no específicos, además de uno o más inhibidores selectivos. Se ha demostrado recientemente que tales combinaciones proporcionan un efecto sinérgico eliminador del tumor en muchos cánceres. La eficacia de tales combinaciones en la inhibición de la actividad de ALK y el crecimiento del tumor *in vivo* se puede evaluar como se describe a continuación.
- 25 Identificación de compuestos inhibidores de quinasa ALK mutante.
- La descripción también proporciona, en parte, un método para determinar si un compuesto inhibe la progresión de un cáncer caracterizado por un polinucleótido de fusión y/o un polipéptido de fusión de EML4-ALK o TFG-ALK, mediante la determinación de si el compuesto inhibe la actividad del polipéptido de fusión de EML4-ALK o TFG-ALK o un polipéptido de quinasa ALK truncado en el cáncer. En algunas realizaciones preferidas, la inhibición de la
- 30 actividad de ALK se determina mediante el examen de una muestra biológica que comprende células de la médula ósea, sangre, efusión pleural, o un tumor. En otra realización preferida, la inhibición de la actividad de ALK se determina utilizando al menos un reactivo específico de un polinucleótido o polipéptido de ALK mutante de la descripción.
- 35 El compuesto sometido a ensayo puede ser cualquier tipo de agente terapéutico o composición como se ha descrito anteriormente. Los métodos para evaluar la eficacia de un compuesto, tanto *in vitro* como *in vivo*, están bien establecidos y son conocidos en la técnica. Por ejemplo, se puede someter a ensayo una composición para determinar su capacidad para inhibir ALK *in vitro* utilizando una célula o extracto celular en el que se activa ALK. Se puede emplear un panel de compuestos para someter a ensayo la especificidad del compuesto para ALK (a diferencia de otras dianas, tales como EGFR o PDGFR).
- 40 Otra técnica para el escrutinio de fármacos que se puede utilizar proporciona el escrutinio de alto rendimiento de compuestos que tienen una afinidad de unión adecuada a una proteína de interés, como se describe en la Solicitud PCT publicada WO84/03564. En este método, tal como se aplica a polipéptidos mutantes de ALK, se sintetiza un gran número de diferentes compuestos de ensayo pequeños sobre un sustrato sólido, tal como alfileres de plástico o alguna otra superficie. Los compuestos de ensayo se hacen reaccionar con el polipéptido de ALK mutante, o
- 45 fragmentos del mismo, y se lavan. El polipéptido mutante unido (p. ej., polipéptido de fusión de EML4-ALK) se detecta a continuación, mediante métodos bien conocidos en la técnica. El polipéptido de ALK mutante purificado también se puede recubrir directamente sobre placas para su uso en las técnicas de escrutinio de fármacos anteriormente mencionadas. Alternativamente, se pueden utilizar anticuerpos no neutralizadores para capturar el péptido e inmovilizarlo sobre un soporte sólido.
- 50 A continuación se puede examinar un compuesto que se ha encontrado que es un inhibidor eficaz de la actividad de ALK *in vitro* para determinar su capacidad para inhibir la progresión de un cáncer que expresa un polipéptido de fusión de EML4-ALK o TFG-ALK y/o un polipéptido de quinasa ALK truncado, *in vivo*, utilizando, p. ej., xenoinjertos de mamífero que albergan tumores humanos, tales como CPCNP. De este modo, se pueden observar los efectos del fármaco en un entorno biológico que se asemeja más estrechamente a un paciente. La capacidad del fármaco
- 55 para modificar la señalización en las células cancerosas o células estromales circundantes se puede determinar mediante análisis con anticuerpos específicos de fosforilación. La eficacia del fármaco en la inducción de la muerte celular o la inhibición de la proliferación celular también se puede observar mediante análisis con marcadores específicos de apoptosis tales como caspasa 3 escindida y PARP escindida. Del mismo modo, se pueden emplear trasplantes de médula ósea de mamíferos (p. ej., ratones) que albergan leucemias humanas que están dirigidas por la proteína ALK mutante. En este procedimiento, las células de médula ósea que se sabe que están dirigidas por la
- 60

quinasa ALK mutante se trasplantan al ratón. Se puede controlar el crecimiento de las células cancerosas. El ratón se puede tratar a continuación con el fármaco, y observar externamente el efecto del tratamiento con el fármaco sobre el fenotipo o la progresión del cáncer. El ratón se sacrifica a continuación, y después se retira la médula ósea trasplantada para el análisis mediante, IHC y transferencia Western, etc.

5 La toxicidad y la eficacia terapéutica de tales compuestos se pueden determinar mediante procedimientos farmacéuticos convencionales en cultivos celulares o animales experimentales, p. ej., para determinar la DL50 (la dosis letal para el 50% de la población) y la DE50 (la dosis terapéuticamente eficaz en el 50% de la población). La razón de la dosis entre los efectos tóxicos y terapéuticos es el índice terapéutico y puede expresarse como la razón DL50/DE50. Se prefieren los compuestos que muestran índices terapéuticos elevados.

10 Los siguientes Ejemplos se proporcionan solo para ilustrar adicionalmente la invención y descripciones, y no se pretende que limiten su alcance, salvo lo dispuesto en las reivindicaciones adjuntas a la misma. La presente invención abarca modificaciones y variaciones de los métodos que se ilustran en la presente memoria que resultarían obvios para un experto normal en la técnica.

Ejemplo 1

15 Identificación de actividad quinasa de ALK en tumores sólidos mediante perfilado de fosfopéptidos global

A. Perfilado de líneas celulares de CPCNP humano.

Se examinó el perfil de fosforilación global de la activación de la quinasa en 22 líneas celulares de CPCNP humano, incluyendo H2228, utilizando una técnica recientemente descrita y poderosa para el aislamiento y la caracterización mediante espectrometría de masas de péptidos modificados a partir de mezclas complejas (la técnica "IAP", véase Rush et al., más arriba). La técnica IAP se llevó a cabo utilizando un anticuerpo específico de fosfotirosina (CELL SIGNALING TECHNOLOGY, INC, Beverly, MA, 2003/04 Núm. de Cat. 9411) para aislar y posteriormente caracterizar, péptidos que contienen fosfotirosina a partir de extractos de las líneas celulares de CPCNP.

Específicamente, se empleó el enfoque IAP para facilitar la identificación de tirosina quinasas responsables de la fosforilación de proteínas en cada una de las líneas celulares de CPCNP. En particular, se consideró la actividad quinasa atípica o inusual.

Cultivo Celular.

Se adquirieron reactivos para cultivo celular de Invitrogen, Inc. Se examinaron un total de 41 líneas celulares de CPCNP humano. Las líneas celulares de CPCNP humano, H520, H838, H1437, H1563, H1568, H1792, H1944, H2170, H2172, H2228, H2347, A549, H441, H1703, H1373, H358 se obtuvieron de la Colección de Cultivos Tipo Americana, y se cultivaron en medio RPMI 1640 con FBS al 10% y se ajustaron para que contuvieran L-glutamina 2 mM, 1,5 g/l de bicarbonato de sodio, 4,5 g/l de glucosa, HEPES 10 mM, piruvato de sodio 1,0 mM, penicilina/estreptomicina. Se adquirieron seis líneas celulares de CPCNP humano adicionales, HCC78, Cal-12T, HCC366, HCC15, HCC44, y Lou-NH91, de DSMZ, y se cultivaron en medio RPMI 1640 que contenía FBS al 10% y penicilina/estreptomicina. Las células se mantuvieron en una incubadora con CO₂ al 5% a 37°C.

35 Para los experimentos de precipitación de inmutafinidad e inmutotransferencia, las células se cultivaron hasta 80% de confluencia y a continuación mueren por inanición en medio RPMI sin FBS durante la noche antes de la recolección.

Inmunoprecipitación de fosfopéptidos.

Se lisaron 100 millones de células en tampón de lisis de urea (Hepes 20 mM, pH 8,0, urea 9 M, vanadato de sodio 1 mM, pirofosfato de sodio 2,5 mM, beta-glicerofosfato 1 mM). El producto lisado se sometió a sonicación y se aclaró por centrifugación. El producto lisado aclarado se redujo mediante DTT y se alquiló con yodoacetamida, como se ha descrito previamente (véase Rush et al., *Nat. Biotechnol.* 23(1): 94-101 (2005)). A continuación las muestras se diluyeron 4 veces con Hepes 20 mM para reducir la concentración de urea a 2 M, y se digirieron con tripsina durante toda la noche a temperatura ambiente con sacudimiento suave.

45 Los péptidos digeridos se purificaron groseramente con columnas Sep-Pak C18, como se ha descrito previamente (véase Rush et al., *más arriba*). El eluato se liofilizó y los péptidos secos se disolvieron en 1,4 ml de tampón MOPS IP (MOPS/NaOH 50 mM pH 7,2, Na₂PO₄ 10 mM, NaCl 50 mM) y la materia insoluble se eliminó por centrifugación. La inmunoprecipitación se realizó a 4°C durante la noche con 160 µg de anticuerpo Phospho-Tyrosine 100 (Cell Signaling Technology) acoplado a cuentas de proteína G agarosa (Roche). Las cuentas se lavaron después 3 veces con 1 ml de tampón MOPS IP y dos veces con 1 ml de dH₂O de grado HPLC en frío. Los fosfopéptidos se hicieron eluir de las cuentas con 60 µl de TFA al 0,1% seguido de una segunda elución con 40 µl TFA al 0,1% y las fracciones se reunieron. Los péptidos eluidos se concentraron utilizando una columna ZipTip (Millipore), y se analizaron con LC-MS/MS. Los espectros de masas se recogieron con un espectrómetro de masas de trampa de iones LTQ (ThermoFinnigan).

Análisis mediante espectrometría de masas LC-MS/MS

Los péptidos del eluato IP (100 µl) se concentraron y se separan de anticuerpo eluido utilizando puntas de extracción Stop and Go (StageTips) (véase Rappsilber et al., *Anal. Chem.*, 75(3): 663-70 (2003)). Los péptidos se hicieron eluir de las microcolumnas con 1 µl de MeCN al 60%, TFA al 0,1% en 7,6 µl de ácido acético al 0,4%/ácido heptafluorobutírico (HFBA) al 0,005%.

Cada muestra de fosfopéptidos se analizó mediante LC-MS por duplicado. Se cargó una columna microcapilar de sílice fundida (125 m x 18 cm) con resina de fase inversa C18 (Magic C18AQ, partículas de 5 µm, tamaño de poro 200 Å, Michrom Bioresources, Auburn, CA). Las muestras (4 µl) se cargaron en esta columna con un muestreador automático (LC Packings Famos, San Francisco, CA) y se hicieron eluir en el espectrómetro de masa por medio de un gradiente lineal de 55 min de acetonitrilo de 7 a 30% en ácido fórmico al 0,1%. El gradiente se liberó a aproximadamente ABC nl/min utilizando una bomba de HPLC binaria (Agilent 1100, Palo Alto, CA) con un divisor de flujo en línea. Se analizó la masa de los iones de los péptidos eluyentes con una trampa de iones lineal híbrida - aparato de resonancia de ión ciclotrón con transformada de Fourier 7 Tesla (LTQ-FT, Thermo Finnigan, San Jose, CA).

Se empleó un método con los siete-principales, por medio del cual se recogieron barridos de MS/MS dependientes de 7 datos en la trampa de iones lineal basándose en las mediciones realizadas durante el barrido de la inspección de MS anterior en la célula ICR, con la trampa de iones lineal y el aparato con transformada de Fourier funcionando al mismo tiempo. Los barridos de MS se realizaron a 375-1800 m/z con una diana de control automático de ganancia (AGC) de 8x10⁶ y una resolución de masa de 105. Para MS/MS el AGC fue de 8x10⁶, el tiempo de exclusión dinámica fue de 25 s, y los iones de una sola carga fueron rechazados por el escrutinio del estado de carga.

Análisis y asignaciones de la base de datos.

Las secuencias de péptidos se asignaron a espectros MS/MS utilizando soporte lógico TurboSequest (v. 27, rev.12) (ThermoFinnigan) y una base de datos de proteína humana IPI compuesta directa/inversas. Los parámetros de búsqueda fueron: tripsina como proteasa; tolerancia de masa del precursor 1,08 Da; modificación estática en cisteína (+57,02146, carboxamidometilación); y modificaciones dinámicas en serina, treonina y tirosina (+79,96633 Da, fosforilación), lisina (+8,01420, 13C615N2), arginina (+6,02013, 13C6) y metionina (+15,99491, oxidación). Se utilizó un enfoque de base de datos objetivo/señuelo para establecer criterios de filtrado de puntuación adecuados, de manera que la tasa de asignación de falsos positivos estimada fuera <1%. Además de superar los umbrales XCorr dependientes de la carga (z = 1, XCorr ≥ 1,5, para z = 2, XCorr ≥ 2,2, para z = 3, XCorr ≥ 3,3), se requirió que las asignaciones contuvieran fosfotirosina, tuvieran una precisión de masa de -5 a +25 ppm, y contuvieran residuos de lisina/arginina todos ligeros o todos pesados.

Las asignaciones que pasaron estos criterios se evaluaron adicionalmente utilizando un programa de cuantificación a la medida, Vista (Bakalarski et al., manuscrito en preparación) para calcular las áreas de los picos y en última instancia, una abundancia relativa entre las formas pesada y ligera de cada péptido. Los péptidos identificados con señal con respecto a ruido en el barrido de MS por debajo de 15 no se tuvieron en cuenta para la cuantificación. Para esos péptidos encontrados solamente en una de las condiciones se utilizó en lugar de eso la razón de señal con respecto a ruido.

Las búsquedas se realizaron frente a la base de datos NCBI humana lanzada el 24 de agosto 2004 que contenía 27.175 proteínas que permitían metionina oxidada (M+16) y fosforilación (Y+80) como modificaciones dinámicas. Todos los espectros de apoyo a la lista final de las secuencias asignadas (no mostrados aquí) fueron revisados por al menos tres científicos para establecer su credibilidad.

El análisis IAP anterior identificó más de 2.000 péptidos no redundantes que contenían fosfotirosina, más de 1.500 sitios de fosfotirosina, y más de 1.000 proteínas con tirosina fosforilada, la mayoría de las cuales son novedosas, a partir de las líneas celulares examinadas (datos no presentados). Se observó que las tirosina quinasas receptoras que se sabe que están implicadas en la señalización de CPCNP tienen tirosina fosforilada en muchas líneas celulares, tales como EGFR, Her2, Her3, de EphA2 y Met. Se observaron altos niveles de fosfopéptidos de EGFR en varias líneas celulares incluyendo HCC827 y H3255, dos líneas celulares conocidas por expresar niveles amplificados de formas activadas genéticamente de EGFR, lo que confirma que el método identifica las tirosina quinasas receptoras que se sabe que son activas en líneas celulares de CPCNP.

Tres líneas de células expresaron tirosina quinasas receptoras no observadas en líneas celulares distintas de CPCNP. Se observaron grandes cantidades de péptidos con tirosina fosforilada de Ros, ALK y PDGFR alfa en las líneas celulares HCC78, H2228, y H1703, respectivamente. La línea celular de CPCNP H2228, que expresa sumamente ALK, se seleccionó para un examen adicional.

B. Perfilado de muestras tumorales de CPCNP humano.

Con posterioridad se aplicó la técnica IAP, sustancialmente como se describe en la Parte A anterior, para examinar los perfiles de fosforilación globales de un panel de 154 muestras de tumores humanos de pacientes con CPCNP. Los tejidos se obtuvieron de la Second Xiangya Hospital, China.

Se cortaron en trozos pequeños muestras de tejido congeladas, se homogeneizaron en tampón de lisis (HEPES 20 mM, pH 8,0, urea 9 M, vanadato de sodio 1 mM, con un suplemento de pirofosfato de sodio 2,5 mM, b-glicerol-fosfato 1 mM, 1 ml de tampón de lisis para 100 mg de tejido congelado) utilizando un Polytron durante 2 periodos de 20 seg. cada vez. El producto homogeneizado se sometió a continuación a sonicación brevemente. El producto

5 lisado aclarado se redujo con DTT y se alquiló con yodoacetamida, como se ha descrito anteriormente (véase Rush et al., *Nat. Biotechnol.* 23(1): 94-101 (2005)). A continuación las muestras se diluyeron 4 veces con Hepes 20 mM para reducir la concentración de Urea a 2 M, y se digirieron por medio de tripsina durante la noche a temperatura ambiente con agitación suave.

10 Los péptidos digeridos se purificaron groseramente con columnas Sep-Pak C18, como se ha descrito anteriormente (véase Rush et al., más arriba). El eluato se liofilizó y los péptidos secos se disolvieron en 1,4 ml de tampón MOPS IP (MOPS/NaOH 50 mM pH 7,2, Na₂PO₄ 10 mM, NaCl 50 mM) y la materia insoluble se eliminó por centrifugación. La inmunoprecipitación se llevó a cabo a 4°C durante la noche con 160 µg de anticuerpo Phospho-Tyrosine 100 (Cell Signaling Technology) acoplado a cuentas de proteína G-agarosa (Roche). Las cuentas se lavaron 3 veces con 1 ml de tampón MOPS IP y dos veces con 1 ml de dH₂O de grado HPLC en frío. Los fosfopéptidos se hicieron eluir

15 de las cuentas con 60 µl de TFA al 0,1% seguido de una segunda elución con 40 µl de TFA al 0,1% y las fracciones se reunieron. Los péptidos eluidos se concentraron utilizando una columna ZipTip (Millipore), y se analizaron con LC-MS/MS. Los espectros de masas se recogieron con un espectrómetro de masas con trampa de iones LTQ (ThermoFinnigan).

20 A continuación se llevó a cabo la inmunoprecipitación de los fosfopéptidos, seguido de análisis de espectrometría LC-MS/MS como se ha descrito anteriormente en la Parte A. La búsqueda en la base de datos y las asignaciones de secuencias se realizaron sustancialmente como se ha descrito anteriormente en la Parte A, pero utilizando la base de datos NCBI humana lanzada el 24 de agosto 2004 que contenía 27.970 proteínas.

25 El análisis IAP anterior identificó más de 2.000 péptidos no redundantes que contenían fosfotirosina, más de 1.500 sitios de fosfotirosina, y más de 1.000 proteínas con tirosina fosforilada de las muestras de tumores humanos examinadas (datos no mostrados). De nuevo se observó que las tirosina quinasas receptoras que se sabía que están implicadas en la señalización de CPCNP tenían tirosina fosforilada en muchos tumores, tales como EGFR, Her2, Her3, de EphA2 y Met. Se observaron de nuevo altos niveles de fosfopéptidos de EGFR en varias muestras tumorales, lo que confirma que el método identifica las tirosina quinasas receptoras que se sabe que son activas en las líneas celulares de CPCNP.

30 Cinco muestras de pacientes expresaron tirosina quinasas receptoras no observadas en líneas celulares y tumores distintos de CPCNP. Se observaron grandes cantidades de péptidos con tirosina fosforilada de ALK en pacientes CS010/11, CS045, y CS110. Estos tres tumores, que expresaban sumamente ALK, se seleccionaron para su examen adicional.

Ejemplo 2

35 Aislamiento y secuenciación de tres genes de fusión de ALK

A. Secuenciación en líneas celulares de CPCNP humano.

Dado el alto nivel de fosforilación de la quinasa ALK detectado en la línea celular de CPCNP H2228, se llevó a cabo una amplificación rápida 5' de los extremos de ADNc en la secuencia que codifica el dominio quinasa de ALK con el fin de determinar si se encontraba presente un transcrito de ALK quimérico.

40 Amplificación rápida de extremos de ADN complementarios

Se utilizó el RNeasy Mini Kit (Qiagen) para extraer el ARN de la línea celular H2228. El ADN se extrajo utilizando el DNeasy Tissue Kit (Qiagen). La amplificación rápida de los extremos del ADNc se realizó utilizando el sistema 5' RACE (Invitrogen) con los cebadores ALK-GSP1 para la síntesis de ADNc y ALK-GSP2 y ALK-GSP3 para una reacción de PCR anidada.

45 5' RACE

La Figura 5 (panel A) muestra la detección del gen de fusión EML4-ALK (variante corta) mediante 5'RACE y la detección del producto de amplificación de PCR después de 2 rondas. El producto de la PCR se purificó con el kit de purificación de PCR (Qiagen) y se secuenció utilizando ALK-GSP3, un secuenciador de ADN automático capilar ABI 3130 (Applied Biosystems). El análisis de secuencia del producto resultante reveló que el dominio quinasa y el extremo C de ALK estaban fusionados al extremo N del gen de EML-4 (véase la Figura 1, panel B). El gen de fusión de EML4-ALK (variante corta) estaba en marco y fusionaba los primeros 233 aminoácidos de EML-4 a los últimos 562 aminoácidos de ALK (véase la Figura 1, panel B). Los genes de EML-4 y ALK están ambos ubicados en el cromosoma 2, de modo que el gen de fusión fue creado por la delección de genes entre estos dos loci.

Se utilizaron los siguientes cebadores:

ALK-GSP1: 5'-GCAGTAGTTGGGGTTGTAGTC (SEQ ID NO: 9)

ALK-GSP2: 5'-GCGGAGCTTGCTCAGCTTGT (SEQ ID NO: 10)

ALK-GSP3: 5'-TGCAGCTCCTGGTGCTTCC (SEQ ID NO: 11)

Análisis mediante PCR

Se realizó un análisis de RT-PCR para confirmar que el extremo N de EML-4 estaba intacto en la proteína de fusión (véase la Figura 6 (panel B)). La primera hebra de ADNc se sintetizó a partir de 2,5 mg de ARN total utilizando el sistema de síntesis de la primera hebra Superscript™ III (Invitrogen) con oligo (dT)₂₀. A continuación, se amplificó el gen de fusión de EML4-ALK con el uso de los pares de cebadores EML-ATG y ALK-GSP3. La fusión recíproca se detectó con el uso de los pares de cebadores EML-4-43 y ALK-GSP3 y EML-4-94 y ALK-GSP3 y EML4-202 y ALK-GSP3. Para la PCR genómica, la amplificación del gen de fusión se realizó utilizando la ADN polimerasa Platinum Taq de alta fidelidad (Invitrogen) con los pares de cebadores EML-4-ATG y ALK-TGA.

Se utilizaron los siguientes cebadores:

ALK-GSP3: 5' - TGCAGCTCCTGGTGCTTCC (SEQ ID NO: 12)

EML4-Atg: 5' - CGCAAGATGGACGGTTTGGC (SEQ ID NO: 13)

EML4-43: 5' - TGTTCAGATCGCCTGTCAGCTCT (SEQ ID NO: 14)

EML4-94: 5' - TGAAATCACTGTGCTAAAGGCGGC (SEQ ID NO: 15)

EML4-202: 5' - AAGCCCTCGAGCAGTTATTTCCAT (SEQ ID NO: 16)

ALK-Tga: 5' - GAATTCCGCCGAGCTCAGGGCCCAG (SEQ ID NO: 17)

Es de destacar que, en la fusión de EML4-ALK (variante corta), el radical de ALK se fusiona al radical de EML-4, precisamente en el mismo punto en el se ha observado ALK en otras fusiones de ALK, tales como la fusión NPM-ALK que se producen en el LACG. El dominio quinasa de ALK en la línea celular H2228 se secuenció adicionalmente a partir de ADN genómico y se encontró que era de tipo salvaje. Por lo tanto, la mutación por delección descubierta en H2228 no afecta al dominio quinasa de ALK. Adicionalmente, EML-4 de tipo salvaje está fosforilada en tirosina solo en un sitio que no está presente en la proteína de fusión de EML4-ALK (variante corta), lo que sugiere que el dominio N-terminal de la bobina en espiral que se conserva en la proteína de fusión (véase la Figura 1A) pueden funcionar dimerizando y activando ALK, así como promoviendo la interacción con ALK de tipo salvaje.

B. Secuenciación en la línea celular de CPCNP humano.

De manera similar, dado el alto nivel de fosforilación de la quinasa ALK detectado en las muestras de tumores de CPCNP humanos de pacientes CS010/11, CS045, y CS110, se llevó a cabo una amplificación rápida 5' de los extremos de ADNc en la secuencia que codifica el dominio quinasa de ALK con el fin de determinar si el transcrito de ALK quimérico estaba presente en estos tumores.

La rápida amplificación de los extremos de ADN complementarios y 5' RACE se llevaron a cabo, sustancialmente como se ha descrito anteriormente en la Parte A, con los cebadores ALK-GSP1 para la síntesis de ADNc y ALK-GSP2 y ALK-GSP3 para una reacción de PCR anidada.

La Figura 5 (panel C) muestra la detección de los genes de fusión de EML4-ALK (variantes tanto corta como larga) mediante 5' RACE en dos muestras de pacientes, y la detección del gen de fusión TFG-ALK en un paciente, y la detección del producto de la amplificación de PCR después de 2 rondas. Los productos de la PCR se purificaron y secuenciaron sustancialmente como se ha descrito anteriormente en la Parte A. El análisis de secuencia de los productos resultantes reveló que el dominio quinasa y el extremo C de ALK se fusionaron al extremo N del gen de EML-4 en dos variantes diferentes (véanse las Figuras 1A-1B, panel B). Los genes de fusión de EML4-ALK estaban en marco y fusionaban los primeros 233 aminoácidos (variante corta) o los primeros 495 aminoácidos (variante larga) de EML-4 a los últimos 562 aminoácidos de ALK (véanse las Figuras 1A-1B, panel B). Los genes de EML-4 y ALK están ubicados ambos en el cromosoma 2, por lo tanto el gen de fusión fue creado mediante delección génica entre estos dos loci. La observación del gen de fusión (variante corta) en el paciente CS045 confirmó el hallazgo de

este gen mutante en la línea celular humana H2228.

El gen de fusión de TFG-ALK también estaba en marco y fusionaba los primeros 138 aminoácidos de TFG a los últimos 562 aminoácidos de ALK (véanse las Figuras 1C, panel B). Los genes de TFG y ALK están localizados en diferentes cromosomas (cromosomas 6 y 2, respectivamente), de este modo el gen de fusión se creó mediante translocación génica entre estos dos loci. Curiosamente, la fusión de TPG a ALK se produjo exactamente en el mismo punto de ALK que se observó para la fusión de las dos variantes de EML4-ALK, lo que indica que el truncamiento de ALK en tumores sólidos en este punto puede ser un evento común.

Se usaron los mismos cebadores descritos en la Parte A anterior. Se llevó a cabo un análisis de RT-PCR, sustancialmente como se ha descrito en la Parte A anterior, para confirmar que el extremo N de EML-4 y TFG están intactos en las proteínas de fusión (véase la Figura 6 (panel B)). Los pares de cebadores para EML-4 y ALK fueron como los descritos en la Parte A anterior. Se utilizó el siguiente par de cebadores para TFG:

TFG-F1: 5'-TTTGTTAATGGCCAGCCAAGACCC-3 (SEQ ID NO: 28)

Es de destacar que, en ambas variantes de fusión de EML4-ALK, el radical de ALK se fusiona al radical de EML-4, precisamente en el mismo punto de ALK que se ha observado en otras fusiones de ALK, tales como la fusión NPM-ALK que se produce en el LACG. Adicionalmente, EML-4 de tipo salvaje está fosforilado en tirosina solo en un sitio que no está presente en las proteínas de fusión de EML4-ALK, lo que sugiere que el dominio de bobina en espiral N-terminal que está conservado en las proteínas de fusión (véanse las Figuras 1A-1B) puede funcionar dimerizando y activando ALK, así como para promoviendo la interacción con ALK de tipo salvaje. También es de destacar que, la fusión del radical de TG a ALK también se produce precisamente en el mismo punto en ALK, y de hecho la fusión de TFG de ALK en este punto se ha descrito en el linfoma humano (véase Hernández *et al.* (2002), más arriba), pero no se ha descrito previamente en tumores sólidos humanos, tales como CPCNP.

Ejemplo 3

Inhibición del crecimiento de tumores sólidos de mamíferos que expresan la fusión de ALK utilizando ARNip

Con el fin de confirmar que las formas truncada/de fusión de ALK están dirigiendo el crecimiento celular y la supervivencia celulares en la línea celular de CPCNP H2228, así como en muestras de tumor de CPCNP de pacientes CS010/11, CS045, y CS110, se puede examinar la capacidad de ARNip (frente a ALK) para inhibir el crecimiento de estas células y tumores.

Se pueden adquirir dúplex ARNip de SMARTpool de ALK (secuencias diana en propiedad - datos no mostrados), p. ej., de Dharmacon Research, Inc. (Lafayette, CO). Se utiliza como control un ARNip de SMARTpool no específico. Las células se transfectan con el ARNip por medio de electroporación. En resumen, se aplican pulsos a 2×10^7 células (H2228) (20 ms; 275V, 20 ms K562; 285V) utilizando un electroporador de onda cuadrada (BTX Genetronics, San Diego, CA), se incuban a temperatura ambiente durante 30 minutos y se transfirieron a matraces T150 con 30 ml de RPMI - 1640/FBS al 10%.

Se determina el número de células viables con el análisis de proliferación de células en disolución CellTiter 96 AQueous One (Promega). La Cl_{50} se calcula mediante el uso del soporte lógico OriginPro 6.1 (OriginLab). El porcentaje de células apoptóticas a las 48 horas se puede determinar mediante el análisis de citometría de flujo Cleaved-Caspase-3 (Cell Signaling Technology).

El análisis por inmunotransferencia revelará que la expresión de ALK es reducida específicamente y significativamente a las 72 horas de la transfección del ARNip a células H2228 o células tumorales de pacientes CS010/11, CS045, y CS110. Se espera que la regulación a la baja de ALK de como resultado una fuerte inhibición del crecimiento celular. También se espera que el tratamiento con ARNip de ALK de como resultado un aumento de la apoptosis de estas células de tumores sólidos. Estos resultados indicarán adicionalmente que las quinasas ALK mutantes/fusionadas en la línea celular H2228 y en tumores de pacientes están dirigiendo la proliferación y el crecimiento de estas células de CPCNP, y que tal crecimiento y tal proliferación pueden ser inhibidos utilizando ARNip para inhibir la expresión y la actividad de la quinasa ALK.

Ejemplo 4

Inhibición del crecimiento de tumores sólidos de mamíferos que expresan la fusión de ALK utilizando WI-131 y/o WI-154

Para confirmar adicionalmente que las proteínas de fusión de ALK mutantes están dirigiendo el crecimiento y la viabilidad de la línea celular de CPCNP H2228 y de las células tumorales de CPCNP de pacientes CS010/11, CS045, y CS110, las células se pueden tratar con un inhibidor selectivo de la quinasa ALK, tal como WI-131 y/o WI-154. WI-131 y W-154 son inhibidores de molécula pequeña de tipo quinazolina de la quinasa ALK, y se ha descrito su actividad contra la proteína de fusión NPM-ALK en el linfoma de células T. Véase Marzec *et al.*, más arriba.

En resumen, se cultivan células de CPCNP, y se lleva a cabo un análisis de inhibición del crecimiento celular con

CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) de acuerdo con la sugerencia del fabricante. Brevemente, se siembran de 1000 a 5000 células sobre placas de 96 pocillos de fondo plano y se cultivan en medio completo con FBS al 10%. Después de 24 horas, se cambia el medio celular por 100 µl de medio de crecimiento completo con FBS al 10% que contiene diversas concentraciones del fármaco, y las células se incuban durante 72 horas más. Cada concentración de fármaco se aplica a pocillos de células por triplicado. Al final de la incubación, se añaden 20 µl de solución CellTiter 96 Aqueous One a cada pocillo, y la placa se incuba durante 1-4 horas. La absorbancia se lee a 490 nm utilizando un lector de microplacas Titan Multiskan Ascent (Titertek Instrument). La inhibición del crecimiento se puede expresar como la media $\pm$ valor DT del porcentaje de la lectura de absorbancia de las células tratadas en comparación con las células no tratadas. El análisis se repite al menos tres veces.

Se espera que tal análisis confirme que las proteínas de fusión de ALK (EML4-ALK (variantes corta y larga), TFG-ALK) están dirigiendo el crecimiento y la supervivencia de un subgrupo de tumores de CPCNP humano en las que estas proteínas mutantes son expresadas, y que tales células se pueden inhibir mediante la inhibición de la actividad de la quinasa ALK de fusión utilizando un inhibidor selectivo, tal como WI-131 y/o WI-154.

Ejemplo 5

Las proteínas de fusión de ALK dirigen el crecimiento y la supervivencia de líneas celulares de mamífero transformadas.

Con el fin de confirmar que la expresión de una o más de las proteínas de fusión de ALK puede transformar las células normales en un fenotipo canceroso, se pueden transformar células 3T# con los constructos de ADNc descritos anteriormente (Ejemplo 2), que expresan las proteínas de fusión de EML4-ALK (variantes corta y larga) o de TFG-ALK, respectivamente.

En resumen, las células se mantienen en medio RPMI-1640 (Invitrogen) con suero bovino fetal (FBS) al 10% (Sigma) y 1,0 ng/ml de IL-3 (R&D Systems). La producción de sobrenadante retroviral y la transducción se llevan a cabo como se ha descrito previamente. Véase Schwaller et al, *EMBO J.* 17(18): 5321-33 (1998). Las células 3T3 son transducidas con sobrenadante retroviral que contiene el vector MSCV-Neo/EML4-ALK (o TFG-ALK) y se seleccionan con G418 (1 mg/ml). A continuación se inspecciona la capacidad de las células transformadas para crecer en agar blando cultivando en placa las células transducidas después de haber lavado las células tres veces en PBS. Si se desea, para las curvas de respuesta a la dosis, las células se tratan con ARNip contra ALK como se ha descrito anteriormente (véase el Ejemplo 3), y se determina el número de células viables con el análisis de proliferación de células en disolución CellTiter 96 Aqueous One (Promega). La CI_{50} se puede calcular utilizando el soporte lógico OriginPro 6.1 (OriginLab). Se puede determinar el porcentaje de células apoptóticas a las 48 horas mediante análisis de citometría de flujo Cleaved-Caspase-3 utilizando un anticuerpo específico para esta diana (Cell Signaling Technology). Tal análisis demostraría que la expresión de la proteína de fusión de EML4-ALK (variante corta o larga) o de la proteína de fusión de TFG-ALK puede transformar las células 3T3 y confirmar la supervivencia y el crecimiento en agar blando cuando estas células son dirigidas por la proteína de fusión de ALK, y adicionalmente que la inhibición de la expresión de ALK en las células transformadas conduce a la disminución de la viabilidad y a un aumento de la apoptosis.

Ejemplo 6

Detección de la expresión de la proteína de fusión de EML4-ALK en tumores sólidos humanos utilizando el análisis FISH

La presencia de la proteína de fusión de EML4-ALK (variante corta) en muestras de tumores de CPCNP humano se detectó utilizando el análisis de hibridación fluorescente *in situ* (FISH), como se ha descrito previamente. Véase, p. ej., Verma et al. *HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES*, Pergamon Press, New York, N.Y. (1988). Se examinaron más de 200 muestras de tumores de CPCNP humanos incluidas en parafina.

Se obtuvo una sonda de translocación por rotura de dos colores para ALK, de Vysis (Vysis, Dowers Grove, IL, USA) y se utilizó de acuerdo con las instrucciones del fabricante con las siguientes modificaciones. En resumen, las secciones de tejidos incluidas en parafina se rehidrataron y se sometieron a recuperación de antígeno con microondas en de tampón de citrato 0,01 M (pH 6,0) durante 11 minutos. Las secciones se digirieron con proteasa (4 mg/ml de pepsina, 2000-3000 U/mg) durante 25 minutos a 37°C, se deshidrataron y se hibridaron con la sonda FISH ajustada a 37°C durante 18 horas. Después del lavado, se aplicó 4',6-diamino-2-fenilindol (DAPI; mg/ml) en un medio de montaje Vectashield (Vector Laboratories, Burlingame, CA) para la contratinción nuclear.

La sonda de translocación para ALK contiene dos sondas marcadas de forma diferente en lados opuestos del punto de rotura del gen de ALK (en el nucleótido 3171) en la secuencia de tipo salvaje (SEQ ID NO: 6). Cuando se hibriden, la región de ALK nativa aparecerá como una señal de fusión de color naranja/verde, mientras la translocación en este locus (como ocurre en los mutantes por delección de EML4-ALK) dará como resultado señales de color naranja y verde separadas. Véase la Figura 6.

El análisis FISH reveló una incidencia relativamente baja de esta mutación de EML4-ALK de variante corta en la población de muestras estudiada (una de cada 229 muestras). Sin embargo, dada la alta incidencia de CPCNP en

todo el mundo (más de 151.000 casos nuevos en los Estados Unidos solo cada año), se espera que haya un número significativo de pacientes que alberguen esta ALK mutante, cuyos pacientes pueden beneficiarse de un régimen terapéutico inhibidor de ALK.

Ejemplo 7

- 5 Detección de la expresión de la proteína de fusión de ALK en tumores sólidos humanos utilizando el análisis por PCR

También se puede detectar la presencia de una o más proteínas de fusión de ALK en una muestra de tumor sólido humano utilizando ya sea la reacción en cadena de la polimerasa (PCR) genómica o con transcriptasa inversa (RT), descrita anteriormente. Véase, p. ej., Cools et al., *N. Engl. J. Med.* 348: 1201-1214 (2003). Brevemente y a modo de
 10 ejemplo, se pueden obtener muestras de tumores sólidos de un paciente que tiene, p. ej., CPCNP, utilizando técnicas convencionales. Se construyen sondas para PCR contra la quinasa ALK truncada o la proteína de fusión de EML4-ALK (variante corta o larga) o la proteína de fusión de TFG-ALK. Se puede utilizar RNeasy Mini Kit (Qiagen) para extraer el ARN de las muestras de tumor. El ADN se puede extraer utilizando DNeasy Tissue Kit (Qiagen). Para la RT-PCR, se sintetiza la primera hebra de ADNc, p. ej., a partir de 2,5 g de ARN total utilizando, p. ej., el sistema
 15 de síntesis de la primera hebra Superscript[™] III (Invitrogen) con oligo (dT)₂₀.

A continuación, se amplifica el gen de fusión de ALK mediante el uso de pares de cebadores, p. ej., EML4-202 y ALK-GSP3 (véase el Ejemplo 2 más arriba). Para la PCR genómica, la amplificación del gen de fusión se puede realizar utilizando ADN polimerasa Platinum Taq de alta fidelidad (Invitrogen) con pares de cebadores, p. ej., EML4-202 y ALK-GSP3 (ver Ejemplo 2, más arriba). Tal análisis identificará un paciente que tenga un tumor sólido
 20 caracterizado por la expresión de la quinasa ALK truncada (y/o una o varias proteínas de fusión de EML4-ALK o una proteína de fusión de TFG-ALK), cuyo paciente es un candidato para el tratamiento utilizando un agente terapéutico inhibidor de ALK, tal como WHI-131 y/o WHI154.

REIVINDICACIONES

1. Un inhibidor de ALK para usar en el tratamiento del cáncer de un mamífero caracterizado por la expresión de un polipéptido de fusión de EML4-ALK.

5 2. Un inhibidor de ALK para usar según la reivindicación 1, en donde dicho cáncer es cáncer de pulmón.

3. Un inhibidor de ALK para usar según la reivindicación 1 ó 2, en donde dicho cáncer es CPCNP.

4. Un inhibidor de ALK para usar según cualquier reivindicación anterior, en donde dicho inhibidor es un oligonucleótido antisentido o ARNip capaz de inhibir la expresión de dicho polipéptido de fusión de EML4-ALK.

10 5. Un inhibidor de ALK para usar según cualquiera de las reivindicaciones 1 a 3, en donde dicho inhibidor es un anticuerpo capaz de bloquear la actividad de dicho polipéptido de fusión de EML4-ALK.

6. Un inhibidor de ALK para usar según cualquiera de las reivindicaciones 1 a 3, en donde dicho inhibidor es una molécula pequeña tal como WHI-131 y / o WHI-154, o sus análogos.

15

Fig.1 A

Fusión de EML4-ALK en la línea celular H2228 y el paciente cs045

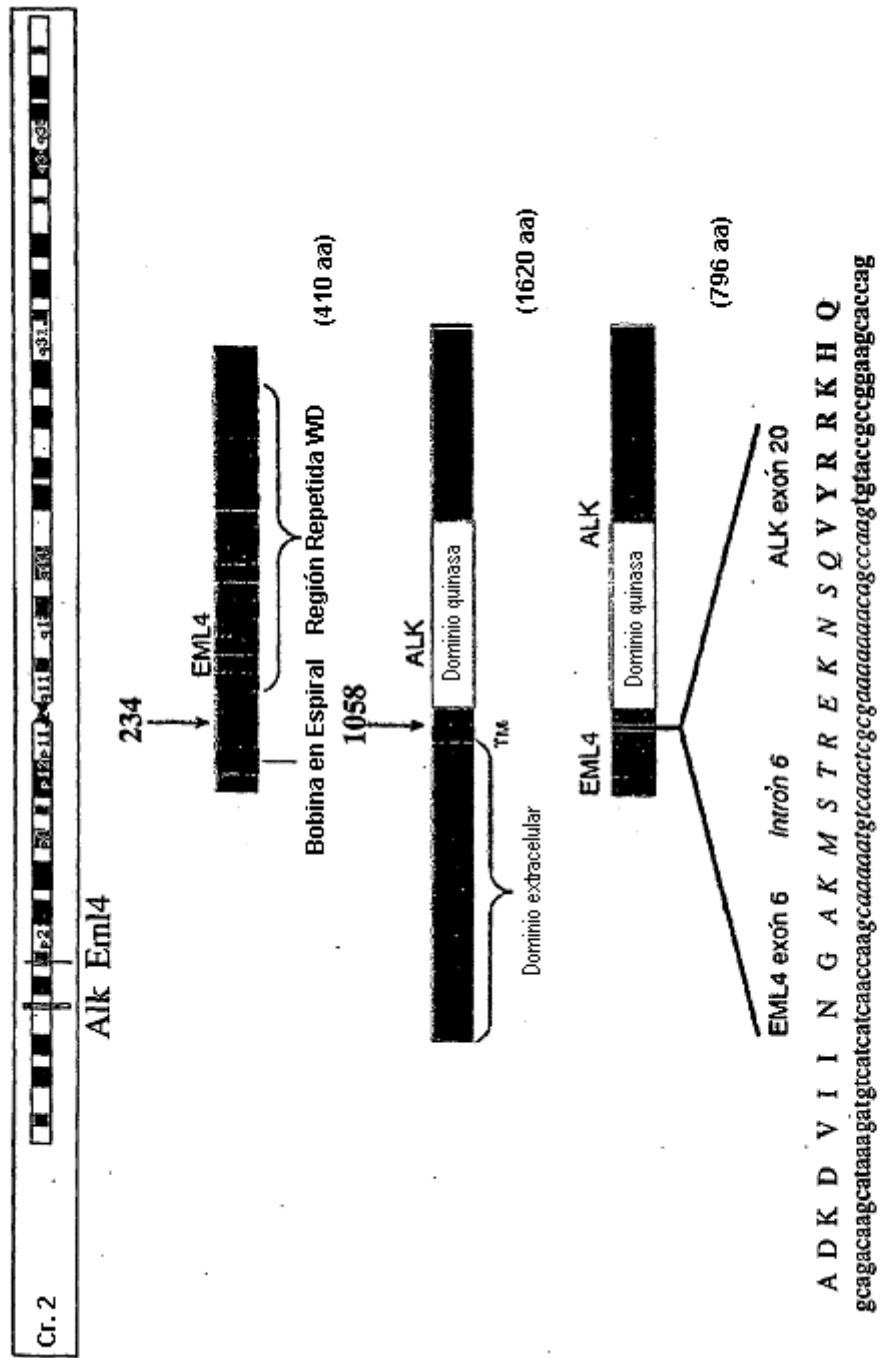


Fig. 1C **Fusión de TFG-ALK en el paciente CS110**

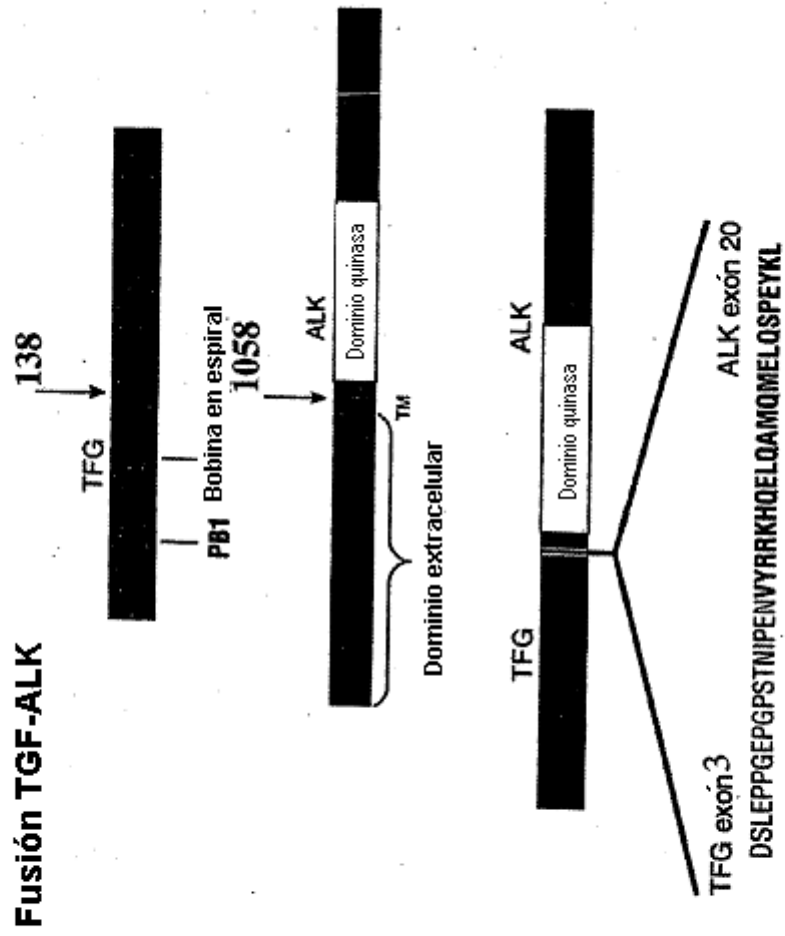


Figura 2A

MDGFAGSLDDSSISAASTSDVQDRLSALESRVQQQEDEITVLKAALADVLRLRLAISEDHVASVKK
SVSSKGQPSRAVPMSCITNGSGANRKPSTSAVSIAGKETLSSAAKSGTEKKKEKPQGQRE
KKEESHNDQSPQIRASPSQPSSQPLQIHRQTPESKNATPTKSIKRPSPAESHNSWENSDD
SRNKLSTPSTPKLIPKVTKTADKHKDVIIINQAKMSTREKNSQVYRRKHQELQAMQMELQSPEY
KLSKLRTSTIMTDYNPNYCFAGKTSSISDLKEVPRKNITLIRGLGHGAFGEVYEGQVSGMPN
DPSPLQVAVKTLPEVCSEQDELDLMEALISKFNHQNIVRCIGVSLQSLPRFILLELM
AGGDLKSFLRETRPRPSQPSSLAMLDLLHVARDIACGCQYLEENHFIHRDIAARNCLL
TCPGPGRVAKIGDFGMARDIYRASYYRKGGCAMLVVKWMPPEAFMEGIFTSKTDTW
SFGVLLWEIFSLGYMPYPSKSNQEVLEFVTSGGRMDPPKNCPPVYRIMTQCWQH
PEDRPNFAILLERIEYCTQDPDVINTALPIEYGPLVEEEEKVPRPKDPEGVPPLLVSQAKRE
EERSPAAPPPLPTSSGKAAKKPTAAEVSVRVRGPAVEGGHVNMAFSQSNPPSELHKVHGS
NKPTSLWNPTYGSWFTEKPTKKNPIAKKEPHDRGNLGLGSCVTPPNVATGRLPGASLLLEPS
SLTANMKEVPLFRLRHFPNGVNYGYQQQGLPLEAATAPGAGHIEDTILSKNSMNQPGP

atggacggtttcgcggcagctctcgatgatagattttctgctgcaagtaattctgatgttcaagatcgctgtcagctc
ttgagtcacgagttcagcaacaagaagatgaaatcactgtgctaaggcggtttgctgatgtttgagg
cgtcttgcattctcgaagatcatgtggtcctcagtgaaaaatcagctcgaagtaaggccaaccaagccctcag
cagttattcccatgtcctgtataaccaatggaagtgtgtcaaacagaaaaacaaagtcataccagtgctgtctca
attgcaggaaaaaagaactcttctcgtctgctcaaaagtgtgacagaaaaaagaagaaaaaccaaggaag
agagaaaaaaagagggaatctcattctaatgatcaaaagtcacaaattcgagcatcaccttctccagccctct
tcaaacctctccaaatacacagacaaactccagaaagcaagaatgctaaccacaaaagcataaaacgacat
caccagctgaaaagtcacataattctgggaaaattcagatgataccgtaataaattgtcgaataaccttca
acacccaattaataccaaaagttacaaaactgcagacaagcataaagatgtcatcatcaaccaag
caaaaatgtcaactcgcgaaaaaaacagccaag
tgtaccgcccgaagcaccaggagctgcaagccatgcagatggagctgcagagccctgagtacaaagctgagcaagct
ccgcaactcgaccatcatgaccgaatacaacccaactactgcttctggtgcaagacccctccatcagtgacc
tgaaggaggtgctcgcggaacacatcacctcattcggttctgggcatggcgctttggggaggtgtatg
aaggccagggtgtccggaatgcccaagcaccgaagccccctgcaagtggctgtgaagacgctgcctgaag
tgtgctctgaacaggacgaactggatttctcatggaagccctgatcatcagcaaatcaaccaccagaac
attgttcgtgcattgggtgagcctgcaatccctgcccgggttcctcctgctggagctcatggcggggg
agacctcaagtccttctccgagagaccgcccctgcagcagccctcctcctggcatgctggacc
ttctgcacgtggctcggaacattgctgtggtgtcagatattggaggaaaaccactcatccaccgagac
attgctgccagaaactgctcttgacctgtccaggccctggaagagtgccaagattggagacttcggga
tgcccgcagacatctacagggcgagctactatagaaggaggctgtgcatgctgccagtttaagtga
tgccccagaggccctcatggaaggaatattcacttctaaaacagacacatggtccttggagtgtgctat
gggaaatcttttcttgatatatgcatatccagcaaaaagcaaccagggaagttctggagttgtcacca
gtggaggccggatggaccaccaagaactgccctggcctgtataccggataatgactcagtgctggc
aacatcagcctgaagacaggcccaacttggcatctttggagaggattgaatactgcacccaggaccc
ggatgtaatacaacccgcttgcgatagaatagttccactgtggaagagggaagagaagtgctgtgagccca
aggacctgaggggttctcctcctcctggtctctcaacaggcaaaacgggaggaggagcgagccagctgccccac
cacctctgctaaccctcctctggcaaggctgcaagaaaacccacagctgcagagggtctctgttcgagtcctagagg
ccggccgtggaagggggacagtgaaatggcattctcagtcacacccctcctcgagttgcaaggtccacggat
ccagaaaacagccacagctgtggaacccaacgtaaggctcctgtttacagagaaaccacaaaagaataatc
ctatagcaagaaggagccacagcaggggtacactggggtggaaggagctgtaactgtcccaactaacgtgca
actgggagactccgggggctcctagagccctctcgtgactgccaataatgaaggaggtaactctgttcag
gctacgtcacttctgtgggaatgtcaattacggctacagcaacagggttgccttagaagccgctactgcccctg
agctggtcattacgaggatacattctgaaaagcaagaatagcatgaaccagcctggccctga

Figura 2B

MDGFAAGSLDDSSISAASDVQDRLSALESRVQQQDEITVYLKAALADVLRLAISEDHVASVKKSVSSKGQ
 PSPRAVIMSCITNGSGANRKPSTSAVSIAGKETLSSAAKSGTEKKKEKPGQREKKEESHNDQSPQIR
 ASPSPQSSQPLQIHRQTPESKNATPTKSIKRPSPAESHNSWENSDDSRNLSKIPSTPKLIPKYTKTADK
 HKDVIINQEGEYIKMFMGRGPITMFIPSDVDNYDDIRTELPEKLKLEWAYGYRGKDCRANYLLPTGEI
 VYFIASVVYLFNYEERTQRHYLGHTDCVKCLAIHPDKIRIATGQIAGVDKDGRLQPHVRVWDSVTLSTL
 QIIGLGTFERGVGCLDFSKADSGVHLCHIDDSNEHMLTVWDWQKKAKGAEIKTTNEVVLAVEFHPTDAN
 TIITCGKSHIFFWTWSGNSLTRKQGIFGKYEKPKFVQCLAFNGNDVLTGDSGGVMLIWSKTTVEPTPG
 KGPKVYRRKHQELQAMQMELOQSPYKLSKLRTSTIMTDYNPNYCFAGKTSSISDLKEVPRKNITLI
 RGLGHGAFGEVYEGQVSGMPNDPSPLQYAVKTLPEVCSEQDELDLMEALIISKFNHQNIVR
 CIGVSLQSLPRFILLELMAGGDLKSFLRETRPRSPSSLAMLDLHVARDIACGCQYLEENH
 FIHRDIAARNCLLTCGPGGRVAKIGDFGMARDIYRASYRKGGCAMLPVKWMPPEAFMEGI
 FTSKTDTSWFSGLLWEIFSLGYMPYPSKSNQEVLEFVTSGGRMDDPPKNCPPVYRIMTQCW
 QHQPEDRPNFAILERIEYCTQDPDVINTALPIEYGPLVEEEKVPVRPKDPEGVPPLLVSQQAKR
 EEERSPAAPPLPTSSGKAAKKPTAAEISVRVPRGPAVEGGHVNMAFSQSNPPSELHKVHGSRNK
 PTLWNPTYGSWFTEKPTKKNPIAKKEPHDRGNLLEGSCTVPPNVATGRLPGASLLLEPSSLTA
 NMKEVPLFRLRHFCGNVNYGYQQQLPLEAATAPGAGHYEDTILKSKNSMNQPGP

ATGGACGGTTTCGCCGGCAGTCTCGATGATAGTATTTCTGCTGCAAGTACTTCTGATGTTCAAGATCG
 CCTGTCAGTCTTGTAGTACAGAGTTCAGCAACAAGAAGATGAAATCACTGTGCTAAAGCGGCTTTG
 GCTGATGTTTTGAGGCGTCTTGAATCTCTGAAGATCATGTGGCCTCAGTGAATAATCAGTCTCAAG
 TAAAGGCCAACCAAGCCCTCGAGCAGTTATCCCATGTCTGTATAACCAATGGAAGTGGTGCAAAAC
 AGAAAACCAAGTCATACAGTGCTGTCTCAATTGCAGGAAAAGAACTCTTTCATCTGCTGCTAAAAG
 TGGTACAGAAAAAAGAAAAGAAAAACCAAGGACAGAGAGAAAAAAGAGGAATCTCATTCTAATG
 ATCAAAGTCCACAAATTGAGCATCACCTTCTCCCCAGCCCTCTTCAACCTCTCCAAATACACAGA
 CAAACTCCAGAAAGCAAGAAATGCTACTCCCAACCAAGCATAAAACGACCATCACCAGCTGAAAAGT
 CACATAATTCTTGGGAAAATTCAGATGATAGCCGTAATAAATTGTGCGAAAATACCTTCAACACCCAAAT
 TAATACCAAAAAGTTACCAAACTGCAGACAAGCATAAAGATGTCATCATCAACCAAGAAGGAGAAATAT
 ATTAATAATGTTTATGCGCGGTGCGGCAATTACCATGTTTCATTCTTCCGATGTTGACAACTATGATGAC
 ATCAGAACCGAACTGCCTCCTGAGAACTCAAACTGGAGTGGGCATATGGTTATCAGGAAAGGACT
 GTAGAGCTAATGTTTACCTTCTCCGACCGGGGAAATAGTTTATTTTCATTGCATCAGTAGTAGTACTAT
 TTAATTATGAGGAGAGAACTCAGCGACACTACCTGGGCCATACAGACTGTGTGAAATGCCTTGCTATA
 CATCTGACAAAATTAGGATTGCAACTGGACAGATAGCTGGCGTGATAAAGATGGAAGGCCTCTAC
 AACCCACGTCAGAGTGTGGGATTCTGTTACTCTATCCACACTGCAGATTATTGGACTTGGCACTTTT
 GAGCGTGGAGTAGGATGCTGGATTTTTCAAAAGCAGATTCAAGGTGTTTATTATGATTATTGATGA
 CTCCAATGAGCATATGCTTACTGTATGGGACTGGCAGAAGAAAGCAAAAGGAGCAGAAATAAGACA
 ACAATGAAGTTGTTTTGGCTGTGGAGTTTCACCCAACAGATGCAATACCATAATTACATGCGGTAA
 ATCTCATATTTTCTTCTGGACCTGGAGCGGCAATCACTAACAAGAAAACAGGGAATTTTGGGAAAT
 ATGAAAAGCCAAAATTTGTGAGTGTGTTAGCATTCTTGGGGAATGGAGATGTTCTTACTGGAGACTCA
 GGTGGAGTCATGCTTATATGGAGCAAACTACTGTAGAGCCACACCTGGGAAAGGACCTAAAGTGT
 ACCGCCGGAAGCAACAGGAGCTGCAAGCCATGCAGATGGAGCTGCAGAGCCCTGAGTACAAG
 CTGAGCAAGCTCCGACCTCGACCATCATGACCGACTACAACCCCACTACTGCTTTGCTGGC
 AAGACCTCCTCCATCAGTGACCTGAAGGAGGTGCCGCGGAAAAACATCACCCCTCATTGCGG
 GTCTGGGCCATGGCGCCTTTGGGGAGGTGTATGAAGGCCAGGTGTCCGGAATGCCCAA
 CGACCCAAAGCCCCCTGCAAGTGGCTGTGAAGACGCTGCCTGAAGTGTGCTCTGAACAGG
 ACGAACTGGATTTCTCATGGAAGCCCTGATCATCAGCAAAATCAACCACCAGAACATTG
 TTCGCTGCATTGGGGTGAGCCTGCAATCCCTGCCCCGTTTCATCTGCTGGAGCTCATG
 GCGGGGGGAGACCTCAAGTCTTCTCCGAGAGACCCGCCCTCGCCCGAGCCAGCCCTC
 CTCCTGGCCATGCTGGACCTTCTGCACGTGGCTCGGGACATTGCCTGTGGCTGTCACT
 ATTTGGAGGAAAACCACTTCATCCACCGAGACATTGCTGCCAGAACTGCCTCTTGACCT

GTCCAGGCCCTGGAAGAGTGGCCAAGATTGGAGACTTCGGGATGGCCCGAGACATCTAC
AGGGCGAGCTACTATAGAAAGGGAGGCTGTGCCATGCTGCCAGTTAAGTGGATGCCCC
AGAGGCCTTCATGGAAGGAATATTCACCTTCTAAAACAGACACATGGTCCTTTGGAGTGC
TGCTATGGGAAATCTTTTCTCTTGGATATATGCCATACCCAGCAAAAGCAACCAGGAAG
TTCTGGAGTTTGTACCAAGTGGAGGCCGGATGGACCCACCCAAGAACTGCCCTGGGCCT
GTATACCGGATAATGACTCAGTGCTGGCAACATCAGCCTGAAGACAGGCCCAACTTTGC
CATCATTTTGGAGAGGATTGAATACTGCACCCAGGACCCGGATGTAATCAACACCGCTTT
GCCGATAGAATATGGTCCACTTGTGGAAGAGGAAGAGAAAAGTGCCTGTGAGGCCCAAGGACC
CTGAGGGGGTTCTCCTCTCCTGGTCTCTCAAAGGCCAAAACGGGAGGAGGAGCGCAGCCCA
GCTGCCCCACCACCTCTGCCTACCACCTCCTCTGGCAAGGCTGCAAAGAAACCCACAGCTGCA
GAGATCTCTGTTTCGAGTCCCTAGAGGGCCGGCCGTGGAAGGGGGACACGTGAATATGGCATT
CTCTCAGTCCAACCTCCTTCGGAGTTGCACAAGGTCCACGGATCCAGAAACAAGCCCACCAG
CTTGTGGAACCCAACGTACGGCTCCTGGTTTACAGAGAAAACCCACCAAAAAGAATAATCCTAT
AGCAAAGAAGGAGCCACACGACAGGGGTAACTGGGGCTGGAGGGAAGCTGTACTGTCCCAC
CTAACGTTGCAACTGGGAGACTTCCGGGGGCCCTCACTGCTCCTAGAGCCCTCTTCGCTGACTG
CCAATATGAAGGAGGTACCTCTGTTCAAGGCTACGTCACTTCCCTTGTGGGAATGTCAATTACG
GCTACCAGCAACAGGGCTTGCCCTTAGAAGCCGCTACTGCCCTGGAGCTGGTCATTACGAGG
ATACCATTCTGAAAAGCAAGAAATAGCATGAACCAGCCTGGGCCCTGA

Figura 2C

MNGQLDLSGKLIKAQLGEDIRRIPIHNEDITYDELVLMMQRVFRGKLLSNDEVTIKYKDEDGDLI
 TIFDSSDLSFAIQCSRLKLTFLVNGQPRPLESSQVKYLRRELIELRNKVNRLDSLEPPGEPGPSTN
 IPENVYRRKHQELQAMQMELOQSPYKLSKLRTSTIMTDYNPNYCFAGKTSSISDLKEVPRKNITLI
 RGLGHGAFGEVYEGQVSGMPNDPSPLQYAVKTLPEVCSEQDELDFLMEALHISKFNHQNIVR
 CIGVSLQSLPRFILLELMAGGDLKSFLRETRPRPSQPSSLAMLDLLHVARDIACGCCQYLEENH
 FIHRDIAARNCLLTCPPGPRVAKIGDFGMARDIYRASYRKGKGCAMLVVKWMPPEAFMEGI
 FTSKTDTSFSGVLLWEIFSLGYMPYPSKSNQEVLEFVTSGGRMDDPKNCPGPVYRIMTQCW
 QHQPEDRPNFAILERIEYCTQDPDVINTALPIEYGPLVEEEKVPVRPKDPEGVPPLLVSQQAQR
 EEERSPAAPPPLPTSSGKAAKKPTAAEISVRVPRGPAVEGGHVNMAFSQSNPPSELHKVHGSRNK
 PTLWNPTYGSWFTEKPTKKNPIAKKEPHDRGNLGLGSCCTVPPNVATGRLPGASLLEPSSLTA
 NMKEVPLFRLRHFCGNVNYGYQQGLPLEAATAPGAGHYEDTILKSKNSMNQPGP

ATGAACGGACAGTTGGATCTAAGTGGGAAGCTAATCATCAAAGCTCAACTTGGGGAGGATATTC
 GGCGAATTCCTATTCATAATGAAGATATTACTTATGATGAATTAGTGCTAATGATGCAACGAGTT
 TTCAGAGGAAAACTTCTGAGTAATGATGAAGTAAACAATAAAGTATAAAGATGAAGATGGAGATC
 TTATAACAATTTTGTAGTTCTGACCTTTCTTTGCAATTCAGTGCAGTAGGATACTGAAACTG
 ACATTATTTGTTAATGGCCAGCCAAGACCCCTTGAATCAAGTCAGGTGAAATATCTCCGTCGAGA
 ACTGATAGAACTTCGAAATAAAGTGAATCGTTTATTGGATAGCTTGGAAACCACCTGGAGAACCA
 GGACCTTCCCAAAATATTCTGAAATGTGTACCGCCGGAAGCAACAGGAGCTGCAAGCCATG
 CAGATGGAGCTGCAGAGCCCTGAGTACAAGCTGAGCAAGCTCCGCACCTCGAACCATCATGAC
 CGACTACAAACCCCACTACTGCTTTGCTGGCAAGACCTCCTCCATCAGTGACCTGAAGGAGGT
 GCCGCGGAAAAACATCACCTCATTCCGGGTCTGGGCCATGGCCGCTTTGGGGAGGTGT
 ATGAAGGCCAGGTGTCCGGAATGCCCAACGACCCCAAGCCCCCTGCAAGTGCTGTGAAG
 ACGCTGCCTGAAGTGTGCTCTGAACAGGACGAACTGGATTTCTCATGGAAGCCCTGAT
 CATCAGCAAATTCACCACCAAGACATTGTTCTGCTGCATTGGGGTGAGCCTGCAATCCC
 TGCCCCGGTTTATCCTGCTGGAGCTCATGGCGGGGGGAGACCTCAAGTCCTTCTCCGA
 GAGACCCGCCCCGCCCCGAGCCAGCCCTCCTCCCTGGCCATGCTGGACCTTCTGCACGT
 GGCTCGGGACATTGCCTGTGGCTGTCACTTTGGAGGAAAAACCACTTCATCCACCGAG
 ACATTGCTGCCAGAACTGCCTCTTGACCTGTCCAGGCCCTGGAAGAGTGCCCAAGATT
 GGAGACTTCGGGATGGCCCCGAGACATCTACAGGGCGAGCTACTATAGAAAGGGAGGCT
 GTGCCATGCTGCCAGTTAAGTGGATGCCCCAGAGGCCCTTCATGGAAGGAATATTCACT
 TCTAAACAGACACATGGTCTTTGGAGTGCTGCTATGGGAAATCTTTTCTCTTGATAT
 ATGCCATACCCAGCAAAAGCAACCAGGAAGTTCTGGAGTTTGTCAACAGTGAGGCCG
 GATGGACCCACCAAGAACTGCCCTGGGCCTGTATACCGGATAATGACTCAGTGCTGGC
 AACATCAGCCTGAAGACAGGCCCAACTTTGCCATCATTTTGGAGAGGATTGAATACTGC
 ACCCAGGACCCGGATGTAATCAACACCGCTTTGCCGATAGAATATGGTCCACTTGTGGAAGA
 GGAAGAGAAAGTGCTGTGAGGCCCAAGGACCTGAGGGGGTCTCTCTCTCTGGTCTCTCA
 ACAGGCAAAACGGGAGGAGGAGCGCAGCCAGCTGCCCAACCACTCTGCCTACCACTCTCT
 CTGGCAAGGCTGCAAAAGAAACCCACAGCTGCAGAGATCTCTGTTGAGTCCCTAGAGGGCCG
 GCCGTGGAAGGGGGACACGTGAATATGGCATTCTCTCAGTCCAACCTCCTTCGGAGTTGCAC
 AAGGTCCACGGATCCAGAAACAAGCCACCAAGCTTGTGGAACCAACGTACGGCTCTGGTTT
 ACAGAGAAACCCACCAAAAGAAATAATCCTATAGCAAAGAAGGAGCCACACGACAGGGGTA

ACCTGGGGCTGGAGGGAAGCTGTACTGTCCACCTAACGTTGCAACTGGGAGACTTCCGGGG
GCCTCACTGCTCCTAGAGCCCTCTTCGCTGACTGCCAATATGAAGGAGGTACCTCTGTTCAAG
CTACGTCACTTCCCTTGTGGGAATGTCAATTACGGCTACCAGCAACAGGGCTTGCCCTTAGAA
GCCGCTACTGCCCCTGGAGCTGGTCATTACGAGGATACCATCTGAAAAGCAAGAATAGCATG
AACCAGCCTGGGCCCTGA

Figura 3A

mdqfagslddsisaastsdvqdrisalesrvqqgedeitvkaaladvlr
 rlaisedhvasvksavsskqgpspravipmscitnqsganrkpshtsavs
 iagketlssaaksqtekkkekpggqrekkeeshsndqspqiraspqps
 sqplqihrtpesknaptksikrpspaekshnswensddsrnklskip
 tpklipkvtktdakhkdviingegeyikmfmrgxpitmf
 psdvdyddirtelppekiklewaygyrgkdcranvylptgeivyfias
 vvlfnyeertqrhylvghtdcvkclaihpdkiriatgqiagvdkdgrplq
 phrvvwdavtltlgiigltfergvgcldfskadsgvhlcviddsnehm
 ltvw dwqkkakgaeikttnvvlavefhptdantiitcgkshiffwtwsg
 nsltrkqgfigkyekpkfvqclafingdvltgdsqgvmlwskttvept
 pgkqpkgyvqiskqikahdgsvftlcqmrngmltgggkdrkiilwdh
 npereievdpdyqtiravaegkadqflvgtsrnfirgtfndgfqievqg
 htdelwglathpfkdllltcaqdrqvcwnsmehrlwtrlvdepgcad
 fhpsgtvvaigthsgrwfvldaetrdlvsihtdgneqlsvmrysidgtfl
 avgshdnyiylvvsengrkysrygrctghssyithldwspdkyimsn
 gdeilywdipngeckliinrsdckdidwttytcvlgfvqvwpegsdgt
 dinalvrshnrkviavaddfckvhlfqypcskakapshkysahsshvtnv
 sfthndshlistggkdmisiiqwkllveklslpnetvadttitkapvste
 sviqsnptpppsqplnetaeesrissptllensleqtvepsedhsee
 eseegsgdlgeplyeepcneiskeqakatllledqqdpsps

Figura 3B

atggacgggttccggcagctcgaatgagatattctgctcaagtccttctgagtcgaagatccctg
 icagctcttgagtcacgagtcagcaacaagaagatgaatcactgtgctaaaggcggccttggctgaf
 gttttgagggcgtcttgaatctcgaagatcatgtgcccctcagtgaaagaaatcagctcgaatgaaggcc
 aaccaaggccctcagcagcttattccatgtctgtatataaccaatggaaagtggtgcaaacagaaaacca
 agtcataccagtgctgtctcaattgcaggaagaaacacitcttcatctgtgctaaagtggtacagaa
 aaaaagaagaagaaacacaaaggacagagagaaaaaagaggaaatctattcaatgatcaaa
 gtccacaaatcagacatcaccttctccacggcccttccacaacctctccaaatcacagacaaac
 cagaaagcaagaaatgctactccacaaagcataaaacgaccatcaccagctgaaagtcacat
 aatttctgggaaatcagatgatagccglaataaattgctgaaatccttcaacac
 ccaaatataaaccaaaagttacaaaactgcagacaagcataaagatgtatcatcaaccaagaaag
 gagaatataaataaattgttatgcggctggcgaattaccatgttacttctccgatgtgacaactatg
 atgacatcagaacggactgctcctgagaagctcaaacggagtgaggcatatgttatcagagaa
 ggactgtagagctaatgttacccttctccgaccggggaatagttattcattgcacagtagtagtact
 atttaattatgaggagagaaactcagcagacacacctggccatacagactgtgtgaatgccttctat
 acatcctgacaaaataggatgcaactggacagatagctggcgtggataaagatggaaaggccctta
 caacccacagtcagaggtgtggattctgttactctatccacactgcagattattggacttggactttgag
 cgtggagtaggagtgcttggaattttcaaaagcagattcaggtgttcaattatgtgtattgatgactccat
 gagcaiatgcttactgtatggactggcagaagaaagcaaaaggagcagaaataaagacuaaaa
 tgaagtgttttggctgtggagtttaccacacagatgcaaataccataattacatgcggtaaatctcata
 tttctctgacactggagcggcaattcactaacaagaaacagggaaattttggaaatatgaaaagc
 caaaatttgtcagtggttagcatttctgggaatggagatgttcttactggagactcaggtggagtcag
 ctatattgagcaaaactactgttagagccacacctgggaaggacataaaggtgtatataaactcag
 caaacaatacaagctcatgtggcagtggtgtacacattgtcagatgagaaatgggatgttataactgga
 ggagggaagacagaaaaataattctgtggatcatgatctgaatccgaaagagaaatagaggttctgta
 tcagtatggcacaatcagagctgtagcagaaggaaggcagatcaatttttagtaggcacatcacgaaact
 ttattttacgagggaacatttaagtggttccaaatagaagtacagggtcacaacagatgagcttgggtctt
 gccacacatccctcaaaagatttgccttgcacatgtgtcagcagaggcaggtgtgcttgggaactcaatg
 gaacacaggctggaatggaccaggctgttagatgaaccaggacactgtgcagattttatccaagtgga

Cagtggtggtccgacaggaagagcactcagcaggtgtgtgttcttggtatgcagaaacacagagatctagttc
 atccacacagacggggaatgaacagctctctctgtagcgtactactcaatagatgtagtacctctctgctgtagga
 tctcatgacaacctttattacctctatgtagtctctgaaaatgggaagaaaatatagcagatlatgggaaggtgcac
 tggacattccagctacatcacacacctgactgtgtccccagacaaacagtatataatgcttaacctcgggagga
 ctatgaaaattattgactggggacattccaaaaggctgcgaaaactaalcaggaaatcgaatgggattgtaaggacat
 gattgggacgacataacctgtgtgtcaggaattcaagtatllgggtgtcgtgccagaaggatctgtagggaca
 gataactaagtacatggtgtgcgtaccccaataagaaagctgtatagctgtgtccgagctcgtttgttaaggctccalc
 ttttcagatccctgtctcaaaagcaaaagctcccgagtcacagatagctgtgccacagacgacatgttaccac
 atgtcagttttactcacaatgacagtcacctgatalcaactgggtggaaaagacatgagcaltacatcagtgga
 aacttgtggaaaagtattcttllggccicagaatgagactgtagcgggatactacttaaccaaggccccgtctc
 ttccactgaaaagtgtcatccaatcaatactccacacccgccctctctcagcccttaaatgagacagctgaa
 gagggaaaagtataaataagcagttctccacactctcggagaaacagcctgggaacaaactgtgggccaaagtga
 agaccacagcgcgagggaggagagtgagagaggcgacggagagacctgtgtgagcctcttlatgaaaggccat
 gcaacgagataaagcaaggagcagcgccaaaagccacctctctggaggaccagcagaagaccttcgcctctg
 tctaaacacctggtctcagtgcaactctttctctcagctgcagtgattttgtatataaggctcaggatlaacagg
 atgggcagtgtagtgagaatcacgttgatgagatttggttccatgtgattgtttttcacaatagctctattttc
 agtctctcaataacagccaaactgttcttgggtgtgtttatgaaaaataaccaacttaactagtagga
 agactgtcaacttaatgtagtcttcaacaattactgtgttacctaaagtgtgtgtataactctgtgaaaacaaaa
 cagcagttgtcattgtatttgaaaaaaaaacccccctgtttatctgaacatgtttcttcaggaaacaaccagggtat
 tcacaaaacactgttactactactggtcagacgtactacttttttttttcttgaaaaaagaaaccggaa
 aaaaaatgtactcttactgagataccccctcaccacaaatgtgtaatggaaaaatttaataagaaaaaactcag
 ttttgcgaagtgcgaatgtgtgtgtcttctttaaagtataatgaggaacagctcccttaattctgtgtgcaact
 ctgtttatcttagaactaaaggagagcattgtgttgaaggactttcaatgtatataaaaaccttcaatactcag
 aatatgtaggattcctccaaggaggtcccttactagcctaa
 acattctcaaatgtttgagattcaagtgaaatgggaaggaaaaccacatgctctttaaacttaactgtaata
 atattctgtctaatttcagtaagcttaactaataattgttccctcagtaataaggagaaatatataatc
 agtaagtttgaatttgaattgtgtgtgtgaaatttattgtgtgtgttgaatttatcacagattatagagggtata
 gatactcatcaaatgtcaaatcttctttttacagaaggtgtggglaacagtcacagcagttttttccaacagc
 atacttaacagacttgcgtgttagcagttttttctgggtggagttgtctgaagctgttaaagtctaagtgtgctat
 cctactcttttgggcaatgcagtattatgcattggaaaaggtatllttttaaagttctgtgtgctagctatgtttc
 agtaacatttccactttaaaggagtaattactgacaaaatatgtatttccatattgtttatacttgattataaaaaagtat
 ttgtttgatttttaactgtctgattgtttgtatactttctattttttgttcacaaatcatgtttagaacattgtgag
 gttagaagagcttaagatgcagcgcgttactgtattgtgccattccaaagtgcacagaactgtcattccctt
 ctaatatctctcagggagtaatacaaacacaggtatttactgaactattgtgtaataatgaaaactccagtgaaatcc
 caaggacatttaacaaatatacttaccagctgtatgtggaagggtgtgggtgtgtgtgaaggggcgagtg
 agacactgtgtatctctagataaaggagatgtaccacgttgaaaatactcagtgtagactctatgtgtat
 aggtatctgtatatttccctttgtttacactgtttaaataaacctcaaaaatgttcttcaaaaagaaagagatt
 ccaagcaacccacttllctcaglatgtatgttctgtacalactatcggagcgcgccagtaaglatcaggcat
 atatactgtctgttagcaatgattattacatcatcagatcagcatgtgtctatactcccgtcaagaaatatactg
 acatgaacaggcagttcttggagaagaaagagactttctttaaagctcggggaaatacagctctcagtgat
 cagcagggagattttttgaggacatcagtcacctttgggggtgtccatgtacaatgagattataatcaltgatac
 tttctgtgtgtatgttcaaaaagacactactaataacgcaggaaggctccagctatttaattgctggcaacactgt
 ttaatgtgacagcttaactctgaataatgtttgtggaaggtgtgtgtgtgttgaagtgtatgattgttttgaaaaa
 ctgtgtglaagaaaatgaatcgaatgtttctgtggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 cgcaaaaattttgaagataaaaaataaataaataaataaataaataaataaataaataaataaataaataaataa

Figura 4A

MGAIGLLNLLPLLLSTAAGSGMGTGQRAGSPAAGSPLOPREPLSYSRLQ
 RKSLAVDFVPSLFRVYARDLLLPSSSELKAGRPEARGLALDCAPLLR
 LLGPAPGVSWTAGSPAPAEARTLSRVLKGGSVKLRRAKQLVLELGEZAI
 LEGCVGPPGEAAVGLLQFNLSLFSWIRQGEGRRLRIRLMPKKASEVGR
 EGRLSAAIRASQPRLLFQIFGTGHSSLESPTNMPSPSPDYFTWNLTWIMK
 DSFFFLSHRSRYGLECSDFDFCELEYSPLHDLRNQSWSWRRIPSEEASQ
 MDLLDGPAGERSKEMPRGSFLLNLSADSKHTILSPWMRSSSEHCTLAVS
 VHRHLQPSGRYIAQLLPHNEAAREILLMPTPGKHGWTVLQGRIGRPDNP
 RVALEYISSGNRSLSAVDFFALKNCSEGTSPGSKMALQSSFTCWNGTVLQ
 NFEDGFCGWTQGLTSPHTPQWQVRLTKDARFQDHDHALLSTTDVPASE
 SATVTSATFPAPIKSSPCELMSWLIRGVLRGNVSLVLVENKTGKEQGRM
 VWHVAAYEGLSLWQWMLPLLDVSDRFLQMVANWQGSRAIVAFDNISI
 SLDCYLTISGEDKILQNTAPKSRNLFERNPNKELKPGENSPTQTFIDPT
 VHWLFTTCGASGPHGPTQAQCNNAYQNSNLSVEVGSEGPLKGIQIWKVPA
 TDYTSISGYGAAGGKGGKNTMHRSHGVSVLGIFNLEKDDMLYILVGQOGE
 DACPSTNQLIQKVICIGENNVIEEIRVNRSVHEWAGGGGGGGGATYVFKM
 KDGVPVPLIIAAGGGGRAYGAXTDTFHPERLENNSSVLGLNGNSGAAGGG
 GGWNNDNTSLWAGKSLQEGATGGHSCPMKKWGWETRGGFGGGGGGCS
 GGGGGGYIGGNAASNDPEMDGEDGVSFISPLGILYTPALKVMEGHGEVN
 IKHYLNCSHCEVDECHMDPESHKVICFDHGTVLAEDEVSCIVSPTPEPH
 LPPLSLILSVVTSALVAALVLAFLAFSGIMIVYRRKHQELQAMQMELOSPYKL
SKLRTSTIMTDYNPNYCFAGKTSSISDLKEVPRKNITLIRGLGHGAPGEV
YEGQVSGMPNDPSPLOVAVKTLPEVCSEODELDPLMEALIISKFNHONIV
RCIGVSLQSLPRFILLELMAGGDLKSLRETRPRSPQSSLAMLDLHVA
RDIAACGQYLEENHFIHRDIAARNCLLTCPGPGRVAKIGDFGNARDIYRA
SYRKGCCAMLPVKWMPPEAFMEGIFTSKTDTSFGVLLWEIFSLGYMPY
PSKSNQEVLEFVTSGRMDPPKNCPPVYRIMTQCHQOPEDRPNFAILL
ERIEYCTQDDVINTALPIEYGLVEEEKVPVRPKDPEGVPPLLVSQQA
KREEERSPAAPFPLTTSSGKAAKKPTAAEVSVRVRPRAVEGGHVNMAF
SQSNPPSELHKVHGSRNKPTSLWNPTYGSWTFEKPTKKNPIAKKEPHDR
GNLGLGSCCTVPPNVATGRLPGLSLLLEPSSLTANMKVPLFLRLRHFFCG
NVNYGYQQGLPLEAATAPGAGHYEDTILSKNSMNQPGP

Figura 4B

1 agagacttgc gcgcacgcac agtcctctgg agatcaggtg gaaggagccg ctgggtacca
 61 aggactgttc agagcctctt cccatctcgg ggagagcgaa gggtaggct gggcccgag
 121 agcagtgtaa acggcctcct ccggcgggat gggagccatc gggctcctgt ggctgctgcc
 181 gctgctgctt tccacggcag ctgtgggctc cgggatgggg accggccagc gcgcgggctc
 241 cccagctgcg gggccgcgcg tgcagccccc ggagccactc agctactcgc gcctgcagag
 301 gaagagtctg gcagttgact tcgtggtgcc ctgctcttc cgtgtctacg cccgggacct
 361 actgctgcca ccatcctcct cggagctgaa ggctggcagg cccgagggcc gggctcgtc
 421 agctctggac tgcgccccgc tgcctcaggtt gctggggccg gcgcccgggg tctcctggac
 481 cgccgggttc ccagccccgg cagaggcccg gacgctgtcc aggtgctga agggcggctc
 541 cgtgcgcaag ctccggcggtg ccaagcagtt ggtgctggag ctggggcagg agggcatctt
 601 ggagggttgc gtcgggcccc ccggggaggc ggctgtgggg ctgctccagt tcaatctcag
 661 cgagctgttc agttggtgga ttccgcaagg cgaaggcgga ctgaggatcc gctgatgac
 721 cgagaagaag gcgtcgggag tgggcagaga ggggaaggctg tccgcggcaa ttcgcgctc
 781 ccagccccgc cttctcttcc agatcttcgg gactggctat agctccttg aatcaccaac
 841 aaacatgcct tctccttctc ctgattattt tacatggaat ctacactgga taatgaaga
 901 ctccttccct ttcctgtctc atcgagccg atatggtctg gagtgcagct ttgaattccc
 961 ctgtgagctg gattattccc ctccactgca tgacctcagg aaccagagct ggtcctggcg
 1021 ccgatccccc tccgaggagg cctccagat ggacttgctg gatgggctg gggcagagcg
 1081 ttctaaggag atgcccagag gctccttctt ccttctcaac acctcagctg actccaagca
 1141 caccatcctg agtcctgga tgaggagcag cagtgcagcac tgcacactgg ccgtctcggg
 1201 gcacaggcac ctgcagccct ctggaaggta cattgcccag ctgctgccc acaacaggagc
 1261 tgcgaagagag atcctcctga tgcctactcc agggaaagcat ggttggaacg tgcctcaggg
 1321 aagaatcggg cgtccagaca acccatttcg agtgccctg gaatacatct ccagtggaaa
 1381 ccgcagcttg tctgcagtg acttcttgc cctgaagaa tgcagtgaa gaacatcccc
 1441 aggtcccaag atggccctgc agagctcctt caattgttg aatgggacag tctccagct

FIGURA 4B (CONT.)

1501 tgggcaggcc tgtgacttcc accaggactg tgcccaggga gaagatgaga gccagatgtg
1561 ccggaaactg cctgtgggtt tttactgcaa ctttgaagat ggcttctgtg gctggaccaca
1621 aggcacactg tcaccccaca ctctcagtg gcaggtcagg accctaaagg atgcccgggt
1681 ccaggaccac caagaccatg ctctatttgt cagtaaccat gatgtcccg cttctgaaag
1741 tgetacagtg accagtgcta cgtttccctg accgatcaag agctctccat gtgagctccg
1801 aatgtctcgg ctcatctgtg gagtcttgag gggaaacgtg tccttgggtg tagtggagaa
1861 caaaaccggg aaggagcaag gcaggatggt ctggcatgtc gccgcctatg aaggcttgag
1921 cctgtgagcag tggatggtgt tgctctcct cgtatgtctc gacaggttct ggctgcagat
1981 ggtgcgcatg tggggacaag gatccagagc catcgtggct tttgacaata tctccatcag
2041 cctggactgc tacctcaaca tttagcggaga ggacaagatc ctgcagaata cagcaccaca
2101 atcaagaana ctgtttgaga gaaacccaaa caaggagctg aaaccggggg aaatttcacc
2161 aagacagacc cccatctttg accctacagt tcattggctg ttcaccacat gtggggccag
2221 ggggccccat gggccacacc aggcacagt caacaacgcc taccagaact ccaacctgag
2281 cgtggagggtg gggagcgagg gcccccgtga aggcattccg atctggaaag tgccagccac
2341 cgacacotac agcatctcgg gctacggagc tgctggcggg aaaggcgga agaaccacat
2401 gatgcggtcc cagggcggtg ctgtgctggg catcttcaac ctggagaagg atgacatgct
2461 gtacatctct gttgggcagc agggagagga cgctgcccc agtacaacc agttaatcca
2521 gaaagtctgc attggagaga acaatgtgat agaagaaga atccgtgtga acagaagcgt
2581 gcattgagtg gcaggagcg gaggaggagg ggggtggagc acctacgtat ttaagatgaa
2641 ggatggagtg ccggtgcccc tgatcattgc agccggaggt ggtggcaggg cctacggggc
2701 caagacagac acgttccacc cagagagact ggagaataac tcctcggttc tagggctaaa
2761 cggcaattcc ggaagccgag gtgtgagggt tggctggaat gataaacatt ccttgcctctg
2821 gggcggaaaa tctttgcagg aggggtgccac cggaggacat tcctgcccc agggcatgaa
2881 gaagtggggg tgggagacaa gagggggttt cggagggggg ggaggggggt gctcctcagg
2941 tggaggaggc ggaagatata tagggggcaa tgcagcctca acaaatgacc ccgaaatgga
3001 tggggaagat ggggtttctt tcactagtc actgggcatc ctgtacaccc cagctttaa
3061 agtgatggaa ggcacgggg aagtgaatat taagcattat ctaaactgca gtoactgtga
3121 ggtagacgaa tgtcacatgg accctgaaag ccacaaggtc atctgcttct gtgaccacgg
3181 gacggtgctg gctgaggatg gctgctcctg cattgtgtca cccaccccg agccacacct
3241 gccactctcg ctgactctct ctgtggtgac ctctgcccgc gtggcgccgc tggctctggc
3301 tttctccggc atcatgattg tgtaccgccc gaagcaccag gactgtcaag ccatgcagat
3361 gtagctgcag agccctgagt acaagctgag caagctccgc acctcgacca tcatgccga
3421 ctacaacccc aactactgct ttgctgqcaa gacctcctcc atcagtgacc tgaaggaggt
3481 gccgcggaaa aacatcacc tcattcgggg tctgggcat ggccctttg gggaggtga
3541 tgaaggccag gtgtccgaa tgcccaacca cccaagcccc ctgcaagtgg ctgtgaagac
3601 cctgcttcaa gtgtgctctg aacagggcga actggttttc ctctggaag ccttgcctat
3661 gacgaatttc aaccaccaga acattgttgc ctgactggg gtgagcctgc aatccctgcc
3721 ccggttcctc ctgctggagc tcattggggg ggggaacctc aagtccttcc tccgagagac
3781 ccgcccctgc ccgagccagc cctcctccct ggcctgctg gaccttctgc actgagctcg
3841 ggcattgccc tttggtctgc agtatttga ggaacccac ttcattccac gacacattgc
3901 tggcagaaca tgcctcttga cctgtccagg ccttggaaag gtggccaaag ttggaggtga
3961 cgggatggcc cgaacatct acaggcgag ctactataga aaggagggct gtccctgct
4021 gccagttaaq tggatgccc cagaggcctt catggaagga atattccatt ctaaaacaga
4081 cactatgccc tttgagtg tgctatgga aatctttct ctggaataa tggcataccc
4141 cagcaaaagc aaccaggaa ttctggaatt tgtaccagt gggggccgga tggaccacc
4201 caagactgc cctgggctg tataccggat aatgactcag tctgqgac atcagctga
4261 cagcagcccc aactttgca tcattttga gggattgaa tactgcacc aggaccgga
4321 tgaatcaaac accgctttgc cgtatgata tggctcact gtggaagag aagagaaagt
4381 gcctgtgagg cccaaggacc ctgagggggt tctcctctc ctggtctctc aacaggcaaa
4441 acgggagggg gaggcagcc cagctgccc accacctct cctaccacct cctctggcaa
4501 ggcgtgcaaa aaaccacag ctgcagaggt ctctgttcga gtccctagag ggcggccgt
4561 ggaaggggga cagtgcaata tggattctc taagtccaac cctccttcgg agttgcacg
4621 ggtccacgga tccagaaata agccacacag ctgttggaac ccaacgtac gctcctggt
4681 tacagagaaa ccacacaaa agataatcc tatagcaag aaggagccac acgagagggg
4741 taacctgggg ctggagggaa cgtgtactgt ccacctaac gttgcaact gggagactcc
4801 ggggggctca ctgctcctag agccctcttc gctgactgcc natatgaag aggtacctct
4861 gttcaggtca gttcacttcc ctgtgggaa tgtcaattac ggtaccagc aacagggct
4921 gcccttagaa gccgctaetg ccctggagc tggctattac gaggatacca ttctgaaaag
4981 caagaatagc atgaaccagc ctgggcccct agctcgtctg caccctcact tctcttctt
5041 gggatcccta agaccgtgga ggaagagag gcaatcnatg gctcctttca caaacagag
5101 accaagtgtc acgttttgtt ttgtgcaaac ctattttgaa gtaccacca aaagctgtg
5161 ttttgaatat gctttgaaa ggttttgagc atgggttcat cctattcttt cgaagagaga
5221 aatatatata aaatgagtg ataatataa ggcagatgt ggttgataa ggtttttatg
5281 catgtttgtt qta

Figura 4C

MNGQLDLSGKLIVKAQLGEDIRRIPIHNEDITYDELVLMMQRVFRGKLLS
NDEVTIKYKDEDGDLITIFDSSDLSFAIQCSRILKLTLEFVNGQPRPLESS
QVKYLRRRELIELRNKVNRLDSLEPPGEPGPSTNIPENDTVDGREEKSAS
DSSGKQSTQVMAASMSAFDPLKNQDEINKNVMSAFGLTDDQVSGPPSAPA
EDRSGTPDSIASSSSAAHPPGVQPQPPYTGAQTQAGQIEGOMYQQYQQQ
AGYGAQPPQAPPQPPQYGIQYSASYSQQTGPQPPQPFQGYGQQPTSQAP
APAFSGQPQQLPAQPPQYQASNYPAQTYTAQTSQPTNYTVAPASQPGMA
PSQPGAYQPRPGFTSLPGSTMTPPPSGPNPYARNRPPFGQYTPQPGPYR

Figura 4D

ATGAACGGACAGTTGGATCTAAGTGGGAAGCTAATCATCAAAGCTCAACTTGGGG
AGGATATTCGGCGAATTCCTATTCATAATGAAGATATTACTTATGATGAATTAGT
GCTAATGATGCAACGAGTTTTCAGAGGAAAACCTCTGAGTAATGATGAAGTAACA
ATAAAGTATAAAGATGAAGATGGAGATCTTATAACAATTTTTGATAGTTCTGACC
TTTCCTTTGCAATTCAGTGCAGTAGGATACTGAAACTGACATTATTTGTTAATGG
CCAGCCAAGACCCCTTGAATCAAGTCAGGTGAAATATCTCCGTCGAGAACTGATA
GAACTTCGAAATAAAGTGAATCGTTTATTGGATAGCTTGGAAACCACCTGGAGAAC
CAGGACCTTCCACCAATATTCCTGAAAATGATACTGTGGATGGTAGGGAAGAAAA
GTCTGCTTCTGATTCTTCTGGAACACAGTCTACTCAGGTTATGGCAGCAAGTATG
TCTGCTTTTGATCCTTTAAAAAACCAGATGAAATCAATAAAAAATGTTATGTCAG
CGTTTGGCTTAAACAGATGATCAGGTTTCAGGGCCACCCAGTGCTCCTGCAGAAGA
TCGTTTCAGGAACACCCGACAGCATTGCTTCCTCCTCCTCAGCAGCTCACCCACCA
GGCGTTCAGCCACAGCAGCCACCATATACAGGAGCTCAGACTCAAGCAGGTCAGA
TTGAAGGTTCAGATGTACCAACAGTACCAGCAACAGGCCGGCTATGGTGCACAGCA
GCCGCAGGCTCCACCTCAGCAGCCTCAACAGTATGGTATTCAGTATTCAGCAAGC
TATAGTCAGCAGACTGGACCTCAACAACCTCAGCAGTTCAGGGATATGGCCAGC
AACCAACTTCCAGGCACCAGCTCCTGCCTTTTCTGGTCAGCCTCAACAACCTGCC

TGCTCAGCCGCCACAGCAGTACCAGGCGAGCAATTATCCTGCACAACTTACACT
 GCCCAAACCTTCTCAGCCTACTAATTATACTGTGGCTCCTGCCTCTCAACCTGGAA
 TGGCTCCAAGCCAACCTGGGGCCTATCAACCAAGACCAGGTTTTACTTCACTTCC
 TGGAAGTACCATGACCCCTCCTCCAAGTGGGCCTAATCCTTATGCGCGTAACCGT
 CCTCCCTTTGGTCAGGGCTATACCCAACCTGGACCTGGTTATCGATAAGGAGGCT
 CCTCTACACCAATTAATGTAGCTGCTAGCTATTGGCCTCCCAAAAGACTCCAGTA
 CTATTTTAATTTGTATTGAAGAAGTTCAGAAATTTAAAAGCAGAGCATTTTTTTAT
 GATATCATTGTTGGTGTAAATTGAAAGTATAATTTGCTGGAACACAAAGACCAAA
 ATGAAAGTTTTTCCTCCCTGCTTAAAAATGTAGCAGCTTCTTAGTTACTTTGGA
 ACACTACTCTTACATGTATAAAGTGATTGACTTGACTTTCTAGCTTCCCTTGTC
 GGAGGATATTAAATGCTAGGGTGAGGTTTAGCCATCTTACTTGGCTTTACTAT
 TAACATGATGTACTAAAGTAGAGCCCTTTGAGAATACAAGATATTATGTATAAAA
 TGTAACACTGATGATAGGTTAATAAAGATGATTGAATCCATTAAAAAAAAAAAA

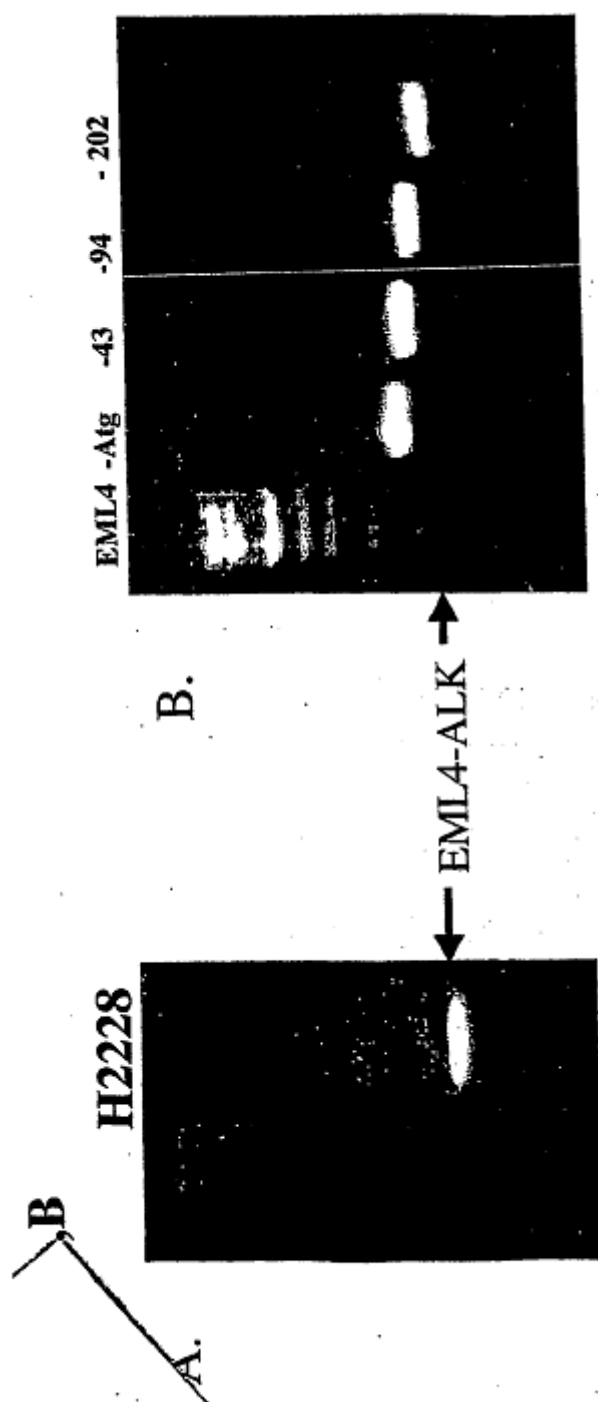


Fig. 5 C,D

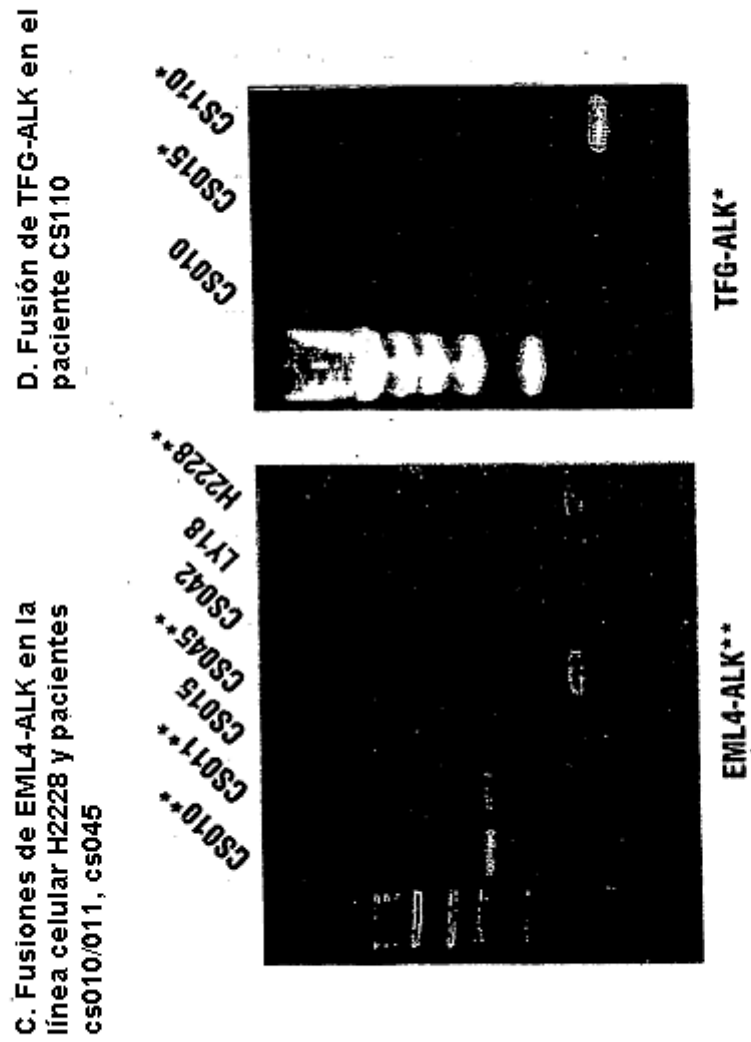


Fig. 6

